AVE Trends in Intelligent Health Letters

Enhanced Pneumonia Detection Through Transfer Learning on Chest Radiograph Images

J. Angelin Jeba[1,*], S. Rubin Bose[2], R. Regin[3], O. Jeba Singh[4], Bushra Rehman[5], R. Nathiya[6]

[1]Department of Electronics and Communication Engineering, S. A. Engineering College, Chennai, Tamil Nadu, India.
[2,3]School of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[4]Center for Academic Research, Alliance University, Bengaluru, Karnataka, India.
[5]Institute of Pathology and Diagnostic Medicine, Khyber Medical University, Hayatabad, Peshawar, Pakistan.
[6]Department of Computer Applications, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu, India.
angelinjeba@saec.ac.in[1], rubinbos@srmist.edu.in[2], reginr@srmist.edu.in[3], jeba.singh@alliance.edu.in[4], drbushra.ipdm@kmu.edu.pk[5], nathiyar@hindustanuniv.ac.in[6]

*Corresponding author

Abstract: Utilising the potential of transfer learning in conjunction with deep convolutional neural networks, this research presents a comprehensive framework for the automated diagnosis of pneumonia based on chest radiographs. Therefore, the suggested methodology uses the VGG19 architecture to extract meaningful features from X-ray images, thereby improving diagnostic accuracy and generalisation. The model's robustness is further improved by extensive data augmentation and iterative fine-tuning, enabling consistent performance despite class imbalances and limited medical data. Experimental findings from a labelled chest X-ray dataset demonstrate the model's capacity to effectively differentiate between normal and pneumonia cases. These results indicate considerable improvements in classification accuracy, sensitivity, and specificity. This technique has the potential to assist clinical workflows, accelerate patient assessment, and promote timely action. It does this by streamlining detection operations and minimising the likelihood of human error. The expansion of algorithmic variety, optimisation of computing efficiency, and integration of multimodal data sources should be the focus of future research initiatives. This will pave the way for more scalable and intelligent diagnostic systems in the healthcare industry.

Keywords: Pneumonia Detection; Chest Radiographs; Transfer Learning; Healthcare Industry; Fine Tuning Strategies; Data Augmentation; Diagnostic Systems; Algorithmic Diversity.

Cite as: J. A. Jeba, S. R. Bose, R. Regin, O. J. Singh, B. Rehman, and R. Nathiya, "Enhanced Pneumonia Detection Through Transfer Learning on Chest Radiograph Images," *AVE Trends in Intelligent Health Letters*, vol. 2, no. 4, pp. 216–227, 2025.

Journal Homepage: https://avepubs.com/user/journals/details/ATIHL

Received on: 19/12/2024, **Revised on:** 07/04/2025, **Accepted on:** 18/07/2025, **Published on:** 09/12/2025

DOI: https://doi.org/10.64091/ATIHL.2025.000209

1. Introduction

Pneumonia remains one of the most pressing health challenges worldwide, contributing significantly to mortality, especially among children and the elderly. Its clinical detection relies heavily on interpreting chest radiographs, a process that can be time-consuming and prone to human error. This has put tremendous pressure on healthcare systems, particularly in regions where access to skilled radiologists is limited, and the sheer volume of cases can overwhelm available resources. Traditional diagnostic workflows often struggle to balance speed, accuracy, and consistency when evaluating chest X-rays for signs of pneumonia.

Subtle visual cues can be easily missed, and variations in clinicians' expertise may lead to inconsistent outcomes. These challenges have driven a growing interest in leveraging advanced computational methods, particularly deep learning techniques, to augment and support clinical decision-making. Convolutional Neural Networks (CNNs) have demonstrated remarkable potential for extracting complex patterns from medical images, but the size and diversity of available datasets often limit success. Recognising these challenges, this research focuses on leveraging the proven capabilities of the VGG19 network and transfer learning to automate pneumonia detection in chest X-rays. The methodology emphasises robust data augmentation, incremental fine-tuning, and class balancing to deliver a solution that is both highly accurate and resilient to real-world limitations in medical data. By introducing advanced AI-driven techniques into diagnostic workflows, the study aims to improve detection accuracy, accelerate patient triage, and reduce the burden on healthcare providers.

2. Literature Review

Recent advancements have positioned deep learning as a transformative tool for medical image analysis, particularly in pneumonia diagnosis. Dey et al. [1] surveyed neural network-based approaches for chest X-ray analysis, emphasising the impact of combining image enhancement techniques with a VGG-based architecture, which achieved an impressive AUC of 0.991 for pneumonia detection. This work underscores the importance of both network selection and preprocessing steps for achieving optimal diagnostic reliability. An et al. [2] conducted a comparative study of the VGG16 and Xception architectures, concluding that VGG16 delivers superior accuracy for pneumonia classification. However, Xception exhibited greater sensitivity to frequent cases in the datasets. Their findings illustrate the critical nature of architecture selection and fine-tuning for effective automated diagnosis. Singh et al. [3] demonstrated the practical benefits of VGG-based CNNs for pneumonia detection, achieving reliable classification accuracy across multiple public datasets and highlighting their adaptability to varying clinical data. Mallidi [4] proposed an ensemble technique, combining GoogLeNet, ResNet-18, and DenseNet-121. Using a weighted-average ensemble method based on several evaluation metrics, their approach achieved 98.81% accuracy on the Kermany dataset. This research demonstrates that ensemble learning and metrics fusion can yield robust performance and greater generalizability.

