Global Scientific JOURNALS

GSJ: Volume 13, Issue 9, September 2025, Online: ISSN 2320-9186
www.globalscientificjournal.com

Smart Pneumonia Diagnosis Model Using Chest X-ray Images: An Ensemble Transfer Learning-Based Approach

Sule Haruna Sani[1a], Njoku Perfect Izuchukwu[1b], Alagbe S. Adebayo[1c], Otubobola Seyi Ayomide[1d], Samuel Edwin Inalegwu[2], Idogho Obaro Princewill[3], Shadi H.E. Ghaleb[4], Ali Shair[5], Matavele Clesia Jamila[6], Onuzuruike Nnenna Favour[7], Ndukwu Samuel Chidera[8], Onuzuruike Obinna Nneji[9], Irabor Jonah[10], Salman Sheini[11], Kulsoom Iqra[12], Hammal Djamel Amine[13], Rachitha Teekshana Siriwardena[14]

1a, c, & d Department of Statistics, Federal University of Agriculture Abeokuta.

1b, 4, 6 & 14 Department: Department of Public Health and Evidence Base Medicine, I.M. Sechenov First Moscow State Medical University.

2 Department of Medical Laboratory Sciences, University of Jos.

3, 7, & 10 Department of Business Administration in Public Health Care, I.M. Sechenov First Moscow State Medical University.

5 Department of pharmacology, I.M. Sechenov First Moscow State Medical University.

8 Department of Radiology, National Orthopedic Hospital, Enugu, Nigeria.

9 Department of Medicine and Surgery, University of Nigeria Enugu Campus, Enugu, Nigeria.

11 Department of Healthcare Administration, I.M. Sechenov First Moscow State Medical University.

12 Department of Public health, Organisation and Sociology of healthcare, Medical- Social Expertise, Chechen State Medical University, Russia.

13 Department of Biotechnology, School of Pharmacy, I.M. Sechenov First Moscow State Medical University.

Abstract

Pneumonia is a major infectious disease that affects the lungs, causing inflammation among under age children and elderly people. This creates a need for optimize medical diagnosis as Traditional approaches that includes image studying, clinical evaluation and examination are not proactive and efficient enough for early detection of this disease. Artificial intelligence and Deep learning algorithms have been proven by researchers to be more robust for tackling this challenge by using a Convolution Neural Networks and Transfer learning models in predicting medical diagnosis outcome with high precision and accuracy. This study proposed an ensemble learning based framework that integrates the predictive power of pretrained base learners like VGG16 and

DenseNet121 in predicting Pneumonia using Chest X-ray images using a weight-based strategy. The study used public dataset that consists of 5856 Chest X-ray images that subdivided into Pneumonia (3883) and Normal (1349) Classes. The Chest X-ray dataset was preprocessed using various strategies that includes image scaling, data-augmentation, class weight allocation to tackle class-imbalance. The VGG16 and DenseNet121 was trained and evaluated using accuracy, precision, recall, f1-score, and Area Under the Curve (AUC) metrics. The study results showed that the base learners attained accuracies of 94.55% and 93.75% respectively. The metrics of these based learners was used to developed weighted strategy using a hyperbolic tangent function, to improve the accuracy of the ensemble learning model. The results shows that the ensemble learning model outperforms the base learners with an accuracy of 95.19%, precision of 96.15%, recall of 96.15%, f1-score of 96.15%, and an AUC of 98.58%. The ensemble model also showed minimal misclassification error, showing the model capacity of attaining higher precision and accuracy over the individual base learners.

