AUTOMATIC DIAGNOSIS OF VARIOUS DISEASE FROM A CHEST X-RAY USING A POWER OF DEEP LEARNING

Arjun Choudhary, Dr. Kalpna Sharma and Dr. Prakash Choudhary

ABSTRACT

Artificial Intelligence has been witnessing a monumental growth in bridging the gap between the capabilities of humans and machines.The diagnosis of pathology in a chest X-ray is generally complicated, even for experienced practitioners. A system that can automatically diagnose the findings in images achieved through X-rays of chest can be useful in the medical examination of the patient as there is a shortage of experienced doctors. Classifying the chest X-ray is a multi-label classification task as a patient may have multiple diseases. In this research, we aim to develop an algorithm using deep learning techniques to identify the condition in the chest X-ray with high accuracy. In this research, we fine-tune a pretrained CNN architecture named DenseNet-121 to extract the features from the chest X-ray and to classify the extracted features into the pathology. The weights of the model are initially set with the weights of a model which is trained on ImageNet then the model is trained on the sample of the" ChestXray14" dataset. They are comprised of node layers, containing an input layer, one or more hidden layers, and an output layer. Each node connects to another and has an associated weight and threshold. If the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data will be passed along to the next layer of the network.

***Index Terms** — Chest X-ray; CNN; CAD; DenseNet; ChestXray14*

Biographical notes:

***Arjun Choudhary, a** Research Scholar, Computer Science Department, Bhagwant University, Ajmer, Rajasthan-305001, India.*

***Dr. Kalpna Sharma**, HOD & Assistant Professor, Computer Science and Engineering Department, Bhagwant University, Ajmer, Rajasthan-305001, India.*

***Dr. Prakash Choudhary,** HOD & Assistant Professor, Computer Science and Engineering Department, National Institute of Technology, Hamirpur, Himachal Pradesh,Hamirpur, HP-177005, India*

1. INTRODUCTION

During the past decades, Deep Learning has proved to be a very powerful tool because of its ability to handle huge amounts of data. The interest to use hidden layers has surpassed traditional techniques, especially in pattern recognition. This is the most popular deep neural networks is Convolutional Neural Networks. The Chest X-ray is the most common and currently the best available method for medical imaging which is an increased demand for diagnosis and treatment. To detect lungs disease chest X-rays plays a vital role. There are various types of chest diseases like Asthma, Chronic Obstructive Pulmonary Disease (COPD), Chronic Bronchitis, Pneumonia tuberculosis and many other diseases which can examine with the help of chest X-rays by experienced radiologists. CNN is a type of deep neural network, and there are many other types. CNNs are more popular because they are very useful in applications to image recognition. More than 65 million people in the world suffer from Chronic Obstructive Pulmonary Disease(COPD). The most common deadly disease is Tuberculosis or TB. More than 10 million people suffer from TB and 1.6 million people die in every year. In the world, 1.6 million people die from lung cancer each year, which is the most common neoplasm. Most of the children are affected by All over the world, nearly 4 million deaths are caused due to chronic respiratory disease. Exposure to toxic smoke is also dangerous and has a hazardous effect on human health. About 2 billion people come in contact with the indoor toxic smoke and around a billion to tobacco smoke. In addition to this inhalation of outdoor pollutants has also harmful effect on the respiratory system of the human body. Because of the industrialization, more people are exposed to the polluted air because of which more people are likely to develop respiratory diseases. CAD (Computer-Aided Design) can help in diagnosing the patients quickly. In recent years computer vision has come a long way, because of the enhancements in deep learning algorithms and improvements in the computational power. CNN has come up as one of the best way for the image classification problems. CNN has proved very good in learning local and global structures in an image. In this study, we will be using CNN to examine the chest X-ray. The task of diagnosing the disease in a chest X-ray can be broadly divided into three parts: pre-processing, feature extraction and classification. The block diagram is shown below in fig
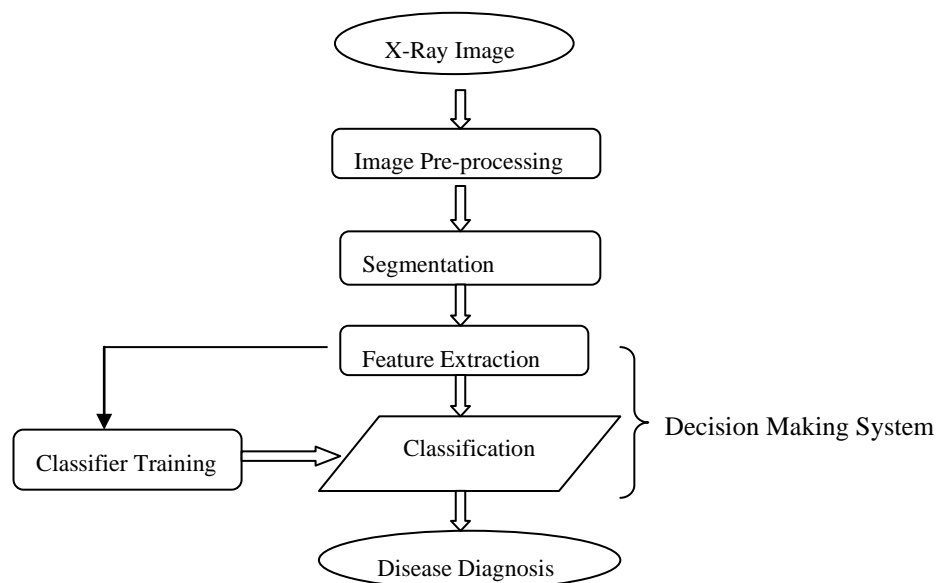

