# COVID-19 and Viral Pneumonia Chest X–ray images Classification using a Deep Ensemble based learning Model



**Lakshmidevi N,**
Assistant Professor, Department of CSE, GMR Institute of Technology, Rajam, Vizianagaram District, Andhra Pradesh. Email ID: lakshmidevi.n@gmrit.edu.in

**MSR Naidu**
Associate Professor, Department of ECE, Aditya Institute of Technology and Management, Tekkali, Srikakulam District, Andhra Pradesh., Email ID: naidu.msr@adityatekkali.edu.in

**Anil Kumar B**
Sr.Assistant Professor, Department of ECE, GMR Institute of Technology, Rajam, Vizianagaram District, Andhra Pradesh., Email ID: anilkumar.b@gmrit.edu.in

## Abstract

A fatal viral disease that has killed thousands and infected millions around the world is the unique Corona virus disease 2019- COVID-19. The usefulness of artificial intelligence in the quick and accurate identification of COVID-19 from X-Ray pictures is demonstrated in this work. It is a difficult issue for both radiologists and medical academics to classify the viral COVID-19. In recent years, the classification of images using deep learning has grown in popularity. CNNs have demonstrated their mastery in extracting a vast array of machine-generated information. By utilising CNN for feature extraction and classifiers for classification on publically available datasets, the chest X-ray pictures are categorised. A Deep Neural Network with an ensemble structure called CoEnNet is trained with various hyperparameters, and its performance is compared. The proposed model must provide notable performance in terms of overall accuracy with 98% when compared to other machine learning and pre-trained CNNs models.

**Keywords:** CNN, Ensemble learning, feature extraction, pre-trained network, transfer learning, covid-19, Viral Pneumonia, Chest X–ray.

**DOI Number: 10.14704/nq.2022.20.13.NQ88080** **Neuro Quantology 2022; 20(13):601-608**

## I.Introduction

Due to its rapid global spread and high level of contagiousness, the World Health Organization (WHO) declared the coronavirus infection (COVID-19) pandemic on March 11, 2020. [1]. Extremely contagious, it can cause multiple organ failure or respiratory distress in severe cases. [2]. The condition can be identified using the reverse transcription polymerase chain reaction (RT-PCR) assay. [3]. Even though RT-PCR is the gold standard for diagnosis, it can be challenging to identify and treat putative patients due to both high false-negative rates and weak sensitivities when COVID-19 individuals are being verified. [4]. Although real-time reverse transcriptase polymerase chain reaction (RT-PCR) has been the gold standard for SARS-CoV-2 diagnosis, the extremely low supply and strict requirements for laboratory environments have caused significant delays in the accurate diagnosis of suspected patients, posing unprecedented challenges in preventing the spread of the infection, particularly in epidemic hotspots.

Chest computed tomography (CT) is a quicker and easier way to make a clinical diagnosis of COVID-19 by combining the patient's clinical symptoms and signs with their recent close contacts, travel history, and laboratory findings. This enables a quick diagnosis as early as possible in the clinical practise. In areas where the pandemic is severe, it is also helpful for immediately isolating sick persons and containing the epidemic. A non-invasive imaging technique called computed tomography (CT) can identify features in the COVID-19-infected lung,

such as bilateral patchy shadows or ground glass opacity (GGO) [5]. In light of this, CT may be a crucial technique for the early discovery and diagnosis of COVID-19 patients. Although CT has many advantages, it may also have imaging characteristics in common with COVID-19 and other types of pneumonia, making automated identification difficult.

Recently, deep learning-based artificial intelligence (AI) technology has demonstrated great success in the field of medical data analysis due to its capacity to extract rich features from multimodal clinical information. [6]. Using thoracic imaging data, deep learning was previously utilised to identify and classify bacterial and viral pneumonia. [7]. Other chest CT imaging characteristics have also been tried to be found. [8]. Although real-time reverse transcriptase polymerase chain reaction (RT-PCR) has been the gold standard for SARS-CoV-2 diagnosis, the extremely limited supply and strict requirements for laboratory environments have caused significant delays in the accurate diagnosis of suspected patients, posing unprecedented difficulties in preventing the spread of the infection, particularly in epidemic hotspots [9].

