Pneumonia Detection Using X-ray Images by Shallow- CNNs

[1]Khushboo Lalwani, [2]Rashmi Upasani

Student
Master of Computer Applications
Medi-Caps University
Indore, India

***Abstract*— Pneumonia is a hazardous interstitial lung disease that is considered life-threatening and is responsible for taking away the lives of infants, children under the age of five and adults over the age of 65. Pneumonia is a respiratory infection that causes inflammation in the air sacs of one or both lungs and is mainly caused by viruses, fungi and bacteria. According to google, over 10 million people in India are diagnosed and affected by CAP every year. WHO (World Health Organization) reported that every one in three deaths is caused by pneumonia in India, which is extremely alarming.**
Detection of pneumonia on early stages can be crucial as it helps in speeding-up the recovery rate in children as well as adults. In order to increase and improve the efficiency of the predicted results, chest x-ray images are proven to be helpful as it helps the doctors to examine your lungs and blood vessels. These images help radiologists to look for white spots in your lungs which indicates infection but to obtain the results that are that are as true as he predicted results, deep learning has changed the game in the field of *medical image analysis* as neural networks are impacting the accuracy of the results drastically since the last few years. A Convolutional Neural Network (CNN or ConvNet) is a type of deep learning network that can learn patterns directly from data. CNNs were created expressly to recognise patterns in images, allowing them to distinguish between objects, classes, and categories. In addition to categorising image data, they can also be used successfully to categorise audio, time-series, and signal data.

To make things less complex and more efficient, we trained the model with less parameters so that it is shallower in depth and easy to train with more *accurate* results. We used a dataset [5] of 5125 x-ray images where 4172 images belong to the training set, 1044 images belong to the validation set and 624 images belong to the test set. This chest x-ray dataset [5] was distributed and labelled between 2 classes, 0 and 1 respectively indicating normal (no pneumonia) and pneumonia. On training this dataset [5], the results received for validation set was that validation loss was brought down to 0.1560 and the validation accuracy received was 0.94 which is equivalent to 94%. In the testing set, the test loss was brought down to 0.4381 and the test accuracy obtained was 0.8324 which is equivalent to 83.24%.

***Index Terms*—CNN, deep learning, image classification, medical image analysis, Pneumonia, x-ray images.**

I. INTRODUCTION

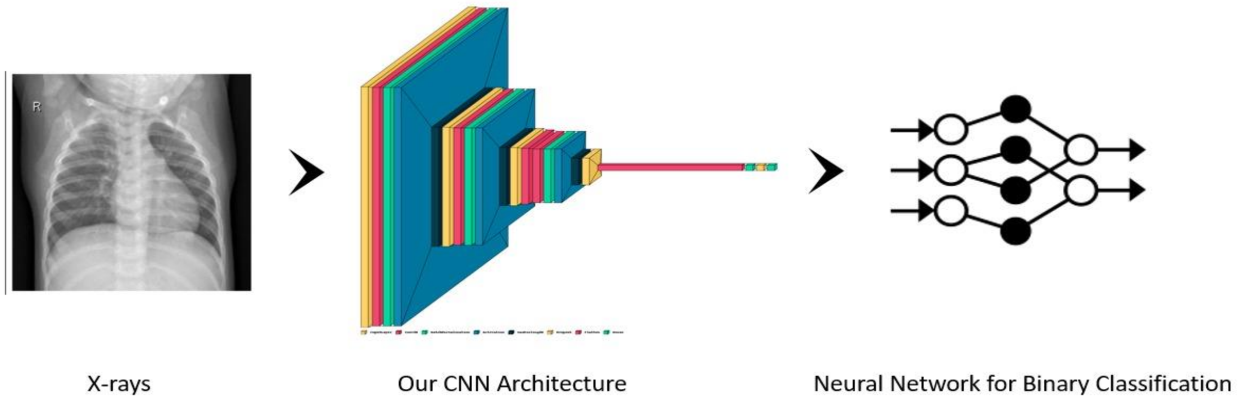


Figure 1: Structure of the method used

All manuscripts must be in English. These guidelines include complete descriptions of the fonts, spacing, and related information for producing your proceedings manuscripts. Please follow them.
Pneumonia is a frequent and serious respiratory infection and lung parenchyma inflammation that affects the lungs severely. It is brought on by a number of infectious pathogenic microorganisms, including bacteria, viruses, fungi, and parasites. Due to pneumonia, your lung tissues start to swell up due to inflammation which causes pus and fluids in your lungs. Streptococcus pneumoniae bacteria, also known as pneumococcal disease, is responsible for CAP (Community-acquired pneumonia) alongside with Mycoplasma pneumoniae, Chlamydophila pneumoniae. Although anybody can get pneumonia, severe instances are more common in new-born, young children, the elderly, and persons with compromised immune systems. Pneumonia symptoms include coughing, fever, breathing problems, exhaustion, and muscular aches. Pneumonia can be categorized in a variety of ways that are often used.

