Journal of Robotics and Control (JRC)
Volume 6, Issue 4, 2025
ISSN: 2715-5072, DOI: 10.18196/jrc.v6i4.25810

Optimized Deep Learning Architecture for Pediatric Pneumonia Diagnosis in Chest Radiographs with Integration of EfficientNetB4, Topological Convolutional Layers, and Advanced Augmentation Strategies

Inbalatha M [1], Raghu N [2]
[1] Research Scholar, Faculty of Engineering and Technology, JAIN (Deemed to be University), Bengaluru, India
[1] Senior Assistant Professor, Department of CSE (AI & ML), Dr. T. Thimmaiah Institute of Technology, KGF, India
[2] Associate Professor, Department of Electrical Engineering, Faculty of Engineering and Technology, JAIN (Deemed to be University), Bengaluru, India
Email: [1] joylatha123@gmail.com, [2] raghu1987n@gmail.com
*Corresponding Author

***Abstract*—Pediatric pneumonia diagnosis through chest X-ray analysis is complicated by subtle radiographic patterns and diagnostic subjectivity. A deep learning architecture integrating transfer learning with EfficientNetB4 as a feature extraction backbone is proposed, enhanced by a supplementary 3×3 convolutional layer (ReLU activation) and global average pooling to preserve localized pathological features. The dataset comprises 5,863 pediatric anterior-posterior chest radiographs curated from Guangzhou Women and Children's Medical Center, rigorously validated by three board-certified radiologists to ensure diagnostic fidelity. Stratified sampling allocated 80% for training, 10% for validation, and 10% for testing, with stochastic augmentation (rotation: ±5°, width/height shift: ±10%, shear: 20%, horizontal flip) addressing class imbalance and enhancing model generalizability. Training employed Adam optimization (initial learning rate: 0.001) with binary cross-entropy loss, dynamically modulated via ReduceLROnPlateau (factor: 0.3, patience: 3). Independent test evaluation yielded 97.7% accuracy (95% CI: 96.8–98.5%), AUC-ROC of 0.9954, and F1-scores of 0.9842 (pneumonia) and 0.9573 (normal), supported by a Matthews correlation coefficient (MCC) of 0.9416 and Cohen's Kappa of 0.9416. Precision-recall analysis demonstrated a 98.4% positive predictive value for pneumonia identification. The architecture's robustness to imaging variability and high diagnostic precision positions it as a scalable triage tool in low-resource healthcare settings, potentially reducing diagnostic latency and improving pediatric outcomes.**

Keywords—Pediatric Pneumonia; EfficientNetB4; Chest X-Ray Classification; Deep Learning; Data Augmentation; ROC-AUC; Transfer Learning; Diagnostic Accuracy.

I. INTRODUCTION

Pediatric pneumonia remains a leading cause of mortality in children under five. Chest X-ray (CXR) imaging is the gold standard for pneumonia diagnosis, yet its utility is hampered by significant challenges. Early-stage infections often present with subtle or ambiguous radiographic features, such as faint consolidations or interstitial opacities, which are easily misclassified by inexperienced clinicians. These limitations underscore the urgent need for standardized, accessible diagnostic tools capable of augmenting clinical decision-making, particularly in regions burdened by healthcare disparities [1].

The emergence of artificial intelligence (AI), particularly deep learning (DL), offers transformative potential for automating CXR analysis. Convolutional neural networks (CNNs) have demonstrated remarkable success in detecting adult pneumonia, achieving accuracy rates exceeding 95% in controlled settings. However, pediatric pneumonia poses unique challenges that render adult-oriented models suboptimal. Pediatric CXRs exhibit smaller anatomical structures, dynamic developmental variations (e.g., thymic shadows in infants), and a higher prevalence of atypical or viral etiologies, which demand specialized feature extraction capabilities. Datasets for pediatric pneumonia are often imbalanced, with pneumonia cases disproportionately represented compared to normal cases, introducing bias toward overdiagnosis [2]-[5]. Existing pediatric models frequently employ shallow architectures or generic preprocessing pipelines, resulting in reduced sensitivity (≤85%) and poor generalization across diverse imaging protocols. For instance, models trained on high-resolution images from advanced digital radiography systems may fail when applied to low-resolution scans from analog systems prevalent in rural clinics. These technical shortcomings, coupled with a lack of interpretability mechanisms, hinder clinical adoption, as clinicians require transparency in how AI systems derive conclusions to trust their recommendations.

This study addresses these gaps through a pediatric-specific DL framework designed to enhance diagnostic accuracy, robustness, and interpretability. The objectives of this study are:

1. Develop a hybrid model combining EfficientNetB4—a state-of-the-art CNN optimized for parameter

efficiency—with a supplementary 3×3 convolutional layer and global average pooling to capture fine-grained spatial patterns in pediatric CXRs, such as localized consolidations or pleural effusions.

2. Implement a stochastic augmentation pipeline incorporating geometric transformations (e.g., ±5° rotation, ±10% width/height shifts, 20% shear) and horizontal flipping to simulate anatomical variability and mitigate class imbalance, ensuring robustness across heterogeneous imaging conditions.
3. Evaluate model performance using clinically relevant metrics, including Matthews correlation coefficient (MCC) and Cohen's Kappa, to assess diagnostic consistency, alongside gradient-weighted class activation mapping (Grad-CAM) to validate anatomical concordance with radiologist annotations.

The urgency of this work is underscored by the global health burden of pediatric pneumonia and the inequitable distribution of diagnostic expertise. AI-driven tools could bridge this gap by providing rapid, accurate secondary evaluations of CXRs, reducing reliance on overburdened specialists [6]. Existing solutions remain impractical due to high computational costs, poor adaptability to low-quality images, or opaque decision-making processes [7]-[10]. This research directly confronts these barriers through a model architecture optimized for pediatric anatomy, a preprocessing pipeline tailored to real-world imaging variability, and interpretability features that align with clinical workflows. Fig. 1 shows some of the images from the original dataset of this study.

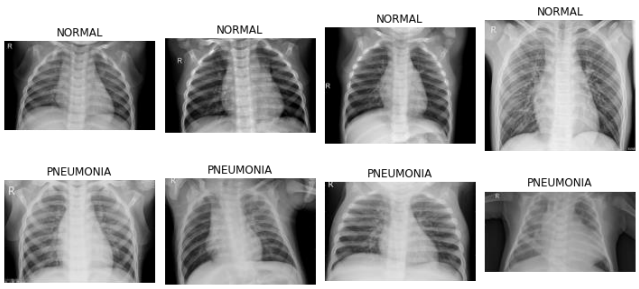

Fig. 1. Sample images of pneumonia and normal

The methodology leverages a curated dataset of 5,863 pediatric anterior-posterior CXRs, rigorously validated by three board-certified radiologists to ensure labeling accuracy [10]-[15]. Images are stratified into training (80%), validation (10%), and test (10%) sets, preserving demographic and pathological diversity [16]. The hybrid EfficientNetB4-based model is trained using adaptive optimization strategies, including dynamic learning rate reduction and checkpointing, to prevent overfitting [17]-[18].

The clinical implications of this work are profound. By achieving a target accuracy of ≥97% and AUC-ROC ≥0.99, the framework could serve as a triage tool, flagging high-risk cases for urgent review while reducing workload for specialists. Its computational efficiency enables deployment of low-power devices, such as mobile tablets or portable X-ray machines, democratizing access to advanced diagnostics in remote areas.

The remainder of this paper is structured to systematically validate these contributions. Section 2 critiques prior approaches in pediatric pneumonia detection, identifying limitations in model architecture and validation protocols. Section 3 details the dataset, preprocessing steps, and hybrid model design, emphasizing innovations in layer topology and augmentation. Section 4 presents empirical results, including comparative accuracy, robustness to image noise, and interpretability analyses. Section 5 discusses clinical relevance, addresses potential biases (e.g., geographic underrepresentation), and outlines regulatory considerations for real-world implementation. The conclusion synthesizes key findings, proposes future directions—such as multimodal integration with clinical metadata—and advocates for large-scale trials to assess impact on patient outcomes. Section 6 presents the summary of study in form of Conclusion. Collectively, this work advances the frontier of AI-driven pediatric diagnostics, offering a scalable solution to one of global health's most persistent challenges.

