POLITECNICO DI TORINO
Repository ISTITUZIONALE

Double-Step deep learning framework to improve wildfire severity classification

Original
Double-Step deep learning framework to improve wildfire severity classification / Monaco, S., Pasini, A., Apiletti, D., Colomba, L., Farasin, A., Garza, P., Baralis, E.. - ELETTRONICO. - (2021). (Workshops of the 24th International Conference on Extending Database Technology/24th International Conference on Database Theory, EDBT-ICDT 2021 Nicosia (Cyprus) March 23-26, 2021).

Availability:
This version is available at: 11583/2871352 since: 2022-04-14T09:30:13Z

Publisher:
OpenProceedings

Published
DOI:

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

26 November 2026

Double-Step deep learning framework to improve wildfire severity classification

Simone Monaco
Politecnico di Torino
Torino, Italy
simone.monaco@studenti.polito.it

Andrea Pasini
Politecnico di Torino
Torino, Italy
andrea.pasini@polito.it

Daniele Apiletti
Politecnico di Torino
Torino, Italy
daniele.apiletti@polito.it

Luca Colomba
Politecnico di Torino
Torino, Italy
luca.colomba@polito.it

Alessandro Farasin
Politecnico di Torino
Torino, Italy
alessandro.farasin@polito.it

Paolo Garza
Politecnico di Torino
Torino, Italy
paolo.garza@polito.it

Elena Baralis
Politecnico di Torino
Torino, Italy
elena.baralis@polito.it

ABSTRACT

Wildfires are dangerous events which cause huge losses under natural, humanitarian and economical perspectives. To contrast their impact, a fast and accurate restoration can be improved through the automatic census of the event in terms of (i) delineation of the affected areas and (ii) estimation of damage severity, using satellite images. This work proposes to extend the state-of-the-art approach, named Double-Step U-Net (DS-UNet), able to automatically detect wildfires in satellite acquisitions and to associate a damage index from a defined scale. As a deep learning network, the DS-UNet model performance is strongly dependent on many factors. We propose to focus on alternatives in its main architecture by designing a configurable Double-Step Framework, which allows inspecting the prediction quality with different loss-functions and convolutional neural networks used as backbones. Experimental results show that the proposed framework yields better performance with up to 6.1% lower RMSE than current state of the art.

1 INTRODUCTION

In the recent years, European countries witnessed an increasing trend in the occurrence of wildfires. According to the annual report of the European Forest Fire Information System, in 2019 more than 1,600 wildfires have been recorded in the European Union: about three times more than the average over the past decade [2, 3]. Those events are causing large losses not only to forests and animals, but also to human lives and cities. The geographical delineation of the affected regions and the estimation of the damage severity are fundamental for planning a proper environment restoration.

The European Union is active in natural disasters monitoring and risk management through the Copernicus Emergency Management Service platform (EMS) [1]: it provides data about past disasters such as forest wildfires and floods. The census of an hazard is usually performed either manually or semi-automatically using in-situ information, images captured from aircrafts and remote-sensing sensors such as satellites. The latter two data sources can be used to develop computer vision systems, mainly based on neural networks, to automatize the entire detection and damage estimation process.

For this purpose, we use satellite images acquired by Copernicus Sentinel-2 mission to automatically identify burnt areas [22] and to assess the damage severity without requiring human efforts. We can identify two different approaches to address this task: (i) assigning a class label to each pixel of the satellite image (i.e., burnt or unburnt), or (ii) an increasing number representing the damage intensity. The former can be modeled with the well-known computer vision task called semantic segmentation, while the latter requires a regression methodology.

The current state of the art proposes a solutions based on Convolutional Neural Networks (CNNs), called Double-Step U-Net (DS-UNet) [7], which involves both binary semantic segmentation and regression to obtain a damage-severity map. Specifically, each pixel is labeled with a numerical value representing the damage level: 0 - No damage, 1 - Negligible to slight damage, 2 - Moderately damaged, 3 - Highly Damaged, and 4 - Completely destroyed. The network is trained according to the official hazard annotations, named grading maps, publicly available on Copernicus EMS [1].

