

**UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO**

**Cooperative Aerial Robotics and Artificial Intelligence
for Wildfire Detection and Monitoring Systems**

Maria João Santos Lopes de Sousa

Supervisor: Doctor Alexandra Bento Moutinho

Co-Supervisor: Doctor Miguel Abrantes de Figueiredo Bernardo de Almeida

Thesis approved in public session to obtain the PhD Degree in

Mechanical Engineering

Jury final classification: Pass with Distinction

2023

**UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO**

**Cooperative Aerial Robotics and Artificial Intelligence
for Wildfire Detection and Monitoring Systems**

Maria João Santos Lopes de Sousa

Supervisor: Doctor Alexandra Bento Moutinho

Co-Supervisor: Doctor Miguel Abrantes de Figueiredo Bernardo de Almeida

**Thesis approved in public session to obtain the PhD Degree in
Mechanical Engineering**

Jury final classification: Pass with Distinction

Jury

Chairperson: Doctor João Miguel da Costa Sousa, Instituto Superior Técnico, Universidade de Lisboa

Members of the Committee:

Doctor Antonios Tsourdos, Centre for Autonomous and Cyberphysical Systems, Cranfield University

Doctor Domingos Xavier Filomeno Carlos Viegas, Faculdade de Ciências e Tecnologia da

Universidade de Coimbra

Doctor Alexandre José Malheiro Bernardino, Instituto Superior Técnico, Universidade de Lisboa

Doctor Alexandra Bento Moutinho, Instituto Superior Técnico, Universidade de Lisboa

Funding Institution

Fundação para a Ciência e a Tecnologia

2023

The author acknowledges the support from Fundação para a Ciência e a Tecnologia (FCT), through the Ph.D. scholarship SFRH/BD/145559/2019. This work was partially supported by FCT, through IDMEC, under project Eye in the Sky, PCIF/SSI/0103/2018, and through IDMEC, under LAETA, project UIDB/50022/2020. This work also received funding from IEEE through the Robotics and Automation Society Special Interest Group on Humanitarian Technology (RAS-SIGHT) and from NVIDIA Corporation, through the GPU Grants Program.

Acknowledgments

This thesis is the culmination of several intense years of research that had a significant impact in my professional development and personal growth. Finishing this PhD degree closes a very important chapter in my individual path, and I owe my sincerest words of gratitude and appreciation to the many people who challenged and championed me throughout this very rewarding journey.

First, I would like to express my deepest appreciation to my PhD advisors, Prof. Alexandra Moutinho and Prof. Miguel Almeida for fostering such a strong spirit of multi-disciplinary research, for all the opportunities to join exciting field experiments and for the countless helpful discussions. I am extremely grateful for their collaboration, encouragement and support in all the bold ideas I dared to chase. That gave me freedom to explore and experiment on the research topics that fascinated me the most, while discovering the kind of scientist I would strive to become.

Many of the principles shaping this thesis were seeded long before this doctoral adventure even started. My great interest in systems engineering and robotics started through the influence of Prof. Jorge Martins, to whom I am deeply grateful for his inspiring and generous teaching in the numerous classes, labs and office hours throughout the many courses that I was so incredibly privileged to take under his guidance. While often things lose their magic as we disentangle their inner-workings, he had the unique ability to uphold it and with captivating enthusiasm bring a clear understanding that opened up immense possibilities for me.

As for my fondness for computational intelligence methods, equally prominent in my work, it is certainly due to the profound impact of the lessons I learned from Prof. Susana Vieira and Prof. João Miguel Sousa. The open-ended nature of the way the content was showcased in their courses remarkably stimulated my curiosity and was decisive for motivating me to later on pursue research avenues in intelligent systems. As my horizons have expanded afterwards, my appreciation for algorithmic elegance in system modeling has only strengthened, which was instilled in me in those early lessons that for sure have their imprint. I am deeply grateful for their influence and teachings, and overjoyed to continue to explore this wonderful research area.

The work I developed was supported by several institutions that undoubtedly elevated this research work. I thank IST and LAETA for the continuous support, particularly crucial in the first steps of research of my PhD. I am also indebted to Fundação para a Ciência e a Tecnologia, as being awarded an individual scholarship to pursue my dream doctoral research was instrumental for the development of this thesis. I am extremely grateful for the many doors it has opened for me and know I will always look back on this period with great pride and appreciation.

I would also like to thank especially the Center of Intelligent Systems (IDMEC) and the Forest Fire Research Center (ADAI) for welcoming throughout this period and providing solutions for the technical and bureaucratic challenges often encountered. I am very thankful for all the advice and support from many students, colleagues, senior researchers, professors and staff. I have always been convinced that there would be no better place to conduct this research work, a conviction which was always reinforced at every instance

I would present at computational intelligence and robotics gatherings, and whenever dipping my toes in wildfire research circles.

Although my experience working on wildfire-related research preceded my PhD, this time around it was remarkably different due to the aftermath of the large fire events of 2017. During these years, the community invested in working on wildfire-related problems which create a research environment that contributed to growing my awareness, knowledge and sensitivity to many aspects of this very complex issue. I thank ADAI, specially Prof. Domingos Xavier Veigas, Miguel Almeida and Luís Mário Ribeiro for their leadership in this space and for convening the research community working on this topic in many courses, workshops and field campaigns, which I benefited tremendously from attending.

In this scope, I had the unique opportunity to collaborate in the project Eye in the Sky, funded by FCT. This participation was essential to a large part of my research work and I am extremely grateful for having had the opportunity to learn immensely during this project and overcoming many challenges along the way. I am indebted to my advisors for this and to FCT for the chance to undertake extensive field work, which enabled me to shape my thesis in the direction that interested me the most, i.e., around real-world data.

For a large part of my PhD I also had the great honor to serve as a volunteer at Climate Change AI, collaborating on amazing projects for the community working at the intersection climate change and machine learning. I am greatly thankful for the opportunity to work with so many bright people across the core team. I would like to specially thank Priya Donti, Lynn Kaack, David Rolnick, Evan Sherwin, Kasia Tokarska and Ján Drgoňa for their trust and mentorship, and to my pals from the summer school team, Hari Prasanna Das, Jeremy Irvin and Olivia Mendivil with whom I have had so much fun in the last couple of years.

I would also like to thank Rafael Cordeiro, who I had the luck and privilege of being office mates with when he was a post-doctoral researcher at IST. I can not thank him enough for sharing his wisdom, experience and tricks with me, and especially for his empathy and help as I suffered through what I now know are the normal hurdles of a PhD experience. He definitely spared me of a lot of troubles through his optimized workflows and sound work practices, and I cherish dearly the laughter in the timely breaks. He is a brilliant researcher, dependable colleague and a trusted friend that I will continue to look up to.

This experience would most definitely not have been possible without the encouragement and support from my dear friend Tomás Hipólito, who has been there for me in my lowest lows and highest highs, from the start until the very end. By virtue of sharing the PhD journey we got to share exciting adventures and bond over frustrations and achievements, while enjoying endless conversations about teamwork, loyalty and leadership. Tomás is a trailblazer in ways the world has yet to value and recognize, his ability to transpire his vulnerabilities is the kind of strength I aspire to have someday, and has been an inspiration in my personal growth over these years.

This long ride has also been shared with Daniela Melo, a friend almost since my start at IST over a decade ago and during the last years in parallel as she pursued her PhD at EPFL. Despite our research topics standing on opposite extremes as they are fire and snow, our connection points and fun trips along the

way have been highlights of this experience and the source of many memories to cherish for life. My perspectives and outlook have largely been enriched by her wise and caring advice in many conversations along the way. I am deeply grateful for the friendship we have shared and certain that this journey would have been much harder without her support.

I could not have undertaken such a long academic journey without the true friendship of many colleagues I had the fortune to cross paths with at IST. I am indebted to Beatriz Arraiano, Afonso Ferreira, Pedro Cabral, Clara Correia, Eduardo Araújo, Ricardo Maia, Joana Correia, Maria Teresa Soveral, Frederico Alves, Patrícia Branquinho, Duarte Magalhães, António Santos, André Antunes, Pedro Fraga, Renato Severiano and Vasco Sampaio. Each, in their own way, has propelled me to overcome many challenges, and their candid words often echo in the back of my mind, having a long-lasting effect on how I lead my life and work.

Lastly, I would like to express my deepest gratitude to my family and lifelong friends who through their unconditional support, inspiration and example have enabled me to pursue this PhD degree and all my wildest aspirations and dreams. It is hardly possible to put into words how much I owe my parents and brother for their support, caring advice and patience while I have been away for over a decade now. Although I have changed and grown in many ways, you are the reason that I have stayed the same on what really matters, and that is my greatest privilege and treasure.

Resumo

Com as alterações climáticas, as condições meteorológicas de incêndio estão a tornar-se mais comuns em várias regiões do mundo. Em resultado, em combinação com as transformações sociais, os incêndios rurais estão a aumentar em frequência e intensidade, podendo resultar em grandes incêndios que excedem a capacidade de resposta operacional das equipas de emergência e criando condições extremas que colocam em risco as populações. A crescente dificuldade em lidar com estes fenómenos evidenciou várias lacunas nas infraestruturas de resposta a desastres, que exigem canais aprimorados de recolha de informações. Neste contexto, o interesse pela utilização de veículos aéreos não-tripulados (VANTs) para vigilância e monitorização ativa de incêndios tem aumentado nos últimos anos.

Esta tese propõe soluções inovadoras para sistemas descentralizados de monitorização ambiental para detecção e monitorização de incêndios rurais, baseados em redes de robôs aéreos que podem fornecer informações em tempo-real, através da utilização de robótica cooperativa, detecção remota e inteligência artificial. O sistema combina sensores estáticos, junto ao solo, e dinâmicos a bordo de plataformas robóticas aéreas, como drones e balões de alta-altitude que permitem agregar várias camadas de observação. A natureza descentralizada, multimodal e dinâmica do sistema proposto permite o seu accionamento em regiões específicas em períodos em que existem previsões de elevado risco de incêndio.

Para abordar os desafios de engenharia de sistemas ciber-físicos tão complexos, esta tese segue uma perspectiva de engenharia de sistemas, enquadrando a investigação particularmente numa abordagem como sistema de sistemas. As contribuições desta tese têm quatro partes. Primeiramente, a investigação foca-se na otimização do dimensionamento e projeto de configurações de frotas heterogêneas de VANTs, bem como em estratégias de coordenação cooperativa para redes multi-VANT resilientes. Em segundo lugar, realiza-se um estudo aprofundado para avaliar câmaras térmicas de infravermelhos e as suas capacidades complementares às câmaras do espectro visível para percepção robótica multimodal. Nesse âmbito, novos conjuntos de dados multimodais e abordagens de modelação de dados para detecção de incêndios rurais são propostas com base em dados de ensaios laboratoriais e de campo. Em terceiro lugar, desenvolvem-se abordagens de sensores inteligentes derivados de dados para detecção e monitorização de incêndios a partir de imagens térmicas e visíveis, usando modelos de lógica difusa e redes neuronais profundas baseados em métodos de aprendizagem automática. Em quarto lugar, são apresentadas abordagens de curadoria de dados para anotar dados de imagens de incêndio para abordar as lacunas em conjuntos de dados de elevada qualidade, que são essenciais para escalar essas soluções para implementações no mundo real.

Palavras-chave: sistemas inteligentes, incêndios rurais, detecção e monitorização de fogo, redes de sensores descentralizadas, percepção robótica multimodal, sistemas multi-agente

Abstract

With climate change, fire weather conditions are becoming more frequent in several regions worldwide. As a result, in combination with societal transformations, rural fires are increasing in frequency and intensity, which can result in large fire events that exceed the operational response capacity of emergency teams and create extreme conditions that endanger populations. The increased difficulty in addressing these phenomena has highlighted several hindrances in the disaster response infrastructure that call for enhanced intelligence gathering pipelines. In this context, the interest in the use of unmanned aerial vehicles (UAVs) for surveillance and active fire monitoring has been growing in recent years.

This thesis proposes novel solutions for decentralized environmental monitoring systems for wildfire detection and monitoring based on networks of aerial robots that can provide enhanced real-time wildfire intelligence by leveraging cooperative robotics, remote sensing and artificial intelligence. The system combines static sensors on the ground and dynamic sensors onboard mobile aerial platforms, e.g., drones and high-altitude balloons, that enable aggregating several observation layers. The decentralized, multimodal and dynamic nature of the system proposed enables its deployment target regions in periods when there are forecasts of increased fire risk.

To address the challenges of engineering such complex cyber-physical systems, this thesis follows a systems engineering perspective, framing the research particularly in a system of systems approach leveraging multi-agent systems and intelligent systems methods. The contributions of this thesis are fourfold. First, the research focuses on optimization of the dimensioning and design of configurations of heterogeneous fleets of UAVs, as well as on cooperative coordination strategies for resilient multi-UAV networks. Second, an in-depth study is undertaken to assess thermal infrared cameras and their complementary capabilities to visible spectrum cameras for multimodal robotic perception. In this scope, new multimodal datasets and novel data modeling approaches for wildfire detection are proposed based on extensive laboratory and field trial experiments. Third, solutions are developed for data-driven intelligent sensors for fire detection and monitoring for thermal and visible images using fuzzy models and deep neural networks based on machine learning methods. Fourth, data curation approaches for annotating fire image data are presented to address the gaps in high-quality datasets, which are essential for scaling these solutions to real-world deployments.

Keywords: intelligent systems, wildfires, fire detection and monitoring, decentralized sensor networks, multimodal robotic perception, multi-agent systems

Contents

List of Tables	xxi
List of Figures	xxiii
I Introduction	1
1 Wildfire Intelligence	3
1.1 Fire Hazard, Risk and Effects	4
1.2 Decision Support Systems for Wildfire Management	7
1.3 Environmental Monitoring for Wildfire Intelligence	8
1.3.1 Environmental monitoring challenges	8
1.3.2 Earth Observation solutions	9
2 Cyber-Physical Systems	11
2.1 Systems Engineering	12
2.2 System of Systems Engineering	14
2.3 Multi-Agent Systems	16
2.4 Intelligent Systems	19
2.4.1 System Modeling	20
2.4.2 Data Modeling	22
3 Proposed Approach	25
3.1 Proposed Systems Architecture	26
3.2 Thesis Objectives	28
3.3 Research Questions and Directions	28
3.4 Outline	33
4 Contributions	35
4.1 Contribution Overview	36
4.2 Decentralized Networked Systems	38
4.3 Multimodal Robotic Perception	39
4.4 Intelligent Fire Detection and Monitoring	40
4.5 Data Curation Approaches	42

II	Decentralized Networked Systems	43
5	Design of UAV Fleets	45
5.1	Introduction	46
5.2	Demand-Driven Network Optimization	48
5.2.1	Problem Statement	48
5.2.2	Demand Modeling	50
5.3	Clustering-based Graph Partitioning	52
5.3.1	Graph Model	53
5.3.2	Gustafson-Kessel Fuzzy Clustering	54
5.3.3	Hierarchical Density-based Clustering	55
5.4	Decentralized distribution of m-UAV fleets	56
5.4.1	Cluster Optimization and Validity Conditions	56
5.4.2	Design of Fleet Configurations	56
5.5	Results	57
5.6	Conclusion	59
6	Coordination of Decentralized Multi-UAV Fleets	61
6.1	Introduction	62
6.1.1	Related Work	63
6.1.2	Proposed Approach	64
6.2	Networked Aerial Systems	64
6.2.1	Problem Statement	65
6.2.2	Mathematical Modeling	65
6.3	Coordination of Multi-UAV Fleets	68
6.3.1	Fuzzy partitioning policy	69
6.3.2	Ant Colony Optimization	70
6.3.3	Cooperative Framework	71
6.4	Experiments	72
6.4.1	Case-study scenario	72
6.4.2	Simulation Setups	73
6.4.3	Performance Measures	73
6.5	Results	74
6.5.1	Clustering-based Graph Partitioning	74
6.5.2	Cooperative Decentralized Ant Colony Optimization	75
6.5.3	Convergence Analysis	75
6.6	Conclusion	76
III	Multimodal Robotic Perception Systems	77
7	Thermal Infrared Sensing	79
7.1	Introduction	80

7.1.1	Related Work	81
7.1.2	Proposed Approach	82
7.2	Thermal Imaging	83
7.2.1	Sensor characteristics: preliminaries	84
7.2.2	Mapping raw digital data to thermal images	86
7.3	Data-driven Thermal Imaging Analysis	93
7.3.1	Thermal Imaging in Fire Scenarios	93
7.3.2	Feature engineering	95
7.4	Controlled Fire Experiments	96
7.4.1	Laboratory Test	97
7.4.2	Caravan Burning Test	97
7.4.3	Summer Festival Trials	98
7.4.4	Mountain Range Field Trials	99
7.5	Results and Discussion	100
7.5.1	Laboratory Test	100
7.5.2	Caravan Burning Test	102
7.5.3	Summer Festival Trials	104
7.5.4	Mountain Range Field Trials	107
7.5.5	Discussion	109
7.6	Conclusion	110
8	MAVFire: Multimodal Dataset	113
8.1	Introduction	114
8.2	Related Work	115
8.3	Data Collection	116
8.4	MAVFire Dataset	118
8.5	Dataset Applications and Limitations	121
8.6	Conclusion	122
IV	Intelligent Fire Detection and Monitoring	123
9	Wildfire Detection Using Transfer Learning	125
9.1	Introduction	126
9.1.1	Wildfires in Wildland-Urban-Interface areas	127
9.1.2	Related work	128
9.1.3	Proposed approach	129
9.2	Deep Neural Networks	130
9.2.1	Convolutional Neural Networks	130
9.2.2	Pretrained models	131
9.2.3	Inception-v3	132
9.2.4	Model evaluation	133
9.3	Databases	134

9.3.1	Portuguese Firefighters Portal Database	134
9.3.2	Corsican Fire Database	134
9.3.3	Other open-source datasets	135
9.4	Methodology	135
9.4.1	Classification problem	136
9.4.2	Data augmentation	136
9.4.3	Database	138
9.4.4	Dataset partition	139
9.4.5	Hyperparameter selection	139
9.5	Results and discussion	140
9.5.1	K-fold cross-validation	140
9.5.2	Performance analysis	141
9.5.3	Classification analysis	143
9.6	Conclusion	148
10	Fire Monitoring with Fuzzy Models	151
10.1	Introduction	152
10.1.1	Related Work	153
10.1.2	Proposed Approach	154
10.2	Experiments and Data Acquisition	155
10.2.1	Thermal Cameras	155
10.2.2	Controlled Fire Experiments	156
10.2.3	Field Trials during a Summer Festival	158
10.2.4	Image Dataset	158
10.3	Thermal Imaging Analysis	162
10.3.1	Tent Burning Example	162
10.3.2	Feature Engineering	163
10.3.3	Data Analysis	164
10.4	Fuzzy Modeling	168
10.4.1	Classification Problem	169
10.4.2	Takagi-Sugeno Fuzzy Model	169
10.4.3	Gustafson-Kessel Fuzzy Clustering	170
10.4.4	Model Evaluation	172
10.4.5	Dataset Division	173
10.4.6	Parameter Estimation	174
10.4.7	Performance Evaluation	175
10.4.8	Model Parameters	176
10.4.9	Training Results	178
10.4.10	Testing Results	179
10.5	Conclusion	181

V	Data Curation Approaches	183
11	Expert-in-the-loop Data Annotation	185
11.1	Introduction	186
11.2	Image-based Wildfire Management Tasks	187
11.3	Proposed Approach	187
11.4	Conclusion and Future Work	189
12	Color-based Linguistic Fuzzy Models	191
12.1	Introduction	192
12.2	Related work	193
12.3	Dataset	194
12.4	Fire Data Annotation Pipeline	195
12.4.1	Problem Formulation	195
12.4.2	Color Feature Engineering	196
12.4.3	Color-based Superpixel Segmentation	196
12.4.4	Interpretable Rule-based Models	198
12.4.5	Implementation details	199
12.5	Experiments	201
12.5.1	Performance Evaluation	201
12.5.2	Results Evaluation	202
12.5.3	Understanding the Architecture	202
12.6	Conclusion	203
VI	Conclusion	205
13	Conclusions and Future Directions	207
13.1	Summary and Conclusions	208
13.2	Future Directions and Outlook	210
13.2.1	From Research to Deployment Roadmap	211
13.2.2	Implementation Considerations: Enablers and Barriers	213
13.2.3	Pathway to Impact	214
	Bibliography	217

List of Tables

2.1	Comparison between intelligent systems modeling approaches.	22
3.1	Outline relating the structure of the thesis with the proposed system.	33
5.1	Optimization results of selected network models for light-load scenarios.	59
5.2	Optimization results of selected network models for heavy-load scenarios.	59
6.1	Model parameters for each type of vehicle k .	73
6.2	Selected optimization results for case study with balanced demand profile.	75
7.1	Summary of the specifications of the thermal cameras.	85
7.2	RGB segmentation thresholds for feature construction.	96
8.1	Summary of experimental trials featuring key characteristics and flight altitude profiles.	119
8.2	Summary of data provided for each trial.	119
9.1	Other open-source datasets used in the literature.	135
9.2	Cropping areas height x width (in pixels).	137
9.3	Model performance evaluation.	140
9.4	Comparison of performance with different threshold values for the CV dataset.	141
10.1	Summary of the specifications of the thermal cameras.	156
10.2	Summary of the experimental tests.	158
10.3	Segmentation thresholds for feature construction.	164
10.4	Confusion matrix categories.	172
10.5	Formulas of the performance measures derived from the confusion matrix.	172
10.6	Summary of the samples in the training dataset.	173
10.7	Summary of the samples in the testing dataset.	173
10.8	Model benchmarking for selected parameters.	175
10.9	Results of the gray heuristic for different threshold configurations.	175
10.10	Results of selected fuzzy models with $\delta = 0.5$.	175
10.11	Cluster centers.	177
10.12	Consequent parameters.	177

12.1	Description of the ruled-based systems designed for fire data annotation.	197
12.2	Ablation study for model comparison.	203

List of Figures

1.1	Qualitative schema relating climate-related drivers of fire occurrences.	5
1.2	Diagram with possible effects of climate extremes for a normal distribution.	5
1.3	Historical wildfire trends in Portugal.	6
1.4	Wildfire management stages: prevention, preparedness, response and recovery.	7
1.5	Current EO capabilities for active fire monitoring from low-earth-orbiting satellites.	8
2.1	Abstract system model representation.	14
2.2	Comparison between system elements and system of systems elements.	15
2.3	Diagram of network topologies.	17
3.1	Diagram of the system architecture and data flows.	26
3.2	Diagram of data flows within the proposed architecture.	27
5.1	Demand-driven clustering for multi-UAV fleets.	46
5.2	Example of demand forecasting for a 5-day time horizon.	47
5.3	Demand modeling with Poisson point process.	48
5.4	Demand distribution comparison for a 5-day period.	49
5.5	Example of the multiple steps of the proposed framework.	58
6.1	Framework for cooperation of multi-UAV systems.	62
6.2	Example of forecasted tasks over a region of interest.	66
6.3	Balanced demand distribution using homogeneous (uniform) point process.	72
6.4	Clustering results with different overlapping degrees.	74
6.5	Convergence analysis for different network configurations.	76
7.1	High-level diagram of the process flow in uncooled microbolometer arrays.	84
7.2	Automatic gain control with linear histogram.	87
7.3	Automatic gain control with linear histogram.	88
7.4	Automatic gain control with plateau equalization algorithm.	89
7.5	Schematic of file structure of FLIR image format.	90
7.6	Examples of thermal imaging color palettes.	92
7.7	Comparison of results of AGC algorithms.	92

7.8	Fire ignition detection with a radiometric thermal camera.	93
7.9	Comparison of temperature and raw data and the evolution of the adaptive color scale.	94
7.10	Division of the color scale.	96
7.11	Laboratory trials with FLIR SC660 mounted on an elevated platform.	97
7.12	Caravan burning trial recorded.	98
7.13	Festival venue and camping areas during the event where the trials were conducted.	99
7.14	Thermal camera setup on tethered helium balloon.	99
7.15	Experiments in mountain range field trials.	100
7.16	Feature-based representation of the sensor response of thermal camera to a fire ignition.	101
7.17	Effect of temperature-dependent non-uniformity correction (NUC).	102
7.18	Filtered sensor response of thermal camera to a fire ignition inside the caravan.	103
7.19	Results for data acquired surveying the festival venue.	105
7.20	Results for data acquired surveying the community kitchens area.	106
7.21	Results for fire detection at long-distance in a mountain range.	107
7.22	Results for fire aftermath captured at long-distance in a mountain range.	108
8.1	MAV-based data collection system.	116
8.2	Data collection setup in field experiments.	118
8.3	Dataset examples with multiple types of experiments.	120
9.1	Comparison between fire hazard assessment and burned area maps.	127
9.2	Example of a convolutional layer.	131
9.3	Parameter sharing scheme.	131
9.4	Inception-v3 network architecture.	132
9.5	Samples of fire (top row) and not fire (bottom row) from PFPDB.	134
9.6	Rescaled samples of <i>fire</i> images from CFDB.	135
9.7	Samples of positives and negatives for both classes.	136
9.8	Preprocessing schema for data augmentation.	137
9.9	Description of cross-validation database.	138
9.10	Description of procedure to partition data for k-fold cross-validation.	139
9.11	Comparison of threshold values using ROC curves.	141
9.12	Transparent confusion matrix depicting the mean fire prediction scores for fold 10.	142
9.13	Histogram of misclassifications for fold 10 (10 iterations).	142
9.14	Successful classifications of <i>fire</i> (left and center) and <i>not fire</i> (right).	143
9.15	False negative classification associated with reduced spatial scale.	144
9.16	Examples of false negative classification associated with fire pattern spatial location.	145
9.17	Comparison of cross-channel variance.	146
9.18	Example of false positives due to sunset (left and center).	146
9.19	Examples of false positives associated with the presence of smoke.	147
9.20	Overall misclassifications sorted by original image.	148

10.1	Schema of the proposed approach.	154
10.2	Thermal radiometric and nonradiometric cameras.	155
10.3	Experimental setup with a drone and static platform.	157
10.4	Experimental setup for the burning of a tent.	157
10.5	Partial view of the camping area at the case study summer festival.	159
10.6	Balloon setup.	159
10.7	Images acquired from the static elevated platform.	160
10.8	Perspectives with 3 different cameras.	160
10.9	Image data from the burning of camping tent.	161
10.10	Heat detection over the camping area captured with the FLIR Vue Pro.	161
10.11	Tent burning experiment: color images (top row), FLIR Vue Pro (bottom row).	162
10.12	Division of the FLIR GrayRed palette into color segmentation classes.	164
10.13	Color segmentation results.	165
10.14	Partial results from the festival tests with FLIR Vue Pro.	166
10.15	Color segmentations results.	167
10.16	Partial results from the festival tests with FLIR Vue Pro.	167
10.17	Sequence of frames taken over one of the community kitchens with the FLIR Vue Pro.	168
10.18	Accuracy results for 500 iterations.	174
10.19	Antecedent membership functions for three clusters.	177
10.20	Training results: drone test.	179
10.21	Training results: festival without fire (baseline).	179
10.22	Testing results: platform (Vue).	180
10.23	Testing results: outdoor festival (test 4).	180
11.1	Samples of fire and not fire instances captured from different points of view.	188
11.2	Expert-in-the-loop system comprising a computational data annotation pipeline.	188
12.1	Overview of the fire data annotation pipeline.	192
12.2	Selected samples of the fire image dataset and respective ground truth masks.	194
12.3	Description of the fire data annotation pipeline.	195
12.4	Fire features in the HSL color space.	196
12.5	Fire features in the YCbCr color space.	197
12.6	Detailed illustration of the superpixel merging procedure.	198
12.7	Comparison of superpixel segmentation with different parameters.	200
12.8	Examples for each step in real-world scenarios.	201
12.9	Examples of limitations in real-world scenes.	203
12.10	Fine-tuning of fire classification threshold for improving the semantic segmentation.	204
13.1	Resources availability for innovation at various TRLs.	211

I | Introduction

Contents

1 Wildfire Intelligence	3
2 Cyber-Physical Systems	11
3 Proposed Approach	25
4 Contributions	35

Summary

The introduction of this thesis comprises two distinct chapters, the first regarding the essential motivation of this work, and the second covering the fundamental principles of the proposed approach. Chapter 1 opens by presenting the current and emergent societal challenges in facing wildfire events under the ongoing climate crisis. Then, follows the demand for enhanced fire intelligence systems and description of current Earth Observation solutions. Chapter 2 introduces the background on cyber-physical systems employed for designing environmental monitoring networks, starting by describing concepts from systems engineering, systems of systems, and multi-agent systems. Chapter 3 - Proposed Approach presents the overview of the proposed architecture to devise decentralized multi-robot networks for wildfire detection and monitoring. Then, the thesis scope and contributions are presented in Chapter 4 - Contributions, along with an outline of the structure of this work.

1 | Wildfire Intelligence

Contents

1.1	Fire Hazard, Risk and Effects	4
1.2	Decision Support Systems for Wildfire Management	7
1.3	Environmental Monitoring for Wildfire Intelligence	8

In this introductory chapter, we start by framing the challenges related to fire hazard in the context of the ongoing climate crisis and present an overview of the present technological infrastructure for fire preparedness and response. Then, follows a description of the state-of-the-art environmental monitoring solutions available for fire-relevant observations, along with its respective advantages, drawbacks and limitations that allow the identification of the current gaps in fire intelligence that guide the motivation of the proposed approach.

1.1 Fire Hazard, Risk and Effects

Fire is an essential process in Earth System dynamics that shapes ecosystems and landscapes worldwide. This complex phenomenon depends on climate, vegetation and land-cover [1], being closely tied to the natural dynamic equilibria between recurrent ecosystem cycles that influence the characteristics of fire regimes, leading specific regions to be inherently more susceptible to fire occurrences [2].

The drivers of wildfire occurrences can be divided into two main categories: natural and anthropogenic [3]. Natural drivers include the climate, vegetation, weather conditions and topography, whereas anthropogenic factors include human-induced ignitions, land use practices and also fire management policies [4]. While some of these drivers are structural and make certain geographic regions across the globe more fire-prone, they vary as a result of human activities over time [5]. The interplay between these various drivers makes wildfire a very complex phenomenon, ever-changing and without a universal solution, requiring multifaceted approaches adapted to local characteristics.

In this introductory section, we start by outlining some grounding concepts [6] used to explain the motivation and specificities related to wildfire applications, before delving into the technological and scientific aspects pertaining to this research.

Fire hazard: Represents the intensity with which an area is likely to burn in case of a fire occurrence based on the fuel structure and topography that influences potential fire behavior, without regard to the state of weather-related variables.

Fire danger: The combination of both constant and variable factors of the fire environment that determine the ease of ignition, rate of spread, difficulty of controlling a wildfire, and fire impacts.

Fire risk: Probability resulting from the combination of the likelihood of a fire occurring and the potential impacts of that fire.

Fire regime: Pattern of natural or human-caused fire activity that characterizes a given area, e.g., the frequency or fire interval, extension of the fire season, and the number, type, and intensity of fires.

The climate type is a structural driver in defining the ability of a given region to carry fire and the inherent productivity of wildfire occurrences [7]. As illustrated in Fig. 1.1a, the tropics and temperate regions have the combination of two important factors: temperature and vegetation. In contrast, desert and arctic regions, due to the low vegetation or the temperatures encountered, are less prone to fire events. In addition, as depicted in Fig. 1.1b, high-mesic regions with high moisture levels and, conversely, low-arid regions with low moisture levels limit the productivity gradient of fire occurrences. These contrasts translate a key relation concerning fire hazard in the sense that either climate limiting or fuel limiting factors influence the potential effects of fire occurrences, making for a wide diversity of fire regimes worldwide. For each region, the fire regime also has a vital role in regulating natural ecosystems and biodiversity, and shaping landscapes by preventing fuel load accumulation, controlling invasive species and improving soils [8].

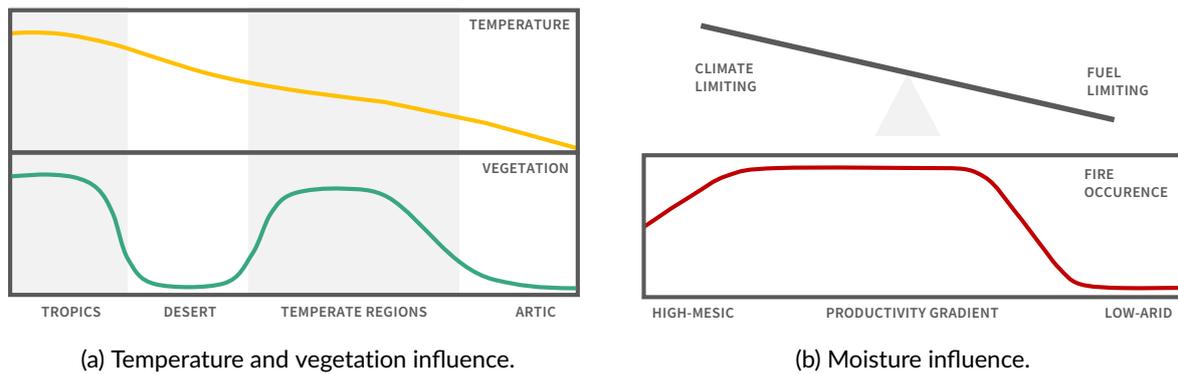


Figure 1.1: Qualitative schema relating climate-related drivers of fire occurrences.

Although fire depends on the confluence of fuel, oxygen, and heat, most ignitions are anthropogenic in nature, driven by accidental, negligent, and intentional human activities that make fire-prone regions increasingly vulnerable on environmental and societal levels. In addition, besides human intervention as an ignition source, fire exclusion practices with prominent fire suppression and fine fuel accumulation result in human-altered fire regimes [9], in which extreme weather events can lead to natural disasters with fires reaching intensity levels that can no longer be directly suppressed or controlled, threatening environmental resilience and posing tremendous risks for populations at the wildland-urban interface [10].

Wildfires and Climate Change

Due to climate change, natural fuels are dryer and fire weather conditions more frequent, which increases ignition propensity, extends fire seasons [11] and causes fire events to be more frequent, intense and severe [12]. With the escalation of these phenomena, fire science and fire management are central for the understanding of these phenomena and addressing the risks and effects posed by these events.

With climate change the probability of occurrence of extreme conditions increases as can be observed in Fig. 1.2 that presents an illustrative diagram showcasing the implications of increase in mean temperature and/or of variance (the combination of both effects is also possible). In these scenarios, the increase in average temperature can lead to the increase in number of days with adverse weather in both cold and hot extremes. For wildfires occurrences, the increase in more hot weather days and more extreme hot weather days can create conditions of very high ignition propensity, in which the occurrence of a fire can quickly escalate to large proportions.

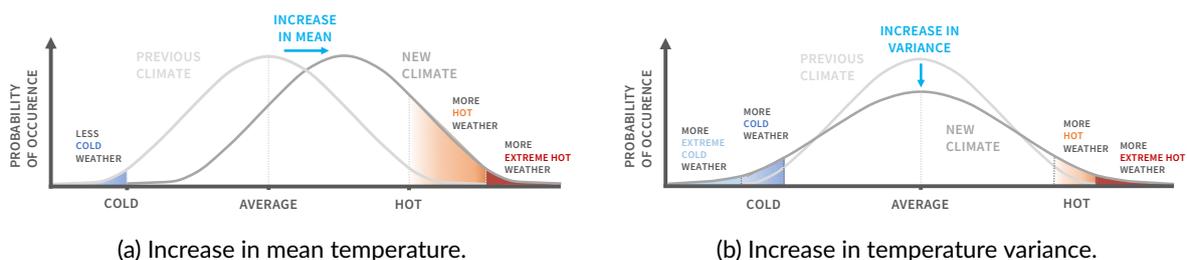


Figure 1.2: Diagram with possible effects on climate extremes. Adapted from [11].

In that sense, it is important to provide an example of the fire activity trends in Portugal, which is a region with high susceptibility to fires given its temperate climate with areas featuring abundant vegetation. In Fig. 1.3, we showcase the historical data collected by ADAI of the past wildfire occurrences recorded on a yearly basis and the respective total area burned per year, as well as the records of the five largest wildfire events per year.

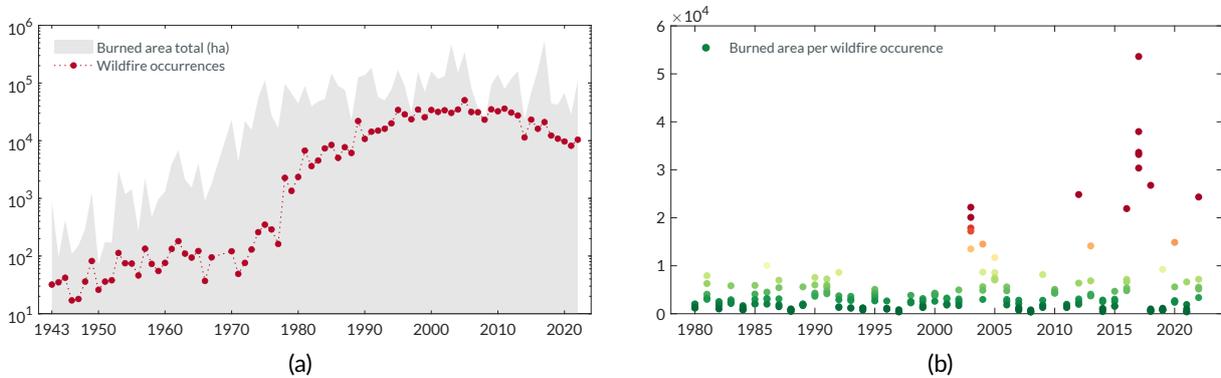


Figure 1.3: Historical wildfire trends in Portugal. Featured data courtesy of ADAI [13].

Analyzing the fire trends over the last decades in Portugal, in Fig. 1.3a, it is possible to observe that there is a steep rise in the period between 1970 and 1990, after which the total number of wildfire occurrences per year has stabilized. Although this evolution denotes the improvement in wildfire record systems, this also reflects the profound socioeconomic transformations over that period that have seen the abandonment of rural regions in favor of urban counterparts, leading also to the widespread halt of areas of the primary sector, e.g., agriculture, forestry, as well as pastoral and livestock activities [14]. In turn, forested areas proliferated undermanaged thereafter, having increased long-term natural fuels accumulation [15].

In addition, as fire suppression strategies have become highly effective in an initial response stage, even small fires that do not present significant danger are extinguished early on. This has led to fire exclusion in territories where low intensity wildfires would have been fundamental to regulate ecosystem cycles and manage fuel accumulation over time, and thereby preventing greater risks and impacts associated with higher fuel loads. Moreover, the nonexistence of land-use by primary activities surrounding villages brings unmanaged forests to the boundary of the wildland-urban interface, therefore putting populations in close danger. These risks are most troubling given the lacking infrastructure for response in rural areas.

With these human-driven changes, while the number of fires has maintained stable, the structural factors for wildfires to potentially be larger have increased gradually season after season, with the natural cycles in vegetation growth in precipitation periods, and the drying of fine fuels in heat waves and drought spells. As a result, when fire weather conditions are present the potential impacts of an ignition can be much larger. This can be observed in Fig. 1.3b, which shows the largest wildfire events for each year, and demonstrates that the magnitude of fires in terms of burned area has been rising, particularly since 2003.

Notwithstanding, it is important to also note that the vast majority of fire ignitions are from causes directly related to the human presence and activities, namely by negligence or through risk behaviors in the use of fire in rural spaces that precipitate the occurrence of wildfires.

1.2 Decision Support Systems for Wildfire Management

Wildfire phenomena affect disproportionately communities that are underserved by a societal unjust valorization of ecosystems services and inadequate land management and land-use policies, which make these communities more susceptible and vulnerable to these events. While preparedness and response improvements can help minimize its impacts, the dire effects are endured by communities far beyond the timescale of a fire event.

While acknowledging that despite the high spatial and temporal uncertainty of these events, it is an expected and necessary phenomena. In that sense, it is crucial to stress the urgency in improving wildfire management across its multiple vectors: i) prevention, ii) preparedness, iii) response and iv) recovery.

To that end, enhanced fire intelligence plays a pivotal role through the emergence of advanced technologies for mapping and real-time monitoring purposes, allowing its incorporation in the decision chains both in risk mitigation policies and prevention actions, improved early-warning and real-time monitoring services, as well as informed recovery planning for building resilient communities.

The advent of new advanced environmental monitoring paradigms can contribute to the several stages of wildfire management, namely in the tasks outlined in Fig. 1.4.

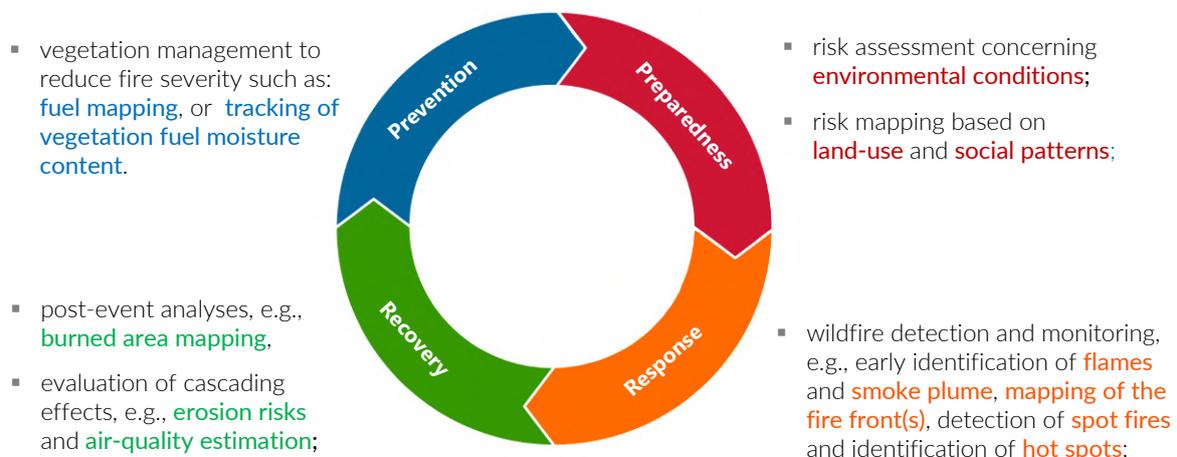


Figure 1.4: Wildfire management stages: prevention, preparedness, response and recovery.

This thesis focuses on the response stage, more specifically in the wildfire detection and monitoring tasks. In this scope, the topic of study centers on addressing the real-time intelligence gathering in the contexts of early detection, active fire monitoring and detection of hotspots. As will be discussed in the next sections, these areas have significant infrastructure and technological gaps that can be unlocked by the new emerging technologies like autonomous robotics and artificial intelligence.

Given the increased urgency for enhanced fire intelligence to tame the dire impacts of these events, this thesis proposes novel environmental monitoring solutions capable of providing systematic observations that can deliver real-time data to decision-makers, civil protection agencies and emergency response teams.

1.3 Environmental Monitoring for Wildfire Intelligence

1.3.1 Environmental monitoring challenges

With wildfires becoming more frequent and severe worldwide, fire monitoring systems face significant difficulties as existing solutions have important limitations in providing comprehensive situational awareness. While ground-based sensors can provide real-time data [16], their area coverage is very limited and infrastructure-wise very expensive to scale to provide wide situational intelligence. In turn, low-altitude unmanned aerial vehicles (UAVs) provide increased flexibility because of their high mobility, but due to their very inefficient flight endurance [17], missions can also only provide localized situational awareness. On a broader scale, satellites in low-earth orbit provide considerable intelligence value and are widely used to inform emergency response teams and civil protection agencies [18], however despite providing a global area coverage its time coverage is extremely limited, providing a limited number of observations from different satellite sources as illustrated in Fig. 1.5.

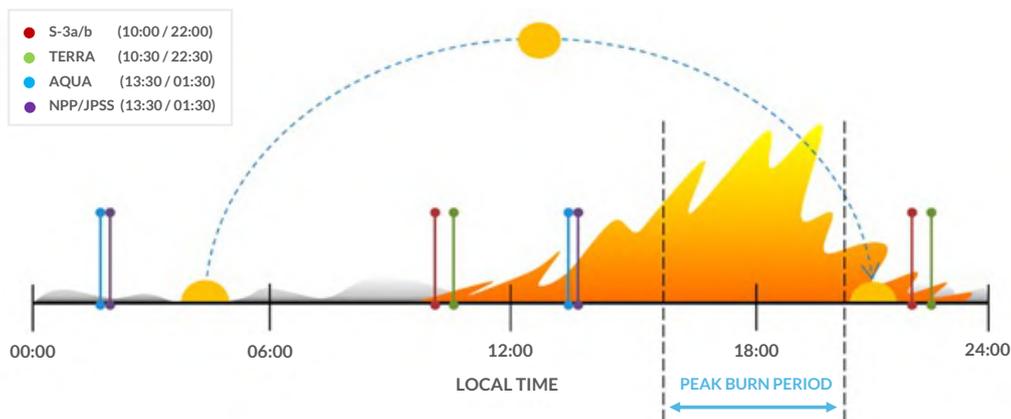


Figure 1.5: Current EO capabilities for active fire monitoring from low-earth-orbiting satellite observations from instruments such as SLSTR (Sentinel-3a/b), MODIS (TERRA/AQUA), and VIIRS (NPP/JPSS) fail to observe wildfires during the most active portions of the day (Adapted from [18]).

Given these challenges, high-altitude pseudo satellites (HAPS) can complement these alternative modalities and leverage their characteristics to tackle several of the current gaps [19], namely by providing a very wide area coverage by operating in the stratosphere at about 20km altitude, bringing benefits in the spatial resolution of imagery but also to its time-resolution as these communications can be performed in real-time. Notably, the former Loon company had implemented large-scale balloon-borne infrastructure for communications provisioning [20].

In this context, the advances in artificial intelligence (AI) can be an enabling technology for situational awareness and decision support systems. Recent works have proposed AI solutions for image-based fire detection tasks, however the quality and limited size of image databases available limit reliability for deployments in real contexts. Although transfer learning and data augmentation techniques have been explored, the lack of large-scale databases for wildfire detection tasks is a known hurdle in developing AI algorithms with suitable generalization, hindering its reliability.

1.3.2 Earth Observation solutions

The evolution of remote sensing technologies led to the development of monitoring systems using satellite-based information to forecast fire hazard and allow risk assessment, e.g., European Forest Fire Information System (EFFIS), European Centre for Medium-Range Weather Forecasts (ECMWF). Nevertheless, while providing wide coverage, the latency and resolution of current satellite solutions is still a roadblock for an application as time-sensitive as wildfire detection.

To mitigate this, aerial systems can provide local coverage with high spatial and temporal resolution, in areas identified as being of risk. Currently a large stake of aerial surveillance missions is performed with piloted aircraft, but the expensive and dangerous nature of this type of operation make this solution unfeasible to scale. Alternatives are unmanned aerial vehicles (UAVs), such as multi-rotors, fixed-wing aircraft or airships used individually or cooperatively. High-altitude balloons are low-cost platforms that can go as high as the stratosphere and provide wider aerial coverage being however wind-driven.

Image-based sensors are widely used for wildfire detection with visible range, thermal and multispectral technologies. Solutions increasingly combine computer vision and intelligent systems, including neural networks, fuzzy clustering, or support vector machines, but are largely limited to a single sensor type. Thermal-based approaches are currently understudied, but have remarkable potential for this application, having renewed interest due to the decrease in equipment costs and weight/size ratio. In turn, methods applied to sensors networks monitor environmental variables, e.g., temperature, humidity, and/or gas levels, which are complementary to information from image-based sensors.

However, with ever-increasing data streams, centralized systems are unable to handle the torrent of information from numerous sources. This requires networks to be designed following decentralized or distributed approaches, and that estimation and inference algorithms comply with these paradigms.

Despite extensive research on this topic, its complexity is typically avoided with domain-specific approaches that are designed for single data types, often reporting high false alarm rates, which hinders its application in real-world contexts.

2 | Cyber-Physical Systems

Contents

2.1 Systems Engineering	12
2.2 System of Systems Engineering	14
2.3 Multi-Agent Systems	16
2.4 Intelligent Systems	19

To bridge the gap between the need for real-time data with high spatiotemporal resolution in specific areas and the current Earth Observation monitoring systems, this work proposes a novel integrated networked system architecture for wildfire detection and monitoring. The dynamic system combines static sensors on the ground and dynamic sensors onboard mobile aerial platforms e.g. drones and high-altitude balloons, that can be deployed to monitor areas of high risk, based on satellite data or on-demand requests. The system relays real-time data for decision-making, enabling enhanced situational-awareness for decision and operational support, e.g. with respect to early warnings and optimization of resource allocation.

This chapter focuses on introducing the background principles and methods used in this thesis for modeling networked cyber-physical systems. It starts by outlining the systems engineering approach employed for the high-level modeling of the proposed networked architecture, which enables a holistic understanding of how this system is designed and the interactions between its several layers. Then, being a networked system composed of numerous entities, the model naturally evolves to a system of systems framework that can leverage multi-agent systems concepts for coordination and cooperation, as well as intelligent systems approaches for optimization and robotic perception.

The study of dynamical systems has continuously aimed to formulate simple, powerful abstractions to model systems by describing its characteristics through establishing relations between its inputs and outputs. This notion has enabled the development of extensive mathematical methods that are the foundations of system modeling that have permeated many engineering disciplines. However, to address increasingly demanding requirements, engineering systems are progressively more complex, often calling for contributions from different fields of expertise. In that sense, the discipline of systems engineering has gained prominence over the last decades for addressing the holistic development of complex problem-oriented systems [21]. In this context, this approach is widely used to model cyber-physical systems, a type of systems designed to leverage computing to interact with the physical environment.

In this work, the backbone of the proposed architecture is a systems engineering perspective, which will provide a template for the system concept, as well as render a roadmap for the structure of this document. For this reason, the following sections aim to describe the objectives and principles of this approach, which are followed by the introduction to the frameworks of system of systems and multi-agent systems, also cornerstones of the intelligent networks proposed.

2.1 Systems Engineering

Systems engineering is a branch of engineering centered on methodical, multidisciplinary approaches for the holistic design, realization, operation, and management of complex problem-driven systems [22]. These approaches deal with a global assessment of problem requirements and the development of a system with the ability to satisfy those needs. Under this framework, complex systems are understood in its complete form as an integration of several subsystems with different functions to reach the required capability.

Systems engineering has unique characteristics that depart considerably from many engineering fields that design and analyze isolated components separately with a single-discipline view. In contrast, rather than a "*divide and conquer*" approach, systems engineering follows more of a "*unite and build*" problem-solving strategy as it is concerned foremost with the big picture and understanding how the parts of the system are interconnected. By considering the contributions of multiple disciplines, e.g., from mechanics, electronics, cybernetics, human-machine interaction to decision-making, this approach can reconcile the whole as more than the sum of its parts. For this reason, this transversal approach can achieve tradeoffs and compromises between the specifications of the several elements of the system to optimize the solutions created. Hence, the reasons why it is a powerful problem-solving and high-level modeling approach.

In the following, we start by establishing the key core notions from systems engineering that motivate the use of its principles in the research that underpins this thesis, as will be explained throughout this chapter.

System: A collection of elements and their interrelations (e.g., physical, functional, and/or behavioral) combined to form a unified whole greater than the sum of the parts, i.e., a system-level capability only achievable by the integration of the various parts that provides a solution to a required need [22].

To address complex systems, systems engineering theory and practice devises systems **concepts**, **principles** and **reasoning tools** that enable breaking down complex problems and building up integrated solutions. From the comprehensive framework the systems thinking philosophy provides, which are extensively detailed in [23], in this introduction we will highlight the ones that are most relevant in this thesis.

Concepts – *abstractions, or general ideas inferred or derived from specific instances.*

Wholeness and Interaction: A holistic understanding of the system as whole and how elements interconnect is fundamental for building desired end-goal functionalities.

State and Behavior: The state is a set of attributes characterizing the system at a given time that can be changed by events over time depending on the behavior of the system, i.e., its response.

Adaptation and Learning: The ability to change operation modes to adjust to environment changes or respond to distinct functional demands, and the capability to improve its effectiveness.

Principles – *general basis for reasoning about systems thinking or systems approaches.*

Abstraction: A focus on essential characteristics is important in problem solving because it allows problem solvers to ignore the nonessential, thus simplifying the problem.

Encapsulation: Hide internal parts and their interactions from the external environment.

Modularity: Unrelated elements should be separated or decoupled into known discrete modules, and related parts of the system should be grouped together with a functional hierarchy.

Parsimony: One should choose the simplest explanation of a phenomenon, such that it requires the fewest assumptions, for requirements, design and operations.

Relations: The interconnections between elements are key to characterize the system, shaping the structure, topology and data flows, e.g., through feedforward and feedback links.

Reasoning Tools – *representations to ease the understanding of system boundaries and interactions.*

Model representations: Mental models and system mappings.

Visual representations: Block diagrams and data flow diagrams.

These notions of systems thinking are harnessed and formalized through descriptive models that embody themselves most of these concepts. To provide a generic example of a model representation, Fig. 2.1 presents a system, subject to a set of requirements, with an abstract rendering of its several modular elements and connecting relations, as well as how it relates to the environment through inputs and outputs.

While systems engineering develops vast frameworks to formally address system synthesis, operation, management and life-cycle planning, in this work the main focus is on system architecture for adaptive environmental monitoring networks and its implications for development of intelligent systems solutions.

The concepts, principles and tools from systems thinking herein presented are widely applied throughout this thesis, and also transfer to the abstractions to be introduced in the remaining of this chapter.

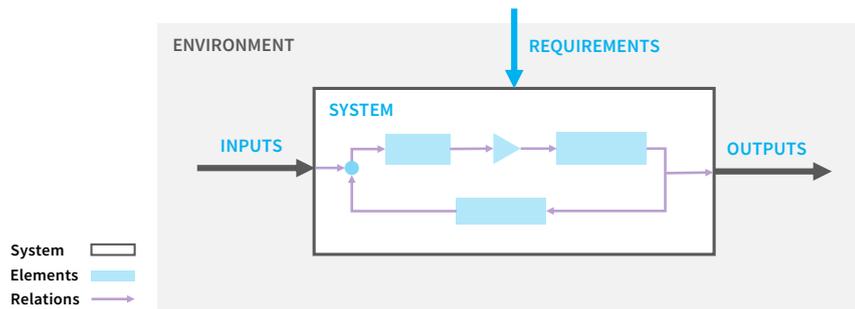


Figure 2.1: Abstract system model representation.

2.2 System of Systems Engineering

With the advent of technologically empowered societies, there is an ever-increasing number of complex systems. That poses both a challenge and an opportunity because novel system architectures are built of intricate subsystems, but also combine a set of several systems — a concept termed **system of systems**. Although a system of systems can be viewed under the lens of systems engineering, the nature of this special type of architecture arises particular differences that warrant further clarification and highlight [24].

On the one hand, systems are designed for providing a desired capability but also to be manageable, which imposes limits to the size and complexity of their architectures. On the other hand, in the real world systems do not operate in a vacuum, but rather exist in an environment with a multitude of monolithic systems, each designed for a narrow purpose. Naturally, as in most cases, no single system can address the extent of the intricacy of complex problems, societies rely on the interactions between several systems.

For these reasons, the nature of interconnection between these systems is not static nor steady but rather highly variable. In this way, to develop specific methods that can model these higher levels of uncertainty, and leverage those to build more adaptable and flexible systems, arises the emerging field of **systems of systems engineering**.

Tackling the development of a **system of systems** (SoS) requires a broader understanding than what the systems engineering approach by definition can provide because this "system" departs from a stand-alone nature built upon self-contained processes. Instead, aspects pertaining to the environment in which it operates, and essentially the emergent behavior resulting from its nature and interactions, become nuclear to addressing a system of systems architecture. For this reason, it is important to clarify a definition for this concept and its characteristics, to then motivate its relevance and purpose in this work.

System of Systems: A collection of connected (in)dependent legacy, evolving, and novel systems, that provide a unique capability, which must cope with uncertainty in the environment and its intrinsic structure by having a high degree of flexibility and adaptability [24].

Noting the greater emphasis on features such as adaptability and flexibility and the interplay between multiple actors, next we delve deeper into the distinctive characteristics of this type of system. Figure 2.2 juxtaposes system of systems core characteristics with monolithic system characteristics based on the differences of the nature of composing parts, i.e., its elements.

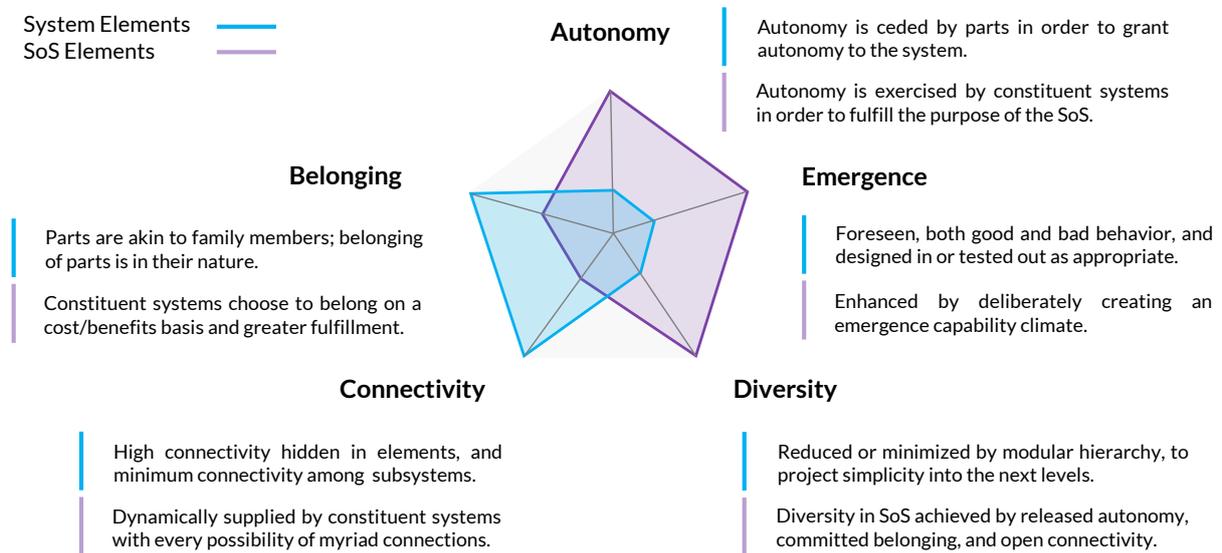


Figure 2.2: Comparison between system elements and system of systems elements based on [25].

While a system of systems must still be understood holistically as a system, addressing its elements, which are also complex systems, requires a more evolved approach. Higher degrees of autonomy, emergence, and diversity will significantly influence the state and behavior of the system. Conversely, wholeness and interaction will be radically impacted by the different natures of belonging and connectivity in SoS.

To address these issues, the nascent field of system of systems engineering (SoSE) specializes in developing multi-faceted systems that incorporate largely autonomous, heterogeneous agents with a variety of objectives and priorities. This type of evolving systems with dynamic connections and emergent behavior dimensions extends beyond the traditional SE approach. Thus, requiring devising adequate architectures, processes, and tools for design, deployment and decision-making for system of systems contexts.

In that sense, next we clarify the scope of both SE and SoSE approaches with the following definitions, and subsequently we reflect in further detail on how these two approaches are explored in this research.

Systems Engineering (SE) is an interdisciplinary engineering field concerning the development of systems that can deliver a system-level capability based on modular, interconnected and interdependent subsystems, which are integrated in an explicitly defined form to result in an expected behavior [22].

Systems of Systems Engineering (SoSE) is a subfield of systems engineering that concerns the development of systems leveraging flexible boundaries and interactions between independent, distributed, and evolving constituent systems and their stakeholders [24].

To provide some intuition of the distinction of both paradigms, we start by translating these concepts into a concrete example of a basic capability, e.g., localization awareness of a robot in an unknown environment in the real-world. A robot in an open unstructured environment, i.e., what is often termed *in the wild*, can hardly operate reliably without external "help". As such, to have an estimate of its own position in the environment it operates, i.e., on an open world map, it is designed to rely on an array of external systems for localization, most notably a global navigation satellite system (GNSS). In this way, this simple example allows grasping that a system-centric approach would be limited to address several complex dimensions of more advanced solutions, namely when the robot would need to navigate a changing environment.

Given that sensor networks must perceive their environment and interact with existing legacy infrastructure, the modeling of this type of systems is best addressed with a System of Systems engineering framework, which relies on the special characteristics and complementary capabilities of a variety of systems, e.g., the multitude of satellite-based services used for remote sensing, weather forecasting and even global positioning systems. Thereby, to design decentralized intelligence architectures that can deliver the necessary awareness features, i.e., self-awareness, collective awareness and situational awareness.

In that sense, for this thesis on decentralized sensor networks for fire detection and monitoring, the use of both SE and SoSE approaches becomes necessary to tackle the main research gaps this work addresses. First, SE is fundamental in designing networks of mobile aerial robots that can perform fire detection and monitoring, and to have a deep understanding of its inherent subsystems. Second, given the nature of the dynamic networks intended the SoS gains more prevalence due to the need to devise algorithms that can confer the necessary adaptability and flexibility features into sensor-driven architectures. Third, in the genesis of the proposed architecture, it is underlying that this system will add complementarity to the existing real-time monitoring gaps. Therefore, incorporating along with legacy infrastructure and other emergent solutions, the next generation of highly interconnected and integrated systems of systems that will provide fire intelligence to support decision-making.

In this context, beyond presenting the system architecture for the proposed decentralized sensor networks, this thesis contributions focus on several subsystem elements, e.g., optimization of cooperative multi-robot systems and multi-modal perception solutions based on intelligent sensors. The next sections, introduce the background on multi-agent systems and intelligent systems, providing the essential concepts for the proposed system architecture and the methods underlying the contributions of this work.

2.3 Multi-Agent Systems

Having discussed the characteristics of systems of systems and emphasized the core role interactions between several systems play, next we cover briefly how these can be addressed. With many interacting systems and stakeholders within SoS, the complexity in driving such systems to deliver on common functionality or achieve a shared goal becomes more involved. To address this issue, we will resort to a system modeling approach that allows handling this type of architectures, i.e., **multi-agent systems**.

Multi-agent systems comprise several interacting agents that operate in a shared environment to perform actions according to their capabilities towards a goal task. This general framework is employed in many fields, e.g., engineering, computer science, and economics [26]. In this thesis, it is particularly suited for modeling networked multi-robot systems and for distributed optimization methods discussed further along. As such, in the following, we introduce the taxonomy to refer to network topology, the paradigms for collective behavior and implications inter-agent communications play in achieving end-goals.

Sensor networks and robot teams can be modeled as multi-agent systems and described using graph theory, where sensor agents are represented as nodes, and communication links are edges establishing the information flow. Communications are critical for energy expenditure, so how and when sensors share information must be optimized to maximize network operation while not compromising the application, i.e., delivering on its capacity to achieve an intended mission goal. In that sense, where the data processing and advanced computing take place, and how data flows throughout the network play a pivotal role.

With respect to network organization, there are three important types to distinguish: centralized, decentralized and distributed [27]. As illustrated in Fig. 2.3, the elements or agents making up those networks can encompass different types and the manner in which these are organized and connected also differs.

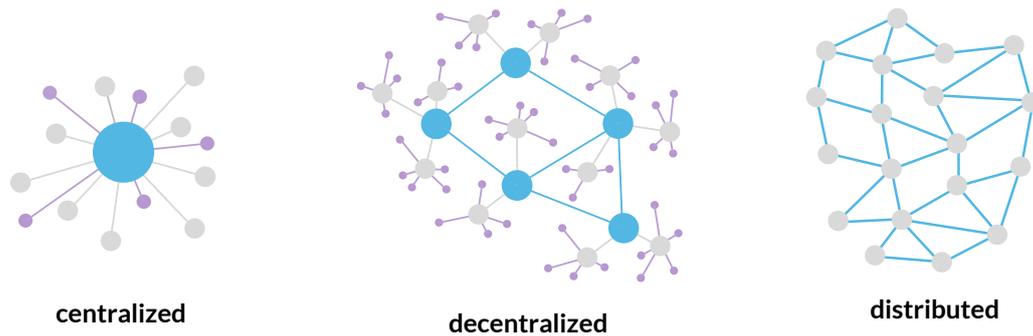


Figure 2.3: Diagram of network topologies.

In that way, the types of heterogeneous agents and the hierarchical structure of the network are instrumental to determine the data processing and communication flows within the system. Consequently, these design choices also have significant implications to features, e.g., autonomy, connectivity and redundancy.

Centralized: Organization follows a hierarchical structure, where a central unit concentrates most resources and controls lower-level elements, which have limited connectivity and autonomy.

Decentralized: Organization follows a fault-tolerant hierarchical structure where control is shared across several agents with high connectivity and autonomy, which allow enhanced awareness through efficient processing of local information.

Distributed: Organization follows a predominantly flat structure where agents with limited autonomy share equal control due to their high connectivity that ensures redundancy in information flow.

Note that decisions on network architecture must be defined by the problem requirements, and each type has advantages and disadvantages, thus being better suited for different applications. Next, we highlight a few attributes to serve as comparison axes between the different types of network architectures.

Attributes — *properties or features that influence system functionality, behavior or performance.*

Control: The ability to coordinate and manage actions of agents within the system through layers of authority and autonomy in decision-making and task execution.

Resilience: The ability to withstand failures or disruptions by adapting to and/or recovering from changes within the system or in its environment and continue to function reliably.

Scalability: The ability to sustain expansion and adaptation to meet evolving application demands, by accommodating structural changes without performance degradation.

Centralized networks benefit from a high degree of control by concentrating authority and resources in a high-level single agent. However, this comes at the expense of resilience as these architectures become more vulnerable due to their single point of failure and limited autonomy of agents at low-level layers of the system. These architectures also face more difficulty in scaling as the central agent needs to manage and control cumbersome volumes of processes and interactions.

Decentralized networks have the advantage of greater control by introducing levels of shared control through diverse decision-making levels and adding redundancy at each level. This change also supports high system resilience by the elimination of single points of failure and increased agent autonomy. In turn, these architectures can be easily scalable, but at the cost of being harder to manage due to high diversity within the system, requiring more involved coordination methods.

Distributed networks scale notably well and have the highest resilience due to its flat structure that democratizes resource allocation, eliminating critical points of failure. Though, system control decreases since the coordination and management becomes more difficult, either because of high or low degrees of autonomy of the agents involved, requiring increased foresight in accounting for emergent behavior.

Bridging this context to the problem of developing environmental monitoring solutions for fire detection and monitoring, a key application requirement justifies the better suitability of decentralized networks. Wildfires are phenomena with high spatial and temporal uncertainty so the adaptability and flexibility of sensor networks with multiple agents will largely benefit from incorporating heterogeneous capabilities. In that sense, a layered organizational structure composed of different types of monitoring solutions is preferable to deal with the demands of this application.

Given the complexity of multi-agent systems operation, the coordination of the system of systems and its elements towards a shared goal introduces new system dynamics, e.g., cooperation and collaboration. In this work both of these workflows are used to address different problems, namely through the population-based optimization methods employed and the proposed coordination strategies.

Coordination: The process of organizing elements of a system, i.e., multiple agents able to provide a desired capability, to work together in an efficient manner.

Collaboration: The process by which elements of a system, i.e., multiple agents, work together to complete a task or reach a goal that is not reachable without the confluence of complementary efforts.

Cooperation: The process by which elements of a system, i.e., multiple agents, work together to achieve a shared aim by compromising on their individual goals towards the collective global benefit.

Coordination of efforts in networked systems is essential since while networks are built to embody redundancy for resilience purposes, the design shall avoid the pitfall of over-dimensioning the system. For this reason, cooperation and collaboration are central in enabling multi-agent systems to achieve the desired flexibility and adaptability stemming from emergent behavior in response to changing environment. The next section, introduces methods from the field of intelligent systems to enable modeling such systems.

2.4 Intelligent Systems

The endeavor of building intelligent systems is by no means a recent undertaking, yet it most likely never felt so within reach as it currently appears so. Data is perceived as the new oil, and the demand for implementation of artificial intelligence embedded in complex systems is propelling fast-paced developments in this area both in academia and industry.

Intelligent systems are an essential part of the advances in autonomous robotics and optimization in networked systems, as well as in supporting decision-making. In this thesis, the objective of developing intelligent systems guides the research questions addressed and dominates the methods explored in a variety of domains, from cooperative robotics to intelligent sensors for wildfire detection.

In this context, a myriad of **computational intelligence** methods [28] are explored to address many of the challenges encountered, delving into the three branches of this field, which can be outlined as follows.

Fuzzy Systems: Models derived from mathematical representations of uncertainty based on fuzzy logic and possibilistic theory, which are able to capture the imprecision in class definitions and non-linear behavior to mimic human-like approximate reasoning, e.g., through rule-based inference.

Neural Networks: Models inspired in the human brain that are designed to leverage gradient-based, probabilistic methods to learn abstract representations from data samples by mapping patterns in relationships between inputs and outputs.

Evolutionary Computing: Mathematical representations and algorithms inspired in biological evolution that use metaheuristics for fast iteration over complex, non-convex search spaces, e.g., population-based, derivative-free optimization techniques.

In the next section, we frame concepts that are widely used in this thesis from a systems modeling perspective, covering the foundations for intelligent data-driven modeling, and the discuss important aspects regarding different flavors of model transparency and interpretability.

2.4.1 System Modeling

For modeling approaches based on fundamental laws, also known as first-principles, the theoretical descriptions are conceived from intuition and interpretation of specific behaviors. In this way, phenomena can be translated to mathematical descriptions, which inherently render the reasoning of the assumptions, being a powerful tool for communication among researchers. This approach has fueled scientific research and propelled knowledge discovery, leading to the uptake of *knowledge-based modeling*.

In engineering, modeling from first-principles is known as *white-box*, i.e. if we think of our model as fictitious box, and were to look inside, it would be transparent to interpretation, and we would recognize the causality and rational in the assumptions and mathematical equations that govern the model. This idea captures the concepts of transparency and interpretability, that will be highlighted in this work. As we proceed, analogous examples will be used to underline these characteristics in different modeling approaches.

When a system is complex in nature or it exhibits highly nonlinear behavior, our understanding can be limited or insufficient to describe its levels of intricacy, which prevents a traditional model derivation. Therefore, complex systems for which there is no prior knowledge call for different modeling techniques, for instance based on empirical data — *data-driven modeling*.

Data-driven models encompass a multitude of approaches ranging from data mining to computational intelligence techniques. In recent years, as a result of the technological advances that enabled large-scale data storage, scientific fields like data mining, knowledge discovery in databases (KDD), or artificial intelligence have been gaining a lot prominence. These fields focus on developing accuracy-oriented models, with an emphasis on computational efficiency since their main strengths are leveraged on big data analysis.

In opposition, the domain of computational discovery, stresses the importance that the knowledge emerging from computational means be communicable. However, the requirement established for the communications is the ability to follow the mathematical formalism used by scientists, e.g. in the form of equations or algorithms. The computational discovery field borrows its intentions from artificial intelligence and cognitive science, to attempt to emulate the process of discovery that could be obtained by humans in an automated way, yet its models rarely translate human reasoning.

To address this issue, fuzzy modeling approaches based on rule-based inference can provide better transparency and interpretation by translating the model to linguistic terms, which can be understood by non-experts. Moreover, fuzzy systems are flexible structures, which can incorporate laws derived from first-principles, prior expert knowledge, empirical data and heuristic rules.

Data-driven fuzzy models are regarded as relevant tool for dynamic modeling of complex systems for which there is no prior knowledge, through the identification of a set of local linear models based on fuzzy logic and approximate reasoning, expressed in the form of heuristic rules. For this reason, in a modeling sense, fuzzy inference systems can be considered a **gray-box** approach. Due to its rule-based nature, this type of systems can provide better insight into the inner-workings of the complex system under study, but also allow obtaining a better understanding of the model limitations. However, to obtain transparent and interpretable fuzzy models that can support a tractable description, the model structure has to be simple in that it follows the law of parsimony.

In light of the recent achievements in deep learning, **learning-based modeling** gained wide prominence. However, despite the high predictive ability of deep neural networks, there are many issues left undressed and open questions with this type of models. In comparison with models derived from first-principles, these models are considered as **black-box**, as they are seen as opaque. The lack of transparency and interpretability of these models hinders an adequate evaluation of model flaws and limitations. For this reason, these models might not be adequate for many real-world tasks. Furthermore, these models are computational expensive to train and run, thus recently there has been increased interest in leverage the learning acquired by deep architectures in models with reduced complexity better suited for implementation on e.g., in mobile devices.

To summarize the different facets of the intelligent systems modeling approaches explored in this thesis, the following definitions juxtapose the concepts introduced. This is followed by a brief overview of the tradeoffs related to model transparency, explainability and interpretability synthesized in Table 2.1.

White-box Models: Based on first principles such as laws from fundamental sciences, e.g., physics, chemistry and biology, or from applied fields like engineering, e.g., thermodynamics or mechanics.

Black-box Models: Based on empirical data through the derivation of mappings between inputs and outputs, without such mappings having a significant meaning, by being impossible to discern or access.

Gray-box Models: Based on expert knowledge or derived from empirical data, or a combination of both, that allows for partial visibility into model inner-workings enabling greater model insight.

Knowledge-based: Describes methods or models built upon theoretical understanding of processes or phenomena – have a pre-established range of applicability and produce deterministic solutions.

Data-driven: Describes methods or models designed to rely on databases and / or data streams for its intrinsic parameterization and operation, leveraging off-line and on-line derivation paradigms.

Learning-based: Describes methods or models that employ computational techniques to derive relations by mapping patterns in empirical data, which can be from simulated experiments or real observations, and produce in nondeterministic solutions.

By virtue of the many fields and schools of thought dealing with complex systems, different notions of the scope of transparency, explainability and interpretability exist in the literature [29]. While there are significant relations between these terms, we attempt to simplify it with the framing as follows.

In an intelligent system **transparency** enables understanding **what** the model does, **explainability** seeks to derive **how** the model operates, and **interpretability** provides insight into **why** the model works (or not).

Table 2.1: Comparison between intelligent systems modeling approaches.

