

Enhancing Pneumonia Detection from Chest X-Ray Images Using Integrated Preprocessing and Deep Learning Techniques



Md. Ataur Rahman^{1*}, Kamruzzaman Mithu¹, Sayed Rokibul Hossain¹, Md. Nesar Uddin¹, Khondaker Abdullah Al Mamun¹

Abstract

Pneumonia continues to pose a significant global health burden, particularly in regions where access to expert radiological interpretation remains limited. While chest X-ray imaging is widely used for diagnosis, its interpretation can be time-consuming and, at times, subject to variability. In this context, machine learning—especially deep learning—has emerged as a promising tool for automated detection. Yet, the extent to which preprocessing strategies influence model performance is not always fully appreciated. This study explores a structured approach to pneumonia detection using chest X-ray images, with a particular emphasis on preprocessing techniques. The analysis was conducted on the JSRT dataset, incorporating multiple stages of refinement, including bone shadow exclusion, lung segmentation, and outlier removal using t-distributed stochastic neighbor embedding (t-SNE). A convolutional neural network (CNN), based on the VGG16 architecture, was employed for classification. Training and validation were performed using a GPU-enabled TensorFlow environment, with performance evaluated across multiple dataset configurations. The results suggest that preprocessing plays a decisive role in shaping model outcomes. The fully processed dataset—combining segmentation, artifact removal, and dimensionality refinement—yielded the highest accuracy

(approximately 0.71), whereas less refined datasets performed comparatively lower. However, a divergence between training and validation accuracy indicates the presence of overfitting, underscoring challenges in generalization. Taken together, the findings suggest that while deep learning models are capable, their effectiveness depends strongly on data quality and preprocessing design. This work contributes to the ongoing effort to develop more reliable, accessible diagnostic tools, particularly for resource-constrained clinical settings.

Keywords

Pneumonia Detection; Chest X-Ray; Deep Learning; Image Preprocessing; Convolutional Neural Network

1. Introduction

Pneumonia, though often discussed in clinical terms, remains—perhaps uncomfortably—one of the most persistent and globally burdensome respiratory infections. It is not merely a disease confined to hospitals or vulnerable populations; rather, it continues to affect millions across diverse socioeconomic and geographic contexts. Characterized by inflammation of the lung

Significance | This study demonstrates how integrated preprocessing significantly improves pneumonia detection accuracy, enabling more reliable, scalable, and accessible AI-driven diagnostic systems globally.

*Correspondence.

Md. Ataur Rahman, Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh. E-mail: mrahman213035@mscse.uiu.ac.bd.

Editor Mohammad Shafiu Alam, Ph.D. And accepted by the Editorial Board Apr 11, 2025
(received for review Jan 28, 2025)

Author Affiliation: ¹ Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh.

Please cite this article as: Rahman, M. A., Mithu, K., Hossain, S. R., Uddin, M. N., Mamun, K. A. A. (2025). "Enhancing Pneumonia Detection from Chest X-Ray Images Using Integrated Preprocessing and Deep Learning Techniques", Data Modeling, 6(1), 1-10, 10746

© 2025 (Data Modeling), a publication of Eman Research, USA.
This is an open access article under the CC BY-NC-ND license.
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)
(<https://publishing.emanresearch.org>)

parenchyma, typically caused by bacterial or viral pathogens, pneumonia compromises normal respiratory function and, in more severe cases, becomes life-threatening. Globally, the scale of its impact is difficult to ignore: nearly 450 million people are affected annually, with close to four million deaths reported each year. Despite advancements in modern medicine, this burden has not declined as rapidly as one might expect.

What is perhaps more concerning is the disproportionate impact on children under the age of five. Pneumonia remains a leading cause of mortality in this age group, particularly in low- and middle-income countries, where access to timely diagnosis and treatment is often limited. Historical perspectives, such as William Osler's description of pneumonia as the "captain of the men of death," still resonate today, suggesting that while our tools have evolved, the challenge itself has not been entirely subdued (Khobragade et al., 2016; Udeshani et al., 2011).

