IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DIAGNOSIS OF COVID-19 FROM CHEST X-RAY IMAGES USING DEEP LEARNING AUTOENCODER

MR.K. NAGARAJU[1], DR. TRYAMBAK A. HIWARKAR[2]

[1]Reseach scholar, Department of Computer Science and Engineering, Sardar Patel University, Sardar Patel Knowledge City, Waraseoni Road, Dongariya, Balaghat Madya Pradesh India-481001,

[2]Professor, Department of Computer Science and Engineering, Sardar Patel University, Sardar Patel Knowledge City, Waraseoni Road, Dongariya, Balaghat Madya Pradesh India-481001

Abstract: One of the most dangerous viruses of the 20th century is COVID-19, automated diagnosis has become one of the most fashionable research topics to achieve faster mass screening. Deep learning-based approaches have been found to be the most promising methods for finding this type of disease. However, this paper proposes a two-step deep CNN-based method to detect COVID-19 from chest X-ray images to achieve optimal performance with limited training images. In the first step, an encoder-decoder-based autoencoder network trained on unsupervised lung X-ray images is proposed, and the network learns to reconstruct the X-ray images. In the second step, an encoder network is proposed, which consists of different layers of the encoder model and then the encoder. Here, the encoder model is initialized with the weights learned in the first step, and the outputs of the different layers of the encoder model are effectively used by combining them into the proposed component network. An intelligent functional redundancy system is implemented in the proposed redundancy network. Finally, the link network of the encoder is trained to extract the detected features from the X-ray images, and the resulting features are used in the classification layers of the proposed architecture. Considering the final classification task, the EfficientNet-B4 network is used in both stages. Full training is performed on datasets covering the following categories: COVID-19, Normal, Bacterial Pneumonia, Viral Pneumonia. The proposed method gives a very satisfactory performance compared to the state-of-the-art methods and achieves 91.23 accuracy in 4-class, 95.72% in 3-class and 98.87% in 2-class classification.

Index Terms **- COVID-19 diagnosis, Medical Image Analysis, X-ray,Neural Network,Autoencoder,Deep Learning.**

I. INTRODUCTION

Novel coronavirus disease 2019, also known as COVID-19, first emerged in Wuhan, Hubei, China, in December 2019 [1] and has since become a global pandemic affecting millions of lives worldwide. COVID-19 is a new severe acute respiratory syndrome coronavirus that mainly affects the lungs of the human body [2]. Researchers have observed ground opacities, compaction and inferior zone predominance on chest radiographs of patients with COVID-19 [3]. Because of these lung imaging features; it has been shown that chest radiography can be used to detect the virus [4] in patients. Deep learning-based methods have been significantly used in lung X-ray related tasks, such as: nodule classification [5], tuberculosis detection [6], rib suppression [7], pneumonia detection [8], and lung segmentation [9] . Although CT images can also be used to detect COVID-19, X-ray technology is cheaper and more widely available than CT imaging technology [10]. X-ray imaging can also be used for faster mass testing, and this is where machine learning

techniques can really help. In addition, X-ray imaging provides a simple interpretation of various chest problems. As such, this study uses X-rays rather than CT images. The current biggest challenge for detecting COVID-19 from chest x-rays using deep learning is the relatively small size of available labeled data. In the deep learning literature, it has been found that in these cases unsupervised learning can be used to learn the representations first, which allows supervised learning to harmonize and generalize even with small data. [11] showed that using a deep convolutional autoencoder for unsupervised learning of image features enabled lung nodule detection with only a small amount of labeled data. [12] proposed that using a multiscale representation learning method through sparse autoencoder networks to capture the internal scales of medical images leads to better performance in the classification task. In pathology detection [13] used a conditional variational autoencoder to learn the distribution of whole image reconstruction and coding, and the encoder part used those learned features later in the classification task. Autoencoder-based reconstruction techniques are already being used to detect COVID-19 in chest CT images. Researchers have successfully used U-Net-based architectures [14] to segment multiple regions of COVID-19 infection in chest CT images. Some studies [15], [16] demonstrated the use of coding networks in their system to classify COVID-19 infection from CT images. The method proposed in [17] uses contrast domain invariant augmentation techniques in the feature discrimination output to further improve their classification performance and make the system more general to detect COVID-19 in CT images. Given the success of deep learning-based methods in chest X-ray tasks, it is natural to use them to classify COVID-19 based on chest X-ray images. A more of research is done in this domin. COVID-Net [18], a deep convolutional neural network trained to classify COVID-19 on a dataset containing three classes (normal, pneumonia, and COVID) in chest X-ray images, achieved a classification accuracy of 93.3%. Another CNN model developed by DarkCovidNet for this task, developed by [19], was trained in both three classes and two classes (COVID and non-COVID) with accuracies of 87.02% and 98.08%, respectively. Another CNN model based on the Xception [20] architecture, called CoroNet [21], was trained for four classes (normal, COVID, bacterial pneumonia, and viral pneumonia), 3 classes, and 2 classes with an accuracy of 89.6%. anyway. 95% and 99%. [23] proposed a method for lung segmentation based on a chest X-ray image and using random patches of the segmented image to train a pre-trained ResNet-18 [22] to classify COVID-19. Using a small dataset containing 50 normal and 50 patient images of COVID-19, [24] trained InceptionV3, ResNet-50, and Inception-ResnetV2 models and achieved accuracies of 97%, 98%, and 87% for the two classes, respectively. Obtaining highly satisfactory COVID-19 image recognition performance from the relatively small amount of available training data remains a difficult and never-ending challenge.