Chest computed tomography (CT) is a quicker and easier way to make a clinical diagnosis of COVID-19 by combining the patient's clinical symptoms and signs with their recent close contacts, travel history, and laboratory findings. This enables a quick diagnosis as early as possible in the clinical practise. In areas where the pandemic is severe, it is also helpful for immediately isolating affected individuals and containing the epidemic [10]. In this study, we will examine the diagnosis and classification of COVID-19 patients' illnesses as well as the use of chest CT imaging to identify infected areas. Because of its proficiency and accuracy in image categorization, convolution neural networks (CNN) are commonly employed. There are three methods for classifying images using CNNs.

1. Modify the output layer of the pre-trained network and freeze the entire feature extraction portion to do transfer learning.
2. Any pre-trained network can be fine-tuned by changing some of its feature extraction layer's layers.
3. Training the ensemble based CNN from scratch

## II.Related Work

The pre-processing, feature extraction, feature dimension reduction, and final point classification are some of the common phases in machine learning algorithms for classification. Professional experience with a specific domain typically determines the feature selection and extraction. For non-professionals, using machine learning techniques in research investigations is a challenging task. The accuracy of classification is impacted by features taken into consideration, therefore feature selection and extraction is a crucial stage in machine learning methodologies.

Two categories can be used to classify an image's features. The first one comes directly from the image itself and is called local features. For instance, textural characteristics, intensity, first- and second-order statistics

The second category is global features, such as fisher vectors, scale-invariant feature transformations, bag-of-words, etc. Large well-labelled datasets with illustrative data distribution features are essential for data driven learning in order to develop more accurate models. [11]. The CNN models that were trained on the source database greatly enhance a variety of image segmentation and classification goals for additional target datasets. [12]. Predictable image classifiers were created for item identification in natural photos, despite the fact that there is a larger semantic difference between medical images and non-medical images. Pre-trained CNNs trained on the ImageNet dataset have been utilised to identify and detect chest disease in X-ray and CT modalities. [13]. Transfer learning refers to the idea that classification knowledge gained from training on a source picture dataset can be applied to additional target images. There are two additional transfer learning methods that can be used. The first step is to classify the photos using a different classifier and extract the features using pre-trained CNN [14].

## III.Proposed Method

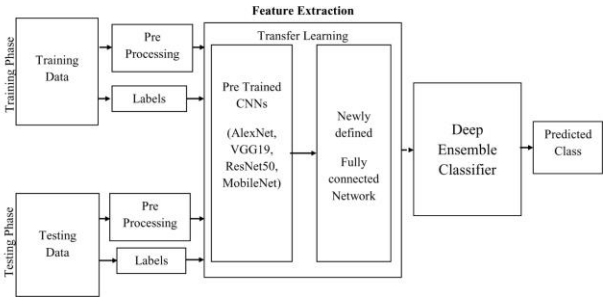



Figure 1: Training Phase and Testing Phase of proposed model

In Figure 1, the training and testing phases of the CovEnNet model are shown. The model is first loaded with pre-processed pictures and labels from the Chest X-Ray dataset, and then the dataset is divided into training, validation, and testing datasets. The training and validation sets are fed through augmentation to fit into the proposed CovEnNet during the training phase. The structure of the proposed architecture is trained with enhanced training data and validated with validation data at each epoch. Using training parameters including the number of epochs, mini batch size, regularisation, learning rate, and other factors, the training is carried out. The testing phase involves putting the trained model to the test on a testing dataset and validating it with performance metrics.

### 3.1. **Dataset**

A deep learning model needs a lot of data to be trained. Such datasets containing a significant amount of labelled data are uncommon in the field of medical imaging.

The dataset for this study was compiled from the 10,316 chest X-ray images of Covid -19 impacted, normal, and pneumonia in the Kaggle repository.

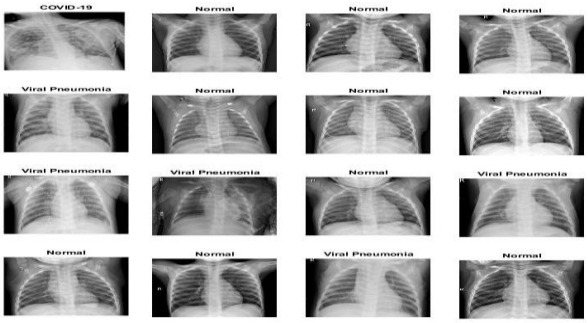



Figure 2: Sample dataset from Chest X-Ray Images

The folders are separated into three categories: training, validation, and testing. To prevent overfitting, 80 percent of the photos are utilised as training data, 10 percent are used as validation, and the final 10 percent are used for testing.



