Journal of Electrical Engineering, Electronics, Control and Computer Science – JEEECCS, Volume 7, Issue 26, pages 33-40, 2021

Childhood Pneumonia Recognition using Convolutional Neural Network from Chest X-ray Images

Diponkor Bala
Department of Computer Science and Engineering
Islamic University
Kushtia-7003, Bangladesh
diponkor.b@gmail.com

Abstract*—In Bangladesh, Pneumonia is accountable for around 28% of mortality among children under the age of five. According to a recent study, one child dies from pneumonia every 39 seconds around the world. Pneumonia occurs because of microscopic organisms, infections or growths, and leaves youngsters battling for breath as their lungs load up with discharge and liquid. Delay in looking for suitable consideration and access to different sources for treatment are the fundamental hazard factors for pneumonia demise in small kids in Bangladesh. To fight this problem, this country needed an easily accessible quick solution. Due to the advancement of technology, Artificial Intelligence has made human life easier by using various machine learning techniques. In the last few years, Convolutional Neural Network has become an effective way for the classification of multiclass images. This paperwork has been developed a system in order to provide a simpler way to detect pneumonia in a short period of time. The model is developed with chest x-ray images taken from the frontal views to identify pneumonia. In this paper, Convolutional Neural Network (CNN) was utilized for the recognition of pneumonia disease. The CNN model is trained on a data set of 5,200 recently collected x-ray images. The dataset was divided into two classes, normal and pneumonia x-ray images. I have trained my model on different size x-ray images to evaluate its performance. This model can successfully detect pneumonia at an accuracy rate of approximately 98.87%. The proposed CNN model improves the pneumonia recognition accuracy than some existing methods.

Keywords-Pneumonia Recognition; Chest X-ray; Classification; Deep Learning; Convolutional Neural Network

I. INTRODUCTION

Pneumonia disease is the largest single risk factor for mortality in children under the age of five and it is a major determinant of child deaths throughout every region of the earth, with Inter Africa and South Asia suffering the effects of the impact. Pneumonia kills more children under the age of five than other diseases like AIDS, malaria, as well as influenza, despite greater focus more on those illnesses in recent times [1]. Pneumonia is a terrible disease in Bangladesh, responsible for almost 28% of child death under the age of 5 years, which is around 50000 children in number, 2500 a day. Every year, 80000 children are referred to hospitals with this viral respiratory disease, despite the fact that the overall patient count is substantially greater. [2]. In Bangladesh, four children are killed by pneumonia every three hours. Bangladesh was placed 14th out of 15 countries for having the highest number of pneumonia deaths, followed by its South Asian countries. India came in second with 127,000 child fatalities, while Pakistan came in third with 58,000. According to the report, pneumonia was the third leading cause of infant death in Bangladesh in 2017. In 2018, four children under the age of five died as a result of this acute disease for every 1,000 live births [3]. According to World Health Organization (WHO), just household pollution is responsible for 4 million premature death each year, around the world, causing pneumonia [4].

The death rate of pneumonia is so high due to wrong or delayed treatment. An accurate diagnosis is necessary for effective pneumonia therapy, although this can be challenging. In a common scenario, doctors often make mistakes to detect pneumonia on time and the response time for this disease is very narrow, less than two days [5]. So, a fast-detecting process is very necessary to ensure proper treatment for a pneumonia patient. Pneumonia is detected by Chest X-Ray, which depends on the availability of expert radiologists or doctors, which leads to a severe chance of mistakes, ignorance, and lack of medical resources and personals. Excess fluid, internal hemorrhage, lung cancer, and other medical disorders can all cause comparable opacities in a CXR [6]. A detailed analysis of the CXR takes a bit of time, and some clinicians must evaluate a high number of CXRs in a short time period, which is one of the key challenges with treatment. A computerized clinical guideline is essential to get rid of those situations. We are developing a model to dig up pneumonia from a chest x-ray efficiently and precisely. Our model's purpose is to predict whether a patient has pneumonia or not so that a clinic can take action as soon as possible and can prevent a child from death.

By deep learning, we can easily learn a machine and can have a machine or software or anything related to deep learning or artificial intelligence to overcome our daily life problems. Deep Neural Network [7] models are usually designed and checked by human experts experimenting with them in real-time, using trial and error method. A Deep Neural Network Architecture is used to reduce time to optimize the classification of recourses which is very crucial for this model. This particular model is designed to classify pneumonia chest x-ray images classification tasks using a convolution neural network algorithm, processing several related information on a given image and detecting relevant data from it. This model is aimed to reduce the processing cost and time comparing other conventional pneumonia classification processes. The deep learning algorithm, using CNN has become common in medical image classification nowadays [8]. CNN exhibits impressive results comparing the human examined one. CNN deep learning architecture has two-part, first extraction and input image encoding using convolutional layers CNN and the second one shows a prediction model for the classification task using a fully connected neural network classifier. CNN models optimize many hyperparameters like convolutional layer numbers, filter numbers, and filter sizes.