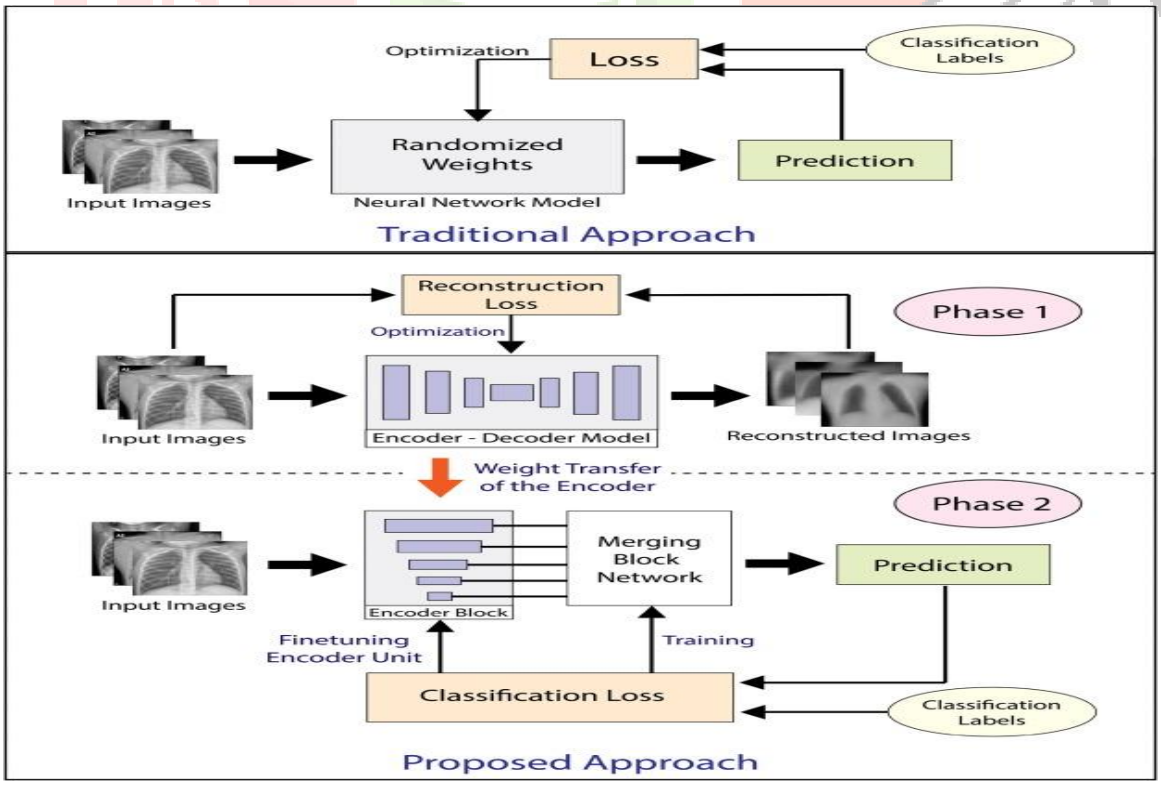

Fig1. The novel approach of the proposed method is presented. In traditional approach images are passed through a randomly initialized neural network model and the model learns to classify the images, in the proposed method there are two phases of training. In the first phase an autoencoder model learns to

reconstruct the input X-ray images. In the second phase the encoder portion of the autoencoder is initialized with the weights learned in phase one and connected to a proposed merging block network and this combined model is trained for the classification task.

II. METHODOLOGY

In the proposed method, both unsupervised and supervised deep neural network architectures are effectively used for the classification of the COVID-19 image. The main blocks involved in the proposed scheme are shown in Figure 2. First, a deep convolutional autoencoder network is designed to perform unsupervised feature extraction from a given chest X-ray image. Next, a supervised deep CNN architecture is designed using the extracted first-stage features, and then these features were used for supervised learning in the classification network. The classification grid consists of individually designed smaller blocks arranged in a tree-style architecture. Both networks are trained on chest X-ray images. One of the main challenges of this work is to deal with the classification task when there is a limited amount of training data, especially in the case of COVID-19. Therefore, to obtain a better trained model, an effective feature extraction step is added before the network training step. We propose to use an unsupervised extraction stage based on an autoencoder-decoder to extract the spatial features of the input image. The motivation for introducing such an additional encoder-decoder before the conventional classification step is its ability to preserve the detailed information of a given image at its different levels. Since in the autoencoder-decoder structure, the given image must be reconstructed using a general optimization model in the output stage, it is assumed that the spatial characteristics of the input image are accurately captured in the encoder stage. Thus, when features are extracted at different levels of the encoder, the extracted features can accurately represent a given class, allowing better interclass discrimination. The proposed system also develops an efficient combination method with unique combination blocks (M-blocks) to effectively utilize the features extracted from different levels of the autoencoder. Using these combined features in a classification network helps achieve better training even on labeled chest X-rays with a small dataset.