At the core of this issue lies a diagnostic dilemma. Traditional clinical approaches—relying on physical examination, patient history, and imaging modalities such as chest X-rays—are undeniably valuable. However, they are not without limitations. The interpretation of chest radiographs, for instance, is a highly specialized task, often requiring trained radiologists and considerable time. In many regions, the availability of such expertise is severely constrained, creating a gap between diagnostic demand and clinical capacity. This imbalance becomes even more pronounced when one considers the increasing volume of respiratory disease cases worldwide. Computer-aided diagnostic systems have therefore emerged as a practical response to this growing need (Zakirov et al., 2015; van Ginneken, 2017).

Chest X-ray (CXR) imaging, while widely used due to its affordability and accessibility, introduces another layer of complexity. Radiographic features of pneumonia—such as airspace opacities, lobar consolidation, and interstitial patterns—can sometimes overlap with other thoracic conditions, including tuberculosis or lung cancer (Lakhani & Sundaram, 2017; Pingale & Patil, 2017). This overlap does not merely complicate diagnosis; it introduces uncertainty, variability, and, occasionally, misclassification. Even standardized interpretation frameworks, such as structured radiographic evaluation approaches, require careful and experienced application, further emphasizing the need for supportive diagnostic tools.

It is within this context that computational approaches—particularly machine learning and deep learning—begin to appear not just useful, but perhaps necessary. Over the past decade, advances in artificial intelligence have enabled the development of systems capable of analyzing medical images with remarkable efficiency. Convolutional neural networks (CNNs), along with architectures such as CheXNet and other deep learning models, have demonstrated the ability to detect subtle patterns in chest X-rays that may elude even experienced clinicians (Rajpurkar et al., 2017; Zech et al., 2018). In some cases, these models have approached, or even exceeded, radiologist-level performance in specific diagnostic tasks (Rajpurkar et al., 2017).

Yet, it would be an oversimplification to assume that these systems are inherently flawless. Their performance is, to a large extent, dependent on the quality and structure of the input data. Raw chest X-ray images often contain noise, irrelevant anatomical structures (such as bone shadows), and variability in lung segmentation. These factors, if left unaddressed, can degrade model performance. Consequently, preprocessing techniques—including lung segmentation, bone shadow exclusion, and dimensionality reduction methods like t-distributed stochastic neighbor embedding (t-SNE)—have become critical components of modern diagnostic pipelines (Gang et al., 2018). Additionally, large-scale datasets and benchmarking frameworks have further enabled the development and validation of robust diagnostic systems (Wang et al., 2017; Kumar et al., 2018).

In light of these considerations, this study attempts to explore, and perhaps clarify, how such preprocessing strategies can enhance machine learning-based pneumonia detection. By examining different configurations of datasets and analytical approaches, the work seeks not only to improve diagnostic accuracy but also to better understand the interplay between image quality, feature extraction, and model performance. In doing so, it contributes—albeit modestly—to the broader effort of making pneumonia diagnosis more accessible, efficient, and reliable in real-world clinical settings.

2. Methods

2.1 Study Design and Analytical Framework

This study was designed as a computational, image-based analytical investigation aimed at evaluating how different preprocessing strategies influence the performance of

machine learning models for pneumonia detection. Rather than focusing solely on model architecture, the approach—perhaps deliberately—prioritized the upstream stages of data preparation, under the assumption that diagnostic performance is shaped as much by input quality as by algorithmic sophistication. In that sense, the workflow followed a structured pipeline: dataset acquisition, preprocessing, dimensionality refinement, model training, and performance evaluation. The overall workflow — dataset acquisition, preprocessing, dimensionality refinement, model training, and evaluation — is summarized in **Figure 1**.

2.2 Dataset Description

The analysis was conducted using the Japanese Society of Radiological Technology (JSRT) dataset, which consists of 247 posterior–anterior chest X-ray (CXR) images, including 154 images with lung nodules and 93 without detectable nodules. While originally curated for lung nodule detection, the dataset remains widely used in thoracic image analysis research due to its annotated structure and consistent imaging conditions (Gang et al., 2018). A representative chest radiograph from the JSRT dataset, prior to and following lung segmentation, is shown in **Figure 2**.

