Pre-print of article that will appear in the 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE).

Available online on IEEE *Xplore* since 5 August 2021.
URL: doi.org/10.1109/FUZZ45933.2021.9494421

Please cite this paper as:

P. Messias, M. J. Sousa, and A. Moutinho. (2021). “Color-based Superpixel Semantic Segmentation for Fire Data Annotation” in *2021 IEEE International Conference on Fuzzy Systems (FUZZ–IEEE)*, Luxembourg, Luxembourg, 2021, pp.1–7. IEEE. doi: 10.1109/FUZZ45933.2021.9494421.

bibtex:
@inproceedings{messias2021color,
title={Color-based Superpixel Semantic Segmentation for Fire Data Annotation},
author={Messias, Pedro and Sousa, Maria Jo{\~a}o and Moutinho, Alexandra},
booktitle={2021 {IEEE} International Conference on Fuzzy Systems ({FUZZ-IEEE})},
pages={1−−7},
publisher={{IEEE}},
year={2021},
doi={10.1109/FUZZ45933.2021.9494421}
}

Color-based Superpixel Semantic Segmentation for Fire Data Annotation

Pedro Messias
Instituto Superior Técnico,
Universidade de Lisboa,
Lisbon, Portugal
pedro.messias@tecnico.ulisboa.pt

Maria João Sousa
IDMEC, Instituto Superior Técnico,
Universidade de Lisboa,
Lisbon, Portugal
maria.joao.sousa@tecnico.ulisboa.pt

Alexandra Moutinho
IDMEC, Instituto Superior Técnico,
Universidade de Lisboa,
Lisbon, Portugal
alexandra.moutinho@tecnico.ulisboa.pt

***Abstract*—Image-based fire detection is a safety-critical task, which requires high-quality datasets to ensure performance guarantees in real scenarios. Automatic fire detection systems are in ever-increasing demand, but the limited number and size of open datasets, and lack of annotations, hinder model development. Solving this issue requires that experts dedicate a significant time to classify and segment fire events in image datasets. Towards building large-scale curated datasets, this paper presents 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 performance of the proposed algorithm is evaluated for relevant scenarios using a publicly available dataset, namely through the assessment of the segmentation quality and the labeling of fire color categories. 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.**

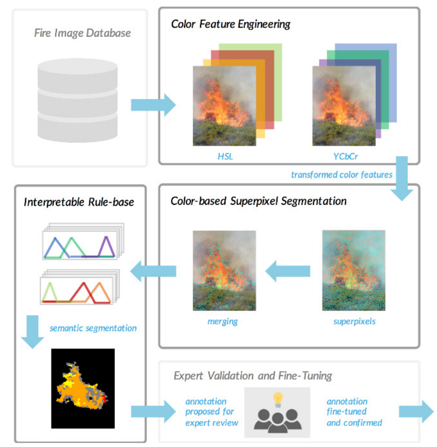


Figure 1: Overview of the fire data annotation pipeline: Color-based features across the HSL and YCbCr color spaces are used to model semantic classes.

I. 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.

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. 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.

This work was supported by FCT – Fundação para a Ciência e a Tecnologia, I.P., through IDMEC, under project Eye in the Sky, PCIF/SSI/0103/2018, and through IDMEC, under LAETA, project UIDB/50022/2020. M. J. Sousa acknowledges the support from FCT, through the Ph.D. Scholarship SFRH/BD/145559/2019, co-funded by the European Social Fund (ESF).

978-1-6654-4407-1/21/$31.00 © 2021 IEEE

II. Related work

Image-based fire detection approaches have widely explored color-based techniques [1], [2], and recent works have adopted learning-based methods [3]–[5]. 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 [6]. Most center on a single application, e.g., ground operations, watchtower systems, or aerial surveillance [7]. 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 [5], [8]. 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 [9], rather than segmentation granularity [2], [10], 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 [11]. 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 could discourage expert commitment in the data curation process, which is avoided using our explainable and interpretable approach.

III. 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 [12]. 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.

Figure 2: Side-by-side samples of fire images (colored) and respective ground truths (binary) of the reduced Fire Image Dateset used, including firefighting and wildland-urban interface elements, and varying visibility conditions, e.g., day, sunset, night and with smoke.