Previous works on semantic segmentation showed that the appropriate configuration of the CNN structure and the choice of loss functions have significant impacts on the final results [12, 15]. In this work we aim to improve the performances of the Double-Step U-Net maintaining the base architecture, composed of two separated CNN modules, but assessing different CNNs and loss functions. Hence, we propose the *Double-Step Framework (DSF)*, a configurable architecture whose modules allow an in-depth analysis on the effects of different loss-functions and CNNs, comparing the results with the baseline in [7].

Our contribution can be summarized as follows: (i) we define the *Double-Step Framework*, inspired by the DS-UNet neural network, and (ii) we show detailed experimental results on classification and regression tasks, comparing the different configurations.

Our paper is organized as follows. Section 2 presents the related works, while Section 3 discusses the neural network model and the proposed variations in terms of deep network backbones

© 2021 Copyright held by the owner/author(s). Published in Proceedings of the 24th International Conference on Extending Database Technology (EDBT), March 23-26, 2021, ISBN 978-3-89318-084-4 on OpenProceedings.org.
Distribution of this paper is permitted under the terms of the Creative Commons license CC-by-nc-nd 4.0.

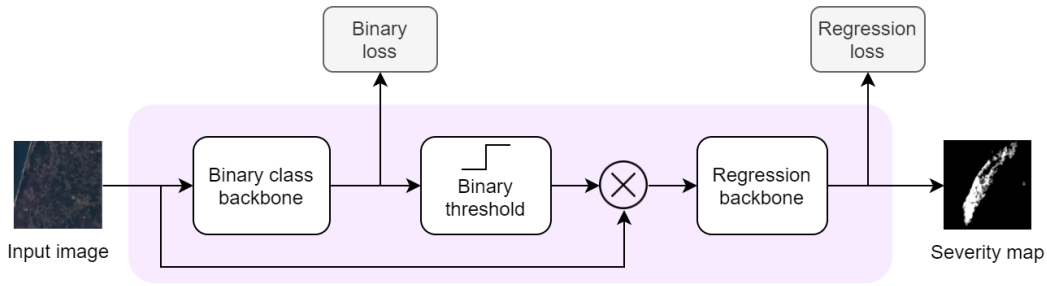


Figure 1: Double-Step Framework architecture.

and loss-functions. Finally, Section 4 shows the experimental results and Section 5 draws conclusions.

2 RELATED WORK

In this section, we firstly review previous works on wildfire prediction and severity classification, then we analyze the state-of-the-art architecture, addressing the semantic segmentation problem. Then we focus on the adopted loss functions, highlighting the differences with the proposed techniques.

Many previous works are used to monitor the evolution of wildfires during the event to support domain experts. Some of these techniques are implemented via deep learning models [5, 18]. Differently, in this paper we are focused on automatic detection of involved areas and damage estimation after the event, by only exploiting post-event satellite images.

The burnt area identification problem is well-known in remote sensing literature: many different approaches have been proposed and recently, machine learning and deep learning-based approaches are being considered, such as [9, 19]. Some mapping operations are performed based on in-situ information, such as the Composite Burned Index (CBI) [14], which are time-consuming and requires evaluations of the soil and vegetation conditions for the entire area of interest (AoI). Other approaches exploit the use of remote sensing techniques and burnt area indexes: satellites collects information across different bandwidths, some of which are sensible to water and vegetation. Specifically, we consider 12 bandwidths available from Sentinel2-L2A products. Burnt area indexes highlight burnt regions by combining specific bandwidths and eventually comparing pre-fire and post-fire acquisitions: Normalized Burn Ratio (NBR) [17], delta Normalized Burn Ratio (dNBR) [16] and Burned Area Index for Sentinel2 (BAIS2) [8] are some examples. Different approaches use such indexes to identify damaged areas and eventually assess the severity level [21].

