A Combined Approach Using Image Processing and Deep Learning to Detect Pneumonia from Chest X-Ray Image

Md. Mehedi Hasan[1], Mir Md. Jahangir Kabir[2], Md. Rakibul Haque[3], Mohiuddin Ahmed[4]
Dept. of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Email: mmehedihasann@gmail.com[1], mmjahangir.kabir@gmail.com[2], rakibulhaq56@gmail.com[3], mohiuddin.nirob.mn@gmail.com[4]
Rajshahi, Bangladesh

Abstract— **Pneumonia is an epidemic disease that is needed to be detected in the early stage to prevent unfortunate deaths. Traditional methods take a lot of time to detect the disease. With the introduction of Medical Imaging, the detection of disease has been accelerated by using chest x-ray image. But it also requires the availability of an expert and experienced radiologist in order to interpret the x-ray image accurately. Sometimes, manual interpretation is affected by various kinds of artifact in X-ray or Optical Coherence Tomography(OCT) images. For this reason in our paper, we have proposed a combined approach using Image Processing and either VGG-16 or VGG-19, variants of Deep Convolutional Neural Network for automatic detection of pneumonia from Chest X-ray image. We have used Mendeley OCT and Chest X-Ray dataset to evaluate our model. We have achieved an accuracy of 96.2% using VGG-16 and 95.9% using VGG-19, both of which outperform transfer learned InceptionV3 benchmark model used on this dataset which has an accuracy of 92.8%.**

Keywords — Pneumonia, Chest X-ray Image, Image Processing, VGG-16, VGG-19, Transfer-Learned InceptionV3.

I. INTRODUCTION

Pneumonia is an infection which creates inflammation in one or both of the lungs. The infection can be caused by bacteria, virus, fungi or any other germs [1]. Pneumonia is one of the main infectious diseases responsible for the high child mortality rate. It kills more than 2 million children under the age of five each year [2]. Apart from the children, more than 1 million adults are hospitalized and 50,000 among them dies every year from pneumonia in the US alone [3]. For this reason, detecting pneumonia in the early stage and providing accurate treatment is a crucial task in the present era.

Traditional methods consume a huge time in predicting the disease. With the introduction of medical imaging, we can detect a larger number of diseases faster and more accurately from the Chest X-ray image, OCT, CT-Scan, MRI Scan, etc. [4]. Chest X-ray is currently the best method available for providing proper diagnosis and detection of pneumonia [5]. However, for providing an accurate interpretation of medical images is a difficult task to perform. For performing the interpretation of X-ray images we need expert radiologist and also the background information of the patient. Also then the interpretation might be wrong because of various kinds of artifacts in the image.

In order to extract information from the X-ray image first, we need to preprocess the image in order to reduce the unnecessary data. Sharma and Raju [6] has used histogram equalization and thresholding in order to detect Pneumonia cloud in Chest X-ray images. In order to find the lungs as largest contour, we need to preprocess the image and sometimes also crop the existing image by a certain percentage.

Convolution Neural Network (CNN) is one of the Deep learning architectures that has significantly used in medical imaging [7] due to its significant performance in classifying image form ImageNet [8]. Convolutional Neural Network uses a wide range of convolution kernel which makes it easy to detect and classify different kind of diseases based on the image. Rajpurkar [9] used a variant of 121 layers DenseNet called ChexNet on Chest X-Ray 14 dataset in order to pneumonia better than radiologist. Kermany et al.[10] used a transfer learned InceptionV3 model to detect pneumonia from Chest X-ray image in Mendeley dataset. The used AlexNet was trained on 1000 categories of the image from ImageNet dataset [8]. However, they didn't perform any kind of image pre-processing on Chest X-ray. For that reason, some of the unnecessary information degraded their performance.

