

Enhanced Forest Fire Detection Using a Combined SVM and Deep CNN Approach

Bendjillali Ridha Ilyas¹, Bendelhoum Mohammed Sofiane¹, Tadjeddine Ali Abderrazak¹, Kamline Miloud²

¹ LSETER Laboratory, Nour Bachir University Center,
El-Bayadh, Algeria

² LTIT Laboratory, Tahri Mohammed University,
Bechar, Algeria

*Corresponding author; Email: r.bendjillali@cu-elbayadh.dz

Article Info

Article history:

Received , 15/03/2024

Revised , 15/09/2024

Accepted , 20/11/2024

Keywords:

Forest Fire Detection
SVM
PNasNet-5
NFNet-F5
CNN

ABSTRACT

Our proposed approach for forest fire detection presents a significant advancement over existing techniques. By integrating SVM-based classification with state-of-the-art deep CNN architectures specifically PNasNet-5 and NFNet-F5, we achieve outstanding accuracy. Notably, our method attains a remarkable 100% training set accuracy on NFNet-F5. This exceptional accuracy enhances early fire prediction and minimizes false alarm rates, contributing significantly to environmental conservation and human life preservation. Moreover, our adept utilization of fine-tuning techniques effectively addresses challenges related to poor generalization and overfitting, thereby further enhancing the overall efficacy of our innovative approach.

I. Introduction

Forests play a pivotal role in upholding the delicate environmental equilibrium of our planet. Nevertheless, forest fires stand as a formidable threat to these invaluable ecosystems. Regrettably, the detection of most forest fires occurs only after they have already spread extensively, rendering them arduous to manage and extinguish. The repercussions of such fires are dire, inflicting irreparable harm to the environment. Additionally, the massive release of carbon dioxide during these fires exacerbates global warming, compounding the climate crisis.

Forest fires yield severe forest degradation and yield far-reaching economic, social, and ecological ramifications. Forests possess unique attributes, including their isolation, uninhabited nature, and the abundance of dry fuel sources like wood and leaves, creating an ideal environment for rapid fire ignition and propagation. Various factors such as human activities like smoking or barbecues and natural causes like high temperatures or concentrated sunlight due to broken glass can ignite these fires. Once ignited, the fires discover ample fuel to sustain and grow, posing a substantial threat to the surrounding environment.

To counteract forest fires, machine learning techniques have been employed through two primary approaches. The initial approach involves image-based recognition where models undergo training to distinguish between images depicting forest fires and those without. Drones equipped with sensors are deployed to capture images of forested regions. Subsequently, machine learning models analyze these images to identify the presence of fire. This real-time data coupled with precise location coordinates facilitates early intervention and firefighting efforts before the fire escalates.

The second approach hinges on sensor-based detection. Strategically placed sensors dispersed throughout the forest monitor various factors including oxygen and carbon monoxide levels as well as temperature and humidity. Machine learning models analyze this sensor data, promptly alerting forest officials to potential forest fires by identifying anomalies. The overarching objective is to detect fires as early as feasible and provide timely notifications to firefighting units, thereby minimizing ensuing damage.

The central objective of this study is to formulate an early detection system for forest fires, enabling swift intervention and diminishing the extent of devastation. Through the amalgamation of sophisticated machine learning techniques, specifically SVM-based segmentation and deep convolutional neural networks (CNNs) namely PNASNet-5 and NFNet-F5, we endeavor to enhance the precision and efficiency of fire detection within forested regions. Our research aspires to make a substantive contribution toward conserving these critical ecosystems and mitigating the deleterious consequences of forest fires.

The paper is organized in the following manner: In Section 2, we provide a succinct explanation of the suggested technique, founded on two principal pillars, namely the detection of forest fires utilizing Support Vector Machine (SVM) and the extraction of features and classification through two deep CNN architectures PNASNet-5 and NFNet-F5. Section 3 encompasses the experimental results while Section 4 encapsulates the paper's conclusion.

II. Research Method

Forests have a pivotal role in maintaining the delicate balance of our ecosystem. Nevertheless, forest fires pose a substantial threat to this equilibrium. Despite the extensive research in the field of forest fire detection, achieving high detection rates across various databases remains a challenging endeavor. This paper introduces a model designed to confront this challenge through the utilization of transfer learning.

