Using Deep Learning and Transfer Learning for Pneumonia Detection

Rajasri Dhonikena[1], Miryala Hasini[2], Devavarapu Sreenivasarao[3], Shaik Nazma[4]

[1,2,4]B. Tech Scholars, Department of Computer Science and Engineering, Sreenidhi Institute of Science and Technology Hyderabad, Telangana, India

[3]Assistant Professor, Department of Computer Science and Engineering, Sreenidhi Institute of Science and Technology Hyderabad, Telangana, India

---

Abstract - *The biggest cause of death in children under five years old worldwide is pneumonia. Radiologists with the necessary training examine chest X-rays to identify pneumonia. But it is time-consuming and exhausting. Medical image analysis has a lot of potential applications for biomedical image diagnostics techniques. This study suggests a model for diagnosing pneumonia that is trained on chest X-ray pictures. Finding a large enough dataset for detection tasks was extremely difficult, hence data augmentation techniques were utilized to expand the training dataset. The models are additionally trained via transfer learning. The proposed approach might help radiologists' clinical judgment and assist in the early diagnosis of illnesses. The model was statistically validated after being checked for overfitting and generalization errors. To evaluate the effectiveness of the suggested model, various metrics including testing accuracy, F1, recall, precision, and AUC score were constructed.*

Key Words*:* *Pneumonia; Chest X-ray images; Convolution Neural Network (CNN); ResNet50; Transfer Learning*

1. INTRODUCTION

Elderly and young children are more likely than others to die from pneumonia. Since few people are aware of it politically and socially, it is known as "the silent killer" in comparison to other diseases. Children's pneumonia is an illness with roots in poverty. Its primary causes include a lack of access to basic healthcare, inadequate education, and inadequate childcare. Adult pneumonia may be viewed as a public health problem that needs more proactive treatment. Two major factors driving academics to increase learning rates are the availability of vast datasets and the development of more potent GPUs. For a variety of medical imaging tasks, such as the diagnosis of pneumonia, the detection of arrhythmias, the classification of skin cancer, the detection of diabetic retinopathy, and the identification of bleeding, resource-intensive models are now feasible with performance superior to that of human experts. Deep learning methods akin to CNN are hence the researchers' chosen technology for disease categorization and diagnosis from medical imaging.

2. SCOPE

For example, chest X-rays, chest MRIs, chest ultrasounds, lung needle biopsies, computed tomography of the lungs, and chest X-rays are all techniques that can be used to diagnose pneumonia. X-rays are the most widely used diagnostic imaging tool. Radiotherapists find it difficult to examine chest X-rays. Pneumonia on the patient's X-ray might occasionally be challenging to diagnose. Disease prediction becomes tough since it is difficult to identify the traits that identify the presence of the disease. The X-ray pictures in the dataset were wrongly categorized mostly due to this. It has been demonstrated in the past that several CAD systems are beneficial in the medical field, particularly in the identification of lung nodules and breast cancer.

The most effective and popular Machine Learning (ML) method for disease diagnosis in general and radiography, in particular, continues to be deep learning. The accuracy of disease prediction using deep learning algorithms has already been demonstrated to be on par with that of a typical radiologist. At this time, trained physicians cannot be replaced by deep learning-based algorithms in medical evaluation. Therefore, deep learning-based computer-aided diagnosis techniques can be employed as an addition to clinical decision-making.