Fig.1. The block diagram of CAD

Pre-processing phase includes resizing of the image, improving contrast, removal of noise flipping of the images and some other techniques. For feature extraction, we will fine tune a pretrained CNN architecture. A CNN uses convolution layers, max pooling layers, activation functions and dropout for extraction the features from an image.

For the classification of the extracted features, we will use an artificial neural network. ANN takes the flattened output of the CNN as an input and passes it through several hidden layers and into the output layer. The number of healthy people is much higher than the diseased people, so the number of samples of healthy X-rays is higher than the diseased Xrays. Furthermore, there is a difference in the number of patients within the diseases, as some conditions are rare compared to others. The development of the system for CAD of chest X-rays poses a problem of an imbalanced dataset. Training a model of imbalanced data has many disadvantages. Suppose if a model learns from a dataset which has 95% of the samples of class 1 and 5% samples of class 2. The model will overfit on class 2 and the model will always predict class 2 as output no matter what input is given. To address this issue we have studied various ways of handling the class imbalance and we have chosen costsensitive learning for this research. Basic concept of the cost-sensitive learning is to penalize the CNN for wrongly classifying the samples of the minority class.

3 LITERATURE REVIEW

Many types of research have been performed around medical imaging and CAD in recent years. Advancement in deep learning algorithms and architectures have accomplished great feats in the tasks related to computer vision. Especially the CNNs have outperformed all of the other techniques in computer vision and enabled computer to process, evaluate images and extract details similar to the human mind. Researches in the field of CAD have been used CNN to their advantage and achieved good results. The existing work on automatic disease detection in chest X-rays have also shown that transfer learning is very handy technique while training a model on a small dataset. Stefan Jaegar et al. proposed a method for automatically diagnosing tuberculosis (commonly known as TB) in chest radiographs [Jaegar et al.(2013)]. In this method, they first identified the lungs using cut segmentation method in which they first used a lung model representing the shape of a normal lung and after that the graph approach was applied in which lung area has to satisfy certain conditions to be considered as lung region. Then form the identified region, features (shape and texture) are extracted. Two types of features sets, which are CBIR-based image features and object detection inspired features, were obtained separately. With the use of extracted features, the chest X-ray is classified as positive or negative for TB. The researchers have used SVM for the classification of the features extracted from the images.