These methodologies showed so far suffer from a strong dependence on the different weather conditions of the satellite acquisitions. Moreover, the usage of indexes to estimate the damage severity level typically requires the manual or semi-manual definition of predefined thresholds that are usually soil-dependent and cannot be easily set. The solutions adopted in this work solve the previously mentioned issues by only including post-fire images and applying a supervised prediction approach on pre-labelled severity maps. Specifically, we apply a semantic segmentation model, combined with a regression one, to derive the final result. Many different semantic segmentation architectures have been proposed in literature [4, 6, 23], but the work in [7] shows that U-Net [20] is a valuable choice for addressing the wildfire damage-severity estimation task.

The state-of-the-art solution proposes a Double-Step U-Net architecture. This double step configuration relies on the Dice loss function to learn predicting the boundaries of wildfires, and on the Mean Squared Error (MSE) function for estimating the final severity level. Many other different loss functions have been proposed in literature [11], and several works showed that a correct choice typically makes a real difference in the results [12].

3 DOUBLE-STEP FRAMEWORK

In this section we define the Double-Step Framework (DSF), with the aim of obtaining a configurable architecture based on the Double-Step U-Net working principles. The proposed framework allows a complete customization of both training loss functions and backbone neural networks. The main building blocks of the DSF are depicted in Figure 1 and their functionalities are described in the following paragraphs.

Binary class backbone. This building block has the task of assigning a binary label (i.e., burnt or unburnt) to each pixel of the input image. Its output is a probability map with values in the range $[0, 1]$.

Binary threshold. The output probabilities of the Binary class backbone are thresholded to obtain the final binary mask, highlighting regions affected by wildfires. The value of the threshold is fixed to 0.5 in all the experiments.

Regression backbone. This step aims at deriving a severity map to specify the damage intensity in range $[0, 4]$ for each pixel. It takes as input the product between the binary mask and the original input image, in order to consider the satellite image information only for the regions that have been classified as burnt by the Binary class backbone. Indeed, accurate binary masks are fundamental to provide only the information related to regions affected by wildfires. False positives (i.e., unburnt areas classified as burnt) have shown to negatively affect the regression quality.

Binary loss. This loss function is exploited to train the Binary class backbone, by comparing its output with ground-truth binary masks.

Regression loss. After the completion of the training process of the Binary class backbone, this loss function is used to train the Regression backbone. During this training phase, the weights of the Binary class backbone are kept constant.

The Binary class backbone, the Regression backbone, and the two loss functions defined for the DSF can be customized to obtain several configurations. In the following, we present the different options available for these configurable modules, dividing the analysis in two parts: (i) backbone architectures, and (ii) loss functions.

Table 1: Loss function selection experiments.

Config. name	Binary loss	Regression loss
BCE-MSE	BCE	MSE
Dice-MSE	Dice	MSE
B+D-MSE	Compound BCE, Dice	MSE
B+S-MSE	Compound BCE, sIoU	MSE
sIoU-sIoU	sIoU	sIoU
sIoU-MSE	sIoU	MSE

3.1 Backbone architectures

The Binary class and the Regression backbones can be implemented with a custom encoder-decoder neural network. We propose three different DSF configurations for these modules, by changing the backbone architectures. Specifically, we selected the following models: U-Net [20], U-Net++ [24], and SegU-Net [13]. When choosing one among the proposed backbone architectures, we use the same one for both the Binary class and the Regression backbone. In the next sections of this paper we refer to these configurations with the names Double-Step U-Net (DS-UNet), Double-Step U-Net++ (DS-UNet++), and Double-Step SegU-Net (DS-SegU).

The state-of-the-art Double-Step U-Net [7] is exactly reproduced by our framework when choosing the DS-UNet configuration. The U-Net in the Binary class backbone is set up with a sigmoid activation function to generate the probability map, while for the Regression backbone we do not use any activation function, since the output values may range in $[0, 4]$.

