Detection of Anomalous Fire using Deep Learning Techniques

Ashish Gupta[1]
[1]Associate Professor, Computer Science and Engineering, Dev Bhoomi Uttarakhand University Dehradun, India, Email:[1]ashishgupta.wit@gmail.com
Gunjan Bhatnagar[2*]
[2]Assistant Professor, Computer Science and Engineering, Dev Bhoomi Uttarakhand University Dehradun, India, Email:[2]csegunjan03@gmail.com[*](Corresponding Author)
Ranjit Kumar[3]
[3]Department of Computer Science, Maharaja Agrasen University, Baddi, Himachal Pradesh, India, Email:[3]ranjitpes@gmail.com, Orcid ID: 0000-0003-3001-7404
Umang Garg[4]
[4]Assistant Professor, Computer Science and Engineering, Graphic Era Hill University, Dehradun, India, Email:[4]umangarg@gmail.com
Dr.Atul Kumar[5]
[5]Associate Professor, Computer Science and Engineering, Shri Ramswaroop Memorial College of Engineering and Management Lucknow,India, Email:[5]atulverma16@gmail.com
Neeraj Panwar[6]
[6]Assistant Professor, School of Computing, Graphic Era Hill University, Dehradun, India, Email:[6]neeraj.pan28@gmail.com

ABSTRACT

One of the most damaging anomalous occurrences is fire. Both human lives and property are severely damaged and destroyed by it. Deep Learning has recently demonstrated promising outcomes in a number of classification and detection research projects. A better technique is still needed for effective fire and smoke detection, though. It is largely because of its ambiguous shape, texture, and potential for appearing in several shapes in day-to-day existence. In order to identify smoke and fire in films or photos, a novel, lightweight, real-time convolutional neural network is proposed in this study. Exiting datasets are constrained or artificially produced for testing.The validation process for this study's challenging planned dataset, which includes the majority of fire and smoke event scenarios, has been completed. An outcome that incorporates bounding box localization of the fire and smoke regions has been obtained on a highly diversified as well as newly introduced and targeted early fire detection image dataset. In comparison to RetinaNet, MobileNet, InceptionNet, and FireNet on the available dataset, the suggested approaches with modified EfficientDet obtained roughly 80% in terms of Average Precision (AP). Future research in this field may benefit from the proposed dataset, and implementing the evaluation by the suggested method to a real-world scenario would be beneficial.

Keywords: Anomalous, Convolutional Neural Network, Fire Detection.
DOI Number:10.14704/nq.2022.20.8.NQ44595 **NeuroQuantology 2022; 20(8): 5674-5689**

1.INTRODUCTION

In many sectors and residential surveillance systems, fire and smoke detection solutions have become essential. It was brought about by a sharp rise in fire accidents. One of the most hazardous natural disasters, fire can swiftly result in serious injury and property destruction. In 2017, 1,319,500 fire events were registered, resulting in 18,000 civilian injuries and almost $23 billion in property losses, according to research by the US Fire Department [1]. Early or real-time fire detection has become extremely important in surveillance systems in order to reduce or avert such accidents. Fire detection not restricted to useful in context of disaster management systems like forests, nuclear plants, but also keep a watch on ecological damages, economic and social damages. As the demand for fire detections are huge across multiple industries, thus there have been various hardware-based fire detection technologies necessitates.

Existing fire alarm systems have been proven to be inefficient in terms of various real-world scenarios. Most widely used fire alarm systems have been based on physical hardware sensors like thermal detectors, smoke detectors and heat/flame detectors. The major drawback of sensor-based system is that, it must be in near proximity to the fire or heat area. However, this makes them inefficient to use in various commonly occurring scenario such as long-distance fire occurrence as shown in Figure 1. Due to this, conventional method has failed to prevent a lot of fire accidents. The solutions to this mostly require the sufficient amount of fire or heat sensations to trigger off the alarm. Moreover, it fails to accurately locate the fire or smoke regions.

Figure 1: Sample images from training data for fire class

Figure 2: Sample images from normal class. These are confusing images which looks like fire or smoke

Due to limitations in detection of fire, researchers have been investigating Computer Vision based methods became alternative to enhance the fire and smoke detection system. Fire is one of the forms of anomaly event. Under network security applications, network anomaly detection plays an important role [2]. Existing vision-based methods were solely focused on colour space transformation for fire region detection [3, 4]. Rule-based approach along with colour space has promising future in providing a better performance. However, such systems are often sensitive to other illuminated objects like street lights. Further methods added to additional features to colour-based methods like area, boundary and motion cues [5, 6] to the decision-making algorithm. These methods have been used a classifiers like Bayes classifier, dual optical flow and multi-expert system, in order to reduce false detection or miss classification. Nevertheless, these solutions are prone to error and fails in multiple complex real-world scenarios as shown in Figure 2. Thus, fire detection is a challenging problem due to the complex nature of the problem. As it does not have definitive shape, region of occurrence, dynamic temporal behaviour, so that to extract feature. The hand-crafted feature selection requires a large amount of domain knowledge.

Recently Deep Learning based methods has shown tremendous advancement in large scale image classification task. This solution eliminates the role of hand-crafted feature selection. The capability of the deep learning-based approaches, to extract features automatically from the raw images. Exploration of high-dimensional features is transformed into training the network using sufficient amount of training data in order to avoid overfitting.

Exiting Deep Learning based methods largely focuses on classification methods [7, 8, 9]. These methods fail to locate the fire regions. Moreover, these solutions are computationally heavy to makes them difficult to make use for real-time fire and smoke detection system, unless and until use a special hardware.

Accuracy-Efficiency trade-off can be observed through Figure 3. Figure 3 shows the relationship between accuracy and operations (G-ops). For high accuracy algorithms (around 80 %) operations would be less i.e., around 20. In contrast, it can be observed that for high operations around 40, accuracy would be around 70 %.