Abhishek Hazra et al. experimented to design a method to automatically diagnose the disease in a chest X-ray using CNN [Choudhary and Hazra (2019)]. In this study, the researcher compared the performance of a pre-trained on ImageNet and after that, it is fine tuned on the chest X-ray dataset. The custom model proposed by the researcher was trained on the dataset of 5606 images and it achieved an accuracy of 83.67%. The VGG16 model achieved better accuracy than custom model on the sample of the chest X-ray 14 dataset. Softmax function was used at the end of the fully connected layer which converts the logits into the probability distribution among the classes. In the

trained model if an image is fed as input, the model shows the top 3 diseases by probability. The researcher used Grad CAM for the visualization and analysis of the prediction results. Yaniv Bar et al. proposed an automated method for diagnosing a chest X-ray for diseases [Bar et al.(2015)] They used a dataset of 443 images which had three types of disease classes. In the proposed method they used a unique technique of pretraining by fusing the features from a decaf CNN and a GIST features. They proposed that the combination of the features results in a much better categorization as the combination of the features would capture the hidden features. The results are improved a little more by the fusion of the Decaf, Decaf and GIST descrip-tors. The method achieved a mean AUC of 0.87 to 0.94 for different disease classes. Alex Krizhevsky et al.(2012) realized that it was time to bring back the branch of deep learning that uses multi-layered neural networks. The availability of large sets of data, to be more specific ImageNet datasets with millions of labeled images and an abundance of computing resources enabled researchers to revive CNNs. Gao Huang et al. performed a detailed study on densely connected convolutional neural net-works and proposed a new type of neural network architecture named DenseNet [Huang et al.()].The paper proposes that the CNN can be made much deeper if the connections between the layers close to the input and the layers close to the output are made shorter. DenseNet exploits this fact and connects each layer to every other layer in a feed forward manner. It proposes to concate-nate the outputs of the previous layers. DenseNets have many plus points over the existing CNN architectures: mitigation of vanishing gradient problem, improved flow of information and reduction in number of parameters. The performance of the network was evaluated using the image recognition tasks using SVHN, CIFAR10, CIFAR-100 and ImageNet datasets. The deeper neural networks had a problem of vanishing gradient as due to a large number of layers the gradient washes out by the time it reaches the end or the beginning. The dense CNNs solve this problem. The layer in DenseNet are narrow as they add the very little number of features and do not change the existing feature maps, a classifier is used to classify the feature maps into a class. A densely connected CNN is divided into dense blocks and a transition layer to connect the dense blocks. The DenseNets outperformed the ResNets in all of the experiments despite having a lower number of parameters. The DenseNets are also less prone to the problem of over-fitting.

4 PROPOSED METHOD
In this work, the various methods were used which are explained below in each sub- section

4.1 Pre-processing of the Dataset
Pre-processing of the data plays an integral part in the performance of a trained model. In this phase, the data gets transformed and encoded into a format that the computer can process. In this phase, we have covered the operations performed on the dataset before feeding it into our model.

4.1.1. Conversion from grayscale to rgb
We have used transfer learning in this research, and the model is pre-trained on the ImageNet, the ImageNet contains RGB images. The images in the "ChestXray14" dataset are in grayscale format i.e., they have a single channel, but our model is designed to classify pictures with three channels, so we need to convert the grayscale image into RGB image. To do so, we have used a simple method in which we have just duplicated the single-channel values into three channels to convert images into RGB format**.**

4.1.2. Image data augmentation

Image data augmentation applies various transformations on images such as flipping, shifting, zooming, etc. Image data augmentation aims to feed different variations of the training set**.**

4.1.3. Splitting the dataset

The dataset which we have used in this research have 5,606 labeled images of the chest X-rays. The dataset is split into three parts: Training data (70%), Validation data(10%) and Test data(20%). Nearly 4 thousand images are used for training the model. The validation data is used for selecting the most appropriate model. The test set is used to evaluate the performance of the trained model.