3. RELATED WORK

Several authors have already presented various biomedical image detection methods. Authors in [1] talked about diagnosing pneumonia. The difficulties with medical imaging technologies were discussed by Razaak et al. [2]. Various approaches were put forth by numerous writers for the identification of numerous diseases [3]. For instance, Andre presented a deep CNN-based Inception v3 architecture in his paper for the classification of skin cancer [4], Milletari also worked on a method for CNN-based prostate detection in MRI volumes [5], Grewal applied deep learning to CT scans to identify brain hemorrhages [6], and Varun developed a method for the detection of diabetic retinopathy [7]. In comparison to DNN, CNN is a lot superior breakthrough since it can easily work with 2-D and 3-D images and extract the features needed to categorize the disease. Because the max-pooling layer in CNN is so effective and because it is coupled with some weights, this is conceivable. As they use gradient-based learning while they are being trained, CNNs also deal with the serious issue of the declining gradient. CNN's identify different chest X-rays and accurately classify them using the information retrieved by the various layers. Some algorithms were devised earlier in the papers of Shin for data mining [10,11] and the paper of Boussaid also proposed that labels are predicted by the extraction of features and application of segmentation techniques from radiology images of chest X-rays [12]. The studies by Shin [10,11] and Boussaid [12] both provided certain algorithms for data mining, and Boussaid's work also suggested that labels are predicted by the extraction of features and application of segmentation techniques from radiology pictures of chest X-rays.

4. DATASET

The initial dataset from the Guangzhou Women and Children's Medical Center had 5836 images altogether. Both healthy individuals and pneumonia patients were depicted in the pictures. There were 1583 photographs of healthy chest X-rays and 4273 images of chest X-rays with pneumonia. A training set with 5136 images and a test set with 700 images were created from the entire dataset. The figure shows the first two example images from the dataset.

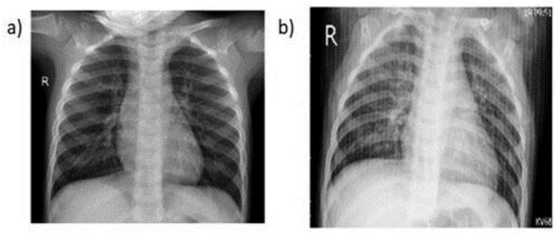

Figure 1: Chest X-ray of a healthy person (a) and a person suffering from pneumonia (b).

5. PROPOSED METHODOLOGY

This research suggests an ideal strategy for detecting pneumonia from chest X-rays. To expand the small dataset, data augmentation techniques were used.

5.1 PRE-PROCESSING AND DATA AUGMENTATION

The input photos were first reduced in size to 224*224 and used to train the model. A sizable dataset is necessary for efficient neural network training. Because the deep networks cannot generalize when trained on a smaller dataset, testing accuracy suffers. One approach to solving this issue is data augmentation, which extends and effectively makes use of the current dataset. There were 1283 healthy chest X-ray images and 3873 pneumonia-infected chest X-ray case images in the training dataset used in this study. Since the dataset already contained adequate photos of the pneumonia case, just the images of the normal case needed to be enhanced twice. There were 3849 normal pictures and 3873 photos of pneumonia following augmentation. Images from the test set were not enhanced.

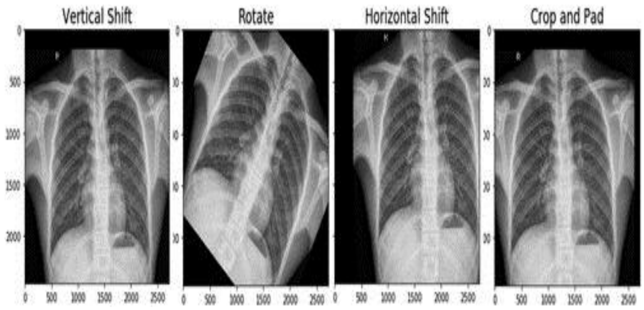

Figure 2: Resultant images after using augmentation techniques.

5.2 METHODS USED

1) CONVOLUTIONAL NEURAL NETWORK

Undoubtedly, the most well-known deep learning architecture is CNNs. They fall under the heading of feed-forward networks. CNN's ability to take locality into account is by far its greatest benefit. The CNN's convolutional layer, which gives the network its name, is its core component. The convolutional layer extracts the characteristics from an input image. Each input map's dimensionality is decreased via the pooling layer while critical data is kept. The network architecture is represented in Figure 1 as a sequence of various convolutional and pooling layers. The softmax layer is utilized for picture classification at the network's end. The main drawback of deeper CNNs is of vanishing gradients.