	WHITE-BOX	GRAY-BOX	BLACK-BOX
FOUNDATIONS	Based on <i>first-principles</i> , e.g., fundamental laws and mathematical equations.	– Based on <i>expert knowledge</i> ; – Based on <i>empirical data</i> ; – Both modalities <i>combined</i> .	Based on <i>empirical data</i> , e.g., simulations or experiments using <i>iterative learning</i> .
TRANSPARENCY	Allows <i>complete visibility</i> and <i>knowledge</i> of model inner workings.	Allows <i>partial visibility</i> to the <i>knowledge base</i> and <i>inference</i> inner workings of the models.	Model structure is <i>opaque</i> or <i>hidden by encapsulation</i> and does not encode knowledge.
EXPLAINABILITY	All model abstractions and inner-workings can be <i>described</i> and <i>justified</i> .	Model parts and structure can be distilled to translate <i>explanations of inner workings</i> .	Model structure captures relations <i>without logical</i> or <i>intelligible meaning</i> .
INTERPRETABILITY	Allows a <i>direct translation</i> of model inner workings and <i>clearly identifying</i> its <i>limits</i> .	Allows understanding the <i>reasoning of patterns modeled</i> and to assess model <i>limitations</i> .	Insight is strictly based on experimentation, and <i>limitations are hard to scope</i> .

As intelligent systems rapidly permeate into many areas of society, e.g. in healthcare, finance, or civil protection, there is an increased need for accountability concerning the manner and extent intelligent systems are used. In that sense, the understanding of model transparency, explainability and interpretability is key towards the deployment of these technologies in safety-critical scenarios such as wildfire response.

2.4.2 Data Modeling

Data processing is a crucial part of modeling based on empirical data and is a cornerstone element of intelligent systems techniques. Data modeling approaches allow harnessing data in different ways as to leverage distinct capabilities between model architectures, e.g., neural networks versus fuzzy models. To take advantage of multiple modeling approaches while maximizing model performance, this thesis leverages methods based on **raw data**, as well as techniques using **derived features**, which incorporate a layer of expert knowledge in data modeling through **feature engineering**.

Raw Data: Concerns unprocessed information stored in digital formats that can be employed in the state encountered at the source, e.g, as found in a database, or from sensor readings.

Derived Features: Concerns data resulting from preprocessing steps that enable amplifying modeling capabilities, obtained by extraction, manipulation, or selection of variables with higher significance.

Feature Engineering: A set of data modeling processes used to improve model performance by reducing the number of variables and identifying the most significant features through feature extraction and/or feature selection techniques.

Feature Extraction: Process of applying data transformations to create more informative features.

Feature Selection: Process of choosing the most relevant subset from the available variables.

On the one hand, featuring engineering techniques allow performing dimensionality reduction, which is crucial for handling high-dimensional data, e.g., multimodal image data. More specifically, in the scope of this research, because it is a hard requirement for autonomous robotics that data is processed efficiently with edge computing devices with limited resources and energy constraints. On the other hand, for creating robotic perception pipelines also benefits significantly from well engineered computer vision algorithms that account for sensor characteristics. As such the inclusion of domain knowledge about the working mechanisms through feature engineering also proves an additional advantage in this context.

Beyond data sources, e.g., images or time-series, which are essential of data-driven processing pipelines, it is also relevant to emphasize the high value of complementing metadata, when it is available or collected. It can have an instrumental role in preprocessing, but also in the whole modeling chain since it can be vital for downstream tasks, e.g., building specific functionalities, performance evaluation and quality assurance.

Metadata access is a key resource when it is collected or available because it allows retrieving information about data collection aspects, e.g., situational context, modes of operation of the sensors, sensor calibration, etc. These factors are of the utmost relevance for applications in the physical world since for such contexts data preprocessing tasks have more stringent demands, e.g., for rigorous matching and combination of data from different sources. In that sense, metadata can allow access to key parameters for incorporation of domain knowledge into data modeling and validation.

In this thesis, capturing and interpreting metadata is particularly relevant in the context of multimodal robotic perception. More specifically, in creating valuable datasets with visible and thermal range imaging data and the corresponding, synchronized pose information derived from position and inertial sensors. The research conducted in this work included these considerations end-to-end. Starting from the design and development of several sensor payloads, to feature engineering for thermal infrared data, and modeling data-driven intelligent sensors for wildfire detection and monitoring with complementary visible and thermal data.

3 | Proposed Approach

Contents

3.1	Proposed Systems Architecture	26
3.2	Thesis Objectives	28
3.3	Research Questions and Directions	28
3.4	Outline	33

The core vision and overarching purpose of this investigation center on contributing to the development of networked environmental monitoring systems for wildfire detection and monitoring that leverage cooperative sensing and autonomous robotics to provide enhanced real-time fire intelligence.

To that end, the research avenues explored in this work provide novel solutions on a variety of methods to address narrow problems, e.g., design of sensor networks, resource optimization, or fire detection. However, the scope of the proposed thesis expands beyond such particular domains. By design there are intentional underlying interconnections that stem from the broader vision of the proposed system architecture, which intrinsically merits paramount emphasis.

For that reason, this doctoral dissertation frames the proposed thesis in the systems engineering perspective under which it was envisioned, and that has guided the research questions that were formulated and addressed. In this chapter, the prime concern is to provide not only the *big picture*, but rather the *full picture* of the directions followed and the goals of the research directions explored and outline the manner in which these are framed within the structure of this document.

3.1 Proposed Systems Architecture

This thesis proposes an integrated system based on satellite and sensor networks data to assist in early detection and monitoring of fire events. The system combines static sensors on the ground and dynamic sensors onboard mobile aerial platforms e.g. drones and high-altitude balloons (HABs), that can be deployed to monitor areas of high risk, based on satellite data.

The decentralized, multimodal and dynamic nature of the system proposed enables its deployment in target regions in specific time-windows where there are forecasts of increased fire risk. The three-layer system comprises network, perception and inference modules.

The first layer corresponds to a decentralized sensor network to be deployed in wildland and wildland-urban-interface regions, that can relay real-time data relative to regions of interest. The second layer addresses the estimation and data aggregation, yielding environment perception states. The final layer employs intelligent soft-sensor approaches, identifying specific areas in the monitored region with higher risk.

The proposed system, depicted in Fig. 3.1 integrates this multi-agent sensor network, robotic perception modules equipped with advanced sensors for data gathering, whose data is then preprocessed and filtered to be used by intelligent systems that can perform fire detection and monitoring or fire risk prediction depending on the context of application. The information generated through this network can be provided to decision-making authorities who subsequently engage with mission control to deploy further aerial means and communicate with the several agents within the network.

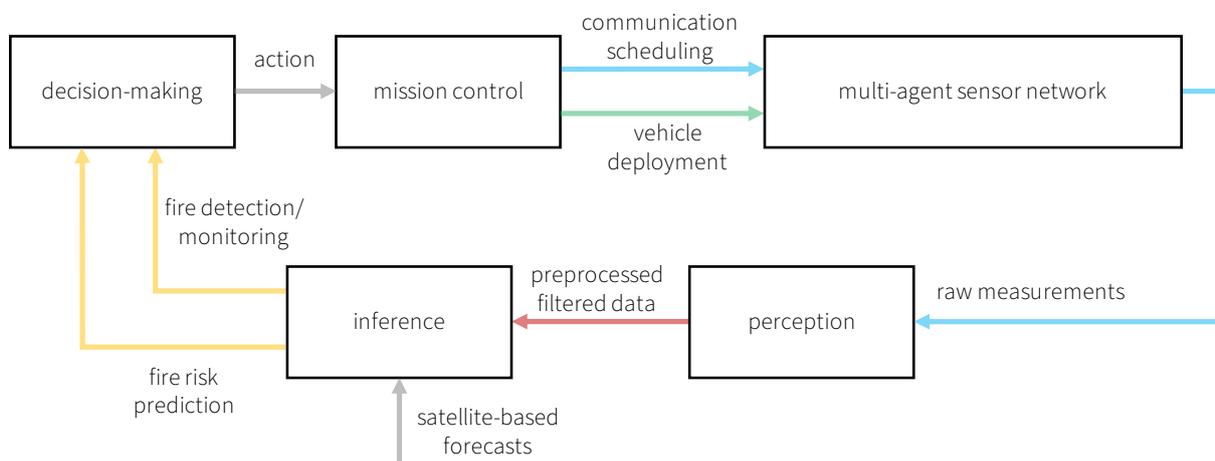


Figure 3.1: Diagram of the system architecture and data flows.

This generic architecture can accommodate a myriad of agents, namely ground-based sensors like static camera networks, heterogeneous aerial robots like multi-rotor drones or flying-wing aircraft operating at low-altitudes, high-altitude balloons acting as high-altitude pseudo-satellites in the stratosphere, as well as integrate data from satellite sources. Fig. 3.2 illustrates the realization of this concept of system of systems for wildfire detection and monitoring applications.

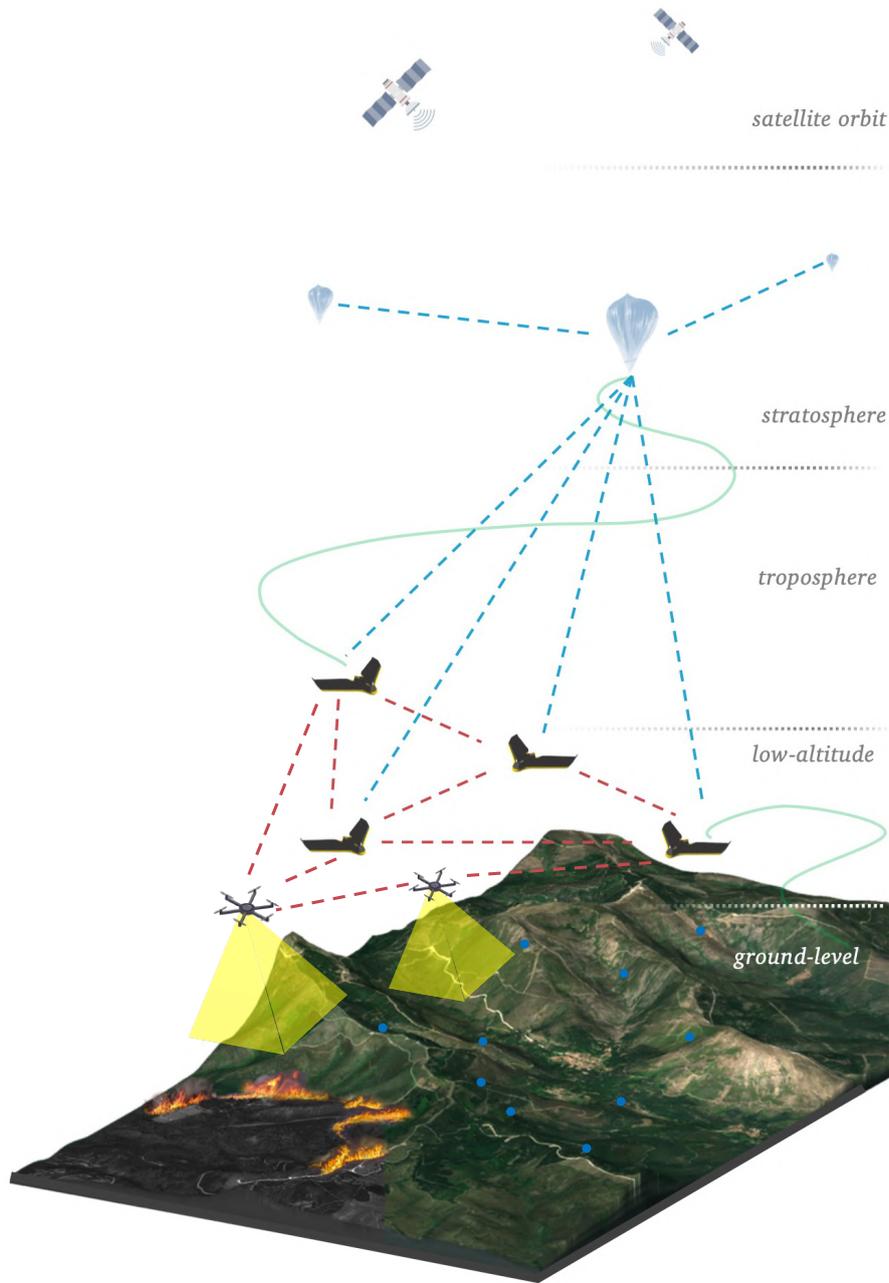


Figure 3.2: Diagram of data flows within the proposed architecture.

Leveraging HABs as high-altitude pseudo-satellites in near-space enables monitoring large-scale events and areas where fighter aircraft do not operate. In this way, this dynamic network can provide real-time Earth Observation data to decision-makers allowing its integration in existing decision support systems to generate early-warnings and optimize resource allocation.

3.2 Thesis Objectives

The principal objective of this thesis is the development of solutions for identification and monitoring of regions with higher fire probability, and for early detection of fire events towards the mitigation of the resulting impacts. Fire ignitions have a high degree of uncertainty, both on spatial and temporal levels, being difficult to pinpoint the ignition sources in a short time-frame. Furthermore, climate change increases the probability of a fire ignition leading to a fire of significant proportions. With the escalation in the initial fire spread propensity of these events, it is necessary to diminish the response time from emergency teams.

To address this issue, this work proposes an integrated model based on satellite and sensor networks data to assist in early detection and monitoring of fire events. The three-layer system with network, perception and inference modules, enables relaying real-time data to decision-makers, enabling the deployment of autonomous aerial vehicles for operational and decision support.

The envisioned system combines static and dynamic sensors, providing multiple layers of perception of the environment. Ground sensors provide data from identified fire risk regions, whereas aerial vehicles empowered with perception systems deliver extensive area coverage.

The decentralized and dynamic nature of the system enables its deployment in wildland-urban interface (WUI) areas in specific time-windows, targeting forecasts of increased fire risk. Moreover, it is intended that the information gathered from this dynamic network can be integrated in existing decision support systems to generate early warnings and optimize resource allocation.

The objective of this thesis is achieved by the following steps:

- Development of decentralized sensor network architectures for environmental monitoring;
- Propose cooperative strategies for engaging heterogeneous robotic platforms;
- Development of laboratory tests and field trials in realistic scenarios for data acquisition for multi-modal robotic perception and development of intelligent systems for fire detection and monitoring;
- Propose intelligent inference systems for fire event detection and monitoring;

3.3 Research Questions and Directions

To investigate on these multiple domains, the proposed approach is underpinned by the systems engineering philosophy. This allows harmonizing the desired integration of solutions developed, but, more importantly, to lay the grounding assumptions and guiding principles that motivate the design choices in the methodologies proposed in this thesis.

In that sense, this research explores four scientific areas, namely **systems engineering**, **networked systems**, **robotic perception** and **computational intelligence**. Each of these four vectors essentially contributing with scientific approaches to address the spatial and temporal complexity and uncertainty associated with real-time detection and monitoring of wildfire events.

To tackle this complex issue, the research directions explored address three major questions. The first concerns the design of decentralized mobile networks. The second focuses on empowering mobile robots with perception capabilities to detect and monitor wildfires. And, the third pertains to scaling data curation of image data towards harnessing learning-based techniques in fire intelligence tasks in a broader sense.

In the following, each of the research questions is formulated and the derived requirements are discussed.

Question 1: How to **devise environmental monitoring networks** that (1) can adapt according to the environment changes, and (2) provide fire intelligence in real-time?

Main goal: Develop systems that embody **high-levels of automation, flexibility** and **versatility**.

Requirements:

- Provide coverage for extensive areas with highly uncertain spatiotemporal event incidence.
- Reconciling the trade-offs between the energy constraints and data-intensive solutions.

To address this question, it is important to recall two of the foundations set forth in the introduction, i.e., the system and system-of-systems concepts. In this thesis, the proposed approaches leverage these frameworks to design the multi-agent networks, i.e., a system-of-systems, able of providing a capability to detect and monitor wildfires. Furthermore, to attain the desired adaptability, flexibility and versatility, the network design will need to reflect the intended levels of autonomy, emergence, diversity, connectivity and belonging of each element in the network itself and its constituent systems, as described in Fig. 2.2.

In that vein, this work investigates techniques to address strategic and tactical decision-making concerning dimensioning and design of aerial networks. To that end, aiming at establishing such mobile infrastructures to handle distinct demand profiles and with versatility to be useful for tasks, e.g., aerial surveillance for early fire detection and active fire monitoring. Along these lines, this research aims to develop solutions that enable deriving the configurations of heterogeneous fleets of aerial vehicles that are optimized for a predicted demand profile known *a priori*.

Moreover, attending to the energetic limitations aerial platforms have, the coverage of large areas requires the deployment of several decentralized multi-UAV fleets. For this reason, the complexity to engage such mobile infrastructures in a demand-driven paradigm calls for explicit schema for coordination of these networks. In this vein, this research studies novel methods to cooperatively coordinate multiple networks of aerial assets, distributed along several regions, while being cognizant of its flight endurance characteristics and realistic constraints concerning range of communication and on-board computing capabilities, which play an important role energy-wise.

In this thesis, Part II - Decentralized Networked Systems is devoted to presenting the solutions proposed to answer Question 1. Additionally, a concise summary highlighting its most significant contributions is presented in the next chapter in section 4.2.

In this context, to empower unmanned aerial vehicles to perform fire intelligence tasks, e.g., surveillance

and active fire monitoring with a high level of automation, these need to rely heavily on artificial intelligence breakthroughs. However, several roadblocks challenge the implementation of these solutions due to their high autonomy requirements and energy-constrained nature. In a sense, the widespread focus on artificial intelligence developments based on large models, to some extent, hinders the development of models suitable for deployment onboard these platforms. For this reason, although artificial intelligence approaches can be an enabling technology that can effectively scale real-time monitoring services and optimize emergency response resources, the design of these solutions needs to be informed by three important vectors : (i) data requirements, (ii) computing constraints and (iii) communications limitations.

Data-driven artificial intelligence is central to both handling multimodal sensor data in real-time and to annotating large amounts of data collected, which are necessary to build robust safety-critical monitoring systems. Nevertheless, these two objectives have distinct implications computation-wise because the first must happen on-board, whereas the second can leverage higher processing capabilities off-board.

In this way, this thesis argues that real-time processing and data annotation shall be tackled in a complementary manner instead of the general practice of solely targeting improvement of overall accuracy. To build wildfire intelligence at the edge, we propose developments on two tracks of solutions: (i) **deployment on the edge** and (ii) **data annotation**. The need for considerable research efforts in these two avenues stems from both having very distinct development requirements and performance evaluation metrics. Whereas the former is driven foremost by timeliness, the latter must emphasize accuracy. Thus, in this work these research problems are posed as two distinct research questions.

Question 2: How to **develop data processing pipelines** that (1) have reduced computational burden, and (2) are robust to failure modes that occur in the real-world?

Main goal: Improve **performance in deployment** conditions.

Requirements:

- Data-driven approaches have significant energy constraints.
- Processing speed is key, yet accuracy can not be overlooked at least on a coarse level.

To answer this second question, it is important to first recognize that the perception autonomy of robotic platforms essentially drives mission performance, being a primary reason for the need for edge computing of onboard sensor data. Nonetheless, in addition to this argument, the communications design is also fundamental for mission endurance as relaying large amounts of data in real-time is unfeasible energy-wise.

For these reasons, this work devotes attention to data-driven intelligent systems methods that can offer state-of-the-art performance in deployment in real-world contexts, while having a computational load amenable for the energy limitations of aerial robots. Though, for data-driven models, the methods are only half of the equation, bringing to the forefront key considerations about the data modalities used, the form in which those are processed, and the implications source data have for model robustness.

Image-based fire detection and monitoring solutions typically explore images from the visible and thermal infrared spectral range because each have its own advantages and disadvantages. This thesis aims to explore multimodal approaches to leverage the capabilities of each of these sensory modalities in a complementary manner so as to augment the scope of applicability and minimize their limitations. However, despite tremendous leaps on data-driven modeling and intelligent inference based on image data in recent years, the lack of high-quality image databases for this application is still a major barrier for harnessing learning-based approaches in this domain.

Data quality and availability are crucial for the development of robust learning-based models that can provide generalization ability to real-world contexts. Given the current gap in this area, this research invested significant time and efforts to develop cost-efficient payloads and perform extensive experimental trials both to study sensor response in fire scenarios and to collect datasets for multimodal robotic perception.

In this context, maintaining a strong emphasis on data processing with edge computing and ways to improve performance in deployment, this research undertakes an in-depth study of thermal cameras. This analysis allows a better characterization of the behavior of these sensors in real fire situations and prevents misconceptions that could hamper their use in robotic perception pipelines. Furthermore, with the growing interest in leveraging aerial robotics for wildfire support operations, one desired result of this work centers on the open release of multimodal datasets for this application.

Concerning the development of data-driven approaches for intelligent inference without significant energy constraints two routes are taken, one targeting visual range data and another focusing on thermal image data. This choice is justified by the fact that visual and thermal imagery have very distinct characteristics and consequently the state-of-the-art methods for handling those data are also rightfully different.

Given that learning-based models depend to a great degree of the quality of data available for deriving such models, the methodologies this thesis explores bears in mind the aforementioned data limitations. For visual range data, this work investigates deep neural network architectures proposing a transfer learning approach leveraging data augmentation techniques. In turn, for thermal imaging data, this research proposes a novel approach using data-driven fuzzy models based on feature engineering.

Towards the goal of improving performance in deployment, these techniques are developed and tested on a wide array of scenarios from real contexts, e.g., fire events, experimental burnings, as well as scenarios without fire sources that are commonly identified as potential causes of misclassifications. Thus, allowing for better generalization and robustness to potential failure modes in the real world. Moreover, an important asset of the methods concerns inference speed, which enables application to real-time data streams, namely onboard mobile platforms using low-cost, commercially-of-the-shelf edge computing devices.

In this document, we address Question 2 in Part III - Multimodal Robotic Perception and Part IV - Intelligent Fire Detection and Monitoring. Further along in Chapter 4 - Contributions, the advances on both dimensions are briefly discussed in sections 4.3 and 4.4, respectively.

While for deployment the development architectures need to compromise on robustness and architectural

parsimony in order to be efficient for edge processing, for data curation the goals and requirements change considerably. In this scope, improving data annotation capacity is essential to build high quality databases that can provide better source data for machine learning solutions. From this perspective, for the third research question, this thesis shifts focus towards how intelligent systems can be an instrumental tool for automating data curation processes.

Question 3: How to **build data curation pipelines** that (1) enable creating large-scale datasets, and (2) include relevant annotations?

Main goal: Improve large-scale **data annotation**.

Requirements:

- Data-driven approaches without significant energy constraints.
- Can prioritize fine-grained performance over processing speed.

Before delving into the third question, an important consideration concerns the complexity in reconciling the need of curating ever-increasing amounts of data while ensuring that the resulting databases have relevant data and annotations for their use in subsequent applications. Artificial intelligence can play a pivotal role in streamlining these procedures, however, the strengths and limitations of black-box and gray-box methods shall be weighed when developing data processing pipelines in this domain.

Despite the renewed interest in data-centric approaches to AI-based data annotation tasks, in many use-cases the incorporation of domain knowledge is not as relevant as in applied sciences and engineering, or in safety-critical applications, e.g., environmental monitoring, maintenance inspections and healthcare. In that sense, black-box approaches present several drawbacks due to their lack of transparency and interpretability, limiting the scope of their application and imposing stricter requirements on data quality.

In this context, this research investigates how expert-in-the-loop approaches can be explored for data annotation in wildfire management tasks, by combining the benefits of semi-automated data annotation with relevant domain knowledge expertise. Along this line, we aim to propose data processing pipelines for fine-grained pixel-wise data annotation, reducing significantly the time that experts dedicate to classify and segment fire events in image datasets, which is extremely time consuming and onerous.

To promote a better multidisciplinary collaboration in the data curation procedure, the approaches explored in this thesis aim to leverage interpretable linguistic models that can be easily adjusted through fine-tunable rules. In this way, besides guiding appropriate data annotations tasks, experts can be involved in the feedback loops to improve algorithms and validate the quality of the resulting annotations.

This thesis addresses Question 3 specifically in Part V - Data Curation Approaches, building upon challenges and learnings identified from the preceding Part IV - Intelligent Fire Detection and Monitoring. Similarly to the previous questions, Chapter 4 - Contributions features a brief summary of the main contributions on this topic in section 4.5.

3.4 Outline

In this Part I - Introduction, the first two chapters started by framing the motivation (Chapter 1) and setting the background for the proposed approach (Chapter 2). Chapter 3 - Proposed Approach outlined the objectives this work sets out to accomplish and research questions addressed, explaining the reasoning underpinning the research directions explored. In the next chapter, Chapter 4 - Contributions, the scientific contributions on each topic are summarized and highlighted.

In a fitting way, the structure of this thesis follows a systemic approach, organizing this body of research in distinct modules as follows.

Part II - Decentralized Networked Systems centers on mobile environmental monitoring networks, focusing on design of demand-driven networks and cooperative coordination strategies for such systems.

Part III - Multimodal Robotic Perception presents the research on multimodal perception using on thermal and visual range data, resulting from real fire experiments done in laboratory and field environments.

Part IV - Intelligent Fire Detection and Monitoring focuses on learning-based intelligent systems approaches for fire detection and monitoring for thermal and visible images.

Part V - Data Curation Approaches tackles data curation approaches for wildfire intelligence in a broader sense using expert-in-the-loop approaches, and presents a proposed technique fire data annotation.

At last, Part VI - Conclusion closes with concluding remarks concerning the principal takeaways from this thesis and points to ensuing directions of future work.

To relate the structure of this document with the proposed system architecture outlined in Figure 3.1, next Table 3.1 showcases which different parts and chapters contribute to each module of the system proposed in this thesis.

Table 3.1: Outline relating the structure of the thesis with the proposed system.

THESIS PARTS	CHAPTERS	SYSTEM ARCHITECTURE
II DECENTRALIZED SENSOR NETWORKS	5 DESIGN OF UAV FLEETS 6 COORDINATION OF DECENTRALIZED MULTI-UAV FLEETS	DECISION-MAKING ; MISSION CONTROL ; MULTI-AGENT SENSOR NETWORK
III MULTIMODAL ROBOTIC PERCEPTION	7 THERMAL INFRARED SENSING 8 MAVFIRE: MULTIMODAL DATASET	PERCEPTION
IV INTELLIGENT FIRE DETECTION AND MONITORING	9 WILDFIRE DETECTION USING TRANSFER LEARNING 10 FIRE MONITORING WITH FUZZY MODELS	PERCEPTION & INFERENCE
V DATA CURATION APPROACHES	11 EXPERT-IN-THE-LOOP DATA ANNOTATION 12 COLOR-BASED LINGUISTIC FUZZY MODELS	INFERENCE

4 | Contributions

Contents

4.1	Contribution Overview	36
4.2	Decentralized Networked Systems	38
4.3	Multimodal Robotic Perception	39
4.4	Intelligent Fire Detection and Monitoring	40
4.5	Data Curation Approaches	42

This chapter presents the contributions of this thesis, by summarizing the breadth of topics explored and listing several publications that resulted from this research. Following an initial overview, each subsequent section delves deeper into the contributions of each chapter, highlighting its most significant contributions and achievements accomplished.

4.1 Contribution Overview

In line with the objectives previously outlined, this research will follow an application focused approach, targeting detection and monitoring of wildfire events. Nonetheless, the some of the solutions proposed and developed are not limited to this application, and the methodologies herein presented can be used in a wider range of robotics-based applications. This results from the systems engineering approach followed, and of the modeling abstractions and methods employed, which to a large extent are based on data-driven methods. As such it is important to highlight that pivotal features of the proposed mobile decentralized networks are its flexibility and versatility. Thus, the proposed approach could translate well to different wildfire management tasks, or to a plethora of emerging applications requiring aerial assets.

The integrated approach contributes to extend the state-of-the-art on four main topics:

- T1 Development of multimodal networks of sensors, improving performance and robustness using decentralized approaches;
- T2 Novel feature engineering approaches for wildfire detection;
- T3 Multimodal data-driven intelligent systems;
- T4 Data curation approaches for annotation of fire image data;

Towards the first topic, T1, research focused on optimization of the dimensioning and design of heterogeneous multi-UAV fleets, as well as on cooperative coordination strategies for resilient multi-UAV networks, which has led to the following works:

M. J. Sousa, A. Moutinho, M. Almeida. (2020). **Decentralized Distribution of UAV Fleets Based on Fuzzy Clustering for Demand-driven Aerial Services**. 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Glasgow, Scotland, July, 2020. doi: 10.1109/FUZZ48607.2020.9177642.

M. J. Sousa, A. Moutinho, M. Almeida. (2023). **Coordination in Decentralized Multi-UAV Fleets for Network Resilience using Adaptive Fuzzy Graphs**. [*journal article in development*]

M. J. Sousa, A. Moutinho, M. Almeida. (2022). **Leveraging High-Altitude Balloons and Mobile Robotics for Wildfire Detection and Monitoring Systems**. In Proceedings of the 25th ESA Symposium on European Rocket and Balloon Programmes and Related Research (ESA-PAC), Biarritz, France, May, 2022.

Concerning the second and third topics, T2 and T3, that are intrinsically related, extensive work is being developed on creating proper multimodal image datasets. This is an ongoing activity throughout this collaborative research, including both data collection from online open-sources and data acquisition in field measurement campaigns. Part of this work contributed to the national research project Eye in the Sky¹.

A. Moutinho, M. J. Sousa, M. Almeida et al. (2022). **Eye in the Sky - Using High-Altitude Balloons for Decision Support in Wildfire Operations**. IX International Conference on Forest Fire Research (ICFFR), Coimbra, Portugal, November, 2022. doi: 10.14195/978-989-26-2298-9_29

¹Eye in the Sky website: eyeinthesky.tecnico.ulisboa.pt

In parallel to these efforts, several iterations of versatile multimodal sensor payloads have been developed for real-time on-line edge computing that will enable autonomous robotics missions for aerial surveillance services. This line of research has generously been supported in form of hardware and seed grants, namely the NVIDIA GPU Grant and IEEE RAS-SIGHT Humanitarian Award, respectively. The current dataset includes thermal and visible range data, acquired at ground level, as well as from airborne platforms, including drones, aircraft and high-altitude balloons.

M. J. Sousa, A. Moutinho, M. Almeida. (2020). **Autonomous multimodal robotic perception system for intelligent wildfire detection and monitoring**. IROS 2020 Workshop on Humanitarian Robotics, October, 2020. [*non-archival contribution, RAS-SIGHT project*]

M. J. Sousa, A. Moutinho, M. Almeida. (2023). **MAVFire: Aerial thermal-visual-inertial-GNSS dataset for MAV-based fire applications with motion-based calibration**. [*journal article in development*]

Furthermore, also regarding topics T2 and T3, this research contributed to the state-of-the-art in multimodal wildfire detection and monitoring by extending the understanding of thermal infrared cameras in fire situations, and proposing data modeling approaches for detection and monitoring applications. Furthermore, this research developed novel data-driven intelligent sensors based on clustering-based fuzzy modeling and deep neural networks, resulting in the following journal publications.

M. J. Sousa, A. Moutinho, M. Almeida. (2019). **Classification of potential fire outbreaks: a fuzzy modeling approach based on thermal images**. Expert Systems with Applications, 129, 216–232. Elsevier. doi: 10.1016/j.eswa.2019.03.030

M. J. Sousa, A. Moutinho, M. Almeida. (2020). **Thermal Infrared Sensing for Near Real-time Data-driven Fire Detection and Monitoring Systems**. Sensors 2020, 20, 6803. doi: 10.3390/s20236803

M. J. Sousa, A. Moutinho, M. Almeida. (2020). **Wildfire detection using transfer learning on augmented datasets**. Expert Systems with Applications, 142, 112975, Elsevier. doi: 10.1016/j.eswa.2019.112975

An intended outcome of this continued effort is the public release of open-source image databases that can empower future advances in data-driven wildfire intelligence tasks. With that aim, in the scope of the fourth topic, T4, this research addressed approaches and methods to scale data annotation using machine learning and computer vision techniques with relevant expert knowledge informing data processing pipelines, as reflected in the following works.

M. J. Sousa, A. Moutinho, M. Almeida. (2020). **Expert-in-the-loop Systems Towards Safety-critical Machine Learning Technology in Wildfire Intelligence**. NeurIPS 2020 Workshop Tackling Climate Change with Machine Learning, December, 2020. [*non-archival*]

P. Messias, M. J. Sousa, A. Moutinho. (2021). **Color-based Superpixel Semantic Segmentation for Fire Data Annotation**. 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Luxembourg, Luxembourg, July, 2021. doi: 10.1109/FUZZ45933.2021.9494421

Part of this research has been presented and formally accepted as a Best Practice Use Case Proposal to integrate the work of the Focus Group on AI for Natural Disaster Management (FG-AI4NDM), organized

by three agencies from the United Nations, namely the International Telecommunication Union (ITU), the World Meteorological Organization (WMO) and United Nations Environment Programme (UNEP). Regarding more directly the third and fourth topics, T3 and T4.

M. J. Sousa, A. Moutinho, M. Almeida. (2022). **Multimodal Databases and Artificial Intelligence for Airborne Wildfire Detection and Monitoring**. Proposal presented at the 6th Meeting of the Focus Group on AI for Natural Disaster Management, Geneva, Switzerland, 7–9 June 2022.

In this context, this work has been presented in the activities of the Focus Group and related events, namely:

M. J. Sousa, A. Moutinho, and M. Almeida. (2022). **Bringing AI to the Edge: Data-driven Approaches for Real-Time Wildfire Detection and Monitoring**, AI for Good Webinar on Artificial Intelligence for Natural Disaster Management, 16th March 2022.

M. J. Sousa, A. Moutinho, and M. Almeida. (2022). **Building wildfire intelligence at the edge: bridging the gap from development to deployment**, EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-12432, doi: 10.5194/egusphere-egu22-12432.

M. J. Sousa, A. Moutinho, and M. Almeida. (2022). **Building wildfire intelligence at the edge: bridging the gap from development to deployment**, UN Environment Webinar on Modern Technologies in Combatting Disasters on a Hotter Planet, 21st September 2022.

The remaining of this chapter covers in greater detail the contributions presented in each part of this thesis, which is divided into chapters corresponding to the most relevant aforementioned publications. The next sections are structured to present a brief context of the problems tackled and highlight the key contributions and relevant interrelations to other chapters presented in this thesis.

4.2 Decentralized Networked Systems

Background

Unmanned aerial vehicles provide a flexible and versatile solution for environmental monitoring, finding a prime application in wildfire detection and monitoring operations. However, due to the nature of fire events, the demand for this type of aerial services (i.e., requests for surveillance tasks) is sparsely distributed and comprises great spatiotemporal uncertainty. For these reasons, prior to the deployment of aerial means, the dimension and configurations of multi-UAV systems need to be properly designed to build adequate capacity for the expected coverage and range requirements.

However, due to the stringent energy limitations of aerial robots, the range of operation is inherently local, which is an hindrance for fire detection and monitoring that requires operation over extensive areas. In this context, it is crucial to design mobile decentralized systems with the ability to adapt its structure in response to the dynamic changes in real-world scenarios. In that sense, this thesis argues that sharing resources between neighboring multi-robot networks may be critical for resilience to demand variations.

Contributions

To address these issues, this thesis presents a demand-driven dimensioning strategy for networks of aerial vehicles based on a hybrid clustering approach, that incorporates heterogeneous vehicle characteristics and energy-related autonomy constraints to determine the type and number of vehicles required to meet a given demand. The approach follows a fuzzy data-driven method that enables generating coverage regions with elastic boundaries that allows for cooperation between neighboring regions to manage high and low demand scenarios, providing improved resilience to demand fluctuations and ensuring a well-conditioned formulation for the optimal resource allocation problem.

In turn, this thesis introduces a coordination framework based on adaptive fuzzy graphs that leverages symbiotic cooperation between neighboring decentralized networks. The cooperation strategy leverages the uncertainty in the boundaries of each region coupled with a decentralized ant colony optimization approach. The proposed fuzzy approach allows adjusting the redundancy level to negotiate the trade-off between mission goals and infrastructure cost. Simulation experiments applied to the task assignment problem allowed evaluating the effect of redundancy for specific tasks and how inter-network trades contribute to a reduced global cost.

These contributions represent novel methods for (i) the design of aerial networks and (ii) the coordination of multiple decentralized networks operating over several regions. These developments have resulted in a paper published in the proceedings of an international conference and are leading to an original journal article currently in preparation. Both these are presented in detail in Chapter 5 and Chapter 6, respectively.

4.3 Multimodal Robotic Perception

Background

The advances in sensor hardware and embedded computing have led to the gradual miniaturization of these technologies and the concomitant decrease in equipment costs. These emerging solutions are enabling real-time processing of high-dimensional data, enabling equipping UAVs with advanced thermal and optical image sensors and powerful computing platforms for on-board data processing. With the increasing interest in leveraging autonomous mobile robots for fire detection and monitoring arises the need to design robotic perception systems that can cope with these extreme environments.

In this context, one of the aims of this thesis is to contribute to the development of autonomous UAV-based systems for fire detection and monitoring, namely in what concerns the integration of thermal cameras in robotic perception for these tasks. Furthermore, given that visual and thermal range data have complementary capabilities, a topic of focus in this thesis centers on the development of multimodal datasets to support further investigation in this area of field robotics.

Contributions

This thesis presents a deep investigation of the sensing capabilities and image processing pipelines of thermal imaging sensors for fire detection applications, and showcases extensive experimental tests carried out with multimodal sensor payloads for data collection purposes, resulting specifically in a curated dataset for UAV-based applications in this domain.

Regarding the study on thermal cameras, we start by presenting an overview of image processing algorithms used in thermal imaging regarding data compression and image enhancement, which are responsible for transforming raw data into color-encoded images. This step is crucial for having an adequate representation and perception of the fire characteristics, such as its dimension and boundaries. Secondly, we introduce a method for data-driven thermal imaging analysis designed for fire situation awareness in robotic perception that includes variables based on raw data, as well as features derived from color-encoded images.

Towards robot autonomy in these tasks, robust perception systems have to handle highly dynamic contexts and extreme values of sensor response without filtering out key data, which poses a challenge as datasets in field robotics with these characteristics are not available in the research community. For this reason, an important research angle was the development of a multimodal dataset was created from controlled burns field experiments comprising several flight profiles and relevant fire scenarios, designed to aggregate a series of sequences that presented challenges for robotics perception pipelines.

On the one hand, a significant contributions of the experiments and studies undertaken, consists of advancing the understanding of the inner-workings of thermal cameras in wildfire scenarios, which is instrumental for the development of sound data modeling approaches. On the other hand, another important contribution resulted from developing a thermal-visual-inertial-GNSS dataset for UAV-based robotics and fire applications, envisioned to support the advancement on intelligent robot autonomy in this domain.

These contributions in multimodal robotic perception resulted in a peer-reviewed research article published in the international journal *Sensors* (MDPI) and have been of a data paper currently in preparation. Both these contributions are presented in extended form in Chapter 7 and Chapter 8, respectively.

4.4 Intelligent Fire Detection and Monitoring

Background

Wildfire detection is a time-critical application as the difficulty to pinpoint ignition locations in a short time-frame often leads to the increase of fire perimeter, larger burned areas and wider fire impacts. This difficulty has motivated considerable interest in the use of image-based sensors for the development of accurate detection and monitoring solutions. However, image can be regarded as high-dimensional data and applications in the wild, a term often used to refer to outdoor, unstructured environments, involve very complex visual understanding tasks. Although intelligent systems are providing tremendous improve-

ments in performance as demonstrated in the current literature, there is still a challenging gap to translate state-of-the-art research results to reliable deployments in real-world contexts.

To address this gap, this work considers solutions to this issue on two different angles by exploring multimodal data and investigating suitable methodologies and algorithms to develop data-driven intelligent systems. To that end, this work leverages contributions from preceding and ongoing research in multimodal data acquisition in fire environments. In addition, for visual range data additional database sources are employed, some available in current scientific literature, while some datasets created specifically for the context and purposes of this research.

Contributions

To address the challenges in visual-based fire detection, this work reviewed state-of-the-art literature to identify common shortcomings of these approaches. Given the a central issue concerns the quality of the databases used for learning-based approaches, namely for developing deep neural network architectures, this work looked into alternative options. To overcome data limitations, this work proposes a transfer learning approach coupled with data augmentation techniques tested with a tenfold cross-validation scheme. This framework enables leveraging an open-source dataset with images from real fire events, which unlike video-based works offers higher variability between samples, allowing evaluating the approach in an extensive set of real scenarios. In addition, our model evaluation included a comprehensive analysis of the patterns causing misclassifications that offered key insights to guide future research in this area.

Concerning thermal imaging data, explores datasets from experimental trials conducted in a laboratory setting and at a summer festival venue under actual operation conditions. Through an in-depth and insightful data analysis of the response of thermal imaging sensors to a fire ignition, this research devised a novel feature engineering process that allows an intuitive understanding of the behavior of thermal cameras in these conditions. To deal with high-dimensional data, the feature construction method follows a statistical color-based approach, which characterizes the dynamic behavior captured in the data using three features. Subsequently, this work leverages these variables in a fuzzy modeling approach that is transparent to interpretation and enables the assessment of the patterns being modeled.

These two avenues of research introduced novel approaches for (i) improving performance evaluation of transfer learning for image-based detection of wildfires using deep neural networks (ii) data-driven fuzzy models for fire detection and monitoring based on thermal images. These contributions have been featured in two peer-reviewed articles published in the international journal *Expert Systems with Applications* (Elsevier). Both of these are presented in detail in [Chapter 9](#) and [Chapter 10](#), respectively.

4.5 Data Curation Approaches

Background

Intelligent fire detection systems are a promising application in ever-increasing demand, but the limited number and size of open datasets, and lack of annotations, hinder model development. Solving these issues requires that experts dedicate significant time to classify and segment fire events in image datasets from multiple perspectives when these exist. From this standpoint, it is inefficient to endeavor such efforts for very narrow use cases. As such, it makes sense to broaden the scope of data annotation procedures to include more contextual information. This is relevant not just for enhancing the scene understanding of particular tasks but also because it opens the applicability of the datasets for a myriad of relevant applications in wildfire management and beyond, e.g., land cover and land use, forestry, and agriculture.

In that sense, this work investigates methods to scale data annotation by means of automated computer vision and intelligent systems techniques. One central aspect for devising image processing pipelines to bridge the gap of automating the image processing concerns the inclusion of rich annotations that are relevant for the target application and advantageous for the learning algorithms and performance evaluation processes. To tackle this problem, our research focuses on devising expert-in-the-loop systems that combine the benefits of semi-automated data annotation with relevant domain knowledge expertise, particularly through the use of interpretable data-driven modeling techniques.

Contributions

This thesis proposes the development of expert-in-the-loop systems that combine the benefits of semi-automated data annotation with relevant domain knowledge expertise and outlines several areas of wildfire management where image-based tasks could be leveraged. Through this preliminary review, we contribute to establish a wider awareness on data curation processes and the requirements for creation of large-scale image databases for relevant wildfire tasks. Therefore, contributing to building foundations that can empower the application of machine learning techniques in wildfire intelligence in real scenarios.

Towards building large-scale curated datasets, we explored a data annotation method that leverages semantic segmentation based on superpixel aggregation and color features. The approach introduces interpretable linguistic models that generate pixel-wise fire segmentation and annotations, which are explainable through simple fine-tunable rules that can support subsequent annotation validation by fire domain experts. The outcomes of this approach pave the way for creating large-scale datasets that can empower future deployments of learning-based architectures in fire detection systems.

These contributions concerning data curation are two fold: (i) propose expert-in-the-loop systems for image-based applications in wildfire management, and (ii) develop a fire data annotation pipeline leveraging a novel algorithm combining color-based clustering segmentation and interpretable linguistic fuzzy models. These contributions have resulted in two peer-reviewed papers to international conferences and are presented in Part VI - Data Curation Approaches.

II | Decentralized Networked Systems

Contents

5	Design of UAV Fleets	45
6	Coordination of Decentralized Multi-UAV Fleets	61

Summary

This part centers on devising mobile environmental monitoring networks, focusing on architectural and coordination aspects of such systems. Chapter 5 proposes a novel optimization approach for dimensioning and design of heterogenous aerial fleets that combine different vehicles types to respond to a variety of realistic demand scenarios. Chapter 6 builds upon the work presented in the previous chapter, by focusing on coordination of multiple decentralized multi-UAV fleets deployed across several regions with variable demand profiles. Both of these chapters explore novel hybrid methods leveraging fuzzy modeling, clustering techniques and ant colony optimization.

5 | Design of UAV Fleets

Contents

5.1	Introduction	46
5.2	Demand-Driven Network Optimization	48
5.3	Clustering-based Graph Partitioning	52
5.4	Decentralized distribution of m-UAV fleets	56
5.5	Results	57
5.6	Conclusion	59

The material included in this chapter was previously featured in a peer-reviewed conference paper, published in the proceedings of the 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) © 2020 IEEE.

Published in:

M. J. Sousa, A. Moutinho, M. Almeida, **Decentralized Distribution of UAV Fleets Based on Fuzzy Clustering for Demand-driven Aerial Services**, in IEEE International Conference on Fuzzy Systems (Institute of Electrical and Electronics Engineers Inc., 2020), vols. 2020-July. doi: 10.1109/FUZZ48607.2020.9177642

5.1 Introduction

The pressing challenges faced in modern societies to address emergency response operations, remote environmental monitoring or on-demand delivery are opening a broad scope of novel civilian applications of unmanned aerial vehicles (UAVs) [30]. For this reason, there has been increased research interest in developing multi-UAV systems to perform autonomous missions, providing a cost-efficient solution to carry out valuable tasks. In that sense, the advances in robotics are paving the way for future aerial networked systems that comprise a resilient mobile infrastructure capable of adapting to dynamic events in real-time.

In this context, the optimization of multi-UAV systems can be casted as a combinatorial problem and several variations of classical formulations have been proposed *e.g.* *Traveling Salesman Problem* [31, 32], *Vehicle Routing Problem* [33], or *Orienteering Problem* [34]. However, although extensive research has been devoted to address this issue from a resource allocation standpoint, in which a set of fixed resources has to be assigned to tasks, these approaches formulate the problem under the assumption that the resources of the system are able to satisfy the tasks to be performed. In that sense, it is implied that the endurance of the available vehicles is sufficient to provide adequate coverage to a fixed area. In general, works do not delve into the question of how to dimension the system, which is essential to ensure the problem is not ill-conditioned.

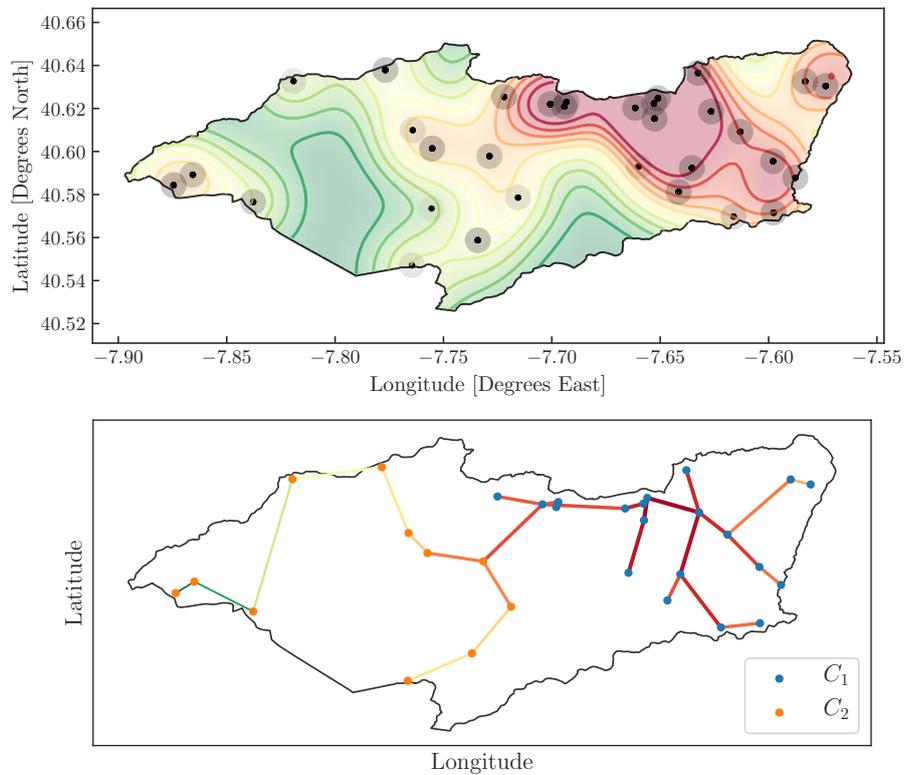


Figure 5.1: Demand-driven clustering for multi-UAV fleets.

Hence, the problem of dimensioning multi-UAV systems requires considering a wider scope, because the optimization of resources in vehicle assignment scenarios encompasses a chain of decisions on different levels and temporal scales:

- i) *strategic level*: long-term design decisions related to the dimension of the fleet, e.g., the number of vehicles in the network, the type of vehicles and desired characteristics;
- ii) *tactical level*: mid-term planning decisions regarding network configurations for different demand scenarios and deployment strategies, targeting availability and costs;
- iii) *operational level*: short-term operational decisions concerning vehicle routing, scheduling strategies and trajectory optimization for mission-oriented performance.

These aspects are highly intertwined because flight time is dependent on vehicle characteristics, as well as deployment and demand locations. These parameters drive how the system is able to respond to fluctuations in the stochastic demand, but are rarely explicitly handled in resource optimization.

To bridge this gap, this paper focuses on strategic and tactical levels, by addressing the dimensioning and design of multi-UAV systems, focusing on demand-driven network optimization based on fuzzy clustering, illustrated in Fig. 5.1.

The proposed approach is threefold: i) it builds decentralized networked systems based on cluster-based partitioning using fuzzy clustering to ensure adequate area coverage in the region of interest; ii) within each cluster, inner-clusters based on hierarchical density structure are extracted to dimension the multi-UAV fleets with adequate vehicle-types; and iii) the fleet configurations are designed to satisfy the demand.

This work contributes to the state-of-the-art by proposing fuzzy partitioning policies that incorporate heterogeneous vehicle characteristics and energy constraints into the cluster optimization process. The principal advantage of this method is that its ability to cope with data uncertainty enables generating coverage regions with elastic boundaries which allows cooperation between neighboring regions to manage high and low demand scenarios. To explore this approach, this work analyzes a case-study focused on

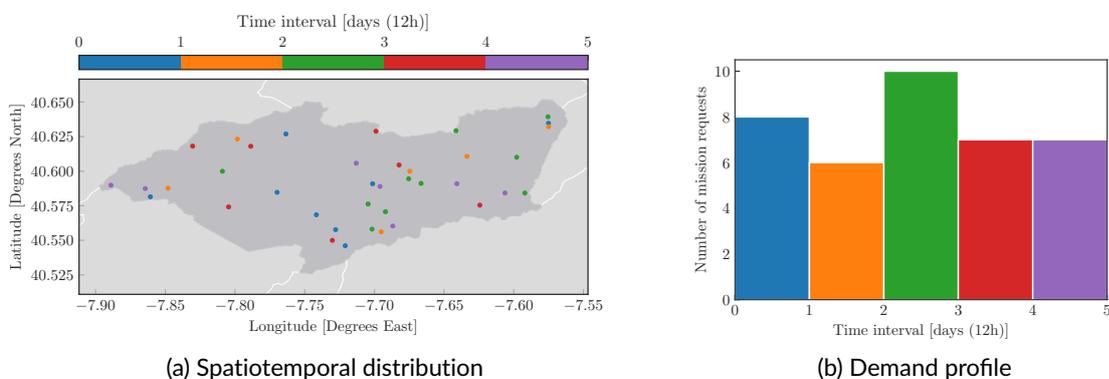


Figure 5.2: Example of demand forecasting for a 5-day time horizon: (a) distribution of aerial services according to homogeneous Poisson point process; (b) demand profile for each time interval, \mathcal{T} , considering each day with 12-hour operational time.

area coverage in wildfire detection and monitoring scenarios, namely in surveillance missions and active fire monitoring. From a resource optimization perspective, the proposed framework for dimensioning and design of UAV fleets simplifies the problem formulation stage, because by following a data-driven method it ensures a well-dimensioned system. The outcomes produce improved resilience to spatial variation in demand, as well as fluctuations in the temporal intensity of the events.

In the sequence, this work is organized as follows: Section 5.2 introduces the problem statement and demand modeling approach. Section 5.3 presents clustering-based partitioning methods. Section 5.4 focuses on the optimization of the decentralized network structure, followed by the analysis of results in Section 5.5. To conclude, Section 5.6 overviews the core takeaways and discusses possible directions of future research.

5.2 Demand-Driven Network Optimization

Networked aerial systems can enable a mobile infrastructure to support a multitude of applications involving the coverage of extensive areas, by having the capability of adapting in response to dynamic events. However, UAVs have stringent energy constraints that limit flight endurance, imposing the need to optimize the distribution of multi-UAV fleets.

Furthermore, considering missions are distributed over large areas and have a great degree of uncertainty on spatial and temporal levels, it is important to establish *a priori* demand models in order to design systems with an adequate amount of vehicles and deployment locations so as to avoid over-dimensioned or under-dimensioned systems, and promote fault-tolerance in dynamic scenarios. To that end, the following establishes the problem statement, as well as the demand modeling approach.

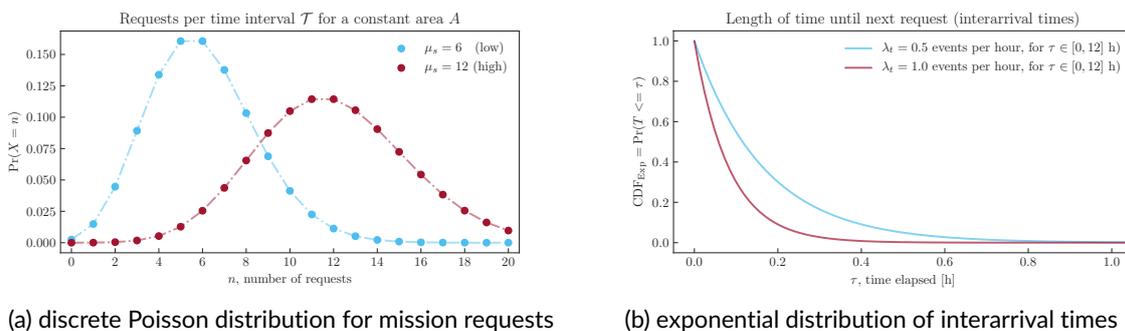


Figure 5.3: Demand modeling with Poisson point process: (a) spatial uncertainty modeled with Poisson distribution represents the variability of the number of requests occurring per time interval in a specific area; (b) temporal uncertainty concerns the variability in the time between requests, modeled with a continuous decaying exponential distribution.

5.2.1 Problem Statement

Consider a heterogeneous fleet, letting $\mathcal{V} = \{1, \dots, K\}$ represent the types of vehicles employed, and \mathcal{V}_k , define the subset of vehicles of type k . The fleet is composed of m_k vehicles of each k type, and the sum of the cardinality of each subset \mathcal{V}_k gives the total number of vehicles of the fleet. In this work, two types

of low-altitude UAVs are considered, namely multi-rotors and fixed-wing drones.

The discrete domain of aerial services tasks can be described using the mathematical formalism from graph theory, where in a graph, $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, the locations are represented by a set of nodes, \mathcal{N} , and the paths between locations are denoted by a set of edges, \mathcal{E} . To formulate the network optimization problem in a decentralized form, the set of locations \mathcal{N} is partitioned into clusters to build partial subgraphs. Let the objective function of the global problem, F_G , be defined by the cumulative sum of the local problems $F_i(x_i)$. The problem can then be stated as:

$$\min_{\mathbf{x}} F_1(x_1) + F_2(x_2) + \dots + F_c(x_c) \quad (5.1)$$

where $\mathbf{x} = [x_1, x_2, \dots, x_c]$ and c denotes the number of clusters. The core advantages of adopting a decentralized approach consist of increased fault-tolerance and flexibility, as well as reduced computational burden. The local cost functions, which incorporate density-based and energy-related components associated to the different vehicles, are formalized further along in section 5.4.

The principal goal is to optimize the type and distribution of UAVs in fleets able to perform on-demand missions that are sparsely distributed over a geographical region, as exemplified in Fig. 5.2a.

Considering that geographical administrative boundaries are highly asymmetric, an unsupervised clustering approach is a well-suited approach to derive the structure of the fleets based on probabilistic demand. To that end, the following presents the demand modeling approach for scenarios of interest.

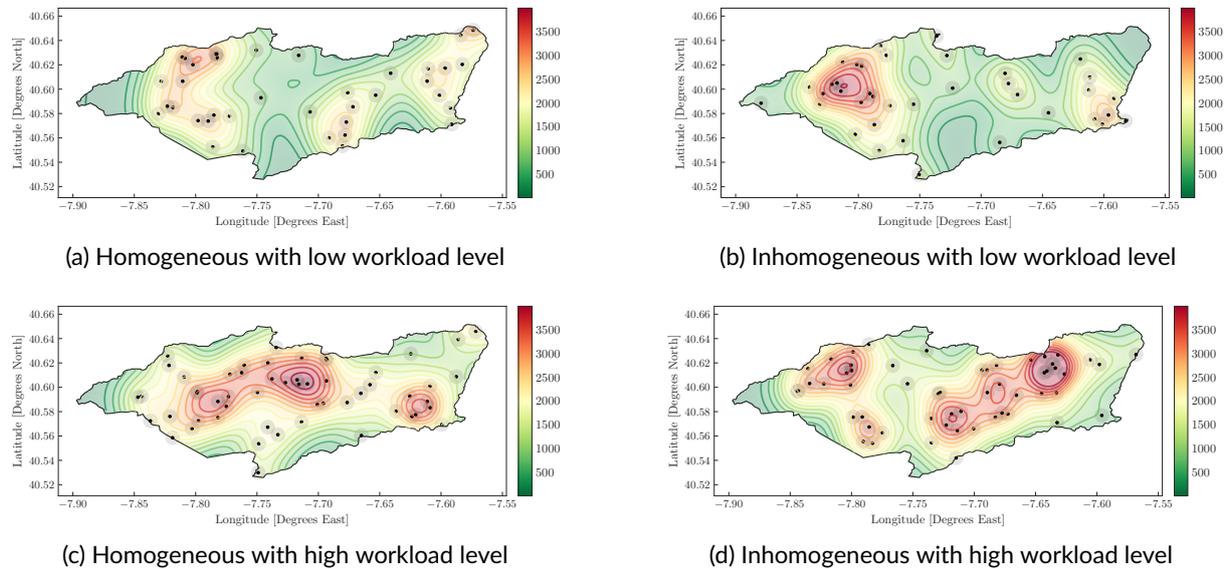


Figure 5.4: Demand distribution comparison for a 5-day period: (a) and (b) homogeneous (uniform), with low and high load, respectively; (c) and (d) nonhomogeneous, with low and high load, respectively;

5.2.2 Demand Modeling

The demand of aerial services can be modeled as a stochastic process describing a spatiotemporal pattern. For that purpose, this work employs a Poisson point process, which is widely used to model random events. Applications include *e.g.* modeling occurrences of natural hazards such as earthquakes or wildfires, on spatial and/or temporal levels, or representing the arrival of customer orders at a service provider [35].

In that sense, it is assumed that the requests, represented by location and request time, $L(x, y, t_R)$, are independent random variables, which for a given time interval, \mathcal{T} , have constant average spatial rates of occurrence for a bounded area, A , and that the average rate (requests per time period) is constant.

The spatial demand is represented according to a discrete Poisson distribution, f_{Pois} , that models the probability of a discrete number of requests, n , occurring in a time interval for a specific bounded area:

$$f_{\text{Pois}}(n; \mu_s \in \mathbb{R}^+) = \Pr(X = n) = \frac{\mu_s^n e^{-\mu_s}}{n!} \quad (5.2)$$

with the constant expected value, μ_s , depending on the spatial intensity of the demand and area size, *i.e.*, $\mu_s = \lambda_s A$. The spatial intensity, λ_s , is the expected average number of tasks per unit area. The spatial locations can be described with *e.g.*, longitude, latitude coordinates or transformed into cartesian space. In this context, the sampling of the Poisson distribution for several time intervals yields a demand profile for a *time horizon* considered in the problem, for instance as is illustrated in Fig. 5.2b, for a forecast with a 5-day time horizon.

The temporal uncertainty of the demand is described by the variability in the interval of time between consecutive requests, *i.e.* the *interarrival times*, T . In a Poisson point process these time increments are independent and identically distributed random variables that follow a continuous decaying exponential distribution. Then, the interarrival times, T , are obtained using the inverse of the cumulative distribution function of the exponential distribution, $[F_{\text{Exp}}]^{-1}$, as follows:

$$\begin{aligned} F_{\text{Exp}}(\tau; \mu_t) &= \Pr(T \leq \tau) = \int_0^\tau \mu_t e^{-\mu_t t} dt = 1 - e^{-\mu_t \tau} \\ [F_{\text{Exp}}(\tau; \mu_t)]^{-1} &= -\frac{\ln(1-\tau)}{\mu_t} = T \end{aligned} \quad (5.3)$$

with μ_t denoting the average rate per time sampling, which is a function of the temporal intensity, λ_t , and time window size, τ . The temporal intensity, describes the average number of tasks per unit time. For instance, taking the example of Fig. 5.2b, a time interval of one day, \mathcal{T} , with 8 mission requests, can have a temporal intensity of 0.5 events per hour for a time window, τ , of 12-hour period. In this way, by finding the number of requests in each day according to the Poisson distribution (5.2), for each event in that interval \mathcal{T} , the interarrival time is determined through (5.3) based on a random uniform distribution, with $\tau \sim \text{Unif}(0, 1)$. Note that the sum of interarrival times can not exceed the defined time interval, \mathcal{T} , so the values are sampled as to not exceed this upper bound.

Load Level

To establish different load levels, high and low workload scenarios are modeled by selecting distinct spatial and temporal intensities. The definition of these parameters is intrinsically related to the process under study and the time horizon considered. Herein, the process will be considered stationary, *i.e.* the average spatial and temporal intensities do not vary throughout the time-horizon (forecast window). This premise is valid assuming demand scenarios that occur in short periods in a specific region. Nevertheless, in reality, spatial and temporal intensities are expected to vary depending on the application, seasonality and geographic region.

Since aerial services are highly constrained by the limited autonomy of UAVs, the way spatial intensity relates to the area covered plays an important role in establishing decentralized multi-UAV networks. Therefore, testing different workload levels is essential for the analysis of the dimensioning problem because it drives the total number of missions that have to be performed and in result influences fleet size and the operation area of the fleet. Fig. 5.3 depicts distributions with different intensity levels used to simulate high and low workloads.

Process Type (homogeneous vs. nonhomogeneous)

The spatial structure of Poisson point processes can have distinct distributions, which for this problem influences how the mission requests are spread over the area of interest. For the homogeneous case, the spatial coordinates (x, y) are generated by a uniform distribution within the limits of the specified bounded area A , defined as a polygon or a multi-polygon. In turn, for the nonhomogeneous or inhomogeneous case, the spatial coordinates can be generated by a spatially varying deterministic intensity function $\Lambda(x, y)$, through a *thinning procedure* of a homogeneous point process of intensity λ_{\max} , where points are eliminated or retained according to a probability which depends on spatial location, $p(x, y)$ [35]. An example of the difference between homogeneous and nonhomogeneous processes is presented in Fig. 5.4 with the following intensity function for the nonhomogeneous case:

$$\Lambda(x, y) = 2(x^2 + y^2) \quad (5.4)$$

$$p(x, y) = \Lambda(x, y) / \lambda_{\max} \quad (5.5)$$

In this context, these differences allow simulating in a generic sense *e.g.* patrolling missions where a large area has to be monitored periodically (homogeneous), or active fire monitoring scenarios where mission requests are more likely to be concentrated in a particular area (nonhomogeneous). Hence, for comparison purposes the thinning process is implemented with a stop criterium to halt when the number of requests matches the load level from the Poisson distribution, enabling the evaluation of different processes with the same load level.

In this way, the parametrization of the *load level* and *process type* allows modeling the demand and generating the mission requests for simulation. Considering different vehicle types have distinct flight endurance

characteristics, the density of mission requests will impact the optimization of the configuration of multi-UAV fleets. For instance, the high maneuverability is one advantage of multi-rotor drones, however this reduces the flight endurance, thus restricting missions to a limited range. Conversely, fixed-wing drones have the benefit of harnessing the aerodynamic lift, which enables longer flights. Therefore, for short-range missions in an area with higher number of missions multi-rotors are more well-suited, whereas for performing long-range missions fixed-wing drones are a better alternative. For these reasons, the way the requests are spread has to be subsequently estimated.

Demand Density

To measure the density of requests per unit area, *i.e.* an estimate of the spatial intensity function of the point pattern, this work uses kernel density estimation (KDE) based on the convolution of isotropic Gaussian kernels [36–38].

Let $\mathbf{L} = \{\ell_1, \dots, \ell_n\}$ denote the request locations in bidimensional (2D) space, belonging to a bounded area A . The fixed-bandwidth kernel density estimate of the intensity function, *i.e.* the local intensity estimate at location p_i , is given by:

$$\hat{\lambda}(p_i) = \frac{1}{nh^2} \sum_{i=1}^n \kappa \left(\frac{p_i - \ell_i}{h} \right) e(p_i)^{-1} \quad p_i \in A \quad (5.6)$$

where κ denotes the 2D Gaussian smoothing kernel, $h > 0$ is the smoothing parameter (*i.e.* the bandwidth), and $e(p_i)$ represents an edge-correction factor [39]. Note that herein the temporal data is not considered for density estimation, as the multi-UAV fleets are to be dimensioned for the period of the time horizon, in this case based on a 5-day forecast. Recalling Fig. 5.4 comparing variable spatial distribution and load levels, the density estimation enables identifying zones with a higher number of mission requests as described by the color schema.

Albeit having a greater computational cost than alternative density estimation techniques, KDE has the benefit of considering the spatial distribution over complete neighborhoods in the region of interest. Conversely, proximity-based core measures tend to produce myopic density estimates, biased by local information. The following sections delve deeper into this issue, and outline the proposed demand-driven clustering approach designed to: *i)* build decentralized multi-UAV networks and *ii)* design the configurations of multi-UAV fleets.

5.3 Clustering-based Graph Partitioning

The problem of deploying UAVs to perform surveillance or monitoring tasks over extensive areas lends itself to be easily represented by a graph. However, if the problem has a high number of aerial services to perform, the subsequent resource allocation problem will become very complex, if feasible at all. Indeed, due to the energy constraints of aerial platforms, the universe of discourse of the entire problem results in many unfeasible solutions in practice.

To handle this issue, a decentralized approach is proposed based on clustering methods, which divides the problem into multiple subgraphs, enabling solving simpler problems in parallel by limiting the size of the search space. From an optimization standpoint the main advantage is that the search of feasible solutions is more effective at a reduced computational burden. In addition, this approach increases control over fleet dimensioning and design, whilst making the decentralized system more flexible and fault-tolerant.

Clustering algorithms are generally unsupervised learning techniques that allow grouping data according to different objectives [40]. This allows dividing the problem space using characteristics intrinsic to the data. Herein, based on the demand-driven modeling approach adopted, the interest centers on centroid-based clustering and hierarchical clustering.

While centroid-based clustering methods, *e.g.* K-means[41, 42] or Fuzzy C-Means [43–45], focus on partitioning the space in a balanced volume per cluster in terms of area coverage, this approach disregards cardinality, shape and density of each cluster, *i.e.* if there are many or few aerial tasks to perform, how are these distributed and concentrated in space, respectively. Conversely, clustering based on distance-based density measures, such as DBSCAN [46, 47] or hierarchical extensions HDBSCAN [48], concentrate on extracting cluster structures without restricting the maximum cluster volume.

In the context of the problem, considering flight endurance limitations, volume-constrained partitioning is critical to ensure adequate area coverage of the region of interest. In turn, to determine the fleet configurations, proximity-based density and hierarchical information are important to select suitable vehicle types. Thus, combining both alternatives is essential, but given the spatiotemporal uncertainty in the data, a soft clustering approach is better suited to address this problem.

In that sense, this work proposes a decentralized distribution framework based on fuzzy clustering, which incorporates density and hierarchical information, that enables dimensioning and designing a flexible multi-UAV fleet system capable to adapt to stochastic demand. More specifically, the first stage consists in a fuzzy partitioning policy based on distance-based fuzzy clustering that encompasses spatial and density information, using the Gustafson-Kessel fuzzy clustering algorithm [49]. Subsequently, the second stage concerns deriving clusters within each main subgraph using HDBSCAN based on proximity-based density information, namely mutual reachability distance and hierarchical structure. The following describes the graph model and the main components of the proposed three-stage clustering algorithm, and how these relate to the proposed framework for dimensioning and design of multi-UAV fleets.

5.3.1 Graph Model

The demand dataset is defined in the LLA (Latitude, Longitude, Altitude) referential and are subsequently converted to the NED (North, East, Down) coordinate system. The demand density at each location is estimated using the KDE method at each service waypoint. Given a set of N samples, and a data vector $\mathbf{z}_k = [X, Y, Z, \hat{\lambda}]^T$, defined by the NED coordinates and KDE-based density, let $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N]$ define the dataset of demand waypoints of the aerial services to be performed. The proposed methodology employs a two-stage clustering algorithm, thus the graph model undergoes transformations throughout

the algorithm. The following definitions relate the key components in this demand-driven approach.

Distance Measures

- Mahalanobis distance is employed in the GK algorithm to allow for clusters with different shapes but identical area;
- Core distance based on the Euclidean distance to the n -th neighbor, is used to compute the mutual reachability distance (MRD), to retrieve proximity-based density estimates and hierarchical structure of the clusters.

Demand Density Estimates

- KDE density conveys the number of missions in the region of interest;
- MRD density translates the proximity of nearby missions;

Further vehicle-related aspects are presented in section 5.4.

5.3.2 Gustafson-Kessel Fuzzy Clustering

To derive fuzzy data partitions from a set of locations \mathcal{N} , the Gustafson-Kessel (GK) fuzzy clustering algorithm clusters each data point based on centroid-based distances, according to a degree of membership, μ_{ik} , forming the fuzzy partition matrix, $\mathbf{U} = [\mu_{ik}]$. This allows locations at the boundary of each clustered region to belong to more than one fuzzy set. The algorithm computes the clusters centers, \mathbf{v}_i , as:

$$\mathbf{v}_i = \frac{\sum_{k=1}^N (\mu_{ik})^m \mathbf{z}_k}{\sum_{k=1}^N (\mu_{ik})^m}, \quad i = 1, 2, \dots, C \quad (5.7)$$

defining the matrix of cluster centers $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_C]$. The overlap between clusters is given by the fuzziness parameter, $m \in [1, \infty)$, with the lower bound equal to 1 corresponding to a hard partition. The number of clusters, C , is defined heuristically as a function of the area to be covered, and the fuzziness parameter m , through a grid search procedure.

The GK algorithm uses an adaptive distance measure based on the Mahalanobis distance, a squared inner-norm, given by:

$$D_{ik\mathbf{A}_i}^2 = (\mathbf{z}_k - \mathbf{v}_i)^T \mathbf{A}_i (\mathbf{z}_k - \mathbf{v}_i) \quad (5.8)$$

where $\mathbf{A}_i = |\mathbf{F}_i|^{\frac{1}{n}} \mathbf{F}_i^{-1}$ is a norm-inducing matrix based on the fuzzy covariance matrix, \mathbf{F}_i , given by:

$$\mathbf{F}_i = \frac{\sum_{k=1}^N (\mu_{ik})^m (\mathbf{z}_k - \mathbf{v}_i)(\mathbf{z}_k - \mathbf{v}_i)^T}{\sum_{k=1}^N (\mu_{ik})^m} \quad (5.9)$$

The GK algorithm, in addition to \mathbf{U} and \mathbf{V} , also uses the matrices $\mathbf{A} = (\mathbf{A}_1, \dots, \mathbf{A}_C)$ as optimization variables, to allow varying the shape of each cluster while maintaining a fixed volume, ensuring each multi-UAV fleet will cover equivalent areas. Thus, the clustering criterion to minimize is given by the objective

function:

$$J(\mathbf{Z}; \mathbf{U}, \mathbf{V}, \mathbf{A}) = \sum_{i=1}^C \sum_{k=1}^N (\mu_{ik})^m D_{ik\mathbf{A}_i}^2 \quad (5.10)$$

The partition matrix \mathbf{U} is updated in an iterative process, by updating the membership degrees, computed as:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^C \left(\frac{D_{ik\mathbf{A}_i}}{D_{jk\mathbf{A}_i}} \right)^{2/(m-1)}} \quad (5.11)$$

halting if the improvement on the cost function satisfies a given tolerance, or if the maximum number of iterations is reached.

By applying this technique it is possible to partition the problem and build decentralized networks with fuzzy boundaries for increased flexibility. In the sequence, within each cluster we aim to extract the demand structure so that multi-UAV fleets are designed with suitable configurations, *i.e.* how many drones and of which type, *e.g.* fixed-wing or multi-rotor.

5.3.3 Hierarchical Density-based Clustering

Considering that multi-rotors have limited flight endurance, these are better suited for short-range missions, even if within that range there exists a high number of service requests. Conversely, fixed-wing drones can serve long-range missions more effectively. In that sense, the nature of the demand influences the selection of type of vehicles within each fleet.

To match the demand structure, *i.e.* if it is sparsely/densely distributed, to better suited fleet configurations, this work employs partly a density-based method, the HDBSCAN. The following describes the central aspects of this method, while subsequent sections focus on the algorithm proposed to address this problem. First, the problem space is transformed to represent density information, through a proximity measure termed mutual reachability distance, d_{mreach} . To that end, for each $\mathbf{z}_k \in \mathbf{Z}$ a density estimate is computed, denominated the core distance to the n -th nearest neighbor [50], which for the sake of simplicity we represent as d_{core} , and without loss of generality since n is an input parameter. With the mutual reachability distance between sample j and k given by:

$$d_{\text{mreach}}(\mathbf{z}_j, \mathbf{z}_k) = \max \{ d_{\text{core}}(\mathbf{z}_j), d_{\text{core}}(\mathbf{z}_k), d(\mathbf{z}_j, \mathbf{z}_k) \} \quad (5.12)$$

a weighted graph based on the MRD is established for \mathbf{Z} , from which a minimum spanning tree (MST) is computed [51].