4.2 CNN

CNNs are very effective at utilizing spatial and temporal dependencies, which makes CNN optimal for images. Because in pictures, the position of the pixels is equally important as their values. Primary CNNs are built using convolutional layers, pooling layers, activation functions, and fully connected layers. A sequence of the above layers can be used to create a CNN architecture. The convolutional layer can be said the core of a CNN, the convolutional layer has learnable filters (kernels) which are also known as the parameters of the convolutional layer. A filter has three dimensions (height, width and depth). At the time of forward pass the filter is kept at a pixel and dot product of the pixel values and filter is taken and the result of element wise dot product is replaced at the center pixel. This process is done at all the pixels, which gives a 2D feature map as output. Padding is often used to keep the dimensions unchanged. The convolutional layer has 4 parameters: (1) Number of filters: It is set according to the dimension of the required output. (2) Stride: It defines how many step a filter moves after each dot product. (3) Amount of zero padding: Decides whether the size will reduce or remain the same. (4) Size of filters. Activation functions are used in CNN after convolutional layer, they are used to bring the values in the feature map between a ranges. There are a wide variety of activation functions available such as sigmoid, tanh, gaussian, rectified linear unit (ReLU) etc. Among these the activation functions the most commonly used activation function after the convolutional layers is ReLU. In Fig.2. The bias values are: $b0 = [1]$ and $b1 = [0]$, which are added after the dot product of the filter and input. ReLU function keeps the positive values as it is and changes the negative values in the feature map to 0. ReLU is computationally efficient and also makes the CNN converge faster. Batch normalization is also used in CNN to normalize the values. Normalization is done according to the statistics of the mini batch. It has certain advantages: decreases the chance of getting large values, chance of over fitting is reduced, learning rate can be increased making the model learn faster. Pooling layer is used in a convolutional neural network to cut down the size of output of convolutional layer. These are used after convolutional layers which causes the reduction in the number of parameters and this decrease leads in rise of the computation efficiency. It has two parameters: size and stride. Commonly used pooling techniques are : Network (CNN/ConvNet) in Deep learning. (a) Max Pooling: It chooses the maximum value from the pool size. (b) Average Pooling: This averages out the values in the pool size. The role of the ConvNet is to reduce the images into a form that is easier to process, without losing features that are critical for getting a good prediction as shown in Fig.3.

