A NEW LIGHT ENSEMBLE DEEP-LEARNING FRAMEWORK TO DETECT FIRE

Ahmed Alsheikhy (1), Tawfeeq Shawly (2), Hossam E. Ahmed (1,3), Hassan Fareed M. Lahza (4)*

(1) *Department of Electrical Engineering, College of Engineering, Northern Border University, Arar 91431*; (2) *Department of Electrical Engineering, Faculty of Engineering at Rabigh, King Abdulaziz University, Jeddah 21589;* (3) *Electrical Engineering Department, Benha Faculty of Engineering, Benha University, Benha;* (4) *Department of Information Systems, College of Computers and Information Systems, Umm Al-Qura University, Makkah 21955,*
(1,2,4)*Saudi Arabia,* (3)*Egypt*

* Corresponding Author, e-mail: aalsheikhy@nbu.edu.sa

Abstract: Fires can cause devastating damage to lands, properties, and humans. Many countries suffer from huge financial losses due to these fires. Therefore, there is a need to implement a practical solution to spot fires effectively and accurately. Deep-learning algorithms and artificial intelligence have been deployed recently in various fields, such as monitoring systems, economics, and detection. This paper proposes a New Light Ensemble Deep-Learning Framework (NLEDLF). This framework consists of two deep-learning technologies, which are a New Generative Adversarial Network (NGAN) and a New Convolutional Neural Network (NCNN). These two tools are incorporated into the framework along with some image preprocessing methods to detect fires using pixels. The proposed framework achieves a reasonable.

Keywords: NGAN, NCNN, NLEDLF, fires, detection, light ensemble, framework.

1. INTRODUCTION

Generative Adversarial Networks (GANs) are an unsupervised learning approach [1, 2]. This approach is considered a powerful tool when dealing with distributions of complex data [1]. Recently, attention has been focused on this approach due to its promising results. GANs were developed in 2014 by Goodfellow and his team. GANs have proved their effectiveness in generating images [1, 3]. However, these networks require a huge number of parameters. Therefore, a huge execution time is possible. Hence, implementing light GANs is highly recommended. GANs consist of two elements, which are a Generator (G) and a Discriminator (D) [4, 5].

The generator produces fake data and feeds the data into the discriminator. The discriminator detects the true data from the false ones. For simplicity, the generator tries to deceive the discriminator [4]. These two components are opponents to each other. The GAN models are widely used due to their ability to generate accurate outputs.

Deep-Learning Technologies (DLTs) can learn effectively to extract data characteristics in spatial and spectral dimensions [6]. This process requires no knowledge of the statistical data of inputs. As in [6-8], concerned teams have figured out how to extract multi-level features from Convolutional Neural Networks (CNNs) properly. However, these CNNs face difficulty in extracting spatial relationships between characteristics of inputs [7]. In addition, typical CNNs involve a large number of parameters [9, 10]. Consequently, the processing time is expected to be higher than usual. Therefore, it is required to have a dependable model that requires a smaller number of parameters and produces reasonable results in less time.

CNNs are utilized mostly in Computer Vision (CV) applications because of their high learning capabilities. In general, any CNN architecture includes multiple layers of convolutional, pooling, activations, and fully connected. Each layer is deployed to perform required tasks, such as extracting features and expanding the learning capacity.

1.1 Research Problem

Some implemented methods in [1, 2, 5], and [7] achieved accuracy between 67% and 96%. However, these results were associated with many parameters and the processing time. Thus, there is a need to develop a model that achieves better results with less time and a smaller number of parameters compared to the existing models. In this article, a light ensemble framework of two deep-learning techniques is proposed. This framework contains new light GANs and new light CNNs.