My latest studies show that latest improvement in deep learning models with the availability of huge datasets with the availability of algorithms taking huge impacts on medical fields which deal with medical images tasks such as skin cancer classification [9], hemorrhage identification [10], arrhythmia detection [11], and diabetic retinopathy detection [12]. So, we build a model which also deals with medical images. I work on chest X-ray images and build a model to classify those images by deep learning to identify whether pneumonia is present or not which takes a few moments to predict by my model but for doctors and clinics, it takes time to identify because this is very complex to identify. My theme was to reduce the timing problem of identifying this disease because this disease takes a very short time to impact on the patient a huge cause most of the patients of this disease is a child below 5 years old. If we cannot take action in the very primary stage, we have to pay worse for it. And by overcoming this problem, we can minimize the percentage of children's death from pneumonia in our country.

The article is structured as follows: The introduction to this study is included in Section I, Section II includes a description of related works, Section III contain the brief description of the dataset, The brief description of the proposed methodology given in Section IV, The essential description of evaluation metrics that we have used for the analysis of the performance of the proposed model is given in section V, Section VI describes results and discussion about this work and finally, the conclusion of this research work has been drawn.

II. Related Works

From the last decades, many researchers have been focusing on the identification of pneumonia disease based on x-ray images of the chest through various methods. The description of some existing method works is given below:

Kermany et al. [13] proposed a deep learning framework based on a transfer learning approach to diagnosing pneumonia disease from chest x-ray images. The experimental dataset consists of 5,232 images of chest x-ray and achieved 92.8% on test data.

Rajpurkar et al. [14] have built a prototype with Chex Net containing 121-layer of CNN. In this system, if an input chest X-ray image is given, the output is going to be the probability of pneumonia in addition to a heat map confining the areas of the image which highly indicate pneumonia. They have used the model Chex Net on ChestX-ray14 dataset containing 112,120 frontal-view chest X-ray images. These images are separately marked with up to 14 non-identical thoracic diseases, including pneumonia. They managed the optimization of this huge deep network using batch normalization and dense connections. There are two limitations in their model. The first one is, only frontal radiographs were submitted to the radiologists and model throughout diagnosis, but to get a proper result up to 15% of precise diagnoses require the lateral view. So, we know that this model can come up with a moderate estimate of performance. Secondly, radiologists and the model itself could not access patient history, as a result, the performance in interpreting chest radiographs has decreased. Their Chex Net achieves an F1 score of 0.435 (95% CI 0.387, 0.481).

Deep Thakkar et al. [15] demonstrated a strategy for diagnosing child pneumonia based on transfer learning. They used 5,000 frontal chest x-ray images. They applied MobileNet and InceptionV3 architecture and achieved 81.4% and 78.4% final test accuracy respectively. Enes AYAN et al. [16] proposed a deep learning method based on a transfer learning technique to diagnose pneumonia cases from chest x-ray images. They used VGG16 and Xception network and achieved 87% and 82% accuracy respectively on test data.

Benjamin Antin et al. [17] developed a pneumonia detection system from chest x-ray images using Supervised Learning. They used the NIH dataset, collected from Kaggle which contains 1,12,120 chest x-ray images of 30,805 distinct patients. Their dataset was labeled by different classes like Pneumonia, Fibrosis, etc., and Healthy (If the patient has no disease of these classes). They use binary classification to detect from X-ray images whether a patient has one of these diseases or not. After trained the entire dataset using 32*32 dimensions as input in their logistic regression, they achieved 0.60 scores on the AUC curve. Using 128*128 dimensions their AUC curve received a 0.58 score. The limitation was that model was unable to do a hyperparameter sweep on images larger than 32x32 and 224x224 where they were able to run logistic regression.

K.C. Santosh et al. [18] presented a concept that demonstrates the symmetry of the right and left lung regions and implemented the chest X-ray system for tuberculosis confirmation. The proposed methodology is based on the analysis of radiological tests that result in symmetrical observations in the lung region. Multilayer network perceptron, random forest, and Bayesian network are the three classifiers that are used here. The accuracy of abnormality detection was 91% and in the ROC curve, the area under was 0.96. They compare the changes between the left and right lung while we think it is not a good approach to identify tuberculosis like this. Both should be checked to identify the disease properly.

A. A. Saraiva et al. [19] proposed a model using Convolution Neural Network that describes a comparative classification of Normal and Pneumonia stages. They compared it to the model which has 92.8% accuracy and their model has 95.30% accuracy. They used a publicly available dataset from Kaggle which contains 5863 images with Pneumonia and Normal Classes. They use 300*300 image dimensions as input of their network and with each Convolutional layer, they used ReLU. After every two convolutional layers, they use Pool Layer (Max Pool). And in the last layer, they use the softmax activation function. Although their proposed model has 95 percent of accuracy it has some limitations. The pattern they followed in the CNN layer will face error from overfitting while a large number of datasets are applied in their network and also it is not very time and cost-effective.