The system under consideration commences by preprocessing the image data and using data augmentation methods such as shearing and flipping in order to produce supplementary images. The images are classified using existing models like PNASNet-5 and NFNet-F5. Figure 1 provides a comprehensive depiction of the suggested paradigm. In order to train the model, the dataset is divided into two parts: 80% for training and 20% for testing. In order to mitigate the prevalent problem of overfitting in deep learning models, data pretreatment and augmentation techniques are used. The model utilizes NFNet-F5 in conjunction with transfer learning whereby images are reduced to dimensions of 331x331 pixels. The process of fine-tuning is crucial in guaranteeing the dependability of the given model. The model under consideration has many layers, each serving distinct purposes. These layers include a convolutional layer, pooling layer, Batch Normalization, ReLU activation layer with Dropout (to address the issue of overfitting), and a softmax layer used for image categorization. The suggested model achieves a remarkable training set accuracy of 98.01% by optimizing the learning rate. Figure 2 classifies image detection into two distinct categories: fire and non-fire. The provided picture is subjected to convolution with several feature detectors, resulting in the generation of a feature map that is then combined to generate a maximum feature pooled map. The NFNet-F5 architecture effectively oversees this process. The NFNet-F5 model is applied to the scaled picture with dimensions of 331x331. The picture is categorized into classes 0 (fire) and 1 (non-fire) using softmax when the probabilities are above the threshold.

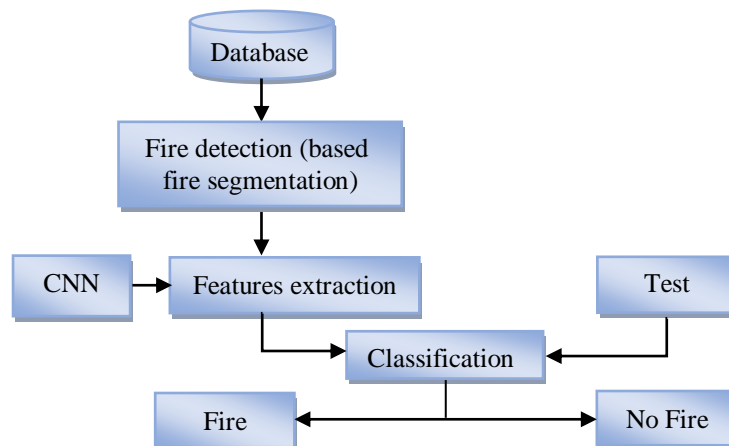


Figure 1. General diagram of a fire detection system.

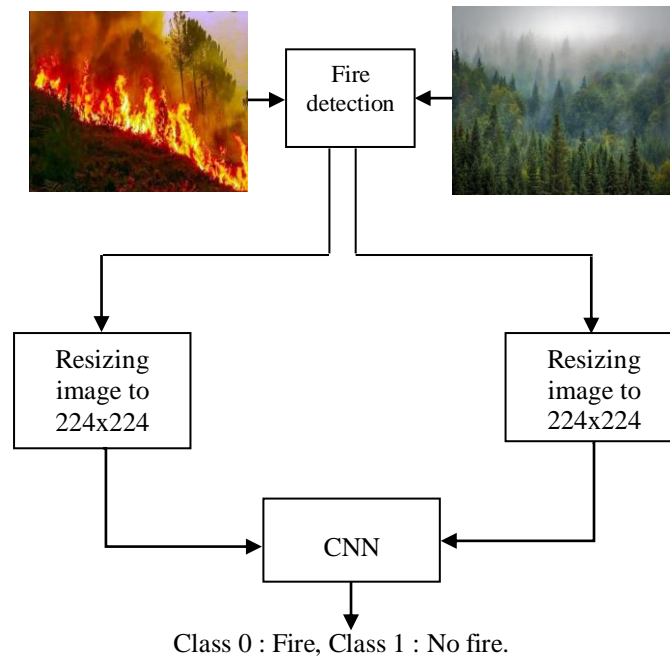


Figure 2. Detecting Images with Class: Fire. No Fire.

II.1. SVM based fire segmentation

Our fire segmentation approach is grounded in supervised segmentation, employing Support Vector Machine (SVM). Our objective is to partition the RGB color space into two distinct subspaces: one containing pixels with fire-like colors and the other encompassing the remainder of the image. The process relies on manually annotated images, which enable us to construct a dataset comprising 12,000 training pixels evenly distributed between fire and non-fire pixels. Each pixel is characterized by its normalized RGB coordinates, and a label of +1 is assigned to fire pixels, while non-fire pixels receive a label of -1. The implementation employs Lib-SVM with a radial basis function as the SVM kernel.

