ENHANCING THE QUALITY OF CNN-BASED BURNED AREA DETECTION IN SATELLITE IMAGERY THROUGH DATA AUGMENTATION

Vik. Hnatushenko[1,3]*, V. Hnatushenko[2], D. Soldatenko[1], C. Heipke[3]

[1] Dept. Information Technologies and Systems, Ukrainian State University of Science and Technologies, Dnipro, Ukraine - vvitagnat,@gmail.com, zeirison@gmail.com
[2] Dept. Information Technologies and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine - Hnatushenko.V.V@nmu.one
[3] Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Germany - heipke@ipi.uni-hannover.de

KEY WORDS: Forest Fire, Satellite Images, Augmentation, CNN

ABSTRACT:

This study aims to enhance the quality of detecting burned areas in satellite imagery using deep learning by optimizing the training dataset volume through the application of various augmentation methods. The study analyzes the impact of image flipping, rotation, and noise addition on the overall accuracy for different classes of burned areas in a forest: fire, burned, smoke and background. Results demonstrate that while single augmentation techniques such as flipping and rotation alone did not result in significant improvements, a combined approach and the addition of noise resulted in an enhancement of the classification accuracy. Moreover, the study shows that augmenting the dataset through the use of multiple augmentation methods concurrently, resulting in a fivefold increase in input data, also enhanced the recognition accuracy. The study also highlights the need for further research in developing more efficient CNN models and in experimenting with additional augmentation methods to improve the accuracy of burned area detection, which would benefit environmental protection and emergency response services.

1. INTRODUCTION

Currently, a pressing issue is the rapid provision of up-to-date and comprehensive spatial information for making optimal management decisions in the forestry industry
(REPORT on a new EU Forest Strategy for 2030 – Sustainable Forest Management in Europe, 2022). Remote sensing data (RSD) represents a most effective tool for addressing many tasks in this field. The modern stage of RSD development from space is characterized by a rapid increase in the number of operating satellites and the simultaneous increase in spatial and temporal resolution, leading to a large increase in data volume. On the one hand, this development allows for a significant expansion of the range of tasks solved using RSD, including forest monitoring; on the other hand, there is a need to develop new algorithms that ensure automated and fast recognition of objects in satellite images. Satellite images and the results of their automated processing (thematic products) constitute one of the main sources of data for operational mapping, the role of which has been continuously increasing in recent years.

Operational maps are designed to address a wide range of tasks, primarily for object inventory, predicting adverse or dangerous processes, monitoring their development, making recommendations and forecasts, selecting control options, and stabilizing or altering the course of a process in various spheres - from ecological situations to political events.

Forest fires occur in all countries with significant forest resources, but the issue is most acute in countries with a continental climate and a predominance of coniferous species in forests, e.g. in the United States, Canada, and Europe. The number of fires and the area they cover in European countries vary widely and depend on many factors: forest territory development, the value of forest resources in the country's economy, technical equipment, and the level of forest protection, among others (Brushlinsky et al., 2022). However, it should be noted that under relatively stable forest vegetation conditions, the overall flammability of forests is primarily determined by weather conditions. Periods of atmospheric and soil moisture deficiency create conditions in which forest fires can easily develop (Kashtan and Hnatushenko, 2022). In times without rain for more than a month, the drought can become critical, and fires can get out of control. In this situation, they quickly cover large areas and can easily provoke an emergency situation.

Operational mapping of forest cover disturbances, associated with both fires and other natural and anthropogenic factors, is an important area of application for RSD. Previously developed methods for recognizing forest cover disturbances in satellite images, based on calculating various vegetation indices and other spectral channel transformations, no longer meet modern requirements for accuracy, speed and automation levels in monitoring systems. In particular, as the spatial resolution of images increases, the importance of texture and geometric features has grown, which are not considered by these methods. Consequently, there is a need to develop new algorithms adaptable to contemporary satellite data. A promising approach involves the application of convolutional neural networks, which in recent years have been considered the most effective image analysis algorithms.

