DEEP LEARNING BASED MODEL FOR THE DETECTION OF PNEUMONIA FROM CHEST X-RAY IMAGES USING RESNET50 AND NEURAL NETWORKS

Surajit Dey
Computer Science and Technology
Dr. B C Roy Polytechnic
Durgapur, India
surajit.dey@bcrec.ac.in

Saikat Chatterjee
Computer Science and Technology
Dr. B C Roy Polytechnic
Durgapur, India
saikat.chatterjee@bcrec.ac.in

Santu Kundu
Computer Science and Technology
Dr. B C Roy Polytechnic
Durgapur, India
santu.kundu@bcrec.ac.in

ABSTRACT— PNEUMONIA IS AN INFECTION OF THE LUNGS CAUSED BY BACTERIA, VIRUSES, FUNGI, OR PARASITES, LEADING TO THE ACCUMULATION OF PUS IN THE AIR SACS, WHICH CAN AFFECT ONE OR BOTH LUNGS. THIS SERIOUS ILLNESS POSES A GLOBAL HEALTH THREAT, WITH EARLY DIAGNOSIS BEING A KEY CHALLENGE. TRADITIONALLY, IT IS DIAGNOSED BY MEDICAL PROFESSIONALS USING CHEST X-RAYS. IN THIS STUDY, A COLLECTION OF X-RAY AND CT SCAN IMAGES IS EMPLOYED TO ENABLE AUTOMATED PNEUMONIA DETECTION. AS THE CONDITION PROGRESSES, PATIENTS EXPERIENCE INCREASING DIFFICULTY BREATHING. MACHINE LEARNING METHODS OFFER POTENTIAL FOR FASTER AND MORE ACCURATE DIAGNOSIS BY APPLYING COMPUTER VISION TECHNIQUES FOR AUTOMATIC DETECTION IN MEDICAL IMAGING.

Keywords— ***Deep Learning, RESNET50, CNN, Pneumonia, Neural Network, X-Ray***

I. INTRODUCTION

Throughout history, epidemics and chronic diseases have claimed the lives of countless individuals, leading to significant crises that often took years to overcome. Infectious diseases within communities have been classified as either pandemics or outbreaks over time. A pandemic occurs when there are more cases of illness, injury, or health issues than expected in a particular region or among a specific population within a certain period, with most cases appearing to be linked. In contrast, an outbreak is more localized and typically causes less public concern. One notable pandemic from the past is pneumonia, a serious illness that leads to numerous health complications. Epidemics and chronic illnesses have taken the lives of many throughout history, resulting in major crises that often-required years to resolve. Over time, infectious diseases in communities have been categorized as either pandemics or outbreaks. A pandemic refers to a situation where there is a higher-than-expected number of cases of disease, injury, or health issues within a specific area or group over a certain period, with most cases being interconnected. On the other hand, an outbreak is more localized and generally causes less widespread concern. Pneumonia, a severe disease that causes various health complications, is an example of a significant pandemic from the past.

Machine learning (ML), deep learning (DL), and statistical methods are highly effective tools in disease diagnostic systems. These techniques are capable of addressing complex vision tasks in medical imaging, including lung disease classification, lung segmentation, and more. Recent advancements in DL have not only matched but often surpassed human performance in many tasks. DL can also be applied to predict treatment outcomes, such as in cancer therapy and clinical studies. The use of labeled data combined with DL algorithms has shown promising results in classifying thoracic diseases using X-ray images. Traditionally, deep neural network (DNN) models have been developed and tested by experts through a time-consuming trial-and-error process requiring significant resources and expertise.

To tackle this problem, a new model is proposed that leverages DNN architecture for efficient and accurate classification. This model is specifically designed for the classification and prediction of pneumonia using chest X-ray (CXR) images. The method is based on neural network (NN) architecture, which employs multiple neurons to link, identify, and extract essential features from the image set. While state-of-the-art models exist, NN provides a similarly targeted architecture for training and testing systems, which has been a key factor in its development. The NN model has inspired DL-based algorithms to become a standard approach for predicting and classifying healthcare image datasets.