Mujahid et al. [5] and Dahmane et al. [6] advanced research on object detection and computational efficiency, proposing anchor-free frameworks and YOLO-based models for lesion detection. Hasan et al. [7] introduced a Fast-YOLO architecture that optimises feature extraction modules and reduces computational demands while achieving superior accuracy and real-time detection suitable for clinical settings. Khan et al. [8] significantly contributed to the field by releasing the Chest X-ray14 dataset, comprising over 112,000 front chest radiographs from 32,717 unique patients, annotated for 14 more common thoracic diseases, including pneumonia. This publicly available dataset catalysed research into large-scale deep learning applications in medical imaging, enabling rigorous benchmarking of CNN architectures, including AlexNet, GoogleNet, ResNet, and VGG16. Their extensive evaluations revealed ResNet's superior ability to capture complex pathological features, achieving one of the best and a very strong state-of-the-art classification performance. Importantly, the author also addressed challenges such as label noise introduced by automated annotation, significant class imbalance across disease categories, and this also required carefully curated validation data splits to avoid biased performance estimations. This dataset and its findings laid the groundwork for subsequent, more successful pneumonia diagnostic algorithms and remain a cornerstone resource for researchers worldwide. Jahnav and Sivasankar [9] built on these foundations, developing CheXNet, a 121-layer densely connected convolutional neural network tailored for pneumonia detection on the Chest X-ray14 dataset.

The model was rigorously trained using transfer learning and advanced optimisation techniques, enabling it to learn discriminative features even with relatively scarce labelled data. Remarkably, CheXNet outperformed a panel of four practising radiologists in F1 score (0.435 vs 0.387), highlighting the transformative potential of AI to augment human expertise in clinical diagnosis. The architecture's design deliberately emphasised depth and connectivity to exploit hierarchical feature representations, thereby enabling robust identification of subtle radiographic patterns associated with pneumonia. Jahnav and Sivasankar [9] underscored the clinical significance of their model not only as a diagnostic aid but also as a tool for resource optimisation in settings with limited radiology expertise. Jha et al. [10] proposed an innovative application of a Siamese network architecture for pneumonia subtype classification, distinguishing viral, bacterial, and normal cases. By employing a dual-input CNN that simultaneously compares paired imaging samples, their model captured minute pathological variations that conventional single-input nets can overlook. Acharya’s approach leveraged contrastive loss to enhance feature embeddings, thereby improving separation between clinically distinct pneumonia etiologies. Their model was trained and validated on approximately 5,600 chest X-rays, demonstrating an outstanding ROC AUC of 0.95, signifying high clinical utility. This work represents an important step toward precision medicine, offering clinicians a tool to tailor treatment strategies based on pneumonia subtyping, which is critical given the differing therapeutic approaches for viral versus bacterial infections.

Rajpurkar et al. [11] synthesised findings from 68 studies investigating deep learning for chest X-ray diagnostics, paying particular attention to pneumonia and COVID-19 applications. Their review highlighted that architectures such as VGG, ResNet, EfficientNet, and DenseNet consistently delivered superior results, often exceeding 90% accuracy when adequate data augmentation and transfer learning were applied. Nevertheless, the review also highlighted persistent challenges, including limited dataset size, heterogeneity in acquisition, and the lack of explainability in black-box CNN models, which hinder clinical trust and acceptance. Hashmi et al. [13] advocated for the development of interpretable AI frameworks, enhanced multi-centre

data collaborations to increase diversity, and the embedding of AI tools within existing clinical workflows to maximise impact. Their comprehensive analysis provides a roadmap for future research priorities to bridge the gap between experimental AI systems and day-to-day radiological practice. The reviewed studies showcase a wide array of deep learning strategies applied to the challenge of pneumonia diagnosis using chest radiographs, highlighting substantial advancements but also identifying areas for future improvement. Salehi et al. [14] introduced the ChestX-Ray14 dataset, providing researchers with over 112,000 labelled images covering 14 thoracic diseases, including pneumonia. This dataset laid the foundation for benchmarking many deep convolutional neural networks, including AlexNet, GoogleNet, ResNet, and VGG16, with ResNet demonstrating superior classification accuracy. Their work also addressed practical challenges such as noisy annotations and pronounced class imbalance, underscoring the importance of data quality for algorithmic success.

Brima et al. [15] synthesised findings from 68 studies on deep learning for chest X-rays, focusing on pneumonia and COVID-19 detection. The review highlighted the consistent effectiveness of VGG, ResNet, EfficientNet, and DenseNet architectures, while also drawing attention to persistent challenges in dataset heterogeneity, model interpretability, and clinical integration. Emphasising the need for explainable AI and collaborative data sharing, this comprehensive analysis sets a roadmap for translating advances in deep learning into practical healthcare applications. Luján-García et al. [12] form a pivotal body of work that shapes current understanding and progress in automated pneumonia diagnosis. From foundational dataset contributions and cutting-edge model developments to meta-analyses of challenges and trends, these studies drive the AI-powered reimagining of clinical radiology. However, further efforts are required to enhance model robustness, improve transparency, and ensure seamless integration across diverse healthcare environments, thereby maximising patient benefit and acceptance among medical professionals. Additional research by Singh et al. [3] and Mallidi [4] introduced novel pooling mechanisms inspired by VGG networks for enhanced pneumonia detection on multitask datasets, while Mujahid et al. [5] and Dahmane et al. [6] explored ensemble approaches and interpretable prototype-based networks, achieving accuracies of upwards of 94% and highlighting the value of hybrid, explainable models. Such efforts collectively address critical limitations of single-model approaches and the need for transparency in automated diagnosis.

3. Methodology

The methodology of this research paper is carefully designed to ensure a comprehensive, effective approach to automated pneumonia detection using chest radiographs, built on advanced deep learning techniques and the principles of transfer learning. Each step of the process is crucial to the solution's overall success and practical viability. The study begins with the collection and curation of a high-quality dataset of chest X-ray images, categorised into normal and pneumonia classes. Recognising that medical imaging data can be inherently imbalanced due to a lower number of disease-positive cases, a robust data augmentation pipeline is applied (Figure 1).