1.2 Motivations and Contributions

The government of the Kingdom of Saudi Arabia launched a promising vision in 2016, which is Vision 2030. This vision is composed of twelve initiatives and their main targets are people, the environment, safety, and the economy. The twelve initiatives depend on digital solutions and the government has transformed most of their services into the digital world. These solutions shall be effective, productive, and reliable. This study aims to provide a novel solution based on artificial intelligence to be used in CV applications. In addition, increasing accuracy and saving lives and properties are considered other targets.

In this research, coming up with a new ensemble solution is intended. This solution shall be light regarding the number of parameters, operations, and processing time. This proposed model is composed of GANs and CNNs. We will call this model the New Light Ensemble Deep-Learning Framework for Fire Detection (NLEDLF). The main contributions are implementing the proposed algorithm NLEDLF, integrating two new deep-learning tools, conducting analysis to evaluate the performance of the presented model, and performing simulation experiments on three fire datasets to detect fires.

This article is organized as follows: Section 2 contains related work and Section 3 offers a complete detail of the proposed approach. The results and discussion are in Section 4, while the conclusion is in Section 5.

2. RELATED WORK

A parallel decoding approach based on GANs to paint images is developed in [1]. This approach contained encoding and parallel decoding networks. The authors deployed a method to distribute weights using a convolutional layer called a rate-adaptive dilated layer. This layer extracts the required features. Two datasets were used to verify the developed method and achieved nearly 91% accuracy. In contrast, the proposed light ensemble framework achieved 96.8% accuracy on three datasets when applied to detect fires. The proposed algorithm includes two new light deep-learning networks based on GANs and CNNs. This model outperforms the implemented approach in [1]. The research in [2] presents a method to improve image resolution using a new bi-cubic interpolation model and incorporated GANs to extract spectral and spatial features after fusing original inputs. The authors reached an average of 92% for the f-score on three utilized datasets, while the proposed framework attained 96.8% for the same performance metric on the same number of datasets. This result shows that the proposed framework provides better outcomes.

In [5] is implemented a semi-supervised GAN classifier based on the Sine-Cosine algorithm to optimize hyperparameters for classification purposes. In addition, the authors employed the Synthetic Minority Oversampling Technique (SMOTE) to resolve the imbalance issue. Six datasets were utilized to evaluate the model and achieved a maximum accuracy of nearly 96%, while the presented framework achieved an average 96.8% accuracy based on the ensemble of two deep-learning tools. A method based on a deep-learning technology for image recognition is presented in [7]. The authors applied a CNN to extract edges, contours, and local characteristics using filters. The authors reached 94% accuracy, whereas the proposed framework achieved promising results with 96.8% accuracy using the ensemble method of two different deep-learning topologies.

3. THE PROPOSED FRAMEWORK

3.1. Problem statement

Computer Vision (CV) has been deployed in various applications, such as pattern recognition, object detection, and disease classification. In addition, CV is applied in numerous solutions to save the environment from natural disasters like fires. Recently, various methods have been developed using GANs or CNNs to serve different purposes. Some methods have taken more processing time and included a vast number of parameters and variables. These parameters may result in taking an additional memory space. Thus, this article proposes a new light ensemble deep-learning framework (NLEDLF). This framework is composed of new GAN and new CNN architectures. These two topologies are integrated together to form the proposed light ensemble framework.

3.2. Datasets

The annotation of any data set plays a significant part in deciding the effectiveness of any suggested methodology. Thus, the quality of images affects the model's performance. To evaluate the presented framework, three datasets are utilized. These datasets include fire and non-fire images. Hence, the proposed model is tested on fire detection. The utilized three datasets are collected from different sources and are publicly available from [11-13]. These datasets include 33,689 images of fire and non-fire with a size of 6.1GB. These images are grouped into three sets, which are training, validation, and testing. The first set contains 70%, while the validation has 10%, and the rest is reserved for testing purposes. Table 1 provides details about the utilized datasets in the three sets.