III. DATASET

In this study, I have used a public dataset that consists of 5,856 frontal chest x-ray images [20]. All the images are in JPEG format. The dataset was developed by "University of California San Diego". The dataset consists of two classes of chest x-ray images which are with pneumonia and without pneumonia or normal. Some sample of normal and pneumonia images is shown in Figure 1. We have used 5,200 images from 5,856 images for the experiment. We split the new dataset into two parts, which are the training set and the test set. The training set consists of 3,330 pneumonia and 770 normal images and the test set consists of 560 pneumonia and 540 normal images. All values are mentioned in Table I below.

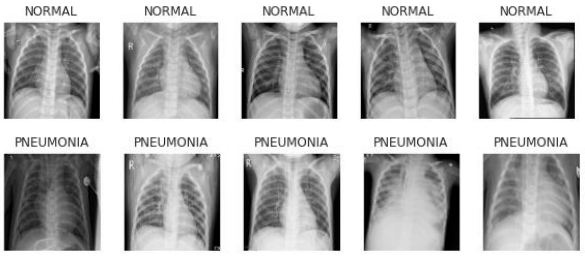

Figure 1. Sample of Normal and Pneumonia images

TABLE I. DATASET DISTRIBUTION

Class	Training set	Test set
Pneumonia	3,330	560
Normal	770	540
Total	4,100	1,100

IV. METHODOLOGY

The proposed methodological steps for the recognition of childhood pneumonia disease is shown in Figure 2.

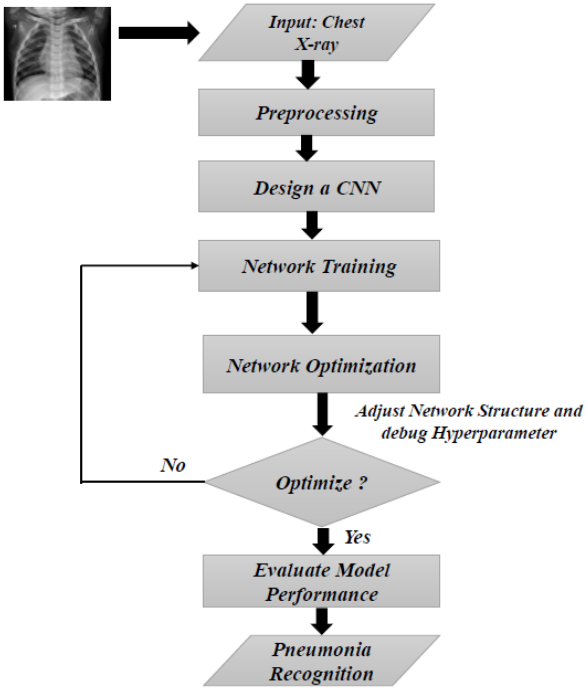

Fig. 2. Methodological steps for Pneumonia Recognition

A. Loading Dataset

I have started to work by loading the dataset and modifying it such that it could be used to properly forecast pneumonia using a convolutional neural network. I have already distinct training and testing set in the provided dataset.

B. Data Augmentation

Augmentation [21] describes the process of extending the size of the dataset used for training. As a result, we may have more variability in our dataset, and it is also utilized to expand the number of training samples in order to prevent overfitting. My data augmentation involves scaling, random shear range of 0.2, random zooming range of 0.2 and random horizontal flips.

C. Convolutional Neural Network Model Description

I have proposed a Convolutional Neural Network [22] architecture for recognizing children's pneumonia disease, which comprises four convolutional layers, four pooling layers, one dropout layer, and a fully connected layer. Figure 3 depicts the key layers of the proposed CNN model, which are then discussed below:

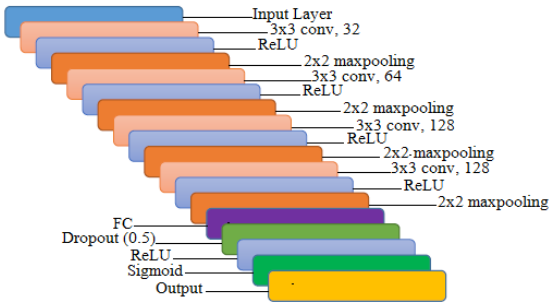

Figure 3. Proposed CNN Model

Input Layer: Through this layer, the model propagates information from the input layer to the hidden layer and then to the output layer.