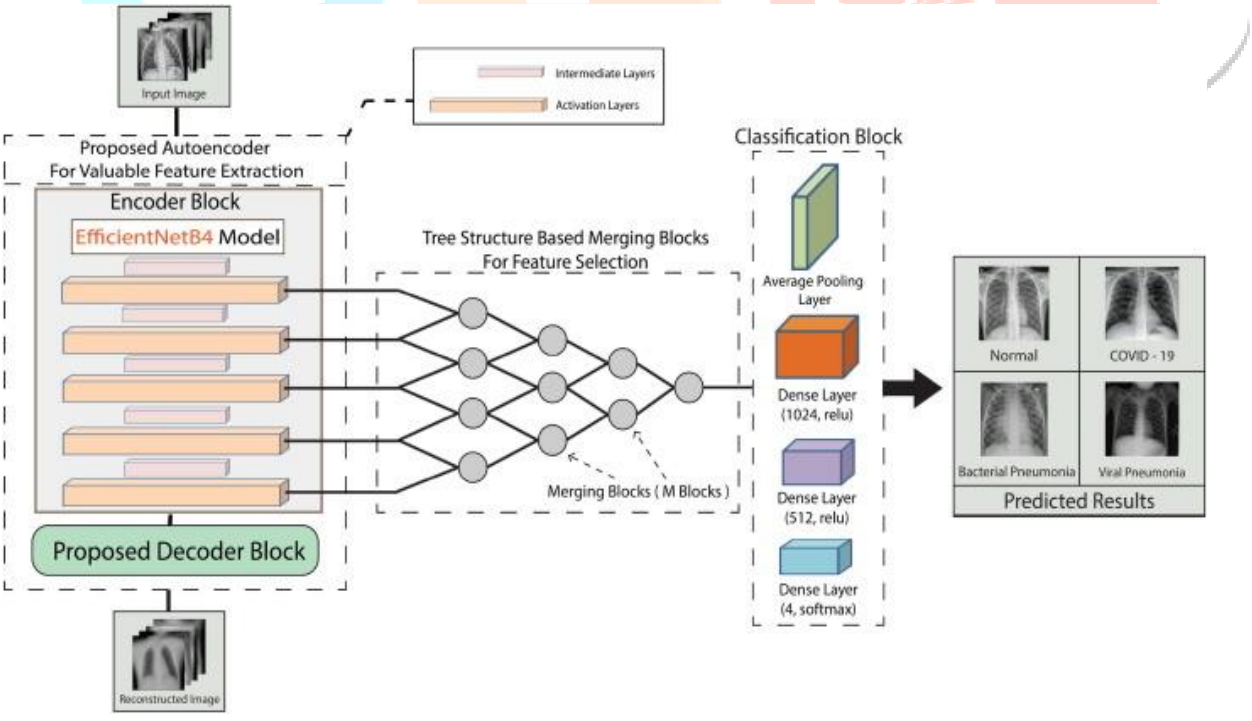

Fig2. The proposed autoencoder framework consists of two-stage training: (1) training the autoencoder network using the input X-ray images and (2) training the intermediate layers of the encoder network (utilizing weights obtained in the first stage) followed by the merging network. Output from the merging network is passed to a classification network that makes the final classification.

The main steps of this methodology are shown in Figure 2. Figure 2 shows the main blocks used in the proposed method, where the first block corresponds to the proposed unsupervised feature extraction step. In this step, the backbone architecture of the EfficientNet-B4 model is used to design an encoder-decoder model that optimizes each given input image and produces a decoded image. In this process, the encoder extracts different information from different perspectives, which is then encoded into the encoder. These different encoder levels, which contain different information, are treated as useful features for use in the next step. In

Figure 2, the next step represents a characteristic fusion block, where the features taken from different levels of the encoder are effectively merged using the proposed fusion blocks. As a result, the features collected from different levels of the encoder are combined into a single feature vector, which is then finally used in the classification layer, as shown in Figure 1. 2. These different training phases are described in the following sections.

