Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 2233–2242

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

FireNet-*v2*: Improved Lightweight Fire Detection Model for Real-Time IoT Applications

Anwer Shees[a], Mohammad Samar Ansari[b,*], Akshay Varshney[c], Mamoona Naveed Asghar[d], Nadia Kanwal[e]

[a]*Z.H. College of Engineering and Technology, Aligarh Muslim University, Aligarh, India*
[b]*Faculty of Science and Engineering, University of Chester, United Kingdom*
[c]*Adobe Systems, India*
[d]*National University of Ireland Galway, Ireland*
[e]*School of Computing and Mathematics, Keele University, United Kingdom*

Abstract

Fire hazards cause huge ecological, social and economical losses in day to day life. Due to the rapid increase in the prevalence of fire accidents, it has become vital to equip the assets with fire prevention systems. There have been numerous researches to build a fire detection model in order to avert such accidents, with recent approaches leveraging the enormous improvements in computer vision deep learning models. However, most deep learning models have to compromise with their performance and accurate detection to maintain a reasonable inference time and parameter count. In this paper, we present a customized lightweight convolution neural network for early detection of fire. By virtue of low parameter count, the proposed model is amenable to embedded applications in real-time fire monitoring equipment, and even upcoming fire monitoring approaches such as unmanned aerial vehicles (drones). The fire detection results show marked improvement over the predecessor low-parameter-count models, while further reducing the number of trainable parameters. The overall accuracy of FireNet-*v2*, which has only 318,460 parameters, was found to be 98.43% when tested over Foggia's dataset.

Keywords: Fire Detection; Convolution Neural Network; FireNet; Deep learning.

1. Introduction

Fire is one among the several hazards which puts up human life at risks and destroy properties [15]. Early detection of fire is as important than ever to prevent from irreparable fire hazards. There are existing methods like traditional fire

* Corresponding author
E-mail address: m.ansari@chester.ac.uk

10.1016/j.procs.2023.01.199

detectors based on thermal approaches, photometry and chemical detection which raise alarm to prevent fire accident to happen. However, the main drawback of these methods is it requires a sufficient amount of smoke or fire to sense and trigger. Moreover, it can not be deployed for outdoor detection such as forest, streets, playgrounds, industries etc. and have high rate of false triggering. With advancement of technologies researchers have contributed with image processing, computer vision and artificial intelligence (AI) techniques which get the better of issue faced by previous methods. Using these state-of-art technologies in fire detectors leads to more robust, fast, and reliable fire detection performance. Furthermore, such detectors are also capable of performing in outdoor environments with less false triggering and faster response times (as the build up of smoke is not an essential requirement for such a detector).

Considering that these days almost every place is under cameras and closed circuit television (CCTV) surveillance, and that the infrastructure like streets, industries, shopping malls, parks, buildings etc. have complete interlinked networking from visualizing the objects and activities through cameras and CCTVs to monitoring with the help of Internet of Things (IoT), the visual-based fire detectors have made it possible to integrate the state of art of vision based fire detectors with the existing technology of video surveillance [13], thereby eliminating the cost and materials used in traditional fire detectors by just utilizing a software (the deep learning model) in video based fire detectors.

The development in digital camera and computer vision technologies, coupled with intelligent video methods is being increasingly used in fire detectors by replacing the hand designed detectors. Initially, hand-designed features for identifying smoke and fire in image and video classification were extracted using colour and form properties of smoke. As we meet a non-stop growth in electronics together with evolution of graphical processing units (GPUs) and high performance processors great enough to provide large amount of computation, we witness increasing advances in artificial intelligence. Deep learning [25] models are constantly used for fire detectors at this age of time which eliminate the need for manual feature extraction and brings an automatic extraction of features directly from images [47]. Also the deep leaning (DL) visual based fire detectors are more accurate, show lower false trigger, are more robust and more reliable.

Contribution and Novelty: In this work we present an improved version of FireNet [19], with less false triggering and significantly lesser number of trainable parameters in the model. The novelty lies in the fact that the proposed lightweight model is tailored from scratch for the specific purpose of fire detection only. We go on to show that even with the significantly reduced parameter count, the proposed model is able to provide better performance than its predecessor, thereby making the proposed model a viable contender for modern day computer vision based embedded fire detectors.

The following is a breakdown of the paper's structure. The second section highlights previous research in the areas of by-hand (manual) feature engineering and deep learning for fire detection models. Section III describes our complete architecture and motivation of using FireNet-*v2* over FireNet, followed by Section IV which includes a detailed description of the datasets and comparison of results with other existing works. This is then followed by Section V which contains a short discussion. Finally, in Section VI, there are some closing observations.

2. Related Works

A lot of study has gone into developing potent and efficient fire detection systems to defend against fire dangers. Both handcrafted techniques and automated fire recognition have been acknowledged by the researchers as a key factors in fire detection methods.