4.3 DenseNet

The DenseNet is considered as a logical extension of the famous architecture named ResNet. It uses the concept of the concatenation of the outputs of the previous layers. DenseNet is gaining popularity because of its improved flow of informa-tion (and gradients), reuse of feature maps, a lesser number of parameters makes the training computationally efficient, and it is also less prone to overfitting. In this research we have used DenseNet-121 architecture as the feature extraction unit. We have fine tuned a DenseNet-121 which was initially trained on ImageNet. In this section we will discuss the features of DenseNet in detail and explain the architecture of the DenseNet. DenseNet has new kind of connections named Dense connections which boosts the flow of gradients and information in the same network. In the dense connection every layer is directly connected to all of the following layers. Fig.4. show how dense connection work. DenseNet is a network architecture where each layer is directly connected to every other layer in a feed-forward fashion (within each dense block). For each layer, the feature maps of all preceding layers are treated as separate inputs whereas its own feature maps are passed on as inputs to all subsequent layers. In Dense Net, each layer obtains additional inputs from all preceding layers and passes on its own feature-maps to all subsequent layers. Concatenation is used. Each layer is receiving a "collective knowledge" from all preceding layers Since each layer receives feature maps from all preceding layers, network can be thinner and compact, i.e. number of channels can be fewer. The growth rate k is the additional number of channels for each layer.
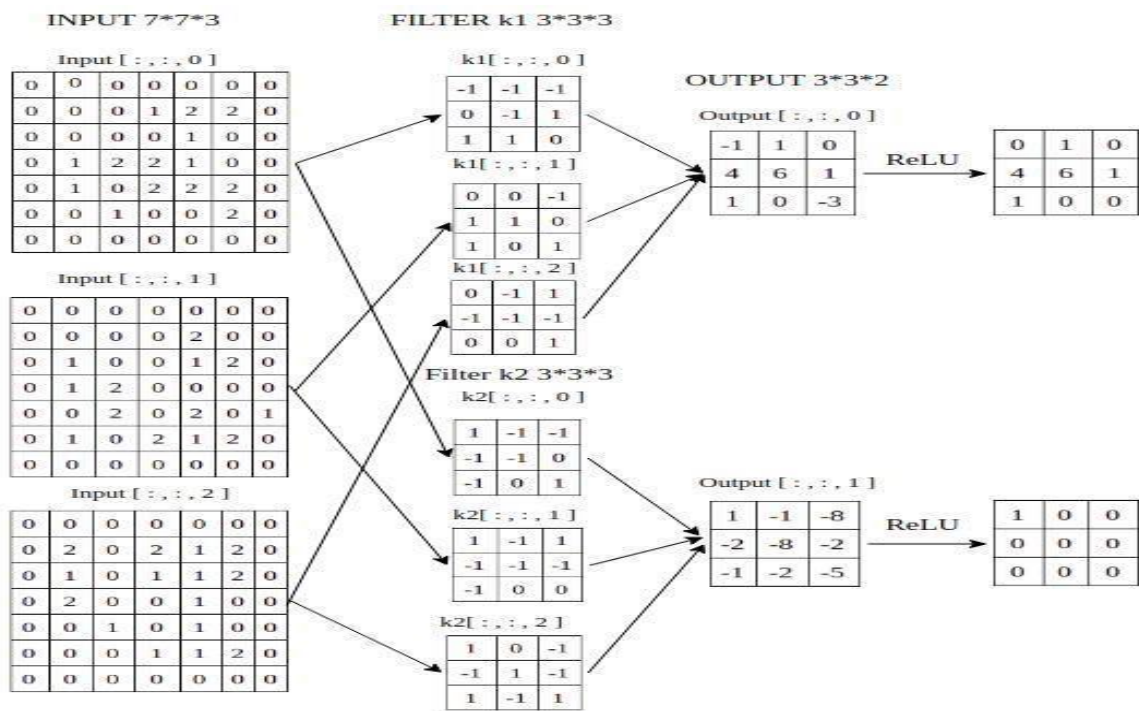
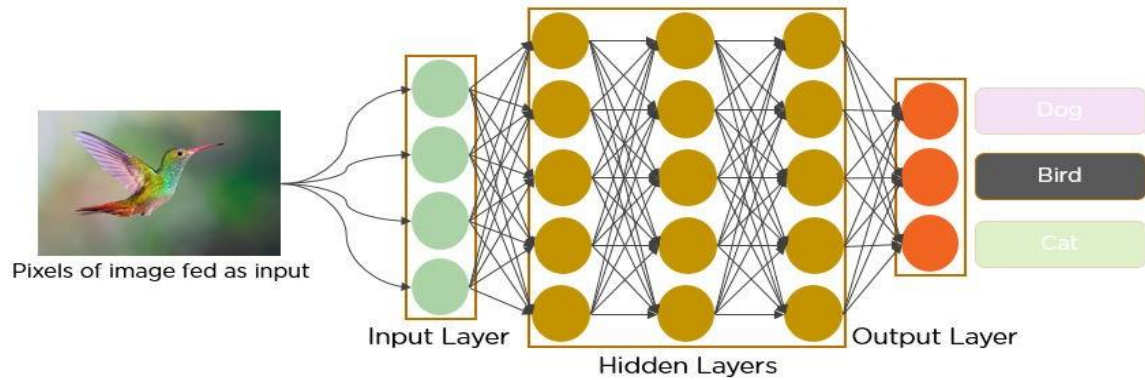Fig.2. Convolution operation with ReLU activation.
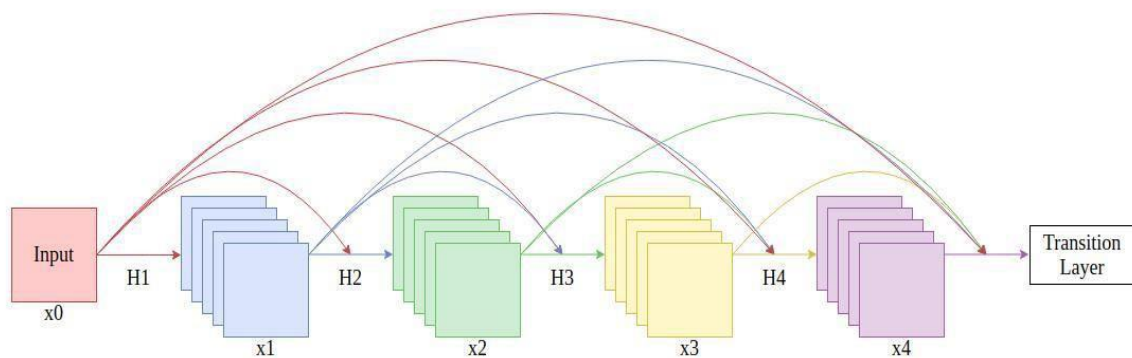
Fig.3. A Convolutional Neural