The DS-UNet++ follows the same working principles and differs only by the selected neural network. Specifically, U-Net++ enhances the structure of the standard U-Net by adding convolutional layers in correspondence of the skip connections between the encoder and the decoder.

Finally, the DS-SegU configuration exploits another variation of the standard U-Net. In particular, with the SegU-Net network, the skip-connections typical of the U-Net are integrated into SegNet [4], which is based on pooling indices to provide information from the encoder to the decoder.

3.2 Loss functions

This section describes the different loss functions that we propose for training the Binary class and the Regression framework. The complete list of configurations is specified in Table 1. The first column of the table provides the configuration name, used in the experiments in Section 4, while the other two columns specify the corresponding Binary and Regression loss.

Binary loss. For the Binary loss function we consider Binary Cross Entropy (BCE), Dice, sIoU, and two compound loss functions (i.e., B+D, B+S). In the following we provide the main characteristics of these loss functions.

The sIoU (soft Intersection over Union) is defined as a per-pixel `AND-like` operation applied between the ground-truth image and the network estimation to get the Intersection, and a per-pixel `OR-like` operation to get the Union. Differently to standard IoU, the sIoU is computed directly on the probability map predicted by the neural network, without discretizing the values to a binary mask. This allows evaluating the actual distance between the prediction and ground truth, for a more effective calculation of gradients. The definition of the sIoU loss function can be formalized as follows:

$$L_{sIoU} = 1 - I_{soft}/U_{soft}$$

where I_{soft} and U_{soft} are the soft intersection and the soft union, respectively. Compound loss functions have shown to be an effective way for training neural networks [15].

They are typically defined as a weighted sum of standard loss functions. In this work we inspected the effectiveness of B+D, defined as $B{+}D = 0.5 \cdot BCE + 0.5 \cdot Dice$, and B+S, defined as $B{+}S = 0.5 \cdot BCE + 0.5 \cdot L_{sIoU}$.

Regression loss. Since the output values of the Regression backbone can range into 5 severity levels, for the regression loss we considered a second set of functions. Specifically, we inspected the results obtained with the Mean Squared Error (MSE), a generalization of the sIoU to a multiclass case, and a combination of the MSE and the F_1 score.

In the case of sIoU, predictions and ground truth are compared by considering separately the pixels corresponding to each severity level. The division of the pixels based on the severity level is made by applying rectangular functions to the matrices. In the case of the network prediction matrix, to avoid defining a sharp selection of the severity levels (i.e., loosing important information for the gradient), we applied smooth rectangular functions. After computing the intersections and the unions between ground truth and predictions, the contribution of each severity level is finally summed up in the final sIoU function.

Let Π_c be a sharp rectangular function that takes the value 1 when the input pixel belongs to class c and 0 otherwise. Let $\tilde{\sigma}_c(x)$ be a smooth rectangular function, defined as $\tilde{\sigma}_c(x) = \sigma(\epsilon - |x - c|)$, where $\epsilon = 0.5$ and σ is the sigmoid function. The sIoU loss function, is defined as:

$$L_{sIoU,reg} = \frac{\sum_c |\Pi_c(Y_{\text{GT}}) \circ \tilde{\sigma}_c(Y_{\text{PR}})|}{\sum_c |\Pi_c(Y_{\text{GT}}) + \tilde{\sigma}_c(Y_{\text{PR}}) - \Pi_c(Y_{\text{GT}}) \circ \tilde{\sigma}_c(Y_{\text{PR}})|},$$

where Y_{GT} is the ground-truth matrix, Y_{PR} are the predictions, and the symbol $\circ$ represents the element-wise product between two matrices. Given this definition, for each class, the intersection is represented by the product between the two matrices and the union is given by their sum minus the intersection.