Table 1: Details of Collected Dataset

| S. No | Type of Disease | No. of Images |
|---|---|---|
| 1. | Normal | 3800 |
| 2. | Covid-19 | 3616 |
| 3. | Viral Pneumonia | 2900 |

### 3.2. **Pre-processing**

Three pre-processing procedures are employed to fit the dataset to the proposed CovidNet model: de-noising, data augmentation, and intensity normalisation. The Dn-CNN and median filters are used to remove noise from the Chest X-Ray dataset, To match the dataset to the suggested CovidNet model, three pre-processing techniques are used: de-noising, data augmentation, and intensity normalisation. The Chest X-Ray dataset is provided in PNG format, and noise is removed using the Dn-CNN and median filters. This de-noising method reduces computation time by removing high frequency distortions and Gaussian noise from images.

Data augmentation is the main component of transfer learning pre-processing. Other strategies include resizing, tossing, situations, incorporating salt and pepper noise, streamlining, Translation, Rotation, and Perspective Transform.

These datasets' intensity levels don't always mean the same thing. The dataset's observations show that the intensity levels of X-Ray images vary among and between people. Therefore, normalisation is crucial for datasets in order to avoid putting the network in a bad state.

min-max normalization is used as pre-process the datasets to balance the intensity value to [0, 1] .

$$y_i = (x_i - \min(x)) / (\max(x) - \min(x)) \quad (1)$$

Where max(x) and min(x) are the maximum and minimum intensity levels for the entire image, and yi is the value of normalised intensity with respect to xi.

### 3.3. CNN Architecture

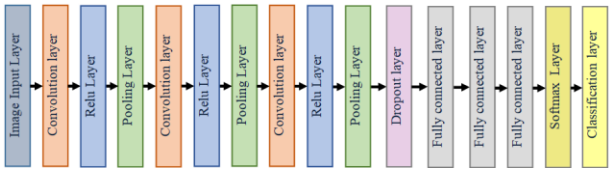



Figure 3: CNN Architecture

The CNN design is depicted in Figure 3 along with an explanation of each layer. The first layer is an image input layer, which accepts an image with a defined input size and performs normalisation.

CovEnNet computes the dot product of the filter and the input throughout the full input image and is composed of three convolution layers, each with a K number of filters and a filter size of M X N. As the filters move horizontally along the input image, the Stride (S) is referred to vertically. Before the filters are applied, padding (P) of the original photographs can be done to keep the information to the edges. Lower-level layers will take on low-level features like lines and edges, whereas higher-level layers will take on high-level features.

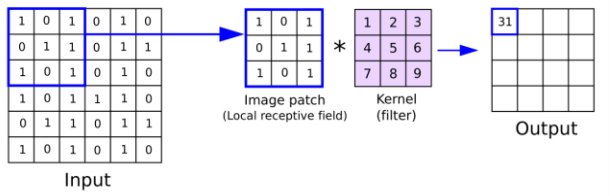



Figure 4: Convolution Layer

Figure 4 shows an example of a convolution layer with kernel and output feature map. A kernel of size 3 X 3 is applied on 6 X 6 input image and got 4 X 4 feature map. The output feature map size depends on the following equation.

$$\frac{W+2P-f}{S}+1 \text{ X } \frac{H+2P-f}{S}+1 \quad (2)$$

Where P is for pooling, F is for kernel size, and S is for stride value, W and H stand for the width and height of the input picture.

The Softmax activation layer will reduce the likelihood of predicted classes from 0 to 1 by half. The sum of all predicted values is one. The output of the softmax layer is computed as follows.

$$y(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{k} e^{z_k}} \quad (3)$$

Figure 5 shows how the chance of any class (j) can be roughly expressed as a function of y (z) over (k) different classes, and their sum equals one. [17].



