Automatic Outdoor Fire Detection Using Deep Learning

Aesr S. Abdalsattar [1], Naji M. Sahib[1]
[1](Department of Computer Science, College of Science, University of Diyala,, Baquba, Iraq)

ABSTRACT : *One of the unlucky phenomena that contribute to environmental disasters is fire, which also poses a serious threat to human safety and life, especially when sensor-based fire detection systems fail to detect it. Therefore, putting inexpensive and efficient sensors in certain locations will greatly speed up the detection of fires. Vision-based fire detection systems take advantage of the three fundamental features of fire: color, movement, and shape (fire shape). Systems for detecting smoke and fire have been developed using security camera footage. In this study, the characteristics of fire photographs are recovered, and inputs are then created. To a deep learning-based classification strategy where the photos are split, given labels, and the form of the flame is extracted before the movement of the flame is recognized. This method is believed to be one of the most contemporary ones in recent years.*

KEYWORDS -*Deep learning, YOLOV5S model, Fire detection, Image collection*

I. INTRODUCTION

The greatest threat to human life comes from abnormal occurrences, which include floods, earthquakes, accidents, battles, medical problems, and fires. These occurrences also pose a threat to public safety and health. The majority of other occurrences are fires[1]. Early smoke or fire detection is essential to enable prompt action to prevent extensive damage. There are numerous techniques and equipment that can be used to spot fire or smoke in a setting. The majority of conventional detection techniques rely on tools and sensors that are designed to pick up the presence of smoke or flame [2]. All conventional methods of fire early detection have significant shortcomings because they are prone to breakdowns, necessitate routine maintenance, and can only detect flames or flames near places where installed in it. Additionally, they do not provide enough details about the fire's direction, initial location, size, and numerous other flaws [3].

To address the shortcomings of conventional approaches, a variety of video and image-based techniques were employed. Cameras, which may be deployed in all indoor environments, are used in these techniques to acquire still photos and video footage. As well as the exterior to spot flames and smoke[4]. using a variety of methods, such as the static decision rule based on color and the dynamic decision rule based on diffusion, to determine whether or not a pixel was smoke [5]. Additionally, there are methods for detecting fire and flames that make advantage of the characteristics of blurring edges and flash, and the video-based luminous flux method was also documented[6].

The main goal of the aforementioned approaches is to develop a rule-based algorithm based on handcrafted features or specific specialized knowledge that, depending on a variety of variables such as composition, texture, color, and motion, can be the key characteristics of smoke and fire. This is a difficult and complicated task. Recently, a different method has been used to advance the detection of fires in videos and images with the aid of deep learning, which represents an emerging concept that supports artificial neural networks and has produced exceptional results in a number of cases. A correct and accurate system that monitors fire as soon as possible is provided in addition to computer vision to overcome the weaknesses in the current organization. This system is able to operate in a variety of environments and can recognize and extract the significant features of color and smoke in photographs and video clips. Convolutional neural networks (CNN), in particular, have produced outstanding results when used to solve challenges involving visual recognition thanks to deep learning [7].

II. RELATED WORKS

In 2016, S. Frizzi et al. [8] The effectiveness of deep learning target detection technology to detect forest fires was demonstrated by the use of the convolutional neural network (CNN) to detect fires by segmenting the full image into numerous small sections.

In 2016, Q. Zhang et al. [9] A whole-image fire classifier was trained using deep correlated convolutional neural networks (CNN), and it was then utilized to foretell forest fires. During training, the suggested method's detection accuracy was 97% and 90%.

In 2018, K. Muhammad et al.[10] The results demonstrated how effectively a Transformational Neural Network (CNN) and synchronized video monitoring apps worked in terms of fire detection. A proposed deep learning-based fire detection technique has been used.

In 2019, Z. Jiao et al. [11] With the aid of YOLOv3, (CNN) was implemented in this work. The test results show that this algorithm has a detection frame rate of more than 3.2 fps and a recognition rate of almost 83%. For applications using UAVs to detect forest fires in real time, this technology has many benefits.

