

Burned Area Detection with Sentinel-2A Data: Using Deep Learning Techniques with eXplainable Artificial Intelligence

Elif Ozlem Yilmaz¹, Taskin Kavzoglu¹

¹Dept. of Geomatic Engineering, Gebze Technical University, 41400, Gebze/Kocaeli, Turkey –
eoyilmaz@gtu.edu.tr; kavzoglu@gtu.edu.tr

Keywords: Explainable Artificial Intelligence, Deep Learning, Wildfire, Remote Sensing, SHAP.

Abstract

Annually, a considerable quantity of forest is burned on a global scale. Therefore, it is essential to obtain precise and fast information regarding the size of burned regions in order to effectively monitor the adverse consequences of wildfires. The objective of this investigation is to indicate the effectiveness and usefulness of a deep learning (DL) architecture, such as Convolutional Neural Networks (CNNs), in the mapping of areas affected by fire, employing an eXplainable artificial intelligence (XAI) algorithm known as SHapley Additive exPlanations (SHAP) with accuracy evaluation criteria. Furthermore, this paper presents the evaluation of the Çanakkale-Kızılkeçili village wildfire. The research investigated the impacts of a variety of spectral indices, including the normalized burn index (NBR), differentiated normalized burn index (dNBR), Green-Red Vegetation Index (GRVI), simple ratio vegetation index (RVI), and normalized difference vegetation index (NDVI). At the end of the training process, the model achieved a training accuracy of approximately 0.99, with model loss values converging to approximately 0.1. The findings of the burned area identification analysis indicate that by incorporating spectral indices as supplementary information, the CNN model achieved a high level of accuracy, with an overall accuracy of 98.88% and a Kappa Coefficient of 0.98. Additionally, the SHAP technique was employed to gain insights into the output of the models. The feature importances of the spectral bands were determined through the SHAP analysis of the CNN model. Hence, the significance of the auxiliary data generated by the NBR, dNBR, and NDVI indices was identified as being the highest among the original bands and auxiliary data employed in this investigation.

1. Introduction

Forest ecosystems play a crucial role in the preservation and regulation of ecological equilibrium, offering a wide range of ecological, and economic benefits to humanity (Lu et al., 2014). Wildfires have been recognized as severe destructive occurrences that disrupt these vital activities. They might cause significant harm to forest ecosystems due to their negative impacts on the environment, and the potential threats they pose to vegetation functionality (Bailon-Ruiz and Lacroix, 2020). Every year, they result in the destruction of around one million hectares of forests on a global scale. In recent times, climatic variations and alterations in land use practices have led to the emergence of unmanageable and exceptionally severe wildfires (Benguerai et al., 2019; Singh, 2022). Acquiring trustworthy and accurate information on areas affected by wildfires is of utmost importance in comprehending the entirety of fire incidents and formulating efficacious procedures for disaster management (Guha et al., 2022). The Mediterranean and Aegean areas of Türkiye are highly susceptible to wildfires because of their specific forest types and climatic conditions (Tonbul et al., 2022). The occurrence of wildfires in Türkiye has been attributed to climate change and the consequent global warming trend (Garg et al., 2023; Tonbul et al., 2023). Therefore, there is a pressing need to develop accurate and reliable burned area maps specifically tailored for Türkiye. These maps will serve as valuable tools to aid decision-makers in their efforts to address this issue effectively. In order to address these issues, it is vital to promote collaborative efforts between researchers, policy makers and local communities to improve wildfire preparedness and response.

The application of remote sensing has also extended to the identification and detection of regions affected by wildfires. The improvement of burned area mapping methodologies has been facilitated by the recent progress in analysis tools (Seddouki et al., 2023). The emergence of deep learning (DL) based

methodologies have demonstrated significant potential in many applications pertaining to the examination of remotely sensed data (Ghali and Akhloufi, 2023; Kavzoglu et al., 2023). DL algorithms could autonomously extract valuable characteristics from the input dataset. Lately, there has been a growing interest in the application of DL techniques for the identification of burned areas, which has become a prominent field of research (Tonbul et al., 2023). Furthermore, Artificial Intelligence (AI) can potentially be implemented to develop strong algorithms that exhibit exceptional predictive capabilities across a wide range of challenging domains. However, these systems frequently include an intricate structure, generally referred to as a "black-box" structure, that might hinder the comprehension of the data (Abdollahi and Pradhan, 2023). Nowadays, DL models that utilize explainable AI techniques could enhance model interpretability. These techniques can provide conclusions that contribute to the improvement of the performance of models (Arrieta et al., 2020). As one of the most widely used techniques, The SHapley Additive exPlanations (SHAP) is utilized as a method for analyzing black box models (Kavzoglu et al., 2021; Yilmaz and Kavzoglu, 2022).