The depth scaling is proven winner to improve model accuracy such as GooleNet (6.8M), SENet (145M), and Gpipe (557M). In other end, efficiency reduces the model size by trading the accuracy. The Handcraft small models such as, SquezeNet, MobileNet etc. The model compression and pruning methods can be used for drawing these results. Recently, NAS based methods are getting more popular for small network search, but difficult for larger networks.

Different types of Model Scaling play a vital role in performance evaluation. The model scaling is a method for improving the performance of CNN model. Typically, model scaling was done in any one of the dimensions from below mentioned three (refer Figure 4). Intuitively, all the three scaling dimensions follow some relationship [4, 5].

a) Increasing the #layers (ResNet 18 – ResNet 200)

b) Increasing the #channels (WideResNet)

c) Increasing the input resolution

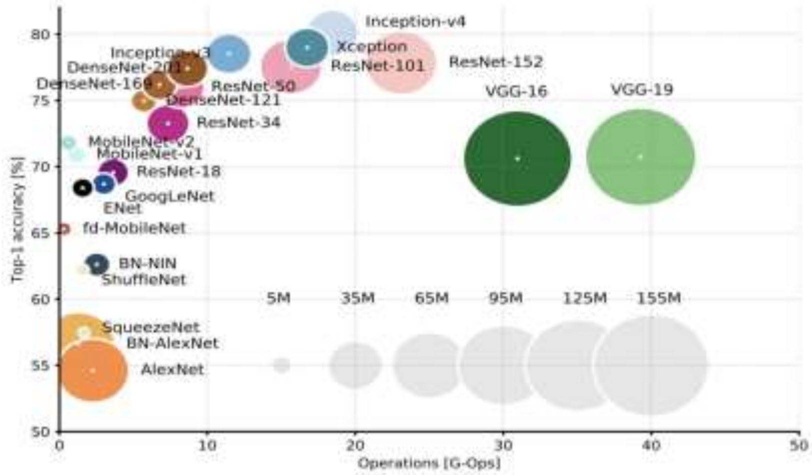


Figure 3: Top 1 Accuracy on ImageNet Dataset [10]

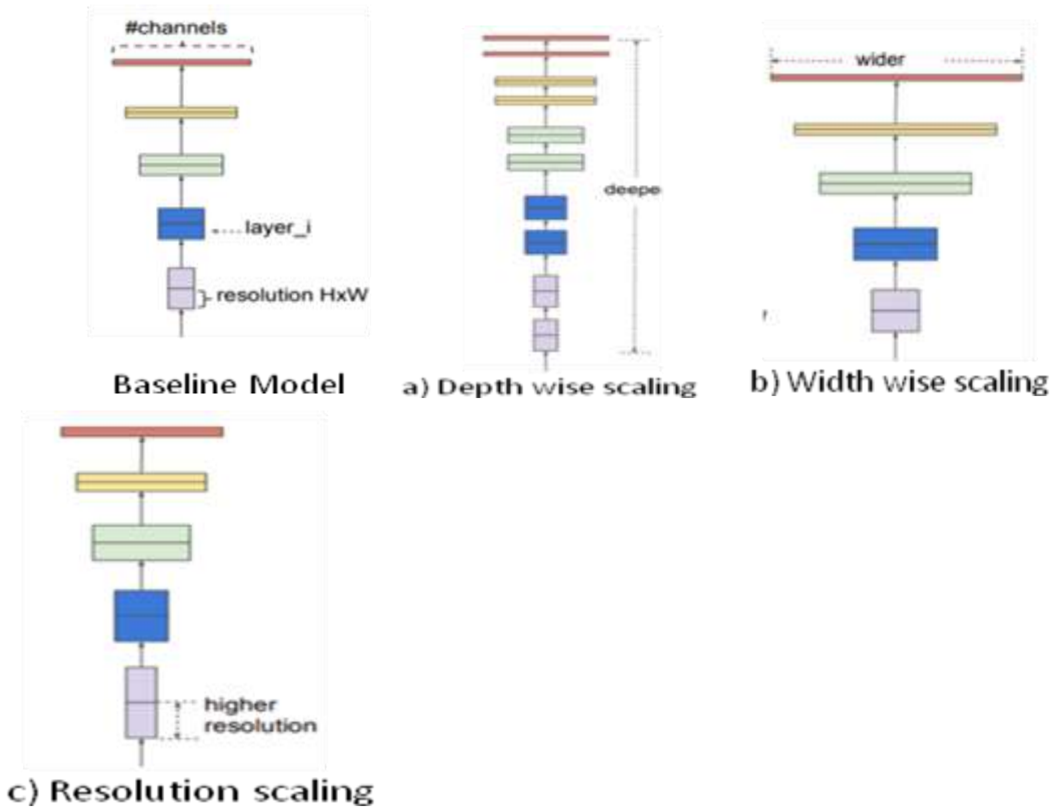


Figure 4: Types of Model Scaling [11]

5677

2.

LITERATURE REVIEW

Detecting fire and smoke is one of the important forms of anomaly detection in videos. There is a dire need to use of modern technological solution for early detection of fire. Due to which, the tremendous damages caused can be reduced. In recent years, research in this field of direction has been exponentially increased.

Earlier literature indicates the fire detection using only vision cameras falls under three main categories. These are pixel-level, blob-level and patch methods. Pixel-level methods uses colour and flickers [3] to detect fire in the scene. But these methods are highly biased as well as it performs inaccurately in real-world scenario. In contrast, blob-level method shows better performance than pixel-level method, as these methods uses blobs as features [3] to detect fire in the scene. The major problem with these methods lies in difficulty to train a classifier as the shape of the fire blob is a lot diverse in nature. Patch-level methods [5] perform better than other two categories but they are very expensive methods. Lastly, recent techniques, can be divided into two main categories, Computer-vision based approaches and Deep Learning based approaches. Presently fire detection method has also been observed in IoT industrial applications [12], in controlled robot [13]. Idris et al. [14] had implemented fire recognition along with alert system in LabVIEW.