In 2021, R.Xu et al. [12] Yolo v5 and Efficient Net, two different learners, were merged to achieve the fire detection process in a recently suggested ensemble learning system for forest fire detection under varying situations. Utilizing the Efficient Net scenario for individual students is an additional strategy for preventing false positives. Without adding additional delay time, false positives dropped by 51.3% and detection performance rose by 2.5% to 10.9%.

III. DEEP LEARNING

In recent years, deep learning has revolutionized almost every industry, including those in automation, robotics, and healthcare. Due to its success Deep learning has emerged as one of the most popular approaches in applications such as language processing, object recognition, image categorization, and matching news, publications, or products of interest to customers.[13]Understanding how to categorize and recognize photographs is one of the most crucial foundational concepts in object detection. Convolutional layers, aggregation layers, and fully connected layers (FCL) are the three primary neural layers that make up a convolutional neural network (CNN) in most cases. The deep learning training process is shown in Figure (1).

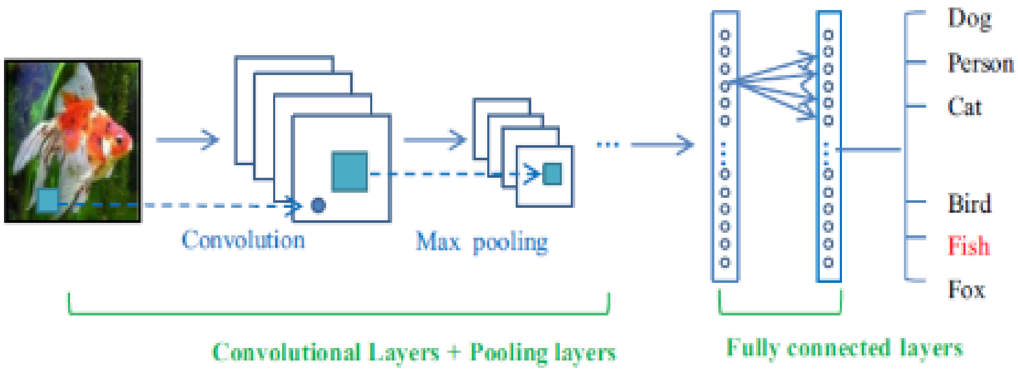

Fig 1. Deep learning training algorithm [7].

IV. BASIC YOLOV5S MODEL

Glenn Jocher published the YOLOv5 model in 2020It is a target identification model that relies on regression. The foundation for the development of the YOLOv5 model was provided by target detection models like YOLOv3 and YOLOv4. The YOLOv5 model maintains detection speed while enhancing detection accuracy in comparison to earlier models. The YOLOv5 model is made up of the YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x structures. The number of convolution kernels used in each of these four systems and bottlenecks in specific areas are the main differences between them. The accuracy of each model's detection improves as the parameters for each of the four models rise. As a result, the models' scale must be severely limited. The YOLOv5s model was selected as the experimental subject in this investigation as a result. Glenn Jocher published the YOLOv5 model in 2020. It is a regression-based target identification model. Target detection models like YOLOv3 and YOLOv4 served as the foundation for the creation of the YOLOv5 model. The YOLOv5 model maintains detection speed while enhancing detection accuracy in comparison to earlier models. The YOLOv5 model is made up of the YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x structures. The key variations among these four systems are the quantity of convolution kernels and bottlenecks in particular regions. The accuracy of each model's detection improves as the parameters for each of the four models rise. As a result, the models' scale must be severely limited. The YOLOv5s model was selected as the experimental subject in this investigation as a result. The detection network processes employing anchor boxes to generate detection boxes that represent the category on the input feature map, location, and level of certainty of the target inside the image. The YOLOv5s model's intricate network structure is depicted in Fig.2 [14].