This study aims to comprehensively evaluate the effectiveness and utility of using Convolutional Neural Networks (CNNs) in conjunction with spectral indices for accurate mapping of wildfire affected areas within the Çanakkale-Kızılkeçili village location of Türkiye. Utilising Sentinel-2 imagery which are publicly accessible, this research attempts to apply a CNN-based classification approach to analyse both pre- and post-wildfire Sentinel-2 image, complemented by supplementary data including spectral indices such as normalized burn index (NBR), differentiated normalized burn index (dNBR), normalized difference vegetation index (NDVI), Green-Red Vegetation Index (GRVI), and simple ratio vegetation index (RVI). By systematically examining the individual contributions of these spectral indices to the predictive accuracy of burned area delineation, the study aims to elucidate the most effective

indicators for wildfire detection and mapping purposes. In addition, the research embarks on a thorough investigation and validation of eXplainable Artificial Intelligence (XAI) techniques to unravel the predictive mechanisms inherent in the CNN-based model, thereby shedding light on the decision-making processes underlying area mapping. Through this approach, the study aims not only to advance the state of the art in wildfire mapping methodologies, but also to contribute to a deeper understanding of the burned intricate interactions between spectral indices, CNNs and XAI techniques in the context of remote sensing applications.

2. Study Area and Materials

The study was conducted in the village of Kızılkçeçili in the province of Çanakkale, Türkiye (Figure 1). This area has a landscape characterized by a combination of agricultural and forested areas, consisting mainly of wheat fields and fruit tree plantations. The wildfire that occurred in Çanakkale on July 16, 2023, resulted in the burning of an area of more than 2,000 hectares.

Main data source in this research was the Sentinel-2 Level 2A images with atmospheric and radiometric corrections acquired by European Space Agency (ESA) on 14 and 24 July 2023. The imagery comprises 13 bands with varying spatial resolutions (i.e., 4 bands at 10 m, 6 bands at 20 m, and 3 bands at 60 m) and a radiometric resolution of 12-bit. In order to prepare Sentinel-2 for processing, which includes layer stacking and pan-sharpening, the 60 m resolution bands have to be eliminated. As a result, the spectral bands with spatial resolutions of 10 m and 20 m were employed in subsequent analyses. The 20-meter-

resolution spectral bands were downscaled to a resolution of 10 m using the Gram-Schmidt pan sharpening method.

Numerous spectral indices have been developed for the purpose of extracting burned areas through the utilization of remotely sensed data. Five popular spectral of RVI, GRVI, NBR, dNBR, and NDVI (Abdelbaki and Udelhoven, 2022; Kavzoglu et al., 2016; Tonbul et al., 2023; Yin et al., 2022) were utilized in this study. In burned area mapping, the Normalized Burn Ratio (NBR) measure demonstrates a notable level of sensitivity in effectively delineating the magnitude of burned regions and assessing the intensity of fire-induced destruction. The assessment of fire impacts involves the quantification of many factors including the decrease in moisture content, and the loss of vegetation. These impacts may be evaluated by analyzing the variations in reflectivity within the near-infrared (NIR) and shortwave infrared (SWIR) bands, relative to the reflectance of unaffected vegetation before the fire incident. The calculation of dNBR values required subtracting the pre-fire NBR spectral index from the post-fire NBR spectral index. Moreover, the RVI, GRVI, and NDVI indices have frequently been employed as analytical tools for examining alterations in vegetation resulting from the forest fires. The formula of the related indices is given in Table 1.

The study used TensorFlow and Keras to perform deep learning tasks, Pandas and Scikit-learn for data pre-processing, and SHAP for XAI-related applications. These tools allowed for a thorough investigation, ensuring the reliability of the results and advancing knowledge in the field.