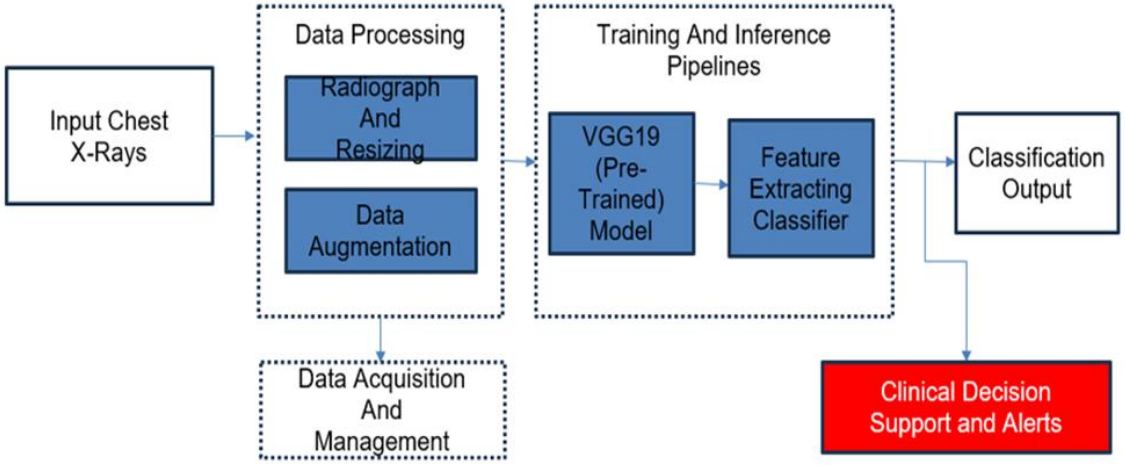

Figure 1: Block diagram

This approach includes operations such as random horizontal and vertical flips to simulate real-world image orientations, controlled rotations to account for minor patient positioning differences, shearing to replicate subtle distortions, and width and height shifts to introduce variability in the region of interest. Such augmentation not only increases the diversity and size of the training set but also reduces the risk of overfitting by forcing the network to learn invariant features rather than memorise the data, thereby enhancing its ability to generalise to unseen cases (Figure 2). At the core of the architecture lies the VGG19 convolutional neural network, a deep, well-established model recognised for its effectiveness in image classification, thanks to its layered structure and small convolutional filters that capture detailed features.

Instead of training the model from scratch, which is computationally expensive and data-hungry, this paper employs transfer learning by initialising the network with weights pre-trained on the vast ImageNet dataset. This strategy leverages learned low-

level features, such as edges and textures, that transfer to medical images, thereby accelerating convergence and improving accuracy. Customisation involves replacing the original classification layers with new fully connected layers tailored for binary pneumonia classification, along with dropout layers that randomly disable neurons during training to prevent overfitting and produce a more robust model. A key aspect of the training strategy is the incremental and selective unfreezing of the VGG19 layers, known as fine-tuning. Initially, only the newly added layers are trained, allowing the model to specialise in pneumonia-specific feature extraction without altering the generic pre-trained filters.

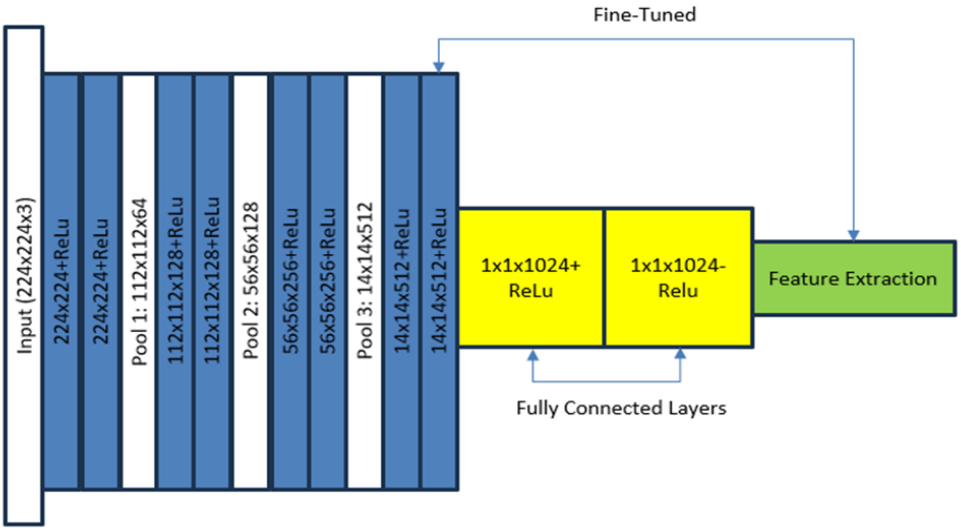

Figure 2: Architecture diagram of the VGG19 model

Subsequently, deeper convolutional blocks are gradually unfrozen and trained with a significantly smaller learning rate, allowing the model to adapt its intermediate representations subtly while preserving foundational features. This layered approach is critical, as it balances the benefits of pre-trained knowledge with the need for task-specific adaptation, optimising performance in the medical imaging domain without sacrificing generalisation. The model is trained using stochastic gradient descent with Nesterov momentum, a technique that accelerates gradient updates by incorporating an anticipatory adjustment, thereby improving convergence speed and stability.

This is complemented by early stopping criteria that monitor validation loss and halt training once the model's performance plateaus or starts to degrade, preventing overfitting and conserving computational resources. Moreover, a learning rate scheduler automatically reduces the learning rate if no improvement is observed over a set number of epochs, ensuring that the model refines its weights with finer adjustments in the later stages of training. Following data collection, data augmentation techniques are applied to the training dataset. Augmentation involves applying transformations such as random scaling, rotation, flipping, and colour jittering to increase the dataset's diversity. This helps the model generalise better to unseen data and improves its robustness. Evaluation of the model is conducted comprehensively on validation and unseen test datasets to provide unbiased estimates of its diagnostic performance.