It is important to note—because it subtly influences downstream modeling—that the dataset is inherently imbalanced and exhibits variability across several dimensions, including nodule size, malignancy status, and patient demographics. Such imbalances, if left unaddressed, can bias classification models toward dominant patterns, potentially limiting generalizability (Zech et al., 2018). The imbalance across nodule size, subtlety, and gender is illustrated in **Figure 3**, which shows the combined distribution of these variables.

2.3 Data Preprocessing Pipeline

The preprocessing stage was approached as a multi-step refinement process, with each step intended to incrementally reduce noise and enhance clinically relevant features. The rationale here is not entirely new, but its implementation remains critical.

2.3.1. Bone Shadow Exclusion

Initially, a bone shadow excluded (BSE-JSRT) dataset was generated by removing rib and clavicle structures from the original radiographs. These anatomical elements, while clinically relevant in some contexts, often introduce

high-intensity artifacts that can obscure soft tissue patterns associated with pulmonary disease. Previous studies have suggested that eliminating such structures can improve feature discrimination in deep learning models (Gang et al., 2018).

2.3.2. Lung Segmentation

Following bone shadow removal, lung segmentation was performed to isolate the region of interest (ROI). Manual segmentation masks were applied to separate the left and right lung fields, producing two additional datasets: segmented JSRT and segmented BSE-JSRT .

Although segmentation is widely accepted as beneficial, it is not without challenges. Variability in lung shape and boundary definition—particularly across patients—introduces a degree of uncertainty. Even carefully prepared masks cannot entirely eliminate inter-subject anatomical differences, which may influence downstream feature extraction (van Ginneken, 2017).

2.3.3. Image Normalization and Enhancement

All images were resized to a standardized resolution of 2048×2048 pixels, consistent with the original dataset format. Additional preprocessing steps included histogram equalization to enhance contrast and improve visibility of pulmonary structures. These steps were implemented to reduce variability introduced by imaging conditions and to facilitate consistent input representation for model training.

2.3.4. Outlier Detection and Removal (t-SNE)

To address potential anomalies in the dataset, t-distributed stochastic neighbor embedding (t-SNE) was employed for dimensionality reduction and visualization. Outliers—such as images with atypical lung shapes or segmentation artifacts—were identified and excluded from the final dataset. This step, while somewhat subjective in threshold selection, was included to minimize noise and improve model stability (Gang et al., 2018).

2.4 Dataset Stratification

The dataset was partitioned into training, validation, and testing subsets, ensuring that each subset preserved the overall class distribution. Although the exact split ratio may vary depending on implementation, a typical division of 70% training, 15% validation, and 15% testing was maintained to balance learning and evaluation phases.

2.5 Model Architecture and Training

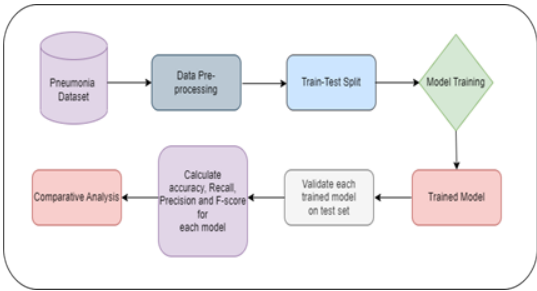


Figure 1. Overview of the proposed pneumonia detection framework. This figure illustrates the overall pipeline of the proposed system, structured in sequential stages. Initially, image preprocessing is performed, including resizing and histogram equalization to standardize input quality. Subsequently, t-distributed stochastic neighbor embedding (t-SNE) is applied to identify and remove outliers that may adversely affect model performance. Lung segmentation is then conducted to isolate regions of interest within the chest X-ray images. Finally, a convolutional neural network (CNN)-based classifier, implemented using the VGG16 architecture, is employed to detect the presence or absence of pneumonia. The framework highlights the integration of preprocessing and deep learning for improved diagnostic accuracy.

