COVID-19 Detection using Medical Images

Shreya Shivhare [1], Roshani Yadav [2], Samiha Javed Rehman [3], Shreishth Agarwal [4] ,Ashutosh singh [5]

*[1,2,3,4] *Department of Computer Science and Engineering, Babu Banarsi Das Institute of Technology and Management, Lucknow , India*

[5] *Department of Computer Science and Engineering, Babu Banarsi Das Institute of Technology and Management, Lucknow , India*

Abstract : COVID-19 is considered to be the most dangerous and deadly disease for the human body caused by the novel coronavirus. In December 2019, the coronavirus spread rapidly around the world, thought to be originated from Wuhan in China and is responsible for a large number of deaths. Earlier detection of the COVID-19 through accurate diagnosis, particularly for the cases with no obvious symptoms, may decrease the patient's death rate. Chest X-ray images are primarily used for the diagnosis of this disease. This has proposed a machine vision approach to detect COVID-19 from the chest X-ray images. The features extracted by the histogram oriented gradient (HOG) and convolutional neural network (CNN) from X-ray images were fused to develop the classification model through training by CNN (VGG Net). Modified anisotropic diffusion filtering (MADF) technique was employed for better edge preservation and reduced noise from the images. A watershed segmentation algorithm was used in order to mark the significant fracture region in the input X-ray images. The testing stage considered generalized data for performance evaluation of the model. Cross-validation analysis revealed that a 5-fold strategy could successfully impair the overfitting problem. K-fold cross-validation demonstrated that the proposed feature fusion technique provided higher accuracy than the individual feature extraction methods, such as HOG or CNN.

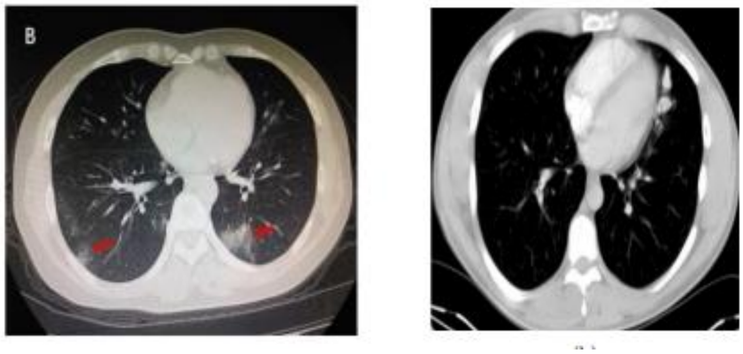
Keywords: X-ray image; convolutional neural network (CNN); histogram oriented gradient (HOG); watershed segmentation , COVID-19.

I. INTRODUCTION

Coronavirus Disease 2019 (COVID-19) is an infectious disease that started to proliferate from Wuhan China, in December 2019[1] . Within a short period of time, this disease is ravaged every corner of the world and the World Health Organization declared this disease as a pandemic on 11 March 2020[2]. This disease is caused by the strain of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). COVID-19 disease is increasing daily due to the lack of quick detection methods. All over the world, a huge number of people died of this disease in 2020. The respiratory tract and lungs are the media where the virus can spread easily. As a result, inflammation occurs, and air sacs can be filled with fluid and discharge. The process is responsible for creating an obstacle in oxygen intake. Quick and accurate detection of the virus is a major challenge for doctors and health professionals around the world in order to reduce the death rate caused by this virus. As a result, it has become important to detect the affected people earlier and isolate them to stop further spreading of this virus. Observing the properties of COVID-19 it can be asserted that this life-threatening virus can unfurl from individual to individual via cough, sneezing, or even close contact. As a result, it has become important to detect the affected people earlier and isolate them to stop further spreading of this virus.

In the past, artificial intelligence (AI) techniques were employed to successfully diagnose Pneumonia either from chest X-ray images or CT [3-5]. The classification methods employed vary from Bayesian function to convolutional neural network (CNN). More recently, CNN has been found to be useful and effective in identifying COVID-19 via image classification. CNN consists of multilayer neural networks, which are highly capable of recognizing the image patterns without conducting diverse preprocessing of the images. Although several CNN models, including Alex Net, Resnet50, VGG16, VGG19, are available, VGG19 demonstrates better performance for the COVID-19 classification [4-7].

The diagnosis of any disease is like the light at the end of the tunnel. In the case of the COVID-19 pandemic, the importance of earlier diagnosis and detecting the disease is beyond measure. The initial focus must be on the data by which we need to efficiently train a model. Due to the disadvantages of RT-PCR, researchers adopted an alternative method which is the use of Artificial Intelligence on chest CT or X-Ray images to diagnose COVID-19. Fundamentally, a chest CT image is an image taken using the computed tomography (CT) scan procedure where X-Ray images are captured from different angles and compiled to form a single image. A depiction of the CT images (COVID-19 infected and Normal) is illustrated in Figure 1.