Researchers from all over the world are actively recommending more accurate and fast diagnostic techniques due to the significant incidence of pediatric pneumonia mortality. Numerous measures have been created as a result of technological advancements, with radiology-based approaches being the most well-liked and effective. Computed tomography (CT), magnetic resonance imaging (MRI), and chest X-ray imaging are examples of diagnostic radiological procedures for pulmonary disorders. The most efficient and successful of these is chest X-ray imaging since it is portable, extensively used in hospitals, and exposes patients to less radiation.
Even the doctors, who are trained professionals and experience in this field, have found it difficult to correctly identify the pneumonia from X-ray scans. This is because different diseases, including lung cancer, have similar regions of information in their X-ray images. As a result, conventional methods for diagnosing pneumonia require a lot of time and effort, and it is not always possible to determine a patient's status as having pneumonia using a standardized procedure. In order to resolve this issue, we have used CNN (convolutional neural network) to detect and diagnose pneumonia through x-ray images to generate the accurate and efficient results of 94% on validation set and 83% on test set. Due to the success rate of deep learning in medical image analysis and results, we used CNN (Convolutional neural networks) which works on large as well as small datasets which are equally useful in image classification and feature extraction. CNNs are made up of several layers of interconnected nodes, each layer carrying out a particular kind of computation. A convolutional layer serves as the first layer in a CNN and typically applies a collection of learnable filters to the input data in order to extract local characteristics from the input image. Pooling procedures, which lower the data's dimensionality, and fully linked layers, which carry out further processing and classification, are frequently used in subsequent layers.
The capacity of CNNs to automatically learn features from raw data, without the need for manual feature engineering, is one of their main advantages. Due to the complexity and difficulty of manually defining the underlying features in tasks like object recognition, image segmentation, and speech recognition, they are particularly effective in these applications. Our CNN is substantially shallower and has significantly less trainable parameters when compared to other research publications in the field of deep learning and the use of convolutional neural networks (CNNs), which are often deep and need significant processing resources and time to train. Our CNN performs almost as well as deeper networks with more parameters while having only 2.6 million parameters and taking less than an hour to train.

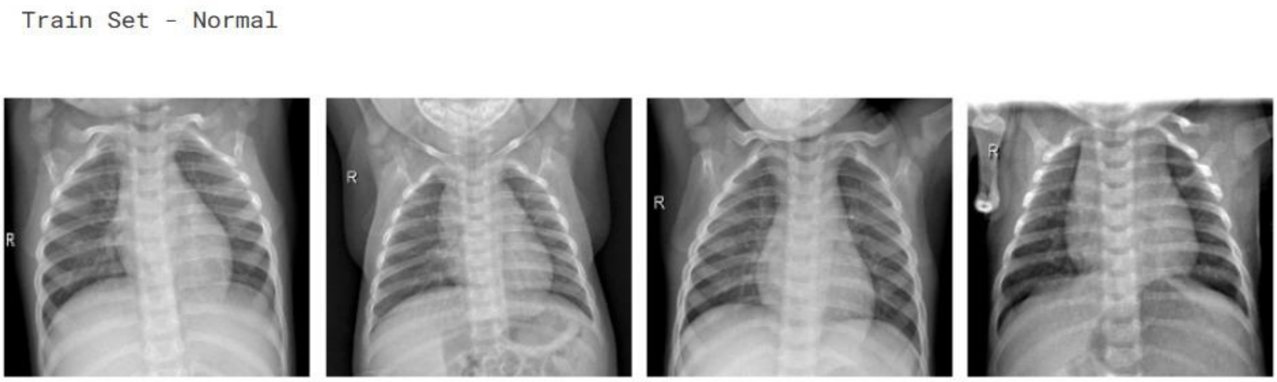


Figure 2: plotted X-ray images for class normal (non-pneumonia)