To extract the cluster structure, a dendrogram is constructed from the MST of graph of the mutual reachability distance, $\mathcal{G}_R = (\mathcal{N}, \mathcal{A})$. Subsequently, based on a specified minimum cluster size the cluster hierarchy is condensed to obtain a locally aggregated version of the demand. However, the hierarchical approach alone does not suit well this problem.

5.4 Decentralized distribution of m -UAV fleets

The first stage of the proposed algorithm (Algorithm 1) consists of the partitioning of the problem space using fuzzy clustering to create decentralized networks described by subgraphs. Secondly, following a decentralized approach, the procedure to extract hierarchical structure of the inner-clusters applies to the set corresponding to each fuzzy cluster, denoted by \mathcal{Z}_i with $i = 1, 2, \dots, C$. Then, having density-based inner-clusters, the main goal is to design the fleet configurations with the appropriate characteristics to respond to demand. To that end, recall the global objective function (5.1). For each operating area, *i.e.* each fuzzy cluster, the local cost, F_i is evaluated by:

$$\min \sum_{k \in \mathcal{V}} \sum_{m \in \mathcal{V}_k} \sum_{(i,j) \in \mathcal{A}} d_{ij} x_{ij}^{k,m} + \sum_{k \in \mathcal{V}} \sum_{m \in \mathcal{V}_k} e_k v_{k,m} \quad (5.13a)$$

$$\text{s.t. } x_{ij}^{k,m} \in \{0, 1\}, \forall (i,j) \in \mathcal{A}, \forall k \in \mathcal{V}, \forall m \in \mathcal{V}_k \quad (5.13b)$$

$$v_{k,m} \in \{0, 1\}, \forall k \in \mathcal{V}, \forall m \in \mathcal{V}_k \quad (5.13c)$$

with the binary decision variables representing the allocation of a vehicle in a inner-cluster, $x_{i,k}^{k,m}$, that is 1 if vehicle m of type k is allocated, or null otherwise, and the deployment of vehicle m of type k , $v_{k,m}$, that is 1 if a vehicle is deployed and assigned a cluster, and zero otherwise. The objective function balances two goals, namely the allocation of best suited vehicles for each demand density structure based on distance, d_{ij} , and the minimization of energy expenditure, accounting for a energy cost, e_k , for each vehicle deployed, $v_{k,m}$.

5.4.1 Cluster Optimization and Validity Conditions

Operational areas can not violate minimum service levels, which relate to the maximum vehicle range and expected demand, λ . Thus, the area of the region of interest, A , has to be divided into equivalent-sized clusters, according to the operational area ratio given by $\gamma = A \cdot \lambda / \text{max}_{\text{range}}$. Thereby, the number of clusters can be computed heuristically, as $C = \gamma \cdot m^s$, with m representing the overlap between clusters, and where s denotes a safety/redundancy parameter. This allows increasing the flexibility and fault-tolerance of the networked system, depending on the application requirements. Since $m \in [1, \infty)$ with 1 corresponding to a hard partition, for $s > 0$ the redundancy of the system is increased, because the number of clusters is overdimensioned. In turn, for $s < 0$ the system will be less flexible as with constant overlap between clusters, the number of clusters decreases, the cluster volume increases. Since the areas to be covered expand, the number of vehicles required to meet the minimum service level also increases, but the ability to share resources does not, so the system will be less capable of adapting to demand fluctuations.

5.4.2 Design of Fleet Configurations

With the minimum spanning tree of the mutual reachability distance, this estimate of local density is leveraged to determine the type of vehicle that should be allocated to that aerial service. As mentioned through-

out this work, multi-rotors (MR) are preferential for short-distance missions, whereas fixed-wing (FW) drones will be preferentially allocated to long-distance tasks. To build solutions for fleet configurations, the average of the MTS weights allows assessing if the tasks are in a dense/sparse area, which enables knowing if a MR or a FW should be selected, respectively. When vehicle maximum flight endurance is exceeded, additional aerial vehicles are added.

Therefore, with this iterative process, solutions for each cluster can be composed by a single type of vehicle, e.g. either fixed-wing or multi-rotor drones, or result in heterogeneous configurations with both types of vehicle in the same multi-UAV fleet. By optimizing the types of vehicles deployed according to the nature of the demand, and by having fuzzy overlap between clusters, the solutions will allow sharing of resources between neighboring regions to manage fluctuations over time in demand, and consequently in the fleet load level.

Algorithm 1: Demand-driven clustering for multi-UAV networks

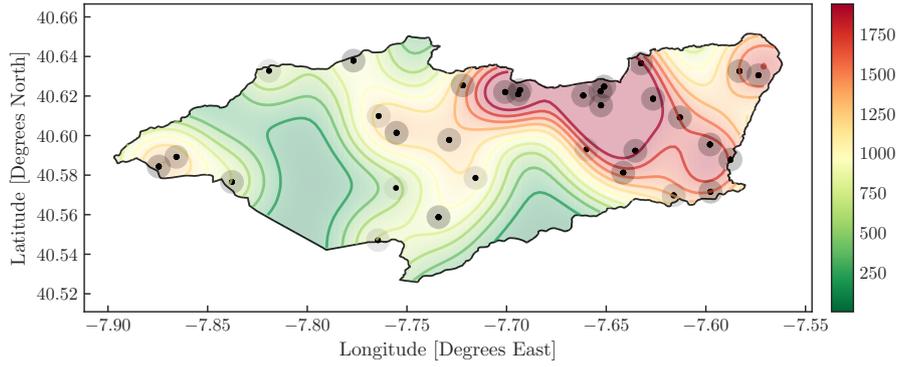
data : $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N]$, dataset with $k = 1, 2, \dots, N$ samples; \mathcal{S} , scenario parameter dictionary
 $\mathbf{z}_k = [X, Y, Z, \hat{\lambda}]^T$ in NED coord. and density estimate;
input : m , overlap degree between clusters; γ , operational area ratio; n , core distance parameter, c_{min} , minimum cluster size
output : \mathbf{Z}_C , clusters, \mathbf{v} , clusters centers, \mathbf{U} , partition matrix;

- 1 initialization
- 2 **stage 1.** build decentralized networks with GK fuzzy clustering (6.3.1)
- 3 **for** \mathbf{Z} **do**
- 4 compute fuzzy partition with C clusters and fuzziness m ;
- 5 \mathbf{v}, \mathbf{U} ,
- 6 **stage 2.** extract inner-cluster density structure with MST (5.3.3)
- 7 **for** \mathbf{Z} **do**
- 8 $d_{core} :=$ distance to the $n - th$ neighbor;
- 9 compute mutual reachability graph (5.12), and derive MST;
- 10 **stage 3.** design fleet configurations with available vehicle-types (5.4.2)
- 11 **foreach** $\mathbf{Z}_i \in \mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_C]$ **do**
- 12 build configuration solutions **if** *feasible* **then**
- 13 compute clusterCost
- 14 evaluate globalCost;

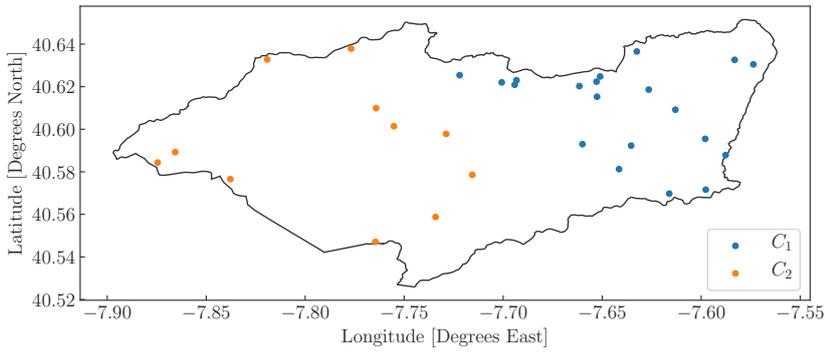
5.5 Results

This section examines a proposed case-study based on real locations from rural areas in the central region of Portugal, where there is typically increased fire hazard. The demand was modeled as outlined in section 5.2. For assessment of the proposed clustering framework, the scenarios presented in Fig. 5.4 were tested, though special attention was given to the analysis of a case resembling an active fire monitoring mission (*i.e.*, inhomogeneous with high workload). For this approach to dimensioning and design of multi-UAV fleets, the deployment/land points are not pre-established because in real contexts these can be executed by mobile operational teams, thereby assumed to be within the range of aerial services. In reality, higher demand density for aerial services in fire monitoring scenarios is mainly due to higher risks for populations in wildland-urban interfaces, or because some areas are more susceptible to phenomena of extreme fire behavior.

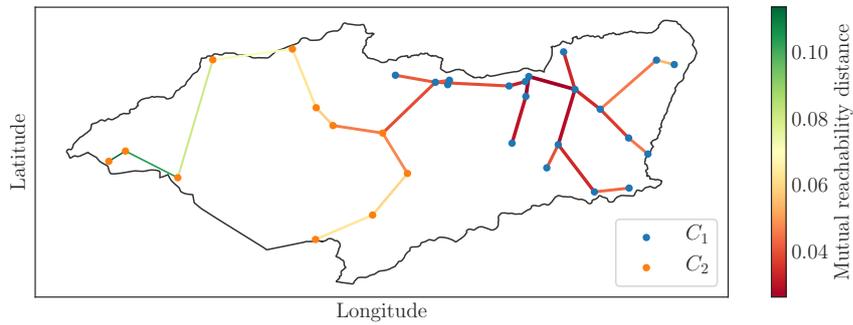
The stages of the proposed algorithm are illustrated in Fig. 5.5, for an example of an inhomogeneous high workload case. Observing Fig. 5.5c representing the GK fuzzy clusters and the global mutual reachability



(a) Inhomogeneous high workload



(b) GK fuzzy clustering



(c) GK clustering and MST of mutual reachability graph

Figure 5.5: Example of the proposed framework: (a) demand modeling simulation for fire monitoring scenario; (b) decentralized network with fuzzy partitioning.

graph, there is an higher density in C_1 in comparison to C_2 . This information is clearly valuable to select the vehicles for each fleet.

Considering fire monitoring missions, a positive safety parameter, s , is advisable, to provide the system additional flexibility and redundancy in emergency operation scenarios that are highly dynamic. With the overlap between clusters, the fleets can share resources as the situation develops. For comparison, herein the situations analyzed consider two clusters. Vehicle characteristics were defined generically with energy costs for MR and FW as 1500 and 1000, respectively. The maximum range was set for MR as 4000 and for FW as 8000.

Tables 5.1 and 5.2 present selected fleet design results, showcasing the influence of varying the degree of

Table 5.1: Optimization results of selected network models for light-load scenarios (homogeneous case).

Clusters (C)	m	Aerial Vehicles C_1	Cost C_1 (F_1)	Aerial Vehicles C_2	Cost C_2 (F_2)	Global Solution (F_G)
2	1.1	MR(1), FW(2)	20959	MR(0), FW(2)	17432	38391
	1.2	MR(1), FW(2)	20959	MR(0), FW(2)	17432	38391
	1.5	MR(0), FW(2)	17306	MR(0), FW(3)	23809	41115

Table 5.2: Optimization results of selected network models for heavy-load scenarios (nonhomogeneous case).

Clusters (C)	m	Aerial Fleet C_1	Cost C_1 (F_1)	Aerial Fleet C_2	Cost C_2 (F_2)	Global Solution (F_G)
2	1.1	MR(1), FW(4)	37872	MR(1), FW(1)	11127	48999
	1.2	MR(1), FW(3)	24554	MR(1), FW(1)	9747	34301
	1.5	MR(1), FW(2)	22794	MR(1), FW(4)	38152	60946

overlap, *i.e.* the redundancy in the system, for light and heavy load scenarios, respectively. While cluster C_1 has a low reachability level (Fig. 5.5c), it has high cardinality, thus a combination of vehicles achieves the best trade-off. In turn, the sparsity in cluster C_2 leads to fleets with mostly fixed-wing drones. Attending to the results in both scenarios, varying the degree of overlap does not evidence particular improvement in light-load cases, but can be beneficial if flexibility is desired. For heavy-load cases, the results demonstrate that increasing redundancy ($m = 1.2$) can create systems that are more efficient, but high fault-tolerance ($m = 1.5$) implies a cost increase.

Without loss of generalization, this demand-driven approach can be applied to a multitude of aerial services, by establishing different parameters according to the application requirements.

5.6 Conclusion

This paper proposes a framework for decentralized distribution of multi-UAV fleets for on-demand aerial services, addressing strategic and tactical decision-making related to dimensioning and design of aerial networks. The proposed clustering-based approach is threefold: *i*) it derives decentralized networked systems using fuzzy clustering to ensure adequate area coverage in the region of interest; *ii*) within each cluster, inner-clusters based on hierarchical density structure are extracted to dimension the multi-UAV systems with adequate vehicle-types; *iii*) fleet configurations are designed with the required number of UAVs to meet the demand. The overarching benefit of handling data uncertainty using a soft clustering approach is that it results in increasing the flexibility and fault-tolerance of the networked system.

Following this dimensioning and design strategy, to delve into actual operationalization of fleet management, multi-UAV mission planning can be framed into well-conditioned resource allocation and scheduling problems. The proposed method with fuzzy boundaries opens opportunities to research cooperative strategies, *e.g.* to balance workload between neighboring regions. Future formulations of the mission planning of multi-UAV fleets shall also evolve to include different priority levels and time-windows constraints, to approximate the study scenarios to decision-making problems faced in real contexts.

6 | Coordination of Decentralized Multi-UAV Fleets

Contents

6.1	Introduction	62
6.2	Networked Aerial Systems	64
6.3	Coordination of Multi-UAV Fleets	68
6.4	Experiments	72
6.5	Results	74
6.6	Conclusion	76

The material included in this chapter is part of a work in preparation to be submitted as a research article to an international peer-reviewed journal.

In Preparation:

M. J. Sousa, A. Moutinho, M. Almeida, **Coordination of Decentralized Multi-UAV Fleets for Network Resilience using Adaptive Fuzzy Graphs.**

6.1 Introduction

The proliferation of unmanned aerial vehicles (UAVs) is driving a broad array of novel civilian applications, e.g., drone delivery, search and rescue, or environmental monitoring [30, 52]. In that sense, there is increasing research on aerial networked systems towards creating resilient mobile infrastructures, that can adapt to dynamic changes in real-time [53, 54]. In general, multi-robot systems encompass single or multi-task robots performing single-robot tasks, or advanced configurations with robot coalitions where single/multi-task robots combine efforts to perform multi-robot tasks [55]. Recently, several contributions have proposed the use of multi-UAV systems to perform collaborative or cooperative tasks [56–58].

This work focuses on a scope of applications in which UAVs have to perform on-demand missions where tasks are distributed over large areas with variable sparsity structure. Example scenarios are, e.g., wild-fire aerial surveillance, or environmental monitoring in remote areas. Due to the energy constraints of aerial vehicles, this leads to the need to have decentralized multi-UAV fleets, which may encompass heterogeneous robots according to demand requirements. In this context, this work targets tactical aspects concerning the coordination of decentralized multi-UAV fleets, by establishing a cooperative framework that devises flexible systems resilient to demand variations. More specifically, the proposed approach explores cooperation between neighboring network entities to enable the sharing resources to address specific tasks. In practice, tasks can be exchanged between different multi-UAV to optimize overall cost. The exchange strategy is realized over an implicit communication method based on ant colony optimization. This novel coordination framework contributes to the state-of-the-art by providing an integrated optimization approach that reconciles the various elements of resilience, i.e. robustness, redundancy and resourcefulness [53, 59–61]. As outlined in Fig. 6.1, this framework (i) devises robust decentralized systems composed of multi-UAV fleets, (ii) allows for cooperation through redundancy within each robot team and also neighboring teams, and (iii) is able to adapt the areas of operation of each system depending on demand variations.

Herein, the proposed method is used for coordination in the scope of the task allocation problem, but without loss of generalization, this optimization approach can be extended to related planning problems in networked multi-robot systems, e.g., task scheduling, vehicle routing and path planning.

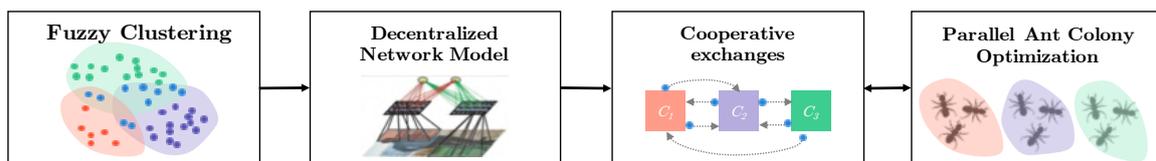


Figure 6.1: Framework for cooperation of multi-UAV systems: decoupled approach based on fuzzy clustering and ant colony optimization, that enables task exchanges between neighboring fleets allowing balancing workload, while considering vehicle motion and energy constraints in the path optimization. First, fuzzy partitioning to build decentralized subgraphs with a reasonable redundancy level. Second, cooperative mechanism allows exchanging workload requests between clusters, i.e. the multi-UAV fleets, which is optimized to derive routes and scheduling for task allocation.

6.1.1 Related Work

Task assignment problems have a wide range of applications, spanning from classical operations research domains to more recent applications in computing systems. In multi-robot systems, the development of coordination strategies can be formally casted in the domain of combinatorial optimization. These are termed multi-robotic task allocation (MRTA) problems and use several variants of the task assignment problem [62]. However, depending on problem complexity, these approaches can have prohibitive computational times to yield an optimal solution, and for NP-hard problems can not be solved in polynomial time. Several relaxation approaches have been proposed to allow solving these problems with approximation methods that provide suboptimality guarantees [63], yet the computational time associated with these algorithms can prevent its use in applications with real-time requirements. To mitigate this and improve robustness, decentralized [64, 65] and distributed [66] optimization approaches have received considerable interest, as these enable diminishing the computational burden significantly, whilst allowing enhanced fault-tolerance.

In that sense, for coordination of large-scale robotic systems, relevant research to address these issues has been developed in the realm of swarm robotics [67]. Since on-board computation and communications have to be used rationally due to the burden these have on energy consumption, centralized methods do not scale well in respect to these aspects. For this reason, swarm intelligence approaches tend to favor decentralized frameworks and bio-inspired algorithms [68, 69], in order to enable real-time realizations of coordinated collective behavior. While from an optimization standpoint metaheuristics do not provide optimality guarantees [70], these powerful algorithms cope well with nonconvex problems, by harnessing exploration and exploitation strategies to escape local minima and navigate the search space to find well-suited solutions with limited computation effort. This opens the spectra of application to problems without convexity guarantees, while proving an efficient solution for real-time deployments.

To decentralize the problem, the global set of tasks has to be divided, which can be framed with the set partitioning problem for mutually exclusive groupings or under the sensor coverage problem to embed redundancy. However, both problems are known to be NP-hard [71]. To address this issue, clustering-based partitioning methods are used in a wide variety of domains [40], and distance-based algorithms are generally well-suited for problems with spatial dependencies.

Regarding cooperative coordination methods, auction or market-based approaches have also received extensive interest [72, 73], being generally derived from classical multi-agent systems methods and decision theory. Negotiation mechanisms are especially important for cooperation of multi-task robots (MT) and in multi-robots tasks (MR) contexts.

Applications for multi-UAV surveillance tend to focus more on operational aspects and motion coordination, typically performing area partitioning using mutually exclusive areas for each robot, and designing algorithms for patrolling [74, 75], ensuring adequate coverage in the areas of interest. Conversely, as mentioned previously, in this work the focus is on a tactical level, focusing on task planning decisions for deployment of multi-UAV fleets in different demand scenarios, targeting both availability and costs;

6.1.2 Proposed Approach

The potential for deployment of multi-UAV systems in safety critical applications, e.g., wildfire surveillance, or search and rescue, demands the development of resilient robotic systems, capable to adapt in a timely manner as events unravel. In that sense, several constraints inherent to these systems and application requirements have to be considered for devising adequate coordination strategies.

To that end, this work proposes a coordination framework for task planning for multi-UAV fleets based on a decentralized optimization approach that integrates redundancy efficiently by allowing exchanges in workload requests between neighboring aerial networks. The core novel idea presented in this work concerns the dynamic exchange of workload requests between neighboring decentralized aerial networks, by allowing the exchange in task allocation for decentralized multi-UAV systems. The contributions of this work are twofold.

First, to embed redundancy in the system without scaling up infrastructure cost excessively, this work employs fuzzy partitions that cope with data uncertainty through a possibilistic approach. Based on the demand-driven method proposed in [76], that dimensions and designs adequate heterogeneous multi-UAV fleets, this work proposes leveraging fuzzy partitions to assign tasks to each local network of UAVs, whilst determining which tasks may be exchanged with neighboring fleets.

Second, to coordinate task planning across fleets, a decentralized ant colony optimization algorithm is devised to allow exchanges between neighboring networks based on implicit communication based on the degrees of membership identified with the fuzzy partitions and pheromone levels. Considering demand for aerial tasks in each area can result in distinct workloads, the ability to evaluate solutions cooperatively allows balancing workload between fleets and minimize global cost.

The investigation of the case-study inspired in real-world scenarios reveals the benefits of increasing redundancy sparingly, as well as promoting cooperation between multi-UAV fleets, demonstrating that improved resilience of the system does not have to escalate costs prohibitively, making the approach well-suited for aerial support operations.

Following this introduction, the remainder of this work is structured as follows: Section 6.2 establishes the problem formulation and specifications. Section 6.3 introduces the fuzzy partitioning schema proposed for decentralization and the cooperative strategy used for coordination. Section 6.4 presents a case-study with several variant configurations along with the performance evaluation measures. Section 6.5 explores the simulation results in detail. Section 6.6 highlights the main conclusions and discusses directions of future work.

6.2 Networked Aerial Systems

Networked aerial systems have particularly demanding energy constraints, with the need to recharge or refuel limiting the range of operation. Moreover, to adapt these systems to perform cooperative tasks,

the flight endurance of the vehicles is severely impacted by the energy drain associated with the communications between network entities. Thus, to leverage different vehicle characteristics, this work includes heterogeneous aerial vehicles, e.g., multi-rotor and fixed-wing drones, to explore distinct advantages in cooperative scenarios.

6.2.1 Problem Statement

Consider a discrete set of aerial services tasks, \mathcal{D} , sparsely distributed over a region of interest, with demand profiles modeled with spatiotemporal point patterns based on Poisson distributions [35], that may vary across different areas. Fig. 6.2 exemplifies this type of scenario, presenting the spatiotemporal distribution of aerial tasks, along with the corresponding demand profiles. It is assumed that several multi-UAV fleets are to be deployed to service these tasks, and have capacity to fulfill the nominal forecasted demand. To handle expected variations in demand, fleets distributed over neighboring areas can cooperate to satisfy mission objectives and balance the workload.

Consider a set the multi-UAV fleets, $\mathcal{C} = \{1, 2, \dots, c\}$ with a given number of fleets, c , associated with demand-based clustering of a set of requested aerial tasks \mathcal{D} . For each fleet, let the K types of vehicles employed be represented through set $\mathcal{V} = \{1, \dots, K\}$, and let set \mathcal{V}_k denote the subset of vehicles of type k . The fleet is composed of a total of m aerial vehicles, given by the sum of the cardinality of each set \mathcal{V}_k , with m_k vehicles of each k type.

The problem is framed in the class Single-Task robot, Single-Robot task Time-extended Assignment (ST-SR-TA) [55], where each robot is allocated to one task at a time, and each task requires one robot to be completed. With the number of tasks, n , exceeding the number of robots, m , this poses as a time-extended assignment, where each drone may have to attend to more than one task, resulting in essence in a scheduling problem, known to be strongly NP-hard [77]. This problem is also related to task allocation in parallel machine scheduling formulations [78].

6.2.2 Mathematical Modeling

For coordination of task planning among the fleets, the objective of the decentralized optimization problem is to minimize the overall cost function, f , given by the cumulative sum of the local cost functions, f_i , of each multi-UAV fleet, $i \in \mathcal{C} = \{1, 2, \dots, c\}$:

$$\min_{\mathbf{x}} f(\mathbf{x}) \triangleq \sum f_i(x_i) \quad (6.1)$$

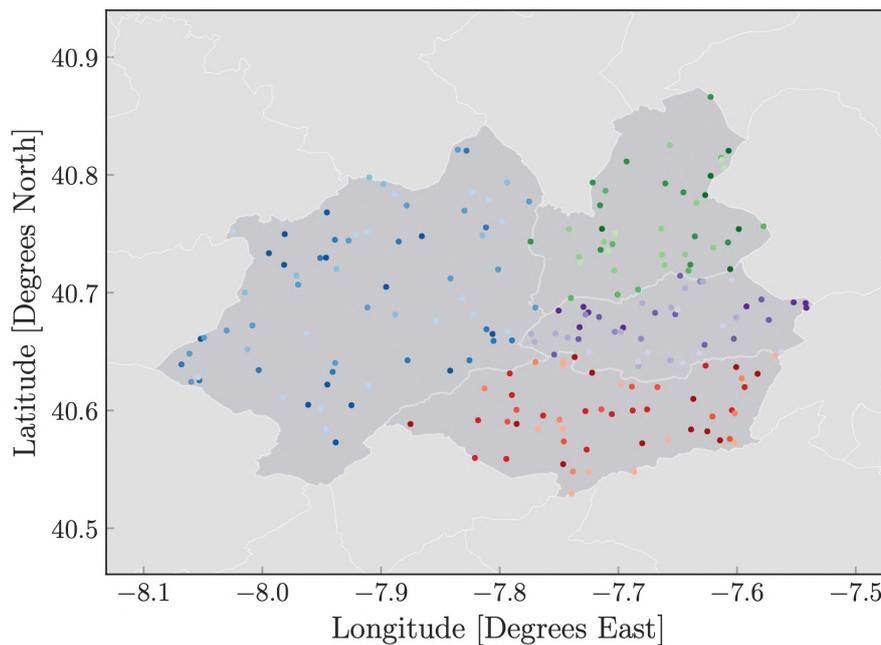
with the decision variables represented as $\mathbf{x} = [x_1, x_2, \dots, x_c]$, where c denotes the number of fleets. The cost of each fleet, f_i results from a time-based distance measure that encodes the distances that need to be travelled for each request.

To address the coordination of the task allocation problem across the decentralized system, the main objective is to enable the cooperation between neighboring fleets. In that sense, the scheduling problem

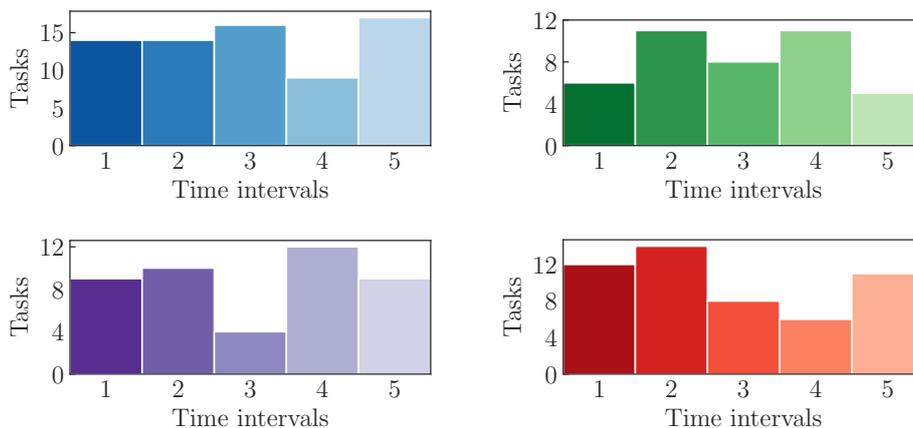
can be described by:

$$R \parallel C_{max} \tag{6.2}$$

that in this context stands for optimizing the makespan of unrelated parallel fleets, *i.e.* to minimize total operation time of the multi-UAV fleets for a given time period \mathcal{T} . This is an important goal for task allocation as responsiveness is one key feature in resilient networks. Furthermore, this objective function serves the intended cooperation strategy, as the minimization of makespan has the effect of balancing the workload across different multi-UAV fleets. In general terms, the processing time of a task can be considered the sum of the average travel time of a given vehicle to a task location and a specified task



(a) Spatiotemporal distribution



(b) Demand profiles

Figure 6.2: Example of forecasted tasks over a region of interest. Illustration of expected variations in demand: a) spatiotemporal distribution with modeling different intensity distributions depending on the area; b) demand profiles forecasted for 5 time intervals. The aerial tasks are color-coded by area and the color contrast represents the corresponding time interval the tasks arrive.

duration, which is dependent on the application.

Assuming that for each mission the UAVs can service more than one task in a sequential order, the optimization of the paths results in a routing problem. Therefore, for each fleet the problem can be approached as an in-schedule dependent (ID) allocation of class ID [ST-SR-TA] [79], that can be modeled as a multi-Vehicle Routing Problem (m -VRP) [80].

To model the discrete domain of aerial services tasks using the mathematical formalism from graph theory, a graph, $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, describes the task locations by the set of nodes, \mathcal{N} , and represents the paths between tasks through the set of arcs, \mathcal{A} . For computational simplicity, the set of nodes includes the set of locations, \mathcal{L} and the set of recharge locations, \mathcal{R} , thus $\mathcal{N} = \mathcal{L} \cup \mathcal{R}$. To consider a heterogeneous fleet, let $\mathcal{V} = \{1, \dots, K\}$ represent the types of vehicles employed, and \mathcal{V}_k , define the subset of vehicles of type k . The fleet is composed by m_k vehicles of each type, and the sum of the cardinality of each set \mathcal{V}_k gives the total number of vehicles of the fleet:

$$\sum_{k=1}^K m_k = m_k \quad (6.3)$$

Then, the static formulation of this energy-constrained heterogeneous vehicle routing problem can be stated as follows:

$$\min \sum_{k \in \mathcal{V}} \sum_{m \in \mathcal{V}_k} \sum_{(i,j) \in \mathcal{A}} d_{ij} x_{ij}^{k,m} + \sum_{k \in \mathcal{V}} \sum_{m \in \mathcal{V}_k} e_k v_{k,m} \quad (6.4a)$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{V}} \sum_{j \in \mathcal{N}} x_{ij}^k = 1, \quad \forall i \in \mathcal{L} \quad (6.4b)$$

$$\sum_{i \in \mathcal{N}} x_{ih}^k - \sum_{j \in \mathcal{N}} x_{hj}^k = 0, \quad \forall h \in \mathcal{L}, \forall k \in \mathcal{V} \quad (6.4c)$$

vehicle constraints :

$$\sum_{(i,j) \in \mathcal{A}} d_i x_{ij}^{k,m} \leq c_k, \quad \forall k \in \mathcal{V}, \forall m \in \mathcal{V}_k \quad (6.4d)$$

$$x_{ij}^k (b_i + t_{ij}^k - b_j) \leq 0, \quad \forall (i,j) \in \mathcal{A}, \forall k \in \mathcal{V} \quad (6.4e)$$

energy constraints :

$$\sum_{i \in \mathcal{N}} x_{i,r}^k = 1, \quad \forall k \in \mathcal{V}, r \in \mathcal{R} \quad (6.4f)$$

$$\sum_{k \in \mathcal{V}} \sum_{(i,j) \in \mathcal{A}} t_{ij}^k x_{ij}^k \leq 0.8 f_k \quad (6.4g)$$

decision variables :

$$x_{ij}^{k,m} \in \{0, 1\}, \quad \forall (i,j) \in \mathcal{A}, \forall k \in \mathcal{V}, \forall m \in \mathcal{V}_k \quad (6.4h)$$

$$v_{k,m} \in \{0, 1\}, \quad \forall k \in \mathcal{V}, \forall m \in \mathcal{V}_k \quad (6.4i)$$

The locations of task requests and start locations are based on the LLA (Latitude, Longitude, Altitude) referential, which are subsequently converted to the NED (North, East, Down) coordinate system. The

distance between nodes i and j , d_{ij} , that has to be traveled by the aerial vehicles, is modeled as a cartesian distance in 3-dimensional space, with respect to the NED reference frame of the position of the waypoints:

$$d_{ij} = \|p_i - p_j\| \quad (6.5)$$

The travel time, t_{ij}^k , is a function of vehicle-type and is computed as $t_{ij}^k = d_{ij}/s_k$, relating the distance, d_{ij} , to a node with the average travel speed, s_k , of the vehicle of type k . Note this naive simplification is proposed for task allocation purposes only, and is employed similarly to flight leg duration estimates common in aircraft scheduling problems [81]. The incorporation of more realistic assumptions and uncertainties should to be integrated downstream in path planning with more accurate vehicle models and constraints.

The objective function balances two objectives, namely the optimal routing of vehicles and the optimal resource allocation. This is performed by minimizing the distance travelled, d_{ij} , and the number of vehicles used through the inclusion of the energy expenditure parameter, accounting for the energy cost, e_k , for each vehicle m of type k deployed, $v_{k,m}$. To guarantee the problem of class single-task robot and single-robot task (ST-SR), constraint (6.4b) specifies that each location can only be visited once by one vehicle, irrespective of k type. In turn, (6.4c) assures that each vehicle arriving at location h , also departs from that location. The vehicle constraint in equation (6.4d) enforces that each vehicle m of type k does not exceed its capacity c_k when serving capacity demand, d_i . Additionally, (6.4e) assures that for each subroute the travel time between node i and j is accounted for. The travel time, t_{ij}^k , is a function of vehicle-type and is computed as $t_{ij}^k = d_{ij}/s_k$, relating the distance, d_{ij} , to a node with the average travel speed of the vehicle, s_k . The energy constraints are defined in equations (6.4f–6.4g). To assure each vehicle returns to a recharge location, constraint (6.4f) states that the final destination has to be a reset node, r_n in \mathcal{R} . For safety reasons, constraint (6.4g) establishes that each route of vehicles of type k can only operate until a vehicle is within 80% of its maximum flight time, f_k . Regarding the binary decision variables (6.4h–6.4i), the allocation of a vehicle to a route, $x_{ij}^{k,m}$, assumes the value of 1 if vehicle m of type k travels from node i to j , and is null otherwise. The deployment of vehicle m of type k , $v_{k,m}$, is 1 if a vehicle is deployed and assigned a route, and zero otherwise.

6.3 Coordination of Multi-UAV Fleets

The coordination of multi-UAV systems requires information exchange between vehicles, but since the energetic autonomy of these vehicles is severely constrained, cooperation algorithms have to consider communication parsimony and efficiency. To address this problem, this work proposes decentralizing the problem using a fuzzy partitioning policy based on distance-based fuzzy clustering. In a first stage, this allows organizing the network supergraph into several subgraphs with less instances, making the problem more tractable. Subsequently, the task allocation problem is addressed using a cooperative algorithm which employs decentralized ant colony optimization. To implement a cooperative strategy that allows the ability of sharing resources, an exchange mechanism which relies on implicit communication is proposed. The sequence details the methods and algorithms of the proposed coordination framework outlined in Fig. 6.1.

6.3.1 Fuzzy partitioning policy

Fuzzy decision-making approaches based on possibility theory provide flexible models that encapsulate the uncertainty that characterizes the existence of ambivalent class definitions [82]. In real scenarios, geographical boundaries are often highly irregular, so the definition of operation areas should follow a data-driven approach. To deal with data uncertainty, this work employs the Gustafson-Kessel (GK) fuzzy clustering algorithm [49], to derive fuzzy partitions from the dataset of aerial tasks \mathcal{D} . The dataset matrix of the set of locations of the aerial tasks, $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N]^T$, is composed of N samples, with each k sample, $\mathbf{z}_k = [X, Y, Z]^T$, describing the coordinates of the task locations. In contrast with hard partitions that divide data into mutually exclusive sets, in fuzzy partitions each data point belongs to a cluster with a degree of membership μ_{ik} , building the fuzzy partition matrix $\mathbf{U} = [\mu_{ik}]$. The overlap between clusters is controlled by the fuzziness parameter, $m \in [1, \infty)$, with the lower bound corresponding to a hard partition. In an iterative procedure the GK computes the distance between each sample, \mathbf{z}_k and the cluster centers, \mathbf{v}_i , through an adaptive distance measure, namely the Mahalanobis distance, a squared inner-norm, given by:

$$D_{ik\mathbf{A}_i}^2 = (\mathbf{z}_k - \mathbf{v}_i)^T \mathbf{A}_i (\mathbf{z}_k - \mathbf{v}_i) \quad (6.6)$$

where $\mathbf{A}_i = |\mathbf{F}_i|^{\frac{1}{n}} \mathbf{F}_i^{-1}$ is a norm-inducing matrix based on the fuzzy covariance matrix, \mathbf{F}_i , given by:

$$\mathbf{F}_i = \frac{\sum_{k=1}^N (\mu_{ik})^m (\mathbf{z}_k - \mathbf{v}_i)(\mathbf{z}_k - \mathbf{v}_i)^T}{\sum_{k=1}^N (\mu_{ik})^m} \quad (6.7)$$

The cluster centers, \mathbf{v}_i , which compose the matrix of cluster centers $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_C]$, are computed as:

$$\mathbf{v}_i = \frac{\sum_{k=1}^N (\mu_{ik})^m \mathbf{z}_k}{\sum_{k=1}^N (\mu_{ik})^m}, \quad i = 1, 2, \dots, c \quad (6.8)$$

with c being the maximum number of clusters, which can be determined heuristically depending on vehicle characteristics and the area to be covered. The fuzziness parameter can be determined according to the level of redundancy intended and an adequate value is typically found through a grid search procedure. The elements of partition matrix \mathbf{U} are updated iteratively based on the adaptive distance measure:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{D_{ik\mathbf{A}_i}}{D_{jk\mathbf{A}_i}} \right)^{2/(m-1)}} \quad (6.9)$$

The clustering procedure minimizes the following objective function:

$$J(\mathbf{Z}; \mathbf{U}, \mathbf{V}, \mathbf{A}) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m D_{ik\mathbf{A}_i}^2 \quad (6.10)$$

which besides using \mathbf{U} and \mathbf{V} , also includes as optimization variables matrices $\mathbf{A} = (\mathbf{A}_1, \dots, \mathbf{A}_C)$, adapting the shape of the cluster, while maintaining the volume constant. The selection of an adaptive distance measure allows each cluster to adapt to the local topological structure of the data, i.e. the sparsity nature of the demand, while fixing the cluster volume guarantees the fleets operate over equivalent areas. The

optimization procedure is stopped when the improvement of the cost function satisfies a given tolerance $||\mathbf{U}^{(k)} - \mathbf{U}^{(k-1)}|| < \epsilon$, or if a specified maximum number of iterations is reached.

6.3.2 Ant Colony Optimization

Ant Colony Optimization (ACO) is a metaheuristic inspired in the natural behavior of ants [83], which is recognized for its merits in compromising between exploration and exploitation in complex nonconvex search spaces. This population-based method relies on probabilistic selection based on a combination of knowledge and experience that enables escaping local minima and taking advantage of the fittest solutions. The ability of combining local search with dispatching rules and employing different updating schemes makes ACO a flexible framework and particularly appealing when dealing with scheduling problems [78].

In this work, ACO is applied in two stages: *i*) vehicle routing and *ii*) coordination between fleets. The algorithm developed runs in parallel in each cluster and comprises a two-step system. First, each ant travels the subnetwork to optimize the routes within each cluster, minimizing the inner-distance between the nodes. Second, each ant optimizes the allocation of resources, minimizing the system response-time, i.e. the time necessary to complete the scheduled tasks.

The proposed framework combines two pheromone updating schema, namely delayed updates in the outer-loop and immediate updates in the inner-loop. The delayed pheromone update acts as a global update at the end of each iteration and is applied with respect to the best solution found, aiming to minimize makespan. In turn, immediate updating is used for the local search procedure, and is employed to iterate through different vehicle routes, favoring the exploration of different paths within the same population.

The route optimization step has the straightforward objective of minimizing the total distance travelled. To that end, the heuristic information, η_{ij} , is encoded and defined in matrix form by:

$$\eta_{ij} = \frac{1}{t_{ij}} \quad (6.11)$$

where t_{ij} represents the euclidean distance between locations i and j according to the NED reference frame defined. The implicit communication between each population of ants is implemented through an immediate pheromone update scheme, where at the end of each iteration the values of the best global solution are updated. The pheromone matrix is updated as follows:

$$\tau_{ij}(t+1) = \tau_{ij}(t) \cdot (1 - \rho_\tau) + \Delta\tau_{ij}^k \quad (6.12)$$

where t stands for the iteration index, and ρ_τ denotes the evaporation coefficient that is applied to each trail at the end of each iteration. The pheromone reinforcement is computed as:

$$\Delta\tau_{ij}^k = \frac{Q}{D_{path}} \quad (6.13)$$

where Q is a constant parameter and D_{path} the total distance travelled in the k th ant path, with τ_{ij}^k being normalized to values in the interval $[0, 1]$.

The time-based scheduling process the heuristic information, γ_{ij}^k , is defined as a function of the traveling time between locations (i, j) for each vehicle of type k :

$$\gamma_{ij}^k = \frac{1}{t_{ij}^k} \quad (6.14)$$

where t_{ij}^k represents the traveling time between locations i and j for each vehicle of type k . Contrary to the communication scheme described previously, in the task allocation step the ants communicate with agents in the same population through a delayed pheromone update. After each ant in the population attempts a set of vehicle routes, the scheduling pheromone matrix, ϕ_{ij} , is computed as:

$$\phi_{ij}(\kappa + 1) = \phi_{ij}(\kappa) \cdot (1 - \rho_\phi) + \Delta\phi_{ij}^k \quad (6.15)$$

where κ stands for the population index, and ρ_ϕ represents the evaporation coefficient that is applied to each allocation attempt to favor the exploration of the new assignments by the other ants in the population. The pheromone deposited is computed as:

$$\Delta\phi_{ij}^k = \frac{Q}{MCT} \quad (6.16)$$

being also normalized to the range from 0 to 1. The parameter Q represents an initialization constant, while MCT is the Maximum Completion Time, i.e. the time at which the last task is completed by the vehicles of each fleet. To compute the cost function of the scheduling optimization, the makespan, C_{max} , is obtained as the maximum value of MCT across all fleets.

6.3.3 Cooperative Framework

The methods presented in the previous sections are instrumental to achieve the proposed resilient coordination approach, as these allow creating decentralized systems with a reasonable level of redundancy ingrained. To implement a cooperative strategy between neighboring multi-UAV fleets, an exchange mechanism allows the dynamic transfer of nodes between iterations and is based on communication between agents in the fleets, that relay information about the desirability of having that node in the subnetwork.

To keep a parsimonious communication regime and limiting the amount of information that has to be exchanged, the algorithm builds an incidence matrix based on the nodes that are signaled in the network as being transferable. These can be thought of as *joker* nodes that can be alienated to achieve an improvement of the local solution, as well as contributing to a better global solution.

The symbiotic cooperation strategy is currently based on the comparison of pheromone entries of fuzzy nodes in each of the subsystems. However, to truly leverage this approach in dynamic scenarios, load balancing and priority-based heuristics should be considered and if relevant embedded in the information exchange.

6.4 Experiments

To explore the flexibility of the proposed framework, the following sections delve into the experiments conducted, which are motivated by a real-world application. First, the case-study scenarios considered are presented, followed by the simulations setups used to explore the proposed algorithms. Note that the demand characteristics and VRP model parameters, e.g., dimension and configurations of the multi-UAV fleets, will be assumed *a priori*, because these can be derived using data-driven approaches. Although intrinsically related, these topics are outside of the scope herein. For further details on the demand modeling approach and demand-driven dimensioning and design of heterogeneous fleets refer to [76].

6.4.1 Case-study scenario

The following sections analyze a proposed case-study focused on wildfire surveillance scenarios. The dataset of aerial tasks is based on real locations from the central region of Portugal, comprised of rural and wildland-urban-interface areas. The dataset was generated using geographical administrative areas, and the demand was modeled using spatiotemporal Poisson point process [76]. To investigate the merits of the proposed approach a scenario with balanced demand across areas is studied, where tasks requests are generated with a homogeneous point process model and similar spatiotemporal intensity.

To illustrate this scenario in an intuitive manner, in Fig. 6.3, we employ kernel density estimation (KDE) to translate the measured spatial intensity of the demand. Fig. 6.3 exemplifies an instantiation of the case of balanced demand, where the tasks are distributed with similar intensity throughout the area of interest.

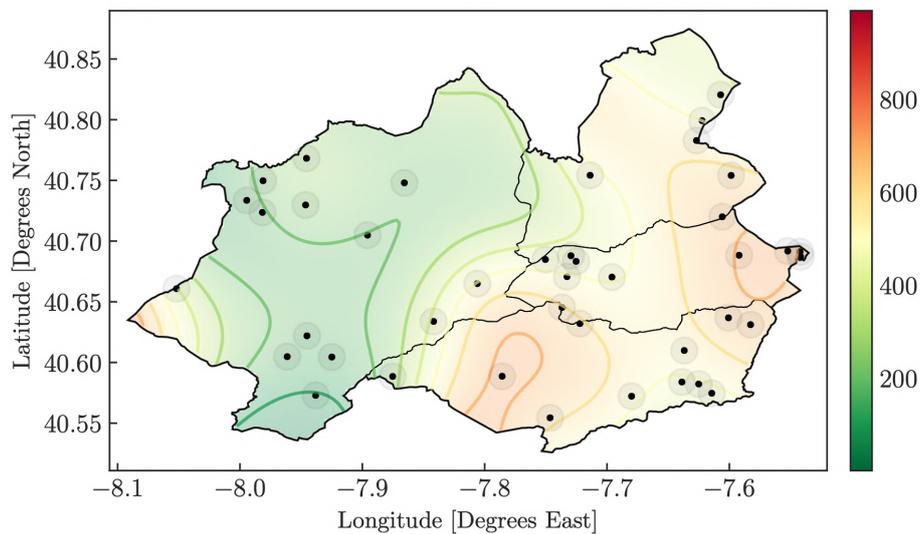


Figure 6.3: Balanced demand distribution using homogeneous (uniform) point process, with similar spatiotemporal intensity levels in each area.

The general characteristics of this case resonates with a broad range of situations encountered in real-world scenarios. However, in the context of this case-study, the main motivation is to assess possible scenarios in wildfire support operations. In this context, this case centers on situations resembling wildfire

detection missions in a prevention stage.

Being wildfire support operations both a safety- and time-critical application, leveraging redundancy can be instrumental for an adequate trade-off between infrastructure costs and mission goals. It becomes pertinent to understand the level of cooperation desirable in these distinct scenarios, and the redundancy necessary to enable such cooperative behavior.

The distance-based measures for optimization defined by the graph model were determined according to the specifications given in Section 6.2.2, and the capacity constrained is implemented as function of the maximum flight time of each platform as described in the same section.

6.4.2 Simulation Setups

The simulation setups of the experiments of the problem addressed were designed to have similar structure regarding the number of multi-UAV fleets. To that end, the number of fleets will be matched by the number of clusters, C . The main objective of the following experiments will be to investigate how variations in redundancy will contribute to the resilience of the network, with respect to its responsive and workload balancing abilities, as well as the overall costs required to satisfy mission goals. The characteristics of the heterogeneous platforms are summarized in Table 6.1.

For comparison purposes, the demand was modeled with the same number of tasks for each configuration tested. Fig. 6.3 showcases instance the resulting demand dataset for a single time interval.

Table 6.1: Model parameters for each type of vehicle k .

parameters	multi-rotor	fixed-wing
energy cost e_k	$n + 4p$	$n + 1p$
flight time f_k	0.5h	2h
travel speed s_k	0.5	2

Regarding the ACO-based algorithm, the number of colonies is defined equal to the number of fleets, and the initialization parameter was defined as: $Q = 200$. The decentralized algorithm considers one ant colony per cluster. Several values of population size were tested, leading to the definition of population of 20 ants per colony. The simulations were performed for $It_{\max} = 1000$ iteration steps, without implementation of early stopping criteria so as to analyze the convergence curves. The simulation tests were performed for a total of 200 runs for each configuration.

6.4.3 Performance Measures

To evaluate the performance of the proposed algorithm, a sensitivity analysis of parameter selection will be performed. The models will be compared based on the optimization of the routes, that aims to minimize (6.4a), as well as the optimization of the makespan, minimized through (6.2). The overarching optimization to this objective concern the performance of the global system given by the sum of local problems as described in (6.1).

Furthermore, to assess the performance of the optimization approach, the convergence of the algorithm will be analyzed as well. This will investigate if the convergence rate is sufficient for direct implementation, or if further parallelization is required on a computational level. The overall evolution of these curves will demonstrate the effect of the selection of optimization parameters in the improvement of the solutions, specifically if an adequate trade-off is achieved between exploration and exploitation of new solutions.

6.5 Results

This section explores the case-study scenario to evaluate the proposed coordination framework for task allocation in multi-UAV fleets. In the sequence, the clustering results will be analyzed for fuzzy partitions with four clusters, since it facilitates the visualization and understanding of the network model. Subsequently, the algorithm will be analyzed under balanced demand conditions, concerning the best solutions found through this approach for the configurations tested. The section concludes with the analysis of convergence curves of the algorithm and the discussion of the insights revealed.

6.5.1 Clustering-based Graph Partitioning

To test the proposed hypothesis of increasing network adaptability by having transferable nodes that can be exchanged to other neighboring subnetworks to improve the overall performance of the system, the fuzzy decentralization strategy is presented for three configurations, which introduce distinct levels of redundancy. Fig. 6.4 shows the clustering results yielded from the demand generated for adjacent areas.

Observing Fig. 6.4a, it is possible to verify that for this case a fuzzy parameter of $m = 1.2$ is equivalent to a hard partition, where each cluster has mutually exclusive sets, and share possible recharge locations, depicted in yellow. In turn, in Fig. 6.4b, by increasing the overlap between clusters with $m = 1.5$, there are four nodes depicted in blue that belong to the overlapping region, thus have a significant membership degree to more than one cluster. By increasing further the degree of overlap, the number of exchangeable tasks increases accordingly, as illustrated in Fig. 6.4c for $m = 1.8$.

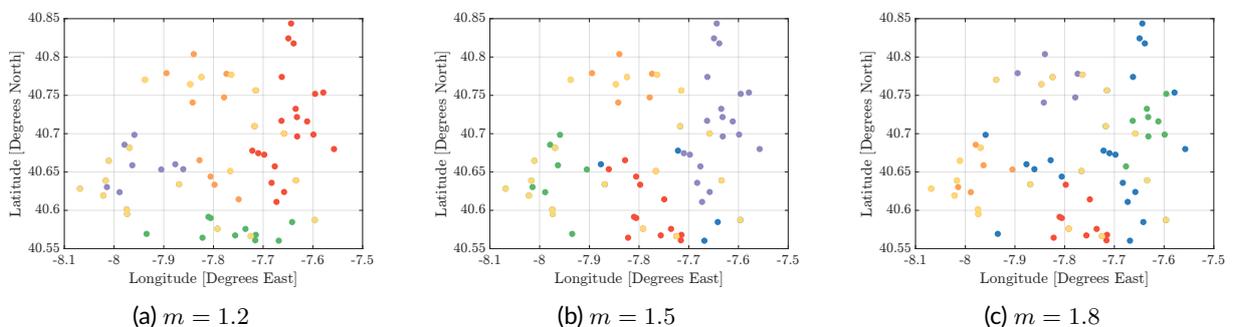


Figure 6.4: Clustering results with different overlapping according to *fuzziness* parameter m for the initial static instance: illustrated in orange, purple, red and green are the nodes of each cluster (C_1, C_2, C_3, C_4); the fuzzy nodes are depicted in blue color on (b) and (c), with a fuzzy membership degree in more that one cluster.

6.5.2 Cooperative Decentralized Ant Colony Optimization

Based on the network graphs described in the previous section, this section presents selected results for the cooperative framework. The experimental results are analyzed for the balanced demand scenario with three levels of redundancy as mentioned previously, by considering the fuzziness parameter $m \in \{1.2, 1.5, 1.8\}$. Recall the multi-UAV fleets are composed of the same number of vehicles available for deployment in all simulations, and the fleet parameters are defined in Table 6.1. For the case with balanced demand (Fig. 6.3), selected solutions obtained are presented in Table 6.2, for different network configurations.

Overall, comparing the global solutions, the graph partitions that enable cooperation present better results, e.g., $m = 1.2$, which in this case is equivalent to a hard partition because no transferable nodes were derived from the demand profile with this fuzziness level. Notably, it is also clear that increasing the level of redundancy and foster collaboration between networks is beneficial cost-wise as solutions for increased m tend to have lower global operation cost for the entire infrastructure (F_G). The makespan, C_{max} , yielded similar results for all the configurations, with the best resulting for $m = 1.5$ but at twofold the number of iterations of the other configurations tested. While this can be indicative that the time allocation can see improvements from longer optimization runs, we can observe that in all setups the algorithm converged faster to a suboptimal solution so the decision for early stopping would be well-suited in practice for implementation. Also important to note is that through increased trading of the tasks among networks, in the case where $m = 1.8$, not only the global cost is the lowest, but the cost of operation of each network is the most homogeneous across networks, revealing adequate workload balancing can be achieved.

6.5.3 Convergence Analysis

In terms of convergence rate, the behavior for the fuzzy partitions was similar, with achieving the best results within less than 400 iterations, for a population size of 20 ants. Fig. 6.5 showcases the results obtained over 10 runs for each network configuration, representing the maximum and minimum values

Table 6.2: Selected optimization results for case study with balanced demand profile.

m	Clusters (C)	Aerial Vehicles C_i	Cost $C_i (F_i)$	Global Solution (F_G)	C_{max}	Iteration
1.2	1	MR(1), FW(2)	57671	228703	2331	180
	2	MR(2), FW(2)	56630			
	3	MR(0), FW(3)	29237			
	4	MR(0), FW(3)	85164			
1.5	1	MR(1), FW(2)	38462	201062	2017	352
	2	MR(2), FW(2)	30057			
	3	MR(0), FW(3)	88449			
	4	MR(0), FW(3)	44092			
1.8	1	MR(1), FW(2)	10461	196727	2331	179
	2	MR(2), FW(2)	26973			
	3	MR(0), FW(3)	28768			
	4	MR(0), FW(3)	36373			

of global cost, F_G .

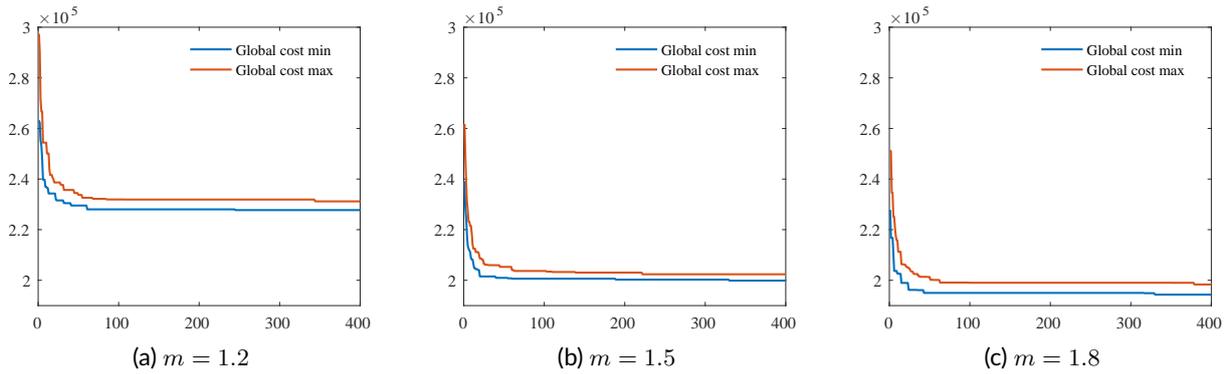


Figure 6.5: Convergence analysis for different network configurations over 10 runs.

The results obtained over a series of 10 runs for each network configuration confirm the fast convergence of the proposed algorithm and demonstrate that increasing the value of the fuzziness parameter translates in lower global costs. Subsequent research should investigate further the behavior in unbalanced scenarios.

6.6 Conclusion

This work focused on tactical aspects concerning the task planning in networks of aerial platforms, proposing a coordination framework that allows cooperation between several neighboring multi-UAV fleets. The proposed approach aims for improved resilience by: i) developing decentralized systems for increased robustness, ii) introducing redundancy sparingly and efficiently, and iii) allowing the operation areas to adapt to fluctuations in demand.

The core novelty introduced concerns the dynamic exchange of workload requests between neighboring decentralized multi-UAV fleets, through a cooperative coordination strategy based on fuzzy clustering and ant colony optimization. The main benefit consist of enabling the sharing of resources between different multi-UAV, while allowing balancing the workload across these decentralized networks to respond to variations in demand.

The case-study developed included the dataset construction from real GPS coordinates of locations in rural areas in the center region of Portugal, and several network configurations were tested, with different redundancy levels for a given fleet configuration. The results demonstrate the advantage of the proposed approach, namely by embedding increased flexibility and adaptability to the aerial resources infrastructure that translates into cost savings for the operation.

While the algorithm was presented for solving several static instances, the insights gained motivate the continuation of the development of this framework, due to the benefits this approach may provide in dynamic scenarios. Future research should consider evolving the formulation to include different priorities and time-based constraints, as well as extend the framework to address complex multi-robot task, by introducing dependency constraints.

III | Multimodal Robotic Perception Systems

Contents

7 Thermal Infrared Sensing	79
8 MAVFire: Multimodal Dataset	113

Summary

This part presents the research on multimodal perception, with an emphasis on thermal and visual range imaging solutions. In this context, several different sensing payloads were developed and tested throughout this work, along with extensive experimental work conducted in laboratory and field tests. In this thesis we select the most relevant contributions of that work. Chapter 7 presents an in-depth study devoted to thermal infrared image systems, focusing on the behavior in fire scenarios, allowing uncovering several specificities of their operation important towards their integration in robotic perception pipelines. Chapter 8 proposes a novel multimodal dataset for UAV-based robotics research and fire applications collected in field experiments that can empower the development of robust robotic navigation solutions and use-cases in fire support operations.

7 | Thermal Infrared Sensing

Contents

7.1	Introduction	80
7.2	Thermal Imaging	83
7.3	Data-driven Thermal Imaging Analysis	93
7.4	Controlled Fire Experiments	96
7.5	Results and Discussion	100
7.6	Conclusion	110

The material included in this chapter was previously featured in a peer-reviewed journal article, published in *Sensors (Switzerland)*, MDPI.

Published in:

M. J. Sousa, A. Moutinho, M. Almeida, **Thermal infrared sensing for near real-time data-driven fire detection and monitoring systems**. *Sensors (Switzerland)*. 20, 1–29 (2020).

7.1 Introduction

With the emergent effects of climate change, several regions worldwide have been undergoing an increasing number of wildfire events, more intense and devastating, and extending fire seasons [84, 85]. In this context, given the high spatial and temporal uncertainty intrinsic of these phenomena, environment monitoring is determinant for firefighting activities to mitigate the consequences of these events.

Currently, limited areas can be monitored through automatic systems in watchtowers. However, by being installed near the ground, hence at low altitudes, these systems have several limitations. Since these are widely based on visible range sensors, clouds can be easily confused with smoke and the sunset or reflections can be mistaken by flames, leading to false alarms. Moreover, as the systems depend on the identification of the smoke column, flames can only be detected when the fire has increased in magnitude. Hence, often preventing early fire detection. Furthermore, solutions based on satellite data have considerable latency, thus also hindering its application for early detection.

Albeit the widespread use of aerial means in prevention and emergency-response to forest fires, piloted aircraft require highly trained personnel and are expensive to operate [86, 87], limiting the number of vehicles that are used for firefighting and surveillance tasks. However, given that these scenarios require the coverage of extensive areas where the environment is highly dynamic, the availability of aerial means is paramount on tactical and operational levels for situation-awareness.

To address this issue, in recent years there has been active research towards developing systems based on unmanned aerial vehicles (UAVs) for fire detection [88, 89] and operational support [90–92], but it has intensified lately as a result of the difficulties faced in the response to large-scale wildfires, which has reinforced the need to detect fires in an early stage, as well as to provide near-real-time monitoring.

In parallel, the recent advances in robotic perception have opened a path towards autonomous robotics by taking advantage of novel sensors and intelligent systems to enable autonomous navigation and exploration [93]. The breakthroughs in sensor technology and embedded computing have resulted in the progressive decrease in weight/size ratio and equipment cost while incorporating powerful computing capabilities. These emerging technologies are enabling real-time processing of high-dimensional data, which allows equipping UAVs with advanced thermal and optical image sensors as well as computing platforms for on-board data processing.

In this context, the general aim of this investigation is to contribute to the development of UAV-based systems for fire detection and monitoring, namely in what concerns the integration of thermal cameras in robotic perception for these tasks.

The main objective of this study is to provide a comprehensive understanding of the behavior of thermal imaging sensors in fire detection scenarios by relating the sensor response and the image processing employed in this type of device. In that sense, this work explores sensory data from two different thermal cameras by linking the raw sensor data to the mapping functions employed to obtain images encoded in pseudocolor. For this purpose, several fire experiments in controlled conditions were performed in

laboratory and field trials covering distinct operating regimes of this type of sensors.

The contributions of this article are threefold: (1) the overview of state-of-the-art image processing algorithms widely used in thermal imaging cameras; (2) the in-depth analysis of the behavior of thermal cameras targeted at fire identification scenarios based on controlled fire experiments; (3) discussion of the implications of the insights exposed and the potential developments on robotic perception towards autonomous UAV-based fire detection and monitoring systems.

Considering the growing availability of low-cost, compact thermal cameras for UAV-borne applications, the uptake of this technology will progressively increase in the near future [94]. Being this topic an active area of research, the relevant insights outlined in this article provide important considerations to guide future research. Hence, also contributing towards the practical implementation of thermal-enabled fire detection and monitoring systems, which bring great potential to minimize the impacts of fire events.

7.1.1 Related Work

The advent of the evolution of remote sensing technologies has led to the continuous improvement of fire hazard identification and risk assessment systems. Over the years several types of platforms have been explored for these solutions such as satellites, high-altitude aircraft, remotely piloted aircraft systems, enabling the assessment of the progression of fire events [17].

Although satellite-based imagery is used by emergency-response agencies to monitor large-scale wildfires that burn over extensive periods, the wait interval for satellite overpass induces a considerable time-delay, which prevents its application in time-sensitive fire detection scenarios such as emergency evacuations or search and rescue operations [95]. Despite its value from a strategic standpoint, for tactical and operational decision support the availability of updated information is crucial. To address this while avoiding the expensive operation costs of piloted aircraft, UAVs are considered as a viable alternative for remote sensing, by providing local coverage with high spatial and temporal resolution.

In previous contributions, wildfire detection applications based on airborne systems have explored visible range, thermal, or multispectral technologies [96]. However, since the radiation emitted from a fire is high in the thermal range, there has been significant interest in the use of thermal infrared bands [17]. In that sense, the following review focuses on contributions employing the thermal range.

Thermal infrared cameras provide sensing capabilities suitable for ongoing environmental monitoring by operating in both daylight and in night conditions. Additionally, in contrast to visible range sensors for which smoke severely affects the ability to detect and track the perimeter of fire fronts, for thermal infrared cameras the impact of smoke presents lower interference.

Regarding the development of automatic algorithms, several contributions have presented advances towards this objective. On the one hand, works on image processing algorithms have focused on the extraction of image descriptors to obtain signals representing fire instances as to detect the presence of a fire by assessing the time-series data [97]. On the other hand, since thermal images adapt according

to the context in the field of view of the camera, considering only the brightness information can lead to false alarms. To address this issue, some approaches combine both spatial and temporal features, i.e. brightness, motion and flicker [98]. In turn, the integration of infrared cameras has been suggested as a way to improve visual range fire surveillance systems, as to harness the advantages of both visual and thermal features, to yield more accurate early fire detection rates [99–101]. More recently, an off-line processing fire-tracking algorithm based on edge-detection has been proposed to process georeferenced thermal images previously acquired using an airborne thermal camera [102].

Note, however, that not all the research efforts previously mentioned focus specifically or solely on wildfire detection. In a wider scope, considering fire detection outdoors, authors have recognized the difficulty in applying image processing algorithms in this setting due to a significant rate of false positives, caused by external factors such as weather conditions, sunlight reflections, or saturation of the infrared sensors caused by other heated objects [103]. Given the challenges faced in real contexts, it becomes rather complex to overcome these limitations with *ad hoc* classical computer vision algorithms.

In alternative to computer vision methods, to improve generalization in dynamic scenarios, intelligent systems approaches have been proposed, namely data-driven models based on feature engineering and fuzzy inference systems [104]. The nonlinear approach has been successfully tested in fire experiments under controlled conditions and validated in highly dynamic environments such as camping sites.

Most of previous approaches were designed for data acquisition purposes and algorithm design, whereas only recently autonomous systems are being developed as result of the increased data processing capabilities aboard aerial vehicles.

7.1.2 Proposed Approach

Thermal cameras have a clear potential for wildfire detection and monitoring tasks, but the path towards its integration in robotic perception pipelines, that are essential for autonomous systems, is still rather understudied. To contribute to narrow this gap, this work conducts a comprehensive study of thermal imaging sensing to extend the understanding of the inner-workings of this technology and how it can be leveraged for wildfire surveillance tasks. For that purpose, it is important to identify the main challenges this work addresses in the following.

First, although thermal imaging cameras are increasingly available for a multitude of industrial domains [105], with expanding model ranges and accessible equipment costs, its application is still limited, which can be attributed to two factors. On the one hand, most applications rely on a human-in-the-loop approach, which typically requires specialized training and technical expertise to interpret the image data, but this is in general based on high-level knowledge oriented to the domain of application. On the other hand, as discussed in the literature review, machine vision approaches depend on feature based approaches derived from image data, which do not take into account the image processing algorithms underlying the output data. However, these abstraction layers hinder the development of automatic algorithms due to the adaptive nonlinear nature that is at the core of the behavior of these systems. Therefore, the knowledge of the

underlying processing methods involved to generate the image output is central to understanding how to leverage this technology in a robotic perception framework. In that sense, the first step in this work concerns the overview of state-of-the-art image processing algorithms employed in most commercially off-the-shelf thermal cameras.

Second, adapting thermal imaging cameras for wildfire detection scenarios differs considerably from general applications e.g. industrial inspection or precision agriculture, in the sense that it deals with extreme temperatures. In this context, the importance of quantitative information is less prevalent than qualitative data because a fire can be identified by high temperature gradients with respect to ambient conditions, thus can be detected using the relative temperature differences in the images. To that effect, having radiometric information is not determinant because the intensity levels can translate the relative difference between objects in the scene. Nonetheless, a correct interpretation of the adaptive algorithms is required, because the color encoding schema adapts to the range of measurements in each instance. For these reasons, after covering the processing algorithms in the first part of this work, we demonstrate the implications of its usage in wildfire detection scenarios. To that end, several fire experiments under controlled conditions were conducted to study the behavior of thermal cameras in those situations, to characterize how the raw sensor data is mapped to visually interpretable pseudocolor images. In this regard, attention is taken to the identification of the saturation levels of this type of sensors.

Third, we discuss the implications of the insights exposed and the potential developments on robotic perception towards autonomous UAV-based fire monitoring systems, namely through the identification of current roadblocks and possible enabling solutions manufacturers should integrate for the widespread use of this technology.

After this introduction, the remainder of this article is structured as follows: Section 7.2 covers the overview of image processing algorithms employed in thermal cameras. Section 7.3 presents a data-driven method based on thermal imaging for fire situation-awareness. Section 7.4 presents the experimental setups and conditions of the laboratory and field trials. Section 7.5 analyzes the results and discusses the implications for the integration of thermal imaging in robotic perception on wildfire surveillance. At last, Section 7.6 presents the conclusions and offers suggestions for future research.

7.2 Thermal Imaging

The basic principle of thermal imaging is based on the concept of sensing the radiance emitted from objects of interest. Note that above absolute zero temperature, i.e. 0K, all bodies emit thermal radiation. In a general definition, what is referred to as thermal range is comprised by radiation with wavelengths in the 10^{-7}m to 10^{-4}m interval of the electromagnetic spectrum. Hence, this gamma includes spectral bands in the ultraviolet, visible, and infrared region of the spectrum.

Although this range is rather ample, usually thermal cameras only cover a part of it, which varies depending on the model and the application for which it is intended. More specifically, the devices employed in this

work are based on uncooled microbolometer detectors that cover part of the infrared band, namely the 7.5 - 13.5 μm spectrum, which is widely common for camera models in the market.

Thermal infrared cameras employ complex signal processing architectures in order to output images that convey relative temperature differences between objects in a scene [106]. This process comprises several stages that are summarized in Figure 7.1, which are briefly described in the sequence.

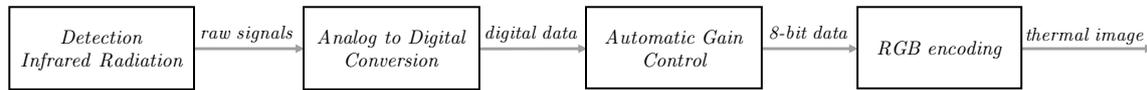


Figure 7.1: High-level diagram of the process flow in uncooled microbolometer arrays.

In the first stage, incident infrared radiation is absorbed, inducing changes of resistance in each microbolometer detector of the focal plane array, which are translated into a time-multiplexed electrical signal by a readout integrated circuit (ROIC) [107]. Then, the array is calibrated automatically at each it is powered to match all microbolometers to the same input/output characteristic function that relates the measured radiance intensity and the output signal. This is performed through a linearization process and temperature compensation of the signals from individual detectors of the array [108]. Second, with these compensated signals, the measurements are transformed into raw pixel values that translate the intensity values that compose a monochromatic image. Raw image data can be subsequently transformed into pseudocolor images through a automatic gain control procedure and RGB encoding according to a user-specified color palette to facilitate interpretation.

Besides sensing the radiation being emitted from objects in the field of view, another important aspect concerning thermal cameras is the ability to measure temperature. While this topic is extensively covered in the literature, with respect to how the incident radiation is transformed into approximate temperature readings [108], in this work we do not delve into this matter for two main reasons. First, with this study we aim to address the general image processing pipeline that is transversal to most thermal cameras, i.e. irrespective of these having radiometric capabilities or being non-radiometric. Second, although in different contexts the correction of temperature values of thermographic images can be performed *a posteriori* using off-line post-processing methods, this requires a known reference in the image content [109]. In the case of wildfire surveillance, if we consider the environment to be open and unknown with regard to the temperature, i.e. without access to external absolute temperature readings, on-line thermal correction of the calibration for real-time applications is not possible.

Taking into account these considerations, the main focus of this work is on how raw digital readings are encoded into pseudocolor images, which can be applicable to image data from both radiometric and non-radiometric thermal cameras.

7.2.1 Sensor characteristics: preliminaries

Thermal imaging systems are optical instruments that are able to generate two-dimensional representations of the surrounding environment as a function of the strength of incoming thermal radiation. These

exteroceptive sensors transform a digitally encoded array of pixels into an image according to the camera perspective projection, which depends on the focal length of the lens.

Digital images are generated according to the characteristics of the camera sensor, namely spatial resolution, temporal resolution and dynamic range [110]. The spatial resolution is intrinsically related to the size of the focal plane array and defines the number of pixels in an image, as well as the corresponding aspect ratio. The temporal resolution is associated with the operating frequency of device, i.e. the frame-rate at which the camera yields image data. In turn, the dynamic range corresponds to the gamma of intensity values represented. Additionally, the resulting images also depend on the automatic gain, which is a fundamental aspect of the manner thermal images are encoded, since most camera models employ automatic gain control, as will be discussed further along.

To explore the effect of these characteristics, this work resorts to two distinct thermal cameras, namely FLIR SC660 and FLIR Vue Pro, that operate in the region denominated as Far Infrared (FIR) or LWIR (Long Wave Infrared). The main specifications of these camera models are presented in Table 7.1.

Table 7.1: Summary of the specifications of the thermal cameras.

	FLIR SC660	FLIR Vue Pro
Spatial Resolution (px)	640 x 480	336 x 256
Temporal Resolution (Hz)	30	8.3
Bit resolution (bit)	14	16
Focal Length (mm)	19	9.0
Horizontal FOV (°)	45	35
Vertical FOV (°)	34	27
Spectral Band (μm)	7.5 ~ 13.5	7.5 ~ 13.5
Measurement Range (°C)	-40 ~ +1500	-60 ~ +150
Size [L x W x H] (mm)	299 x 144 x 147	63 x 44.4 x 44.4
Weight (g)	1800	92.1 - 113.4

Regarding the field of view (FOV) specifications and focal length, note that these cameras have distinct characteristics, which will influence the data recorded. Moreover, spatial resolution, temporal resolution and dynamic range also vary between both models, which is an aspect to take into consideration. In the case of the FLIR SC660 the automatic gain can be adjusted to different configurations to change the dynamic range.

In addition, having in mind the payload budget of small UAVs, comparing the weight and size of both cameras, only FLIR Vue Pro has suitable characteristics to be taken onboard a small UAV. However, given the exploratory basis of this work, the radiometric capabilities of FLIR SC660 are valuable for the study of these sensors in extreme conditions inherent to fire scenarios.

Although this work explores these cameras, other alternative models have been released more recently,

with higher spatial resolution and gimbal integration, albeit at superior costs. Notwithstanding, from the standpoint of hardware and firmware available the update does not have a significant impact on thermal imaging results, since the major benefits of the newer models lie on facilitated integration for deployment on UAVs or graphic user interfaces for control from mobile devices. Thus, the analysis presented herein also applies to different thermal cameras.

In the sequence, the image processing methods for generating thermal images are described, through examples explaining the mapping algorithms that perform the transformation of raw digital data into pseudocolor encoded images.

7.2.2 Mapping raw digital data to thermal images

Thermal cameras are in general single-band in the sense these only produce monochromatic images, which subsequently undergo a sequence of image processing steps, to transform the raw digital data into pseudocolor images to highlight the details of the scene context. This processing pipeline can be implemented either on-device for storage or digital output, or on external software.

Currently, commercially off-the-shelf devices already provide a variety of color palettes to enhance the visual interpretation of the amounts of radiance captured by the sensors. However, to design intelligent algorithms for autonomous systems, the color encoding schema has to be well-suited for the robotic perception approach, which is essential to fulfill the application requirements. For this reason, this also requires a deeper understanding of the image processing pipeline to leverage the potential of this type of sensors for novel applications.

