Pneumonia Detection from Chest X-ray using Neural Networks

Ankita Bandal
bandal.a@northeastern.edu

Abhishek Sinha
sinha.ab@northeastern.edu

Vivekbhai Ashokbhai Radadiya
radadiya.vi@northeastern.edu

***Abstract*—This study builds a neural network-based system that detects pneumonia in chest X-ray images. The project applies and assesses several architectures, such as CNN, ResNet, VGGNet, and Inception, to categorize X-ray images as either normal or pneumonia-infected using convolutional neural networks (CNNs) and transfer learning. The models are trained and validated on a dataset of 5,856 labeled chest X-ray images. ResNet is the recommended model for automated pneumonia diagnosis because it had the best validation accuracy (94.01%) and the best recall among the architectures examined. The study emphasizes how sophisticated neural networks can help medical professionals detect diseases accurately and efficiently.**

I. INTRODUCTION

A. Overview

The goal of this project is to create a reliable neural network-based system that can identify chest X-ray images as either normal or infected with pneumonia, assisting in prompt and precise diagnosis. Millions of people worldwide suffer from pneumonia, a serious respiratory infection that needs to be diagnosed quickly to lower morbidity and mortality. Neural networks, especially Convolutional Neural Networks (CNN), have become extremely effective medical imaging tools due to deep learning advancements, allowing for accurate disease detection and classification.

B. Motivation

Pneumonia remains a major global health concern, causing significant morbidity and mortality, especially among children and the elderly. Early and accurate diagnosis is critical for effective treatment, yet many healthcare systems face challenges such as limited access to radiologists, high patient volumes, and resource constraints. Traditional diagnosis relies heavily on manual interpretation of chest X-rays, which is time-consuming, prone to human error, and inconsistent in low-resource settings.

Advancements in artificial intelligence and deep learning offer a transformative solution, providing automated, efficient, and reliable disease detection. Leveraging neural networks for pneumonia classification addresses these challenges by reducing diagnostic workloads, improving accuracy, and increasing accessibility in underdeveloped regions. This project is motivated by the potential of AI to bridge healthcare gaps, enhance diagnostic reliability, and contribute to better patient outcomes, especially in resource-limited areas.

C. Approach

The project follows a systematic approach to develop an automated pneumonia detection system. It starts with preparing the dataset by resizing images to consistent dimensions and applying techniques like rotation, flipping, and zooming to make the model more adaptable and avoid overfitting. Various CNN architectures, including ResNet, VGGNet, and Inception, are explored to identify the most effective model for this task. Pre-trained weights through transfer learning are used to build on existing knowledge, speeding up training and enhancing accuracy for medical image classification.

Models are trained using binary cross-entropy loss, with optimizers like Adam or RMSprop to ensure effective learning. To avoid overtraining, early stopping, and model checkpointing are applied. Performance is assessed by splitting data into training and testing sets, and key metrics like accuracy, precision, and recall are measured. The model with the best validation results is chosen to deliver a reliable and scalable solution for pneumonia detection.

Figure 1 illustrates the workflow of the proposed methodology. The dataset utilized in this project is sourced from Mendeley Data, comprising 5,856 annotated chest X-ray images categorized into "Pneumonia" and "Normal" classes. Dataset Link: https://data.mendeley.com/datasets/rscbjbr9sj/3

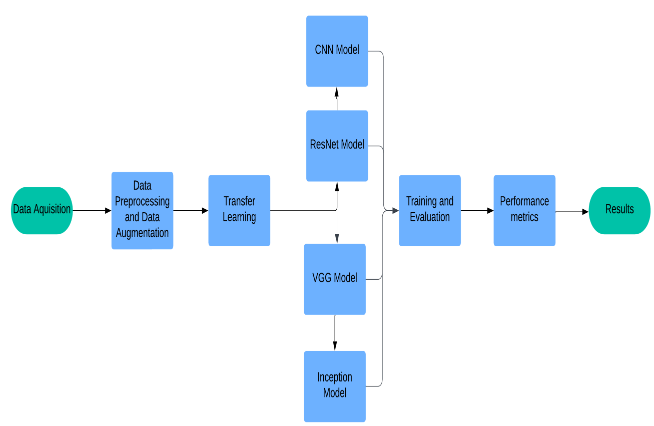


Fig. 1: Pneumonia detection Methodology

II. BACKGROUND

In recent years, advancements in deep learning techniques have driven significant progress in medical imaging, including pneumonia diagnosis. Jain et al. [1] leveraged convolutional neural networks (CNNs) with transfer learning to enhance accuracy in detecting pneumonia from chest X-rays. By utilizing pre-trained models, they achieved better generalization across diverse datasets while reducing training costs. However, the study noted the potential for further optimization of architectures and parameters, which remains an open area for exploration.

