Section A-Research paper ISSN 2063-5346

EB DETECTION OF COVID-19 FROM CHEST X-RAY IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

Sunaina Mehta¹, Charanjiv Singh²

¹Student, M.Tech, Department of Computer Science and Engineering, Punjabi University, Patiala, 147001, India

²Assistant Professor, Department of Computer Science and Engineering, Punjabi University, Patiala, 147001,

India

DOI: 10.48047/ecb/2023.12.si4.1536

Abstract:

The COVID-19 pandemic has a significant impact on the world, putting the health and lives of millions of people in danger. The difficulty in diagnosing COVID-19 at an early stage is a challenging task because many patients could not exhibit any symptoms. The pathogenic laboratory test like RT-PCR are inaccurate they can produce falsenegative results, which increases the risk of human life. The chest X-ray is most easy to capture and low cost radiological images used for detecting COVID-19. In this paper the Convolutional Neural Networks (CNNs) approach is applied over a deep learning system to extract features and detect COVID-19 from chest X-ray images. The datasets used in experiments consists of 3000 chest X-ray images, view images of 1000 normal, 1000 pneumonia, 1000 COVID-19 diseases from the various sources. In this paper the CXR images are resize, gaussian noise without clipping is added to the images for making the proposed model to be more robust and to improve ability for generalizing new data, denoising with DnCNN sigma25 is used to remove potentially clean images in the hidden layer. Then deep learning feature extraction techniques with Local Binary Pattern (LBP) and Gray Level Co-Occurrence Matrix (GLCM) are carried out for extracting the structural and textual characteristics of the chest X-ray images. This study evaluates the performance of trained model on unseen data and testing set is used to correctly identify the COVID-19 positive and negative cases from CXR images. The pretrained model is compare with other classification models to achieve the better performance parameters in COVID-19 detection. The results of proposed model shows the accuracy, F1 Score, Positive Predicted Value (PPV), Specificity, Sensitivity of 99%, 98%, 97%, 97%, 96% among all other classifier models in COVID-19 detection.

Keywords: COVID19, Chest X-Ray, Convolution Neural Networks

1. Introduction

In December 2019, Wuhan, China, reported the discovery of COVID-19, a corona viral disease that affects people's respiratory systems. Early symptoms of COVID-19 include fever, coughing, exhaustion, and muscle discomfort, and patients have abnormalities visible on their chest images on the CT and X-ray

machines. The first method that significantly contributes to the diagnosis of COVID- 19 disease is X-ray analysis [1]. Although the current gold standard for disease diagnosis is reverse transcription-polymerase chain reaction (RT-PCR), molecular testing of respiratory tract materials is also strongly advised because it enables laboratory confirmation of infections. A substantial difficulty has been created by the dramatic spread of COVID-19 because there aren't enough laboratory kits available[2]. As a result, during the COVID-19 outbreak, radiological examinations have gained more appeal as a means of diagnosing illnesses[3].

Although computed tomography (CT) scans have been shown to be more reliable, the rising patient population and ensuing increase in radiological examines make it impractical to rely on chest CT scans for each patient from the time of diagnosis until discharge. Chest X-rays (CXR) are a more practical choice for COVID-19 detection because a heavy reliance on CT scans will place a large burden on radiology departments. [4].

Machine learning algorithms, also known as classifiers, may accept input data, analyse it using statistical analysis, and forecast the outcome based on the type of data provided. They are used for resolving the issues including sentiment analysis, speech recognition, predictions, and picture and voice identification. The advancement of machine learning offers many benefits for making healthcare decisions and creating computer- aided systems[5].

The numerous studies on the detection of COVID-19 using artificial intelligence approaches. Recent studies have looked at the use of convolutional neural networks (CNNs) to automate the detection process in X-ray imaging which has emerged as a viable method for COVID-19 early detection. This study aims to evaluate the effectiveness of the most recent pretrained Convolutional Neural Networks (CNNs) in autonomously diagnosing COVID-19 using chest X-rays (CXRs). Chest X-ray picture data has been used in some deep learning-related studies to identify the disease. Deep Learning techniques that guarantees a higher degree of accuracy in terms of disease prediction and detection.

