Fire Detection Effect of Data Size Using on Deeplearning Techniques

M. Sc. Student Wijdan Yunes[1]

Department of Computer Science
College Of Computer and Mathematics
University of Mosul – Mosul - Iraq
Email: wijdan.20csp70@student.uomosul.edu.iq

Prof. Dr. Khalil Ibrahem Alsaif[2]

Department of Computer Science
College Of Computer and Mathematics
University of Mosul – Mosul - Iraq
Email : Khalil_alsaif@uomosul.edu.iq

Fires are considered one of the unfortunate phenomena that cause disasters in the environment, and it also represents a great threat to lives and human safety, especially when this fire is not detected by sensor-based fire detection systems. Therefore, equipping those places with cheap and effective sensors is very useful in fire detection speed.

Image-based smoke and fire detection systems that use surveillance cameras have been designed, as vision-based fire detection systems benefit from three basic characteristics of fire: color, movement and shape (the shape of fire).

In this research, the features are extracted from the fire images, then an input to one of the classification methods, which depends on the method of deep learning using the Convolutional Neural Network (CNN), which is one of the modern technologies in recent years, the movement of the flame is detected after segmenting the images and then labeling them in order to extracting the shape of the flame. In general, this work aims, to detect fire and flames in images by acquiring video files using regular surveillance cameras, and then the process of classification and detection is carried out with the applications of deep learning algorithms (classification using deep learning method).

The proposed algorithm provided positive results when it was adopted in video files containing real flames that were entered into the approved file so that through the proposed algorithm, the expected fires would be discovered. A very high percentage of alerting about the presence of a flame was achieved, as the error rate did not exceed 2% in different and varying video files in the location of the flame, in addition to its variation in size and shape.

Keywords: Image processing, object tracking, fire detection, deep learning.

تأثير حجم البيانات على كشف اللهب باستخدام التعلم العميق

طالبة الماجستير وجدان يونس عبد ا. د. خليل ابراهيم السيف
كلية علوم الحاسوب والرياضيات
جامعة الموصل, الموصل , العراق

الملخص

تعتبر الحرائق واحدة من الظواهر المؤسفة التي تسبب كوارث في البيئة وهو ايضا يمثل تهديدا كبيرا على الارواح وسلامة الانسان بالأخص عندما يكون هذا الحريق غير مكتشف من قبل انظمة كشف الحرائق القائمة على جهاز الاستشعار , ولذلك يعد تجهيز تلك الأماكن بأجهزة استشعار رخيصة الثمن وفعالة مفيد جدا في سرعة اكتشاف الحريق .

تم العمل على تصميم انظمة كشف الدخان والحرائق القائمة على الصور التي تستخدم كاميرات المراقبة حيث تستفيد انظمة الكشف عن الحرائق القائمة على الرؤية من ثلاث صفات اساسية للنار :- اللون , الحركة والشكل (شكل النار) في هذا البحث يتم استخراج الميزات (Features) من صور الحريق ومن ثم تكون مدخلات الى احد طرق التصنيف والتي تعتمد طريقة التعلم العميق (deep learning) باستخدام الشبكة العصبية التلافيفية CNN والتي تعتبر من التقنيات الحديثة في السنوات الاخيرة حيث يتم الكشف عن حركة اللهب بعد تجزئة الصور ومن ثم وسمها (Labelling) واستخراج شكل اللهب .

يهدف هذا العمل بصورة عامة الى الكشف عن النار واللهب في الصور من خلال اكتساب ملفات فيديوية باعتماد كآمرات المراقبة الاعتيادية، ومن ثم تتم عملية التصنيف والكشف بتطبيقات خوارزميات التعلم العميق (classification using deep learning method).

وقد قدمت الخوارزمية المقترحة نتائج ايجابية عند اعتمادها في ملفات فديوية تحتوي على لهب حقيقي ادخلت الى الملف المعتمد ليتم من خلال الخوارزمية المقترحة اكتشاف الحرائق المتوقع نشوبها . تم تحقيق نسبة عالية جدا في التنبيه بوجود لهب حيث كانت نسبة الخطأ لا تتجاوز 2% في ملفات فديوية مختلفة ومتباينة في موقع اللهب اضافة الى تباينه في الحجم والشكل .

الكلمات المفتاحية: معالجة الصور, تتبع الاهداف , اكتشاف النار , التعلم العميق.

1- Introduction:

The field of disaster management is an important field in relation to the life of organisms on Earth, and this topic is of interest to many researchers in the world. Disasters are generally classified into two categories. There are technological disasters, such as matters related to hazardous materials or nuclear power plants, and others. The other type is known as natural disasters such as floods, earthquakes, forest fires and others. The phases of disaster management include: Preparedness, mitigation and response. This data can be processed by using data analysis techniques, where data processing techniques include information extraction, retrieval, purification of noise, data mining and decision-making[1][2].