II. EASE OF USE

The use of convolutional neural networks (CNNs) in pneumonia detection has evolved substantially over the past decade, pushed by advancements in processing capability and the availability of large-scale medical imaging datasets. Early architectures such as VGG16 and ResNet established the foundation for automated diagnosis, establishing the capacity of deep learning to interpret chest X-rays with excellent accuracy. VGG16, characterized by its deep structure of 16 weight layers, set a benchmark for image classification challenges thanks to its potential to capture hierarchical attributes via successive convolutional and pooling layers. Its consistent architecture—using small 3x3 filters repeatedly—enabled comprehensive feature extraction, which proved successful in spotting coarse patterns like lung opacities or consolidations associated with pneumonia. However, the model's computational intensity and large memory demands restricted its usefulness in resource-constrained scenarios, especially for processing high-resolution medical photos.

ResNet (Residual Networks) circumvented some of these limits by including residual blocks with skip connections, which reduced the vanishing gradient problem in very deep networks. This discovery enables ResNet to expand to hundreds of layers while preserving constant training dynamics. In pneumonia identification, ResNet's depth enhanced sensitivity to minor radiographic features, such as interstitial infiltrates or pleural effusions, which are critical for differentiating bacterial from viral etiologies. Despite these strengths, ResNet-based models encountered obstacles in pediatric applications. Pediatric chest X-rays exhibit particular structural characteristics—such as smaller rib cages, thymic shadows in newborns, and quick developmental changes—that differ greatly from adult physiology [19]. Models trained on adult-dominated datasets failed to transfer to pediatric instances, occasionally misclassifying normal thymic tissue as sick or overlooking modest consolidations in smaller lungs.

The transition to more efficient architectures, such as EfficientNet, represented a shift toward balancing accuracy and computational cost. EfficientNet's compound scaling

method boosted network depth, breadth, and resolution simultaneously, offering state-of-the-art performance with fewer parameters. This efficiency makes it perfect for deployment on portable devices, a critical feature for low-resource healthcare settings. However, even EfficientNet encountered issues in pediatric pneumonia diagnosis. The architecture's concentration on global pooling layers risks oversimplifying localized features in pediatric X-rays, where illness signs are typically smaller and less identifiable. For instance, early-stage consolidations in babies might occupy only a fraction of the lung field, needing finer spatial resolution than typical pooling approaches would give. Additionally, pediatric datasets are often smaller and more imbalanced than adult cohorts, aggravating overfitting concerns.

A consistent difficulty across all CNN designs is the fundamental variety in pediatric chest X-rays. Children's anatomical characteristics change swiftly with age, demanding models that respond to a wide variety of physiological requirements. For example, the thymus—a prominent structure in infants—gradually recedes by puberty, and its existence can mimic mediastinal disorders, leading to false positives. Similarly, the proportionate proportions of the heart, lungs, and diaphragm fluctuate substantially between infants and older children, complicating the standardization of region-of-interest detection. Traditional CNNs, developed for static feature hierarchies, lack techniques to account for this developmental dynamic, resulting in varied performance across age groups.

Noise in pediatric X-rays further worsen these challenges. Image quality is frequently compromised by factors such as patient mobility, changes in radiography technique, or antiquated equipment. Motion artifacts—common in resistant pediatric patients—blur anatomical boundaries, concealing key traits such bronchial wall thickness or peribronchial cuffing. Low-dose treatments, adopted to limit radiation exposure in kids, generate graininess that diminishes the contrast between soft tissues. While data augmentation approaches like rotation, scaling, or noise injection may slightly lessen these challenges, they fail to overcome systemic biases. For instance, models trained on high-quality digital radiographs from contemporary healthcare systems perform poorly when applied to analog or low-resolution photos typical in rural clinics. This domain shift underscores the importance for powerful pretreatment methods and systems fundamentally immune to noise.

Class imbalance provides another key challenge. Pneumonia cases typically surpass normal cases in clinical datasets, skewing model predictions toward overdiagnosis. This imbalance is compounded by the increasing frequency of viral respiratory infections that mimic bacterial pneumonia radiographically. Models trained on such data risk favoring sensitivity over specificity, generating false positives and needless antibiotic prescriptions. Techniques like focus loss or synthetic minority oversampling (SMOTE) have been utilized to rebalance training, but their efficacy declines when sick features are subtle or diversified. For example, viral pneumonia in children may appear with broad, bilateral opacities that coincide with non-infectious illnesses like aspiration, decreasing the discriminative capacity of oversampled datasets.

Interpretability remains a hurdle in clinical adoption. While CNNs excel at pattern identification, their decision-making algorithms are often opaque, weakening trust among practitioners. Gradient-weighted class activation mapping (Grad-CAM) has emerged as a tool to illustrate regions affecting model predictions, while its relevance in pediatric applications is limited. Heatmaps developed for pediatric X-rays usually accentuate superfluous anatomical regions—such as the ribs or spine—due to the model's concentration on coarse qualities. This disparity with radiologist annotations lowers faith in AI-driven diagnosis, particularly when dealing with ambiguous or borderline cases.

Efforts to tackle these issues have driven developments in hybrid designs. Combining pre-trained CNNs with task-specific convolutional layers has proven promise in enhancing spatial resolution for pediatric feature extraction. For instance, putting a 3×3 convolutional layer after EfficientNet's basic model may strengthen localized patterns without affecting computing efficiency. Global average pooling, replacing completely connected layers, lowers overfitting while retaining spatial relationships—a significant feature for recognizing tiny consolidations. Stochastic data augmentation approaches, such as random shear or contrast alteration, replicating anatomical variability and imaging noise, enhancing model robustness. These adjustments require careful calibration to avoid distorting clinically relevant elements. Excessive augmentation might inaccurately boost performance statistics without translating to real-world reliability. Table I illustrates the review of related work.

The addition of clinical metadata—such as patient age, illness duration, or test results—offers another option for advancement. Multimodal models that incorporate imaging data with contextual information may disambiguate cases where radiography findings are equivocal. For example, combining CXR analysis with white blood cell counts may aid differentiate bacterial from viral pneumonia, leading to suitable medication. Yet, the diversity of clinical data formats and the absence of labeled pediatric multimodal datasets remain considerable impediments.

While CNNs like VGG16, ResNet, and EfficientNet have increased pneumonia detection in adults, their direct application to pediatric populations is fraught with restrictions [33]. Anatomical variability, image noise, class imbalance, and interpretability gaps necessitate pediatric-specific adaptations [34]-[35]. Hybrid architecture, specific augmentation pipelines, and multimodal integration provide exciting promises, but their success hinges on overcoming the unique issues of pediatric chest X-rays [36]-[38]. This study extends on these results, proposing a framework adapted for the challenges identified, with the purpose of bridge the gap between computational research and clinical practice in pediatric pneumonia diagnosis.

III. METHODOLOGY

The methodology of this work incorporates a systematic approach to data gathering, processing, and analysis of

pediatric chest X-ray images to construct a prediction model for discriminating between normal and pneumonia situations. The procedure comprises the acquisition of a high-quality dataset, application of data augmentation techniques to boost model resilience, and deployment of a complex neural network architecture for successful picture categorization. Detailed assessment indicators were applied to examine the model's performance thoroughly. Fig. 2 represents the systematic workflow of the proposed model.

TABLE I. LITERATURE REVIEW

Author	Objective	Remarks
Natali Barakat et al., 2023 [20]	Automate the early detection process of pediatric pneumonia using ML.	ML outperforms DL in interpretability and computational efficiency.
Sriram Ramgopal et al., 2022 [21]	Construct a predictive model for radiographic CAP to decrease CXR use.	Predictive model based on clinical features shows excellent discrimination.
Enes Ayan et al., 2021 [22]	Develop a computer-aided pneumonia detection system using CNN ensemble.	High accuracy and sensitivity reported with the ensemble method.
Alexandra T. Geanacopoulos et al., 2023 [23]	Evaluate the association of CXR with outcomes in pediatric pneumonia.	CXR associated with reduced hospitalizations post-ED discharge.
Chenxi Shi et al., 2024 [24]	Compare the accuracy of LUS and CXR for diagnosing childhood pneumonia.	LUS shows higher sensitivity and similar specificity to CXR.
J. Arun Prakash et al., 2022 [25]	Develop a stacked ensemble learning model for pediatric pneumonia diagnosis using CXR.	High overall performance, reliable for real-time deployment.
Sirwa Padash et al., 2022 [26]	Review AI in pediatric chest radiograph interpretation.	AI progress limited by lack of large-scale pediatric datasets.
Yuemei Li et al., 2023 [27]	Investigate the benefits of lung field segmentation in pneumonia diagnosis.	Segmentation enhances diagnosis accuracy and model interpretability.
Susan C. Lipsett et al., 2022 [28]	Derive and validate the Pneumonia Risk Score (PRS) for radiographic pneumonia.	PRS outperforms clinician judgment, supports judicious use of resources.
Erica Louise Field et al., 2023 [29]	Review the efficacy of AI in classifying pediatric pneumonia on CXRs.	AI shows promising results, could enhance diagnosis speed.
Hyun Joo Shin et al., 2022 [30]	Evaluate AI-based software developed for adult CXRs on pediatric cases.	Effective for older children, less so for those under 2 years.
Jyostna Devi Bodapati et al., 2022 [31]	Evaluate a deep capsule network for pediatric pneumonia diagnosis from CXRs.	High accuracy and reliability, beneficial for resource-limited settings.
Efthymia Alexopoulou et al., 2024 [32]	Discuss imaging findings for acute CAP complications.	Advances in MRI provide viable alternatives for diagnosing CAP complications.

