

# Assessing the impact of the loss function, architecture and image type for Deep Learning-based wildfire segmentation

J. F. Ciprián-Sánchez<sup>1</sup>, Gilberto Ochoa Ruiz<sup>1</sup>, Lucile Rossi<sup>2</sup>

<sup>1</sup>Tecnologico de Monterrey, School of Engineering and Sciences, Mexico

<sup>2</sup>Università di Corsica, UMR CNRS 6134 SPE, France

A01373326@itesm.mx, gilberto.ochoa@tec.mx, rossil@univ-corse.fr

## Abstract

*Wildfires stand as one of the most relevant natural disasters worldwide, particularly more so due to the effect of global warming and its impact on various societal and environmental levels. In this regard, a significant amount of research has been done to apply traditional computer vision techniques, using several imaging modalities and technologies to address this problem. Although there is work regarding Deep Learning (DL)-based fire segmentation, it is currently unclear whether the architecture of a model, its loss function, or the image type employed (visible, infrared, or fused) has the most impact on the fire segmentation results. In the present work, we evaluate different combinations of SOTA DL architectures, loss functions, and types of images to identify the parameters most relevant for the improvement of the segmentation results. Finally, we benchmark the generated combinations to identify the top-performing one and compare it to traditional fire segmentation techniques. To the best of our knowledge, this is the first work that evaluates the impact of the architecture, loss function, and image type in the performance of DL-based wildfire segmentation models.*

## 1. Introduction

Wildfires can occur naturally or due to human activities and have the potential to get out of control and have a significant impact on the environment, properties, and lives. We can see examples of the latter in the terrible damage caused by the Australian wildfires of 2019 and 2020 that took the lives of at least 28 people [14] and the devastating 2020 wildfire season in California in which 6.7 million acres burned [9]. Tasks such as wildfire detection, segmentation, and characterization are relevant as geometric features of a fire are necessary to understand and model the events that develop during its propagation [13].

Fire segmentation is a relevant task as it allows the detection of the distribution region of the flame. The latter enables a quick location of specific areas of interest [2]. The main advantage of the segmentation of RGB images is the accurate detection and localization of objects in a single operation [4].

The fusion of visible and infrared images has the potential to improve the robustness, accuracy, and reliability of fire pixel detection systems [16]; however, DL-based methods for wildfire segmentation with visible-infrared fused information have not yet been investigated. Works such as the one by Nematlidi et al. [7] and Toulouse [10] address visible-infrared image fusion for fire imagery, with the FIRE-GAN model [3] being one of the only works addressing a DL-based approach for the said task. Furthermore, it is still unclear if the inclusion of fused information allows for a significant improvement in the segmentation performance of a model or if factors such as the architecture and loss function play a more relevant role in the said performance.

The main research problem of this work focuses on the evaluation of different combinations of DL architectures, loss functions, and types of images (visible, infrared, and fused), identifying the best performing combinations and the elements that display the most impact in the segmentation performance as measured by selected metrics. Finally, we benchmark the top combination against traditional fire segmentation techniques.

The paper proceeds as follows. Section 2 describes the dataset and methods and states the technical contribution of this work. Section 3 presents the results. Finally, Section 4 shows the conclusions and potential future work avenues.

## 2. Data and methods

For the present paper, we employ the visible-infrared image pairs of the Corsican Fire Database, first presented by Toulouse et al. [12]. This dataset contains 640 pairs of visible and near-infrared (NIR) fire images, alongside their

corresponding ground truths for fire region segmentation.

Figure 1 displays a sample visible-NIR image pair from the Corsican Fire Database with its corresponding ground truth.

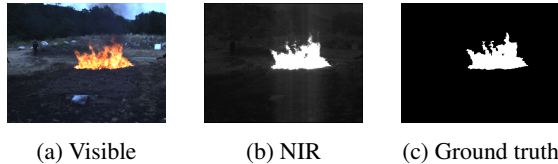


Figure 1: Sample images of the Corsican Fire Database.