In recent years, intelligent processing capabilities have increased and monitoring devices have become smarter than they used to be. Monitoring devices depend on monitoring any early and important information that can cause major disasters [3]. Where the chances of major disasters can be reduced to the least possible and the least possible losses. Fires are the most common accidents in the world that can be detected in the early stages. In the United States alone, a total of 1345,500 fires occur, according to reports prepared by (The National Fire Protection Association: NFPA) in 2015, resulting in material losses estimated at 14.3 billion dollars, 15,700 injuries and 3280 deaths, and the incidence of home fires was 78% [4].

Early detection of fires is of paramount importance to disaster management systems, in order to avoid such disasters. Therefore, researchers have discovered different methods for detecting fires, such as conventional fire alarm systems and systems based on optical sensors. Conventional fire systems rely on optical or ionic sensors, need a location close to the fire, and therefore fail to provide additional information such as fire size, location, and degree of combustion. Additionally, such systems rely on human intervention, such as visiting the fire site to confirm the fire in the event of any fire alarm.

To deal with above limitations, several optical sensor-based fire detection systems have been introduced[5]. The majority of research is directed towards fire alarm systems, which rely on vision sensors through cameras. The visible-light camera flame detection method consists of three levels, including pixel level, binary data level, and correction level. The binary data level showed better performance than the pixel level, but there is a problem that the classifiers are difficult to train due to the many shapes of flame points[6].The use of fire detection systems based on optical sensors has several advantages[1]:

- As the cost is low.
- Provides monitoring of large areas and leads to a relatively quick response.
- The occurrence of the fire is confirmed without the need to visit the site of the accident.
- In addition, these systems provide flexibility in detecting smoke and flames and give accurate details about the fire such as size, location and degree.

Despite the many advantages of these systems in detecting fires, they suffer from many problems, including the complexity of scenes during observation due to the presence of people and objects that look like fire, and the irregular lighting leads to a decrease in the quality of the images that are captured. Researchers seek to solve these problems to achieve the best level of disaster management. Sending all data to multiple cameras during monitoring is an impractical measure due to network restrictions due to the momentum of data flow through it. In addition to the above, the fire alert system needs an independent and reliable communication medium to send data through, so that the disaster management team can deal with it as soon as possible. All the methods that were dealt with in previous research to detect fires are based on vision using color model, movement, and spatial-temporal features to detect flames[1].

Image processing is done by searching for moving pixels and applying a color model to search for fire. Modern deep learning method is used to train data and train discriminant classifiers to detect fires. One of the methods of deep learning is the deep neural network (CNN), which has effectively contributed to the detection of fires[1]. Techniques have been developed based on the spatial features of the fire zone as the fastest zone-based convolutional neural network (R-CNN) to detect suspected and non-burning areas. The features are summarized within the bounding boxes in successive frames and assembled by (Long Short_Term Memory:LSTM) to determine whether there is a fire in a short period. Then the decisions for successive short periods are combined with the selection of the final decision in the long term [7].

2- Historical overview and related works:

Technological development has led to the emergence of various sensors for various applications. In 2004, C.-B. Liu et al. [8] examined three different models, including spectral, spatial and temporal of fire zone in the images. An algorithm was proposed to detect the spatial model of the spatial structure within the fire zone and an algorithm to detect the temporal model of the time signatures of the fire zone. B.U. Toreyin et al. in 2006 [9] developed an algorithm for detecting flames in color video. The results showed the success of the proposed method, as the false alarms from ordinary objects with colors similar to fire were significantly reduced. Another model was proposed, by G. Marbach et al. In 2006 [10] adopting the YUV model, which represents an images processing technology for automatic detection of fire in real time for video images. The results of this study revealed the success of the proposed method for detecting fire in video images under a variety of conditions. In 2008, another method was presented by Z. Zhang et al. [11] for detecting forest fires using wave analysis or rapid Fourier analysis FFT, and a new method was developed that combines the two methods for early detection of fire, and the results of the study achieved its goal. Also, B.C. Ko and others in 2009 developed visual sensors for early warning of fires using the SVM model, and experimental results showed that this method is better than the previous methods and has a robustness towards noise such as smoke [12]. In 2009, T. Çelik and H. Demirel [13] developed and investigated the YCbCr model for lighting and coloring components to classify flame images and achieved results with a higher detection rate, lower false alarm rate, and a higher than 90% rate of fire detection. And in 2010 P.V.K. Borges et al.[14] developed computer vision-based fire detection algorithms to model color in videos, and low-level specific features of fire zones such as aberration, color, roughness, and area size for fire recognition. This method achieved an average false positive rate of 0.68% and an average false negative rate of 0.028%. In 2014, I. Mehmood et al. developed wireless sensor devices through which imaging inside the human body and managing sensor data through a mobile cloud using wireless capsule endoscopy (WCE). The results of this study contributed to reducing the cost of information transmission and saving time for data analysts for remote browsing [15]. In 2015, A. Sorbara et al. developed sensors to detect obstacles[16] and in 2015 the vision-based sensors were studied where candidate fire zone are identified, color analysis and then fire behavior is modeled by using different spatial-temporal features using a two-class SVM classifier to classify areas, the results showed the superiority of the proposed method over other algorithms [17]. In 2016, Q. Zhang et al.[18] proposed a fire detection method based on a non-fixed camera mounted on a drone (UAVs) to detect forest fires by training a whole-image fire classifier using deep correlated convolutional neural networks (CNN), it was obtained on detection accuracy 97% and 90% in training. In 2018, K. Muhammad et al. [19] developed effective fire detection using a Transformational Neural Network (CNN) and coordinate video surveillance applications, and the results showed the accuracy and efficiency of detection in this way. A proposed deep learning based fire detection method adopted In 2019 by B. Kim et al.[7] using Faster R-CNN that uses a video sequence simulating real fire detection process and the proposed method achieved excellent fire detection accuracy and false alarm reduction. In 2021, R.Xu et al. [20] proposed a new ensemble learning method for forest fire detection in different scenarios in which two individual learners Yolo v5 and Efficient Net. were combined to achieve the fire detection process. Another scenario is the individual learner Efficient Net to avoid false positives. Detection performance was improved by 2.5% to 10.9% and false positives were reduced by 51.3% without any additional delay time.

3- Deep learning for object detection:

Deep learning has penetrated all fields and revolutionized the automation and robotics industries and in the medical field, in recent years. Deep learning has become one of the commonly used techniques due to the results this technology has achieved in applications of language processing, object detection, images classification, and matching news, publications, or products of interest to users[21]. One of the most important basics of detecting things is the classification and detection of images.

A traditional convolutional neural network (CNN) consists of three main neural layers namely, convolutional layers, aggregation layers, and fully connected layers (FCL). Figure 1 shows the training process using deep learning.

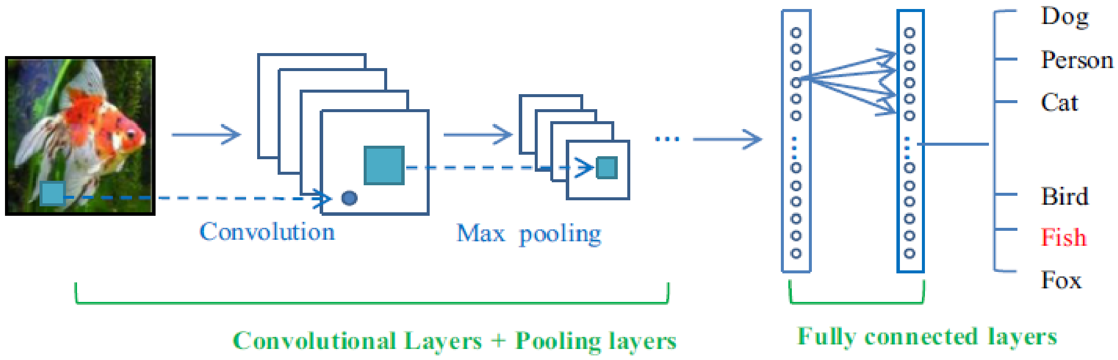


Figure 1: Deep learning training algorithm

One method of deep learning is Single Shot Detection (SSD) that adds many layers and thus increases the features of the final network and makes the detection process easier. And the other way is Faster Region based Convolutional Neural Networks (Faster R-CNN)

It is a standardized, fast and accurate way to discover objects using a convolutional neural network. Another deep learning technology, You Only Look Once (YOLO) network, which appeared in 2015 and was practically adopted in 2016, was developed by Joseph Redmon who introduced a comprehensive network[22]. The training process using SSD includes matching and is divided into negative matching and positive matching, followed by the hard passive mining process, then the data is increased to increase the accuracy and finally the final detection stage. Figure 2 shows the SSD model.

