

Explainable deep learning framework for classifying pathogen-induced chest infections with fuzzy severity reasoning

Nirmala Bai Limba Naik^{1,2}, Manju Devi¹

¹Department of Electronics and Communication Engineering, The Oxford College of Engineering affiliated to Visvesvaraya Technological University, Belagavi, India

²Department of Electronics and Instrumentation Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, India

Article Info

Article history:

Received Nov 10, 2025

Revised Jul 20, 2026

Accepted Aug 12, 2026

Keywords:

Chest X-ray

Deep learning

Fuzzy reasoning system

Model explainability

Pathogen-induced disease

ABSTRACT

Accurate classification of pathogen-induced chest infections is a vital step in enabling rapid diagnosis in the context of a high clinical burden scenario. This paper proposes an explainable deep learning (DL) framework that employs a hybrid approach by integrating an efficient classification model with both decision-level and feature-level interpretability. A neural search architecture is adopted to develop a fine-tuned EfficientNet model for multi-class classification of chest radiographs subjected to seven classes. To enhance clinical trust, a Mamdani-type fuzzy-logic-based severity reasoning layer is introduced to provide contextual assessments of disease severity and possible co-infections. Furthermore, feature-level explanations are enabled using gradient-weighted class activation mapping (Grad-CAM) to visualize important regions that contribute to each prediction. The model is validated on a comprehensive dataset, and the experimental outcome demonstrated the effectiveness of the proposed framework with classification accuracy of 96% and sensitivity of 95%.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Nirmala Bai Limba Naik

Department of Electronics and Communication Engineering, The Oxford College of Engineering affiliated to Visvesvaraya Technological University

10th Milestone, Hosur Rd, Bommanahalli, Bengaluru, Karnataka-560068, India

Email: nirmala.pari@gmail.com

1. INTRODUCTION

Chest infections represent a major cause of disease and mortality worldwide, with millions of reported cases each year and a significant burden on healthcare systems, and these infections are caused by diverse pathogens including viruses, bacteria, fungi, and parasites that often result in severe respiratory complications [1]. In recent times, viral chest infections such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), Middle East respiratory syndrome coronavirus (MERS-CoV), and influenza gained global attention due to rapid transmission and high mortality rates [2], [3]. Therefore, early identification of chest infections remains essential; however, determining the exact infection type or subtype is quite challenging even for experienced radiologists because of overlapping clinical symptoms and radiographic manifestations, including cough, fever, chest pain, and shortness of breath, particularly between viral and bacterial conditions [4]. The conventional diagnosis is dependent on clinical history, laboratory investigations, and chest X-ray imaging [5]; however, this combined approach faces practical constraints since laboratory tests are time-consuming, costly, and often inaccessible in low-resource settings, and polymerase chain reaction (PCR)-based testing reported false-negative results during early stages of infection [6]. Furthermore, radiographic interpretation is heavily dependent on human expertise, which may

introduce inter-observer variability and inconsistent diagnostic decisions. To overcome these limitations, deep learning (DL) approaches, especially convolutional neural networks (CNNs), were applied for automated chest image analysis [7], [8]; however, most existing DL models require large annotated datasets, remain sensitive to class imbalance, and operate as black-box systems, and the absence of interpretability restricts clinician understanding of prediction reasoning, thereby raising ethical and regulatory concerns in high-risk medical environments [9], [10].

In past recent years, several research works examined chest infection diagnosis from radiographic images, where most proposed approaches were dependent on transfer learning-based CNNs. A significant portion of existing studies focused on coronavirus disease 2019 (COVID-19) and viral pneumonia detection using chest X-ray images. The authors [11], [12] applied CNN-based transfer learning models for classification of COVID-19, bacterial pneumonia, viral pneumonia, and normal cases and reported high accuracy under limited class settings; however, the classification performance reduced when the problem was extended to multi-class scenarios. Ensemble CNN strategies were also explored, where Rajpoot *et al.* [13] integrated multiple deep architectures and emphasized feature-level explainability using LIME, SHAP, and gradient-weighted class activation mapping (Grad-CAM)++, although the framework remained dependent on single-label prediction and did not explicitly address diagnostic uncertainty. The discrimination of overlapping viral infections such as COVID-19, Middle East respiratory syndrome (MERS), and severe acute respiratory syndrome (SARS) was addressed [14], yet the study remained constrained to viral subclasses and controlled datasets. There are also several works focuses on pneumonia classification and clinical applicability. The work of Gu *et al.* [15] introduced a computer-aided diagnosis system for pediatric pneumonia that applied lung segmentation followed by CNN feature extraction and support vector machine (SVM) classification, whereas Barakat *et al.* [16] highlighted diagnostic complications and inter-observer variability in pediatric chest X-ray interpretation, which emphasized the dependency on human expertise. Bakir *et al.* [17] evaluated deep features extracted from different CNN architectures and reported reduced performance in multi-class settings compared to binary classification, reflecting the increased complexity of generalized chest infection diagnosis. Tuberculosis detection from chest X-ray images received focused attention in recent studies, where Kotei and Thirunavukarasu [18] combined a data-efficient image transformer with a CNN architecture for tuberculosis diagnosis, and Induri *et al.* [19] applied a ResNet50-based DL method and reported strong performance. In contrast, studies addressing fungal pneumonia and co-infection scenarios remained limited, with Koesmarsono *et al.* [20] examining chest X-rays of immunocompromised patients and reporting a weak correlation between radiographic signs and confirmed fungal infections, and Wu *et al.* [21] explored computed tomography (CT)-based indicators for bacterial and fungal co-infection in pneumocystis pneumonia. Overall, existing studies mainly addressed isolated disease categories using black-box prediction models, and very few works supported multi-class chest infection diagnosis with explicit reasoning for co-infections, diagnostic uncertainty, and decision-level interpretability using chest radiographs.