Deep learning (DL) plays a crucial role in predicting classification outcomes. Various DL models are commonly used in the healthcare sector for disease prediction, but their effectiveness depends on the type of data involved. In this study, the dataset consists of images, making the neural network (NN) model a suitable choice for the experiment. NN models are particularly effective for analyzing data with a consistent grid structure, such as radiographs, as they are inspired by the organization of the human visual system. These models are designed to automatically and flexibly learn spatial hierarchies of features, from basic to complex structures.

The NN architecture typically consists of three hierarchical layers: convolution, pooling, and fully connected layers. The convolution and pooling layers are responsible for feature extraction, while the fully connected layer links the nodes and transfers the extracted features to the output layer to generate classification results. Several NN architectures are available for classification tasks, including Visual Geometry Group (VGG), ResNet, and Google Net, which have demonstrated high accuracy in various fields.

This paper utilizes the RESNET50 transfer learning model for feature extraction due to its flexibility in handling smaller strides and window sizes. RESNET50, with its 50 deep layers, is particularly well-suited for larger datasets, offering better performance. The proposed approach employs this DL-based model to enable quick and early pneumonia detection, using RESNET50 to optimize both accuracy and computational efficiency by keeping the number of layers at 50 to minimize processing time.

II. Present state-of-the-art

This section discusses the use of pre-trained and ensemble models for predicting pneumonia in chest X-ray (CXR) images. Sirazitdinov et al. [5] introduced a hybrid ensemble model for pneumonia prediction, utilizing the Adam optimizer and a batch size of 8. The model was configured with a learning rate of 0.0001, a training-testing ratio of 75:25, and an input image size of 512x512. The results are summarized by the recall, precision, and F1-score values, which were found to be 0.284, 0.288, and 0.286, respectively.

Ahmad et al. [12] employed a deep convolutional neural network (CNN) approach for feature extraction from the chest X-ray (CXR) radiograph dataset. The data were classified based on the Area Under the Curve (AUC), reflecting the severity of the patients' health conditions, with a reported AUC of 0.98. Meanwhile, Rajpurkar et al. [3] developed a 121-layer CheXNet model for pneumonia prediction using CXR data. The results were assessed using the F1-score and compared to an average median F1-score of 0.387, which is significantly lower than the performance achieved by the CheXNet algorithm.

Zech et al. [13] introduced a CNN-based method for diagnosing pneumonia and noted that their model could potentially overstate the accuracy of real-world pneumonia predictions, achieving an AUC of 0.931. Rahimzadeh et al. [14] developed a Deep CNN framework that combines the ResNet50V2 and XceptionNet models, demonstrating a high accuracy of 91.4% when compared to other existing models. Additionally, Ieracitano et al. [15] designed a CNN algorithm specifically for the identification of pneumonia.

This model extracts features from chest X-ray data and fuzzy images. Additionally, it was compared to existing algorithms and achieved an accuracy rate of up to 81%. Zhang et al. [16] examined various AI-based algorithms for pneumonia detection and concluded that Inf-Net algorithms might outperform others after a comprehensive analysis of the current methods. Kundu et al. [17] utilized an ensemble of three CNN frameworks—ResNet, DenseNet, and GoogleNet—to diagnose pneumonia, yielding promising results compared to earlier design approaches. The research utilized two datasets, showing accuracy rates of 87.02% for the Radiological Society of North America (RSNA) dataset and 98.8% for Kermany's dataset, respectively.

Yaseliani et al. [18] introduced an ensemble hybrid deep learning system that integrates support vector machines (SVM), radial basis functions, and logistic regression, utilizing three distinct classification processes. The first step involves a fully connected layer for image categorization, followed by weight adjustments to extract features from the images. Finally, the computer-aided model is applied to classify chest X-ray (CXR) images. Meanwhile, Mabrouk et al. [19] enhanced a deep learning model by merging DenseNet169, MobileNetV2, and Vision Transformer models for pneumonia prediction using CXR images. These three models were used to extract features from the images, which were then employed in the experimental analysis.