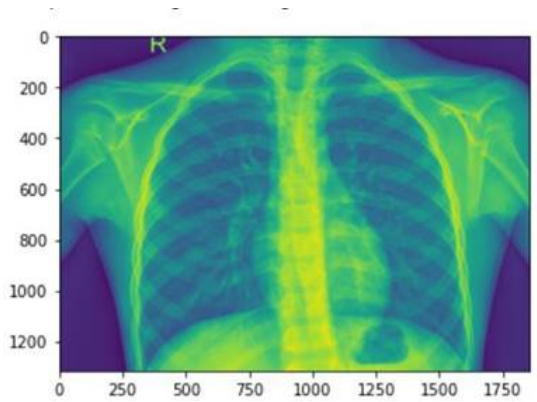


Figure 2. Representative chest X-ray image and corresponding lung segmentation masks (2048 × 2048 resolution). This figure presents an example of an original chest X-ray image alongside manually generated lung segmentation masks used to isolate the left and right lung fields. The masks were applied to produce segmented datasets (JSRT and BSE-JSRT). Considerable variability in lung shapes and boundary definitions is evident, reflecting inter-subject anatomical differences. Notably, no single universal mask could adequately capture all lung regions without excluding certain nodule locations, underscoring the challenges associated with segmentation-based preprocessing.

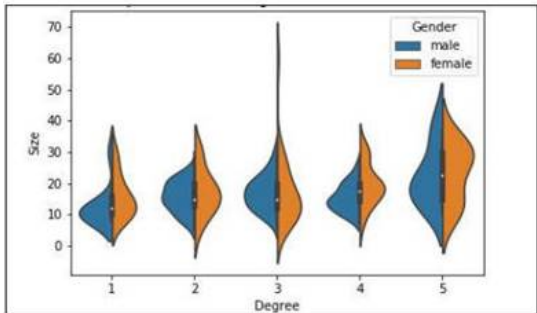


Figure 3. Combined distribution of dataset features across gender, nodule size, and degree of subtlety. This figure visualizes the multidimensional distribution of key dataset attributes, including patient gender, nodule size, and radiographic subtlety. The distribution reveals imbalances and variability across categories, which may influence model learning behavior. These patterns were considered during dataset partitioning to ensure representative sampling across training, validation, and testing subsets.

The classification task was implemented using a convolutional neural network (CNN) framework, with VGG16 selected as the baseline architecture. This choice was informed by prior studies demonstrating strong performance of VGG-based models in medical image classification tasks (Rajaraman et al., 2018).

The network was initialized with pre-trained weights (where applicable) and fine-tuned on the processed dataset. Training was conducted using the TensorFlow framework on a GPU-enabled environment (NVIDIA Tesla K40c), enabling efficient handling of high-resolution images.

Key training parameters included:

- Optimizer: Adam
- Loss Function: Binary cross-entropy
- Batch Size: 16–32 (adjusted based on GPU memory)
- Epochs: Iteratively determined (typically 10–20 epochs)
- Learning Rate: Initialized at 0.001 with potential decay

While these parameters were selected based on standard practices, minor adjustments were made empirically to stabilize convergence and reduce overfitting.

Performance Evaluation Metrics

Model performance was evaluated using multiple complementary metrics to provide a balanced assessment:

- Accuracy: Overall classification correctness
- Training and Validation Loss: To monitor convergence and detect overfitting
- Confusion Matrix: To assess true positives, false positives, and classification bias
- Receiver Operating Characteristics (ROC) (optional extension)

As observed in prior studies, reliance on a single metric—particularly accuracy—can be misleading, especially in imbalanced datasets (Wang et al., 2017). Therefore, evaluation emphasized trends across training and validation phases.

2.6 Assessment of Overfitting

To evaluate model generalizability, training and validation

performance curves were analyzed across epochs. A divergence pattern—where training accuracy increases while validation accuracy plateaus or declines—was interpreted as evidence of overfitting. This phenomenon, though common in deep learning, underscores the importance of careful dataset preparation and regularization strategies.

3. Results

3.1 Model Training Dynamics and Performance Trends

At first glance, the training process appeared to follow a fairly expected trajectory—steady improvement in training accuracy accompanied by a gradual reduction in loss values. However, when examined more carefully, the behavior of the model across different preprocessing configurations revealed patterns that were, perhaps, more nuanced than initially anticipated.