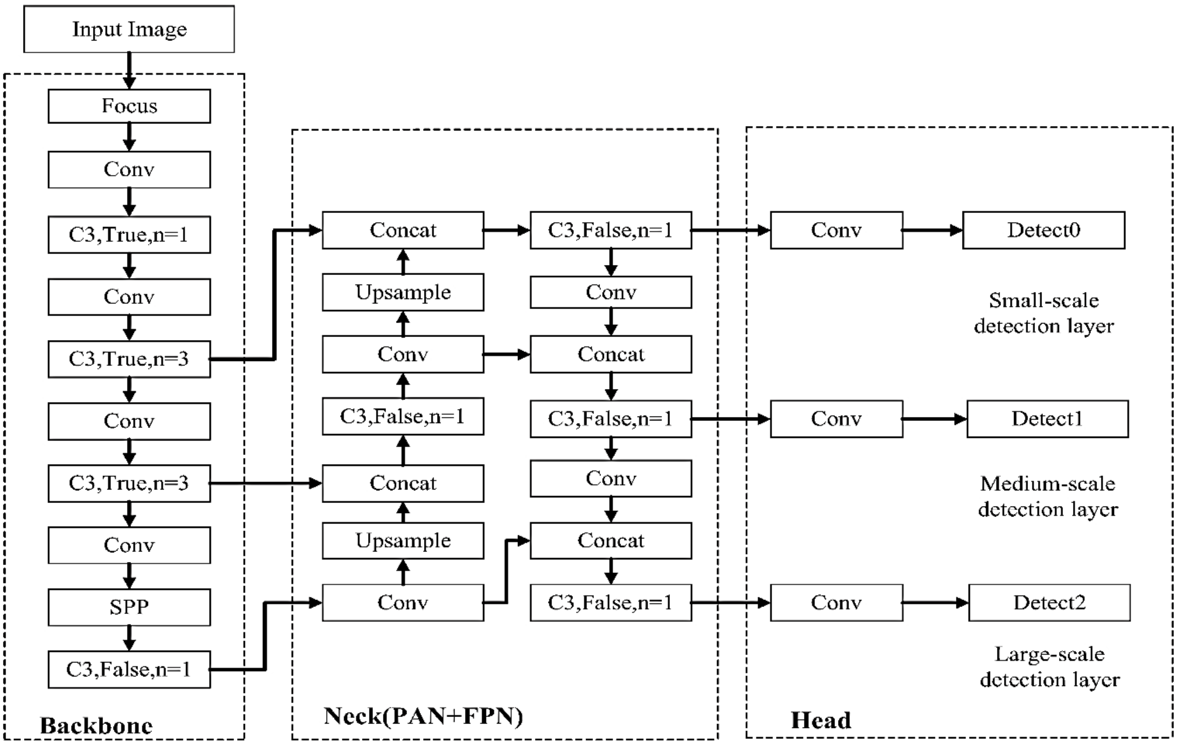

Fig 2. Structure of the YOLOv5s model's overall.

V. CONVERT DATA ANNOTATION

A dataset of photos with labels (Data Annotation) that include the coordinates of the bounding squares must be used to train the YOLOv5s algorithm. Drawing a bounding box by hand for every image that the system will train on is challenging since this requires using a dataset with comprehensive annotation.

The challenge is made more difficult by the fact that some of the photographs were taken under real-world circumstances, making some of the faces blurry, concealed, rotated, or wearing various forms of masks. The suggested system would learn the incorrect characteristics, leading to false detection and a reduction in the precision and accuracy of the system if annotations are incorrect in certain photos. As a result, precise training of the suggested system is essential, and the process of annotating is crucial during the deployment stage.

The YOLOv5s algorithm only accepts annotations with the TXT extension. However, some of the photos were in XML format since the two databases were combined. Some of the photos also had inadequate or inaccurate annotations. As a result, these annotations needed to be processed before the algorithm was applied to the photos. In this study, the code was implemented in algorithm to overcome the issues with annotation.