(a) (b)

Figure 1 : CT Images (a)Covid-19 (b)Normal

The proposed work here provides an intelligent machine learning architecture in order to detect COVID-19 disease using chest X-ray images. The method proposes a novel fusion of features extracted by histogram-oriented gradient (HOG) and CNN and classification by CNN. Furthermore, a modified anisotropic diffusion filtering (MADF) technique was applied to eliminate multiplicative speckle noise from the test images. The watershed segmentation technique was used to identify the fractured lung regions, which could further provide evidence for the COVID-19 attacked lungs .

The remaining sections of this research work are formulated as follows: A comprehensive study on the Related Work done is presented in Section 2. Section 3 describes the methodology used to identify COVID-19 from X-ray image datasets by classifying it into the COVID-19 and other classes. Section 4 presents the results with analysis achieved from this work. Finally, conclusions drawn from this research and the future work are presented in Section 5.

II. RELATED WORK

In recent months, researchers have investigated and analyzed chest X-ray images using deep learning algorithms to detect COVID-19. First, the images are preprocessed using the CNN technique for extracting better features, which are fed in deep learning algorithms for image classification.

Ahammed et al. [8] proposed a deep neural network based system where CNN provided high accuracy (94.03%). The authors trained the system with normal, pneumonia and COVID-19 patient's chest X-ray images. The limitation of the work was that a dataset with only 285 images was used for developing the system, and this small number of data was not perfect for training a deep learning-based system for the COVID-19 prediction.
Chowdhury et al. [9] worked with chest X-ray images to develop a novel framework named PDCOVID Net based on parallel-dilated CNN. In the proposed method, the authors used a dilated convolution in the parallel stack that could capture and stretch necessary features for obtaining a detection accuracy of 96.58%.
Abbas et al. [10] proposed and validated a deep convolutional neural network called decompose, transfer, and compose (DeTraC) to detect COVID-19 patients from their chest X-ray images. They proposed a decomposition mechanism to check irregularities from the dataset by investigating class boundaries for obtaining a high accuracy (93.1%) and sensitivity (100%).
Azemin et al. [11] used a deep learning method based on the ResNet-101 CNN model. In their proposed method, thousands of images were used in the pre-trained phase to recognize meaningful objects and retrained to detect abnormality in the chest X-ray images. The accuracy of this method was only 71.9%.
Wang et al. [12] have developed a transfer learning method (Xception model) using deep learning models for diagnosing COVID-19. The proposed method showed 96.75% diagnostics accuracy. Furthermore, Deep features and machine learning classification (Xception + SVM) were also employed to develop an efficient diagnostic method for improving the accuracy of the Xception model by 2.58%.
From the result, the authors claimed that their proposed method attained higher classification accuracy and efficient diagnostic performance of the COVID-19. However, the authors have not compared their results with the existing similar works. Sahlol et al. [13] proposed an improved hybrid classification approach using CNNs and marine predators algorithm for classifying COVID-19 images, which were obtained from international cardiothoracic radiologists. Inception architecture of CNNs was employed to extract features, and a swarm-based marine predators algorithm was used to select the most relevant features from the images. However, the research work did not consider any fusion approach to improve the classification and feature extraction of the COVID-19 images.

III. METHODOLOGY

The chest X-ray images of the patients were acquired and stored in a commonplace. The images were categorized as either COVID-19-positive or negative as a reference to evaluate the performance of the intelligent system. In this work, three standard datasets were employed to validate the system's performance.

The proposed work here provides an intelligent machine learning architecture in order to detect COVID-19 disease using chest X-ray images. The method proposes a novel fusion of features extracted by histogram-oriented gradient (HOG) and CNN and classification by CNN. Furthermore, a modified anisotropic diffusion filtering (MADF) technique was applied to eliminate multiplicative speckle noise from the test images. First, the HOG technique was used to extract a feature vector from the X-ray COVID-19 dataset. Then the CNN method was used to extract another feature vector from the same images. These two features were fused and used as the input to train the classification model. The number of features extracted by one technique was not large enough to accurately identify COVID-19.

However, the fusion approach of extracting features by two different techniques could provide a large number of features for accurate identification. Fusion was considered as a concatenation between the two individual vectors in this context. Speckle-affected and low quality X-ray images along with good quality images were used in our experiment for conducting tests. If training and testing are performed with only selected good quality X-ray images in an ideal situation, the output accuracy may be found higher. However, this does not represent a real-life scenario, where the image database would be a mix of both good- and poor-quality images. Therefore, this approach of using different quality images would test how well the system can react to such real-life situations. A modified anisotropic diffusion filtering technique was employed to remove multiplicative speckle noise from the test images. The application of these techniques could effectively overcome the limitations in input image quality. Next, the feature extraction was carried out on the test images. Finally, the CNN classifier performed a classification of X-ray images to identify whether it was COVID-19 or not.