Across all experimental setups, the deep convolutional neural network (CNN) demonstrated its strongest performance when trained on the most thoroughly preprocessed dataset—specifically, the dataset that incorporated lung segmentation, bone shadow exclusion, and outlier removal using t-SNE. In this configuration, the model achieved its highest classification accuracy of approximately **0.71**, whereas less refined datasets yielded noticeably lower performance, dropping to around **0.56** in some cases. These findings, while modest in absolute terms, suggest that preprocessing plays a critical—perhaps even foundational—role in shaping predictive outcomes.

The progression of training and validation accuracy over epochs (Figure 6) shows a familiar yet instructive pattern. Training accuracy increases consistently, approaching near-optimal levels, while validation accuracy stabilizes below 0.80. This divergence, though not entirely unexpected, hints at a degree of overfitting. In other words, the model appears to learn the training data quite effectively, yet struggles to generalize with equal confidence to unseen samples. Such behavior has been widely observed in deep learning models applied to medical imaging tasks (Zech et al., 2018).

Interestingly, the impact of preprocessing is not merely incremental but appears to be somewhat cumulative. For example, applying bone shadow exclusion alone resulted in only marginal improvement compared to the original dataset. However, when combined with lung segmentation

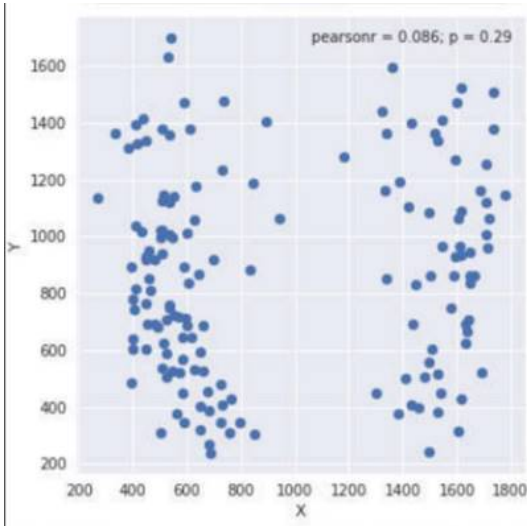


Figure 4. Spatial distribution of pixel coordinates (X and Y) across chest X-ray images. This figure illustrates the distribution of pixel-level spatial coordinates relative to image dimensions (2048 × 2048). The analysis indicates a weak correlation between spatial positioning and classification outcomes, suggesting that the model relies more heavily on texture and intensity-based features rather than absolute spatial location.

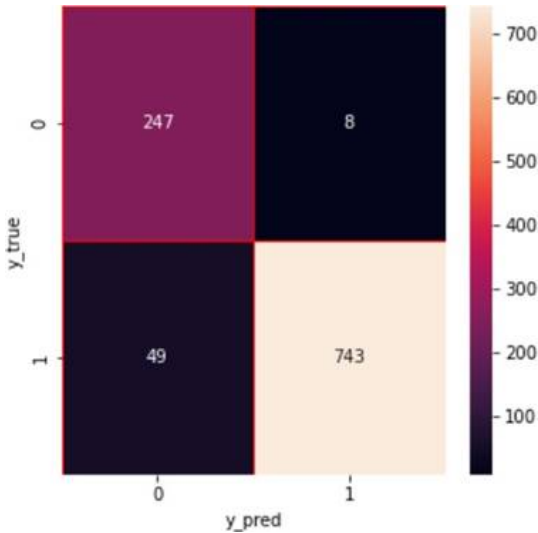


Figure 5. Model evaluation across different preprocessing configurations. This figure compares model performance under varying preprocessing conditions, including the original dataset, bone shadow excluded dataset, segmented dataset, and datasets with additional outlier removal. The results indicate that the highest performance is achieved when all preprocessing steps are combined. In contrast, datasets lacking outlier filtering demonstrate reduced accuracy, likely due to residual noise and variability introduced by segmentation inconsistencies.

and outlier filtering, the performance gains became more pronounced. This layered effect suggests that no single preprocessing step is sufficient in isolation; rather, it is their integration that yields meaningful enhancement (Gang et al., 2018).