Recent technological breakthroughs have resulted in a range of sensors for various purposes. For viewing of the human body's interior they used wireless capsule sensors [26], obstacle detection sensors for vehicles [4], and sensors that detect fire [22]. Most fire detection sensors used in modern-day solutions, such as ion-based, infrared, and optical sensors, must be in close proximity to the site/source of heat, radiation (infra-red), fire, and/or smoke to activate, making them unsuitable for critical settings. Such approaches also pose the issues related to false triggering [23]. Vision-based sensors are a popular alternative to these sensors because they offer many benefits over conventional sensors, such as reduced costs, faster reaction times, wider surveillance coverage, and little (or none) human intervention, which eliminates the requirement to get to the site where the fire alarm was sounded [11]. Despite the fact that vision-based sensors offer a number of promising characteristics, they do have significant drawbacks, such as dependence on scene complexity & changing lighting conditions, and poor camera image quality because of network issues. Early on, the researchers focused on the flame detection's motion and colour features to build customised algorithms for fire detection. Thou-Ho et al [9] showed real flame identification requires both chromatic and dynamic aspects of flames and smoke. Authors in [8] used two separate colour schemes to identify fire scenarios from only-smoke situations, and to make the categorization more reliable, they used fuzzy logic concepts to distinguish fire scenes from other fire-like scenes and images. In [40], the authors explored the YCbCr colour model created new criteria for separating luminance and chrominance components, resulting in rule-based flame pixel classification. Authors in [14] looked into another colour model *viz.* YUV, with motion for pixel categorization into fire or non-fire potential. In certain publications, in addition to the colour feature, motion has been used as a criteria for detecting fire. The fire and smoke have both static and dynamic characteristics which were employed by Rafiee et al [32]. However, owing to the existence of additional background items with comparable colour attributes to the fire pixels, the false negative rate remains an issue. Aside from colour models, particular low-level fire zone features such as skewness, colour, roughness, and area size, and so on, have been utilised to determine frame-to-frame variations, which may be used in conjunction with a Bayes classifier to classify fire [7]. [31] presents another approach that uses a lookup table to locate fire zones and validate them using temporal variation. This approach uses heuristic characteristics to reduce the likelihood of returning the same outcomes when the input data is changed.

These manually-engineered fire detection algorithms are computationally inexpensive and may readily be implemented on resource-constrained embedded hardware such as the Raspberry Pi with a reasonable performance in terms of frame rates, but they have the disadvantage of requiring human extraction of features from raw fire-scene images. This flaw makes manual feature-engineering time-consuming and quite often inefficient, especially when the dataset contains a large number of images. In light of this, handcrafted vision-based fire detection systems are being replaced by DL-based techniques which have increasingly replaced alternatives with lesser accuracy and a higher rate of false triggering, and perform better across a range of parameters. This improved performance can be credited to the ability of DL models to automatically extract characteristics from raw images of fire scenes. The handmade procedures, on the other hand, need greater attention because the characteristics from the incoming images must be manually extracted. Thus, combining these more accurate vision-based fire detection technologies might result in significantly more potent and effective fire alarm systems when compared with traditional sensor-based techniques.

Deep learning approaches offer the benefit of automatically extracting meaningful features from the data presented, thereby making the overall process more efficient and less operator-oriented, and significantly enhancing the state-of-art (SoA) in image classification and object recognition procedures [24, 47]. More recently, there has been a significant amount of research efforts directed towards the development of lightweight deep learning models targeted towards deployment on resource-constrained edge devices [43]. Typical examples of such lightweight deep learning models can be seen in the areas of automated disease detection [41], image forgery detection [1], vehicle trajectory prediction [21] , and a variety of other real-life applications [43].

For fire detection, a sizeable number of deep learning algorithms have been offered in the technical literature. In their work on forest fire detection, Zhang et al. [48] used a appropriately-tuned pre-trained CNN, referred to as 'AlexNet' [24], to detect fire patches, whereas researchers in [36] proposed a CNN-based technique that used baseline architectures VGG16 [37] and Resnet50 [16]. However, both the research models are inappropriate for in-the-field fire detection applications where low-cost, resource-constrained hardware is usually the only hardware available. Different CNN variations, such as AlexNet [27], SqueezeNet [28], GoogleNet [29], and MobileNetV2 [30] were fine-tuned by Muhammad et al. They utilised Foggia's dataset [12] as the training dataset in works based on [27, 28, 29], whereas the training dataset in [30] was a combination of [10] and [12].