Table 1. Details of the applied three sets

Number of images					
Training		Validation		Testing	
Fire	Non-fire	Fire	Non-fire	Fire	Non-fire
15,254	8,328	681	329	5,935	3162

3.3. The Proposed Model

This subsection offers an intensive detail of the proposed model, which is NLEDLF. Figure 1 illustrates a general block diagram of the proposed framework. The block diagram is composed of four components. The first component is the preprocessing part, where any image is processed to remove any noise and rescale it to a common size which is 640 x 640 pixels. Then the processed image or a set of images is/are normalized. The second component is the new GAN, and the third component is the new CNN. Lastly, both components, GAN and CNN, are integrated and concatenated together and fed into a set of layers as shown in Figure 1. In this figure, the (+) operator denotes the concatenating process.

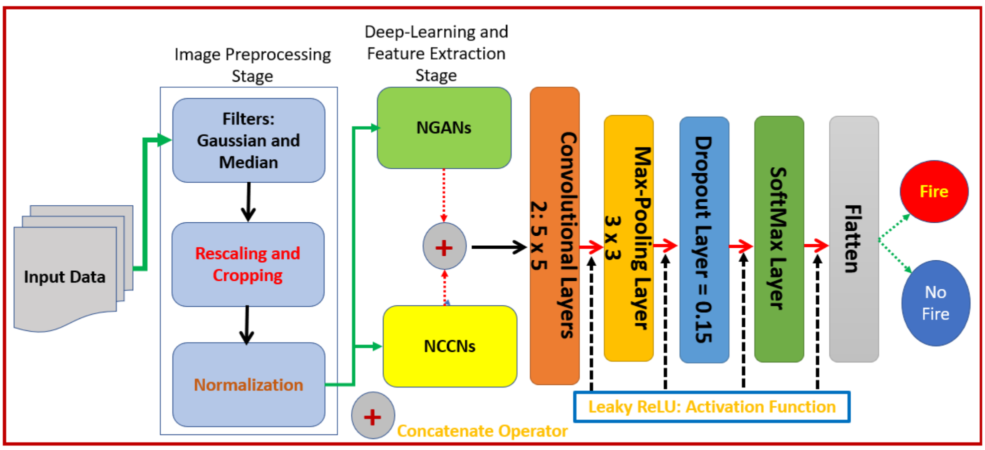

Figure 1. The flowchart of the proposed model

These layers are two convolutional layers, one Max-Pooling layer, a Dropout layer, a SoftMax layer, and a flattened layer. The convolutional layers are used to enhance the model's capabilities to extract spatial and spectral characteristics and to simplify the design as well. The Max-Pooling layer is utilized to reduce the size of a resultant feature map into 3 x 3. The Soft-Max layer is used to generate the outputs and it contains 3 fully connected layers. The first two fully connected layers have 100 nodes, while the last one contains just two nodes. Table 2 provides information about the hyperparameters of these layers. ADAM optimizer is applied in this framework. In addition, the utilized learning rate is 0.1, and 32 and 16 are the sizes of the applied batches in the proposed method.

Table 2. The applied configurations

The name of the layer and defined parameters	*Values*
Learning rate (L)	*0.1*
Batch size (bs)	*32*
Optimizer	*Adam*
Convolutional	*256 filters, 5 x 5 of kernel size, padding = same, 4 number of strides, and 200 neurons*
Max-Pooling	*3 x 3 and 3 number of strides.*
Dropout	*0.15*
Number of epochs (Noe)	*35*
Number of iterations (noi)	*4000*
Activation function	*Leaky ReLU*

Figures 2 and 3 depict internal architectures of the developed NGAN and NCCN models and Table 3 shows the hyperparameter configurations inside the NGAN model. Figure 3 displays a structure of the developed NCNN model, while its hyperparameters are listed in Table 4. NGAN and NCNN models provide the deep-learning processes to distinguish between pixels with fires or any components of fires, which are smoke, flames, and pixels that contain no fires or their elements. In Figures 2 and 3, the activation function is deployed to all convolutional and pooling layers. In addition, every batch is normalized in all convolutional layers. In Figure 3, the fully connected layers include 50 nodes.