The five models of the YOLOv5 algorithm's fifth iteration differ in each model's speed, size, accuracy, and layer depth. Version YOLOv5s was touted by the YOLOv5 algorithm's creators as being the quickest of the other versions.

The finding that expanding the model's size would not significantly increase the model's efficiency led to the choice of the modest model YOLOv5s for the suggested system. However, because the model's goal is to identify just two distinct categories, complexity will increase and detection speed will decrease (with or without a mask). The simplest, fastest, and shortest model should be selected as a result, especially because YOLOv5's model has a strong real-time detection speed of 155 FPS.

VI. PROPOSED APPROACH

In this work, we create a new dataset from merge two dataset naming; fir and smoke dataset as the standard dataset “https://www.kaggle.com/datasets/dataclusterlabs/fire-and-smoke-dataset”. The second dataset is the standard fire and gun dataset “https://www.kaggle.com/datasets/atulyakumar98/fire-and-gun-dataset”

The fire and smoke dataset is that. This incredibly difficult dataset contains more than 7000+ original Fire and Smoke photographs that were crowdsourced from more than 400+ urban and rural regions. Datacluster Labs' machine vision experts carefully evaluate and validate each image. Additionally, it is not enough to only identify the flame when detecting complex fire conditions. With the information on flame, smoke, and 5,000 images stripped from gunshots The experiment's fire dataset was created with the goal of resolving issues with existing datasets. It covered car fires, building fires, industrial fires, and urban fires, etc. Figure (3) is an illustration of a recent dataset:

Fig 3. example of a new dataset.

VII. SECOND PROPOSED SYSTEM TESTING

Testing the created system proposal is the procedure's final phase. Three test data sets were used at this point to assess the proposed system. The system processes test data made up of information that is comparable to the training data. Following that, the system compares the parameter value to the training value. Since there are three datasets, the algorithm classifies the Face or Face Mask object for each of them. Use the suggested technique to tell who is wearing a face mask from who is not. Test data from all 10630 pictures in the combined data set as well as test data from each individual data set were used to assess the proposed system. Consequently, the system was trained on three sets of data using three different models.

VIII. RESULTS OF THE ONLINE APPROACH

Each object image must have a label for object detection models to properly process the photos. The fundamental information in the image collection is represented by annotations. They include the necessary text information that identifies a specific image as well as the bounding box coordinates, class label (the name of the class), and class label information. Each fire and its area are identified in the image using the annotation process.

IX. DATA ANNOTATION

One of the most well-liked computer vision methods is object detection and data annotation because of its adaptability. The goal of fire and non-fire object identification models is to locate related items in photographs and divide them into two groups. Additionally, they train an object identification model to recognize all variations of these items' faces, whether or not they are covered by a mask, and accurately categorize them. Figure (4) below uses the preceding photos as models to demonstrate how to find a fire.

