International Journal of Computer Sciences and Engineering
Vol.**11**, Issue **5**, pp.**01-05**, May **2023**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

Research Paper

Pneumonia Detection on Chest X-ray Images Using Hybrid Convolution Neural Networks

Pradeep Rao K.B.[1*], H Manoj T. Gadiyar[2], Guruprasad[3], Basavaraj N.[4], Dhanush T.M.[5], Mugodera Madhukumara[6]

[1,2,3,4,5,6]Dept. of Computer Science and Engineering, Sri Dharmasthala Manjunatheshwara Institute of Technology, Ujire-574240, India

**Corresponding Author: kbpradeeprao@gmail.com*

Received: 25/Mar/2023; **Accepted**: 09/May/2023; **Published**: 31/May/2023. **DOI**: https://doi.org/10.26438/ijcse/v11i5.15

Abstract: Pneumonia primarily affects individuals who are either older than 65 years or younger than five years. Timely identification and prompt treatment of pneumonia can significantly improve the chances of survival for individuals. Pneumonia detection often involves extensive analysis of Chest X-ray images. Recent research indicates that the utilization of deep learning technique holds significant promise in the accurate identification and diagnosis of pneumonia. A novel approach is proposed in this research, where a hybrid Convolutional Neural Network is introduced for the purpose of pneumonia detection in chest X-ray images. In this approach, initially images of Chest X-ray are gathered and preprocessed. Later feature extraction was done using VGG16 and VGG19 model. After training and testing Machine Learning (ML) classifiers, an ensemble classifier was created for classification of pneumonia. Experiment results shows that ensemble classifier outperforms existing state of art methods by exhibiting superior accuracy and recall performance.

Keywords: Pneumonia,Chest X-ray, VGG16, VGG19, Ensemble Classifier, Convolutional Neural Network

1. Introduction

Pneumonia is an illness characterized by the infection and inflammation of lungs. It occurs due to the presence of infection in respiratory system. Fever, shivering, cough, lack of appetite and pain while breathing are the commonly observed symptoms associated with pneumonia. Pneumonia is generally observed among people older than 65 years and children below 5 years. Pneumonia affects approximately 7.7 of the global population on an annual basis. According to statistics, the number of deaths caused by this illness has sharply grown in recent years. Many children and old people die due to this disease. Pneumonia is primarily caused by fungi, bacteria and viruses. Early identification of pneumonia has the potential to greatly increase survival rates of individuals.

Various techniques are employed for the detection of pneumonia. Recent research study suggests that Chest X-ray (CXR) is most viable tool for detecting pneumonia. Chest X-ray images are analyzed by radiologist for detecting pneumonia. Detecting pneumonia accurately from CXR images requires considerable amount of time and effort for human beings. Recently machine learning(ML) is widely used for detecting different diseases such as breast cancer and heart disease. An ML model can be trained using CXR images to develop a system capable of detecting pneumonia. This can assist radiologist in detection of pneumonia. In this research, convolution neural network model (CNN) is proposed which identifies pneumonia from CXR images. Medical practitioners can use this model for detection of pneumonia.

The subsequent sections of the paper are structured as follows. Section 2 summarizes various research activities carried out for detection of pneumonia. Section 3 contains the methodology and pseudo code. The architecture of pneumonia detection system is outlined in section 4. Section 5 contains findings and analysis. Conclusion of research work is highlighted in Section 6.

2. Related Work

In recent years, deep learning methods have been increasingly employed by researchers to identify diseases in chest X-ray images. Jain et al. analyzed six models for detection of pneumonia. The models were VGG16, VGG19, ResNet50, Inception-v3, CNN Model 1, CNN Model 2. Out of six models, CNN Model 2 provided good performance with validation accuracy of 92.31% and recall of 98% [1].

Gupta et al. addressed the issue of pneumonia detection by developing an ensemble model. Transfer learning was used for development of model. Five neural networks were used for developing ensemble model. The developed model provided good performance with 96.4 % accuracy and 99.62% recall [2].