Metrics such as accuracy measure overall correctness, while loss indicates the degree of prediction error. Sensitivity (recall) and specificity provide insights into the model's ability to correctly identify pneumonia-positive cases and avoid false negatives, crucial for clinical relevance. The model’s weights and architecture are saved in the industry-standard Keras format, facilitating reproducibility and easing future deployment in clinical environments. Once the model is trained, it is evaluated on the test dataset to assess its performance in terms of detection accuracy and speed. Metrics such as mAP (mean Average Precision) are computed to quantify the model's performance. The model's predictions, including precision, recall, and F1-score, are analysed to assess its effectiveness in detecting chest abnormalities. Here are some key mathematical equations relevant to training and inference in your pneumonia detection model using VGG19 and transfer learning.

3.1. Loss Function

The model is trained by minimising the categorical cross-entropy loss function, defined as:

$$L = \sum_{i=1}^{C} \quad y \log \log (y) \qquad (1)$$

Where C is the number of classes (2 for pneumonia detection: pneumonia, normal), yi is the binary indicator (0 or 1) if class label i is the correct classification for the input, and y^i is the predicted probability of the input belonging to class ii. Equation (1) represents the categorical cross-entropy loss, a fundamental metric in multi-class classification tasks such as pneumonia detection. It quantifies the divergence between the true class labels of chest X-ray images and the model's predicted probabilities. When the model assigns a low probability to the correct class (pneumonia or normal), the loss increases, indicating low confidence in the prediction.

3.2. Softmax Activation Function

The network's final layer applies the SoftMax function to output probabilities over classes:

$$\sigma(z) = \frac{e\hat{}z}{\sum_{k=1}^{C} e\hat{}zk} \quad (2)$$

Where zj is the raw output (logit) score for class j, equation (2) represents the Softmax activation function, which transforms the raw output scores (logits) of the pneumonia detection model into normalised probability values across multiple classes—typically pneumonia and normal. Each output value ranges between 0 and 1, with the total probability summing to 1, allowing clear interpretation of prediction confidence. In the context of pneumonia classification, Softmax assigns a higher probability to the most probable class while reducing the influence of incorrect classes, ensuring accurate and interpretable predictions for clinical decision support.

3.3. Weight Upgrades Using Stochastic Gradient Descent with Nesterov Momentum

During backpropagation, model weights are updated to minimise loss using the gradient descent rule with Nesterov momentum:

$$v(t+1) = \mu v(t) - \eta \nabla L(\theta t + \mu vt) \quad (3)$$

Where θt are weights at iteration t*t*, v*t* the velocity, μ the momentum coefficient (e.g., 0.9), η the learning rate, and $\nabla L(\cdot)$ the gradient of loss with respect to weights. Equation (3) shows the need for weights when using Stochastic Gradient Descent (SGD) for pneumonia detection, driven by the imbalance and variability in medical image data. In real datasets, the number of pneumonia-positive and normal chest X-rays often differs significantly, which can bias the model toward the majority class. Applying class weights ensures that the model assigns equal importance to both classes by penalising misclassifications of minority classes more heavily.

3.4. Learning Rate Scheduling

Learning rate η is adapted during training to improve convergence:

$$\eta = \eta o * \text{decay}_{\text{factor}} \quad (4)$$

Where η is the initial learning rate, decay_factor <1 reduces the learning rate, and k is the number of decay steps. Equation (4) represents the learning rate decay strategy, which progressively reduces the learning rate (η) as training advances. In the pneumonia detection model, this approach allows the network to make larger parameter updates in the initial stages—helping it quickly learn key features such as lung texture and opacity patterns—and then take smaller, refined steps later to fine-tune its accuracy. The initial learning rate (η_0) enables rapid progress, while the decay factor ensures stability and prevents overshooting during optimisation. This gradual reduction is crucial for achieving balanced convergence, resulting in a more reliable and precise pneumonia classification model.

3.5. Accuracy Metric

The effectiveness of the model during validation and testing is evaluated using accuracy:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

Equation (5) depicts the classification accuracy, which measures the overall effectiveness of the pneumonia detection model by calculating the ratio of correctly predicted cases to the total number of samples.

4. Experimental Setup

4.1. Training

The experimental setup for training the pneumonia detection model using VGG19 transfer learning begins with the careful organisation and preparation of the chest X-ray datasets. These datasets comprise thousands of labelled images, differentiated into pneumonia-positive and normal cases, sourced from publicly available medical imaging repositories. Each image undergoes an essential preprocessing step: resizing and normalisation to conform to the VGG19 architecture's input requirements, ensuring uniformity and enhancing feature extraction effectiveness. To address the challenge of limited and imbalanced medical data, a comprehensive data augmentation strategy is employed. This involves applying random geometric transformations, such as horizontal and vertical flipping, rotations within a 30-degree range, zooming between 80% and 120%, and brightness adjustments between 90% and 110%. By artificially diversifying the training samples, augmentation not only increases the dataset size but also encourages the model to learn robust, invariant features, which are necessary for reliable real-world performance.