The initial segmentation outcome is provided by the SVM classifier, separating fire pixels from non-fire pixels. However, this initial classification may include false positives, which could be isolated fire-colored pixels from objects resembling fire or misclassified bright regions, such as the sky. To enhance fire detection accuracy, we introduce post-processing operations.

Firstly, we apply morphological opening to the SVM output using a 3x3-pixel square structuring element to eliminate scattered false positives. Subsequently, we tackle the challenge of removing bright false positive pixels while preserving the bright fire cores. Conventional thresholding based solely on pixel intensity would indiscriminately eliminate both the fire cores and other bright non-fire regions, which doesn't align with our objectives. Instead, we leverage the distinct RGB characteristics of fire pixels, which exhibit high red, moderate green, and low blue values.

To identify bright false positive pixels, we introduce a two-pass binary attribute filter with an acceptance criterion defined as:

$$\lambda \times \max_intensity \leq (\max_intensity - \text{mean_intensity}) \quad (1)$$

Here, λ is a threshold given as a percentage of the maximum pixel intensity. Nodes can refer to either image regions or individual image pixels. During the first iteration, we preserve areas that exhibit brightness levels above the predetermined threshold. Subsequently, in the subsequent iteration, we maintain individual pixels that surpass the threshold inside the regions that are preserved. The SVM output is further filtered to reduce the bright false positives by removing the filtered pixels. Based on empirical evidence, setting a threshold of 80% yields favorable outcomes by decreasing the occurrence of false positives while maintaining the accuracy of real positives.

According to Figure 3, the segmentation outcomes achieved at various phases of our suggested technique are shown, using two photos that match to specified burnings. The findings illustrate the efficacy of our approach and underscore the significant enhancement attained by including post-treatments.

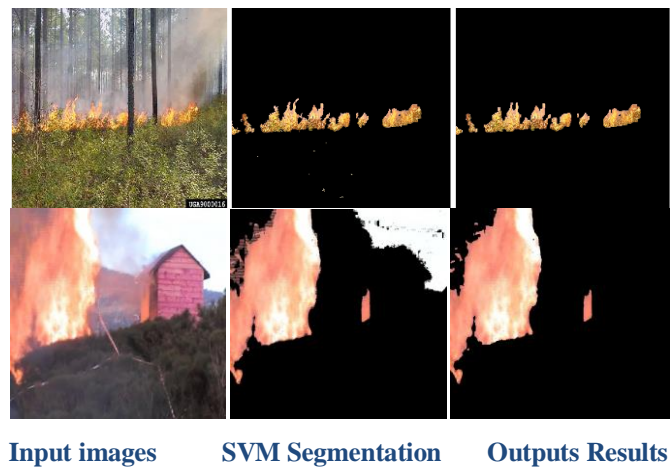


Figure 3. Some results obtained with the proposed method.

II.2. Classification using Deep Convolution Neural Network

1) Normalizer-Free Network (NFNet-F5)

In this study, we implemented a cutting-edge architecture known as the Normalizer-Free Network (NFNet-F5). NFNet-F5 stands out for its ability to achieve state-of-the-art performance without relying on batch normalization, a feature that addresses the limitations of traditional normalization techniques. By eliminating batch normalization, NFNet-F5 reduces computational overhead and improves training stability, particularly for large-scale datasets.

Figure 4 provides a visual representation of the NFNet-F5 architecture, which was specifically employed in this study.