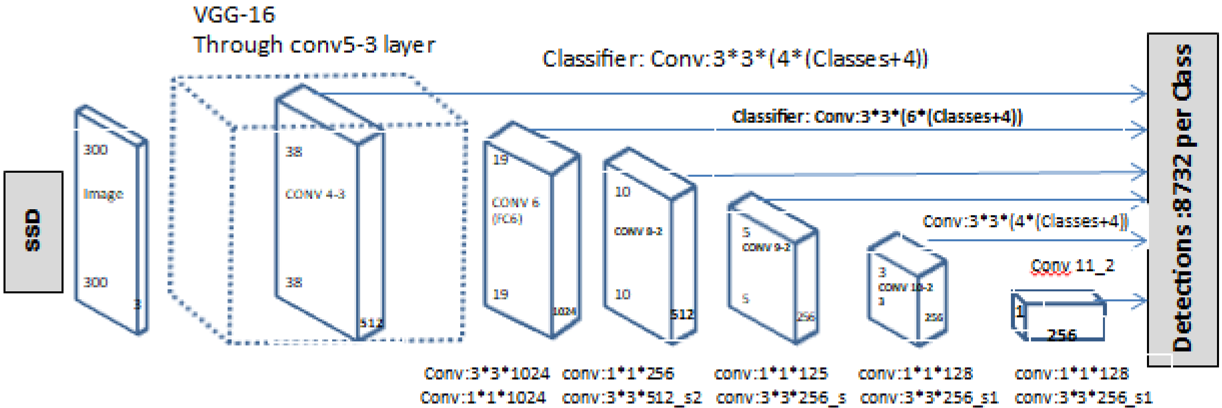


Figure 2: the SSD model

The zone- based (R-CNN) method uses object segmentation to detect, which is done using a selective search algorithm, extracting distinct features and then feeding the extracted features into the SVM (support vector machine) to identify objects. Although this technology achieves, it suffers from a lot defects. The classification of the proposals of nearly 2000 regions makes the training process take a long time, which makes the implementation in real time impossible, because the test of each image takes 47 seconds approximately to execute using medium specification computers[22]. Figure (3) shows the Faster R-CNN model. It was developed into Faster R-CNN, and most of the problems suffered by the previous method were overcome and classification performance improved [22].

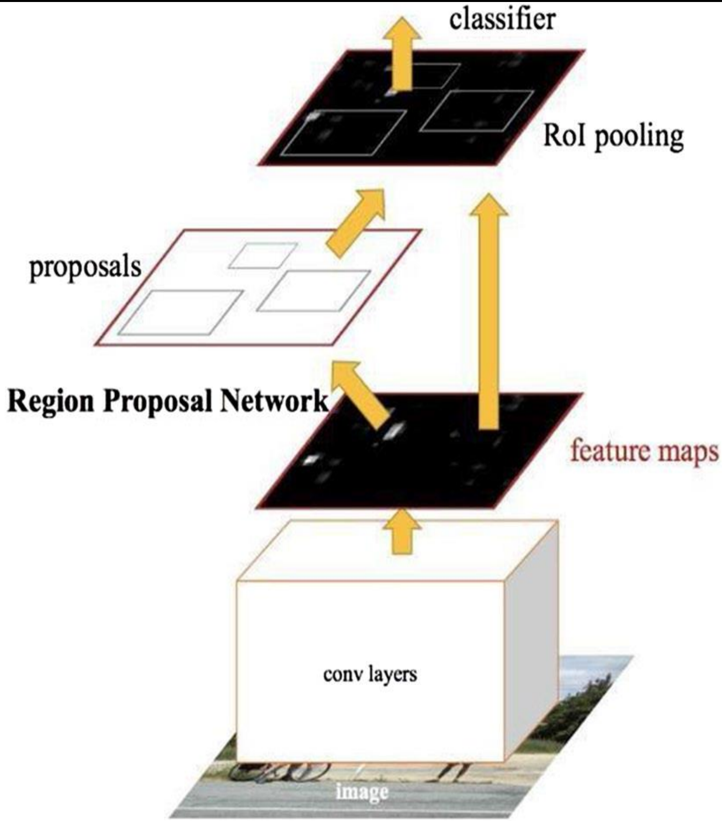


Figure (3): the Faster R-CNN model

In recent years, the YOLO algorithm you look only once at the picture to predict what things are and where they are- has been developed, which is a single convolutional network. This network has the ability to predict multiple frame boxes and objects probabilities in a clear and simultaneous manner, as it is currently one of the accurate objects detection algorithms. This algorithm is used to combine the target location with recognition on target within a comprehensive detection process because YOLO network trains the whole image to improve detection performance [23]. The YOLO model excels the previous models with real-time target detection. YOLO and CNN apply one to the whole image ,which divides the image into networks. The bounding boxes prediction and the relevant confidence score are calculated for each network. These bounding boxes are analyzed by the predicted confidence score. The YOLO structure comprises 24 convolutional layers and two fully connected layers[24][25]. In addition to multiple convolutional layers, other combined and connected layers. In the 1x1 and 3x3 convolutional layer an alternate convolution kernel is used to extract multiple features. The network's job is to configure it to use a pre-training model for the ImageNet dataset. The original image size is set to 448×448 and the pixel value is set to [-1, 1]. Convolutional layers extract the target feature and then reduce the dimensions of that feature using an assembly layer. The final step in the network structure is the fully connected layer (FCL) during which the object's class and coordinates are predicted [23]. The structure of the YOLO network is shown in figure (4).

