Baltic J. Modern Computing, Vol. 12 (2024), No. 4, 443-453
https://doi.org/10.22364/bjmc.2024.12.4.06

Comparison of Two Deep Learning Models to Determine Burned Forest Areas from Sentinel-2 Imagery

Ahmet KILIC [(1)], Bahadir KULAVUZ [(2)], Tolga BAKIRMAN [(2)], Bulent BAYRAM [(2)]

[(1)] Yildiz Technical University, Graduate School of Science and Engineering, Data Science and Big Data Program, Istanbul-Türkiye,
[(2)] Yildiz Technical University, Department of Geomatics Engineering, Istanbul-Türkiye

ahmet.kilic2@std.yildiz.edu.tr, bahadir.kulavuz.@yildiz.edu.tr, bakirman@yildiz.edu.tr, bayram@yildiz.edu.tr

ORCID 0009-0000-2285-0549, ORCID 0009-0007-2320-6350, ORCID 0000-0001-7828-9666, ORCID 0000-0002-4248-116X

Abstract. Precise mapping of burned forest areas is essential for monitoring the effects of wildfires and supporting forest management efforts, particularly given the increasing frequency of wildfires driven by climate change. In this study, two image datasets were generated from Sentinel-2 imagery using RGB (Red, Green, Blue) and RGNIR (Red, Green, Near-Infrared) bands to evaluate the effectiveness of these spectral bands for semantic segmentation of burned areas. The U-Net and Feature Pyramid Network (FPN) models were compared for binary segmentation of burned regions using Sentinel-2 satellite data and the Satellite Burned Area Dataset. The U-Net model, utilizing the RGNIR band combination, outperformed the FPN model, achieving an Intersection over Union (IoU) score of 0.7601 and an overall accuracy of 90.92% across 138 test images. These findings underscore U-Net's capacity to extract sufficient spectral and spatial features even with limited training data, providing an efficient method for large-scale mapping of burned areas.

Keywords: Deep learning, U-Net, Semantic segmentation, burned area mapping, Sentinel-2, Remote Sensing

1. Introduction

Forest ecosystems, covering approximately 30% of the Earth's land surface, sequester 45% of the carbon stored in terrestrial ecosystems (Baldrian et al., 2023). However, global forested areas are in decline due to factors such as urbanization and the expansion of agricultural land, both of which contribute to the exacerbation of climate change.

In recent years, the incidence of large forest fires has risen across Mediterranean countries and Europe, where forests cover more than one-third of the total land area. While extreme weather conditions are considered a key factor in the occurrence of local fires in the short term, it is evident that climate change significantly increases the risk of forest fires. Beyond the immediate environmental destruction, these fires have profound ecological and economic consequences for ecosystems, biodiversity, and natural resources (WEB, a). According to Copernicus, 2023 has been recorded as the warmest year in Europe (WEB, b). The relationship between forest fire risk and climate change is complex; however, when these trends intersect with ignition sources and the availability of fuel, large-scale fires, particularly wildfires, are likely to have severe impacts on societies, economies, ecosystems, and the nature of forest fires themselves (WEB, c).

The rapid, cost-effective, and accurate identification of burned forest areas is essential for effective disaster management, ecological restoration, and understanding the broader impacts of global climate change. Remote sensing technology serves as one of the most vital tools for the swift, economical, and timely monitoring, localization, and mapping of burned areas across large geographic scales.