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. 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.

IV. 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. 3 are twofold: *i)* pixel-wise segmentation of fire and *ii)* description of the fire color category.

A. 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

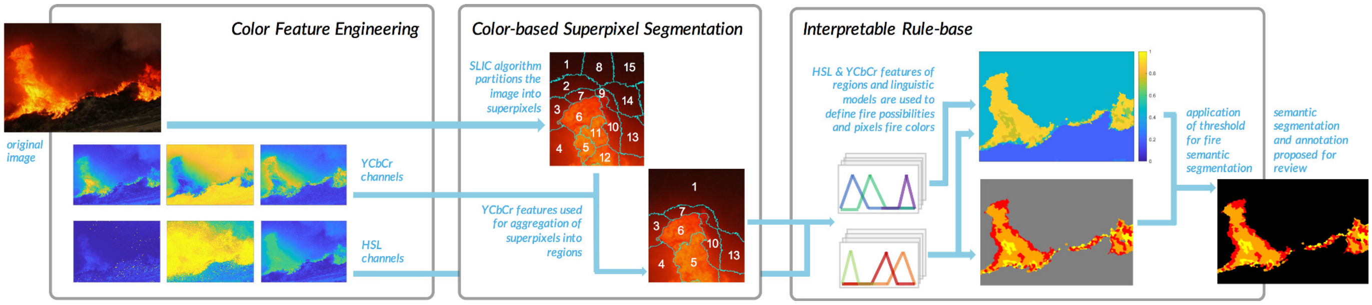


Figure 3: **Fire Data Annotation Pipeline.** The proposed architecture is comprised of three core parts: *i)* Color Feature Engineering (Section IV-B), *ii)* Color-based Superpixel Segmentation (Section IV-C), and *iii)* Interpretable Rule-base (Section IV-D). The first layer transforms the image data 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.

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 IV-B - IV-D), and the implementation details are discussed in Section IV-E.

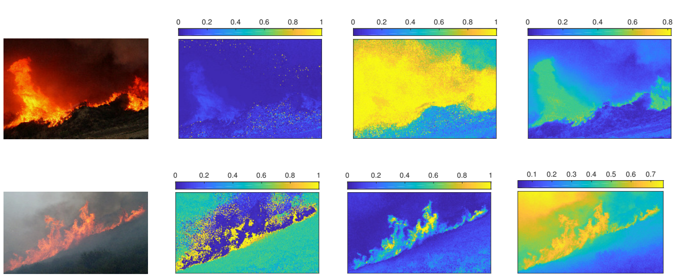
B. Color Feature Engineering

1) 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. 4). In these situations, the YCbCr color space allows for an easier flame segmentation (Fig. 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.

C. Color-based Superpixel Segmentation

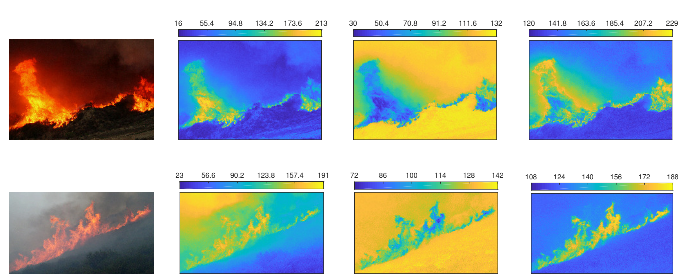
The use of superpixels allows to segment an image by clustering pixels based on their color and proximity. This technique can be very useful since superpixels can carry more information than pixels [13], adhere better to the image boundaries and reduce the complexity of several image processing operations [14], thus decreasing processing times.

1) Superpixel Algorithm.: The superpixels are generated using the simple linear iterative clustering (SLIC) algorithm [14] 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



(a) Original (b) Hue (c) Saturation (d) Lightness

Figure 4: **Fire Features in HSL.** Visual display of each HSL color channel for two different images.



(a) Original (b) Luminance (c) Chrom. Blue (d) Chrom. Red

Figure 5: **Fire Features in YCbCr.** Visual display of each YCbCr color channel for two different images.