III. Materials and Methods

This section outlines the methodology employed for pneumonia prediction in chest X-ray (CXR) images, along with the dataset utilized for the experiment. In this study, ResNet50, a pre-trained feature extractor, is employed for feature extraction ResNet50 is a neural network (NN)-based architecture that garnered significant attention following the 2014 ILSVRC (ImageNet) challenge [1]. It is recognized as one of the most effective architectures for image classification. Notably, it utilizes 3x3 filter convolutions and features the same max pooling and padding layers with a 2x2 filter, avoiding the need for an excessive number of hyper-parameters. The arrangement of convolutional and max pooling layers is consistent throughout the model. Ultimately, the output comprises two fully connected softmax layers [20][21].

A. *Database description*

The first dataset was obtained from Kaggle [22] and is organized into two main directories: a training folder and a testing folder. Each of these directories contains two subdirectories—one for pneumonia X-ray radiographs and another for CXR radiographs of normal lungs. A total of 5,856 anteroposterior CXR images were carefully selected from historical records of pediatric patients aged between 1 to 5 years [20]. Two labels, pneumonia and normal, were used to categorize all the folders, applied to the pneumonia X-ray images as well. The initial data classifications were refined and combined, resulting in the entire dataset being divided into 70% for training and 30% for testing. Consequently, the distribution of images in the dataset includes 5,216 X-ray images for training and 640 radiographs for testing the system. This dataset is publicly available and consists solely of JPEG images.

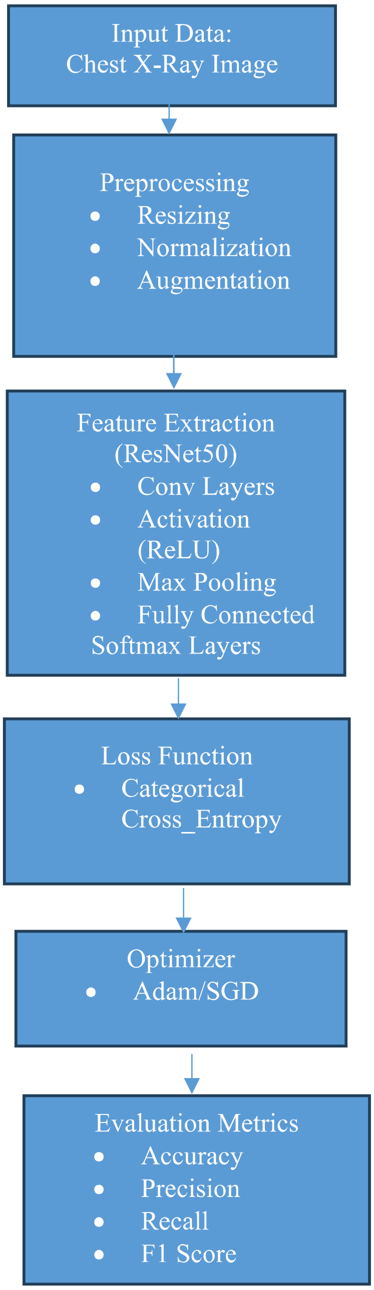

Fig. 1 Proposed deep learning model for pneumonia prediction.

The dataset obtained from Kaggle [24] and is organized into two main directories: training and testing. Each directory includes three subfolders: pneumonia, COVID-19, and normal. In total, the dataset comprises 6,432 JPEG-format radiographs, with 80% allocated for training and 20% for testing. This dataset was sourced from Kaggle. Specifically, the training and testing folders contain 3,418 pneumonia images, 1,266 normal images, and 460 COVID-19 images. In the training folder alone, there are 855 pneumonia images, 317 normal images, and 116 COVID-19 images. The first dataset is utilized for binary classification, while the second dataset is used for multi-class classification.