The raw intensity levels are given by a digital number assigned by the sensor analog to digital converter, which can be of 14-bit or 16-bit order depending on the sensor bit resolution. In this work, we will employ both these alternatives as specified in Table 10.1, but since these algorithms apply in the same way to both versions, the following examples showcase only the processing of raw data in 14-bit space. Further along in the analyses of the sensor response to fire scenarios both cases will be covered in detail. Note that the bit resolution relates to the temperature range of the camera and for camera models with different modes such as the FLIR SC660, the intensity level values will also be influenced by the camera configuration, i.e. the high-gain or low-gain modes.

To obtain a thermal image in pseudocolor, the image processing pipeline is divided into two main steps: (1) application of a data compression technique denominated automatic gain control; (2) application of the color palette specified yielding images with three channels corresponding to the RGB color-space representation.

Automatic Gain Control (AGC) is a histogram-based technique that performs the transformation between raw data formats to 8-bit image data. This processing method is responsible for data compression, which implies a considerable loss of information. For the 16-bit case, from a range of possible values of 0 to 65535, the resulting image will be represented with values in the 0 to 255 interval. To counteract the

decrease in detail, the AGC algorithms are designed to enhance the image contrast and brightness in order to highlight the scene context.

The following sections cover with illustrative examples the main variants of AGC algorithms implemented in thermal cameras that practitioners should be aware off to leverage this technology. Then, the color mapping schema and several color palettes available are presented in Section 7.2.2.

Histogram-based automatic gain control

AGC methods are typically variants of histogram-based operations widely used in computer vision for contrast enhancement, e.g. histogram equalization [110]. However, in thermal imaging AGC also implies data compression between raw formats (e.g. 14-bit or 16-bit) to display-ready data (8-bit).

In classical histogram equalization the nonlinear mapping used for contrast enhancement is derived directly from the cumulative distribution function (*cdf*) of the raw intensity values. This approach allows achieving an approximate linear *cdf* on the compressed 8-bit data, yielding an image with intensity values spread across the full extent of the available 8-bit range. Figure 7.2 illustrates this AGC procedure for a 14-bit raw image converted to 8-bit range.

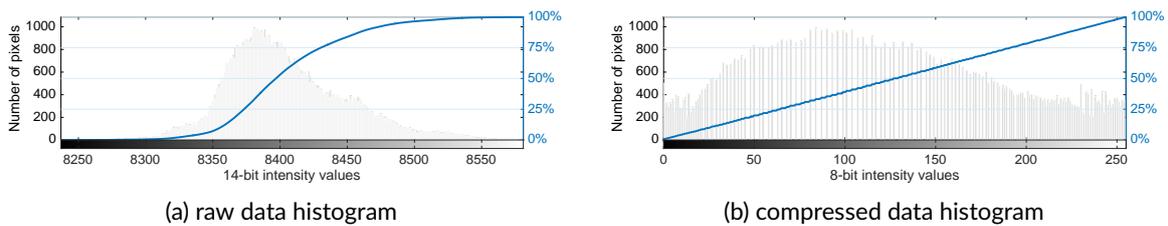


Figure 7.2: Automatic gain control with linear histogram.

Note that although the bit resolution of the 14-bit sensor represents values up to 16383, for environments with ambient temperatures around 20°C the raw data captured is represented in a narrow band of the full range, as can be observed in Figure 7.2. Therefore, compression and contrast enhancement play a pivotal role in the encoding of thermal images. However, note also that enhancement operations in thermal images artificially distort the data, meaning that the physical correlation that relates the radiant flux from infrared radiation and pixel intensity is lost.

Alternatively, for cases where it is important to preserve the correspondence between pixel intensity and temperature of objects for instance, a "linear" mapping function is better suited. The linear approach also relies on the *cdf* to define the image transformation table (ITT), by defining the slope and clipping points of the resulting nonlinear mapping function. Figure 7.3 depicts the application of this AGC algorithm to the previous example.

In addition to the *cdf*, the linear transformation, given by $T(x) = mx + b$, requires setting the midpoint of the Image Transfer Table, ITT_{mid} , which is normally the average value of the range of the 8-bit range (128), and the tail rejection percentage $r\%$. These parameters are used to determine the points on the *cdf*

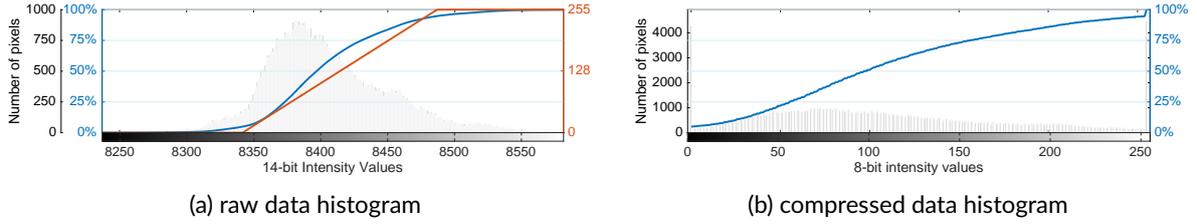


Figure 7.3: Automatic gain control with linear histogram.

employed to define the linear equation and respective slope, m , and intercept, b , as:

$$b = \text{ITT}_{\text{mid}} - \text{mean}(x_{100-r\%}, x_{r\%}) \cdot m \quad (7.1)$$

$$m = 255 / (x_{100-r\%} - x_{r\%}) \quad (7.2)$$

with x representing the raw values indexed in the x-axis of the histogram. The image transformation function, $T'(x)$, is defined for the gamma of values of the histogram by clipping the values to the lower and upper bounds of the output domain $[0, 255]$, as can be observed in Figure 7.3. Note that the clipping of the increases the absolute frequency of 0 and 255 intensities in the 8-bit histogram. Subsequently the data are converted through the ITT look-up table.

In practice, the ITT midpoint influences the brightness of the image in the sense that increasing or lowering the midpoint shifts the equation horizontally to the left or right, respectively. As a result this clips the data to zero at a corresponding lower or higher raw value, and vice-versa for the 255 upper bound. This aspect is especially relevant because for fire detection scenarios which present high temperature gradients both low and high raw values are important for situation-awareness. Likewise, the definition of the tail rejection percentage follows the same principle. Therefore, parameter tuning should be approached with caution as not to discard relevant data.

Nonetheless, the linear algorithm is not the most used AGC method, because this compression technique implies a considerable loss of detail. To avoid this limitation, thermal cameras usually employ the plateau equalization algorithm that aims to balance the distribution of all intensity levels in the image scene, thereby enhancing image contrast and highlighting differences in temperature.

The plateau equalization algorithm [111] implements a nonlinear mapping that compromises between histogram projection and histogram equalization. The concept is to bound the representation of the different intensity levels to a defined threshold termed as plateau value, P , while limiting the slope of the transformation function through the maximum gain value, G , which by default is set to 1. However, for images with low dynamic range, where a small interval of values have high bin counts, the plateau equalization algorithm may yield less intensity values than the 256 available in 8-bit space. To ensure the entire contrast depth is leveraged, the maximum gain sets the upper limit of gain that can be used to stretch the data to the full extent of the 8-bit range. Considering fire detection applications exhibit most likely high dynamic range situations, this parameter will not be tuned in the development of this work.

First, the image histogram is clipped according to the plateau value, and represented through the effective count, c_x , for each raw intensity level x , which conditions the absolute frequencies not to exceed the prescribed plateau value, P , as:

$$c_x \equiv \min(C_x, P) \quad (7.3)$$

where C_x represents the original count of pixels having raw intensity value x . The plateau equalization mapping function is based on the cumulative distribution function of the clipped histogram values, denoted by cdf' , which is computed according to:

$$cdf'(x) = \sum_{j=0}^x c_j, \quad 0 \leq k < 2^N \quad (7.4)$$

with N representing the exponent corresponding to the original bit resolution of the sensor. Then, the transformation function based on plateau equalization for a 8-bit compression is defined as:

$$T_{PE}(x) = \left\lfloor \frac{255 \cdot cdf'(x)}{cdf'(2^N)} \right\rfloor \quad (7.5)$$

where $\lfloor \cdot \rfloor$ represents the truncation operator to the next lower integer. Note that if the plateau value equals the maximum absolute frequency in the original histogram, this algorithm is equivalent to histogram equalization. In turn, if the threshold is defined as 1, this algorithm behaves as histogram projection. The plateau value is by default established as a percentage of the maximum value of the bin count of the histogram, e.g. 7%, but varies depending on the camera model and specifications. Figure 7.4 illustrates the application of the plateau equalization algorithm for the previous example.

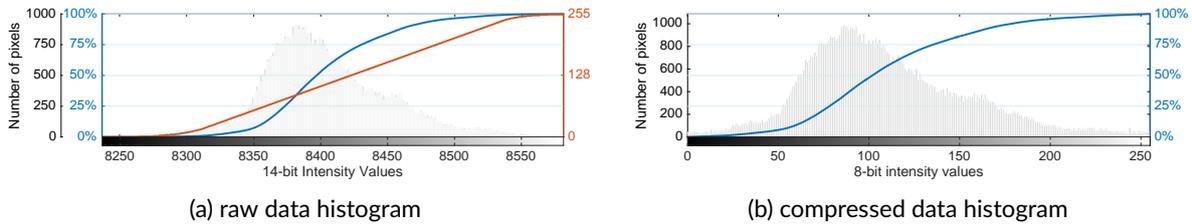


Figure 7.4: Automatic gain control with plateau equalization algorithm.

Besides these algorithms some cameras also have the information-based equalization variant that combines the plateau equalization algorithm with basic image enhancement techniques to yield more detail from scene context, irrespective of the data distribution. This method is not implemented in FLIR ResearchIR but was implemented herein for illustration purposes.

Information-based equalization aims to allocate a proportional amount of the dynamic range to different parts of scene to capture the most information from the scene context, irrespective of these being represented with a large part of the image such as the background, or by small areas with slight variations in temperature in the foreground. This method performs this by using a high pass filter to highlight detail in the image, which is subtracted from the original image before application of plateau equalization. The low pass histogram is subsequently modified by increasing the bin count of the pixels in the high pass im-

age. Hence, increasing image contrast whilst also including greater detail. For instance, in a fire scenario, this may be useful to distinguish the fire while also being able to discern if there are people in the scene. Since this is an extreme case, with high dynamic range in the raw data, the effect of the compression to 8-bit data implies a significant loss of detail, therefore making more important these image enhancement techniques to highlight even subtle variations in temperature.

While in controlled environments adequate tuning of this type of algorithms can yield better results, in outdoor contexts adjusting these methods for a robust and consistent performance becomes very complex, because the uncertainty and measurement errors associated with these open and unknown environments is greater due to the distinct emissivity properties of the multitude of heterogeneous materials that can be encountered [108]. In this way, by distorting the actual measurements these contrast enhancement techniques also introduce greater inaccuracy in the image representation, which can hinder the development of robust robotic perception methods.

Thermal Imaging Metadata

In addition to the image data, thermal imaging files also store different types of metadata encoded under standard metadata formats e.g. TIFF, Exif and XMP. By accessing this information through adequate interfaces several valuable parameters can be retrieved. Besides encoding the standard properties of digital camera e.g. focal length, focal plane resolution and GPS coordinates, the manufactures can also store camera calibration parameters relevant for conversion between raw intensity values and temperature values.

Herein, we explore relevant metadata tags which help to shed light into the working mechanisms of these sensors, providing important insight into useful parameters for robotic integration. More specifically, since we tackle the relation between raw data and RGB encoded data, only the parameters that influence this transformation will be addressed.

The metadata was retrieved from raw video files and image files encoded in "SEQ" or "TIFF" and a proprietary format from FLIR that encodes both JPEG and metadata as outlined in Figure 7.5.

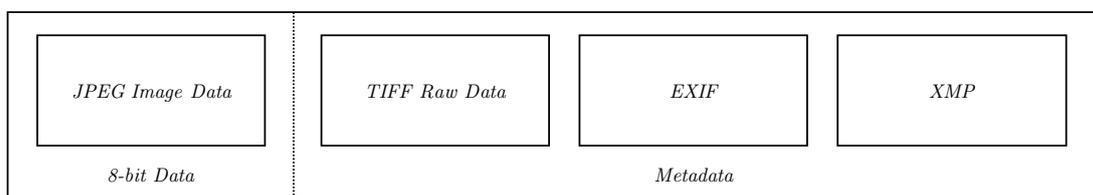


Figure 7.5: Schematic of file structure of FLIR image format.

By analyzing the metadata of the image and video files, two Exif tags were identified that are closely related to the histogram-based algorithms presented in Section 7.2.2, namely Raw Value Average and Raw Value Range. The range of raw values and average values are encoded separately in the metadata but may not correspond to the values that would be computed from the raw frames for every instance. The reason for this is probably the proprietary in-camera processing to deal with noise and other types of outliers. For FLIR SC660 the calculated values do not match the encoded values in the files, whereas

for FLIR Vue Pro both match exactly for data recorded in image mode. This indicates that the update of these values also varies depending on the device and the mode of capture as it is related to the firmware. With the average and the range of raw values from the metadata tags, the maximum and minimum values of the color scale are computed as follows:

$$\text{Raw Value Max} = \text{Raw Value Average} + \frac{\text{Raw Value Range}}{2} \quad (7.6)$$

$$\text{Raw Value Min} = \text{Raw Value Average} - \frac{\text{Raw Value Range}}{2} \quad (7.7)$$

However, note that if these metadata tags are not available the maximum and minimum values of the raw data can be computed directly online as observed in the histogram-based algorithms presented in Section 7.2.2, by extracting maximum and minimum values directly from the raw frames.

Given that the histogram-based mapping functions and the respective color encoding adapt according to the scene context, these parameters are essential to understand how the color scale adjusts over time. Thus, these will be explored next in the data-driven analysis approach proposed in Section 7.3.

Considering that fire surveillance applications are an extreme case with high dynamic range, it is important to evaluate how these techniques behave in such scenarios. Thus, to delve into this issue, in the sequence the color encoding used to enhance the interpretation of thermal data is presented, along with the comparison of an example for a controlled burn performed in a real-world context.

Color mapping

Following the conversion to 8-bit image format, the data is represented in the 0 to 255 range. To encode these values in a RGB color space representation, the cameras offer several color palettes with distinct characteristics, which are adequate for different applications. In cases where the cameras provide the raw data in addition to the RGB encoded images, this data can be post-processed with different color palettes for further analysis.

The color palette is a discrete set of color samples composed by values for each color channel and is defined in a lookup table (LUT) in the camera firmware or external software. In practice, this discrete sequence of values provides a continuous representation of the mapping of values in the image. Figure 7.6 illustrates this for some widely available color palettes, presenting the set of discrete colors that form the color mapping applied to the 8-bit data. The color sequences are depicted matching the increasing order of bit values, meaning these are attributed from lower values to higher values of sensed radiant flux. Notwithstanding, recall that the color scale is adaptive, thus the color assignment also depends on the AGC algorithm and for this reason the full spectrum of colors is not necessarily used in each image.

The different types of colormaps depicted can be advantageous in different applications. For instance, the WhiteHot or Ironbow, employ a sequential colormap with an uniform distribution between two main colors. In turn, Lava uses the contrast between several colors to enhance subtle differences between tem-

peratures of objects in the scene, whereas the GrayRed alternative employs a divergent color distribution to highlight large temperature gradients in a scene.

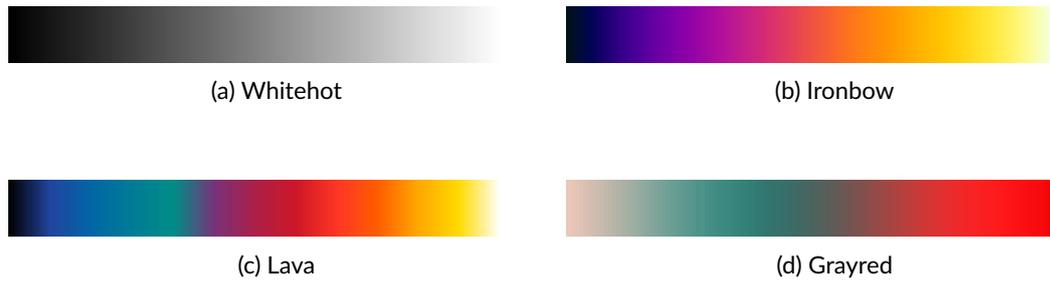


Figure 7.6: Examples of thermal imaging color palettes.

Since herein the principal interest is to adapt thermal imaging sensors for wildfire detection and monitoring, the GrayRed color palette was selected. As depicted in Fig. 7.6d, this palette applies high-contrast colors with a divergent color scheme, which is useful to draw attention to the hottest objects in the scene. Furthermore, an extensive color-based data analysis also employing this colormap in the detection of fire situations was presented in previous work [104]. Thus, in the interest of extending the scope of this investigation to the analysis of both color encoded images and the raw data, the same color palette will be used in this article.

For comparison of alternative automatic gain control techniques, Figure 7.7 demonstrates the resulting images obtained with the different histogram-based methods, for a controlled burn in a real context. The linear and plateau equalization (PE) methods follow the principles explained previously, whereas two additional FLIR proprietary methods available through FLIR ResearchIR are also presented, namely Digital Detail Enhancement (DDE) and Advanced Plateau Equalization (ACE).

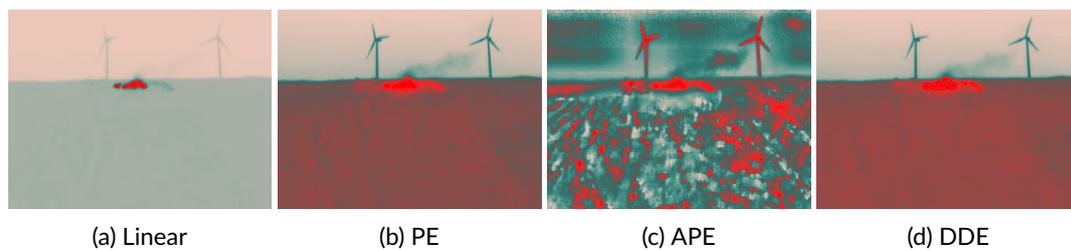


Figure 7.7: Comparison of results of AGC algorithms: (a) linear, (b) plateau equalization (PE), (c) Active Plateau Equalization (APE), (d) digital detail enhancement (DDE).

From observing these examples it becomes clear that advanced contrast enhancement techniques degrade the ability to distinguish the fire from the surroundings. Note that these algorithms were applied with the default configurations, i.e. in its less intense mode, so further tuning of the parameters would deteriorate the image interpretation even further.

In addition these core AGC methods, the image processing pipeline can also include for some cameras an array of proprietary features such as Active Contrast Enhancement (ACE) and Smart Scene Optimization (SSO). These algorithms also tune the image output and their parameters can be adjusted manually using manufacturer software applications, but in this study the default values were used.

7.3 Data-driven Thermal Imaging Analysis

Since fires are sources of extreme heat with a strong emission of infrared radiation, thermal imaging sensors can potentially bring significant advantages for fire surveillance applications. However, to implement autonomous systems for early-detection and monitoring, i.e. independent of human-in-the-loop approaches, the robotic perception algorithms have to be data-driven.

To provide a preliminary intuition on the behavior of this type of sensors in the advent of a fire ignition, this section starts by introducing an experimental example of a fire detection instance in a controlled environment with static image capture conditions. Subsequently, a feature engineering approach is designed for data-driven situational-awareness in fire scenarios.

7.3.1 Thermal Imaging in Fire Scenarios

The following example presents an illustrative demonstration relating image response, the temperature profile of the scene, and raw measurements to offer some insight into the sensor behavior. The experiment was performed in a laboratory environment, with direct line-of-sight between the fire and the thermal camera, in this case FLIR SC660. Figure 7.8 presents selected samples from the ignition of a straw fuel bed at different time instances. The corresponding temperature profile, which is represented in Figure 7.9, was extracted from the radiometric data provided through FLIR ResearchIR software. At this juncture, this example is presented as a preliminary analysis. For further information concerning the experimental setup refer to Section 10.2, where the experiments are covered in full detail.

Observing the samples depicted in Figure 7.8, it should be noted that although the strongest source of thermal radiation is depicted in red, this does not mean it is necessarily fire. Hence, it would be a naive approach to focus only on the detection of the hottest objects in the scene. This is caused by the adaptive nature of the histogram-based algorithms in the image processing pipeline.

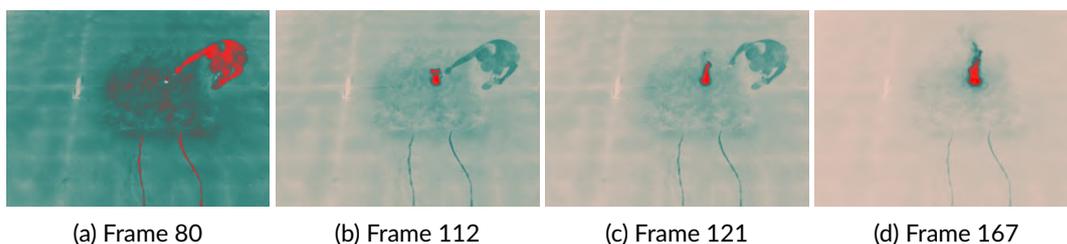


Figure 7.8: Fire ignition detection with a radiometric thermal camera.

In the first frame, right before the ignition, the majority of the frame has green colored pixels, whereas from the second frame onwards, as the fire starts the contrast progressively shifts. In result, the objects with temperatures close to the ambient temperature are subsequently represented by pixels with a whiter shade. It is important to notice that in this context the color scale that translates the differences in temperature adjusts with respect to the object in the scene with the maximum temperature. This effect can be observed by comparing the first and second frames. Whereas in the first the person can be observed

in red, in the second the person is depicted with a dark green. This happens because radiation being emitted by the greatest heat source is measured by the sensor with a much higher intensity than the radiation emitted by the person. While the temperature of the person remained unaltered, the range of raw measurements expanded, prompting a significant adjustment in the color scale. The underlying reason for the color evolution is that at this stage the body temperature is much closer to the ambient temperature than to the temperature of the fire. Comparing the first and last frames highlights that as the temperature differences increase the contrast in the image adjusts accordingly.

The temperature profile of the experiment is illustrated in Figure 7.9, along with the raw statistics computed from the data frames. Note that the temperature profile depicts the maximum and minimum values of temperature extracted from the pixels in the image, as well as the average temperature computed with the maximum and minimum values. Note that the first increase in maximum temperature around frame 80 concerns an ignition that did not fully develop, so a re-ignition was promptly conducted, which can be observed by the increase in maximum temperature from around frame 110 onwards.

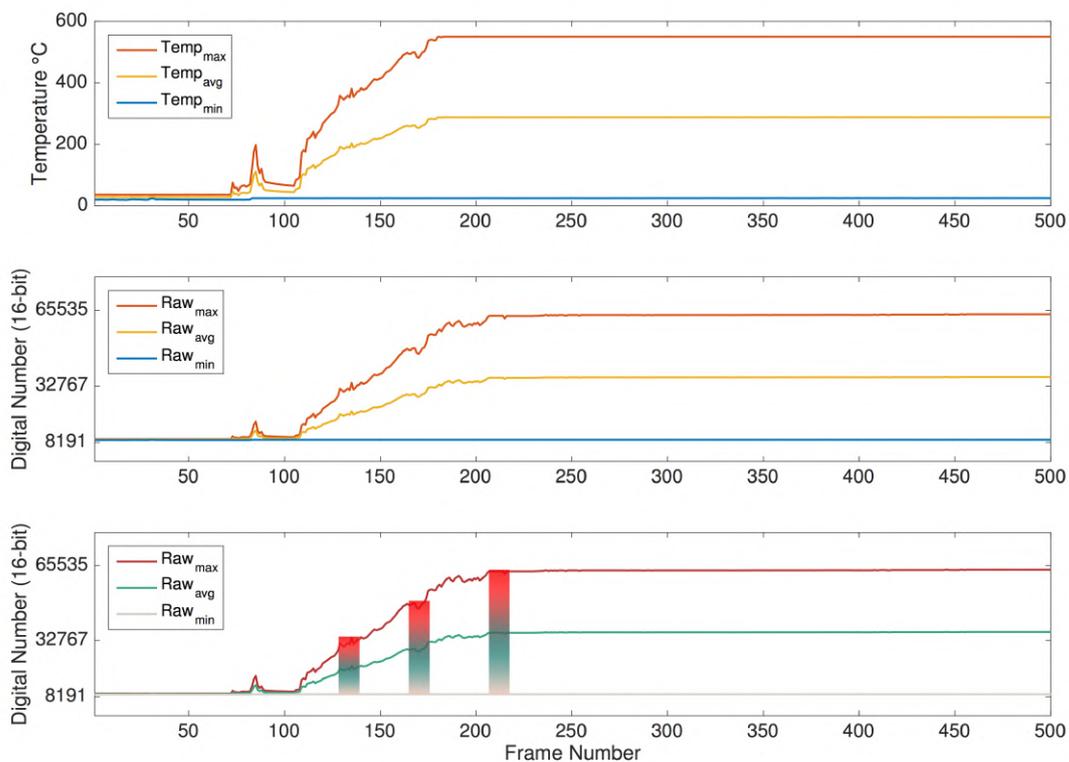


Figure 7.9: Comparison of temperature and raw data and the evolution of the adaptive color scale.

In juxtaposition, Figure 7.9 also represents the corresponding evolution of the raw values, including the maximum and minimum raw values extracted from the image, as well as the average between these two metrics, to illustrate how the color scale adjusts as the dynamic range extends. Note that this depiction does not take into account the application of histogram-based algorithms, in the sense that it only translates a linear application of the GrayRed colormap to the raw values. Nevertheless, it provides a broad intuition regarding the adaptive color scale, while relating the raw measurements and the temperature data with the pseudocolor images presented in Figure 7.8.

Since for this trial the camera was configured to measure up to 500°C, it is noticeable by observing both graphs, that the sensor saturates when that temperature is reached.

The analysis of these variables highlights the importance of testing the thermal imaging cameras in real operation scenarios and characterizing the behavior of these sensors in extreme conditions, as it may exhibit differences depending on the camera models. In the case of the FLIR Vue Pro in image mode the values in the metadata match the maximum and minimum values of the raw frames.

Therefore, in order to employ thermal imaging sensors for situation-awareness in fire detection and monitoring scenarios, it is important to translate the insights from this analysis into meaningful features for robotic perception, as explored will be in the sequence.

7.3.2 Feature engineering

Autonomous robotics requires real-time processing of sensor data for navigation and guidance, e.g., to enable sensor-driven planning, as it is crucial for the robot to have precise estimates of its position and attitude, as well as its surroundings, in order to plan its trajectory online. Even though emerging on-board computing platforms have higher processing capabilities, to incorporate powerful exteroceptive sensors capable of generating high-resolution samples of the surrounding environment, it is essential to handle the burden of data dimensionality. Furthermore, considering the required high sampling frequencies suitable for these tasks, mobile robotics applications are increasingly data-intensive, thus feature engineering plays a prominent role to derive efficient algorithms.

In light of the insights exposed in the previous sections, the following presents a feature engineering approach for situation-awareness using thermal imaging, that combines feature extraction from raw data and pseudocolor images, to yield relevant features for robotic perception in fire scenarios.

Raw-based features

The feature extraction approach for raw data is based on the construction of features that aim to preserve the intuition of the actual working mechanism of thermal imaging sensors and the histogram algorithms of the image processing pipeline. Recalling the examples presented in Section 7.2.2, the raw-based features will explore the maximum, minimum and average value of the range of the raw data frames. As explained in the metadata description in Section 7.2.2, these variables can be computed according to equations (7.6) and (7.7), but since these do not always match the metadata tags for the cameras tested, it is advisable to extract these variables online from the raw data frames.

Color-based features

To extract features that translate the evolution of the color statistics of the image in fire scenarios, as observed in Figure 7.8, this work employs the color segmentation heuristic proposed in [104]. The approach divides the GrayRed color palette into three parts according to the three main colors that compose this

palette: gray, green and red. The complete color scale that comprises 120 distinct color representations, defined in the RGB color space, is partitioned as illustrated in Figure 7.10.



Figure 7.10: Division of the color scale.

The first segment represented in gray corresponds to lower raw values, i.e. lower temperatures, and is defined by 18 color levels, which represent 15% of the color scale. The green segment is defined by 48 color levels, that represent the mid-range of raw values, corresponding to 40% of the color palette. Since this color palette is designed to draw attention to the hottest elements in a scene, the largest portion corresponding to 45% of the color scale is dedicated to 54 red color levels. Although these color exhibit high contrast between them, color gradients between each color are low. The segmentation is performed according to the parameters in RGB color space defined in Table 7.2.

Table 7.2: RGB segmentation thresholds for feature construction.

features	gray			green			red		
	R	G	B	R	G	B	R	G	B
upper limit	253	199	185	143	169	157	255	73	71
lower limit	149	171	160	98	90	86	103	89	85

In contrast with strictly data-based approaches, the proposed feature engineering method allows the incorporation of expert knowledge about the sensor by deriving features from the information retrieved from raw data, in combination with the data-driven color segmentation heuristic that describes the behavior of the sensor over time. Furthermore, these features are designed with data interpretability in mind, which is especially important for safety-critical applications such as wildfire detection and monitoring systems.

In order to study the response of thermal imaging sensors, a comprehensive set of experiments was conducted, for analysis of the sensor image processing pipeline and its application in real scenarios, which are presented in the sequence.

7.4 Controlled Fire Experiments

This section presents a set of controlled fire experiments performed in laboratory and field contexts. The trials comprise the burning of wild fuels such as straw, common heather (*Calluna vulgaris*) or Baccharis (*Baccharis trimeris*), as well as artificial materials present in a caravan that was also experimentally burned. In the following, each section describes the experimental setup and conditions used in the trials, as well as the means employed for data acquisition.

7.4.1 Laboratory Test

The laboratory experiment enables testing the response of the camera to a fire ignition in a controlled environment, allowing a high-resolution sampling of the phenomena under environmental static conditions. Since the test was performed indoors there was no influence of wind. The horizontal fuel bed was composed by straw with a fuel load of 600g/m^2 and a moisture content of 13%. The environmental temperature and relative humidity were 20°C and 78%, respectively. By using these testing conditions the behavior of the sensor and the subsequent image processing is not affected by external factors.

For data acquisition purposes, the FLIR SC660 is positioned on top of an elevated platform with direct line-of-sight to the fire as depicted in Figure 7.11. The line of sight makes an angle of 45° with the horizontal plane. In this setup, the data are processed and recorded in real-time by FLIR ResearchIR software, through a cable connection to a desktop computer.



(a) experimental setup



(b) top view of fire in straw fuel bed

Figure 7.11: Laboratory trials with FLIR SC660 mounted on an elevated platform.

The objective of this type of test centers on capturing the transition to a fire scenario, thus it requires a high frame rate. In this case, the device was configured to 15Hz frame rate.

7.4.2 Caravan Burning Test

The caravan burning test comprises the burn of man-made fuels, unlike the remaining trials which only include natural fuels. This scenario is important to take into account especially for airborne fire surveillance in wildland-urban-interface regions. In addition, due to the presence of some materials like plastic, the temperatures expected are much higher than those registered for wild fuel burning. The test was performed with average values of relative humidity of 75% and temperature of 23°C . The wind velocity was below 2m/s .

Akin to the previous trial, the device used for data acquisition was the FLIR SC660 operated in connection with FLIR ResearchIR software with a 15 Hz frame rate. In this case the camera was mounted on a tripod at ground level at approximately 10 m from the caravan, which is depicted in Figure 7.12.

To study this case, as fire related temperatures reach higher levels, the automatic gain control configuration



(a) before flashover



(b) after flashover

Figure 7.12: Caravan burning trial recorded.

was changed mid-trial to expand the measurement range to cover higher temperatures.

In this trial, unlike the previous example, there is not direct line-of-sight to the fire at the point of ignition, which occurred in the interior of the caravan, as can be observed in the visible range image presented in Figure 7.12a. Indeed, only after flashover, i.e. the point where the heat built-up reaches the maximum, the fire starts to be visible in the visible range.

7.4.3 Summer Festival Trials

To design algorithms with robustness in real-world contexts, it is crucial to analyze data acquired in highly dynamic scenarios. In that sense, a set of tests was performed at the venue of a festival outdoors in midsummer, to emulate conditions of real operation of fire safety systems. The trials were conducted in a region of dry weather in the summer, with strong winds and high temperatures, when conditions of ignition propensity were consequently high as well. More specifically, during the period of the trials the following meteorological conditions were registered: minimum humidity of 17%, maximum wind speed of 21km/h and temperatures up to 36°C.

The tests performed encompass the aerial surveying of the festival area at several different locations. Figure 7.13 illustrates part of the venue surroundings, and provides some detail on the nature of the camping areas. As can be observed, the festival occurred in the fire prone area, where the vegetation of trees and shrubs are very combustible. Besides the vegetation, the inclusion of more combustible materials such as tents, caravans increases fire risk which is exacerbated by the human presence, raising fire safety concerns.

To provide aerial monitoring for several hours as to detect possible fire ignitions, the thermal camera was installed on a tethered helium balloon, as depicted in Figure 10.6. With this setup, the experiments were conducted at various locations of the venue to test distinct conditions. There were no fire events in these field tests, so for the purposes of this study, the fire in the rocket stoves of the community kitchens was used instead.



(a) festival venue



(b) camping areas

Figure 7.13: Festival venue and camping areas during the event where the trials were conducted.

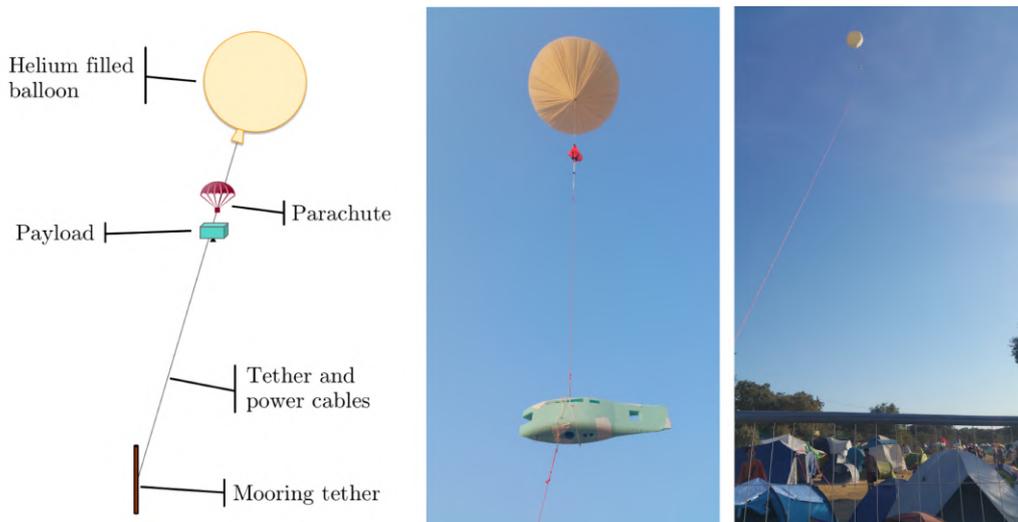


Figure 7.14: Thermal camera setup on tethered helium balloon.

Concerning image acquisition, due to the limited payload budget of the balloon, only FLIR Vue Pro has suitable weight to be installed onboard. The camera was programmed to capture images with a 5s time-lapse, recording in the camera storage both the raw data and the pseudocolor images encoded with the GrayRed color palette.

In contrast to previous trials, where the acquisition conditions were static, since these experiments employ an aerial platform, without active means of actuation for stabilization, the dynamics of the balloon influence the image acquisition by varying the image field of view as a result of wind disturbances.

7.4.4 Mountain Range Field Trials

In addition to the previous experiments, several tests were performed in a real-world scenario in mountain range field trials, to test the capabilities of the sensors in a long-distance monitoring scenario. For these tests, the main interest is to evaluate the performance of a compact thermal camera, namely FLIR Vue Pro, for future airborne monitoring operations. The trials were undertaken at an altitude of over 1000 m with the following meteorological conditions: wind speed of range of 13-16km/h from NW, average relative humidity of 45%, and average temperature of 21°C.

Figure 7.15 illustrates the experimental conditions of the trials. The field experimental burns were conducted in a hill with average slope of 25%, and comprised the burn of shrubland vegetation, including common heather (*Calluna vulgaris*) and Baccharis (*Baccharis trimerera*).



Figure 7.15: Experiments in mountain range field trials.

Concerning the image acquisition, in the trials both thermal cameras were employed, but given the aforementioned objective of these experiments in particular, this work only explores the data from FLIR Vue Pro. The images were captured at ground level from across the valley at around the same altitude as the field burnings, with direct line-of-sight between the camera and the fire at a distance of approximately 600m. The data acquisition was configured to record the data on the camera storage, saving both the raw data and RGB encoded images with the GrayRed color palette at a rate of 1 frame per second.

7.5 Results and Discussion

This section presents the results of the application of the data-driven thermal imaging analysis to the controlled fire experiments performed in laboratory and real contexts, according to the specifications presented in Section 10.2. As previously introduced, the proposed approach explores raw-based features, as well as the color-based features derived from the segmentation heuristic proposed in Section 7.3.

7.5.1 Laboratory Test

In the laboratory trial, under static conditions of image acquisition, it was possible to record with FLIR SC660 the transition from a situation without fire to a fire scenario, through a high-rate sampling which includes the point of ignition. This enables capturing the influence this phenomenon produces in the data acquired.

Figure 7.16 depicts the evolution of the raw-based features that illustrate how the dynamic range expands as the fire develops. In turn, it also demonstrates intuitively how the variation of the dynamic range influences the color statistics of the images. The sample frames concern the transient stage immediately before and after the start of the fire. Note that the first small peak around frame 80 occurs in a first attempt at the ignition, but as the fire does not fully develop, the straw was reignited. The fast response

of the sensor indicates that it is sensitive enough for this type of application.

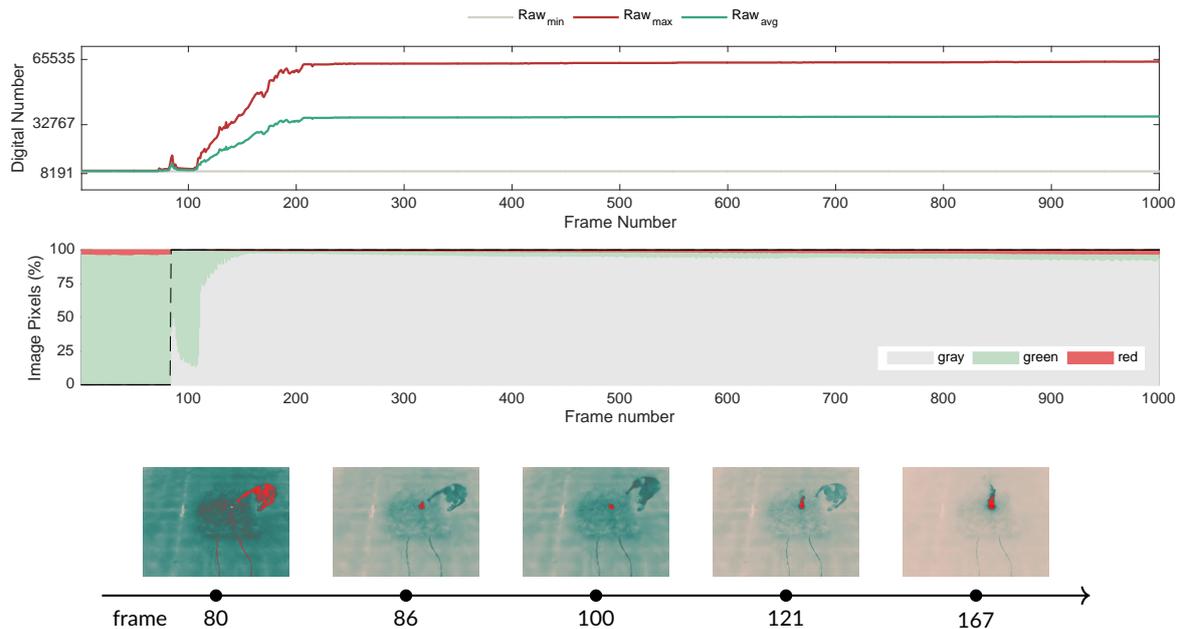


Figure 7.16: Feature-based representation of the sensor response of thermal camera to a fire ignition.

Attending to Figure 7.16, it is possible to identify that the sensor undergoes three distinct states: i) before the point of ignition; ii) transient state; iii) steady-state. As indicated in the figure, at first the environment is in equilibrium, with generally green levels corresponding to the mid-range. Note that the dynamic range in this state is narrow since the incoming radiation originates predominantly from objects at ambient temperature.

In the second stage, which is prompted by the fire ignition, the camera detects different strengths of incoming radiation corresponding to a severe temperature gradient and as a result the dynamic range starts increasing. This drives the adaptation of the color representation, according to the histogram-based algorithms presented in Section 7.2.2. Consequently, the increase in percentage of gray pixels is prompted by the fire ignition, which represented in red since it is the hottest object in the scene and has strong emission of infrared radiation, whereas objects at ambient temperature remain at the previous raw level.

Third, a new equilibrium state emerges after the sensor stabilizes shortly after the beginning of the fire, approximately at frame 167. However, in these circumstances due to the extended dynamic range the color statistics have a considerably different representation with a high percentage of grey shades in image context.

For this case, it is important to take into account that this experiment captures the image sequences with a frame rate of 15Hz, thus it is possible to record the transition between the equilibrium states. For cases with lower sampling frequency, the data may not capture this behavior, but only similar conditions resembling the first and third state.

Note also that although the dynamic range adapts over time its measurement range is not altered, as such

the complete bit resolution is used in this configuration. Accordingly, since the measurement range is set to $-40 \sim +550^{\circ}\text{C}$, the raw measurements saturate at a digital number that would correspond to these temperatures, in this case these match the full bit resolution from 0 to 65535.

7.5.2 Caravan Burning Test

In the caravan burning test performed in an outdoor environment, as described in the experimental setup, the image acquisition conditions were also maintained static, with the FLIR SC660 configured to capture image sequences at 15 frames per second. While this setup also allows for a high sampling of the phenomenon, in this case the fire ignition is not in direct line of sight, because it occurs inside the caravan. For this reason, it is also necessary to characterize how these circumstances impact the behavior of the thermal cameras in such scenarios. Moreover, the measurement range of the camera was adjusted mid-trial to encompass an extended gamma of temperatures. Consequently, this allows the evaluation of how the scaling of the bit resolution affects the raw digital data values.

However, before delving into the trial results, the effect of automatic temperature-dependent non-uniformity correction (NUC) should be addressed. This method consists of an internal process employed in thermal cameras to deal with spatial non-uniformity (NU) noise, which produces a fixed-pattern over the thermal images, that varies in intensity due to the internal temperature instability of such cameras [106]. To that end, on occasion the camera performs a NUC reset, which results in a disturbance in the data acquisition process. Figure 7.17 exemplifies an instance illustrating this behavior along with relevant sample frames.

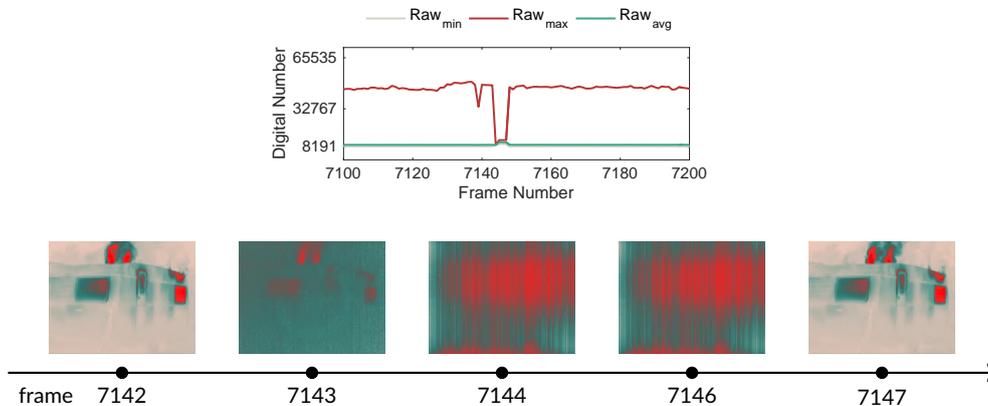


Figure 7.17: Effect of temperature-dependent non-uniformity correction (NUC).

Attending to the selected window of the raw-based features depicted in Figure 7.17, it can be observed that when the camera triggers the NUC reset, in frame 7143 the raw values exhibit a sudden drop, which continues for several instants until frame 7147. Given that this behavior is sporadic but it is expected to occur, as was verified for the case in the caravan burning test, appropriate filtering of this type of frames should be included in the image processing pipeline in automatic systems, particularly if integrated in autonomous navigation of robotic platforms for early detection and monitoring. In this work, a median filter was used with a window of 15 frames, which introduces a one second delay in the case of FLIR SC660, since this camera operates at 15Hz.

Figure 7.18 illustrates the variation of the features extracted from the raw data, in comparison with the color features derived with the color segmentation heuristic, for images encoded with a linear histogram approach and the plateau equalization algorithm in the second and third graphs, respectively. The data presented in the graphs was pre-processed to filter out the effect NUC frames. In this example we present the sensor response for two automatic gain control methods: linear (L) and the plateau equalization (PE).

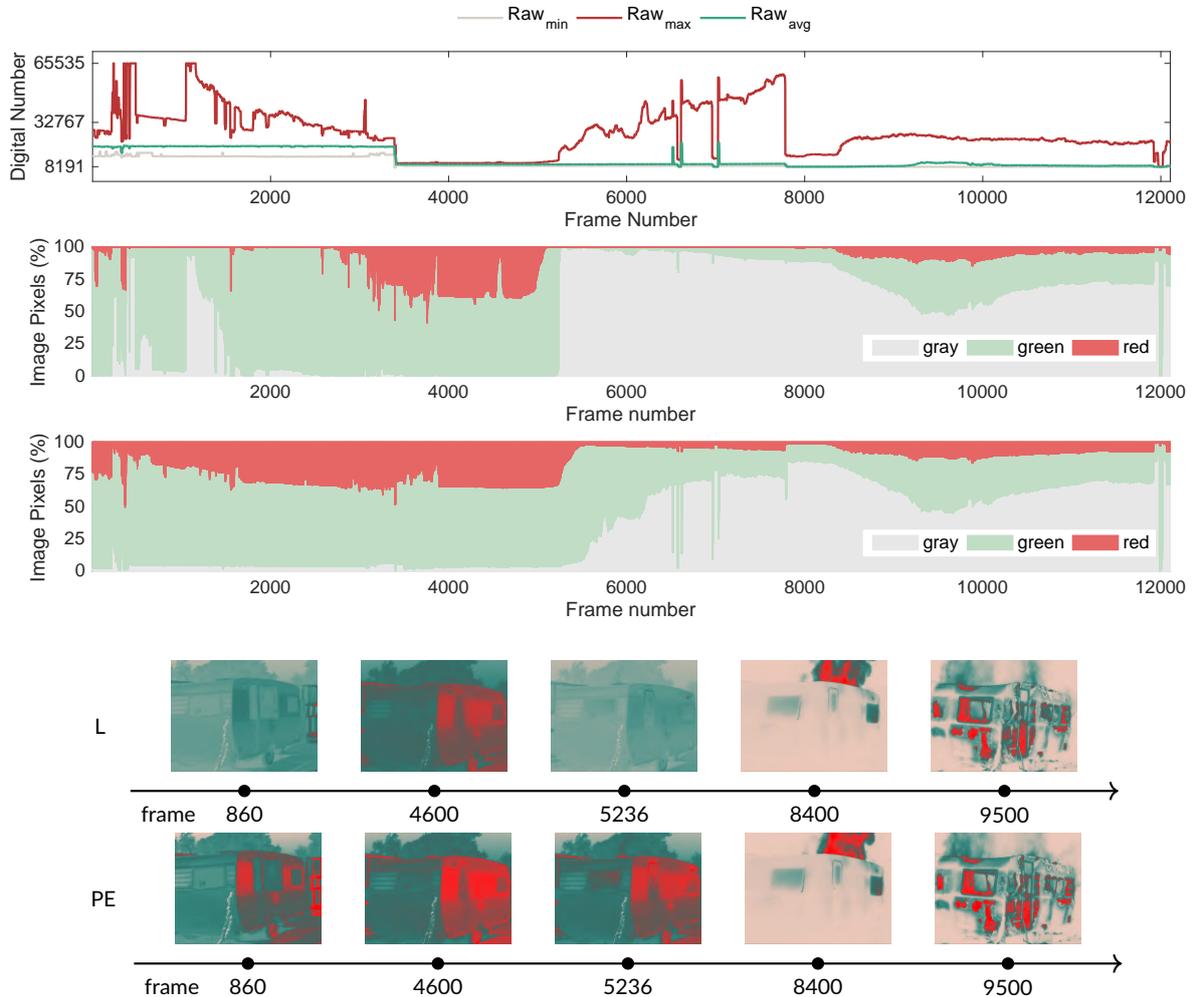


Figure 7.18: Filtered sensor response of thermal camera to a fire ignition inside the caravan: (a) raw variables; (b) color-based features with linear AGC (L); (c) color-based features with plateau equalization AGC (PE).

Observing Figure 7.18, the adjustment of the measurement range from $-40 \sim +550$ to the upper temperature limit of $+1500$ occurs at frame 3406 and has a prominent impact in the raw data, which can be identified by the steep drop in the raw features at frame 3406. Notably, the scaling of the bit resolution amplifies the measurement range but at first this results in a narrow distribution of the raw features at lower bit values.

However, as the fire starts around frame 5100, the dynamic range begins to expand. With the evolution of the intensity of the fire the dynamic range extends to comprehend higher values within the available resolution, reaching a maximum of 52343 at frame 7736, corresponding to a maximum raw value of 60978 and minimum raw value of 8635.

Subsequently, the sudden drop at frame 7782 indicates a second adjustment of the measurement range. The flashover point occurs at frame 8310 and as a result the fire starts burning the exterior of the caravan with greater intensity as can be seen in the corresponding frames (8400 and 9500). In addition, it can also be verified that with the fire visible in plain sight the sensor response tends to stabilize.

Concerning the color-based features, the effect of the fire starts to be detected by the sensor around frame 5236, as indicated by the increase in the gray pixel percentage, as had been verified in the previous test. This occurs precisely when the fire burns a small hole on the upper right side of the exterior of the caravan, as depicted in frame 5236. In frame 8400, it can be observed the effect after flashover, where the fire can be seen in plain sight, causing the contrast in the scene to considerably increase. Even with adjustments in the measurement range mid-trial the behavior of the color segmentation heuristic is generally stable.

Comparing the response in terms of color features for the linear (L) and plateau equalization (PE) algorithms, it becomes clear that the linear approach can detect the beginning of the fire in a prompt manner as can be observed by the abrupt increase of the percentage of gray pixels. Conversely, for plateau equalization algorithm the response occurs more gradually but it is still possible to detect the fire with this color segmentation heuristic. However, this difference in terms of image-based features is an important aspect to consider when developing intelligent fire detection systems to operate in real-time.

Following the analyses presented regarding tests performed in controlled environments, to validate the insights exposed, the next sections concern field trials undertaken in real contexts.

7.5.3 Summer Festival Trials

In the tests performed at the summer festival, FLIR Vue Pro was installed onboard a tethered helium balloon as illustrated previously in Figure 10.6. Although the movement of the balloon enabled surveying a wider area of the venue, for several periods of up to 2 hours long, given that this unactuated aerial platform moves according to the direction of the wind, the image capture is considerably less stable since the field of view of the camera drifts due to wind disturbances. Due to payload restrictions there were also no means of camera stabilization, therefore the sensor response for this set of data will inherently be more stochastic.

Notwithstanding, since the proposed feature engineering approach does not rely on temporal dependencies, but strictly features extracted from each frame, theoretically the proposed method is still viable for images acquired from mobile platforms. Hence, the analysis of these experiments is paramount to validate the applicability of this technique for integration in robotic perception pipelines for mobile platforms in the context of wildfire surveillance tasks.

To that end, from several tests carried out at different parts of the festival venue, two examples were selected to showcase representative situations encountered in real contexts. The first example, presented in Figure 7.19, comprises a test in which fire situations were not detected, whereas the second case,

depicted in Figure 7.20, covers areas without fire and also parts of the venue where fire is detected, namely in the community kitchens where rocket stoves were installed. The images were recorded with a 5 second time-lapse between frames at an altitude from the ground of approximately 80m.

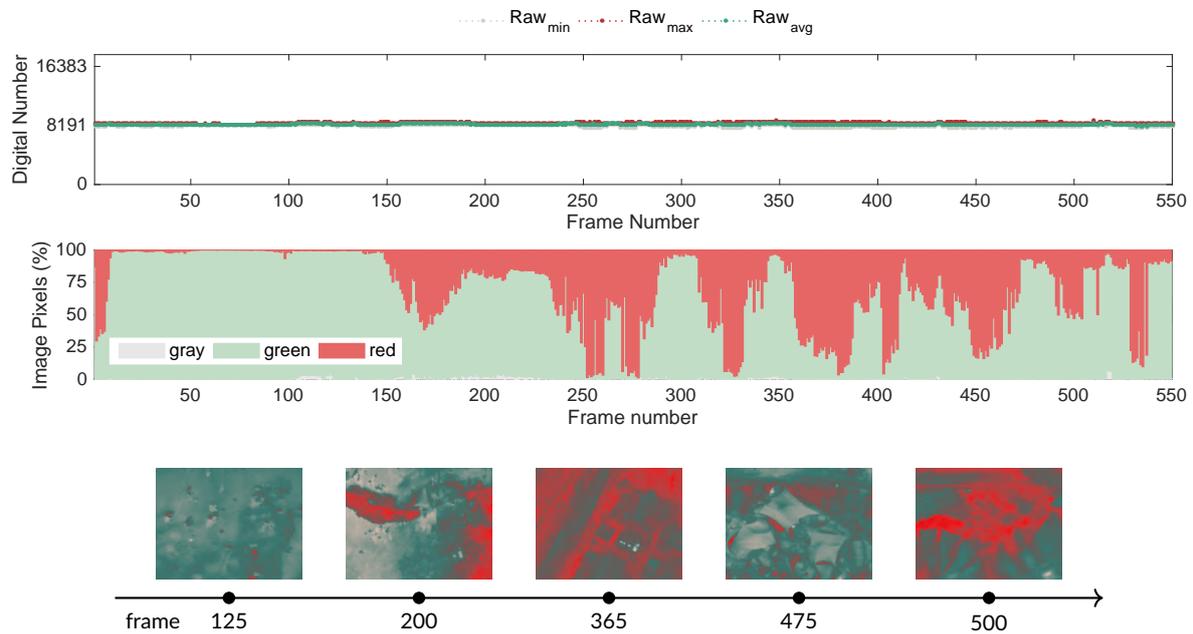


Figure 7.19: Results for data acquired surveying the festival venue.

Observing Figure 7.19 in regard to the evolution of the raw-based features, it becomes clear that under normal conditions the sensor of the camera does not exhibit significant variations. Considering that the measurement range would allow sensing incoming infrared radiation up to a dynamic range of 16383, i.e. the upper bound saturation level, the data demonstrates that in scenarios without fire, the maximum and minimum raw values do not exceed 8841 or go below 7812, respectively. Note also that in these trials ambient temperatures reach up to 36°C , so it is expected that the surface of objects captured in the field of view of the camera may exceed that value due to the different absorption properties of materials in such a heterogeneous context. Hence, including the raw-based features in detection and monitoring algorithms is rather promising to avoid false alarms.

Conversely, since the festival took place in the summer, it is also normal to capture objects for cooling purposes that are at temperatures below the average ambient temperature. While this does not seem to have a considerable effect in the magnitude of the raw values, this effect can be observed attending to the representation of the color-based features over time. More specifically, as different areas of the venue are surveyed, the color statistics vary considerably, influencing predominantly the red and green pixels percentages. For instance, attending to frame 365, it is possible to observe that a small set of items at cool temperatures cause the adaptation of the color representation, making the surroundings become depicted in red, even though there was not a sudden increase in temperature. In this way, since the histogram-based algorithms promote highly nonlinear effects in image data, color-based features require careful interpretation.

In turn, the sample frames presented also demonstrate that the appearance of red hotspots can be illusive,

as it does not necessarily mean that fire sources exist in the scene, but rather that these areas are slightly above the average temperature. The representation of the color features takes this into account since it considers in the color segmentation that the upper 45% of the scale is represented by red shaded tones, as illustrated in Figure 7.10. The uneven distribution of colors in the GrayRed color palette also justifies the behavior of the sensor mentioned in previous comment, namely regarding how the presence of cool objects in the scene easily turn red the most dominant color in the image.

Concerning the second case represented in Figure 7.20, which includes situations at the venue where fire is detected, the behavior of the sensor is consistent with the controlled fire scenarios previously discussed. In normal situations, the raw features follow the same behavior of the baseline case presented in Figure 7.19, but when covering the community kitchen area the dynamic range expands due to the presence of fire sources. More specifically, regarding the complete sequence, the maximum raw value reaches 12477, the minimum raw value was 7918 and the maximum amplitude of the dynamic range, recorded in a fire detection instance was 7290, which is well above the variations registered in the baseline case.

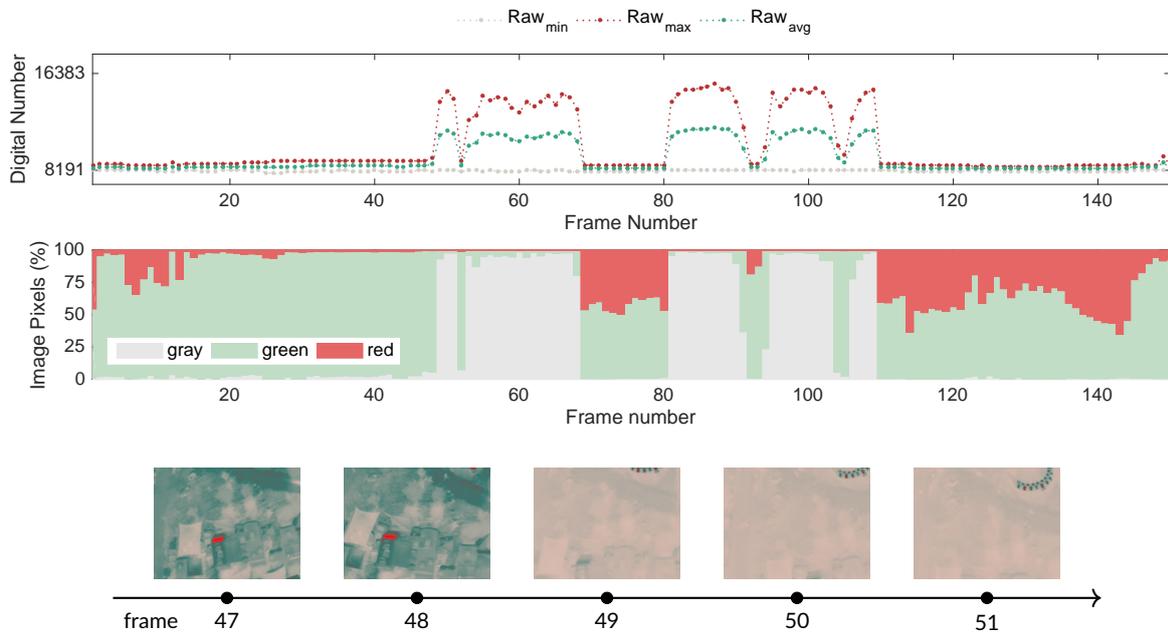


Figure 7.20: Results for data acquired surveying the community kitchens area.

Accordingly, the color representation also suffers an abrupt adjustment, with the surroundings becoming depicted mostly by gray pixels. This effect can be attested in the accompanying sample frames illustrating the first detection instance, captured while covering the community kitchens area. Since fire was strictly forbidden in the festival premises due to fire safety concerns, with exception for specific fires lit and controlled by the organization for scenic purposes and the community kitchens area, the remaining detection points in this sequence also concern the same area illustrated in Figure 7.20.

Considering that in these field trials the images were captured from a considerably longer distance than the laboratory tests, the data demonstrates that the color scale adapts even when hot objects have low spatial resolutions, which is essential to allow early fire detection in aerial surveillance scenarios. Therefore, confirming the validity of the proposed approach for fire detection using mobile platforms.

Furthermore, recalling that FLIR Vue Pro has a more limited measurement range, with an upper saturation level at 150°C, this may raise some concerns regarding false alarms. However, under real operation conditions no false detections were encountered.

7.5.4 Mountain Range Field Trials

In the field trials conducted in a mountain range, the images were captured from a longer distance than in the previous tests. The purpose of these tests is to evaluate a key issue regarding the use of compact thermal cameras, namely concerning its capabilities when operating outside the nominal conditions specified in the official hardware datasheets. The field trials performed allow the evaluation of the behavior of the camera in a long-distance image capture setup, to assess if the signal-to-noise ratio in this scenario is sufficient for fire detection applications. In this case, the tests were performed with FLIR Vue Pro configured to record images with a 1s time-lapse, at a distance of approximately 600m, as specified in the description of the experimental setup (Section 7.4.4).

From the several tests performed at the mountain range, two cases were selected for analysis, encompassing a situation with an active fire, which is presented in Figure 7.21, and another in the aftermath of a fire, depicted in Figure 7.22.

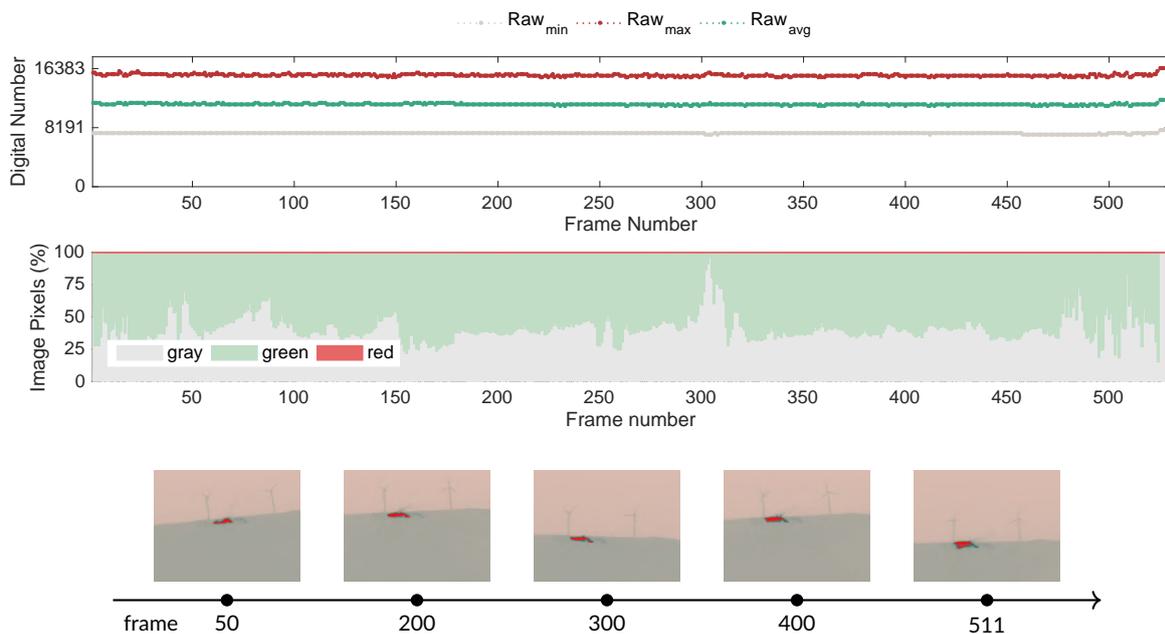


Figure 7.21: Results for fire detection at long-distance in a mountain range.

Regarding the monitoring of an active fire scenario, it can be observed in Figure 7.21 that the statistics of the raw variables do not exhibit particular variation, remaining with an extended dynamic range for the several minutes featured in this trial. Notably, despite the long-distance the sensed infrared radiation does not suggest a significant signal attenuation, as the sensor bit resolution is used to a great extent. Indeed, the maximum raw value registered is the sensor upper saturation level at 16383, while the minimum raw value recorded is 7235, in this case corresponding to pixels representing the sky, as can be confirmed in

the sample frames.

For this case, the evolution of the color-based features does not reveal particular changes and is consistent with the insights exposed in the previous trials. The sensor response stays generally constant with slight variations attributed to the change of field of view of the camera. The gray and green levels dominate the image context and a reduced pixel percentage of red pixels depicts the area of the fire as can be observed in the sequence of frames selected.

Although FLIR Vue Pro was not specifically designed for this type of task, these results validate the capability to detect fire sources at long-distance. In turn, an important aspect to take into account when considering extended distances is that due to the limited resolution of the focal plane array, and consequently the output image resolution, with the increase in distance the minimum size of the fire that can be detected naturally increases also.

Since wildfires often extend over large periods and can be inactive after firefighting operations take place but re-ignite before the fire event is completely extinct, the monitoring of areas in the fire perimeter is important to prevent rekindles. In this context, the ground is usually at higher temperatures, adding difficulty in pinpointing potential reignition spots.

The following test performed in the field trials targeting the aftermath of a controlled fire, which is depicted in Figure 7.22, aims to uncover how these circumstances differ from the previous cases studied.

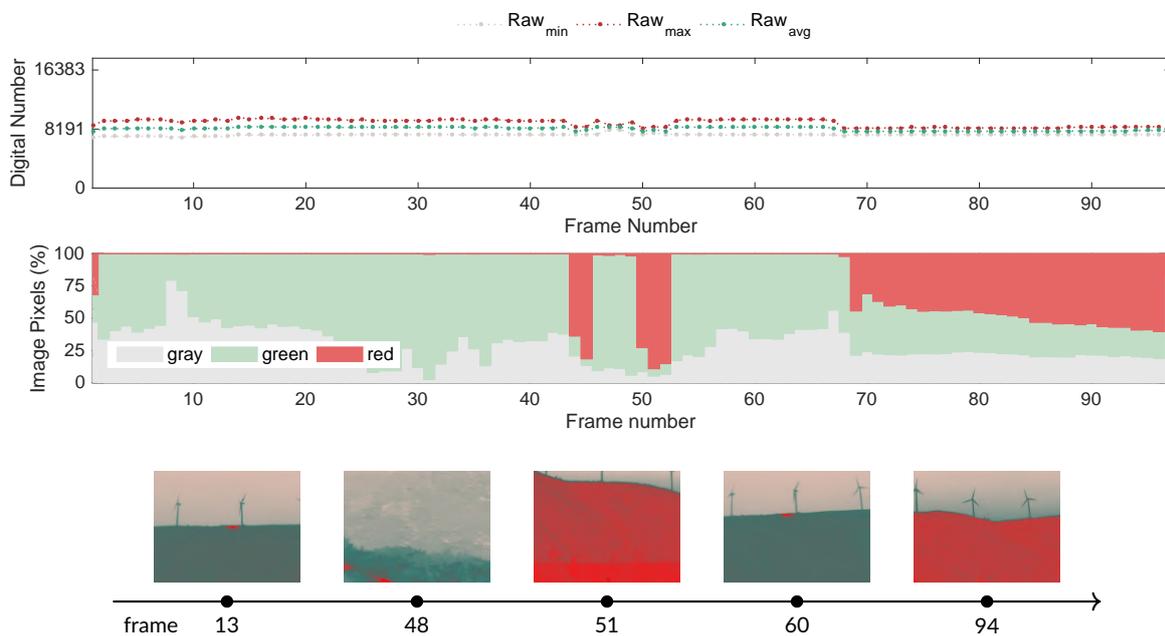


Figure 7.22: Results for fire aftermath captured at long-distance in a mountain range.

Observing the time-series of the raw-based features, it is immediate to notice that the dynamic range is reduced once again, but with a visible gap between minimum average and maximum values, unlike the baseline case encountered in the festival trials. In contrast, in this case the maximum raw value extends as high as 9782, the minimum raw value remains around 8141, and when the dynamic range is greater at 2426, the average raw value is 8569.

Concerning the adaptation of the color-based features it should be noted that the field of view of the camera was changed considerably to capture distinct parts of the monitoring area, as illustrated by the set of sample frames presented in Figure 7.22. In this case, the green and gray percentages are predominant when capturing the burned area, which is depicted in red in frames 13 and 60 because it is at a higher temperature than the surroundings. The sky area remains depicted in gray, however, in contrast with the situation with an active fire, the ground is represented by a considerable darker tone of green. In that sense, the color-based features alone do not evidence a difference from the previous case, which highlights that the color segmentation heuristic has to be fine-tuned depending on the image capture setup, or defined with more restrictive gamma for the green class. Interestingly, when the burned area is not in the field of view of the camera, the ground is generally depicted in red, as can be observed in frames 51 or 94.

In turn, when facing downward on the valley, as illustrated in frame 48, the distance influences the color representation, with closer objects represented by the bottom part of the image, and the gray part concerning the mountain on the other side of the valley. This result shows that in addition to being capable of detecting fire at long-distance, the thermal sensitivity of this sensor is also able to convey depth perception. While this does not play a significant role in the detection in itself, it has interesting implications for autonomous navigation by robotic platforms.

7.5.5 Discussion

From the analyses presented throughout this study it becomes evident that thermal cameras can potentially bring improved reliability for fire detection and monitoring systems on several aspects in which visible range systems have limitations. On the one hand, the proposed data-driven feature engineering approach that combines expert knowledge on the sensor working principles with color-based features derived from processed image data demonstrated to be robust to false alarms, even when tested in highly dynamic environments. On the other hand, it was possible to validate the applicability of commercially-off-the-shelf camera models in long-distance scenarios for fire detection applications, as verified the compact onboard thermal camera.

However, adapting thermal imaging cameras for wildfire detection and monitoring systems entails an adequate understanding of the application requirements and capabilities of each device, which changes depending on manufacturers and camera models. Furthermore, considering UAV-based applications, the integration of this type of devices in robotic perception frameworks is not straightforward, requiring knowledge of the underlying processing image processing algorithms, which are highly nonlinear. For this reason, based on insights gained from the comprehensive analyses presented several remarks deserve to be highlighted.

This work explored a series of automatic gain control methods available in the firmware and software of thermal cameras, which are based on histogram-based algorithms for data compression and contrast enhancement. In particular, testing the applicability of several alternatives in fire scenarios, the results revealed that advanced contrast enhancement techniques degrade the ability to distinguish the fire from

the surroundings, preventing an accurate identification of the burning area. Indeed, the linear approach registered the best results, as identified in Figure 7.7.

Since histogram-based algorithms promote highly nonlinear effects in image data, color-based features require careful interpretation, because tracking the hottest areas can be misleading depending on the scene context. In that sense, extending state-of-the-art color-based approaches [104] by including also raw-based data provides enhanced situation-awareness, thus potentially enabling improvements in robotic perception.

By testing the proposed method in distinct scenarios, under static image acquisition conditions and also from onboard mobile platforms (e.g. an tethered helium balloon), it was possible to validate the approach in both contexts. Nonetheless, for integration in robotic platforms it is advisable to employ means of camera stabilization to minimize the effect of the external disturbances due to the movement of the platform and consequent changes in the field of view of the camera.

In turn, the operating distance conditions of these sensors play an important role when applying heuristic methods and as such the performance shall be fine-tuned depending on the image capture setup. Moreover, the characterization of the device in terms of sensor response is fundamental for this type of application, because although manufacturers define nominal operating conditions, since fire is a source of extreme temperatures, it is possible to successfully detect fire at long distances as well, since a high signal-to-noise ratio is preserved.

Regarding the integration in robotic perception pipelines for autonomous navigation of UAVs, it was also verified that the thermal sensitivity of these sensors is also able to convey depth perception. However, in order to apply the methodology proposed in this work, the camera model has to provide access to the raw data online, so that it can be incorporated and used in the image processing pipeline of the robotic perception framework.

7.6 Conclusion

The study conducted in this work provides a comprehensive introduction to adapting thermal imaging sensors for wildfire detecting and monitoring systems. Since these phenomena are increasing in intensity and frequency there is increased urgency in the early detection, monitoring and surveillance of these events. For these reasons, leveraging mobile robotics in this type of extreme environments requires devising application-specific techniques.

For that purpose, this paper explored the foundations of image processing of thermal imaging, covering the most used automatic gain control methods and the color mapping schemas responsible for generating thermal images encoded in pseudocolor.

The proposed data-driven feature engineering approach combining expert knowledge about thermal imaging sensors as well as color-based segmentation heuristics offers improvements in situation-awareness in comparison to existing methods focusing only on one data type.

The study of the response of thermal imaging sensor in fire scenarios as well as situations without fire considering raw sensor data as well as color-based features revealed fundamental to provide a better intuition of the working-mechanisms of this sensors and the expected behavior under such conditions. By evaluating how the features evolve over time, these analyses provided a clear understanding of the nonlinear adaptive color scale, providing important insights relative to the sensing capabilities of commercially-off-the-shelf thermal cameras. Thus, establishing a solid foundation for the development of robotic perception algorithms.

Considering wildfire detection and monitoring are safety-critical applications, there is a great benefit in exploring highly interpretable feature engineering approaches, in order to subsequently combine these methods with intelligent soft sensor approaches, enabling better model understanding and interpretation.

8 | MAVFire: Multimodal Dataset

Contents

8.1	Introduction	114
8.2	Related Work	115
8.3	Data Collection	116
8.4	MAVFire Dataset	118
8.5	Dataset Applications and Limitations	121
8.6	Conclusion	122

The material included in this chapter is part of a work in preparation to be submitted as a research article to an international peer-reviewed journal.

In Preparation:

M. J. Sousa, A. Moutinho, M. Almeida, **MAVFire: Thermal-visual-inertial-GNSS dataset for aerial robotics and fire applications.**

8.1 Introduction

The operational value of aerial means in wildfire support operations has motivated the increased use of unmanned aerial vehicles for applications, e.g., aerial surveillance, fire detection, and mapping of fire fronts [112]. Due to their low cost and high maneuverability, micro-aerial vehicles (MAVs) are becoming an emerging solution to address these tasks, providing significant lower risks than manned aircraft [113] while enabling a higher decentralization of efforts, which is a prime concern for extensive area coverage.