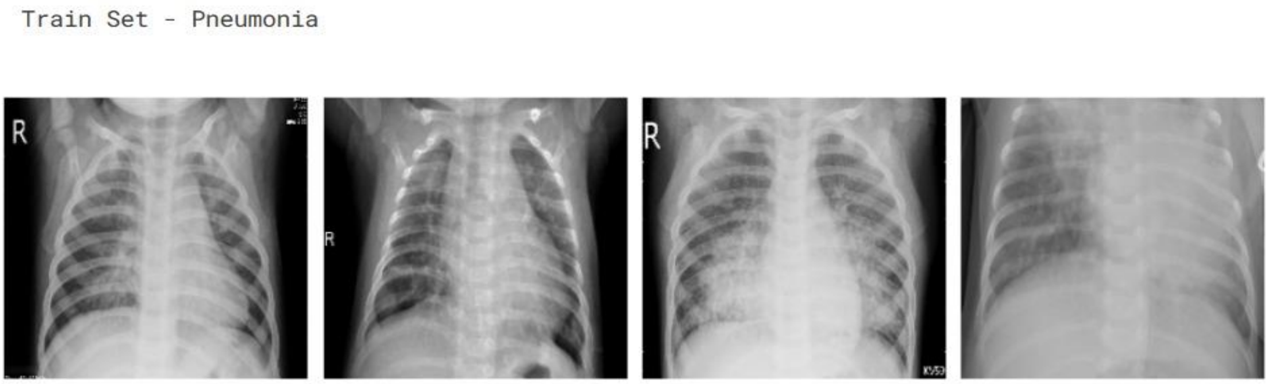


Figure 3: plotted X-ray images for class pneumonia

II. Related Work

Some related work done in the field of pneumonia detection using chest X-ray images by CNN (convolutional neural network) are as follows:

- In Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network by Dejun Zhang et al, they utilized the Dynamic Histogram Equalization (DHE) technique to enhance the image contrast. Numerous analyses of various input shapes and loss functions were offered to demonstrate the performance of our suggested model. [2]
- In Pneumonia Detection Using CNN based Feature Extraction by Dimpy Varshni et al, the customized model was combination of CNN based feature-extraction and supervised classifier algorithm resulted in optimal solution for classifying abnormal (Pneumonia labelled) and normal Chest X-Ray images primarily due to the substantive features provided by Dense Nets followed by optimal hyper-parameter values of SVM (Support Vector Machine) classifier. [3]

- In Pneumonia Detection Using Convolutional Neural Networks (CNNs) by Rachna Jain et al. it was concluded that CNN classifier which used data augmentation technique can, therefore, be effectively used by medical officers for diagnostic purposes for early detection of pneumonia in children as well as adults. [1]
- Pneumonia Detection Using an Improved Algorithm Based on Faster R-CNN (Regions with convolutional neural network) by Shangjie Yao et al. This paper uses Faster R-CNN model to suggest an enhanced approach for pneumonia identification. To increase accuracy and efficiency, the suggested algorithm incorporates a feature pyramid network and a region proposal network. [8]
- A two-stage deep learning model for the diagnosis of pneumonia was suggested in a study that was published in the Journal of Biomedical Science and Engineering. A support vector machine (SVM) was utilized for classification in the second stage and a CNN for feature extraction in the first stage [9].
- In CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning by Pranav Rajpurkar, they add a fully connected layer that generates a 14-dimensional output in place of CheXNet's last fully connected layer, and then they apply an element-wise sigmoid nonlinearity.[4]

.
The amount of convolution layers used in all of the deep learning architectures mentioned above that are based on or centered around convolutional neural networks is enormous, which makes the network deeper but also increases the number of trainable parameters, making the model difficult to train and fine-tune and increasing the computational cost of the process. Our approach requires less memory during training, which makes it simpler to train and also yields outcomes at a comparable level.

III. MATERIALS AND METHODS

I. How CNNS works?

Convolutional neural networks (CNNs) are a specific sort of neural network that often consist of multiple layers. The first layer is the convolution layer, which scans the input according to its dimensions while performing convolution operations using filters. This layer outputs a feature map or activation map and has hyperparameters like filter size and stride.
The pooling layer, which often comes after a convolution layer and does down sampling, is the second layer. Common pooling techniques that use the maximum and average values are max pooling and average pooling. The fully connected layer, which operates on a flattened input and connects each input to all neurons, is often the final layer of a CNN architecture. Fully connected layers, if they exist, are often employed towards the ending of the network, and can enhance goals like test results.
In summary,

- A CNN architecture, in its most basic form, is a collection of Layers that turn an image volume into an output volume, which could contain the class scores.
- CNN architectures are made up of various Layers, such as pooling, fully connected, and convolution.
- Each Layer receives a 3D volume input and transforms it into an output 3D volume using a differentiable function.
- Depending on the type of Layer, each Layer may or may not have parameters. Convolution and fully connected Layers, for example, include parameters, whereas ReLU and pooling Layers do not.
- Each Layer may or may not have its own set of hyperparameters.