Recent developments have also highlighted the robustness of transfer learning in handling complex medical imaging tasks. Rahman et al. [2] used pre-trained models like VGG16 and ResNet50 to classify pneumonia cases effectively. Their approach tackled challenges like dataset imbalance through transfer learning but revealed the risk of overfitting with limited data. This insight emphasizes the importance of strategies like data augmentation, which the current project incorporates.

Custom CNN architectures have also shown promise, as demonstrated by Zhang et al. [3]. Their work achieved a balance between accuracy and efficiency by designing a tailored CNN model. However, the study's reliance on large datasets for training limits its flexibility in other imaging contexts. This reinforces the need for adaptable models, a focus of this project as it evaluates architectures like ResNet and VGG.

Lastly, Bhatt et al.[4] and Shah et al. [4] explored ensemble CNN models, combining multiple architectures to improve pneumonia detection. Their approach enhanced recall and minimized false negatives, making it more reliable for clinical use. Despite these gains, the added complexity and computational demands pose challenges. The current study addresses these trade-offs by comparing the efficiency of single and ensemble models to develop a robust solution.

III. APPROACH

A. Data

The dataset for this project was taken from the Mendeley Data repository and contained a total of 5,856 chest X-ray images, divided into two categories: "Pneumonia" and "Normal". The "Pneumonia" class had 4,273 images, while the "Normal" class had only 1,583, resulting in a significant class imbalance. To address this issue and ensure fair representation, a subset of 1,583 images was randomly chosen from the "Pneumonia" class, making the dataset balanced with an equal number of images for both classes. Balancing was crucial to prevent the model from being biased toward the majority class.

The dataset was then split into training and testing sets in an 80:20 ratio. The training data was further divided into training and validation sets using a validation split during augmentation. This allowed us to ensure that the model was trained on diverse data while also validating its performance on unseen examples during training. Each image was labeled appropriately, and the entire dataset was prepared for input into the neural network models. Fig. 2 represents Chest X-ray images from the dataset, showcasing both Normal and Pneumonia infected X-Rays.

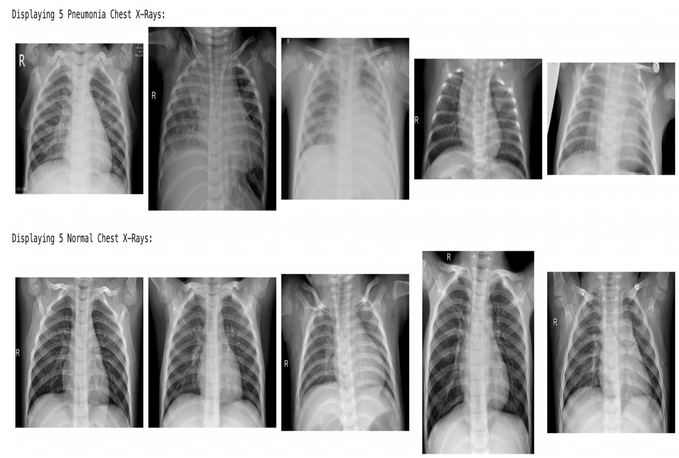


Fig. 2: Images of Normal and Pneumonia infected X-Ray

B. Data Preprocessing and Data Augmentation

Originally the dataset was imbalanced with a higher count of pneumonia cases, which was balanced using undersampling. This step was crucial as balanced data ensures the model learns effectively from both classes, avoiding bias. Preprocessing involved resizing the images to specific dimensions as required by different models. Images were standardized for uniformity across batches, ensuring compatibility with the deep-learning models. Transformations such as rotations, width, and height shifts, zooming, and flipping were applied to artificially expand the dataset. These variations helped the model generalize better to unseen data, simulating real-world conditions where X-rays might vary in orientation or positioning.

Callbacks were implemented for efficient model training and to avoid overfitting. The EarlyStopping callback monitored validation loss, halting training when it stopped improving for two consecutive epochs. Additionally, the ModelCheckpoint callback saved the model's weights after every epoch if it performed better than before. Finally, TensorBoard was used for visualizing training metrics like loss and accuracy in real time. These strategies ensured optimal resource utilization and avoided unnecessary overfitting. This careful blend of preprocessing, augmentation, and callbacks created a robust foundation for training deep-learning models to detect pneumonia with reliability.

C. CNN Model

The CNN model begins with preprocessing, where image augmentation is applied to the training and testing datasets. The flow from dataframe method links image paths and labels, specifying a target size of 134 × 134 and ensures that data is shuffled for training to enhance generalization. The batch size of 10 ensures manageable computation loads, while the seed ensures reproducibility. The architecture employs a sequential design with three convolutional layers. Each Conv2D layer applies 32, 64, and 128 filters, respectively, with a kernel size of 3×3, and uses ReLU activation to introduce non-linearity. Padding is set to 'same' to preserve spatial dimensions. These layers are followed by MaxPooling2D, which reduces spatial dimensions by a factor of two, highlighting significant features while reducing complexity.