Keyword(s): Pneumonia Disease, Transfer learning, Ensemble learning Models

1.0 Introduction

Pneumonia can be described as an acute infectious disease that triggers inflammation in the lungs, and it has been reported to be a major cause of morbidity and mortality for under-age children and the elderly population (Rudan et al. 2013; Theodoratou et al. 2014; Bakare et al. 2020; Reddy et al. 2025). Addressing this infectious disease using traditional diagnostic methods involves a combination of physical clinical examination, imaging studies, and computed tomography (Wankhede et al., 2024; Ren et al., 2025). These methods are time-consuming and resource-intensive and are not free from possible human error. Artificial Intelligence algorithms have been used as a transformative tool for bridging the gap for precise detection and diagnosis of Pneumonia (Chumbita et al. 2020; Wankhede et al. 2024). Deep learning algorithms, in particular Convolutional Neural Networks and Pretrained models, have shown valuable performance in terms of high accuracy and precision in detecting Pneumonia using an image dataset. However, these algorithms often face challenges like overfitting and model performance bias due to class imbalance. To address this gap, ensemble learning strategies were explored using the predictive power of multiple base learners in improving their accuracy and robustness. Several studies have shown the superiority of ensemble learning frameworks in outperforming both custom CNN and Pretrained Models in image classification tasks for medical diagnostics (Anitha et al. 2021; Seung et al. 2024). This research employs a novel ensemble approach that incorporates a customized weight strategy based on learners, specifically DenseNet121 and VGG16, for pneumonia diagnosis using Chest X-ray images. The weighted strategy evaluates the base learners' performance metrics as a yardstick for allocating weight penalties for improving the ensemble learning prediction. The main aim of the proposed system is to show the effectiveness of ensemble learning in achieving accurate and precise diagnosis for Pneumonia using Chest X-ray images.

The remaining part of this study includes Section 2.0, which discusses recent studies that used deep learning pretrained CNN models in detecting pneumonia, Section 3.0 focuses on the research methods and algorithms used in delivering the study objectives, Section 4.0 discusses the result findings, and Section 5.0 highlights the study conclusion.

2.0 Literature Review

This section of the study examines recent studies that applied both Traditional Machine learning and Deep learning-based models, alongside transfer learning and Ensemble models in predicting patients' Pneumonia diagnosis status using X-ray images. Seung et al. (2024) proposed an ensemble-based learning approach for Pneumonia diagnosis using a Machine learning-based approach using models such as XGBOOST, SVM, CatBoost, LGBM, KNN, RF, and MLP. The study found that the ensemble algorithm was the best model with an accuracy score of 90%. Dubbaka et al. (2023) utilized a deep learning Convolutional neural network model in predicting patients' Pneumonia diagnosis status. The authors trained a CNN over 3,000 X-ray chest images, and it attained an accuracy score of about 95%. Karan J. & Nishtha H. (2018) proposed a Big Data Deep Learning Framework for Pneumonia diagnosis using several Machine learning algorithms, which include RF, Adaboost, SVM, Logistic Regression, Naïve Bayes, DT, and Neural Network. The study results reveal that CNN attained the highest accuracy score of 84% indicating the proficiency of these models in handling image classification tasks. RF, DT, Adaboost, Logistic Regression, and Naïve Bayes all had an accuracy of 82%, 77%, 78%, 77%, 72% respectively. Shagun & Kalpna (2023) utilize a hybrid deep learning approach integrating Neural network architecture with Pretrained Networks like VGG16 for predicting Pneumonia. The study evaluated several combinations of CNN with VGG16, SVM with VGG16, KNN with VGG16, RF with VGG16, and NB with VGG16, and it attained an accuracy score of 92.15%, 91.5%, 91%, 87.2%, and 84.8% respectively. Rachna et al. (2020) applied four (VGG16, VGG19, ResNet50, and Inception-v3) transfer-learning models for Pneumonia Detection in chest X-ray images. The study results review these models had an accuracy score of 87.28%, 88.46%, 77.56% and 70.99% respectively, on the test X-ray images. Uzair et al. (2020) used a VGG16 CNN architecture for Predicting Pneumonia from Chest X-rays. The study results review that the model attained a classification accuracy of 96.6%, sensitivity of 98.1%, specificity of 92.4%, precision of 97.2%, and a F1 Score of 97.6%. Veeranjaneyulu et al. (2020) used a CNN model for the identification and Classification of patients' Pneumonia, and it attained an accuracy score of 97.73% for the binary classification case and 91.17% for a multi-class task. Halit et al. (2023) proposed a hybrid-based approach, which involves the combination of these models: ANN, RestNet, InceptionV3, and MobileNet. The study explored the combination of these models in predicting Pneumonia diagnosis for both binary and multi-class related tasks, and it observed a maximum accuracy score of (95.67% MobileNet-ANN) and 81.57% (ResNet-ANN).

3.0 Methodology

3.1 Dataset Description

The image dataset used in this research was extracted from the study by Kermany and Goldbaum (2018), and it contains Chest X-ray images of pediatric patients aged one to five years. The dataset consists of 5856 X-ray Chest images, which are subdivided into two diagnosis classes: Pneumonia and Normal. The training dataset consists of 3883 pneumonia cases and 1349 normal cases, while the test dataset contains 234 normal cases and 390 Pneumonia cases.