In recent years, it has been demonstrated that deep learning-based models—and more specifically convolutional neural networks (CNN)—outperform the traditional AI approaches in the majority of computer vision and medical image analysis tasks. CNN have been applied to a variety of issues, including classification, segmentation, face recognition, super-resolution, and image enhancement. This paper focuses on identifying COVID-19 infection in chest X-ray pictures using deep learning techniques.

Abiyev et al. [6] CNN models were used to classify the data using chest x-ray images of pneumonia and for the normal; the accuracy was 89.57% and 70% of the data was used for training. Chouhan et al. [7] used deep learning algorithms to identify the image of pneumonia from datasets images of typical, virus- and bacteria-caused pneumonia. They proposed a deep learning framework for pneumonia detection using transfer learning concept. In order to train five separate models, the image characteristics were retrieved using CNN models based on ImageNet. Then, using an ensemble model, all of the pretrained models were integrated. They were able to classify the data with 96.39% accuracy.

The paper is organized as follows. Section 2 presents the comparative analysis of COVID-19 X-rays images using machine learning and deep learning techniques, Section 3 describes the steps for proposed classification model. Section 4 discusses the materials and methodology used for feature extraction on data

set of chest X-ray images and comparison with previous study and proposed model for evaluating the performance parameters on different classification models with graphical analysis Section 5 presents the conclusion of this study.

2. Comparative Analysis of COVID-19 X-ray Images using machine learning and deep learning approach

X-ray modality is the first procedure to diagnose COVID19, which has the advantage of being inexpensive and low-risk from radiation hazards to human health. It takes a lot of time and effort for the radiologist to carefully identify the white areas on these images that are filled with pus and water. Another condition, like lung TB, may be misdiagnosed as COVID-19 by a radiologist or specialist physician [8].

Machine Learning (ML) approaches such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and random forests have been widely employed to effectively address a variety of computer science issues, from image processing to bio-informatics. The bulk of researchers, used machine learning approach for x-ray imaging as an open source data sets to forecast COVID-19 in people. Using a machine learning approach, we analyse the literature and find that there are certain problems with these medical images since they are more prone to noises like Gaussian noise and poison noise. The system's classification accuracy is directly impacted by these disturbances, which in turn makes it more difficult to extract information from large dataset.

Many deep learning techniques must be used in some of the most recent research on COVID-19 diagnosis. Deep learning was shown to play a substantial impact in the identification of COVID-19 affected patients using open source data sets of chest X-ray images. Convolutional neural networks (CNNs) have recently been proved to perform better than conventional artificial intelligence (AI) methods in the majority of computer vision and medical image processing tasks. CNNs have been used in a variety of problems, including classification, segmentation, face recognition, super-resolution, and image enhancement. We apply deep learning approaches to recognize COVID-19 and extract characteristics from chest X-ray images, which has considerably enhanced the diagnostic performance of X-ray.

Ioanni et al. [9] assess the most modern CNN architectures, including VGG19, MobileNet v2, Xception, Inception-v2 that have based on graphical analysis has been employed in recent years for medical picture Categorization.

Transfer learning is used by the author because effective at finding numerous anomalies in small medical i mage data sets. They used 1,442 patient X ray data sets 1200 individuals COVID-19 cases, 714 instances of bacterial and viral pneumonia and 714 cases of pneumonia caused by other pathogens. The findings demonstrate that MobileNetv2 and VGG19 provide the most accurate categorization among the remaining CNNs. While MobileNetv2 performs better in terms of sensitivity and specificity (reaching 99.10 and 97.09% respectively), VGG19 exceeds the other techniques in terms of accuracy (reaching 98.75%).

3. The Proposed Model

The proposed system for COVID-19 detection focuses on the chest X-ray images as a pre-processing phase. Further denoising of images and feature extraction using LBP and GLCM algorithm can be applied to trained deep learning model. The classification criteria is able to obtain the of best performance results in efficiently detecting the COVID-19 using deep learning techniques. The pre-processing phase involves several steps to prepare the COVID-19 images for classification. These steps include image resizing, noise addition, denoising using a deep learning pre-trained network, and feature extraction using LBP and GLCM methods. Figure. 1 shows the fundamental overview of the proposed model, which comprises of seven essential steps.