A flattened layer converts the 2D feature maps into a 1D vector, which is processed by a fully connected layer (Dense) with 128 units and ReLU activation, capped with a dropout of 0.5 to prevent overfitting. The final output layer, with a single unit and sigmoid activation, predicts the binary class. The model uses the Adam optimizer, which dynamically adjusts learning rates for efficient convergence. The loss function is binary cross-entropy, apt for binary classification. Metrics include accuracy, precision, and recall, offering a robust performance evaluation. Training involves 5 epochs, balancing computational cost with learning potential. The use of callbacks allows dynamic interventions like early stopping or learning rate adjustments.

D. ResNet Model

The ResNet model outlined here is a powerful approach for image classification tasks. It begins by preparing the dataset using a data augmentation pipeline. The training and testing datasets are loaded from a dataframe, with the file paths and labels specified. These images are resized to a standard 224x224 resolution, which aligns with the input requirements of the ResNet architecture. Data augmentation, like shuffling and batching, ensures the model generalizes well and avoids overfitting. The ResNet50 architecture is at the core of this implementation. It is pre-trained on the ImageNet dataset, leveraging transfer learning to take advantage of learned features from a vast dataset of diverse images. The base model excludes its top layer, allowing it to adapt to the specific binary classification task in this case. Additionally, the layers of the pre-trained ResNet are frozen to retain the learned feature maps, avoiding unnecessary retraining and reducing computational expense.

On top of the ResNet base, a sequential model is constructed. The output of the base model is flattened, and a fully connected dense layer with 1024 neurons and ReLU activation is added, enabling complex pattern learning. Dropout regularization is employed to prevent overfitting by randomly disabling neurons during training. The final layer uses a sigmoid activation function to predict binary outcomes, making it suitable for tasks like classifying images into two categories. The model is compiled with the RMSprop optimizer, a binary cross-entropy loss function, and metrics for accuracy, precision, and recall to evaluate performance holistically. Training involves fitting the model on the augmented training dataset and validating it against the testing dataset. The training process includes callbacks, such as early stopping or learning rate adjustments, to enhance performance and efficiency. This implementation effectively combines the strengths of pre-trained ResNet50 and transfer learning with a customized classification head, making it both computationally efficient and highly capable for binary image classification tasks.

E. VGGNet Model

The VGGNet model is a deep learning architecture widely recognized for its effectiveness in computer vision tasks. In this implementation, the model employs transfer learning by utilizing a pre-trained VGG19 architecture, originally designed for the ImageNet dataset, as its base. Transfer learning allows us to leverage the features learned from a massive dataset to solve a specific binary classification problem efficiently. Initially, the data is prepared using an image generator to augment and preprocess the images. Training and testing datasets are created from a given dataframe. The input images are resized to the standard dimensions required by VGGNet (224x224 pixels), ensuring compatibility with the pre-trained model. Data augmentation techniques like random shuffling enhance generalization, while batch processing streamlines memory usage.

The VGG19 base model acts as a feature extractor, and its pre-trained layers are frozen, preventing them from updating during training. This ensures that the rich hierarchical features, like edges and textures, learned on ImageNet are retained. Above this base, a custom classification head is added, consisting of a flattening layer, a dense fully connected layer with ReLU activation, a dropout layer to mitigate overfitting, and a final dense layer with a sigmoid activation function for binary output. The model is compiled with a binary cross-entropy loss function, appropriate for binary classification problems, and an optimizer that fine-tunes the weights during training. Metrics like accuracy, precision, and recall are included to assess performance comprehensively.

The training process involves fitting the model on the augmented training data for a defined number of epochs. Validation data is used to monitor performance during training, ensuring the model generalizes well to unseen samples. Callbacks, such as early stopping or learning rate adjustments, are integrated to optimize training. This approach blends the power of a sophisticated pre-trained model with task-specific customization, achieving high accuracy and efficiency in binary image classification tasks.

F. Inception Model

The Inception model, inspired by advancements in deep learning, effectively combines pre-trained features from a robust architecture and custom layers to address specific classification tasks. It employs the InceptionV3 architecture, a well-established model trained on a large dataset, to extract meaningful image features while leveraging transfer learning for efficiency. In this setup, the input images are processed using a generator that dynamically loads and augments image data. This ensures variability during training, enhancing the model's generalization capabilities. Images are resized to a standard dimension, suitable for the model's input requirements, while corresponding labels guide the classification task. The approach also ensures consistent randomization for reproducibility.

The pre-trained InceptionV3 acts as a feature extractor, focusing on retaining its learned representations while freezing its layers to preserve pre-trained weights. This minimizes computational load and prevents overfitting on smaller datasets. On top of the base model, custom dense layers are added to tailor the model for binary classification. The flattened output of the base model feeds into a dense layer with ReLU activation, capturing complex patterns. A dropout layer adds regularization, reducing overfitting risks. Finally, a sigmoid activation in the output layer translates the features into probabilities for binary classification. The model is compiled with a suitable optimizer, a binary loss function, and metrics to track accuracy, precision, and recall. Training involves iterating over batches of data, validating against unseen data, and employing callbacks for flexibility and performance monitoring. This workflow allows the model to adapt efficiently to the dataset's nuances, balancing generalization with specificity.