This paper proposes explainable DL framework for multi-class classification of pathogen-induced chest infections using chest radiographs. The proposed framework focuses on addressing the limitations of existing black-box DL models by integrating classification performance with both feature-level and decision-level interpretability for effective clinical decision support. The main contributions of the proposed research work are as follows: i) a construction of a comprehensive seven-class chest X-ray dataset by integrating multiple publicly available datasets with common and rare pathogen-induced chest infections; ii) a fine-tuned EfficientNet-based classification architecture is introduced, where the network is optimized through a lightweight neural architecture search (NAS) process with explicit focus on classification accuracy and computational efficiency; iii) a fuzzy logic-based severity reasoning mechanism is proposed to model decision-level interpretability, and feature-level interpretability is presented using Grad-CAM mechanism.

2. METHOD

The primary objective of the system is to develop a DL framework that classifies multiple chest infections with high accuracy and at the same time provides transparent and interpretable decision support. The overall architecture of the proposed methodology is illustrated in Figure 1, and the modelling of the framework is organized systematically into four core components: i) image preprocessing, ii) dataset partitioning, iii) chest infection classification using a NAS-driven EfficientNetB0 CNN model, and iv) explainable inference with comprehensive model evaluation. However, unlike conventional black-box artificial intelligence (AI) models, the proposed framework integrates explainability features that indicate the reasoning behind model predictions, and such transparency becomes important in hospital environments where diagnostic decisions directly affect patient outcomes. In addition, the system is capable of identifying severity levels and generating clinically relevant interpretations, which can function as a clinical decision

support mechanism for early identification and classification of pathogen-induced diseases, and it can assist in hospital resource optimization by prioritizing urgent and high-risk cases.

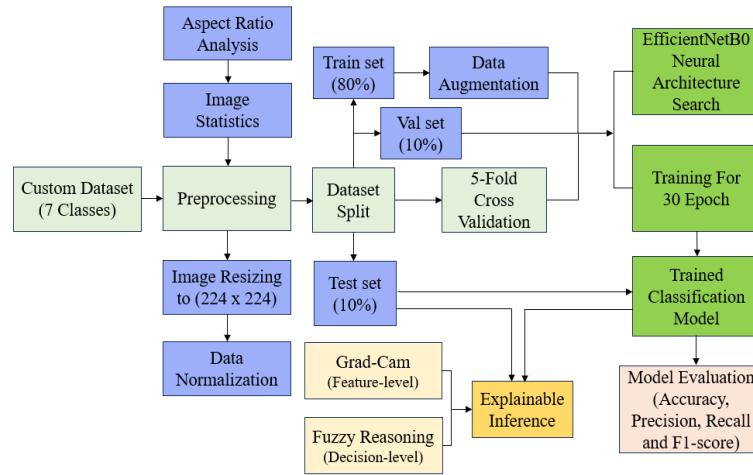


Figure 1. The workflow of the proposed methodology

2.1. Dataset description

The study introduces a custom multi-class dataset constructed by combining multiple publicly available chest radiograph datasets obtained from authentic sources on Kaggle and Mendeley, with the objective of incorporating a broader spectrum of common as well as rare pathogens responsible for severe chest infections. The integration of these datasets is clinically acceptable because all selected sources consist of frontal chest X-ray images acquired under comparable radiographic views and standard diagnostic protocols routinely adopted in practice, and to reduce inter-dataset variability, all images undergo uniform preprocessing including resizing and intensity normalization to maintain consistency across sources. For the COVID-19, normal, and viral pneumonia classes, the study uses the widely referenced COVID-19 radiography database [22], which contains 3,616 COVID-19 X-rays, 10,192 normal images, and 1,345 viral pneumonia samples. The bacterial pneumonia class includes 2,780 chest X-rays from the chest X-ray pneumonia dataset [23], which also provides 1,493 additional Viral Pneumonia images that are merged with the earlier set to form a unified viral pneumonia category. To ensure sufficient representation of tuberculosis, data are integrated from the tuberculosis chest X-ray database [24] consisting of 700 images and the tuberculosis chest X-rays dataset [25] from Mendeley, contributing 2,494 additional samples. For less common respiratory infections such as SARS and MERS, images are obtained from the SARS-MERS X-ray image dataset [26], which includes 134 SARS and 144 MERS X-rays. The inclusion of these diverse classes addresses the diagnostic complexity associated with pathogen-induced chest infections and accounts for clinically important but less frequently encountered diseases, thereby reflecting the heterogeneity observed in real-world clinical settings. Table 1 presents the class-wise distribution of images, and Figure 2 illustrates representative samples from the custom dataset for both normal and infected cases.

Table 1. Class-wise image distribution in the curated chest X-ray dataset

Class	Image count	Common pathogen(s)
COVID-19 [22]	3,616	SARS-CoV-2
Normal [22]	10,192	N/A
Viral pneumonia [22], [23]	2,838	Influenza A/B, RSV, adenovirus
Bacterial pneumonia [23]	2,780	Haemophilus influenzae
Tuberculosis [24], [25]	3,194	Mycobacterium tuberculosis
MERS [26]	144	MERS-CoV
SARS [26]	134	SARS-CoV
Total	22,898	

2.2. Preprocessing

The preprocessing stage begins with analysis of image aspect ratios, followed by resizing all radiographic images to a uniform resolution of 224×224 pixels, and pixel intensity normalization scales the values to the range of $[0, 1]$ to maintain consistency and preserve clinically relevant features. The complete dataset of 22,898 chest X-ray images is partitioned using stratified sampling into 80% training

(18,552 images), 10% validation (2,059 images), and 10% testing (2,287 images) to ensure balanced class distribution across subsets. Data augmentation, including mild brightness and contrast adjustments, is applied exclusively to the training subset to improve generalization; however, to address class imbalance, class-weighted categorical cross-entropy loss assigns higher importance to minority categories such as SARS and MERS, and additional augmentation is applied more intensively to these minority classes to reduce bias toward majority groups. Furthermore, five-fold stratified cross-validation is conducted on the training subset to enhance robustness, whereas the independent test set remains completely unseen during model development and evaluation.

