An efficient deep belief network for Detection of Coronavirus Disease COVID-19

Shaymaa Adnan Abdulrahman[1,2,*] , Abdel-Badeeh M. Salem[2]

[1]Department of computer Engineering, Imam Ja'afar Al-Sadiq University, Baghdad, Iraq
PhD Student at Ain Shams University, Egypt ; Shaymaaa416@gmail.com

[2]Department of Computer &Information Science, Ain Shams University, Cairo, Egypt
; absalem@cis.asu.edu.eg

Abstract

COVID-19 infection is one of the most dangerous respiratory viruses and the early detection of this disease reduces the speed of its spread among people . The goal of this virus is to infect the lung by creating white patchy shadows inside the lungs. This paper presents an intelligent method based on the deep learning technique to analyze the medical images of respiratory diseases . Two data set was used in this experiment first dataset is normal lungs taken from Kaggle data repository. While abnormal lungs taken from (https://github.com/muhammedtalo/COVID-19) .The results show that the proposed system identifies the COVID-19 cases with an accuracy of 90%.

Keywords: COVID-19 , machine learning , deep learning , X-ray , Image processing

1.Introduction

COVID-19 virus is nowadays a global pandemic and has caused many deaths in the world [1] . One of the problems caused by a virus is transmission of the disease from one person to another . The symptoms caused by the sick person are diseases of the respiratory system , as well as severe pneumonia and it may lead to respiratory failure [2] . Among the people most at risk of death are those who have heart disease , high blood pressure and diabetes according to the survey conducted by F. Zhou et al. [2] . This survey was 813 patients in Wuhan Chinese Hospital . Anyone who has this serious disease has special symptoms for this disease , which is the high temperature , Coughing , Phlegm and fatigue [3] . The doctor can determine the injury by symptoms of the affected patient and also by taking X-ray of the chest area .Through the X-ray images of affected lung contain multiple white patchy shadows [4] .

There are anumber of techniques that have been applied within the biometric area , especially the diagnosis of diseases . Researchers use artificial intelligence with medical images to distinguish between healthy and infected lung . deep belief network is an important AI tool which is applied for detection of diseases from the X-ray medical images .

According to (Tulin Ozturka,etal)[5] was prposed (DarkCovidNet) model to provide accurate diagnostics for classification . applied A COVID-19 X-ray image as dataset .This dataset conisist of 43 female and 82 male positive cases . CNN approach used for extracts features . DarkCovidNet model achieved 87.02% of accuracy. While (Mohamed et al)[6] analyzed seven deep learning models and studied performance for each model when applied two categories of classification such as binary and multiclass . Four classification used such as (Naive Bayes, Random forests., Artificial neural network and Support Vector Machine) In machine learning approache . CNN model gives the highest detection rate when used (CSE-CIC-IDS2018) as dataset . Data mining model used for the prediction of (COVID-19) infected patients. of South Korea was proposed by (L. J. etal)[7] . Six algorithms such as (SVM, logistic regression , K-nearest neighbor , naive Bayes, random forest and decision tree) were applied directly on the data-set . The results was that the model with decision tree data mining algorithm is more efficient to predict of the infected patients from COVID-19 with 99.85% of accuracy . While (Shuai Wang et al) [8] used 1,065 CT images of COVID-19 cases . accuracy , Sensitivity, PPV .NPV , Specificity and Area Under Curve (AUC) used as classification .

The main objective of this study is to identify the COVID-19 viruses from the chest X-ray image of a patient. Data-set was collected from (https://github.com/muhammedtalo/COVID-19) [5] . Several stages were used such as pre-processing when Applied Filter like (Gaussian filter) to reduce the artifact in this study and improve the accuracy level further , Segmentation process based on the gray level thresholding, to extract the important portion from the medical X-ray images . Finally used deep belief network architecture to classify normal and COVID-19 patients .

The structure of the paper is organized as follows. Section 2 Proposed Method , Data Acquisition , Pre-processing, Segmentation , Feature Extraction, 3. Section Result & Analysis , Section 4 Conclusion .