In this paper, we have first performed contrast limited adaptive histogram equalization on Chest X-ray image from Mendeley dataset, then we have vertically cropped 6% from both the sides of the image. After that, we have found the two largest contours and their leftmost and rightmost points. We have then performed vertical cropping along with those point in order to reduce irrelevant data. Finally, we have fed the preprocessed image into VGG-16 and VGG-19 architectures, one of the variants of CNN proposed for ImageNet challenge. We have achieved a better result in predicting pneumonia from chest x-ray image than other state-of-the-art methods.

The remainder of the paper is organized as follows. Section II reviews the Image preprocessing and the architecture we have used for the detection of pneumonia. In section III we have discussed the performance of our architecture in details. Finally, we have drawn some conclusion based on the performance of our architecture.

II. METHODOLOGY

The approach shown in figure 1 for detecting pneumonia from Chest X-ray image can be broken down into two broad parts. One is Image processing for converting the raw Chest X-ray image (CXR) and the second one is to train the VGG-16 and VGG-19 network in order to detect pneumonia from the preprocessed image.

A. Dataset Description

We have used the publicly available Mendeley's Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification dataset [11]. The dataset contains 5,856 Chest X-ray Images of children among which 4,273 images are characterized by pneumonia infected images and 1,583 are characterized as normal images. We have then split the dataset by 80:20 training and testing ratio. After that training data have been also split into 80:20 training and validation ratio.

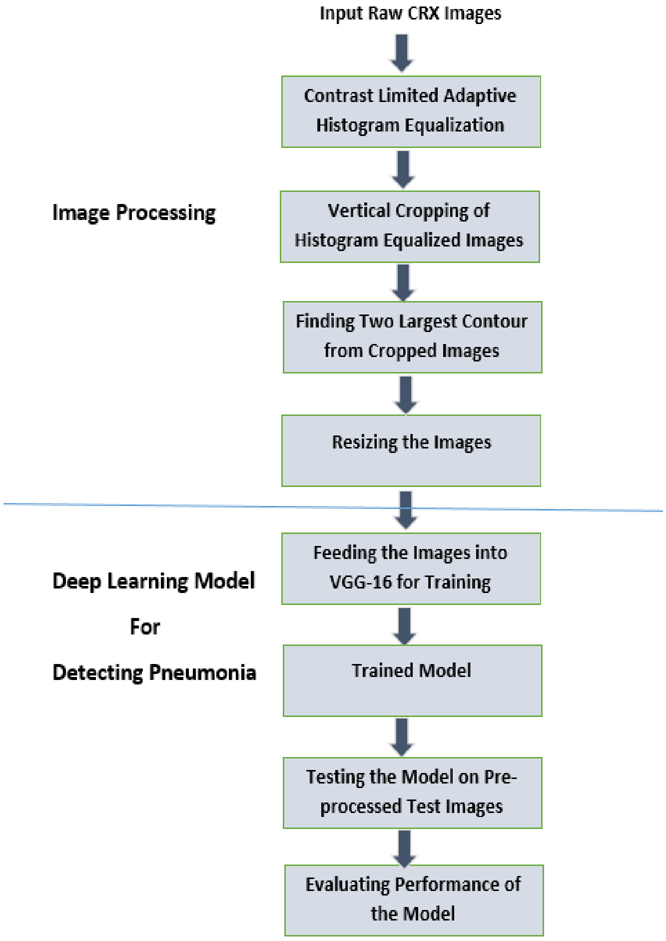


Fig. 1. An overview of the proposed approach

B. Image Processing

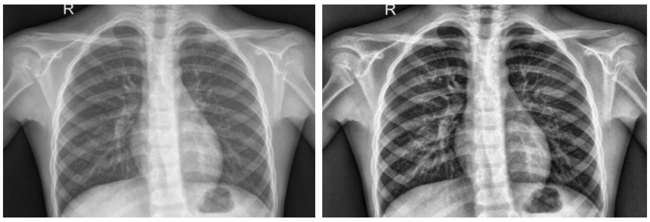
The image processing was mainly performed in order to crop the lung nodule area more accurately as lung nodule area is a haze in the chest x-ray image. The image processing task can be further divided into four major part. The major parts of the image processing are elaborated below.