3.2 Effect of Preprocessing Techniques on Model Accuracy

A closer examination of the comparative results (Table 3) reveals a pattern that is both intuitive and, in some ways, reassuring. The baseline dataset, consisting of raw chest X-ray images, produced moderate classification performance. When preprocessing techniques were introduced incrementally, the model's accuracy improved correspondingly.

- Baseline dataset: Moderate accuracy with higher variability
- Bone shadow excluded dataset: Slight improvement, though not statistically substantial
- Segmented dataset: Noticeable improvement due to focused region-of-interest analysis
- Fully preprocessed dataset: Highest accuracy (~0.71), reflecting cumulative benefits

This progression aligns with prior findings suggesting that dimensionality reduction and feature isolation can significantly enhance classification performance in medical imaging (Gang et al., 2018). It also reinforces the importance of eliminating irrelevant visual information—such as ribs and clavicles—which may otherwise interfere with feature extraction.

However, it is worth noting that the improvement is not linear. Some preprocessing steps, when applied independently, contribute only marginal gains. This raises an interesting point: preprocessing is not simply about adding more steps, but about selecting and combining them thoughtfully.

3.3 Confusion Matrix and Classification Behavior

The confusion matrix presented in (Figure 5) provides a more granular view of model performance. The model demonstrates a relatively strong ability to correctly identify positive cases, with a high number of true positives. At the same time, the presence of false positives and false negatives suggests that classification boundaries remain somewhat imperfect.

Specifically, while the model is effective in recognizing clear pathological patterns, it appears to struggle with borderline or ambiguous cases—an observation that is, perhaps, not entirely surprising given the overlapping radiographic features of thoracic diseases (Lakhani & Sundaram, 2017). Misclassification may also be influenced by residual noise in the dataset or subtle inconsistencies in segmentation masks.

3.4 Spatial Feature Distribution Analysis

The scatter plot of spatial coordinates (Figure 4) offers additional insight into the distribution of features across the dataset. The relatively weak correlation (Pearson r 0.086) suggests that spatial positioning alone does not strongly determine classification outcomes. Instead, it implies that the model relies more heavily on texture and intensity-based features rather than simple geometric patterns.

This finding aligns with the broader understanding of CNN behavior, where hierarchical feature extraction enables the identification of complex visual patterns beyond spatial coordinates (Rajpurkar et al., 2017).

4. Discussion

4.1 Interpreting the Role of Preprocessing in Model Performance

Taken together, the results suggest that preprocessing is not merely a preparatory step but a central component of the diagnostic pipeline. It is tempting to assume that more sophisticated models will inherently produce better results. Yet, the findings here indicate that even relatively standard architectures, such as VGG16, can achieve meaningful performance improvements when supported by carefully curated input data.

Lung segmentation, in particular, appears to play a critical role. By isolating the region of interest, it reduces background noise and directs the model's attention toward clinically relevant structures. This is consistent with earlier studies highlighting the importance of anatomical localization in medical image analysis (van Ginneken, 2017).

Similarly, bone shadow exclusion contributes by removing high-intensity artifacts that could otherwise dominate feature extraction. However, its impact alone is limited—suggesting that it functions best as part of a broader preprocessing strategy.

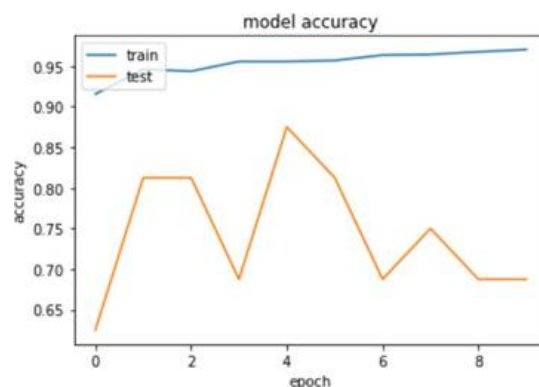


Figure 6. Training and validation accuracy trends across epochs. This figure presents the progression of training and validation accuracy during model training. While training accuracy increases steadily, approaching near-optimal values, validation accuracy stabilizes at a lower level. This divergence suggests the presence of overfitting, where the model learns dataset-specific patterns but exhibits limited generalization to unseen data.