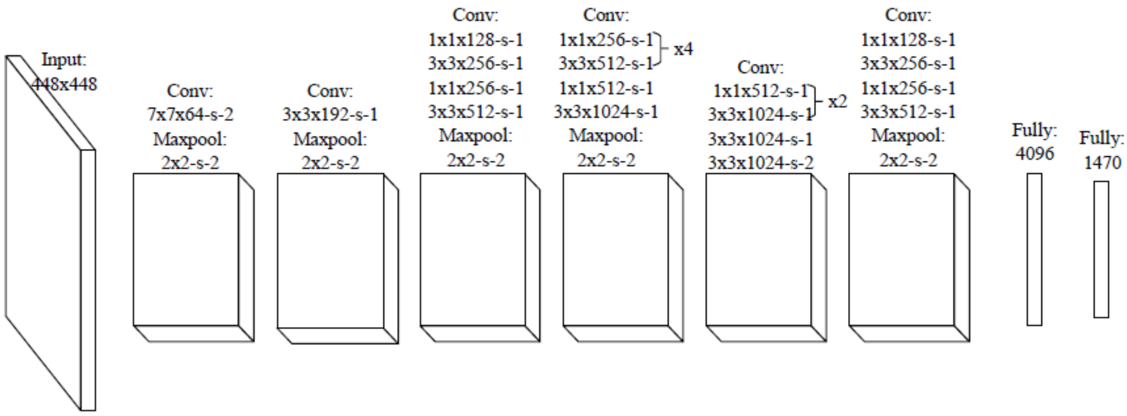


Figure (4):YOLO network architecture

YOLO takes advantage of a number of design principles to create one-shot architectures to achieve objects discovery. The structures of this network are very large, reaching 240 MB when copied YOLO v3 and due to the computational complexity, the inference speed decreases when processing the edge. Tiny YOLO was developed to address this problem, which contributed to reducing the size of the models. YOLO Nano has been developed recently, that an embedded network with a micro-architecture dedicated to micro-architectural designs, to detect embedded objects and bodies[26].

4- Practical Implementation steps:

Configuration and training are carried out by means of fire and various internal and external fire images and by using a wide range of videos of fire, smoke and other passive clips in different environments, whether indoors such as the home and office or outdoors such as parking areas and in the storage area. The work included two main phases, the training and testing phase. 984 images were used for training, many of which contain flames in different condition. The flame area is determined by labeling the image and specifying the location within a rectangle. Followed by the flame detection stage and then the test is conducted. The test included several real images in different conditions, with or without lighting, and random images were entered from the set of images approved for training. Figure (5) displays images that were used in the test phase.

(A) (B)

Figure (5): Sample images used in the test

5- Results and discussion:

Initialization and training of data is done by using fire and various internal and external fire images and using a wide range of videos of fire, smoke and other passive clips in different environments, whether indoors such as home and office or outdoors such as parking areas and in the storage area. After the data was trained and the test was conducted, the results obtained were as shown in Figure (6) and for cases of the error rate for fire identification (Iteration). The results showed that the lowest rate of fire recognition was when Iteration = 10, where only two images were recognized. When Iteration raises the number of images were recognized proportionally increased . The effect of two factors was studied: the accuracy scale (RMSE) and the training error rate (Loss) clearly seen in Table (1). The results concluded that the worst value of the accuracy scale was

8.30 when Iteration 30, and the best value of the accuracy scale was obtained at Iteration 90. It also appears that the highest training error rate is 69 at Iteration 30, while at Iteration 90 the best error rate was achieved and equal to 0.3 .