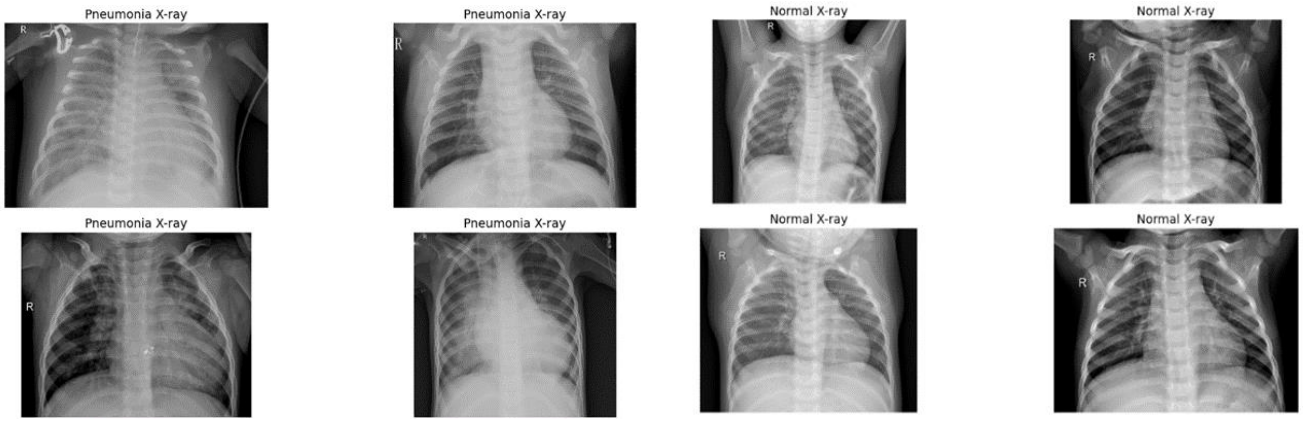

Figure 3: (a) Pneumonia X-rays dataset (b) Normal X-rays dataset

Figure 3 illustrates that the collected dataset consists of chest X-ray images categorised into Normal and Pneumonia classes. Normal X-rays show clear lung fields without opacities or fluid accumulation, representing healthy respiratory conditions. Pneumonia X-rays exhibit visible lung infiltrates and dense regions indicating infection. The dataset includes varied age groups and imaging conditions to ensure model robustness. This diverse data supports reliable training and evaluation of pneumonia detection models. The training phase of this pneumonia detection model utilised a high-performance computing system to handle the computational demands of deep learning. The hardware setup included an Intel Core i5 processor (12th Gen or equivalent) to ensure efficient execution of training scripts, paired with at least 16 GB of DDR4 RAM to manage large image batches and data augmentations without bottlenecks. A 512 GB SSD provided fast read/write speeds, crucial for loading large chest X-ray datasets and saving model checkpoints. For accelerated training, a dedicated GPU supporting CUDA (such as an NVIDIA RTX 3060 or higher) is recommended, as it provides the parallelism needed for the convolution operations in the VGG19 architecture.

4.2. Evaluation

Once model training is complete, the pneumonia detection system undergoes rigorous evaluation using a designated test dataset that was not used in training. This ensures an unbiased assessment of the model's true predictive performance. Key evaluation metrics computed include accuracy, precision, recall (sensitivity), and the F1-score, which collectively provide a comprehensive understanding of the model's diagnostic capabilities. Accuracy reflects the overall correctness of predictions, while precision measures the reliability of pneumonia positive classifications, and recall indicates the model's effectiveness in identifying actual pneumonia cases. The F1-score balances precision and recall, providing insight into the model's ability to manage false positives and false negatives, which are critical in medical diagnosis. In addition, the loss function values calculated during evaluation quantify the difference between predicted outputs and true labels, guiding further optimisation if necessary. This comprehensive evaluation framework enables a reliable assessment of the model's strengths and limitations before deployment. Following successful evaluation, the trained model can be integrated into clinical or health IT systems,

facilitating real-time pneumonia detection from chest radiographs and supporting timely clinical decision-making through automated alerts and diagnostic support.

4.3. Implementation

The pneumonia detection model leverages advanced transfer learning and deep convolutional neural networks to efficiently analyse chest radiographs. The system first preprocesses the input images by normalising and resizing them to ensure consistency across samples. Data augmentation artificially expands the dataset, thereby improving model robustness. The VGG19 model, pre-trained on the ImageNet dataset, serves as the core feature extractor, reducing the need for extensive training from scratch and enabling faster convergence with limited medical imaging data. Custom classifier layers, such as dense or SVM layers, are appended to perform the final binary classification of pneumonia presence. The model is integrated into clinical workflows, providing real-time diagnostic support and alerting healthcare professionals to enable timely intervention.

5. Results and Discussions

The VGG19-based pneumonia detection model processes input chest X-ray images through a series of pre-processing steps, including normalisation and resizing to fit the network's requirements. During training, the model leverages transfer learning by using pre-trained ImageNet weights, enabling it to extract meaningful features from radiographs without extensive training from scratch.

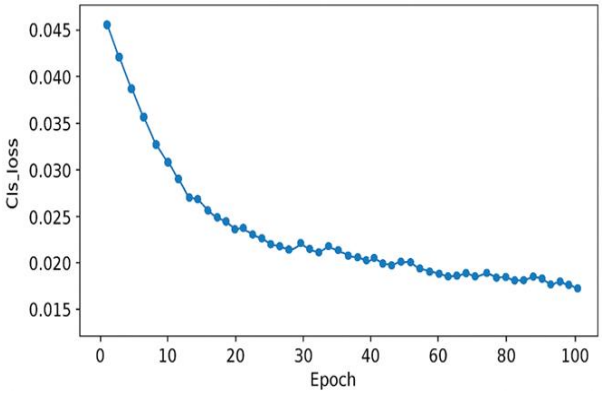

Figure 4: Object loss for pneumonia model

The training pipeline incorporates iterative weight optimisation to minimise a cross-entropy loss function, while data augmentation techniques enhance model generalizability. After training, the model performs inference, classifying new chest X-rays as pneumonia or normal by passing them through the fine-tuned VGG19 and subsequent classification layers. Performance metrics such as loss, accuracy, precision, recall, and mean average precision are used to evaluate the model's robustness and accuracy. The final output is designed to support clinical decisions by providing clear predictions and alerts for pneumonia detection. Figure 4 illustrates the variation of classification loss during the training phase of the pneumonia detection model.