Fig.4. Multiple-Connections between densely connected layers.

KEY CONTRIBUTIONS OF THE DENSENET ARCHITECTURE

Alleviates vanishing gradient problem

Stronger feature propagation

Feature reuse

Reduced parameter count So, it have higher computational efficiency and memory efficiency. The following figure shows the concept of concatenation during forward propagation.

5. SIMULATION AND RESULTS

In this section, we will discuss about the parameters of the different techniques that are used in this research. This section also elaborates the implementation details of the research.

5.1 Dataset

The" ChestXray14" dataset was released by NIH. The sample of this dataset is used in this research. The sample dataset consists of 5,606 labeled chest X-ray images. A chest X-ray can have multiple diseases. The dataset consists of chest Xrays belonging to 14 diseases, the names of the diseases are:

Pneumonia, Fibrosis, Edema, Cardiomegaly, Atelectasis, Emphysema, Pneumothorax, Infiltration, Consolidation, Effusion, Mass, Hernia, Nodule and Pleural Thickening.

5.2 Preprocessing

In pre-processing we have firstly resized the images to 224*224 resolution. After resizing the images we have applied image data augmentation in with the images are flipped horizontally and zoomed. For image augmentation we have used Image Data Generator class of the Keras library with parameters set to "horizontalflip=True", zoom range=0.2". The pre-processing phase also includes applying CLAHE (for local contrast enhancement) on the X-ray images. The Fig. 4. shows the result of applying CLAHE on a chest X-ray image. For applying CLAHE we have used equalize adapthist class of the scikit-image library with the parame-ters:
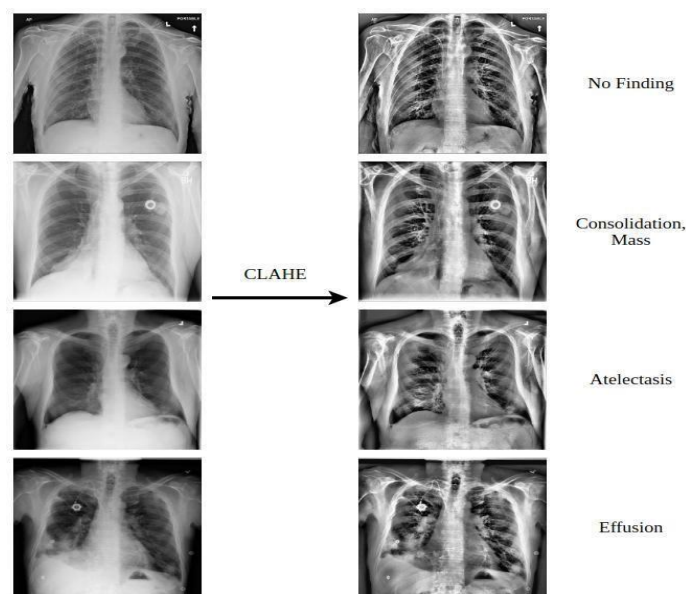