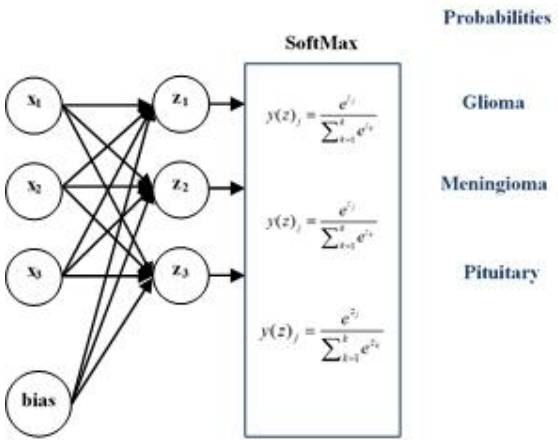



Figure 5: Softmax layer

## IV.Results and Discussions

Validation measures are employed to assess the performance of various deep learning networks. The following validation indicators were taken into account while assessing how well the proposed CovEnNet model performed. The performance measures based on TP, FP, TN, and FN are as follows.

$$\text{Accuracy} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (4)$$

$$\text{Sensitivity} = \frac{T_P}{T_P + F_N} \quad (5)$$

$$\text{Specificity} = \frac{T_N}{T_N + F_P} \quad (6)$$

$$\text{Precision} = \frac{T_P}{T_P + F_P} \quad (7)$$

$$\text{F1-Score} = 2\frac{PPV * TPR}{PPV + TPR} \quad (8)$$

Table 2 displays the overall performance classification for normal, COVID-19, and viral pneumonia in terms of Accuracy, Sensitivity, Specificity, Precision, and F1-Score. In comparison to previous deep learning networks as Alex-Net, VGG-16, VGG-19, Resnet 101, Resnet 18, MobileNetV2, DenseNet201, and CheXNet, our proposed technique demonstrated the greatest results. Additionally, our strategy considerably increased the sensitivity to the COVID-19 example. Fig. 6 displays a graph of loss and accuracy.

Here x-axis refers to the number of model training epochs which refer to the number of training cycles through the full dataset and y-axis indicates the loss and accuracy respectively. If the accuracy graph is observed closely, in the beginning, validation accuracy is higher than the training accuracy for some epochs.

Both the validation and training accuracy curves follow an upward trend as the number of epochs increases.

Here x-axis refers to the number of model training epochs which refer to the number of training cycles through the full dataset and y-axis indicates the loss and accuracy respectively. If the accuracy graph is observed closely, in the beginning, validation accuracy is higher than the training accuracy for some epochs.

Both the validation and training accuracy curves follow an upward trend as the number of epochs increases.

Here x-axis refers to the number of model training epochs which refer to the number of training cycles through the full dataset and y-axis indicates the loss and accuracy respectively. If the accuracy graph is observed closely, in the beginning, validation accuracy is higher than the training accuracy for some epochs.

Both the validation and training accuracy curves follow an upward trend as the number of epochs increases.

Here x-axis refers to the number of model training epochs which refer to the number of training cycles through the full dataset and y-axis indicates the loss and accuracy respectively. If the accuracy graph is observed closely, in the beginning, validation accuracy is higher than the training accuracy for some epochs

Here x-axis refers to the number of model training epochs which refer to the number of training cycles through the full dataset and y-axis indicates the loss and accuracy respectively. If the accuracy graph is observed closely, in the beginning, validation accuracy is higher than the training accuracy for some epochs

The Y-axis shows the loss and the X-axis the accuracy. The X-axis represents the number of model training epochs, or training cycles through the entire data set. With more epochs, both the training accuracy and the validation curve show an increased trend. The loss rate gradually decreases until it eventually ceases. On the other hand, Fig. 7 displays the suggested model's confusion matrix, which demonstrates how many test images are correctly identified. For a total of 60 photos, we manually separated 20 images from each class. 60 photos were used, and 59 were properly predicted, achieving a remarkable performance accuracy of 98.3%. Individual comparisons of performance metric parameters including accuracy, sensitivity, specificity, precision, and F1-score of normal are shown in Tables 4,5,6,7, and 8, COVID-19 and viral pneumonia which were 98.3%, 100%, 97.5% ,95.2% and 97.9. The comparative performance for different CNNs with proposed model (Covid Net) is shown in Table. 9.

| S.NO | Transfer learning Model | Accuracy (%) | Sensitivity (%) | Specificity (%) | Precision (%) | F1 Score |
|---|---|---|---|---|---|---|
| 1. | Alex Net | 62.7 | 52 | 91.5 | 82 | 63.5 |
| 2. | VGG16 | 95.3 | 92 | 99.95 | 96.73 | 94.3 |
| 3. | VGG19 | 97.5 | 92 | 99.1 | 92 | 92 |
| 4. | ResNet101 | 95.7 | 89.6 | 89.7 | 92.3 | 90.9 |
| 5 | Resnet18 | 96.4 | 96.4 | 96.5 | 90.5 | 93.3 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 6. | MobileNetV2 | 96.1 | 96 | 96.3 | 90.5 | 93.2 |
| 7. | DenseNet201 | 98.7 | 99.1 | 99 | 95.5 | 97.3 |
| 8. | Chex Net | 98.1 | 96.9 | 97 | 93.4 | 95.6 |
| 9. | **CovEnNet** | **99.3** | **100** | **99.5** | **97.2** | **99.9** |