Conventional methods focus on finding the salient features in the image using colour and brightness in the image. Celik et al [3] used a simple rule-based approach of frame differencing between consecutive frames. They used RGB and HSV (hue, saturation, value)colour models for analyzing the patterns to detect flames. In Ko et al [4], they used similar approach by converting images into YCrCb domain to analyze luminance and chrominance model separately [8]. Another colour-based model is explored by Wang et al [6] using HIS (hue, intensity, saturation) colour model to calculate dispersion of the fire area.

The colour-space based methods are highly vulnerable to variety of real-world natural setting. These methods fail to differentiate between similar colour patches which are very common to occur like artificial lighting, shadow and similar coloured objects. The quality of results highly depended on the limited predefined features to make use for classification. This makes hand-engineering task more tedious and inefficient to work in real-world scenarios.

Deep Learning based approaches eliminates the manual intervention in the process of feature extraction. Deep learning techniques have tremendous improvements in many complex vision-based tasks like large scale image classification. Use of deep learning, in the research field of fire and smoke detection are still at its nascent stages.

In recent year, a handful of deep learning methods for fire classification along with simultaneously in detection have been published in literature. The most of the work has been published by the image classification networks using transfer learning and fine-tuning the network. Khan et al [15] fine-tuned AlexNet CNN modified it for fire classification problems by identifying fire patches during training the model. Later, these learned patches are post-processed to accurately localize fire regions and remove false patches. Muhammad et al. [15-17] proposed multiple fine-tuned CNNs like AlexNet, GoogleNet, and SqueezeNet for the similar problem. In their research study, applied Foggia et. al [18] on training videos for training and testing these fine-tuned models. The given dataset is very limited in number and since this dataset consists of long videos which makes it less diverse. Training with such dataset might make network over fitted on such repetitive scene properties, resulting poor performance on real-world diverse dataset. These networks have larger memory footprint as they have large number of trained parameters [19, 20]. Moreover, these networks are computationally heavy for real-time usage. In order to address such limitation Jadon et al [8] proposed a FireNet, which is a novel light-weight CNN architecture. In this study, they have improved and extended the training data by web scrapping images and videos. However, this network is small sized classification network. It fails to localize the region of fire in the scene. In current era, detection and classification would be made using CNN and Haar classifier [24].

Baseline fine-tuned methods like AlexNet, VGG16, GoogleNet or Resnet are computationally heavy methods which makes it difficult for real-time usage. These methods require incredibly large amount of training data because they have large number of trained parameters.

Latest method EfficientNet became popular in market because of its better efficiency compared to earlier used backbones [21]. It's noticeable model efficiency in presence of network scaling due to width, depth and resolution [11]. The feature network design [22] had introduced to fuse multi-scale features at different levels. The mobile-size models [23] used as a baseline which later scale up to obtain better accuracy. Further, the weighted bi-directional feature pyramid network (BiFPN) and compound scaling consistently achieve better efficiency in EfficientDet [25]. The scaling method works well in MobileNets baseline network strengthening, so that applied in neural architecture search [23, 26, 27]. Learning object detection capability achieved as a outperforming best detector by [28]. Howard et al. [29] proposed a new MobileNetarchitechture. This architechture based on deapthwise separable convolution. A dense-MobileNet models have been introduced as a novel image classification approach.

To test the efficiency of loss in images, Lin et al. [30] created and trained RetinaNet. This network found to be a simple dense detector. RetinaNet was also applied in evaluation on blind and visually impaired persons navigation assistance. Wu et al. [30] has observed that, feature pyramid network (FPN) which is a RetinaNet's backbone network. It is a more adaptive neural network architecture that can recognise multiresolution feature objects.

Despite of Deep Learning performing extremely well in image classification tasks, but object localization was certainly a challenging problem. In recent, various object detection techniques like RCNN (Region based CNN) family largely solved the problem of learning a region proposal which likely has objects and later on fitting a regressing network for localizing the bounding box detection. This technique is accurate but computationally heavy. Later, single pipeline for localizing and classification has been proposed in YOLO (You Only Look Once) and SSD (Single Shot Detector) family of networks. In their study, for detection and classification, backbone network remains same. They divide the image into logical grids and tried to localize the object in all the given grids with pre-defined anchor boxes.

Gonzalez et al. [29] used 2-stream CNN network for fire and smoke classification. The detection method and classification are combined with a single stage de-convolutional layer, which results in a map like feature space. However, the results are tested on fewer dataset.

Recently, Kim and Lee [30] proposed a detection-based system for fire and smoke also with classification. In this study, they had used spatial and temporal features combined with logical voting for fire and smoke detection. Faster RCNN was used as feature extractor while using the backbone as ResNet 101. Further, a LSTM network is trained by combining learned features to understand long-term dynamics of fire and smoke regions. At the end, to remove remaining false positive a logical voting mechanism is used to finally classify them into fire and smoke region in the image [31, 32]. However, FRCNN (Faster Region based CNN) and RNN (Recurrent Neural Network) are both computationally expensive methods to be used for real-time usage.

The existing detection methods are not real-time in nature as FRCNN with backbone ResNet are computationally heavy detector. However, in practical scenario, fire and smoke detection both would run in real-time or pseudo real time system. The inherent assumption of training LSTM is always depending on the notion of the time. It fails when the testing video does not inherit the similar time frame of fire and smoke. The logical voting is heuristically determined and has a hardcoded time for voting. There has been a lack of diversified real-world fire localization test dataset. Thus, to address the issues in this research field, in current study investigated a Convolutional neural network (CNN) based deep features for early fire detection in surveillance networks. Therefore, this paper mainly focuses on, the lack of deep learning-based fire and smoke localization methods, the lack of well diversified dataset for training a novel CNN network, and the lack of real-time CNN methods for early fire and smoke detection. Further this study has been investigated towards reduction in easy confusion with bright object such as; sun, lamps, shadow and bright regions, so that result accuracy would be increased.

