Automated Diagnosis of Pneumonia using CNN and Transfer Learning Approaches

A. Madhavi1*, Marisetti Sathya Abhijna1, Peddanna Sumanjali1, Shane Supraja1, Manish Ghalwan2, M. Rathna Chary3

1Department of CSE (AIML), GRIET, Hyderabad, Telangana State, India
2Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, 248007, India
3KG Reddy College of Engineering & Technology

Abstract. Pneumonia is one of the most deadly diseases, especially for children below 5 years of age. To detect pneumonia radiologists, have to observe the chest x-ray and he/she has to update the doctor correctly which sometimes may not be accurate due to human error. The main objective of this paper is to identify if the person has Pneumonia or not with high accuracy. Automated diagnosis of pneumonia can be done with the help of CNN and Transfer Learning Approaches so that the person can get treatment as early as possible. The dataset used here is the chest X-ray (CXR) dataset based on a chest X-Ray scan database from paediatric patients from one to five years of age at the Guangzhou Women and Children’s Medical Centre. Deep Learning (CNN) and Transfer Learning Techniques along with Ensemble Learning have been implemented concluded that CNN achieved an accuracy of 89%, the Transfer Learning model achieved an accuracy of 93% and the ensemble model got an accuracy of 92%. Even though the highest accuracy is for the Transfer Learning model, considering all the other metrics like Recall, Support, and score, Ensemble has exhibited the best results.

1 Introduction

An illness called pneumonia, causing inflammation in the air sacs, pneumonia can also be brought on by other microorganisms and specific chemicals. When someone breathes, air passes through their airways and into their alveoli, where it is converted into oxygen and injected into their circulation. In pneumonia, the alveoli swell up and fill with fluid or pus, impairing lung function and resulting in symptoms including coughing, fever, breathing difficulties, and chest pain. Types of pneumonia are as follows:

- The type of pneumonia known as bacterial pneumonia is brought on by bacteria infecting the lungs. Streptococcus pneumonia is the most frequent bacteria that causes this kind of pneumonia.
- Viral pneumonia occasionally causes serious respiratory distress and consequences, particularly in those with compromised immune systems, elderly people, and people with underlying medical disorders. Laboratory testing, a physical examination, and a

* Corresponding author: arelli.madhavi1988@gmail.com
chest X-ray are used to diagnose viral pneumonia.

- Pneumocystis Pneumonia (PCP), also known as fungal lung infection or fungal pneumonia, is a kind of pneumonia brought on by certain fungi. Pneumocystis, the most frequent fungus linked to pneumonia, primarily affects those with compromised immune systems, such as those with HIV/AIDS or receiving immunosuppressive medication.

2 Literature survey

Two convolutional neural networks (CNN) designs were proposed by Harsh and team [1]. The 5863 Chest X-ray pictures in the dataset were downloaded from Kaggle. The study found that CNN outperformed the other models when a dropout layer was trained using augmented data.

Sammy and team [2] proposed a CNN adaptive Deep-learning model to identify pneumonia. Data augmentation was employed for an intensified dataset, overfitting was prevented by pooling, and the number of parameters was decreased. The system was assessed using a dataset of 28,000 chest X-ray images from the Radiological Society of North America (RSNA). The analysis found that the GoogleNet and LeNet models performed the best.

A confidence-aware anomaly detection model (CAAD) was proposed by Zhang and team [3]. The goal is to use the CAAD model to separate all cases of viral pneumonia from all cases of non-viral pneumonia in chest X-ray pictures. The X-VIRAL, X-COVID, and X-ray datasets were utilised. The investigation came to the conclusion that the classification of viral pneumonia was successfully accomplished by anomaly detection.

The VGG19 network was employed by Rajasenbagam and team [4] for the quick identification of pneumonia. Creating a Deep Convolutional Neural Network model to identify pneumonia from chest X-rays was the goal. The study indicated that the proposed DCNN was the best option for using chest X-ray images to diagnose pneumonia infection.

Deep-learning methodologies were proposed for the auto-detection of pneumonia by Bhattacharyya and his team [5]. Approaches for image segmentation and classification based on DL were used. To obtain the lung images, the raw X-ray images were segmented using the Conditional Generative Adversarial Network (C-GAN). The study found that, of all the evaluated supervised machine learning algorithms, the C-GAN model produced the best results.