The last loss function we considered is inspired from the fact that the second network is designed for a regression task, but actually the final result admit a set of classes. Hence we built a function both penalizing the distance from the ground truth and favouring the consistency with the real classes. The two contributions are provided by the MSE loss and the F_1 score the result obtain on the 5 classes, multiplied together following:

$$L_{MSE \cdot \text{F}_1} = L_{MSE} \cdot (1 - \text{F}_1).$$

4 EXPERIMENTAL RESULTS

In this section we provide the evaluation of the proposed Double-Step Framework, by inspecting the results with all the configurations described in Section 3. We also show a detailed comparison with other standard encoder-decoder architectures.

The next subsections are organized as follows. Section 4.1 describes the analyzed dataset, Section 4.2 outlines the experimental setting, while Section 4.3 provides the results to assess the modules of the DSF. Finally, Section 4.4 compares our framework with other single-step architectures.

4.1 Dataset analysis

The experimental setting adopted in this paper follows the same dataset preparation as in [7]. Specifically, the satellite images are

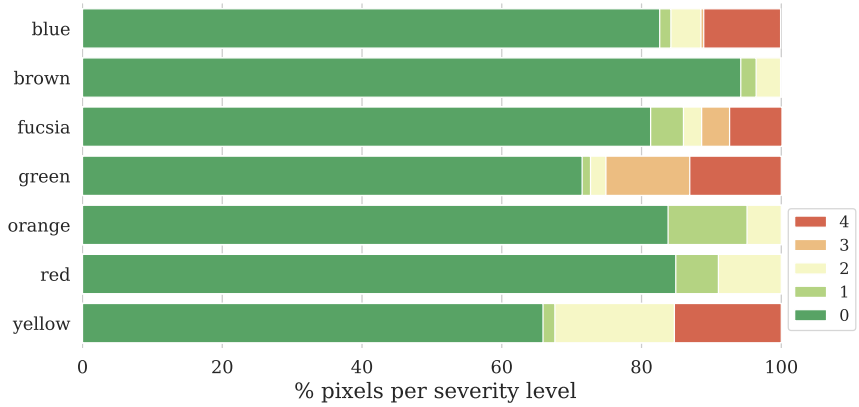


Figure 2: Distribution of the 5 severity levels for each fold.

Table 2: IoU of burnt class for the Binary classification backbone.

Model	BCE	Dice ([7])	B+D	B+S	sIoU
DS-UNet	**0.80**	0.58	0.58	0.38	0.39
DS-UNet++	**0.79**	0.47	0.50	0.37	0.30
DS-SegU	**0.63**	0.24	0.19	0.15	0.14

extracted from the Copernicus Emergency Management Service dataset (Copernicus EMS) [1], focusing on the samples acquired by Sentinel2 (L2A products). The satellite acquisitions represent terrain areas with matrices of variable size (approximately 5000 × 5000) and 12 channels (for the different acquisition bandwidths). Each sample is manually annotated with pixel-wise ground-truth severity levels corresponding to the damage intensity caused by the wildfire. The number of severity levels is 5 (i.e., from 0 for no damage, to 4 for completely destroyed).

The images are provided to the neural networks under analysis by tiling them into squares with size 480 × 480. Indeed, their original size is too large for being consumed by these deep learning models. After excluding the samples without burnt regions, the dataset contains a total of 135 tiles. These data are then distributed into 7 different folds based on the geographical proximity of the analyzed regions (i.e., close regions typically share the same morphology).

The percentage of pixels of the 5 severity levels in the 7 dataset folds is provided in Figure 2. The plot shows that the class 0 (i.e., no damage) is predominant over all the others. Moreover, different folds present significantly different distributions of the severity levels, which confirms the difficulty of the prediction task.

4.2 Experimental setting

Motivated by the small dataset size and the unbalanced classes, data augmentation techniques have been performed to change the variability of the training data at each epoch, applying random rotations, horizontal/vertical flips, and random shears.

After applying data augmentation, we run a cross-validation for each model under analysis. At each iteration, five folds out of seven are used for training, 1 for validation (i.e., to enable early stopping), and 1 for testing. The early stopping process is configured with patience 5 and a tolerance of 0.01 on the loss function.