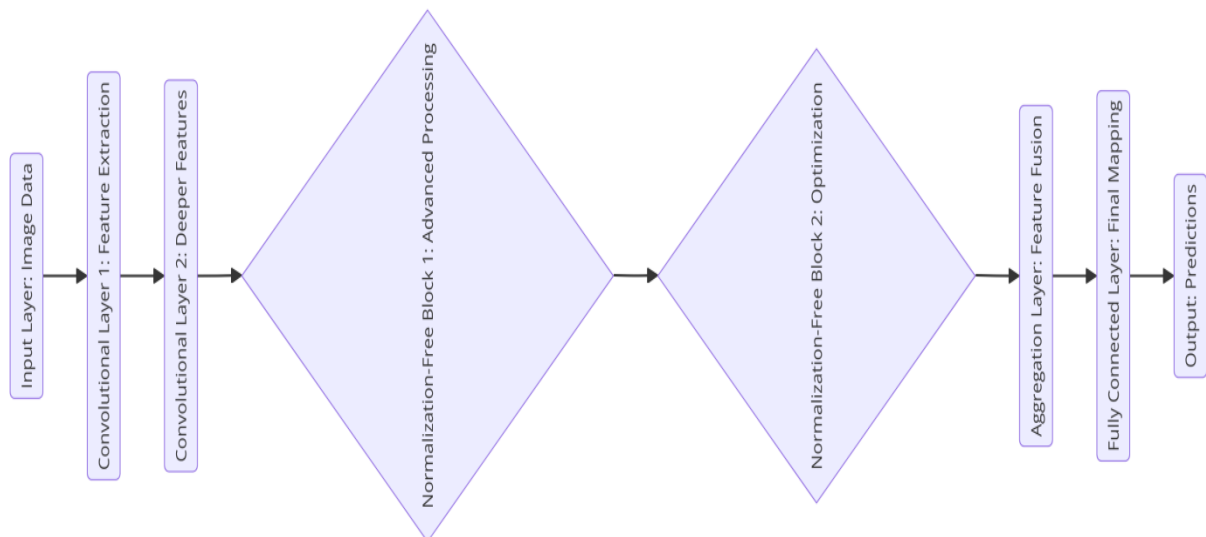


Figure 4. The proposed NFNet-F5 architecture.

The input images, with a resolution of [331 x 331], are processed through a series of convolutional layers designed to effectively extract hierarchical features. To introduce non-linearity, the Scaled Exponential Linear Unit (SELU) activation function is utilized, ensuring robust gradient flow during training. Unlike conventional architectures that depend heavily on normalization layers, NFNet-F5 employs adaptive gradient clipping (AGC), which stabilizes the training process without the need for batch normalization.

This innovative approach not only enhances the network's scalability but also minimizes memory usage and computational requirements, making it a highly efficient solution for complex image processing tasks. By streamlining the architecture and leveraging advanced training techniques, NFNet-F5 achieves superior performance while maintaining computational efficiency.

2) PNasNet-5 Network

In this study, we utilized **PNasNet-5 (Progressive Neural Architecture Search Network)**, an advanced convolutional neural network designed through neural architecture search (NAS). PNasNet-5 represents a significant evolution in deep learning by employing automated optimization to identify the most efficient and high-performing network structures for image classification tasks.

2.1) Key Architectural Features:

2.1.1) Input Size:

PNasNet-5 processes images with an input resolution of [331 x 331], allowing the model to capture finer details compared to standard architectures like ResNet, which commonly use [224 x 224] inputs.

2.1.2) Building Blocks:

- **Normal Cells:** Retain the spatial resolution of the input and extract feature hierarchies.
- **Reduction Cells:** Halve the spatial resolution while increasing feature abstraction, improving computational efficiency.

2.1.3) Optimized Operations:

The network employs **depthwise separable convolutions**, a lightweight alternative to traditional convolutions that reduce the number of parameters and computation time without compromising performance.

2.1.4) Output Layer:

PNasNet-5 concludes with a **fully connected layer of 1000 nodes**, followed by a **softmax activation function** to predict class probabilities.

2.1.5) Advantages Over Traditional Architectures:

- **Efficiency:** PNasNet-5 reduces computational costs by utilizing operations optimized through NAS.
- **Scalability:** Its progressive design enables better performance when scaled to handle larger datasets or higher-resolution images.
- **Performance:** The architecture consistently demonstrates state-of-the-art results on benchmark datasets, outperforming many manually designed networks.

Figure 5 illustrates the architecture of PNasNet-5, showcasing its structured progression from input to classification. This automated and highly efficient approach highlights the potential of NAS in advancing deep learning frameworks for image processing tasks.