Convolutional Layer: Our CNN model makes use of a mathematical procedure known as convolution. The fundamental goal of the convolution operation is to use a feature extraction technique to explore features in our images and place those in a feature vector [23]. This convolution operation is executed by using a two-dimensional matrix ($k_1 x k_2$) known as the kernel k, which scans all the data matrices available. Here let's take an input image I convoluted by k, the operation is expressed as:

$$(I * k)ij = \sum_{m=-\frac{k_1}{2}}^{\frac{k_1}{2}} \sum_{n=-\frac{k_2}{2}}^{\frac{k_2}{2}} I(i-m, j-n)\, k(m,n) \qquad (1)$$

In this study, the feature extractors are composed of cov3x3, 32; cov3x3, 64; cov3x3, 128; cov3x3, 128.The size of feature maps for the convolution operation are 148x148x32, 72x72x64, 34x34x128 and 15x15x128.

Activation Layer: The ReLU activation function [24] is applied in CNN models to enhance non-linearity. At the Convolution layers, the ReLU activation function was applied. If the value is greater than zero, the input is passed; else, it returns 0. The mathematical formulation of the ReLU activation function is as follows:

$$ReLU(x) = \max(0, x) \qquad (2)$$

where x is the neuron's input value.

Pooling Layer: The pooling layer helps to minimize overfitting by eliminating a few parameters while executing a pooling operation to obtain a pooled matrix. Downsampling is another term for pooling. Max pooling [25] is a technique for focusing just on the necessary information in an image by extracting mainly greater yields from just a part of the input by using an empty feature extraction technique and also decreases the dimensionality of the feature matrix layouts, which accelerates network time. In this study, the size of the feature maps for pooling operation is 74x74x32, 36x36x64, 17x17x128 and 7x7x128. The mathematical expression of maxpooling operation can be defined as:

$$N_{out} = floor\left(\frac{N_{in} - F}{S}\right) + 1 \qquad (3)$$

Where N_{in} = The size of the input image
F = The kernel size
and S = The stride size

Dropout Layer: The dropout layer [26] is utilized after convolution layers to prevent overfitting. In the proposed CNN, I have applied a single dropout layer with a dropping rate of50%, which indicates that 50% of random neurons are turned off to avoid overfitting of my model.

Fully Connected Layer: After layers of convolution and max pooling, I have obtained a set of feature cards. We connect them into one vector and this vector will be fed into the fully connected network. A fully connected layer [23] is utilized to give the result of classification. In this study, I carried out a flattening procedure and it'll be transformed into a 512x1 vector and sent to the dense layer along with some predefined parameters.

Output Layer: The final predicted class is achieved through this layer. The sigmoid function [27] is the core of the probabilistic technique, and it performs best when classifying between two classes. In order to binary classification, the sigmoid function is used and the sigmoid makes sure that the sum of the probability of the outcome is one. For these reasons, I have used is the sigmoid activation function for the proposed CNN. The linear input data is transformed into a probability array of all two classes via sigmoid, and the probability of occurrence is tested against the actual output. The sigmoid function is defined as follows:

$$f(x) = \frac{1}{1 + e^{-(x)}} \qquad (4)$$

where x is the inputs from the previous fully connected layers.

D. Network Compiling and Training

After successfully build the CNN model, it is ready to compile. For the calculation of the model loss, I have used the binary crossentropy loss function [28]. I used the Stochastic Gradient Descent method [29] based adaptive moment estimation (Adam) optimizer [30] as my optimization method to quickly and efficiently reach convergence of my model. The total number of trainable parameters that must be determined is 3,453,121 and there is no non-trainable parameters in the proposed CNN architecture.

After successfully compile the CNN model, the model is trained. During model training, hyperparameters [31] such as learning rate, batch size, and the number of epochs play an important role. I have set the value learning rate as 0.001, the batch size

is 50 and the number of epochs is up to 100. The training dataset was used to train our model, and the test dataset was utilized to validate my model. I performed my experiment on compiler Jupiter Notebook integrates with some numbers of packages with an "Intel® Core™ i7-10510U (1.8 GHz, up to 4.9 GHz, 8 MB cache, 4 cores) powered with NVIDIA GeForce MX330 (2GB), 16GB DDR4 RAM" and Windows 10 based 64-bit operating system. I have used Google Colabratory for training and testing data. After that, I examined my proposed model is based on predictions made by my trained model on the test dataset and obtained its accuracy, precision, recall, and F1-score to evaluate the proposed model's performance.

V. Performance Metrics

To estimate the performance of the proposed model, four different measurements which are accuracy, precision, recall, and F1 score are utilized to measure the performance of the model. The confusion matrix can calculate these four measurements. Definition of the four measurements and confusion matrix is illustrated below:

A. Confusion Matrix

The confusion matrix [32] is utilized to illustrate the performance for a machine learning classification problem whose output can be two or more classes. Pneumonia classification is a multiple classification problem that is suitable to utilize confusion matrix to measure performance. The confusion matrix includes four significant parts which are true positive (TP), true negative (TN), false negative (FN), and false positive (FP). This matrix can be utilized to measure accuracy, recall, precision, and F1 score.