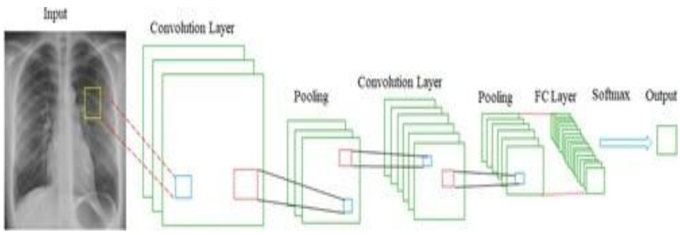

Figure 3: Architecture of the CNN.

2) TRANSFER LEARNING

In general, larger datasets are needed to train CNNs. CNN performs badly in generalization when trained on smaller datasets. In these situations, transfer learning is an option. The process of transfer learning is depicted in Figure where knowledge gathered by the model while resolving one problem is applied to resolve another.

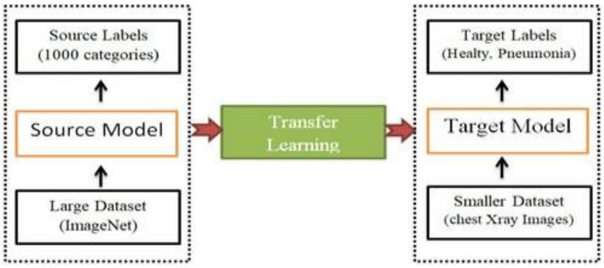

Figure 4: The mechanism of transfer learning.

DENSENET121

DenseNet (Dense Convolutional Network) is an architecture that focuses on making deep learning networks go deeper while also making them more efficient to train by using shorter connections between the layers. DenseNet is a convolutional neural network in which each layer is connected to all other layers deeper in the network, i.e. the first layer is connected to the second, third, fourth, and so on, and the second layer is connected to the third, fourth, fifth, and so on. This is done to allow for maximum information flow between network layers. To maintain the feed-forward nature, each layer receives input from all previous layers and passes on its feature maps to all subsequent layers.

VGG16

2014's ILSVR (Imagenet) competition was won using the convolution neural net (CNN) architecture VGG16. One of the best vision model architectures to date, according to many. The most distinctive feature of VGG16 is that, rather than concentrating on having many hyper-parameters, they concentrated on having convolution layers of a 3x3 filter with a stride 1 and always used the same padding and maxpool layer of a 2x2 filter with a stride 2. Convolution and max pool layers are arranged in this manner continuously throughout the entire architecture. It finishes with two FC (completely connected layers) and a softmax for output. There are 16 layers with weights, as indicated by the 16 in VGG16. There are around 138 million parameters in this network, making it a sizable network.

RESNET50

A common starting point for transfer learning is ResNet-50, a shrunken-down version of ResNet 152. There are five stages in the ResNet-50 model, each of which has a convolution and identity block. The identity blocks and each convolution block have three convolution layers. The ResNet-50 has more than 23 million trainable parameters.

INCEPTION V3

Convolutional neural network Inception v3 was developed as a plugin for GoogLeNet and is used to support object detection and picture analysis. The Google Inception Convolutional Neural Network, which was first presented during the ImageNet Recognition Challenge, is in its third iteration. Inceptionv3 was created to enable deeper networks without allowing the number of parameters to become unmanageably large; it contains "under 25 million parameters," as opposed to 60 million for AlexNet.

Inception aids in the classification of items in the field of computer vision, much like ImageNet can be seen as a database of categorized visual objects. Numerous applications have utilized the Inceptionv3 architecture, frequently using "pre-trained" data from ImageNet.

6. CONCLUSIONS AND FUTURE SCOPE

Pneumonia is a life-threatening infectious disease. For patients over 75 years, the mortality rate of pneumonia is 24.8%. In this paper, an algorithm that can further support the computer-aided diagnosis of pneumonia has been proposed. The deep residual network, proposed in the paper, has a more complex structure but fewer parameters and higher accuracy. Furthermore, this model was scaled up efficiently using the method of compound scaling. Data augmentation and transfer learning have also been used to tackle the obstacle of the insufficient training dataset. Different scores, such as recall, precision, and accuracy, were computed to prove the robustness of the model. The future works involve developing an algorithm that can localize the parts of the lung affected by pneumonia.