Transfer learning and ensemble approaches for pneumonia classification were proposed by Kumar and team [6]. To develop a Deep Transfer Learning-based Ensemble model for the early detection of COVID-19 infection, EfficientNet, GoogLeNet, and XceptionNet were employed. The study comes to the conclusion that the novel deep transfer learning model is capable of identifying X-ray chest pictures as discriminating normal, COVID-19 (+), pneumonia, and tuberculosis-infected cases.

In order to diagnose pneumonia, Tang and team [7] recommended Quantitative and Qualitative analysis of 18 Deep Convolutional Neural Networks in addition to Transfer Learning. A total of 700 CXR images make up the dataset, of which 280 were used for training, 120 for validation, and 300 for testing. The study recommends VGG-16 and SqueezeNet as additional tools for the diagnosis of pneumonia.

VGG19 was recommended by Nilanjan and team [8] for the detection of pneumonia. A modified VGG19 network using the Ensemble Feature Scheme (EFS) is suggested. There were 7150 CXR pictures total in the collection. This study provided proof that the suggested DLS would function properly on the test photos.

To identify pneumonia, Rachna and team [9] used 6 different transfer learning models. The goal was to use deep learning to use chest X-rays to detect pneumonia. To lessen
overfitting, dropout was used in models VGG16, VGG19, and ResNet. About 15 million photos in 22,000 different categories make up the dataset. The study came to the conclusion that medical officers might utilise the CNN Model 2 and VGG19 models to diagnose pneumonia early on.

A good deep learning model for pneumonia diagnosis was proposed by Saikrishna and his team [10]. Convolutional neural network (CNN) deep learning models were used to find the best model for detecting pneumonia. Chest X-rays make up the dataset, which was taken from Kaggle. This study explored a number of methods, including ResNet50, Inception, Modified CNN, ChexNet, R-CNN, CNN, Mask-RCNN, Transfer Learning (CNN), and Dual Net architecture. According to the study's findings, Mask-RCNN demonstrated superior performance in the most accurate pneumonia identification.

The model for detecting pneumonia combining CXR pictures and CT images techniques was described by Varalakshmi and team [11]. With the help of CXR and CT scans, features from Transfer Learning, and Haralick, the objective was to locate COVID-19. Effective transfer learning and fine-tuning procedures were examined rather than starting from scratch to create the CNN model. According to the study's findings, the suggested CT+X-Ray images had a 93% accuracy rate using transfer learning.

The development of a solid model for the categorization of features for the Covid-19 analysis was the focus of Sitaula and team [12] effort. The goal was to use a variety of deep visual, word-based criteria to categorise chest X-ray pictures for COVID-19 diagnosis. Bag of Deep Visual Words is a strategy that uses deep features. The four categories in Dataset 4 are Covid, Normal and Pneumonia.

Ensemble Deep Learning was proposed by Tang and team [13] for the detection of COVID-19 and pneumonia patients. There are 13,975 CXR pictures in the COVIDx dataset. The study found that the EDL-Ensemble Deep Learning technique, which promises high accuracy and outperforms individual deep learning models.

The CovXNet-multi-dilation convolutional neural network with transferable multi-receptive feature optimisation was proposed by Mahmud and team [14] for the identification of pneumonia from chest X-ray pictures. To identify anomalous locations, gradient-based discriminative localization was developed. There are 5856 photos in the datasets. The study found that the CovXNet is highly extensible and suitable for usage in a variety of CV applications, and the stacking algorithm produced effective and improved outcomes.

CNN-PCA-ELM + (CLAHE) was recommended by Nahiduzzaman and his team [15] as a classification method for pneumonia. The work in this study involved creating three different models: the Extreme Learning Machine (ELM), the CNN-PCA-ELM hybrid convolutional neural network-based feature extraction method, and the CNN-PCA-ELM with Contrast-Enhanced by Contrast-Limited Adaptive Histogram Equalisation (CLAHE) technique. According to the results of the investigation, the upright ELM model with CLAHE and hybrid CNN-PCA acquired high performance.

Treating pneumonia using the Ensemble model in Chest X-Ray Images was proposed by Ayan and team [16]. In this study, transfer learning was used to train seven separate models, and the top three models were pooled to create an ensemble model. The study found that class distribution in a balanced dataset is significant and that the ensemble model outperformed the CNN model in terms of accuracy.