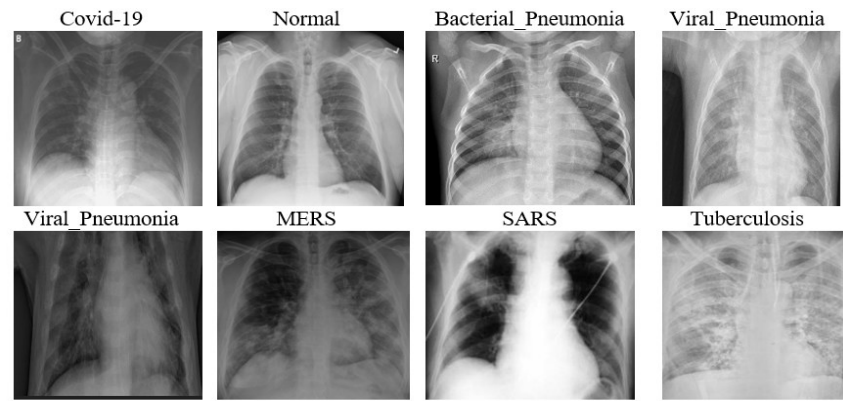


Figure 2. A sample visualization of random images from each class

2.3. Neural architecture search-driven classification model

Selecting an appropriate CNN architecture for clinical deployment requires consideration of classification accuracy, computational complexity, and inference speed; therefore, an empirical comparison is conducted among widely used architectures including VGG16, ResNet50, InceptionV3, MobileNetV2, DenseNet121, and EfficientNetB0, where each model is trained for 30 epochs using the same training and validation splits to ensure fair and consistent evaluation. The comparative results are presented in Figure 3, which illustrates the relationship between classification accuracy and model complexity measured in millions of parameters. To determine the final architecture, a multi-criteria decision scheme assigns weighted importance to four evaluation factors, namely classification accuracy (35%), inference speed (30%), parameter efficiency (25%), and model interpretability (10%), thereby enabling balanced architectural selection rather than relying on accuracy alone.

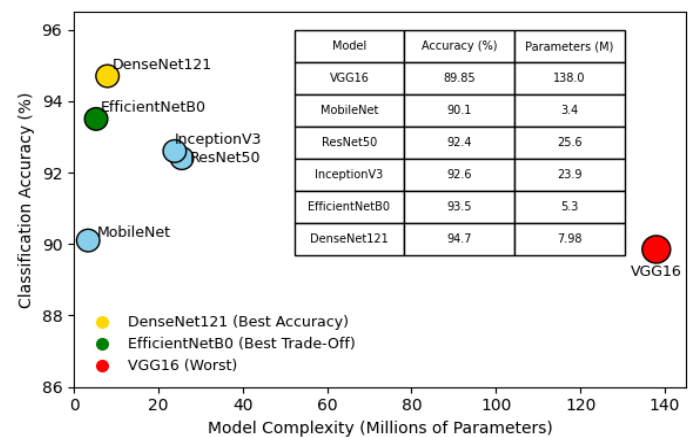


Figure 3. Preliminary backbone architecture comparison for model selection

Based on the multi-criteria evaluation, EfficientNetB0 provides the most balanced trade-off between classification accuracy, computational complexity, and inference efficiency; therefore, a lightweight NAS strategy is applied to refine the classifier head while keeping the convolutional backbone frozen. As

illustrated in Figure 4, a predefined search space is considered, including dense layer units {256, 512, 1024}, dropout rates between 0.2 and 0.5, optimizers {SGD, Adam, RMSProp, Adamax}, learning rates {0.1, 0.01, 0.001}, L2 regularization values between 0.01 and 0.02, and activation functions {ReLU, Swish}. Each sampled configuration is trained for 30 epochs and evaluated using validation accuracy and loss, and random search is adopted to explore the hyperparameter combinations efficiently without exhaustive grid-based evaluation, where candidate models are iteratively assessed until stable validation performance is observed. The final selected architecture consists of a 256-unit dense layer followed by batch normalization and dropout of 0.45, with Adamax optimizer using learning rate 0.001, L2 regularization of 0.016, ReLU activation, global average pooling for feature aggregation, and batch size of 16. The NAS refinement process is executed on an NVIDIA GeForce RTX 4070 GPU with 8 GB VRAM and 16 GB system RAM (AMD Ryzen 9 7940HS), where 40 candidate configurations are evaluated during random search and the complete optimization requires approximately 6.5 hours of computational time.