To enhance the reliability of the results, cross-validation is run 5 times for each model configuration. All the evaluation metrics are computed separately for each run and each cross-validation iteration, then averaged to obtain the final scores.

The output of the analyzed neural networks is evaluated in a (i) regression fashion, and a (ii) classification fashion. The first case exploits the Root Mean Squared Error (RMSE) to verify the quality of the predictions. Due to dataset imbalancing, the RMSE is computed separately for each severity level for a proper evaluation. Specifically, given a severity level, we compute the RMSE between all the ground-truth pixels with that value and the neural network predictions.

Since severity levels in the ground-truth annotations are provided in the form of discrete numbers, we also applied a classification metric for the evaluation. Specifically, we computed the Intersection over Union (IoU) between ground truth and the predictions discretized to integer values. Similarly to the RMSE evaluation, the IoU is computed separately for each severity level.

4.3 Loss function selection

We begin the assessment of the Double-Step Framework by focusing on the Binary classification backbone. To this aim, Table 2 evaluates the Binary classification backbone by providing the IoU of the burnt class. This phase inspects the ability of the network in distinguishing between burnt and undamaged areas, regardless of the severity levels. The results clearly show that the BCE loss function brings an important advantage with respect to the others, reaching 0.80 IoU for the DS-UNet and 0.79 for the DS-UNet++. The Dice loss function, exploited in the original Double-Step U-Net, compares to BCE with moderately lower results for the DS-UNet and the DS-UNet++ (0.58 and 0.47 respectively) and a very low score (i.e., 0.24) for the DS-SegU.

Table 3: Results on burnt-areas only, with different loss functions.

Metric	Model	BCE-MSE	Dice-MSE [7]	B+D-MSE	B+S-MSE	BCE-MSE·F_1	sIoU-sIoU	sIoU-MSE
avg RMSE	DS-UNet	**1.08**	1.15	1.13	1.27	1.12	1.64	1.31
	DS-UNet++	**1.10**	1.28	1.19	1.35	1.14	2.31	1.28
	DS-SegU	1.45	1.60	1.73	1.79	**1.38**	2.50	1.79
avg IoU	DS-UNet	**0.16**	0.13	0.14	0.14	0.13	0.10	0.12
	DS-UNet++	**0.16**	0.11	0.13	0.13	0.14	0.15	0.11
	DS-SegU	0.12	0.14	0.14	**0.15**	0.14	0.14	0.13

Table 4: Architecture selection results (RMSE).

Severity	DS-UNet BCE-MSE	DS-UNet++ BCE-MSE	DS-SegU BCE-MSE·F_1	Unet++ MSE	PSPNet MSE	SegU-Net MSE
0	0.30	0.33	**0.23**	1.04	1.14	0.39
1	1.09	1.00	**0.79**	1.16	1.37	0.91
2	1.04	0.95	1.09	**0.93**	1.21	1.11
3	0.96	0.97	1.33	**0.91**	1.09	1.44
4	**1.25**	1.50	2.33	1.35	1.38	2.14
avg (1-4)	**1.08**	1.10	1.38	1.09	1.26	1.40

Table 5: Architecture selection results (IoU).

Severity	DS-UNet BCE-MSE	DS-UNet++ BCE-MSE	DS-SegU B+S-MSE	Unet++ MSE	PSPNet MSE	SegU-Net MSE
0	**0.95**	0.94	0.68	0.00	0.00	0.82
1	0.11	**0.13**	0.08	0.01	0.01	0.09
2	**0.22**	0.21	0.07	0.19	0.11	0.14
3	0.03	0.07	**0.28**	0.01	0.01	0.08
4	**0.28**	0.21	0.14	0.14	0.16	0.06
avg (1-4)	**0,16**	**0,16**	0,15	0,09	0,07	0,09

Motivated by these results, we inspect the ability of the Double-Step Framework in distinguishing the different severity levels for burnt regions. To this aim, we computed the RMSE and the IoU, averaged for the levels in range [1, 4]. Level 0 is excluded by the average, since it represents the majority class, describing unburnt regions. Table 3 provides the results for all the configurations proposed in Section 3. Both the loss functions and the Binary/Regression backbones are evaluated at this step.