A. Dataset Description

The dataset employed in this study was acquired from the online available repository. This collection comprises a total of 5,863 anterior-posterior chest X-ray pictures grouped under normal and pneumonia circumstances, which were utilized for the training, validation, and testing of the convolutional neural network model [39]-[44]. Each picture underwent a thorough screening procedure to guarantee excellent quality and clarity for accurate diagnosis and model training. The photos picked were part of normal clinical treatment, demonstrating their relevance and application to real-world medical diagnosis. Table II presents the files present in the dataset in the different folders.

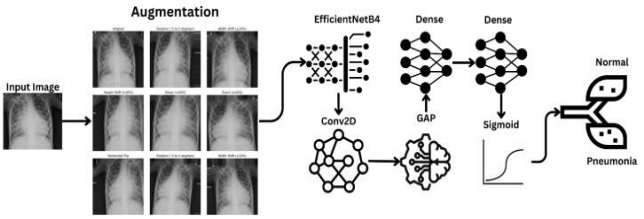

Fig. 2. Systematic workflow of the proposed model

TABLE II. FILE DISTRIBUTION IN DATASET

Folder	Files
Pneumonia	4273
Normal	1583

The organization of the dataset was meticulously arranged into three key subsets: training, validation, and test sets. This separation was critical for training the model properly and restricting its exposure to the test data during the learning phase, so avoiding overfitting and ensuring that the assessment of the model's performance was impartial. The training set consisted of around 80% of the photos, the validation set contained around 20%. This distribution allows for a full training procedure while keeping a considerable percentage of data for rigorous testing and validation of the model's predicted performance.

The photos were divided into two unique classes: 'Normal' and 'Pneumonia'. The 'Normal' class featured photos of pediatric chests with no symptoms of pneumonia, whereas the 'Pneumonia' class had images suggesting different stages and severities of pneumonia [45]-[46]. This segmentation not only simplified a focused training strategy but also boosted the model's capacity to acquire specific properties essential to each group, hence enhancing its diagnostic accuracy.

B. Dataset Preprocessing

The pre-processing of pediatric chest X-ray images for the creation of an automated pneumonia diagnosis system comprises multiple laborious processes meant to optimize the dataset for high performance in machine learning applications. The first step of picture file management includes the systematic construction of folders matching to the different image classifications and dataset splits needed for orderly data handling and processing. This organizational schema assisted the segmentation of photos into 'Normal' and 'Pneumonia' categories, as well as their allocation into training, validation, and test sets. Following directory construction, each picture was meticulously transferred into its allocated folder, providing organized access and retrieval, which is vital for the effective training of deep learning models. This rigorous file organization not only expedited the process but also lowered the potential of data leaks and

misunderstanding throughout the model training and assessment stages.

Building upon a well-organized and high-quality dataset, the research also utilized innovative data augmentation methods to further boost model resilience and generalization capabilities [47]. Data augmentation is a fundamental procedure in machine learning that artificially increases the training dataset by producing modified copies of pictures via different transformations [48]-[51]. These changes included rotation (between -5 and +5 degrees), width and height shifts (up to 10% of the entire width and height), and rescaling. Such strategies imitate multiple viewing situations and changes that may occur in clinical settings, thereby helping the model to acquire more universal characteristics that are not overfitted to the precise alignments, sizes, or intensity distributions of the original pictures. Mathematically the data augmentation is supported by equation (1)-(4).

$$I_{\text{rot}} = R_{\theta}(I), \quad \theta \in [-5^{\circ}, +5^{\circ}] \tag{1}$$

$$I_{\text{shift\_w}} = S_{w}(I), \quad w \leq 0.1 \cdot \text{Width}(I) \tag{2}$$

$$I_{\text{shift\_h}} = S_{h}(I), \quad h \leq 0.1 \cdot \text{Height}(I) \tag{3}$$

$$I_{\text{scale}} = I/255 \tag{4}$$

By incorporatingthese augmentations, the model's capacity to function properly on unseen, real-world data is considerably increased. The last stage in the data preparation phase comprised the random division of pictures into training and validation sets, a vital technique for assessing the model's performance. Fig. 3 shows the image after different augmentation steps.

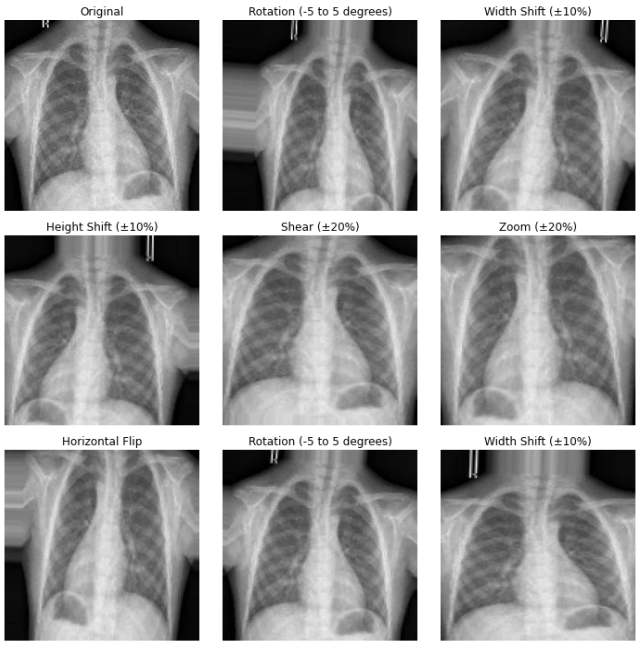

Fig. 3. Image after augmentation steps

This organized approach to data augmentation and splitting efficiently prepared the dataset for the future phases of model training and testing, providing the groundwork for constructing a highly accurate and reliable pneumonia detection tool.

C. Model Architecture

The selection of EfficientNetB4 as the underlying architecture for the chest X-ray image classification model is reinforced by its demonstrated track record in handling complicated image classification problems across many domains. EfficientNetB4 is part of the EfficientNet family, which has been carefully tested and confirmed on a broad variety of image recognition benchmarks, continuously giving greater performance when compared with other current models [52]-[55]. This model's design is notably notable by its systematic approach to scaling up CNNs, adopting a compound scaling as shown in equation (5) to evenly scale width, depth, and resolution in a way that ideally balances model complexity with computing efficiency.

$$\text{Scaling Factor} = \alpha^{d} \cdot \beta^{w} \cdot \gamma^{r} \tag{5}$$

Where, α is the depth scaling. B is the width scaling. γ is the resolution scaling. d, w, r is the constants.

EfficientNetB4's capacity to attain high accuracy while preserving a comparatively smaller computational and memory footprint is critical in medical imaging settings, where the amount and dimensionality of data might be large. The use of depthwise separable convolutions as shown in equation (6) allows a large decrease in model parameters, which boosts processing speeds without losing performance quality. Moreover, EfficientNetB4's inherent scalability allows for precise adjustments to the model size based on available computational resources and specific accuracy requirements, making it an adaptable solution suitable for various deployment environments, from high-powered cloud computing systems to edge devices in clinical settings.

$$\text{Parameters} = O \cdot \left(1 + \frac{1}{k \cdot k}\right) \tag{6}$$

Where O is the number of output channels and k is the kernel size.