The European Space Agency's Copernicus programme provides multispectral data that has proven highly effective in detecting land surface changes, including those resulting from forest fires (Zhang et al., 2022). Compared to traditional machine learning (ML) models, deep learning (DL) models offer several advantages for forest fire detection (Maskouni and Seydi, 2021). Numerous DL-based studies have been proposed for the segmentation of burned forest areas using Sentinel-2 imagery. For instance, Knopp et al. (2020) developed a new training and validation dataset for the semantic segmentation of burned areas, employing the U-Net architecture, achieving an overall accuracy (OA) of 0.98. Similarly, Hu et al. (2021) explored the potential of DL methods for mapping burned areas using Sentinel-2 and Landsat-8 datasets. Seydi et al. (2021) introduced a framework for burned area mapping based on the Deep Siamese Morphological Neural Network (DSMNN-Net) and heterogeneous datasets, proposing the Burnt-Net framework with Sentinel-2 data, which yielded an OA exceeding 97%. Cambrin et al. (2023) contributed a new dataset of pre- and post-fire Sentinel-2 L2A acquisitions of California forest fires, providing three baselines: spectral index analyses, SegFormer, and U-Net models, with the U-Net achieving an Intersection over Union (IoU) score of 0.577 ± 0.002 without specific pre-training. Sui et al. (2024) proposed BiAU-Net, a novel U-Net-based model for burnt area segmentation using pre- and post-fire Sentinel-2 imagery across five independent regions. Han et al. (2024) developed the Burned Area-Burn Severity (BA-BS) dataset and introduced a multilevel feature fusion mechanism, achieving a 0.781 IoU score using Google Earth Engine with multispectral Landsat 8 data. Lee et al. (2025) proposed an automatic forest fire detection approach using false-colour RNG (B4, B8, and B3) Sentinel-2 images and the HRNet model, achieving an IoU score of 89.40.

The results of burned area mapping using deep learning (DL) approaches demonstrate the strong potential of DL models for effectively identifying burned regions. A review of the literature reveals that DL methods have consistently produced highly promising outcomes. However, one of the primary challenges in this domain is the limited availability of open-access datasets, which remains a significant barrier to further improving these results. Although the Satellite Burned Area Dataset (SBAD) was originally developed for burn severity mapping, its class imbalance posed difficulties for our analysis, leading us to modify it for binary semantic segmentation. Additionally, we focused on two specific combinations of three Sentinel-2 bands and utilized U-Net and

Feature Pyramid Network (FPN) models to assess the accuracy of the dataset with a limited number of samples. In this context, our key contributions are as follows:

- We modified the SBAD open-access dataset for binary segmentation of burned forest areas.
- We evaluated the RGB and RGNIR Sentinel-2 band combinations for burned area segmentation using the modified SBAD dataset.
- We assessed the binary segmentation performance of the FPN and U-Net models on the altered SBAD dataset.

2. Material and Methods

2.1. Satellite Burned Area Dataset (SBAD)

In this study, the open-access Satellite Burned Area Dataset (SBAD) was utilized for the binary segmentation of burned areas in an initial phase aimed at determining burn severity, using data derived from Sentinel-2 imagery (WEB, g). The dataset consists of 73 atmospherically corrected post-fire Sentinel-2 images, each containing 13 spectral bands. These images cover approximately 19,000 km² across various morphologies and terrain types throughout Europe, collected between 2017 and 2019. The dataset includes labelled data for five distinct burn severity levels, with unburned areas accounting for 91.9% of all image pixels. For burned areas, class 1 covers about 12%, while classes 2 through 4 account for roughly 30% of the total, with individual class proportions of 1%, 2.25%, 2.35%, and 2.5%, respectively (Colomba et al., 2022). Table 1 provides details of the class categories and corresponding pixel values, while Figure 1 illustrates an example of the image-label pair. The dataset is labelled with six classes, including one background class and five classes representing varying levels of burn severity.