B. Precision

Precision [32] is proposed to illustrate the percentage of images which are actually positive, out of the positive classes which are predicted correctly. It can be defined as:

$$Recall = \frac{TP}{TP + FN} \qquad (5)$$

C. Recall

The recall [32] is utilized to illustrate the percentage of images which are predicted correctly, out of the positive classes. It can be defined as:

$$Precision = \frac{TP}{TP + FP} \qquad (6)$$

D. Accuracy

Accuracy [32] is a common measure in machine learning and is measured by the ratio of correctly classified samples to the total number of samples in the dataset. It can be defined as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad (7)$$

E. F1 score

The F1-score [32] (sometimes known as the balanced F-beta score) is a single metric that combines both precisions and recalls via their harmonic mean. It can be defined as:

$$F1 = \frac{2 * Recall * Precision}{Recall + Precision} \qquad (8)$$

F. ROC Curve

Another common metric for machine learning classifiers is Receiver Operating Characteristic (ROC) curves [33]. The ROC curve depicts a model's performance across all categorization levels. The curve represents the two parameters TPR (y-axis) and FPR (x-Axis) at different threshold values. ROC curves can thus be used to analyze the threshold-dependent trade-off between TPR and FPR. Both TAR and FAR lies in the interval of 0.0 to 1.0. A perfect classification system would achieve a TAR of 1.0 and a FAR of 0.0 at a given threshold.

VI. Results and Discussion

I conducted experiments on my proposed model 10 times to validate and evaluate its effectiveness. Each process took approximately 4 hours. We repeatedly changed hyperparameters and parameters for the best possible result. I kept trying until I was satisfied with my result. This report consists of the most valid output I received.

I have developed my model and tested it on a small dataset of 5,200 images. The whole dataset was divided into two classes pneumonia and normal. I have used a deep convolution neural network to obtain the output shown in Figure 4 and Figure 5. Training loss = 0.0556, training accuracy = 0.9802, validation loss = 0.034, and validation accuracy = 0.9887 are the final results.

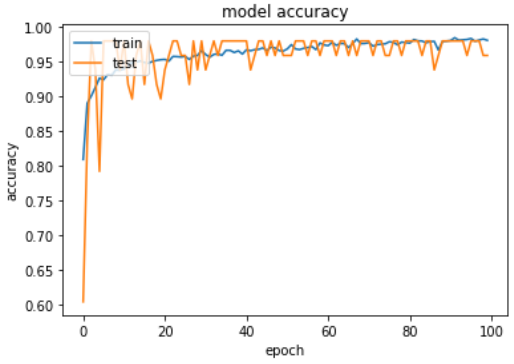

Fig. 4. Model Accuracy on 150x150x3 image size

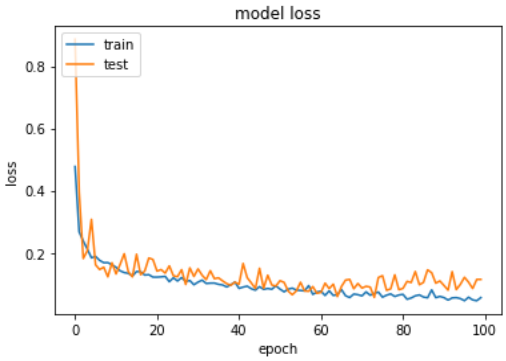

Figure 5. Model Loss on 150x150x3 image size

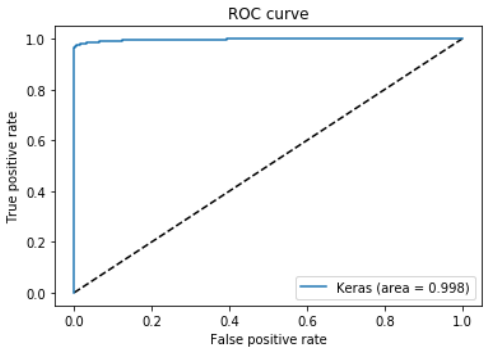

Figure 6: ROC curve on 150x150x3 image size

Here I used different sizes of images to test our model as the CNN framework always requires images with fixed sizes. I reshaped the x-ray images in 100x100x3, 150x150x3, 200x200x3, 250x250x3 and 300x300x3 sizes. Each took almost 4 hours to train. I acquired the confusion matrix and ROC curve for all of the picture sizes listed above in order to evaluate and validate the performance of the proposed architecture. The performance of all sizes images is shown in Table II and Figure 7-11. The classification report and confusion matrix are shown in Table III and Figure 12 respectively.

From the repeated experiment I noticed that the larger the image size is the lesser the validation accuracy was obtained. And finally, I got the best result when image size 150x150x3 was used. But other changes were very near. No significant change was noticed during multiple experiments. With 150x150x3 we obtained 98.87 percent validation accuracy with a minimal loss of 0.034.