Figure 1: Proposed classification model

4. Materials and Methodology

In this section, we provide a brief overview of the strategy employed to accomplish the study's goals for COVID-19 images on preprocessing phase and classification phase. The proposed method's diagram can be shown in Figure 1. The methodology's steps for proposed system are follows as :

4.1 Data Acquisition: Chest X-ray image Datasets

The first step is to get the COVID-19 images from a trustworthy source. The images can be acquire from public databases, hospitals, or medical centers. It is important to ensure that the images are of good quality and representative of the actual COVID-19 cases. In our study, the datasets were collected from the COVID-19 radiography database by

Kaggle Community of chest X-ray images from along with Normal, Pneumonia, COVID-19 images https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database which was prepared by the tawsifur rahman The database has 3616 COVID-19 positive cases along with 10,192 Normal, 6012 instances of Lung capacity (Non- COVID lung infection), and 1345 Viral Pneumonia images and corresponding lung masks. In our study we have taken 3000 chest X-ray images.Then, the datasets are divided into for training and to test the classifiers. Figure. 2 shows an example of chest X-ray image Datasets .



Figure 2: Sample of chest X-ray Images

4.2 Preprocessing:

Preprocessing is an essential stage in image analysis since it has a big impact on how accurate the results are. It could be seen that all the acquired X-ray images are of variable shapes and sizes, contained various sources of interference such as sensor noise, error transmission which increases the difficulty in effective classification. In order to effectively perform classification tasks, these X-Ray images are preprocessed using data resizing, shuffling, and normalization. The first input size of all of the images are 299 \times 299 in the Portable Network Graphics (PNG) file type. In the preprocessing stage, these images are converted into the size of 128 \times 128 dataset to reduce the computational complexity analysis and enhance model performance for training and testing. The addition of noise enhances the resilience and venerability of the model to new data .

4.3 Denoising using Deep learning method:

Denoising is the advanced technique that is used for removing unwanted noise from an image. Deep learning-based denoising methods have shown to be highly effective in removing noise from images while preserving important features. A pretrained deep learning denoising method to remove the noise from the preprocessed imperfect COVID-19 images. The main focus on pre-processing part to reduce the impact of Gaussian noise.

4.4 Feature Extraction using GLCM and LBP:

The features are extracted from the preprocessed and denoised images. Two common feature extraction methods used in image analysis are Local Binary Patterns (LBP) and Gray Level Co-occurrence Matrix (GLCM). LBP is a computationally simple method that extracts local patterns from an image by comparing the intensity of each pixel

with its surrounding pixels. LBP method evaluate results in binary number and creates labels for image pixels. It is text measurement method independent of gray level. LBP is an effective approach in extracting features from an image by analyzing the local structure and pattern variations. On the other hand, GLCM is a matrix that contains information about the co-occurrence of pixel intensities in an image.

GLCM captures the spatial relationships between pixels by computing the co-occurrence probabilities of pairs of intensity values at different pixel distances and angles. GLCM extracts texture-based features such as contrast, entropy, variance, and other features effectively. These features are useful in analyzing the structural and textural characteristics of COVID-19 images. The basic advantages are describing the shape and contour properties of an image. Fig. 3 shows an results of GLCM and LBP feature extracted from an x-ray image of our proposed model.



Figure 3: Results of Original, Noisy Filtered, LBP (Normal Images)



Figure 4: Results of Original, Noisy, Filtered, LBP (COVID Images)



Figure 5: Results of Original, Noisy, Filtered, LBP (Pneumonia Images)

4.5 Separation of Dataset for training and testing:

The dataset was divided into different sections in order to evaluate how well the trained model performs on fresh data. While the training set is used to develop the model, the testing set is used to evaluate the model's performance on data that has never been seen before. To avoid the model being overfitting to the training data

and to ensure that it can generalize to new data, the dataset is split into training and testing sets.

4.6 Convolutional Neural Network (CNN) Training:

A CNN is an effective deep learning method that is trained on the preprocessed and featured extracted COVID-19 images [11] . In CNN, 784 neurons of hidden layers are utilized and 3 output layer neurons are used. Then, in hidden layer neurons, relu activate function was used and softmax activate function was employed in output layer neurons.