Table 1. Classes in SBAD dataset (Colomba et al., 2022)

Class	**1 Background**	**2 Unburned**	**3 Low Severity**	**4 Moderate-Low Severity**	**5 Moderate-High Severity**	**6 Hight Severity**
Class Value [0-255]	0	32	64	128	192	255

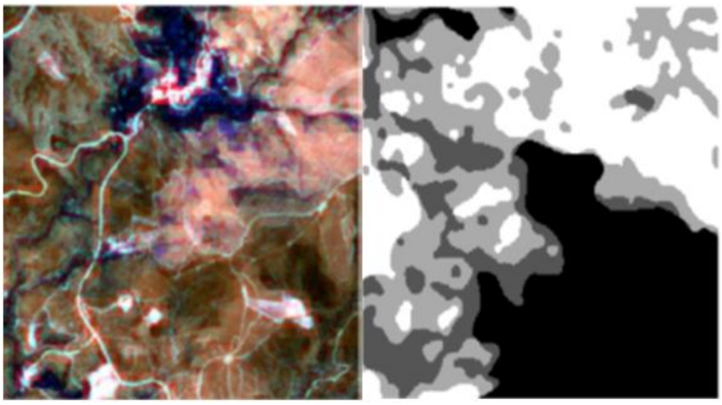

Figure 1. Original and labelled image in 256x256 pixel size

In the initial stage of the study, the suitability of image patch sizes for the Feature Pyramid Network (FPN) and U-Net architectures was assessed. To achieve this, the maximum and minimum dimensions of the image patches generated from the 73 Sentinel-2 images in the dataset were analyzed, as presented in Table 2. Based on this analysis, it was determined that the optimal size for both image and label data should be set to 256x256 pixels. This decision was informed by the range of pixel sizes observed across the 73 images, ensuring that the selected patch size accommodates the spatial characteristics of the dataset for efficient processing by the FPN and U-Net networks.

Table 2. Minimum-maximum width and length values of the original dataset

Minimum Length	Maximum Length	Minimum Width	Maximum Width
261	313	4449	3910

In this study, two distinct image datasets were generated from Sentinel-2 imagery, each with a 256x256 pixel resolution: (i) true color images consisting of Red, Green, and Blue (RGB) bands (Figure 2-a), and (ii) false color images utilizing Red, Green, and Near-Infrared (RGNIR) bands (Figure 2-b). These datasets were created to evaluate the effectiveness of different spectral bands for the semantic segmentation of burned areas. Given that binary segmentation was the primary objective of this study, the six original classes were consolidated into two categories: burned and unburned areas. In the segmented images, yellow represents burned areas, while purple denotes unburned regions (Figure 2-c).

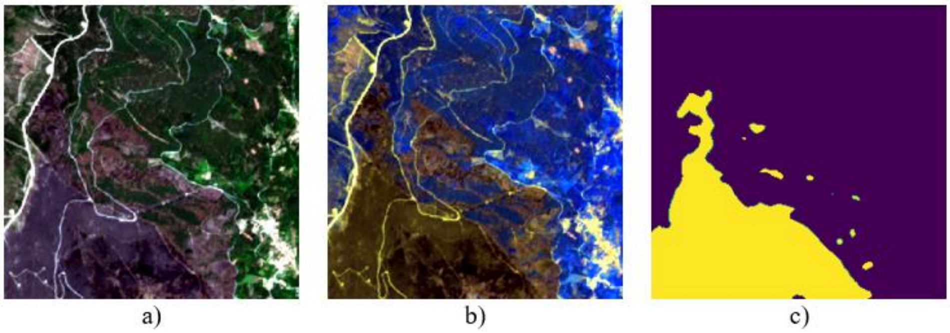

Figure 2. a) RGB image, b) RGNIR image and c) label

To enhance the dataset's accuracy, images with more than 98% unburned area were excluded from the original dataset. This refinement resulted in a total of 1,094 image patches, each with a 256x256 pixel resolution. These patches were then reclassified into two categories for binary segmentation: the first, second, and third original classes were combined into a single "unburned" category, while the fourth, fifth, and sixth classes were consolidated into a "burned" category (see Table 1). Table 3 provides the distribution of these classes based on grey values. After this reclassification, images with more than 98% unburned area were removed again. The final dataset comprised 690 images, which were divided into 552 patches for training and 138 patches for testing.