Table 1.0: Dataset summary

Class	Training Dataset	Testing dataset
Normal	1349	234
Pneumonia	3883	390
Total	5232	624

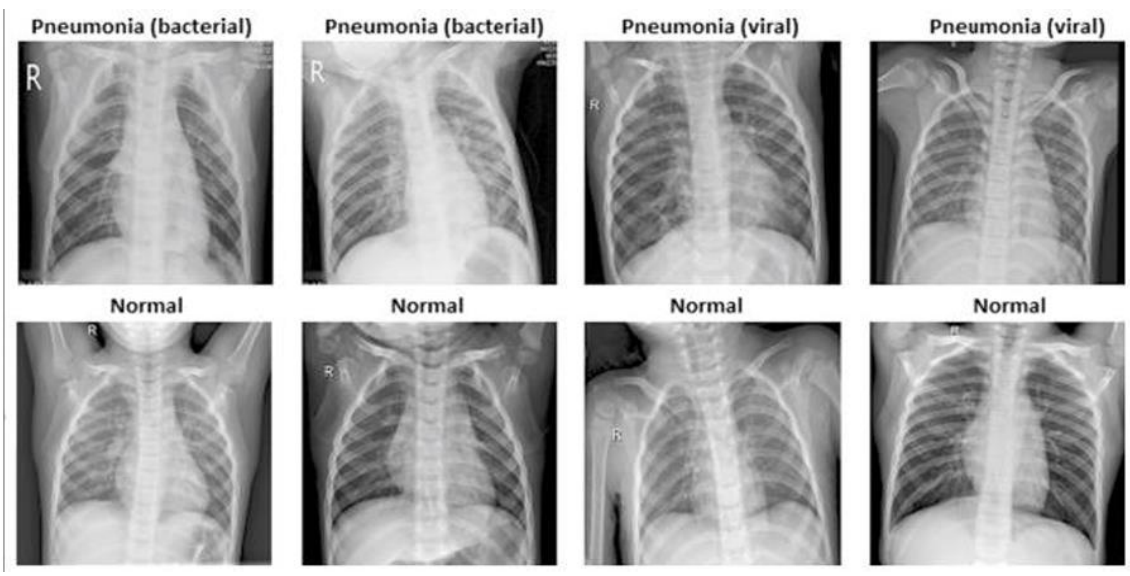

Figure 1.0: Chest-Xray Images

3.2 Data Preprocessing

The Data preprocessing technique employed in this study involves Image scaling, data augmentation, and Data splitting for Model training. The train and test Chest X-ray Images were imported using the Python *ImageDataGenerator()* method, and the train and test image pixels were normalized between [0,1]. Data augmentation was performed on both the train and validation (20% of the train dataset) images. This includes rotating the images randomly by 150, shifting the images vertically/horizontally by 10%, applying a random zoom in/out of 10%, and flipping the image horizontally. After creating the Data-augmentation instance, the train, validation, and test images were converted to a Data generator. For the train and validation dataset, the image target size was placed as (224, 224), batch size per iteration was 128, the class mode = binary, and the images were shuffled. A similar procedure was initiated for the test dataset; however, the shuffle parameter was set as False to ensure alignment between the predicted images and true images.

3.3 Transfer Learning Models

3.3.1 DenseNet121

The DenseNet architecture is based on a concatenated structure that connects the preceding layers of the network using feature maps. This structure supports feature reuse, a strong flow of gradient, and reduced need for more parameters (Kumar, R. 2020; Huang et al. 2017). DenseNet121, which is a variant of the DenseNet models, is a pretrained Convolution neural model that consists of 121 layers. This variant has been reported to be efficient in Glaucoma Detection, Skin Cancer Detection, and Pneumonia Detection (Chakraborty et al. 2024; Hamsalekha R. 2024; Gundabatini, 2024).