Albeit, the use of machine learning to process RSD also has limitations. First and foremost, a large training dataset is required to create a good model. Unfortunately, many applied areas of machine learning do not necessarily have access to big data, including to some extent satellite image analysis. In this case, overfitting can easily occur. Overfitting refers to the phenomenon where a network learns a function with high variance to model the training data very well, while subsequently not

* Corresponding author

https://doi.org/10.5194/isprs-archives-XLVIII-1-W2-2023-1749-2023

being able to model the actual test data. One way to avoid overfitting (Vinaykumar et al., 2023) is to increase the amount of training data.

There are several methodological directions for reducing overfitting (Heipke and Rottensteiner, 2020). One strategy focuses on altering the model's architecture itself. For example, dropout regularization, and batch normalization have been developed to extend deep learning for applications with smaller training data volumes. Also, transfer learning and the use of additional data, e.g. existing maps, as well as the use of simulated training data shows promising results.

Another strategy involves data augmentation, which increases the volume of data used for model training by transforming the original data (Shorten and Khoshgoftaar, 2019, Yang et al., 2022, Chlap et al., 2021, Lindner et al., 2019). The most commonly used image transformations are listed below:

- **Flipping** Flip the image horizontally or vertically.
- **Rotation** Rotate the image by an arbitrary angle.
- **Cropping** Cut out a portion from the source image and resize the cropped image to a specified resolution.
- **Shifting** Shift the image left, right, up or down.
- **Color changes** Modify coefficients of intensity, hue, and saturation in color space.
- **Noise** Add random perturbations (noise) to the RGB channels of each pixel. Gaussian noise is typically used (Shijie et al., 2017).

Data augmentation encompasses a broad range of methods that operate at the data level rather than the model architecture level. This can improve training by artificially creating more diverse samples with balanced classes in the training dataset. In other words, the training data should meet two requirements, i.e., adequate diversity and size, both of which can be achieved through data augmentation. A study (Shijie et al., 2017, Perez and Wang, 2017, Tran et al., 2017, Hnatushenko and Zhernovyi, 2020, Hnatushenko et al., 2023) has shown that applying data augmentation methods improved the accuracy of results by $4-8\%$.

The aim of this article is to develop and study an approach for improving the accuracy of fire and burned area recognition in satellite images using convolutional neural networks (CNNs) by optimizing the training dataset selection via data augmentation.

2. MATERIALS AND METHODS

The ability to accurately classify and analyze satellite images is essential for various applications such as disaster management, environmental monitoring, and resource management. In this study, the main focus is on the classification (Soldatenko and Hnatushenko, 2023) of the images into the classes fire, burned, and smoke (besides, we need to deal with the background class). The research uses an image of the region of Tizi Ouzou, Algeria, taken on August 12, 2021, during the forest fire season. The image was acquired from the MultiSpectral Instrument (MSI) onboard the Sentinel-2 satellite (NSSDC identifier: 2017-013A, Satellite number: B). We use the three RGB channels RGB and processing level-2A.

In order to improve classification performance (Kantarcioglu et al., 2023), augmentation methods are employed. These methods increase the amount of data available for training the neural network. The basic dataset, prior to augmentation, consists of images with dimensions of 128×128. After each augmentation the number of such images is increased, depending on on the technique used. For any augmentation, all images of the original dataset were transformed in the same way, and the original images were used along with the transformed ones. The number of images for each class and experiment (see below) is shown in Table (1).

Experiment	Fire		Burned		Smoke	
	train	test	train	test	train	test
1	6	3	12	4	12	6
2	12	6	24	8	24	12
3	12	6	24	8	24	12
4	18	9	36	12	36	18
5	24	12	48	16	48	24
6	12	6	24	8	24	12
7	12	6	24	8	24	12

Table 1. Number of training and test images for each experiment and class.

The three band image (B04, B03, and B02) is shown in Figure (1).Ground truth was collected manually and was subsequently compared to the automatically derived results in order to assess the quality of the proposed method.