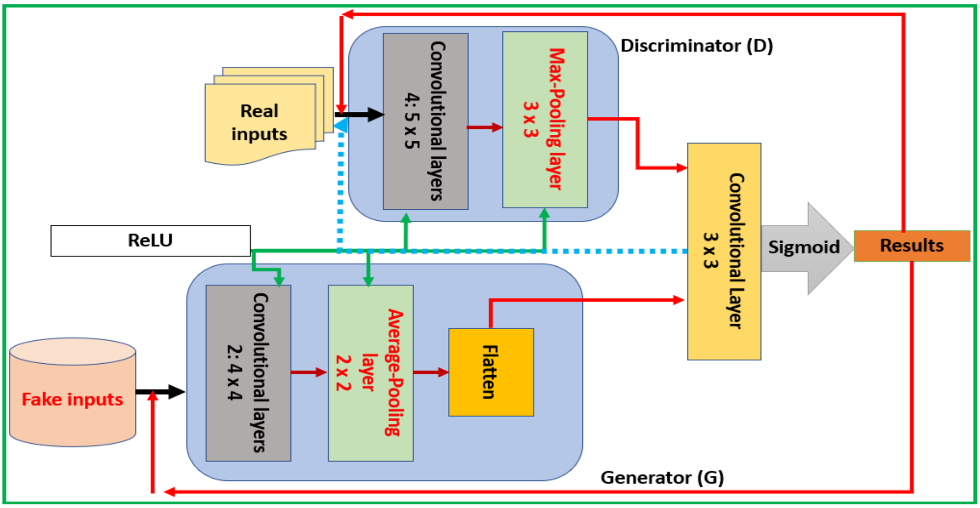

Figure 2. The internal structure of the implemented NGAN model

Table 3. The applied configurations of NGAN

The name of the layer and defined parameters	*Values*
Learning rate (NL)	*0.001*
Batch size (Nbs)	*16*
Convolutional	*64 filters, 3 x 3 of kernel size, padding = same, 1 number of strides, and 50 neurons*
Optimizer	*Adam*
Activation function	*Sigmoid*
Number of epochs (Noe)	*42*
Number of iterations (noi)	*1500*
Generator	
Convolutional	*64 filters, 4 x 4 of kernel size, padding = same, 2 number of strides, and 100 neurons*
Average-Pooling	*2 x 2 and 2 number of strides.*
Activation function	*ReLU*
Discriminator	
Convolutional	*192 filters, 5 x 5 of kernel size, padding = same, 3 number of strides, and 150 neurons*
Max-Pooling	*2 x 2 and 2 number of strides.*
Activation function	*ReLU*