IV. RESULTS AND ANALYSIS

A. Filtering Performance: The proposed method used COVID-19 X-ray images as test data with different speckles, noises and resolutions. To work with meaningful features, information preservation and noise reduction are the prerequisite conditions to fulfill. The current system used modified anisotropic diffusion filtering (MADF) at the image preprocessing stage. The performance measurement in MADF was assessed using three evaluation metrics, namely Signal-to-noise ratio (SNR), minimum square error (MSE) and edge preservation factors (EPF) . Higher values of SNR and EPF represent more noise reduction and much edge details preservation, respectively. On the other hand, the minimum MSE value indicates less error between the input and filtered images. Classification models were run 10 times, and the highest values were reported for the performance metrics.

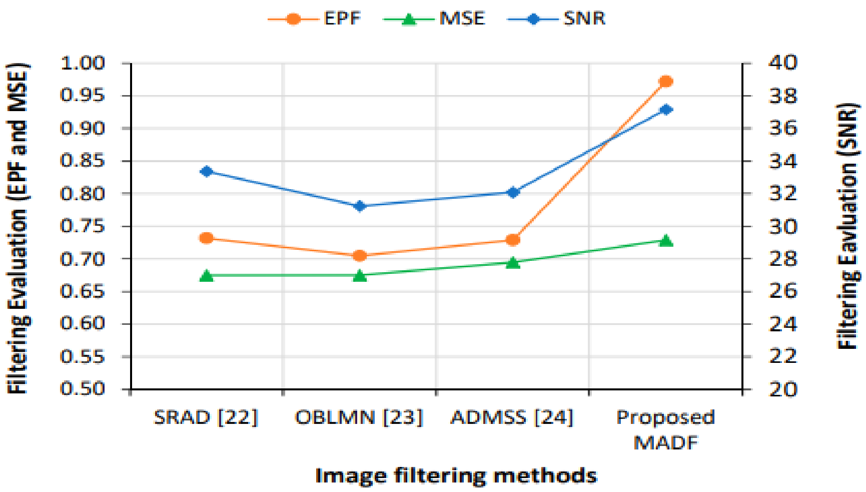


Figure 2 : Performance analysis of MADF technique

B. Feature Extraction Performance: CNN used extracted features to train before classifying. Test features were also found from the test images using different pre-trained models to measure the performance of CNN models. Nowadays, CNN uses different pre-trained models like Alex Net, ResNet50, VGG-16, VGG-19 and ResNet50 to extract features from the training and test data sets. All these models produce similar results for standard training and testing data. It was apparent that VGG19, which was proposed in this work, achieved better accuracy and specificity than the other CNN models, although ResNet50 showed the best performance in terms of sensitivity.

C. Classification Performance: This work proposed a fusion of feature vectors obtained by a combination of HOG and CNN techniques. This fusion vector was deliberated as the final input for the training and test datasets. Figure 3 presents a comparative study of different feature extraction approaches. The performances of different individual feature extraction techniques were less satisfactory than the fusion approach. This demonstrated that the proposed approach could classify COVID-19 cases more accurately than the single feature extraction approaches.

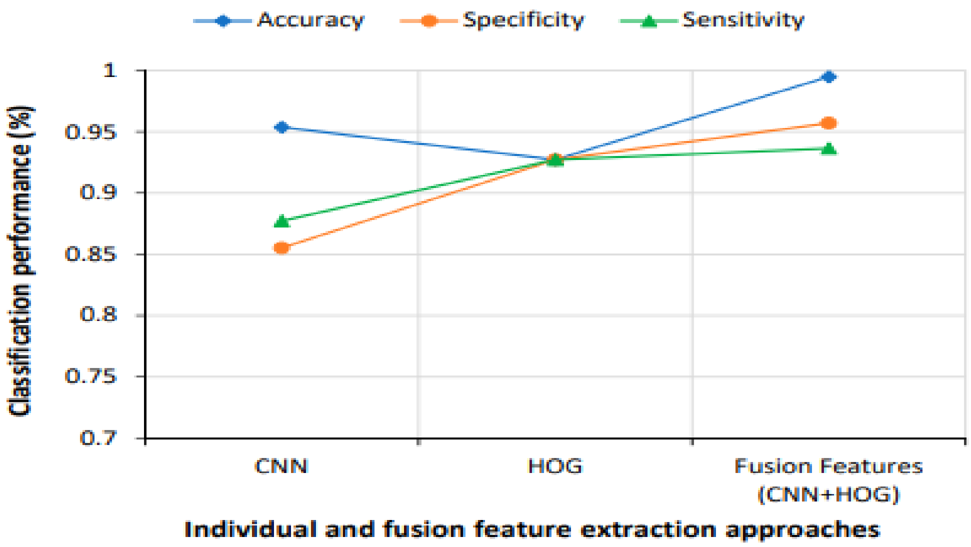


Figure 3 : Comparative results of individuals and fusion features.