3. EFFICIENTDET PROBLEM FORMULATION

The current research study used the capabilities of recently proposed Neural Network Architecture Search techniques for identifying a novel CNN architecture for fire detection. Proposed method changes the existing detection architecture and form a lighter version of the same without affecting and compromising the quality of results. A less accuracy supervised method necessary to train fire region attention network. This would generate Deep Learning based synthetic images blending for data augmentation. Later, would start with lightweight detector such as tiny-Yolo with pre-trained features.

The study mainly focused on extracting deep attention feature and near confusing features. While studying these features using dimensionality reduction methods, possibility this helps in their relevance to use and finding correlation between them. Indirectly, this will help to choose different loss function to penalize false positives.

Evaluation of such systems needs to be done only on real-world scenario. Exiting datasets has been limited or synthetically generated for testing. For validation of results, a real-world challenging dataset has been required which covers most of the scenarios of fire events. Second, objective of our research is to generate a highly diverse dataset which includes bounding box localization of the fire regions. This proposed dataset could be an asset for future research in this domain. For quantitatively validating the result metrics like Precision-Recall curves and Mean Average Precision (MAP) has been used.

As discussed in previous sections that this study has been started with Deep Learning based approach. A typical Convolutional neural network layer consists of 3 main operations. The convolutional filter, activation function and pooling operation. Convolutional filter varies in its shape and number of filters based on the receptive field dimension defined by network architecture. Activation functions add nonlinearity to the system. Most widely used activation function is Rectified Linear Unit (Relu). Pooling operation captures high responsive regions in the given layer. Thus, it helps to reduce the dimensionality and introduces translational and scale invariance. These three operations called a layer, are stacked on top of each other to form a complete network architecture. The stacking arrangement is done in such a way that output of a given layer is fed to another layer in the network. After the stacks of convolutional layers few fully connected layers are used to as a classification layer to understand high-level abstractions of the training data. At the top layer the final classification or detection heads are placed for desired results. The weights of the network are randomly initialized. The network is trained will large amount of data in order to update weights automatically. The weights of the network are updated by universal back-propagation method. A standard back-propagation method uses stochastic gradient decent algorithm to dynamically update the network weights. A suitable loss function like Binary cross entropy is obvious choice for a binary classification problem.

3.1 Network modification details

Many network considered either Hardswish Activation or Leaky Relu function. In present study,
Hardswish Activation has been replaced with Leaky Relu. The Leaky Relu function can be represented by,

$$f(x)=\begin{cases}0.1x & for \quad x<0\\ x & for \quad x\geq 0\end{cases} \tag{1}$$

Further, Hardswish Activation function can be represented by, $f(x)=\frac{x}{1+e^{-x}} \quad for \quad 0\leq x$ (2)

In this case, constant sharp curve of Leaky Relu function advantageous in classification work which mainly helps in performance improvement. It has been observed that while adding a Leaky Relu helps to separate the dominant color from the background. The results have been showed with the help of anchor box technique. The anchor boxes were chosen based on weighted k-means clustering. Further, it was chosen dynamically based on the given batch size. Weighted K-means clustering can be represented as distance of $P^{(t)}$ with set of observations as (y1, y2, …………….yn),

$$P_i^{(t)}=\left\{y_p:\left\|y_p-n_i^t\right\|^2\leq\left\|y_p-n_j^t\right\|^2\ \forall i,j,1\leq i,j\leq q\right\} \tag{3}$$

where, every yp mapped to distance of P(t) and nit step modification has been mentioned below. The modified step of K-means is ni(1) ,……,nk(1) given by,

$$n_i^{(t+1)}=\frac{1}{\left|P_i^{(t)}\right|}\sum_{y_j\in P_i^{(t)}}y_j \tag{4}$$

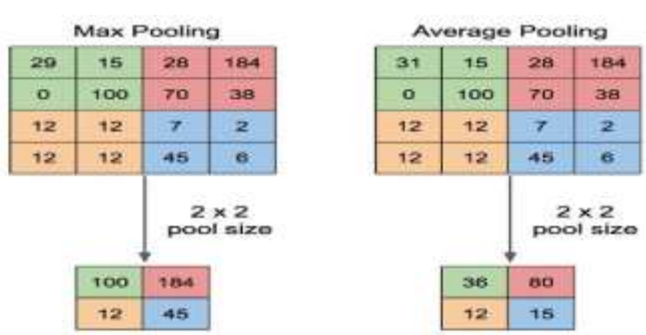
There are two common functions used in the pooling operation, average and maximum, shown in Figure 5.

Average Pooling: Calculate the average value for each patch on the feature map.

$$k'_{2\times 1}=AVG\left(\sum_{c=1}^{4}p_n.q_c\right) \tag{5}$$

Maximum Pooling (or Max Pooling): Calculate the maximum value for each patch of the feature map.

$$k^1_{2\times 2}=MAX\left(\sum_{n=1}^{4}p_n.q_n\right) \tag{6}$$



a) b)

Figure 5:a) Max pooling, b) Average pooling

For networks to generalize for the colour, in this study replaced max-pooling with average pooling. Also, the global average pooling is being used at the input layers to make the network more robust.

3.2 Training modifications details

A small resolution centric hyper-parameters training handled in sophisticated way. The variety of augmentation mainly imposed for object detection i.e. bounding box transform, color space etc.. The generated augmentation either for the lower resolution or for optimized the resolution. The batch size certainly was increased for better normalization. The learning rate schedule was changed to drop early, based on the loss calculation. Further, performance improvement it may require to downsize the input layer of the network. Therefore, maintaining the aspect ratio will be a key task in whole training modification process.