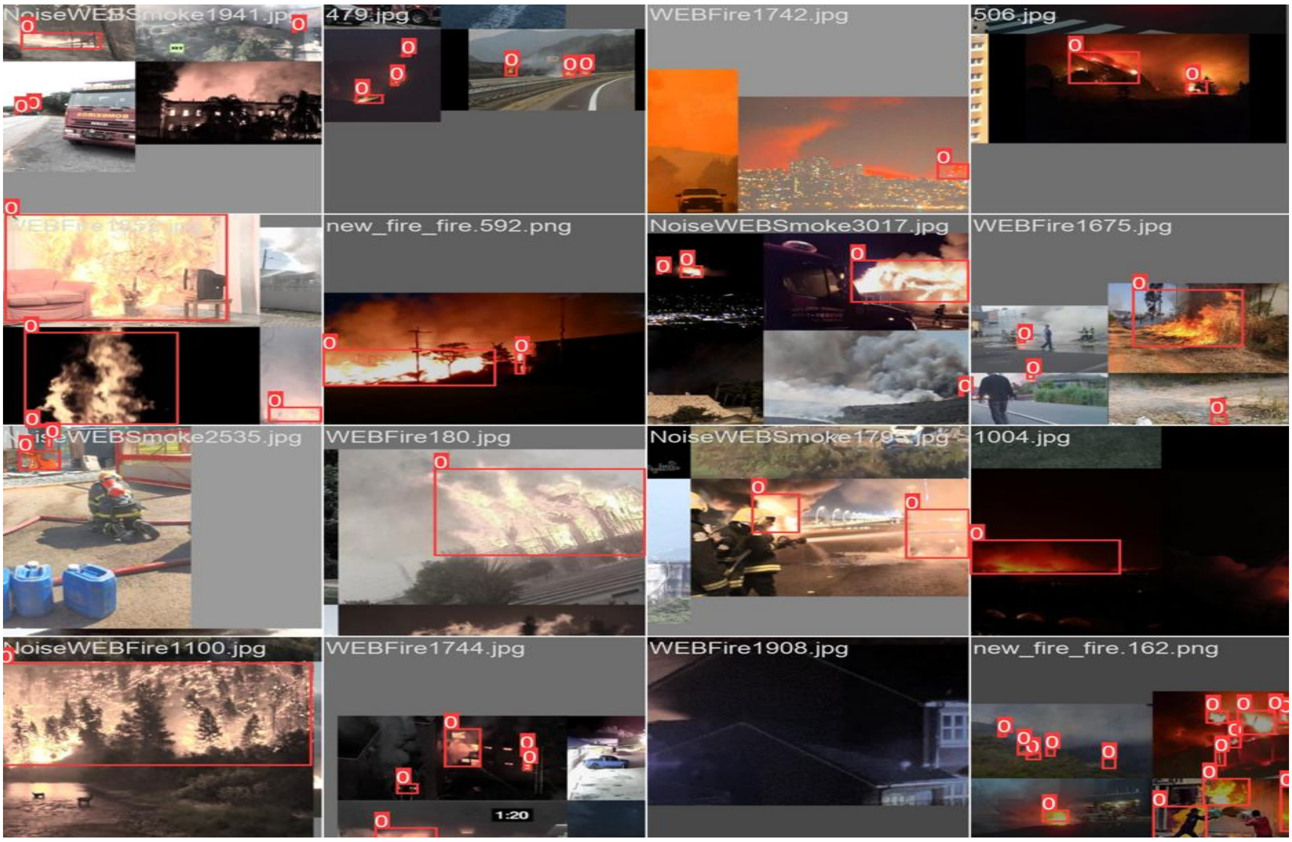

Fig 4. photos as models to demonstrate how to find a fire.

X. MODEL TRAINING RESULTS

The training outcomes vary depending on how many epochs were run for each dataset. For each dataset, four training cycles with varying numbers of epochs were run, and the results are as follows. Objective loss quantifies the probability that an item exists in a proposed area of interest. If objectivity is high, the picture window will likely include an item and suggest labels like those in Figure (5). The classification loss indicates how well the algorithm can identify the correct class for a given object. Less loss during training and validation means lower values for the box, objectless, and classification losses, which enhances the results and increases the model's efficiency. A rise in the accuracy, recall, and mean average precision statistics, on the other hand, denotes a stronger prediction and raises the model's usefulness.