The results clearly show that the BCE-MSE loss function configuration is able to achieve the best results according to RMSE. The only difference is for the DS-SegU, which reach better result with the BCE-MSE·F_1 loss function. For what concerns IoU, the BCE-MSE confirms its first place for the DS-UNet and DS-UNet++, while the loss functions including the sIoU for the Binary classification backbone, namely the combo loss B+S-MSE, achieve a better score for the DS-SegU. Among the three proposed DSF architectures, the DS-UNet with BCE-MSE presents the best result in terms of avg RMSE, while the DS-UNet and DS-UNet++ with BCE-MSE achieve the best IoU.

4.4 Architecture comparison

We complete our experimental results by comparing the prediction quality of the Double-Step Framework with other single-step neural networks. In the following, for the DS-UNet, the DS-UNet++, and the DS-SegU, we only show the results with the best overall loss function configurations for each network. The other neural networks analyzed in this section are the Unet++, PSPNet, and SegU-Net. All of them are trained by means of the MSE loss function. PSPNet is considered as example of a more complex neural network with respect to the other ones. Indeed, this model exploits multiple pyramidal pooling filters to capture features at different resolutions. In our case we used a PSP layer including pooling kernels with size 1, 2, 3, 6 and the ResNet18 [10] as backbone. We did not use deeper ResNet models due to possible underfitting issues (caused by the small size of the analyzed dataset).

Table 4 and 5 show the complete set of results, analyzed with RMSE and IoU respectively. The first five lines of the two tables present the scores separately for the severity levels. The final line provides the average score excluding level 0 (i.e., undamaged regions).

According to the average RMSE (Table 4), the best model is the DS-UNet, with value 1.08. Despite this result, it only reach the best score for level 4 regions with respect to other models. Looking at average IoU scores (Table 5), the DS-UNet obtains again the best result, together with the DS-UNet++ when considering the average score.

For the majority class (i.e., level 0), SegNet achieves a very low RMSE (0.01). However, this model makes important errors on all

the other severity levels, probably because the model is incline to predict class 0 most of the times.

5 CONCLUSIONS AND FUTURE WORKS

The objective of this paper was to define a complete experimental setting to compare different architectures for wildfire severity prediction. We defined a Double-Step Framework, with customizable loss-functions and network backbones. Different backbones and loss functions have been evaluated according to RMSE and IoU, showing that the Double-Step Framework tends to give more accurate results with respect to single-step neural networks.

ACKNOWLEDGMENT

The research leading to these results has been partially supported by the SmartData@PoliTO center for Big Data and Machine Learning technologies, and the HPC@PoliTO center for High Performance Computing. The authors are grateful to Moreno La Quatra for his help in exploiting the HPC resources.

REFERENCES

[1] 2009 (accessed November 9, 2020). *European Union. Copernicus Emergency Management Service. 2020.* https://emergency.copernicus.eu/

[2] 2019. Euronews. https://www.euronews.com/2019/08/15/there-have-been-three-times-more-wildfires-in-the-eu-so-far-this-year. Accessed: 2020-12-03.

[3] 2019. European Forest Fire Information System (EFFIS) - Annual Reports. https://effis.jrc.ec.europa.eu/reports-and-publications/annual-fire-reports/. Accessed: 2020-12-03.

[4] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence* 39, 12 (2017), 2481–2495.

[5] Yifang Ban, Puzhao Zhang, Andrea Nascetti, Alexandre R Bevington, and Michael A Wulder. 2020. Near Real-Time Wildfire Progression Monitoring with Sentinel-1 SAR Time Series and Deep Learning. *Scientific Reports* 10, 1 (2020), 1–15.