Prominent designs for semantic segmentation is encoder–decoder (ED) model architectures, which have lately been emerging as the most fitting choice. In this types of models, initially the encoder is employed to create a feature map (high-dimensional entity) from the input pictures using convolution and interspersed layers performing the 'pooling' operation. Second, the decoder is supposed to provide a decoding of the provided features (generated by the encoder) and thereby defines a mask for the object of interest [44], which is made up of unpooling and deconvolution layers. Numerous research based on the ED structure to segment different objects have been undertaken as a result of the successful completion of various tasks, such as image segmentation. Zhang et al.[18], for example, established an effective DL model for fire detection in forest scenes, and identification utilizing the popular U-Net and lightweight SqueezeNet DL architectures. Researchers in [2] integrated the ED with a DL- based fire detection model because of its strong performance in forest fire detection and segmentation tasks. Using a limited dataset (only 419 images from the CorsicanFire dataset), this model earned an FM-score of more than 97 percent during the training process, and 91 percent in the test phase. Authors in [6] presented the UUNet deep concatenative architecture as a refinement of U-Net. In this method, binary and multiclass U-Net techniques were integrated. The work in [45] proposed a fire detection method for forest scenes utilizing CNNs and a general meta approach (ensemble learning). In other works, two independent DL-based object detectors, the recently introduced YOLOv5 [20] and high-performance EfficientDet [39], as well as an EfficientNet [38]-based classifier, are integrated to accurately localise and identify fires in varied situations.

3. Proposed Model

Since the previous decade, the majority of research has concentrated on classic feature identification/extraction approaches for fire detection in different settings (forest videos, CCTV footage, etc.). The main drawbacks of such approaches are their lengthy feature engineering process and limited fire detection performance. Such approaches also produce a large amount of false alarms, especially when using shadows, variable illumination, and recolored objects in surveillance. There are numerous deep learning architectures for early fire detection extensively to deal with such challenges. The success of these deep learning techniques can be attributed to their inherently nature, which imparts the capability to learn extremely strong characteristics from raw/unprocessed data automatically. The three well-known processing layers that make up a standard CNN architecture are: 1) a convolution layer, in which different kernels are used on the input data to produce various activation maps (commonly known as feature maps); 2) a layer performing the pooling operation, that selects maximal activation based on a small neighbourhood of activation maps acquired from the preceding layer performing the convolution operation – the purpose is to obtain the dual objectives of dimensionality reduction and translation invariance; and finally 3) a fully-connected (dense) layer, that generates a global representation of the incoming data by modelling high-level information. We have witnessed a number of CNN variants being employed to increase the accuracy of the fire detection operation, and reduce the percentage of false/incorrect alerts, inspired by recent developments in edge computing capabilities and the promise of deep feature extraction [48, 37, 16, 27, 28, 29, 30, 12]. However, due to the heavyweight architecture of these CNN models they are difficult to deploy on low cost hardware. For this purpose, we present a lightweight neural network FireNet-*v2* which while having a significantly low parameter count than its predecessor (i.e. FireNet [19]) maintains its fire detection performance.

1) Motivation for FireNet-*v2*

Model selection has always been an important task for different types of applications. For such cases where minor delays can result in significant human and economic losses, a fast and accurate model is imperative. FireNet [19] was initially devised as an improvement over other available CNN models. The biggest advantage was its shallowness in contrast to prior DL-based fire detection methods which are typically large convolutional neural networks that can detect fire in real time at the rate of atleast 24 fps (or possibly higher). It was also appropriate for mobile and embedded applications, and performed well in real-time (continuous) fire detection applications. The model operates at a quite high frame rate of >24 frames per second on resource-constrained, economically inexpensive, single-board computing platforms (one pertinent example is the Raspberry Pi 3B). The motivation of devising FireNet-*v2* compared to FireNet is its decreased model size both in terms of (i) number of computations required, and (ii) number of trainable parameters, with the advantage of better accuracy. As will be explained in detail in this paper, FireNet-*v2* outperforms the original FireNet while having a reduced parameter count (by approximately half). Concretely, the modifications in our work from FireNet is as follows:

- Number of filters in first, second and third convolutional layer are 15, 20 and 30 respectively whereas in FireNet, it was 16, 32 and 64 respectively – this results in a significant reduction of the trainable parameters.
- Activation funtion used in last convolutional layer and both the inner Dense layers are 'Sigmoid'. In contrast, 'ReLU' was used in FireNet.
- Instead of 70% percent and 30% percent, we chose a 90% percent and 10% split across the train–testing sets – this enabled the proposed FireNet-*v2* model to learn on a larger assortment of the fire data.