Fig.5. Effects of applying CACHE in chest X-ray

(1) kernel size = 28*28, the kernel size represents the size of the region of the image which comes under one histogram. (2) clip limit = 0.05, the clip limit controls the amount of contrast and higher values mean more contrast. (3) nbins = 256, nbins shows the number of bins in a histogram among which the equalization occurs. In Fig. 5. We can see that there is an enhancement in the visual quality of the chest X-ray image. Preparing the dataset for CNN: We have split the dataset randomly into three parts training data, validation data and testing data. Training data is 70%, validation data is 10% and the test data is 20% of the total im-ages. For splitting the dataset we have used train test split method of the sklearn library. After splitting the dataset we have created three separate data generators (for training, validation and testing) using the data from the csv file which contains the mapping from the image indices to the disease labels. Training: For fine tuning the pre-trained model we feed the train data generator of pre-processed images into the model and monitor the performance of the model on the validation data generator based on certain metrics and choose the best model according to the values of the metrics. Finally, we check the performance of the selected model on the test data generator and get the results. Optimizer : When we use class weights the range of the

loss changes, the optimizer like SGD(stochastic gradient descent) whose step size rely on the gradient may tend to fail. In the Adam optimizer the step size is not dependent on the gradient so we have used Adam optimizer for this classification task.

Metrics: The metrics on which we evaluate our model are:

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

Loss: It tells the difference between the predicted value by the model and the original value of the sample. If it is 0 then the model is perfect higher values mean more deviation from the actual labels.

AUC: It tells us about the area enclosed under the ROC(receiver operating characterstic) curve. ROC curve is a graphical plot between true positive rate(TPR) and false positive rate(FPR).

The values of AUC varies from 0 to 1, higher the value of AUC means a better model. We have trained all the 6 models on the sample dataset of ChestX-ray14 for 29 epochs with the default learning rate of Adam optimizer and tracked the progress of the model over 29 epochs on the basis of accuracy, loss and auc. The following table shows the values of the metrics of the models after training each of them for 29 epochs.
Table.1 shows the various metrics over training and validation. The accuracy and auc on the test data shown in Table.2

Table. 1. The values of various metrics over training and validation.

Metric / Model	1	2	3	4(a)	4(b)	4(c)
Training accuracy	92.57	88.83	90.37	90.45	90.14	91.26
Training loss	0.1768	1.6984	0.2554	0.4050	0.4099	0.5343
Training auc	0.8202	0.6337	0.6283	0.7976	0.7993	0.7923
Validation accuracy	92.60	88.24	89.99	90.46	90.16	91.22
Validation loss	0.2223	1.7335	0.3138	0.2619	0.3007	0.4052
Validation auc	0.8211	0.6301	0.6298	0.7983	0.7964	0.7923

Table. 2. Accuracy and AUC on the test data

Metric / Model	1	2	3	4(a)	4(b)	4(c)
Testing accuracy	93.39	90.57	91.13	90.55	88.61	74.96
Testing AUC	0.8269	0.6535	0.6526	0.8121	0.7540	0.6983

6. CONCLUSION

This research shows how the deeper CNNs help with better feature extraction from the chest X-rays. CLACHE is used for the local histogram of the X-ray images. CLACHE makes images clearer and it would be very helpful for the doctors. CAD has made tremendous progress, but the scope of CAD in chest X-rays remains limited because of the scarcity of large and well-labeled X-ray datasets. Another big problem limits the performance of CAD as compared to practitioners. The CAD systems do not know the medical history of the patient, whereas the practitioners get the help of the patient's old records. Despite the limits of convolutional neural networks, however, there's no denying that they have caused a revolution in artificial intelligence. Today, CNN's are used in many computer vision applications such as facial recognition, image search, and editing, augmented reality, and more. The advances in convolutional neural networks shows that, our achievements are remarkable and useful, But, still we are very far from replicating the key components of human intelligence.

REFERENCES

[1]. Sheffield, European Respiratory Society, (2017): The Global Impact of Respiratory Disease – Second Edition Forum of International Respiratory Societies Print ISBN:9781849840873; eISBN:9781849840880

[2]. Jaegar, S, et al.(2013): Automatic screening for tuberculosis in chest radiographs: a survey. Quantitative imaging in medicine and surgery, 3(2):89, 2013.