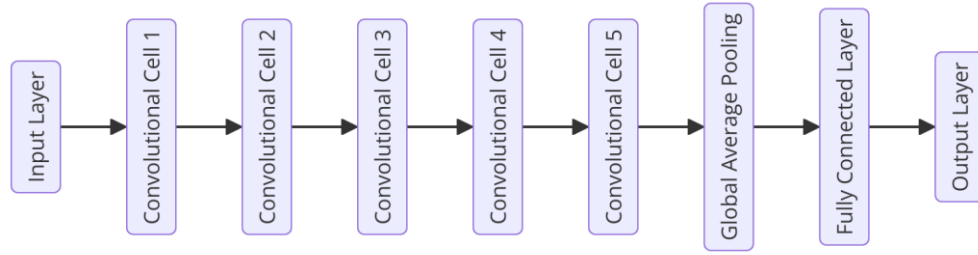


Figure 5. The proposed PNasNet-5 architecture.

III. Results and Discussion

III.1. Dataset and Experimental Setup

In this section, we provide comprehensive details regarding the dataset, evaluation metrics, data splitting, and experimental setup.

- **Dataset Description**

To conduct a comprehensive study, we assembled a diverse image database representing various scenarios encountered in forest fire detection. This dataset consists of 760 meticulously annotated images collected from multiple sources, including Forestrypictures.org [15] and WildlandFire.com [16]. These images depict fires occurring at different times of the day and night in various regions of Northern America and Europe. Additionally, the dataset includes images extracted from prescribed burning experiments conducted in Vigan, France, offering a wide range of fire scenarios.



Figure 6. A preview image of the database.

- **Data Preprocessing**

All images were resized uniformly to dimensions of **331x331 pixels** to ensure compatibility with the PNasNet-5 and NFNNet-F5 architectures. Data augmentation techniques were applied to enhance the dataset's diversity and mitigate overfitting.

- **Dataset Splitting**

The dataset was divided into two subsets: a training set (80% of the images) and a testing set (20% of the images). This division followed an 80:20 ratio, ensuring an adequate volume of data for model training and evaluation.

- **Experimental Setup**

The experiments were conducted on a 64-bit PC equipped with a 3.0 GHz Intel Core i7 processor and 16 GB of RAM. MATLAB R2022b was used as the primary platform for implementation.

III.2. Output Result

Before discussing the results of the NFNNet-F5 architecture, it is important to understand three key evaluation metrics: Accuracy, Precision, and Recall.

Accuracy: The accuracy of a model is determined by evaluating the proportion of properly categorized samples in relation to the total number of samples. The equation represents it:

$$Accuracy = \frac{(True\ Positives + True\ Negatives)}{(True\ Positives + True\ Negatives + False\ Positives + False\ Negatives)} \quad (2)$$

Precision: Precision is a metric that measures the model's capacity to accurately detect positive instances, such as fire images, among all occurrences that are predicted as positive. The calculation is performed using the following equation:

$$Precision = \frac{True\ Positives}{(True\ Positives + False\ Positives)} \tag{3}$$

Recall: Recall, frequently referred to as sensitivity or true positive rate, quantifies the model's capacity to accurately detect all positive cases among the total number of positive instances. The equation provides the specified value:

$$Recall = \frac{True\ Positives}{(True\ Positives + False\ Negatives)} \tag{4}$$

III.3. Results and Evaluation

In this section, we focus on presenting the results of our experiments and discussing their implications.

- **Results of NFNet-F5 Architecture for training**

Table I summarizes the results obtained using the NFNet-F5 architecture. It includes accuracy, precision, and recall metrics over different epochs and iterations.

Table 1. Shows the results of NFNet-F5 architecture.

Epoch	Iteration	Time elapsed (Min)	Accuracy %	Precision %	Recall %
1	8	2.57	80.20	81.14	78.40
5	40	11.45	85.25	86.46	84.83
10	80	20.24	94.17	95.98	92.10
15	120	24.20	96.01	97.67	94.23
20	160	33.44	96.01	97.49	94.65

The results obtained from the model demonstrate promising performance in classifying fire and non-fire images. The accuracy ranges from 80.20% to 96.01%, indicating that the model is reasonably good at distinguishing between fire and non-fire instances. However, it is important to note that accuracy alone may not provide a complete assessment of the model's effectiveness, as it can be influenced by class imbalances within the dataset.

Precision values range from 81.14% to 97.49%, indicating that the model has a relatively high ability to accurately identify true fire instances while minimizing the number of false positives. This suggests that the model exhibits precision in its predictions, providing reliable results for fire detection.