Table 2: Performance metrics of different deep learning networks in terms of Accuracy, Sensitivity, Specificity, Precision and F1 Score

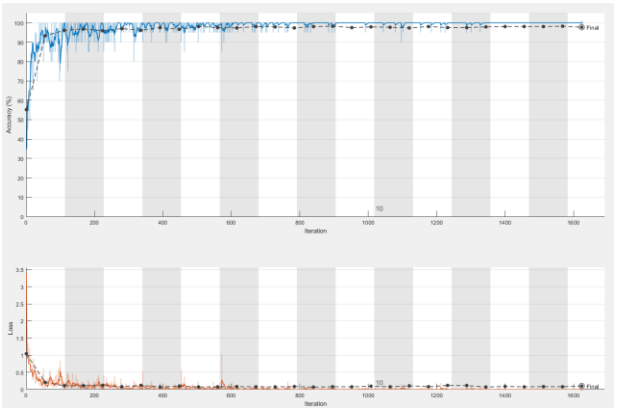
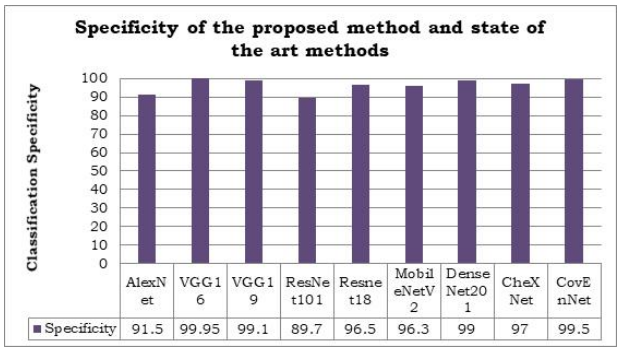
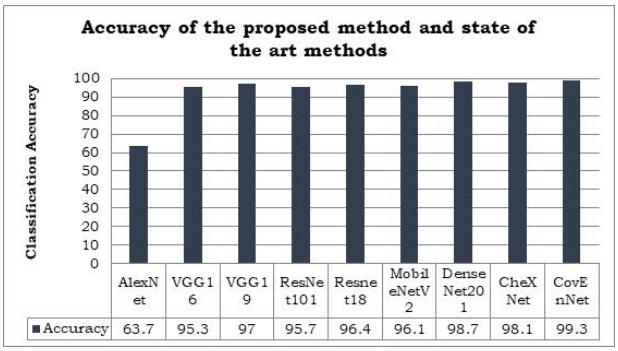


Figure 6: Accuracy and Loss graphs of CovEnNet



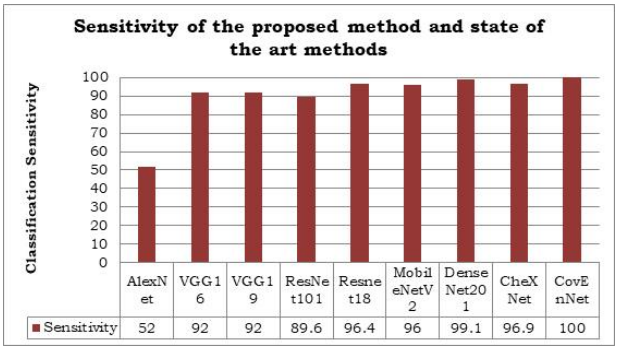

| | AlexNet | VGG16 | VGG19 | ResNet101 | Resnet18 | MobileNetV2 | DenseNet201 | CheXNet | CovEnNet |
|---|---|---|---|---|---|---|---|---|---|
| Specificity | 91.5 | 99.95 | 99.1 | 89.7 | 96.5 | 96.3 | 99 | 97 | 99.5 |

Table 6: Comparing Specificity of the proposed model (CovEnNet) with different deep learning networks