Table (1) Relationship between the accuracy scale (RMSE) and the training error rate (Loss)

Epoch	Iteration learning	Time Elapsed (hh:mm:ss)	Mini-batch RMSE	Mini-batch Loss
1	1	00:00:03	7.90	62.3
2	90	00:05:08	1.85	3.4
3	150	00:08:22	1.79	3.2
.	.	.	.	.
.	.	.	.	.
8	450	00:24:26	1.27	1,6
10	610	00:33:04	1.58	2,5

Iteration 10 (1)

Epoch	Iteration learning	Time Elapsed (hh:mm:ss)	Mini-batch RMSE	Mini-batch Loss
1	1	00:00:03	8.15	66.4
2	90	00:04:24	1.83	3.3
3	150	00:07:13	1.64	2.7
.	.	.	.	.
.	.	.	.	.
14	810	00:39:27	1.35	1.8
15	900	00:43:52	1.19	1.4

Iteration 15 (2)

Epoch	Iteration learning	Time Elapsed (hh:mm:ss)	Mini-batch RMSE	Mini-batch Loss
1	1	00:00:03	8.29	68.7
2	90	00:04:27	1.67	2.8
.	.	.	.	.
14	810	00:38:55	1.22	1.5
.	.	.	.	.
20	1220	00:58:44	1.45	2.1

Iteration 20 (3)

Epoch	Iteration learning	Time Elapsed (hh:mm:ss)	Mini-batch RMSE	Mini-batch Loss
1	1	00:00:03	8.15	66.4
2	90	00:04:21	1.65	2.7
.	.	.	.	.
21	1230	01:00:56	1.05	1.1
.	.	.	.	.
25	1525	01:16:02	1.34	1.8

Iteration 25 (4)

Epoch	Iteration learning	Time Elapsed (hh:mm:ss)	Mini-batch RMSE	Mini-batch Loss
1	1	00:00:02	8.30	69.0
2	90	00:04:23	1.75	3.0
.	.	.	.	.
27	1590	1:19:02	1.06	1.1
.	.	.	.	.
30	1830	01:30:29	1.16	1.3

Iteration 30 (5)

Epoch	Iteration learning	Time Elapsed (hh:mm:ss)	Mini-batch RMSE	Mini-batch Loss
1	1	00:00:02	8.0	63.9
2	90	00:04:33	1.72	3.0
.	.	.	.	.
.	.	.	.	.
59	3540	02:58:36	0.92	0.8
60	3600	03:01:38	0.82	0.7

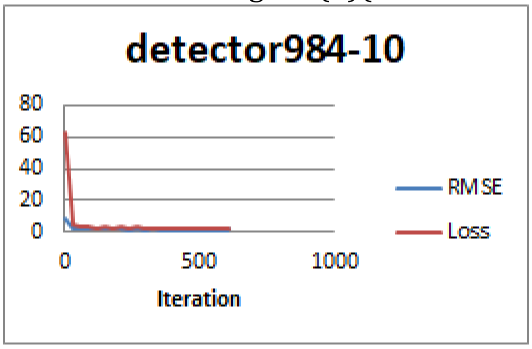
Iteration 60 (6)

Epoch	Iteration learning	Time Elapsed (hh:mm:ss)	Mini-batch RMSE	Mini-batch Loss
1	1	00:00:03	8.25	68.1
2	90	00:04:40	1.70	2.9
.	.	.	.	.
.	.	.	.	.
86	5190	06:26:10	0.58	0.3
.	.	.	.	.
90	5490	06:41:08	0.69	0.5

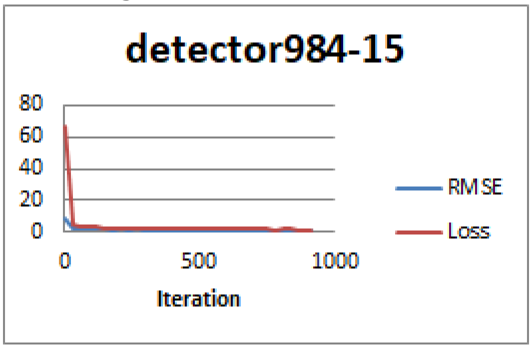
Iteration 90 (7)

ISSN: 2795-7640

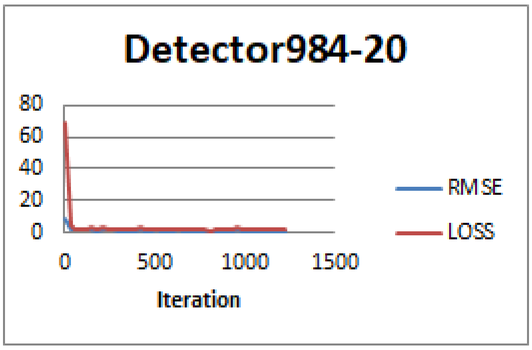
Figure (6)(1,2,3,4,5,6,7): Training data results



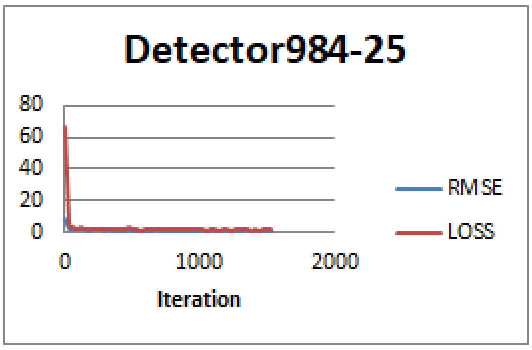
(1)