2. The Proposed Method

The COVID-19 detection depend on the chest (X-ray image) is a complicated images identification problems due to varying image quality across equipment used and patient. Figure 1 . refer to proposed work .

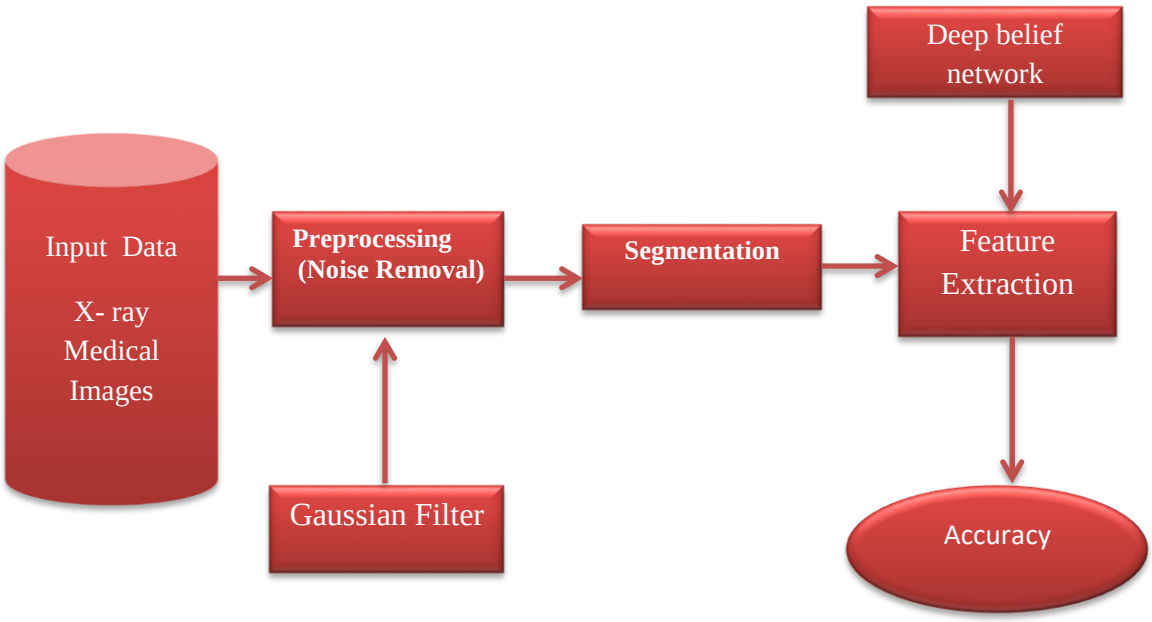

Figure 1: structure of proposed work

2.1 Data Acquisition

In this search used two dataset , first data is chest X-ray image dataset taken from kaggle. All medical images are for healthy people without lung disease . While second dataset (A COVID-19 X-ray image database) was developed by Cohen JP [9] using images, from various open access sources . This dataset consist of 125 X- ray images diagnosed with COVID-19. please see : (https://github.com/muhammedtalo/COVID-19) . Table 1 refer to Clinical symptoms of COVID-19 patients .

Table 1 : Clinical symptoms of COVID-19

Feature	Total Patients
Male	43
Female	82
average age of these subjects	approximately 55 years.

3.2 Pre-processing

Pre-processing step considered is essential in classification techniques to enhance the system accuracy. In pre-processing step we have taken away the noise present in the medical images . I n this case 2D Gaussian filter was applied on X-ray images . This is because these images are affected by the amount of noise present on the image sensitivity . The equation 1 represents the mathematical representation of a Gaussian 2D filter .

$$Gu(xi , yi) = \frac{1}{\sqrt{2\pi\delta}} \exp(-(x^2 + y^2)/2\delta^2) \ldots\ldots\ldots\ldots (1)$$

δ^2 refer to size of the filter kernel (-1<= xi , yi<=1) and is the variance . when applied this filter , signal to noise ratio has been enhanced [10] . Figure 2 refer to Remove noise from chest X-ray image by using Gaussian filter. First image is the original image with (6) dB SNR value, while second image is the filtered image with (10) dB SNR value.