1) Contrast Limited Adaptive Histogram Equalization (CLAHE)

Histogram equalization is a technique used adjusting image intensity for enhancing the contrast of the image. An input image, r is converted to an image s using the following transformation function T [12].

$$S=T(r) = (L-1)\sum_{j=0}^{k} p(r_j) \tag{1}$$

Here L is the number of levels used in the image and $p(r_j)$ is the probability of j'th pixel in input image r. This transformation provides a uniform distribution in the output image. However normal histogram equalization sometimes results in washing out the information on white pixels because the contrast gets much higher in that region. For this reason, we have used contrast limited adaptive histogram equalization. Contrast limited adaptive histogram equalization [13] solves the problem using 8×8 tiles and applying histogram equalization to all and by setting the maximum contrast limit.



(a) (b)

Fig. 2. (a) Raw chest X-ray Image (b) Image after performing CLAHE.

From figure 2 we can clearly see that after applying CLAHE, we got more information in the image about the lung nodule area.

2) Vertical Cropping of Histogram Equalized Image

After performing contrast limited histogram equalization we have vertically cropped the equalized image by a margin of 6% on both sides as shown in figure 3. The margin was a tradeoff between detecting false contour as lung module and discarding information. We kept the margin as low as possible in order to avoid any loss of information and preventing from detecting the false contours

Fig. 3. Performing 6% vertical Cropping on both sides.

3) Finding Two Largest Contour from Cropped Image

Contour can be explained as a simple curve joining all the continuous points in the boundary [14]. Generally, in the CXR images, the lung nodules are the largest two contour. The goal of this step is to find the two lung nodule. But, in 155 images among the 5856 images, it became difficult to find the two nodular areas as largest contours. In these cases, the images were replaced by 6% cropped images manually. Besides these cases, the lung modules were successfully detected as the largest contour automatically by the combined approach. After detecting the largest two contours, we found the leftmost and rightmost points of the contour in order to perform vertical cropping as we only need the lung nodules in order to detect pneumonia cloud for pneumonia detection. The largest two contour of a random chest x-ray image is shown in figure 4.

Fig. 4. The Largest Two Contour and cropped along the vertical axis of the rightmost and leftmost points of contours.

C. VGG-16

VGG-16 is one of the variants of the convolutional neural network. Simonyan and Zisserman [15] have first introduced VGG-16 for ImageNet Large Scale Visual Recognition challenge [8]. VGG-16 has 16 weighted layers [15]. It has been trained on the ImageNet database and it has performed better than LeNet-5 [16] and AlexNet [17].

VGG-16 has a total of 13 convolution layers, 5 pooling layers and 3 fully connected layers stacked together. In the convolution layer, it always uses a 3×3 filter with a stride of 1 and it uses the same padding in pooling layers 2× 2 with a stride of 2. Relu [18] is mainly used in VGG-16 as an activation function for hidden layers [15].

The original VGG-16 was trained with the input of size $224 \times 224 \times 3$ [14]. We have modified the VGG-16 architecture so that it can be trained with input size $128 \times 128 \times 1$. This is mainly done because in the image processing step we have resized all our variable-sized image into a fixed size of $128 \times 128 \times 1$. As Chest X-ray image is a gray-scale image, so it has only one channel compared to the 3 channel in color image of ImageNet [8] dataset. Originally VGG-16 has 1000 nodes on the output but we have kept only two output node with softmax activation function as our problem is binary classification.