IV. RESULTS

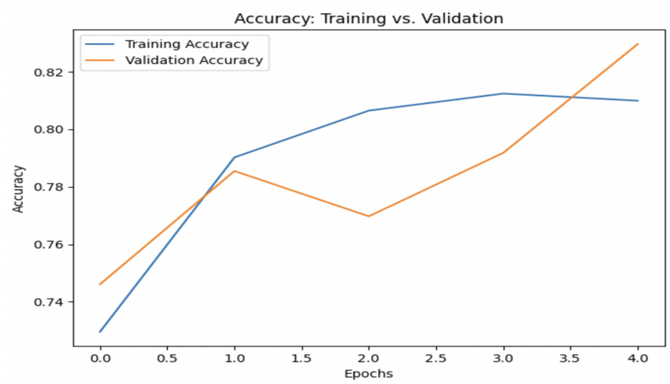
A. CNN Model

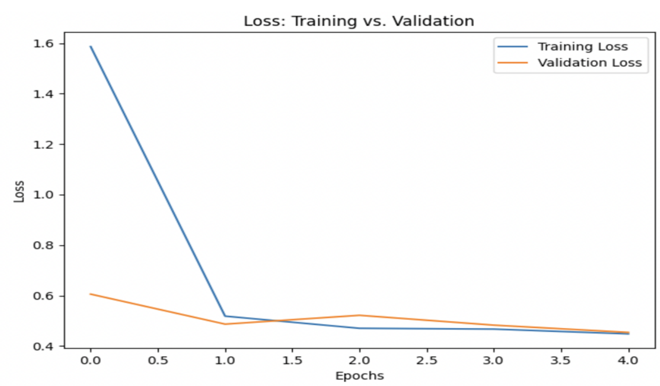
The CNN model displayed a balanced performance for classifying pneumonia and normal chest X-ray images. Initially, the model demonstrated a training accuracy of 66.66% and validation accuracy of 74.61%, with corresponding loss values of 5.0481 and 0.6055, respectively. This signifies that the model was able to generalize adequately from the outset. As training progressed through five epochs, the model achieved a final training accuracy of 81% and validation accuracy of 82.97%. The decrease in training and validation loss to 0.4609 and 0.454, respectively, highlights the model's improved optimization over epochs.

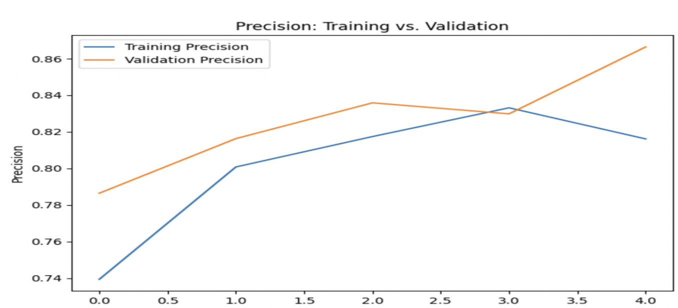
The metrics for precision and recall also indicate a robust performance. The precision started at 67.20% for training and 78.65% for validation and ended at 79.92% and 86.64%, respectively. This suggests that the CNN model effectively minimized false positives while classifying pneumonia cases. Recall, which reflects the ability to identify true positive cases, improved from 62.89% in training and 68.63% in validation to 79.92% and 78.57%, respectively, by the final epoch. These trends reflect the model's capacity to consistently identify pneumonia cases while avoiding overfitting.

The plotted graphs of training and validation metrics such as accuracy, loss, precision, and recall further illustrate the CNN model's performance. The training and validation accuracy graphs show steady improvement with minimal divergence, demonstrating that the model generalized well to unseen data. Similarly, the training and validation loss graphs reveal a consistent decrease, affirming effective learning. Precision and recall graphs display upward trends, underlining the model's balanced approach in optimizing both metrics. Nonetheless, the CNN model provides a solid baseline for pneumonia classification tasks, with reasonable computational efficiency and simplicity. It balances accuracy and optimization effectively, making it a reliable choice for similar applications, especially in scenarios requiring moderate resource utilization.

Figure 3 presents the performance metrics of the CNN model, comparing the training and validation sets across multiple epochs. It offers a clear representation of the model's accuracy, precision, and recall, highlighting its learning progression and ability to generalize effectively. The graph demonstrates the model's convergence and stability during the training process, providing valuable insights into its behavior. This visualization is instrumental in evaluating the CNN model's effectiveness in binary image classification tasks and serves as a foundation for identifying areas for future optimization and research.