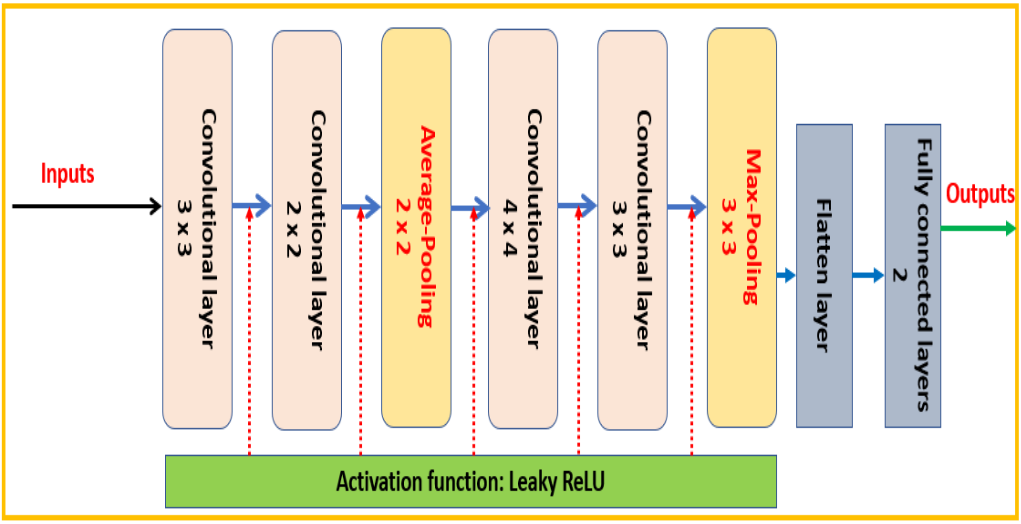

Figure 3. The internal structure of the implemented NCNN model

Table 4. Settings of NCNN

The name of layer and defined parameters	*Values*
Learning rate (L)	*0.0001*
Batch size (bs)	*16*
Optimizer	*Adam*
Convolutional	*128 filters, 3 x 3 of kernel size, padding = same, 2 number of strides, and 100 neurons*
Convolutional	*64 filters, 2 x 2 of kernel size, padding = same, 2 number of strides, and 100 neurons*
Average-Pooling	*2 x 2 and 2 number of strides.*
Convolutional	*512 filters, 4 x 4 of kernel size, padding = same, 3 number of strides, and 200 neurons*
Convolutional	*192 filters, 3 x 3 of kernel size, padding = same, 3 number of strides, and 150 neurons*
Max-Pooling	*3 x 3 and 2 number of strides.*
Dropout	*0.3*
Number of epochs (Noe)	*75*
Number of iterations (noi)	*1500*
Activation function	*Leaky ReLU*

In the proposed framework, data augmentation during the training stage is performed to make the model robust and able to generate reasonable findings. During this process, the brightness of inputs is improved to let the framework captures diverse characteristics and resolves the overfitting problem. The model can detect small fires, smoke, or flames. The learning processes are improved since the framework deploys three sizes of the learning parameter as shown in Tables 2, 3, and 4. The deployed layers of convolutional, pooling, and fully connected are chosen to minimize the complexity of the model. The different sizes of developed layers simplify the framework. Each layer comes with its own parameters as shown in the previous three Tables from 2 to 4.

Various quantities are computed to evaluate the framework performance. These quantities are True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), Precision (PCS) and is calculated as in equation (1):

$$\text{PCS} = \frac{TP}{(TP+FP)} \tag{1}$$

Recall (RCA), this metric is computed as shown in equation (2):

$$\text{RCA} = \frac{TP}{(TP+FN)} \tag{2}$$

Accuracy (ACU) and is evaluated as in equation (3):

$$\text{ACU} = \frac{(TP+TN)}{(TP+TN+FN+FP)} \tag{3}$$

4. RESULTS AND DISCUSSION

The proposed framework was evaluated on a machine that contains MATLAB software. The MATLAB version was R2017b. The specifications of the hosting machine are an Intel core I7, 8^{th} generation with 2GHz and 16 RAM. Unfortunately, an installed GPU was malfunctioning. The evaluation procedures were conducted to analyze results and discover any indicators that could influence the framework. The performance metrics were analysed using three fire datasets as stated earlier. In this evaluation process, 0.01 was the assigned weight decay and the momentum was set to 0.75.

As shown in Table 1, the training set contains 23,582 images of different sizes. Table 5 displays the average obtained values of the considered performance quantities. 96.8% is the average accuracy, while 97.76% was the maximum one. The use of global and local characteristics extraction allowed the framework to be able to distinguish efficiently between fire and non-fire regions. The processing time per input was improved by nearly 23% when an image was resized to 332 x 332 pixels. However, this reduction had a small negative impact on the classification since the resolution became smaller and the framework faced difficulties in extracting the features. The proposed framework reached its maximum values at 96.94%, 96.15%, and 97.76% for precision, recall, and accuracy, respectively. Figure 4 depicts the obtained accuracy and loss values for 50 epochs with an outer learning rate = 0.001. For 50 epochs, the number of iterations was 1550 and this number was automatically generated by the system. The validation process occurred every 30 iterations and it is marked by black dots and dashed lines.