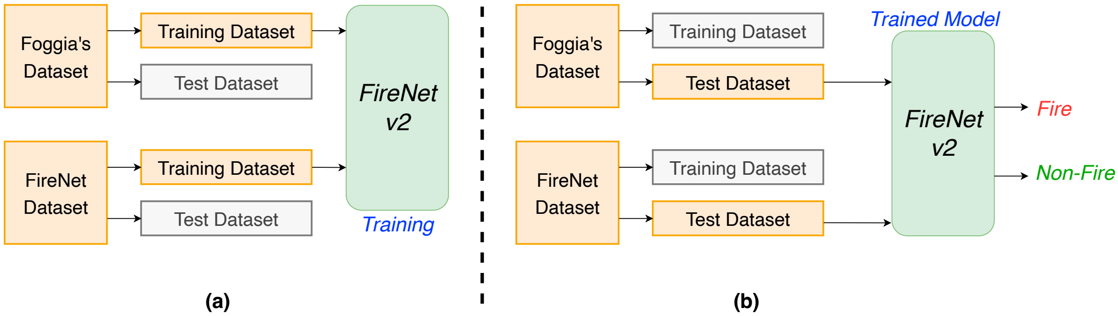


Fig. 1. Training and test phases of the proposed FireNet-*v2*, on the two datasets (Foggia's dataset and FireNet dataset)

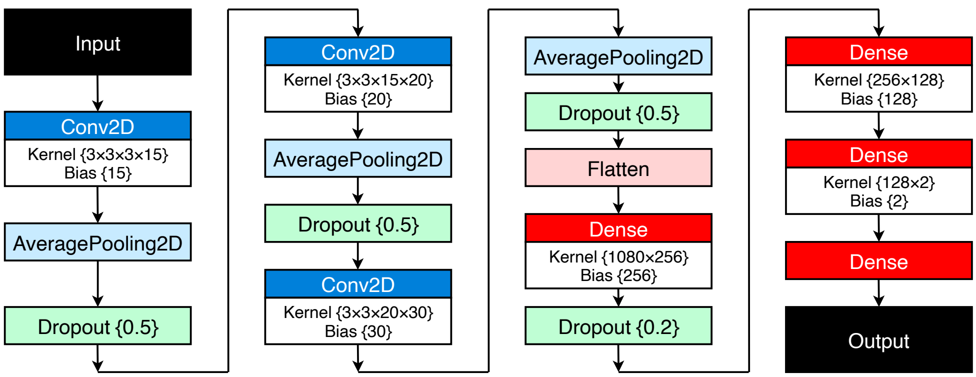


Fig. 2. Architecture of the proposed FireNet-*v2*

2) FireNet-*v2* Architecture

In our proposed model, we used the first (input) layer taking image dimensions of 64×64×3 as first convolution layer. Cascade of convolution, dropout, and average pooling layers are among the intermediate layers with ReLU (rectified linear unit) and Sigmoid being the most used activation functions. All the three convolutional layers are coupled with average pooling. We have used 15 filters in the initial layer, 20 filters in the subsequent (second) layer, 30 filters in the third layer and dropout of value 0.5 with all the three convolutional layers while maintaining a constant kernel size i.e 3×3. Image input features were mapped using convolution layers. The filter dimensions for the convolutional layers are set at [3×3]. A Flatten layer and two Dense layers with 256 and 128 neurons each appear after the convolutional layers, having 'Sigmoid' as activation function. Also, a dropout of value 0.2 is used after the initial Dense layer. The fully-connected Dense layer with the 'Softmax' output has 2 neurons and is the prediction layer outputting 'Fire' and 'Non-Fire' signals (only one of these will be high at any given time). The train and test phases of the approach are presented in Fig. 1, and the complete DL model architecture diagram is shown in Fig. 2 which is comprised of 14 layers. There are 318,460 trainable parameters in total.

4. Results

This section offers information on the dataset utilised, the experiments carried out, and the results.

4.1. Dataset

Although deep learning models for computer vision applications have excelled over the last decade for a considerable number of real-world use-cases, the performance is dependent to a significant extent upon the quality and quantity

Table 1. Specifications of the trained FireNet-*v2* model

(a) Details of Parameters

Parameter	Value
Number of parameters	318,460
Batch size	32
Epoch	100
Validation split	0.1
optimizer	Adam

(b) Activation Functions

Layer	Activation
1st conv layer	Relu
2nd conv layer	Relu
3rd conv layer	Sigmoid
Dense layer	Sigmoid
Output layer	Softmax

of data available for the model training. For fire detection DL models, there is a dearth of high quality datasets. One commonly used dataset in the research works on fire detection models is the Foggia's dataset[1], which contains 31 fire and non-fire videos. From the video clips available in the dataset, we extracted 363 images with fire and 3021 images without fire (non-fire images) for training and testing of FireNet-*v2* model.