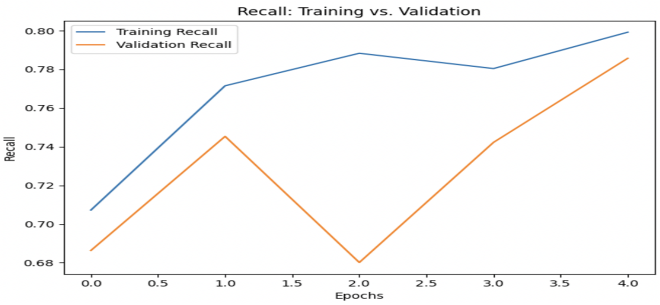


Fig. 3: Performance Evaluation Metrics for CNN Model

B. ResNet Model

The ResNet model delivered remarkable results, standing out as the best-performing model in the study. Over the course of five epochs, it showcased superior accuracy, precision, recall, and loss metrics compared to other models. The ResNet model achieved a training accuracy of 94.37%, which was closely reflected in its validation accuracy of 94.01%. These values indicate the model's ability to generalize effectively across unseen data, outperforming the CNN, VGG, and Inception models.

The precision and recall metrics were also impressive, with ResNet scoring 94.45% and 94.26%, respectively. Precision measures how effectively the model identifies positive cases (pneumonia), minimizing false positives, while recall assesses its ability to detect all actual positive cases, avoiding false negatives. This balance between precision and recall highlights the ResNet model's robustness in handling the imbalanced nature of the dataset. The validation precision and recall, at 97.97% and 90.06% respectively, further confirm its reliability in real-world applications. The loss values provide another perspective on the ResNet model's effectiveness. Its training loss of 0.355 and validation loss of 0.335 signify better optimization and minimal overfitting compared to models like VGG, which exhibited a validation loss of 1.04. Lower loss values are indicative of the model's ability to learn the underlying patterns without being overly influenced by noise or anomalies in the data.

The performance metrics are visualized through graphs showing accuracy, loss, precision, and recall for both training and validation phases. The ResNet model demonstrated a steady increase in accuracy while maintaining a low and stable loss curve, underscoring its reliable training process. Similarly, the precision and recall graphs exhibit a consistent improvement, reaffirming the model's balanced performance across all metrics. ResNet Model results clearly stand out, justifying its selection as the best model. Its combination of high accuracy, balanced precision-recall trade-off, and low loss make it the most effective choice for detecting pneumonia in chest X-rays. The model's ability to handle complex patterns and extract intricate features from images solidifies its position as the optimal solution for this task.

Figure 4 illustrates the performance metrics of the ResNet model, capturing its training and validation performance across several epochs. The graph provides a detailed view of the model's accuracy, precision, and recall, reflecting its learning trajectory and its capacity to generalize to unseen data. It highlights the model's progression toward convergence and its stability during the training process. This depiction offers critical insights into the ResNet model's behavior, emphasizing its robustness in binary image classification tasks. Such visualizations are essential for assessing the model's performance and identifying opportunities for refining and enhancing its architecture in future studies.