2.1 Pre-Processing

Before X-ray images are used in deep nerve models, the images undergo a two-step preprocessing: resizing and normalization. The input images are resized to 256 x 256 square images containing three channels. Min-max normalization is then applied to the transformed input images. These speeds up the training and helps the model converge more easily. Proposed Unsupervised Feature Learning Architecture The first stage of the system is an automatic encoder that is trained on unlabeled chest X-ray images and learns to reconstruct the input images. Autoencoding algorithms are able to use an unsupervised learning method to automatically learn features from unlabeled data [25] and are particularly useful in the field of medical image analysis with sparsely labeled data [11]. An autoencoder consists of two parts: an encoder and a decoder. An encoder learns how to efficiently compress and encode a representation of a dataset. And the decoder part learns to take that encoded data and reconstruct it as a representation that is as close as possible to the original input data. When choosing this model, it is important to note that in the first stage it is used to extract features based on encoder-decoder and in the last stage the same architecture works as the basic classification network. One of the goals is to choose a classification architecture such that two separate architectures are not required for these two different phases, which unnecessarily increases the computer load. Therefore, in this case, the goal was to choose a single classification network that could fulfil both goals. In deep convolutional autoencoders, the encoder part is made by stacking convolutional layers followed by stacking the layers. As a result, the resolution of the input image gradually decreases and the number of channels increases. This feature is similar to conventional CNN architectures used in classification tasks. Because of this similarity, a traditional classifier architecture can be used to implement the encoder block. Among the different types of deep convolutional neural networks, the effective network proposed by [26] carefully balances the depth, width, and resolution of the network to achieve a better classification result. The Efficient Net architecture offers a combined scaling system that uniformly scales all dimensions of depth/width/resolution. Such fusion scaling has the advantage of focusing on more relevant regions with more object details and can significantly improve the classification performance compared to the result achieved by one-dimensional scaling methods [26]. For this purpose, various available types of deep convolutional neural network architecture are tested, and the EfficientNet-B4 [26] model proved to be the best in terms of accuracy. In the Results and Simulation Section, this is a study of how the performance would have changed if another state-of-the-art architecture such as InceptionV3, Resnet50, VGG11 etc. would be used in the proposed system. In the first step, the Efficient Net-B4 model is used as an encoder-decoder to obtain optimal weights using different training images. Once the weights of this cipher-decoder block are optimized, these weights are used as output weights in a later stage where the classification task is performed. And the same encoder-decoder network is trained under supervision at that time. Therefore, using a more efficient accuracy model, such as the EfficientNet-B4 model, reduces the computational complexity because it is used both as an encoder block and later as a classification block. The EfficientNet-B4 network balances these tasks without compromising performance accuracy. The EfficientNet-B4 model was initialized using ImageNet pre-trained weights [27] because the dataset used here is relatively small to use without ImageNet weights. The fully connected layers at the bottom of the network were omitted and the output was taken from the last convolution block to be used as coded data for the autoencoder. For an input image size (256, 256, 3), the encoder network produces coded shape data (8, 8, 1792). The next part of the autoencoder is the decoder. The decoder module was designed to reconstruct the original input image size (256, 256, 3) based on the encoded data of the (8, 8, 1792) format. This is anti-model behaviour of the encoder. The traditional CNN architecture does not perform such an operation, and as a result, the decoder is designed according to the proposed scheme to reconstruct the input image from the encoded data produced by the EfficientNet-B4 encoder model. More analytical details about the decoder can be found in [28]. The decoder model consisted of five blocks, where each block started with a transposed convolution layer that removed the image 2 times. This was followed by a convolution layer with the same number of filters as the transposed convolution layer. The detailed architecture is shown in the figure. 3 and all layers used in the decoder model are shown in Table 1 with their respective output formats. At the end of the decoded model is a convolutional layer with the same number of channels as the input image. This layer had the ReLU activation function. ReLU is a scaled exponential linear unit activation function. It is defined as.

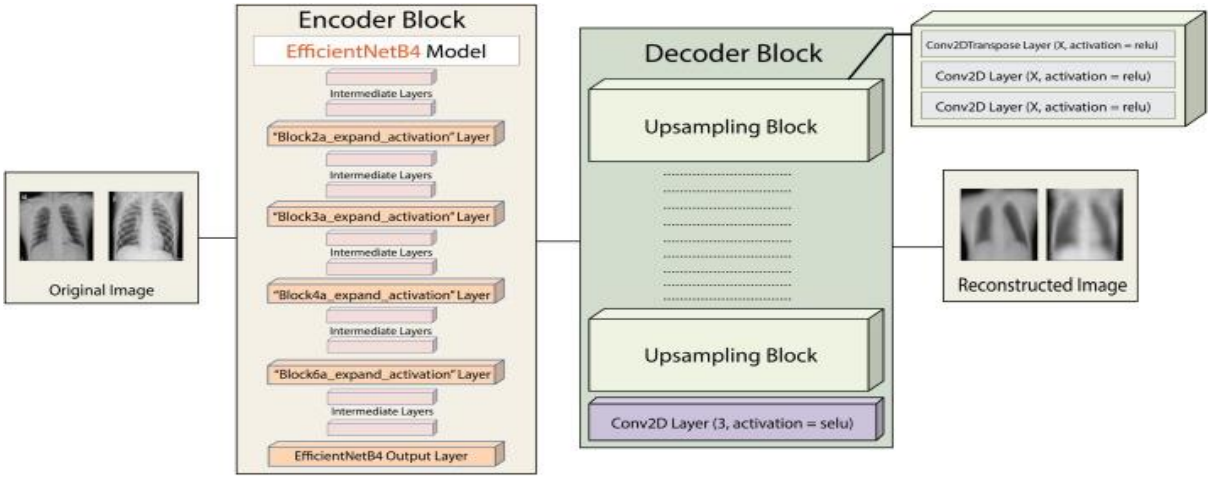

Fig3. Model architecture of the proposed deep convolutional autoencoder.

Table 1 – The layers and their corresponding output shape for the proposed autoencoder model.

Encoder Feature Extraction Layers		Decoder Layers			
Layer (type)	Output Shape	Layer (type)	Output Shape	Layer (type)	Output Shape
Block2a expand activation" Laye	128,128,144)	1.Conv2D Transpose Layer1.	(16,16,512)	8.Conv2D Layer	(64,64,256)
"Block3a expand activation" Layer	(64,64,192	2.Conv2D Layer	(16,16,512)	9.Conv2D Layer	(64,64,256)
"Block4a expand activation" Layer	(32,32,336	3.Conv2D Layer	(16,16,512)	10. Conv2D Transpose layer	(128,128,128)
"Block6a expand activation" Laye	(16,16,960)	4.Conv2D Transpose Layer	(32,32,256)	11.Conv2D Layer	(128,128,128)
EfficientB4 Output Layer	(8,8,1792)	5.Conv2D Layer 6.Conv2D Layer 7.Conv2D Transpose Layer	(32,32,256), (64,64,256)	12.Conv2D Transpose Layer 13.Conv2D Layer 14.Conv2D Laye	(256,256,64) (256,256,64) (256,256,3)

Although the X-ray image reconstructed from that network is not used directly, it is an important by-product of the proposed architecture. Without this reconstructed X-ray image, it is not possible to train an encoder with a small dataset to learn relevant features. The quality of the reconstructed X-ray image also shows how well the autoencoder network converges. If the autoencoding network is properly trained, it helps the encoder to retain the detailed information of the images in different layers, which can later be used in the classification task. Since the autoencoder model learns to reconstruct the input image, it does not need a label, and the entire pixel space of the input image serves as labels. So even with a small amount of data, and also with other unlabeled chest X-ray data, this network can be trained and approximated. In the process of creating encoded data useful for reconstruction, the encoder model succeeds in preserving the data from different perspectives by learning the unique features of the images in the dataset, and these features can then be used for classification purposes.