Furthermore, the FireNet-*v2* model was also trained and tested on the dataset compiled for FireNet[19], which was a very realistic dataset made up of 46 videos of fire scenes and 16 videos where fire was not in the scene (non-fire). We extracted 1118 images of fire and 1301 images without fire (non-fire images) from the video clips in the dataset used in FireNet[19]. It needs to be mentioned here that although the training images extracted from the FireNet dataset are lesser in number (as compared to the Foggia dataset images), they contains very diverse and realistic images, and are therefore harder to train the model on. This observation shall be validated subsequently in a later section, where it is shown that the accuracy obtained from the FireNet-*v2* model is higher for the samples in Foggia's dataset as compared to the FireNet dataset.

4.2. Training Details

The model is trained over 100 training epochs with a low learning rate to ensure that the majority of previously learned information is preserved in the network. The learning parameters are moderately updated by the pre-trained model in order to get optimal performance on the target dataset. Table 1 lists the many hyperparameters and activation functions we employed for FireNet-*v2*. The choice of the Softmax activation instead of the Sigmoid for the last layer was made because of the fact that the model is designed to have outputs in the form of 'one-hot' encoding i.e. there are 2 outputs and only one of them is high for any given test image (fire/non-fire). The Softmax activation will ensure that the sum of the assigned probabilities for the 2 output classes is equal to unity, and therefore, in order to increase the estimated probability of a particular class, the model is forced to correspondingly decrease the estimated probabilities of the other classes (and vice-versa). Therefore, using this approach, there is a clearer fire/non-fire classification. If a Sigmoid was used, there would be only 1 output, which would have shown the probability that the given image was fire/non-fire. Lastly, since in binary classification both Sigmoid and Softmax functions are essentially the same, there is no 'performance degradation' using Softmax over Sigmoid. We utilised the Adaptive Moment Estimation (Adam) optimizer[2] with a batch size of 32. The most significant benefit is the shallow network and, as a result, the minimal number of trainable parameters (318,460). The configuration of the system used for the training was quite modest with an Intel® Core™ i7-7700HQ CPU@2.80GHz processor coupled with 16 GB RAM and an Nvidia GeForce GTX 1050 GPU running Windows 10 Home Single Language. Model training, and testing, were performed using Python 3.9 and Keras 2.4.0 running in a 64-Bit Anaconda Navigator environment.

[1] https://mivia.unisa.it/datasets/video-analysis-datasets/fire-detection-dataset/

[2] The choice of the Adam optimizer was inspired by the following observation: Adam adds momentum and bias-correction to the conventional RMSprop optimizer. To a certain extent, Adadelta, RMSprop, and Adam are very comparable algorithms with somewhat similar performance in most circumstances. However, the inlusion of bias-correction tends to help the Adam algorithm to marginally outperform RMSprop towards the close of the optimization process (when the gradients tend to get meager). Therefore, Adam may prove be the best overall candidate in most use-cases [33].

Truth		Fire	No Fire
	Fire	314	49
	No Fire	4	3017
		Predicted Label	

Truth		Fire	No Fire
	Fire	553	40
	No Fire	4	274
		Predicted Label	

Fig. 3. Confusion Matrix for the FireNet-*v2* model over Foggia's dataset (left), and FireNet dataset (right)

Fig. 4. Sample results of FireNet-*v2* model over both datasets. The images shown above were correctly classified as either fire or non-fire images.

4.3. Model Performance

The FireNet-*v2* model was trained and tested on the two datasets mentioned above, in separate instances. The predictions are classified into False Positives (FP), True Positives (TP), False Negatives (FN), and True Negatives (TN). In Fig. 3, we have shown confusion matrix of Foggia's dataset and dataset-*v2* respectively. This matrix shows the number of true and false prediction of fire and without-fire images. As can be seen from Fig. 3, FireNet-*v2* performs well on both the datasets, with the percentages of the FNs and FPs both being low. The percentage of FNs is slightly higher in the case of the FireNet-*v2* model being used for FireNet dataset. This is in agreement with the observation mentioned heretofore – the images in FireNet dataset are more challenging than the images contained in Foggia's dataset. The accuracy of DL model for the predictions can be estimated as in (1):