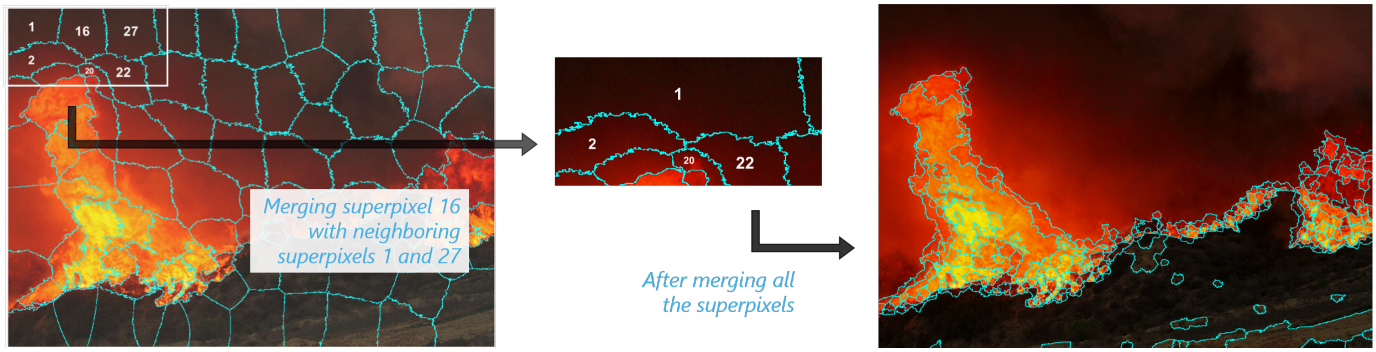


Figure 6: Zoomed area shows how similar superpixels 1, 16 and 27 merge, creating region 1. Merging results in a region-defined image.

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 $j \in [1, N_{\mathrm{sp}}]$.

2) 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 allow the separation of fire from other instances like smoke. Adjacent superpixels, j and i, are compared based on their mean color ($Y_{j,i}^{\mathrm{sp}}$, $Cb_{j,i}^{\mathrm{sp}}$, $Cr_{j,i}^{\mathrm{sp}}$) and merged if each entry of the pairwise difference is lower or equal to a threshold $\{0.034, 0.1, 0.03\}$, as follows:

$$| Y_j^{\mathrm{sp}} - Y_i^{\mathrm{sp}} | \leq 0.034 \tag{1}$$

$$| Cb_j^{\mathrm{sp}} - Cb_i^{\mathrm{sp}} | \leq 0.1 \tag{2}$$

$$| Cr_j^{\mathrm{sp}} - Cr_i^{\mathrm{sp}} | \leq 0.03 \tag{3}$$

To exemplify this procedure, Fig. 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.

D. Interpretable Rule-based Models

1) Color Perception: Humans describe colors using linguistic terms like red, green or pink, but color perception might differ from person to person [15]. 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.

2) 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 build with Mamdani-type fuzzy inference systems [16], 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 I. 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.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 $\{\mathcal{C}_{\mathrm{red}}, \mathcal{C}_{\mathrm{orange}}, \mathcal{C}_{\mathrm{yellow}}, \mathcal{C}_{\mathrm{other}}\}$. 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.

E. Implementation details

The proposed algorithm allows the flexible selection of three parameters. For segmentation (Section IV-C), the number of superpixels, N_{sp}, and the compactness, C, which are defined within the algorithm. For classification (Section IV-D), a threshold is applied to the outputs generated in the semantic

Table I: **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.

	Input			Output		
Model	Variable	Linguistic terms	Parameters	Variable	Linguistic terms	Output parameters
HSL	Hue	red1, orange, yellow, other, red2	$[0, 0, 0.03, 0.055]$; $[0.04, 0.09, 0.133]$; $[0.11, 0.16, 0.2]$; $[0.17, 0.25, 0.87, 0.96]$; $[0.9, 0.97, 1, 1]$;	fire possibility	low, high	$[0, 0, 0.3, 0.5]$; $[0.4, 0.7, 1, 1]$;
	Saturation	low, high	$[0, 0, 0.4, 0.65]$; $[0.545, 0.75, 1, 1]$;	fire color	red, orange, yellow, other	$[0.5, 1, 1.75]$; $[1.25, 2, 2.75]$; $[2.25, 3, 3.75]$; $[3.25, 4, 4.5]$;
	Lightness	low, medium, high	$[0, 0, 0.23, 0.39]$; $[0.23, 0.427, 0.85, 0.96]$; $[0.94, 0.965, 1, 1]$;			
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	$[0, 0, 0.49, 0.625]$; $[0.58, 0.647, 0.69, 0.78]$; $[0.71, 0.826, 1, 1]$;			

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.