4. EXPERIMENTAL ANALYSIS

The experiments have been conducted using a system with the specifications: NVidia RTX 2080) with 10 GB onboard memory using a deep learning framework and Ubuntu OS16.04 installed on an Intel Core i5 CPU with 64 GB RAM. A total of 68,457 images were used in the experiments; these were obtained from well-known fire datasets including those of Foggia et al. [17] with 62,690 frames. The training and testing phases of the experiments, has followed the experimental strategy of [17], where 20% and 80% of the data are used for training and testing, respectively. This strategy has been implemented in this study with trained proposed modified EfficientDet model having with 5,000 fire images and 5,061 non-fire images, then resulting in a training dataset of total 10,319 images. Therefore, the proposed network with only 2-classes i.e., fire and not fire class. Data sets is one of the important components to evaluate the performance of any given system. Evaluating the proposed algorithm against a standard dataset is one of the challenging tasks. In proposed datasets, all the images are unique and captured by real humans. Thus, this dataset is the most challenging and diversified dataset ever created. This hand-crafted test dataset was prepared to understand the generalization of the trained model. These include average 2 boxes per image of different scale and aspect ratio.

While doing analysis on this dataset, loss function has been used as binary cross-entropy. Further the optimizer has considering as RMSProp with starting learning rate of 0.001s. The number of epochs 300 has taken into account. The details of the results using the different fire datasets as well as their comparison with state-of-the-art techniques, has been introduced in subsequent sections.

5. RESULTS AND DISCUSSIONS

In this section, four different networks have been compared with EfficientDet with the help of Average precision (AP) indicator. InceptionNet [7], FireNet [8], MobileNet [29], and RetinaNet [30] these four networks are versatile networks chosen for comparative study.

In this study, verified the loss function (FL) proposed by Lin et.al [32] on our proposed early fire and smoke dataset, shown in Figure 6. This FL was evaluated by below Eq. (7). This introduces FL starting from the cross entropy (CE) loss for binary classification.

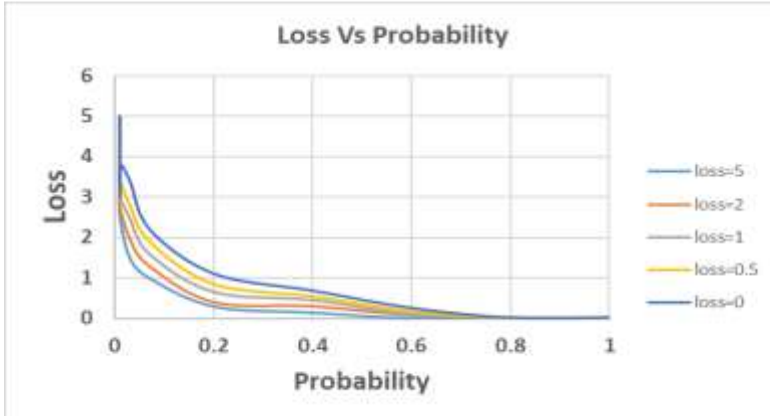


Figure 6:Loss function on proposed dataset

Khan et. al. [7] and FireNet [8] presented methods based on classification used leave-one-out approach to get for each class. In contrast to these algorithms, modified EfficientDet more focused on detection level. In this study, Average precision (AP) indicator has been considered for quantitative analysis. Results have been obtained and listed in Table 1 for proposed dataset compared with Foggia dataset. Activation mapping has been exploited, to get the approximate bounding box.

Table 1 Comparison of modified EfficientDet with another algorithm

Method	Early Fire and Smoke (proposed)		Foggia datas et	
	AP@50	AP@75	AP@50	AP@75
Khan et. al [7] InceptionNet	53.41	50.63	65.23	61.28
FireNet [8]	68.46	57.94	73.23	70.65
MobileNet[29]	65.45	53.67	68.87	63.98
RetinaNet[30]	67.89	63.48	70.12	66.78
Modified EfficientDet	73.35	70.78	81.92	78.23

Table 1 indicates modified EfficientDet results are improved compared to other algorithms. At present, AP at 50 and AP at 75 both have been compared. Modified EfficientDet result obtained considering proposed Early Fire dataset, have been around 73 % and 71 % comparable to around 68 %, 63 % for RetinaNet. Followed by, around 68 %, 58 % and around 66 %, 54 % for FireNet and MobileNet, respectively. In contrast, while considering Foggia dataset, obtained results have been around 82 %, 78 % comparable to around 70 %, 67 % for RetinaNet. While, around 69 %, 64 %, around 65 %, 61 % and around 73 %, 71 % for MobileNet, InceptionNet and FireNet respectively. On both dataset among all algorithms, modified EfficientDet well performed while, InceptionNet down performed.

5.1 Accuracy Improvement

The main challenge in obtained results would be the accuracy on lower resolution i.e. 384x384, which significantly affect. The performance gains on fine-tunning on higher resolution (640x640) were not directly proportional to lower resolution (384x384) due to lowered accuracy. So, it's necessary for optimal Non-Maximum Suppression (NMS) threshold. Since the detection boxes were in arbitrary aspect ratio, and hence choosing the NMS threshold plays critical role. This step helped in removing duplicate boxes. The used PR relationship with respect to confidence in Precision-Recall curves (PRC) curve analysis, to choose optimal NMS and Confidence threshold.

5.2 Performance Improvement

Most important task to improve the performance had been fixed with an input resolution to 384x384. Although, keeping aspects ratio same throughout the batch network. Thus, obtained results shown in Table 2 provided with timing details of EfficientDet pipeline with batch no. 5. All the numbers in Table 2 are in millisecond. The pipeline modification has been handled effectively in different means. One, resizing very essential first step by keeping in mind, aspect ratio should be maintained constant then followed by normalization. Second, floating operations should convert Float32 to Float16 operation; so that computation time would be fast. Third, batch-wise non-maximal separation should be kept around 0.4.