D. VGG-19

VGG-19 is an extension of VGG-16 with some minimum changes made. It was created by the same group that created VGG-16 to improve the ImageNet challenge error rate. It has 19 layers altogether. 64 filter convolutional layer x 2 + 128 filter convolutional layer x 2, 256 filter convolutional layer x 4, 512 filter convolutional layer x 8, fully connected layer x 2 and 1 output layer. Originally it had 1000 output nodes for ImageNet challenge, but we have changed the output nodes to 2 for fulfilling our purpose

III. Experimental Analysis

A. Design of Experiment

We have implemented our network in Keras and trained our model on Kaggle and online computing service. We have trained the model for 100 epoch with a batch size of 256. We stopped the training after 100 epoch as there was no decrease in validation loss. We have used Adam optimizer in order to minimize our loss function. The loss function used for this problem was categorical cross-entropy. We have also used 50% dropout in order to avoid overfitting. Figure 5 shows the accuracy and loss curves of the training.

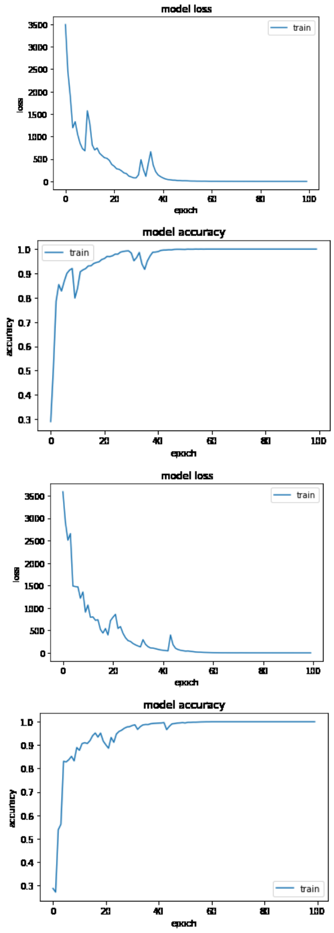


Fig. 5. Training and loss curve of VGG-16 and VGG-19. First two graph corresponds to VGG-16 and last two graph corresponds to VGG-19

B. Result Analysis

We have used five accuracy metric based on the confusion matrix in order to evaluate our result. They are precision, specificity, sensitivity, F1-Score, and accuracy.

TABLE I. Evaluation of our proposed approaches on Mendeley dataset for pneumonia detection.

Accuracy Metric	VGG-16	VGG-19
Precision	0.977	0.971
Sensitivity	0.970	0.974
Specificity	0.939	0.919
F1-score	0.973	0.972
Accuracy	0.962	0.959

From the above table, we can see that we have achieved an accuracy of 96.2% using VGG-16 and 95.9% using VGG-19 which is better than the current state of the art method such as transfer learned InceptionV3 for this dataset. A detailed comparison in figure 6 between our proposed approach and current state of the art method.

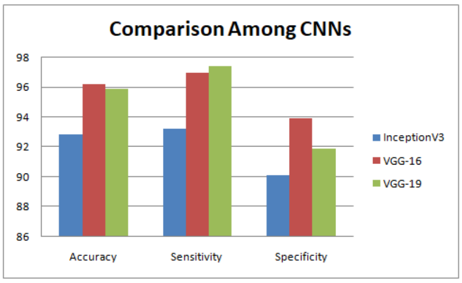


Fig. 6. Comparison between our proposed approaches and inception v3.

We have also used ROC curve in order to measure the performance of our proposed approaches which are shown in figure 7.

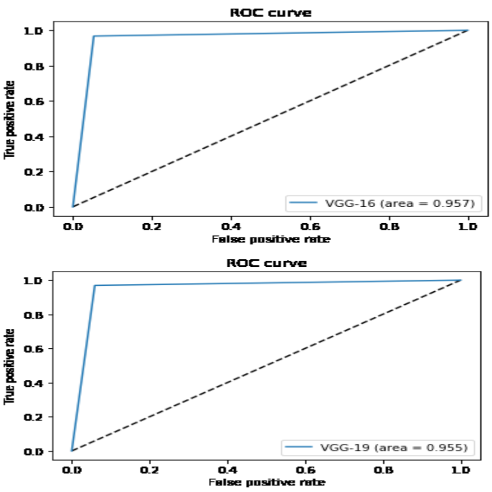


Fig. 7. ROC curve for VGG-16 and VGG-19.