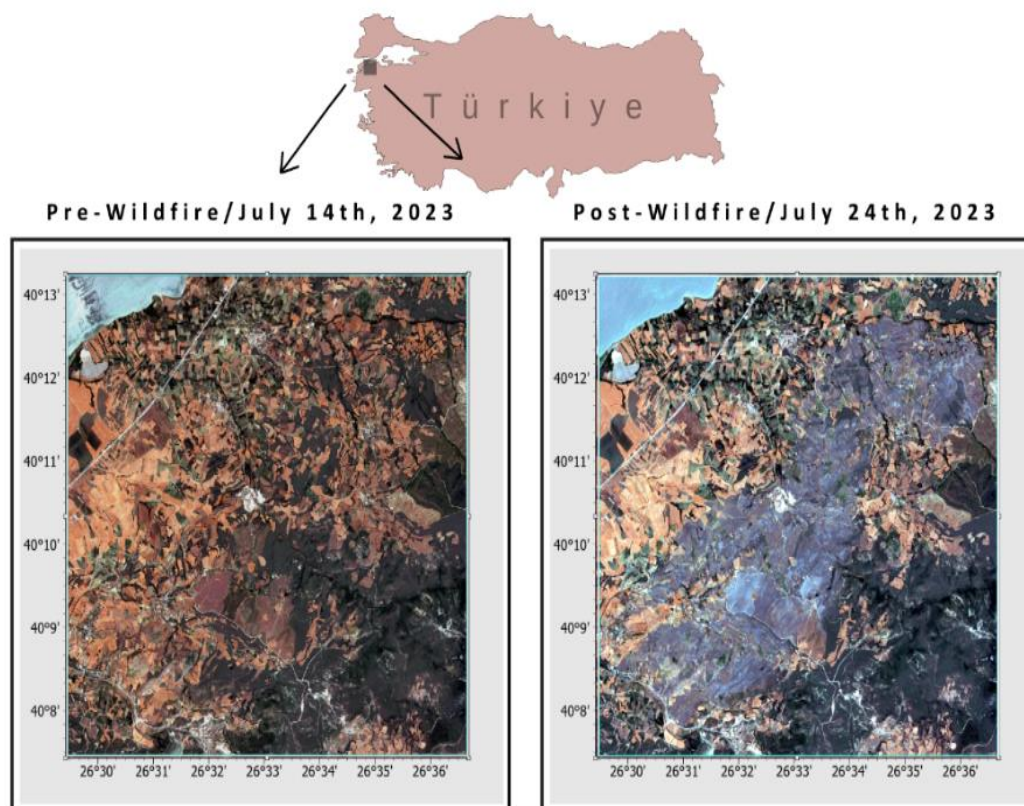


Figure 1. The study area in the village of Kızılkçeçili in the province of Çanakkale, Türkiye.

Spectral Index Name	Abbreviation	Formula
Simple Ratio Vegetation Index	RVI	$\frac{NIR}{Red}$ (1)
Green-Red Vegetation Index	GRVI	$\frac{Green-Red}{Green+Red}$ (2)
Normalized Difference Vegetation Index	NDVI	$\frac{NIR-Red}{NIR+Red}$ (3)
Normalized Burn Index	NBR	$\frac{NIR-SWIR}{NIR+SWIR}$ (4)
Differentiated Normalized Burn Index	dNBR	$NBR_{pre} - NBR_{post}$ (5)

Table 1. Formulation of spectral indexes.

3. Methodology

Before the training stage, a dataset containing 1200 samples was generated to be used in the training process. To ensure the training of the CNN model and mitigate any biases in the training dataset, an equivalent number of samples were selected from the areas determined as affected and unaffected by the wildfire. The severity thresholds from the fire effects monitoring and inventory protocol (FIREMON) report are applied to distinguish burned areas from unburned areas in the collected samples (Tonbul et al., 2022). Samples with a dNBR index above 0.099 were labeled as burned areas, while values below were labeled as unburned areas. Besides, the training dataset was divided such that 70% of the total pixels were allocated for training, while the remaining pixels were split between testing and validation, each accounting for an equal proportion. Other process steps carried out in the study are shown in Figure 2.