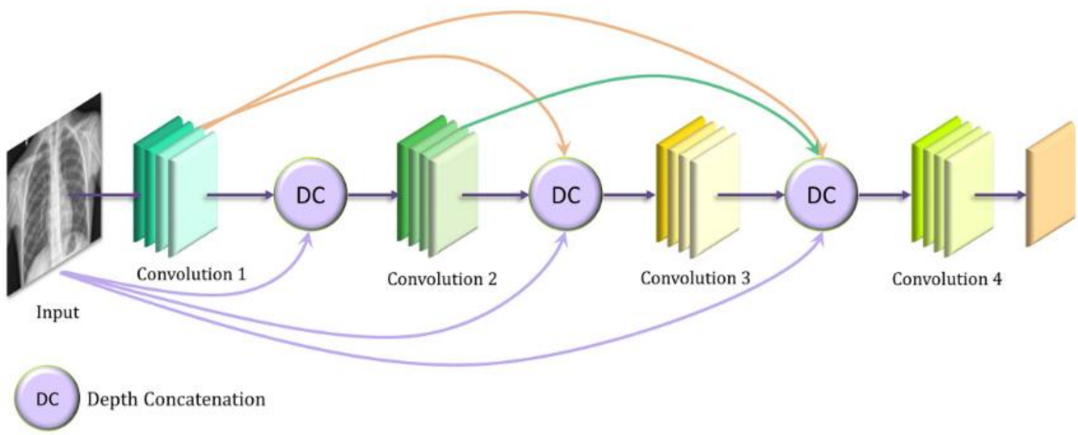

Figure 2.0: DenseNet Architecture

3.3.2 VGG16

VGG16 is a pretrained Deep learning convolutional neural Network made of 16 layers with learnable weights, with a fixed input size of (224 X 224) pixels. The network architecture is made of filters (3 by 3) with a stride parameter of 1, a ReLU activation, and a Max-pooling of stride 2 (Vijay et al. 2017). The network-stacked architecture setup enables the model to detect complex patterns during image classification. VGG16 has been reported to be proficient in medical image classification, which includes brain tumor detection, lung cancer using CT scans, and kidney disease (Khan & Rafath, 2024; Mukesh et al., 2024**;** Rohith et al., 2023).

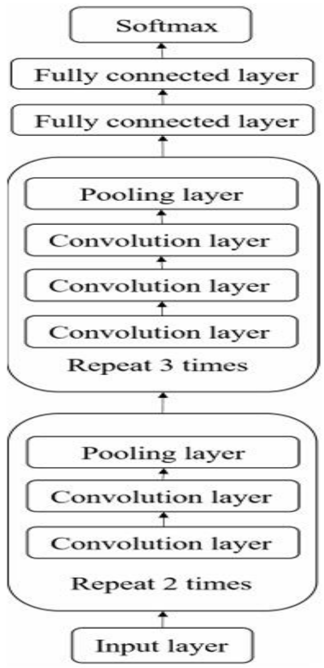

Figure 3.0: VGG architecture

3.4 Proposed Ensemble Model

The ensemble learning approach used in this research integrates the probability predictions of the diagnosis classes (Pneumonia vs Normal) based on learners (DesNet121 & VGG16) using a weighted average strategy. This strategy has been described as an efficient approach when dealing with an imbalanced dataset (Kundu et al., 2021). The implementation process of the ensemble learning model is described in the section below.

Stage I: The probability scores of DesNet121 and VGG16 are extracted to estimate the weight allocation for the base learners.

Stage II: The prediction of the base models is compared to the true label (Diagnosis Class) to estimate the model accuracy, precision, recall, f1-score, and AUC. This will be extracted using a custom Python function, and the base learners' performance metrics are stored in an Array similar to this $A^i = [acc^i, prec^i, recall^i, f1^i, auc^i]$.

Stage III: The weight assigned to each base classifier is computed using the hyperbolic tangent function as stated in Equation 1 below.

$$w^{(i)} = \sum_{x \epsilon A^{(i)}} \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad \text{I}$$

Where x represents the base learner performance metrics with range between [0,1] and the range of the tangent function is [0, 0.762]. The function which is monotonic nature assigns more priority to the base learner with a higher metric.

Stage IV: This computed weight $w^{(i)}$ is multiplied by the corresponding predicted probability scores by base learners using Equation II stated below.

$$\text{ensemble\_prob}_j = \frac{\sum_i w^{(i)} p_j^i}{\sum_i w^{(i)}} \quad \text{II}$$

Where p_j^i = {a, 1-a}, and a ≤ 1 and the ensemble probability for the sample is $ensemble\_prob_j$. The ensemble model makes a prediction using Equation III stated below.

$$\text{prediction}_j = \text{argmax}\left(\text{ensemble}_{\text{prob}_j}\right) \quad .\text{III}$$

3.5 Model Evaluation Metrics

This section discusses the evaluation metrics used to assess the proficiency of both the transfer learning models and the ensemble models. It includes the accuracy, precision, recall, F1-score, and Area under the curve. These metrics are discussed in detail in the section below.