Being fire events hazardous, highly dynamic, environments, flying aerial robots in these contexts poses tremendous challenges, making robot autonomy a prevalent matter, not only for the automation of safety-critical tasks but foremost due to the inherent difficulties in performing complex missions in adverse conditions.

MAVs navigating in these contexts often face paradoxical scenarios because although fire events can be identified either by detecting fire or smoke, e.g., in long-distance operations or fire front mapping, the latter often prevents the identification of the former in the visible spectrum [112, 114]. To address this issue, with the developments in lightweight, compact thermal cameras, LWIR sensors are starting to become increasingly used onboard remotely controlled MAVs in fire support operations. However, due to the higher costs of these devices, comprehensive datasets with thermal images are still not widely disseminated. Although there is emerging research in aerial pedestrian tracking [115], wildlife monitoring [116], or precision agriculture [117], the thermal data characteristics for these use-cases differ considerably from fire applications. Most uses of thermal cameras deal with identifying small temperature gradients and enhancing these relative radiometric differences with high contrast, a task thermal cameras excel at owing to the employed automatic gain control (AGC) algorithms [108]. However, in fire scenarios there are high temperature gradients which result in high variability and extreme values of sensor response, normally reaching saturation levels of these sensors [118].

Towards increasing the levels of robot autonomy, perception systems need to be designed considering these specific data characteristics to ensure robust deployments. However, building adequate test datasets is difficult, time-consuming, and expensive to collect. To address this gap, this article introduces the MAV-Fire dataset, a multi-view thermal-visual-inertial-GNSS dataset for MAV-based fire applications.

Our data collection efforts in building this field robotics dataset aim to contribute to support and further the research and development of autonomous robotics solutions, namely for fire support operations. The main contributions of the present dataset are threefold:

- **multi-sensor data** encompassing thermal radiometric and AGC-based RGB encoded thermal data, visual images, inertial measurement units (IMU), and GNSS receivers (one RTK), with variable sampling frequencies;
- **multi-view data** acquired from an airborne vehicle, namely a MAV flying at a comprehensive range of flight profiles, and two ground stations, that allow the further derivation of fire characteristics and its relation to aerial-captured data in different modalities;

- novel **multimodal dataset** for MAV-based robotics and fire applications based on field robotics experiments that can empower the development of robust robotic navigation solutions and use-cases in fire support operations.

The rest of this paper is structured as follows. Section 8.2 presents an overview of related work. Section 8.3 describes the data acquisition systems and experiments. Section 8.4 covers the detailed description of the proposed MAVFire dataset. In Section 8.5, we reflect on possible applications of this dataset and discuss known limitations, and Section 8.6 provides concluding remarks and reflects on directions of future work.

8.2 Related Work

This section discusses MAV-based advances in fire applications and delves into thermal-enabled localization and mapping for aerial robots. This is followed by a more narrow discussion focused on multi-sensor calibration.

Fire Applications: Detection, Monitoring, Mapping

MAV-based fire applications have notably been an active field of interest over the years, with research spanning from the development of automatic perception systems for fire detection, monitoring [119], or fire front mapping [102], to cooperative strategies for multi-vehicle missions [120–122]. Although previous efforts have focused primarily on evolving fire situational awareness [123], on-board processing solutions and decision autonomy for guidance and navigation are generally still underdeveloped in field deployments. The progress in cognitive robotics promises to enable missions with a greater degree of robot autonomy with active perception in sensor-driven planning, and risk-aware systems.

Thermal-enabled Localization and Mapping

Towards autonomous navigation in field robotics, perception systems have evolved to perform VIO and VSLAM in generic environments [124–127]. However, in many situations where risk to humans motivates robotic solutions, the operating conditions can be very challenging, which can degrade the performance significantly. Multimodal perception frameworks leveraging thermal LWIR data have been proposed to tackle low illumination conditions [128–130] and the presence of visual obscurants [131, 132], such as in night operations or in search and rescue scenarios in underground mines.

Although LWIR sensors are growing in applicability, open datasets containing thermal data for field robotics are still scarce and designed for specific use-cases [116, 133, 134]. Moreover, thermal radiometric data are not always distributed or representative of the full range of operation of these sensors, hindering its use for applications that deal with extreme radiation sources such as fire scenarios. Additionally, with the fire inducing sensor saturation, AGC algorithms stretch the pseudocolor RGB encoding over the full extent of the dynamic range of the 14-bit or 16-bit resolution. Conversely, the high compression to 8-bit range

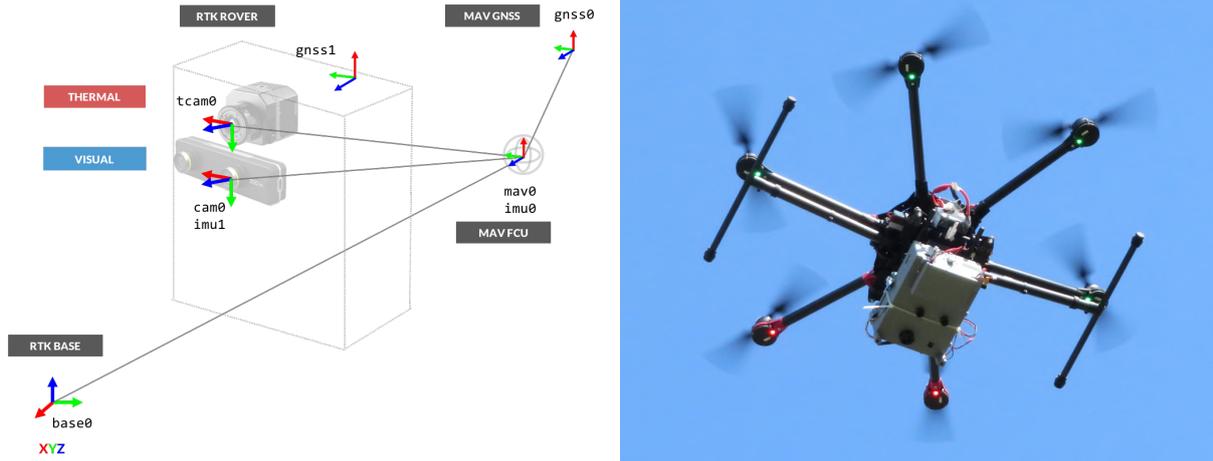


Figure 8.1: **MAV-based data collection.** The multimodal sensor suite schematic (left) showcases sensor coordinate frames and extrinsic relations – each frame is labelled with the respective identifier, e.g. $mav0$, $cam0$, $imu0$, etc.). Airborne sensor suite onboard the drone (right).

causes severe feature loss in image content, removing identifiable landmarks from the image [118]. To address this issue, since the estimation of accurate robot localization and derivation of the environment structure is critical to map the fire location and spread, this work proposes a novel thermal-visual-inertial-GNSS dataset that enables harnessing the advantages of the several sensor modalities for improved reliability.

Multi-sensor Calibration

To combine different sensor modalities, it is important to align asynchronous measurements with distinct sampling frequencies and to perform spatio-temporal calibration between the multiple sensing systems. Although multi-sensor calibration and time synchronization have been extensively studied, traditional calibration methods using noninformative algorithms often require previous knowledge of the operation environment and are generally time-consuming, onerous and cumbersome to perform in the field [133, 135] in a pre-deployment stage. These processes become increasingly unfeasible when including additional sensing stations/vehicles that would require lengthy calibration procedures. Furthermore, since calibration priors accumulate significant error over time [136], the validity of such would degrade considerably, thus not justifying the repetition of these procedures to achieve a suboptimal result. Motivated by this, in our dataset we leave multi-sensor calibration to be further explored with data-driven calibration methods.

8.3 Data Collection

For data acquisition, a set of field tests was performed to collect a comprehensive dataset of multimodal aerial data during several MAV flights over controlled burns, which was complemented with additional views from ground-based stations. The experiments aim to aggregate a series of scenarios relevant for MAV-based fire applications. These trials were performed with specific known conditions, to allow an accurate evaluation of how robotic perception approaches cope with these challenging scenarios. In that

sense, these trials allow characterizing the sensor response in fire scenarios and assessing the impact of the altitude on spatial resolution and on visual and thermal data measurements. The next sections present the details of the data collection systems and field tests performed.

Systems Overview

Controlled burns field tests were performed in an airfield. A commercially available MAV platform was used, namely DJI S900 hexarotor, equipped with a Pixhawk 2.1 Cube Black flight control unit (FCU) and a GNSS receiver, connected to the ground control station via telemetry radio. To achieve the goals of the field tests, the MAV was teleoperated via radio control to ensure the desired altitude levels were reached and that an immediate controlled landing could be performed if required by the airfield control tower. Onboard data collection included all flight log variables recorded by the FCU. For the MAVFire dataset, the IMU and GNSS data were selected, each with the respective sampling of 25 Hz and 5 Hz.

The aerial sensor payload was designed as an independent unit to be carried on-board the multicopter in a leveled, downward facing pose (Fig. 8.1). The payload includes a Jetson Nano on-board computer, connected via USB3.0 to a Stereolabs ZED Mini stereo visual range camera with internal IMU, sampled at 15 Hz and 500 Hz, respectively. Note that the right-hand side camera of the visible stereo-pair was not recorded. For thermal data acquisition a FLIR Vue Pro camera was used. The thermal camera was juxtaposed with the Stereolabs camera, geometrically aligned with the left camera of the stereo pair and captured images at 5 Hz. For a benchmarking test, this camera was connected to a Pixhawk FCU to record synchronized attitude and GNSS-RTK georeferencing data, with a 5 Hz sampling frequency.

In addition to the multimodal payload onboard the MAV, two stationary ground stations were equipped with thermal and visual range cameras, one station at ground level and the second placed on an elevated platform. The visual sensors used are generic video cameras, namely Sony FDR-AX53 and Sony HXR-NX30F, recording at 24 Hz and 25 Hz, respectively. For thermal data acquisition, the ground stations used FLIR SC640 and SC660 cameras, both connected independently to desktop computers with the required proprietary data acquisition software (FLIR ResarchIR), recording at 15 Hz.

Experimental Setup

The experimental trials conducted in a controlled airfield allowed the MAV to safely perform a wide variety of altitude profiles and horizontal paths that would allow testing the multiple sensors. Several flights exceeded the regulated 120 m altitude, being the maximum altitude reached ≈ 250 m. Given the complementary nature of multimodal sensors, it was particularly important to design experimental setups that allowed to collect representative data of challenging instances to enable the understanding of how different data modes can potentially influence the perception element in autonomous robotics tasks.

To complement the data acquired with the hexarotor depicted in Fig. 8.1, the two additional ground stations were placed around the working area where the fires were being set up, adding two additional viewpoints equipped with both thermal and visible range cameras. The position of such ground stations is



Figure 8.2: Data collection in field experiments from multimodal systems, including airborne and ground stations.

illustrated in Fig. 8.2, respectively in the left and right side of the image demonstrating the working area. The two ground stations were installed at closer distances to the controlled burns to capture relevant fire characteristics that can be derived using the nearby vertical markers.

For each trial, one or multiple fire ignitions were held using a mixture of dry shrubs as fuel (moisture content $\approx 12\%$), under favorable weather conditions with high visibility and minor wind influence (average 1m/s). The experiments were carried out so that during each trial the fire had various specific known perimeters and flame heights and lengths, enforced by physical constraints and by refueling as needed, respectively. The set of experiments comprehends small fires with a wide variety of spatial and temporal characteristics.

8.4 MAVFire Dataset

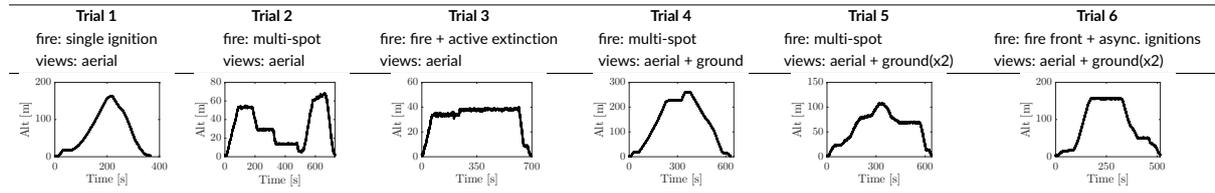
Dataset Description

The MAVFire dataset results from a set of six independent experimental trials. These experiments were performed over the course of two days and at different moments of the day. As described in the previous section, we employed a total of three independent multimodal sensor systems, one placed on the MAV (`mav0`) for aerial data acquisition and two stationary ground stations (`gs1` and `gs2`). Note that the order by which these experiments were performed does not relate to the order by which they appear in this text, nor with their numeric identifier.

The experiments conducted comprise a comprehensive array of situations as is summarized in Table 8.1, e.g., single fire ignitions, multiple fires, with synchronous and asynchronous ignitions, as well as varied

spatial configurations. In addition, the flights performed allowed the introduction of variability in the altitude profiles (showcased in Table 8.1) but also purposeful movement on the horizontal plane.

Table 8.1: Summary of experimental trials featuring key characteristics and flight altitude profiles.



A fundamental goal of the experiments was to characterize the multimodal sensor responses in multiple relevant conditions while onboard the MAV (see Table 8.1, trials 1, 2, and 3). The tests then evolve to more complex scenarios (see Table 8.1, trials 4, 5, and 6), which in addition to the MAV aerial view also allow the introduction of multi-view localization and mapping problems, by the addition of the ground stations data. The complete summary of which sensor data are provided for each dataset sequence is compiled in Table 8.2, where the sensor identifiers relate to the devices presented in Section 8.3.

Table 8.2: Summary of data provided for each trial.

	mav0						gs1		gs2			
	gnss0	imu0	cam0	imu1	tcam0	gnss1	cam1	tcam1	cam2	tcam2	gnss2	imu2
Trial 1	✓	✓	✓	✓	✓							
Trial 2	✓	✓	✓	✓	✓							
Trial 3	✓	✓	✓	✓	✓	✓						
Trial 4	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
Trial 5	✓	✓	✓	✓	✓				✓	✓	✓	✓
Trial 6	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓

Data Modalities

The cornerstones of this dataset with respect to the state-of-the-art are its rich thermal and visual data elements, which combined with GNSS and inertial data can allow significant improvements in navigation and mapping tasks since the perception layer is crucial for both. However, significant progress has yet to be made particularly in integration of thermal cameras. As such, to promote further avenues of research, the raw data are included. For data accessibility reasons, we also provide two pre-processed RGB encoded modalities of the raw thermal frames. Fig. 8.3 showcases a set of samples to exemplify the distinctions between the data types: visual and thermal RGB-encoded modalities. To that end, the raw thermal data were processed with two distinct automatic gain control algorithms and encoded with two different color palettes. The first uses a linear AGC algorithm and applies a divergent color palette (greyred) for RGB encoding. The second employs a plateau equalization algorithm for AGC, followed by application of a sequential and monochromatic grayscale color palette. For a fully-detailed, in-depth explanation of these procedures please refer to [118, 137], which also demonstrates the effectiveness of the linear + greyred approach for fire applications. Conversely, since the grayscale approach is widely used for encoding ther-

mal data used in robotics applications, it justifies its inclusion in the dataset as its applications extend beyond fire-tasks.



Figure 8.3: **Dataset example:** Samples from trial 6 featuring three different view points. The trial starts with a fire front and multiple additional fire spots are ignited asynchronously. The top row illustrates visible range data, whereas the bottom row, juxtaposes for comparison the two rgb encoded thermal frames. Note that the grayscale sample from the aerial view clearly depicts the fires as seemingly larger in spatial size.

Data Formats

The heterogeneous sensors used for data collection yielded data in a wide variety of formats. This results in the need for a series of data retrieval and pre-processing procedures to generate a uniformly formatted dataset. Since this dataset can be of use to several communities, e.g., ranging from robotics and artificial intelligence to fire domain fields, it is important that the available data uses common formats, e.g., `png`, `tiff`, and `csv`. All data pertaining to relevant device configurations, sensor parameters or general trial information, are encoded in `yaml` files for convenience of importing such data in subsequent post-processing or development pipelines. All time-series data are provided in `data.csv` files for each sensor, where the first column stores the timestamp in Unix time in second units but encoded with higher sub-second precision. The variables are generally in SI units, unless specified otherwise, e.g., in pose or orientation data. For image-based data, all files are named with the respective timestamp in Unix time in second units with respect to the device clock. This means that two sensors, e.g., `cam0` and `tcam0`, can have image files with the same name. However, one should not assume *a priori* these are from the same time instant, as sensors are not synchronized. If this is relevant for the application, appropriate temporal sensor calibration to estimate the appropriate timing offsets between data sources should be conducted. The files related to each sensor are stored separately to promote the modularity within the dataset, but are also compiled on a system-level basis in `bag` files.

Dataset Structure

The dataset is organized by trial, comprising six distinct data sequences. For each, there are subfolders for each `<system_id>` (e.g., `mav0`, `gs0`), under which the sensor subfolders are organized (e.g., `cam0`, `tcam0`, `imu0`, etc.). Within each sensor subfolder there are `csv` and `yaml` as previously outlined. For camera sen-

sors (e.g., `cam0`, `tcam0`), the raw image data are provided in `raw_frames` subfolders, in their respective formats, as well as a `raw_frames.csv`. Additionally, thermal data which involved a further processing step described in Section 8.4, features the preprocessed data in subfolders with the following name formatting, `encoded_<agc_alg>_<color_palette>`, yielding then:

- `[raw_frames]`
- `[encoded_linear_greyred]`
- `[encoded_plateau_eq_grayscale]`
- `[raw_frames.csv]`

Besides the information contained in the dataset directory, for each `<system_id>` a bag file is provided.

8.5 Dataset Applications and Limitations

While the MAVFire dataset aims to address an existing gap in field robotics research for fire applications, and more generally in thermal-enabled robotic perception, the use of this resource can extend to a variety of subtopics in robotics research and beyond.

(Multi-)Sensor Calibration

By providing the raw data without pre-processing it with a given multi-sensor calibration approach, this dataset can support the development and benchmark of novel approaches to problems in this realm, from the improvement of single sensor calibration approaches that can have a great impact in improving image registration and georeferencing, to the calibration of the full array of sensors presented.

Robot Robust Navigation

As noted in the introduction, multimodal perception is increasingly seen as a pathway towards achieving robust navigation and dealing with challenging and hazardous applications. However, dealing with specialized sensors, be them thermal, or multispectral in general, results in different challenges as the field is still in an early stage. With the decreasing costs of these technologies and its uptake in robotics research, by providing situations with extreme sensor measurements, this dataset can play a role in increasing the robustness of robot navigation.

Fire Detection, Localization, and Mapping

To extend the use of the dataset beyond specific robotics inner-workings, the dataset was built with higher resolution than possibly was needed for perception in visible range. That option emerged from recognizing that the image content includes a diverse set of elements, e.g., fire, smoke, car, truck, human/operators, road, vegetation, which are relevant for data annotation tasks or could be useful also for developing intelligent algorithms for other robot-assisted applications involving detection, localization and mapping.

Known Issues and Future Directions

Although comprehensive for multimodal perception in general, the current dataset still lacks some instances with higher scale of the scenarios as the intent is its application to fire support applications. However, the sensors used saturate in these conditions, so a similar behaviour is expected in scenarios with larger burns and higher fire intensity.

Concerning data acquisition, a key takeaway from our experiments was that if we wanted to pursue higher sampling data collection, higher processing power from the on-board computer and on-board storage units would be necessary. In that regard, some missing data is to be expected in the data sequences, as sometimes the writing speed and recording processes do not complete fast enough to ensure perfectly sampled sequences.

8.6 Conclusion

This paper presents the MAVFire dataset, an aerial thermal-visual-inertial-GNSS dataset for MAV-based robotics and fire applications, envisioned to support the advancement on intelligent robot autonomy, particularly in this domain. The multimodal dataset was created from controlled burns field experiments comprising several flight profiles and relevant fire scenarios, designed to aggregate a series of sequences that presented challenges for robotics perception pipelines.

With the growing interest in leveraging aerial robotics for wildfire support operations, we believe that the release of this dataset is an important stepping stone towards robot autonomy in these environmentally challenging scenarios and will be of great value for testing and evaluating different algorithms for field robotics deployments.

IV | Intelligent Fire Detection and Monitoring

Contents

9 Wildfire Detection Using Transfer Learning	125
10 Fire Monitoring with Fuzzy Models	151

Summary

This part focuses on intelligent systems approaches for wildfire detection and monitoring, and is organized in two independent chapters exploring two distinct data-driven methods. Chapter 9 explores a wildfire detection solution for visible range data based on deep neural networks that harnesses a transfer learning approach on augmented datasets. In turn, Chapter 10 is dedicated to devising intelligent techniques for fire monitoring using thermal data, namely using a data-driven fuzzy modeling approach based on fuzzy clustering.

9 | Fire Detection with Transfer Learning

Contents

9.1	Introduction	126
9.2	Deep Neural Networks	130
9.3	Databases	134
9.4	Methodology	135
9.5	Results and discussion	140
9.6	Conclusion	148

The material included in this chapter was previously featured in a peer-reviewed journal article, published in *Expert Systems with Applications*, Elsevier.

Published in:

M. J. Sousa, A. Moutinho, M. Almeida, **Wildfire detection using transfer learning on augmented datasets.** *Expert Systems with Applications.* 142 (2020). doi: 10.1016/j.eswa.2019.112975

9.1 Introduction

Over the years, with the evolution of remote sensing technology, there has been continued interest in applying computer vision methods to fire hazard identification and risk assessment research. The applications include, e.g., forestry inventory [138], fuel mapping [139], burned area maps [140], and wildfire detection [141].

Recently, in the wake of the advances in the field of artificial neural networks, deep neural models have become the state-of-the-art in a variety of computer vision problems. The current success of this type of architectures in highly complex tasks has opened up the range of their potential use cases and is paving the way for the application of these models to real-world challenges.

The renewed interest in deep learning architectures such as convolutional neural networks (CNNs) came as a result of two determinant factors. First, storing data has become inexpensive and is no longer a twenty-first century problem. On the other hand, it is not humanly possible to perform data analysis on all data without using artificial intelligence. Second, a process as data-intensive as image processing using CNNs became viable by the parallel computation power provided by graphics processing units, which accelerate the learning and inference stages.

However, although deep learning models have achieved state-of-the-art results in a variety of closed-set problems, the application of this approach to open-set problems is still in early stages. More specifically, a known drawback of this approach concerns requiring large datasets for model development, hence the limited amount of data for specific applications is a common hurdle in applying deep learning approaches to real problems.

In this context, several recent works propose the use of these architectures for early fire detection, which is an application with high potential benefits as fire hazard has severe impacts on a global scale. Conversely, image databases for fire detection for benchmark is still an evolving topic, since there is not yet a large-scale dataset for this problem, but there is a growing interest in that direction namely with several contributions in the literature on a smaller scale and open or public-domain sources.

The main objective of this investigation is to further the development of expert systems for early detection of fire events. In that sense, this work centers on exploring a transfer learning technique based on deep neural networks, to handle the limitations of current open-source databases for this application. Being wildfire detection a time-critical application understanding model limitations is paramount for the implementation of intelligent systems in real contexts. For this reason, this article undertakes an in-depth study of the challenges facing the application of this type of models, namely in wildland-urban interface regions.

Since the rapid detection of an original fire event ignition significantly increases the efficiency of a first intervention, early fire detection systems can help avoid the occurrence of large burnt areas. In this way, the potential impact of the integration of intelligent algorithms in decision support systems for firefighting and civil protection as well as fire self-protection systems with automatic detection can have a significant contribution to mitigate the social, cultural, environmental and economic effects associated with wildfires.

9.1.1 Wildfires in Wildland-Urban-Interface areas

Wildfires are recurrent natural disasters, which have a brutal impact on the environment and natural ecosystems. These phenomena have drastic effects on communities on social and economic levels, since these can lead to the loss of lives and material damages.

With the increasing effects of climate change, the conditions for fast initial spread rate are more frequent, driving a rise in the number of events that escape to the initial firefighting. These conditions also aggravate the severity of fire events, leading to extensive spread with multiple fire fronts and episodes of extreme fire behavior.

Recently, massive wildfires hit different regions across the globe, with a particular incidence in wildland-urban-interface (WUI) areas. At the end of 2017, a series of fires struck Southern California, with over 100 kha burned. The Thomas Fire advanced for over a month, only being extinguished after 39 days, and accounting for 114,078 ha of area burned [142]. In 2018, the situation was similar in Northern California with fire events such as the Mendocino Complex Fire reaching almost 200 kha [143].

In Portugal, the scenario was likewise dramatic with more than 250 kha of area burned and an unprecedented number of fatalities with more than 115 deaths, only in two fire occasions - Pedrógão Grande in 17Jun2017 [144] and Fire Events of 15Oct2017 [145]. To illustrate the extent of the affected areas, Fig. 9.1a depicts the fire hazard report released by ICNF (Institute of Conservation of Nature and Forest) in February of 2017 [146], whereas Fig. 9.1b illustrates the burned area over the course of the same year [147].

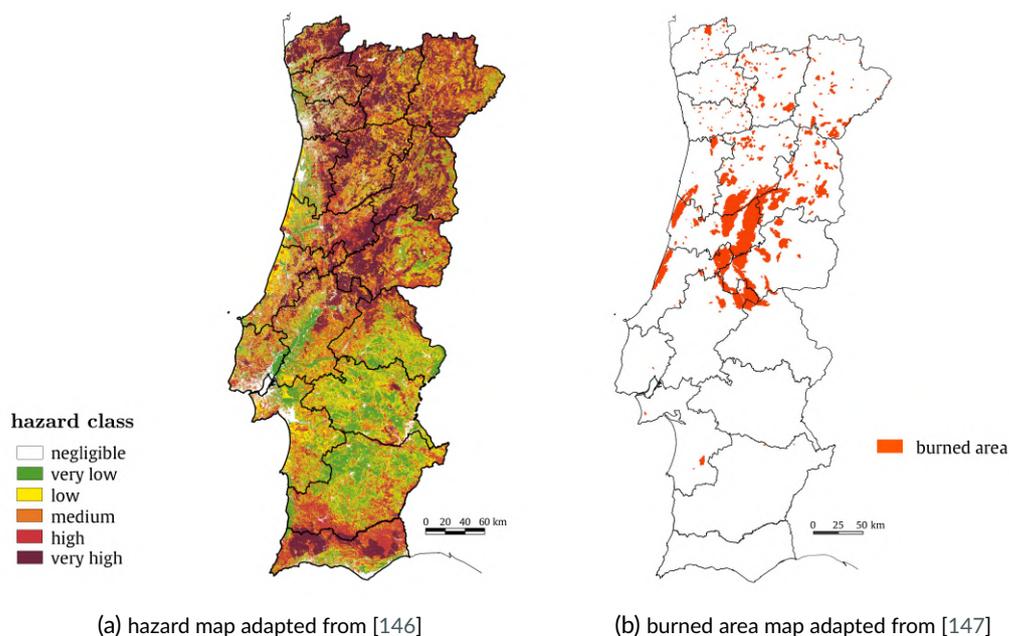


Figure 9.1: Comparison between fire hazard assessment and burned area maps.

Despite the identification of areas with elevated fire hazard, and meteorological reports warning of time periods of critical danger for the occurrence of fire events, the extensive devastation caused by these incidents is of major concern, as Fig. 9.1b demonstrates. Since a fire occurrence is by nature a process

with a high degree of uncertainty, both on spatial and temporal levels, it is difficult to pinpoint the ignition locations in a short time-frame. Therefore, the communication of the detection of fire outbreaks is key to allow a prompt action from emergency response teams on the ground and minimizing its consequences [148, 149]. In that sense, since the greater concern lies in WUI areas, the focus of this article is on vision-based fire detection in rural environments.

9.1.2 Related work

Image-based algorithms for fire detection have been extensively explored in the computer vision domain. There is a plethora of different approaches that propose color models for flame pixel classification using RGB [150], YCbCr [151], HSI [152] or the YUV [153] color spaces. Additionally, other approaches employ temporal and spatial wavelet analysis [154, 155].

Although classical computer vision algorithms can register high success percentages in classification, by being based on fine-tuned heuristics they lack the generalization to be successfully applied to new datasets and different scenarios [156]. To mitigate this limitation, researchers increasingly employ computer vision algorithms in tandem with intelligent systems approaches. As noted by Cetin et al. in their review of video-based fire detection [94], from 2007 onwards most of the approaches integrate intelligent systems, e.g., fuzzy c-means clustering for selection of smoke-regions [157], and support vector machine classifiers [158].

Recently, with the uprise of the prominence of convolutional neural networks (CNNs) more contributions employ these architectures. The results reported for fire detection using CNNs demonstrate the merit of the approach, with high success percentages. Frizzi et al. propose a CNN architecture similar to LeNet-5 but that incorporates dropout and LeakyReLU activations [159]. However, despite reporting a 97.9% on a testing set of 5585 images, the article does not provide access to the database neither presents a proper characterization of the datasets used, thus the generalization of the model is unclear. Recent works also propose CNN architectures inspired by state-of-the-art models like AlexNet, and employing video-based datasets [148], or propose solutions with smaller datasets by exploring transfer learning approaches, using pretrained models such as VGG, or ResNet [160].

While deep learning (DL) models have proven efficiency in several closed-set problems, the application of these techniques to open-set problems is still in early stages. The limited amount of data for the fire detection problem is still a major challenge for training deep neural networks, and is a common hurdle in applying DL approaches in a real context. Previous works focused on fire detection also faced this limitation, with most resorting to images downloaded from web and social media platforms, which allow for keyword searches, such as Google or Flickr [160, 161]. Alternatively, other works resort to video datasets available in the literature [148, 162].

However, the aforementioned studies built datasets of fire instances irrespective of the scene context, e.g., fires in urban scenarios, in riots, industrial fires, or even fires in an indoor setting. Trained on these scenarios, the deep neural models learn from samples that are not representative of wildfire detection,

e.g. forest and vegetation, characteristic to rural regions.

While these contributions report high success percentages, the use of imbalanced datasets and/or video-based datasets provides a limited variety between the samples, which undermines the significance of the results. In this way, the performance may not offer guarantees regarding generalization to different scenarios.

9.1.3 Proposed approach

Wildfire detection poses challenging problems that differ from general fire detection developed for image surveillance systems in indoor environments. In outdoor settings, the fire detection task inherently becomes more difficult due to e.g. the environment being highly dynamic, having varying light conditions or reflections on natural or man-made objects, or seasonal changes. This means that both ad-hoc methods and data-driven expert systems developed for video surveillance in indoor environments are not adequate for forest fire detection, because in a different context these would not match the same performance guarantees.

Conversely, computational vision methods require extensive time to develop robust feature extraction pipelines, whereas application-oriented deep neural models require large-scale datasets, which hinders achieving accurate frameworks that can be valuable for operational use. Although recent works in current literature based on deep learning architectures demonstrate improvements regarding accuracy and computational efficiency, there is still a lack of understanding of the limitations of these models in real scenarios.

To address that gap, this work centers on addressing wildfire detection in wildland-urban-interface areas, and studying the challenges specific to this type of task. To overcome data limitations, we sourced databases from different stakeholders to build an adequate database for training and performance evaluation. Moreover, we devised a data augmentation procedure to introduce variations in translation, distortion and scale, increasing the difficulty of the classification task.

This paper proposes a transfer learning approach to wildfire identification in WUI areas, using the Inception-v3 model pretrained on ImageNet [163]. This model is retrained on an open-source database, the Portuguese Firefighters Portal Database, and evaluated with tenfold cross-validation. Furthermore, the Corsican Fire Database [164] is used as a benchmark for testing the best model obtained. This paper performs an in-depth study of the results by highlighting limitations in a number of common situations that must be addressed and solved before the application of deep neural models in a real context. The main advantages of the proposed framework are that it allows targeting this application with smaller datasets and with reduced computational load, while the testing setup using cross-validation enables testing a state-of-the-art deep neural architecture in a diverse and representative set of data.

The main contributions of this work are threefold: (1) identify the difficulties and shortcomings of state-of-the-art fire detection research, and assess issues related to the quality of the databases; (2) leverage

open-source or public datasets of real fire events using a transfer learning approach coupled with data augmentation techniques tested using tenfold cross-validation; and (3) present a comprehensive study of the causes of misclassification, revealing important insights for the development of future work.

This paper is organized as follows: Section 9.2 presents the Inception-v3 model used. Section 9.3 introduces the open-source datasets employed. Section 9.4 presents the methodology followed to formulate the classification problem, perform data augmentation, and model evaluation. In Section 9.5 the results are analyzed and discussed and Section 9.6 presents the conclusions.

9.2 Deep Neural Networks

Deep neural networks (DNNs), in a modeling sense, are a class of nonlinear models capable of learning multiple levels of abstract representations from raw data without the use of expert knowledge. These models can have different learning schemes, but all make use of a set of methods known as deep learning [165]. Unlike classical neural networks, which require prior feature extraction, DNNs handle the feature extraction within the model. DNNs obtain the *deep* denomination as a result of stacking multilayers, and are capable of dealing with raw high-dimensional data, e.g. images, text, sound, or video.

9.2.1 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a class of feedforward DNNs inspired in the animal visual cortex. The term was first introduced by LeCun when he proposed LeNet-5 [166], inspired by the neocognitron proposed in 1980 by Fukushima [167]. LeNet-5 comprises convolutional and pooling modules for feature extraction, fully-connected layers, and a softmax classifier, which represent the fundamental building blocks of CNNs.

The structure of the fully-connected (FC) layer is based on the architecture of classic neural network models like the multi-layer perceptron. In fully-connected layers, the neurons in adjacent layers have pairwise connections, but despite their name, neurons within a single layer do not share connections between them. Due to their amount of connections FC layers are very computationally expensive. Therefore, these are used sparingly in CNNs, and usually at the penultimate layer before the classifier.

Convolutional layers are the main building block of CNNs and are responsible for extracting representations from the data as it flows through the network. As illustrated in Fig. 9.2, this type of layer computes the dot product between patches of the input volume and the weights of the filters. The dimension of the input and output volumes reflects the usual application of this type of layer that reduces data in spatial size and increases its depth.

Due to their parameter sharing scheme (Fig. 9.3), where neurons in each depth slice share the same weights, the network has translation invariance. The output volume of this layer, see Fig. 9.2, stacks the activation maps of the different slices of neurons, which extract distinct features.

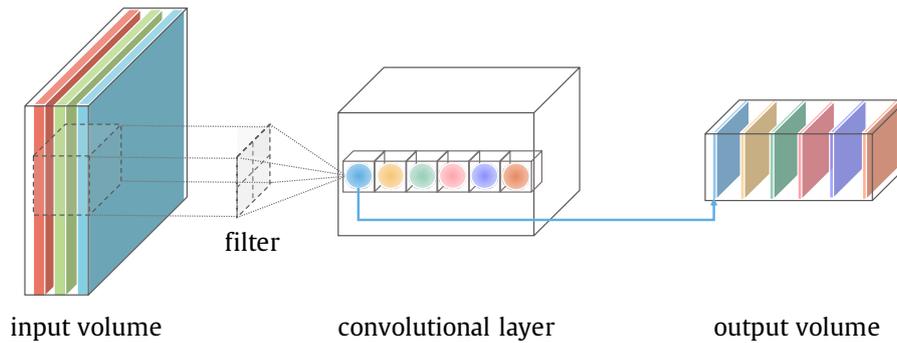


Figure 9.2: Example of a convolutional layer.

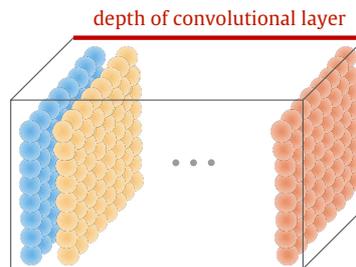


Figure 9.3: Parameter sharing scheme.

Pooling layers are also employed for subsampling of the feature maps to extract the most relevant features at specific points of the network.

In the last few years, the resurgence of the interest in CNNs came about after AlexNet [168], a deep convolutional neural network, won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) contest [169]. The model proposed a combination of strategies to make training faster using Rectified Linear Units (ReLUs) instead of sigmoidal activation functions, and leveraged GPU parallel computation for time efficient training. Additionally, to reduce overfitting, the model employed data augmentation techniques in order to generate more training samples, and a new regularization method called dropout. This marked a pivotal point in this field, to the extent that most of these techniques have become a standard for training DL models. On the footsteps of this success, in the following years, the top ranks were dominated by other CNN models, e.g. GoogleLeNet [170], VGG [171] and ResNet [172].

9.2.2 Pretrained models

The top-tier results achieved using CNNs in computer vision tasks were accomplished with large datasets, e.g., benchmark datasets such as ImageNet [173], CIFAR [174], MNIST [175]. This point is critical to the performance and versatility of DL models. On the one hand, since neural networks learn from examples they inherently become more accurate and robust when they have a larger set to learn from. On the other hand, having a comprehensive input space is also of concern because models are more likely to fail in the testing stage when they are given new inputs out of the input space. Therefore, when working with supervised learning models, it is imperative to be aware that these models excel at generalization but are

not able to extrapolate for data for which they have not learned to extract representations.

However, by highlighting the importance of dataset size, it is critical to point out that applications to real-world problems are often hindered by insufficient data. Although these are growing by the day, large datasets often take years to build, and even on a small scale, data acquisition depends on the collaboration between stakeholders from different fields. Moreover, while data is a 21st century commodity, there are still many challenges regarding data collection, compliance with privacy laws, and open-access. To increase the amount of data and overcome such limitations, several DL techniques will be employed including transfer learning and data augmentation.

9.2.3 Inception-v3

The Inception-v3 model (Fig. 9.4) is a deep neural network that is built modularly, by means of a stacking of parallel convolution structures known as Inception modules [163]. These employ several different convolution factorization schema depending on their relative depth location in the network and provide different hierarchical levels of feature extraction. Inception-v3 has 5 types of layers for feature extraction purposes, i.e., fully-connected, convolutional, two types of pooling layers (max and average), the dropout layer, and a softmax classifier. The diagram with the compressed view of the network divides it into two parts, the first responsible for extracting spatial features (A-F) and the second corresponding to the classifier.

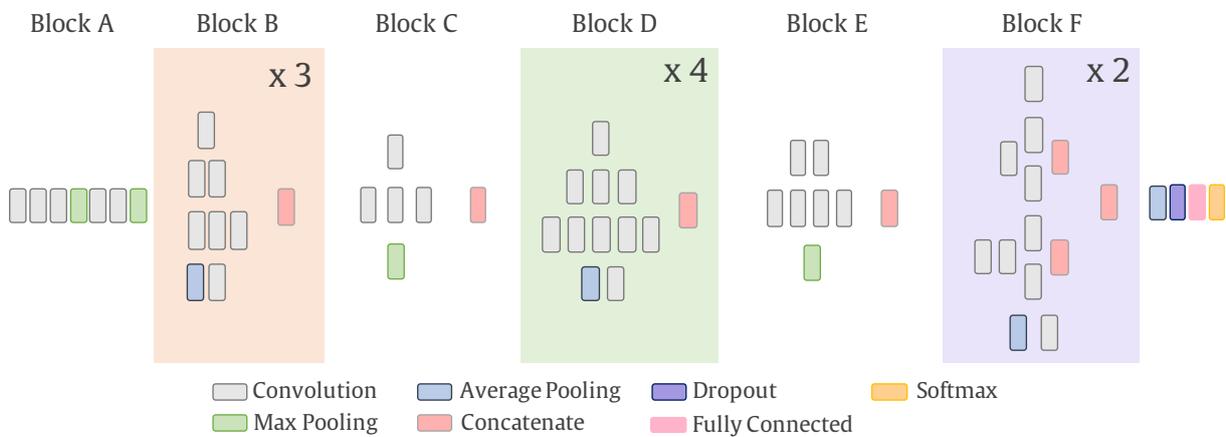


Figure 9.4: Inception-v3 network architecture.

To perform transfer learning, the classifier part of the network will be retrained. The softmax is the score function responsible for generating the predictions of the model. This function is given by:

$$\hat{y}_k = \frac{\exp(z_k)}{\sum_i^K \exp(z_i)} \quad , \quad k = 1, \dots, K \quad (9.1)$$

with k representing the index of the class label. For every input of the output layer, z_k , the softmax will generate K class prediction scores, \hat{y} . The numerator acts the same way as a sigmoid function, by squashing the inputs to values between 0 and 1, while the denominator normalizes the continuous-valued output by limiting the sum of all the class probabilities to be bounded between 0 and 1 as well.

The formulation of the optimization problem for readjusting the parameters of the last layers consists in finding the weights and biases that define the mapping between the inputs and the correct outputs. To that end, the objective function will minimize the error between the prediction of the model and the ground truth labels.

The loss function quantifies the distance between the predicted outputs, \hat{y} , and the ground truth labels, y . For this model, the loss is computed using the cross-entropy function, which is calculated as:

$$L(\mathbf{y}, \hat{\mathbf{y}}) = - \sum_k y_k \ln(\hat{y}_k) \quad (9.2)$$

9.2.4 Model evaluation

The performance benefits obtained by training with more samples, coupled with the computational burden associated with training DL models, usually dictates performing a single split into training/validation/test sets.

While cross-validation is used extensively for evaluation of shallow models, it is seldom applied to deep models, since the computational cost increases k -fold. However, considering that a transfer learning approach enables retraining with a small dataset with a less demanding computational load, cross-validation becomes a viable option.

For small datasets, performing a single split into training/test, e.g. 70/30 subsets, entails testing the model on only 30% of an already restricted number of samples, which could give place to a relatively noisy estimate of model performance [176].

In opposition, by dividing the data into k -folds, the data are fit in $K - 1$ parts of the entire dataset and validated with 1 alternating fold. For each iteration, this yields K different models, and the performance measures are averaged over all folds, giving a more representative estimate of the classifier performance [177].

The most used metrics to rank classifiers are scalar performance measures e.g. accuracy, precision, sensitivity, specificity, or F-score. These performance measures are computed based on a confusion matrix which accounts for the number of true positives, false negatives, true negatives and false positives. However, by relying on a fixed threshold and being independent of class priors, these measures can be misleading when working with skewed datasets.

Conversely, recalling the softmax classifier evaluates the degree of confidence in classification according to a posterior probability, i.e., it generates a continuous-valued output, this enables using an alternative performance metric, the receiver operating characteristic (ROC) curve.

ROC curves represent the trade-off between hit rates and false alarm rates, i.e., the compromise between true positive rate and false negative rate [178]. Albeit its original use in signal detection theory for binary classification problems, they have for long been widely adopted by the machine learning field to evaluate classifier performance.

The prediction scores are discretized using variable threshold values, which allows tuning the latter to optimize classification performance.

9.3 Databases

The following sections present the datasets used in this work and their characteristics. Moreover, section 9.3.3 references other open-source databases available in the literature.

9.3.1 Portuguese Firefighters Portal Database

A Portuguese media outlet dedicated to supporting the work of firefighters provides on the webpage Bombeiros.pt an image database categorized in fields associated with firefighting activities such as: fire events, accidents, aerial support vehicles, firehouses, emergencies, etc. Herein we refer to this database as the Portuguese Firefighters Portal Database (PFPDB). The data are publicly available online and new images are regularly published, offering at the time of this writing 25,157 images and 454 subcategories [179]. For this work, the subfields related to forest and rural fires were selected, corresponding to a total of 3,900 distinct images.



Figure 9.5: Samples of fire (top row) and not fire (bottom row) from PFPDB.

Note that the samples depicted in Fig. 9.5 have been cropped and reshaped to squares from the originals available at the website. The latter are watermarked and have varying portrait and landscape aspect ratio. The preprocessing steps to work around this constraint are detailed further along in Section 9.4.2.

9.3.2 Corsican Fire Database

The Corsican Fire Database (CFDB) is an online image database to support wildfire research, which provides a testing dataset for comparison of computer vision algorithms [164]. It strives to be an evolving dataset, and as of March of 2018 is composed by 500 visible range images of wildfire, 100 pairs of visible and near infrared images, and also 5 multi-modal sequences of 30 pairs each.

In contrast with the database presented in Section 9.3.1, this dataset is composed solely by fire images. Samples of this database (see Fig. 9.6) will be used for testing in the performance evaluation stage, despite



Figure 9.6: Rescaled samples of *fire* images from CFDB.

not having situations without fire.

9.3.3 Other open-source datasets

The fact there is a lack of large-scale datasets for fire detection publicly available poses a problem for the comparison of approaches. Besides the datasets previously mentioned, there are other datasets employed in related works that are available open-source, such as the examples from the literature presented in Table 9.1.

Table 9.1: Other open-source datasets used in the literature.

Datasets	Source	Total Images	<i>fire</i>	<i>not fire</i>
MIVIA [162]	Video	64085	12596	51489
BoWFire [180]	Images	226	119	107

Note that video-based sources account for consecutive frames, thus the samples vary less than when dealing with image-based datasets. This is an important distinction, since training models with very similar samples can result in overfitting.

These datasets concern fire detection in general therefore were not deemed adequate for this study, since it is aimed at exploring fire detection in rural environments and identifying possible limitations in these scenarios.

9.4 Methodology

Transfer Learning (TL) is a deep learning method that consists of leveraging a previously optimized network, a pretrained model, and use it to learn an entirely new task, using a different dataset. Currently, this technique is extensively used to address tasks where data availability is still scarce [181], enabling novel applications in other scientific domains such as remote sensing.

Following a transfer learning approach, this work employs the Inception-v3 model pretrained on ImageNet, and retrains the classifier part to perform single-image classification under a supervised learning scheme.

9.4.1 Classification problem

The task of fire detection is formulated as a binary classification problem alike preceding works [148, 159, 160]. The images will be classified individually into one of two classes: *fire* or *not fire*. As a starting point, we will set the threshold value for the prediction scores at 0.5. Further along, after performance evaluation, we will reassess this value by comparing it with the optimum threshold obtained from the ROC curve.

Before delving into the details of our dataset, we present the criterion adopted to establish the ground truth of each image. Using visible range sensors, a fire event can be identified by spotting a fire flame, a smoke curtain or both. At this stage, the criterion for fire detection only includes situations in which a fire flame is visible in plain sight. Let the definition of the classes be defined as follows:

- *fire* (positives): flame is visible in the image; and
- *not fire* (negatives): flame is not visible in the image.

Fig. 9.7 exemplifies instances of each class.

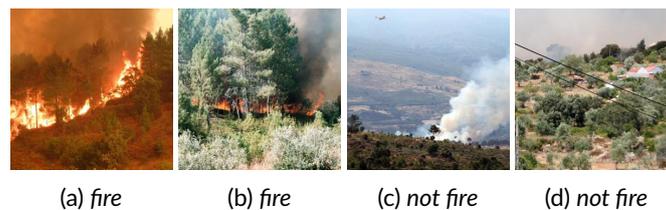


Figure 9.7: Samples of positives and negatives for both classes.

The selection of this criterion was influenced by three factors. On the one hand, previous approaches to fire detection using CNNs also focus on flame detection [148, 159, 160]. This happens because algorithms that rely on smoke detection usually present other issues such as an elevated false alarm rate. While in some situations, detecting a fire event through the detection of smoke could provide an earlier detection, conversely the potential for false alarms could make the classifier unreliable in a real context. A third factor concerns data limitations, since classifying images with only smoke as positives would lead to a unbalanced database.

9.4.2 Data augmentation

The need to remove the logo in every image of the firefighter's database, paired with the objective of having a larger and balanced training dataset, motivated the application of a data augmentation process.

The main steps of the algorithm are summarized in Fig. 9.8, as well as an illustrative example of the pre-processing schema. After first removing the borders of the images, the image logo is located with an algorithm based on color segmentation and feature descriptors. To reduce the number of features, a color segmentation algorithm targets the red color of the Bombeiros.pt logo. This process is performed on the logo image and the original image, narrowing down the search space significantly. Subsequently, features

are extracted on both segmented images using speeded up robust features (SURF) [182]. The relative location of the points in the original image that match the features gives the position of the logo.

The images in the database have either landscape or portrait aspect ratio, and the logo is located on one of the corners, which results in 8 possible cropping configurations. For each image, the position of the logo is obtained in 3 steps:

1. Assess the type of aspect ratio by comparing image height and width dimensions.
2. Perform red segmentation of the image logo and original image.
3. Match SURF features to extract the logo position in the original image.

Then, the original image is cropped in three configurations to create the same number of new images. The cropped areas for landscape and portrait aspect ratios are defined in Table 9.2.

Table 9.2: Cropping areas height x width (in pixels).

Aspect ratio	Red area	Yellow area	Blue area
Landscape	395x700	465x465	395x395
Portrait	625x470	470x470	395x395

As a final preprocessing step, the new images are rescaled to 299x299 to match the size of the input layer of Inception-v3.

To illustrate this process, take the example of the augmentation performed on the original image of Fig. 9.8b. To obtain the first image, the area signaled with the interrupted red line is cropped to create a new image. Since the aspect ratio is rescaled from a rectangle to a square, this transformation introduces a new sample with distortion. For the second and third images, the areas are the ones depicted in yellow and blue respectively. Due to some overlap in the areas of these two images, this will produce a translation effect in the resulting images. Conversely, since the crops have different areas and will all be reshaped to 299x299, this will also result in changes in scale.

All in all, this augmentation procedure introduces variances in distortion, translation and scale.

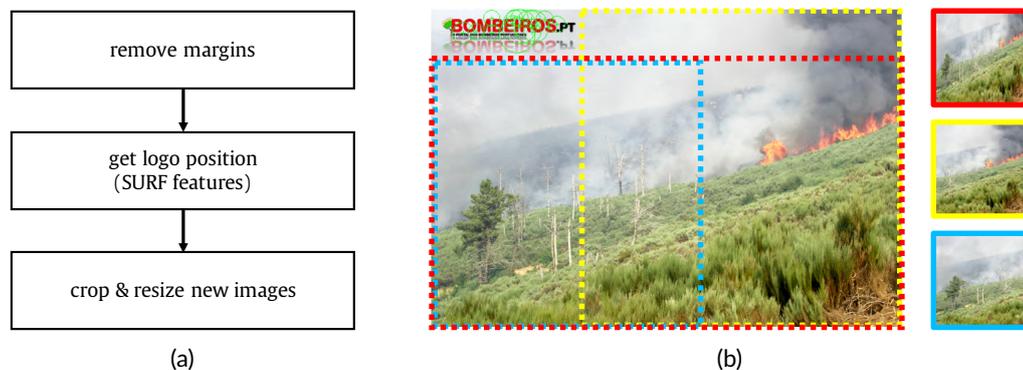


Figure 9.8: Preprocessing schema to remove the image logo and to perform data augmentation: a) algorithm; b) example of cropping and resizing to generate three new images of size 299x299.

9.4.3 Database

The data augmentation technique described previously was applied to the Portuguese Firefighters Portal Database, creating a larger source of images for training. From this new set, images were selected with discretion with the overarching objective to ensure variety of the samples. The resulting database has predominantly images of WUI scenarios captured at varying distances, as well as instances of subclasses such as firefighters, fire engines and aerial support vehicles, as previously illustrated in Fig. 9.5.

To establish the ground truth, the images were regarded as single instances, irrespective of how they were obtained. Once again, considering the original image depicted in Fig. 9.8b, the red and yellow images have the ground truth label *fire*, whereas the blue has the ground truth label *not fire*.

Due to the characteristics of the PFPDB database, which has predominantly fire scenarios, the data augmentation process yielded a skewed distribution between classes with a ratio of approximately 65/35 of *fire* and *not fire* images respectively.

However, since the performance measures widely used in the deep learning field are dependent of class prior probabilities, having a balanced dataset is a factor of critical importance. To even out the number of samples between classes, additional images were sourced online. Images hand-picked online were analyzed case-by-case and manually cropped at random to introduce variations in scale, translation and distortion.

Fig. 9.9, summarizes the process of creating the resulting database, herein referred to as the cross-validation database (CVDB), which is composed by 3623 *fire* images and 3551 *not fire*.

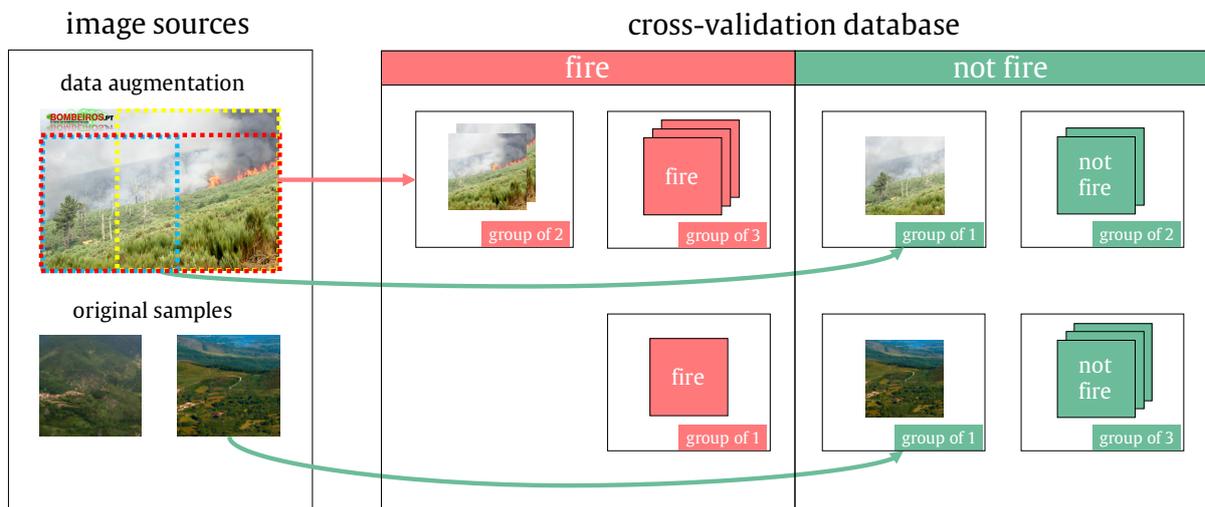


Figure 9.9: Description of cross-validation database.

Note that, as depicted in Fig. 9.9, the images derived from the same original image are indicated as forming a group. This is an important consideration for cross-validation as will be described further along. These groups of images can have 1, 2 or 3 images and can be of three types:

- fire: all sub-images with *fire* ground truth

- not fire: all sub-images have *not fire* ground truth
- mixed: sub-images can have different ground truths

This information will be relevant in the analysis done in Section 9.5.3.

9.4.4 Dataset partition

Typically, the partition algorithms for dividing datasets into folds for cross-validation simply split the data randomly in approximately equal-sized parts. However, as consequence of the effort to increase the number of samples in the database, and because each image is treated individually as if it was an independent observation, it is necessary to ensure that images derived from the same original image are allocated to the same fold. Otherwise, similar samples such as the red and yellow images from Fig. 9.8b could end up on the training and testing subsets respectively, thus undermining completely the validity of the results.

To overcome this issue, the partition algorithm devised, illustrated in Fig. 9.10, starts by grouping images derived from the same original image that share the same class. The folds are created sequentially and separately for each class with a 50/50 class distribution. Then groups of images from the *fire* class are assigned into equal-sized folds. Before allocating a group of images from the *not fire* list to a certain fold, the existing folds are searched for the existence of a group from the same original of the *fire* class. If a match is found, the *not fire* group is directly assigned into the correct fold, otherwise it can be added to the current fold being created. This procedure ensures a balanced class distribution in every fold and that images derived from the same original image are not in the training and testing sets simultaneously.

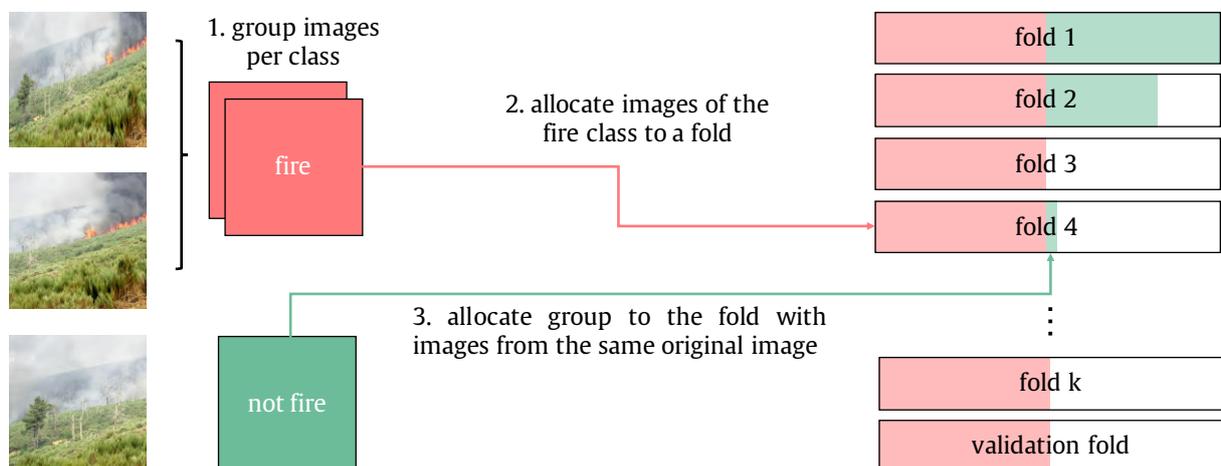


Figure 9.10: Description of procedure to partition data for k-fold cross-validation.

This partition algorithm was used to create 10 folds for cross-validation (6000 images), plus an extra fold of 600 images to be used as the validation set for hyperparameter selection.

9.4.5 Hyperparameter selection

For hyperparameter selection, 21 values of learning rates were tested between 0.001 and 0.1. For each, training batch size was changed between 100 and 200, but it did not yield benefits in classification per-

formance, hence the batch size was fixed at 100 to speed up training. From the results of these iterations the learning rate was fixed at 0.05 and the number of epochs at 10, so as not to overfit the models. The models were trained using the stochastic gradient descent (SGD) optimizer.

9.5 Results and discussion

The following sections present the results of retraining Inception-v3 with the cross-validation database, and the best model obtained is evaluated with the Corsican Fire Database (CFDB). Subsequently, an in-depth analysis is presented of the limitations associated with the application of this transfer learning approach, with a comprehensive analysis of the misclassifications and its possible causes.

9.5.1 K-fold cross-validation

The dataset used for cross-validation employs 10 folds of 600 images each (CV dataset), evenly split between classes, randomly sampled from the cross-validation database. The Inception-v3 model is trained using the data in $K - 1$ folds and the data in 1 fold is hold out for testing. The Inception-v3 model was retrained 10 times for each training set yielding a total of 100 models. The classifier is evaluated by averaging the results of the performance measures of all the models, which are presented in Table 9.3, with the mean and standard deviation of each performance measure. Table 9.3 also presents the performance for the best model tested on the CFDB.

Table 9.3: Model performance evaluation.

Dataset	Accuracy(%)	Precision(%)	Sensitivity(%)	Specificity(%)
CV dataset	93.60±1.57	94.12±2.45	93.13±3.28	94.07±2.78
CFDB*	98.6	100	98.6	-

*for the best model only

Overall, based on these performance metrics, it is possible to infer that the model is able to generalize well for unseen data, considering the results achieved for both datasets. Nevertheless, while achieving best accuracy performance is important, when possible, it is also desirable to minimize the rate of false alarms (maximize specificity). Therefore we will use the ROC curve to assess if a better trade-off between hit rates and false alarms could be achieved.

Observing the ROC curves for a fixed threshold at 0.5 and variable threshold, depicted in Fig. 9.11, it becomes evident that the dominating curve is for variable threshold with a greater area under the curve (AUC). Notwithstanding, by calculating the best trade-off between hit rates and false alarms, represented by the optimum point in the curve, the value of the threshold for this point is 0.5099. The minute difference between threshold values does not suggest a significant improvement in performance, which is confirmed by the results presented in Table 9.4.

Considering an accuracy mean error rate of 6.4% corresponds to an average of 40 images misclassified in 600, it still warrants further investigation to understand its possible causes.

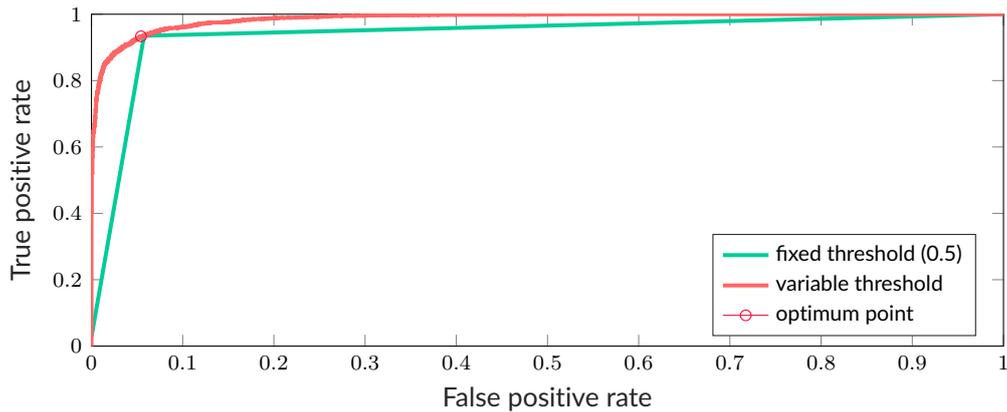


Figure 9.11: Comparison of threshold values using ROC curves.

Table 9.4: Comparison of performance with different threshold values for the CV dataset.

Threshold	Accuracy(%)	Precision (%)	Sensitivity(%)	Specificity(%)
0.5000	93.60±1.57	94.12±2.45	93.13±3.28	94.07±2.78
0.5099	93.60±1.57	94.30±2.39	92.92±3.32	94.28±2.67

9.5.2 Performance analysis

On the one hand, scalar performance measures give a tangible metric for ranking classifiers, but are too opaque to give insight about the fitness of the model. On the other hand, ROC curves address the same ranking objective by considering the prediction scores instead of the final binary classification, enabling a better fine-tuning of model classifications by varying the decision threshold to attain the desired balance between hits and false alarms. Notwithstanding, to understand the fitness of our classifier we delve beyond these methods.

Taking a step back, we start by analyzing the prediction scores of the model and check how far off those are relative to the ground truth labels. For that purpose, we designed a transparent confusion matrix, which can be observed in Fig. 9.12. To follow the structure of a confusion matrix, images were sorted by class, with the positive class, in our case *fire*, in the first column. To facilitate discerning the separate classes, images of the *fire* class are depicted in red and from *not fire* class in green. The symbols representing the hits or misses are indicated in the legend.

Despite the expected variations of performance of each fold, we study fold 10 to exemplify the reasoning of the analysis. Observing Fig. 9.12, most of the images are predicted with a high confidence level. However, focusing on the area of the threshold value at 0.5, there are several borderline instances, where some are on average well classified and others not.

Due to the variance between iterations, it matters to quantify deviations, but a box plot visualization would be difficult to inspect for 600 instances at a time. For this reason, we opted for a histogram with the number of repetitions each image is misclassified, displayed in Fig. 9.13. For ease of visualization, errors that occurred in more than 5 of the iterations are illustrated in red, whereas more seldom errors are depicted in green.

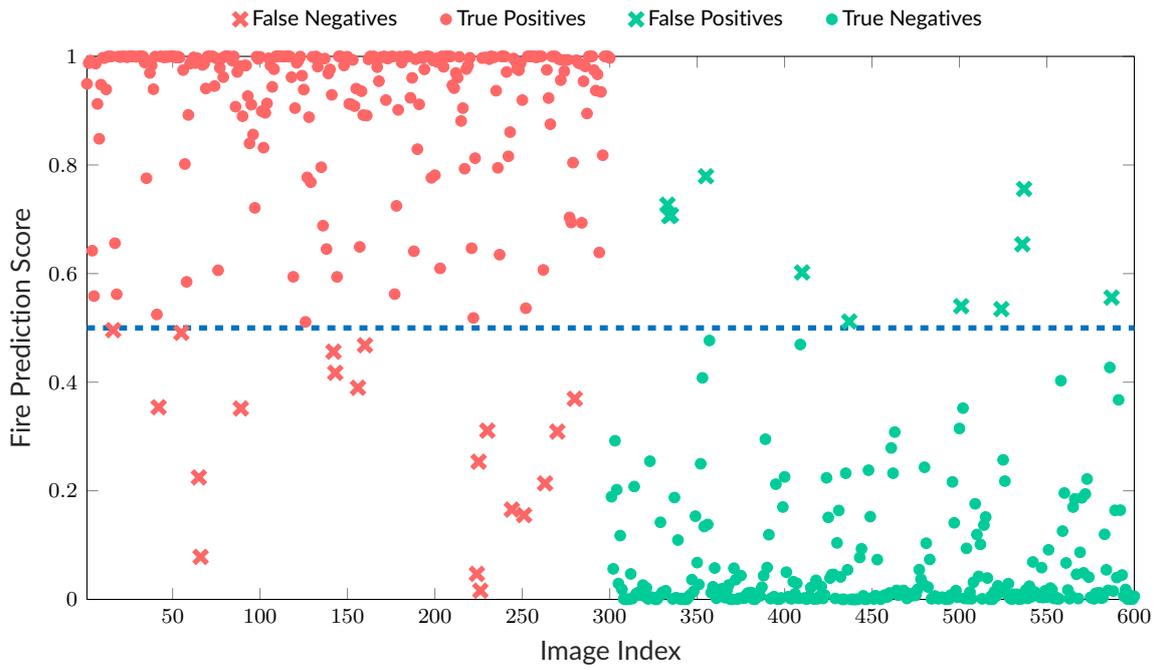


Figure 9.12: Transparent confusion matrix depicting the mean fire prediction scores for fold 10.

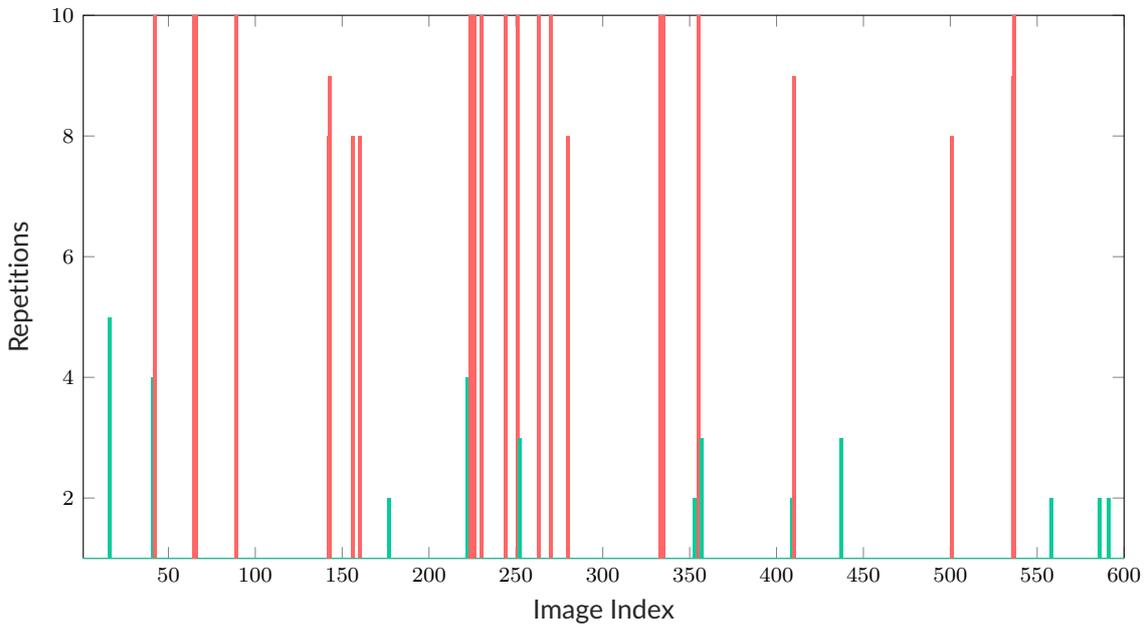


Figure 9.13: Histogram of misclassifications for fold 10 (10 iterations).

Attending to the results illustrated in Fig. 9.13, the images where the network exhibits difficulties in classification are usually the same across iterations. Then if most of the misclassifications are recurrent, it matters to understand if there are limitations associated with these errors inherent to the model or to the data.

9.5.3 Classification analysis

With a 93.6% classification accuracy, the model generalization is high, even so considering the wide variety of situations encompassed by the database, which has a total of 6000 images from different fire events. The following sections present examples of real scenarios that demonstrate the fitness of the classifier and its limitations, covering with special attention both false negative and false positive misclassifications.

The focus of the analysis lies on groups of 3 images for two reasons: it is the most represented arrangement in our database, and more data provides greater insight, than comparing isolated instances. The samples presented throughout Section 9.5.3 juxtapose the ground truth of each image, with the values of the mean and standard deviation of the fire prediction scores.

Successful fire classification

This section highlights examples of correct classifications of groups of three images with mixed classes, which were derived from the same original image but have different ground truths, as described previously in Section 9.4.3.

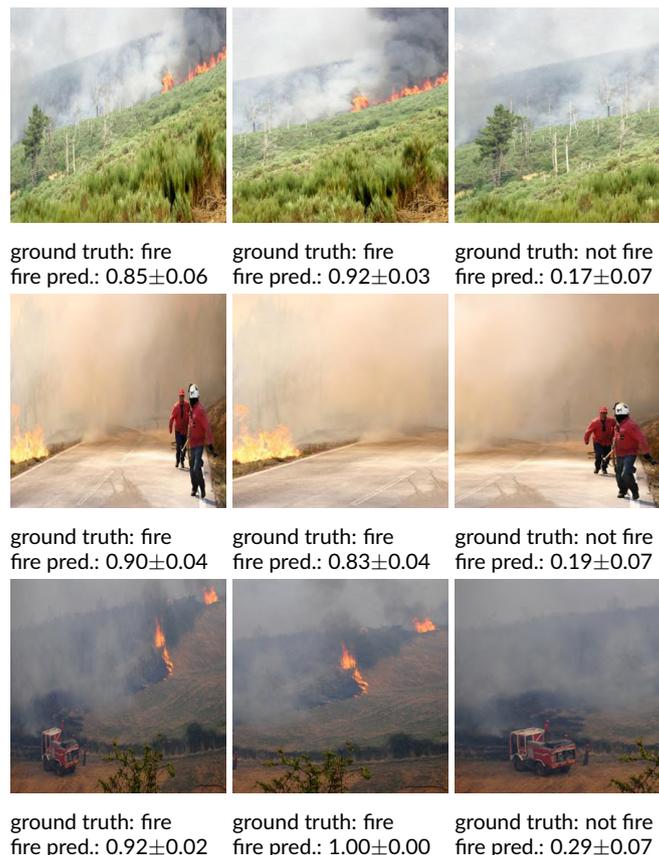


Figure 9.14: Successful classifications of *fire* (left and center) and *not fire*(right).

As can be observed in Fig. 9.14, considering a threshold value of 0.5, it is possible to attest that fire is correctly classified in different scene contexts, including in the presence of elements such as firefighters or firetrucks. These examples represent appreciable difficulty because the augmented images are very

similar between them but have opposite ground truths. The prediction scores demonstrate the network was able to learn fire patterns and generalize well in these conditions.

False negatives

For a fire detection classifier, false negatives should be kept to a minimum, and therefore the analysis of these cases can bring to light the limitations of the model as well as of the training itself. In this section we analyze a series of examples, to bring attention to some patterns that were identified.

By analyzing the batch of false negatives, two causes for the misclassification can be suggested: reduced spatial scale and spatial location.

Reduced spatial scale

The pixel scale at which fire can be detected by the network is a factor of paramount importance. In the examples in Fig. 9.15, in which the fire is captured at long distances, most of the images were misclassified, or were close to the decision threshold. Nevertheless, considering that other images in the database with similar conditions were correctly classified, further tests would need to be done to assess the degree to which variance in scale affects classification performance.

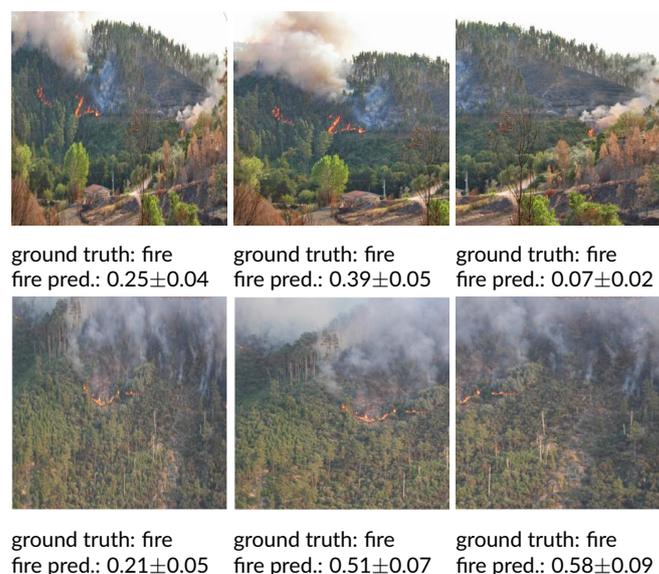


Figure 9.15: False negative classification associated with reduced spatial scale.

In the future, evaluating the smallest pixel area that can be detected by this type of models plays a crucial role in determining the altitude of operation of monitoring systems.

Pattern spatial location

Comparing the images in Fig. 9.16 it is noticeable that images with fire patterns around the borders are misclassified in opposition to their counterparts, which are correctly classified with a high degree of confidence. This type of misclassifications occur for the images on the right of the top two rows, and the one

at the center of the bottom row.



Figure 9.16: Examples of false negative classification associated with fire pattern spatial location on the image borders.

Despite the fact that CNN models acquire translation invariance through their parameter sharing scheme (Fig. 9.3), the incorrect outputs indicate the model is not learning this capability after all. However, keeping in mind neural networks are not able to generalize for situations for which they were not trained, the reason for failure in this scenarios could be lack of representativity of similar situations in the training data.

To ascertain if this was the case for our database, we computed the standard deviation across the RGB color channels. Since fire is characterized by having high values in the red channel and lower values in the blue and green, the standard deviation between these variables gives a straightforward way to measure the pattern distribution.

Following this reasoning, we computed the cross-channel variance for the false negatives which met this misclassification category. Fig 9.17 displays the comparison between the distribution for the false negatives and the whole training set. The fact that fire patterns are less present around the borders in Fig. 9.17b confirms our hypothesis. To minimize this limitation, functions responsible for partitioning the data for cross-validation should incorporate this measure to ensure a better distribution.