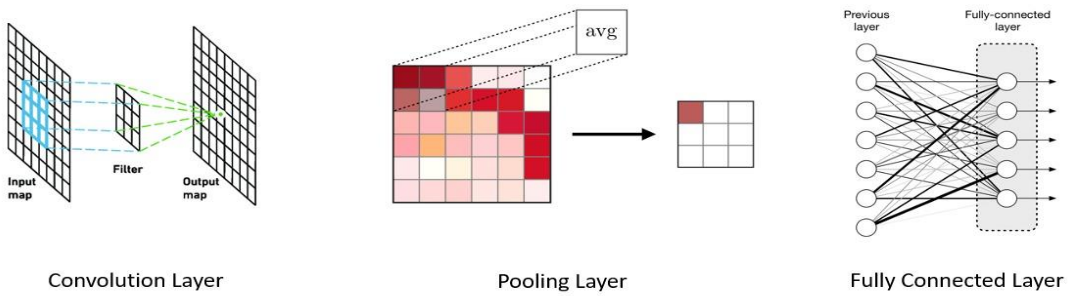


Figure 4: Types of Layers in Convolutional Neural Network

II. How to tune CNNs to get optimized output

Data augmentation is a technique to create new, synthetic data from existing data. A process known as "data augmentation" that may be used to artificially increase the size of a dataset. The original data is exposed to a number of alterations in order to achieve this. Establishing the architecture involves deciding on the number of layers, the type of activation functions to be used, the number of neurons in each layer, and the connections between them. Hyperparameters are settings that are specified explicitly by the user rather than learned from data. Examples of hyperparameters include learning rate, epochs, batch size, and regularization strength. The process of methodically adjusting the settings of these hyperparameters and assessing the model's performance on a validation set in order to discover the ideal combination is known as hyperparameter tuning.

III. Utilization of CNNs in the field of Pneumonia detection using Chest X-ray images:

Due to their capacity to automatically identify pertinent features from images, CNNs are frequently utilized in the medical industry to perform image classification tasks.

CNNs have demonstrated encouraging results in reaching high accuracy and sensitivity in recognizing patterns linked with pneumonia in the case of pneumonia detection utilizing X-rays. CNN can efficiently collect spatial characteristics within the X-ray image by using convolutional and pooling layers, while preventing overfitting by using dropout and batch normalization

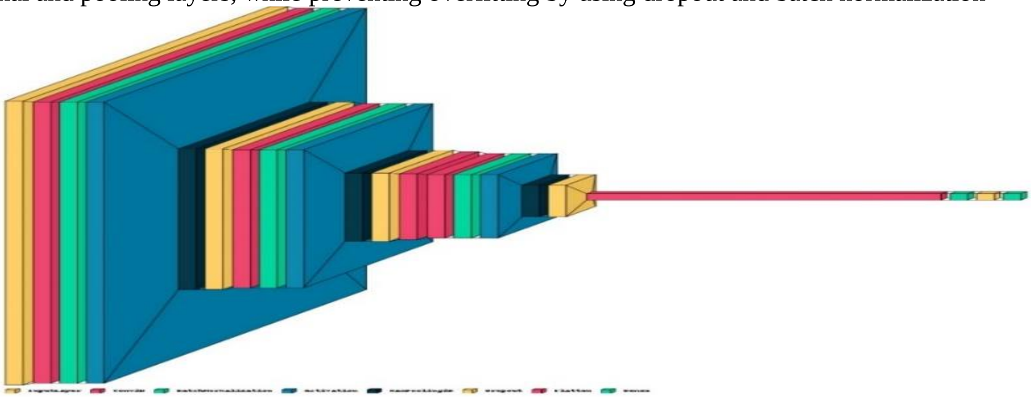


Figure 5: Visual representation of proposed model

.

IV. IMPLEMENTATIONS AND RESULTS

I. Implementing our proposed Convolutional Neural Network

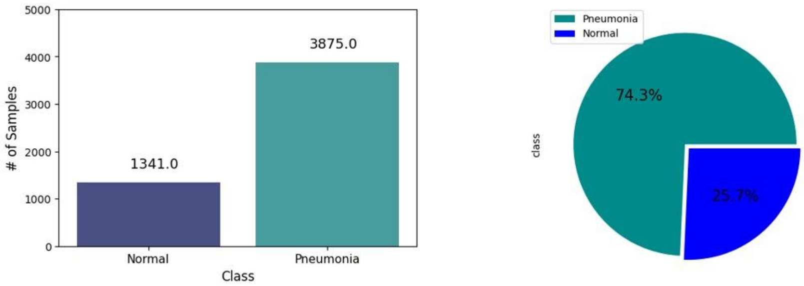


Figure 6: Distribution of images in both classes in dataset

• **Analyzing the dataset:**

Our research utilized a dataset [5],[10] of 5,863 JPEG images of chest X-rays, with 2 categories: Pneumonia and Normal. The images were selected from pediatric patients aged one to five years old who received routine clinical care at Guangzhou Women and Children's Medical Centre. Our dataset [5],[10] was split into 3 parts: Training set (4172 images), Validation set (1044 images), and Test set (624 images). The distribution of images belonging to each class was plotted to visualize the data distribution.