Authors [17] explored the application of Transfer Learning (TL) in automated medical image analysis, highlighting its effectiveness in various tasks. TL models like AlexNet, ResNet, VGGNet, and GoogleNet prove valuable for enhancing medical image analysis. Authors [18] presented data-driven prediction techniques using ARIMA and LSTM to forecast COVID-19 cases and deaths. It uses statistical measures to assess accuracy and aims to assist several countries in managing the pandemic. Authors [19] discussed the
importance of quick COVID-19 detection using Chest X-ray images and presented a DL method achieving 99% accuracy in binary classification of COVID-19 cases using open-source datasets. Authors [20] highlighted the significance of ML in prediction, pattern recognition and error reduction across diverse fields, emphasizing the impact of AI in broad domain. Authors [21] discussed the use of Scanning Electron Microscopy (SEM) for material characterization and how Python programming is employed to process SEM images, including histogram equalization and morphological operations for accurate analysis.

3 Proposed method

3.1 Problem statement

Bacterial pneumonia is more hazardous than virus pneumonia. It is the leading cause of infectious death. People who smoke, age above 65, or survive with lung or heart diseases are prone to pneumonia more. To detect pneumonia radiologists, must observe the chest x-ray and he/she must update the doctor correctly which sometimes may not be accurate due to human error. The objective of this model is to classify if the person has Pneumonia or not with high accuracy. Automated diagnosis of pneumonia can be done with the help of CNN and Transfer Learning Approaches so that the person can get treatment as early as possible.

3.2 Objectives

- To develop a deep learning framework to automatically diagnose pneumonia using chest X-ray images and to classify the result as normal cases or pneumonia cases.
- Timely Automated detection of pneumonia in children can help to fast-track the recovery.

3.3 Proposed method

This work uses deep learning models such as CNN and transfer learning. The CNN can train the images of chest X-rays and then it can predict with good accuracy. Transfer learning is reusing the pre-trained model that is trained on a large dataset and using it accordingly to our model.

3.4 Architecture of the proposed work

3.4.1 Data collection

Data is been collected from Kaggle. This dataset is a collection of anterior and posterior chest X-rays of kids and women in Guangzhou which were scanned for quality check and omitted the degrading images.

3.4.2 Data pre-processing

It is a technique where data is pre-processed before training the model, which increases and modifies the original data for better results. Data augmentation is basically two types where there is augmented data and synthetic data. In this work, data augmentation is performed...
and uses the augmented dataset for training the model. In data augmentation, the given image is set as input, and the augmentation techniques are applied like rescaling, zooming, brightness range, shift range and rotation.

**Fig. 1.** Architecture diagram of pneumonia classification.
3.4.3 Training and testing

Training a model involves building the architecture for the model including the layers and weights and classifying the results into respective classes, input of training model will be the augmented data. The first approach is multiple layers make up a CNN which processes the input image to carry out various operations like feature extraction, down-sampling, and classification. The second approach is using Transfer learning, used exception, VGG19, and Mobile Net. Combining the three models and building an ensemble model in transfer learning, the basic knowledge is extracted. Some examples are VGG16, VGG19, ResNet, MobileNet. The third approach is Ensemble Learning, Three models of transfer learning were combined into a single model. Ensemble learning is a combination of multiple classifiers and predicts the class by taking the prediction of all the classifiers into consideration.

3.4.4 Classification

Input is given as an image of the test set which contains the images of pneumonia and normal cases and is classified into respective classes.
4 Results and discussions

4.1 Description of the dataset

This work is implemented using chest X-Ray images (pneumonia) dataset which contains around 5800 images that were divided into the train, test, and validation images. Around 3000 of the images are in the train folder. The dataset is a mixture of Pneumonia which is again classified into Bacterial and Viral, Normal cases.

4.2 Experimental results

4.2.1 Confusion matrices

To evaluate the algorithm's accuracy in classification issues, it is frequently utilized. A table with four distinct values is referred to as a confusion matrix.