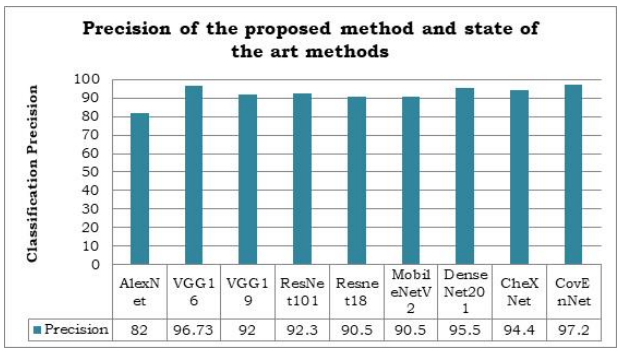

| | AlexNet | VGG16 | VGG19 | ResNet101 | Resnet18 | MobileNetV2 | DenseNet201 | CheXNet | CovEnNet |
|---|---|---|---|---|---|---|---|---|---|
| Accuracy | 63.7 | 95.3 | 97 | 95.7 | 96.4 | 96.1 | 98.7 | 98.1 | 99.3 |

Table 4: Comparing the Accuracy of the proposed model (CovEnNet) with different deep learning networks



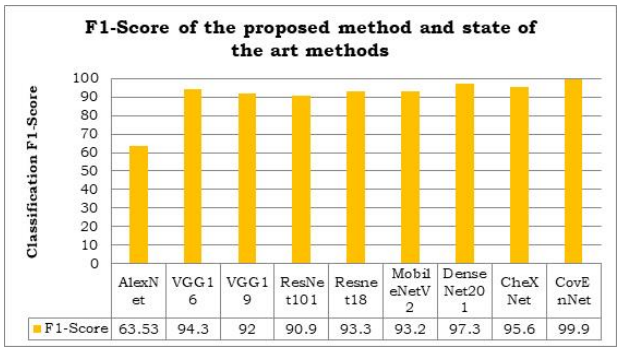

| | AlexNet | VGG16 | VGG19 | ResNet101 | Resnet18 | MobileNetV2 | DenseNet201 | CheXNet | CovEnNet |
|---|---|---|---|---|---|---|---|---|---|
| Sensitivity | 52 | 92 | 92 | 89.6 | 96.4 | 96 | 99.1 | 96.9 | 100 |

Table 5: Comparing Sensitivity of the proposed model (CovEnNet) with different deep learning networks



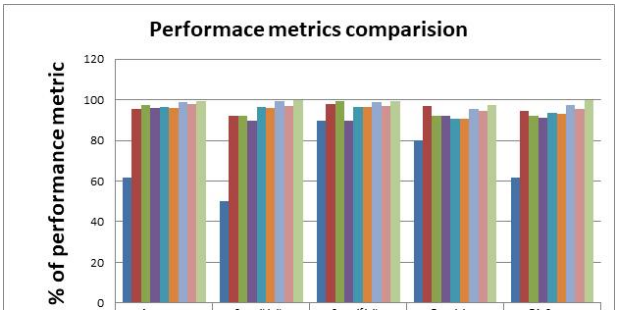

| | AlexNet | VGG16 | VGG19 | ResNet101 | Resnet18 | MobileNetV2 | DenseNet201 | CheXNet | CovEnNet |
|---|---|---|---|---|---|---|---|---|---|
| Precision | 82 | 96.73 | 92 | 92.3 | 90.5 | 90.5 | 95.5 | 94.4 | 97.2 |

Figure 7: Comparing Precision of the proposed model(CovEnNet) with different deep learning networks




| | AlexNet | VGG16 | VGG19 | ResNet101 | Resnet18 | MobileNetV2 | DenseNet201 | CheXNet | CovEnNet |
|---|---|---|---|---|---|---|---|---|---|
| F1-Score | 63.53 | 94.3 | 92 | 90.9 | 93.3 | 93.2 | 97.3 | 95.6 | 99.9 |

Figure 8: Comparing F1-Score of proposed model (CovEnNet) with different deep learning networks




| | Accuracy | Sensitivity | Specificity | Precision | F1-Score |
|---|---|---|---|---|---|
| AlexNet | 61.7 | 50 | 89.5 | 80 | 61.53 |
| VGG16 | 95.3 | 92 | 97.95 | 96.73 | 94.3 |
| VGG19 | 97.5 | 92 | 99.1 | 92 | 92 |
| ResNet101 | 95.7 | 89.6 | 89.7 | 92.3 | 90.9 |
| ResNet18 | 96.4 | 96.4 | 96.5 | 90.5 | 93.3 |
| MobileNetV2 | 96.1 | 96 | 96.3 | 90.5 | 93.2 |
| DenseNet201 | 98.7 | 99.1 | 99 | 95.5 | 97.3 |
| CheXNet | 98.1 | 96.9 | 97 | 94.4 | 95.6 |
| CovEnNet | 99.3 | 100 | 99.5 | 97.2 | 99.9 |