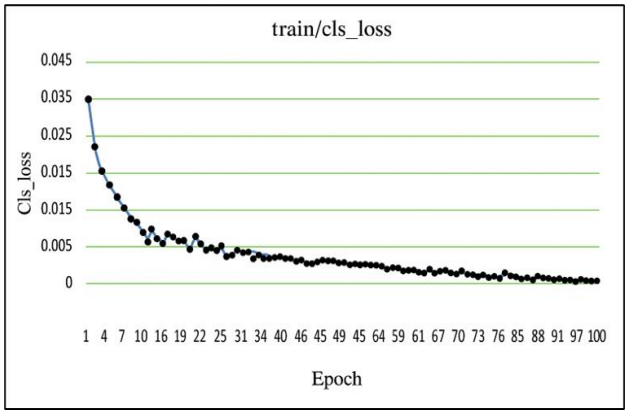

Figure 5: Class loss for pneumonia model

The Cls_loss represents how well the model distinguishes between pneumonia and normal chest X-ray images. As the number of epochs increases, the loss value shows a consistent downward trend, indicating that the model is learning effectively and improving its predictive accuracy. The gradual stabilisation of loss after around 70 epochs suggests that the model has reached near-optimal convergence, with minimal further improvement in training performance. Figure 5 represents how accurately the network distinguishes between normal and pneumonia-affected chest X-rays. It measures the difference between the model's predicted output and the actual class label provided in the dataset. A lower loss value indicates that the model's predictions are closer to the true results, meaning it is learning effectively. During training, this loss is minimised via backpropagation, enabling the model to adjust its internal weights and improve prediction accuracy. The class loss ensures the model improves at identifying pneumonia cases while reducing false classifications of normal X-rays, ultimately enhancing the overall reliability of the diagnostic system.

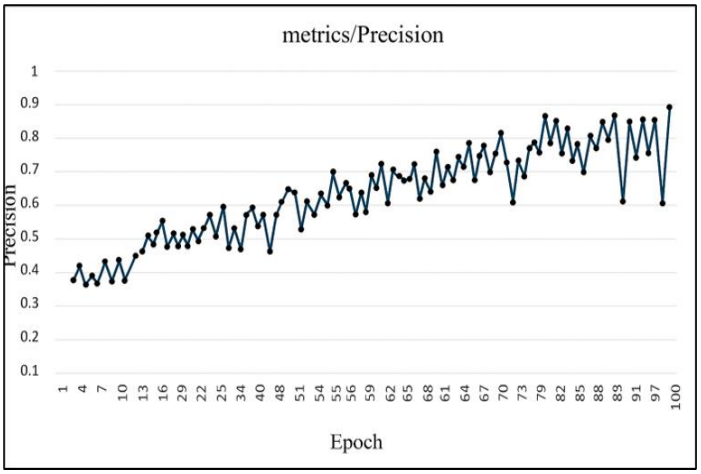

Figure 6: Precision

Figure 6 illustrates the precision metrics across multiple training epochs. Precision measures the model's ability to correctly identify pneumonia cases among all predicted positive cases. The steady upward trend in precision indicates that the model becomes increasingly accurate in detecting pneumonia as training progresses. The slight fluctuations near the end suggest adaptive learning behaviour but overall stability in prediction quality.

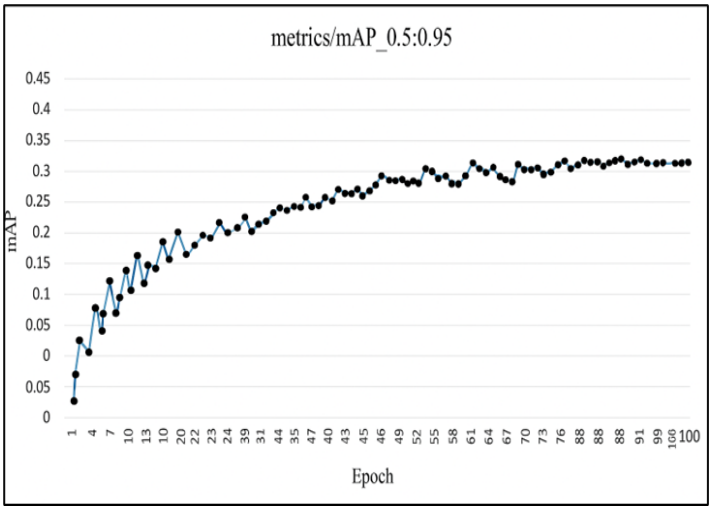

Figure 7: Mean average precision for pneumonia

Figure 7 represents a performance metric used to evaluate how accurately a model detects and classifies objects or patterns—in this case, pneumonia in chest X-ray images. It combines both precision and recall to provide a balanced measure of the model's effectiveness. Precision indicates how many of the detected pneumonia cases are actually correct, while recall measures how many of the actual pneumonia cases were successfully identified. The model first calculates Average Precision (AP) for each class using the area under the precision-recall curve, then averages these values to obtain the mAP score. A higher mAP

value indicates that the model is performing well, accurately identifying pneumonia cases with fewer false positives and false negatives. Overall, mAP gives a comprehensive view of how well the model balances detection accuracy and consistency across all classes (Figure 8).