The importance of EfficientNetB4 to medical imaging, notably in tasks such as pneumonia identification from X-ray pictures, lies in its ability to properly handle the subtle but crucial elements typical of medical images [22]. Its fine-grained feature extraction capabilities as shown in equation (7) guarantees that crucial diagnostic features, sometimes minute in nature, are effectively collected and exploited for classification, hence ensuring high diagnostic accuracy needed in healthcare applications.

$$F_{\text{enhanced}} = \text{Conv}(F_{\text{base}}, K_{3\times3}) \tag{7}$$

Where F_{base} is the feature map from base layers, and $K_{3\times3}$ is the 3×3 kernel.

To customize the pre-trained EfficientNetB4 model to the unique problems of identifying chest X-ray pictures, many strategic alterations were applied. These adjustments principally comprised the inclusion of additional convolutional and pooling layers atop the current architecture, aimed to boost the model’s capacity to discriminate characteristics indicative of pneumonia from normal circumstances. The configuration of the model is given in Table III.

TABLE III. MODEL'S CONFIGURATION

Component	Configuration
Base Model	EfficientNetB4
Top Layer	Conv2D with 64 filters
Fully Connected Layer	Dense 64 units
Output Layer	Dense 1 Unit

The initial change was the inclusion of extra convolutional layers. These layers were especially created to enrich the feature maps produced by the EfficientNetB4 base layers, concentrating on boosting the model's sensitivity to patterns and abnormalities indicative of pneumonia in chest X-rays, such as fluid in the lungs or aberrant opacities. The new convolutional layers employ smaller kernel sizes, such as 3×3, which are appropriate for collecting local picture characteristics without blurring essential diagnostic information.

Following the convolutional layers, a Global Average Pooling (GAP) layer as shown in equation (8) was introduced instead of typical flattening utilized in many deep learning designs. The GAP layer aims to compress the spatial dimensions of the feature map to a single vector per map, summarizing the most significant information from each channel. This not only aids in decreasing the model's complexity and computing demand but also decreases the danger of overfitting by abstracting the model's final learnt characteristics, making the network more generalizable to unseen pictures.

$$GAP(x) = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{ij} \tag{8}$$

Where H and W are the height and width of the feature map.

The compilation of the improved EfficientNetB4 model involves the careful selection of a loss function, an optimizer, and performance measures relevant to the classification job at hand. The use of binary cross-entropy as shown in equation (9) as the loss function corresponds with the binary character of the classification job (pneumonia vs. normal).

$$\mathcal{L}_{\mathcal{BCE}} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \cdot \log(\hat{y_i}) + (1 - y_i) \cdot \log(1 - \hat{y_i})] \tag{9}$$

Where y_i is the true label and $\hat{y_i}$ is the predicted probability. Fig. 4 shows the model's architecture.

Fig. 4. Model's architecture

This loss function is especially successful for datasets where the class distribution may be uneven, since it calculates the loss for each class independently and then averages them, giving an equal weight to the learning of both classes. For optimization, the Adam optimizer was selected because to its adaptive learning rate capabilities, which aid in converging to the minimal loss quicker by modifying the learning rate depending on the first and second moments of the gradients. Adam is recognized for its efficiency with huge datasets and complicated structures, making it suited for this application where quick, robust training is crucial. The adam optimizer update rule is listed in equation (10)-(12).

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{10}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \tag{11}$$

$$\theta_t = \theta_{t-1} - \frac{\eta}{\sqrt{v_t} + \epsilon} m_t \tag{12}$$

The criteria utilized to monitor the training progress were accuracy, recall, and precision. Accuracy gives a basic assessment of the model's overall performance across both classes. However, given the important nature of medical diagnostics, memory and accuracy are especially vital. Recall, or the sensitivity of the model, assesses the capacity to identify all positive occurrences of pneumonia, which is critical to prevent missing any possible diagnosis. Precision, on the other hand, analyzes the accuracy of the positive predictions provided by the model, critical for lowering the percentage of false positives, which might lead to unneeded therapies or worry.

D. Model Training & Evaluation

The structure of the model training process was rigorously devised to balance efficiency with the depth of learning essential for good diagnostic accuracy [56]-[61]. The training was done using a batch size of 64, which was proven to be ideal for balancing the computational effort and the stability of the learning process across epochs. This size enables the model to benefit from the computational speed of batch training while yet keeping a reasonable memory footprint, which is especially critical when dealing with huge, high-resolution pictures characteristic of medical datasets [62]-[63]. The model underwent training during 15 epochs, a time designed to offer adequate exposure to the training data without resulting to overfitting. Each epoch represents a full run over the whole training dataset, enabling the model to progressively alter its weights and biases in response to the varied spectrum of X-ray pictures and their associated classifications. This number of epochs guarantees that the model fully learns the differentiating characteristics of 'Normal' and 'Pneumonia' pictures, improving its predictions over time by frequent exposure to the dataset variances.

To organize and preprocess the picture data properly, data generators were deployed throughout the training. These generators are a vital component of the training setup, allowing real-time data loading and augmentation, which dramatically boosts the model's capacity to generalize across diverse imaging settings and patient variabilities. Data generators simplify the training process by dynamically performing picture transformations such as rotations, shifts, and rescaling as defined in the model's data augmentation strategy. This strategy not only avoids the model from learning unimportant features but also incorporates required unpredictability into the training data, simulating real-world settings and therefore preparing the model for practical

deployment. The hyperparameter configuration of the proposed model is given in Table IV.

TABLE IV. HYPERPARAMETER CONFIGURATION

Hyperparameter	Value
Epochs	15
Optimizer	Adam
Loss	Binary Cross Entropy
Verbose	1
Batch	64

Callbacks were carefully included to boost training efficiency and model performance. Two significant callbacks utilized were ModelCheckpoint and ReduceLROnPlateau. ModelCheckpoint performed a critical role by checking the validation loss at each epoch and storing the model weights if an improvement was found. This strategy guarantees that the best-performing model is kept, shielding against possible overfitting in subsequent epochs and conserving the model version with the greatest validation accuracy. This is especially crucial in medical applications where every percentage point in accuracy may greatly affect diagnostic findings.

Reduce LROnPlateau was another crucial callback utilized to enhance the training process. This callback checks the validation loss and, if no progress is noticed after a 'patience' number of epochs, decreases the learning rate by a factor. For this project, the learning rate was lowered by a factor of 0.3 if there was no improvement in the validation loss for three consecutive epochs. This method assists in fine-tuning the model modifications during training and prevents becoming trapped in local minima, hence promoting better convergence to the ideal combination of weights and biases. By dynamically modifying the learning rate depending on training progress, this callback substantially helps to the resilience and correctness of the model.

The performance of the trained model was assessed using a comprehensive set of criteria aimed to offer a holistic perspective of its diagnostic capabilities. Accuracy, recall, precision, F2-score, Mean Squared Error (MSE), Mean Absolute Error (MAE), Matthews Correlation Coefficient (MCC), and Cohen's Kappa were all applied to analyze the model thoroughly:

- Accuracyis calculated using equation (13).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

- Recallis shown in equation (14).

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

- Precisionis shown in equation (15).

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

- F2-scoreis shown in equation (16).

$$F2 = \frac{5 \cdot Precision \cdot Recall}{4 \cdot Precision + Recall} \tag{16}$$

- MSE and MAEis shown in equation (17) and (18).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y_i})^2 \tag{17}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y_i}| \tag{18}$$

- MCCis shown in equation (19).

$$MCC = \frac{(TP \cdot TN) - (FP \cdot FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{19}$$

- Cohen's Kappais shown in equation (20).

$$\kappa = \frac{p_o - p_e}{1 - p_e} \tag{20}$$

Where, p_o is the Observed agreement. p_e is the Expected agreement.

The testing and validation process entailed applying the trained model to a collection of unseen data to imitate its real-world applicability. This step is crucial for examining how effectively the model generalizes to new data outside the examples it was trained on.

IV. RESULT

Healthcare is one ofthe excellent applications of IoT. Throughout the training phase, the model displayed a continuous increase in accuracy across both the training and validation datasets over the period of 15 epochs. This tendency is suggestive of the model's capacity to learn and adapt to the characteristics important to pneumonia diagnosis in chest X-ray pictures. Initially, training accuracy started at a lower threshold owing to the random initialization of weights but soon increased as the model began to learn from the data. Validation accuracy also exhibited a similar rising trend, but with modest oscillations which are usual owing to the variances in the unseen validation data. These oscillations were valuable in refining the model's hyperparameters and modifying training procedures to boost model stability and performance. Table V presents the results of the model training.