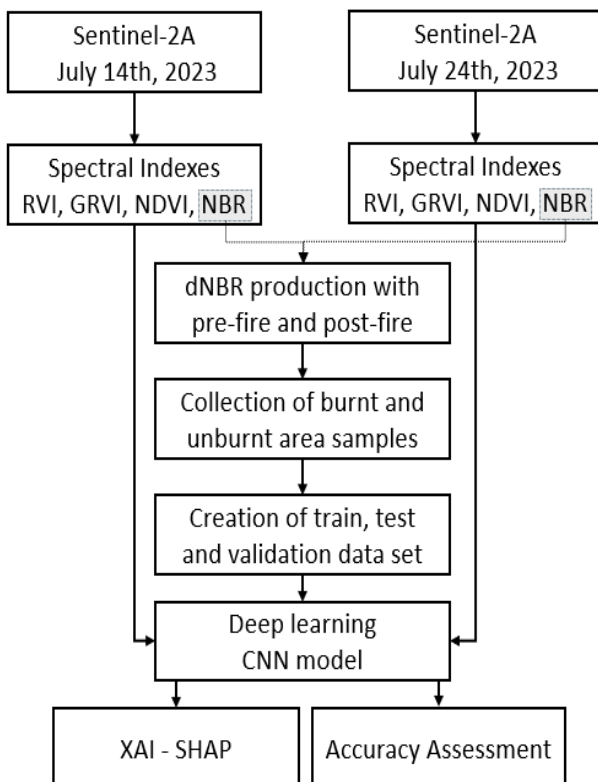


Figure 2. The workflow of the study.

3.1 Deep Learning Models

Convolutional Neural Networks (CNNs) have been widely regarded as the prevailing methodology in the field of DL and

have been extensively employed to address intricate challenges. CNN models have gained significant popularity in the domain of remotely sensed imagery for tasks consisting of object detection, image classification, and object recognition (Kavzoglu and Yilmaz, 2022). This preference can be attributed to their capacity to facilitate completely automated procedures, including extraction of features. The architecture of CNN models comprises several layers, such as convolutional, flattening, dense, pooling, and dropout layers (Yilmaz et al., 2022). The process of extracting feature maps is accomplished using a convolutional layer, whilst the dense layer establishes connections between neurons in the layers. The pooling layer serves to remove extraneous information while retaining the salient components. The flattening layer is responsible for transforming the characteristics of multiple dimensions into a unified vector space. Furthermore, the use of dropout layers in CNN models serves to mitigate the likelihood of overfitting to happen. Alongside this particular architecture, many training parameters, including as optimization methods, learning rate and, activation functions, are employed in CNN algorithms during the process of training.

3.2 Explainable Artificial Intelligence

The intricate nature of DL algorithms renders them challenging to explicate and assess their results. The process of elucidating model outputs aids in the identification of influential characteristics that impact the model, hence enhancing not only the dependability but also the resilience of prediction performance (Xu et al., 2023). Explainable AI (XAI) encompasses several methodologies aimed at elucidating the inner workings of deep learning models. One such technique, rooted in game theory, seeks to provide insights into the performance of these models. To achieve this objective, the researchers commonly employed the SHapley Additive exPlanations (SHAP) approach to provide explanations and interpretations of the outputs generated by the DL model (Moradi et al., 2022). The approach employed in this study involves the utilization of an additive feature importance strategy, wherein the model's output is expressed as the linear summation of the input parameters of the model. The methodology aims to establish an interpretable way for analysis (Kavzoglu et al., 2021).

The deep SHAP methodology is an explainability method suitable for models employing a neural network architecture. The neural network explainability technique is the most efficient in terms of speed. It involves the utilization of a modified version of the deep lift algorithm, which incorporates the SHAP framework. DeepExplainer is an enhanced version of DeepLIFT (Deep Learning Important Features), which is utilized to decompose the output prediction of a neural network on a specific input by backpropagating the contributions of all neurons in the network to each feature of the input (Temenos et al., 2023).

4. Results

In this study, the performance of a CNN classifier was evaluated by its classification accuracy while employing several spectral bands and indices, namely GRVI, NDVI, NBR, RVI, and dNBR, for the creation of wildfire maps. The CNN model contained an estimated 5,000 trainable parameters. It had several layers such as convolutional, flatten, dense, batch normalization, and dropout layers. The objective was to quantify the area affected by fire within the research zone. The proper hyperparameters for the wildfire detection were found using the grid-search method. Particularly, It settled on a batch size of 32 and regulated the number of epochs at 500. The sigmoid activation function, Adagrad optimizer, and Mean Squared Error loss function were all employed in the CNN model training. For mitigating the issues of underfitting and overfitting, the loss and accuracy metrics were computed for the dataset, encompassing both the training and validation samples. These indicators were utilized to construct CNN models that exhibited superior predictive capabilities. The use of the backpropagation technique can lead to a reduction in loss values. This is accomplished by the iterative adjustment of model weights, known as fine-tuning. The CNN models conducted training and the outcomes of the learning curve metrics (Figure 3) where the validation and training accuracies of the model produced anticipated accuracies of over 97%.