Figure 2: Removal of noise by using 2D Gaussian filter. First image is the original image second image is the filtered image

3.3 Segmentation

Usually segmentation refer to the identification of objects represented in images . In segmentation stage , some mathematical application are used . These arithmetic methods depend on template matching, deformable models , statistical model , neuronal network Segmentation refers to the process of identifying and isolate the surface and region of the digital images which corresponds to the structural units . Segmentation may also depend on different types of features that are contained in the images , either colors or texture . Segmentation process have been developed to segment the image [11] . they are depend on the two properties such as discontinuity and similarity . In discontinuity depend on partition and subdivision is carried out based on abrupt changes in intensity level or grey level of an images . Various types of segmentation were applied to X-ray Image Dataset . Segmentation can be classified into the following categories such as Segmentation by using Edge Detection ,Segmentation by using Thresholding , Segmentation by using Region based , Segmentation by using Feature based Clustering .

Segmentation by using Edge Detection , it is first type of segmentation stage that was used in our experience . In this type of image segmentation method the first step is edge detection. It divides an images into object and its background [12][13]. There are two methods used to detect the edge , namely gray histogram and gradient . Usually edge detection operators can divided into two categories such as first order derivative operators , also second order derivative operators. Considered Second order operators give better results . Segmentation by using Thresholding can be applied either global or local. This type is the simplest way to segmentation images . depend on the intensity levels and is called as threshold based[21] . Usually globally threshold (adaptive thresholding) is distinguishing between object and background pixels through compare with threshold values chosen and use binary partition to segment the images . In this method , the threshold value differs over the images depend on the local features of the subdivide region in the x-ray images .

in general Segmentation is the methods of dividing an images into some sections with related properties such as contrast , brightness , gray levels, color . In this step , the segmentation process is on the basis chest X-ray images based on the gray levels thresholding as the, lungs of a COVID_19 patients contains many white patches. The goal of this stage is to transform all gray pixels into white pixels and remaining into black pixels , and at the end of this step , all areas of gray are separated from other . Figure 3 refer to two section , the lower section is for the infected lung , while the lower section is for lung of people who are not infected with the virus .

Figure 3: Segmentation process of COVID-19 and normal lungs

3.4 Feature Extraction

There are different type of deep learning algorithms such as convolutional neural network , deep conventional extreme machine learning , deep belief network , deep neural network , recurrent neural network [14] . Each type of network has its own mathematical operation and statistical properties that are different from the other . Deep belief network was applied in this step . Usually deep learning consist of feature detector, units arranged in multiple layers . Simple feature is extract through lower layers and put this feature into higher layers to extract more complex feature . This deep belief network model has visible layers and also hidden layers . The goal of this network is to discover the most complex feature which can reveal , hidden information and higher order correlations , of the data. Through figure 4 left part of figure vi,(i=1,2,3,4,……..) refer to vector of visible layers . While hi(i=1,2,3,4,….) refer to vector of the hidden layers . In addition , right part of figure is consist of a visible layer and a hidden layer [15] [16] . Also The number of the visible units in the lower network equals to the number of the hidden units in next higher network . In the training process of deep belief network model composed learning the networks one by one during which the learned features of one deep belief network are put into the next network as the ,input data.