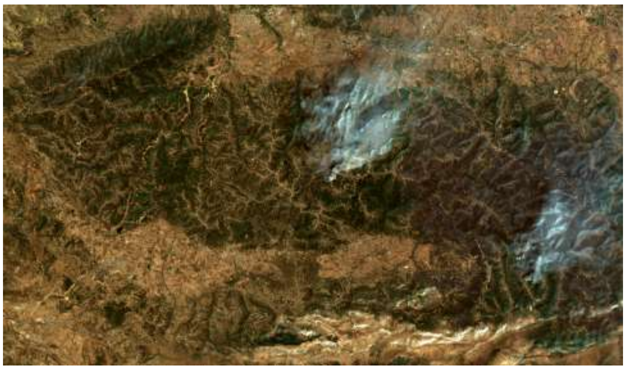

Figure 1. The satellite image used in this study

A Convolutional Neural Network (CNN) (Hnatushenko et al., 2023) is used for classification in this research. The model is designed for segmentation tasks, and it comprises an encoder, a middle block, and a decoder. The architecture utilizes Xception blocks in the encoder and middle block with depthwise separable convolutions, allowing for a more efficient computation.

The encoder consists of several convolutional layers with batch normalization and ReLU activation functions, followed by Xception blocks. The middle block contains one additional Xception block. The decoder is made up of a series of transposed convolutional layers with ReLU activation functions, followed by standard convolutional layers with ReLU activation. The output layer is a single convolutional layer followed by a softmax activation function.

The model is designed in such a way that it connects the encoder and decoder layers by concatenating the output from the encoder layers with the output from the corresponding decoder layers, and then applying the appropriate convolutional layers to the concatenated output. The output is then resized to the desired dimensions using bilinear interpolation.

3. EXPERIMENTS

In this section, we describe the different experiments we have carried out to study the performance of data augmentation.

Experiment 1. The first experiment focused on assessing the performance of the CNN in recognizing burned areas using the original set of training data. The average overall accuracy as well as the Intersection over Union (IoU) scores for the test data were calculated for the classes fire, burned and smoke.

The results are visualized in Figure (2). Despite promising results, the model struggled to recognize certain areas, particularly those under smoke; areas showing fire are too small to be visible. We reached an IoU of 64.4% and an overall accuracy of 64.1%. These results and their visualization constitute a solid starting point for the subsequent experiments, which explore the impact of data augmentation and should help identify areas for further improvement.

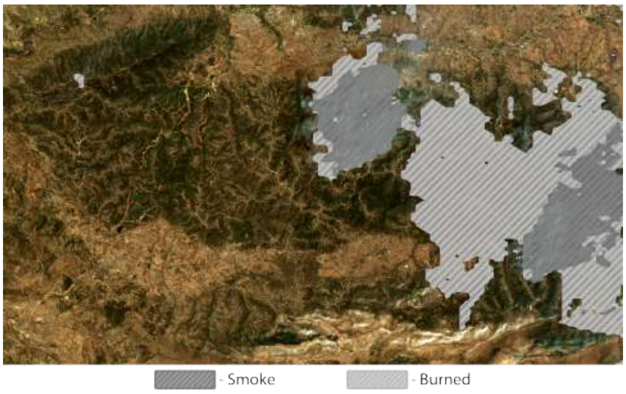

Figure 2. The classification result of Experiment 1

3.1 Flipping

Flipping is one of the simplest data augmentation methods that can be used for images. It involves mirroring the image along the vertical or horizontal axis.

Mathematically, flipping (Hussain et al., 2017) an image can be described as follows: the initial image with dimensions (H, W) is denoted as $I(x, y)$, where x denotes the column and y denotes the row, and the flipped image with dimensions (H, W) is denoted as $I'(x', y')$. $I'(x', y')$ can be obtained using the following formula (1), for horizontal flipping:

$$I'\left(x', y'\right) = I\left(W - x' - 1, y'\right), \tag{1}$$

where x', y' = pixel coordinates in the flipped image
W = width of the original image

One can conclude from this formula that each pixel on the flipped image $I'(x', y')$ can be obtained by flipping the initial image $I(x, y)$ along the central vertical axis. An example of horizontal image flipping is shown in Figure (3).