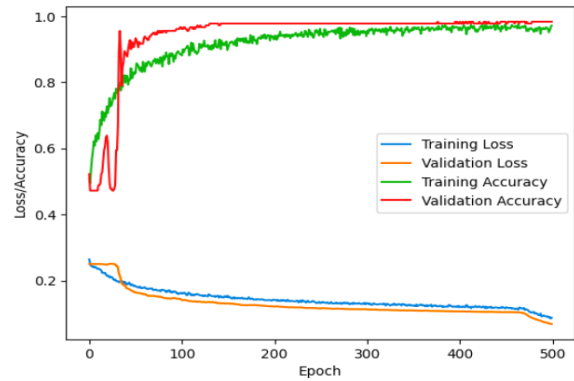


Figure 3. The graph of learning curves by DL model.

To assess the predictive capabilities of the proposed model, six commonly accepted evaluation criteria were employed. These metrics include overall accuracy, Kappa coefficient, area under the receiver-operating curve (AUC ROC), along with precision, recall and F1 Score (Table 2). Based on the predicted results, the CNN model demonstrated an overall classification accuracy of 98.88% and a Kappa coefficient value of 0.98 for Sentinel-2A imagery. Also, the CNN model yielded a recall value of 98.02%, a precision value of 100.00%, and an F1 Score value of 99.00%. In addition, the CNN model exhibited an AUC value of 0.99 for the dataset. To provide a comprehensive visual evaluation of the geographical distribution of the burned regions within the selected study region, a thematic map was produced for the dataset through the utilization of a DL model (Figure 4). Based on the results, the constructed dataset yielded successful results in the process of delineating regions affected by wildfire, although slight discrepancies in certain aspects, which may be resulting from image resolutions and properties. The CNN model demonstrated a high level of accuracy in classifying burned pixels.

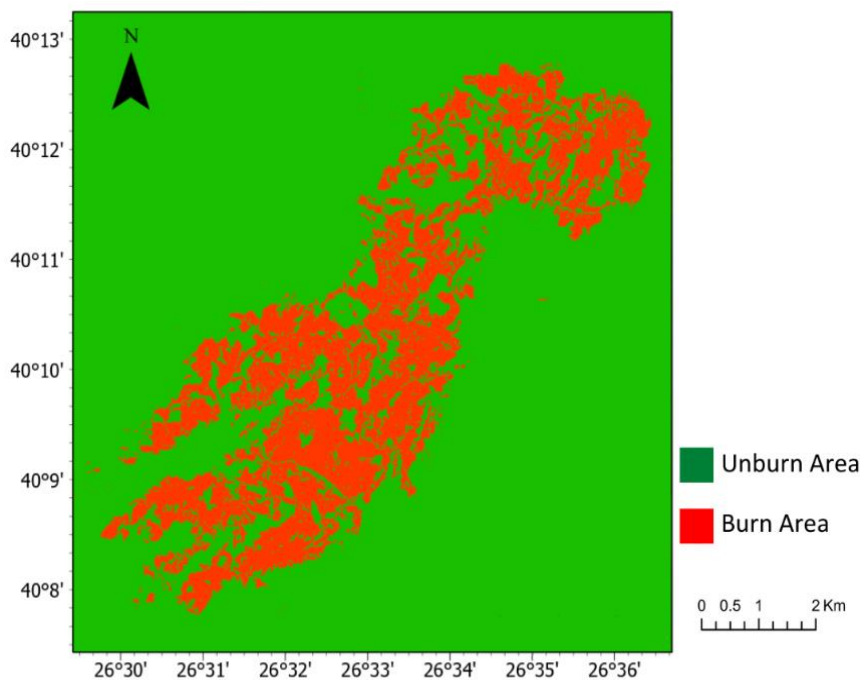


Figure 4. The burned map generated with CNN classifier for the study area.