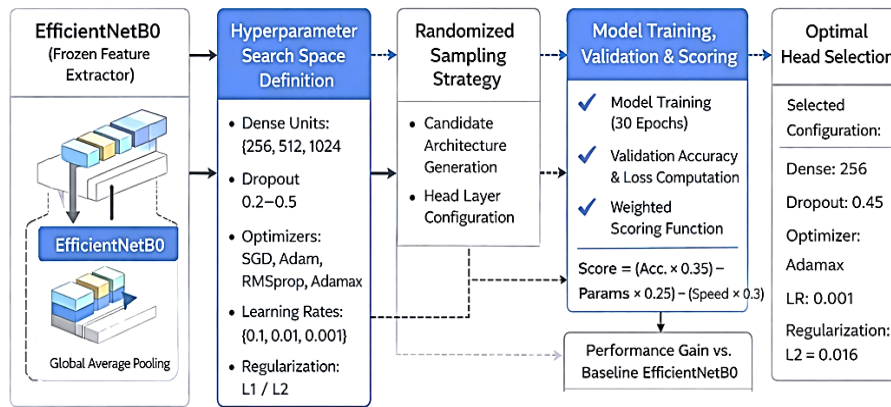


Figure 4. NAS workflow for EfficientNetB0 optimization in multi-class chest infection diagnosis

2.4. Proposed explainable inference framework

One of the biggest challenges in using DL in clinical settings is that they often act like a “black box” because they give us a prediction, but no clear reasoning behind it. In real-world medical scenarios, doctors often face cases that present features of multiple disease classes, or borderline cases, and are subject to diagnostic uncertainty. In this situation, they need a reasoning process that incorporates the degree of certainty and the possibility of co-infection. To understand in a better manner, let us consider the model output a vector of predicted class probabilities P expressed as (1).

$$P = [p_1, p_2, p_3, \dots, p_n], \text{ where } \sum_{i=1}^n p_i = 1 \quad (1)$$

Where p_i represents the model-assigned probability (also referred to as prediction confidence score) for class i , out of n disease classes. The existing approaches generally select the most probable class $\hat{y} = \arg \max_i (p_i)$ as the final decision. However, this strategy disregards the diagnostic details, which are small but important when two or more classes have closely matched probabilities or when the highest probability itself lacks sufficient confidence.

In order to address this issue, the proposed study develops a two-fold explainable inference framework that combines fuzzy-based decision-level reasoning and Grad-CAM-based feature-level visual explanation to help clinicians make informed and trustable decisions. The proposed fuzzy reasoning system adopts a Mamdani-type fuzzy inference system which aims to map the confidence outputs of a classification model into linguistically interpretable severity labels by integrating primary and secondary prediction scores. Let $P = [p_1, p_2, \dots, p_n]$ be the vector of class probabilities output by the SoftMax layer of the classification model, such that: $\sum_{i=1}^n p_i = 1$, $0 \leq p_i \leq 1$, $\forall i$, where, $n=7$ denotes the number of chest infection classes. Unlike conventional classification systems that depend only on the maximum class probability for decision-making, the proposed fuzzy system utilizes both the highest p_{max} and second-highest prediction p_{second} probabilities to infer diagnostic uncertainty and co-infections. The two most significant values p_{max} and p_{second} from vector P are extracted as (2) and (3).

$$p_{max} = \max(p_i) \quad (2)$$

$$p_{second} = \max(\{p_j | p_j \neq p_{max}\}) \quad (3)$$

The value of p_{max} (i.e., top confidence score) and p_{second} are fuzzified using predefined membership functions and evaluated against a rule base to infer the severity of the condition such as severe, moderate, mild/needs review, and ambiguous. The proposed approach analyzes the joint distribution of the top predictions to reason about diagnostic certainty and uncertainty, and it generates human-readable interpretations that reflect competing class probabilities rather than definitive co-infection confirmation. Table 2 presents a representative subset of the fuzzy rule set employed in the framework. The threshold intervals (≥ 0.90 , $0.70 - 0.89$, $0.40 - 0.69$) are determined through empirical validation-set probability analysis, where predictions with $p_{max} \geq 0.90$ consistently correspond to stable and confident classifications, whereas lower probability ranges indicate increased misclassification risk and inter-class overlap. It is important to note that the fuzzy reasoning mechanism models confidence patterns in classification outcomes and highlights cases that may require further clinical review instead of claiming absolute diagnostic certainty. For feature-level explanation, the framework applies Grad-CAM to generate activation heatmaps over input chest X-ray images, thereby enabling visual interpretation of the image regions that contribute most to the model's classification decision.

Table 2. Representative subset of the fuzzy rule set employed

p_{max}	p_{second}	Severity level	Interpretation logic
≥ 0.90	< 0.10	Severe	High certainty, no significant second class
$0.70 - 0.89$	≥ 0.30	Moderate	Considerable confidence but possible co-infection
$0.40 - 0.69$	≥ 0.10	Mild/needs review	Borderline case with moderate second-class influence
< 0.40	Any	Ambiguous	Low confidence, system unsure, manual review recommended

3. RESULTS AND DISCUSSION

This section discusses the outcomes derived from experimental evaluation of proposed multi-class classification framework integrated with fuzzy-based severity reasoning. The entire computational framework was developed and executed in Python using the Anaconda distribution. The evaluation is carried out in three stages, viz. i) classification performance on a custom dataset comprising seven distinct chest infection classes, ii) fuzzy-based severity estimation driven by model confidence and AI interpretability, and iii) comparative analysis with relevant existing approaches in the literature.

3.1. Classification model performance

The performance of the proposed NAS-driven classification model shown in Figure 5 demonstrates the model's learning capability and its stabilized generalization on the validation data. Figure 5(a) shows the model's accuracy progression for both training and validation phases for 30 epochs. As can be seen, the model has achieved near-perfect accuracy $\sim 99.9\%$ on the training set and obtained $\sim 96.4\%$ accuracy on the validation set. Similarly, Figure 5(b) presents the learning curve in terms of loss (error rate) for both training and validation sets over 30 epochs.

