International Journal of Engineering &Technology, 7 (4) (2018) 6934-6936

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.17955
Research paper

Vision-based smoke detector

Abdullah Ali Mohammed Noman [1] *, Zaw Zaw Htike [1]

[1] International Islamic University Malaysia Jalan Gombak Selangor
**Corresponding author E-mail: abdullah.saif26@gmail.com*

Abstract

Previous studies have documented the significant applications of the electronic smoke detector. With the capabilities of vision based fire detection and increase in the number of surveillance cameras, a lesser attention is given to the vision-based type smoke detector. Moreover, some drawbacks have been identified in the accuracy and efficiency of smoke detection. The present study proposes a vision based smoke detector to overcome the shortcomings of the current traditional electronic and vision based smoke detectors. A Convolutional Neural Network is used to classify the smoke regions. After testing the proposed method, the accuracy was approximately 94%. When a modern approach of object detection is used to support image classifying, its accuracy increases by 96%.

Keywords: *Convolutional Neural Network; Rectified Linear Unit; You Only Look Once; Artificial Neural Network.*

1. Introduction

In the last few decades, fires and smoke were the leading cause of death and property destruction with the estimated cost burden of more than $27.2 billion globally. Fire disasters affect the surrounding environment, provoking air pollution. Many of the fire-related deaths are caused by a suffocation from the fire smoke. However, when smoke detectors were introduced and popular, accidents caused by fire and smoke decreased because the fire detectors assisted in the early detection of fire. The traditional smoke detectors commonly used nowadays are manufactured with an ionization sensor capable of triggering the alarm when smoke exists. But there are many drawbacks in the traditional smoke detector that perverts it from detecting the smoke with high accuracy. During the weather disturbances, or when the ceiling is very high, the detector might be incapable of detecting the smoke during a fire outbreak. Another consideration is that the smoke density might not be enough to trigger the smoke alarm. To save a life, and protect the environment from the fire outbreak, there is a need to come up with a novel technology to improve the performances of the fire detectors. The present study proposes the vision-based smoke detector, capable of detecting smoke rapidly to develop safer working environment, and preserve the environmental sustainability. Electronic smoke detectors are widely used in the industries, schools, and homes, however, there are many drawbacks to such systems. In many cases, the smoke might not reach the detectors if the roof is high. Moreover, the detectors might not sense the smoke of the newly developing fire. For example, the smoke particles might get colder, or smoke might be blown by air preventing it from reaching the ceiling. Thus, this type of smoke detector requires a different amount of smoke density to be alerted; if the smoke density is not enough, the smoke detector might not detect it [1]. To overcome these drawbacks, previous studies have focused on a vision-based smoke detection. However, many of these studies focus on color features shape and temperature characteristics. Segmentation using K-Means Algorithm is used for smoke detection where pixels are divided into clusters, and clusters with the same color are identified as smoke. Fig. 1 shows the algorithm of smoke detection, starting from input camera to the output frames, which contains smoke detection [2]. Another algorithm is type-2 fuzzy clustering for smoke detection.

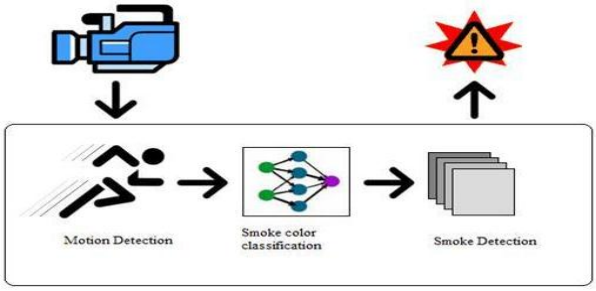


Fig. 1: Algorithm of K-Means Smoke Detection.

This classifier is used to separate smoke from fire pixels, and placed into classes. Then, clusters are prepared as a model of smoke or fire [3] to develop a convolutional neural network (CNN) classification algorithms. Then, a deep learning trains a given data to predict object classification. In this study, the proposed method is to use CNN to detect smoke particles in given frames or videos.
The rest of the paper is organized as follows: Section II presents a review of the literature relevant to the vision-based smoke detector. Section III explains the CNN principles. Section IV explains the YOLO CNN model. Section V presents the results with a discussion on the efficacy of the new system. Section VI provides the conclusion, limitation and recommendation.