We have found that a deep convolutional network performs better in pneumonia detection if we first pre-process the CXR images in order to bring out more information.

IV. Conclusion

Pneumonia is a potential threat in the present era due to the increase in air pollution. In order to provide proper treatment to this disease, it needs to be detected in the early stage. Chest X-Ray images provide a way to detect pneumonia fast. But, due to the shortage of radiologists, an automatic detection system that can predict pneumonia precisely is needed. In this paper, we have used VGG-16 and VGG-19 network followed by a pervasive image processing to detect pneumonia which has performed 3.4% and 3.1% respectively better than transfer learned InceptionV3. In the future, we will like to try other architectures containing convolution network in order to build a more accurate pneumonia detection system

References

[1]. White, R. J., Blainey, A. D., Harrison, K. J., & Clarke, S. K. , "Causes of pneumonia presenting to a district general hospital." *Thorax* 36.8 (1981): 566-570.

[2]. Wardlaw, T., Salama, P., Johansson, E. W., & Mason, E., "Pneumonia: the leading killer of children." *The Lancet* 368.9541 (2006): 1048-1050

[3]. CDC, https://www.cdc.gov/features/ pneumonia/index.html, Last Accessed – 13th March ,2019.

[4]. Leutner, C. C., Gieseke, J., Lutterbey, G., Kuhl, C. K., Glasmacher, A., Wardelmann, E., Theisen,A. & Schild, H. H., "MR imaging of pneumonia in immunocompromised patients: comparison with helical CT." *American Journal of Roentgenology* 175.2 (2000): 391-397.

[5]. World Health Organization. *Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children*. No. WHO/V&B/01.35. Geneva: World Health Organization, 2001.

[6]. Sharma, A., Raju, D., & Ranjan, S., Detection of pneumonia clouds in chest X-ray using image processing approach." *2017 Nirma University International Conference on Engineering (NUiCONE)*. IEEE, 2017.

[7]. Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J., "Convolutional neural networks for medical image analysis: Full training or fine tuning?." *IEEE transactions on medical imaging* 35.5 (2016): 1299-1312.

[8]. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L., "Imagenet: A large-scale hierarchical image database." (2009): 248-255.

[9]. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K. and Lungren, M.P., "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." *arXiv preprint arXiv:1711.05225* (2017).

[10]. Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C.C., Liang, H., Baxter, S.L., McKeown, A., Yang, G., Wu, X., Yan, F. and Dong, J., "Identifying medical diagnoses and treatable diseases by image-based deep learning." *Cell* 172.5 (2018): 1122-1131.

[11]. Kermany, D. Zhang, K., Goldbaum, M., "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification", MendeleyData ,v2 ,2018.

[12]. Gonzalez, R. C., "Digital Image Processing/Richard E." Woods. Interscience, NY (2017), ch. 3, pp. 126-127.

[13]. Reza, A. M. ,"Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement." *Journal of VLSI signal processing systems for signal, image and video technology* 38.1 (2004): 35-44.

[14]. Arbelaez, P., Maire, M., Fowlkes, C., & Malik, J. "Contour detection and hierarchical image segmentation." *IEEE transactions on pattern analysis and machine intelligence* 33.5 (2011): 898-916.

[15]. Simonyan, K., & Zisserman, A., "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).

[16]. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P., "Gradient-based learning applied to document recognition." *Proceedings of the IEEE* 86.11 (1998): 2278-2324.

[17]. Krizhevsky, A., Sutskever, I., & Hinton, G. E., "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.

[18]. Yarotsky, D., "Error bounds for approximations with deep ReLU networks." *Neural Networks* 94 (2017): 103-114.