2.2 Proposed Classification Architecture

the next part of the study was to develop a convolutional neural network architecture for a supervised learning system to identify patients with COVID-19 based on chest X-ray images. At this stage, a classification network is required, where the problems to deal with are of 2, 3 or 4 categories. For this task, as mentioned above, the outputs of the different levels of the encoder were separated using the weights learned by our autoencoder in the previous step. The features of this coding network were then taken and passed through the classification network. The different parts of the classification network are described below.

2.3 Functional Discharge Phase

The encoder was the EfficientNet-B4 model trained in the previous step. As this model learned to reconstruct chest X-ray images, the middle layers of this model had valuable features that could be used in the classification task. In general, CNN models use concatenation techniques followed by a convolution operation to reduce the size of the input image. Therefore, the outputs of the intermediate layers of each block were extracted from the encoder model after merging. The characteristics can be extracted from the number of different layers of the encoder model. However, if too many layers are tried to aggregate or reduce information from different layers into a single channel, it may require more steps, and if too few layers are tried, it may not retain information. Therefore, five (5) number of layers is chosen in this work considering that it is the most suitable number.

2.4 Combining Blocks (M-Blocks)

For our classification network, data is taken from different layers of the encoder and reduced to a single channel so that the classification task can be performed. In this case, one important task is to reduce these five layers of data into one layer, and for this, a unique block called M-blocks is developed, which intelligently combines the characteristics of the two layers of the encoder and then these functions together, also from other M The block takes two 3D tensor inputs, the first of which is twice the height and width of the second input. The first input is then passed through a pooling layer that uses a (2,2) filter and averages each window value. The output tensor of this stage has the same height and width as the second input, and the two tensors are then joined on the channel axis. This coupled tensor is then convolved with a window size of (1,1) and a number of filters equal to the number of the second input channel. As a result, the output format of each M-block is the same as the format of the other input, but contains the functions of both inputs. The structure of this block is shown in detail in Figure 4. Thus, this block combines the features of the other two layers and then learns new functions on top of them with convolutional layers. Figure 5.

2.5 Tree Structured Feature Merging Network

The classification network is made by a combination of an encoder model and M-blocks. The M-blocks of the classification network are connected in a tree network-style architecture as shown in Figure 5. For N number of feature extraction layers, the network has (n-1) stages, each with one less M-block than the previous one. stage the last stage of this network has one M block. Five feature extraction layers were used in this study and this network has four M-blocks, the first stage takes inputs from the encoder intermediate layers and has five blocks. The second stage of M-blocks takes the input characteristics of the first stage and has four blocks. So, it continues until the last step, where there is one M-block whose output form is (8, 8, 1792). This tensor is then subjected to global average summation and produces a feature vector of size 1792. These features are then fed through two fully connected dense neural networks, one with 1024 neurons and the other with 512 neurons. Softmax activation is then performed and classification based on that prediction is performed.

III. Result and Discussion

This section presents the performance of the proposed method considering different classification cases and different performance measurement criteria. The results obtained by the proposed method are compared with the results obtained by some state-of-the-art methods. Then, first the dataset and then the results with detailed analysis and comments are presented. The proposed model is trained and tested with 5-fold cross-validation data containing 3 classes. Accuracy, sensitivity, F1 score, and precision are calculated as performance measures for each test set and can be seen in Table 2.

Table 2. Precision, Recall, F1-score and Accuracy across all 3 classes for the 5 folds of data.

Folds	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
Fold 1	98.05	97.97	97.96	97.97
Fold 2	95.93	95.93	95.93	95.93
Fold 3	95.57	95.52	95.53	95.93
Fold 4	97.13	97.12	97.11	97.12
Fold 5	95.58	95.47	95.47	95.47
Average	96.45	96.41	96.39	96.41

From Table 2 it can be seen that the model got the highest accuracy of 97.97% from fold 1 and the average accuracy for all the 5 folds is 96.41%. The model had accuracy in the range of 95.47% to 97.97% for all of the folds of data. Even the lowest accuracy of 95.47% is still quite high. The same performance metrics are also generated in a class-wise basis for all of the folds. The class-wise result for fold-1 can be seen in the

Table 3.

Table 3. Precision, Sensitivity, F1-score and Accuracy of the 3 classes for Fold 1.

Class	Precision (%)	Sensitivity (%)	F1-score (%)	Accuracy (%)
COVID19	98.79	100	99.39	100
Normal	100	93.90	96.86	93.90
Pneumonia	95.35	100	97.62	100

As evident from Table 3, the model performed exceptionally well in the COVID19 and Pneumonia class getting an accuracy of 100% for both of these classes. While for the Normal class it gets an accuracy of 93.90%. These claims are further supported by the confusion matrix generated for each of the folds. The confusion matrix for fold-1 and fold-2 are presented in Fig. 7. From Fig. 7 it can be observed that the model accurately predicted all the COVID-19 class images. But some of the Normal class images were classified as the Pneumonia class. While for fold-2, some of the Pneumonia class images were classified as the Normal class.