2. Related works

Smoke and fire detection, developed with different algorithms, have been deeply investigated in the literature. A study on smoke detection proposed a technique of frames segmentation, using frame subtraction, and smoke particles are extracted by chromatic features. With the aid of the dynamic properties, the algorithm checks if the detected smoke is a real or not [4]. Another vision of smoke detection was from the support vector machine algorithm along the

Copyright ©2018 Abdullah Ali Mohammed Noman, Zaw Zaw Htike. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

extraction of wavelet properties from approximate constant [5]. The author stated this system worked with an accuracy of 85% using the block-based approach [5]. Moreover, an image enhancement fuzzy logic, dynamic, and static features are integrated with smoke. These features are normalized and grouped in a single vector using the support vector machine model. The performance is of high accuracy, however, it is characterized by false alarm detection [6]. In addition, analysis of chromaticity showed that it could be used to obtain the greyscale color video from each pixel. Moreover, harmonic motion detection technique is introduced to obtain the characteristic motion of the smoke particles detected for the thickness of smoke boundary [7]. All the models discussed require strong features of smoke to reduce a false alarm. In this article, a CNN model is introduced for faster and more accuracy of smoke detection.

3. Convolutional neural network

The CNN is a deep learning method used in machine learning. Analysis of [8] CNN shows that it delivers the best performances than other classifications as it is a part of Artificial Neural Network very effective in image processing. Accordingly, the convolutional layer uses several feature maps as input convolved to detect feature layer, producing a feature map layer as output [9]. The convolution pooling layers perform feature image extraction, such as edges, eyes, shapes and many other features. LeNet5 was the first popular neural network that helped in developing deep learning successfully after many iterations. Based on its effectiveness, LeNet5 was first used for character recognition. Fig. 2 shows CNN layers.

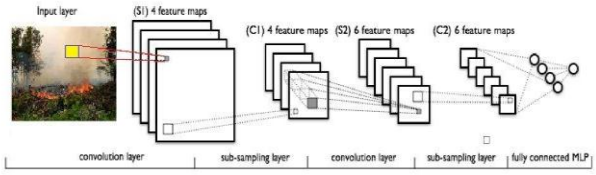


Fig. 2: CNN with Multiple Layers.

3.1. Convolutional neural layer

The convolutional neural layer is the main layer of CNN that consists of a set of small neurons where filters serve as a feature detector sometimes call filter detector from the original input image. The filer convolves over the pixels of the original images, and the output matrix is called the feature map. Multiples feature maps are called convolutional layer. The better feature map will be developed and different feature detector will come up using the convolutional layer. Equation (1) shows the convolutional operation, where f is the input frame and g is the feature detector.

$$(f * g)(t) = \int_{-\infty}^{\infty} f(T)g(t-T)dT$$

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau \quad (1)$$

Each layer in the convolutional neural networks acts as s filter for the specific character or patterns. The first layer of the convolutional layer is used to recognize and encode the abstract of the characters. For more difficult characters, further layers are aimed to classify the CNN as the CNN classification run faster when selecting a minimized network, such as the nine-layer CNN [8]. However, it was quite difficult to detect smoke region correctly because of the nature of smoke, thus, the training set must be developed for better accuracy. When compared this method with other algorithms over the smoke videos, three activation functions from CNN are identified [10]. These functions are RELU, Sigmoid, and Leaky ReLu, acting as a cutting tool for detecting video-based smoke. Taking Relu as an example, the values are lesser than zero and making the values greater than zero unchanged. These functions have improved the performances of the system marginally. Fig 3 provides the illustration of these functions.

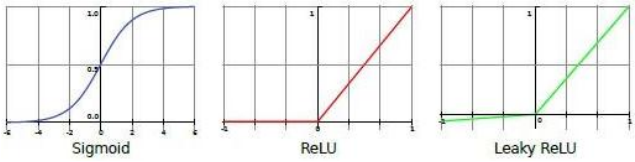


Fig. 3: Sigmoid, RELU and Leakly Relu Activation Function.