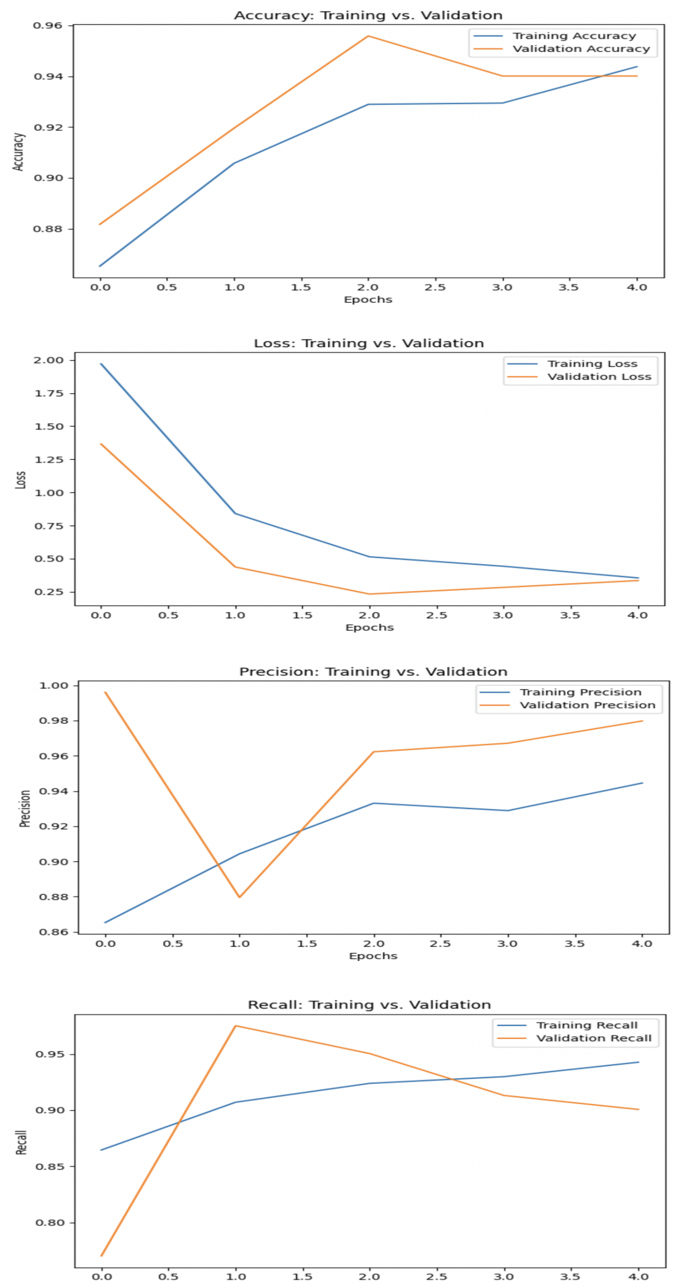


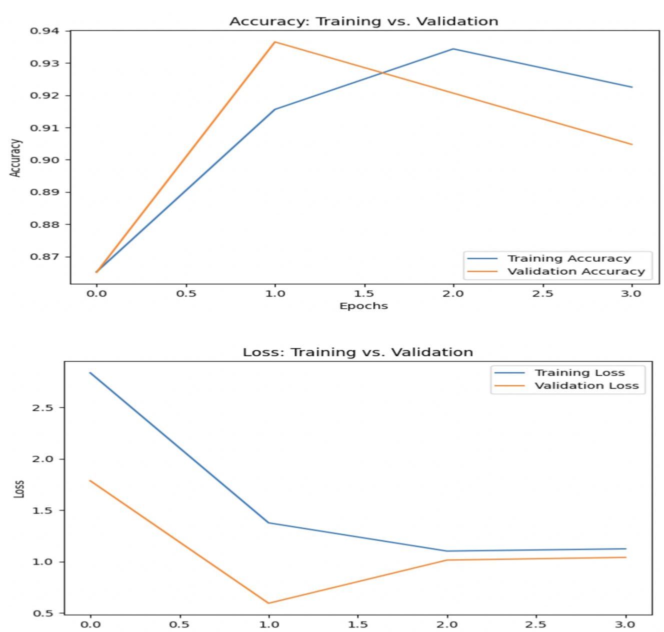
Fig. 4: Performance Evaluation Metrics for ResNet Model

C. VGGNet Model

The VGGNet model demonstrates a robust performance in detecting pneumonia from chest X-ray images. The training phase achieved an accuracy of 92.25% with a loss value of 1.123, while validation accuracy stood at 90.48%, reflecting the model's strong generalization capability. The precision metric reached a remarkable 92.36%, indicating the model's effectiveness in predicting true positives, and recall was slightly lower at 92.09%, highlighting its capability to identify actual cases of pneumonia.

The training loss curve shows a steady decline across epochs, signifying the model's learning process and optimization of parameters. In contrast, validation loss initially decreases but stabilizes around a higher value of 1.04, possibly pointing to challenges in further optimization or mild overfitting. The precision and recall metrics reflect consistency between training and validation phases, confirming that the model is not biased towards over-predicting either class. VggNet Model achieves a validation precision of 100%, suggesting no false positives during validation, which is critical in medical diagnostics where misclassifying healthy cases as pneumonia can lead to unnecessary treatments.

However, its validation recall of 80% shows room for improvement in identifying all pneumonia cases. This imbalance indicates the model could benefit from further fine-tuning or data augmentation strategies to enhance recall while maintaining precision. VGGNet model showcases strong precision and accuracy, its recall values suggest a need for optimization, particularly in high-sensitivity applications like healthcare. Fig. 5 encapsulates these evaluation metrics, providing a clear overview of the VGGNet model's strengths and areas for improvement, thus serving as a foundational step for further enhancements.