Sharma et al. detected pneumonia using two CNN models. Dropout layer was used in first model, whereas in second model it was absent. Overfitting was done using various augmentation techniques. Model having combination of augmentation and dropout layer gave 90.68% accuracy [3].

Chowdhury et al. used CXR images for detection of pneumonia by training four models. While training the models, image augmentation was applied. Models used for training were AlexNet, ResNet18, DenseNet201 and SqueezeNet. Experiments results showed that DenseNet201 gave good performance by giving an accuracy of 98% during classification of pneumonia[4].

Rajpurkar et al. detected pneumonia using a CNN based approach called CheXNet. 121 layers were used in CheXNet. Around 1 lakh chest X-ray images were used for training the model. CheXNet gave good performance while comparing with radiological pneumonia detection[5].

Sain et al. used chest X-ray images for extracting features and to detect pneumonia. CNN based method achieved excellent classifier performance[6].

Wu et al. introduced a hybrid model named ACNN-RF for identifying pneumonia from images of chest X-ray. Model utilized a combination of adaptive average filtration CNN and Random Forest(RF). Initially, noise was removed from chest X-ray image by applying adaptive filtration. Later features were extracted using CNN model. In last step, for identifying pneumonia RF model was trained using extracted features. The model gave good performance by providing 97% accuracy [7].

3. Methodology

3.1 Implementation

1) Data Acquisition

For this research study, a publicly available dataset of CXR images (5856) are used [8]. Images are classified into two classes. CNN model is trained using 80% of the data and tested using 20% of the data. To tackle the problem of imbalanced data, an equal proportion of images from every class are chosen.

2) Data Preprocessing

Some preprocessing was performed on the dataset before training. During data preprocessing, initially images are resized to required dimensions. Later resized images are rescaled.

3) Feature Extraction using Pre-Trained VGG Models

For extracting features, VGG models are used. To VGG models, images are fed as input. VGG19 networks accept input images of dimension 224 × 224 pixels. VGG16 and VGG19 run concurrently by pooling together the extracted features from CXR images. Final layer extracts features of dimensions 7×7×512. By combining the extracted features, a resulting output with dimensions of 7x7x1024 is obtained through feature concatenation. Five ML classifiers are employed to capture the most informative CXR image features that accurately depicts the characteristics of the input images.

4) Training and Testing of ML Classifiers

All the features are combined and fed to various ML classifiers. The performance of the classifiers trained on the training set is assessed by evaluating them on test set.

5) Creation of an Ensemble Classifier

Ensemble classifier is created using trained classifiers. CXR images are classified using probability vector. Each classifier is assigned weights.

6) Hyperparameter Tuning

Hyperparameter tuning was performed on both the individual ML classifiers and the ensemble model. In order to strike a balance between model complexity and training set error, specific hyperparameters were fine-tuned for each of the 5 machine learning classifiers. Each ML classifier is tested with different weights. Finally, overall performance of ensemble model was optimized.

3.2 Pseudocode

Algorithm 1: Hybrid CNN Model for Diagnosing Pneumonia
Input: Chest X-ray dataset
Output: Pneumonia detecting model

1. Resize and rescale CXR images.
2. Extract the features from CXR images using VGG-16 and VGG-19 models.
3. Replace the FC layers at the end of the VGG models with ML classifiers to classify CXR images as either pneumonia or non-pneumonia.
4. Train and test multiple ML classifiers (such as KNN, LR, NB, RF, and SVM) on the feature maps extracted from the VGG models. Use an 80/20 split of the dataset for training and testing.
5. Create an ensemble classifier by combining the predictions of the ML classifiers. The ensemble classifier's output is based on the predictions of all models, with a majority of pneumonia predictions leading to a pneumonia classification.
6. Examine the effectiveness of different ML classifiers as well as ensemble classifier. Then determine the model that achieves highest accuracy.
7. To make forecast on new chest X-ray images, choose the model that gives highest performance.