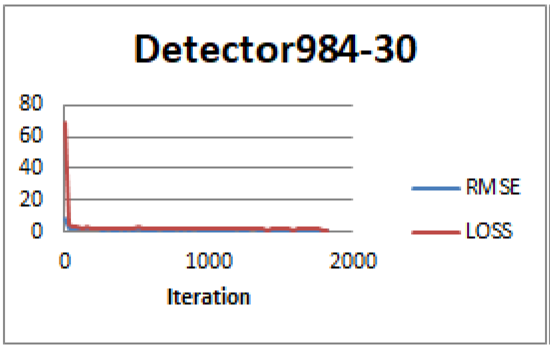
(2)



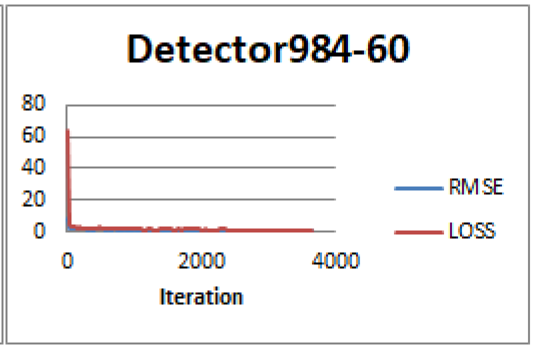
(3)



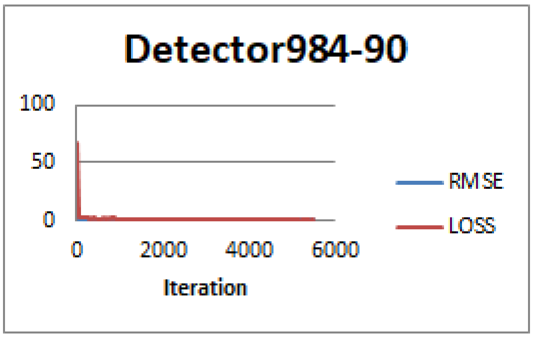
(4)



(5)



(6)



(7)

6- Conclusion:

This study achieved good results by detecting fires, using regular surveillance cameras, and with high accuracy, using deep learning. The work included the application of the automatic labeling method for the image data of fires in order to identify and label important areas in the images, which include the shape of fire and flames in the images. The latest deep learning convolutional neural networks (YOLOv2) has also been applied as a classification tool to distinguish and define fire and flame zones in images with high accuracy. The proposed fire detection system has been studied on a set of contrasting positive and negative images gives a

very high percentage of alerting about the presence of a flame was achieved, as the error rate did not exceed 2% in different and varying video files in the location of the flame, in addition to its variation in size and shape in addition to different environments, whether indoors such as home and office or outdoors such as parking areas and in the storage area . and compare results using accurate tables and graphs.

Reference:

1. K. Muhammad, J. Ahmad, and S. W. Baik, "Early fire detection using convolutional neural networks during surveillance for effective disaster management," *Neurocomputing*, vol. 288, pp. 30–42, 2018, doi: 10.1016/j.neucom.2017.04.083.
2. V. Hristidis, S. C. Chen, T. Li, S. Luis, and Y. Deng, "Survey of data management and analysis in disaster situations," *J. Syst. Softw.*, vol. 83, no. 10, pp. 1701–1714, 2010, doi: 10.1016/j.jss.2010.04.065.
3. K. Muhammad, J. Ahmad, I. Mehmood, S. Rho, and S. W. Baik, "Convolutional Neural Networks Based Fire Detection in Surveillance Videos," *IEEE Access*, vol. 6, pp. 18174–18183, 2018, doi: 10.1109/ACCESS.2018.2812835.
4. J. Choi and J. Y. C. B, "An Integrated Framework for 24-hours Fire," no. Vl, pp. 463–479, 2016, doi: 10.1007/978-3-319-48881-3.
5. B. C. Ko, S. J. Ham, and J. Y. Nam, "Modeling and formalization of fuzzy finite automata for detection of irregular fire flames," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 12, pp. 1903–1912, 2011, doi: 10.1109/TCSVT.2011.2157190.
6. K. M. Douglas and R. M. Sutton, "An Auto-adaptive Edge Detection Algorithm for Flame and Fire Image Processing," *Eur. J. Soc. Psychol.*, vol. 40, no. 2, pp. 366–374, 2010.
7. B. Kim and J. Lee, "A video-based fire detection using deep learning models," *Appl. Sci.*, vol. 9, no. 14, 2019, doi: 10.3390/app9142862.
8. C. Liu and N. Ahuja, "Vision based fire detection," no. 1.
9. B. U. Töreyin, Y. Dedeoğlu, U. Güdükbay, and A. E. Çetin, "Computer vision based method for real-time fire and flame detection," *Pattern Recognit. Lett.*, vol. 27, no. 1, pp. 49–58, 2006, doi: 10.1016/j.patrec.2005.06.015.
10. [G. Marbach, M. Loepfe, and T. Brupbacher, "An image processing technique for fire detection in video images," *Fire Saf. J.*, vol. 41, no. 4, pp. 285–289, 2006, doi: 10.1016/j.firesaf.2006.02.001.
11. Z. Zhang, J. Zhao, D. Zhang, C. Qu, Y. Ke, and B. Cai, "Contour based forest fire detection using FFT and wavelet," *Proc. - Int. Conf. Comput. Sci. Softw. Eng. CSSE 2008*, vol. 1, pp. 760–763, 2008, doi: 10.1109/CSSE.2008.837.
12. B. C. Ko, K. H. Cheong, and J. Y. Nam, "Fire detection based on vision sensor and support vector machines," *Fire Saf. J.*, vol. 44, no. 3, pp. 322–329, 2009, doi: 10.1016/j.firesaf.2008.07.006.
13. T. Çelik and H. Demirel, "Fire detection in video sequences using a generic color model," *Fire Saf. J.*, vol. 44, no. 2, pp. 147–158, 2009, doi: 10.1016/j.firesaf.2008.05.005.
14. P. V. K. Borges and E. Izquierdo, "A probabilistic approach for vision-based fire detection in videos," *IEEE Trans.*

Circuits Syst. Video Technol., vol. 20, no. 5, pp. 721–731, 2010, doi: 10.1109/TCSVT.2010.2045813.

15. I. Mehmood, M. Sajjad, and S. W. Baik, *Mobile-cloud assisted video summarization framework for efficient management of remote sensing data generated by wireless capsule sensors*, vol. 14, no. 9. 2014.
16. A. Sorbara, E. Zereik, M. Bibuli, G. Bruzzone, and M. Caccia, "Low cost optronic obstacle detection sensor for unmanned surface vehicles," *SAS 2015 - 2015 IEEE Sensors Appl. Symp. Proc.*, 2015, doi: 10.1109/SAS.2015.7133652.
17. K. Dimitropoulos, P. Barmpoutis, and N. Grammalidis, "Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 2, pp. 339–351, 2015, doi: 10.1109/TCSVT.2014.2339592.
18. Q. Zhang, J. Xu, L. Xu, and H. Guo, "Deep Convolutional Neural Networks for Forest Fire Detection," no. Ifmeita, pp. 568–575, 2016, doi: 10.2991/ifmeita-16.2016.105.
19. K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, "Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 49, no. 7, pp. 1419–1434, 2019, doi: 10.1109/TSMC.2018.2830099.
20. R. Xu, H. Lin, K. Lu, L. Cao, and Y. Liu, "A forest fire detection system based on ensemble learning," *Forests*, vol. 12, no. 2, pp. 1–17, 2021, doi: 10.3390/f12020217.
21. Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
22. S. Srivastava, A. V. Divekar, C. Anilkumar, I. Naik, V. Kulkarni, and V. Pattabiraman, "Comparative analysis of deep learning image detection algorithms," *J. Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00434-w.
23. C. Y. Cao, J. C. Zheng, Y. Q. Huang, J. Liu, and C. F. Yang, "Investigation of a promoted you only look once algorithm and its application in traffic flow monitoring," *Appl. Sci.*, vol. 9, no. 17, pp. 1–14, 2019, doi: 10.3390/app9173619.
24. S. Shinde, A. Kothari, and V. Gupta, "YOLO based Human Action Recognition and Localization," *Procedia Comput. Sci.*, vol. 133, no. 2018, pp. 831–838, 2018, doi: 10.1016/j.procs.2018.07.112.
25. G. Plastiras, C. Kyrkou, and T. Theocharides, "You Only Look Once: Unified, Real-Time Object Detection," *ACM Int. Conf. Proceeding Ser.*, 2018, doi: 10.1145/3243394.3243692.
26. A. Wong, M. Famuori, M. J. Shafiee, F. Li, B. Chwyl, and J. Chung, "YOLO Nano: A Highly Compact You only Look Once Convolutional Neural Network for Object Detection," *Proc. - 5th Work. Energy Effic. Mach. Learn. Cogn. Comput. EMC2-NIPS 2019*, pp. 22–25, 2019, doi: 10.1109/EMC2-NIPS53020.2019.00013.