3.5.1 Accuracy

Accuracy can be defined as the ratio of correct predictions to the total number of instances in the dataset.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

3.5.2 Recall

Recall can be defined as ratio correct positive instances captured to the total number of positive instances prediction (Lydia & Patricia, 2024).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

3.5.3 Precision

Precision can be defined as the ratio of positive prediction to the total number of positive instances in the dataset.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

3.5.4 F1-score

F1-score can be defined as a harmonic mean that place a balance between precision and recall particularly in cases of class-imbalance.

$$\text{f1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

3.5.5 Area Under Curve

AUC measures the model's ability to distinguish between the positive and negative instances efficiently across different classification thresholds (Lobo et al., 2007). An AUC value below ranges from [0,1] and a value closer to 1 indicates that the model attains a perfect discrimination between the predicted classes.

4.0 Result Findings

This section of this study discusses the performance metrics of the transfer-learning Convolution Neural Networks and Ensemble model on the test Chest X-ray images relating to Pneumonia diagnosis. These models were all trained using a single batch size of 128 and a learning rate of 0.0001, which was increased after every 3 epochs. The study utilized a ReLU activation for all the transfer learning architectures, including the Ensemble model, and an output function of sigmoid. The sigmoid function was considered because our prediction task is a binary classification task. A binary cross-entropy loss function was used to measure the training loss of the models.

Table 3.0: Comparative Model performance on the test set

Model	Accuracy	Precision	Recall	F1-score	AUC
DenseNet121	94.55%	94.95%	96.41%	95.67%	98.61%
VGG16	93.75%	94.21%	95.90%	95.04%	98.07%
Ensemble Model	95.19%	96.15%	96.15%	96.15%	98.58%

Table 4.0 shows the transfer learning performance on the unseen X-ray images, and the results indicate that the Ensemble Model attained the highest classification accuracy score of 95.19%, with a precision of 96.15%, recall of 96.15%, f1-score of 96.15% and the Area under the curve was 98.58%. This result indicates that the Ensemble model generalized well over the unseen images in predicting the patients' Pneumonia status using the Chest X-ray images. DenseNet121 had an accuracy score of 94.55% with a precision of 94.95%, recall of 96.41%, f1-score of 95.67% and an AUC of 98.61%. The VGG16 model showed a lower accuracy score of 93.75% with an AUC of 98.07%. Figure 5.0 shows the Transfer learning models' confusion matrix, and it shows that for DenseNet121, there is only 3.2% misclassification for patients without Pneumonia and 2.2% for patients with the disease. VGG16 showed a misclassification rate of 3.7% for patients without Pneumonia and 2.6% for patients with the disease. This model happens to show a higher misclassification rate among the proposed models in this study. The Ensemble transfer learning model showed the lowest misclassification rate, with 2.4% for patients with Pneumonia and 2.4% for patients without the disease. The ensemble transfer learning model AUC (98.58%) shows that the model is robust for distinguishing between the two diagnosis classes.