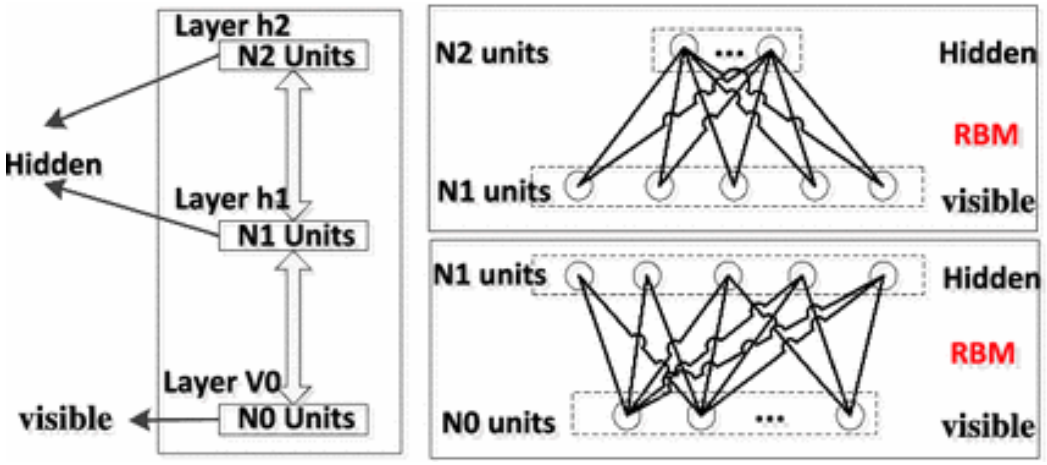

Figure 4: architecture of Deep belief network

3.4.1 Deep belief network with C-SLBP

The process of applying Deep belief network to realistic-sized images is very difficult. This is because pixie level COVID-19 images are high-dimensional which will cause high computationally complexity to the training process algorithm . Through these problems and complication in medical images like (COVID-19 images) it was suggested C-SLBP algorithm .

Center Symmetric -Local Binary Pattern (C-S LBP) proposed by Heikkila [17] . Local binary pattern LBP is an feature descriptor which has been proven, to be effective at texture. feature description . C-SLBP refer to it encode the change of the images from four different, direction with center symmetric principle . The C-SLBP features can be described with Equation . (2)

$$\text{C-SLBPR,N,T} \quad (X,Y)=\sum_{i=0}^{(\frac{N}{2})} \mathbf{S}\,(\mathbf{n_i}-\mathbf{n}_{\mathbf{i}+\frac{\mathbf{N}}{\mathbf{2}}})\mathbf{2^i} \qquad \text{……..(2)}$$

If S(x) =1 … X> T else otherwise , where N refer to the pixel numbers on a circle of radius R , in addition (ni) and (n i+(N/2)) correspond to the gray value of center symmetric area pixels[18][19] . In this case , the threshold for T is set to change the intensity of the image and for this is its durability enhancement on flat image regions . And by

comparing traditional LBP and C-SLBP , it find that the last (C-SLBP) has fewer dimensions and lower computational complexity also more robust to noise interference . Therefore C-SLBP is applied for feature extraction to preserve more useful data of the images and reduce the impact of noise such as pose variation. As shown in figure 5 , first step , the local feature of the input image is extracted by C-SLBP. Second step , the obtained feature is feed into the deep belief network instead of original COVID images as the input of the visible layers. While third step, training process the DBN from the bottom layers to the top layers[22,23.24, 25]. The goal of this process is, training the first layer's network ,parameters and use its output. as the second layer's input data. and so on . To obtain the best net, parameters the Back Propagation method is applied to, fine tune the trained Deep belief network . The final deep belief network approach consist of three layers and also number of iterations ,for each layers is twenty .