Table 2 Timing details of EfficientDet pipeline with batch=5 on Nvidia RTX 2080 Ti

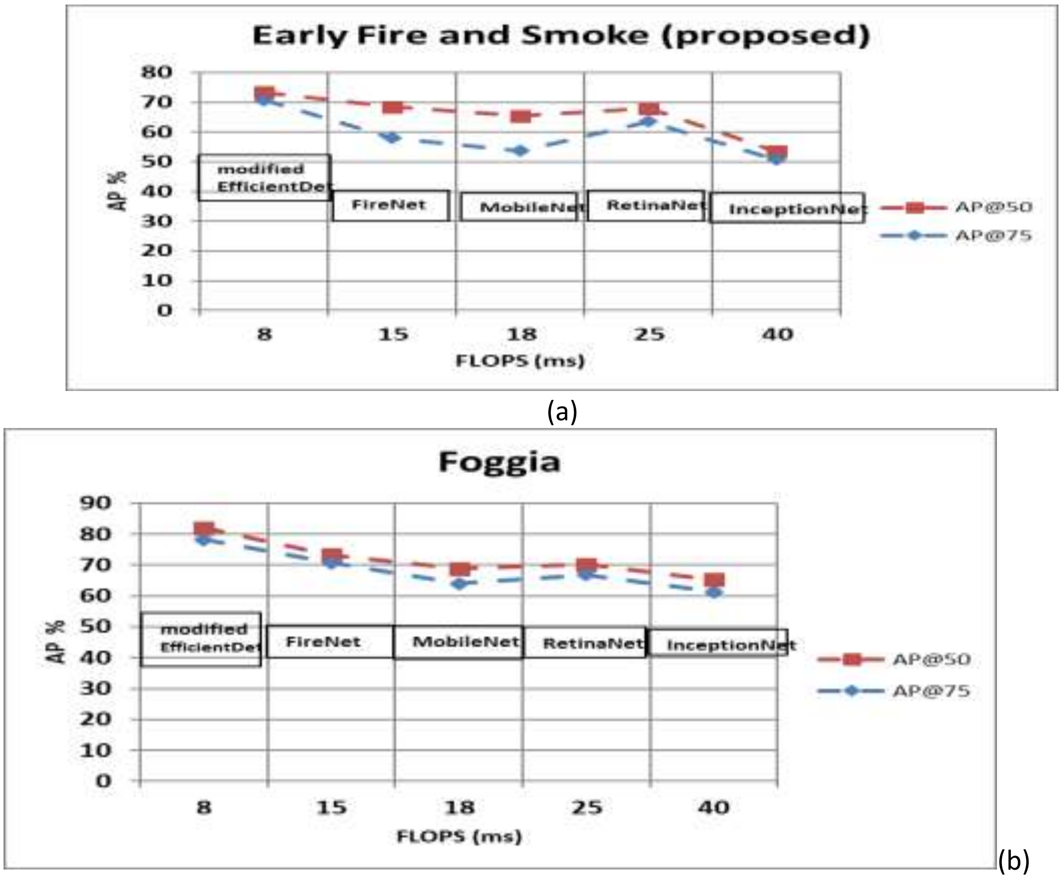
Pre-resizing	Norm (p1)	Resize + padding (p2)	Total preprocess (p1+p2)	GPU copy + float conv.	Inference	Post process	Parsing output	Total time/ batch	Avg. time / image
512x288	10.25	4.04	14.37	2.45	30.37	10.32	0.05	7.78	11
384x216	5.51	2.41	8.03	1.63	30.06	2.54	0.04	2.47	8
256x144	2.44	0.96	3.46	1.02	30.11	0.70	0.01	5.41	7

All numbers are in millisecond (ms) p1: normalized value

GPU copy + float conv.: For processing floating conversion, need to copy the images from CPU to GPU, this takes some time.
Inference: need to copy the images from CPU to GPU, this takes some time.

Post process: after inference we need to do pros-processing like non-max-suppression and box filtering that also takes time. Table 2 depicted the performance results of EfficientDet in terms of time details. The pre-sizing image (kept aspect ratio constants) at baseline has been categorized as high (512x288), medium (384x216) and low (256x144) image size quality for performance analysis. In comparison to high and low image scenario, medium size average time found superior i.e. 8 ms, also of improved accuracy. There would be a trade-off between accuracy and performance. In case of medium size image, time and resolution, both have been improved.

Figure 7 shows the comparison of FLOPS and AP for obtained results on proposed dataset and Foggia dataset using different networks. It clearly shows that, the improved FLOPS as well as AP values, EfficientDet works well on proposed and Foggia dataset. Figure 8 shows the obtained results on real-time image of proposed dataset. It also shows the detection bounding box to fire and smoke dataset images.



(a)

(b)

Figure 7 (a, b): Comparison of FLOPS and AP obtained for proposed dataset and Foggia dataset using different networks