Table 3. Binary class definition

	Unburned	**Burned**
Merged Classes	0, 32, 64	128, 192, 255

2.2. Deep Learning Models

In this study, the Segmentation Models Library (WEB, d) was employed to assess the binary semantic segmentation capabilities for detecting burned areas. This library, compatible with TensorFlow, provides access to various pre-trained models, including U-Net, Feature Pyramid Network (FPN), Linknet, and PSPNet. The U-Net and FPN models were selected for this research due to their demonstrated success in the literature. Given that deep learning (DL) models represent a state-of-the-art approach for burned area mapping and offer substantial improvements over traditional methods, both U-Net and FPN were utilized to evaluate the modified Satellite Burned Area Dataset (SBAD).

U-Net is a deep learning architecture specifically designed for biomedical image segmentation. Its encoder-decoder structure allows it to achieve high accuracy even with limited training data. U-Net is particularly notable for its pixel-based classification capability. A key feature of U-Net is its use of skip connections, which link high-resolution feature maps from the encoder side to the up-sampling layers on the decoder side. This design facilitates the recovery of details lost at lower resolutions, resulting in more precise pixel-level classification (Ronneberger et al., 2015; WEB, e).

The Feature Pyramid Network (FPN) is a deep learning architecture utilized in image processing tasks such as object detection and segmentation. FPN is designed to effectively identify both small and large objects by leveraging multi-scale features. A distinguishing feature of FPN is its pyramid structure, which enables the simultaneous access of information at various resolutions. This is achieved through the Bottom-Up Pathway, which reduces image size to capture attribute data, and the Top-Down Pathway combined with Lateral Connections, which enlarges image size or maintains consistent dimensions to retrieve attribute data. This architecture aims to integrate both explicit features and hidden information (Kim et al., 2018; WEB. f).

The hyperparameters used for training both U-Net and FPN models are detailed in Table 4. During training, the dataset was augmented using the augmentation library, with the specific augmentation techniques listed in Table 5.

Table 4. The hyperparameters used for the training of U-Net and FPN

Hyperparameters	**U-Net**	**FPN**
Encoder	Resnet18	Resnet18
Train Batch Size	16, 8	16, 8
Test Batch Size	16,8	16, 8
Learning Rate	0,0001	0,0001
Epoch	100	100
Loss	Dice Loss + Binary Focal Loss	Dice Loss + Binary Focal Loss
Optimizer	Adam	Adam

Table 5. Utilised augmentation types

Augmentation Type	**Value**	**Augmentation Type**	**Value**
Horizontal Flip	0.5	CLAHE	1
Shift Scale Rotate	0,5	Random Brightness	1
Pad If needed	256x256	Random Gamma	1
Random Crop	256x256	Sharpen	1
Gaussian Noise	0,2	Blur	1
Perspective	0,5	Motion Blur	1

3. Results and Discussion

The Tesla P100 16GB GPU, available to users on Kaggle, was utilized for training in this study. The performance of U-Net and FPN architectures was compared across different band combinations using consistent parameters. The test results indicated that RGNIR images yielded better results for both U-Net and FPN models in detecting burned and unburned areas compared to RGB images (see Tables 6 and 7). To enhance model performance and ensure that burned areas are effectively learned, weights trained on the ImageNet dataset (WEB, h) were employed, and the ResNet-18 encoder was selected for this purpose.

Table 6. Test results for RGB image dataset

Model	**Encoder**	**Loss**	**IoU**	**F1 Score**	**Precision**	**Recall**	**Train Batch**	**Test Batch**
U-Net	Resnet-18	0,7889	**0,4573**	**0,6241**	0,7208	**0,5535**	16	16
U-Net	Resnet-18	0,9337	0,3856	0,5502	**0,7218**	0,4627	8	8
FPN	Resnet-18	0,7778	0,4232	0,5918	0,7204	0,5048	16	16
FPN	Resnet-18	0,8486	0,3841	0,5498	0,7246	0,4562	8	8