The formula for vertical flipping (2) is not much different and reads as follows:

$$I'\left(x', y'\right) = I\left(x', H - y' - 1\right), \tag{2}$$

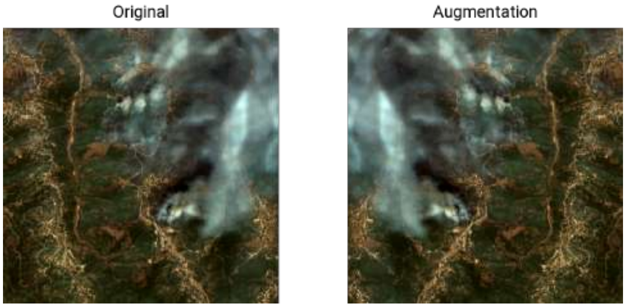

Figure 3. Example of horizontal flipping

where x', y' = new pixel coordinates
x, y = initial pixel coordinates
H = height of the original image

Similar to the case of horizontal flipping, each pixel on the vertically reflected image $I'(x', y')$ can be obtained by flipping the initial image $I(x, y)$ along the horizontal axis. An example of vertical flipping is shown in Figure (4).

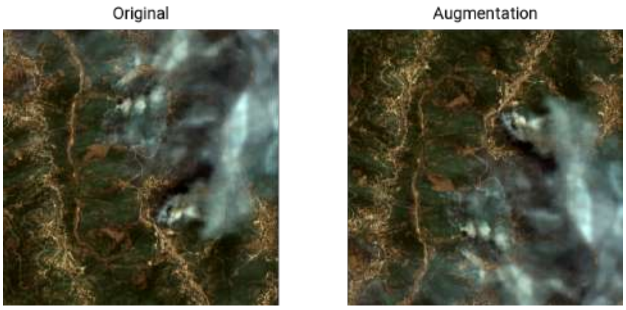

Figure 4. Example of vertical flipping

In the case of images with multiple channels, these formulae need to be applied to each channel separately.

Experiment 2. The second experiment involved the use of horizontal flipping augmentation, applied to the original dataset.

In this experiment, the different classes were recognized reasonably well, as illustrated in Figure (5). Again, pixels of the class fire were too scattered to be visible. The obtained IoU value was 64.8%, while the overall accuracy equaled 71.7%. However, the model identified many false smoky areas, and numerous small regions not belonging to the class burned were detected also.

Comparing these results with those of Experiment 1, the model's ability to recognize burned areas appears to have improved. However, the accuracy of smoky areas and the elimination of small false-alarm areas still require enhancement.

Experiment 3. The third experiment focused on the application of vertical flipping augmentation, applied to the original dataset.

Comparing the results of Experiment 3 with Experiments 1 and 2 in Figure (6), it is evident that the recognition quality has improved, particularly in identifying the large burned area. The resulting Intersection over Union (IoU) was 66.3%, and the

Figure 5. Results of Experiment 2 and comparison with Experiment 1

Figure 6. Results of Experiment 3 and comparison with Experiment 1

overall accuracy equaled 74.1%. However, the model's performance in smoky areas and the elimination of small false-alarms still needs to be improved

Experiment 4. The fourth experiment investigated the effects of applying both vertical and horizontal flipping augmentations separately. Combining these two flipping techniques is aimed to further enhance the diversity of the training data and improve the model's generalization capabilities.

In this experiment, the recognition of the main burned area is significantly improved, and the small false-alarm areas were almost completely eliminated, as illustrated in Figure (7). As before, pixels of the class fire were too scattered to be visible. An IoU of 70.4% and the overall accuracy equaled 90.8%. Although the smoky area recognition was not as accurate as in Experiment 1, it was considerably better than in Experiments 2 and 3.

Figure 7. Results of Experiment 4 and comparison with Experiment 1

Comparing the results of Experiment 4 with the previous experiments, it is clear that the combination of vertical and horizontal flipping augmentations has positively impacted the model's performance. The improvement can be observed in the accurate recognition of the main burned area, the elimination of small false-alarm areas, and the enhanced smoky area recognition compared to Experiments 2 and 3.

Experiment 5. In the fifth experiment, the model was trained using a combination of original images, images with horizontal flipping, images with vertical flipping, and images with both horizontal and vertical flipping applied simultaneously. Note that the latter augmentation is equivalent to a rotation of 180 degrees. This comprehensive approach was aimed at maximizing the benefits of flipping augmentation, providing the model with an even more diverse set of training data.