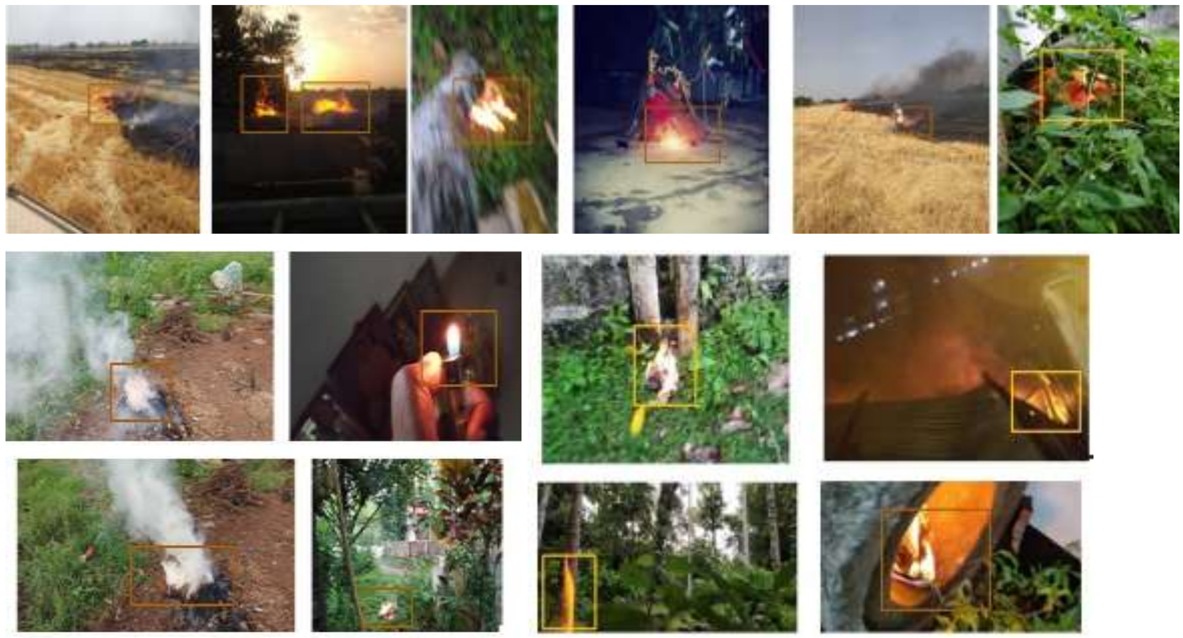


Figure 8 Detection of fire and smoke in proposed dataset

6. CONCLUSION

With the recent advancements in processing capabilities and advancement in deep learning shows promising results in surveillance system for identification of different abnormal events like fire, accidents, attacks etc. Fire is considered as one of the most dangerous anomaly events. Traditional vision-based methods do not show promising results to rely on real-world setting. It is largely due to complex nature of the problem that fire can occur in many different forms. Moreover, traditional methods fail in the case of confusion cases like shadows, fire like objects, sun, bright streetlights etc.

Presented recent Deep Learning based methods used for fire classification. These methods are better than conventional computer-vision methods. Most of these methods are computationally heavy and have very high memory footprint which makes them infeasible to use on embedded devices. The usefulness of fire detection system is only when it can be deployed for real-time usage. It is very important to localize the exact fire to make it a meaningful application. This study mainly focuses on object detection or anomaly detection algorithm rather classification only method. Study also devises a lightweight Convolutional Neural Network, CNN which performs on real-time detection. Compared to existing literature, proposed dataset solely generates a challenging and bigger real-world dataset or benchmarks. Foremost and very important advantage of proposed dataset would be the deficiency of image datasets for EARLY fire in existing datasets. The proposed dataset not only satisfies this criterion but also believe this dataset will become the basis for future research in this research field.

Experimental results have been obtained and compared with popular network like; MobileNet, RetinaNet, FireNet, InceptionNet. It shows that Google's latest models modified EfficientDet performs very well among other three netweok techniques, on proposed dataset compared to Foggia dataset. These results have been obtained with taken into consideration of Average Precision (AP) as an indicator. On proposed dataset, it shows around 74 % and 71 % for AP@50 and AP@75 respectively. Qualitative analysis inferred that these networks fail to classify cases where fire like objects is present.

Presently, hand-crafted features such as color, shape and patches are not sufficient for real-world scenarios. Therefore, the present work would be mostly suitable for a combined classification and localization method in early fire anomalous detection in real-time scenarios.

REFERENCES

[1] Liu CB, Ahuja N. Vision based fire detection. 2004. In Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004. 2004 Aug 26 (Vol. 4, pp. 134-137). IEEE.

[2] Kurniabudi, K., Purnama, B., Sharipuddin, S., Darmawijoyo, D., Stiawan, D., Samsuryadi, S., Heryanto, A. and Budiarto, R., 2019. Network anomaly detection research: a survey. Indonesian Journal of Electrical Engineering and Informatics (IJEEI), 7(1), pp.37-50.

[3] Celik T, Demirel H, Ozkaramanli H, Uyguroglu M. 2007. Fire detection using statistical color model in video sequences. Journal of Visual Communication and Image Representation. 2007 Apr 1;18(2):176-85.

[4] Ko BC, Ham SJ, Nam JY. 2011. Modeling and formalization of fuzzy finite automata for detection of irregular fire flames. IEEE Transactions on Circuits and Systems for Video Technology. 2011 May 19;21(12):1903-12.

[5] Choi J, Choi JY. 2015. Patch-based fire detection with online outlier learning. In 2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) 2015 Aug 25 (pp. 1-6). IEEE.

[6] Wang T, Shi L, Yuan P, Bu L, Hou X. 2017. A new fire detection method based on flame color dispersion and similarity in consecutive frames. In2017 Chinese Automation Congress (CAC) 2017 Oct 20 (pp. 151-156). IEEE.

[7] Muhammad, K. J. Ahmad, Z. Lv, P. Bellavista, P. Yang and S. W. Baik, 2019. Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications. IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 49, no. 7, pp. 1419-1434, July 2019, doi: 10.1109/TSMC.2018.2830099.

[8] Jadon A, Omama M, Varshney A, Ansari MS, Sharma R. 2019. FireNet: a specialized lightweight fire & smoke detection model for real-time IoT applications. arXiv preprint arXiv:1905.11922. 2019 May 28.

[9] Chi R, Lu ZM, Ji QG. 2016. Real-time multi-feature based fire flame detection in video. IET Image Processing. 2016 Sep 2;11(1):31-7.