4. Proposed Architecture

Figure 1 shows the architecture of the pneumonia detection system. During first step, CXR images are collected and preprocessed. Preprocessing involves image resizing and rescaling. Later images are divided into training and test dataset. VGG16 and VGG19 networks are utilized to extract features from CXR images. Features are then combined. Concatenated features are given as input to 5 machine learning classifiers. Predictions of different ML classifiers are combined to create an ensemble classifier. In final step, after evaluating the performance of ensemble classifier the best model is chosen for predicting new unseen chest X-ray image.

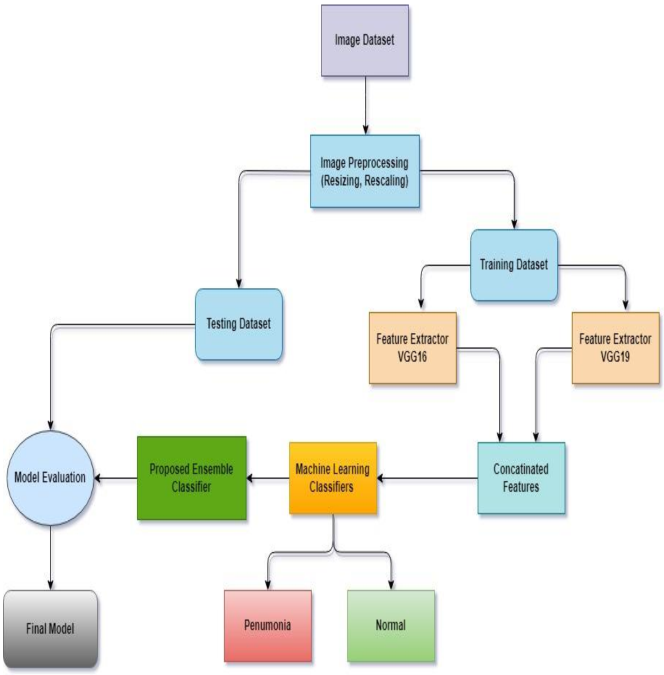

Figure 1. Proposed System Architecture

5. Results and Discussions

The research work incorporated a dataset comprising 5856 images with 4686 images allocated for training purposes and 1170 images allocated for testing. The majority of CXR images, specifically 80% are utilized for training purposes. Rest of the CXR images are used for testing purpose. CXR images are classified as Normal and Abnormal. Abnormal indicates presence of pneumonia. Details of dataset are given in Table 1.

Table 1. Details of Dataset

Classification	Training Data	Test Data
Normal	1267	316
Abnormal	3419	854

Table 2 shows the accuracy of different algorithms. Logistic regression surpasses other algorithms and achieves an impressive accuracy of 91.73%. Naive bayes provides least accuracy among 5 algorithms.

Table 2. Analysis of Different Algorithms

Algorithms	Accuracy
SVM	90.06%
KNN	89.98%
Logistic Regression	91.73%
Random Forest	80%
Naïve Bayes	74.8%

Figure 2 shows the accuracy comparison of five algorithms.

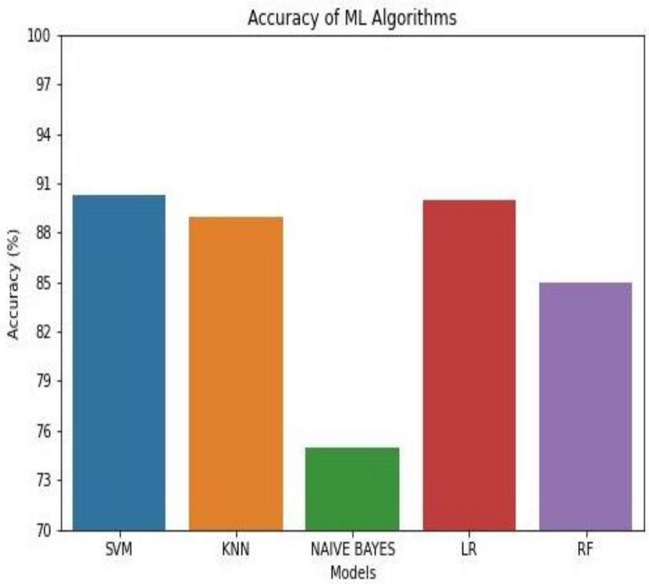

Figure 2. Analysis of Machine Learning Algorithms

The performances of ML algorithms were evaluated using five parameters. Accuracy denotes the percentage of correctly labeled images. Precision, recall and F1-score represents the model's capacity to accurately detect instances of pneumonia as normal and abnormal cases. Specificity represents the model's ability to correctly identify negative cases. The ensemble classifiers performance metrics is shown in Table 3. Ensemble classifier provides recall of 99% and accuracy of 90.34%.