Accuracy Metrics	
Overall Accuracy (%)	98.88
Precision (%)	100.00
Recall (%)	98.02
F1 Score (%)	99.00
Kappa Coef.	0.98
AUC ROC	0.99

Table 2. Accuracy assessment for burned area detection.

By integrating global interpretations derived from the SHAP deep explanation function, a beeswarm SHAP graph (Figure 5) is produced to evaluate the relative relevance of the spectral bands and indices. The vertical axis on the figure represents the spectral bands and indexes contained within the dataset, while the horizontal axis represents the calculated Shapley values.

The color scheme (red and blue) employed in the figure serves to indicate the extent to which the spectral bands contribute to the identification of burned zones. The effect of the NBR index, derived from the NIR and SWIR spectral bands of the Sentinel-2A image acquired on 24 July, had a more noticeable influence on the model output compared to the other spectral bands and indices. Besides, the SHAP analysis revealed that the dNBR index produced from both the pre- and post-images of Sentinel-2A notably improved the ability to identify fire-affected regions. It can be further assumed that the NDVI index derived from the post-image serves as important auxiliary data that help in the identification of fire-affected areas. The NDVI index has low values because of reduced vegetation density in areas affected by the wildfires. Therefore, its contribution to the DL model has provided a large negative contribution to the detection of wildfire areas. In addition to the effectiveness of the spectral indices in detecting burned areas, the spectral bands of the post-imaging have a significant impact on the accuracy of burned area detection.

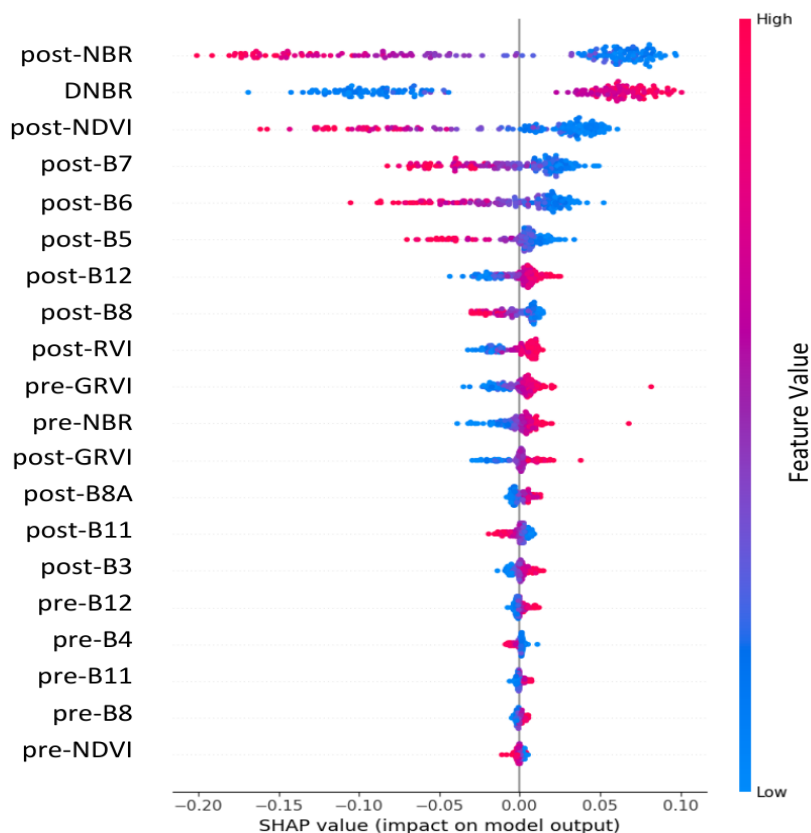


Figure 5. Beeswarm SHAP graph illustrating the feature importance for the DL model. Note that pre and post indicate the bands and indexes estimated for the image before and after the forest fire, respectively.

In addition to the globally evaluated Beeswarm SHAP graph, an additional analysis was performed using a force plot of SHAP at the local scale, obtained by selecting a sample from the dataset (Figure 6). The purpose of this plot is to illustrate how the attributes of a data point contribute to the model prediction. The classification of the selected sample as either burned or unburned is influenced by characteristics including DNBR,

Post-NBR, Post-NDVI and Post-B8A. These investigations provide important insights into the complex dynamics of the model, revealing the precise characteristics that influence its predictions at global and local scales. By clarifying the role of specific features in the model's decisions, these results improve our understanding of the underlying mechanisms involved in the detection and classification of burned regions.