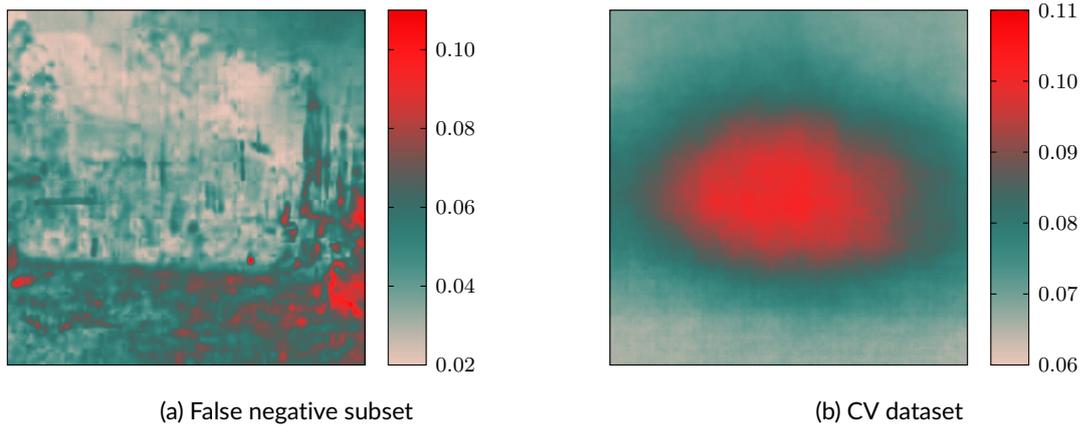


Figure 9.17: Comparison of cross-channel variance.

False positives

The rate of false alarms is an important concern for fire detection problem, because it weights heavily on the reliability of the classifier, and its applicability in real scenarios.

In a general sense most of the false positives could be attributed to two different situations described in the sequence.

Sunset

Sunset contexts, where there is a pattern of high intensity in the images, often cause classification errors. Note that all the images in Fig. 9.18 should have a fire prediction score below 0.5, but for the first two images the presence of the sun leads to errors in classification.

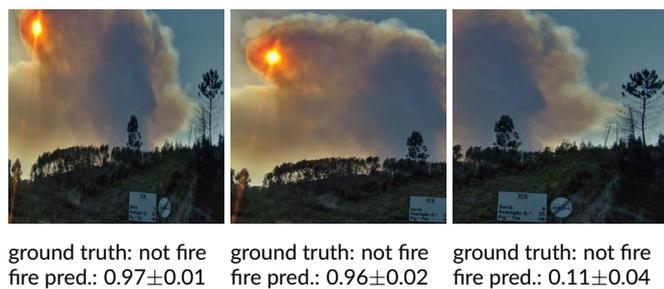


Figure 9.18: Example of false positives due to sunset (left and center).

Since CNNs only acquire color invariance through the training data, the retraining approach might be constrained to this caveat as a consequence of the previous optimization procedure. Despite having 106 images of sunsets in our dataset, the retraining of the last layer was not enough for the network to learn to classify correctly this type of instances. A solution that can be explored is to fully train the network if more data is available, and expose the model to more samples of sunset settings.

Smoke curtain

In the formulation of the classification problem, the ground truth established considered an image as belonging to the *fire* class only if a flame is visible in plain sight. However, the presence of a smoke curtain in the images revealed to be a common cause of false positives, as can be observed in Fig. 9.19.

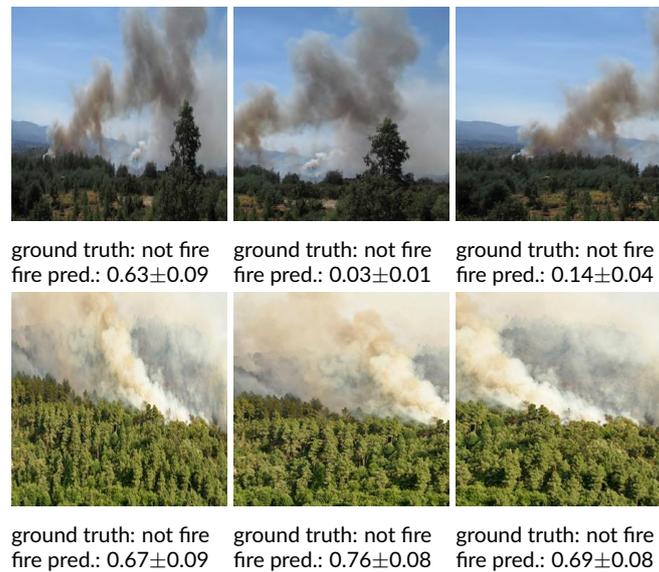


Figure 9.19: Examples of false positives associated with the presence of smoke.

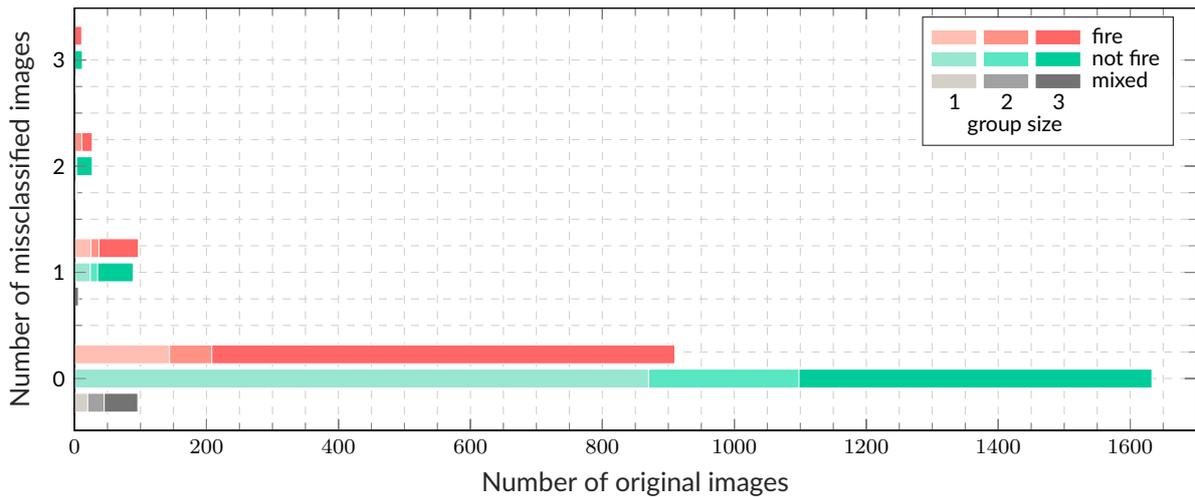
This effect was to be expected because when the network learns high-level features of the *fire* class it associates, to some extent, the patterns of fire and smoke.

Overall performance

Although the classification approach followed was based on single image classification, in the previous section it was identified that frequently in a group of images derived from the same original image, not all images were misclassified.

Therefore, since we have many instances of groups with 3 images, it may be beneficial to approach classification as a group instead of individual instances. To assess the performance of the model in terms of the original images, we quantified the number of misclassifications per original image, as can be observed in Fig. 9.20.

Considering there are groups with 3 images that have only one misclassification, the results could be further improved if the algorithm considered the group has a whole and not as independent samples. This indicates that while the variances to distortion, scale and translation provide benefits in generalization, for the final classification the performance would improve if evaluated considering the ground truth of the original data.



of k-fold cross-validation on augmented data increases size of the dataset, hence enabling testing the approach in a representative set of data rather than a single split where only part of the data would be tested; (4) using image data provides higher variability between samples than video-based images taken from high frame rate sources.

From the thorough analysis of the misclassifications, four patterns emerged as the main causes of error. Reduced spatial scale of the flame as well as the spatial location of the fire patterns originated false negative situations, whereas sunset scenarios and smoke curtains were a common cause of false positives. Further research is required to improve the classification of images with these characteristics, as well as to extend the fire classification to situations with smoke, both being crucial for application of these models in real scenarios.

The difficulty of wildfire detection at reduced spatial scales may be inherited from the deep neural network architecture, thus evaluating the smallest pixel area that can be detected by this type of models will be fundamental for determining the altitude of operation of monitoring systems. Moreover, based on the insights gained concerning the spatial location samples, the functions responsible for partitioning the data for cross-validation should incorporate a cross-channel variance measure introduced to ensure a better spatial distribution in the training datasets. Furthermore, sunset or smoke misclassifications are to be expected due to limitations in the training dataset. To address this issue, the evolution of current datasets should aim at improving the variety of situations included. As noticed in the experiments, the fire class already associates, to some extent, the patterns of fire and smoke, thus a multi-class formulation could benefit the classification performance. While the variances to distortion, scale and translation provide benefits in generalization in training, for the final classification the performance would improve if evaluated considering the ground truth of the original data. In practice, the image classification pipeline may produce more reliable results considering a set of consecutive frames within a time-window.

In light of the insights revealed, this study opens new avenues of future research that should be considered towards improving model generalization while working with the evolving datasets. First, a possible extension of this work may include developing algorithms to balance dataset partitions for cross-validation using sub-classes, e.g., smoke, sunset, mountain range, firefighting vehicles. Second, data augmentation procedures could extend the current approach by improving the representativeness of the spatial distribution of the fire instances, based on the cross-channel variance measure introduced. Third, towards real-world application, the problem formulation could be extended to encompass multiple classes, namely both fire and smoke instances.

Finally, the effort to build benchmark datasets not only limited to visual range data, but also encompassing multimodal data, is a continuing and evolving activity in the community of wildfire research. With the emergence of new databases current approaches in intelligent systems can also aim to include extensions to explore spatiotemporal dependencies, as well as develop hybrid learning systems that leverage multimodal data for improved generalization and reliability in real scenarios.

Fire Monitoring with Fuzzy Models

Contents

10.1 Introduction	152
10.2 Experiments and Data Acquisition	155
10.3 Thermal Imaging Analysis	162
10.4 Fuzzy Modeling	168
10.5 Conclusion	181

The material included in this chapter was previously featured in a peer-reviewed journal article, published in Expert Systems with Applications, Elsevier.

Published in:

M. J. Sousa, A. Moutinho, M. Almeida, **Classification of potential fire outbreaks: A fuzzy modeling approach based on thermal images**. Expert Systems with Applications. 129, 216–232 (2019). doi: 10.1016/j.eswa.2019.03.030

10.1 Introduction

Fire outbreaks can have devastating consequences both on social and economic levels, leading to extensive material damages, and in extreme situations to the loss of lives [183]. While in urban settings regulations establish that buildings must follow mandatory fire safety measures and protocols [184], in temporary settlements like campsites these rules are not so demanding.

Temporary campsites, associated with recreational activities as camping, outdoor festivals, or pilgrimage activities, present high fire hazard. The combination of combustion inducing equipment and highly combustible paraphernalia associated with camping activity, can lead to fire incidents [185]. These events present a considerable danger for public safety and often require the evacuation of thousands of people. In addition, these places are commonly situated in the wildland-urban-interface, which can expose the participants to wildfires that spur from outside the venue premises [186, 187].

In turn, due to ongoing conflicts, violence and persecution, by the end of 2017 there were over 68 million forcibly displaced people worldwide [188]. The migration crisis leads to overcrowding in refugee camps and shelters, which results in poor and unsafe living conditions. In these settlements, individuals are also at risk, as a result of inadequate fire safety conditions and the difficulty of implementing strict fire prevention guidelines [189]. Moreover, the human presence acts as another source of potential fire causes either by negligence or risk behaviors. These factors increase fire risk and make campsites a hotbed for the occurrence and rapid propagation of fires [190, 191].

For these reasons, there is an urgent need of alternative fire prevention systems that can mitigate the consequences of fire outbreaks in this type of environments.

In that sense, the advances in computer vision and computational intelligence open opportunities for the development of intelligent systems that can provide early fire warnings without requiring permanent human monitoring.

Regarding sensor-based fire detection, the type of sensors employed depends considerably on the characteristics of the application. Whereas indoor solutions make wide use of smoke and gas sensors [192], outdoors the use of image surveillance systems is predominant [17, 193]. In that concern, in light of the recent advances in sensor technology, thermal cameras have now more competitive costs and are increasingly being applied in industrial applications [194–196]. Moreover, with an expanding range of low-weight models available in the market, which can easily be installed onboard aerial platforms, applications to detection of wildfires are also a hot field of interest [197].

In this context, this investigation centers on the development of solutions for the early detection of fire events towards the mitigation of the consequences of these incidents, both on socio-economic and environmental levels.

To that end, this work focuses on analyzing nonradiometric thermal imaging data (RGB encoded images) for detection of fire incidents in an initial stage. The principal objective centers on the development of

an expert system that can explore the advantages of thermal imaging sensors, aimed at the application in dynamic environments e.g. campsites.

The proposed solution may be used as a decision support system for surveillance teams of campsites or other delimited locations, providing automatic alerts of identified high-risk situations.

10.1.1 Related Work

Vision-based fire detection has been an extensively researched topic, with approaches that explore visible as well as thermal range images [94]. The type of sensory data is a major driver in the selection of the methodologies to use, and depends widely on the scope of application desired. From literature review, the main research paths taken to address fire detection focus on two objectives: detection of flames and/or smoke. The following aims to briefly lay out the main methodologies studied in previous works.

In the computer vision domain, contributions propose a wide variety of color-based models [193, 198–200], which are frequently complemented by motion [201], or temporal features [202].

However, while many approaches based on classical computer vision algorithms report high success rates, these lack the generalization to be successfully applied to broader contexts. This represents an important problem for the development of automatic fire detection systems that can be reliable for application in real-world scenarios.

To address this issue, intelligent systems methods are increasingly used to complement computer vision techniques [94]. The solutions proposed encompassing the use of expert systems include neural networks [148, 203], fuzzy clustering [204], optimization algorithms [205], and support vector machines [204, 206].

Conversely, in an attempt to obtain better model generalization, recent works tend to rely solely on artificial intelligence techniques such as deep neural networks [148, 203]. Although being promising, this class of models requires considerable amounts of data to be effective, which is still a difficulty when following this methodology.

Regarding uses of thermal infrared images, there have been some research efforts [98, 99, 207], but the acquisition cost of thermal imaging systems has been a roadblock for the widespread use of this technology. Nonetheless, today, the decrease in cost coupled with the versatility of new thermal cameras available in the market, can make this technology an emerging trend in the near future [197].

Previous contributions focus mainly on wildfire detection systems and report a significant rate of false positives, due to variations in light conditions (e.g. sunlight, reflections), the detection of objects at high temperatures, or the false alarms caused by weather conditions [103].

Furthermore, since fire is a source of extreme temperatures, thermal imaging sensors are a natural choice for a fire detection application. Considering that thermal cameras have sensing capabilities that cover part of the infrared band, this type of sensory data can offer great advantages for an earlier detection. Nevertheless, devising a robust computer vision algorithm for this type of data is not straightforward,

because the color of the images captured by these sensors vary with the relative temperatures between objects in a scene. This fact introduces high variability in the data, because colors depicting an object vary over time, depending on the scene context.

10.1.2 Proposed Approach

Considering that the variety of thermal cameras available in the market can have different capabilities, with radiometric models that provide temperature measurements, and nonradiometric models, which do not provide temperature measurements, it is important to develop general algorithms that can work with a wide range of thermal imaging sensors.

Independently of its radiometric capabilities, the user may have access to the sensor raw data or simply to RGB encoded images (according to the color palette chosen), depending on the camera operation mode. For real-time applications, like fire detection or surveillance, in most cases it is only this RGB encoded image output (with or without a temperature-color scale) that is available.

In this way, this study focuses on the development of a classification model based on RGB encoded images captured with thermal cameras for early fire detection at highly dynamic environments like campsites. With that objective in mind, a series of experiments were conducted under controlled fire conditions in a laboratory setting, as well as in the real world context of a campsite during a summer festival.

The data acquired are presented in detail, and a feature engineering process is devised for the analysis of the response of thermal imaging sensors to a fire ignition and the subsequent development of a classification algorithm. Since fire detection algorithms are frequently prone to false alarms, to deal with the high inner-class variability in the data, an intelligent systems approach is followed.

This paper proposes a fuzzy modeling approach based on features constructed from thermal images, which is transparent to interpretation and enables the assessment of how the sensor response influences image variability and its effects in the performance of the classification algorithm for fire detection. Figure 10.1 presents a high-level schema of the proposed approach. By relying on qualitative information only, the proposed approach is suited to be applied to both radiometric and nonradiometric image systems.

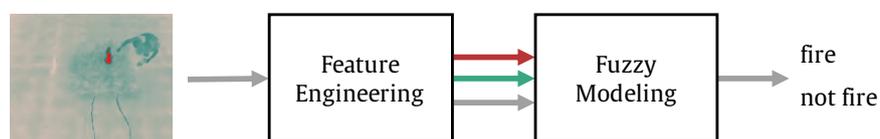


Figure 10.1: Schema of the proposed approach.

This paper is organized as follows: Section 10.2 presents the experiments and data acquired. Section 10.3 introduces a feature engineering process using a color-based approach. Section 10.4 presents the fuzzy inference system used for modeling the sensor behavior, as well as the performance measures employed for model evaluation. In section 10.4.6 the results are analyzed and discussed, and section 10.5 presents the conclusions and suggestions of future work.

10.2 Experiments and Data Acquisition

This section presents the experiments conducted for data acquisition purposes, which includes controlled fire experiments performed in a laboratorial setting, as well as field trials in a real context, namely at the venue of a summer music festival. The experiments conducted in laboratory cover three distinct scenarios of fire risk situations in camping parks. These situations encompass burning of forest fuels like straw, but also include combustible materials characteristic of the camping context. In this section, the environments and setups of each experiment are described in detail, covering the hardware used for each test and the conditions of image capture. Subsequently, the results from data acquisition are presented, with the characterization of the datasets used in this work.

In the following, before delving into the description of the experiments, some preliminary information about thermal imaging systems is required for the understanding of the full reach of the proposed approach.

10.2.1 Thermal Cameras

Although thermal cameras can have very distinct sensory capabilities, these imaging systems can be categorized into two classes: radiometric or nonradiometric. This denomination concerns a characteristic that is intrinsic to each camera model, and refers to its ability to measure thermal radiation, and based on its calibration parameters convert these measurements to temperature values. Figure 10.2 illustrates the difference between thermal radiometric cameras and nonradiometric, regarding the data each type of system can provide.

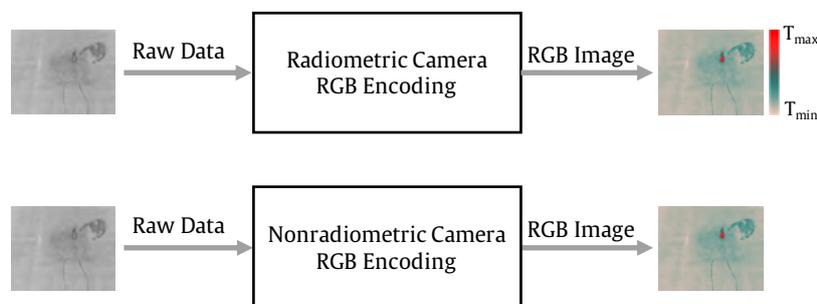


Figure 10.2: Thermal radiometric and nonradiometric cameras.

Note only radiometric cameras provide the temperature scale corresponding to the color values of each pixel. Hence, when thermal imaging systems are not empowered with radiometry capabilities, that represents a significant limitation, because the image data is deprived of quantitative information.

Having direct access to an absolute temperature measurement could give clear information to detect a fire ignition, for example. However, not all cameras possess this capability, but all of these, radiometric or not, provide RGB encoded images in real-time.

The RGB encoded images outputted by the cameras are the result of internal proprietary operations undisclosed by the manufacturer. The colors in the output images accessible to the camera end-user change

depending on the relative temperatures of the scene captured. For nonradiometric camera models the relation between pixel color and temperature is not provided by the manufacturer. For this reason, an image-based solution is necessary for application in real scenarios.

In that sense, this work focuses on developing an intelligent system based on features constructed from RGB encoded images, meaning the aim is to be able to handle thermal imaging data independently of the radiometric capabilities of the camera or lack thereof. By relying only on qualitative information, the proposed approach is suited to be applied in both radiometric and nonradiometric image systems.

10.2.2 Controlled Fire Experiments

Tests concerning fire risk situations that can be encountered on campsites were carried out in indoor laboratory conditions. Using a controlled environment in these experiments allowed for the acquisition of RGB encoded video from static and moving platforms, as well as from different types of thermal cameras, namely FLIR SC660 and FLIR Vue Pro. The main specifications of these two cameras are summarized in Table 10.1.

Table 10.1: Summary of the specifications of the thermal cameras.

	FLIR SC660	FLIR Vue Pro
Size [L x W x H] (mm)	299 x 144 x 147	63 x 44.4 x 44.4
Weight (g)	1800	92.1 - 113.4
Sensor Resolution (px)	640 x 480	336 x 256
Focal Length (mm)	19	9.0
Horizontal FOV (°)	45	35
Vertical FOV (°)	34	27
Spectral Band (μm)	7.5 - 13.5	7.5 - 13.5

In what concerns the spatial resolution and the camera field of view (FOV), note that these cameras have distinct characteristics, thus this will influence the data recorded. Additionally, considering the acquisition of image from a drone, and having in mind the payload budget of these aerial platforms, comparing the weight and size of both cameras, only FLIR Vue Pro has suitable weight and size to be taken onboard a multicopter drone.

Two tests were performed to capture images of straw burning, and to evaluate the response time of the thermal cameras when the fire source is in direct view. For these tests a straw load of $0.6\text{kg}\cdot\text{m}^{-2}$ was used, promoting an average flame height of 30cm. In the first test, both cameras were placed on the static platform. To evaluate the implications of using a moving aerial platform, in the second test, a multicopter was equipped with the FLIR Vue Pro, and short duration flights were performed to record fire ignitions and initial spread. Figure 10.3 depicts this experimental setup, with the drone on the top of the image, and the static platform on the right.

For the situations of controlled fire in the laboratory tests, the data recorded using FLIR SC660 from



Figure 10.3: Experimental setup with a drone and static platform.

the static platform were used as ground truth, due to its radiometric capabilities. This is a reasonable assumption to make, since the basis of this study does not rely on temperature measurements but on image color variations inherent to temperature variations.

The third experimental setup concerns the burning of a camping tent, as depicted in fig. 10.4. For this test, both thermal cameras were placed on the elevated platform, in static conditions.



Figure 10.4: Experimental setup for the burning of a tent.

This test enables the study of fire detection when the fire source is not in plain sight. In this type of situation, it is to be expected that the use of thermal infrared sensing will bring advantages relative to

visible range solutions, by providing an earlier detection.

Note that, in these tests, the average ambient temperature was around 20°C and fuel moisture content was 13%. It will be expected that in a fire situation, with more severe conditions, as is typical in the months of higher fire risk, the fire spread will be faster ($>$ ISI - initial spread index of the Canadian Forest Fire Weather Index).

The complete list of tests performed is summarized in Table 10.2, which includes the controlled fire experiments described in this section, as well as the tests performed during a summer festival that will be presented in the following.

Table 10.2: Summary of the experimental tests.

ID	Description	FLIR	Conditions	Total Frames		FLIR	Conditions	Total Frames	
		SC660	(Platform)	<i>not fire</i>	<i>fire</i>	Vue Pro	(Platform)	<i>not fire</i>	<i>fire</i>
test 1	straw burning	✓	static (platform)	302	1698	✓	static (platform)	302	1698
test 2	straw burning	✓	static (platform)	83	1917	✓	moving (drone)	82	1918
test 3	tent burning	✓	static (platform)	159	1841	✓	static (platform)	159	1841
test 4	summer festival	-	-	-	-	✓	moving (balloon)	1274	42

10.2.3 Field Trials during a Summer Festival

Since campsites are highly dynamic environments, it is crucial to evaluate the thermal sensory data under real operation conditions, so as to be aware of potential false alarms at the initial stage of the algorithm design process.

To that end, field trials were conducted during an outdoor festival midsummer, when conditions of ignition propensity were high due to dry weather, strong winds and high temperatures. During the days of these trials were registered the minimum humidity of 17%, the maximum wind speed of 21 km/h and temperatures of up to 36°C.

As can be observed in Figure 10.5, the natural surroundings coupled with the abundance of combustible material raise fire safety concerns.

For these tests, data were recorded with resource to a tethered balloon at a significantly higher altitude than the previous examples, as illustrated in Figure 10.6.

The data acquisition was performed at different locations of the venue, for periods up to 2 hours, to test experimental conditions covering areas with different characteristics, allowing the collection of a comprehensive and diverse dataset. From these tests, four distinct image sequences were used, in specific two sets of trials for model development, and another two sets of trials for model validation.

10.2.4 Image Dataset

Thermal imaging sensors typically record monochromatic images using a single-channel, which are subsequently encoded in pseudocolor to highlight differences in temperature. For this work, the GrayRed color



Figure 10.5: Partial view of the camping area at the case study summer festival.

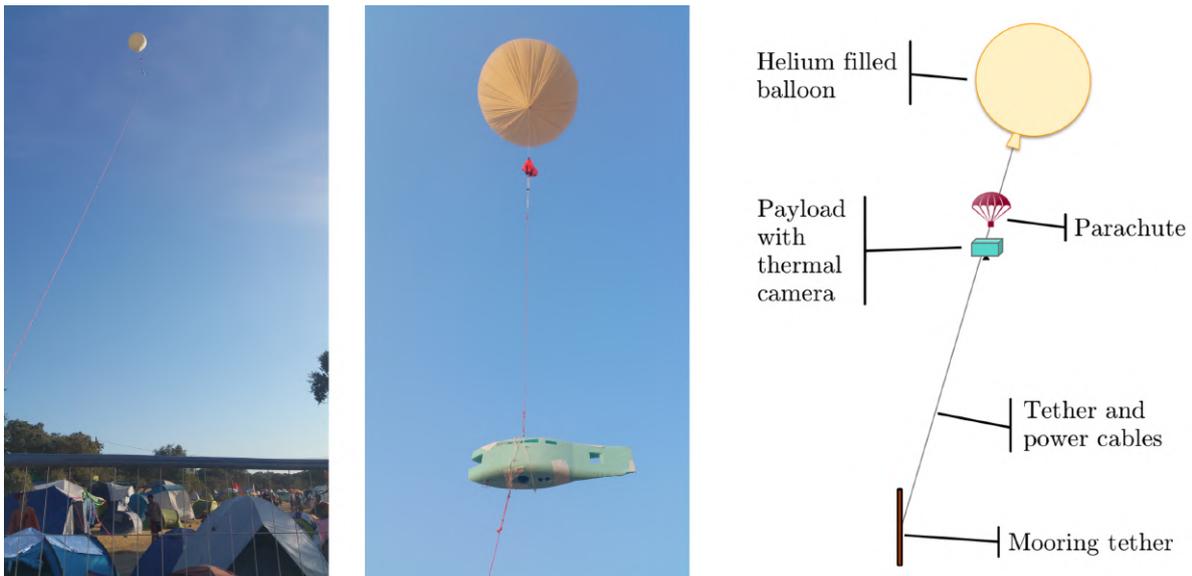


Figure 10.6: Balloon setup.

palette was selected from the list provided by FLIR in the camera firmware, which helps to draw attention to the hottest objects in a scene, by applying high-contrast color values with a divergent color scheme.

While a priori it is unclear if employing this color palette specifically will yield any benefit in classification performance, FLIR preset filters will be used as it facilitates the development of future work, providing a basis for comparison of results. For the same reason, this color palette was applied to the raw data acquired with FLIR SC660, using FLIR ResearchIR software.

Straw Tests

For the controlled ignition of straw, two types of tests were conducted, namely tests 1 and 2. For the first case, videos were recorded with both thermal cameras from a static elevated platform. Fig. 10.7 depicts the data captured with these devices.

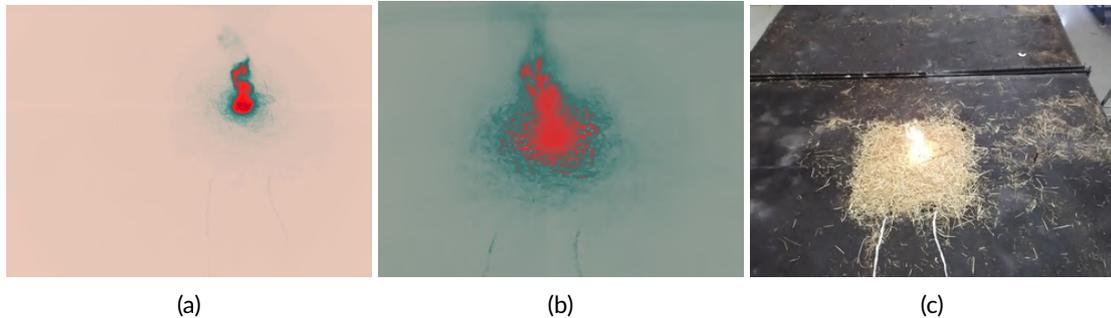


Figure 10.7: Images acquired from the static elevated platform with (a) FLIR SC660, (b) FLIR Vue Pro and (c) visible range camera coupled to FLIR Vue Pro

For the second test concerning the burning of straw (test 2 in Table 10.2), image acquisition was performed from a moving platform using a small unmanned aerial vehicle (UAV). As presented before in figure 10.3, one camera was onboard the multirotor, while the other was positioned on a static elevated platform. Fig. 10.8 illustrates the data from the different points of view.

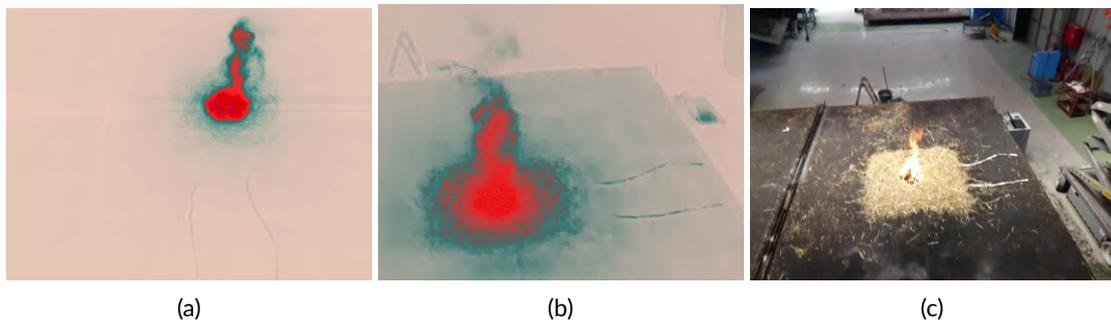


Figure 10.8: Perspectives with 3 different cameras: (a) FLIR SC660, (b) FLIR Vue Pro and (c) visible range camera coupled to FLIR Vue Pro

Both cases presented so far comprise the same type of test, and data were recorded with both cameras, with the only difference being the point of view for the second case. These instances account for situations where the fire is visible to the naked-eye, hence could, in theory, be detected using visible range cameras also. Notwithstanding, one of the main advantages of using thermal imaging is the ability to see beyond the visible spectrum, thus being capable of detecting sources of thermal radiation to which there is no direct view. The following example aims to harness this advantage in another set of conditions.

Tent Test

In a camping environment, fires can originate from inside tents which means it is important to study these cases due to their elevated propagation potential. An important aspect of this type of instances is that the fire sources are obstructed by a somewhat opaque material. To evaluate this situation, a test was

performed with both cameras in static conditions on an elevated platform, as illustrated previously in fig. 10.4. Fig. 10.9 illustrates the differences of data recorded with these devices.

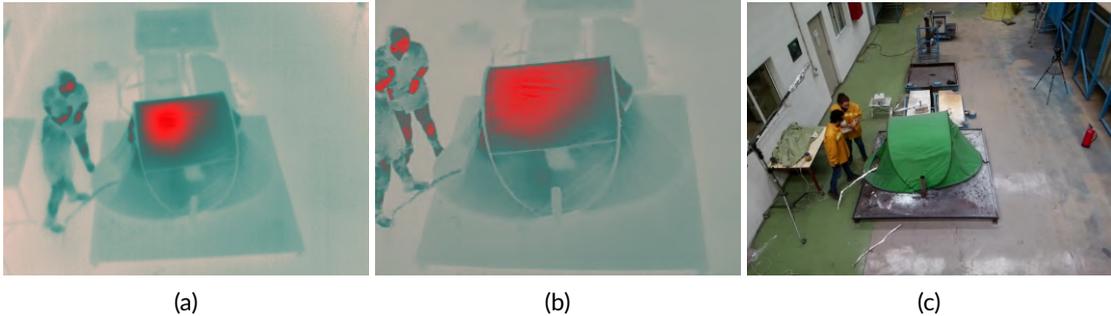


Figure 10.9: Image data from the burning of camping tent with: (a) FLIR SC660, (b) FLIR Vue Pro and (c) visible range camera.

Comparing with the previous tests, now, at an early stage, the area in red does not depict the actual fire, but the heat built-up inside the tent. Whereas this is visible in Figs. 10.9a and 10.9b, for the visible range image (Fig. 10.9c) the fire will only be observable when it starts consuming the exterior of the tent.

Summer Festival Test

For the dataset acquired in the field tests, the images were captured with a 5 second time lapse from a helium balloon, as explained previously. Being this an unactuated aerial platform, the field of view of the camera varies considerably with the effect of the wind. In addition, since the image sequences are taken from a higher altitude, images in this subset can be of difficult interpretation in some cases. Fig. 10.10 illustrates samples that cover different areas of the venue, and where one can recognize tents, caravans, roads, people, trees and rocket stoves in a community kitchen area (semi-circle configuration).

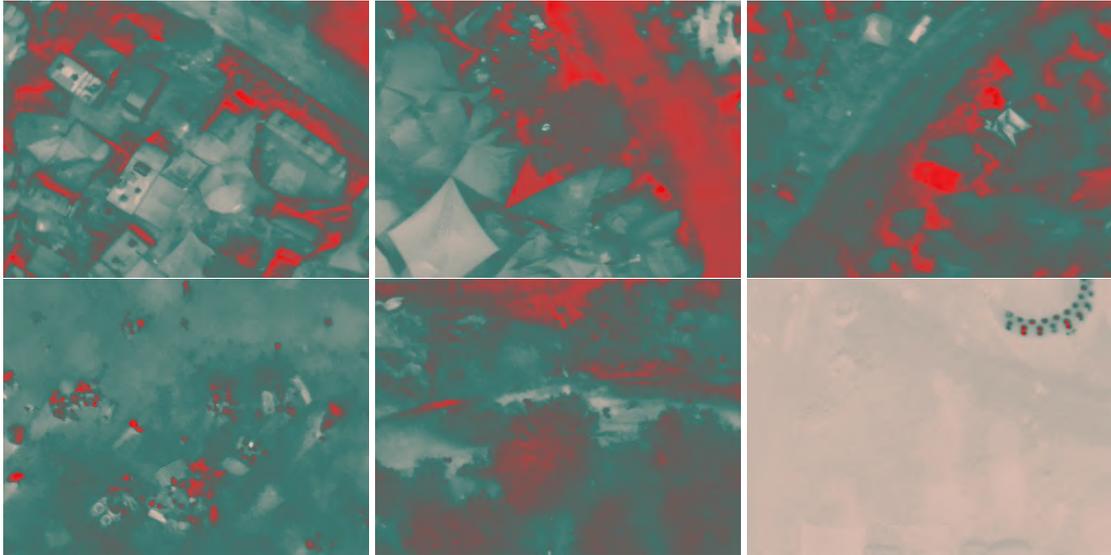


Figure 10.10: Heat detection over the camping area captured with the FLIR Vue Pro.

Note that fire was strictly forbidden on the summer festival premises, except for the community kitchen rocket stoves (depicted in the last image of the bottom row), as well as specific fires lit and controlled by the organization for scenic purposes.

As can be observed in the samples, although the strongest source of thermal radiation is depicted in red, this does not mean it is necessarily fire. Hence, it would be a naive approach to focus only on the detection of the hottest objects in the scene.

In the following sections, to clear misconceptions of this sort, aspects related to thermal images are explored in greater depth.

10.3 Thermal Imaging Analysis

Thermal images in pseudocolor are used to highlight differences in temperature, by mapping the intensity levels according to the color scale of a specified palette. However, this is an adaptive process, in which the color scale adjusts depending on the temperature of objects in the field of view at a given time.

To assess the advantages and limitations of thermal imaging for a fire detection application, it is important to have an understanding of how the thermal sensor responds in the event of a fire ignition.

In that sense, this section starts with an example, exploring on a high-level the tent burning case, which will provide a pathway for the intuitive interpretation of the working of thermal infrared sensors. The insights gained lead to the reasoning that underlies the feature construction process, which in turn will enable further analysis.

10.3.1 Tent Burning Example

This example covers the tent burning test presented previously, and aims to characterize the sensor response to a fire ignition. The study of this case performed in a controlled environment facilitates the interpretation, whilst also giving insight to situations where the fire is not directly visible from the start.

To that end, Figure 10.11 depicts the experiment at different moments since the ignition point, and presents thermal imaging data from FLIR Vue Pro, which are complemented with visible range images, to provide a better understanding of the scene context.

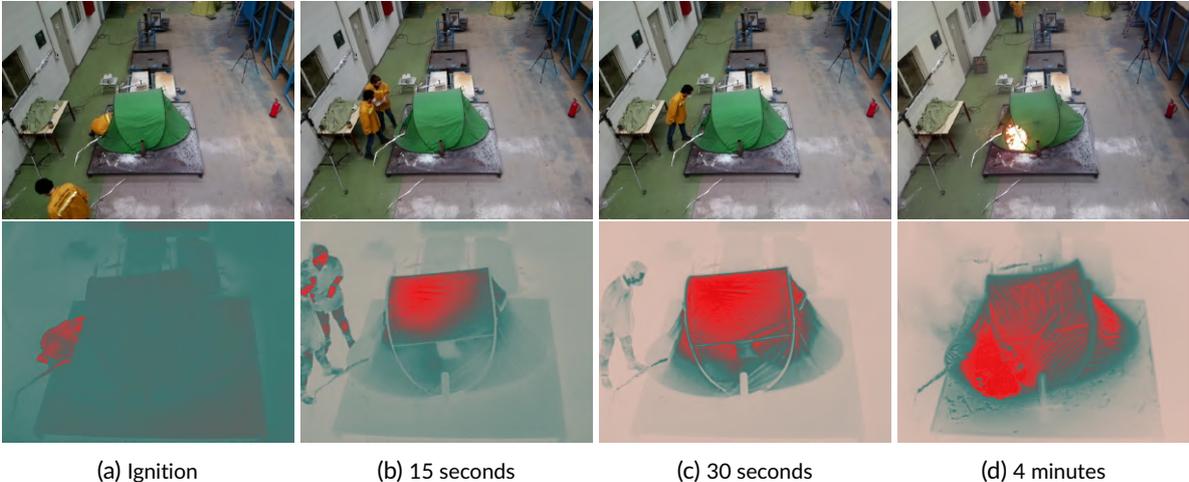


Figure 10.11: Tent burning experiment: color images (top row), FLIR Vue Pro (bottom row).

By observing the image sequence, it is important to notice that the color scale that depicts the differences in temperature adjusts with respect to the object in the scene with the maximum temperature. This adjustment is gradual and can be easily observed in the first three frames, corresponding to the first 30 seconds after the ignition. Whereas, in the first frame the person is represented in red, in the second the person is increasingly depicted with a dark green. In the third frame, the person in the scene is mostly represented with grayish tones. This happens because radiation being emitted by the greatest heat source is measured by the sensor with a much higher intensity than the radiation emitted by the person. While the temperature of the person remained unaltered, the color scale of the camera suffered an adjustment. This makes perfect sense, because now the body temperature is much closer to the ambient temperature than to the temperature of the fire.

In this way, comparing the color spectra of the first and the last frames, it becomes evident that tracking red objects may not be appropriate for a fire detection algorithm based on thermal images encoded with this color palette. However, given the alterations observed in the surroundings, the color change suggests that this type of characteristic can be leveraged to identify the presence of a fire source, for instance by monitoring the shift between green and gray.

Having understood how the changes in the environment influence the data, it is important to design an approach well suited for the task of fire detection. Moreover, taking into account high-dimensional data may be processed onboard in embedded systems, computational complexity is a consideration for the implementation.

Attending to the complete sequence, as the fire develops, the red area of the image increases, while the green shades tend to disappear, being replaced by gray tones instead. Noticing this pattern in the image statistics, we propose a color-based approach to study the dynamic behavior of the sensor. In that sense, the following section describes the feature construction method devised for that analysis.

10.3.2 Feature Engineering

Feature engineering is an essential aspect of designing intelligent systems and plays a crucial role when working with high-dimensional data. To handle the computational burden, dimensionality reduction using feature engineering methods follows two approaches: feature extraction and feature selection. Both approaches attempt to find the set of variables that maximizes the model performance [208]. The main difference between the two is that while feature selection involves choosing the most relevant subset of the available variables, feature extraction consists in applying transformations to construct new features that contain the most useful information [209].

The field of computer vision deals with high-dimensional data, so extensive research has been devoted to feature extraction methods, e.g., for edge detection, region segmentation, and detection of points of interest. In this work, to study the behavior of thermal imaging sensors, and its potential for a fire detection application with expert systems, we will adopt a feature extraction methodology, through color segmentation.

To devise a color segmentation heuristic, the color palette used in the thermal images should be taken into consideration. In this section, we describe the GrayRed color palette and lay out the method of feature construction. This color palette is defined by 120 distinct color values, and depicts a scale of increasing temperature, from gray to red. As Fig. 10.12 illustrates, the color palette was divided into three segments.



Figure 10.12: Division of the FLIR GrayRed palette into color segmentation classes.

The gray segment includes 18 color levels, which represent 15% of the full scale. The mid-range is defined by 48 colors of green tones, which correspond to 40% of the scale. The remainder is the red part, that comprises the last 45% of the color palette, and is defined by 54 color levels. Table 10.3 presents the RGB thresholds established to define the color segmentation heuristic.

Table 10.3: Segmentation thresholds for feature construction.

classes	gray			green			red		
channels	R	G	B	R	G	B	R	G	B
upper limit	253	199	185	143	169	157	255	73	71
lower limit	149	171	160	98	90	86	103	89	85

Although the colors in Table 10.3 are represented in the RGB color space, the segmentation into the three classes (gray, green, red) will help to convey how the color distribution varies over time. In opposition, if this analysis was to be done in terms of absolute values of each color channel (red, green, and blue), the result would be less intuitive and of complex interpretation.

Furthermore, instead of evaluating the color distribution by histogram analysis, the percentage of pixels in each of the segmentation classes will be monitored. Hence, the features that will describe the adaptation of the sensor are the percentage of pixels in red, green and gray.

To assess the merits and limitations of thermal imaging for a fire detection application, the color segmentation heuristic will be evaluated for the situations of the experimental tests.

10.3.3 Data Analysis

This section explores the data obtained for the different scenarios. To that end, the results are analyzed with two distinct approaches. The first addresses how the levels of each segmentation class vary over time, to clarify the RGB encoding done by the cameras. In turn, the second considers the image as a whole, to understand the change in image statistics. In the following, the results of all the tests are presented side-by-side, since it facilitates their interpretation. Yet, by comparing data of images from the two cameras, the effect of the difference in spatial resolution should be taken into account.

Starting with Figures 10.13a and 10.13b, these correspond to the first straw example (test 1) captured with FLIR SC660 and FLIR Vue Pro, respectively. Observing Fig. 10.13a it should be noted that there is

an adjustment in the sensor prior to the ignition. This behavior was caused by the entrance of a person in the field of view of the camera, and demonstrates the environment and the sensitivity of this type of sensors have a considerable effect in the data.

Notwithstanding, observing both graphs (Figs. 10.13a and 10.13b), it is evident the ignition provokes significant changes in the gray and green levels. Considering the ignition occurs around frame 300, it is also noticeable that the FLIR Vue Pro responds with some delay.

Observing the graphs from Figures 10.13e and 10.13f (tent example, test 3), the change in green and gray levels happens, once again, in an almost symmetric fashion. But, for this instance, the adaptation of the color scale to the fire incident takes effect more slowly. The increase in the percentage of red pixels occurs gradually but still at relatively low levels.

Looking at both of these tests, in which images were captured from a static platform, the variation of the color levels develops in an incremental manner, so it would be difficult to establish a threshold based on the change in consecutive frames. Additionally, the percentage of pixels in gray seems the most relevant feature to identify the occurrence of a fire outbreak. Still, the image does not adjust so abruptly in all

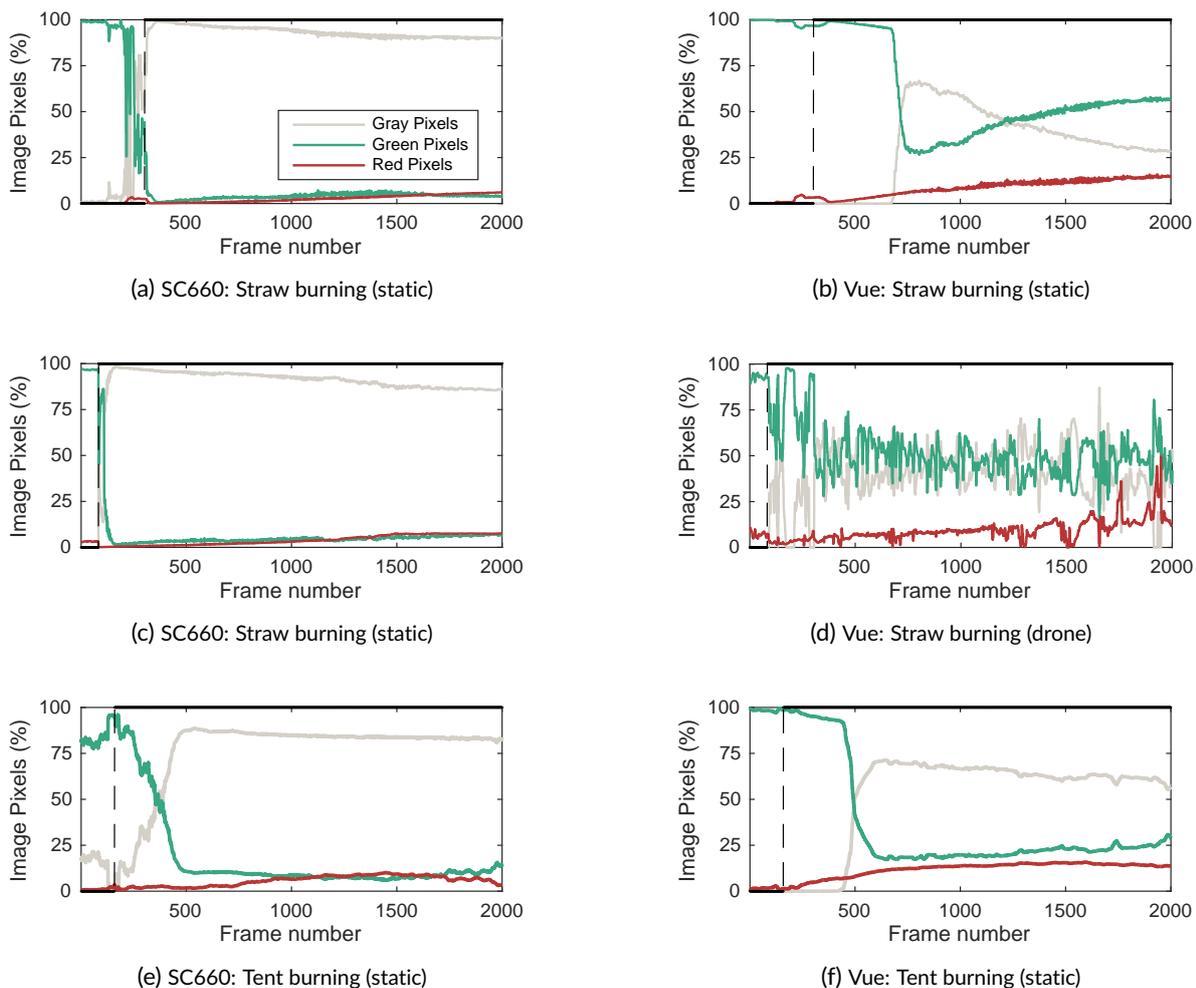


Figure 10.13: Color segmentation results, where the ground truth is represented by the black line, and - - depicts the point of ignition.

cases, so if a low percentage of gray pixels is defined as a threshold, the classification algorithm could be prone to false alarms.

While the results of the examples analyzed so far were obtained in static conditions, introducing the movement of an aerial platform would cause far greater inconsistency in the color levels. This can be observed by comparing Figures 10.13c and 10.13d, which depict the second straw example (test 2) captured from a static platform and a drone, respectively.

In turn, under real operating conditions, this stochastic behavior is confirmed in the tests performed with the tethered balloon. Figure 10.14 presents a selection of the results from the field trials performed, depicting a baseline situation as well as a situation of potential fire detection.

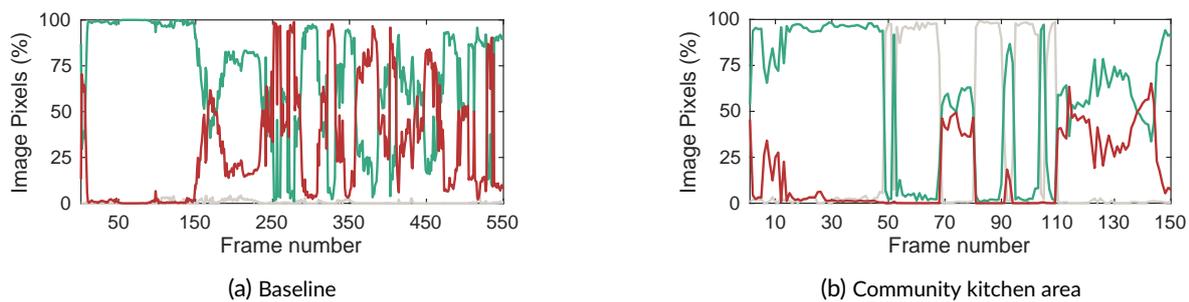


Figure 10.14: Partial results from the festival tests with FLIR Vue Pro.

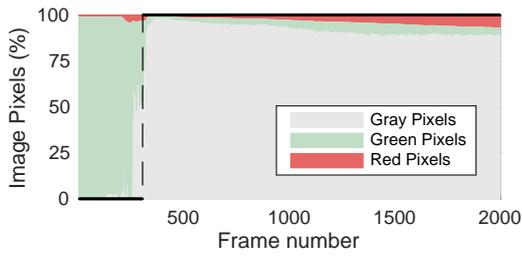
Examining the graphs from Figures 10.14a and 10.14b, relative to the tests at the summer festival with the tethered balloon, the movement of the camera, as well as the set 5 second time lapse, have a considerable effect on the sensor response. Now, the variation is no longer incremental, and consecutive frames display abrupt changes in the percentage of either of the segmented colors.

Overall, the principal insight gained from this analysis is that in response to a fire ignition the gray and green percentages adjust in a complementary way, as the color scale adjusts to the hottest objects in the scene. Furthermore, since the ignition starts with a small area in the image, the red percentage is usually the one with least influence for a fire detection heuristic.

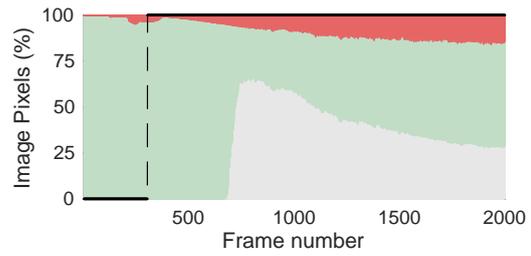
A second approach allows for better intuitions regarding the image statistics. With a different type of visualization, the emphasis is no longer on how the color levels change relative to each other, but rather to the overall aspect of the images. Considering a color segmentation heuristic has to encompass in its thresholds a wide range of scenarios, this enables the overview of percentage ranges for each situation, as can be observed in Figure 10.15.

Observing the results from the static tests, the threshold for gray pixels could have to be established at 50%. But for the case from Fig. 10.15b and Fig. 10.15d would be misclassified. Lowering the threshold value to classify these situations correctly could make the algorithm prone to misclassifications.

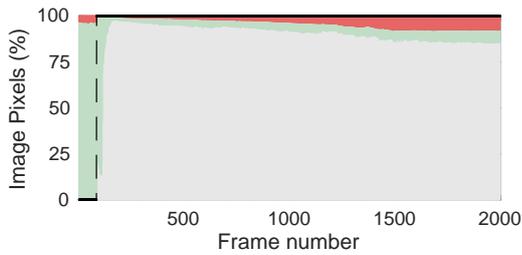
In addition to the previous tests, the segmentation heuristic was also tested on a dataset acquired at the summer festival, which is partially presented in Fig. 10.16. Fig. 10.16a illustrates the baseline situation where the color levels are considerably stochastic, with high red and green color levels, which characterizes



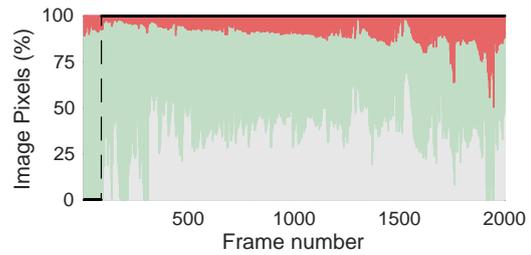
(a) SC660: Straw burning (static)



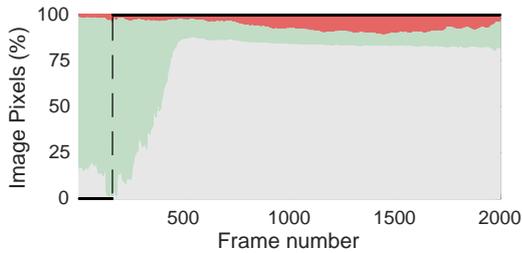
(b) Vue: Straw burning (static)



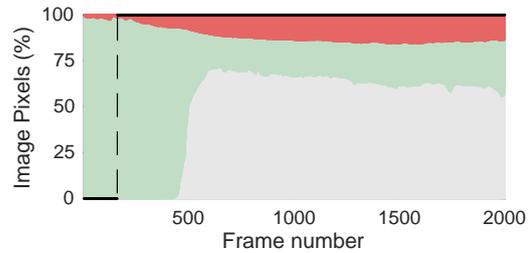
(c) SC660: Straw burning (static)



(d) Vue: Straw burning (drone)



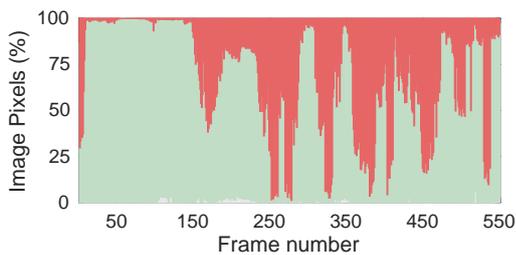
(e) SC660: Tent burning (static)



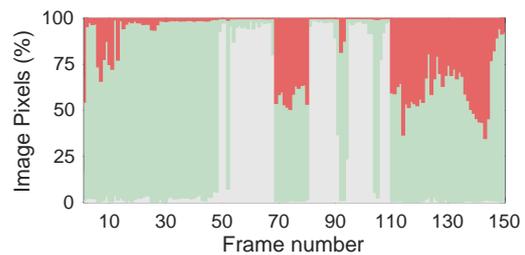
(f) Vue: Tent burning (static)

Figure 10.15: Color segmentation results, where the ground truth is represented by the black line, and - - depicts the point of ignition.

the normal situation. Fig. 10.16b depicts in more detail a period where fire situations are detected.



(a) Baseline



(b) Community kitchen area

Figure 10.16: Partial results from the festival tests with FLIR Vue Pro.

To illustrate that the presence of fire can cause abrupt transitions in the color levels in consecutive frames, a few selected samples are shown in Fig. 10.17.

By inspecting the corresponding sequence of frames presented in Figure 10.17, what stands out is that the color scale adjusts when the same area enters the FOV of the camera. With knowledge of the venue premises, it is possible to confirm this place as the community kitchen area with rocket stoves. This is an

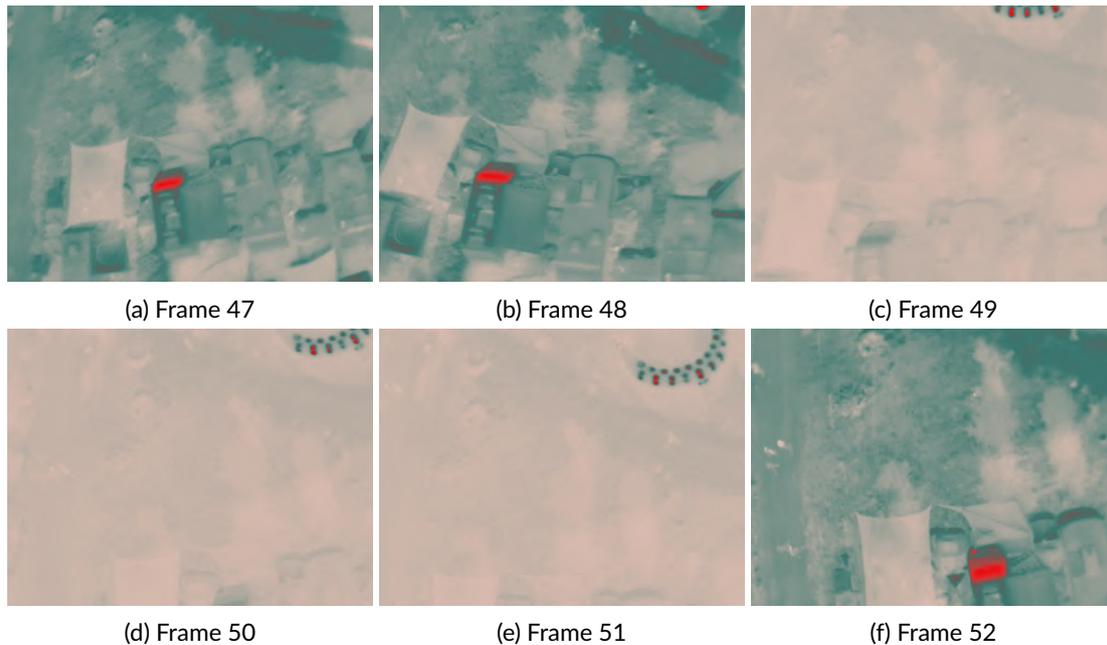


Figure 10.17: Sequence of frames taken over one of the community kitchens with the FLIR Vue Pro.

area where fire detection is to be expected, in addition to other heat sources such as objects at elevated temperatures.

Regarding the color scale, this example, once again, confirms that when the temperature differences in the scene increase, the percentage of green in the image reduces considerably. However, considering this camera was programmed to capture images with a 5 second time-lapse, it is unknown the time the firmware takes to adjust the color scale. Additionally, taking into account the altitude at which the images were captured, this result proves that the color scale adapts even when hot objects have low spatial resolutions.

As has been demonstrated, the variability inherent to dealing with thermal images, which depict relative temperature differences, is a considerable challenge to the development of a robust classification algorithm.

Although a set of crisp heuristic rules could be fine-tuned to work for the situations presented, that type of algorithm would have limited generalization. As the tests from the festival illustrated, real conditions account for much higher variability in image content. Thus, to address this issue we propose an expert system based on fuzzy inference.

10.4 Fuzzy Modeling

Fuzzy modeling is an important method for modeling complex nonlinear behavior in systems for which there is no a priori knowledge of the system or its dynamics can not be derived from first-principles. This type of inference system is based on a rule-base that establishes input-output relations, and allows the interpretation of the rules when model complexity is simple. For this reason, these models are considered

transparent to interpretation when compared to black-box models, and are usually known as gray-box models [210].

The purpose of using a fuzzy modeling approach in this study is two-fold. First, the aim is to investigate the behavior of the sensor and its response to a fire ignition. To that end, a gray-box approach that is considered transparent and can be interpreted may provide greater insight to the inner-workings of thermal cameras than, e.g., black-box models like neural networks. Second, the main objective of this work is to develop a solution that can be applied in real contexts, where operation conditions increase the stochastic nature of the data acquired. The changes in the field of view of the camera and variable capture intervals, can result in a noisy model output, therefore a classification approach should be followed.

This section begins with the formulation of the classification problem to be solved, followed by the description of the fuzzy modeling methodology employed. Subsequently, the performance measures used for parameter estimation and model validation are presented.

10.4.1 Classification Problem

The problem of detecting fire outbreaks is formulated as a binary classification problem, where the negative class is *not fire* (0) and the positive class is *fire* (1). Given a set of N samples, a data vector $\mathbf{z}_k = [\mathbf{x}_k, y_k]$ is defined by 3 input variables $[x_1, x_2, x_3]$, the three features (red, green and gray percentages), and the output, y . The classification threshold applied to the model output is set at 0.5.

10.4.2 Takagi-Sugeno Fuzzy Model

The Takagi-Sugeno (TS) fuzzy model [211] is a rule-based inference system that translates relationships between inputs and outputs by means of "if-then" rules, in which the fuzzy antecedents are formed by logical rules, and the consequent part is a crisp mathematical function.

For each input variable x_j ($j = 1, 2, \dots, n$), the input space is divided into fuzzy regions, which define the antecedent fuzzy sets, A_{ij} , and the consequent functions, $f_i(x)$, describe the behaviour of the system within a given region. Thus, a rule R_i can be expressed as:

$$R_i: \quad \text{If } x_1 \text{ is } A_{i1} \text{ and } x_2 \text{ is } A_{i2} \dots \text{ and } x_n \text{ is } A_{in} \\ \text{then } y_i = f_i(\mathbf{x}), \quad i = 1, 2, \dots, C$$

with C standing for the number of rules. In this work, the fuzzification process follows a data-driven approach, by means of fuzzy clustering in the product space of the inputs and outputs, where the number of rules is given by the number of clusters. In this way, the crisp input variables are mapped to the antecedent fuzzy sets, A_{ij} , represented by membership functions, $\mu_{A_{ij}}(x)$.

For TS-type 1 models, the consequent part of the rules is given by first-order linear functions of type:

$$f_i(\mathbf{x}) = \mathbf{a}_i^T \mathbf{x} + b_i, \quad i = 1, 2, \dots, C \quad (10.1)$$

where parameters \mathbf{a}_i and b_i are determined by least-squares estimation methods in the input-output space.

The inference engine evaluates the degree of fulfillment of the antecedent, i.e., the firing strength of each rule, β_i . To implement the logical connectives of the rules, in this case the "and" operator, the product operator is used:

$$\beta_i = \prod_{j=1}^n \mu_{A_{ij}}(\mathbf{x}) \quad (10.2)$$

where $\mu_{A_{ij}}(\mathbf{x}) : \mathbb{R} \mapsto [0, 1]$ is the membership function of fuzzy set A_{ij} .

In the defuzzification step, the model output, \hat{y} , derives a nonlinear regression model from a collection of local linear models, by combining the contribution of each of the rules. This is performed by the center-of-gravity method, given by the weighted average of its firing strength, β_i , with the consequent functions, $f_i(\mathbf{x})$:

$$\hat{y}(\mathbf{x}) = \frac{\sum_{i=1}^C \beta_i f_i(\mathbf{x})}{\sum_{i=1}^C \beta_i} \quad (10.3)$$

Since this yields a continuous-valued output, for the purpose of a classification task, the crisp values will be subsequently transformed into classification classes, which in the scope of this work is a binary output.

Having explained the basis of the TS model employed, the following describes in further detail the method of fuzzy clustering used for estimating model parameters, namely the Gustafson-Kessel clustering algorithm [49].

10.4.3 Gustafson-Kessel Fuzzy Clustering

In fuzzy models derived from data, the first step for generating the rule-base is to apply a fuzzy clustering algorithm to divide a given dataset $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N]$ into fuzzy subsets. In fuzzy partitions each data point belongs to a cluster according to a degree of membership, μ_{ik} , which enables coping with data uncertainty.

In this way, the fuzzy partition matrix, $\mathbf{U} = [\mu_{ik}]$, defines the degree of membership for each sample in the dataset \mathbf{Z} , and the cluster centers, \mathbf{v}_i , are computed as:

$$\mathbf{v}_i = \frac{\sum_{k=1}^N (\mu_{ik})^m \mathbf{z}_k}{\sum_{k=1}^N (\mu_{ik})^m}, \quad i = 1, 2, \dots, C \quad (10.4)$$

defining the matrix of cluster centers $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_C]$. The fuzziness parameter, $m \in [1, \infty)$, represents the degree of overlap between clusters, with 1 corresponding to hard partitioning where there is no uncertainty in the data.

In this work, TS fuzzy models were constructed using the Gustafson-Kessel (GK) clustering algorithm, which is an extension of the well-known fuzzy c-means (FCM) algorithm [44]. Unlike the FCM, which imposes one shape for all clusters regardless of their actual nature, the GK algorithm uses an adaptive distance measure based on the Mahalanobis distance [49]. For each cluster, the distance function used is

a squared inner-norm, given by:

$$d_{ik\mathbf{A}_i}^2 = (\mathbf{z}_k - \mathbf{v}_i)^T \mathbf{A}_i (\mathbf{z}_k - \mathbf{v}_i) \quad (10.5)$$

where \mathbf{A}_i is a norm-inducing matrix, based on the cluster covariance matrix:

$$\mathbf{A}_i = |\mathbf{F}_i|^{\frac{1}{n}} \mathbf{F}_i^{-1} \quad (10.6)$$

with the fuzzy covariance matrix given by:

$$\mathbf{F}_i = \frac{\sum_{k=1}^N (\mu_{ik})^m (\mathbf{z}_k - \mathbf{v}_i)(\mathbf{z}_k - \mathbf{v}_i)^T}{\sum_{k=1}^N (\mu_{ik})^m} \quad (10.7)$$

In the GK algorithm, matrices $\mathbf{A} = (\mathbf{A}_1, \dots, \mathbf{A}_C)$ are used as optimization variables, thus allowing each cluster to adapt the distance norm to the local topological structure of the data. Allowing the matrix \mathbf{A}_i to vary with its determinant fixed (10.6) corresponds to optimizing the shape of the cluster, while its volume remains constant.

The clustering criterion to minimize is defined by the following objective function:

$$J(\mathbf{Z}; \mathbf{U}, \mathbf{V}, \mathbf{A}) = \sum_{i=1}^C \sum_{k=1}^N (\mu_{ik})^m d_{ik\mathbf{A}_i}^2 \quad (10.8)$$

At the end of each iteration, the partition matrix \mathbf{U} is updated, according to:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^C \left(\frac{d_{ik\mathbf{A}_i}}{d_{jk\mathbf{A}_i}} \right)^{2/(m-1)}} \quad (10.9)$$

until the cost function improvement is less than a specified tolerance, or the maximum number of iterations is reached. For a formal description of the individual steps of the algorithm refer to [82, 210].

Once \mathbf{U} has been determined, the multidimensional fuzzy sets defined point-wise in the rows of the partition matrix are projected onto the antecedent variables, x_j , to obtain the point-wise definition of fuzzy set A_{ij} .

$$\mu_{A_{ij}}(x_{jk}) = \text{proj}_j(\mu_{ik}) \quad (10.10)$$

Then, the antecedent membership functions, $\mu_{A_{ij}}(x_j)$, are approximated by parametric functions, allowing a continuous representation in the domain of the input space.

The estimation of the consequent parameters is usually performed by least-squares (LS) methods [210]. The ordinary least-squares (OLS) method was selected, because it has the benefit of allowing the local interpretation of the consequent parameters, and yielded accurate results for this classification task.

The OLS algorithm estimates the consequent parameters by solving a set of C weighted least-squares problems, where the consequents, f_i , are local models that describe the nonlinear behavior of the system. The consequent parameters are defined by $\theta_i = [\mathbf{a}_i^T, b_i]^T$, and the regressors, given by the pattern matrix

\mathbf{X} , are denoted in extended form by $\mathbf{X}_e = [\mathbf{X}, 1]$. To weigh the degree of activation of each rule, \mathbf{W}_i is defined by a diagonal matrix with the firing strength of each rule, β_i , computed with the membership degrees of the fuzzy partition matrix, as given in (10.2). The OLS formulation is then defined as:

$$\mathbf{y} = \mathbf{X}_e \boldsymbol{\theta} + \epsilon \quad (10.11)$$

with ϵ representing the prediction error.

The least-squares solution is given by:

$$\boldsymbol{\theta}_i = [\mathbf{X}_e^T \mathbf{W}_i \mathbf{X}_e]^{-1} \mathbf{X}_e^T \mathbf{W}_i \mathbf{y} \quad (10.12)$$

if the columns of \mathbf{X}_e are linearly independent and the membership degrees, μ_{ik} , are positive for all samples. For ill-conditioned problems, least-squares estimation is solved in a similar way but the solution is not unique.

10.4.4 Model Evaluation

For estimation of the model parameters, a validity measures approach is followed to determine the adequate number of clusters and fuzziness parameter. For this problem, model selection will favor both classification performance and interpretability.

To rank classifiers, the most widely used metrics are scalar performance measures based on the confusion matrix, namely accuracy, precision, sensitivity or true positive rate (TPR), specificity and false positive rate (FPR). In this way, being the *fire* class the positive class, the images are classified as described in Table 10.4.

Table 10.4: Confusion matrix categories.

Category	Description
True Positive (TP)	classified as <i>fire</i> , ground truth <i>fire</i>
False Negative (FN)	classified as <i>not fire</i> , ground truth <i>fire</i>
False Positive (FP)	classified as <i>fire</i> , ground truth <i>not fire</i>
True Negative (TN)	classified as <i>not fire</i> , ground truth <i>not fire</i>

Table 10.5 presents the formulas of the performance measures employed.

Table 10.5: Formulas of the performance measures derived from the confusion matrix.

Accuracy	Precision	Specificity	Sensitivity	TPR	FPR
$\frac{TP+TN}{TP+FN+FP+TN}$	$\frac{TP}{TP+FP}$	$\frac{TN}{TN+FP}$	$\frac{TP}{TP+FN}$	$\frac{TP}{TP+FN}$	$\frac{FP}{FP+TN}$

However, for binary classifiers, by relying on a fixed classification threshold and being independent of class priors, these measures can be misleading when working with unbalanced datasets. Taking this into account, and recalling the model output (10.3) yields a continuous-valued output, this enables using the

area under the receiver operating characteristic curve (AUC) [212], which allows the comparison of models with variable classification thresholds.

Regarding interpretability, it is key to have in mind that in fuzzy models derived from data, the number of rules is defined by the number of clusters. Thus, the model structure should be kept simple, with a low number of rules, for ease of interpretation.

10.4.5 Dataset Division

From the database acquired in the experimental trials (Table 10.2), a subset was selected to create a better balanced dataset for this classification task, resulting in a total of 5315 samples of the *fire* class, and 3762 of *not fire*. This dataset was split into two distinct subsets with approximately 65% for training the models and 35% for testing. The data from the experiments presented in the previous sections was divided as described in Table 10.6 for the training dataset, and in Table 10.7 for the testing dataset. Note that while the same trial may be used for training and testing, the camera model and the respective frames are considerably different, as demonstrated in the data analysis in section 10.3.3.

The training dataset is evenly divided, and includes an instance of each of tests presented beforehand. To enable the models to learn an extended operating range of the sensor, i.e., to cover the complete extent of the input space, situations from a real-context were included in the training set. From the several tests performed at a festival, the data were split between the training and test datasets, and two additional tests (5 and 6) in similar dynamic scenarios were included in the training set.

Table 10.6: Summary of the samples in the training dataset.

ID	Description	<i>not fire</i>	<i>fire</i>
Test 1	SC660: Straw burning (platform)	302	698
Test 2	Vue: Straw burning (drone)	82	1418
Test 3	SC660: Tent burning (platform)	159	641
Test 3	Vue: Tent burning (platform)	159	401
Test 4	Vue: Summer festival (balloon)	616	-
Test 5	Street (static)	1114	-
Test 6	Aerial view (moving)	287	-
Training dataset (%)		2719 (46.3%)	3158 (53.7%)

For the testing dataset, the cases selected include the situations that are more relevant to evaluate the performance of the model, using for validation part of the trials at the festival under the real operation conditions.

Table 10.7: Summary of the samples in the testing dataset.

ID	description	<i>not fire</i>	<i>fire</i>
Test 1	Vue: Straw burning (platform)	302	1698
Test 2	SC660: Straw burning (platform)	83	417
Test 4	Vue: Summer festival (balloon)	658	42
Testing dataset (%)		1043 (32.6%)	2157 (67.4%)

10.4.6 Parameter Estimation

To define the model structure a grid search was performed by varying the number of clusters, C , between 2 and 8, and the fuzziness parameter, m , between 1.2 and 3.5 with 0.1 increments. Figure 10.18 presents the accuracy results obtained for 500 iterations using the ordinary least-squares (OLS) method for the estimation of the consequent parameters.

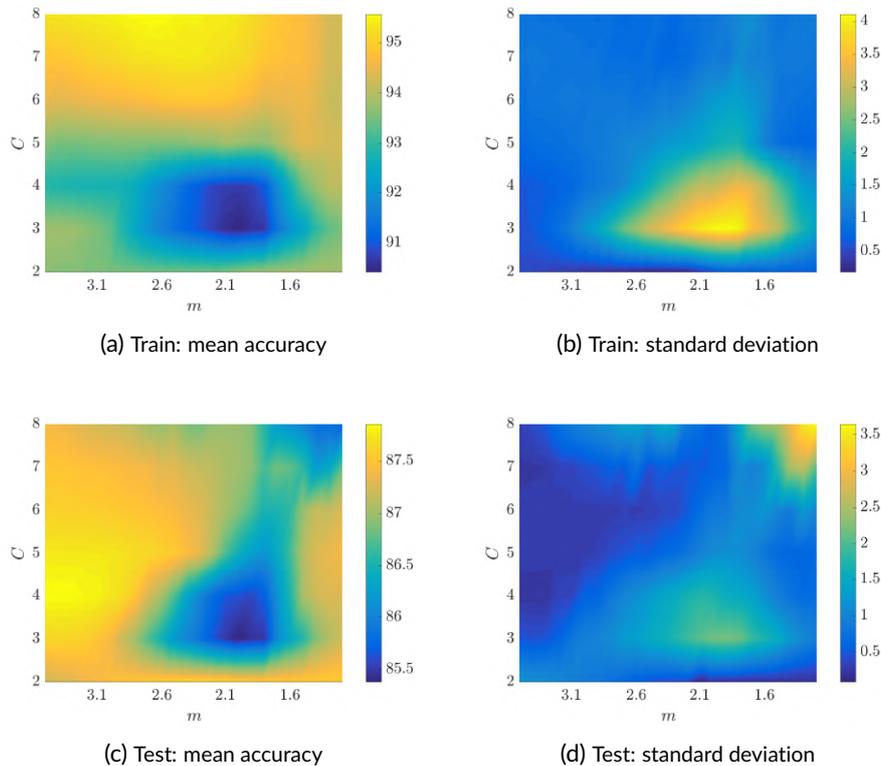


Figure 10.18: Accuracy results for 500 iterations.