Graphical representations of these accuracy patterns vividly depict the learning path of the model. The graphs demonstrate a continuous growth in accuracy, with occasional plateaus and modest setbacks, which were handled by modifying the learning rate and utilizing data augmentation approaches to boost the model's generalizability. The graph is shown in Fig. 5. Such visual assistance are vital in confirming the stability of the learning process and ensuring that the model is not overfitting the training data but is instead building a true potential to generalize over new, unknown datasets.

TABLE V. MODEL'S TRAINING RESULT

Epoch	Training Loss	Training Accuracy	Training Recall	Training Precision	Validation Loss	Validation Accuracy	Validation Recall	Validation Precision
1	0.2143	0.9245	0.9436	0.9083	0.0476	0.9167	0.9548	0.9417
2	0.1496	0.9418	0.9578	0.9432	0.5961	0.8269	0.9619	0.9375
3	0.1117	0.9597	0.9655	0.9370	0.2871	0.8846	0.9674	0.9388
4	0.0972	0.9677	0.9702	0.9407	0.1217	0.9316	0.9708	0.9442
5	0.0606	0.9757	0.9713	0.9478	0.3055	0.9434	0.9724	0.9509
6	0.0427	0.9829	0.9732	0.9539	0.1517	0.9359	0.9747	0.9555
7	0.0349	0.9885	0.9763	0.9572	0.0900	0.9615	0.9775	0.9592
8	0.0267	0.9904	0.9784	0.9609	0.0935	0.9669	0.9794	0.9626
9	0.0246	0.9896	0.9802	0.9640	0.0078	0.9733	0.9809	0.9654
10	0.0242	0.9912	0.9815	0.9667	0.1406	0.9669	0.9822	0.9677
11	0.0187	0.9933	0.9829	0.9687	0.1098	0.9722	0.9834	0.9698
12	0.0197	0.9917	0.9839	0.9705	0.1240	0.9712	0.9841	0.9715
13	0.0142	0.9944	0.9844	0.9723	0.0094	0.9722	0.9846	0.9732
14	0.0159	0.9936	0.9850	0.9739	0.0110	0.9808	0.9854	0.9746
15	0.0138	0.9947	0.9857	0.9753	0.0003	0.9786	0.9860	0.9759

Fig. 5. Training and validation accuracy

The accuracy, recall, and F1-score for the classes 'Normal' and 'Pneumonia' were computed to offer a complete evaluation of the model's performance. For the 'Pneumonia' class, strong recall rates were especially stressed, given the important necessity to properly identify all suspected instances of pneumonia. The model obtained high recall, guaranteeing that it barely missed any genuine instances of pneumonia, thereby confirming its applicability in medical contexts where failure to identify a disease might have serious implications. The accuracy measure, although likewise robust, was calibrated against the recall to ensure that the model would not unnecessarily forecast pneumonia where it did not present, thereby preventing undue concern or therapies for patients. Table VI represents the classification report of the proposed model.

TABLE VI. CLASSIFICATION REPORT

Class	Precision	Recall	F1-Score	Support
Normal	0.9589	0.9558	0.9573	317
Pneumonia	0.9836	0.9848	0.9842	855
Macro Avg	0.9713	0.9703	0.9708	1172
Weighted Avg	0.9769	0.9770	0.9770	1172

In the context of medical image analysis, these measures underline the model's dependability and accuracy in diagnosis. High F1-scores across all classes demonstrate that the model offers a balanced approach between accuracy and recall, successfully regulating the trade-offs between both measures, which is critical in medical diagnosis. Fig. 6 presents the classification report of the proposed model.

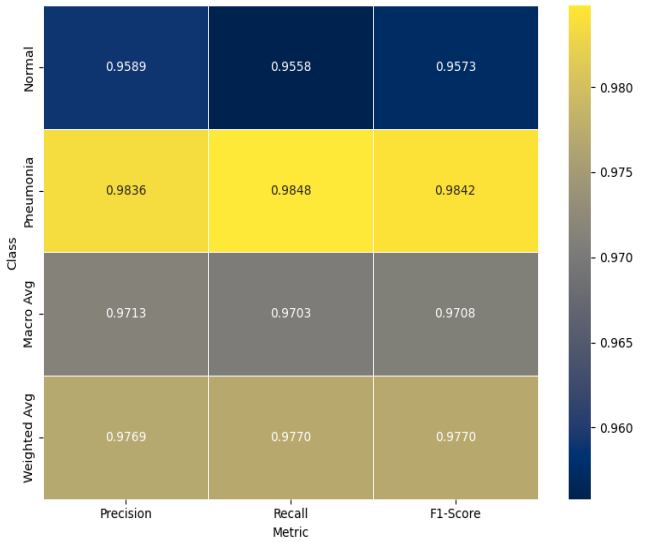

Fig. 6. Classification report

Further analysis was offered by sophisticated metrics such as the F2-Score, Matthews Correlation Coefficient (MCC), and Cohen's Kappa. The F2-Score, which weights recall more than accuracy, was especially beneficial in this medical application, showing the model's effectiveness in recognizing positive instances of pneumonia. The MCC gave an overall quality of the binary classifications, and its high value indicated a significant connection between the observed and anticipated classes. Cohen's Kappa also revealed a high degree of consistency and dependability in the model's predictions, allowing for any chance agreement. The advanced metrics are shown in the Table VII, and Fig. 7.

Mean Squared Error (MSE) and Mean Absolute Error (MAE) were utilized to measure the average magnitude of the model's errors in predictions, with both metrics showing low values, indicating minor deviations from the actual values, which highlights the model's accuracy in clinical predictions.

The error metrics is shown in the Table VIII. Fig. 8 represents the error metrics.

TABLE VII. ADVANCED METRICS

Metric	Value
F2-Score (Weighted)	0.9770
Matthews Correlation Coefficient (MCC)	0.9416
Cohen's Kappa	0.9416

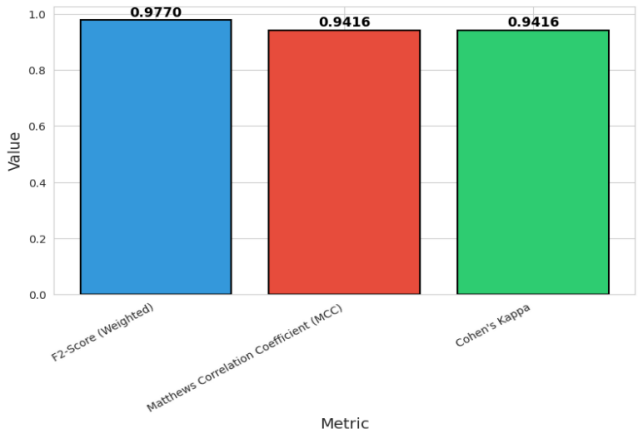

Fig. 7. Advanced metrics

TABLE VIII. ERROR METRICS

Metric	Value
Mean Squared Error (MSE)	0.0230
Root Mean Squared Error (RMSE)	0.1518
Mean Absolute Error (MAE)	0.0230

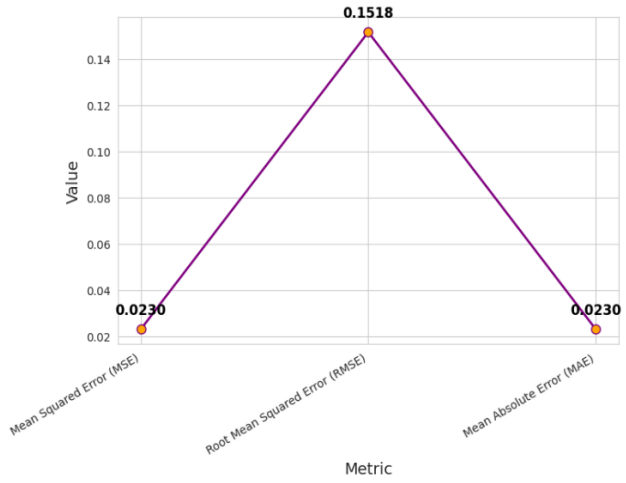

Fig. 8. Error Metrics

The confusion matrix for the model exhibited an amazing number of true positives and true negatives, notably for the 'Pneumonia' class, which is critical for a medical diagnostic tool. The matrix also exhibited very few false positives and false negatives, showing the model's effective separation between 'Normal' and 'Pneumonia' instances. This interpretation is crucial since it immediately reflects the model's practical application in a clinical environment, ensuring that patients get prompt and correct diagnoses. The confusion matrix is shown in Fig. 9.