Table 3. Performance Metrics

Criterion	Value
Precision	86%
Recall	99%
F1-score	92%
Specificity	73%
Accuracy	90.34%

The application was developed using the nodeJs framework. Python, HTML, CSS and JavaScript were used for developing pneumonia detection system. Figure 3 presents the confusion matrix results for the Logistic Regression model's performance.

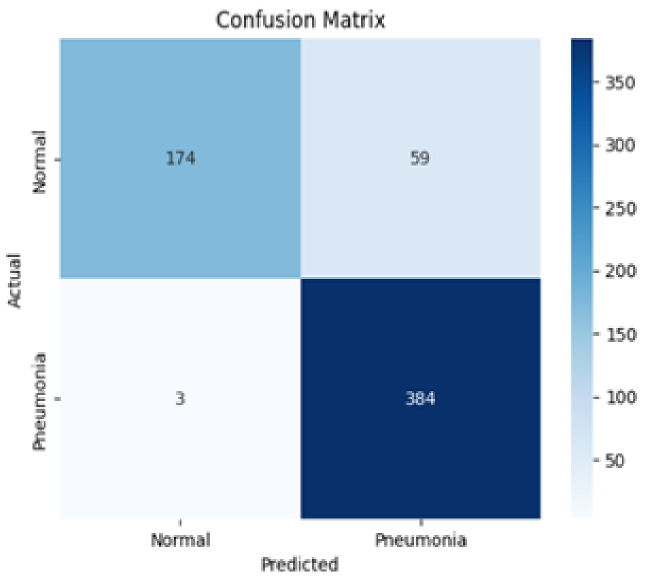

Figure 3. Confusion Matrix of Logistic Regression

6. Conclusion and Future Scope

Pneumonia claims the lives of numerous children and elderly individuals annually, making it a perilous disease. If pneumonia is detected at early stage, we can save many people's life. Researchers have employed CXR images to identify the cases of pneumonia. In this study, CXR images are analyzed for detection of pneumonia by using hybrid VGG-based model. Ensemble classifier is created by combining the predictions of ML classifier. After evaluating the performance of the ensemble classifier best performing model is used for prediction. Experiment results shows that the ensemble classifier performs better than its counter candidates providing better accuracy. As a future work, performance of the CNN architecture can be further enhanced by tuning the hyper parameters. Mobile application can be developed for assisting radiologists in detecting pneumonia.

References

[1]. R Jain, P Nagrath, G Kataria, V S Kaushik, D J Hemanth, "Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning," *Measurement*,Vol.**165**,**2020**.

[2]. Chouhan V, Singh S K, Khamparia A, Gupta D, Tiwari P, Moreira C, De Albuquerque," A novel transfer learning based approach for pneumonia detection in chest X-ray images", *Applied Sciences*, Vol. **10**, Issu**e 2**, pp. **1-17**,**2020**.

[3]. Sharma H, Jain JS, Bansal P, Gupta S, "Feature Extraction and Classification of Chest X-ray images Using CNN to Detect Pneumonia,"*International Conference on Cloud Computing, Data Science & Engineering* , pp.**227-231**,**2020** .

[4]. Rahman T, Chowdhury ME, Khandakar A, Islam KR, Islam KF, Mahbub ZB, Kadir MA, Kashem S , "Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray," *Applied Sciences*, Vol. **10**, Issue **9**, pp. **1-17**,**2020**. https://doi.org/10.3390/app10093233

[5]. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz C, Shpanskaya K, Lungren MP, "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," pp. **1-7**,**2017**.