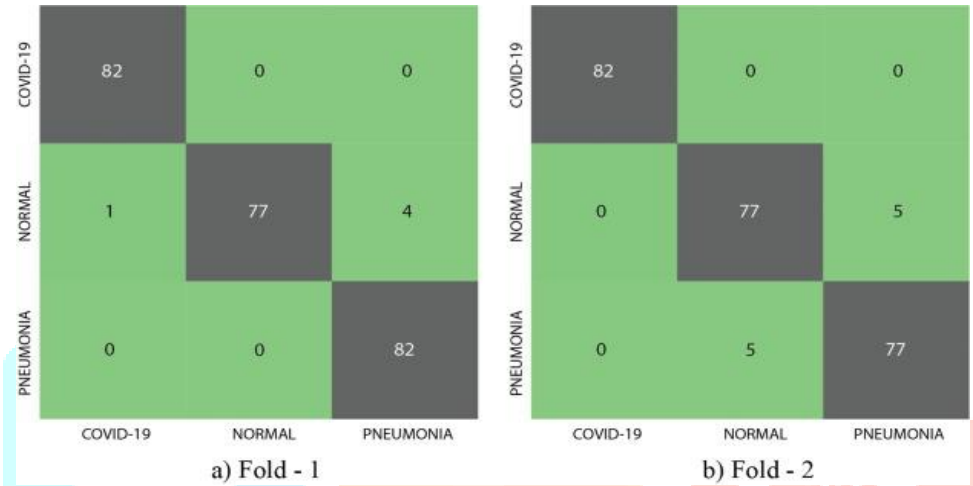

Fig7.Confusion Matrix of the Test Set for 3-class Dataset.

The model is also trained on a 4-class dataset to separately classify bacterial pneumonia and viral pneum-Monia. The same performance metrics from the 3-class setup is used in this case as well. The result for cross validation testing is presented in Table 4 and the class-wise result is presented in Table 5.

Table 4. Precision, Recall, F1-score and Accuracy across all 4 classes for the 5 folds of data

Folds	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
Fold 1	91.46	91.46	91.46	91.46
Fold 2	92.07	92.07	92.07	92.07
Fold 3	89.33	89.33	89.33	89.33
Fold 4	90.12	90.12	90.12	90.12
Fold 5	87.65	87.65	87.65	87.65
Average	90.13	90.13	90.13	90.13

Table 5. Class wise result for 4-class dataset of the best performing Fold.

Class	Precision (%)	Sensitivity (%)	F1-Score (%)	Accuracy (%)
Bacterial Pneumonia	89.74	85.37	87.5	87.5
COVID19	100	100	100	100
Normal	96.25	93.9	95.06	95.06
Viral Pneumonia	84.09	90.24	87.06	87.06

from these tables, it can be seen that the model gave consistent performance in all of the folds and from the class-wise results it can be seen that the model exceptionally well for the COVID-19 class and reasonably well for the Normal class. The performance dropped a bit when differentiating between bacterial pneumonia and viral pneumonia class. On average, for this 4 class dataset, the model achieved a classification accuracy

of 90.13% for the fivefold cross validated data. This experimentation was done to see if the model can generalize for all kinds of low data irrespective of the data source and even if the data are very similar.
Even under these conditions, the model acquired an average accuracy of 90.13% which is a relatively good performance. The model is trained on a 2-class dataset as well. This dataset was derived from the 3-class dataset where the Normal and Pneumonia classes were labelled as non-Covid19. The evaluation metric is the same for this task as well. This detailed result is presented in Table 6.

Table 6. Precision, Sensitivity, F1-score and Accuracy of the 2 classes for Fold 1.

Class	Precision (%)	Sensitivity (%)	F1-score (%)	Accuracy (%)
COVID19	97.62	100	98.8	100
Non-Covid19	100	98.78	99.39	98.78
Average	99.19	99.39	99.19	99.39

From Table 6 it can be seen that the proposed method performed well on both the classes with an average accuracy of 99.39%. These performances on both the 4 class and 2 class datasets can be further inspected with the confusion matrices presented in Fig. 8. As can be observed from the confusion matrix of Fig. 8, that in the case of the two-class dataset almost all the test images were classified correctly except for two Non-COVID images.