To detect pneumonia, I built a CNN architecture using chest x-ray images obtained from frontal views.

TABLE II. MODEL PERFORMANCE ON DIFFERENT IMAGE SIZES

Image Size	Training Accuracy	Testing Accuracy
100	0.9701	0.9754
150	0.9802	0.9887
200	0.9687	0.9703
250	0.9543	0.9634
300	0.9502	0.9630
Average	0.9647	0.9721

TABLE III. CLASSIFICATION REPORT

	Precision	Recall	F1-score	Support
Pneumonia	0.97	1.00	0.98	560
Normal	1.00	0.97	0.98	540
Accuracy			0.98	1100
Macro avg	0.98	0.98	0.98	1100
Weighted avg	0.98	0.98	0.98	1100

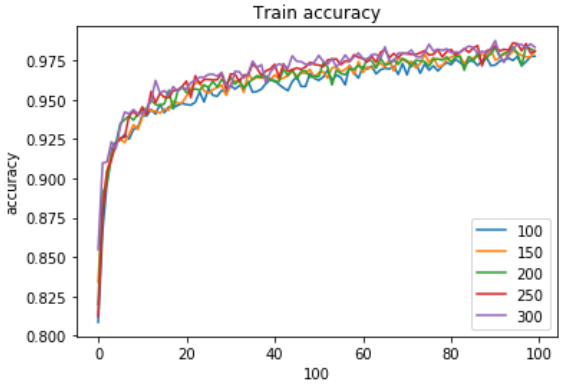

Fig. 7. Training Accuracy on all mentioned data sizes

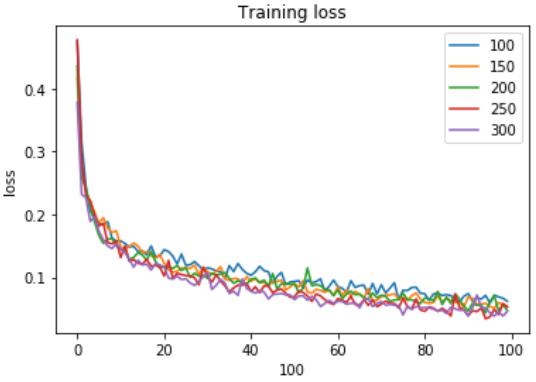

Fig. 8. Training Loss on all mentioned data sizes

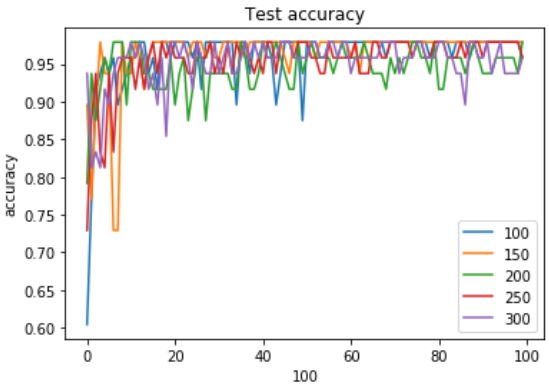

Fig. 9. Test Accuracy on all mentioned data sizes

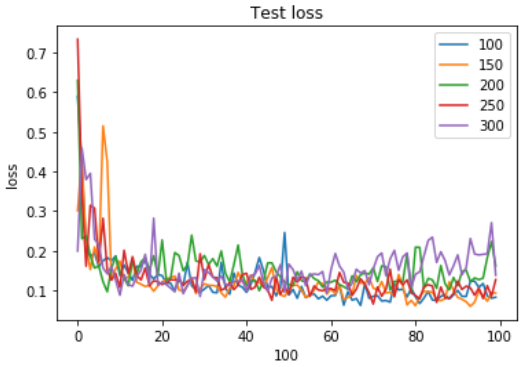

Fig. 10. Test Loss on all mentioned data sizes

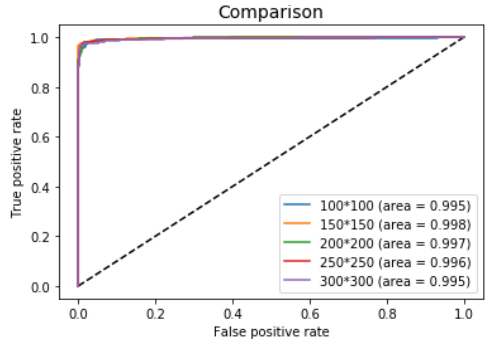

Fig. 11. ROC Curve on all mentioned data sizes

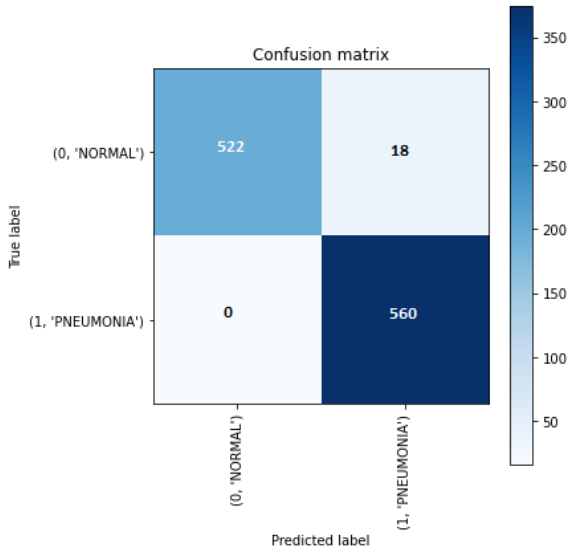

Fig. 12. Confusion Matrix

The design begins with transforming X-ray images into smaller sizes. Then, to identify and diagnose pneumonia, I employed a convolution neural network architecture that extracted features from images. After a couple of experiments, I obtained my expected output which was higher than the other approaches. I repeated my experiment several times to check the accuracy and every time I received the same result. For validating the performance I trained the model on different sizes of x-ray images and obtained similar results.

The comparative study of the proposed CNN method with some state-of-the-art methods is given in table IV.

TABLE IV. COMPARISONS WITH PREVIOUS WORKS

Authors	Methods	Accuracy (%)
Deep Thakkar *et al.* [15]	Transfer Learning	81.4%
Kermany *et al.* [13]	Transfer Learning	92.8%
K.C. Santosh *et al.* [18]	Random Forest	91%
A. A. Saraiva *et al.* [19]	CNN	95.30%
Proposed	CNN	98.87%

Table IV has shown that the proposed CNN model gives higher accuracy than the others.

I anticipate that if I have large amounts of data and can train the model using radiological data from pneumonia patients and normal people from various locations, I will be able to make significant upgrades to this model.

ACKNOWLEDGMENT

I would want to express my heartfelt gratitude to all of the faculty members of the Computer Science and Engineering, Islamic University-Bangladesh, and most importantly my beloved parents for their continuous support.

CONCLUSION