Table 5. The achieved average results of the performance indicators

Performance quantities	Achieved results
Precision	96.35%
Recall	95.23%
Accuracy	96.8%

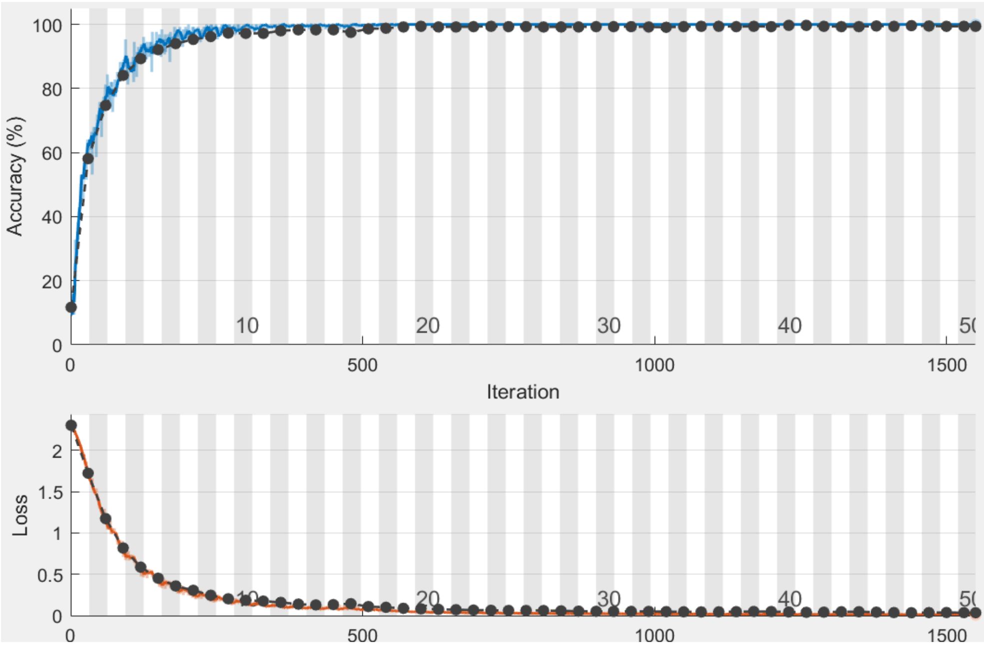

Figure 4. The obtained accuracy and loss charts

The accuracy became steady after almost 20 epochs and reached its maximum value at 97.76%, while the loss function became stable at 30 epochs and its value was 0.1. Table 6 lists the parameters of the model complexity and their evaluated results. These parameters are the processing time in seconds per input, the total number of utilized parameters, and the total number of Floating-Point Operations per Second (FLOPS). These metrics were evaluated using 640 x 640 as the input size for each image. The number of FLOPS and parameters were found in millions. Figure 5 illustrates samples of the obtained results for fire detection.