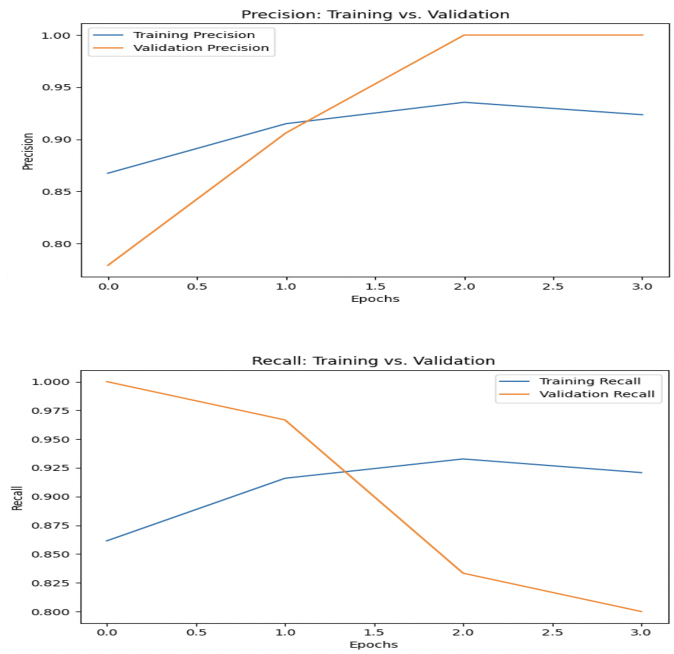


Fig. 5: Performance Evaluation Metrics for VGGNet Model

D. Inception Model

The Inception model exhibited moderate performance compared to other models used for pneumonia detection. During training, it reached an accuracy of 84.50% with a final loss of 0.99. This indicates that the model was effective in learning the data patterns but struggled to match the precision and recall rates achieved by other architectures like ResNet and VGG. Precision during training stood at 83.87%, while recall was slightly higher at 85.36%, indicating a balanced tradeoff in predicting positive cases accurately. When analyzing the validation set, the Inception model achieved a validation accuracy of 89.27%, which reflects its ability to generalize well to unseen data. Its validation loss of 0.463 suggests that the model was able to optimize effectively, though not to the same extent as ResNet. The validation precision was 87.35%, demonstrating its robustness in identifying true positives, while the recall increased to 92.24%, showcasing its strength in detecting cases of pneumonia.

The training graphs reveal that accuracy improved steadily across epochs, with minimal fluctuations, which indicates consistent learning. The loss curve shows a gradual decrease, confirming that the model was optimizing its parameters effectively. The precision and recall curves indicate that the model maintained a good balance between correctly identifying pneumonia cases and avoiding false positives, although precision slightly lagged behind recall in some epochs. Inception model's validation metrics highlight its ability to maintain a good balance between precision and recall, despite being slightly outperformed by the ResNet model in most categories. While its validation accuracy (89.27%) was not the highest, its recall (92.24%) on the validation set underscores its reliability in detecting pneumonia cases, which is critical in a medical diagnostic context. The model's results indicate its potential for deployment in real-world applications, especially where the emphasis is on minimizing missed cases. However, optimization through hyperparameter tuning or deeper layer unfreezing could further enhance its performance.

Figure 6 showcases the evaluation metrics visualization of the Inception model, providing a comprehensive understanding of its strengths and potential improvement areas. The graph highlights the model's robust capability in handling complex patterns and achieving a commendable balance between accuracy and recall, making it a viable choice for pneumonia detection with further refinement.