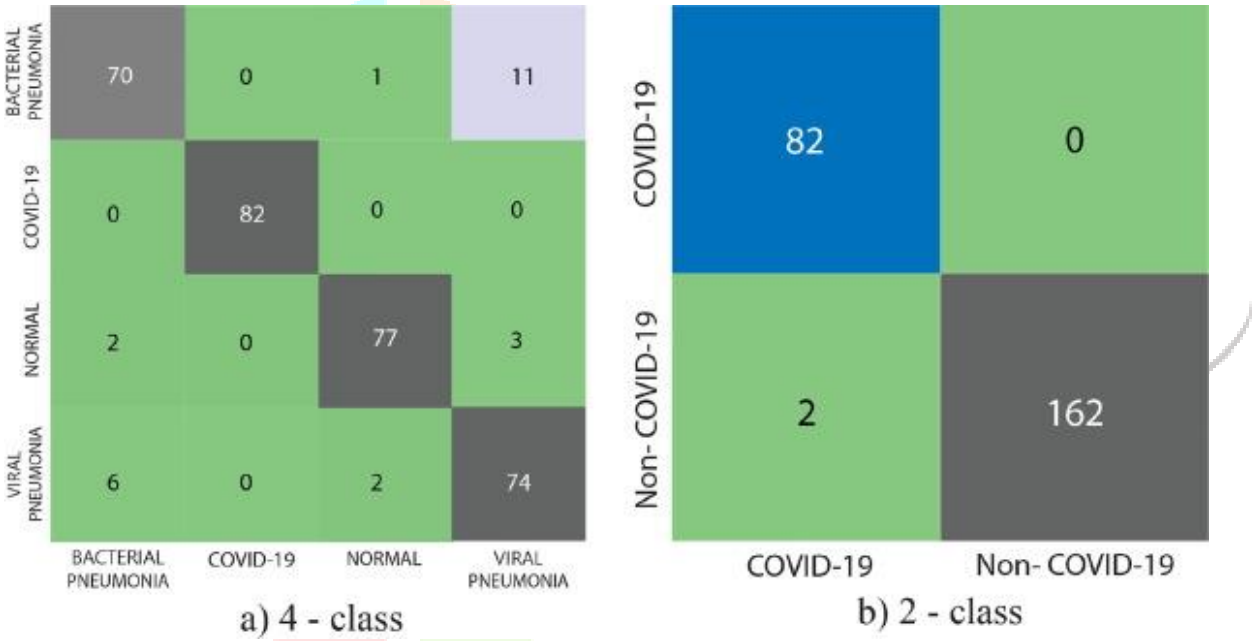

Fig. 8. Confusion Matrices of the Test Set for 4-class and 2-class Dataset.
As mentioned in the methodology section, the EfficientNet-B4 model was used as the encoder network in this study. But other classification networks, such as Resnet-50, InceptionV3, and the other variants of Efficient Net were also tried as the encoder network and their results on the 4 class and 3 class datasets are compared in Table 7.
Table 7. Comparisons of different models for 3-class and 4-class classification using our scheme.

Classification Type	Model Accuracy (%)						
	Efficient Net B1	Efficient Net B2	Efficient Net B3	Efficient Net B4	Inception V3	Resnet 50	Vgg-11
3-class	96.75	97.56	97.56	97.97	97.15	96.75	95.53
4-class	90.85	91.31	92.07	92.38	88.11	89.33	86.89

From Table 7 it can be observed that even though EfficientNet-B4 performed the best, the other models also provide similar performance which is further proof to the credibility and robustness of the proposed scheme. However, in the results section, in order to report the results in all tables, EfficientNet-B4 is used in the proposed method as the encoder network. To further justify the use of EfficinetNet-B4 model as the feature extractor, the Cohen's Kappa score and the Mattheus Correlation Coefficient for the models were evaluated in the 4-class classification scheme using the proposed method. The detailed result of this analysis is presented in Table 8.

Table 8. Cohen's Kappa score and Mattheus Correlation Coefficient of different models for the 4-class classification using the proposed method.

Model	Cohen's Kappa Score	Mattheus Correlation Coefficient
EfficientNet-B4	0.8861	0.8867
ResNet-50	0.7723	0.7727
InceptionNet-V3	0.8292	0.8321
Vgg-11	0.8252	0.8253

To evaluate the effectiveness of the proposed method its results were compared with a simple Efficient Net-B4 classification network pretrained on ImageNet weights. This comparison is presented in Table 9. From the results it can be observed that the use of the autoencoder network coupled with the merging block resulted in a performance improvement and this improvement can be specially seen in case of the four-class dataset where the classification task becomes much more difficult. To further evaluate the performance of the proposed methodology statistical significance test was performed on the two methods mentioned in Table 9. McNemar's test [32] and Wilcoxon signed ranked test [33] are the two statistical tests that were performed for this purpose. The statistical significance tests are performed on the prediction of the two methods mentioned in Table 9. The prediction of each model on the 326 test set images are compared to the ground label of each of these images and a binary label with correct/incorrect decision is generated based on this comparison. There are two distributions of this binary variable for the two models and the disagreement between the two methods is used as the variable for these statistical significance tests. The test tries to see if it is possible to reject the null hypothesis which states that there is no difference in the disagreement between the two methods. The results of these tests are presented in Table 10. It can be observed from the results that the P-value of McNemar's test for the 4-class classification scheme was 0.83825 and for the 3-class classification scheme it was 0.68309. The P-value for the Wilcoxon signed ranked test for the 4-class classification scheme was 0.638 and for the 3-class classification scheme, it was 0.084. As for both the test in both classification schemes the P-value was very close to 0.5 it can be inferred that the proposed methodology produced some degree of statistically significant results.

IV. Conclusion

More than six months have passed since the beginning of the Covid-19 pandemic, and now an automated system is needed to detect COVID-19 based on chest X-rays. This research was conducted with the goal of developing a deep learning-based system that can generalize even with small data. It is shown that the proposed training program uses an unsupervised image reconstruction step in the weight initialization step of the encoder model, and the proposed encoder fusion network extracts features from different layers of the encoder network and learns to combine them effectively. supervised training method that can provide very satisfactory consistent results even on a very small data set. It can effectively handle both binary and multiclass problems. Therefore, it is expected that when a large dataset for this task becomes publicly available, this model will be able to generalize even better. In addition, the network was designed so that both the feature extraction and classification phases used the same EfficientNet-B4 core network. This resulted in more efficient computation and faster convergence.