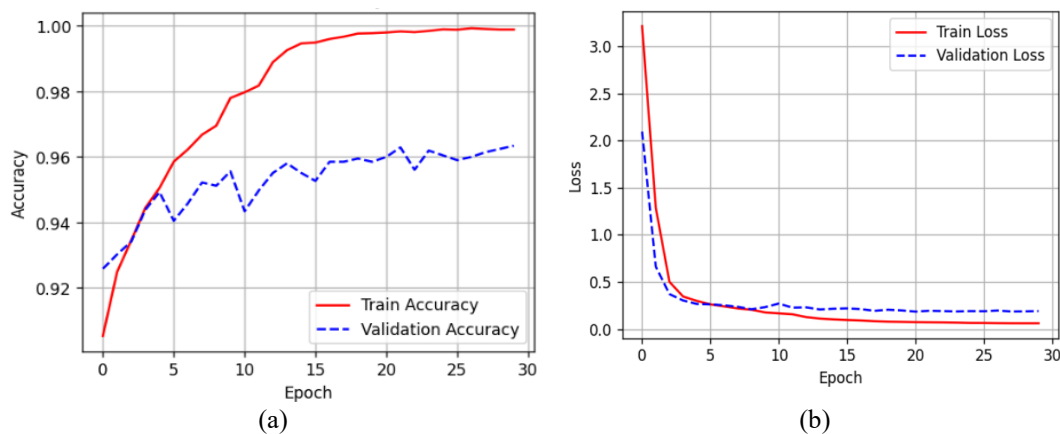


Figure 5. Training and validation learning curve of the model with (a) accuracy plot and (b) loss plot

The learning curves shown in Figure 5 indicate that the model does not suffer from severe overfitting, which can be attributed to the incorporation of regularization, dropout, and learning rate scheduling mechanisms that support a balanced training process. The proposed classification framework achieves an overall test accuracy of 96%, thereby demonstrating strong predictive performance as well as stable generalization capability even under dataset class imbalance conditions. When compared with the baseline EfficientNetB0 achieving 93.5%, as shown in Figure 3, the NAS-refined EfficientNetB0 improves performance to 96.0%, which reflects the contribution of the lightweight NAS-based hyperparameter optimization in enhancing discriminative capacity. The detailed classification report in Table 3 further confirms robustness across all seven classes in terms of precision, recall, and F1-score. Figure 6 presents the confusion matrix of the proposed model to illustrate class-wise prediction distribution and misclassification patterns.

Table 3. Classification report

Class	Precision (%)	Recall (%)	F1-score (%)	Support
Bacterial_pneumonia	85	86	86	278
COVID-19	99	99	99	361
MERS	85	79	81	14
Normal	100	100	100	1,019
SARS	100	85	92	13
Tuberculosis	100	100	100	319
Viral_pneumonia	85	85	85	283

True Label	Bacterial_Pneumonia	240	0	0	0	0	0	38
	Covid-19	0	359	0	1	0	0	1
	MERS	0	3	11	0	0	0	0
	Normal	0	0	0	1016	0	0	3
	SARS	0	0	2	0	11	0	0
	Tuberculosis	0	0	0	0	0	319	0
	Viral_Pneumonia	42	0	0	1	0	0	240
		Bacterial_Pneumonia	Covid-19	MERS	Normal	SARS	Tuberculosis	Viral_Pneumonia
		Predicted Label						

Figure 6. Analysis of confusion matrix

The confusion matrix shows very useful inferences about the true positives and the misclassification rate of the model. Most classes exhibit high true positive counts, particularly normal and tuberculosis, where near-perfect classification was achieved. Misclassifications primarily occur between bacterial pneumonia and viral pneumonia, which is clinically reasonable due to overlapping radiographic patterns. Minor confusion is also observed between SARS and COVID-19, reflecting similarity in viral presentation patterns in chest radiographs. Importantly, no systematic bias toward majority classes is observed, indicating effective handling of class imbalance. The model has maintained consistent and reliable performance for all the classes, even though there is less observation sample for some classes. A 5-fold stratified cross-validation is shown in Table 4 to further validate the generalization ability and robustness of the proposed classification model in different data partitions. Based on the cross-validation outcome, it can be seen that the model has achieved higher weighted precision of $95.56\% \pm 0.19$, along with an effective weighted recall of $95.50\% \pm 0.18$, and weighted F1-score of $95.47\% \pm 0.19$.

The Grad-CAM visualizations presented in Table 5 include representative examples from COVID-19, tuberculosis, bacterial pneumonia, and SARS, demonstrating consistent feature-level interpretability across multiple infection categories. Grad-CAM is adopted as the primary explanation technique because of its compatibility with CNN architectures and its ability to generate spatially coherent activation heatmaps aligned with convolutional feature representations. In contrast, model-agnostic approaches such as local interpretable model-agnostic explanations (LIME) and Shapley additive explanations (SHAP) depend on input perturbation mechanisms, which may lead to spatially fragmented

explanations in high-resolution radiographic images. The highlighted activation regions are examined with reference to established radiological patterns and correspond to clinically relevant pathological areas, including pulmonary opacities, consolidation regions, and peripheral infiltrates, thereby supporting the anatomical plausibility of the model predictions. The fuzzy reasoning component further complements this feature-level explanation by providing decision-level interpretability, forming a dual-layer explanation framework within the proposed system.

The fuzzy severity reasoning mechanism aligns with real-world diagnostic practice, where radiological interpretation inherently involves varying degrees of confidence rather than binary outcomes. By incorporating both primary and secondary class probabilities, the framework captures diagnostic ambiguity and potential overlap between infections, reflecting practical clinical reasoning. From a clinical workflow perspective, the severity categorization may assist in triage and prioritization. High-confidence “severe” predictions can be prioritized for immediate attention, while “mild/needs review” or “ambiguous” cases may be directed for further assessment or confirmatory testing, thereby supporting optimized resource allocation in high-demand settings.