• **Data augmentation:**

The procedures for augmentation include rescaling each image's pixel values by dividing by 255, which normalizes the pixel values to be between 0 and 1. the images are randomly zoomed in to a maximum of 10% of their original size. randomly selecting the images' width and height by a maximum of 10% of their original dimensions. By adding additional variety to the training data, these augmentation strategies can improve the model's ability to identify patterns in the data and lessen overfitting.

• **Setting the Call-backs:**

We created an Early Stopping call-back method that keeps track of the validation loss, or 'val_loss'. If the validation loss does not decrease over enough epochs, the training will end. In other words, training will end early to avoid overfitting if the validation loss does not reduce over time, and best weights will restore the best model's weights during training. We implemented The ReduceLROnPlateau call-back function is used to lower the optimizer's learning rate. The same validation loss metric as early stopping is monitored by it. The optimizer's learning rate will be lowered by a factor specified in the model if the validation loss does not decrease for the specified number of epochs.

• **Defining the CNN architecture:**

- The input layer takes an image with the shape (IMG_SIZE, IMG_SIZE, 3), where 3 stands for the number of color channels. IMG_SIZE if of 224, which makes the final input size (224 x 224 x 3).
- A convolutional layer with 16 filters, a 3x3 kernel size, and "valid" padding make up the first block. Convolution is followed by batch normalization, ReLU activation functions, and max pooling with a 2x2 pool size. The max pooling layer is followed by a dropout layer with a 0.2 dropout rate.
- The convolutional layer in the second block has 32 filters instead of the first block's 16.
- The third block consists of two convolutional layers placed one after the other, each with 64 filters, a 3x3 kernel, and "valid" padding. After each convolution, batch normalization and ReLU activation functions are used, and then max pooling with a 2x2 pool size is performed. After the max pooling layer, a dropout layer with a 0.4 dropout rate is added.
- The output of the convolutional blocks is flattened before being sent through a layer with 64 neurons and a ReLU activation function that is fully linked. After the completely connected layer, a dropout layer with a dropout rate of 0.5 is added.

- The final layer consists of a single neuron with a sigmoid activation function, which outputs a value between 0 and 1. This output represents the predicted probability of the input image being in the positive class (Pneumonia).

• **Training the model:**
For training the model, we choose the loss function, optimizer, and evaluation metric. In binary classification, the binary cross-entropy loss function is frequently used. Since binary accuracy is the parameter used to assess the model's performance during training, the Adam optimizer is frequently used with a learning rate of 3e-5 to train deep neural networks so we used the same learning rate. The model is fitted to the training data in batches during training, and backpropagation is used to update the model's weights. The training process continues for the specified number of epochs which are 50 in our case until the early stopping criterion is met.

• **Testing Methodology:**
By comparing the predicted labels to the actual labels in a dataset, a confusion matrix is a table that summarizes the effectiveness of a classification model. It consists of four main components: true positives, false positives, true negatives, and false negatives. These components provide insight into the accuracy and effectiveness of a classification model.
To further analyze the result of the algorithm’s performance on test set, we plotted confusion matrix and calculated precision, recall and f1 score. Our model achieved precision of 0.92 on class “normal” and 0.80 on class” pneumonia” and the recall of 0.59 and 0.97 on both respective classes (out of /1).

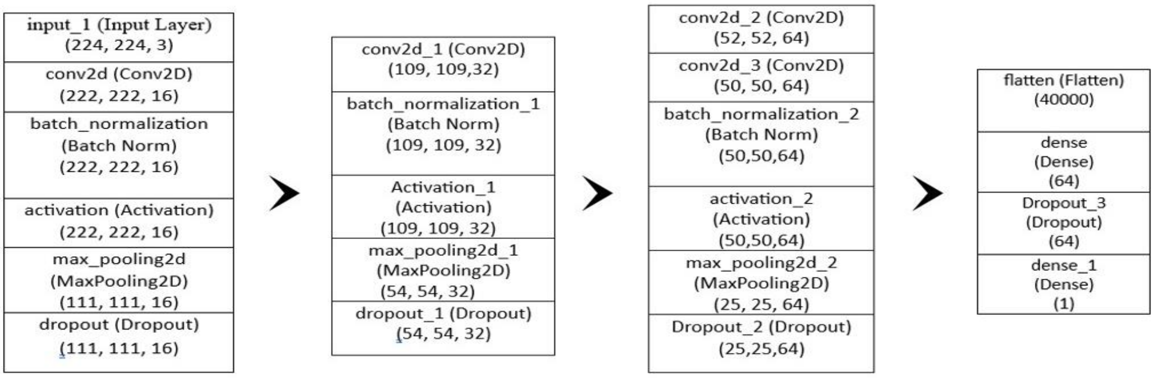
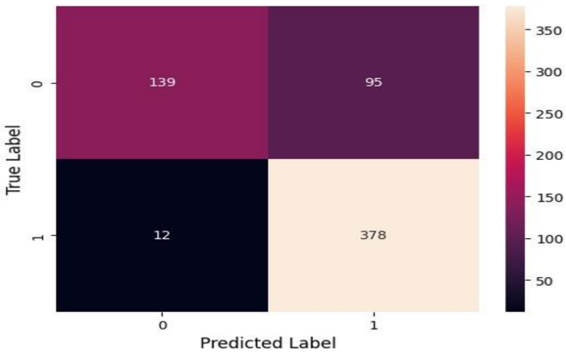