In this paper, I presented a model for separating positive and negative pneumonia data from a set of x-ray images. My model differs from others that depend mostly on the transfer learning approach. The proposed model may also be used to recognize and categorize x-ray images of lung cancer and pneumonia. I hope this model will be very useful in rural areas where expert doctors are not available and people are deprived of medical help. My future plan is to collect a larger dataset consisting of chest X-rays of patients and normal people from different parts of our country and test my model on it. As nowadays people are more comfortable with mobile devices, I am planning to build a mobile application that will be easily accessible to them. My CNN model can be used for the mobile application-based pneumonia recognition system.

REFERENCES

[1] *Pneumonia, the forgotten killer of children. Geneva: UNICEF/ WHO; 2006. Available at: http://whqlibdoc.who.int/ publications/2006/ 9280640489_eng.pdf*

[2] [Online Access] https://www.icddrb.org/news-and-events/press-corner/media-resources/pneumonia-and-other-respiratory-diseases

[3] [Online Access] https://www.dhakatribune.com/bangladesh/2019/11/13/report-pneumonia-kills-4-children-every-3-hours-in-bangladesh

[4] [Online Access] https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health

[5] Ferdous, Farzana & Ahmed, Shahnawaz & Das, Sumon & Chisti, Mohammod & Nasrin, Dilruba & Kotloff, Karen & Levine, Myron & Nataro, James & Ma, Enbo & Muhsen, Khitam & Wagatsuma, Yukiko & Ahmed, Tahmeed & Faruque, Abu. (2018). Pneumonia mortality and healthcare utilization in young children in rural Bangladesh: A

prospective verbal autopsy study. Tropical Medicine and Health. 46. 10.1186/s41182-018-0099-4.

[6] [Online Available] https://www.kaggle.com/c/rsna-pneumonia-detection-challenge

[7] Huang Yi, Sun Shiyu, Duan Xiusheng and Chen Zhigang, "A study on Deep Neural Networks framework," *2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, 2016, pp. 1519-1522, doi: 10.1109/IMCEC.2016.7867471.

[8] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng and M. Chen, "Medical image classification with convolutional neural network," *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*, 2014, pp. 844-848, doi: 10.1109/ICARCV.2014.7064414.

[9] E. Andre, K. Brett, A Roberto et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[10] M. Grewal, M. M. Srivastava, P. Kumar and S. Varadarajan, "RADnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 281-284, doi: 10.1109/ISBI.2018.8363574.

[11] Rajpurkar, P., Hannun, A.Y., Haghpanahi, M., Bourn, C., & Ng, A. (2017). Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. *ArXiv, abs/1707.01836*.