1) 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. 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. 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. 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.

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.

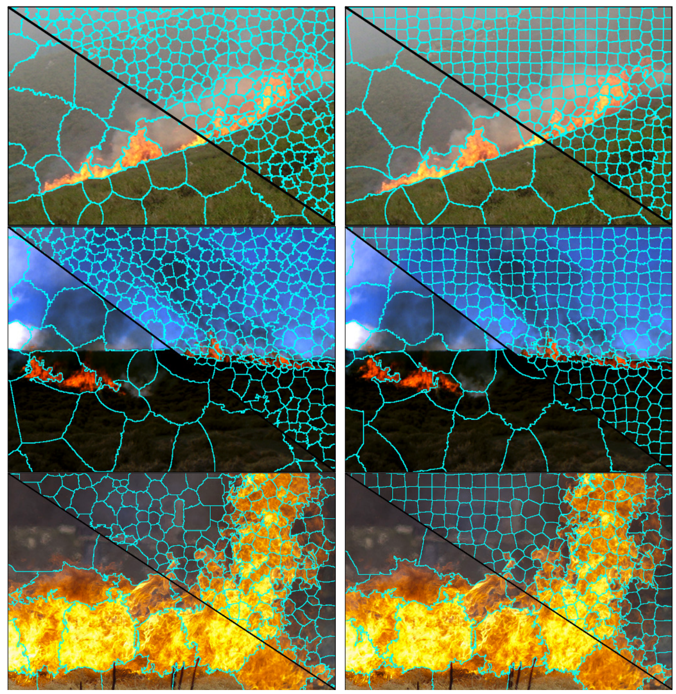


Figure 7: Visual comparison between different N_{sp} and C. The N_{sp} in the lower left corner of each image is 100 and 2000 in the upper right corner. Images on the left have a C equal to 1 and images on the right to 20.

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.

V. Experiments

A. Performance Evaluation

The Fire Image Dataset (Section III) is a subset of the Corsican Fire Database (CFDB), which provides pixel-level segmentations annotated by a single expert as reported in [12]. 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.

1) 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.

2) 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.

B. Results Evaluation

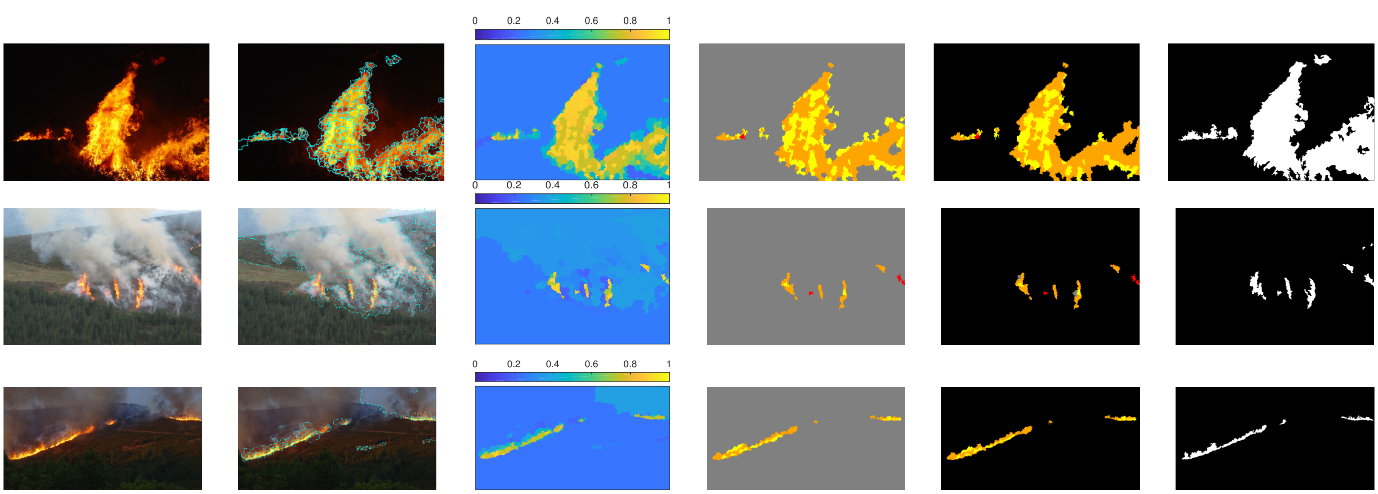
To demonstrate the capabilities of the proposed algorithm, a set of representative examples was selected for analysis. The examples depicted in Fig. 8, showcase the performance of the internal steps of the semantic segmentation for three distinct scenarios. The figure represents the pipeline by the first 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. 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.