The Receiver Operating Characteristic (ROC) curve was generated to assess the model's diagnostic performance. Thecurve revealed a large coverage area, showing the model's outstanding discriminative strength between the classes. The Area Under the Curve (AUC) was also reported and found to be quite high, underlining the model's efficiency in differentiating between 'Normal' and 'Pneumonia' with high sensitivity and specificity. The ROC curve and its related AUC is shown in Fig. 10.

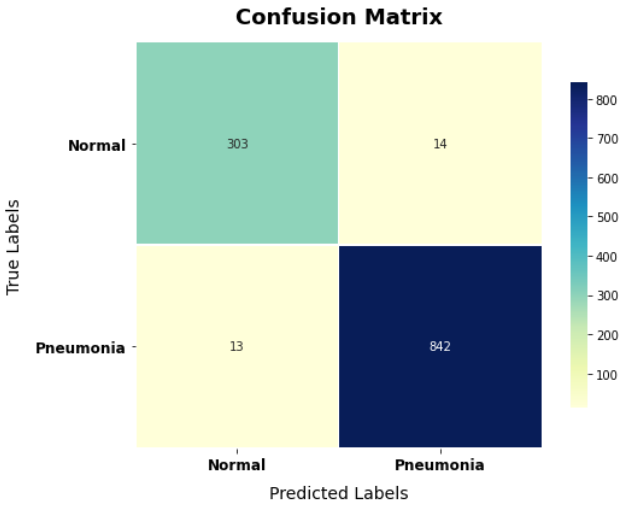

Fig. 9. Confusion matrix

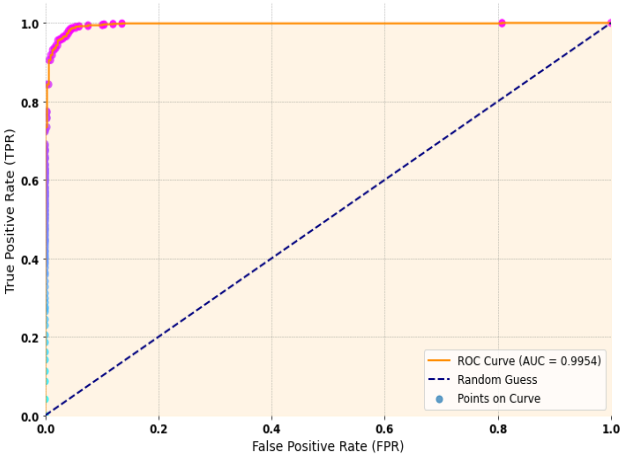

Fig. 10. ROC-AUC Curve

Examples of chest X-ray pictures from the test set were exhibited, showing examples where the model correctly detected the proper categories, along with instances of misclassifications. These visualizations help not only to show the model's prediction capabilities but also to give real-world instances of its use in medical diagnostics.

An examination of misclassified photos identified similar patterns or traits that lead to prediction mistakes as shown in Fig. 11. These results were analyzed to identify probable explanations for the inaccuracies, such as picture quality, uncommon presentations of pneumonia, or overlapping symptoms with other illnesses, and their implications for clinical usage.

Unexpected outcomes or abnormalities in the model's predictions were emphasized, with comments on probable causes for these occurrences, such as data anomalies or model sensitivity. This conversation is critical for improving the

model and preparing it for implementation in varied clinical situations. Table IX illustrates the comparison of proposed model with existing techniques.

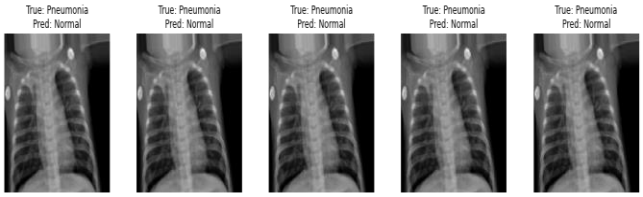

Fig. 11. Misclassified instances

TABLE IX. COMPARATIVE STUDY

Author	Techniques	Accuracy
Shadi A. Aljawarneh et. al (2023) [07]	Enhanced CNN model	92.4%
Parul Nasra (2024) [08]	CNN	91%
Ruchika Bhuria and Sheifali Gupta (2024) [09]	DenseNet-161	89.26%
Mutammimul Ula et. al (2022) [10]	Multilayer Perceptron	92.8%
Pranathi Patel and Hiriyanna GS (2022) [11]	CNN and Data Augmentation	81.89%
Yuting Yang et. al (2022) [12]	Deep Learning	83.968%
Pham Ngoc Ha et. al (2023) [13]	VIT-B/16	94%
Stephen George et. al (2022) [14]	CNN	93%
Anjana Jalnnavar et. al (2023) [15]	ResNet50V2	90%
Ashrafee, Md. Iftid et. al (2023) [16]	CNN	90.75%
Abhishek Dixit et. al (2024) [17]	LO-CNN-LSTM	92.82%
Proposed Methodology	EfficientNetB4	97.7%

V. DISCUSSION

The discussion portion of this paper critically assesses the results, contextualizing the performance of a deep learning model developed to identify between 'Normal' and 'Pneumonia' chest X-ray pictures. The model's strong training and validation accuracies reflect its robust learning capabilities and stability over epochs, indicating good feature extraction and pattern identification from complicated medical pictures. The steady increase in performance indicators during the training period confirms the selected architecture and learning technique, which are especially tuned to handle the subtle needs of medical image processing. Precision, recall, and F1-scores for both classes, especially the high recall for the 'Pneumonia' class, highlight the model's clinical relevance, as it reliably identifies most positive cases, thereby minimizing the risk of false negatives, which are critically important in medical diagnostics to ensure that no condition goes undiagnosed. The deployment of advanced measures such as the F2-Score, Matthews Correlation Coefficient, and Cohen's Kappa enables a wider evaluation of the model's diagnostic accuracy, dependability, and consistency. These measures emphasize the model's capacity to perform well across varied ailments and patient demographics, further proving its potential value in real-world clinical settings. The low values of Mean Squared Error and Mean Absolute Error underscore the model's precision in prediction, validating its applicability for clinical use where accuracy is crucial.

The confusion matrix and the comprehensive classification reports give deep insights into the model's performance, displaying a high number of true positives and true negatives while retaining a low rate of false positives and negatives. This balance is critical for medical applications since the cost of misdiagnosis may be considerable, both in terms of patient outcomes and healthcare resources. Moreover, the ROC curve and the related AUC indicate the model's outstanding discriminative power, suggesting not only its efficacy in categorizing the pictures properly but also its ability to retain this performance over a variety of decision thresholds. This flexibility enables for modifications in clinical settings, where various conditions may call for emphasizing sensitivity over specificity, or vice versa. Visual representations of predictions and misclassifications in the model predictions and error analysis sections not only illustrate the model's practical application but also provide insights into the types of errors it makes, such as those due to overlapping symptoms with other conditions or anomalies in image quality. Understanding these trends is critical for further improving the model and preparing it for efficient deployment in a varied variety of clinical situations.

The discussion of unexpected outcomes or anomalies in the model's predictions underscores the complicated nature of medical diagnostics and the difficulty inherent in designing AI systems that can adapt to the unpredictability in real-world data. These anomalies may inspire deeper inspection of the data and model behavior, perhaps leading to further refinements or tweaks to the model's training and generalization capabilities. Moreover, the relevance of these results on the study's overall conclusions is non-trivial, as they underline the necessity for constant review and adaption of the model, particularly when implemented in varied geographical or demographical situations. This flexibility is critical to maintain the generalizability and reliability of the model, since the features of patient populations may considerably impact diagnostic algorithms.

The discussion of this research not only validates the model's performance and possible usefulness in clinical settings but also underlines the intricacies and obstacles of incorporating AI into healthcare diagnostics. The high degree of accuracy and dependability revealed by the model offers a solid basis for its application as a diagnostic tool, possibly boosting patient outcomes via speedier and more accurate identification of pneumonia from chest X-ray images. However, it also emphasizes the necessity for ongoing testing, monitoring, and adaptation of the model to meet the diverse and evolving needs of healthcare providers and patients alike, ensuring that the technology remains a beneficial adjunct to clinical expertise rather than a standalone solution.