The coronavirus pandemic has stretched the healthcare systems in every country in the world to its limit as they had to deal with a large number of deaths. Early detection of the COVID-19 in a faster, easier, and cheaper way can help in saving lives and reduce the burden on healthcare professionals. Artificial intelligence can play a big role in identifying COVID-19 by applying image processing techniques to X-ray images. This work designed and developed an intelligent system for the COVID-19 identification with high accuracy and minimum complexity by combining the features extracted by histogram-oriented gradient (HOG) features and convolutional neural network (CNN). Suitable feature selection and classification are absolutely vital in the COVID-19 detection using chest X-ray images. Chest X-ray images were entered into the system in order to produce the output of the marked lung significant region, which was used to identify COVID-19. The proposed feature fusion system showed a higher classification accuracy than the accuracies obtained by using features obtained by individual feature extraction techniques, such as HOG and CNN. Furthermore, the proposed fusion technique was validated with higher accuracies using generalization and k-fold validation techniques.

One of the limitations of this work was the imbalance of data in the datasets used for training and testing. In general, balanced data set with an equal number of normal and COVID-19 X-ray images makes the model building more comfortable, and the developed model can provide better prediction accuracy. Furthermore, the classification algorithm finds it easier to learn from a balanced dataset. This could be possible in the future when the continuous effort of collecting X-ray images will create a large balanced dataset to determine if deep learning can provide a solution in the fight against the COVID-19 pandemic.

VI. REFERENCES

[1]. C. Qin, L. Zhou, Z. Hu, S. Zhang, S. Yang, Y. Tao, C. Xie, K. Ma, K. Shang, W. Wang et al., "Dysregulation of immune response in patients with covid-19 in wuhan, china," Clinical Infectious Diseases, 2020.

[2]. "Who director-general's opening remarks at the media briefing on covid-19-11 march 2020." [Online]. Available:https://www.who.int/dg/speeches/detail/whodirector-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020.

[3]. Verma, S.G.; Prakash, S. Pneumonia classification using deep learning in healthcare. Int. J. Innov. Technol. Explor. Eng. .

[4]. Xiaowei, X.; Xiangao, J.; Chunlian, M.; Peng, D.; Xukun, L.; Shuangzhi, L.; Liang, Y.; Qin, N.; Yanfei, C.; Junwei, S.; et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia.

[5]. Shi, F.; Wang, J.; Shi, J.; Wu, Z.; Wang, Q.; Tang, Z.; He, K.; Shi, Y.; Shen, D. Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19. IEEE Rev. Biomed. Eng. 2020.
[6]. Wang, L.; Wong, A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest X-ray images. arXiv 2020, arXiv:2003.09871.
[7]. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 2014, arXiv:1409.1556.
[8]. Ahammed, K.; Satu, M.S.; Abedin, M.Z.; Rahaman, M.A.; Islam, S.M.S. Early Detection of Coronavirus Cases Using Chest X-ray Images Employing Machine Learning and Deep Learning approaches , medRxiv 2020. medRxiv 2020.06.07.
[9].Chowdhury, N.K.; Rahman, M.M.; Kabir, M.A. PDCOVIDNet: A parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images. Health Inf. Sci. Syst. 2020, 8.
[10].Abbas, A.; Abdelsamea, M.M.; Gaber, M.M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. Appl. Intell. 2021, 51, 854–864.
[11].Che Azemin, M.Z.; Hassan, R.; Mohd Tamrin, M.I.; Md Ali, M.A. COVID-19 Deep Learning Prediction Model Using Publicly Available Radiologist-Adjudicated Chest X-Ray Images as Training Data: Preliminary Findings. Int. J. Biomed. Imaging 2020, 2020.
[12].Wang, D.; Mo, J.; Zhou, G.; Xu, L.; Liu, Y. An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest X-ray images. PLoS ONE 2020, 15, e0242535.
[13].Sahlol, A.T.; Yousri, D.; Ewees, A.A.; Al-Qaness, M.A.; Damasevicius, R.; Abd Elaziz, M. COVID-19 image classification using deep features and fractional-order marine predators algorithm. Sci. Rep. 2020, 10.