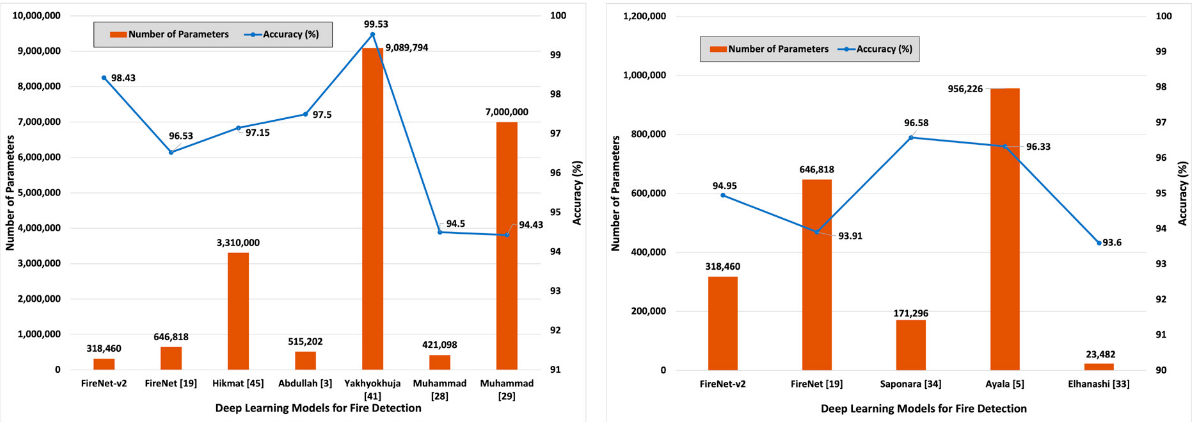
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Fig. 4 presents some samples of images from the two datasets that are correctly classified by FireNet-*v2*. Samples of both 'fire' and 'no-fire' images are included. Similarly, Fig. 5 presents some samples of images from the two datasets that are correctly classified by FireNet-*v2*.

5. Discussion

This section compares the proposed FireNet-*v2* to existing SoA high performance fire identification systems and also the details of the training performed. The most significant benefit is the significantly shallow architecture and, as a result, the minimal count of trainable parameters (318,460). There exist superior performance fire detection methods which are readily accessible in the literature, and this must be acknowledged. However, by virtue of the heavyweight architecture there arise limitations to deploy from the respective technical paper into actual commercial use. Detailed comparative results are presented next.

Fig. 5. Sample of incorrect results from FireNet-*v2* model over both datasets. The images shown above were incorrectly classified.



(a) (b)

Fig. 6. Comparison of accuracy and number of parameters of FireNet-*v2* with available counterparts over: (a) Foggia's dataset; (b) FireNet dataset

5.1. Comparative Results

The results of quantitative comparisons between the available approaches and the one proposed in this work are contained in this section. Table 2 and Table 3 compare the accuracy and number of parameters of the various techniques in both the datasets. The comparison methods that are now available are carefully chosen underlying the dataset, publication year, and features utilised. Existing approaches, for instance, are based on various deep learning architectures with recent literature models. From the given results of Table 2, our accuracy is 98.43%, which is better than most of the counterparts, and is only less than one existing method [42], and has the least number of parameters among all. In fact, although the model in [42] has a better accuracy, it is able to do so at the cost of approximately 9 million parameters. FireNet-*v2* on the other hand, provides an almost similar accuracy with only approximately 300K parameters. Another comparable contender is the model in [46], as they have architect lightweight neural network and also compared their model with famous networks like AlexNet [27], VGG16 [37], ResNet50 [16], MobileNetV1 [17], and NASNetMobile [49]. The suggested model produces more accurate results according to the observations. However, our model has outperformed [46] in terms of (higher) accuracy and (reduced) number of trainable parameters.

The second dataset, FireNet dataset, is smaller but more challenging. There are 871 images in this collection, 593 of which feature fire and 278 of which are fire-resembling images featuring sunsets, lights which appear to be fire-like, and so on. Table 3 shows the outcomes of our technique on this dataset, as well as comparisons to other methods. Also included in Table 3 are all recent publications pertinent to fire detection in which the same data is utilised, to the best of our knowledge [35, 5, 34, 3].

Furthermore, the performance comparison of FireNet-*v2* with the available counterparts, on both the Foggia's dataset and FireNet dataset, is depicted in Fig. 6(a) and Fig. 6(b) respectively. It can be readily observed from Fig. 6(a) that FireNet-*v2* is able to provide a higher accuracy compared to all counterparts (except [42] which provides a better accuracy but at the expense of approximately 28× the parameter count).

Table 2. Comparison of testing accuracy of FireNet-*v2* with other existing works on Foggia's dataset

Works	Accuracy(%)	Number of parameters
FireNet-*v2*	98.43	318,460
FireNet [19]	96.53	646,818
Hikmat [46]	97.15	3.31 million
Abdullah [3]	97.50	515,202
Yakhyokhuja [42]	99.53	9,089,794
Muhammad [28]	94.50	421,098
Muhammad [29]	94.43	7 million

Table 3. Comparison of testing accuracy of FireNet-*v2* model with other existing works on FireNet dataset

Work	Accuracy(%)	Number of parameters
FireNet-*v2*	94.95	318,460
FireNet [19]	93.91	646,818
Saponara [35]	96.58	>171,296
Ayala [5]	96.33	956,226
Elhanashi [34]	93.60	23,482

6. Conclusion

This paper presented a lightweight deep learning model for categorising fire and without-fire images. The proposed FireNet-*v2* is an improved version of FireNet [19], and contains a significantly lesser number of trainable parameters as compared to its predecessor. The prposed model was tested on two different datasets (Foggia's dataset and FireNet dataset), and the prediction accuracy of the proposed model was found to be better than existing low-parameter-count models. For instance, FireNet-*v2* provided an accuracy of 98.43% using only 318,460 parameters as compared to 96.53% accuracy of FireNet[19] using 646,818 parameters.