VI. CONCLUSION

This work has proved the efficiency and dependability of a deep learning model particularly developed to categorize chest X-ray pictures into 'Normal' and 'Pneumonia' categories. Through rigorous training and validation methods, the model has demonstrated great accuracy and the capacity to successfully discriminate between the two groups, showing its potential as a helpful tool in medical diagnostics.

The design of the model, based on the EfficientNetB4 framework, was rigorously selected and tuned to adapt to the specific problems provided by medical imaging, notably the subtleties and essential features evident in chest X-rays.

Throughout the training period, the model demonstrated steady increases in accuracy and stability, supported by both training and validation measures. The use of sophisticated classification measures such as accuracy, recall, and F1-scores proved the model's capacity to achieve a balanced performance, assuring both high sensitivity in recognizing pneumonia cases and specificity in confirming normal instances. This balance is critical in medical contexts where the expense of false negatives may be life-threatening and the consequences of false positives can lead to undue stress and therapy. The model's solid performance was further underlined by sophisticated measures such as the F2-Score, Matthews Correlation Coefficient, and Cohen's Kappa, which validated its dependability and diagnostic accuracy. The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) provided additional evidence of the model's discriminative power, offering flexibility in clinical application through adjustable threshold settings that can be tailored to specific medical needs or patient populations.

Moreover, the model confirmed its actual application via the study of a confusion matrix and thorough classification reports, which indicated a high proportion of true positive and true negative predictions. These findings are promising since they validate the model's capabilities to perform successfully within a clinical context, giving vital help to medical professionals in detecting pneumonia. However, the research also identified and addressed significant limits and anomalies discovered throughout the model assessment. The discovery of recurrent patterns in misclassified photos provided a better understanding of the model's limits, pointing to opportunities for additional development, such as strengthening its capacity to manage image abnormalities and differences in illness presentation. The discussion of unexpected outcomes underlined the necessity for continual model review and adaption to preserve its accuracy and dependability across varied clinical situations.

The outcomes from this research imply that with continuous improvement and validation, the deep learning model has the potential to serve as a substantial assist in the diagnosing process of pneumonia from chest X-rays. It provides a promising supplement to existing diagnostic procedures, possibly leading to quicker and more accurate diagnoses, better patient care, and optimum allocation of medical resources. Future work will concentrate on increasing the dataset, including multi-center data to strengthen the model's generalizability, and performing real-time clinical trials to further confirm its efficacy and usability in a larger clinical setting.

REFERENCES

[1] N. P. O'Grady and S. S. Kadri, "After the COVID-19 Crisis: Update on Complex Infectious Disease Issues in the Intensive Care Unit," *Infectious Disease Clinics of North America*, vol. 36, no. 4, 2022.

[2] S. Sodero, N. Dhungana, and F. Sandoe, "Vital mobilities of medical oxygen: Theorising oxygen justice," *Social Science & Medicine*, vol. 364, p. 117464, 2025.

[3] S. Mehra, S. K. Singh, S. Kumar, H. Dubey, V. Arya, and B. B. Gupta, "Transformative Approach for Prediction of Lung Disease Using Lung X-ray Images," *Research Square*, 2024.

[4] D. Kermany, K. Zhang, and M. Goldbaum, "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification," *Mendeley Data*, 2018.

[5] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, "A comprehensive survey of image augmentation techniques for deep learning," *Pattern Recognition*, vol. 137, p. 109347, 2023.

[6] E. A. Alabdulqader, M. Umer, K. Alnowaiser, H. Wang, A. A. Alarfaj, and I. Ashraf, "Image Processing-based Resource-Efficient Transfer Learning Approach for Cancer Detection Employing Local Binary Pattern Features," *Mobile Networks and Applications*, vol. 29, pp. 1351-1367, 2024.

[7] S. A. Aljawarneh and R. Al-Quraan, "Pneumonia detection using enhanced convolutional neural network model on chest x-ray images," *Big Data*, vol. 13, no. 1, pp. 16-29, 2025.

[8] P. Nasra, "Convolution Neural Network: A Proposed Method for Detecting Pneumonia from X-Ray Images," *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*, pp. 1692-1697, 2024.

[9] R. Bhuria and S. Gupta, "Innovative AI Solutions for Pneumonia Detection: Exploring DenseNet-161 in Medical Imaging," *2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI)*, pp. 638-643, 2024.

[10] M. Ula, M. Muhathir, and I. Sahputra, "Optimization of multilayer perceptron hyperparameter in classifying pneumonia disease through X-Ray images with speeded-up robust features extraction method," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 10, pp. 203-210, 2022.

[11] P. Patel and G. S. Hiriyanna, "Pneumonia detection in x-ray chest images based on cnn and data augmentation," *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 7, pp. 4639-4654, 2022.

[12] Y. Yang and G. Mei, "Pneumonia recognition by deep learning: A comparative investigation," *Applied Sciences*, vol. 12, no. 9, p. 4334, 2022.

[13] P. N. Ha, A. Doucet, and G. S. Tran, "Vision Transformer for Pneumonia Classification in X-ray Images," *Proceedings of the 2023 8th International Conference on Intelligent Information Technology*, pp. 185-192, 2023.

[14] S. George, N. G, S. Saifullah, and I. R. Praveen Joe, "An Efficient Neural Network Model For Pneumonia Detection," *2022 International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS)*, pp. 1-7, 2022.

[15] A. Jalnnavar, S. G. Kanakareddi, V. S. Handur, A. Chikaraddi, and S. Giraddi, "Disease Prediction Model Using Deep Transfer Learning," *International Conference on Intelligent Computing Systems and Applications*, pp. 13-29, 2023.

[16] M. I. Ashrafee, K. B. Sourav, M. K. Dolna, and S. Haque. *Pneumonia Disease detection using the convolutional neural network*. Brac University, 2023.

[17] A. Dixit, A. Mani, and S. Gorbachev, "Optimizing Medical Image Classification: A Hybrid Approach with Integrated Processing and Deep Learning," *2024 Asia Pacific Conference on Innovation in Technology (APCIT)*, pp. 1-4, 2024.

[18] M. Trivedi and A. Gupta, "A lightweight deep learning architecture for the automatic detection of pneumonia using chest X-ray images," *Multimedia Tools and Applications*, vol. 81, no. 4, pp. 5515-5536, 2022.

[19] M. Tomar, "The Chest X-ray," *IAP Specialty Series on Pediatric Cardiology*, p. 78, 2021.

[20] N. Barakat, M. Awad, and B. A. Abu-Nabah, "A machine learning approach on chest X-rays for pediatric pneumonia detection," *Digital Health*, vol. 9, 2023.

[21] S. Ramgopal, L. Ambroggio, D. Lorenz, S. S. Shah, R. M. Ruddy, and T. A. Florin, "A prediction model for pediatric radiographic pneumonia," *Pediatrics*, vol. 149, no. 1, p. e2021051405, 2022.

[22] E. Ayan, B. Karabulut, and H. M. Ünver, "Diagnosis of pediatric pneumonia with ensemble of deep convolutional neural networks in

chest x-ray images," *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 2123-2139, 2022.

[23] A. T. Geanacopoulos, M. I. Neuman, S. C. Lipsett, M. C. Monuteaux, and K. A. Michelson, "Association of chest radiography with outcomes in pediatric pneumonia: a population-based study," *Hospital pediatrics*, vol. 13, no. 7, pp. 614-623, 2023.

[24] C. Shi, X. Xu, and Y. Xu, "Systematic review and meta-analysis of the accuracy of lung ultrasound and chest radiography in diagnosing community acquired pneumonia in children," *Pediatric pulmonology*, vol. 59, no. 12, pp. 3130-3147, 2024.

[25] J. A. Prakash, V. Ravi, V. Sowmya, and K. P. Soman, "Stacked ensemble learning based on deep convolutional neural networks for pediatric pneumonia diagnosis using chest X-ray images," *Neural Computing and Applications*, vol. 35, no. 11, pp. 8259-8279, 2023.

[26] S. Padash, M. R. Mohebbian, S. J. Adams, R. D. Henderson, and P. Babyn, "Pediatric chest radiograph interpretation: how far has artificial intelligence come? A systematic literature review," *Pediatric Radiology*, vol. 52, no. 8, pp. 1568-1580, 2022.

[27] Y. Li, L. Zhang, H. Yu, J. Wang, S. Wang, J. Liu, and Q. Zheng, "A comprehensive segmentation of chest X-ray improves deep learning–based WHO radiologically confirmed pneumonia diagnosis in children," *European Radiology*, vol. 34, no. 5, pp. 3471-3482, 2024.