3.2. Pooling layer

Pooling layer is the second most important layers in CNN because it helps to extract the characteristic of complex cells. The pooling layer minimizes the neurons taken from neighboring pixels to reduce the size of the input layer. [11] stated that pooling layer is used to get spatial invariance that minimizes the feature map resolution [11]. Pooling features demonstrate a good performance for extracting invariances from pictures compared to others line subsampling methods, however its recognition rates have not yet improved [12].

3.3. Regularization

After pooling layer, a fully connected layer is created where all features after convolution and pooling are connected, creating an important feature image vector. Data augmentation is a way of increasing the size of the dataset, transforming the image without changing its label. In a study by [13], snipping and rotating the frames randomly increase the size of the dataset [13]. This method has reduced an overlapping of data significantly.

4. Proposed model

In this paper, a YOLO CNN is proposed for smoke detection. This model predicts all bounding boxes in an image and uses all CNN features to predict all the image classes. The model converts the image into (x *x) cells, and if the center of this cell is pointing to a certain class of an object, the cell has the capacity to classifying that object. Every cell has B surrounding, contour along the confidence factor, which is important to determine a correct classification of the object. Each surrounding box has five variables that reruns. Two variables represent the center of the box, and two other variables represent the image dimensions and the remaining one variable is for the confidence level of prediction.

4.1. Network design

This model is developed with CNN, and this is the first convolutional layers coded as features of fully connected output layers. The proposed network use nine convolution layers. The feature map for this convolution is 3 x 3 layer and the max pooling is a 2 x 2 layer. Last convolution layer is 1 x 1 feature map that depresses the shape to 13 x 13 x 125. From Fig. 4, 13 x 13 is the last convolution layer size, however the number 125 is from the bounding box. Each cell predicts 5 bounding boxes and each bounding box reruns 5 variables explained in the first section. Considering the number of classes to be 20, the equation 5x (5 + 20) will be equal 125. Thus, this simple equation is used to determine feature map size (filter). Since smoke is the only class of detection, if B = 1, the feature map (filter) is 30. From Fig. 4, the input images are resized to 416 x 416 to fit the model without any overfitting issues.

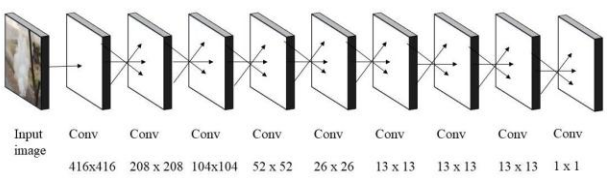


Fig. 4: Nine Convolutional Neural Network.

$$5 \times (B + 20) \quad (2)$$

4.2. Training dataset

For training the dataset, 14.6% of images were downloaded from ImageNet.org. However, not all images were related to smoke. Thus, a total of 2702 images were captured from real smoke frames. For every image, an annotation file is required to locate the smoke region, referring to image labeling. Thus, the training is performed using Core™ i7 4710HQ Processor with 3.5 GHz. However the training process was very slow as it took more than 24 hours to complete. Therefore, a graphical processing unit is used to speed up the process. NIVIDA GTX 860M with 640 cores could finish the training within 6 hours of using 300 epochs. The loss in the model is 1.01 with a 98.99 % training accuracy.

5. Result

The first 395 images trained model has delivered good results. Since the number of images is not enough, false positive detection is recorded. Fig. 5 shows the cases of false positive smoke detection. However, when we trained the model with 2702 images, the result was more accurate and stable (Fig. 6). The trained model has been tested in real time using two embedded systems (GEFOREC NVIDIA 860M GPU and RaspberryPi3). The performance was clearly noticed when the GPU achieved a rate of 30 frames per seconds(fps) with LOGITECH CAMERA. However, the RaspberryPi3 rate is very slow achieving 0.3 frames per second. Although the speed is slow, the system managed to detect smoke images as long it exists in the frames.



Fig. 5: False Positive Detection.

Nevertheless, there were cases where false positive detection was recorded despite a huge dataset. To address this problem, the threshold value is increased to remove non-smoke particles; the higher the threshold value, the higher the accuracy. However, a careful selection of the threshold value is still essential because if threshold value increases beyond a certain limit, no smoke is detected.



Fig. 6: Smoke Detection After 2702 Trained Images.