The recognition results, presented in Figure (8), show that the class burned was well recognized, and the smoky areas were more accurately identified compared to all previous experiments. The experiment resulted in an IoU of 77.5% and the overall accuracy equaled 95.3%. Additionally, the small false-alarm areas were significantly reduced. Notably, this was the only experiment in which the class fire class (see red areas in the figure) was recognized.

Figure 8. Results of Experiment 5 and comparison with Experiment 1

3.2 Rotation

Rotation is also used to increase the size of the training set and improve the model's ability in generalization.

Mathematically, the rotation of an image (Krell and Kim, 2017) can be described as follows: let $I(x, y)$ be an image with pixel coordinates (x, y), where x denotes the column and y denotes the row. Let $I'(x', y')$ be the rotated image with pixel coordinates (x', y'), where x' denotes the column and y' denotes the row .

To compute the rotation of an image by an angle θ in a clockwise direction around the center of the image, the following formulae (3) are used:

$$x' = (x - x_c)\cos\theta - (y - y_c)\sin\theta + x_c, \tag{3a}$$

$$y' = (x - x_c)\sin\theta + (y - y_c)\cos\theta + y_c, \tag{3b}$$

where x_c, y_c = coordinates of the center of the image
x, y = pixel coordinates on the image
θ = rotation angle

To apply the rotation, a random rotation angle θ can be selected from a certain range.

An example of image rotation in Figure (9).

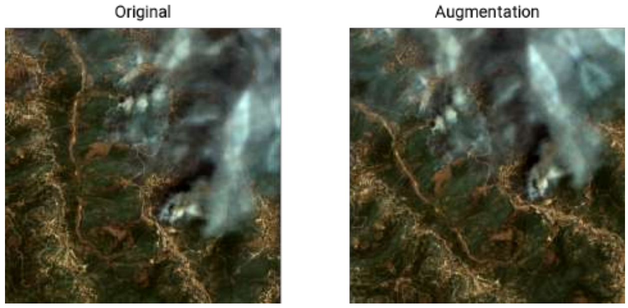

Figure 9. Example of rotating an image by 35 degrees

Experiment 6. In the sixth experiment, the model was trained using a random angle for rotation augmentation selected from the range between 10 and 35 degrees. This approach aimed to explore the impact of introducing different rotation angles on the model's performance in recognizing burned and smoky areas.

The recognition results, illustrated in Figure (10), show a significant number of small falsely recognized regions and parts of the smoky area being incorrectly identified as burned. Actual fire is again not visible. The results show an IoU of 69.4% and the overall accuracy equaled 68.7%.

Figure 10. Results of Experiment 6 and comparison with Experiment 1

These results are quite similar to those of Experiments 2 and 3, suggesting that the rotation augmentation at the specified angles did not provide the desired improvement in the model's recognition capability. Consequently, the model's recognition quality requires further refinement to address these issues.

3.3 Noise

Also, adding noise is a data augmentation method that is being used to improve machine learning models. Gaussian noise is one type of noise that can be added to the data.

The formula (4) for adding Gaussian noise (Huang et al., 2019) to a signal is as follows:

$$x_{\text{noisy}} = x + \epsilon, \tag{4}$$

where x = original signal
ϵ = Gaussian noise with a mean of 0 and standard deviation σ

The choice of the standard deviation σ depends on the problem domain and the level of noise that needs to be added to the data.

The standard deviation of Gaussian noise is determined by the following formula (5):

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2}, \tag{5}$$

where x_i = value of the signal
N = number of values in the signal
μ = mean of the signal

For color images, noise is added to each channel separately, with potentially different standard deviations. An example of adding Gaussian noise to an image can be seen in Figure (11).