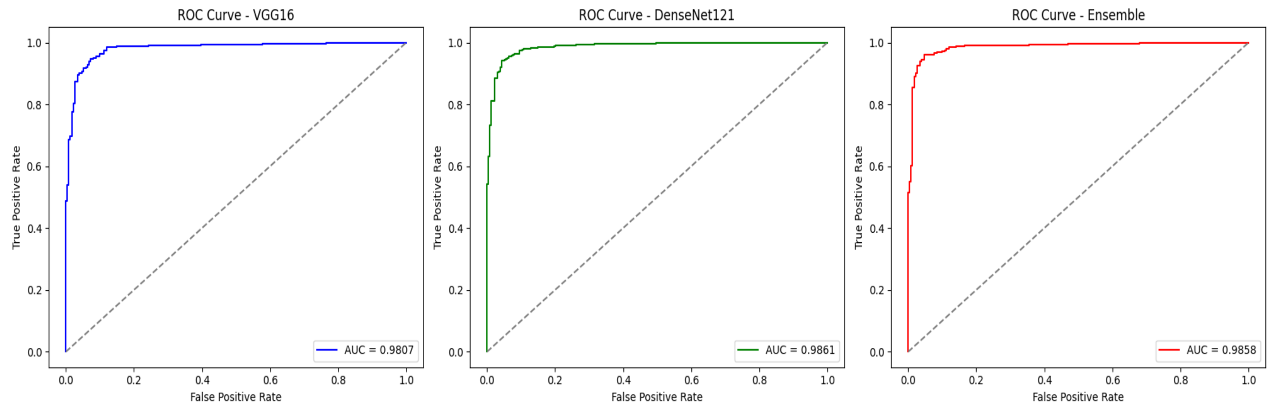

Figure 4.0: ROC Curve of the model

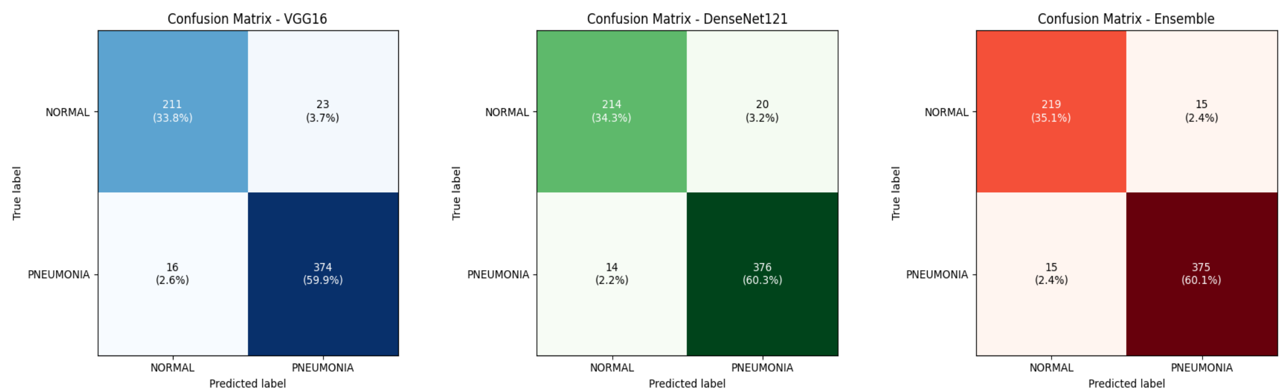

Figure 5.0: Model Confusion Matrix

Table 4.0: Summary report from recent studies that used Transfer learning for Pneumonia detection

Author(s)	**DL algorithm(s)**	**Performance**
Ayan et al. (2019)	Xception and VGG16	VGGI6 accuracy: 87%, Xception accuracy: 82%
Yakur et al. (2021)	VGG16	Accuracy, recall, precision and f1-score 90.54%, 98.7%, 87%, and 92.9% respectively.
Jain et al. (2020)	VGG19, ResNet50, VGG16, two customized models and InceptionV3	Accuracy of VGG19: 88.46%,
Knok et al. (2019)	CNN	Accuracy: 94%

Zhi-Peng Jiang et al. (2021)	LeNet AlexNet, GoogLeNet and VGG16	Best Model Google Net with an accuracy score of 89.5%, recall of 89.6%, and precision of 89.3%
Md. Maniruzzaman et al. (2024)	VGG-16, VGG-19, ResNet-50, Inception-V3, and Xception	Best Model Xception with an accuracy of 95.06%, AUC of 94.23%, recall of 97.43%, specificity of 91.02%.
Alam et al. (2023)	RVCNet	Accuracy 91.27%, Recall of 98.30% and Specificity of 90.48%
Proposed method	Ensemble Model (VGG16+DenseNet121)	Accuracy 95.19%, Recall 96.15%, Precision 96.15%, f1-score 96.15%, AUC 98.58%.

Table 4.0 shows the comparison of the proposed ensemble model with recent algorithms utilized in the context of predicting Pneumonia. It should be noted that our proposed method, which integrates the use of VGG16 and DenseNet121, has base learners that outperform the other methods in terms of accuracy (95.19%), recall (96.15%), precision (96.15%), and AUC (98.58%). These other researchers applied a single CNN architecture in detecting the Pneumonia diagnosis status using Chest X-ray images (Ayan et al., 2019; Yakur et al., 2021; Jain et al., 2020; Knok et al., 2019; Zhi-Peng Jiang et al., 2021; Md. Maniruzzaman et al., 2024; Alam et al., 2024), and the results show that Ensemble models have potential in accurately optimizing Pneumonia diagnosis.