Table 6. The evaluated model complexity

Processing time	FLOPS (M)	Number of parameters (M)
5.35s	21.82	34.62

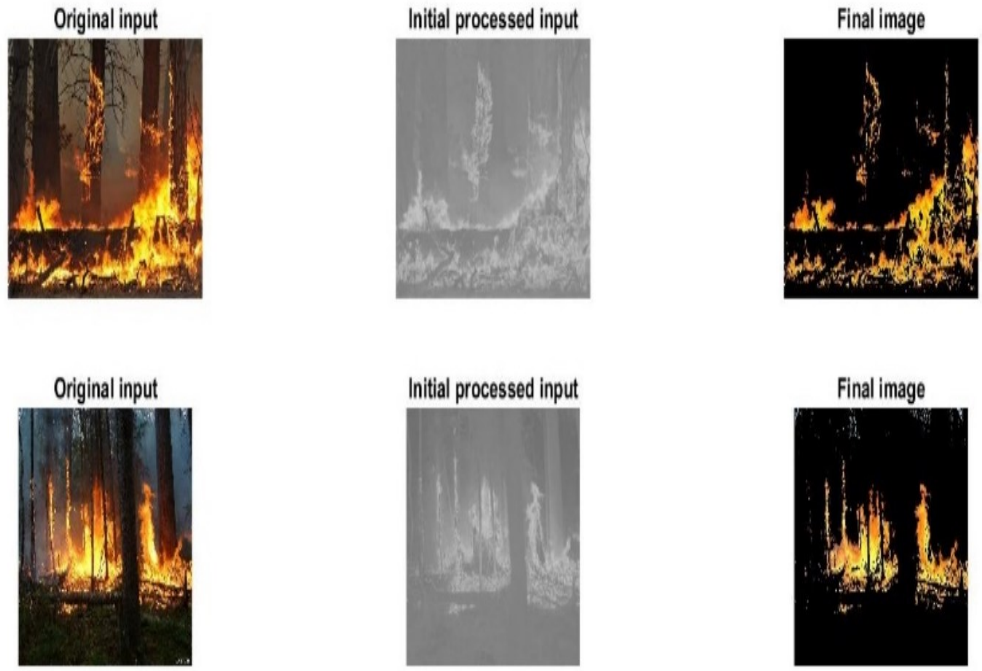

Figure 5. The outputs of the proposed framework

Table 7 shows a comparative evaluation between some the of developed models in [14-17], and the proposed framework regarding the deployed tools and achieved accuracy. The proposed framework outperforms other approaches.

Table 7. The comparative outcomes

Works	Deployed tool	Accuracy
[14], 2023	YOLOv5	60.4%
[15], 2023	Hybrid machine learning methods	92%
[16], 2023	Ensemble Lightweight YOLOX-L and Defogging Method	86.13%
[17], 2023	Staged YOLO Model and Ensemble CNN	85%
The proposed framework	The New Light Ensemble Deep-Learning Framework (NLEDLF).	96.78%

4. CONCLUSION

This article proposes a New Light Ensemble Deep-Learning Framework (NLEDLF). This model achieves better results than other implemented methods regarding accuracy when applied to detect fires. Two new ensemble deep-learning technologies are developed to extract spatial and spectral local and global features. The limitations of this model are some small fires could be misdetected and the processing time is considered relatively high.

Future work can be planned to minimize the execution time to use the model as a real-time application and improve its capabilities to detect all fires without leaving anything.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the approval and the support of this research study by grant no. ENGA-2023-12-1974 from the Deanship of Scientific Research at Northern Border University, Arar, K.S.A.

REFERENCES

[1] Gritsyuk, K. M., Gritsyuk, V. I. Convolutional and long short-term memory neural networks based models for remaining useful life prediction. *International Journal on Information Technologies and Security*, vol. 14, no. 1, 2022, pp. 61-76.

[2] Tsonkov, N., Petrov, K. Valcheva, T.B. Adoption of information technologies for Black Sea region municipalities' smart development. *International Journal on Information Technologies and Security*, vol. 14, no. 1, 2022, pp. 87-96.

[3] Kravets, O. Ja., Aksenov, I. A., Rahman, P. A. Algorithmization of analytical methods for finding motion vectors when processing image series, *International Journal on Information Technologies and Security*, vol. 14, no. 2, 2022, pp.39-50.

[4] Battur, R., Jagadisha N. A performance aware contents-based image retrieval (CBIR) techniques, *International Journal on Information Technologies and Security*, vol. 14, no. 2, 2022, pp.87-98.