Figure 6: Proposed CNN Training Model

The images are passed through the CNN to learn the features and classify the images into COVID-19 positive or negative categories. The preprocessing and feature extraction of COVID-19 from the input images were both done using a CNN.

Input Layer (Flatten): This layer takes the input data and flattens it into a 1-dimensional array.

Hidden Layer (Dense): This layer consists of 784 neurons and uses the ReLU activation function. The number of neurons in this layer is set to the same as the input size (784 in this case), which is often used for fully connected layers.

Output Layer (Dense): This layer consists of 3 neurons, representing the number of classes or categories in your problem. It uses the softmax activation function to produce probability distributions over the classes.

The model is compiled with the following settings:

Loss Function: Sparse Categorical Crossentropy- suitable for multi-class classification problems where the target labels are integers.

Optimizer: Adam - a popular optimization algorithm known for its efficiency and robustness.

Metrics: Accuracy - used to evaluate the performance of the model during training and testing.

4.7 Performance Evaluation:

A number of metrics, including fine tuning and frozen layer in terms of classification, are used to evaluate CNN's performance. The CNN is utilized to investigate several test datasets to assess its generalization potential. The various metrics such as aforementioned criteria, including sensitivity, accuracy, specificity, recall, precision, and F1 score are used in this study to evaluate the model's performance in classifying COVID-19 images.

4.8 Performance Metrics

Various deep learning methods can be applied to assess the effectiveness of the classification model for COVID-19 identification from chest X-ray pictures. The trained model is retrieved using performance analysis based on the confusion matrix. The confusion matrix give the four outcomes of false negative (FN), false positive (FP), true negative (TN), and true positive (TP). The representation of both FNs and FPs may have a negative effect on medical decisions. An FP result is produced when someone is mistakenly assigned to a class, for as when a healthy person is incorrectly labelled as a COVID-19 patient. A FN occur when a person who should belong to a certain class is actually excluded from it. The performance of the various networks was assessed on the test set by calculating the accuracy (Acc), F1 score, precision (PPV), specificity (Spc), sensitivity (Sen), as quantitative assessment indices [12].

Mathematically calculated formulas are defined as:

$$Accuracy = \frac{TN+TP}{TN+FP+TP+FN}$$
(1)

Positive Predicted Value (PPV) =
$$\frac{TP}{TP+FP}$$
 (2)

F1 Score = $2 * \frac{Precision*Recall}{Precision+Recall}$	(3)	Section A-Research paper ISSN 2063-5346
Sensitivity = $\frac{TP}{TP+FN}$	(4)	
Specificity = $\frac{TN}{TN+FP}$	(5)	

5. Experimental Results

The implementation can be done using Python 3.10.3 programming language. In preprocessing stage, the acquired image data set's of 299×299 Portable Network Graphics (PNG) files are downsized to 128×128 before being fed to the neural network. These reduced images are applied for the denoising and feature extraction. The dataset is split into two groups for training (70%) and testing (30%) purposes, with the first group being used for training and the latter for testing the outcomes of the final evaluation. The parameters that are applied to the trained model in terms of classification is given below in Table 1.

Table 1: CNN Training Parameters

Parameter	Value
Training Data Size (%)	70
Testing Data Size (%)	30
Hidden Layer Neurons	784
Ouput Layer Neurons	3
Hidden Layer Activation Function	Relu
Ouput Layer Activation Function	Softmax
Optimizer	Adam
Loss	Cross Entropy
Performance Metrics	Accuracy
Epochs	25

The convolutional layer for deep learning model was built using Tensorflow. While Kera is used for preprocessing the images. The model employs the two activation functions ReLU and Softmax. The Rectified Linear Unit activation (ReLU) function are used for activating all hidden layers. The proposed model is compile for 25 epochs and set to ADAM optimizer for optimizing the loss function. The pretrained proposed model represent 0 as COVID-19, 1 as Normal, 2 as pneumonia for true label and predicted label .The extended overview of the classification performance of the proposed COVID-19 detection model is provided by confusion matrix in Figure 7. The confusion matrix display the consistency between predicted and true label results. This shows the evaluated pretrained model performance as well in classifying the effectiveness of identifying COVID-19 positive and negative cases.