Table 4. Cross-validation performance metrics across 5 folds

Fold	Accuracy	Macro precision	Macro recall	Macro F1	Weighted precision	Weighted recall	Weighted F1
1	0.9528	0.8901	0.8875	0.8869	0.9532	0.9528	0.9526
2	0.9561	0.9367	0.9181	0.9254	0.9563	0.9561	0.9561
3	0.9550	0.9242	0.8763	0.8960	0.9570	0.9550	0.9545
4	0.9572	0.9422	0.9191	0.9298	0.9574	0.9572	0.9572
5	0.9537	0.9444	0.8698	0.9010	0.9539	0.9537	0.9534
Mean±Std	0.9550±0.0018	0.9275±0.0223	0.894±0.0232	0.907±0.0188	0.955±0.0019	0.955±0.0018	0.954±0.0019

Table 5. A representative case scenario of the proposed fuzzy reasoning and Grad-CAM

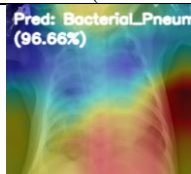
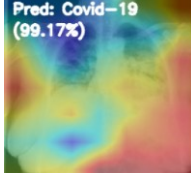
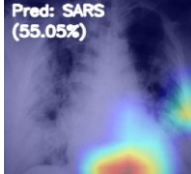
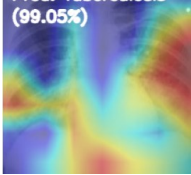
Input image	Fuzzy-reasoning(decision-level)	Grad-CAM (feature-level)
	Primary class: bacterial pneumonia Confidence: 99.37% Secondary class: viral_pneumonia (0.34%) Severity level: severe Interpretation: you got bacterial infection. The model is highly confident in Bacterial_Pneumonia. No significant sign of co-infection.	 Pred: Bacterial_Pneum (96.66%)
	Primary class: COVID-19 Confidence: 99.17% Secondary class: normal (0.50%) Severity level: severe Interpretation: you got viral infection. The model is highly confident in COVID-19. No significant sign of co-infection.	 Pred: Covid-19 (99.17%)
	Primary class: SARS Confidence: 55.05% Secondary class: COVID-19 (39.18%) Severity level: mild/needs review Interpretation: you got viral infection. However, model confidence in SARS is uncertain. Consider further evaluation	 Pred: SARS (55.05%)
	Primary class: tuberculosis Confidence: 82.54% Secondary class: COVID-19 (14.58%) Severity level: moderate Interpretation: you have a bacterial infection, tuberculosis. Minor indication of COVID-19. Suggest monitoring for symptoms related to viral infection.	 Pred: Tuberculosis (82.54%)
	Primary class: tuberculosis Confidence: 99.05% Secondary class: normal (0.23%) Severity level: severe Interpretation: you got bacterial infection. The model is highly confident in tuberculosis. No significant sign of co-infection	 Pred: Tuberculosis (99.05%)

Table 6 presents the comparative evaluation between the proposed approach and previously reported methods. Unlike several comparative studies that restrict evaluation to binary or ternary classification settings, the proposed framework is trained and validated on a 7-category custom dataset to address multi-class diagnostic complexity. In addition, to maintain consistency with existing literature, performance is also reported under a 3-class configuration using the widely adopted COVID-19 radiology database. The analysis shows that the proposed system effectively outperformed existing approaches.

Table 6. Comparative analysis of existing methods and the proposed approach

Reference	Dataset	Classes	Techniques	Accuracy (%)
[27]	Public dataset with 4 classes (7,132 chest X-ray images)	COVID-19, pneumonia, tuberculosis, normal	DL and explainable artificial intelligence (XAI)	94.31
[28]	Public dataset with 3 classes (14,693 chest X-ray images)	COVID-19, pneumonia, tuberculosis, normal	SMOTE and DL	95.7
[29]	COVID-19 radiography dataset	COVID-19, normal, pneumonia	DL, transfer learning models	91.11
[30]	COVID-19 radiography dataset	COVID-19, normal, pneumonia	DeTraC deep CNN	93.1
Proposed (3-class)	COVID-19 radiography dataset	COVID-19, normal, pneumonia	EfficientNetB0, NAS, Fuzzy logic, Grad-CAM	99.9
Proposed (7-class)	Custom dataset (21,581 images)	COVID-19, normal, tuberculosis, bacterial pneumonia, viral pneumonia, SARS, MERS	EfficientNetB0, NAS, Fuzzy logic, Grad-CAM	96

3.2. Discussion

Based on the experimental findings, the proposed NAS-driven EfficientNet framework shows strong performance in multi-class chest infection classification together with consistent explainable inference. The model achieves high overall accuracy and stable cross-validation results, while maintaining interpretability through fuzzy reasoning at the decision level and Grad-CAM at the feature level. The performance improvement is attributed to the balanced parameter efficiency of EfficientNetB0, the lightweight NAS-based refinement of the classifier head over the vanilla baseline, and standardized preprocessing with class-weighted learning that ensures stable training under class imbalance. Unlike studies limited to binary or few-class settings, the framework operates in a seven-class configuration including rare infections such as SARS and MERS, demonstrating feasibility in complex multi-pathogen diagnostic scenarios.