Table 7. Test results for RGNIR image dataset

Model	**Encoder**	**Loss**	**IoU**	**F1 Score**	**Precision**	**Recall**	**Train Batch**	**Test Batch**
U-Net	Resnet-18	0,3438	**0,7601**	**0,8632**	**0,8829**	0,8455	16	16
U-Net	Resnet-18	0,4114	0,7510	0,8562	0,8418	**0,8757**	8	8
FPN	Resnet-18	0,3811	0,7523	0,8579	0,8465	0,8709	16	16
FPN	Resnet-18	0,3748	0,7510	0,8562	0,8640	0,8522	8	8

Comparing Tables 6 and 7, it is evident that the RGB band combination yields significantly lower accuracies compared to the RGNIR band combination. The Near-Infrared (NIR) band provides distinct spectral properties for burned areas that differ from the healthy green cover, allowing for more accurate segmentation of burned regions. While the segmentation accuracy between the U-Net and FPN models did not show substantial differences, the U-Net model with the RGNIR band combination achieved a higher Intersection over Union (IoU) score of 0.7601. This suggests superior performance by U-Net in accurately delineating the boundaries of burned areas. Furthermore, the U-Net model demonstrated a strong overall accuracy of 90.92% when tested on 138 Sentinel-2 images using the RGNIR band combination.

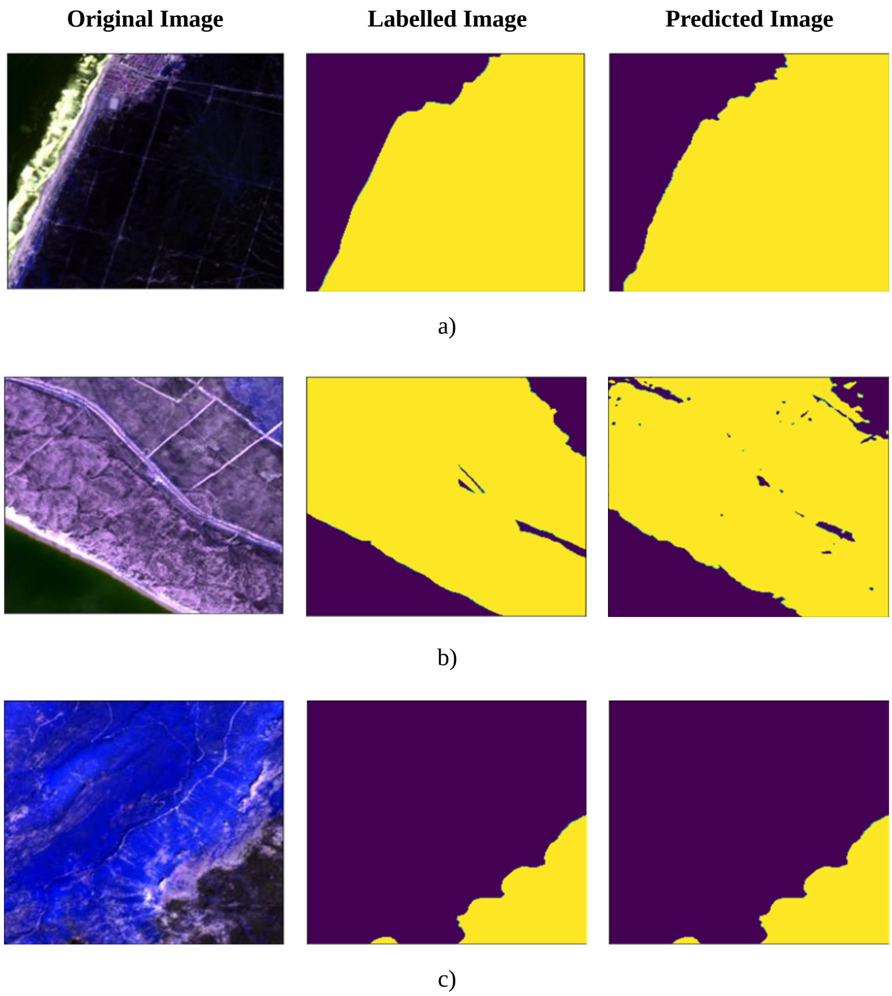

Figure 3. U-Net results for the RGNIR image dataset. The yellow and purple colors represent burned and unburned classes, respectively. a) IoU: 96.85 b) IoU: 95.64 c) IoU: 93.93