7. RESULTS

ACCURACY

In statistics, accuracy is an important measure to estimate the performance of a classifier. Accuracy is the proportion of the sum of TP (true positive) and TN (True negative) over a total number of predictions that is the sum of TP, TN, FP, and FN which is as follows.

$$accuracy = \frac{TP+TN}{total\ number\ of\ predictions\ = \{TP+TN+FP+FN\}}$$

Table 1: Training accuracy and Testing accuracy of different models.

Models	Training Accuracy	Testing Accuracy
CNN	93.23	82.53
DENSENET121	79.70	76.76
VGG16	81.77	64.74
RESNET50	74.10	62.50
INCEPTION V3	89.53	80.29

Comparing different models

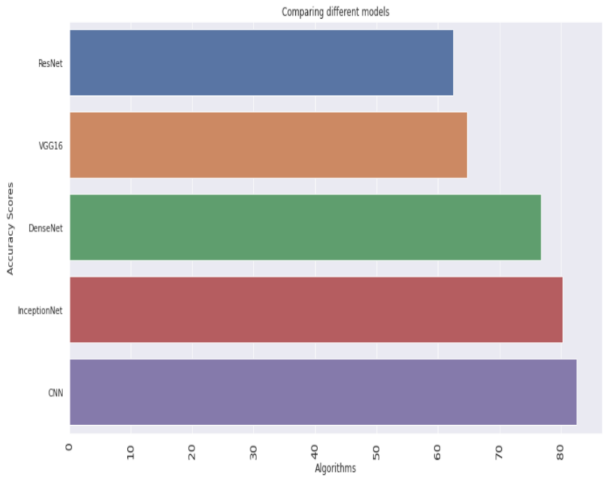

Figure 5: Comparing testing accuracy of different models

References

1. Gonçalves-Pereira J, Conceição C, Póvoa P. Community-acquired pneumonia: identification and evaluation of nonresponders. Ther Adv Infect Dis. 2013;1(1):5–17. [Crossref], [Google Scholar]

2. Razzak MI, Naz S, Zaib A. Deep learning for medical image processing: overview, challenges, and the future. 2017. [Crossref], [Google Scholar]

3. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annual review of biomedical engineering. 2017;19:221–248. doi:10.1146/annurev-bioeng-071516-044442. [Crossref], [Web of Science ®], [Google Scholar]

4. Andre E, Brett K, Roberto A, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–118. [Crossref], [Web of Science ®], [Google Scholar]

5. Milletari F, Navab N, Ahmadi S. V-net: fully convolutional neural networks for volumetric medical image segmentation. 2016 fourth International Conference on 3D Vision (3DV); Stanford, CA; 2016:565–571. doi:10.1109/3DV.2016.79. [Crossref], [Google Scholar]

6. Grewal M, Srivastava MM, Kumar P, et al. 2017. Radiologist level accuracy using deep learning for hemorrhage detection in CT scans. arXiv preprint arXiv:1710.04934. [Crossref], [Google Scholar]

7. Varun G, Lily P, Marc C, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2017;316(22):2402–2410. [Web of Science ®], [Google Scholar]

8. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology. 2017;284(2):574–582. [Crossref], [Web of Science ®], [Google Scholar]

9. Bar Y, Diamant I, Wolf L, et al. Chest pathology detection using deep learning with non-medical training. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI); New York (NY); 2015:294–297. doi:10.1109/ISBI.2015.7163871. [Crossref], [Google Scholar]

10. Shin HC, Lu L, Kim L, et al. Interleaved text/image deep mining on a very large-scale radiology database. Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR); Boston (MA); June 2015. [Crossref], [Google Scholar]

11. Shin HC, Lu L, Kim L, et al. Interleaved text/image deep mining on a large-scale radiology database for automated image interpretation. J Mach Learn Res. 2016;17(40):1–31. [Google Scholar]

12. Boussaid H, Kokkinos I. Fast and exact: ADMM-based discriminative shape segmentation with loopy part models. Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR); Columbus (OH); June 2014. [Crossref], [Google Scholar]