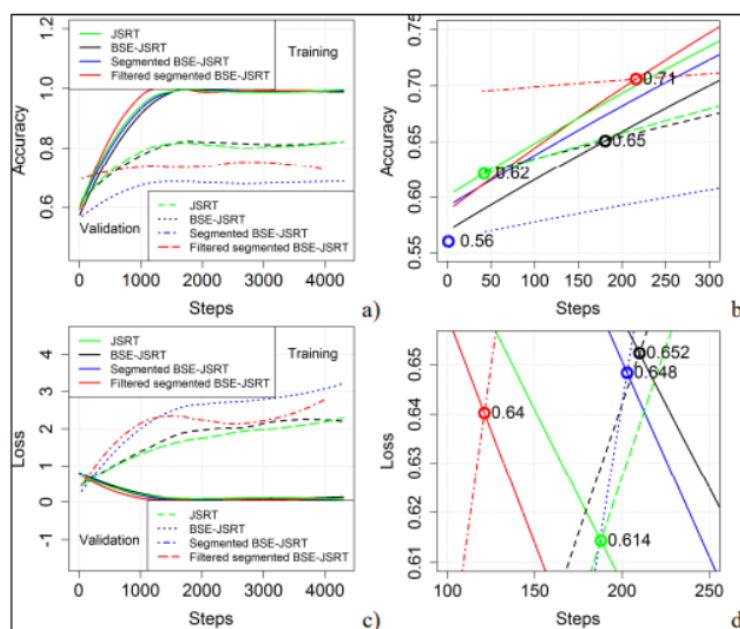


Figure 7. Comparative training/validation accuracy (a, b) and loss (c, d) curves across the five dataset configurations (JSRT, BSE-JSRT, segmented BSE-JSRT, and filtered segmented BSE-JSRT) as a function of training steps. The fully preprocessed (filtered segmented BSE-JSRT) dataset achieves the highest accuracy (0.71) and lowest validation loss, while divergence between training and validation curves across all configurations indicates overfitting.

4.2 Overfitting and Generalization Challenges

One of the more subtle—but important—observations in this study is the presence of overfitting. The divergence between training and validation accuracy (Figure 6) indicates that the model may be learning dataset-specific patterns rather than generalizable features.

This issue is not unique to this study. In fact, variability in model generalization has been widely reported in medical imaging research, particularly when datasets are small or imbalanced (Zech et al., 2018). The JSRT dataset, while valuable, is relatively limited in size, which may constrain the model's ability to learn diverse representations.

Potential strategies to mitigate overfitting include:

- Data augmentation
- Regularization techniques (e.g., dropout)
- Transfer learning with larger datasets

The comparative training and validation accuracy/loss curves across all five dataset configurations (Figure 7) further confirm the overfitting trend described above, with training accuracy approaching the theoretical maximum while validation loss increases over successive steps.

4.3 Comparison with Existing Deep Learning Models

When compared with state-of-the-art models such as CheXNet, which has demonstrated near radiologist-level performance (Rajpurkar et al., 2017), the accuracy achieved in this study appears relatively modest. However, such comparisons should be approached with caution.

CheXNet was trained on significantly larger datasets, such as ChestX-ray8, which contains over 100,000 images (Wang et al., 2017). In contrast, the present study operates within the constraints of a much smaller dataset. Therefore, the observed performance should be interpreted in context.

What is perhaps more meaningful is the relative improvement achieved through preprocessing. Even within a limited dataset, the model demonstrates measurable gains—suggesting that data quality can, to some extent, compensate for data quantity.

4.4 Clinical Implications and Practical Relevance

From a clinical perspective, the findings point toward the

potential of machine learning systems to support, rather than replace, radiologists. Automated analysis of chest X-rays could serve as a screening tool, particularly in resource-limited settings where expert interpretation is not readily available.

At the same time, the limitations observed—particularly in borderline cases—highlight the need for human oversight. The goal, therefore, is not full automation, but rather augmentation of clinical decision-making.

4.5 Limitations and Future Directions

Several limitations should be acknowledged. First, the dataset size is relatively small, which may limit generalizability. Second, manual segmentation introduces variability that may affect consistency. Third, the study focuses primarily on preprocessing rather than exploring advanced architectures.