Typically, deep learning models utilize all 12 Sentinel-2 bands. However, this study found that the U-Net model provided reliable results with a minimal band combination. Given the limited training data, an IoU score of 0.7601 is considered high, indicating that the model effectively captured the spatial boundaries of burned areas with precision, even with a relatively small dataset. This suggests that the U-Net model achieved a well-generalized solution despite data constraints, successfully learning meaningful spectral and spatial features and accurately detecting burned areas with fewer inputs. Sample results of the U-Net model and its performance with a batch size of 16 are illustrated in Figure 3.

4. Conclusion

The results of the study demonstrate that the U-Net model can achieve reliable outcomes using RGBNIR Sentinel-2 RGNIR bands and a limited amount of training data. This underscores the robustness of deep learning models like U-Net, which are capable of generalizing effectively even with scarce labelled datasets. This capability is particularly valuable for remote sensing applications where labelled data is often limited.

Accurate mapping of burned forest areas is becoming increasingly vital in the context of climate change and its profound effects on global ecosystems. As climate change amplifies the frequency and intensity of wildfires, the timely and precise detection of burned areas is essential for monitoring ecological recovery and evaluating carbon emissions. From a forest management standpoint, the ability to swiftly and accurately map burned regions supports informed decision-making for restoration efforts. Additionally, as forest ecosystems face mounting threats from both climate change and human activities, the insights derived from deep learning-based burned area mapping can inform sustainable management practices, fostering healthier forests and mitigating risks associated with future fires. The integration of advanced remote sensing technologies with climate and forest management strategies will be critical in adapting to and reducing the impacts of climate change while safeguarding ecosystems and communities alike.

To enhance the accuracy of burned area mapping, future plans include developing new datasets that are freely available to the research community can help address data scarcity and improve model performance, producing synthetic datasets can supplement real data, providing additional training examples and improving model robustness and incorporating more spectral bands, such as Short Wave Infrared (SWIR), can offer more detailed information and improve the precision of burned area detection.

These strategies aim to refine the accuracy and effectiveness of burned area mapping techniques.

Acknowledgement

This study was supported by the SILVA NYMPHA (Sustainable Use and Smart Management of Forests) project, a Türkiye-Poland bilateral cooperation initiative, funded by the Turkish National Funding Authority (TUBITAK) under grant agreement

No. 122N254. We would like to express our sincere gratitude to TUBITAK for their financial support, which made this research possible. During the preparation of this work the authors used ChatGPT in order to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Baldrian, P., López-Mondéjar, R., Kohout, P. (2023) Forest microbiome and global change. *Nat Rev Microbiol*, **21**, 487–501

Cambrin, D. R., Colomba, L., Garza, P. (2023) CaBuAr: California Burned Areas dataset for delineation, *IEEE Geoscience and Remote Sensing Magazine*, **11**, 3

Colomba, L., Farasin, A., Monaco, S., Greco, S., Garza, P., Apiletti, D., ... Cerquitelli, T. (2022). A Dataset for Burned Area Delineation and Severity Estimation from Satellite Imagery, CIKM '22: *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pp. 3893 – 3897

Han Y., Zheng C., Liu X., Tian Y., Dong Z. (2024) Burned Area and Burn Severity Mapping With a Transformer-Based Change Detection Model, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **17**, pp. 13866-13880, 2024

Hu, X. Ban, Y. Nascetti, A. (2021). Uni-Temporal Multispectral Imagery for Burned Area Mapping with Deep Learning. *Remote Sens*. 2021, **13**, 1509

Kim, S. W., Kook, H. K., Sun, J. Y., Kang, M. C., Ko, S. J. (2018) Parallel feature pyramid network for object detection. In *Proceedings of the European Conference on computer vision (ECCV)*, pp. 234-250.