The training results show that the models with highest mean accuracy have increased complexity with 6 to 8 clusters. However, by observing the distribution of the results it should be noted that the area with lowest mean accuracy in training (Fig. 10.18a), is the one with highest standard deviation (Fig. 10.18b). This indicates there are low complexity models which can achieve high accuracy as well.

Regarding the test results, models with 4 clusters have the highest mean accuracy (Fig. 10.18c), but considering the values of standard deviation (Fig. 10.18d), models with 3 and 8 rules can have similar performance. Taking these aspects into account, a selection of models with good performance is presented in Table 10.8.

While these measures give an indication of the accuracy of the models, the best models were selected considering also the output response. Overall, the models achieve similar performance in training and generalization wise.

This section starts by presenting the classification performance of models selected with 3, 4, and 8 clusters, which are compared with a univariate threshold heuristic, as suggested in section 10.3.3. Subsequently,

Table 10.8: Model benchmarking for selected parameters (mean and standard deviation values of a total of 500 iterations).

Dataset	C	m	Accuracy (%)	Precision (%)	Specificity (%)	Sensitivity (%)	AUC
Train	3	2.1	90.51±3.94	88.43±6.95	84.53±10.29	95.67±1.57	0.9607±0.0246
	4	3.2	92.72±0.73	90.93±1.72	88.81±2.36	96.09±0.82	0.9700±0.0074
	8	2.6	95.55±0.80	96.22±1.50	95.61±1.88	95.51±0.51	0.9841±0.0062
Test	3	2.1	85.48±2.07	96.11±6.95	92.83±7.14	81.93±0.40	0.9133±0.0174
	4	3.2	87.81±0.34	99.83±1.72	99.70±0.87	82.06±0.31	0.9256±0.0169
	8	2.6	86.95±1.38	99.85±1.50	99.74±0.87	80.77±2.03	0.9646±0.0094

section 10.4.8, explores the transparency of a selected fuzzy model with three rules to provide a better understanding of the behavior of thermal cameras, presenting in detail the parameters of the model. Moreover, the outputs of the fuzzy models before applying classification thresholds are analyzed for a set of training and testing cases.

10.4.7 Performance Evaluation

To assess the models proposed following the fuzzy modeling approach, the classification performance will be compared to a classifier with a linear heuristic threshold based on the gray percentage, using the dataset division presented in section 10.4.4.

Based on the observation on the sensor response using the proposed features, as mentioned in section 10.3.3, a simple classifier could be devised by establishing a heuristic to detect a fire outbreak when the gray percentage exceeds a given threshold (δ). The sensitivity analysis of this approach is presented in Table 10.9.

Table 10.10 presents the classification results of three fuzzy models selected. These models have very similar performance in training and generalization wise, despite having a different model structure.

Table 10.9: Results of the gray heuristic for different thresholds (δ), and optimum threshold values (δ^*) for the datasets previously used.

Dataset	Samples	δ^*	δ	Accuracy (%)	Precision (%)	Specificity (%)	Sensitivity (%)	TPR (%)	FPR (%)	AUC
Train	3157 (F)	0.232	0.500	75.91	97.18	98.09	56.79	56.79	1.91	0.9613
	2720 (NF)		0.232	94.18	96.13	95.66	92.84	92.84	4.34	
			0.065	80.42	74.26	60.88	97.24	97.24	39.12	
Test	2157 (F)	0.065	0.500	58.34	100.0	100.0	38.20	38.20	0.00	0.8606
	1043 (NF)		0.232	87.09	99.89	99.81	80.95	80.95	0.19	
			0.065	88.00	99.78	99.62	82.38	82.38	0.38	

Table 10.10: Results of selected fuzzy models with $\delta = 0.5$.

Dataset	Samples	C	m	Accuracy (%)	Precision (%)	Specificity (%)	Sensitivity (%)	TPR (%)	FPR (%)	AUC
Train	3157 (F)	3	2.1	94.33	95.26	94.56	94.14	94.14	5.44	0.9818
	2720 (NF)	4	3.2	94.93	96.79	96.40	93.66	93.66	3.60	0.9825
		8	2.6	96.19	97.38	97.02	95.47	95.47	2.98	0.9909
Test	2157 (F)	3	2.1	87.50	99.89	99.81	81.55	81.55	0.19	0.9186
	1043 (NF)	4	3.2	87.34	99.89	99.81	81.32	81.32	0.19	0.9591
		8	2.6	86.66	99.94	99.90	80.25	80.25	0.10	0.9740

Considering the heuristic based on the gray levels, comparing the results of Table 10.9 for different thresholds, it is clear this approach is sensitive to the threshold values. Furthermore, while in the datasets employed the threshold defined at 0.232 works well akin to the results presented for the fuzzy models, this linear approach does not offer the same robustness. In different scenarios, the threshold required for satisfactory classification performance could easily vary, as the lower AUC for the test dataset indicates.

Regarding the proposed fuzzy modeling approach, as indicated in Table 10.10 the models achieve good training performance, with a slight decrease in testing but on the same order for all models. Therefore, this limitation should not be simply attributed to generalization ability of the models.

Attending to the test results, it is important to take into account that the data from the controlled fire experiments were captured using a frame rate of 15 fps. Having a high frame rate enables getting the full picture of the behavior of the response of the thermal sensor to a fire ignition, but hinders the classification performance of the classifier in situations where the sensor has a delay in response.

From literature review, there are various computer vision and expert systems approaches taken to the fire detection problem, but most employ algorithms in indoor spaces, which is not the focus of this application. Considering works aiming towards fire detection in large outdoor environments using visible range images, [213] proposed a hierarchical pixel-based algorithm, obtaining a 93.98% TPR and a 4.56% FPR. [214] propose an algorithm coupling logistic regression and temporal smoothing using features of size, motion, and color information, achieving a 96.39% TPR and a 0.00% FPR. [99] present an algorithm using both thermal and visible range sensor data, based on computer vision techniques such as dynamic background subtraction and histogram-based segmentation, obtaining a 91.06% TPR and a 1.35% FPR.

While these works attain satisfactory results, their effectiveness is mostly demonstrated in images where fire is detected at a close-range. Towards applications that require detection at long-distance, namely using unmanned aerial systems, individual pixel-level classification approaches may be difficult to implement when covering large areas, because the features in image data become more difficult to discern with increasing distance.

A key objective of the proposed fuzzy modeling approach is to gain insight into the sensor response. Therefore, for a better trade-off between classification performance and interpretability, the model selected for analysis has 3 clusters and a fuzziness parameter of 2.1. Since the performance measures of this model are within the range of the standard deviation presented in Table 10.8, the performance is not significantly compromised by the decrease in model complexity. The response of this model is compared with the models with 4 and 8 clusters presented in Table 10.10 further along.

10.4.8 Model Parameters

The model obtained by means of fuzzy clustering derived the antecedent membership functions illustrated in Fig. 10.19, with the cluster centers presented in Table 10.11.

The clusters identify three distinct fuzzy regions with a great degree of overlapping, which means the

Table 10.11: Cluster centers.

	red	green	gray
cluster	x_1	x_2	x_3
1	3.468	13.704	82.828
2	8.490	47.633	43.877
3	28.563	65.046	6.391

nonlinear relationships would not be trivial to obtain from simple heuristics. For cluster 1, the core of the membership functions (MFs) is defined for low values of red and green, and high values of gray percentages. For cluster 2, the core of the red MF is defined for low values, and is in the mid-range for green and gray, with the support of these two MFs covering the full range with some degree of membership. For cluster 3, data points of the red and green features can take any value and low gray percentages.

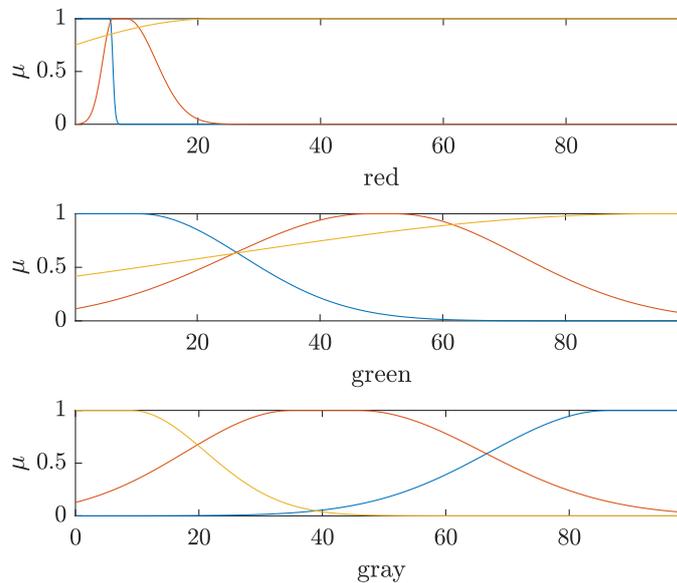


Figure 10.19: Antecedent membership functions for cluster 1 (blue), cluster 2 (orange) and cluster 3 (yellow).

The consequent parameters obtained through least-squares estimation are summarized in Table 10.12.

Table 10.12: Consequent parameters.

rule	red	green	gray	offset
1	$1.1 \cdot 10^{-2}$	$4.2 \cdot 10^{-3}$	$1.0 \cdot 10^{-2}$	0.0
2	$6.2 \cdot 10^{-3}$	$8.7 \cdot 10^{-3}$	$1.1 \cdot 10^{-2}$	0.0
3	$-3.5 \cdot 10^{-3}$	$1.0 \cdot 10^{-3}$	$2.1 \cdot 10^{-2}$	0.0

For rule 1, the variables with strongest influence are the red and gray percentages, with the green having lower order. For rules 2 and 3, the gray percentage has the strongest contribution. Since the MFs are defined for positive values, the negative influence of the red parameter in rule 3 may indicate this rule could be associated with *not fire*, the negative class (0).

The fact that the gray percentage has a strong contribution in all the rules confirms this variable may be associated to *fire* instances as discussed in the analysis in section 10.3.3. Nonetheless, the influence of the red and green percentages in the three rules is evidence a multivariate algorithm may be more appropriate to model this system for a fire detection application.

The fuzzy model is derived for detecting fire outbreaks is defined as:

R_1 : **If** x_1 **is** A_{11} **and** x_2 **is** A_{12} **and** x_3 **is** A_{13} **then**:

$$f_1(\mathbf{x}) = 1.1 \cdot 10^{-2}x_1 + 4.2 \cdot 10^{-3}x_2 + 1.0 \cdot 10^{-2}x_3$$

R_2 : **If** x_1 **is** A_{21} **and** x_2 **is** A_{22} **and** x_3 **is** A_{23} **then**:

$$f_2(\mathbf{x}) = 6.2 \cdot 10^{-3}x_1 + 8.7 \cdot 10^{-3}x_2 + 1.1 \cdot 10^{-2}x_3$$

R_3 : **If** x_1 **is** A_{31} **and** x_2 **is** A_{32} **and** x_3 **is** A_{33} **then**:

$$f_3(\mathbf{x}) = -3.5 \cdot 10^{-3}x_1 + 1.0 \cdot 10^{-3}x_2 + 2.1 \cdot 10^{-2}x_3$$

To compute the model output, the influence of each rule is weighted with the degree of fulfillment as given in eq. (10.3):

$$\hat{y}(\mathbf{x}) = \frac{\beta_1 f_1 + \beta_2 f_2 + \beta_3 f_3}{\beta_1 + \beta_2 + \beta_3}$$

In the following, the model output, \hat{y} , will be assessed and compared with the output of the models with 4 and 8 rules presented in table 10.10.

10.4.9 Training Results

In addition to the model evaluation results presented in Table 10.10, the output of the models will be analyzed before applying the classification threshold of 0.5.

From the training dataset, two cases were selected, namely the drone test in a controlled fire environment that depicts a fire situation, and a test without fire from the summer festival. Fig. 10.20 presents the output results juxtaposed with the samples over time.

Observing Fig. 10.20, given the high accuracy of the models for this test, a fuzzy modeling approach clearly outperforms a heuristic based on a threshold of 0.5 (δ) applied to the gray variable. Whereas for the heuristic most of the samples would be misclassified, for the output of the fuzzy models most of the samples are correctly classified, after applying a 0.5 threshold. Additionally, the fuzzy models cope better with the stochastic variation of the samples, leading to a more stable classification output.

For a trial performed at the summer festival, in a situation without fire, the output results are illustrated in Fig. 10.21.

This example illustrates the classification ability of the fuzzy models for situations in which the gray percentages have low representativity or are negligible, which are modeled by rule 3. While this rule could be difficult to understand due to the high overlap in the antecedent membership functions, this example

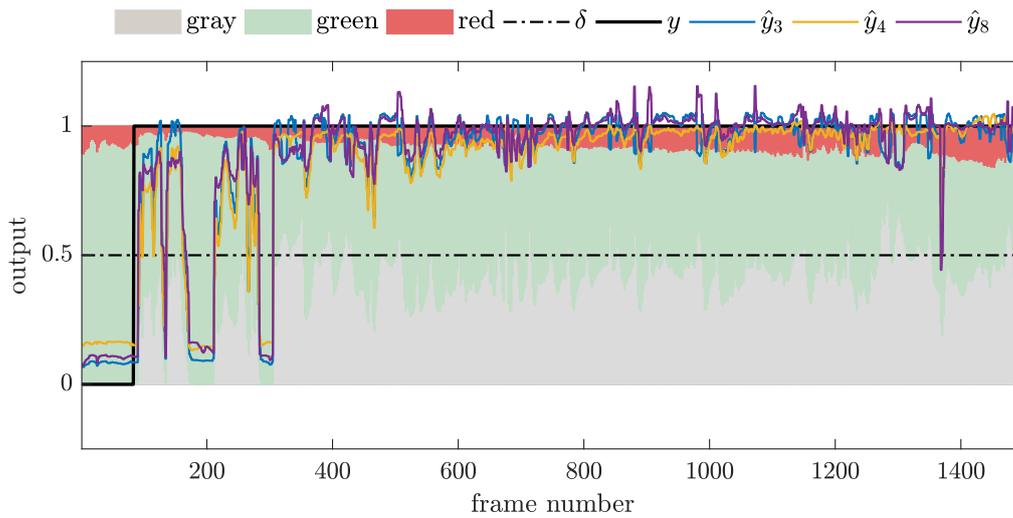


Figure 10.20: Training results: drone test.

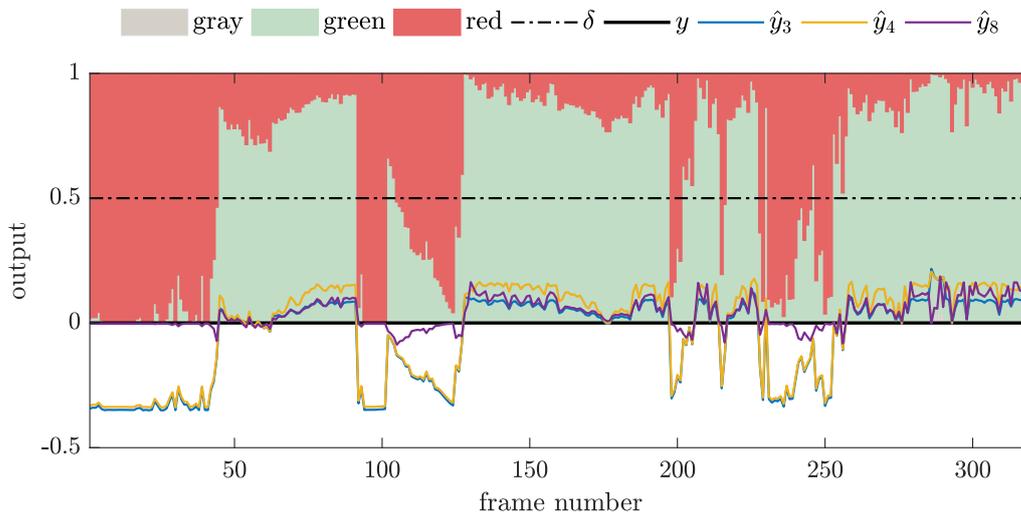


Figure 10.21: Training results: festival without fire (baseline).

clarifies that it has a transparent and intuitive interpretation.

10.4.10 Testing Results

To assess the generalization of the models, two cases will be discussed: a controlled fire case where images were captured from a static platform, and a trial from the festival for validation with a real context case.

Fig. 10.22 depicts the results for a controlled fire case under static conditions. In this case, the delay in response of the sensor causes a considerable part of the samples to be misclassified (approximately from 302 to 698), but afterwards the fuzzy models are able to make accurate classifications.

Moreover, the model with 8 rules starts having worse performance than the models with 3 and 4 rules, with the decrease in gray percentage. Once again, for this case, a heuristic based on just the gray variable

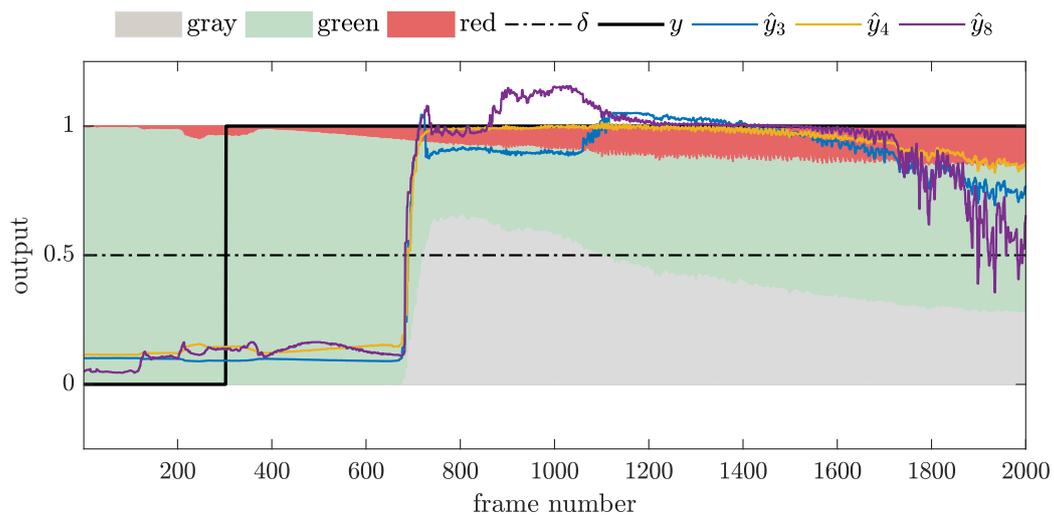


Figure 10.22: Testing results: platform (Vue).

would fail to classify most instances.

For the validation of the proposed models under real conditions, the model output is evaluated on a trial from the festival where fire is detected, as illustrated in Fig. 10.23.

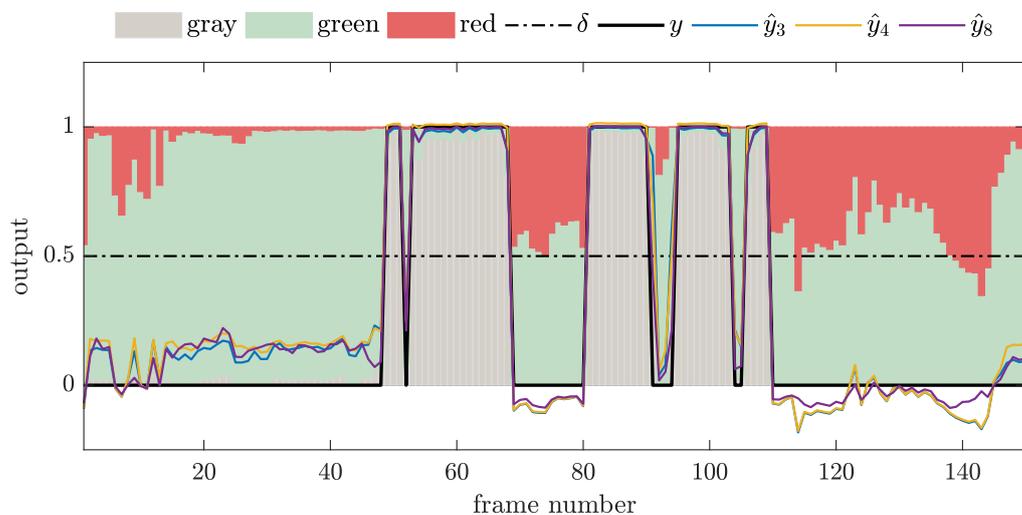


Figure 10.23: Testing results: outdoor festival (test 4).

As was to be expected, when there is a sudden increase in the gray percentage the models identify the presence of fire. This behavior is associated with rule 1, as can be inferred from the MFs (Fig. 10.19) and consequents (Table 10.12).

Considering the output results from the tests presented in this section, the classification is not impacted significantly by the decrease in model complexity, with the model with 3 rules performing on par with the models with 4 and 8 rules.

The examples demonstrate that a transparent fuzzy model with three rules is capable of classifying fire

outbreaks despite the adaptive nature of the color scale. Moreover, the transparency of the model enabled the identification of three fuzzy regions derived from data that are easy to interpret and correlate with the input variables for a set of different situations.

Attending to these results, as well as the overall performance (Table 10.10), the fuzzy models are able to generalize well from known patterns and classify accurately fire outbreaks in dynamic scenarios in real contexts.

10.5 Conclusion

The in-depth analysis of thermal imaging data provided key insights on the inner-workings of thermal cameras, which were instrumental for feature engineering and the construction of features relevant to the fire detection problem.

The proposed fuzzy modeling approach derived from data achieved high performance and generalization, with low complexity models. The transparency of the models created enabled a clear understanding of the data patterns modeled and can be useful in the development of future work with this type of sensors. One of the main advantages of employing this data-driven approach is that it is a flexible method that allows model derivation based on data from different cameras (radiometric or nonradiometric), under different setups and conditions.

Moreover, the features constructed from the thermal images demonstrate this type of data can be valuable for the detection of fire outbreaks. Evaluating by the high specificity and precision of the models, false positives are not a concern for this case, and the lower sensitivity is due to great extent to a 20 second delay in the response of the sensor, which does not compromise the potential application, given that the detection of a fire event, even with this delay, would still be valuable in a real application.

While some delay may be expected in thermal imaging sensors, the capture modes, such as video versus single frames with a specified time-lapse, have a significant influence in the adaptation of the filter parameters, which are based on the camera firmware and calibration parameters. Further investigation should be conducted to characterize this type of sensors, namely in what concerns saturation levels and the color encoding schema.

Considering there are still many degrees of freedom to explore in this application, it is worthwhile to employ soft sensors approaches using thermal imaging sensors. In this respect, the fuzzy modeling approach, due to its interpretability and generalization, provides clear advantages if compared to other research avenues.

V | Data Curation Approaches

Contents

11 Expert-in-the-loop Data Annotation	185
12 Color-based Linguistic Fuzzy Models	191

Summary

This part delves into data curation approaches, opening the scope of this issue to more broadly address fire management tasks in general, while also providing an example of its application to fire data annotation. Chapter 11 proposes an expert-in-the-loop approach to address automated data curation towards scaling the application of data-driven methods in safety-critical technologies in fire intelligence domains. Chapter 12 showcases a purpose-built data annotation pipeline for fire images that combines clustering-based algorithms with interpretable linguistic fuzzy models derived from expert knowledge.

11 | Expert-in-the-loop Data Annotation

Contents

11.1 Introduction	186
11.2 Image-based Wildfire Management Tasks	187
11.3 Proposed Approach	187
11.4 Conclusion and Future Work	189

The material included in this chapter was previously first introduced and presented as a proposal, with double-blind peer-review, at the Tackling Climate Change with Machine Learning workshop at NeurIPS 2020.

Presented at:

M. J. Sousa, A. Moutinho, M. Almeida, **Expert-in-the-loop Systems Towards Safety-critical Machine Learning Technology in Wildfire Intelligence**. NeurIPS 2020 Workshop on Tackling Climate Change with Machine Learning. (2020).

11.1 Introduction

Wildfires are a recurrent natural hazard on a global scale that has a brutal impact on the environment and natural ecosystems, which can lead to disasters with dire impacts on communities [215]. As a result of climate change, fire events are becoming more frequent and severe, with meteorological conditions of high ignition propensity being more frequently met, leading to increased fire spotting, and rapid spread. In addition, these conditions are verified over longer periods, extending fire seasons in several regions worldwide.

To prevent and mitigate the devastating effects of wildfire events, it becomes urgent to detect fires in an early stage and to monitor wildfires in near-real-time as the events unravel, providing enhanced situational awareness for decision-making and operational teams. In that sense, wildfire intelligence plays a pivotal role, especially for high-risk areas such as wildland-urban-interface regions [216] and large-scale wildfires [217]. For these reasons, there is a current demand for improvements and increased levels of automation in the stages of pre-fire event, firefighting, and aftermath.

In this context, the breakthroughs in machine learning (ML) can be an enabling technology towards the integration of artificial intelligence products in current decision support systems. More specifically, related works have proposed ML solutions for image-based fire detection tasks [218, 219], however the quality and limited size of image databases available often do not offer generalization guarantees for reliable deployments in real contexts [114]. Although transfer learning and data augmentation techniques have been explored in related work, the limitations in interpretability and transparency of black-box models prevent effectively fine-tuning these models to solve existing shortcomings [114]. The lack of large-scale databases for wildfire detection and monitoring tasks is a known hurdle in developing machine learning algorithms with adequate generalization.

Considering that employing ML solutions for wildfire intelligence involves deployments in safety-critical applications, the robustness and reliability of the models have yet to undergo significant developments. Conversely, this may be achieved through high-quality data curation, despite not alleviating black-box limitations, or through the exploration of algorithms with increased explainability, fine-tuning ability and interpretability. However, both these scenarios call for the input of expert knowledge to develop accurately annotated data, which is particularly nuanced in the field of wildfire management and operations, and requires domain expertise as there is also high data uncertainty.

To address these issues, we propose the development of expert-in-the-loop systems that combine the automation advantages provided by machine learning with the introduction of relevant domain knowledge expertise. The main objective of the proposed approach is the development of semi-automated software tools that can support data curation by wildfire experts. This solution can allow a better handling of data uncertainty and improve the quality of data sources, with the ultimate objective of enabling the development of large-scale databases for wildfire-related problems. More importantly the expert-in-the-loop approach allows involving domain experts and end-users (*e.g.*, researchers and public agencies) in

the development procedure, thus improving the relevance of machine learning applications developed for wildfire intelligence in real contexts.

11.2 Image-based Wildfire Management Tasks

Image data for wildfire-related tasks can have a broad spectrum of characteristics and modalities as these can be collected *e.g.*, from satellites, aerial vehicles, watchtowers, or ground teams. The latency associated with these data types also varies and consequently, so does its timescales of application. From this plurality arises a great breadth of opportunities for ML approaches to deal with high-dimensional data, but also a great challenge for data curation. Herein, we outline several relevant tasks followed by a brief data description.

Tasks Wildfire management involves four main stages: *i)* prevention, *ii)* preparedness, *iii)* response and *iv)* recovery. The following tasks can extend to several of these stages and involve processing of large amounts of image data, thus having a high potential interest and benefit from the usage of ML.

- Risk assessment concerning environmental conditions and risk mapping based on land use and social patterns;
- Vegetation management to reduce fire severity, *e.g.* fuel mapping, monitoring of fuel breaks, or tracking of vegetation fuel moisture content;
- Wildfire detection and monitoring, *e.g.*, early identification of flames and smoke plume, mapping of the firefront(s), early detection of spot fires and identification of hot spots;
- Post-event analyses, *e.g.*, mapping burned areas, evaluation of possible subsequent cascading effects, *e.g.*, erosion risks, and air quality estimation based on remote sensing.

Data The evolving datasets being developed comprise multimodal image data, currently in the visible and thermal infrared bands, including images captured from ground teams, watchtowers, aircraft and high-altitude balloons, which are exemplified in Fig. 11.1 (for visible range instances only). The image samples comprise a diverse collection of situations, with most concerning real wildfire events and field experimental burns. To create a balanced and robust set of data, factors that may induce misclassifications are also included such as clouds, fog or sunsets, as well as firefighting vehicles, power lines and various types of operational teams.

11.3 Proposed Approach

The development of expert-in-the-loop systems aims to bridge the gap between machine learning automation (*e.g.* in classification, segmentation, or detection tasks) and the inclusion of relevant domain knowledge expertise, so that data curation is relevant for wildfire intelligence in real-world scenarios and



Figure 11.1: Samples of fire and not fire instances captured from ground teams, drones and high-altitude balloons, along with aerial vehicles and teams in operational missions in a wide variety of scenarios and lighting conditions.

wildfire science research. Previous contributions in the literature have favored techniques that are well-suited for real-time deployments despite having limitations in performance, in lack of transparency and interpretability. However, those limitations hamper considerably the reliability and acceptance of end-users of such techniques, hindering the deployment in safety-critical applications in real contexts.

This novel approach in wildfire related applications leverages the potential of exploring machine learning and computer vision, and intelligent systems methods that have not been particularly designed for online computational performance (in terms of speed/energy) in real-time deployments, but are rather accurate despite computationally heavy and/or time-consuming. To harness the advantages of this approach, this project aims to design expert-in-the-loop systems and develop software tools that can introduce domain knowledge into the data curation and task design processes. To that end, this approach can be outlined as exemplified in the diagram in Fig. 11.2.

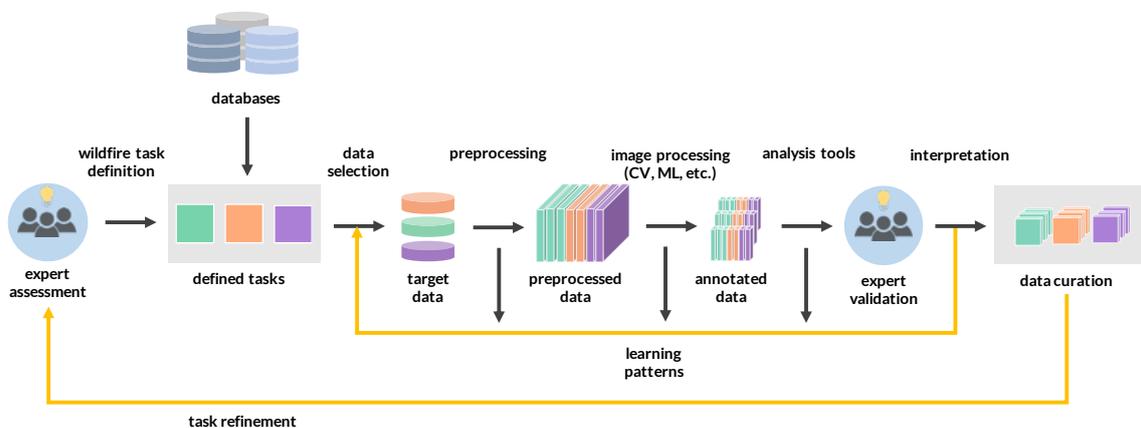


Figure 11.2: Expert-in-the-loop system comprising a computational data annotation pipeline with expert feedback.

Domain expertise is introduced in two stages through two feedback loops: 1) task definition/refinement – where experts define relevant tasks and refine these based on the results of the data curation process (outer loop); and 2) expert validation and interpretation - where at the end of automated processing pipeline, experts enable re-iteration and learning patterns on verified outputs (inner loop). Depending on the task at hand, feature extraction, feature selection, detection and pixel-level segmentation techniques used for semi-automated annotation can resort to several computer vision, machine learning and a broad scope of intelligent systems approaches. Successful implementations of this approach should aim

for obtaining increasingly accurate fine-grained outputs validated with cohorts of experts, which shall be quantified with relevant evaluation metrics for quantitative benchmarking.

The key benefit of this data curation approach is to yield fine-grained annotated data sources for the development of large-scale databases for wildfire-related problems. By being validated by domain experts, it will also improve the relevance of subsequent machine learning applications developed for end-users (e.g. wildfire management, firefighting and civil protection agencies, and researchers working on wildfire-related topics).

11.4 Conclusion and Future Work

The solutions developed through the proposed approach will be an important stepping stone for data curation and creating large-scale datasets for wildfire detection and monitoring. These datasets will open opportunities for leveraging ML technologies in this context, as well as pave the way for relevant multi-view data association [220] and autonomous robotics tasks in this domain [221]. ML along with emerging technologies such as unmanned aerial vehicles and cube-sat systems are important enabling technologies for near-real-time wildfire intelligence, which will have an essential role in decision support systems with crucial impacts in the safety of at risk populations and environment protection.

Broader Impact

Real-time early fire detection and monitoring systems can prevent the loss of natural ecosystems responsible for climate regulation through carbon sequestration, can help to avoid the occurrence of large burnt areas, and the emission of greenhouse gases. Therefore, the integration, at different timescales, of data-driven intelligent systems in decision support systems for firefighting and civil protection can contribute to mitigate the social, cultural, environmental and economic effects associated with wildfires, contributing to the United Nations Sustainable Development Goals.

12 | Color-based Linguistic Fuzzy Models

Contents

12.1 Introduction	192
12.2 Related work	193
12.3 Dataset	194
12.4 Fire Data Annotation Pipeline	195
12.5 Experiments	201
12.6 Conclusion	203

The material included in this chapter was previously featured in a peer-reviewed conference paper, published in the proceedings of the 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) © 2021 IEEE.

Published in:

P. Messias, M. J. Sousa, A. Moutinho, **Color-based Superpixel Semantic Segmentation for Fire Data Annotation**, in IEEE International Conference on Fuzzy Systems (Institute of Electrical and Electronics Engineers Inc., 2021), vols. 2021-July. doi: 10.1109/FUZZ45933.2021.9494421

12.1 Introduction

The challenges in facing wildfires are increasing the demand for automatic surveillance systems. However, current systems are typically prone to false alarms in real deployments, thus unreliable to response teams because even false positives hinder the confidence in these solutions. Although several research efforts have attempted to leverage the advances in deep learning in classification and segmentation tasks, the performance and evaluation are hampered by the quality of existing image datasets. Limitations include, e.g., the low number of samples and quality of image data, lack of representativity of real-world situations, or being heavily based on video frames, reducing variability. Moreover, datasets are often missing annotations, limiting both the performance and the information extracted for real operations.

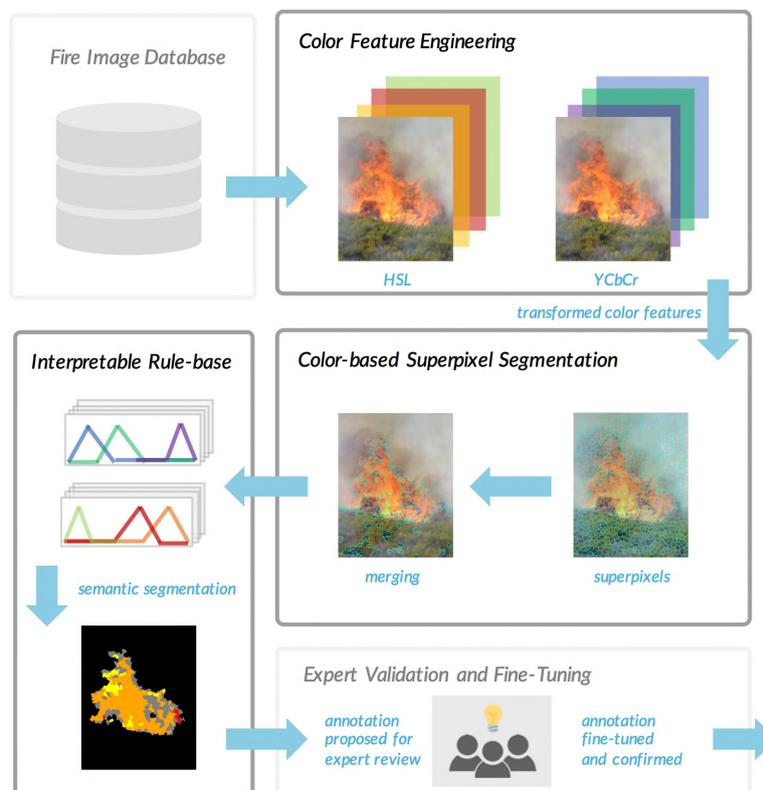


Figure 12.1: Overview of the fire data annotation pipeline: Color-based features across the HSL and YCbCr color spaces are used to model semantic classes.

To address this issue, this paper shifts focus from the usual fire detection architectures, to instead target the development of a pipeline for fire data annotation through semantic segmentation. Towards creating large-scale high-quality annotated datasets, this paper presents a method to generate fine-grained fire segmentations along with fire color labels, which can subsequently be validated by fire domain experts to produce reliable ground truth data. The proposed approach, depicted in Fig. 12.1, leverages the feature rich representations of fire, namely in the HSL and YCbCr color spaces, that allow an insightful interpretation that informs the color-based superpixel segmentation. Subsequently, interpretable rule-based linguistic models are employed for pixel-wise classification, using statistical color attributes to infer which superpixels correspond to fire or not, and generate semantic labels descriptive of the colors represented. Our

approach is evaluated against the Corsican Fire Database, demonstrating excellent results in fine-grained segmentations. The proposed solution can cope with a wide array of real-contexts: e.g., with fire at long distances, with firefighters in the field of operations, and in smoke situations. Where the performance might exhibit some limitations, the interpretability and flexibility of this approach come into play, allowing the expert to intuitively fine-tune the model output to improve the segmentation, and thus guaranteeing the expert confidence in the annotation process. This novel approach extends the state-of-the-art on fine-grained fire semantic segmentation for data annotation, which we believe is an instrumental step towards building large-scale datasets for safety-critical deployments.

12.2 Related work

Image-based fire detection approaches have widely explored color-based techniques [222, 223], and recent works have adopted learning-based methods [224–226]. Herein, we concentrate on image data from the visible spectrum because its widespread availability allows a broader deployment. Contributions in this domain targeting fire event detection generally design techniques for either flame or smoke [227]. Most center on a single application, e.g., ground operations, watchtower systems, or aerial surveillance [228]. That means methods are considerably different, use distinct types of data sources, thus making these hard to compare generalization-wise. For learning-based approaches, the scarce availability of curated datasets currently presents a significant roadblock for scaling up their applications [226, 229]. Aiming for better generalization and broader context applicability, this work pivots the attention to approaches that enable building large-scale datasets, focusing on flame detection.

Data annotation is paramount for improving algorithm performance and embedding relevant information for this application. However, the tendency in fire detection works has favored computational efficiency [230], rather than segmentation granularity [223, 231], hence being blindsided to its key role. In contrast, our objective is to prioritize fine-grained segmentation to streamline the data annotation process. Although state-of-the-art methods can support this goal, the task of defining ground truths for fire data with meaningful information involves a high level of complexity and nuance.

Describing colors of fire in the wild calls for expert knowledge from fire domain specialists because it involves a deep understanding of the natural and artificial fuels burning [232]. The role of this characterization has utmost importance both in wildfire science and wildfire operations, e.g., as it relates to fire intensity and severity. However, hand-crafted pixel-wise annotation is extremely time consuming, therefore semantic segmentation tools are crucial to simplify this procedure and assist experts in developing ground truths.

Model interpretability. Being the experts involved in the validation of algorithm outputs, the fine-tuning capacity and interpretability of the models are essential to leveraging their relevant inputs. For this reason, in this work, we privilege interpretable ruled-based models that are fine-tunable if and whenever needed. In this way, these systems provide higher flexibility in an early stage than black-box models, which could require extensive manual correction cycles until starting generating high-quality outputs. That problem



Figure 12.2: Side-by-side samples of fire images (colored) and respective ground truths (binary) of the reduced Fire Image Datasets used, including firefighting and wildland-urban interface elements, and varying visibility conditions, e.g., day, sunset, night and with smoke.

could discourage expert commitment in the data curation process, which is avoided using our explainable and interpretable approach.

12.3 Dataset

The Corsican Fire Database (CFDB) is a public database of wildfire images with several labels that allows the evaluation and comparison of algorithms related to wildfire detection [233]. As of this writing, despite the authors intention for an evolving dataset, where users can upload new data for categorization, it remains composed of initial data, attesting to the difficulty in developing these repositories. The database comprises 500 images in the visible spectrum, 100 pairs of visible and near-infrared images, and 5 multi-modal sequences with visible and near-infrared pairs. We benchmark our approach using a part of this database.

The Fire Image Dataset. is a reduced dataset was used with 207 images from the 500 visible subset of the CFDB to create a representative dataset with different scenarios (Fig. 12.2).

Samples were excluded due to low resolution, noise or image artifacts, or pixelization that reduced the granularity of fire features. From these 207 images, 50 were used to develop and test the algorithm while the remaining were used only for testing. Images without fire were not added since important real-world contexts, e.g., sunsets, and presence of firefighting means are already portrayed in this dataset.

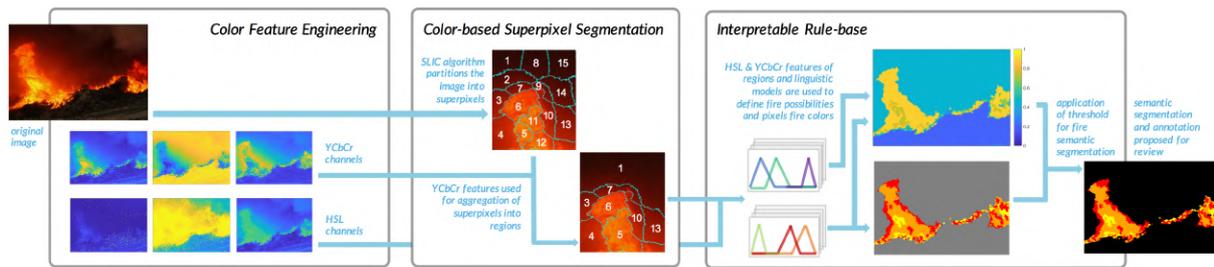


Figure 12.3: **Fire Data Annotation Pipeline.** The proposed architecture is comprised of three core parts: i) Color Feature Engineering (Section 12.4.2), ii) Color-based Superpixel Segmentation (Section 12.4.3), and iii) Interpretable Rule-base (Section 12.4.4). The first layer transforms the image data to the HSL and YCbCr color spaces, better suited for fire instances. The second layer uses a purpose-built segmentation method that generates superpixels and merges them into regions according to statistical color-based features. Then, the interpretable rule-base employs linguistic models to classify the merged regions, yielding the pixel-wise segmentation of fire and the corresponding color labels for each region. Subsequently, the results can be reviewed by experts to validate and fine-tune new ground truths for fire image data.

12.4 Fire Data Annotation Pipeline

Data annotation can be addressed using semantic segmentation methods. In this work, we pose this problem under a two-stage approach: i) *segmentation* and ii) *classification*. The first concerns the segmentation of image data that can be handled with, e.g., partitioning algorithms. The second deals with the classification of segmented regions through assignment to categories according to their attributes. The objectives of the proposed fire data annotation pipeline outlined in Fig. 12.3 are twofold: i) pixel-wise segmentation of fire and ii) description of the fire color category.

12.4.1 Problem Formulation

Consider a sample image, I , encoded in a preset color space domain defined as $\mathcal{D} \in \mathbb{R}^c$, where c represents the number of channels. Let X represent the space of image pixels and x represent a pixel of the image, $x \in X$. To address the first objective of performing a pixel-wise segmentation of fire in an image, let \mathcal{F} represent the set of pixels belonging to the fire class, and \mathcal{N} consist of the pixels that do not depict fire. The *ground truth* defines the expert validated data, where both classes are, in this case binary and mutually exclusive, i.e. $x \in \mathcal{F}$ or $x \in \mathcal{N}$. Regarding the second objective of describing the color of fire, we model four color categories, namely red, orange, yellow and other. Note that, unlike the first case, sets of pixels belonging to a color subset $\{\mathcal{C}_{\text{red}}, \mathcal{C}_{\text{orange}}, \mathcal{C}_{\text{yellow}}, \mathcal{C}_{\text{other}}\}$ are harder to define in a crisp way, thus are modeled with fuzzy sets.

The following sections detail the three layers of the proposed architecture (Section 12.4.2 - 12.4.4), and the implementation details are discussed in Section 12.4.5.

12.4.2 Color Feature Engineering

Seeing Fire Across Color Spaces.

As the objective is to segment the flame based on color, choosing relevant color spaces can be very helpful when defining parameters that correspond to fire colors, thus, leading to better results. This is a particularly important step for images with similar fire colors in non-fire regions and when there is smoke over the flame, decreasing the perception of the fire colors even for human annotation. For these reasons, the color spaces used are the HSL and the YCbCr. The HSL color space is easy to use and works well in scenarios with a high contrast between the flame and the background. The saturation and lightness channels are more intuitive in defining fire colors and separating them from dark smoke (high saturation) and clouds (low lightness). Moreover, the hue channel makes it easier to specify the range of colors for the linguistic terms (e.g., red, orange, yellow). However, this color space is more challenging when it comes to images with smoke in the scene or regions with similar colors to fire colors (Fig. 12.4). In these situations, the YCbCr color space allows for an easier flame segmentation (Fig. 12.5) because of its ability to separate the colors, but it is less interpretable than the HSL making its development more complex. Color-based features derived from these datasets are employed in the two stages of the proposed pipeline, namely in the segmentation and classification parts.

12.4.3 Color-based Superpixel Segmentation

The use of superpixels allows to segment an image by clustering pixels based on their color and proximity. This technique can be very useful since superpixels can carry more information than pixels [234], adhere better to the image boundaries and reduce the complexity of several image processing operations [235], thus decreasing processing times.

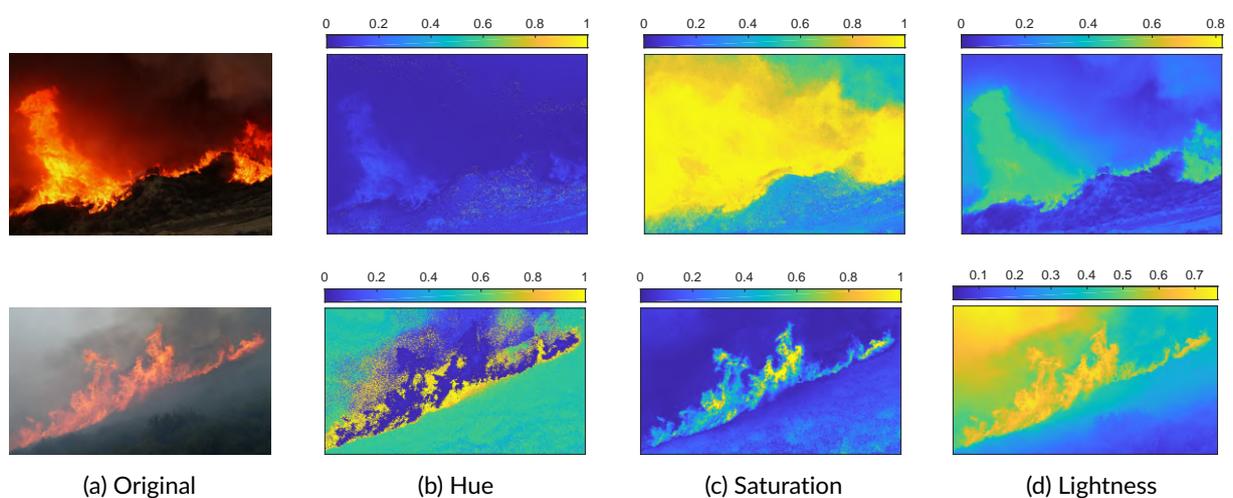


Figure 12.4: Fire features in the HSL color space. Visual display of each HSL color channel for two different images.

Table 12.1: **Interpretable Linguistic Models.** Description of the ruled-based systems designed for fire data annotation, using the HSL and YCbCr color-features. The classification models have three inputs and one common output. Each input/output is defined with different linguistic terms and parameters, forming membership functions. There are triangular (3 parameters) and trapezoidal (4 parameters) membership functions. The HSL model has one additional output, fire color, to generate fire color category.

Model	Input			Output		
	Variable	Linguistic terms	Parameters	Variable	Linguistic terms	Output parameters
HSL	Hue	red1, orange, yellow,	[0, 0, 0.03, 0.055]; [0.04, 0.09, 0.133]; [0.11, 0.16, 0.2];	fire possibility	low, high	[0, 0, 0.3, 0.5]; [0.4, 0.7, 1, 1];
		other, red2	[0.17, 0.25, 0.87, 0.96]; [0.9, 0.97, 1, 1];			
	Saturation	low, high	[0, 0, 0.4, 0.65]; [0.545, 0.75, 1, 1];	fire color	red, orange, yellow, other	[0.5, 1, 1.75]; [1.25, 2, 2.75];
	Lightness	low, medium, high	[0, 0, 0.23, 0.39]; [0.23, 0.427, 0.85, 0.96]; [0.94, 0.965, 1, 1];			[2.25, 3, 3.75]; [3.25, 4, 4.5];
YCbCr	Luminance	low, medium, medium high, high	[0, 0, 0.365, 0.49]; [0.457, 0.5, 0.548, 0.594]; [0.548, 0.63, 0.72, 0.776]; [0.73, 0.8, 1, 1];	fire possibility	low, medium, medium high, high	[0, 0, 0.23, 0.33]; [0.27, 0.35, 0.53, 0.65]; [0.6, 0.65, 0.75, 0.8]; [0.75, 0.83, 1, 1];
	Chrominance Blue	low, medium, high	[0, 0, 0.435, 0.56]; [0.47, 0.527, 0.58, 0.6]; [0.446, 0.69, 1, 1];			
	Chrominance Red	low, medium, high	[0, 0, 0.49, 0.625]; [0.58, 0.647, 0.69, 0.78]; [0.71, 0.826, 1, 1];			

Superpixel Algorithm.

The superpixels are generated using the simple linear iterative clustering (SLIC) algorithm [235] that allows the specification of both the desired *number of superpixels*, N_{sp} , and their *compactness*, C . The number of superpixels drives the granularity of the image partitioning, being higher with the increase in N_{sp} . In turn, the compactness controls the shape of each element, with higher values creating more regularly shaped superpixels (square like), and lower values creating more irregular shapes, which adhere better to intricate image boundaries. Furthermore, the superpixels are defined as vectors in \mathbb{R}^3 in both HSL and YCbCr color spaces, where each entry corresponds to the mean color of each channel ($H_j^{sp}, S_j^{sp}, L_j^{sp}$ or $Y_j^{sp}, Cb_j^{sp}, Cr_j^{sp}$) of the superpixel sp , with $j \in [1, N_{sp}]$.

Merging Superpixels.

We propose a procedure that merges similar superpixels into regions. This method combines neighboring superpixels that register a similar color shade. This is performed using the YCbCr color features as these

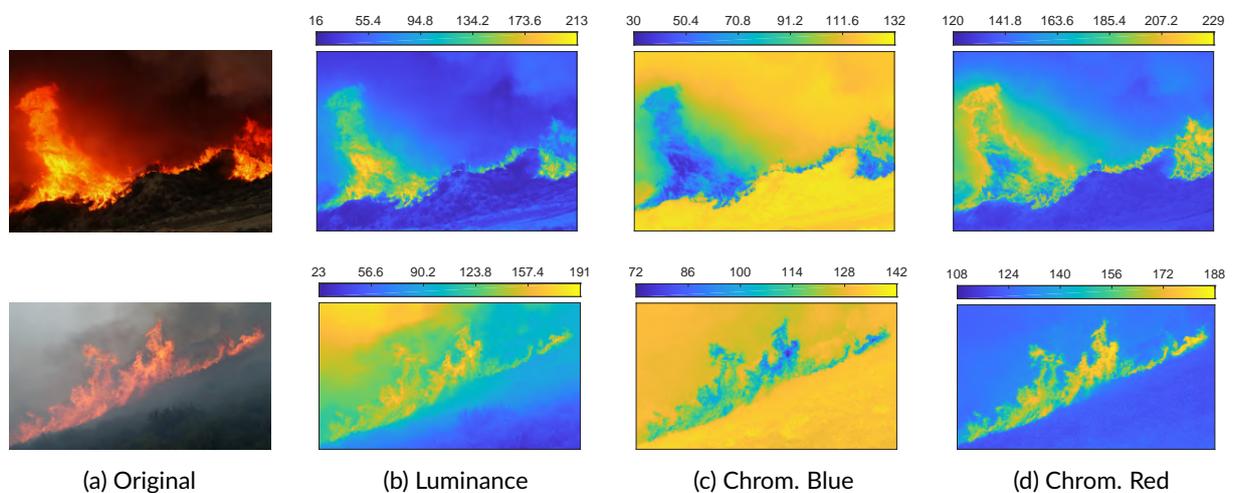


Figure 12.5: **Fire features in the YCbCr color space.** Visual display of each YCbCr color channel for two different images.

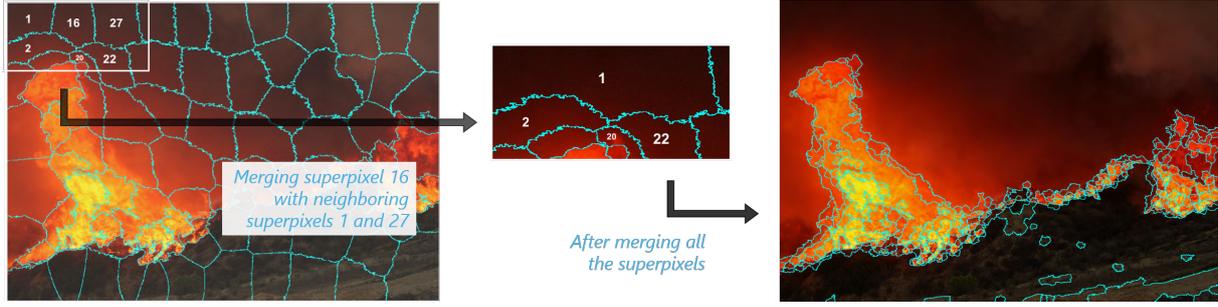


Figure 12.6: Detailed illustration of the superpixel merging procedure. Zoomed area shows how similar superpixels 1, 16 and 27 merge, creating region 1. Merging results in a region-defined image.

allow the separation of fire from other instances like smoke. Adjacent superpixels, j and i , are compared based on their mean color $(Y_{j,i}^{SP}, Cb_{j,i}^{SP}, Cr_{j,i}^{SP})$ and merged if each entry of the pairwise difference is lower or equal to a threshold $\{0.034, 0.1, 0.03\}$, as follows:

$$| Y_j^{SP} - Y_i^{SP} | \leq 0.034 \quad (12.1)$$

$$| Cb_j^{SP} - Cb_i^{SP} | \leq 0.1 \quad (12.2)$$

$$| Cr_j^{SP} - Cr_i^{SP} | \leq 0.03 \quad (12.3)$$

To exemplify this procedure, Fig. 12.6 illustrates the steps of the process and the final result of merging the superpixels. Such operation allows the number of regions to decrease and, consequently, the processing time, and eliminate certain superpixels that could be classified as fire colour.

12.4.4 Interpretable Rule-based Models

Color Perception

Humans describe colors using linguistic terms like red, green or pink, but color perception might differ from person to person [236]. Likewise, image retrieval for search-based analyses is also based on key categorical terms, namely color. Relevant fire characteristics are related to their color so the annotation of these attributes is fundamental towards creating large-scale datasets with relevant information that can be used in a wide range of fire detection scenarios.

Linguistic Models Architecture

In this work, we propose interpretable linguistic models, designed for fire segmentation and classification of the superpixel regions yielded from the previous step. The rule-based architecture is built with Mamdani-type fuzzy inference systems [237], that describe the rules of the knowledge base with linguistic terms. This architecture incorporates uncertainty and is able to bridge the gap between semantic description of colors and its numerical parametrization. The concept of the rules relies on the association of the linguistic terms between both the modeling of the color-based features and the categorization of colors, with the underlying range of values. Our approach leverages two complementary models, developed for

the HSL and YCbCr color spaces and outlined in Table 12.1. Both models are defined with three inputs, corresponding to the mean color of each channel for every region. The two models output a fire possibility per region, that is leveraged to perform the classification of the merged superpixels and achieve the pixel-wise segmentation of fire in the images. In addition, the HSL model is able to describe fire color categories to perform semantic segmentation of the colors in the image. The proposed architecture may integrate both models in the data annotation pipeline (Fig.12.3), combining the HSL and YCbCr using a weighted average or maximum operators, to generate a segmentation of fire. In parallel, the HSL model generates fire color categories that can achieve the final semantic layers that also describe the color of the fire.

The interpretable rule-base is built with simple and intuitive triangular and trapezoidal parametric functions, that model the corresponding linguistic terms for inputs and outputs as follows. The HSL model uses terms describing levels of hue, saturation and lightness to model the fire possibility, i.e. low or high, and the fire color category, which classifies a region to a corresponding color subset $\{C_{red}, C_{orange}, C_{yellow}, C_{other}\}$. The YCbCr model employs terms describing degrees of luminance, chrominance blue and chrominance red. The knowledge-base comprises fourteen rules. The outputs of the rules are combined and transformed to a crisp value representing the fire color possibility. The models generate continuous-valued outputs that are converted to multi-class labels through the application of rounding thresholds. For classification of a region as being fire the decision threshold, δ , is usually applied at 0.5. This value can be fine-tuned in the validation procedure as will be discussed further along.

12.4.5 Implementation details

The proposed algorithm allows the flexible selection of three parameters. For segmentation (Section 12.4.3), the number of superpixels, N_{sp} , and the compactness, C , which are defined within the algorithm. For classification (Section 12.4.4), a threshold is applied to the outputs generated in the semantic classification of fire and the multi-class color categories. These parameters are established in the algorithm but can be easily fine-tuned by experts upon validation.

Parameter selection

Concerning the segmentation part, the selected number of superpixels must ensure that the algorithm can achieve a fine-grained segmentation. The number of superpixels drives the quality of the segmentation in this regard. Since, by nature, flame shapes are very irregular and the image data might contain regions of interest that are captured at long distances, if N_{sp} is too small the image partitioning results in larger superpixels that do not adhere exclusively to the flames. This behavior is illustrated in Fig. 12.7, where in the lower left corner of the samples presented we can distinctly observe that the superpixels can capture the flames but also aggregate other information nearby. This would inherently degrade the quality of the segmentation, but more importantly, it could prevent an accurate semantic segmentation because a misleading mean color value of the superpixel could result in its misclassification. However, selecting higher values of N_{sp} results in a larger number of increasingly small superpixels, as depicted in Fig. 12.7, which

are harder to classify using the mean color statistics as these capture less context information. The value established by default in our algorithm is 1000 as it is considered an adequate trade-off between these factors. Regarding the compactness, since it controls the shape of the superpixels, it is particularly relevant when segmenting irregular shapes like fire. The influence of varying this parameter can be observed notably in Fig. 12.7, by comparing the samples on the upper right corner of left and right side images. The effect of enforcing a higher compactness (depicted on the right side) could result in less fine-grained semantic segmentation for both fire and fire colors. For this reason, the value of C was established as 1, because it is the lowest value possible, making superpixels adhere better to irregular boundaries.

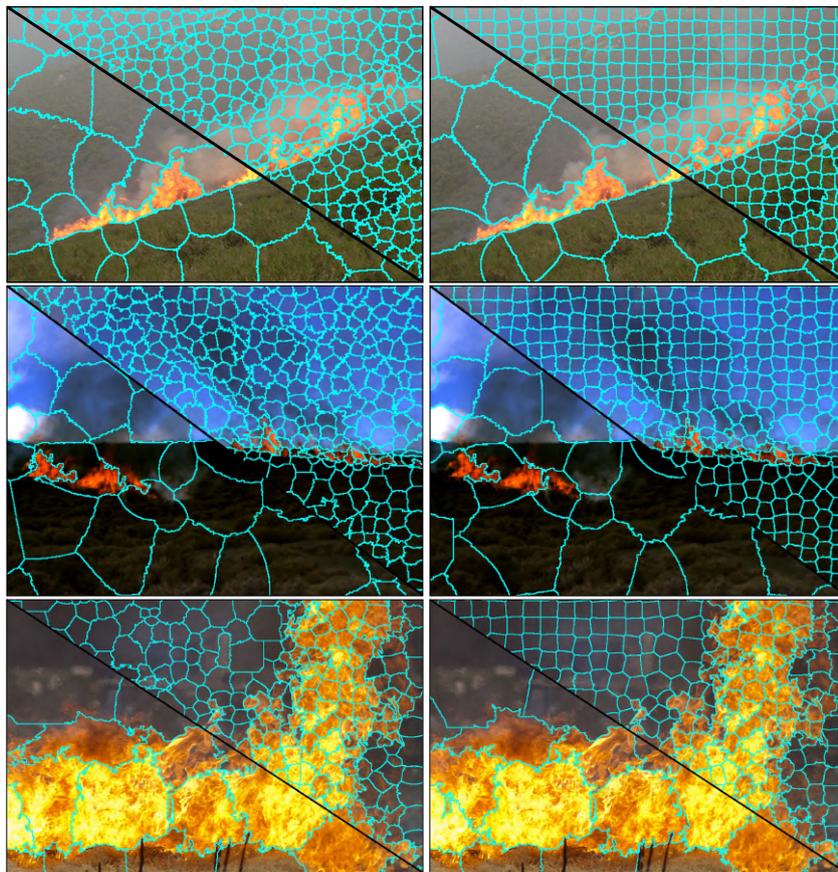


Figure 12.7: Comparison between different N_{sp} and C . The N_{sp} in the lower left corner of each image is 100 and 2000 in the upper right corner. Images on the left have a C equal to 1 and images on the right to 20.

Considering the proposed pipeline integrates two ruled-based models, there are two possible values for the classification of the input in terms of the possibility of corresponding to a color similar to fire colors. The ensemble classifier uses as decision function the weighted average or maximum operators, being possible to define the weights and the strategy to apply. The proposed solution is intended to be fine-tunable, so while a threshold on fire classification is set typically at 0.5, it can be adjusted if required. Our experiments use a $\delta = 0.5$ threshold for the fire segmentation, except when specified otherwise.

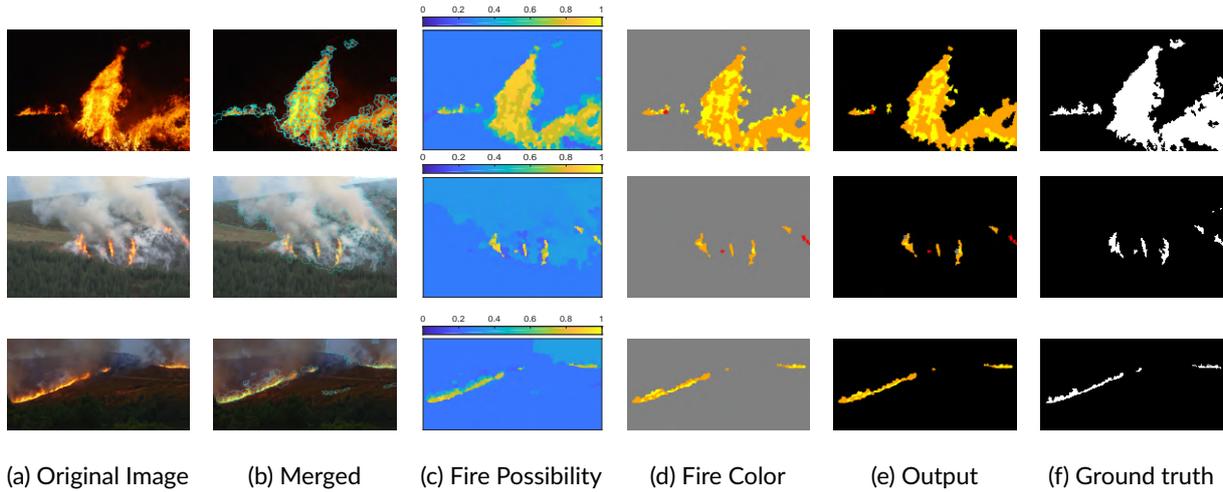


Figure 12.8: Examples for each step in real-world scenarios.

12.5 Experiments

12.5.1 Performance Evaluation

The Fire Image Dataset (Section 12.3) is a subset of the Corsican Fire Database (CFDB), which provides pixel-level segmentations annotated by a single expert as reported in [233]. While the definition of the regions corresponding to fire are complex to define and encompass a great degree of uncertainty, herein we consider this unique source our ground truth.

Baseline

To evaluate the proposed architecture we assess several baseline approaches. The first and simpler ones are based on the HSL model and the YCbCr model separately. We test the integrated approach, with the combination of the two models with a Weighted average ensemble classifier, considering several weight distributions. Subsequently, we also present results for the multi-model approach with an ensemble classification based on the Max Value operator.

Performance Measures

To evaluate the performance of the segmentation algorithms, we will use a set of different metrics that enable the comparison with the ground truth data. For model assessment we will focus on three metrics namely: Accuracy, Intersection over Union (IoU), and Dice coefficient. The metrics consider: true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). The words true and false refer to whether the positives and negatives were correctly classified. For the fire data annotation purpose we are interested in obtaining segmentations with rich information. Considering the outputs will be subsequently validated by experts, the oversegmentation while not ideal is not concerning, since these can be easily corrected through fine-tuning or by specifying which regions of superpixels were misclassified and annotating the true label. In the following, we evaluate cases of interest showcasing the capabilities of the

proposed method, along with the identification of the most challenging scenarios that are likely to require expert annotation.

12.5.2 Results Evaluation

To demonstrate the capabilities of the proposed algorithm, a set of representative examples was selected for analysis. The examples depicted in Fig. 12.8, showcase the performance of the internal steps of the semantic segmentation for three distinct scenarios. The figure represents the pipeline by the first presenting the original of each image, followed by the merged superpixel segmentation. This step already was designed to target the separation from fire from the surrounding context so it already gives a close and granular partitioning of the image, capturing the fire instances even when these are at long distances and covered by smoke. Next, the ruled-based models generate the classification of the superpixels according to fire possibility, with higher values (brighter in the image), representing the superpixels that will be classified with the fire label. The HSL linguistic model also classifies the fire colors in the image, which relates to fire intensity and severity. By comparison with the original image, it can be observed the superpixel approach combined with the fire color classification is able to identify all possible colors in the first two sample images, and correctly attributes the other color label to the remainder of the scene. By intersecting both the fire classification with the color classification, we can automatically generate fire data annotations that closely approximate the expert annotated data, but with richer information concerning the color characteristics.

Regarding limitations of the color-based approach, as expected scene objects resembling fire colors are more likely to be captured by the propose segmentation. However, such cases as depicted in Fig. 12.9, like the sunset and firefighting elements are normally likely to require expert annotation, or need to be complemented by other algorithms for detection of other objects in the scene.

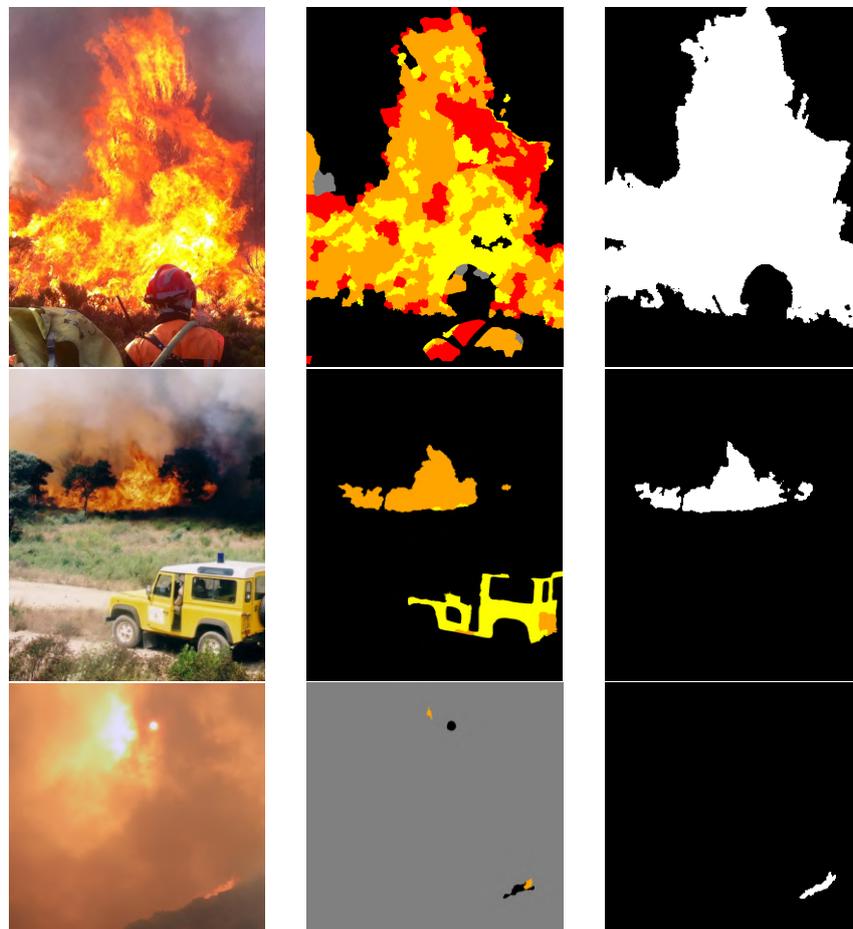
12.5.3 Understanding the Architecture

While most parameters in the architecture are generically preset, the proposed architecture is interpretable and can be fine-tuned according to the more complex scenarios. For instance, for the image represented in Fig. 12.4 obscured by smoke, the fine-tuning of the classification threshold can improve the final semantic segmentation as is illustrated in Fig. 12.10.

The design process described throughout this paper was evaluated with the incremental assembly of its building blocks. Table 12.2 presents the ablation study with the overall results for each model, including variants of the pipeline proposed for semantic segmentation of fire. As represented in bold, the best results were achieved for the multi-model approach using the maximum value operator and by lowering the classification threshold to 0.4.

Table 12.2: Ablation study for model comparison.

Model		δ	Accuracy	IoU	Dice
HSL	HSL model	0.5	93.39	62.49	73.92
YCbCr	YCbCr model	0.5	91.58	56.68	69.82
Weighted	0.4 HSL + 0.6 YCbCr	0.5	93.16	61.52	73.66
	0.3 HSL + 0.7 YCbCr	0.5	93.01	61.00	73.29
	0.2 HSL + 0.8 YCbCr	0.5	92.81	60.13	72.63
Max Value	max(HSL,YCbCr)	0.5	93.47	66.51	77.59
		0.4	94.04	73.53	82.63

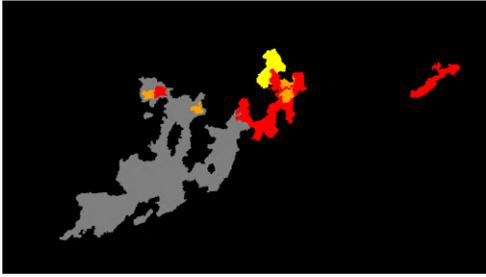


(a) Original (b) Output (c) Ground Truth

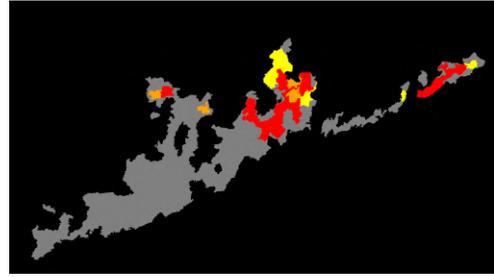
Figure 12.9: Examples of limitations in real-world scenes.

12.6 Conclusion

This paper introduces a novel pipeline for fire data annotation that enables semantic segmentation of fire and the pixel-wise annotation of relevant color characteristics. The architecture proposed leverages a purpose-built algorithm that combines color-based superpixel segmentation with interpretable rule-based



(a) Threshold, $\delta = 0.5$



(b) Threshold, $\delta = 0.4$

Figure 12.10: Fine-tuning of fire classification threshold for improving the semantic segmentation.

models that allow generating pixel-wise semantic labels of fire and of fire colors in the images. We demonstrate that the proposed approach is able to obtain fine-grained semantic segmentations and discuss the limitation in challenging real-world scenarios. Our solution has key advantages due to its fine-tuning ability and interpretable nature, which enables the close involvement of fire domain experts in the validation of new ground truth data for a broad array of fire detection applications. Our approach to fire data annotation aims to streamline this procedure, towards the creation of high-quality large-scale datasets that can allow robust deployments in safety-critical real-world scenarios. Future work will encompass development of automated methods for generating the desired color palettes to ensure a close-set segmentation task, as well as evolving the color modeling approach, namely using fuzzy clustering-based and fuzzy granular modeling methods.

VI | Conclusion

Contents

13 Conclusions and Future Directions	207
---	------------

Summary

The final part of this thesis closes with the conclusions and reflects on future directions stemming from this work and beyond this research. Chapter 13 starts by presenting a summary highlighting the main contributions and key takeaways. Subsequently, future work is discussed with considerations covering future research, development and deployment, as well as the outlook for implementation and the pathway to impact of the proposed contributions.

13 | Conclusions and Future Directions

Contents

13.1 Summary and Conclusions	208
13.2 Future Directions and Outlook	210

This thesis covers the multidisciplinary body of research conducted on the development of aerial networked environmental monitoring systems for wildfire detection and monitoring. The multifaceted nature of this work resulted in substantial contributions along the main research questions explored, which are summarized and highlighted in this final chapter. Herein, we draw the main takeaways from this investigation and discuss broader considerations, such as directions of future work and the general outlook for bringing this type of system to real-world deployments.

13.1 Summary and Conclusions

This thesis presented a comprehensive approach to devise intelligent wildfire detection and monitoring systems based on decentralized networks of aerial robots. Several advances were achieved namely in the scopes of optimization of cooperative robotics systems and intelligent systems for data-driven wildfire detection and monitoring. This was accomplished through research on the four main topics set forth in the introduction:

- T1 Development of multimodal networks of sensors, improving performance and robustness using decentralized approaches;
- T2 Novel feature engineering approaches for wildfire detection;
- T3 Multimodal data-driven intelligent systems;
- T4 Data curation approaches for annotation of fire image data;

Part II - Decentralized Networked Systems focused on approaches to devise environmental monitoring networks that can adapt according to the environment changes, and provide wildfire intelligence in real-time. In this context, a primary concern centered on developing systems that would embody high levels of automation, flexibility and versatility. To that end, the proposed methods tackled the design of aerial networks (Chapter 5) and the coordination of multiple decentralized networks operating over several regions (Chapter 6).