References

[1] Abbas, M.N., Ansari, M.S., Asghar, M.N., Kanwal, N., O'Neill, T., Lee, B., 2021. Lightweight deep learning model for detection of copy-move image forgery with post-processed attacks, in: 2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI), IEEE. pp. 000125–000130.

[2] Akhloufi, M.A.Tokime, Elassady, R., H, . Wildland fires detection and segmentation using deep learning. pattern recognition and tracking. xxix. Int. Soc. Opt. Photonics Proc. SPIE 2018, , 10649, 106490B.

[3] Altowaijri, A.H., Alfaifi, M.S., Alshawi, T.A., Alshebeil, S.A., . A privacy-preserving iot-based fire detector. IEEE Access , 99.

[4] A.Sorbara, E.Zereik, M.Bibuli, G.Bruzzone, M.Caccia, 2015. Low cost optronic obstacle detection sensor for unmanned surface vehicles. Sensors Applications Symposium , 1–6.

[5] Ayala, A., Lima, E., Fernandes, B., Bezerra, B.L., Cruz, F., 2019. Lightweight and efficient octave convolutional neural network for fire recognition , 1–6.

[6] Bochkov, Kataeva, V., L.Y., . wuunet: Advanced fully convolutional neural network for multiclass fire segmentation. Symmetry 2021 .

[7] Borges, P., Izquierdo, E., 2010. A probabilistic approach for vision-based fire detection in videos," ieee transactions on circuits and systems for video technology. IEEE transactions on reliability 20, 721–731.

[8] Çelik, T., Özkaramanlı, H., Demirel, H., 2007. Fire and smoke detection without sensors: Image processing based approach , 1794–1798.

[9] Chen, T.H., Wu, P.H., Chiou, Y.C., 2004. "an early fire-detection method based on image processing". International Conference on Image Processing, 2004 3, 1707–1710.

[10] Chino, D.Y., Avalhais, L.P., Rodrigues, J.F., Traina, A.J., 2015. Bowfire: detection of fire in still images by integrating pixel color and texture analysis, in: 2015 28th SIBGRAPI conference on graphics, patterns and images, IEEE. pp. 95–102.

[11] Dimitropoulos, K., Barmpoutis, P., Grammalidis, N., 2015. Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection,. IEEE transactions on circuits and systems for video technology 25, 339–351.

[12] Foggia, P., Saggese, A., Vento, M., 2015. Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion. IEEE Transactions on Circuits and Systems for Video Technology 25, 1545–1556.

[13] Gagliardi, A.Saponara, S., . Distributed video antifire surveillance system based on iot embedded computing nodes. Springer LNEE 627 , 405—411.

[14] G.Marbach, M.Loepfe, T.Brupbacher, 2006. An image processing technique for fire detection in video images. Fire safety journal 41, 285–289.

[15] Hall, J.R., 2014. The total cost of fire in the united states. National Fire Protection Association, Quincy .

[16] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. in Proceedings of the IEEE conference on computer vision and pattern recognition, , 770–778.

[17] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 .

[18] J., Z., Zhu, H., Wang, P., X., L., 2021. Att squeeze u-net: A lightweight network for forest fire detection and recognition. IEEE Access 2021 .

[19] Jadon, A., Omama, M., Varshney, A., Ansari, M.S., Sharma, R., 2019. Firenet: a specialized lightweight fire & smoke detection model for real-time iot applications. arXiv preprint arXiv:1905.11922 .

[20] Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, Kwon, Y., TaoXie, Fang, J., imyhxy, Michael, K., Lorna, V, A., Montes, D., Nadar, J., Laughing, tkianai, yxNONG, Skalski, P., Wang, Z., Hogan, A., Fati, C., Mammana, L., AlexWang1900, Patel, D., Yiwei, D., You, F., Hajek, J., Diaconu, L., Minh, M.T., 2022. ultralytics/yolov5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference. URL: https://doi.org/10.5281/zenodo.6222936, doi:10.5281/zenodo.6222936.

[21] Katariya, V., Baharani, M., Morris, N., Shoghli, O., Tabkhi, H., 2022. Deeptrack: Lightweight deep learning for vehicle trajectory prediction in highways. IEEE Transactions on Intelligent Transportation Systems .

[22] Ko, B.C., Cheong, K.H., Nam, J.Y., 2009. Fire detection based on vision sensor and support vector machines. Fire Safety Journal 44, 322–329.