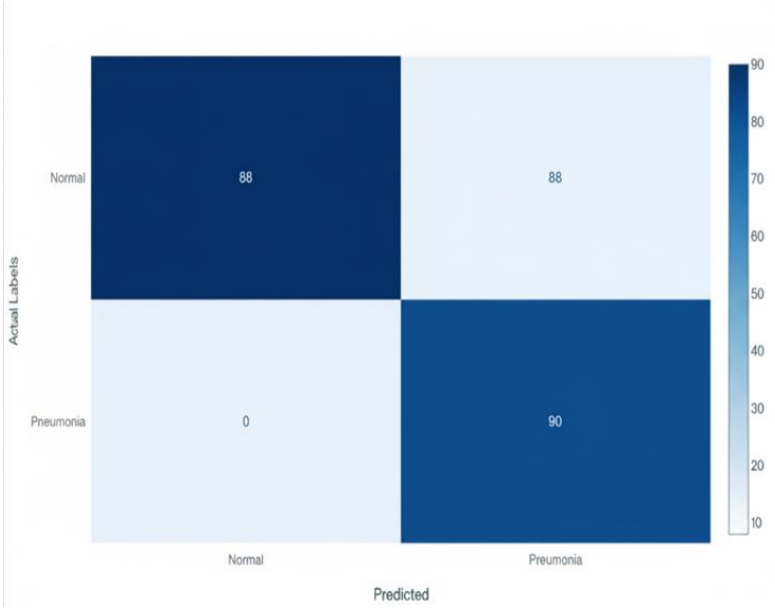

Figure 8: Confusion matrix

A confusion matrix for the VGG19-based pneumonia detection model provides a detailed breakdown of its classification performance by categorising predictions into True Positives (correctly identified pneumonia cases), False Positives (normal cases incorrectly classified as pneumonia), True Negatives (correctly identified normal cases), and False Negatives (pneumonia cases missed by the model). This matrix is essential for evaluating the model's ability to distinguish between pneumonia and non-pneumonia chest X-rays. From the confusion matrix, key metrics such as precision, recall (sensitivity), and F1 score are derived, offering critical insights into the model's diagnostic strengths and weaknesses (Figure 9).

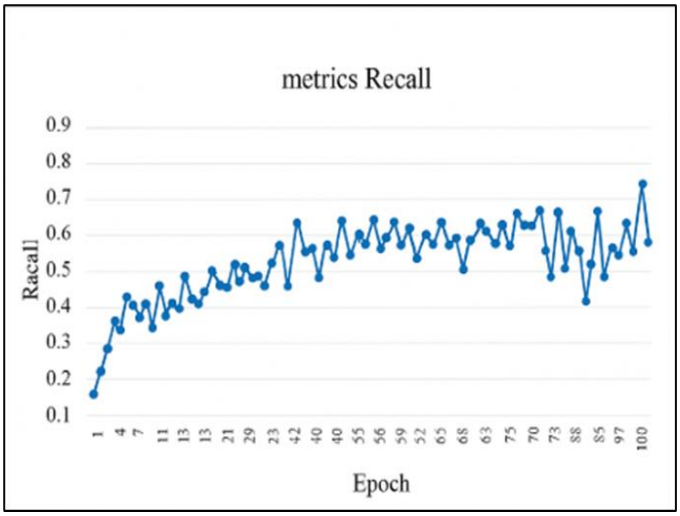

Figure 9: Recall curve

The results of the VGG19-based pneumonia detection model demonstrate strong capabilities in accurately classifying chest radiographs as pneumonia or normal. Leveraging transfer learning, the model benefits from robust feature extraction, resulting in high accuracy and reliable diagnostic predictions. Performance depends on factors such as the quality and diversity of the training data, the VGG19 architecture, and carefully tuned training parameters, including data augmentation and learning rate scheduling. Overall, the system offers a promising AI-driven solution that supports timely, precise identification of pneumonia,

which is crucial for enhancing patient care and resource allocation in healthcare settings. The results are when the user uploads an X-ray image, and the model predicts Normal.

(a) (b)

Figure 10: Normal X-rays

The output displays the results when the user uploads an X-ray image, and the predicted output is Abnormal. Figure 10 shows the output from normal chest X-rays using the VGG19 model; the network identifies clear lung fields, free of opacities or abnormal patterns. The model classifies these images as “Normal” with high confidence, indicating the absence of pneumonia-related features. The activation maps reveal minimal feature responses in lung regions, confirming the presence of normal structures. This accurate classification helps ensure reliable differentiation between healthy and infected cases, supporting precise medical diagnosis. The visualisation of intermediate layers shows that VGG19 effectively ignores irrelevant regions, such as bones and background noise, and focuses mainly on lung textures. The smooth activation patterns further validate that no pathological features are detected. This consistent behaviour across multiple test samples demonstrates the model's robustness in recognising healthy cases. Overall, it reinforces the network’s capability to maintain high specificity in normal chest X-ray classification.

(a) (b)

Figure 11: Pneumonia X-rays

Figure 11 illustrates the VGG19 model's output on pneumonia chest X-rays, showing strong activation in regions with visible lung opacities, inflammation, or consolidation patterns. The model classifies these images as “Pneumonia” with high confidence, indicating the presence of infection-related abnormalities. The feature maps highlight dense or cloudy lung areas,

which are key indicators of pneumonia. This output aids in early detection and supports clinicians in making faster, more accurate diagnostic decisions. The application of transfer learning with the VGG19 model for automated pneumonia detection offers remarkable potential in medical imaging diagnostics. This approach harnesses the strength of deep convolutional neural networks to accurately analyse chest radiographs and identify pneumonia cases with high precision. By leveraging pre-trained weights from large image datasets, the model adapts efficiently to the specific task, enabling reliable and fast classification even with relatively limited medical data. The integration of advanced data augmentation and fine-tuning techniques enhances the model's ability to generalise across diverse patient populations.

6. Conclusion

Using the VGG19 architecture for transfer learning has been highly effective for classifying chest X-ray images of pneumonia. VGG19 can extract rich, useful visual features such as edges, textures, and spatial patterns by leveraging pre-trained convolutional layers trained on vast image datasets. When fine-tuned on datasets specific to pneumonia, these features greatly improve classification accuracy while reducing the need for extensive training data and computing power. The model consistently achieves high precision and recall, indicating it can reliably identify true pneumonia cases while avoiding false positives and false negatives. This is very important for making clinical decisions. Optimisation methods, including data augmentation, learning rate scheduling, and early stopping, further improve the model's performance. Data augmentation makes the dataset more diverse and less likely to overfit, while learning rate scheduling makes sure that training converges quickly and steadily. Early stopping stops training when performance levels off, which makes it easier to generalise to new data. These methods work together to make the system more durable and ensure it behaves consistently across all imaging situations.