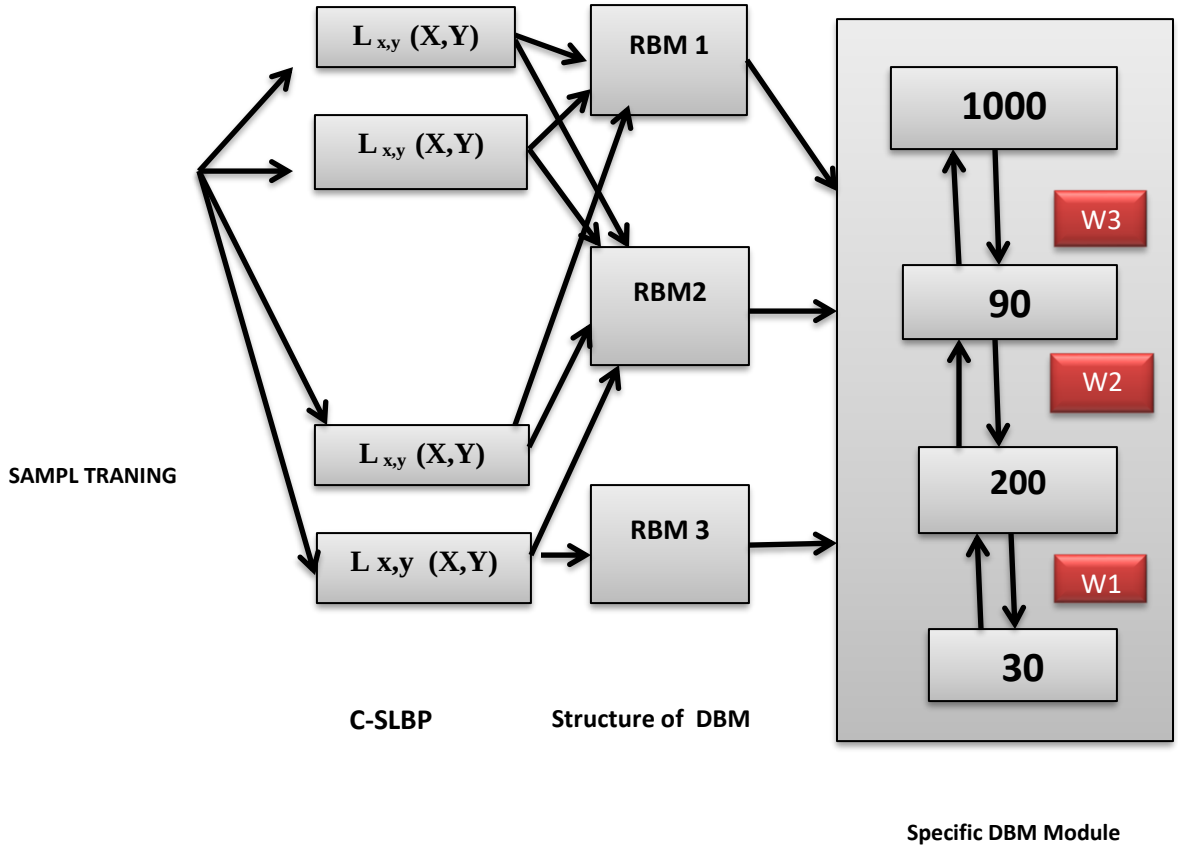

Figure 5: deep belief network with C-SLBP

4. Result & Analysis

Through our study of this research to identify the presence of Covid-19 dieses that affects the lung . Medical images were used such as chest x-ray images according to https://github.com/muhammedtalo/COVID-19. Usually the identification operation is based on the, presence of a white patchy, shadow in the lungs, .In order for to process the detection through these images (medical images) . It performed a processing of these images we did a process of processing these images through the process of removed the noise by using filter (Gaussian filter) . After that all the medical images were divided by segmenting the important lungs, section from the others. A threshold based segmentation is done to separate gray pixel from the others.

After the segmentation process , clearly separated the gray pixels from the white pixels and also clear that the number of gray pixels in Covid-19 affected lungs is very less compared to the normal lungs. Medical images were obtained for normal lungs for the purpose of conducting the segmentation process and making a comparison between the healthy

lung and the affected lung through (chest X-ray image dataset taken from kaggle.) . Features were extracted from , the segmented images using, the above mentioned deep belief network architecture. In the experiment, the data-set was divided into two group where 70% for training and the remaining 30% was applied for testing. Table 2 refer to the result of proposed system through training and testing process . To analyses the performance of our proposed system approach , it measures different parameters such as (accuracy, false positive, and false negative) by using the equations 3, 4, and 5 respectively.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{.....(3)}$$

$$\text{False Positive} = \frac{FP}{TP+TN+FP+FN} \quad \text{.....(4)}$$

$$\text{False Negative} = \frac{FN}{TP+TN+FP+FN} \quad \text{.....(5)}$$

Where TP, TN, FP, and FN refer to the number of True Positive case , True Negative case , False Positive case , and False Negative cases respectively.