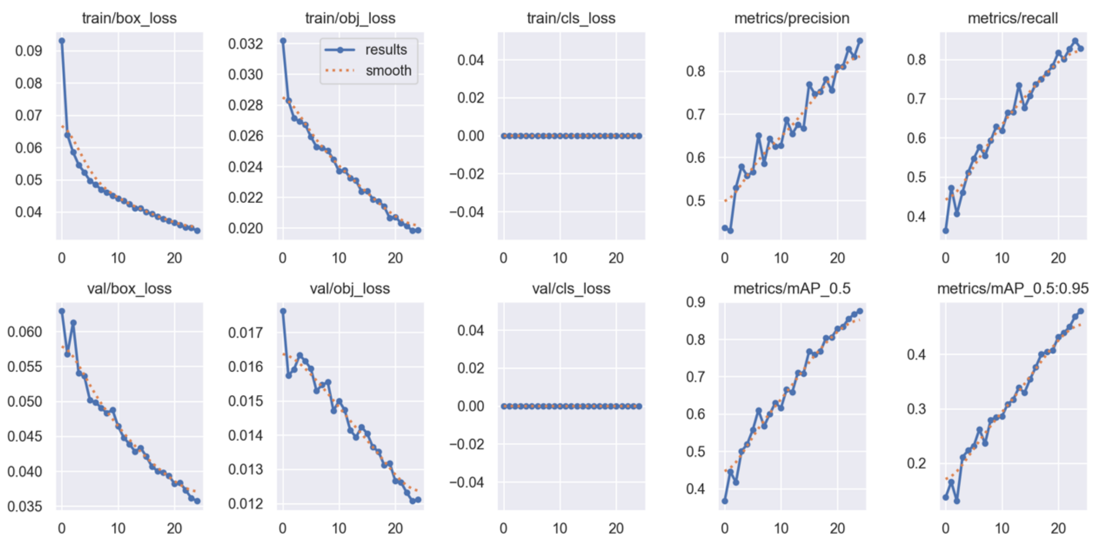

Fig 5. Best results of training the proposed system.

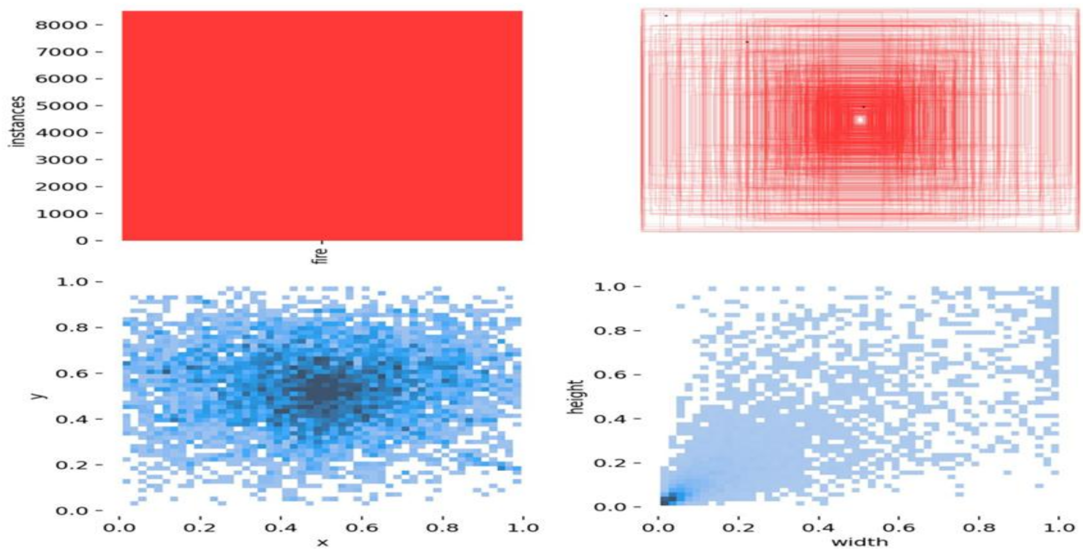

Fig 6. labels of fire images

XI. CONCLUSION

The detection of fire images is the focus of this thesis. Our models base two approaches which are frames images and real-time testing, and considering the outcomes, it came to the following conclusions:

1. The proposed methodology is a robust approach that utilizes modern technologies and computer facilities to detect and classify fires accurately and rapidly.