The system is tested in a real-time with different threshold values, and a performance of the system is evaluated using the equation (3).

$$\left(\frac{\text{number of no smoke frames}}{\text{total frames}} - 1\right) \times 100 \quad (3)$$

Lastly, the vision-based smoke detector successfully detects smoke with a threshold value of 0.4 giving an accuracy of 98%. Table 1 shows the accuracy of the model with respect to the value of the chosen threshold. Table 1 summarizes the overall system design.

Table 1: System Accuracy with A Different Threshold Value

Threshold value	Total frames	Frames detected with no smoke	Accuracy %
0.15	40000	3500	91.25
0.2	40000	2000	95
0.3	11000	94	98.64
0.35	10000	96	99.04
0.47	4000	0	100

6. Limitation

In future work, integration of multiple object detection algorithm is required, and their combination into one model will make the detect object as smoke. Moreover, the time taken to train GPU and test the model could be examined deeply using different models. Since the frame rate is slow, the model can be implemented in FPGA to achieve a better frame rate.

7. Conclusion

In conclusion, the main vision based smoke detector concept has been studied and explained. A lot of problems have been identified while completing this article. The first problem involved obtaining the correct dataset to be trained. Addressing this problem involves transferring some images from a video to frames. Second, the version of the libraries, image processing, and machine learning used for this article are very sensitive to older and newer versions. In some cases, older versions might work with newer version of another library, and it took a lot of time to address this issue. Moreover, training a system takes many hours depending on the size of dataset and number of iteration. Overall, the system performance was high and the results were very robust with 98% level of accuracy.

References

[1] P. A. Panel, "PS-Tools."
[2] M. Shrivastava and P. Matlani, "A SMOKE DETECTION ALGORITHM BASED ON K-MEANS SEGMENTATION Head Of Department (CSE) Guru Ghasidas University, koni, Bilaspur (C. G), India Ph. D Scholar, Guru Ghasidas University, koni, Bilaspur (C. G), India," 2016.
[3] H. D. Duong, D. D. Nguyen, L. T. Ngo, and D. T. Tinh, "On Approach to Vision Based Fire Detection Based on Type-2 Fuzzy Clustering," *Ieee*, pp. 51–56, 2011. https://doi.org/10.1109/SoCPaR.2011.6089094.
[4] C. Thou-Ho, Y. Yen-Hui, H. Shi-Feng, and Y. Yan-Ting, "The smoke detection for early fire-alarming system based on video processing," *Int. Conf. Intell. Inf. Hiding Multimed. Signal Process.*, pp. 427–430, 2006.
[5] J. Gubbi, S. Marusic, and M. Palaniswami, "Smoke detection in video using wavelets and support vector machines," *Fire Saf. J.*, vol. 44, no. 8, pp. 1110–1115, 2009. https://doi.org/10.1016/j.firesaf.2009.08.003.
[6] Y. Wang, "Smoke recognition based on machine vision," *Proc. - 2016 IEEE Int. Symp. Comput. Consum. Control. IS3C 2016*, pp. 668–671, 2016.
[7] L. Miao, Y. Chen, and A. Wang, "Video smoke detection algorithm using dark channel priori," *Proc. 33rd Chinese Control Conf. CCC 2014*, pp. 7405–7408, 2014. https://doi.org/10.1109/ChiCC.2014.6896230.
[8] S. Frizzi, R. Kaabi, M. Bouchouicha, J. M. Ginoux, E. Moreau, and F. Fnaiech, "Convolutional neural network for video fire and smoke detection," *IECON Proc. (Industrial Electron. Conf.*, pp. 877–882, 2016. https://doi.org/10.1109/IECON.2016.7793196.
[9] J. Wu, "Introduction to Convolutional Neural Networks," pp. 1–31, 2017.
[10] C.-C. J. Kuo, "Understanding Convolutional Neural Networks with A Mathematical Model," pp. 1–21, 2016.
[11] D. Scherer, A. Müller, and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6354 LNCS, no. PART 3, pp. 92–101, 2010. https://doi.org/10.1007/978-3-642-15825-4_10.
[12] Q. Pan, C. Darabos, and J. Moore, *Artificial Neural Networks - ICANN 2010*, vol. 7246. 2012.
[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Adv. Neural Inf. Process. Syst.*, pp. 1–9, 2012.