This transfer-learning-based method provides quick, reliable, and automated diagnostic support, helping doctors make decisions with all the information they need. Such systems can help identify problems early, thereby greatly improving patient outcomes by enabling quick intervention and treatment. VGG19 is also a good choice for other medical imaging applications beyond pneumonia detection, as it can be modified and expanded. Future improvements might involve more fine-tuning of higher-level layers and optimising the architecture to make it more efficient and less expensive to run. Adding more classes and severity levels to the model could help with tailored treatment planning. Lightweight model variants would enable their use in real time on edge devices in resource-constrained environments. Adding more varied datasets, possibly with the help of advanced GAN-based methods, can also improve generalisation. Lastly, easy integration with clinical workflows through user-friendly interfaces and connections to hospital information systems will make the product as useful as possible in the real world and have the greatest impact on healthcare.

Acknowledgement: The authors sincerely acknowledge the academic support, research facilities, and encouragement provided by S. A. Engineering College, SRM Institute of Science and Technology, Alliance University, Khyber Medical University, and Hindustan Institute of Technology and Science.

Data Availability Statement: The present study employs datasets associated with enhanced pneumonia detection using transfer learning approaches on chest radiograph images. The data supporting the conclusions of this research can be made available by the authors upon reasonable request.

Funding Statement: The authors confirm that this study was carried out without any external funding or financial assistance from public, private, or commercial sources.

Conflicts of Interest Statement: The authors collectively declare that there are no competing interests that could have influenced the research process or the reporting of results. All information sources have been properly acknowledged.

Ethics and Consent Statement: Appropriate ethical approval was obtained before initiating the study, and informed consent was acquired from all participating organizations and individuals involved in the data collection process.

References

1. N. Dey, Y. D. Zhang, V. Rajinikanth, R. Pugalenthi, and N. S. M. Raja, "Customized VGG19 architecture for pneumonia detection in chest X-rays," *Pattern Recognition Letters*, vol. 143, no. 3, pp. 67–74, 2021.
2. Q. An, W. Chen, and W. Shao, "A Deep Convolutional Neural Network for Pneumonia Detection in X-ray Images with Attention Ensemble," *Diagnostics*, vol. 14, no. 4, p. 390, 2024.
3. S. Singh, M. Kumar, A. Kumar, B. K. Verma, K. Abhishek, and S. Selvarajan, "Efficient pneumonia detection using Vision Transformers on chest X-rays," *Scientific Reports*, vol. 14, no. 1, p. 2487, 2024.
4. S. K. R. Mallidi, "Enhancing pneumonia diagnosis and severity assessment through deep learning: A comprehensive

approach integrating CNN classification and infection segmentation," *arXiv Preprint*, 2025. Available: https:// arxiv. org/abs/2502.06735 [Accessed by 10/10/2025].
5. M. Mujahid, F. Rustam, P. Chakrabarti, B. Mallampati, I. De La Torre Diez, P. Gali, V. Chunduri, and I. Ashraf, "Pneumonia detection on chest X-rays from Xception-based transfer learning and logistic regression," *Technol Health Care*, vol. 32, no. 6, pp. 3847–3870, 2024.
6. O. Dahmane, M. Khelifi, M. Beladgham, and I. Kadri, "Pneumonia detection based on transfer learning and a combination of VGG19 and a CNN built from scratch," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 3, pp. 1469–1480, 2021.
7. M. R. Hasan, S. M. A. Ullah, and S. M. R. Islam, "Recent Advancement of Deep Learning Techniques for Pneumonia Prediction from Chest X-Ray Image," *Medical Reports*, vol. 7, no. 5, pp. 1-15, 2024.
8. M. M. Khan, M. S. Ul Islam, A. A. Siddiqui, and M. T. Qadri, "Dual deterministic model based on deep neural network for the classification of pneumonia," *Intelligent Decision Technologies*, vol. 17, no. 3, pp. 641-654, 2015.
9. Y. Jahnav and C. Sivasankar, "Comparing VGG19 and VGG16 for improved pneumonia detection accuracy," *AIP Conference Proceedings*, vol. 3300, no. 1, p. 020025, 2025.
10. A. Jha, E. John, and T. Banerjee, "Transfer Learning for COVID-19 and Pneumonia Detection using Chest X-Rays," *2022 IEEE 65th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Fukuoka, Japan, 2022.
11. P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng, "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," *arXiv Preprint*, 2017. Available: https://arxiv.org/abs/1711.05225 [Accessed by 14/10/2024].
12. J. E. Luján-García, C. Yáñez-Márquez, Y. Villuendas-Rey, and O. Camacho-Nieto, "A Transfer Learning Method for Pneumonia Classification and Visualization," *Applied Sciences*, vol. 10, no. 8, pp. 1-18, 2020.
13. M. F. Hashmi, S. Katiyar, A. G. Keskar, N. D. Bokde, and Z. W. Geem, "Efficient pneumonia detection in chest X-ray images using deep transfer learning," *Diagnostics*, vol. 10, no. 6, pp. 1-23, 2020.
14. M. Salehi, R. Mohammadi, H. Ghaffari, N. Sadighi, and R. Reiazi, "Automated detection of pneumonia cases using deep transfer learning with paediatric chest X-ray images," *Br J Radiol*, vol. 94, no. 1121, pp. 1-8, 2021.
15. Y. Brima, M. Atemkeng, S. T. Djiokap, J. Ebiele, and F. Tchakounté, "Transfer Learning for the Detection and Diagnosis of Types of Pneumonia including Pneumonia Induced by COVID-19 from Chest X-ray Images," *Diagnostics*, vol. 11, no. 8, pp. 1-30, 2021.