[10] Kawatsu C, Koss F, Gillies A, Zhao A, Crossman J, Purman B, Stone D, Dahn D. 2017. Gesture recognition for robotic control using deep learning. InNDIA Ground Vehicle Systems Engineering and Technology Symposium 2017.

[11] Raghu M, Poole B, Kleinberg J, Ganguli S, Sohl-Dickstein J. 2017. On the expressive power of deep neural networks. Ininternational conference on machine learning 2017 Jul 17 (pp. 2847-2854). PMLR.

[12] Alqourabah, H., Muneer, A. and Fati, S.M., 2020. A Smart Fire Detection System using IoT Technology With Automatic Water Sprinkler. International Journal of Electrical and Computer Engineering (IJECE), 9(4).

[13] Taha, I.A. and Marhoon, H.M., 2018. Implementation of controlled robot for fire detection and extinguish to closed areas based on Arduino. Telkomnika, 16(2), pp.654-664.

[14] Idris, F., Hashim, N., Kadmin, A.F. and Yee, L.B., 2019. Intelligent fire detection and alert system using labVIEW.

International Journal of Electrical and Computer Engineering, 9(3), p.1842.

[15] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho and S. W. Baik, 2018a. Convolutional Neural Networks Based Fire Detection in Surveillance Videos. IEEE Access, vol. 6, pp. 18174-18183, 2018, doi: 10.1109/ACCESS.2018.2812835.

[16] Muhammad K, Ahmad J, Baik SW. 2018b. Early fire detection using convolutional neural networks during surveillance for effective disaster management. Neurocomputing. 2018 May 2;288:30-42.

[17] Muhammad K, Ahmad J, Lv Z, Bellavista P, Yang P, Baik SW. 2018c . Efficient deep CNN-based fire detection and localization in video surveillance applications. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2018 Jun 14;49(7):1419-34.

[18] 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. 2015 Jan 19;25(9):1545-56.

[19] P. V. K. Borges and E. Izquierdo. 2010. A Probabilistic Approach for Vision-Based Fire Detection in Videos. IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 5, pp. 721-731, May 2010, doi: 10.1109/TCSVT.2010.2045813.

[20] Mueller M, Karasev P, Kolesov I, Tannenbaum A. 2013. Optical flow estimation for flame detection in videos. IEEE Transactions on image processing. 2013 Apr 16;22(7):2786-97.

[21] Tan M, Le Q. 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. InInternational Conference on Machine Learning 2019 May 24 (pp. 6105-6114). PMLR.

[22] Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. 2017. Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 2117-2125).

[23] Tan, M., Chen, B., Pang, R., Vasudevan, V., Sandler, M., Howard, A. and Le, Q.V., 2019. Mnasnet: Platformaware neural architecture search for mobile. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2820-2828).

[24] Useche, P., Jimenez-Moreno, R. and Baquero, J.M., 2020. Algorithm of detection, classification and gripping of occluded objects by CNN techniques and Haar classifiers. International Journal of Electrical and Computer Engineering, 10(5), p.4712.

[25] Tan, M., Pang, R. and 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).

[26] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition 2018 (pp. 45104520).

[27] Singh, B., Marks, T.K., Jones, M., Tuzel, O. and Shao, M., 2016. A multi-stream bi-directional recurrent neural network for fine-grained action detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1961-1970).

[28] Zoph B, Cubuk ED, Ghiasi G, Lin TY, Shlens J, Le QV. 2020. Learning data augmentation strategies for object detection. InEuropean Conference on Computer Vision 2020 Aug 23 (pp. 566-583). Springer, Cham.

[29] Andrew G. Howard, Menglong Zhu, Bo Chen ,DmitryKalenichenko,Weijun

Wang, Tobias Weyand ,Marco Andreetto, Hartwig Adam ;2017.MobileNets: Efficient Convolutional Neural Networks for Mobile Vision
Applications. Ar.Xiv.org

[30] Mouna Aff1, · RiadhAyachi ,Yahia Said1, EdwigePissaloux , Mohamed Atri,2020,Springer neural processing letter, An Evaluation of RetinaNet on Indoor Object Detection for Blind and Visually Impaired Persons Assistance Navigation.

[31] Gonzalez et al., 2017. Accurate fire detection through fully convolutional network," 7th Latin American Conference on Networked and Electronic Media (LACNEM 2017), Valparaiso, Chile, 2017, pp. 1-6, doi: 10.1049/ic.2017.0026.

[32] Lin et al. 2020 Focal Loss for Dense Object Detection, IEEE transactions on pattern analysis and machine intelligence, vol. 42, no. 2, February 2020

[33] Kim B, Lee J. 2019. A video-based fire detection using deep learning models. Applied Sciences. 2019 Jan;9(14):2862.

[34] C. Szegedy et al., 2015.Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.

[35] K. He, X. Zhang, S. Ren and J. Sun. 2016. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[36] Christian Esposito, Massimo Ficco, Brij Bhooshan Gupta, Blockchain-based authentication and authorization for smart city applications, Information Processing & Management, Volume 58, Issue 2, 2021,

[37] Daming Li, Lianbing Deng, Brij Bhooshan Gupta, Haoxiang Wang, Chang Choi, A novel CNN based security guaranteed image watermarking generation scenario for smart city applications, Information Sciences, Volume 479, 2019, Pages 432-447,

[38] C. L. Stergiou, K. E. Psannis and B. B. Gupta, "IoT-Based Big Data Secure Management in the Fog Over a 6G Wireless Network," in IEEE Internet of Things Journal, vol. 8, no. 7, pp. 5164-5171, 1 April1, 2021, doi: 10.1109/JIOT.2020.3033131.

[39] Gupta, BB, Quamara, M. An overview of Internet of Things (IoT): Architectural aspects, challenges, and protocols. Concurrency ComputatPractExper. 2020; 32:e4946. https://doi.org/10.1002/cpe.4946