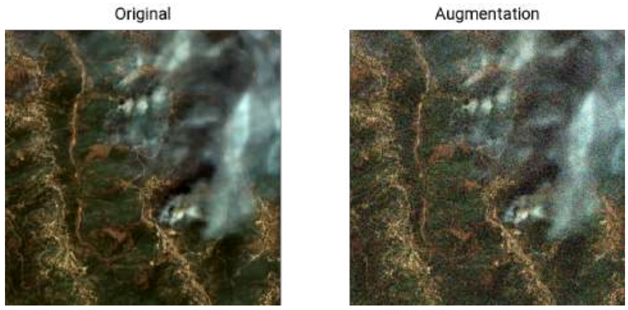

Figure 11. An example of an image distorted with Gaussian noise

Experiment 7. In the seventh experiment, Gaussian noise with a variance randomly selected from the interval 10-50 was added as an augmentation method.

The results, visualized in Figure (12), demonstrate good recognition of the main burned area, akin to the results of Experiment 4. An IoU of 73.7% and the overall accuracy equaled 92.6%. However, there is still an issue with the recognition of the smoky area, and some falsely recognized regions persist, similar to Experiment 5. Also, fire pixels are again not visible.

Figure 12. Results of Experiment 7 and comparison with Experiment 1

3.4 Experiments conclusion

This study aimed to assess the performance of a CNN in recognizing burned areas using various data augmentation techniques. Seven experiments were conducted to explore the impact of different augmentation methods. The overall accuracy and the IoU are given in Table (2).

Exp	Fire		Burned		Smoke	
	IoU	Acc	IoU	Acc	IoU	Acc
1	52.6	52.4	70.7	70.3	69.9	69.6
2	61.1	67.6	64.2	71.0	69.1	76.5
3	59.7	66.7	79.4	88.8	59.8	66.8
4	60.5	79.7	78.5	98.4	72.2	94.3
5	74.5	93.3	75.9	94.6	82.2	98.0
6	56.5	55.9	82.2	81.3	69.5	68.8
7	67.4	85.2	80.9	98.0	72.7	94.6

Table 2. Average IoU (Intersection over Union) and Overall Accuracy in percent for all experiments.

The following findings were obtained:

- **Experiment 1** served as a baseline, using the original dataset. The model exhibited promising performance, but struggled with recognizing areas under the smoke and other environmental factors.
- **Experiment 2** employed horizontal flipping, which improved the model's recognition of burned areas, but led to false alarms for smoky areas and numerous small, falsely recognized regions.
- **Experiment 3** used vertical flipping, resulting in better recognition of large burned areas compared to Experiment 2, but also suffered from false alarms for smoky areas and small, falsely recognized regions.
- **Experiment 4** applied both vertical and horizontal flipping augmentation separately, which reduced the small, falsely recognized areas and improved the recognition of smoky areas compared to Experiments 2 and 3.
- **Experiment 5** incorporated original images and images with horizontal, vertical, and simultaneous horizontal and vertical flipping. The model's performance significantly improved, addressing the issues related to smoky area recognition, the elimination of small, falsely recognized areas, and successfully recognizing the class fire.
- **Experiment 6** used rotation augmentation of 10 and 35 degrees, which did not provide the desired improvement in recognition capabilities, yielding results similar to Experiments 2 and 3.
- **Experiment 7** introduced Gaussian noise with a variance ranging between 10-50 as an augmentation method, yielding good recognition of the main burned area, but with issues in recognizing smoky areas and some falsely recognized regions persisting.

Overall, the experiments demonstrated that data augmentation techniques can improve the performance of the CNN in recognizing burned areas.

4. CONCLUSION

The study examined the impact of commonly used data augmentation techniques (flipping, rotation, and noise addition), on the accuracy of detecting three classes occurring in the context of forest fires: fire, burned and smoke. The results indicate that while single augmentation techniques such as flipping and rotation alone did not result in significant improvements, combining those techniques (which resulted in a larger set of training data also) and the addition of noise resulted in an enhancement of the classification accuracy. Experiment 5, which combined original images and images with horizontal, vertical, and simultaneous horizontal and vertical flipping (i.e. a rotation by 180 degrees), yielded the best results in terms of addressing the challenges related to smoky area recognition and the elimination of small, falsely recognized areas. However, further improvements are still needed, suggesting that alternative augmentation techniques, model adjustments, or training strategies might be necessary to achieve better overall recognition results. Also, it must be kept in mind that only a single Sentinel-2 image and a relatively small dataset were used as input. Therefore, a generalization of the achieved results is not possible at this stage, and more experiments are needed to confirm the current findings.