Table 2 shows the obtained results for various parameters of our proposed model. The parameters include accuracy, PPV, F1 Score, Sensitivity and Specificity.

Table 2: Parameters Values for Proposed Model

Factors	Values

Accuracy	0.9919
PPV	1.0
F1 Score	0.9919
Sensitivity	0.9940
Specificity	1.0

We have acquired very promising results with respect to accuracy 0.9919, PPV 1.0, a specific F1 score around 0.9919, Sensitivity to 0.9940 and Specificity 1.0. The proposed model parameter values are used to evaluate the performance of various classification models. The trained model classify the COVID-19 images into two classes such as COVID-19 positive and negative. The proposed model achieved higher accuracy percentage of 99 and F1 score is 98 approximately among other models. Furthermore, the proposed model graphical evaluation on different classification models are achieved as better results for detection of COVID-19. The parameters values of proposed model are used to evaluate graphical performance of various classification models.

5.1 Comparison to Related Works

We compare the proposed model with studies of Mohammed et al. [13] used chest X-ray images to extract best features for classification process.

Table 3: Comparison Parameters of different CNNs by [13] for COVID-19 with best values of VGG16 in bold (%) with our proposed model

Classifier	Acc	<i>F</i> 1	PPV	Spc	Sen
InceptionV3	98.39	96.97	96.39	98.69	97.56
Xception	97.43	95.24	93.02	97.38	97.56
InceptionResNetV2	94.21	88.46	93.24	97.82	84.15
MobileNet	98.39	97.01	95.29	98.25	98.78
VGG16	98.71	97.59	96.43	98.69	98.78
DenseNet169	98.07	96.39	95.24	98.25	97.56
NasNetLarge	96.14	92.68	92.68	97.38	92.68
DenseNet121	95.50	91.14	94.73	98.25	87.81

Table 3 presents the performance of each CNN model concerning COVID-19. The parameters such as accuracy, F1, PPV, Specificity, Sensitivity are used to extract high-grade features from different classification models. The graphical results are analyzes by comparing the eight well known pretrained CNN models namely InceptionV3, Xception, IncepResNetV2, MobileNet, VGG16, DenseNet169, NasNetLarge, DenseNet121 to achieved an better performance results with our proposed model.



Figure 8: Accuracy comparison of different models

In figure 8 the comparison of accuracy of different models is shown. The x-axis demonstrate the different models like InceptionV3, Xception, IncepResNetV2, MobileNet, VGG16, DenseNet169, NasNetLarge, DenseNet121 and the proposed model. The y-axis represents the percentage of accuracy. From the above graph, it is observed that the InceptionV3 model achieved accuracy of 97% and VGG16 model has attained accuracy percentage of 98. It seems from the figure that the proposed model is more accurate as compared to other models. The proposed algorithm achieved accuracy percentage of 99% approximately.



Figure 9: F1 Score comparison of different models

The given figure 9 represents the comparison between F1 score on various model. The bar graph demonstrates the different algorithms and the percentage of F1 score. The results obtained from the bar graph demonstrate that the different model has achieved different percentage of F1 score. The proposed model attained percentage of F1 score is 98 approximately. Also, the proposed model performed better than other model.



Figure 10: PPV Comparison of different models

The above figure 10, represents the PPV comparison on various models . On the horizontal axis, different models are shown in figure and the vertical axis illustrates the percentage of PPV. In above graph, VGG16 has attained PVV percentage of 85% and the NasNetLarge achieved lowest PVV percentage. Moreover, the proposed has attained highest percentage of PVV 97% approximately



Figure 11: Specificity Comparison of different models

In figure 11, it illustrates the comparison of Specificity on different model. The bar graph represents the different models and the specificity percentage. In this bar graph, there was a slightest difference of specificity percentage in the various models like InceptiveV3, IncepResNetV2, MobileNet, VGG16, DenseNet169, NasNetLarge,

DenseNet121 and the proposed model. It seems from the figure that the proposed model has achieved specificity percentage of 97% approximately. It is observed that the proposed model performed well in terms of specificity than other models.