[6]. Stephen O, Sain M, Maduh UJ, Jeong DU, "An efficient deep learning approach to pneumonia classification in healthcare," Journal of healthcare engineering, Vol. **2019**, pp. **1-7**, **2019**.

[7]. Wu H, Xie P, Zhang H, Li D, Cheng M , "Predict pneumonia with chest X-ray images based on convolutional deep neural learning networks," *Journal of Intelligent & Fuzzy Systems*, Vol. **39**, Issue **3**, pp. **2893-2907**, **2020**. 10.3233/JIFS-191438

[8]. Kermany D, Zhang K, Goldbaum M, "Labeled optical coherence tomography (oct) and chest x-ray images for classification," pp. **651**, **2018**.

[9]. Yaseliani M, Hamadani AZ, Maghsoodi AI, Mosavi A., "Pneumonia Detection Proposing a Hybrid Deep Convolutional Neural Network Based on Two Parallel Visual Geometry Group Architectures and Machine Learning Classifiers," Vol. **10**,pp.**62110-62128**,**2022**.

[10]. Mabrouk A, Díaz Redondo RP, Dahou A, Abd Elaziz M, Kayed M,. "Pneumonia Detection on Chest X-ray Images Using Ensemble of Deep Convolutional Neural Networks," *Applied Science*, Vol. **12**, Issue **13**, pp. **1-15**, **2022**. https://doi.org/10.3390/app12136448

[11]. Militante S V, Dionisio N V, Sibbaluca B G, "Pneumonia and COVID-19 detection using convolutional neural networks," International Conference on Vocational Education and Electrical Engineering, pp. **1-6**,**2020**.

[12]. Sirish Kaushik V, Nayyar A, Kataria G, Jain R, "Pneumonia detection using convolutional neural networks (CNNs)," *International Conference on Computing, Communications, and Cyber-Security*, pp. **471-483**,**2020**.

[13]. Bibek Ranjan Ghosh, Siddhartha Banerjee, Ayush Nandi, Arya Panja, Pritam Mondal, Arunava Das, "Automated Covid-19 Detection System with CNN using Chest X-Ray and CT-Scans," *International Journal of Computer Engineering*, Vol.**10**, Issue.**5**, pp.**47-52, 2022.**

[14]. Amanze B C, Asogwa D C, Chukwuneke C I, "An Android Mobile Expert System for the Diagnosi of Pneumonia with Object-Oriented Methodology," *International Journal of Computer Engineering*, Vol.**7**, Issue.**4**, pp.**745-758, 2019**.

AUTHORS PROFILE

Mr. Pradeep Rao K.B has completed Bachelor of Engineering in CSE from KVG College of Engineering, Sullia in 2014. He has completed his M.Tech in CSE from NMAMIT, Nitte in 2016. His area of interests includes machine learning and image processing.

Dr. H. Manoj T. Gadiyar has completed his B.E in the year 2006 and M. Tech in the year 2015 from VTU Belagavi. He has been conferred with PhD degree in Faculty of Computer and Information Sciences for the thesis titled "Novel Approach to Provide Context Aware Privacy for Consumer using IoT and Cloud" during 22nd Annual convocation by VTU on 24th February 2023. His areas of Interest are Cloud Computing, IoT and AI. He has 17 national and international research publication to his credit. He is the life member of ISTE, Kannada Sahitya Parishad and IIP.

Guruprasad is currently studying BE in CSE at SDM Institute of Technology, Ujire. He completed his schooling from Government High School, Brahmavara, Udupi.

Basavaraj N is currently studying BE in CSE at SDM Institute of Technology, Ujire. He completed his schooling from Poorna Pragna Vidyalaya, Kollegal.

Dhanush T M is currently studying BE in CSE at SDM Institute of Technology, Ujire. He completed his schooling from Kalpataru Central School, Tiptur.

Mugodera Madhukumara is currently studying BE in CSE at SDM Institute of Technology, Ujire. He completed his schooling from Shree Taralabaalu Siddeshwara High School, Thoolahalli.