Table 9: Performance metrics comparison of proposed model (CovEnNet) with different Deep learning Networks

## V.Conclusion

Eight different popular and previously reported efficient CNN based deep learning algorithms were trained, validated and tested for classifying normal, Covid-19 and Viral pneumonia patients using chest X-ray images. Our proposed model i.e., CovEnNet outperforms other deep learning networks and it is best suitable for detection of Covid-19, as it is specially designed for detection of Covid-19. This study proposes a CAD approach for classifying Chest X-Ray Images into distinct kinds (Covid-19, Viral Pneumonia, Normal) using a deep neural network. The proposed deep neural network (CovEnNet) is made up of 16 layers, including an image input layer that processes images before feeding them through convolution layers and activation functions. In addition, to avoid overfitting, one dropout layer is used, followed by three fully linked layers. A softmax layer and a classification layer were used to provide the expected class label for probability distribution and classification. The Kaggle repository provided the 2D dataset. The suggested deep learning architecture outperforms state-of-the-art models, achieving the greatest accuracy of 99.3 percent.

## References

[1]. 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/who-director-general-s-opening-remarks-at-themedia-briefing-on-covid-19-11-march-2020

[2]. Li, Lin, et al. "Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT." Radiology (2020).

[3]. X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, and J. Liu, ``Chest CT for typical 2019-nCoV pneumonia: Relationship to negative RT-PCR testing,'' Radiology, vol. 12, Feb. 2020, Art. no. 200343.

[4]. T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia, ``Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases,'' Radiology, vol. 26, Feb. 2020, Art. no. 200642.

[5]. D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, and Z. Peng, ``Clinical characteristics of 138 hospitalized patients with 2019 novel Coronavirus Infected pneumonia in Wuhan, China,'' JAMA, vol. 323, no. 11, p. 1061, Mar. 2020.

[6]. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, ``A survey on deep learning in medical image analysis,'' Med. Image Anal., vol. 42, pp. 6088, Dec. 2017.

[7]. S. Rajaraman, S. Candemir, I. Kim, G. Thoma, and S. Antani, ``Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs,'' Appl. Sci., vol. 8, no. 10, p. 1715, Sep. 2018.

[8]. M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, ``Lung pattern classication for interstitial lung diseases using a deep convolutional neural network,'' IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 12071216, May 2016.

[9]. J. Lei, J. Li, X. Li, and X. Qi, "CT imaging of the 2019 Novel Coronavirus (2019-nCoV) pneumonia," Radiology, vol. 295, no. 1, 2020, Art. no. 200236.

[10]. F. Song et al., "Emerging 2019 Novel Coronavirus (2019-nCoV) pneumonia," Radiology, vol. 295, no. 1, 2020, Art. no. 200274.

[11]. Hoo-Chang Shin, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, Ronald M. Summers, Deep convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 35, NO. 5, MAY 2016, pp. 1285-1298

[12]. Abiwinanda, N., Hanif, M., Hesaputra, S.T., Handayani, A., Mengko, T.R., 2019. Brain Tumor Classification Using Convolutional Neural Network, World Congress on Medical Physics and Biomedical Engineering 2018. Springer, pp. 183-189.

[13]. H. Shin et al., "Interleaved text/image deep mining on a large-scale radiology image database," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 1090–1099.

[14]. Shin, H.-C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R.M., 2016. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Transactions on Medical Imaging 35, 1285-1298.

[15]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. Adv. Neural Inf. Process. Syst. (NIPS), Jan. 2012, pp. 1097-1105.

[16]. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, ``Gradient-based learning applied to document recognition,'' Proc. IEEE, vol. 86, no. 11, pp. 2278_2324, Nov. 1998. https://doi.org/10.1109/5.726791

[17]. N. M. Nasrabadi, "Pattern recognition and machine learning," J. Electron. Imag., vol. 16, no. 4, 2007, Art. no. 049901. https://doi.org/10.1117/1.2819119