B. Methodology

The proposed deep learning (DL) model is divided into several components: data collection, preprocessing, feature extraction [25], training, testing, classification, and pneumonia prediction, as illustrated in Fig. 1. Data preprocessing is performed to balance and normalize the dataset, ensuring that the data is adjusted to a normalized range of [0-255]. Following this, the input data is supplied to the ResNet50 model for feature extraction, which effectively extracts relevant features from the images for the classification phase of the prediction process. The ResNet50 utilizes a bottleneck architecture to reduce computational complexity. Each residual block consists of three layers:

- 1x1 Convolution (reduces the number of channels/dimensions)
- 3x3 Convolution (operates on the reduced number of dimensions)
- 1x1 Convolution (restores the original dimensions)

The "50" in ResNet50 refers to the total number of layers, including the convolutional, pooling, and fully connected layers. The network is divided into several stages, each containing multiple residual blocks, with increasing depth as the network progresses. The model is implemented using the Google collab, employing the Adam optimizer with a learning rate of 0.0001 and utilizing the ReLU activation function. The training of the entire model is conducted on the training dataset, comprising 70% of the total data, while 30% is reserved for test data set.

IV. Performance Evalutation

The effectiveness of the proposed model is evaluated through several metrics, such as accuracy, precision, F1-score, and recall, as detailed in Table 1. To understand these metrics, it's important to clarify four key terms: "false positive," "true positive," "false negative," and "true negative."

- A **false positive (FP)** occurs when samples from the negative class are incorrectly classified as belonging to the positive class.
- A **true positive (TP)** refers to samples that are accurately classified as positive and actually belong to the positive class.
- **False negatives (FN)** are samples that belong to the positive class but are incorrectly predicted to be in the negative class.
- A **true negative (TN)** indicates samples that are correctly identified as belonging to the negative class.

The definitions of the metrics utilized for predicting the proposed model's performance are provided below.

$$Acuracy = (TP + TN)/\,(TP + TN + FP + FN) \quad (1)$$

$$Recall = TP/\,(TP + FN) \quad (2)$$

$$Precision = TP/(TP + FP) \quad (3)$$

$$F1 - score = 2/((1/Precision) + (1/Recall)) \quad (4)$$

Accuracy is a metric that reflects the proportion of correct predictions made by the algorithm. However, in the case of unbalanced datasets, a high accuracy rate does not necessarily imply that the model can effectively distinguish between different categories. In the field of healthcare, image classification can be applied to all classes within a dataset. Here, the values of "recall" and "precision" offer insights into the model's performance.

"Precision" measures the accuracy of positive predictions, representing the ratio of correctly predicted positive cases to the total predictions made. On the other hand, "recall" indicates the percentage of true positives that the model successfully identifies. The "F1-score" serves to harmonize "precision" and "recall," considering both false negatives (FNs) and false positives (FPs). It evaluates extreme values of "recall" and "precision," where an increase in one may lead to a decrease in the other.

To accurately differentiate between healthy and diseased patients in diagnostic image identification, it is crucial to examine various evaluation metrics rather than relying solely on accuracy.

Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

A. Results and Discussion

This subsection presents a comprehensive analysis of the results achieved with the proposed model for predicting pneumonia.

- ***NN Classifier with RESNET50:*** *The proposed model demonstrates superior performance in pneumonia prediction, as evidenced by various performance metrics showing minimal error values. Specifically, for dataset, the model achieved an overall accuracy of 50.00%, with a precision of 0.50 and a recall of 1.0, resulting in an F1-score of 0.670, as detailed in Table 1. For dataset 2, the model showed an overall accuracy of 94.5%, along with a precision of 0.954, recall of 0.954, and an F1-score of 0.954.*

TABLE 1. PERFORMANCE OUTCOME OF PROPOSED WORK USING GIVEN DATASET

Sl no	*precision*	*recall*	*f1-score*	*support*
	0.50	1.00	0.67	10
Accuracy	0.50		20	
macro avg	0.25	0.50	0.33	20
weighted avg	0.25	0.50	0.33	20

The results are illustrated in ROC graphs, with Fig. 2 representing the model accuracy and Fig. 3 depicting the model loss, highlighting the model's ability to separate the classes. The ROC curve visualizes the relationship between recall and precision, serving as a graphical evaluation tool for binary classification. In contrast to other methods that provide a single performance value, the ROC curve is widely recognized for identifying effective classifiers; the closer the curve is to the top left corner of the graph, the better the model's performance.