The technical contribution of this work is the evaluation of three DL architectures proposed in the current state-of-the-art for visible image-based wildfire segmentation with three loss functions and four types of images (visible, NIR, and fused images generated by two methods). We obtain the source visible and NIR images from the Corsican Fire Database. We then evaluate the segmentation results of the 36 resulting combinations with three selected metrics, identifying the best performing one. In the following subsections, we describe the employed DL architectures, loss functions, and fused images.

### 2.1. Loss functions

We evaluate the performance of three loss functions: the Dice loss, the Focal Tversky loss, and the Mixed Focal loss. The Dice loss is an adaptation of the Dice similarity coefficient, a common metric in the field of computer vision to assess the similarity between two images [5]. The Focal Tversky loss is an adaptation of the Tversky loss that weights down regions that are easy to classify in favour of more difficult ones [15]. Finally, the Mixed Focal loss is proposed by Yeung et al. [15] to handle input and output imbalance; it is a compound loss function derived from variants of the Focal loss and Focal Dice loss functions.

### 2.2. Deep Learning architectures

The first architecture we evaluate is the one proposed by Akhloufi et al. [1]. The authors present a U-Net-based model. The second architecture is the one proposed by Choi et al. [2] that displays a FusionNet-like structure, with the addition of residual blocks to increase FusionNet’s ensemble effect through skip connections. Choi et al. train the network with visible images and a mean squared error (MSE) loss function. Finally, the last architecture we test is the one proposed by Frizzi et al. [4]. The authors implement VGG16 backbone as an encoder, substituting the fully connected layers with a convolution step that serves as a connection to the decoding phase. The decoding step is comprised of transpose convolutions and skip connections in a U-Net-like fashion.

Table 1: Top three best performing combinations per metric.

Metric	Value	Arch.	Loss	Image
MCC	<b>0.9252</b>	<b>Akhloufi</b>	<b>Dice</b>	<b>Visible</b>
	0.9248	Choi	F. Tversky	Visible
	0.9231	Akhloufi	F. Tversky	Visible
F1	<b>0.9323</b>	<b>Akhloufi</b>	<b>Dice</b>	<b>Visible</b>
	0.9274	Akhloufi	F. Tversky	Visible
	0.9265	Choi	F. Tversky	Visible
HAF	<b>0.9098</b>	<b>Akhloufi</b>	<b>Dice</b>	<b>Visible</b>
	0.9068	Choi	F. Tversky	Visible
	0.8957	Akhloufi	F. Tversky	Visible

### 2.3. Fused images

We employ fused images generated by the VGG19 fusion framework proposed by Li et al. [6] and of the previously proposed FIRE-GAN model [3]. Figure 2 displays sample fused images obtained through both methods from the source images presented in Figure 1.



Figure 2: Sample visible-NIR fused images for the VGG19 method in 2a and for the FIRE-GAN model in 2b

## 3. Results

All 36 resulting combinations are trained with 100 epochs, Adam optimizer, batch size of 4, and a learning rate of 0.0001. We apply the benchmarking method that Toulouse et al. [11] use to evaluate their proposed fire segmentation technique, which employs the Matthews Correlation Coefficient (MCC), the F1 score, and the Hafiane quality index (HAF) as evaluation metrics, and compare the results of the best performing one with the best traditional fire segmentation method identified by Toulouse et al. Table 1 displays the top three performing combinations per metric.

We can observe that the best scoring combination for all metrics is the one with the architecture proposed by Akhloufi et al. [1] with the Dice loss and visible images. In Figure 3 we show the results for every parameter to identify the ones that display less variability.