The recall values range from 78.40% to 94.65%, indicating that the model captures a significant proportion of actual fire instances. A higher recall value implies that the model has a lower rate of false negatives, meaning it successfully identifies a large number of true fire instances.

- **Results of PNasNet-5 Architecture for training**

Table II presents the results obtained using the PNasNet-5 architecture, showcasing continuous improvement in performance across iterations and epochs.

Table 2. Shows the results of PNasNet-5 Architecture.

Epoch	Iteration	Time elapsed (Min)	Accuracy %	Precision %	Recall %
1	8	5.19	74.58	68.21	77.85
5	40	14.55	83.14	80.43	85.72
10	80	22.10	87.09	85.91	88.33
15	120	25.85	99.52	98.71	100
20	160	34.20	100	100	100

The results presented in the table indicate a continuous improvement in the model's performance over the epochs and iterations. It is evident that the accuracy of the model significantly increases with the progress of the iterations, reaching 100% in the twentieth iteration. This indicates that the model is capable of correctly classifying the data with 100% accuracy.

In terms of precision, we observe a similar trend where it also increases with the advancement of iterations. It starts at 68.21% in the first iteration and increases to 100% in the twentieth iteration. This indicates the model's ability to correctly identify positive instances with precision and reduce the number of false positives.

Regarding recall, we see that the model also demonstrates excellent performance. It starts at 77.85% in the first iteration and increases to 100% in the twentieth iteration. This indicates the model's ability to detect and classify all positive instances correctly without missing any true positives.

Overall, the table showcases a positive progression of the model's performance over the iterations and training. The evaluation of the results should be done based on the specific variables of each project and the additional metrics required to achieve specific research objectives.

- **Results of NFNet-F5 and PNasNet-5 Architecture for testing**

The table III above summarizes the key performance metrics for NFNet-F5 and PNasNet-5 during the testing phase. Here is a brief discussion of the findings:

Table 3. Results of NFNet-F5 and PNasNet-5 Architecture for testing.

Metric	NFNet-F5 Testing	PNasNet-5 Testing
Accuracy (%)	95.53	97.54
Precision (%)	95.14	97.20
Recall (%)	94.23	95.25

Accuracy: PNasNet-5 outperforms NFNet-F5 in terms of accuracy, achieving an accuracy rate of 97.54% on the testing dataset. NFNet-F5, while still reasonably accurate at 95.53%, falls short in comparison.

Precision: PNasNet-5 demonstrates higher precision during testing, accurately identifying true fire instances while minimizing false positives. NFNet-F5 exhibits lower precision at 95.14% in comparison to PNasNet-5: 97.20%.

Recall: PNasNet-5 excels in recall, capturing a higher proportion of actual fire instances without missing true positives, achieving a recall rate of 95.25%. NFNet-F5 has a lower recall rate of 94.23%.

PNasNet-5 consistently outperforms NFNet-F5 in terms of accuracy, precision, and recall during the testing phase, making it a superior choice for forest fire detection. These results highlight PNasNet-5 reliability and effectiveness in accurately identifying and classifying fire instances in images. However, it's important to consider specific deployment requirements and computational resources when choosing a model for real-world applications.

III. Conclusion

In conclusion, our study presents a transfer learning-based approach combined with SVM-based fire segmentation for early detection of forest fires. The proposed model demonstrates effective classification of images into fire and non-fire classes, with the assistance of NFNet-F5 and PNasNet-5 models. Notably, PNasNet-5 achieves outstanding accuracy, accurately predicting the type of image fed into the model with a remarkable train and test accuracy of 97.54%. Fine-tuning techniques have been employed to mitigate overfitting issues, enhancing the model's reliability.

The developed model holds great potential for real-time application in preventing and mitigating the impact of forest fires on the ecosystem. By enabling early detection and classification of fire incidents, it offers a valuable tool for prompt response and intervention. This can aid in minimizing the devastating consequences of forest fires, preserving biodiversity, and safeguarding the delicate balance of the ecosystem.

Further research and development in this field can focus on refining the model's performance by exploring different network architectures, optimizing hyperparameters, and expanding the dataset. Additionally, incorporating real-time data feeds and integrating the model into an automated fire detection system could enhance its practicality and impact in real-world scenarios.