C. 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. 4 obscured by smoke, the fine-tuning of the classification threshold can improve the final semantic segmentation as is illustrated in Fig. 10.

The design process described throughout this paper was evaluated with the incremental assembly of its building blocks. Table II 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



(a) Original Image (b) Merged (c) Fire Possibility (d) Fire Color (e) Output (f) Ground truth

Figure 8: Examples for each step in real-world scenarios

Table II: **Ablation Study.**

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 9: Examples of limitations in real-world scenes.

(a) Threshold, $\delta = 0.5$ (b) Threshold, $\delta = 0.4$

Figure 10: Fine-tuning of fire classification threshold for improving the semantic segmentation.

maximum value operator and by lowering the classification threshold to 0.4.

VI. 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 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.

References

[1] 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*, vol. 11, pp. 1419–1425, 2017.

[2] T. Toulouse, L. Rossi, M. Akhloufi, T. Celik, and X. Maldague, "Benchmarking of wildland fire colour segmentation algorithms," *IET Image Processing*, vol. 9, no. 12, pp. 1064–1072, 2015.

[3] "Intelligent and vision-based fire detection systems: A survey," *Image and Vision Computing*, vol. 91, p. 103803, 2019.

[4] "Image fire detection algorithms based on convolutional neural networks," *Case Studies in Thermal Engineering*, vol. 19, p. 100625, 2020.

[5] M. J. Sousa, A. Moutinho, and M. Almeida, "Wildfire detection using transfer learning on augmented datasets," *Expert Systems with Applications*, vol. 142, p. 112975, 2020.

[6] "Wildfire smoke detection based on local extremal region segmentation and surveillance," *Fire Safety Journal*, vol. 85, pp. 50 – 58, 2016.

[7] C. Kyrkou and T. Theocharides, "Deep-learning-based aerial image classification for emergency response applications using unmanned aerial vehicles." in *CVPR Workshops*, 2019, pp. 517–525.

[8] "Wildland forest fire smoke detection based on faster r-cnn using synthetic smoke images," *Procedia Engineering*, vol. 211, pp. 441 – 446, 2018, 2017 8th International Conference on Fire Science and Fire Protection Engineering (ICFSFPE 2017).

[9] 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*, vol. 15, no. 5, pp. 3113–3122, 2019.

[10] H. Harkat, J. Nascimento, and A. Bernardino, "Fire segmentation using a DeepLabv3+ architecture," in *Image and Signal Processing for Remote Sensing XXVI*, vol. 11533, International Society for Optics and Photonics. SPIE, 2020, pp. 134 – 145.

[11] J. Dehaan, P. Kirk, and D. Icove, *Kirk's fire investigation*. Pearson-Prentice Hall, 05 2011.

[12] 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*, vol. 92, pp. 188–194, 2017.

[13] P. Neubert and P. Protzel, "Superpixel benchmark and comparison," in *Proceedings Forum Bildverarbeitung*, 2012, pp. 1–12.

[14] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Ssstrunk, "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 2274–2282, 2012.

[15] J. Chamorro-Martínez and J. M. Keller, "Granular modelling of fuzzy color categories," *IEEE Transactions on Fuzzy Systems*, pp. 1–1, 2019.

[16] E. H. Mamdani, "Application of fuzzy logic to approximate reasoning using linguistic synthesis," *IEEE Transactions on Computers*, no. 12, pp. 1182–1191, 1977.