Next, we can see that the architecture by Akhloufi et al. [1] and the Focal Tversky loss present by far the most robust results, displaying little variability. In contrast, the results grouped by image type present very similar results and vari-

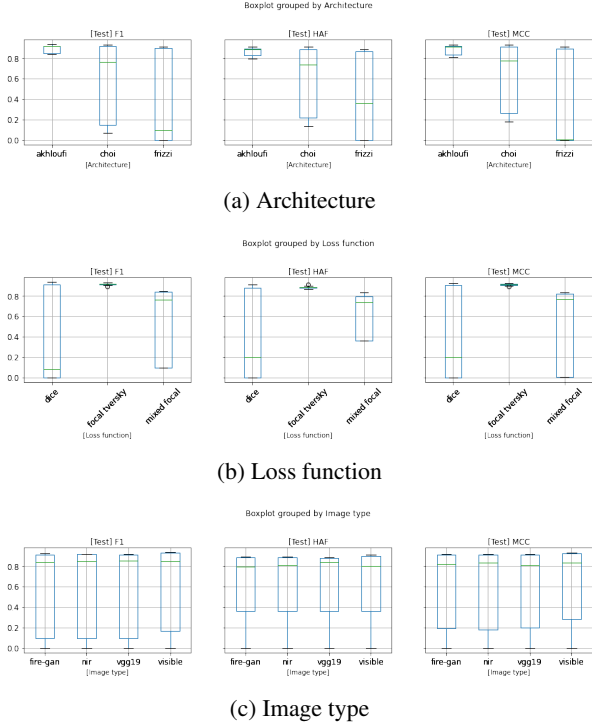


Figure 3: Results per architecture, loss function, and image type.

ability margins. The latter suggests that the performance is more dependant on the architecture and loss function.

We then analyze the obtained data to explore the correlation between the variables (loss functions, architectures, and image types) and the results per metric. In Figure 4, we present the Pearson correlation matrix between the different parameters and the target variables MCC, F1 and HAF.

First, we observe that the target variables are highly correlated with each other. The latter means that in future works, employing just one metric would suffice for performance evaluation. We can also see that the Akhloufi architecture and the Focal Tversky loss display a high positive correlation with the three evaluation metrics. In contrast, the correlation of all image types with the said metrics is close to zero, with the visible images displaying a very weak positive correlation.

Finally, taking into account the robustness of each parameter, as well as their correlation with the performance metrics, we select the Akhloufi architecture coupled with the Focal Tversky loss and visible images as the best one for further comparison with the best traditional fire segmentation technique found by Toulouse et al. [11]. Table 2 displays the results for this comparison.

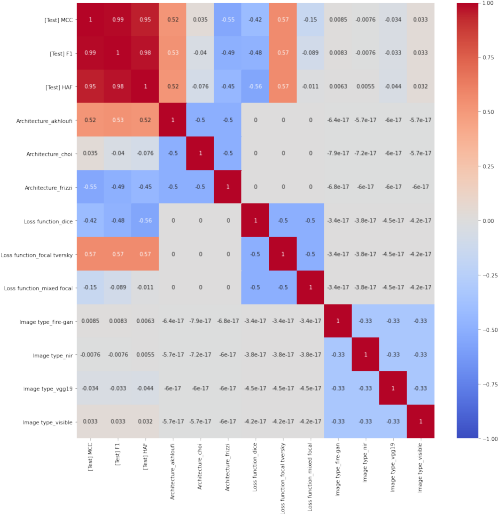


Figure 4: Correlation matrix of the parameters and target variables MCC, F1 and HAF.

Table 2: Comparison between best found combination and the best traditional segmentation method per metric.

Metric	Method	Value
MCC	<b>Akhloufi + F. Tversky + visible</b>	<b>0.92</b>
	Phillips et al. [8]	0.81
F1	<b>Akhloufi + F. Tversky + visible</b>	<b>0.92</b>
	Phillips et al. [8]	0.82
HAF	<b>Akhloufi + F. Tversky + visible</b>	<b>0.89</b>
	Phillips et al. [8]	0.75

## 4. Conclusion and future work

In this work, we can observe that for the images of the Corsican Fire Database, the architecture and loss function of a DL model for fire segmentation appear to be more relevant for its performance than the image type. Since the visible and FIRE-GAN images are color ones, and the NIR and VGG19 ones are in a grayscale format, the presence of color does not appear to make a significant difference for DL methods. The latter is a relevant difference against traditional fire segmentation methods, in which color is one of the most important factors. We consider the combination of *Akhloufi + Focal Tversky + visible* images the best performing one. This combination clearly outperforms the best traditional fire segmentation method.