[12] Gulshan, Varun & Peng, Lily & Coram, Marc & Stumpe, Martin & Wu, Derek & Narayanaswamy, Arunachalam & Venugopalan, Subhashini & Widner, Kasumi & Madams, Tom & Cuadros, Jorge & Kim, Ramasamy & Raman, Rajiv & Nelson, Philip & Mega, Jessica & Webster, Dale. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA. 316. 10.1001/jama.2016.17216.

[13] Kermany, Daniel & Goldbaum, Michael & Cai, Wenjia & Valentim, Carolina & Liang, Hui-Ying & Baxter, Sally & McKeown, Alex & Yang, Ge & Wu, Xiaokang & Yan, Fangbing & Dong, Justin & Prasadha, Made & Pei, Jacqueline & Ting, Magdalena & Zhu, Jie & Li, Christina & Hewett, Sierra & Dong, Jason & Ziyar, Ian & Zhang, Kang. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell. 172. 1122-1131.e9. 10.1016/j.cell.2018.02.010.

[14] Rajpurkar, Pranav & Irvin, Jeremy & Zhu, Kaylie & Yang, Brandon & Mehta, Hershel & Duan, Tony & Ding, Daisy & Bagul, Aarti & Langlotz, Curtis & Shpanskaya, Katie & Lungren, Matthew & Ng, Andrew. (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.

[15] Thakker, Deep & Shah, Vidit & Rele, Jash & Shah, Vivek & Khanapuri, Jayashree. (2019). Diagnosing Child Pneumonia using Transfer Learning.

[16] E. Ayan and H. M. Ünver, "Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning," *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, 2019, pp. 1-5, doi: 10.1109/EBBT.2019.8741582.

[17] Antin, B., Kravitz, J., & Martayan, E. (2017). Detecting Pneumonia in Chest X-Rays with Supervised Learning.

[18] K. C. Santosh and S. Antani, "Automated Chest X-Ray Screening: Can Lung Region Symmetry Help Detect Pulmonary Abnormalities?," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 5, pp. 1168-1177, May 2018, doi: 10.1109/TMI.2017.2775636.

[19] Saraiva, Arata & Ferreira, N. & Sousa, Luciano & Carvalho da Costa, Nator & Sousa, José & Santos, D. & Valente, Antonio & Soares, Salviano. (2019). Classification of Images of Childhood Pneumonia using Convolutional Neural Networks. 112-119. 10.5220/0007404301120119.

[20] Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification", Mendeley Data, V2, doi: 10.17632/rscbjbr9sj.2

[21] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," *2018 International Interdisciplinary PhD Workshop (IIPhDW)*, 2018, pp. 117-122, doi: 10.1109/IIPHDW.2018.8388338.

[22] R. Chauhan, K. K. Ghanshala and R. C. Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition," *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, 2018, pp. 278-282, doi: 10.1109/ICSCCC.2018.8703316.

[23] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

[24] Agarap, A.F. (2018). Deep Learning using Rectified Linear Units (ReLU). *ArXiv, abs/1803.08375*.

[25] J. Nagi et al., "Max-pooling convolutional neural networks for vision-based hand gesture recognition," 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 2011, pp. 342-347, doi: 10.1109/ICSIPA.2011.6144164.

[26] Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting.. Journal of Machine Learning Research, 15, 1929-1958.

[27] K. Hara and K. Nakayamma, "Comparison of activation functions in multilayer neural network for pattern classification," *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94)*, 1994, pp. 2997-3002 vol.5, doi: 10.1109/ICNN.1994.374710.

[28] Mannor, Shie & Peleg, Dori & Rubinstein, Reuven. (2005). The cross entropy method for classification. 561-568. 10.1145/1102351.1102422

[29] G. Cong, O. Bhardwaj and M. Feng, "An Efficient, Distributed Stochastic Gradient Descent Algorithm for Deep-Learning Applications," *2017 46th International Conference on Parallel Processing (ICPP)*, 2017, pp. 11-20, doi: 10.1109/ICPP.2017.10.

[30] Kingma, D. P. & Ba, J. (2014). Adam: A Method for Stochastic Optimization.. CoRR, abs/1412.6980.

[31] Z. Li, L. Jin, C. Yang and Z. Zhong, "Hyperparameter search for deep convolutional neural network using effect factors," *2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)*, 2015, pp. 782-786, doi: 10.1109/ChinaSIP.2015.7230511.

[32] [Online Access] https://medium.com/analytics-vidhya/confusion-matrix-accuracy-precision-recall-f1-score-ade299cf63cd.

[33] E. A. P. A., S. K. Sugathan and A. T. S. Ho, "Receiver Operating Characteristic (ROC) Graph to Determine the Most Suitable Pairs Analysis Threshold Value," *Advances in Electrical and Electronics Engineering - IAENG Special Edition of the World Congress on Engineering and Computer Science 2008*, 2008, pp. 224-230, doi: 10.1109/WCECS.2008.35.