Figure 7: Proposed CNN architecture with 3 convolutional blocks with output sizes



Confusion Matrix

	precision	recall	f1-score	support
0	0.92	0.59	0.72	234
1	0.80	0.97	0.88	390
accuracy			0.83	624
macro avg	0.86	0.78	0.80	624
weighted avg	0.84	0.83	0.82	624

Precision, Recall , f1 score

Figure 8: Model Evaluation Metrices

The AUC (Area Under the Curve) curve is a graphical representation of the performance of a binary classification model. For different classification thresholds, it plots the true positive rate (sensitivity) compared to the false positive rate (1-specificity). No matter what classification criterion is selected, a higher AUC value suggests that the model is performing better overall, and our model achieved AUC score of 0.90.

II. Analyzing the performance architecture of our algorithm:

Our model underwent 50 epochs of training on the training data with a batch size of 32. A learning rate of 0.00003 was utilized using the Adam optimizer. To avoid overfitting, early halting and plateau call-backs were used. On the training data, the model had a binary accuracy of 96.03%, and on the validation data, it had an accuracy of 94.25%. Each time the validation loss did not decrease after two epochs, the learning rate was decreased by a factor of 10, our model stopped after 11 epochs as it faced a plateau in training and also the performance reached at a desirable level. Overall, our model completed the task successfully and accurately on both the training and validation data.
The total time taken by our model was around 24 minutes which is much faster than many deep CNNs which gives the same level of performance but takes much longer to train. We plotted the graphs of how loss decreased during the training process and how the accuracy increased during training process in Fig.9.

Our validation accuracy, precision, recall and AUC (Area Under Curve) with respect to other CNNs used for the same purpose of pneumonia detection. As shown below in the Fig.10, our model has the best validation accuracy, precision and recall in comparison to other models [1],[2],[3],[8], values are sorted in ascending order:

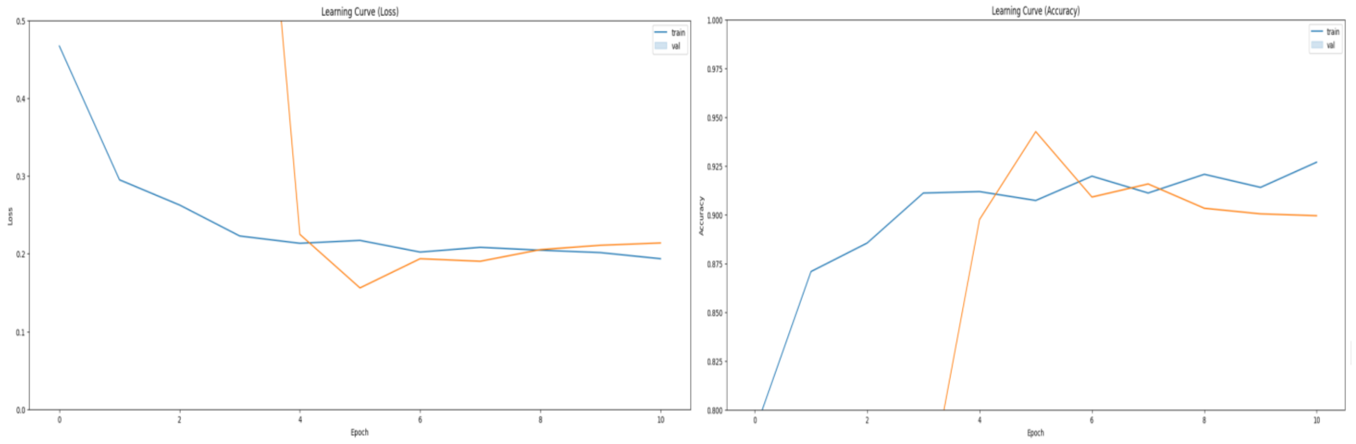


Figure 9: Decreasing loss plotted against epochs in the left plot and Increasing accuracy plotted against epochs in the right