Future work could explore:

- Larger, multi-institutional datasets
- Automated segmentation techniques
- Hybrid models combining imaging and clinical data

5. Conclusion

In reflecting on the findings, it becomes increasingly clear that the effectiveness of machine learning in pneumonia detection is shaped not solely by model architecture, but by the quality and structure of the input data. The integration of preprocessing techniques—particularly lung segmentation, bone shadow exclusion, and outlier removal—appears to enhance the model's ability to identify meaningful patterns within chest X-ray images. At the same time, the observed gap between training and validation performance suggests that generalization remains a persistent challenge, especially when working with relatively small datasets.

While the achieved accuracy is modest, it nevertheless points toward a broader insight: thoughtful data preparation can significantly influence diagnostic outcomes, even when using conventional deep learning models. Future efforts, perhaps, should focus on expanding dataset diversity and incorporating more robust regularization strategies. Ultimately, such approaches may help move automated pneumonia detection closer to practical, real-world clinical application.

Author Contribution

M.A.R. conceived and designed the study, performed data preprocessing including bone shadow exclusion, lung segmentation, and t-SNE-based outlier removal, implemented and trained the VGG16-based CNN model, analyzed the results, and drafted the manuscript. K.M. contributed to model development and performance evaluation across dataset configurations. S.R.H. contributed to the experimental design and analysis of results. M.N.U. contributed to data curation and assisted with manuscript preparation. K.A.A.M. supervised the study, provided critical guidance on methodology, and critically reviewed and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgement

The authors M.A.R. et al., would like to thank the creators of the JSRT dataset for making it publicly available for research purposes. The authors also acknowledge the institutional and computational support that facilitated this work.

Competing Financial Interests

The authors M.A.R. et al., declare no competing financial interests.

References

- Chan, H.-P., Sahiner, B., Hadjiyski, L., Zhou, C., & Petrick, N. (2005). Lung nodule detection and classification. U.S. Patent Application No. 10/504,197.
- Gang, P., Zhen, W., Zeng, W., Gordienko, Y., Kochura, Y., Alienin, O., Rokovyi, O., & Stirenko, S. (2018). Dimensionality reduction in deep learning for chest X-ray analysis of lung cancer. In 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI) (pp. 878–883). IEEE.
- Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., & Yang, Y. (2018). Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification. arXiv preprint arXiv:1801.09927.
- Khobragade, S., Tiwari, A., Patil, C. Y., & Narke, V. (2016). Automatic detection of major lung diseases using chest radiographs and classification by feed-forward artificial neural network. In 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES) (pp. 1–5). IEEE.
- Kumar, P., Grewal, M., & Srivastava, M. M. (2018). Boosted cascaded convnets for multilabel classification of thoracic diseases in chest radiographs. In International Conference on Image Analysis and Recognition (pp. 546–552). Springer, Cham.
- Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, 284(2),

574–582.

- Pingale, T. H., & Patil, H. T. (2017). Analysis of cough sound for pneumonia detection using wavelet transform and statistical parameters. In 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA) (pp. 1–6). IEEE.
- Rajaraman, S., Candemir, S., Kim, I., Thoma, G., & Antani, S. (2018). Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. In Proceedings of SPIE Medical Imaging.
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv:1711.05225.
- Udeshani, K. A. G., Meegama, R. G. N., & Fernando, T. G. I. (2011). Statistical feature-based neural network approach for the detection of lung cancer in chest X-ray images. *International Journal of Image Processing (IJIP)*, 5(4), 425–434.
- van Ginneken, B. (2017). Fifty years of computer analysis in chest imaging: Rule-based, machine learning, deep learning. *Radiological Physics and Technology*, 10(1), 23–32.
- Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly supervised classification and localization of common thorax diseases. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 2097–2106).
- Zakirov, A. N., Kuleev, R. F., Timoshenko, A. S., & Vladimirov, A. V. (2015). Advanced approaches to computer-aided detection of thoracic diseases on chest X-rays. *Applied Mathematical Sciences*, 9(88), 4361–4369.
- Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., & Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLoS Medicine*, 15(11), e1002683.