Figure 12: Sensitivity comparison of different models

The given figure 12 illustrates the comparison of sensitivity on various models. The x-axis shows the various models and the y-axis shows the percentage of sensitivity. The x-axis demonstrates that how much percentage is attained by different models. It is clearly observed that the IncepResNetV2 has achieved lowest percentage of 82% approximately and the proposed algorithm utilized highest sensitivity percentage i.e. 96 approximately.

6. Conclusions

The COVID-19 pandemic has significantly challenged to healthcare systems around the world, and medical imaging has emerged as a crucial tool for treating and identifying COVID-19 patients. In this paper, we proposed a model for detecting COVID-19 with the deep learning techniques to extract the most useful features from Chest X-ray images. The preprocessing phase include resizing for all images to same size, adding Gaussian noise, denoising of images using pretrained trained network to remove noise, and extracting feature from the preprocessed denoised images using LBP and GLCM methods to analyze the structural and textual characteristics the chest X-ray images. The training set is used to trained CNN model on unseen data for classifying the COVID-19 images either positive or negative. The proposed model evaluate the performance parameters values of accuracy 0.9919, PPV 1.0, specific F1 score 0.9919 , Sensitivity to 0.9940 and Specificity 1.0. The graphical results of proposed model shows the accuracy 99%, PPV 97% , F1 Score 98% , Sensitivity 96% and Specificity 97% respectively with different pretrained CNN models. This study present that deep learning techniques are significant marker for extracting the COVID-19 disease.

References

- [1] S. Hussain and A. Khan, "Coronavirus Disease Analysis using Chest X- ray Images and a Novel Deep Convolutional Neural Network," no. April, 2020, doi: 10.13140/RG.2.2.35868.64646.
- [2] M. Chung, A. Bernheim, X. Mei et al., "CT imaging features of 2019 novel coronavirus (2019-nCoV)," Radiology, vol. 295, no. 1, pp. 202–207, 2020.
- [3] H. Shi, X. Han, N. Jiang et al., "Radiological findings from 81 patients with COVID-19 pneumonia in

Wuhan, China: a descriptive study," The Lancet Infectious Diseases, vol. 20, no. 4, pp. 425–434, 2020.

- [4] P. Sun, X. Lu, C. Xu, W. Sun, and B. Pan, "Understanding of COVID-19 based on current evidence," Journal of Medical Virology, vol. 92, no. 6, pp. 548–551, 2020.
- [5] P. K. Sethy, S. K. Behera, P. K. Ratha, and P. Biswas, "Detection of coronavirus Disease (COVID-19) based on Deep Features and Support Vector Machine," Int. J. Math. Eng. Manag. Sci., vol. 5, no. 4, pp. 643–651, 2020, doi: 10.33889/ijmems.2020.5.4.052.
- [6] R. H. Abiyev and M. K. S. Ma'aitah, "Deep convolutional neural networks for chest diseases detection," J. Healthc. Eng., vol. 2018, 2018.
- [7] V. Chouhan et al., "A novel transfer learning based approach for pneumonia detection in chest X-ray images," Appl. Sci., vol. 10, no. 2, p. 559, 2020.
- [8] Das, A.K., Ghosh, S., Thunder, S. et al. Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network. Pattern Anal Applic 24, 1111–1124 (2021). https://doi.org/10.1007/s10044-021-00970-4
- [9] Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43(2):635–640
- [10] M. E. H. Chowdhury, T. Rahman, A. Khandakar et al., "Can AI help in screening viral and COVID-19 pneumonia?," vol. 8, pp. 132665–132676, 2020.
- [11] Sarki, Rubina, et al. "Automated detection of COVID-19 through the convolutional neural network using chest x-ray images." Plos one 17.1 (2022): e0262052.
- [12] D. M. Powers, "Evaluation: from precision, recall and F- measure to ROC, informedness, markedness and correlation,"2020, https://arxiv.org/abs/2010.16061.
- [13] Mundher Mohammed Taresh, Ningbo Zhu Talal Ahmed Ali Ali, Assad Shakir Hameed and Modhi Lafta Mutar "Transfer Learning to Detect COVID-19 Automatically from X-Ray Images Using Convolutional Neural Networks" International Journal of Biomedical Imaging, Volume-2021, ArticleID-828404, https://doi.org/10.1155/2021/8828404