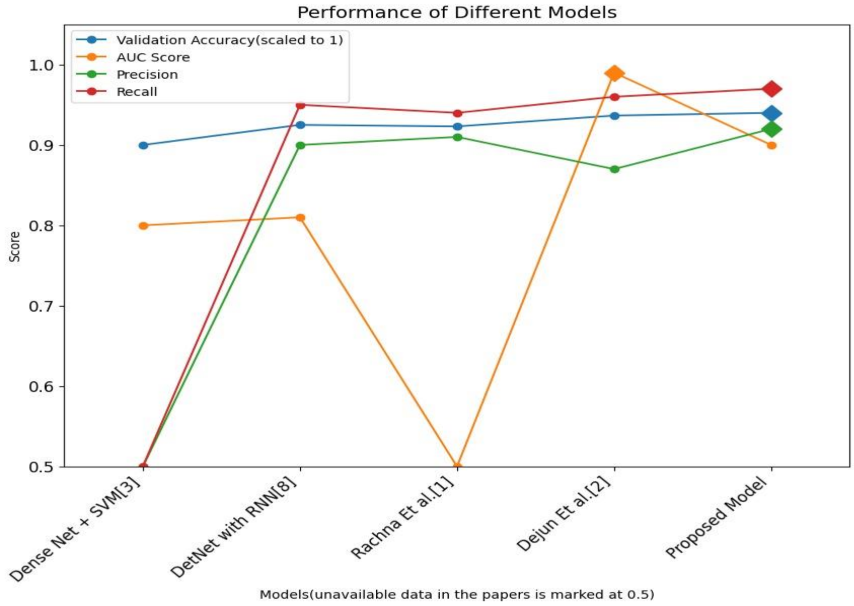


Figure 10: Maximum val. acc., precision, recall achieved by our model, Graph in ascending order. Maximum values marked by diamond

III. Comparative Analysis

Our model has significantly fewer trainable parameters than other commonly used CNNs, including the state-of-the-art ResNet-50 (Residual Networks of depth 50) [11]. In fact, our model is only 1/16th the size of ResNet-50[11]. The chart in Fig.11 below provides a comparison of the parameter counts for various modern CNNs, which serve as the basis for pneumonia detection tasks. This Fig.11 chart highlights the efficiency and practicality of our model for real-world applications.

In comparison to other models used in related works, our proposed model achieves higher training accuracy with a score of 96%, compared to 91% and 95% accuracy for Dense Net + SVM [3] and Rachna et al. [1], respectively. Regarding validation accuracy, our model achieves a score of 94%, which is comparable to Dejun et al. [2] and Rachna et al. [1], who achieve accuracies of 93.66% and 92.31%, respectively.

In a previous study, Customs CNN [2] utilized image enhancement techniques to improve learning, but the model overfit, resulting in an AUC score of 0.99. In contrast, our 3-convolutional block CNN achieves an AUC score of 0.90, indicating that it has not overfitted. Despite being a shallower network, our proposed model outperforms the deeper DenseNet121 + SVM [3] approach, which achieved an AUC score of 0.81.

In terms of precision and recall, our model performs comparatively well among its deeper and heavier counterparts, achieving a precision of 0.92 and a recall of 0.97. While DetNet with RCCN [8] performs well with a precision of 0.92 and recall of 0.95, other models such as Rachna et al. [1] and Dejun et al. [2] achieve recalls of 0.94 and 0.96, respectively. Overall, our proposed model stands out by a very close margin in terms of its performance compared to other models.

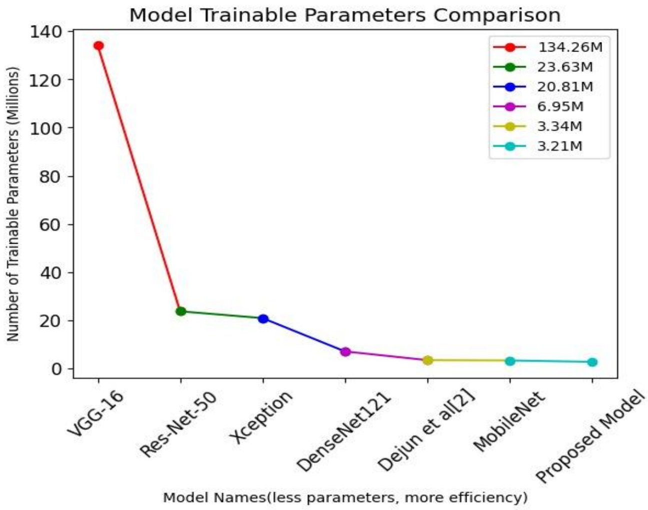


Figure 11: Chart showcasing efficiency of our model through parameter comparison

Table 1: Comparative Analysis Table (Information not provided is marked by -)