2. The choice of a data collection encompassing all cases is one of the challenges the suggested system confronts. One data set was not enough to be used. The model must be trained on pictures that had more problems than the examples in the data set in order to handle varied obstacles. As a result, two datasets with a focus on the face mask and face were combined. For each data collection, three models of the system were created. Separately. The suggested model effectively overcame all obstacles that might prevent the discovery of the face mask for the new merged data set, including angles, distortion, the kind of mask, and the degree of the presence of…

3. To properly train the model, the data set's annotations and pictures must be used. One of the challenges with this model is that if the labels include mistakes or there is a shortage of data, the model will be incorrectly trained, which will result in decreased efficiency. As a result, more effort was put into changing the annotations' XML extension to a TXT extension.

7. The suggested approach is quite effective in detecting both fire and no fire, and it may be applied to real-world situations. The initial options for places where the suggested approach might be helpful include supermarkets, schools, universities, hospitals, and airports.

REFERENCES

[1] R. L. P. Custer, *Fire detection: the state-of-the-art, vol. 839.* US Department of Commerce, National Bureau of Standards, 1974 *Note that the journal title, volume number and issue number are set in italics.*

[2] A. Namozov and Y. Im Cho, “An efficient deep learning algorithm for fire and smoke detection with limited data,” Adv. Electr. Comput. Eng., vol. 18, no. 4, pp. 121–128, 2018.

[3] A. Gaur, A. Singh, A. Kumar, A. Kumar, and K. Kapoor, *“Video flame and smoke based fire detection algorithms: A literature review,” Fire Technol., vol. 56,* pp. 1943–1980, 2020.

[4] T.-H. Chen, Y.-H. Yin, S.-F. Huang, and Y.-T. Ye, *“The smoke detection for early fire-alarming system base on video processing,” in 2006 international conference on intelligent information hiding and multimedia*, 2006, pp. 427–430.

[5] B. U. Töreyin, Y. Dedeoğlu, U. Güdükbay, and A. E. Cetin, *“Computer vision based method for real-time fire and flame detection,” Pattern Recognit. Lett., vol. 27,* no. 1, pp. 49–58, 2006.

[6] M. Mueller, P. Karasev, I. Kolesov, and A. Tannenbaum, *“Optical flow estimation for flame detection in videos,” IEEE Trans. image Process., vol. 22, no. 7*, pp. 2786–2797, 2013.

[7] M. S. S. W. Yunes and K. I. Alsaif, *“Fire Detection Effect of Data Size Using on Deeplearning Techniques,” Eurasian J. Eng. Technol., vol. 6,* pp. 42–53, 2022.

[8] S. Frizzi, R. Kaabi, M. Bouchouicha, J.-M. Ginoux, E. Moreau, and F. Fnaiech, *“Convolutional neural network for video fire and smoke detection,” in IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*, 2016, pp. 877–882.

[9] Q. Zhang, G. Lin, Y. Zhang, G. Xu, and J. Wang*, “Wildland forest fire smoke detection based on faster R-CNN using synthetic smoke images,” Procedia Eng., vol. 211,* pp. 441–446, 2018.

[10] R. K. Mohammed, *“A real-time forest fire and smoke detection system using deep learning,” Int. J. Nonlinear Anal. Appl, vol. 13,* pp. 2008–6822, 2022.

[11] Z. Jiao et al., *“A deep learning based forest fire detection approach using UAV and YOLOv3,” in 2019 1st International conference on industrial artificial intelligence (IAI),* 2019, pp. 1–5.

[12] R. Xu, H. Lin, K. Lu, L. Cao, and Y. Liu, *“A forest fire detection system based on ensemble learning,” Forests, vol. 12, no. 2*, p. 217, 2021.

[13] Y. LeCun, Y. Bengio, and G. Hinton, *“Deep learning,” Nature, vol. 521, no. 7553,* pp. 436–444, 2015.

[14] G. Wen et al., *“YOLOv5s-CA: A Modified YOLOv5s Network with Coordinate Attention for Underwater Target Detection,” Sensors, vol. 23, no. 7,* p. 3367, 2023.