5.0 Conclusion

This study focuses on the potential of Ensemble learning models using a weighted strategy backed by base convolution learners in detecting pediatric pneumonia using chest X-ray image datasets. This proposed system is designed to mitigate the gap in early detection of Pneumonia using optimized algorithms such as Ensemble learning strategies. This ensemble framework developed in this study considers the predicted probabilities from two base learners, which are VGG16 and DenseNet121, to form a weighted average ensemble strategy. This strategy utilized the performance metrics of the base learners, which include accuracy, precision, f1-score, recall, and AUC, in allocating these weight penalties using a hyperbolic tangent function. This proposed method yielded an accuracy of 95.19% and a recall of 96.15%, an F1-score of 96.15%, and an AUC of 98.58%, showing the ability of ensemble learning models to attain better generalization over unseen Chest-X-ray images for Pneumonia diagnosis. Compared with recent studies that used different transfer learning algorithms, we observed that our proposed method has an optimal and robust solution for clinicians for handling Pneumonia diagnostics using chest X-ray images. We recommend future study to explore the model generalization for more diverse Big-dataset, multiclassification use-cases, and also consider the real-time deployment of this model to optimize Patient diagnostics in real-time.

6.0 References

Kumar, R. (2020). Adding Binary Search Connections to Improve DenseNet Performance. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3545071

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017, July). Densely Connected Convolutional Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/cvpr.2017.243

Kundu R, Das R, Geem ZW, Han G-T, Sarkar R (2021) Pneumonia detection in chest X ray images using an ensemble of deep learning models. PLoS ONE 16(9): e0256630. https://doi. org/10.1371/journal.pone.0256630

D.S. Kermany, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell vol. 172 (5) (Feb. 2018) 1122–1131.e9, https:// doi.org/10.1016/j.cell.2018.02.010.

E. Ayan, H.M. Ünver, Diagnosis of pneumonia from chest x-ray images using deep learning, Sci. Meet. Electr.-Electron. Biomed. Eng. Comput. Sci. (EBBT) (2019) 1–5.

Y. and A. S and U. R and S.G. Thakur Saurabh and Goplani, Chest X-Ray Images Based Automated Detection of Pneumonia Using Transfer Learning and CNN, in Proceedings of International Conference on Artificial Intelligence and Applications, M. and B. V. E. and S.R. Bansal Poonam and Tushir, Ed., Singapore: Springer Singapore, 2021, pp. 329–335.

R. Jain, P. Nagrath, G. Kataria, V. Sirish Kaushik, D. Jude Hemanth, Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning, Measurement vol. 165 (Dec. 2020) 108046, https://doi.org/10.1016/J. MEASUREMENT.2020.108046.

Z. Knok, K. Pap, M. Hrnˇ ci´ c, Implementation of intelligent model for pneumonia detection, Teh. ˇ cki Glas. vol. 13 (4) (Dec. 2019) 315–322, https://doi.org/ 10.31803/tg-20191023102807.

R. Siddiqi, 2019, Automated Pneumonia Diagnosis Using a Customized Sequential Convolutional Neural Network, in Proceedings of the 2019 3rd International Conference on Deep Learning Technologies, in ICDLT 2019. New York, NY, USA: Association for Computing Machinery, 2019, pp. 64–70. doi: 10.1145/ 3342999.3343001.

Dubbaka Megha Sai Reddy, Chigullapalli Prem Sai, Gopu Joseph Rahul Reddy, Mr.Shaik Khasim Saheb (2023). Predicting Pneumonia Using on Chest X-ray with Deep Learning. Retrieved from: https://ijrar.org/papers/IJRAR23B3264.pdf

Uzair SHAH a, Alaa ABD-ALRAZEQ a, Tanvir ALAM a, Mowafa HOUSEH a and Zubair SHAH (2020). An Efficient Method to Predict Pneumonia from Chest X-Rays Using Deep Learning Approach. Retrieved from: https://www.researchgate.net/profile/Tanvir-Alam-2/publication/342639390_An_Efficient_Method_to_Predict_Pneumonia_from_Chest_X-

Rays_Using_Deep_Learning_Approach/links/5fd77bd3299bf140880aa790/An-Efficient-Method-to-Predict-Pneumonia-from-Chest-X-Rays-Using-Deep-Learning-Approach.pdf

Rohith, S., Prakash, M. S., Anitha, R., Kumar, K. S., & Yogeswara Sai, K. (2023). Detection of Brain Tumor using VGG16. *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, 1400–1405. https://doi.org/10.1109/icces57224.2023.10192639

Karan Jakhar & Nishtha Hooda (2018). Big Data Deep Learning Framework using Keras: A Case Study of Pneumonia Prediction. Retrieved from: https://www.researchgate.net/profile/Nishtha-Hooda/publication/329705983_Big_Data_Deep_Learning_Framework_using_Keras_A_Case_Study_of_Pneumonia_Prediction/links/5c17380f92851c39ebf2e737/Big-Data-Deep-Learning-Framework-using-Keras-A-Case-Study-of-Pneumonia-Prediction.pdf

Shagun S. & Kalpna G. (2023). A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks.