REFERENCES

Brushlinsky, N., Sokolov, S., Wagner, P., Messerschmidt, B., 2022. World fire statistics, ctif, international association of fire and rescue services. Technical Report 27.

Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L., Haworth, A., 2021. A review of medical image data augmentation techniques for deep learning applications. *Journal of Medical Imaging and Radiation Oncology*, 65(5), 545–563.

Heipke, C., Rottensteiner, F., 2020. Deep learning for geometric and semantic tasks in photogrammetry and remote sensing. *Geo-spatial Information Science*, 23(1), 10–19.

Hnatushenko, V., Hnatushenko, V., Kashtan, V., 2023. Detection of Forest Fire Consequences on Satellite Images using a Neural Network. *43. Wissenschaftlich-Technische Jahrestagung der DGPF*, 31, 29–36. https://www.dgpf.de/src/tagung/jt2023/proceedings/paper/15_dgpf2023_Hnatushenko_et_al.pdf.

Hnatushenko, V., Zhernovyi, V., 2020. Method of improving instance segmentation for very high resolution remote sensing imagery using deep learning. *Data Stream Mining & Processing: Third International Conference, DSMP 2020, Lviv, Ukraine, August 21–25, 2020, Proceedings 3*, Springer, 323–333.

Huang, L., Pan, W., Zhang, Y., Qian, L., Gao, N., Wu, Y., 2019. Data augmentation for deep learning-based radio modulation classification. *IEEE access*, 8, 1498–1506.

Hussain, Z., Gimenez, F., Yi, D., Rubin, D., 2017. Differential data augmentation techniques for medical imaging classification tasks. *AMIA annual symposium proceedings*, 2017, American Medical Informatics Association, 979.

Kantarcioglu, O., Schindler, K., Kocaman, S., 2023. Forest Fire Susceptibility Assessment with Machine Learning Methods in North-East Turkiye. *The International Archives of* *the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-1-2023, 161–167. https://isprs-archives.copernicus.org/articles/XLVIII-M-1-2023/161/2023/.

Kashtan, V., Hnatushenko, V., 2022. Deep learning technology for automatic burned area extraction using satellite high spatial resolution images. *Lecture Notes in Data Engineering, Computational Intelligence, and Decision Making: 2022 International Scientific Conference" Intellectual Systems of Decision-Making and Problems of Computational Intelligence", Proceedings*, Springer, 664–685.

Krell, M. M., Kim, S. K., 2017. Rotational data augmentation for electroencephalographic data. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 471–474.

Lindner, L., Narnhofer, D., Weber, M., Gsaxner, C., Kolodziej, M., Egger, J., 2019. Using synthetic training data for deep learning-based gbm segmentation. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 6724–6729.

Perez, L., Wang, J., 2017. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*.

REPORT on a new EU Forest Strategy for 2030 – Sustainable Forest Management in Europe, 2022. Technical Report A9-0225/2022.

Shijie, J., Ping, W., Peiyi, J., Siping, H., 2017. Research on data augmentation for image classification based on convolution neural networks. *2017 Chinese automation congress (CAC)*, IEEE, 4165–4170.

Shorten, C., Khoshgoftaar, T. M., 2019. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1), 1–48.

Soldatenko, D., Hnatushenko, V., 2023. Investigation of the impact of primary data processing on the results of neural network training for satellite imagery recognition. Springer International Publishing, Cham, 525–546.

Tran, T., Pham, T., Carneiro, G., Palmer, L., Reid, I., 2017. A bayesian data augmentation approach for learning deep models. *Advances in neural information processing systems*, 30.

Vinaykumar, V., Babu, J. A., Frnda, J., 2023. Optimal guidance whale optimization algorithm and hybrid deep learning networks for land use land cover classification. *EURASIP Journal on Advances in Signal Processing*, 2023(1), 13.

Yang, S., Xiao, W., Zhang, M., Guo, S., Zhao, J., Shen, F., 2022. Image data augmentation for deep learning: A survey. *arXiv preprint arXiv:2204.08610*.