[23] Ko, B.C., Ham, S.J., Nam, J.Y., 2011. Modeling and formalization of fuzzy finite automata for detection of irregular fire flames. IEEE Transactions on Circuits and Systems for Video Technology 21, 1903–1912.

[24] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25.

[25] LeCun, Y., Bengio, Y., Hinton, G., . Deep learning. Nature 521(7553) , 436—444.

[26] Mehmood, I., Sajjad, M., Baik, S.W., 2014. Mobile-cloud assisted video summarization framework for efficient management of remote sensing data generated by wireless capsule sensors. Sensors 14, 17112–17145.

[27] Muhammad, K., Ahmad, J., Baik, S.W., 2018a. Early fire detection using convolutional neural networks during surveillance for effective disaster management. Neurocomputing 288, 30–42.

[28] Muhammad, K., Ahmad, J., Lv, Z., Bellavista, P., Yang, P., Baik, S.W., 2018b. Efficient deep cnn-based fire detection and localization in video surveillance applications. IEEE Transactions on Systems, Man, and Cybernetics: Systems 49, 1419–1434.

[29] Muhammad, K., Ahmad, J., Mehmood, I., Rho, S., Baik, S.W., 2018c. Convolutional neural networks based fire detection in surveillance videos. IEEE Access 6, 18174–18183.

[30] Muhammad, K., Khan, S., Elhoseny, M., Ahmed, S.H., Baik, S.W., 2019. Efficient fire detection for uncertain surveillance environment. IEEE Transactions on Industrial Informatics, .

[31] Phillips Iii, W., Shah, M., da Vitoria Lobo, N., 2002. Flame recognition in video. Pattern recognition letters 23, 319–327.

[32] Rafiee, A., Dianat, R., Jamshidi, M., Tavakoli, R., Abbaspour, S., 2011. Fire and smoke detection using wavelet analysis and disorder characteristics, in: 2011 3rd International conference on computer research and development, IEEE. pp. 262–265.

[33] Ruder, S., 2016. An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747 .

[34] Saponara, S., Elhanashi, A., Gagliardi, A., 2020. Exploiting r-cnn for video smoke/fire sensing in antifire surveillance indoor and outdoor systems for smart cities, in: 2020 IEEE International Conference on Smart Computing (SMARTCOMP), IEEE. pp. 392–397.

[35] Saponara, S., Elhanashi, A., Gagliardi, A., 2021. Real-time video fire/smoke detection based on cnn in antifire surveillance systems. Journal of Real-Time Image Processing 18, 889–900.

[36] Sharma, J., Granmo, O.C., Goodwin, M., Fidje, J.T., 2017. Deep convolutional neural networks for fire detection in images,. in International Conference on Engineering Applications of Neural Networks , 183–193.

[37] Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 .

[38] Tan, M., Le, Q., 2019. Efficientnet: Rethinking model scaling for convolutional neural networks , 6105–6114.

[39] Tan, M., Pang, R., Le, Q.V., 2020. Efficientdet: Scalable and efficient object detection, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 10781–10790.

[40] T.Celik, H.Demirel, 2009. Fire detection in video sequences using a generic color model. Fire Safety Journal 44, 147–158.

[41] Tiwari, S., Jain, A., 2022. A lightweight capsule network architecture for detection of covid-19 from lung ct scans. International Journal of Imaging Systems and Technology 32, 419–434.

[42] Valikhujaev, Y., Abdusalomov, A., Cho, Y.I., 2020. Automatic fire and smoke detection method for surveillance systems based on dilated cnns. Atmosphere 11, 1241.

[43] Wang, C.H., Huang, K.Y., Yao, Y., Chen, J.C., Shuai, H.H., Cheng, W.H., 2022. Lightweight deep learning: An overview. IEEE Consumer Electronics Magazine .

[44] Xing, Zhong, Y., Zhong, L., X., 2020. An encoder-decoder network based fcn architecture for semantic segmentation. Wirel. Commun. Mob. Comput. 2020 .

[45] Xu, R., Lin, H., Lu, K., Cao, L., Liu, Y., 2021. A forest fire detection system based on ensemble learning. Forests 12, 217.

[46] Yar, H., Hussain, T., Khan, Z.A., Koundal, D., Lee, M.Y., Baik, S.W., 2021. Vision sensor-based real-time fire detection in resource-constrained iot environments. Computational intelligence and neuroscience 2021.

[47] Zaidi, S.S.A., Ansari, M.S., Aslam, A., Kanwal, N., Asghar, M., Lee, B., 2022. A survey of modern deep learning based object detection models. Digital Signal Processing , 103514.

[48] Zhang, Q., Xu, J., Xu, L., Guo, H., 2016. Deep convolutional neural networks for forest fire detection .

[49] Zoph, B., Vasudevan, V., Shlens, J., Le, Q.V., 2018. Learning transferable architectures for scalable image recognition , 8697–8710.