Veeranjaneyulu Naralasetti1, Reshmi Khadherbhi Shaik2, Gayatri Katepalli1, Jyostna Devi Bodapati (2020). Deep Learning Models for Pneumonia Identification and Classification Based on X-Ray Images.

Halit Bakir, Semih Oktay, Timuçin Emre Tabaru (2023). DETECTION OF PNEUMONIA FROM X-RAY IMAGES USING DEEP LEARNING TECHNIQUES. Retrieved from: https://dergipark.org.tr/en/download/article-file/2833235

Rachna Jain, Preeti Nagrath, Gaurav Kataria, V. Sirish Kaushik and D. Jude Hemanth (2020). Pneumonia Detection in chest X-ray images using Convolutional Neural Networks and Transfer Learning.

Lydia Fischer and Patricia Wollstadt (2024). Precision and Recall Reject Curves. Retrieved from: https://arxiv.org/pdf/2308.08381

Jude Chukwura Obi (2024). A comparative study of several classification metrics and their performances on data

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2007). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, *17*(2), 145–151. https://doi.org/10.1111/j.1466-8238.2007.00358.x

Vijay Badrinarayanan, Alex Kendall , and Roberto Cipolla (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. Retrieved from: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7803544

Mukesh Kumar, C., & Sree Sankar, J. (2024). Comparative Analysis of Convolutional Neural Networks for Brain Tumor Detection: A Study of VGG16, ResNet, Inception, and DenseNet Models. *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 41–46. https://doi.org/10.1109/icaaic60222.2024.10575770

Chakraborty, S., Roy, A., Pramanik, P., Valenkova, D., & Sarkar, R. (2024). *A Dual Attention-aided DenseNet-121 for Classification of Glaucoma from Fundus Images* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2406.15113

Hamsalekha R. (2024). Fusion-Based Deep Learning Approach for Skin Cancer Detection Using DenseNet-121 and Transfer Learning. *Journal of Electrical Systems*, *20*(11s), 3182–3188. https://doi.org/10.52783/jes.8044

Gundabatini, S. G. (2024). Pneumonia Detection using CNN, Resnet and DenseNet. *International Journal for Research in Applied Science and Engineering Technology*, *12*(3), 1680–1686. https://doi.org/10.22214/ijraset.2024.59145

Reddy, R., Natarajan, E., Priya, S., & Asif, M. (2025). Next-Gen Pneumonia Detection: A Review of Secure AI, CNN Models, and Federated Learning. *International Journal For Multidisciplinary Research*, *7*(3). https://doi.org/10.36948/ijfmr.2025.v07i03.45905

Theodoratou, E., McAllister, D. A., Reed, C., Adeloye, D. O., Rudan, I., Muhe, L. M., Madhi, S. A., Campbell, H., & Nair, H. (2014). Global, regional, and national estimates of pneumonia burden in HIV-infected children in 2010: a meta-analysis and modelling study. *The Lancet Infectious Diseases*, *14*(12), 1250–1258. https://doi.org/10.1016/s1473-3099(14)70990-9

Rudan, I., O'brien, K. L., Nair, H., Liu, L., Theodoratou, E., Qazi, S., Lukšić, I., Walker, C. L. F., Black, R. E., Campbell, H., & Group, C. H. E. R. (2013). Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. In *Journal of global health 3 (1), 2013*.

Bakare, O. O., Fadaka, A. O., Klein, A., Keyster, M., & Pretorius, A. (2020). Diagnostic approaches of pneumonia for commercial-scale biomedical applications: an overview. *All Life*, *13*(1), 532–547. https://doi.org/10.1080/26895293.2020.1826363

Wankhede, S. S., Adhale, M. P., Hiwase, S. T., Khedikar, S. D., & Vaitage, T. S. (2024). Pneumonia Diagnosis and Analysis System with Machine Learning. In *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*.

Ren, J., & Mariano, V. (2025). A neural network for disease recognition of radiological images of pneumonia. *International Journal of Security and Networks*, *20*(1), 23–31. https://doi.org/10.1504/ijsn.2025.145034

Seung Min Baik, Kyung Sook Hong, Jae-Myeong Lee, Dong Jin Park (2024). Integrating ensemble and machine learning models for early prediction of pneumonia mortality using laboratory tests. https://doi.org/10.1016/j.heliyon.2024.e34525