[28] S. C. Lipsett, A. W. Hirsch, M. C. Monuteaux, R. G. Bachur, and M. I. Neuman, "Development of the novel pneumonia risk score to predict radiographic pneumonia in children," *The Pediatric Infectious Disease Journal*, vol. 41, no. 1, pp. 24-30, 2022.

[29] E. L. Field, W. Tam, N. Moore, and M. McEntee, "Efficacy of artificial intelligence in the categorisation of paediatric pneumonia on chest radiographs: a systematic review," *Children*, vol. 10, no. 3, p. 576, 2023.

[30] H. J. Shin, N. H. Son, M. J. Kim, and E. K. Kim, "Diagnostic performance of artificial intelligence approved for adults for the interpretation of pediatric chest radiographs," *Scientific reports*, vol. 12, no. 1, p. 10215, 2022.

[31] J. D. Bodapati and V. N. Rohith, "ChxCapsNet: Deep capsule network with transfer learning for evaluating pneumonia in paediatric chest radiographs," *Measurement*, vol. 188, p. 110491, 2022.

[32] E. Alexopoulou *et al*., "Imaging of Acute Complications of Community-Acquired Pneumonia in the Paediatric Population—From Chest Radiography to MRI," *Children*, vol. 11, no. 1, p. 122, 2024.

[33] I. Mese, C. A. Mese, U. Demirsoy, and Y. Anik, "Innovative advances in pediatric radiology: computed tomography reconstruction techniques, photon-counting detector computed tomography, and beyond," *Pediatric Radiology*, vol. 54, no. 1, pp. 1-11, 2024.

[34] J. A. Prakash, C. R. Asswin, V. Ravi, V. Sowmya, and K. P. Soman, "Pediatric pneumonia diagnosis using stacked ensemble learning on multi-model deep CNN architectures," *Multimedia tools and applications*, vol. 82, no. 14, pp. 21311-21351, 2023.

[35] A. Gupta, P. Sheth, and P. Xie, "Neural architecture search for pneumonia diagnosis from chest X-rays," *Scientific reports*, vol. 12, no. 1, p. 11309, 2022.

[36] V. Ravi, "Deep fine-tuned efficientnetv2 ensemble deep learning approach for pediatric pneumonia detection using chest radiographs," *Journal of Intelligent & Fuzzy Systems*, 2024.

[37] U. Bal, A. Bal, Ö. T. Moral, F. Düzgün, and N. Gürbüz, "A deep learning feature extraction-based hybrid approach for detecting pediatric pneumonia in chest X-ray images," *Physical and Engineering Sciences in Medicine*, vol. 47, no. 1, pp. 109-117, 2024.

[38] M. Ali *et al*., "Pneumonia Detection Using Chest Radiographs With Novel EfficientNetV2L Model," *IEEE Access*, vol. 12, pp. 34691-34707, 2024.

[39] 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. 1, no. 4, pp. 1589-1601, 2024.

[40] S. Sharma and K. Guleria, "A systematic literature review on deep learning approaches for pneumonia detection using chest X-ray images," *Multimedia Tools and Applications*, vol. 83, no. 8, pp. 24101-24151, 2024.

[41] X. H. Lan, Y. X. Zhang, W. H. Yuan, F. Shi, and W. L. Guo, "Image-based deep learning in diagnosing mycoplasma pneumonia on pediatric chest X-rays," *BMC pediatrics*, vol. 24, no. 1, p. 720, 2024.

[42] P. Szepesi and L. Szilágyi, "Detection of pneumonia using convolutional neural networks and deep learning," *Biocybernetics and biomedical engineering*, vol. 42, no. 3, pp. 1012-1022, 2022.

[43] M. Kaya, "Feature fusion-based ensemble CNN learning optimization for automated detection of pediatric pneumonia," *Biomedical Signal Processing and Control*, vol. 87, p. 105472, 2024.

[44] D. S. Kermany *et al*., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122-1131.e9, 2018.

[45] D. Nessipkhanov, V. Davletova, N. Kurmanbekkyzy, and B. Omarov, "Deep CNN for the Identification of Pneumonia Respiratory Disease in Chest X-Ray Imagery," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 10, 2023.

[46] V. Arya and T. Kumar, "Enhancing image for CNN-based diagnostic of pediatric pneumonia through chest radiographs," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 2, 2023.

[47] D. Yao, Z. Xu, Y. Lin, and Y. Zhan, "Accurate and intelligent diagnosis of pediatric pneumonia using X-ray images and blood testing data," *Frontiers in bioengineering and biotechnology*, vol. 11, p. 1058888, 2023.

[48] N. Shilpa, W. Ayeesha Banu, and P. B. Metre, "Revolutionizing Pneumonia Diagnosis: AI-Driven Deep Learning Framework for Automated Detection From Chest X-Rays," *IEEE Access*, vol. 12, pp. 171601-171616, 2024.

[49] S. S. Khan, A. U. H. Rupak, W. W. Faieaz, S. Jannat, N. N. I. Prova, and A. K. Gupta, "Advances in Medical Imaging: Deep Learning Strategies for Pneumonia Identification in Chest X-rays," *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1-7, 2024.

[50] A. P. J *et al*., "Transfer learning approach for pediatric pneumonia diagnosis using channel attention deep CNN architectures," *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106416, 2023.

[51] Z. Pan, H. Wang, J. Wan, L. Zhang, J. Huang, and Y. Shen, "Efficient federated learning for pediatric pneumonia on chest X-ray classification," *Scientific Reports*, vol. 14, no. 1, p. 23272, 2024.

[52] M. Yaseliani, A. Z. Hamadani, A. I. Maghsoodi, and A. Mosavi, "Pneumonia Detection Proposing a Hybrid Deep Convolutional Neural Network Based on Two Parallel Visual Geometry Group Architectures and Machine Learning Classifiers," *IEEE Access*, vol. 10, pp. 62110-62128, 2022.

[53] S. Domínguez-Rodríguez *et al*., "Testing the performance, adequacy, and applicability of an artificial intelligence model for pediatric pneumonia diagnosis," *Computer Methods and Programs in Biomedicine*, vol. 242, p. 107765, 2023.

[54] A. E. Yildirim and M. Canayaz, "A novel deep learning-based approach for prediction of neonatal respiratory disorders from chest X-ray images," *Biocybernetics and Biomedical Engineering*, vol. 43, no. 4, pp. 635-655, 2023.

[55] M. Neshat, M. Ahmed, H. Askari, M. Thilakaratne, and S. Mirjalili, "Hybrid Inception Architecture with Residual Connection: Fine-tuned Inception-ResNet Deep Learning Model for Lung Inflammation Diagnosis from Chest Radiographs," *Procedia Computer Science*, vol. 235, pp. 1841-1850, 2024.

[56] A. B. Godbin and S. G. Jasmine, "Pediatric Pneumonia Detection in Chest X-ray Images: A Deep Feature Analysis Approach Enhanced with LightGBM," *2024 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, pp. 1-6, 2024.

[57] O. Güler and K. Polat, "Classification performance of deep transfer learning methods for pneumonia detection from chest x-ray images," *Journal of Artificial Intelligence and Systems*, vol. 4, no. 1, pp. 107-126, 2022.

[58] S. S. Vemuri, S. Kotha, S. Voruganti, P. R. Kunam, A. Nandini, and T. K. Mishra, "Pneumonia Detection from X-Ray Images Using Deep Transfer Learning," *International Conference on Computing, Communication and Learning*, pp. 205-219, 2024.

[59] A. Panwar and S. Gupta, "Deep learning models for early detection of pneumonia using chest X-ray images," *Proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences*, pp. 701-709, 2022.

[60] O. Papadimitriou, A. Kanavos, and M. Maragoudakis, "Automated Pneumonia Detection from Chest X-Ray Images Using Deep Convolutional Neural Networks," *2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA)*, pp. 1-4, 2023.

[61] S. Sajed, A. Sanati, J. E. Garcia, H. Rostami, A. Keshavarz, and A. Teixeira, "The effectiveness of deep learning vs. traditional methods for lung disease diagnosis using chest X-ray images: A systematic review," *Applied Soft Computing*, vol. 47, p. 110817, 2023.

[62] M. Rameez, S. Kumar K, and M. J. Logashree, "Pneumonia Detection using Machine Learning Techniques," *2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 1803-1808, 2024.

[63] 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.