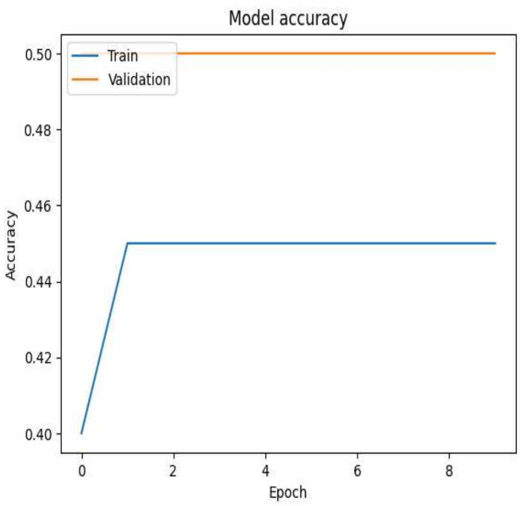

Fig 2: ResNet50 Model Accuracy

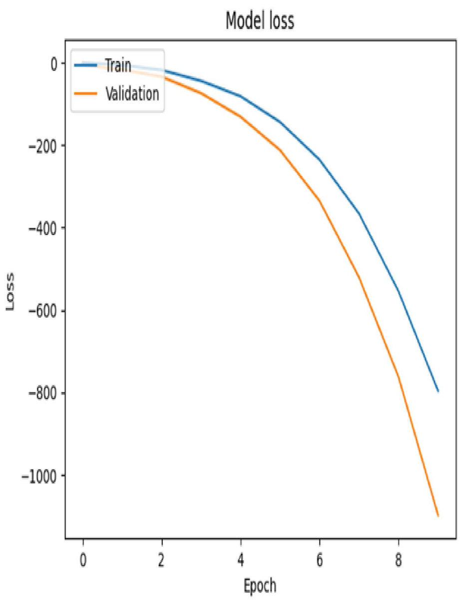

Fig 3: ResNet50 Model Loss

Table2: RESNET50 EVALUATION METRICS OF PRECISION, RECALL, F1-SCORE AND SUPPORT

Models	***Dataset***			
	Accuracy	***Precision***	***Recall***	***F1-score***
NN with RESNET 50	*50.00%*	*0.50*	*1.0*	*0.670*

$$\begin{bmatrix} 234 & 0 \\ 390 & 0 \end{bmatrix}$$

Fig 4: Confusion Matrix

V. CONCLUSION AND FUTURE SCOPE

This study confirms that the NN with RESNET50 model delivers the moderate accuracy and overall performance for pneumonia prediction. In future work, enhancing the model's performance could be achieved by expanding the dataset, incorporating data augmentation, and increasing the number of hidden layers to allow for deeper convolutions in the model, potentially leading to even greater accuracy, particularly for complex cases such as COVID-19 and pneumonia.