However, the proposed work also has its own limitations. First, the dataset is formed by merging multiple public repositories, which may introduce heterogeneity in imaging protocols, acquisition settings, and population characteristics, and although preprocessing standardization is applied, residual variability may still influence model behavior. Moreover, rare classes such as SARS and MERS contain limited samples due to restricted public availability, which may affect statistical robustness. The evaluation is conducted on benchmark datasets rather than prospectively collected clinical data; therefore, external validation on independent and geographically diverse datasets is required before real-world deployment. In addition, the fuzzy reasoning rules are empirically defined and not adaptively learned, and although they capture prediction uncertainty, a data-driven adaptive mechanism may further refine decision boundaries.

Future work will focus on external validation across independent hospital datasets to assess generalizability under varied imaging conditions. The integration of electronic health records (EHR) and laboratory findings may further enhance multimodal diagnostic reasoning. Exploration of adaptive fuzzy rule learning and optimization techniques is also planned to improve dynamic uncertainty modeling. Furthermore, deployment optimization for edge devices and low-resource clinical settings will be investigated to support scalable and accessible AI-assisted diagnosis.

4. CONCLUSION

This paper has presented the development of an explainable DL framework that uses chest X-rays to classify pathogen-induced chest infections into multiple classes. The study also provided a dataset of a total of 7 class labels. The major contribution was the development of the fine-tuned EfficientNet model, under the NAS method, which achieved an accurate classification outcome with 96% accuracy and did not show higher false positives while being computationally efficient. Another important contribution of this work is the dual interpretability framework, which employs a fuzzy logic-driven severity assessment reasoning mechanism for accurate interpretation at the decision-level and feature-level model interpretability. Furthermore, this proposed work not only advances classification accuracy but also emphasizes the need for transparency and diagnostic reliability in AI-Assisted diagnosis. Future work will concentrate on enhancing our dataset such that it includes a greater number of pathogen classes. Additionally, the study can be extended to investigate

temporal progression, multimodal imaging, and clinician feedback loops for reducing a system's immediacy and developing trust in clinical deployment.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Nirmala Bai Limba Naik	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Manju Devi		✓				✓		✓	✓	✓	✓	✓		

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [NBLN], upon reasonable request.

REFERENCES

- [1] R. G. Carbone, F. Puppo, E. Mattar, A. C. Roden, and N. Hirani, "Acute and chronic eosinophilic pneumonia: an overview," *Frontiers in Medicine*, vol. 11, Apr. 2024, doi: 10.3389/fmed.2024.1355247.
- [2] D. S. Hui *et al.*, "The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—the latest 2019 novel coronavirus outbreak in Wuhan, China," *International Journal of Infectious Diseases*, vol. 91, pp. 264–266, Feb. 2020, doi: 10.1016/j.ijid.2020.01.009.
- [3] C. A. Kauffman and A. N. Malani, "Fungal infections associated with contaminated steroid injections," *Microbiology Spectrum*, vol. 4, no. 2, Mar. 2016, doi: 10.1128/microbiolspec.EI10-0005-2015.
- [4] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi, "Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning," *Medical Image Analysis*, vol. 65, Oct. 2020, doi: 10.1016/j.media.2020.101794.
- [5] C. M. Chu, "Role of lopinavir/ritonavir in the treatment of SARS: initial virological and clinical findings," *Thorax*, vol. 59, no. 3, pp. 252–256, Mar. 2004, doi: 10.1136/thorax.2003.012658.
- [6] A. Tahamtan and A. Ardebili, "Real-time RT-PCR in COVID-19 detection: issues affecting the results," *Expert Review of Molecular Diagnostics*, vol. 20, no. 5, pp. 453–454, May 2020, doi: 10.1080/14737159.2020.1757437.
- [7] A. Esteva *et al.*, "Deep learning-enabled medical computer vision," *npj Digital Medicine*, vol. 4, no. 1, Jan. 2021, doi: 10.1038/s41746-020-00376-2.
- [8] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
- [9] Z. C. Lipton, "The mythos of model interpretability," *Communications of the ACM*, vol. 61, no. 10, pp. 36–43, Sep. 2018, doi: 10.1145/3233231.
- [10] K. Yang, J. Fang, X. Zhang, Y. Bai, and H. Ji, "Development and validation of interpretable machine learning models to predict 30-day mortality in patients with intracerebral hemorrhage," *Scientific Reports*, Jul. 2026, doi: 10.1038/s41598-026-61923-7.
- [11] H. Kör, H. Erbay, and A. H. Yurttakal, "Diagnosing and differentiating viral pneumonia and COVID-19 using X-ray images," *Multimedia Tools and Applications*, vol. 81, no. 27, pp. 39041–39057, Nov. 2022, doi: 10.1007/s11042-022-13071-z.
- [12] A. U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, and P. S. Yakoi, "Pneumonia classification using deep learning from chest X-ray images during COVID-19," *Cognitive Computation*, vol. 16, no. 4, pp. 1589–1601, Jul. 2024, doi: 10.1007/s12559-020-09787-5.
- [13] R. Rajpoot, M. Gour, S. Jain, and V. B. Semwal, "Integrated ensemble CNN and explainable AI for COVID-19 diagnosis from CT scan and X-ray images," *Scientific Reports*, vol. 14, no. 1, Oct. 2024, doi: 10.1038/s41598-024-75915-y.
- [14] A. M. Tahir *et al.*, "COVID-19 infection localization and severity grading from chest X-ray images," *Computers in Biology and Medicine*, vol. 139, Dec. 2021, doi: 10.1016/j.combiomed.2021.105002.
- [15] X. Gu, L. Pan, H. Liang, and R. Yang, "Classification of bacterial and viral childhood pneumonia using deep learning in chest radiography," *ACM International Conference Proceeding Series*, pp. 88–93, 2018, doi: 10.1145/3195588.3195597.