Knopp, L. Wieland, M. Rättich, M. Martinis, **S.** (2020**)** A Deep Learning Approach for Burned Area Segmentation with Sentinel-2 Data. *Remote Sens*. 2020, **12**, 2422

Lee, D. Son, S. Bae, J. Park, S. Seo, J. Seo, D. Lee, Y., Kim, J. (2024). Single-Temporal Sentinel-2 for Analyzing Burned Area Detection Methods: A Study of 14 Cases in Republic of Korea Considering Land Cover. *Remote Sens*. 2024, **16**, 884

Maskouni, F.H., Seydi, S.T. (2021). Forest Burned Area Mapping Using Bi-Temporal Sentinel-2 Imagery Based on a Convolutional Neural Network: Case Study in Golestan Forest. *Eng. Proc.* 2021, **10**, 6

Qiao, L. Yuan, W. Tang, L. (2024). DCP-Net: An Efficient Image Segmentation Model for Forest Wildfires. *Forests* 2024, **15**, 947

Ronneberger, O., Fischer, P., Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015:* 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241)

Seydi, S.T., Hasanlou, M., Chanussot, J. (2021). DSMNN-Net: A Deep Siamese Morphological Neural Network Model for Burned Area Mapping Using Multispectral Sentinel-2 and Hyperspectral PRISMA Images. *Remote Sens*. 2021, **13**, 5138

Seydi, S.T., Hasanlou, M., Chanussot, J. (2022). Burnt-Net: Wildfire burned area mapping with single post-fire Sentinel-2 data and deep learning morphological neural network, *Ecological Indicators*, **140**, 2022, 108999

Sui, T., Huang, Q., Wu Mingda, Wu, M., Zhang, Z. (2024), BiAU-Net: Wildfire burnt area mapping using bi-temporal Sentinel-2 imagery and U-Net with attention mechanism, *International Journal of Applied Earth Observation and Geoinformation*, **132**, 2024, 104034

Ribeiro, T.F.R., Silva, F., Moreira, J., Costa, R.L.C. (2023). Burned area semantic segmentation: A novel dataset and evaluation using convolutional networks, *ISPRS Journal of Photogrammetry and Remote Sensing*, **202**, pp. 565-580.

WEB (a). https://www.euronews.com/2023/12/06/copernicus-2023-officially-hottest-year-in-recorded-history (last accessed on 05.09.2024)

WEB (b). https://www.euronews.com/2023/12/06/copernicus-2023-officially-hottest-year-in-recorded-history (last accessed on 05.09.2024).

WEB (c). https://www.undrr.org/media/47703/download (last accessed on 06.09.2024).

WEB (d). https://github.com/qubvel/segmentation_models?tab=readme-ov-file

WEB (e). https://towardsdatascience.com/review-fpn-feature-pyramid-network-object-detection-262fc7482610 (last accessed on 06.09.2024).

WEB (f). https://arxiv.org/pdf/1505.04597 (last accessed on 11.09.2024).

WEB (g). Sentinel-2 mission guide. https://sentinel.esa.int/web/sentinel/missions/ sentinel-2. (last accessed on 12.09.2024)

WEB (h). https://www.image-net.org/ (last accessed on 11.09.2024).

Zhang P., Hu X., Ban Y. (2022), Wildfire-S1S2-Canada: A Large-Scale Sentinel-1/2 Wildfire Burned Area Mapping Dataset Based on the 2017–2019 Wildfires in Canada, *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, Kuala Lumpur, Malaysia, 2022, pp. 7954-7957.

Received September 15, 2024, accepted November 5, 2024