References

[1] El Asnaoui K, Chawki Y, Idri A (2021) Automated methods for detection and classification pneumonia based on x-ray images using deep learning. In: Artificial intelligence and blockchain for future cyber security applications, Springer, Cham, pp 257-284.
[2] Sarangi PK, Guleria K, Prasad D, Verma DK (2021) Stock movement prediction using neuro genetic hybrid approach and impact on growth trend due to COVID-19. International Journal of Networking and Virtual Organisations. 25(3-4):333-52.
[3] Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz C, Shpanskaya K, Lungren MP (2017) Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.
[4] Singh S, Tripathi BK(2022) Pneumonia classification using quaternion deep learning. Multimedia Tools and Applications. 81(2):1743-64
[5] Sirazitdinov I, Kholiavchenko M, Mustafaev T, Yixuan Y, Kuleev R, Ibragimov B (2019) Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. Computers & electrical engineering. 78:388-99
[6] Sharma, A., Guleria, K., & Goyal, N. (2021) Prediction of Diabetes Disease Using Machine Learning Model. In: International Conference on Communication, Computing and Electronics Systems, Springer, Singapore, pp 683-692.
[7] Stephen O, Sain M, Maduh UJ, Jeong DU (2019) An efficient deep learning approach to pneumonia classification in healthcare. Journal of healthcare engineering..
[8] Saravagi D, Agrawal S, Saravagi M, Rahman MH (2022) Diagnosis of Lumbar Spondylolisthesis Using a Pruned CNN Model. Computational and Mathematical Methods in Medicine.
[9] Hasan MD, Ahmed S, Abdullah ZM, Monirujjaman Khan M, Anand D, Singh A, AlZain M, Masud M (2021) Deep learning approaches for detecting pneumonia in COVID-19 patients by analyzing chest X-ray images. Mathematical Problems in Engineering.
[10] Gayathri JL, Abraham B, Sujarani MS, Nair MS (2022) A computer-aided diagnosis system for the classification of COVID-19 and non-COVID-19 pneumonia on chest X-ray images by integrating CNN with sparse autoencoder and feed forward neural network. Computers in Biology and Medicine. 141:105134.
[11] Sharma S, Khanra P, Ramkumar KR (2021) Performance Analysis of Biomass Energy using Machine and Deep Learning Approaches. In: Journal of Physics: Conference Series. 2089(1): 012003
[12] Ahmad J, Saudagar AK, Malik KM, Ahmad W, Khan MB, Hasanat MH, AlTameem A, AlKhathami M, Sajjad M (2022) Disease Progression Detection via Deep Sequence Learning of Successive Radiographic Scans. International journal of environmental research and public health. 19(1):480.
[13] Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK (2018) Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. PLoS medicine. 15(11):e1002683.
[14] Rahimzadeh, M., and Attar, A. (2020) A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. Informatics in medicine unlocked. 19:100360.
[15] Ieracitano, C., Mammone, N., Versaci, M., Varone, G., Ali, A. R., Armentano, A., and Morabito, F. C. (2022) A Fuzzy-enhanced Deep Learning Approach for Early Detection of Covid-19 Pneumonia from Portable Chest X-Ray Images. Neurocomputing. 202-215.
[16] Zhang, F. (2021) Application of machine learning in CT images and X-rays of COVID-19 pneumonia. Medicine, 100(36).
[17] Kundu R, Das R, Geem Z.W, Han G.T, and Sarkar R, (2021) Pneumonia detection in chest X-ray images using an ensemble of deep learning models. PloS one. 16(9): e0256630.
[18] M. Yaseliani, AZ. Hamadani, AI. Maghsoodi, and A. Mosavi, (2022) Pneumonia Detection Proposing a Hybrid Deep Convolutional Neural Network Based on Two Parallel Visual Geometry Group Architectures and Machine Learning Classifiers, IEEE Access
[19] A. Mabrouk, RP. Díaz Redondo, A. Dahou, M. Abd Elaziz, M. Kayed, (2022) Pneumonia Detection on Chest X-ray Images Using Ensemble of Deep Convolutional Neural Networks, Applied Sciences, 12(13):6448.
[20] Hammoudi K, Benhabiles H, Melkemi M, Dornaika F, Arganda-Carreras I, Collard D, and Scherpereel A, (2021) Deep learning on chest X-ray images to detect and evaluate pneumonia cases at the era of COVID-19. Journal of medical systems. 45(7):1-10.
[21] Toğaçar M, Ergen B, Cömert Z, and Özyurt F, (2020) A deep feature learning model for pneumonia detection applying a combination of mRMR feature selection and machine learning models. Irbm. 41(4): 212-222.
[22]ChestX-Rayimages(Kaggle). https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia
[23] Alqudah A.M, Qazan S, and Masad I.S, (2021) Artificial intelligence framework for efficient detection and classification of pneumonia using chest radiography images. Journal of Medical and Biological Engineering. 41(5): 599-609
[24] Chest X-ray (Covid-19 & Pneumonia). https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia
[25] A. Kaur, K. Guleria, and N. Kumar Trivedi, (2021) Feature selection in machine learning: Methods and comparison, In: 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp 789-795.
[26] S. Bhushan, M. Alshehri, N. Agarwal N, I. Keshta, J. Rajpurohit, A. Abugabah, (2022) A Novel Approach to Face Pattern Analysis, Electronics, 11(3):444.