![Fig. 4. CNN confusion matrix.](image)

![Fig. 5. VGG19 confusion matrix.](image)

Here even though the ensemble model just got an accuracy of 92%. It has classified images into a good number of True Positives and True Negatives. Ensemble model has 190 True positives and 353 True negatives. In the case of XceptionNet, it has 169 True Positives and 368 True Negatives.
**4.2.2 Evaluation report**

A summary of the evaluation metrics used to assess the model's performance, including accuracy, precision, recall, and F1 score. Accuracy is the ratio of accurately predicted instances to the total number of instances. Precision is the ratio of genuine positive predictions to total projected positives. It determines how many of the projected positive events are actually positive. Recall (Sensitivity or True Positive Rate): The proportion of
true positive predictions to total real positives. It determines how many of the actual positive events were correctly predicted. F1-Score is the harmonic mean of precision and recall. It provides a balanced measure that accounts for both false positives and false negatives. Specificity (genuine Negative Rate): The proportion of genuine negative predictions to total actual negatives.

Fig. 9. CNN report.

Fig. 10. VGG19 report.

Fig. 11. MobileNet report.

The F1 score for normal case and pneumonia case prediction is highest for the Ensemble model Figure 4.10 which is 0.82 and 0.90, with precision for normal case prediction being highest for Xception Net which is 0.97, and for pneumonia, case highest for the ensemble model which is 0.89, Recall for normal case and pneumonia case is being highest for the ensemble model that is 0.81 and 0.91 respectively. Support for normal case and pneumonia case prediction is 234 and 390 respectively. A binary classifier system's performance is graphically represented by a ROC (Receiver Operating Characteristic) curve as the discrimination threshold is changed.
4.2.3 ROC curves

**Fig. 12.** Xception report.

<table>
<thead>
<tr>
<th></th>
<th>precision</th>
<th>recall</th>
<th>f1-score</th>
<th>support</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.97</td>
<td>0.61</td>
<td>0.75</td>
<td>234</td>
</tr>
<tr>
<td>1</td>
<td>0.81</td>
<td>0.99</td>
<td>0.89</td>
<td>390</td>
</tr>
</tbody>
</table>

**Fig. 13.** Ensemble report.

<table>
<thead>
<tr>
<th></th>
<th>precision</th>
<th>recall</th>
<th>f1-score</th>
<th>support</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.88</td>
<td>0.72</td>
<td>0.80</td>
<td>234</td>
</tr>
<tr>
<td>1</td>
<td>0.85</td>
<td>0.94</td>
<td>0.89</td>
<td>390</td>
</tr>
</tbody>
</table>

**Fig. 14.** CNN ROC curve.

**Fig. 15.** VGG19 ROC curve.
Here are some common interpretations of ROC curves:

- This indicates a high TPR and low FPR, meaning that the model is correctly identifying the most positive instances while making fewer false positive predictions.
- The diagonal line from the bottom-left to the top-right corner of the plot represents the performance of a random classifier, which has an equal chance of correctly predicting positive or negative instances. A classification model that performs worse than the random classifier will have a ROC curve below this line.
- An AUC value between 0.5 and 1.0 represents varying degrees of classification accuracy.

The area under the ROC curve is the key metric for the evaluation of the model. The more the area indicates the better curve. Therefore, from the study, we conclude that the Ensemble model
4.3 Significance of proposed methods with its advantages

4.3.1 Early stopping

Early stopping allows the model to train until it reaches an optimal level of performance on a validation set, after which training is stopped to prevent the model from overfitting to the training data.

4.3.2 Fine-tuning

When fine-tuning a pre-trained model, it is common practice to freeze some of the layers and only update the weights of the remaining layers. Freezing layers means that their weights are not updated during training.

5 Conclusion and future enhancements

We have used CNN and transfer learning approaches for automated diagnosis of pneumonia we got an accuracy of 89% using CNN, 85% using VGG19, 91% using MobieNet. In transfer learning, we got the highest accuracy of 93% using XceptionNet. At last, we proposed the Ensemble model by integrating 3 different models and got an accuracy of 92%. Automated diagnosis of pneumonia can be done with the help of CNN and Transfer Learning Approaches so that the person can get treatment as early as possible with high accuracy which is the main aim of this paper. Hence, we can conclude that the Xception model in transfer learning has shown good accuracy whereas in the ensemble model even though there is a significant difference in accuracy but the precision and recall were high for the ensemble model. Automated diagnosis of pneumonia by using this ensemble model helps the patients to be treated at the early stages which satisfies the objective of this work.

References

2. S. V. Militante, N. V. Dionisio, B. G. Sibbaluca, Pneumonia Detection through Adaptive Deep Learning Models of Convolutional Neural Networks, In the proceedings 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia (2020)
7. L. S. Chow, G. S. Tang, M. I. Solihin, , SN Comp 4, 141 (2023)