[6] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. 2017. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* (2017).

[7] Alessandro Farasin, Luca Colomba, and Paolo Garza. 2020. Double-Step U-Net: A Deep Learning-Based Approach for the Estimation of Wildfire Damage Severity through Sentinel-2 Satellite Data. *Applied Sciences* 10, 12 (2020), 4332.

[8] Federico Filipponi. 2018. BAIS2: burned area index for Sentinel-2. In *Multidisciplinary Digital Publishing Institute Proceedings*, Vol. 2. 364.

[9] Leonardo A Hardtke, Paula D Blanco, Hector F del Valle, Graciela I Metternicht, and Walter F Sione. 2015. Semi-automated mapping of burned areas in semi-arid ecosystems using MODIS time-series imagery. *International Journal of Applied Earth Observation and Geoinformation* 38 (2015), 25–35.

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[11] Shruti Jadon. 2020. A survey of loss functions for semantic segmentation. *arXiv preprint arXiv:2006.14822* (2020).

[12] Shruti Jadon, Owen P Leary, Ian Pan, Tyler J Harder, David W Wright, Lisa H Merck, and Derek L Merck. 2020. A comparative study of 2D image segmentation algorithms for traumatic brain lesions using CT data from the ProTECTIII multicenter clinical trial. In *Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications*, Vol. 11318. International Society for Optics and Photonics, 113180Q.

[13] Uday Kamal, Thamidul Islam Tonmoy, Sowmitra Das, and Md Kamrul Hasan. 2019. Automatic traffic sign detection and recognition using SegU-Net and a modified Tversky loss function with L1-constraint. *IEEE Transactions on Intelligent Transportation Systems* 21, 4 (2019), 1467–1479.

[14] Carl H Key and Nathan C Benson. 2006. Landscape assessment (LA). *In: Lutes, Duncan C.; Keane, Robert E.; Caratti, John F.; Key, Carl H.; Benson, Nathan C.; Sutherland, Steve; Gangi, Larry J. 2006. FIREMON: Fire effects monitoring and inventory system. Gen. Tech. Rep. RMRS-GTR-164-CD. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. LA-1-55* 164 (2006).

[15] Jun Ma. 2020. Segmentation Loss Odyssey. *arXiv preprint arXiv:2005.13449* (2020).

[16] Jay D Miller and Andrea E Thode. 2007. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sensing of Environment* 109, 1 (2007), 66–80.

[17] Gabriel Navarro, Isabel Caballero, Gustavo Silva, Pedro-Cecilio Parra, Águeda Vázquez, and Rui Caldeira. 2017. Evaluation of forest fire on Madeira Island using Sentinel-2A MSI imagery. *International Journal of Applied Earth Observation and Geoinformation* 58 (2017), 97–106.

[18] Miguel M Pinto, Renata Libonati, Ricardo M Trigo, Isabel F Trigo, and Carlos C DaCamara. 2020. A deep learning approach for mapping and dating burned areas using temporal sequences of satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing* 160 (2020), 260–274.

[19] Ruben Ramo and Emilio Chuvieco. 2017. Developing a random forest algorithm for MODIS global burned area classification. *Remote Sensing* 9, 11 (2017), 1193.

[20] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*. Springer, 234–241.

[21] David P Roy, Luigi Boschetti, and Simon N Trigg. 2006. Remote sensing of fire severity: assessing the performance of the normalized burn ratio. *IEEE Geoscience and Remote Sensing Letters* 3, 1 (2006), 112–116.

[22] Dimitris Stavrakoudis, Thomas Katagis, Chara Minakou, and Ioannis Z Gitas. 2019. Towards a fully automatic processing chain for operationally mapping burned areas countrywide exploiting Sentinel-2 imagery. In *RSCy2019*, Vol. 11174. 1117405.

[23] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. 2017. Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2881–2890.

[24] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. 2018. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 3–11.