V. References

[1] Z. Wu, J.M. McGoogan Characteristics of and important lessons from the coronavirus disease (covid-19) outbreak in china: summary of a report of 72 314 cases from the chinese center for disease control and prevention JAMA, 323 (2020) (2019), pp. 1239-1242

[2] P. Verdecchia, C. Cavallini, A. Spanevello, F. Angeli The pivotal link between ace2 deficiency and sars-cov-2 infection Eur J Internal Med (2020)

[3] M.-Y. Ng, E.Y. Lee, J. Yang, F. Yang, X. Li, H. Wang, et al.Imaging profile of the covid-19 infection: radiologic findings and literature review Radiol: Cardiothoracic Imaging, 2 (2020), Article e200034

[4] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, et al.Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in china: a report of 1014 cases

[5] C. Wang, A. Elazab, J. Wu, Q. Hu Lung nodule classification using deep feature fusion in chest radiography Comput Med Imaging Graph, 57 (2017), pp. 10-18

[6] H. Kim, S. Hwang, Scale-invariant feature learning using deconvolutional neural networks for weakly-supervised semantic segmentation, arXiv preprint arXiv:1602.04984 (2016).

[7] E. Soleymanpour, H.R. Pourreza, et al.Fully automatic lung segmentation and rib suppression methods to improve nodule detection in chest radiographsJ Med Signals Sensors, 1 (2011), p. 191

[8] M. Chen, X. Shi, Y. Zhang, D. Wu, M. Guizani Deep features learning for medical image analysis with convolutional autoencoder neural network IEEE Trans Big Data (2017)
[9] H. Uzunova, S. Schultz, H. Handels, J. Ehrhardt Unsupervised pathology detection in medical images using conditional variational autoencoders Int J Comput Assist Radiol Surg, 14 (3) (2019), pp. 451-461
[10] Chen X, Yao L, Zhang Y, Residual attention u-net for automated multi-class segmentation of covid-19 chest ct images, arXiv preprint arXiv:2004.05645 (2020).
[11] A. Amyar, R. Modzelewski, H. Li, S. Ruan Multi-task deep learning based ct imaging analysis for covid-19 pneumonia: classification and segmentation Comput Biol Med, 126 (2020), p. 104037
[12] Z. Wang, Q. Liu, Q. Dou,Contrastive cross-site learning with redesigned net for covid-19 ct classification IEEE J Biomed Health Inf, 24 (2020), pp. 2806-2813
[13] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U. Rajendra Acharya Automated detection of covid-19 cases using deep neural networks with X-ray images Comput Biol Med, 121 (2020), p. 103792
[14] A.I. Khan, J.L. Shah, M.M. Bhat Coronet: A deep neural network for detection and diagnosis of covid-19 from chest X-ray images Comput Methods Programs Biomed, 196 (2020), p. 105581
[15] K. He, X. Zhang, S. Ren, J. Sun Deep residual learning for image recognition Proceedings of the IEEE conference on computer vision and pattern recognition (2016), pp. 770-778
[16] Narin A, Kaya C, Pamuk Z, Automatic detection of coronavirus disease (covid-19) using X-ray images and deep convolutional neural networks, arXiv preprint arXiv:2003.10849 (2020).
[17] Tan M, Le Q, Efficientnet: Rethinking model scaling for convolutional neural networks, in: International Conference on Machine Learning, PMLR, 2019, pp. 6105–6114.
[18] A. Lucas, M. Iliadis, R. Molina, A.K. Katsaggelos Using deep neural networks for inverse problems in imaging:beyond analytical methods IEEE Signal Process Mag,35 (2018),pp.20-36
[19] Pedamonti D, Comparison of non-linear activation functions for deep neural networks on mnist classification task, arXiv preprint arXiv:1804.02763 (2018).
[20] D.S. Kermany, M. Goldbaum, W. Cai, C.C.S. Valentim, H. Liang, S.L. Baxter, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning Cell, 172 (2018), pp. 1122-1131.e9
[21] Cohen JP, Morrison P, Dao L, Covid-19 image data collection, arXiv preprint arXiv:2003.11597 (2020).
[22] T.G. Dietterich Approximate statistical tests for comparing supervised classification learning algorithmsNeural Comput, 10 (1998), pp. 1895-1923
[23] Lacoste A, Laviolette F, Marchand M, Bayesian comparison of machine learning algorithms on single and multiple datasets, in: Artificial Intelligence and Statistics, PMLR, 2012, pp. 665–675.
[24] I.D. Apostolopoulos, T.A. Mpesiana Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks Phys Eng Sci Med (2020), p. 1
[25] P.K. Sethy, S.K. Behera Detection of coronavirus disease (covid-19) based on deep features Preprints, 2020030300 (2020), p. 2020
[26] Hemdan EE-D, Shouman MA, Karar ME, Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in X-ray images, arXiv preprint arXiv:2003.11055 (2020).
[27] T. Mahmud, M.A. Rahman, S.A. Fattah Covxnet: A multi-dilation convolutional neural network for automatic covid-19 and other pneumonia detection from chest X-ray images with transferable multi- receptive feature optimization Comput Biol Med, 122 (2020), p. 103869
[28] A. Abbas, M.M. Abdelsamea, M.M. Gaber Classification of covid-19 in chest X-ray images using detract deep convolutional neural network Appl Intel, 51 (2021), pp. 854-864