- [16] N. Barakat, M. Awad, and B. A. A. -Nabah, "A machine learning approach on chest X-rays for pediatric pneumonia detection," *Digital Health*, vol. 9, Jan. 2023, doi: 10.1177/20552076231180008.
- [17] H. Bakir, S. Oktay, and E. Tabaru, "Detection of pneumonia from X-ray images using deep learning techniques," *Journal of Scientific Reports-A*, pp. 419–440, 2023, doi: 10.59313/jsr-a.1219363.
- [18] E. Kotei and R. Thirunavukarasu, "Tuberculosis detection from chest X-ray image modalities based on transformer and convolutional neural network," *IEEE Access*, vol. 12, pp. 97417–97427, 2024, doi: 10.1109/ACCESS.2024.3428446.
- [19] S. Induri, M. R. Durgasree, and B. Sukumar, "Automated tuberculosis detection from chest x-rays using a ResNet50 architecture," *Milestone Transactions on Medical Technometrics*, vol. 3, no. 1, pp. 133–144, 2025.
- [20] B. Koesmarsono, A. Widyoningroem, W. Ferriastuti, and A. D. W. Widodo, "Correlation between chest X-ray with fungal pneumonia in immunocompromised patients," *International Journal of Research Publications*, vol. 105, no. 1, Jul. 2022, doi: 10.47119/IJRP1001051720223634.
- [21] H.-Y. Wu *et al.*, "Identifying predictors for bacterial and fungal coinfection on chest computed tomography in patients with *Pneumocystis pneumonia*," *Journal of Microbiology, Immunology and Infection*, vol. 54, no. 4, pp. 701–709, Aug. 2021, doi: 10.1016/j.jmii.2020.06.007.
- [22] T. Rahman *et al.*, "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images," *Computers in Biology and Medicine*, vol. 132, May 2021, doi: 10.1016/j.combiomed.2021.104319.
- [23] D. Kermany, K. Zhang, and M. Goldbaum, "Labeled optical coherence tomography (OCT) and chest X-ray images for classification," *Mendeley Data*, 2018, doi: 10.17632/rscbjbr9sj.2.
- [24] T. Rahman *et al.*, "Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization," *IEEE Access*, vol. 8, pp. 191586–191601, 2020, doi: 10.1109/ACCESS.2020.3031384.
- [25] S. Kiran and D. I. Jabeen, "Dataset of tuberculosis chest X-rays images," *Mendeley Data V2*, 2024, doi: 10.17632/8j2g3csprk.2.
- [26] A. M. Tahir *et al.*, "Deep learning for reliable classification of COVID-19, MERS, and SARS from chest X-ray images," *Cognitive Computation*, vol. 14, no. 5, pp. 1752–1772, Sep. 2022, doi: 10.1007/s12559-021-09955-1.
- [27] M. Bhandari, T. B. Shahi, B. Siku, and A. Neupane, "Explanatory classification of CXR images into COVID-19, pneumonia and tuberculosis using deep learning and XAI," *Computers in Biology and Medicine*, vol. 150, Nov. 2022, doi: 10.1016/j.combiomed.2022.106156.
- [28] L. Venkataramana, D. V. V. Prasad, S. Saraswathi, C. M. Mithumary, R. Karthikeyan, and N. Monika, "Classification of COVID-19 from tuberculosis and pneumonia using deep learning techniques," *Medical and Biological Engineering and Computing*, vol. 60, no. 9, pp. 2681–2691, Sep. 2022, doi: 10.1007/s11517-022-02632-x.
- [29] X. Liu, W. Wu, J. C.-Wei Lin, and S. Liu, "A deep learning model for diagnosing COVID-19 and pneumonia through X-ray," *Current Medical Imaging Reviews*, vol. 19, no. 4, Apr. 2023, doi: 10.2174/1573405618666220610093740.
- [30] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," *Applied Intelligence*, vol. 51, no. 2, pp. 854–864, Feb. 2021, doi: 10.1007/s10489-020-01829-7.

BIOGRAPHIES OF AUTHORS



Nirmala Bai Limba Naik     received the B.E. degree in Instrumentation Technology and M.Tech. degree in Power Electronics from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India. She is currently pursuing the Ph.D. degree in Electronics and Communication Engineering with VTU. She is research scholar in the Department of Electronics and Communication Engineering, The Oxford College of Engineering, Bengaluru. She is an assistant professor in the Department of Electronics and Instrumentation Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India. She has published papers in various reputed international journals and conferences. She has earned a patent. Her current research interests include biomedical image processing, machine learning, deep learning, artificial intelligence, and computer vision for improving the efficiency in disease detection in patients affected by pathogens. She can be contacted at email: nirmala.pari@gmail.com.



Dr. Manju Devi     is the professor and head of the Department of Electronics and Communication Engineering at The Oxford College of Engineering, Bengaluru. She has 28 years of teaching experience and 16 years of research experience. She has completed a Ph.D. in VLSI Design, VTU, M.Tech. (Electronics) VTU, and B.E. (Electronics and Communication Engineering). She is an approved guide in VTU, and three research scholars are pursuing Ph.D. and 7 scholars have completed under her guidance. She has published 7 patents, author of 1 book, 7 book chapters, and 80 research papers in national and international Journals. She is appointed as a BOE member in VTU and a BOS member, VTU nominee at BNMIT for M.Tech. in VLSI Design. She is also a senior member of IEEE and a Life member of the Indian Society for Technical Education (ISTE) and a life member of the Institution of Electronics and Telecommunication Engineers (IETE). She is also a photonic executive member of the photonic society and a mentor of many students for the internship program conducted under the photonic society. She can be contacted at email: manju3devi@gmail.com.