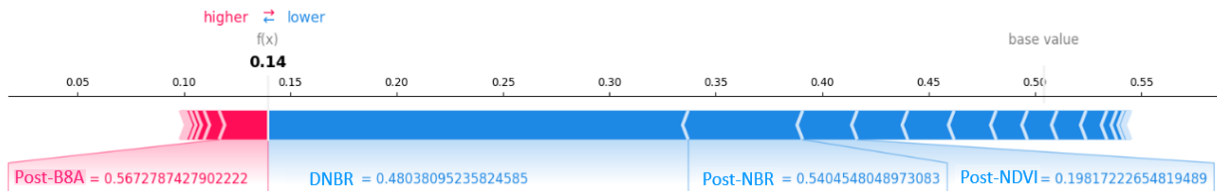


Figure 6. Force plot SHAP plot showing feature importance at local scale for the DL model. Note that pre and post indicate the bands and indexes estimated for the image before and after the forest fire, respectively.

5. Conclusion

Both global warming and poor urban/regional planning have increased the risk of wildfires in Türkiye's forested lands. In the summer of 2023, huge forest areas in Türkiye experienced significant damage due to the occurrence of wildfires in the Mediterranean, Aegean, and Marmara regions of Türkiye. The aim of this research is to present the efficient use of a DL model to map the locations of the wildfires that occurred in July 2023 in Çanakkale province of Türkiye. Moreover, the study intends to evaluate the impact of the indices used as ancillary data on the identification of burned areas using XAI techniques. The findings of this study, carried out with DL and XAI, are poised to provide valuable guidance to administrators and government organizations in defining fire perimeters for controlling post-fire damage. The CNN model could achieve a high level of accuracy, around 98%, in the detection of burned areas. It was also observed that the implementation of spectral indices had a substantial impact or contribution in the differentiation between pixels with unburned and burned in Sentinel-2A imagery. NBR and dNBR indexes estimated from pre- and post-images were the most effective parameters in the model as the XAI algorithm of SHAP analysis revealed. Moreover, spectral bands of the image acquired after the wildfire were more effective in the model compared to the other spectral bands. This study presents critical evidence for the effectiveness and suitability of DL models with XAI approaches in the context of wildfire damage identification and management studies.