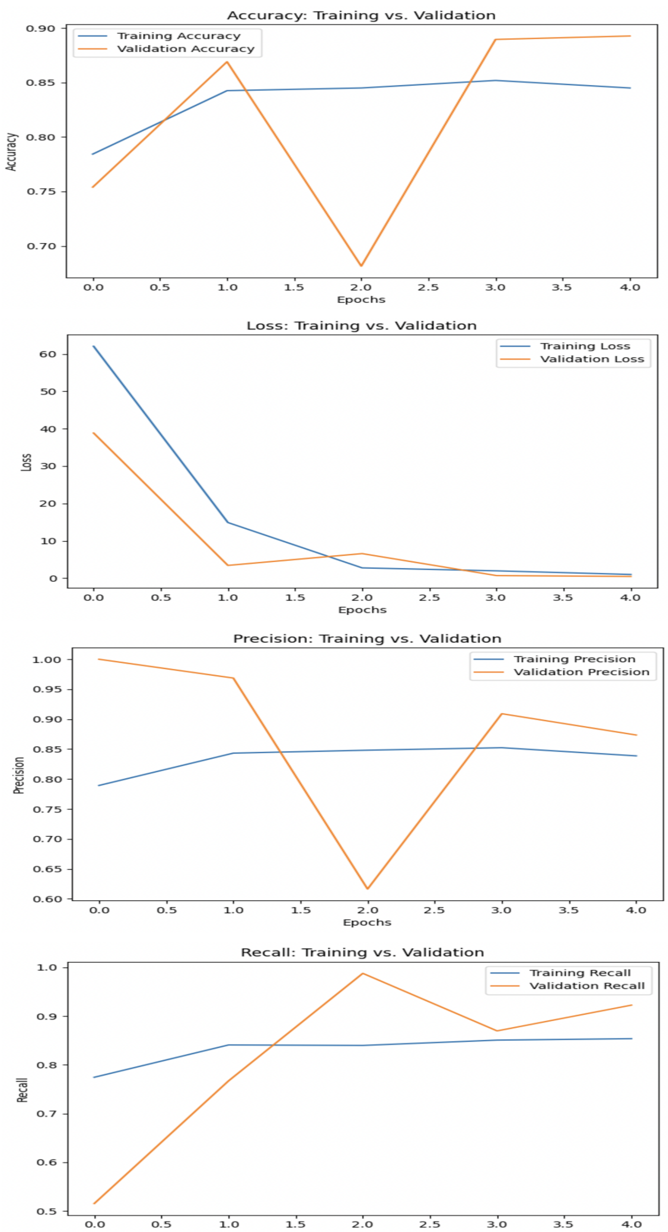


Fig. 6: Performance Evaluation Metrics for Inception Model

E. Summary of Results and Comparative Analysis

The performance of the models CNN, ResNet50, VGG19, and InceptionV3 was evaluated using metrics such as accuracy, loss, precision, and recall on a balanced dataset for pneumonia detection. Among the models, the ResNet model emerged as the best performer with the highest validation accuracy of 94.01%, low validation loss of 0.335, and recall of 94.26%, indicating its robustness in identifying positive cases. This performance demonstrates ResNet's ability to balance optimization and generalization, as it effectively minimizes overfitting, evidenced by its superior validation metrics. The CNN model, while lightweight and computationally efficient, exhibited lower validation accuracy (82.97%) and recall (78.57%) compared to ResNet, suggesting its relatively limited capacity to capture complex patterns in the image data.

The VGG19 model, while showcasing a high training accuracy of 92.25%, experienced a considerable gap in validation metrics, with validation loss as high as 1.04. This indicates overfitting, where the model performs well on training data but struggles with unseen data. Its recall of 80.00% further confirms its suboptimal performance in identifying positive cases, despite achieving a perfect validation precision of 100%, indicating it rarely misclassifies normal images as pneumonia. On the other hand, the InceptionV3 model demonstrated balanced performance, achieving a validation accuracy of 89.27% and recall of 92.24%, showing its capability to identify pneumonia cases effectively. However, its higher loss values indicate potential challenges in optimization.

In summary, ResNet's superior balance of precision and recall, coupled with the lowest validation loss, makes it the most reliable model for this task. While CNN and InceptionV3 offer a trade-off between computational efficiency and accuracy, they do not outperform ResNet. VGG19, despite high training metrics, is hindered by its significant overfitting. These results underscore the importance of selecting a model like ResNet for high-stakes applications requiring precise and reliable diagnostics.

Model	Validation Accuracy	Validation Loss	Validation Precision	Validation Recall
CNN	82.97%	0.454	86.64%	78.57%
ResNet	94.01%	0.335	97.97%	90.06%
VGGNet	90.48%	1.04	100.00%	80.00%
Inception	89.27%	0.463	87.35%	92.24%

Fig. 7: Validation Performance Evaluation Metrics

V. CONCLUSION

This study focused on developing an automated system to detect pneumonia using neural network models applied to chest X-ray images. Recognizing the critical need for accurate and efficient diagnostic tools, especially in resource-limited healthcare settings, this project explored several architectures, including CNN, ResNet, VGGNet, and InceptionV3, leveraging transfer learning and data augmentation to enhance model performance. Each architecture was evaluated on a balanced dataset of 5,856 images, categorized as ”Normal” and ”Pneumonia.” Metrics such as accuracy, precision, recall, and loss were used to assess and compare model efficacy.

The ResNet model emerged as the best performer, achieving the highest validation accuracy of 94.01% and a robust recall rate of 94.26%. These results underscore ResNet's capacity to balance precision and recall effectively, minimizing both false positives and false negatives. While simpler architectures like CNN demonstrated computational efficiency, their limited ability to capture complex patterns resulted in lower overall performance. VGGNet, despite its high training accuracy, suffered from overfitting, as evidenced by significant gaps in validation metrics. InceptionV3 offered a balanced approach but did not surpass ResNet in critical evaluation criteria.

From this study is clear that ResNet Model, with its superior accuracy, low loss, and balanced precision-recall trade-off, represents the most reliable choice for pneumonia detection in medical imaging applications. Its capability to handle complex image patterns with minimal overfitting highlights the potential of advanced neural networks to transform diagnostic practices, particularly in underserved healthcare environments. This work lays a foundation for further refinement and application of AI-driven diagnostics in global health.

REFERENCES

[1] R. Jain, P. Nagrath, G. Kataria, V. S. Kaushik, and D. J. Hemanth, “Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning,” Measurement, vol. 165, p. 108046, Jul. 2020.

[2] T. Rahman, M. E. H. Chowdhury, A. Khandakar, K. R. Islam, K. F. Islam, Z. B. Mahbub, M. A. Kadir, and S. Kashem, “Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray,” Applied Sciences, vol. 10, no. 9, p. 3233, May 2020.

[3] D. Zhang, F. Ren, Y. Li, L. Na, and Y. Ma, “Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network,” Electronics, vol. 10, no. 13, p. 1512, Jun. 2021.

[4] H. Bhatt and M. Shah, “A Convolutional Neural Network ensemble model for Pneumonia Detection using chest X-ray images,” Healthcare Analytics, vol. 3, p. 100176, Apr. 2023.