References

- [1] D. Kolaric, K. Skala, and A. Dubravic, "Integrated system for forest fire early detection and management," *Periodicum Biologorum*, vol.110, no. 2, pp. 205–211, 2008.
- [2] J. R. Martinez-de Dios, B. C. Arrue, A. Ollero, L. Merino, and F. Gomez-Rodriguez, "Computer vision techniques for forest fire perception," *Image and Vision Computing*, vol. 26, no. 4, pp. 550–562, 2008.
- [3] Faming Gong, Chuantao Li, Wenjuan Gong, Xin Li, Xiangbing Yuan, Yuhui Ma and Tao Song, "A real-time fire detection method from video with Multifeature Fusion," *Computational Intelligence and Neuroscience*, Hindawi, 2019.
- [4] Zhentian Jiao, Youmin Zhang, Jing Xin, Lingxia Mu, Yingmin Yi, Han Liu and Ding Liu, "A Deep Learning based forest fire detection approach using UAV and YOLOv3," 1st International Conference on Industrial Artificial Intelligence (IAI), 2019.
- [5] B. R. Ilyas, B. Mohammed, M. Khaled and K. Miloud, "Enhanced Face Recognition System Based on Deep CNN," *2019 6th International Conference on Image and Signal Processing and their Applications (ISPA)*, Mostaganem, Algeria, 2019, pp. 1-6, doi: 10.1109/ISPA48434.2019.8966797.
- [6] B. R. Ilyas, T. A. Abderrazak, B. M. Sofiane, B. Bahidja, H. Imane and K. Miloud, "A Robust-Facial Expressions Recognition System using Deep Learning Architectures," *2023 International Conference on Decision Aid Sciences and Applications (DASA)*, Annaba, Algeria, 2023, pp. 541-546, doi: 10.1109/DASA59624.2023.10286798.
- [7] B. R. Ilyas, B. Mohammed, M. Khaled, A. T. Ahmed and A. Ihsen, "Facial Expression Recognition Based on DWT Feature for Deep CNN," *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, Paris, France, 2019, pp. 344-348, doi: 10.1109/CoDIT.2019.8820410.
- [8] R. I. Bendjillali, M. S. Bendelhoum, A. A. Tadjeddine, and M. Kamline, "Deep Learning-Powered Beamforming for 5G Massive MIMO Systems", *JTIT*, vol. 4, no. 4, pp. 38–45, Oct. 2023.
- [9] Bendjillali, R.I.; Beladgham, M.; Merit, K.; Taleb-Ahmed, A. Improved Facial Expression Recognition Based on DWT Feature for Deep CNN. *Electronics* 2019, 8, 324.
- [10] BENDJILLALI Ridha Ilyas, BELADGHAM Mohammed, MERIT Khaled, Abdelmalik Taleb Ahmed. «Illumination-Robust Face Recognition Based on Deep Convolutional Neural Networks Architectures » *IJECS*, vol. 18, no. 2, 2020.
- [11] K. Miloud, A. M. Lakhdar and B. R. Ilyas, "Arabic Handwriting Recognition System Based on Genetic Algorithm and Deep CNN Architectures," *2021 International Conference on Decision Aid Sciences and Application (DASA)*, Sakheer, Bahrain, 2021, pp. 583-586, doi: 10.1109/DASA53625.2021.9682380.
- [12] M. Rahul, K. Shiva Saketh, A. Sanjeet and N. Srinivas Naik, "Early Detection of Forest Fire using Deep Learning," *2020 IEEE REGION 10 CONFERENCE (TENCON)*, Osaka, Japan, 2020, pp. 1136-1140, doi: 10.1109/TENCON50793.2020.9293722.
- [13] S. Liu And W. Deng, "Very Deep Convolutional Neural Network-Based Image Classification Using Small Training Sample Size," In *2015 3rd Iapra Asian Conference On Pattern Recognition (Acpr)*. Ieee, 2015, Pp. 730–734.
- [14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, And Jian Sun. Deep Residual Learning For Image Recognition. In *Proceedings Of The Ieee Conference On Computer Vision And Pattern Recognition*, Pages 770–778, 2016.
- [15] Forestry Images, "Forest Health, Natural Resources, Fire, Trees, Wildlife, Silviculture Photos," [Http://Www.Forestryimages.Org/](http://www.forestryimages.org/), 2010.
- [16] Wildland Fire, "Home Of The Wildland Firefighter," [Http://Www.Wildlandfire.Com/](http://www.wildlandfire.com/), 2010.