[3]. Choudhary, P; Hazra, A, (2019): Chest disease radiography in twofold: using convolutional neural networks and transfer learning. Evolving Systems, pages 2097-2106, 2017.

[4]. Bar,Y, et al.(2015): Chest pathology detection using deep learning with non-medical training. In 2015 IEEE 12th international symposium on biomedical imaging (ISBI), pages 294–297, IEEE.

[5]. Alex Krizhevsky et al.(2012) Introduction to Convolutional Neural

[6]. Goyal, S.B., Bedi, P., Rajawat, A.S., Shaw, R.N., Ghosh, A. (2022). Smart Luminaires for Commercial Building by Application of Daylight Harvesting Systems. In: Bianchini, M., Piuri, V., Das, S., Shaw, R.N. (eds) Advanced Computing and Intelligent Technologies. Lecture Notes in Networks and Systems, vol 218. Springer, Singapore. https://doi.org/10.1007/978-981-16-2164-2_24

[7]. Rajawat, A.S., Rawat, R., Barhanpurkar, K., Shaw, R.N., Ghosh, A. (2021). Blockchain-Based Model for Expanding IoT Device Data Security. In: Bansal, J.C., Fung, L.C.C., Simic, M., Ghosh, A. (eds) Advances in Applications of Data-Driven Computing. Advances in Intelligent Systems and Computing, vol 1319. Springer, Singapore. https://doi.org/10.1007/978-981-33-6919-1_5

[8]. Gite S.N, Dharmadhikari D.D, Ram Kumar," Educational Decision Making Based On GIS" International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-1, Issue-1, April 2012.

[9]. Rajawat, A.S., Rawat, R., Shaw, R.N., Ghosh, A. (2021). Cyber Physical System Fraud Analysis by Mobile Robot. In: Bianchini, M., Simic, M., Ghosh, A., Shaw, R.N. (eds) Machine Learning for Robotics Applications. Studies in Computational Intelligence, vol 960. Springer, Singapore. https://doi.org/10.1007/978-981-16-0598-7_4

[10]. Ram Kumar, Gunja Varshney , Tourism Crisis Evaluation Using Fuzzy Artificial Neural network, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issue-NCAI2011, June 2011

[11]. Gite S.N, Dharmadhikari D.D, Ram Kumar," Educational Decision Making Based On GIS" International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-1, Issue-1, April 2012.

[12]. Ram Kumar, Sarvesh Kumar, Kolte V. S.," A Model for Intrusion Detection Based on Undefined Distance", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1 Issue-5, November 201

[13]. Rawat, R., Rajawat, A.S., Mahor, V., Shaw, R.N., Ghosh, A. (2021). Surveillance Robot in Cyber Intelligence for Vulnerability Detection. In: Bianchini, M., Simic, M., Ghosh, A., Shaw, R.N. (eds) Machine Learning for Robotics Applications. Studies in Computational Intelligence, vol 960. Springer, Singapore. https://doi.org/10.1007/978-981-16-0598-7_9

[14]. Ram Kumar, Jasvinder Pal Singh, Gaurav Srivastava, "A Survey Paper on Altered Fingerprint Identification & Classification" International Journal of Electronics Communication and Computer Engineering Volume 3, Issue 5, ISSN (Online): 2249–071X, ISSN (Print): 2278– 4209

[15]. Bedi, P., Goyal, S.B., Rajawat, A.S., Shaw, R.N., Ghosh, A. (2022). Application of AI/IoT for Smart Renewable Energy Management in Smart Cities. In: Piuri, V., Shaw, R.N., Ghosh, A., Islam, R. (eds) AI and IoT for Smart City Applications. Studies in Computational Intelligence, vol 1002. Springer, Singapore. https://doi.org/10.1007/978-981-16-7498-3_8

[16]. Kumar, R., Singh, J.P., Srivastava, G. (2014). Altered Fingerprint Identification and Classification Using SP Detection and Fuzzy Classification. In: , et al. Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012. Advances in Intelligent Systems and Computing, vol 236. Springer, New Delhi. https://doi.org/10.1007/978-81-322-1602-5_139