[5] Al-Ragehi, A., Abdulkadir, S. J., Muneer, A., Sadeq, S., Al-Tashi, Q. Hyper-parameter optimization of semi-supervised GANs based-sine cosine algorithm for multimedia datasets, *Computers, Materials and Continua*, vol. 73, no. 1, 2022, pp. 2169-2187.

[6] Zhang, E., Zhang, J., Bai, J., Bian, J., Fang, S., Zhan, T., Feng, M. Attention-embedded triple-fusion branch CNN for hyperspectral image classification. *Remote Sensing*, vol. 15, no. 2150, 2023, pp. 1-22.

[7] Sathya D., Primya T., Vinothini S., Priya J., Jagadeesan D. Smart health system using stacking ensemble classification algorithm. *International Journal on Information Technologies and Security*, vol. 14, no. 3, 2022, pp. 67-78.

[8] Tamire, N. A., Kim, H. D. Effective video scene analysis for a nanosatellite based on an onboard deep learning method. *Remote Sensing*, vol. 15, no. 2143, 2023, pp. 1-10.

[9] Chui, K. T., Gupta, B. B., Ruiz, M. T., Arya, V., Alhalabi, W., Zamzami, I. F. A Convolutional Neural Network-based feature extraction and weighted twin support vector machine algorithm for context-aware human activity recognition. *Electronics*, vol. 12, no. 1915, 2023, pp. 1-15.

[10] Talha, M., Sarfraz, M., Rahman, A., Ghauri, S. A., Mohammad, R. M., Krishnasamy, G., Alkharraa, M. Voting-based deep Convolutional Neural Networks (VB-DCNNs) for M-QAM and M-PSK signals classification. *Electronics*, vol. 12, no. 1913, 2023, pp. 1-17.

[11] https://github.com/gaiasd/DFireDataset.

[12] https://ieee-dataport.org/open-access/large-scale-dataset-active-fire-detectionsegmentation-landsat-8.

[13] https://www.kaggle.com/datasets/dataclusterlabs/fire-and-smoke-dataset.

[14] Li, J., Xu, R., Liu, Y. An improved forest fire and smoke detection model based on YOLOv5. *Forests*, vol. 14, no. 833, 2023, pp. 1-13.

[15] Shao, Y., Feng, Z., Cao, M., Wang, W., Sun, L., Yang, X., Ma, T., Guo, Z., Fahad, S., Liu, X., Wang, Z. An ensemble model for forest fire occurrence mapping in China. *Forests*, vol. 14, no. 704, 2023, pp. 1-15.

[16] Huang, J., He, Z., Guan, Y., Zhang, H. Real-time Forest fire detection by ensemble lightweight YOLOX-L and defogging method. *Sensors*, vol. 13, no. 1894, 2023, pp. 1-15.

[17] Bahhar, C., Ksibi, A., Ayadi, M., Jamjoom, M. M., Ullah, Z., Soufiene, B. O., Sakli, H. Wildfire and smoke detection using staged YOLO model and ensemble CNN. *Electronics*, vol. 12, no. 228, 2023, pp. 1-15.

Information about the authors:

Ahmed A. Alsheikhy is an associate professor and works at the Electrical Engineering Department at Northern Border University. His interest in research includes Artificial Intelligence, performance evaluation, medical applications, networks, and cyber security.

Tawfeeq Shawly is an assistant professor and works at the Electrical Engineering Department at King Abdulaziz University, Rabigh branch. He works on cyber-physical and security systems, artificial intelligence, and networks.

Hossam E. Ahmed is an assistant professor at the Electrical Engineering Department at Northern Border University, Arar, Saudi Arabia. In addition, is a faculty member of the Electrical Engineering Department in the Faculty of Engineering at Benha University, Benha, Egypt. His research interests are communication, electronics, and computer applications.

Hassan Fareed M. Lahza is an assistant professor and works in the Information Systems Department at Umm Al-Qura University. He works on networks, artificial intelligence, and information systems.

Manuscript received on 17 August 2023