References

- Abdelbaki, A., Udelhoven, T., 2022. A review of hybrid approaches for quantitative assessment of crop traits using optical remote sensing: Research trends and future directions. *Remote Sens.*, 14(15), 3515. doi.org/10.3390/rs14153515.
- Abdollahi, A., Pradhan, B., 2023. Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model. *Sci. Total Environ.*, 879, 163004. doi.org/10.1016/j.scitotenv.2023.163004.
- Bailon-Ruiz, R., Lacroix, S., 2020. Wildfire remote sensing with UAVs: A review from the autonomy point of view. 2020 ICUAS, Athens, Greece, 20033821. doi.org/10.1109/ICUAS48674.2020.9213986.
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., 2020. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion.*, 58, 82–115. doi.org/10.1016/j.inffus.2019.12.012.
- Benguerai, A., Benabdeli, K., Harizia, A., 2019. Forest fire risk assessment model using remote sensing and GIS techniques in northwest Algeria. *Acta Silv. et Lignaria Hung.*, 15(1), 9–21.
- Garg, U., Kukreti, V., Pundir, R. S., Manchanda, M., Gupta, N., 2023. Prediction of Türkiye Forest fire using random forest regressor. ICIDCA, Uttarakhand, India, 22963610. doi.org/10.1109/ICIDCA56705.2023.10099565
- Ghali, R., Akhloufi, M. A., 2023. Deep learning approaches for wildland fires remote sensing: Classification, detection, and segmentation. *Remote Sens.*, 15(7), 1821. doi.org/10.3390/rs15071821.
- Guha, S., Jana, R. K., Sanyal, M. K., 2022. Artificial neural network approaches for disaster management: A literature review. *Int. J. Disaster Risk Sci.*, 81, 103276. doi.org/10.1016/j.ijdr.2022.103276.
- Kavzoglu, T., Erdemir, M.Y., Tonbul, H., 2016. Evaluating performances of spectral indices for burned area mapping using object-based image analysis. In: Spatial Accuracy, Bailly, J.-S., Griffith, D., Josselin, D., (Eds.) International Spatial Accuracy Research Association: Montpellier, France, 162–168.
- Kavzoglu, T., Teke, A., Yilmaz, E.O., 2021. Shared blocks-based ensemble deep learning for shallow landslide susceptibility mapping. *Remote Sens.*, 13(23), 4776. doi.org/10.3390/rs13234776.
- Kavzoğlu, T., Yılmaz, E. Ö., 2022. Analysis of patch and sample size effects for 2D-3D CNN models using multiplatform dataset: Hyperspectral image classification of ROSIS and Jilin-1 GP01 imagery. *Turk. J. Electr. Eng. Comput. Sci.*, 30(6), 2124–2144. doi.org/10.55730/1300-0632.3929.
- Kavzoglu, T., Yilmaz E. O., Colkesen, I., Sefercik U.G., Gazioglu, E., 2023. Detection and monitoring of mucilage formations using pixel based Convolutional Neural Networks: The case study of Izmit Gulf, Turkey. In: *Mucilage Problem in The Sea of Marmara*, Albay M., (Eds.), 11, 233–251.
- Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., Moran, E., 2014. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *Int. J. Digit. Earth*, 9(1), 63–105. doi.org/10.1080/17538947.2014.990526.
- Moradi, L., Kalantar, B., Zaryabi, E. H., Halin, A.A., Ueda, N., 2022. On the use of XAI for CNN model interpretation: A remote sensing case study. CSDE, Gold Coast, Australia, 22929265. doi.org/10.1109/CSDE56538.2022.10089337.
- Seddouki, M., Benayad, M., Aamir, Z., Tahiri, M., Maanan, M., Rhinane, H., 2023. Using machine learning coupled with remote sensing for forest fire susceptibility mapping. case study Tetouan Province, northern Morocco. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-4/W6-2022, 333–342. doi.org/10.5194/isprs-archives-XLVIII-4-W6-2022-333-2023.

Singh, S., 2022. Forest fire emissions: A contribution to global climate change. *Front. For. Glob. Change*, 5(925480), 1-11. doi.org/10.3389/ffgc.2022.925480.

Temenos, A., Temenos, N., Kaselimi, M., Doulamis, A. Doulamis, N., 2023. Interpretable deep learning framework for land use and land cover classification in remote sensing using SHAP. *IEEE Geosci. Remote Sens.*, 20, 8500105. doi.org/10.1109/LGRS.2023.3251652.

Tonbul, H., Colkesen, I., Kavzoglu, T., 2022. Pixel- and object-based ensemble learning for forest burn severity using USGS FIREMON and Mediterranean condition dNBRs in Aegean ecosystem (Türkiye). *Adv. Space Res.*, 69(10), 3609–3632. doi.org/10.1016/j.asr.2022.02.051.

Tonbul, H., Yilmaz, E.O., Kavzoglu, T., 2023. Comparative analysis of deep learning and machine learning models for burned area estimation using Sentinel-2 image: A case study in Muğla-Bodrum, Türkiye. *RAST*, İstanbul, Türkiye. doi.org/10.1109/RAST57548.2023.10197926.

Xu, Y., Bai, T., Yu, W., Chang, S., Atkinson, P. M. and Ghamisi, P., 2023. AI security for Geoscience and Remote Sensing: Challenges and future trends. *IEEE Geosci. Remote Sens. Mag.*, 11(2), 60–85. doi.org/10.1109/MGRS.2023.3272825.

Yilmaz, E. O., Teke, A., Kavzoglu, T., 2022. Performance evaluation of depthwise separable CNN and random forest algorithms for landslide susceptibility prediction. *IGARSS, Kuala Lumpur, Malaysia*, 5477-5480. doi.org/10.1109/IGARSS46834.2022.9883954.

Yilmaz, E.O., Kavzoglu, T., 2022 Classification of Jilin-1 GP01 hyperspectral image using machine learning techniques with explainable artificial intelligence. 5th Intercontinental Geoinformation Days, New Delhi, India, 145-148.

Yin, G., Verger, A., Descals, A., Filella, I., Peñuelas, J., 2022. A broadband green-red vegetation index for monitoring gross primary production phenology. *J. Remote Sens.*, 764982. doi.org/10.34133/2022/976498.