	Comparison Metrics			
Models	Training Accuracy	Validation Accuracy	AUC score	Precision/Recall
Proposed Model	**96.03**	**94.00**	**0.90**	**0.92/0.97**
Dense Net + SVM[3]	91.0	90.0	0.80	–
Dejun et al.[2]	–	93.66	0.99	0.87/0.96
Rachna et al.[1]	95.00	92.31	–	0.92/0.94
DetNet with RCNN[8]	–	92.5	0.81	0.90/0.95

IV. Final result after applying the trained algorithm on the test set

After training the model on training and validation dataset [5],[10], it achieved a respectable accuracy of 94.25% and the loss brought down to 0.156. To evaluate the model's performance in real-world situations, it was tested on a completely new and unseen dataset of 624 X-rays. The model performed well on the test dataset, achieving an accuracy of 83% and a test loss of 0.438. This is a remarkable accomplishment for a shallow CNN, especially when compared to its deeper counterparts.

Table 2: Result table

Validation Loss	0.156
Validation Accuracy	0.9425
Validation Accuracy in %	94.25%
Test Loss	0.438
Test Accuracy	0.8324
Validation Accuracy in %	83.24%

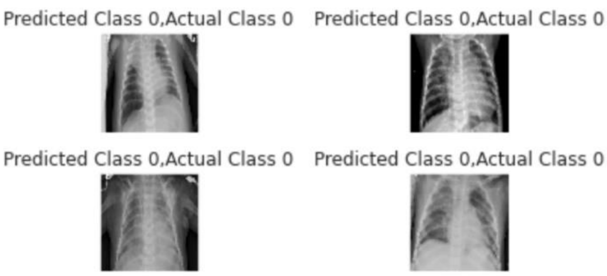


Figure 12: Prediction of the model

V. CONCLUSION AND FUTURE WORK

Convolutional Neural Networks (CNNs) have proven to be highly advantageous in medical image analysis. In particular, deep learning algorithms have the potential to significantly enhance diagnostic accuracy for medical practitioners. Our research findings confirm the usefulness of CNNs for detecting pneumonia using X-ray images.

Compared to previously used CNN models in pneumonia detection, our CNN has a relatively shallow depth, with only 2.6 million parameters and 3 convolutional blocks. Despite this, our CNN achieved impressive accuracy rates of 94% on validation and 83% on the test set. Precision and recall metrics greater than 0.9 were also achieved, and the model required minimal training time.
Our findings suggest that our CNN represents a fast and efficient model for detecting pneumonia using chest X-ray images. Future improvements could include training the model on larger, high-quality datasets, exploring different activation functions, experimenting with various layer combinations, and implementing different optimization techniques. These optimizations could potentially improve the model's performance and accuracy

REFERENCES:

1. Rachna Jain, Preeti Nagrath, Gaurav Kataria, V. Sirish Kaushik, D. Jude Hemanth, "Pneumonia Detection Using Convolutional Neural Networks (CNNs)" Conference: Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019), pp 471 – 472, April 2020.
2. Dejun Zhang, Fuquan Ren, Yushuang Li, Lei Na and Yue Ma, "Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network" Electronics, EISSN 2079-9292, Vol. 10, issue 13, Article no. 1512, pp – 2,3,4,5, 12, 13 23rd June 2021.
3. Dimpy Varshni, K. Thakral, Lucky Agarwal, Rahul Nijhawan, A. Mittal, "Pneumonia Detection Using CNN based Feature Extraction" Conference: 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT) Coimbatore, India, pp 4-5, 1st Feb 2019.
4. Pranav Rajpurkar et al, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning." Stanford ML lab, print id arxiv – 1711.05225, pp 3- 5, 25th Dec 2017.
5. Labelled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification. Mendeley data, Applied Sciences section Ver. 2, 6th Jan 2018.
6. Daniel S. Kermany et al, "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," Cell, Vol 172, issue 5, pp 1-2, 22nd Feb 2018.
7. Sammy V. Militante, Brandon G. Sibbaluca, "Pneumonia Detection Using Convolutional Neural Networks" INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH Vol 9, issue 4, pp 1335-1336, April 2020.
8. Shangjie Yao Yaowu Chen and Rongxin Jiang, "Pneumonia Detection Using an Improved Algorithm Based on Faster R-CNN", Hindawi Computational and Mathematical Methods in Medicine, Volume 2021, Article ID 8854892, pp 9 – 12, 21st April 2021.
9. Hassan Ahmed El Shenbary, Ebeid A Ebeid, Dumitru Baleanu, "COVID-19 classification using hybrid deep learning and standard feature extraction techniques", Indonesian Journal of Electrical Engineering and Computer Science Vol. 29, No. 3, pp. 1787- 1789, 30th Oct,2022
10. Paul Mooney, "Chest Xray Pneumonia Dataset", Kaggle.
11. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition", print id - arXiv:1512.03385, 10th Dec, 2015