Table 2 Performance analysis of the proposed system during training and testing

Parameter	Training	Testing
Accuracy	100	90%
False Positive	0	5
False Negative	0	6

Through the above Table , the process of comparing the proposed work with the previous existing work according to [20] we note that the proposed system is based on deep belief network with C-SLBP is able to detect the presence of corona virus in the lung and accuracy was 90% in addition less false positive and false negative rates . In [20] achieved less a curacy compared to proposed method .In addition another reason in previous work , their experiment was applied directly without working to remove noise from the image and segmentation approach .

5. Conclusion

Nowadays researchers are trying to detect to prevent the spread of corona virus infection . Currently , limiting the spread of the virus is a complex and difficult process . This experiment have presented a deep belief network architecture to classify normal and COVID-19 patients by using medical images of chest X-ray images. The suggested method was good , by comparing our work with the previous work , the result was good , and the accuracy was 90% . These result obtained in our study will help doctors or the health institution to make a more accurate decision .

References

[1] M. A. Mohammed et al., "Benchmarking Methodology for Selection of Optimal COVID-19 Diagnostic Model Based on Entropy and TOPSIS Methods," in IEEE Access, vol. 8, pp. 99115-99131, 2020, doi: 10.1109/ACCESS.2020.2995597.

[2] Zhou, Fei and Yu, Ting and Du, Ronghui and Fan, Guohui and Liu, Ying and Liu, Zhibo and Xiang, Jie andWang, Yeming and Song, Bin and Gu, Xiaoying and others, 2020. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet .
[3] Narin, Ali and Kaya, Ceren and Pamuk, Ziynet, 2020. Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. arXiv preprint arXiv:2003.10849 .
[4] Xu, Zhe and Shi, Lei and Wang, Yijin and Zhang, Jiyuan and Huang, Lei and Zhang, Chao and Liu, Shuhong and Zhao, Peng and Liu, Hongxia and Zhu, Li and others, 2020. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. The Lancet respiratory medicine .
[5] Tulin Ozturk , Muhammed Talo , Eylul Azra Yildirim , Ulas Baran Baloglu , Ozal Yildirim , U. Rajendra Acharya ," Automated detection of COVID-19 cases using deep neural networks with X-ray images" Computers in Biology and Medicine ,2020 , https://doi.org/10.1016/j.compbiomed.2020.103792
[6] Mohamed Amine Ferrag, Leandros Maglaras, Sotiris Moschoyiannis, and Helge Janicke , Deep Learning for Cyber Security Intrusion Detection: Approaches, Datasets, and Comparative Study , Journal of Information Security and Applications vol 50 , 2020
[7] L. J. Muhammad Md ,Milon Islam , Sani Sharif Usman , Safial Islam Ayon ," Predictive Data Mining Models for Novel Coronavirus (COVID-19) Infected Patients' Recovery" SN Computer Science , 2020
[8] Shuai Wang, Bo Kang , Jinlu Ma , Xianjun Zeng5, Mingming Xiao1, Jia Guo , Mengjiao Cai, Jingyi Yang, Yaodong Li, Xiangfei Meng2 , Bo Xu1," A deep learning algorithm using CT images to screen for Corona Virus Disease(COVID-19) , 2020
[9] Cohen JP (2020) COVID-19 image data collection. https://github.com/ieee8023/COVID-chestxray-dataset.
[10] Basu, Mitra ," Gaussian-based edge-detection methods-a survey" vol 32 ,number 3 , pp 252-260 , 2002