The proposed methods demonstrated novel ways of designing systems for resource-constrained scenarios like wildfire monitoring where the extensive spatial coverage needs to be reconciled with the realistic limitations of available aerial surveillance means. Moreover, as the implementation of these solutions grows over the next decade, it is important that systems are highly flexible, adaptive and versatile so that the required infrastructure investments are the most effective. In that sense, this thesis argues that building for redundancy while not overdimensioning systems is a key consideration in the development of such environmental monitoring networks. Furthermore, with management challenges increasing due to the interaction between several complex systems of systems, the coordination strategies relying on cooperative and collaborative approaches are essential for the establishing of resilient operations.

Part III - Multimodal Robotic Perception delved into the investigation of multimodal sensor solutions for data processing pipelines that have reduced computational burden, and are robust to failure modes that occur in the real-world. This required extensive experiments to assess commercial-off-the-shelf cameras that could provide adequate wildfire detection capabilities in deployment conditions. In this context, this thesis presents an in-depth study of the capabilities of thermal cameras in fire situations focused on its application in autonomous robotics implementation (Chapter 7). Moreover, in the same vein, a novel multimodal dataset is presented resulting from field experiments (Chapter 8).

The analyses of experimental data were essential for extending the knowledge on the inner-workings of thermal cameras, understanding how to implement data processing schema that can provide performance

guarantees in real scenarios, and devise novel data modeling approaches for fire detection applications. In this context, the benefits of leveraging multimodal sensor payloads for UAV-based fire detection and monitoring are underscored throughout this thesis, due to the complementary capabilities of thermal and visible range sensors and their ability to provide different detection modes. These are crucial features for diversifying the situations where these solutions can be effective, e.g., early detection, active fire monitoring and monitoring of hotspots. One key takeaway from this process was the instrumental role color encoding schema have in processing raw thermal data that affect the reliability of all downstream tasks (e.g., georeferencing and fire front mapping). Additionally, harnessing different types of color encoding schema and color palettes can be a valuable asset for robot navigation in hazardous environments like wildfire support operations. In that sense, the MAVFire dataset developed in this thesis aims to contribute to increase the resources available to the robotics community for enabling further work in applications related to wildfires, which are still a prevalent research gap.

Part IV - Intelligent Fire Detection and Monitoring was devoted to developing data-driven intelligent systems based on thermal (Chapter 9) and visible range data (Chapter 10). For this purpose, the main concern was again to develop data processing pipelines that have reduced computational burden, and are robust to failure modes that occur in the real-world. With the goal of improving performance in deployment conditions, the solutions also had to tradeoff between having energy constraints and allowing for fast processing speed so that these could be implemented onboard aerial vehicles.

Concerning thermal imaging data, knowledge of the sensor inner-working was fundamental to select appropriate data collection settings and to devise feature engineering techniques. This allowed the data modeling steps to be more effective and as a result the clustering-based fuzzy models developed based on that data have a low computational burden and high interpretability. A key takeaway in this regard is that the access of thermal image data in their raw format is a requirement for the implementation of the proposed solutions. Therefore, future dataset releases for this application should ensure access to the raw formats. Regarding the development of data processing pipelines for visual-range data, while the approach using transfer learning was generally successful it also allowed for the identification of several limitations stemming from the quality of the datasets openly available for this purpose. To address that gap, the last part of this work also presented methods towards scaling database development efforts.

Part V - Data Curation Approaches investigated methods to scale data annotation by means of automated computer vision and intelligent systems techniques (Chapter 11 and Chapter 12). In this case, the principal aim was to build data curation pipelines that would enable creating large-scale datasets, and include relevant annotations in a streamlined manner. In contrast with the methods proposed in Part IV, in this context where data processing is performed offline, the goals prioritized fine-grained annotation performance over processing speed, while not imposing particular energy constraints on the methods employed.

To address this issue, a key takeaway from this work concerns the relevance of including domain scientists with wildfire expertise in the data annotation processes, via expert-in-the-loop approaches as the ones proposed in this thesis. This is essential in the processes of scoping applications where data-driven

services leveraging intelligent systems will be of most value, as well as in validating data annotations generated by those automated systems, thus producing sound datasets for this field. Another important consideration is the role interpretability and explainability in bridging the understanding between domains and in facilitating the identification of model limitations or flaws early on in the development process. In that sense, approaches that incorporate linguistic elements for description of the models help make the models more accessible for a wide range of end-users.

Overall, the research resulting from this work contributed to advance several dimensions in wildfire detection and monitoring systems. Subsequently, this work can be further developed to be integrated in decision support systems for firefighters and civil protection agencies, which can lead to a better resource allocation. Being wildfire detection and monitoring a very complex application, many of the research avenues remain open, with new questions also stemming from this investigation for future work.

13.2 Future Directions and Outlook

Wildfire intelligence and advanced technologies are pivotal for addressing the ongoing and growing challenges in wildfire management, and the emergence of novel solutions with scalability, flexibility and adaptability is crucial for sustainable implementation. The major gaps in early-warning systems and real-time monitoring suffer greatly from the disproportional imbalance between available resources for preparedness and response and the wide-ranging threat this hazard carries despite its seasonal intermittence.

In this context, autonomous robotics, intelligent systems and large-scale optimization techniques can contribute to address numerous logistical, operational and decision-making bottlenecks that hinder an efficient allocation of available resources and the sustained expansion of wildfire safety infrastructure to protect underserved communities in rural and isolated areas.

With the integration of increasingly sophisticated technologies, such complexity needs to be addressed by evolving paradigms in systems engineering, particularly, in systems of systems engineering. This is required in order to continuously reconcile the plethora of existing legacy infrastructures, meet the impending needs for modernization and upgrades in systems and processes, and the emergence of novel solutions that must offer performance guarantees in safety-critical contexts.

At the heart of such developments are data-driven technologies, resources and infrastructure to support decision-making informed and grounded by scientific evidence. The natural disasters related to large wildfire events in recent years have spurred an increased awareness towards the necessity of building bridges between multidisciplinary fields to address this complex problem, e.g., academia, industry, national agencies, policy-makers, first responders and civil society.

In this context, a decisive catalyzer of such developments has been the significant public funding committed towards scientific Research, Development and Innovation and the investment in upgrading national Defense and Civil infrastructures and operations for wildfire preparedness, response, and civil protection. However, the strategic approach remains primarily human-centric, requiring an extensive mobilization

of resources ad hoc – as a function of forecasted risk – as opposed to human-in-the-loop approaches that leverage data-centric solutions. In that sense, there remains a long path ahead for transitioning from human-informed operations to data-informed operations.

Albeit the significant confluence of efforts in recent years from public, private and governmental sectors to advance solutions leveraging data, there are still difficulties in aggregating relevant databases and provide ample data accessibility. Although, European- and National-level policies for scientific dissemination already impose relevant guidelines for open data sharing practices, even when implemented, these result in self-contained and sparsely hosted data sources, which are hard to relate and integrate with ease.

To overcome these issues, beyond a sustained continuity of ongoing research and technological innovation programs, a joint strategy on data governance for wildfire management is key to achieve a critical mass, involving the vast and diverse breadth of cross-cutting fields and stakeholders engaged in this area. This leap forward is essential for maturing robotics and intelligent systems solutions for this area of application.

13.2.1 From Research to Deployment Roadmap

To reflect on the future developments, it is useful to frame this work in the Technology Readiness Level (TRL) framework [238], depicted in Fig. 13.1 along with a scheme relating those with resources, stakeholders and funding. This thesis has proposed several contributions in early stage research and field testing, comprehending TRL 1-4, with different levels achieved for each of the modules of the system architecture.

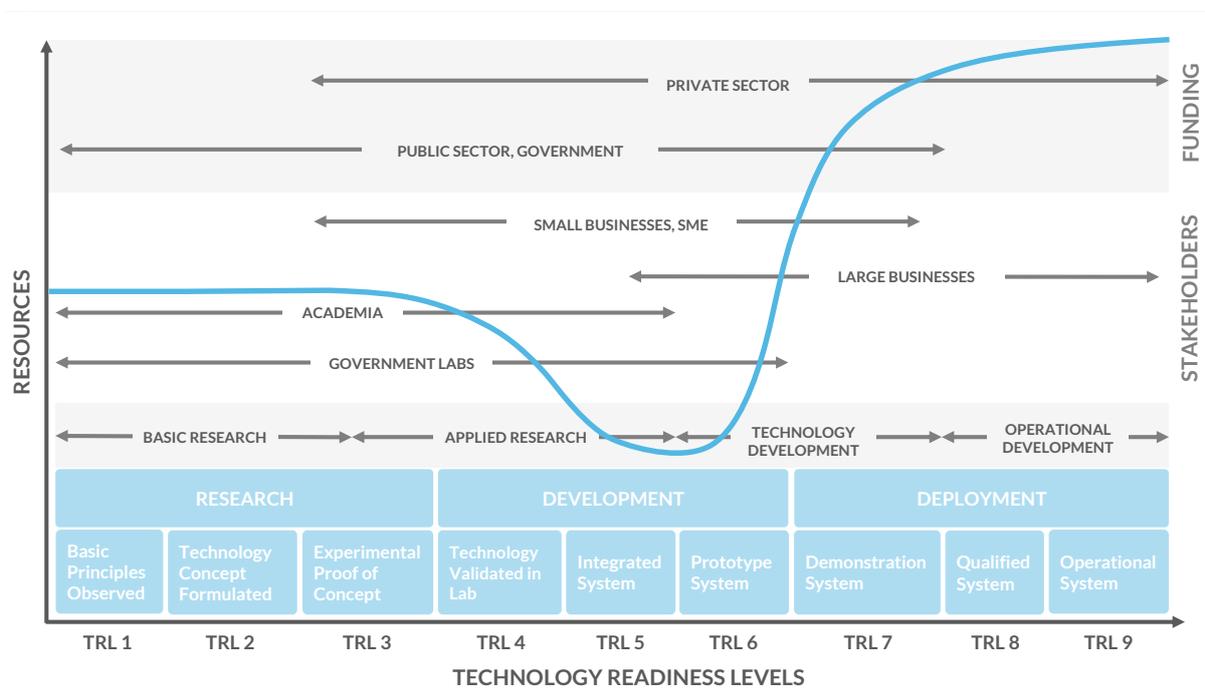


Figure 13.1: Resources availability for innovation at various TRLs, depicting “the valley of death” around TRL 4-6. Adapted from [239].

On the one hand, for some of the proposed methodological and algorithmic advances, the progress centered on developing new concepts and validating them in simulation. On the other hand, research into

intelligent data-driven sensors already achieved promising results with real-world data, either from real fire events, or collected in dedicated laboratory tests and experimental burns in field trials.

With respect to research avenues to be explored in the future, extending the current computational intelligence approaches has extensive opportunities to be leveraged further. Specific examples include the following topics: (i) planning and design of UAV fleet operations; (ii) improving robustness, explainability and interpretability of intelligent sensors; (iii) sensor-driven robot autonomy; (iv) expert-in-the-loop data annotation for creating large-scale datasets for wildfire management tasks.

An important consideration taken into account when choosing evolutionary computing approaches for large-scale optimization problems pertained that as the problem statements converge to approximate more realistic scenarios the complexity of the problems will also increase considerably. In that sense, the methodologies are well-suited such discrete optimization problems. Relevant elements to incorporate for aerial surveillance planning include, but are not limited to, stationary ground-based camera networks, road network and surveillance patrols.

With the aim of using intelligent systems in time- and safety-critical situations, the role of robustness, explainability and interpretability play a more prominent role in ensuring that algorithms, models and systems have appropriate traceability for accountability concerns. In that vein, the extension beyond the core scientific areas addressed in this thesis to adjacent fields like uncertainty quantification and risk assessment, human-computer interaction and decision support systems are relevant additional directions in future research.

For sensor-driven robot autonomy, significant developments can be explored in, e.g., multimodal perception in dynamic environments, cognitive robotics and decision-making and multi-robot risk-aware planning under uncertainty. These areas require increased computational-intensive tasks and communications. Therefore, the dimensions of how to devise algorithms and robotics pipelines that can reconcile fast and high-frequency execution with edge computing under energy constraints are key topics for future work.

In addition to the aforementioned necessary efforts in building database resources and infrastructures that can support data-centric services, scaling the capacity to annotate large amounts of data will then be fundamental. To that end, given that relevant datasets will often need to combine remote sensing and in-situ measurements to develop sound ground truth data, involving domain experts in devising such annotation schema will have a prominent effect for attaining high-quality end results. In that sense, expert-in-the-loop approaches harnessing automatic machine learning and computer vision methods and incorporating active learning should be explored to make these onerous tasks less cumbersome and more worthwhile.

Concerning development and deployment, “*the valley of death*” phase (see Fig. 13.1) represents a significant hurdle for the progress towards implementation. This stage requires increased synergies between stakeholders in the public and private sectors for co-investment and co-creation of technological solutions. At this juncture, aspects that transcend the scientific and technical sphere come into consideration, namely competing societal development priorities, raw materials availability and supply chain bottlenecks, market

demand and convergence to data-centric ecosystems, regulatory and operational aspects concerning the use of drones in civil applications and the use of data and artificial intelligence in public interest domains.

13.2.2 Implementation Considerations: Enablers and Barriers

Several trends over the last decade have contributed to a growing array of enabling technologies to support the implementation of cyber-physical systems that can provide enhanced environmental monitoring. The evolution of devices for remote sensing, the boom in computing, data and artificial intelligence, the breakthroughs in aerospace industry – from commercially-of-the-shelf drone technology to more sustainable satellite launch programs – and the important policy agenda for the green economy.

To large extent, such developments have even led to a fair share of technopositivism, though recent events and the current state of affairs call for a reflection of important enablers and barriers at stake that affect the ongoing development and future conjectures.

Recent years have seen a global pandemic lead to massive lockdowns that slowed down the technological industries, and the aftermath of those events is still very present in the economic environment with looming recession or stagnant growth perspectives worldwide.

The pressures of war in the geopolitical context also added to the economic strains particularly due to the energy and international trade bottlenecks. Furthermore, the use of aerial robotics in warfare has scaled up in recent conflicts, leading to less availability of drone technologies for research and development purposes or for civilian applications. Additionally, the long-standing semiconductor shortage has led to an upturn in the decrease in costs in most electronics-based resources, impacting the cost and actual market availability of all derived devices.

The lack of hardware availability in the last few years has slowed down developments in the robotics community, given that lead times to acquire single-board edge computing devices can take several months and prices for such devices or standard flight controllers have increased twofold or more. This has considerable implications particularly for multi-robot systems development due the multiplication of hardware resources that is necessary for integration stages. The increase in upfront costs leads to greater barriers for organizations with stringent funding constraints like academic laboratories or small businesses, which are predominantly the driving force in the development stage.

At the same time, the sustained commitment from the open-source robotics community, through organizations like Open Robotics and Dronecode Foundation, continue to have a pivotal role in advancing the implementation of robotics technologies. These collaborations towards building shared resources and tools have established cornerstone open standards that are instrumental for interoperability between stakeholders. These standards are endorsed and widely adopted by large semiconductor manufacturers, hardware and software companies, and drone engineering companies.

The principal barrier faced for deployment and implementation is the current regulatory aspect for the operation of unmanned aerial systems. Aerial safety agencies worldwide are evolving regulations to ac-

commodate the safe operation of these systems in the airspaces traditionally used by crewed aircraft, e.g., in the scope of U-space. However, this does not include the sphere of operations in the stratosphere, where traditional aircraft do not fly, but high-altitude pseudo-satellites (HAPS) can have an important role. In this context, the HAPS Alliance established an association of HAPS manufacturers and end-users, including telecommunications, technology, aviation and aerospace enterprises, e.g., AIRBUS, Aerostar and Nokia. The HAPS Alliance advocates for the inclusion of HAPS in the new regulatory frameworks, which is essential to unlock investment from the private sector to propel the development and deployment stages.

Concerning the uptake of drone and HAPS technologies in wildfire detection monitoring tasks, there has been initial efforts over the last few years. Under the umbrella of air force operations for wildfire surveillance, there have been significant efforts in implementing drone-based solutions. However, given the level of maturity of the field some solutions have hit setbacks, and the lack of regulations to allow service providers to market beyond-line-of-sight missions is still a barrier for the growth of these segments. The leap through “*the valley of death*” is the immediate hurdle towards implementation, and collaborations in developing and testing an integrated system and a feasible prototype system are crucial next steps.

13.2.3 Pathway to Impact

Regarding research impacts, the investigation on both i) data-focused methods and ii) deployment-focused methods can enable making significant strides towards reliable deployments of AI approaches to address wildfire detection and monitoring tasks. On the one hand, an important contribution of this work will be leveraging the data collection from field experiments and the proposed data curation methods to provide annotated datasets to be open sourced to the R&D community. On the other hand, our aim to develop methods ready to be efficiently deployed on aerial robots will open a path for approaching the data-driven fire detection methods with application constraints as the driving aspects considered for establishing appropriate model requirements, which has seen limited study in the research community so far.

In terms of future implementation, the rapid detection of an original fire event ignition significantly increases the efficiency of a first intervention to avoid the occurrence of large burnt areas with the dramatic social, cultural, environmental and economic impacts associated. However, with the drawbacks associated with fire suppression practices for the prolonged accumulation of wild fuels, increased wildfire intelligence can have a central role in decision-making. In that sense, data-driven services can better inform emergency response strategies to allow an intelligent use of fire to increase territorial resilience and prevent the occurrence of large fire events.

The proposed system may be included as a module for wildfire detection of decision support systems with better performance when compared with the systems currently used, in terms of early detection and of false alerts. Additionally, considering that spotting is one of the most relevant mechanisms of fire spread in the largest wind-driven wildfire events registered, the automatic detection of spot fires at non-short distances is much more efficient since the focus of attention is typically on the original occurrence. Hence, an early intervention on spot fires may avoid significant impacts.

This thesis aimed to contribute to pressing societal challenges by addressing wildfire response issues related to three objectives of the United Nations Sustainable Development Goals (2030 Agenda) [240]:

- **Goal 11:** Make cities and human settlements inclusive, safe, resilient and sustainable — as lives and goods may be saved by an early intervention because firefighting operations will be more efficient, thus being able of saving people and communities;
- **Goal 13:** Take urgent action to combat climate change and its impacts — as the burned area and the consequent impacts may be reduced;
- **Goal 15:** Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss — since most wildfires have an anthropogenic origin and the burnt area and fire intensity registered currently in some countries like Portugal is not natural.

Bibliography

- [1] A. C. Scott, D M J, W. J. Bond, S. J. Pyne, and M. E. Alexander. *Fire on Earth*. John Wiley & Sons, Nashville, TN, Jan. 2014.
- [2] J. E. Keeley, W. J. Bond, R. A. Bradstock, J. G. Pausas, and P. W. Rundel. *Fire in Mediterranean Ecosystems*. Cambridge University Press, Dec. 2011. doi: 10.1017/cbo9781139033091.
- [3] E. Chuvieco, M. L. Pettinari, N. Koutsias, M. Forkel, S. Hantson, and M. Turco. Human and climate drivers of global biomass burning variability. *Science of The Total Environment*, 779:146361, July 2021. doi: 10.1016/j.scitotenv.2021.146361.
- [4] D. M. J. S. Bowman, C. A. Kolden, J. T. Abatzoglou, F. H. Johnston, G. R. van der Werf, and M. Flannigan. Vegetation fires in the anthropocene. *Nature Reviews Earth & Environment*, 1(10):500–515, Aug. 2020. doi: 10.1038/s43017-020-0085-3.
- [5] S. J. Pyne. *The Pyrocene: How We Created an Age of Fire, and What Happens Next*. University of California Press, 2022. ISBN 978-0520391635.
- [6] CIFFC Training Working Group. Canadian wildland fire management glossary, 2020.
- [7] M. A. Krawchuk and M. A. Moritz. Constraints on global fire activity vary across a resource gradient. *Ecology*, 92(1):121–132, Jan. 2011. doi: 10.1890/09-1843.1.
- [8] V. Fernández-García, E. Marcos, P. Z. Fulé, O. Reyes, V. M. Santana, and L. Calvo. Fire regimes shape diversity and traits of vegetation under different climatic conditions. *Science of The Total Environment*, 716:137137, May 2020. doi: 10.1016/j.scitotenv.2020.137137.
- [9] D. M. J. S. Bowman, J. Balch, P. Artaxo, W. J. Bond, M. A. Cochrane, C. M. D’Antonio, R. DeFries, F. H. Johnston, J. E. Keeley, M. A. Krawchuk, C. A. Kull, M. Mack, M. A. Moritz, S. Pyne, C. I. Roos, A. C. Scott, N. S. Sodhi, and T. W. Swetnam. The human dimension of fire regimes on earth. *Journal of Biogeography*, 38(12):2223–2236, Sept. 2011. doi: 10.1111/j.1365-2699.2011.02595.x.
- [10] A. Ganteaume, R. Barbero, M. Jappiot, and E. Maillé. Understanding future changes to fires in

- southern Europe and their impacts on the wildland-urban interface. *Journal of Safety Science and Resilience*, 2(1):20–29, 2021. ISSN 2666-4496. doi: 10.1016/j.jnlssr.2021.01.001.
- [11] P. Arias, N. Bellouin, E. Coppola, R. Jones, G. Krinner, J. Marotzke, V. Naik, M. Palmer, G.-K. Plattner, J. Rogelj, M. Rojas, J. Sillmann, T. Storelvmo, P. Thorne, B. Trewin, K. Achuta Rao, B. Adhikary, R. Allan, K. Armour, G. Bala, R. Barimalala, S. Berger, J. Canadell, C. Cassou, A. Cherchi, W. Collins, W. Collins, S. Connors, S. Corti, F. Cruz, F. Dentener, C. Dereczynski, A. Di Luca, A. Diongue Niang, F. Doblas-Reyes, A. Dosio, H. Douville, F. Engelbrecht, V. Eyring, E. Fischer, P. Forster, B. Fox-Kemper, J. Fuglestvedt, J. Fyfe, N. Gillett, L. Goldfarb, I. Gorodetskaya, J. Gutierrez, R. Hamdi, E. Hawkins, H. Hewitt, P. Hope, A. Islam, C. Jones, D. Kaufman, R. Kopp, Y. Kosaka, J. Kossin, S. Krakovska, J.-Y. Lee, J. Li, T. Mauritsen, T. Maycock, M. Meinshausen, S.-K. Min, P. Monteiro, T. Ngo-Duc, F. Otto, I. Pinto, A. Pirani, K. Raghavan, R. Ranasinghe, A. Ruane, L. Ruiz, J.-B. Sallée, B. Samset, S. Sathyendranath, S. Seneviratne, A. Sörensson, S. Szopa, I. Takayabu, A.-M. Tréguier, B. van den Hurk, R. Vautard, K. von Schuckmann, S. Zaehle, X. Zhang, and K. Zickfeld. *IPCC Technical Summary*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021. doi: 10.1017/9781009157896.002.
- [12] M. W. Jones, J. T. Abatzoglou, S. Veraverbeke, N. Andela, G. Lasslop, M. Forkel, A. J. P. Smith, C. Burton, R. A. Betts, G. R. van der Werf, S. Sitch, J. G. Canadell, C. Santín, C. Kolden, S. H. Doerr, and C. Le Quéré. Global and regional trends and drivers of fire under climate change. *Reviews of Geophysics*, 60(3):e2020RG000726, 2022. doi: 10.1029/2020RG000726.
- [13] ADAI, Forest Fire Research Centre. Historical data of wildfire events in Portugal from 1943-2022, 2023.
- [14] L. Lourenço. Forest fires in continental Portugal. Result of profound alterations in society and territorial consequences. *Méditerranée. Revue géographique des pays méditerranéens/Journal of Mediterranean Geography*, (130), Sept. 2018. doi: 10.4000/mediterranee.9958.
- [15] S. Gómez-González, F. Ojeda, and P. M. Fernandes. Portugal and chile: Longing for sustainable forestry while rising from the ashes. *Environmental Science & Policy*, 81:104–107, Mar. 2018. doi: 10.1016/j.envsci.2017.11.006.
- [16] M. Bahrepour, N. Meratnia, and P. J. Havinga. Automatic fire detection: A survey from wireless sensor network perspective. *Pervasive System Group, Univeristy of Twente*, 2008.
- [17] R. Allison, J. Johnston, G. Craig, and S. Jennings. Airborne optical and thermal remote sensing for wildfire detection and monitoring. *Sensors*, 16(8):1310, Aug. 2016. doi: 10.3390/s16081310.
- [18] J. M. Johnston, N. Jackson, C. McFayden, L. Ngo Phong, B. Lawrence, D. Davignon, M. J. Wooster, H. van Mierlo, D. K. Thompson, A. S. Cantin, et al. Development of the user requirements for the canadian wildfiresat satellite mission. *Sensors*, 20(18):5081, 2020.
- [19] P. Di Vito, D. Fischer, M. Spada, R. Rinaldo, and L. Duquerooy. HAPs operations and service provision in critical scenarios. In *2018 SpaceOps Conference*, page 2504, 2018.

- [20] A. Alexander, M. Alvidrez, W. Beg, Z. Benezet-Parsons, L. Bouygues, S. Coriell, J. Dempsey, U. Dogruer, B. Freedman, J. Grazer, D. Herbert, B. Ho, R. Lange, M. Lelièvre, D. Nguyen, J. Nutzmann, K. Riordan, K. Roach, J. Rousseau, R. Schlaefli, S. Truslow, , and P. Heninwolf. *Loon Library: Lessons from Building Loon's Stratospheric Communications Service*, 2021.
- [21] R. Haberfellner, O. de Weck, E. Fricke, and S. Vössner. *Systems Engineering*. Springer International Publishing, 2019. doi: 10.1007/978-3-030-13431-0.
- [22] S. J. Kapurch. *NASA Systems Engineering Handbook*. Diane Publishing, 2010.
- [23] SEBoK Editorial Board. The Guide to the Systems Engineering Body of Knowledge (SEBoK), v. 2.7, 2022. URL www.sebokwiki.org. Accessed on 2023-01-06. BKCASE is managed and maintained by the Stevens Institute of Technology Systems Engineering Research Center, the International Council on Systems Engineering, and the Institute of Electrical and Electronics Engineers Systems Council.
- [24] C. B. Nielsen, P. G. Larsen, J. Fitzgerald, J. Woodcock, and J. Peleska. Systems of Systems engineering: Basic Concepts, Model-based Techniques, and Research Directions. *ACM Computing Surveys (CSUR)*, 48(2):1–41, 2015.
- [25] J. Boardman and B. Sauser. System of Systems – the meaning of of. In *2006 IEEE/SMC International Conference on System of Systems Engineering*, pages 6–pp. IEEE, 2006.
- [26] M. W. N. R. Jennings. *Agent Technology: Foundations, Applications, and Markets*. Springer Berlin Heidelberg, 1 edition, 1998. ISBN 978-3-540-63591-8.
- [27] K. Kurose, James & Ross. *Computer Networking: A Top-Down Approach*. Pearson, 8 edition, 2021. ISBN 978-0136681557.
- [28] D. B. Fogel, D. Liu, and J. M. Keller. *Fundamentals of Computational Intelligence*. Wiley, June 2016. ISBN 978-1119214342.
- [29] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, 58: 82–115, 2020. ISSN 1566-2535. doi: 10.1016/j.inffus.2019.12.012.
- [30] A. Otto, N. Agatz, J. Campbell, B. Golden, and E. Pesch. Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: A survey. *Networks*, 72(4):411–458, 2018. doi: 10.1002/net.21818.
- [31] K. Savla, E. Frazzoli, and F. Bullo. On the point-to-point and traveling salesperson problems for Dubins' vehicle. In *Proceedings of the 2005, American Control Conference, 2005.*, pages 786–791. IEEE, 2005.
- [32] P. Váña and J. Faigl. On the Dubins Traveling Salesman Problem with Neighborhoods. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4029–4034. IEEE, 2015.

- [33] F. Bullo, E. Frazzoli, M. Pavone, K. Savla, and S. L. Smith. Dynamic vehicle routing for robotic systems. *Proceedings of the IEEE*, 99(9):1482–1504, 2011. ISSN 0018-9219. doi: 10.1109/JPROC.2011.2158181.
- [34] J. Faigl, P. Váňa, and R. Pěnička. Multi-vehicle close enough orienteering problem with bézier curves for multi-rotor aerial vehicles. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 3039–3044. IEEE, 2019.
- [35] A. Baddeley, E. Rubak, and R. Turner. *Spatial point patterns: methodology and applications with R*. Chapman and Hall/CRC, 2015.
- [36] M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, 27(3):832–837, 1956. doi: 10.1214/aoms/1177728190.
- [37] E. Parzen. On estimation of a probability density function and mode. *Ann. Math. Statist.*, 33(3): 1065–1076, 1962. doi: 10.1214/aoms/1177704472.
- [38] C. M. Bishop. *Pattern Recognition and Machine learning*. Information Science and Statistics. Springer, New York, NY, 2006.
- [39] P. Diggle. A kernel method for smoothing point process data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 34(2):138–147, 1985.
- [40] J. Abonyi and B. Feil. *Cluster analysis for data mining and system identification*. Springer Science & Business Media, 2007.
- [41] S. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982. Originally proposed in an unpublished Bell Laboratories Technical Note (1957).
- [42] J. MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297. Oakland, CA, USA, 1967.
- [43] J. C. Dunn. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3(3):32–57, Jan. 1973. doi: 10.1080/01969727308546046.
- [44] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer Academic Publishers, USA, 1981. ISBN 0306406713.
- [45] J. C. Bezdek, R. Ehrlich, and W. Full. FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, 10(2):191–203, 1984. ISSN 0098-3004. doi: 10.1016/0098-3004(84)90020-7.
- [46] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A Density-Based Algorithm for Discovering Clusters a Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD'96*, page 226–231. AAAI Press, 1996.

- [47] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Trans. Database Syst.*, 42(3), 2017. ISSN 0362-5915. doi: 10.1145/3068335.
- [48] R. J. G. B. Campello, D. Moulavi, and J. Sander. Density-based clustering based on hierarchical density estimates. In J. Pei, V. S. Tseng, L. Cao, H. Motoda, and G. Xu, editors, *Advances in Knowledge Discovery and Data Mining*, pages 160–172, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg. ISBN 978-3-642-37456-2.
- [49] D. Gustafson and W. Kessel. Fuzzy clustering with a fuzzy covariance matrix. In *1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*, pages 761–766, 1978. doi: 10.1109/CDC.1978.268028.
- [50] R. J. Campello, D. Moulavi, A. Zimek, and J. Sander. Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 10(1):1–51, 2015.
- [51] R. C. Prim. Shortest connection networks and some generalizations. *The Bell System Technical Journal*, 36(6):1389–1401, 1957.
- [52] E. Yanmaz, S. Yahyanejad, B. Rinner, H. Hellwagner, and C. Bettstetter. Drone networks: Communications, coordination, and sensing. *Ad Hoc Networks*, 68:1–15, 2018.
- [53] G.-Z. Yang, J. Bellingham, P. E. Dupont, P. Fischer, L. Floridi, R. Full, N. Jacobstein, V. Kumar, M. McNutt, R. Merrifield, et al. The grand challenges of science robotics. *Science Robotics*, 3(14), 2018. doi: 10.1126/scirobotics.aar7650.
- [54] M. V. Chester and B. Allenby. Toward adaptive infrastructure: flexibility and agility in a non-stationarity age. *Sustainable and Resilient Infrastructure*, 4(4):173–191, 2019.
- [55] B. P. Gerkey and M. J. Matarić. A formal analysis and taxonomy of task allocation in multi-robot systems. *The International Journal of Robotics Research*, 23(9):939–954, 2004.
- [56] M. Bernard, K. Kondak, I. Maza, and A. Ollero. Autonomous transportation and deployment with aerial robots for search and rescue missions. *Journal of Field Robotics*, 28(6):914–931, 2011.
- [57] V. Spurný, T. Báča, M. Saska, R. Pěnička, T. Krajník, J. Thomas, D. Thakur, G. Loianno, and V. Kumar. Cooperative autonomous search, grasping, and delivering in a treasure hunt scenario by a team of unmanned aerial vehicles. *Journal of Field Robotics*, 36(1):125–148, 2019.
- [58] I. Palunko, P. Cruz, and R. Fierro. Agile load transportation: Safe and efficient load manipulation with aerial robots. *IEEE Robotics & Automation Magazine*, 19(3):69–79, 2012.
- [59] K. Alexis. Towards a science of resilient robotic autonomy. *arXiv preprint arXiv:2004.02403*, 2020.
- [60] L. Howell. Global risks 2013. World Economic Forum, 2013.

- [61] C. Folke, S. R. Carpenter, B. Walker, M. Scheffer, T. Chapin, and J. Rockström. Resilience thinking: integrating resilience, adaptability and transformability. *Ecology and Society*, 15(4), 2010.
- [62] Z. Yan, N. Jouandeau, and A. A. Cherif. A survey and analysis of multi-robot coordination. *International Journal of Advanced Robotic Systems*, 10(12):399, 2013.
- [63] D. B. Shmoys and É. Tardos. An approximation algorithm for the generalized assignment problem. *Mathematical Programming*, 62(1-3):461–474, 1993.
- [64] H.-L. Choi, L. Brunet, and J. P. How. Consensus-based decentralized auctions for robust task allocation. *IEEE Transactions on Robotics*, 25(4):912–926, 2009.
- [65] P. Segui-Gasco, H.-S. Shin, A. Tsourdos, and V. Segui. Decentralised submodular multi-robot task allocation. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2829–2834. IEEE, 2015.
- [66] S. Martinez, J. Cortes, and F. Bullo. Motion coordination with distributed information. *IEEE Control Systems Magazine*, 27(4):75–88, Aug 2007. ISSN 1066-033X. doi: 10.1109/MCS.2007.384124.
- [67] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo. Swarm robotics: a review from the swarm engineering perspective. *Swarm Intelligence*, 7(1):1–41, 2013.
- [68] M. Saska, V. Vonásek, J. Chudoba, J. Thomas, G. Loianno, and V. Kumar. Swarm distribution and deployment for cooperative surveillance by micro-aerial vehicles. *Journal of Intelligent & Robotic Systems*, 84(1-4):469–492, 2016.
- [69] H. A. Kurdi, E. Aloboud, M. Alalwan, S. Alhassan, E. Alotaibi, G. Bautista, and J. P. How. Autonomous task allocation for multi-uav systems based on the locust elastic behavior. *Applied Soft Computing*, 71:110–126, 2018.
- [70] I. Boussaïd, J. Lepagnot, and P. Siarry. A survey on optimization metaheuristics. *Information Sciences*, 237:82–117, 2013. ISSN 0020-0255. Prediction, Control and Diagnosis using Advanced Neural Computations.
- [71] B. Korte, J. Vygen, B. Korte, and J. Vygen. *Combinatorial optimization*, volume 2. Springer, 2012.
- [72] M. Berhault, H. Huang, P. Keskinocak, S. Koenig, W. Elmaghraby, P. Griffin, and A. Kleywegt. Robot exploration with combinatorial auctions. In *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*, volume 2, pages 1957–1962. IEEE, 2003.
- [73] M. B. Dias, R. Zlot, N. Kalra, and A. Stentz. Market-based multirobot coordination: A survey and analysis. *Proceedings of the IEEE*, 94(7):1257–1270, 2006.
- [74] I. Maza and A. Ollero. Multiple uav cooperative searching operation using polygon area decomposition and efficient coverage algorithms. In *Distributed Autonomous Robotic Systems 6*, pages 221–230. Springer, 2007.

- [75] J. J. Acevedo, B. C. Arrue, J. M. Diaz-Bañez, I. Ventura, I. Maza, and A. Ollero. One-to-one coordination algorithm for decentralized area partition in surveillance missions with a team of aerial robots. *Journal of Intelligent & Robotic Systems*, 74(1-2):269–285, 2014.
- [76] M. J. Sousa, A. Moutinho, and M. Almeida. Decentralized distribution of UAV fleets based on fuzzy clustering for demand-driven aerial services. In *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–8. IEEE, July 2020. doi: 10.1109/fuzz48607.2020.9177642.
- [77] J. Bruno, E. G. Coffman Jr, and R. Sethi. Scheduling independent tasks to reduce mean finishing time. *Communications of the ACM*, 17(7):382–387, 1974.
- [78] M. L. Pinedo. *Scheduling: Theory, Algorithms and Systems*. Springer US, 2012. doi: 10.1007/978-1-4614-2361-4.
- [79] G. A. Korsah, A. Stentz, and M. B. Dias. A comprehensive taxonomy for multi-robot task allocation. *The International Journal of Robotics Research*, 32(12):1495–1512, 2013.
- [80] P. Toth and D. Vigo. *Vehicle Routing: Problems, Methods, and Applications*. SIAM, 2014.
- [81] M. L. Pinedo. *Planning and Scheduling in Manufacturing and Services*. Springer New York, 2009. doi: 10.1007/978-1-4419-0910-7.
- [82] J. M. C. Sousa and U. Kaymak. *Fuzzy decision making in modeling and control, World Scientific Series in Robotics and Intelligent Systems*, volume 27. World Scientific, 2002. ISBN 9810248776.
- [83] M. Dorigo, G. D. Caro, and L. M. Gambardella. Ant algorithms for discrete optimization. *Artificial life*, 5(2):137–172, 1999.
- [84] J. San-Miguel-Ayanz, T. Durrant, R. Boca, G. Libertà, A. Branco, D. de Rigo, D. Ferrari, P. Maianti, T. A. Vivancos, et al. Forest fires in europe, middle east and north africa 2018. *Publications Office of the European Union*, 2019.
- [85] Spreading like wildfire. *Nature Climate Change*, 7(11):755–755, Nov. 2017. doi: 10.1038/nclimate3432.
- [86] J. R. Molina, A. González-Cabán, and F. R. y Silva. Potential effects of climate change on fire behavior, economic susceptibility and suppression costs in mediterranean ecosystems: Córdoba province, spain. *Forests*, 10(8):679, Aug. 2019. doi: 10.3390/f10080679.
- [87] B. R. Christensen. Use of UAV or remotely piloted aircraft and forward-looking infrared in forest, rural and wildland fire management: evaluation using simple economic analysis. *New Zealand Journal of Forestry Science*, 45(1), Sept. 2015. doi: 10.1186/s40490-015-0044-9.
- [88] J. R. M. de Dios, L. Merino, A. Ollero, L. M. Ribeiro, and X. Viegas. Multi-UAV experiments: Application to forest fires. In *Springer Tracts in Advanced Robotics*, pages 207–228. Springer Berlin Heidelberg. doi: 10.1007/978-3-540-73958-6_8.

- [89] L. Merino, F. Caballero, J. R. M. de Dios, I. Maza, and A. Ollero. An unmanned aircraft system for automatic forest fire monitoring and measurement. *Journal of Intelligent & Robotic Systems*, 65(1-4): 533–548, Aug. 2011. doi: 10.1007/s10846-011-9560-x.
- [90] V. G. Ambrosia and E. Hinkley. NASA science serving society: Improving capabilities for fire characterization to effect reduction in disaster losses. In *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2008. doi: 10.1109/igarss.2008.4779800.
- [91] V. G. Ambrosia, S. Wegener, T. Zajkowski, D. V. Sullivan, S. Buechel, F. Enomoto, B. Lobitz, S. Johan, J. Brass, and E. Hinkley. The ikhana unmanned airborne system (UAS) western states fire imaging missions: from concept to reality (2006–2010). *Geocarto International*, 26(2):85–101, Apr. 2011. doi: 10.1080/10106049.2010.539302.
- [92] E. Pastor, C. Barrado, P. Royo, E. Santamaria, J. Lopez, and E. Salami. Architecture for a helicopter-based unmanned aerial systems wildfire surveillance system. *Geocarto International*, 26(2):113–131, Apr. 2011. doi: 10.1080/10106049.2010.531769.
- [93] G.-Z. Yang, J. Bellingham, P. E. Dupont, P. Fischer, L. Floridi, R. Full, N. Jacobstein, V. Kumar, M. McNutt, R. Merrifield, B. J. Nelson, B. Scassellati, M. Taddeo, R. Taylor, M. Veloso, Z. L. Wang, and R. Wood. The grand challenges of Science robotics. *Science Robotics*, 3(14):eaar7650, Jan. 2018. doi: 10.1126/scirobotics.aar7650.
- [94] A. E. Çetin, K. Dimitropoulos, B. Gouverneur, N. Grammalidis, O. Günay, Y. H. Habiboğlu, B. U. Töreyn, and S. Verstockt. Video fire detection – review. *Digital Signal Processing*, 23(6):1827–1843, dec 2013. doi: 10.1016/j.dsp.2013.07.003.
- [95] C. D. Lippitt, D. A. Stow, and L. L. Coulter. *Time-Sensitive Remote Sensing*. Springer, 2015.
- [96] C. Yuan, Y. Zhang, and Z. Liu. A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. *Canadian Journal of Forest Research*, 45(7):783–792, July 2015. doi: 10.1139/cjfr-2014-0347.
- [97] I. Bosch, S. Gomez, R. Molina, and R. Miralles. Object discrimination by infrared image processing. In J. Mira, J. M. Ferrández, J. R. Álvarez, F. de la Paz, and F. J. Toledo, editors, *Bioinspired Applications in Artificial and Natural Computation. IWINAC 2009. Lecture Notes in Computer Science*, pages 30–40, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. ISBN 978-3-642-02267-8. doi: 10.1007/978-3-642-02267-8_4.
- [98] B. U. Töreyn, R. G. Cinbis, Y. Dedeoglu, and A. E. Cetin. Fire detection in infrared video using wavelet analysis. *Optical Engineering*, 46(6):46–46–9, 2007. doi: 10.1117/1.2748752.
- [99] S. Verstockt, A. Vanoosthuysen, S. Van Hoecke, P. Lambert, and R. Van de Walle. Multi-sensor fire detection by fusing visual and non-visual flame features. In A. Elmoataz, O. Lezoray, F. Nouboud, D. Mammass, and J. Meunier, editors, *Image and Signal Processing*, pages 333–341, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. ISBN 978-3-642-13681-8.

- [100] J. M. de Dios, B. Arrue, A. Ollero, L. Merino, and F. Gómez-Rodríguez. Computer vision techniques for forest fire perception. *Image and Vision Computing*, 26(4):550–562, Apr. 2008. doi: 10.1016/j.imavis.2007.07.002.
- [101] I. Bosch, A. Serrano, and L. Vergara. Multisensor network system for wildfire detection using infrared image processing. *The Scientific World Journal*, 2013:1–10, 2013. doi: 10.1155/2013/402196.
- [102] M. M. Valero, O. Rios, E. Pastor, and E. Planas. Automated location of active fire perimeters in aerial infrared imaging using unsupervised edge detectors. *International Journal of Wildland Fire*, 27(4): 241–256, 2018. doi: 10.1071/wf17093.
- [103] P. C. Mengod, J. A. T. Bravo, and L. L. Sardá. The influence of external factors on false alarms in an infrared fire detection system. *International Journal of Wildland Fire*, 24(2):261–266, 2015. doi: 10.1071/wf13200.
- [104] M. J. Sousa, A. Moutinho, and M. Almeida. Classification of potential fire outbreaks: A fuzzy modeling approach based on thermal images. *Expert Systems with Applications*, 129:216–232, Sept. 2019. doi: 10.1016/j.eswa.2019.03.030.
- [105] R. Gade and T. B. Moeslund. Thermal cameras and applications: a survey. *Machine Vision and Applications*, 25(1):245–262, Nov. 2013. doi: 10.1007/s00138-013-0570-5.
- [106] A. Wolf, J. Pezoa, and M. Figueroa. Modeling and compensating temperature-dependent non-uniformity noise in IR microbolometer cameras. *Sensors*, 16(7):1121, July 2016. doi: 10.3390/s16071121.
- [107] A. Rogalski. *Infrared Detectors*. CRC Press, nov 2010. ISBN 9781420076714.
- [108] W. Minkina and S. Dudzik. *Infrared Thermography: Errors and Uncertainties*. John Wiley & Sons, 2009.
- [109] V. Sagan, M. Maimaitijiang, P. Sidike, K. Eblimit, K. Peterson, S. Hartling, F. Esposito, K. Khanal, M. Newcomb, D. Pauli, R. Ward, F. Fritschi, N. Shakoar, and T. Mockler. UAV-based high resolution thermal imaging for vegetation monitoring, and plant phenotyping using ICI 8640 p, FLIR vue pro r 640, and thermoMap cameras. *Remote Sensing*, 11(3):330, Feb. 2019. doi: 10.3390/rs11030330.
- [110] C. Solomon and T. Breckon. *Fundamentals of Digital Image Processing: A practical approach with examples in Matlab*. John Wiley & Sons, 2011.
- [111] V. E. Vickers. Plateau equalization algorithm for real-time display of high-quality infrared imagery. *Optical Engineering*, 35(7):1921, July 1996. doi: 10.1117/1.601006.
- [112] R. S. Allison, J. M. Johnston, G. Craig, and S. Jennings. Airborne optical and thermal remote sensing for wildfire detection and monitoring. *Sensors*, 16(8):1310, 2016.

- [113] C. Yuan, Y. Zhang, and Z. Liu. A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. *Canadian Journal of Forest Research*, 45(7):783–792, 2015.
- [114] M. J. Sousa, A. Moutinho, and M. Almeida. Wildfire detection using transfer learning on augmented datasets. *Expert Systems with Applications*, 142:112975, 2020.
- [115] J. Portmann, S. Lynen, M. Chli, and R. Siegwart. People detection and tracking from aerial thermal views. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1794–1800. IEEE, 2014.
- [116] E. Bondi, R. Jain, P. Aggrawal, S. Anand, R. Hannaford, A. Kapoor, J. Piavis, S. Shah, L. Joppa, B. Dilkina, et al. BIRDSAI: A dataset for detection and tracking in aerial thermal infrared videos. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1747–1756, 2020.
- [117] S. Khanal, J. Fulton, and S. Shearer. An overview of current and potential applications of thermal remote sensing in precision agriculture. *Computers and Electronics in Agriculture*, 139:22–32, 2017.
- [118] M. J. Sousa, A. Moutinho, and M. Almeida. Thermal infrared sensing for near real-time data-driven fire detection and monitoring systems. *Sensors*, 20(23):6803, 2020.
- [119] L. Merino, F. Caballero, J. R. Martínez-de Dios, I. Maza, and A. Ollero. An unmanned aircraft system for automatic forest fire monitoring and measurement. *Journal of Intelligent & Robotic Systems*, 65(1):533–548, 2012.
- [120] L. Merino, F. Caballero, J. R. Martínez-de Dios, J. Ferruz, and A. Ollero. A cooperative perception system for multiple UAVs: Application to automatic detection of forest fires. *Journal of Field Robotics*, 23(3-4):165–184, 2006.
- [121] K. Alexis, G. Nikolakopoulos, A. Tzes, and L. Dritsas. Coordination of helicopter UAVs for aerial forest-fire surveillance. In *Applications of Intelligent Control to Engineering Systems*, pages 169–193. Springer, 2009.
- [122] H. X. Pham, H. M. La, D. Feil-Seifer, and M. Deans. A distributed control framework for a team of unmanned aerial vehicles for dynamic wildfire tracking. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6648–6653. IEEE, 2017.
- [123] R. Bailon-Ruiz and S. Lacroix. Wildfire remote sensing with UAVs: A review from the autonomy point of view. In *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 412–420. IEEE, 2020.
- [124] M. Bloesch, S. Omari, M. Hutter, and R. Siegwart. Robust visual inertial odometry using a direct EKF-based approach. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 298–304. IEEE, 2015.

- [125] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale. Keyframe-based visual-inertial odometry using nonlinear optimization. *The International Journal of Robotics Research*, 34(3):314–334, 2015.
- [126] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós. ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5):1147–1163, 2015. doi: 10.1109/TRO.2015.2463671.
- [127] R. Mur-Artal and J. D. Tardós. ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras. *IEEE Transactions on Robotics*, 33(5):1255–1262, 2017. doi: 10.1109/TRO.2017.2705103.
- [128] L. Chen, L. Sun, T. Yang, L. Fan, K. Huang, and Z. Xuanyuan. RGB-T SLAM: A flexible SLAM framework by combining appearance and thermal information. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5682–5687. IEEE, 2017.
- [129] S. Daftry, M. Das, J. Delaune, C. Sorice, R. Hewitt, S. Reddy, D. Lytle, E. Gu, and L. Matthies. Robust vision-based autonomous navigation, mapping and landing for MAVs at night. In *International Symposium on Experimental Robotics*, pages 232–242. Springer, 2018.
- [130] J. Delaune, R. Hewitt, L. Lytle, C. Sorice, R. Thakker, and L. Matthies. Thermal-inertial odometry for autonomous flight throughout the night. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1122–1128. IEEE, 2019.
- [131] S. Khattak, C. Papachristos, and K. Alexis. Keyframe-based thermal-inertial odometry. *Journal of Field Robotics*, 37(4):552–579, 2020.
- [132] S. Khattak, H. Nguyen, F. Mascarich, T. Dang, and K. Alexis. Complementary multi-modal sensor fusion for resilient robot pose estimation in subterranean environments. In *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 1024–1029. IEEE, 2020.
- [133] C. Papachristos, F. Mascarich, and K. Alexis. Thermal-inertial localization for autonomous navigation of aerial robots through obscurants. In *2018 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 394–399. IEEE, 2018.
- [134] S. S. Shivakumar, N. Rodrigues, A. Zhou, I. D. Miller, V. Kumar, and C. J. Taylor. PST900: RGB-Thermal calibration, dataset and segmentation network. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9441–9447. IEEE, 2020.
- [135] P. Furgale, J. Rehder, and R. Siegwart. Unified temporal and spatial calibration for multi-sensor systems. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1280–1286. IEEE, 2013.
- [136] T. Schneider, M. Li, C. Cadena, J. Nieto, and R. Siegwart. Observability-aware self-calibration of visual and inertial sensors for ego-motion estimation. *IEEE Sensors Journal*, 19(10):3846–3860, 2019.

- [137] M. J. a. Sousa, A. Moutinho, and M. Almeida. Classification of potential fire outbreaks: A fuzzy modeling approach based on thermal images. *Expert Systems with Applications*, 129:216–232, 2019.
- [138] J. P. Dandois and E. C. Ellis. High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. *Remote Sensing of Environment*, 136:259–276, 2013. ISSN 0034-4257. doi: 10.1016/j.rse.2013.04.005.
- [139] L. Boschetti, D. P. Roy, C. O. Justice, and M. L. Humber. Modis&landsat fusion for large area 30m burned area mapping. *Remote Sensing of Environment*, 161:27–42, 2015. ISSN 0034-4257. doi: 10.1016/j.rse.2015.01.022.
- [140] A. Alonso-Benito, P. A. Hernandez-Leal, M. Arbelo, A. Gonzalez-Calvo, J. A. Moreno-Ruiz, and J. R. Garcia-Lazaro. Satellite image based methods for fuels maps updating. In C. M. U. Neale and A. Maltese, editors, *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII*. SPIE, oct 2016. doi: 10.1117/12.2241990.
- [141] C. Yuan, Z. Liu, and Y. Zhang. Aerial images-based forest fire detection for firefighting using optical remote sensing techniques and unmanned aerial vehicles. *Journal of Intelligent & Robotic Systems*, 88(2-4):635–654, jan 2017. doi: 10.1007/s10846-016-0464-7.
- [142] State of California. Thomas Fire Incident Information, 2019. URL <https://fire.ca.gov/incident/?incident=d28bc34e-73a8-454d-9e55-dea7bdd40bee>. Accessed: 2019-07-16.
- [143] National Wildfire Coordinating Group. Mendocino Complex Incident Information, 2019. URL <https://www.fire.ca.gov/incidents/2018/7/27/ranch-fire-mendocino-complex/>. Accessed: 2019-07-16.
- [144] D. X. Viegas, M. F. Almeida, L. M. Ribeiro, J. Raposo, M. T. Viegas, R. Oliveira, D. Alves, C. Pinto, H. Jorge, A. Rodrigues, D. Lucas, S. Lopes, and L. F. Silva. O complexo de incêndios de pedrógão grande e concelhos limítrofes, iniciado a 17 de junho de 2017. Technical report, Forest Fire Research Center, Association for the Development of Industrial Aerodynamics, University of Coimbra, 2017. [The complex of fires of Pedrógão Grande and bordering counties, begun on June 17, 2017 - written in Portuguese].
- [145] D. X. Viegas, M. F. Almeida, L. M. Ribeiro, J. Raposo, M. T. Viegas, R. Oliveira, D. Alves, C. Pinto, A. Rodrigues, C. Ribeiro, S. Lopes, H. Jorge, and C. X. Viegas. Análise dos incêndios florestais ocorridos a 15 de outubro de 2017. Technical report, Forest Fire Research Center, Association for the Development of Industrial Aerodynamics, University of Coimbra, 2019. [Analysis of forest fires occurred on October 15, 2017 - written in Portuguese].
- [146] ICNF. Hazard maps of Forest Fire: Hazard classes, 2017. URL <http://www2.icnf.pt/portal/florestas/dfci/inc/cartografia/cartografia-risco-classes-perigosidade>. Accessed: 2018-03-29.

- [147] ICNF. Forest Fires: burned area maps, 2017. URL <http://www2.icnf.pt/portal/florestas/dfci/inc/mapas>. Accessed: 2018-03-29.
- [148] K. Muhammad, J. Ahmad, and S. W. Baik. Early fire detection using convolutional neural networks during surveillance for effective disaster management. *Neurocomputing*, 288:30–42, may 2018. doi: 10.1016/j.neucom.2017.04.083.
- [149] M. Bilbao, J. D. Ser, C. Perfecto, S. Salcedo-Sanz, and J. Portilla-Figueras. Cost-efficient deployment of multi-hop wireless networks over disaster areas using multi-objective meta-heuristics. *Neurocomputing*, 271:18–27, 2018. ISSN 0925-2312. doi: 10.1016/j.neucom.2016.11.097.
- [150] T.-H. Chen, P.-H. Wu, and Y.-C. Chiou. An early fire-detection method based on image processing. In *Image Processing, 2004. ICIP '04. 2004 International Conference on*, volume 3, pages 1707–1710 Vol. 3, Oct 2004. doi: 10.1109/ICIP.2004.1421401.
- [151] T. Çelik and H. Demirel. Fire detection in video sequences using a generic color model. *Fire Safety Journal*, 44(2):147–158, feb 2009. doi: 10.1016/j.firesaf.2008.05.005.
- [152] W.-B. Horng, J.-W. Peng, and C.-Y. Chen. A new image-based real-time flame detection method using color analysis. In *Proceedings. 2005 IEEE Networking, Sensing and Control, 2005*. IEEE, 2005. doi: 10.1109/icnsc.2005.1461169.
- [153] G. Marbach, M. Loepfe, and T. Brupbacher. An image processing technique for fire detection in video images. *Fire Safety Journal*, 41(4):285–289, jun 2006. doi: 10.1016/j.firesaf.2006.02.001.
- [154] B. U. Töreyn. Fire detection in infrared video using wavelet analysis. *Optical Engineering*, 46(6): 067204, jun 2007. doi: 10.1117/1.2748752.
- [155] S. Ye, Z. Bai, H. Chen, R. Bohush, and S. Ablameyko. An effective algorithm to detect both smoke and flame using color and wavelet analysis. *Pattern Recognition and Image Analysis*, 27(1):131–138, jan 2017. doi: 10.1134/s1054661817010138.
- [156] C. R. Steffens, R. N. Rodrigues, and S. S. da Costa Botelho. Non-stationary VFD evaluation kit: Dataset and metrics to fuel video-based fire detection development. In *Communications in Computer and Information Science*, pages 135–151. Springer International Publishing, 2016. doi: 10.1007/978-3-319-47247-8_9.
- [157] T. Chen, Y. Yin, S. Huang, and Y. Ye. The smoke detection for early fire-alarming system base on video processing. In *2006 International Conference on Intelligent Information Hiding and Multimedia*. IEEE, dec 2006. doi: 10.1109/iih-msp.2006.265033.
- [158] B. C. Ko, K.-H. Cheong, and J.-Y. Nam. Fire detection based on vision sensor and support vector machines. *Fire Safety Journal*, 44(3):322–329, apr 2009. doi: 10.1016/j.firesaf.2008.07.006.
- [159] S. Frizzi, R. Kaabi, M. Bouchouicha, J.-M. Ginoux, E. Moreau, and F. Fnaiech. Convolutional neural network for video fire and smoke detection. In *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, oct 2016. doi: 10.1109/iecon.2016.7793196.

- [160] J. Sharma, O.-C. Granmo, M. Goodwin, and J. T. Fidge. Deep convolutional neural networks for fire detection in images. In *Engineering Applications of Neural Networks*, pages 183–193. Springer International Publishing, 2017. doi: 10.1007/978-3-319-65172-9_16.
- [161] M. V. N. Bedo, W. D. de Oliveira, M. T. Cazzolato, A. F. Costa, G. Blanco, J. F. Rodrigues, A. J. M. Traina, and C. Traina. Fire detection from social media images by means of instance-based learning. In *Enterprise Information Systems*, pages 23–44. Springer International Publishing, 2015. doi: 10.1007/978-3-319-29133-8_2.
- [162] P. Foggia, A. Saggese, and M. Vento. Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(9):1545–1556, Sep. 2015. ISSN 1051-8215. doi: 10.1109/TCSVT.2015.2392531.
- [163] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2818–2826, 2016.
- [164] T. Toulouse, L. Rossi, A. Campana, T. Celik, and M. A. Akhloufi. Computer vision for wildfire research: An evolving image dataset for processing and analysis. *Fire Safety Journal*, 92:188–194, sep 2017. doi: 10.1016/j.firesaf.2017.06.012.
- [165] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
- [166] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. doi: 10.1109/5.726791.
- [167] K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4):193–202, apr 1980. doi: 10.1007/bf00344251.
- [168] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [169] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, apr 2015. doi: 10.1007/s11263-015-0816-y.
- [170] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, jun 2015. doi: 10.1109/cvpr.2015.7298594.
- [171] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.

- [172] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, June 2016. doi: 10.1109/CVPR.2016.90.
- [173] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 248–255. IEEE, 2009.
- [174] A. Krizhevsky. Learning multiple layers of features from tiny images. Technical report, University of Toronto, 2009.
- [175] Y. LeCun. The MNIST database of handwritten digits. <http://yann.lecun.com/exdb/mnist/>, 1998.
- [176] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006. ISBN 0387310738.
- [177] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer New York, 2009. doi: 10.1007/978-0-387-84858-7.
- [178] T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, jun 2006. doi: 10.1016/j.patrec.2005.10.010.
- [179] Bombeiros.pt. Gallery, 2018. URL <https://www.bombeiros.pt/galeria/>. Accessed: 2018-03-29.
- [180] D. Y. T. Chino, L. P. S. Avalhais, J. F. Rodrigues, and A. J. M. Traina. Bowfire: Detection of fire in still images by integrating pixel color and texture analysis. In *2015 28th SIBGRAPI Conference on Graphics, Patterns and Images*, pages 95–102, Aug 2015. doi: 10.1109/SIBGRAPI.2015.19.
- [181] L. Mihalkova and R. Mooney. Transfer learning from minimal target data by mapping across relational domains. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI'09)*, pages 1163–1168, Pasadena, CA, July 2009.
- [182] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3):346–359, jun 2008. doi: 10.1016/j.cviu.2007.09.014.
- [183] J. F. Shroder and D. Paton, editors. *Wildfire Hazards, Risks and Disasters*. Elsevier, 2015. ISBN 978-0-12-410434-1. doi: 10.1016/c2012-0-03331-5.
- [184] M. Kobes, I. Helsloot, B. de Vries, and J. G. Post. Building safety and human behaviour in fire: A literature review. *Fire Safety Journal*, 45(1):1–11, 2010. ISSN 0379-7112. doi: <https://doi.org/10.1016/j.firesaf.2009.08.005>.
- [185] M. Almeida, J. R. Azinheira, J. Barata, K. Bousson, R. Ervilha, M. Martins, A. Moutinho, J. C. Pereira, J. C. Pinto, L. M. Ribeiro, J. Silva, and D. X. Viegas. Analysis of Fire Hazard in Campsite Areas. *Fire Technology*, 53(2):553–575, mar 2017. ISSN 15728099. doi: 10.1007/s10694-016-0591-5.

- [186] M. Planelles, R. Limón, and J. Martín-Arroyo. Authorities optimistic about blaze at Spain's Doñana National Park. *El País*, 2017. URL <https://elpais.com/elpais/2017/06/26/inenglish/1498459870{%5F}951340.html>.
- [187] A. Sandford. Thousands evacuated as new wildfires hit south of France. *Euronews*, 2017. URL <https://www.euronews.com/2017/07/26/thousands-evacuated-as-new-wildfires-hit-south-of-france>.
- [188] United Nations High Commissioner for Refugees. Global Trends: Forced Displacement in 2017, 2017. URL <http://www.refworld.org/docid/5b2d1a867.html>.
- [189] Y. Kazerooni, A. Gyedu, G. Burnham, B. Nwomeh, A. Charles, B. Mishra, S. S. Kuah, A. L. Kushner, and B. T. Stewart. Fires in refugee and displaced persons settlements: The current situation and opportunities to improve fire prevention and control. *Burns*, 42(5):1036–1046, 2016. ISSN 18791409. doi: 10.1016/j.burns.2015.11.008.
- [190] S. Mckenzie. Thousands flee fire at refugee camp in Greece. *CNN*, 2016. URL <https://edition.cnn.com/2016/09/20/europe/fire-refugee-camp-lesbos/index.html>.
- [191] BBC. Syrian refugee camp in Lebanon destroyed by fire. *BBC*, 2017. URL <https://www.bbc.com/news/world-middle-east-40474357>.
- [192] J. Fonollosa, A. Sol³rzano, and S. Marco. Chemical sensor systems and associated algorithms for fire detection: A review. *Sensors*, 18(2), 2018. ISSN 1424-8220. doi: 10.3390/s18020553.
- [193] H. Cruz, M. Eckert, J. Meneses, and J. F. Martínez. Efficient forest fire detection index for application in Unmanned Aerial Systems (UASs). *Sensors (Switzerland)*, 16(6), 2016. ISSN 14248220. doi: 10.3390/s16060893.
- [194] R. Gade and T. B. Moeslund. Thermal cameras and applications: a survey. *Machine Vision and Applications*, 25(1):245–262, Jan 2014. ISSN 1432-1769. doi: 10.1007/s00138-013-0570-5.
- [195] R. Usamentiaga, P. Venegas, J. Guerediaga, L. Vega, J. Molleda, and F. G. Bulnes. Infrared thermography for temperature measurement and non-destructive testing. *Sensors*, 14(7):12305–12348, 2014. ISSN 1424-8220. doi: 10.3390/s140712305.
- [196] R. Vadivambal and D. S. Jayas. Applications of thermal imaging in agriculture and food industry—a review. *Food and Bioprocess Technology*, 4(2):186–199, Feb 2011. ISSN 1935-5149. doi: 10.1007/s11947-010-0333-5.
- [197] M. M. Valero, D. Jimenez, B. Butler, C. Mata, O. Rios, E. Pastor, and E. Planas. On the use of compact thermal cameras for quantitative wildfire monitoring. In *Advances in Forest Fire Research 2018*, pages 1077–1086. Imprensa da Universidade de Coimbra, 2018. ISBN 978-989-26-16-506 (PDF). doi: 10.14195/978-989-26-16-506_119.

- [198] T.-H. Chen, P.-H. Wu, and Y.-C. Chiou. An Early Fire-Detection Method Based on Image Processing Thou-Ho (Chao-Ho) Chen, Ping-Hsueh Wu, and Yung-Chuen Chiou. In *Image Processing, 2004. ICIP'04. 2004 International Conference on*, volume 3, pages 1707–1710. IEEE, 2004. ISBN 0780385543.
- [199] T. Celik, H. Demirel, H. Ozkaramanli, and M. Uyguroglu. Fire Detection in Video Sequences using Statistical Color Model. *Fire Safety Journal*, 44(2):213–216, 2006.
- [200] W.-B. Horng, J.-W. Peng, and C.-Y. Chen. A new image-based real-time flame detection method using color analysis. In *Networking, Sensing and Control, 2005. Proceedings. 2005 IEEE*, pages 100–105. IEEE, 2005.
- [201] C. Yuan, Z. Liu, and Y. Zhang. Aerial Images-Based Forest Fire Detection for Firefighting Using Optical Remote Sensing Techniques and Unmanned Aerial Vehicles. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 88(2-4):635–654, 2017. ISSN 15730409. doi: 10.1007/s10846-016-0464-7.
- [202] G. Marbach, M. Loepfe, and T. Brupbacher. An image processing technique for fire detection in video images. *Fire Safety Journal*, 41(4):285–289, 2006. ISSN 03797112. doi: 10.1016/j.firesaf.2006.02.001.
- [203] J. Sharma, O. C. Granmo, M. Goodwin, and J. T. Fidje. Deep convolutional neural networks for fire detection in images. In *Communications in Computer and Information Science*, volume 744, pages 183–193. Springer, 2017. ISBN 9783319651712. doi: 10.1007/978-3-319-65172-9_16.
- [204] T. X. Tung and J. M. Kim. An effective four-stage smoke-detection algorithm using video images for early fire-alarm systems. *Fire Safety Journal*, 46(5):276–282, 2011. ISSN 03797112. doi: 10.1016/j.firesaf.2011.03.003.
- [205] A. Khatami, S. Mirghasemi, A. Khosravi, C. P. Lim, and S. Nahavandi. A new PSO-based approach to fire flame detection using K-Medoids clustering. *Expert Systems With Applications*, 68:69–80, 2016. ISSN 0957-4174. doi: 10.1016/j.eswa.2016.09.021.
- [206] B. C. Ko, K. H. Cheong, and J. Y. Nam. Fire detection based on vision sensor and support vector machines. *Fire Safety Journal*, 44(3):322–329, 2009. ISSN 03797112. doi: 10.1016/j.firesaf.2008.07.006.
- [207] I. Bosch, S. Gomez, R. Molina, and R. Miralles. Object discrimination by infrared image processing. In J. Mira, J. M. Ferrández, J. R. Álvarez, F. de la Paz, and F. J. Toledo, editors, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 5602 LNCS, pages 30–40, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. ISBN 3642022669. doi: 10.1007/978-3-642-02267-8_4.
- [208] I. Guyon and A. Elisseeff. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research (JMLR)*, 3(3):1157–1182, 2003. ISSN 00032670. doi: 10.1016/j.aca.2011.07.027.

- [209] I. Guyon and A. Elisseeff. *An Introduction to Feature Extraction*, pages 1–25. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-35487-1. doi: 10.1007/978-3-540-35488-8_1.
- [210] R. Babuška. *Fuzzy Modeling for Control*, volume 12. Kluwer Academic Publishers, 1998. ISBN 0792381548. doi: 10.1007/978-94-011-4868-9.
- [211] T. Takagi and M. Sugeno. Fuzzy Identification of Systems and Its Applications to Modeling and Control. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-15(1):116–132, jan 1985. ISSN 21682909. doi: 10.1109/TSMC.1985.6313399.
- [212] A. E. Bradley. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159, 1997. ISSN 0031-3203. doi: 10.1016/S0031-3203(96)00142-2.
- [213] Y. Zhao, G. Tang, and M. Xu. Hierarchical detection of wildfire flame video from pixel level to semantic level. *Expert Systems with Applications*, 42(8):4097–4104, 2015. ISSN 09574174. doi: 10.1016/j.eswa.2015.01.018.
- [214] S. G. Kong, D. Jin, S. Li, and H. Kim. Fast fire flame detection in surveillance video using logistic regression and temporal smoothing. *Fire Safety Journal*, 79:37–43, 2016. ISSN 03797112. doi: 10.1016/j.firesaf.2015.11.015.
- [215] D. Paton, P. Buergelt, S. McCaffrey, and F. Tedim, editors. *Wildfire Hazards, Risks, and Disasters*. Elsevier, United Kingdom, oct 2014. ISBN 9780124104341.
- [216] D. E. Calkin, J. D. Cohen, M. A. Finney, and M. P. Thompson. How risk management can prevent future wildfire disasters in the wildland-urban interface. *Proceedings of the National Academy of Sciences*, 111(2):746–751, 2014.
- [217] F. Tedim, V. Leone, M. Amraoui, C. Bouillon, M. R. Coughlan, G. M. Delogu, P. M. Fernandes, C. Ferreira, S. McCaffrey, T. K. McGee, et al. Defining extreme wildfire events: difficulties, challenges, and impacts. *Fire*, 1(1):9, 2018.
- [218] P. Jain, S. C. Coogan, S. G. Subramanian, M. Crowley, S. W. Taylor, and M. D. Flannigan. A review of machine learning applications in wildfire science and management. *Environmental Reviews*, 2020. doi: 10.1139/er-2020-0019.
- [219] A. Gaur, A. Singh, A. Kumar, A. Kumar, and K. Kapoor. Video flame and smoke based fire detection algorithms: A literature review. *Fire Technology*, 2020.
- [220] D. Pachauri, R. Kondor, and V. Singh. Solving the multi-way matching problem by permutation synchronization. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 1860–1868. Curran Associates, Inc., 2013.

- [221] J. R. Martínez-de Dios, L. Merino, A. Ollero, L. M. Ribeiro, and X. Viegas. *Multi-UAV Experiments: Application to Forest Fires*, pages 207–228. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 978-3-540-73958-6. doi: 10.1007/978-3-540-73958-6_8.
- [222] X. Han, J. Jin, and M. W. et al. Video fire detection based on gaussian mixture model and multi-color features. *Signal, Image and Video Processing*, 11:1419â–1425, 2017. doi: 10.1007/s11760-017-1102-y.
- [223] T. Toulouse, L. Rossi, M. Akhloufi, T. Celik, and X. Maldague. Benchmarking of wildland fire colour segmentation algorithms. *IET Image Processing*, 9(12):1064–1072, 2015. doi: 10.1049/iet-ipr.2014.0935.
- [224] F. Bu and M. S. Gharajeh. Intelligent and vision-based fire detection systems: A survey. *Image and Vision Computing*, 91:103803, 2019. ISSN 0262-8856. doi: <https://doi.org/10.1016/j.imavis.2019.08.007>.
- [225] P. Li and W. Zhao. Image fire detection algorithms based on convolutional neural networks. *Case Studies in Thermal Engineering*, 19:100625, 2020. ISSN 2214-157X. doi: <https://doi.org/10.1016/j.csite.2020.100625>.
- [226] M. J. Sousa, A. Moutinho, and M. Almeida. Wildfire detection using transfer learning on augmented datasets. *Expert Systems with Applications*, 142:112975, 2020. doi: 10.1016/j.eswa.2019.112975.
- [227] Z. Zhou, Y. Shi, Z. Gao, and S. Li. Wildfire smoke detection based on local extremal region segmentation and surveillance. *Fire Safety Journal*, 85:50–58, 2016. ISSN 0379-7112. doi: <https://doi.org/10.1016/j.firesaf.2016.08.004>.
- [228] C. Kyrkou and T. Theodorides. Deep-learning-based aerial image classification for emergency response applications using unmanned aerial vehicles. In *CVPR Workshops*, pages 517–525, 2019.
- [229] Q. xing Zhang, G. hua Lin, Y. ming Zhang, G. Xu, and J. jun Wang. Wildland forest fire smoke detection based on faster r-cnn using synthetic smoke images. *Procedia Engineering*, 211:441–446, 2018. ISSN 1877-7058. doi: <https://doi.org/10.1016/j.proeng.2017.12.034>. 2017 8th International Conference on Fire Science and Fire Protection Engineering (ICFSFPE 2017).
- [230] K. Muhammad, S. Khan, M. Elhoseny, S. Hassan Ahmed, and S. Wook Baik. Efficient fire detection for uncertain surveillance environment. *IEEE Transactions on Industrial Informatics*, 15(5):3113–3122, 2019. doi: 10.1109/TII.2019.2897594.
- [231] H. Harkat, J. Nascimento, and A. Bernardino. Fire segmentation using a DeepLabv3+ architecture. In *Image and Signal Processing for Remote Sensing XXVI*, volume 11533, pages 134 – 145. International Society for Optics and Photonics, SPIE, 2020. doi: 10.1117/12.2573902.
- [232] J. Dehaan, P. Kirk, and D. Icove. *Kirk's fire investigation*. Pearson-Prentice Hall, 7th edition edition, 05 2011. ISBN ISBN-10: 0135082633 ISBN-13: 978-0135082638.

- [233] T. Toulouse, L. Rossi, A. Campana, T. Celik, and M. A. Akhloufi. Computer vision for wildfire research: An evolving image dataset for processing and analysis. *Fire Safety Journal*, 92:188–194, 2017.
- [234] P. Neubert and P. Protzel. Superpixel benchmark and comparison. In *Proceedings Forum Bildverarbeitung*, pages 1–12, 2012.
- [235] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. ŠŤEsstrunk. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34:2274–2282, 2012. doi: 10.1109/TPAMI.2012.120.
- [236] J. Chamorro-Martínez and J. M. Keller. Granular modelling of fuzzy color categories. *IEEE Transactions on Fuzzy Systems*, pages 1–1, 2019. doi: 10.1109/tfuzz.2019.2923966.
- [237] E. H. Mamdani. Application of fuzzy logic to approximate reasoning using linguistic synthesis. *IEEE Transactions on Computers*, (12):1182–1191, 1977. doi: 10.1109/TC.1977.1674779.
- [238] J. C. Mankins et al. Technology readiness levels. *White Paper, April*, 6(1995):1995, 1995.
- [239] J. Hensen, R. Loonen, M. Archontiki, and M. Kanellis. Using building simulation for moving innovations across the “valley of death”. *REHVA Journal*, 52:58–62, 05 2015.
- [240] United Nations. Transforming our world: the 2030 agenda for sustainable development, 2015. URL <https://wedocs.unep.org/20.500.11822/9814>.