Finally, the NIR and fused images are expected to provide an advantage when the smoke occludes significant portions of the fire regions. The development of fire image datasets with these types of images could allow for the comparison of the segmentation performance between images with significant smoke occlusion and less challenging ones.

## References

- [1] Moulay A. Akhlooui, Roger Booto Tokime, and Hassan Elasadly. Wildland fires detection and segmentation using deep learning. In Mohammad S. Alam, editor, *Pattern Recognition and Tracking XXIX*, volume 10649, pages 86 – 97. International Society for Optics and Photonics, SPIE, 2018. 2
- [2] Han-Soo Choi, Myeongho Jeon, Kyungmin Song, and Myungjoo Kang. Semantic fire segmentation model based on convolutional neural network for outdoor image. *Fire Technology*, Jan 2021. 1, 2
- [3] J. F. Ciprián-Sánchez, G. Ochoa-Ruiz, M. Gonzalez-Mendoza, and L. Rossi. Fire-gan: A novel deep learning-based infrared-visible fusion method for wildfire imagery, 2021. 1, 2
- [4] Sebastien Frizzi, Moez Bouchouicha, Jean-Marc Ginoux, Eric Moreau, and Mounir Sayadi. Convolutional neural network for smoke and fire semantic segmentation. *IET Image Processing*, n/a(n/a), 2021. 1, 2
- [5] S. Jadon. A survey of loss functions for semantic segmentation. In *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–7, 2020. 2
- [6] H. Li, X. Wu, and J. Kittler. Infrared and visible image fusion using a deep learning framework. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 2705–2710, 2018. 2
- [7] Siva Mouni Nimalidinne and Deep Gupta. Nonsubsampled contourlet domain visible and infrared image fusion framework for fire detection using pulse coupled neural network and spatial fuzzy clustering. *Fire Safety Journal*, 101:84 – 101, 2018. 1
- [8] Walter Phillips III, Mubarak Shah, and Niels da Vitoria Lobo. Flame recognition in video. *Pattern Recognition Letters*, 23(1):319–327, 2002. 3
- [9] The Visual and Data Journalism Team. California and oregon 2020 wildfires in maps, graphics and images. <https://www.bbc.com/news/world-us-canada-54180049>, Sep 2020. Accessed: 2020-09-30. 1
- [10] Tom Toulouse. *Estimation par stéréovision multimodale de caractéristiques géométriques d’un feu de végétation en propagation*. PhD thesis, Université De Corse – Pasquale Paoli, 11 2015. 1
- [11] Tom Toulouse, Lucile Rossi, Moulay Akhlooui, Turgay Celik, and Xavier Maldague. Benchmarking of wildland fire colour segmentation algorithms. *IET Image Processing*, 9(12):1064–1072, 2015. 2, 3
- [12] Tom Toulouse, Lucile Rossi, Antoine Campana, Turgay Celik, and Moulay A. Akhlooui. Computer vision for wildfire research: An evolving image dataset for processing and analysis. *Fire Safety Journal*, 92:188 – 194, 2017. 1
- [13] Tom Toulouse, Lucile Rossi, Turgay Celik, and Moulay Akhlooui. Automatic fire pixel detection using image processing: a comparative analysis of rule-based and machine learning-based methods. *Signal, Image and Video Processing*, 10(4):647–654, Apr 2016. 1
- [14] Jessie Yeung. Australia’s deadly wildfires are showing no signs of stopping. here’s what you need to know. <https://edition.cnn.com/2020/01/01/australia/australia-fires-explainer-intl-hnk-scli/index.html>, Jan 2020. Accessed: 2020-04-10. 1
- [15] Michael Yeung, Evis Sala, Carola-Bibiane Schönlieb, and Leonardo Rundo. A mixed focal loss function for handling class imbalanced medical image segmentation, 2021. 2
- [16] Chi Yuan, Youmin Zhang, and Zhixiang Liu. A survey on technologies for automatic forest fire monitoring, detection and fighting using uavs and remote sensing techniques. *Canadian Journal of Forest Research*, 45:150312143318009, 03 2015. 1