[11] Mohammed, M.A., Abdulkareem, K.H., Mostafa, S.A., Ghani, M.K.A., Maashi, M.S., Garcia-Zapirain, B., Oleagordia, I., Alhakami, H. and AL-Dhief, F.T., 2020. Voice Pathology Detection and Classification Using Convolutional Neural Network Model. Applied Sciences, 10(11), p.3723.
[12] Obaid, O.I., Mohammed, M.A., Ghani, M.K.A., Mostafa, A. and Taha, F., 2018. Evaluating the performance of machine learning techniques in the classification of Wisconsin Breast Cancer. International Journal of Engineering & Technology, 7(4.36), pp.160-166.
[13] Mohammed, M.A., Abd Ghani, M.K., Arunkumar, N.A., Mostafa, S.A., Abdullah, M.K. and Burhanuddin, M.A., 2018. Trainable model for segmenting and identifying Nasopharyngeal carcinoma. Computers & Electrical Engineering, 71, pp.372-387.
[14] Singh Anuj , Kumar Gupta, Bhupendra, " A novel approach for breast cancer detection and segmentation in a mammogram " Procedia Computer Science, vol 54 , pp 676-682, 2015 .
[15] Shaymaa Adnan Abdulrahman , Wael Khalifa , Mohamed Roushdy , Abdel Badeeh M. Salem , "Comparative study for 8 computational intelligence algorithms for human identification " 2020 ,https://doi.org/10.1016/j.cosrev.2020.100237
[16] Zebari, D.A., Haron, H., Zeebaree, S.R. and Zeebaree, D.Q., 2019, April. Enhance the Mammogram Images for Both Segmentation and Feature Extraction Using Wavelet Transform. In 2019 International Conference on Advanced Science and Engineering (ICOASE) (pp. 100-105). IEEE.
[17] Coates A, Ng AY (2011) Selecting receptive fields in deep networks. Adv Neural Inf Process Syst, 2528–2536
[18] Arunkumar, N., Mohammed, M.A., Mostafa, S.A., Ibrahim, D.A., Rodrigues, J.J. and de Albuquerque, V.H.C., 2020. Fully automatic model-based segmentation and classification approach for MRI brain tumor using artificial neural networks. Concurrency and Computation: Practice and Experience, 32(1), p.e4962.
[19] Khatami Amin , Khosravi, Abbas , Lim, Chee Peng , Nahavandi, Saeid " A wavelet deep belief network-based classifier for medical images, International Conference on Neural Information Processing, pp467-474 , 2016
[20]Wang, Linda and Wong, Alexander, 2020. COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images. arXiv preprint arXiv:2003.09871
[21] Abd Ghani, M.K., Mohammed, M.A., Arunkumar, N., Mostafa, S.A., Ibrahim, D.A., Abdullah, M.K., Jaber, M.M., Abdulhay, E., Ramirez-Gonzalez, G. and Burhanuddin, M.A., 2020. Decision-level fusion scheme for nasopharyngeal carcinoma identification using machine learning techniques. Neural Computing and Applications, 32(3), pp.625-638.
[22] Mohammed, M.A., Al-Khateeb, B., Rashid, A.N., Ibrahim, D.A., Abd Ghani, M.K. and Mostafa, S.A., 2018. Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images. Computers & Electrical Engineering, 70, pp.871-882.

[23] Mohammed, M.A., Abd Ghani, M.K., Hamed, R.I. and Ibrahim, D.A., 2017. Analysis of an electronic methods for nasopharyngeal carcinoma: Prevalence, diagnosis, challenges and technologies. Journal of Computational Science, 21, pp.241-254.
[24] Arunkumar, N., Mohammed, M.A., Abd Ghani, M.K., Ibrahim, D.A., Abdulhay, E., Ramirez-Gonzalez, G. and de Albuquerque, V.H.C., 2019. K-means clustering and neural network for object detecting and identifying abnormality of brain tumor. Soft Computing, 23(19), pp.9083-9096.
[25] Abdulkareem, K.H., Mohammed, M.A., Gunasekaran, S.S., Al-Mhiqani, M.N., Mutlag, A.A., Mostafa, S.A., Ali, N.S. and Ibrahim, D.A., 2019. A review of Fog computing and machine learning: Concepts, applications, challenges, and open issues. IEEE Access, 7, pp.153123-153140.