Int. J. Appl. Math. Comput. Sci., 2024, Vol. 34, No. 4, 679–699
DOI: 10.61822/amcs-2024-0045

PNEUMONIA DETECTION: A COMPREHENSIVE STUDY OF DIVERSE NEURAL NETWORK ARCHITECTURES USING CHEST X-RAYS

WAJAHAT AKBAR[a], ABDULLAH SOOMRO[b], ALTAF HUSSAIN[c], TARIQ HUSSAIN[d,*], FARMAN ALI[e], MUHAMMAD INAM UL HAQ[c], RAAZ WAHEEB ATTAR[f], AHMED ALHOMOUD[g], AHMAD ALI ALZUBI[h], REEM ALSAGRI[i]

[a] School of Electronic and Control Engineering
Chang'an University, Beilin, Xi'an, Shaanxi, 710061, China

[b] Department of Computer Science
The Islamia University of Bahawalpur, Rabia Hall Rd, Bahawalpur, 63100, Pakistan

[c] Department of Computer Science and Bioinformatics
Khushal Khan Khattak University Karak, Methawalah, Karak, 27200, Pakistan

[d] School of Computer Science and Technology
Zhejiang Gongshang University, 18 Xuezheng St., Hangzhou, 310018, China
e-mail: uom.tariq@gmail.com

[e] Department of Applied AI
Sungkyunkwan University, Seoul, 03063, South Korea

[f] Management Department
Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh, 11671, Saudi Arabia

[g] Department of Computer Science
Northern Border University, Arar, 91431, Saudi Arabia

[h] Department of Computer Science
King Saud University, Riyadh, 11495, Saudi Arabia

[i] Department of Software Engineering
University of Hafr Al Batin, Hafar Al Batin, 39524, Saudi Arabia

Pneumonia is of deep concern in healthcare worldwide, being the most deadly infectious disease, especially among children. Chest radiographs are crucial for detecting it. However, certain vulnerable groups exhibit heightened susceptibility, emphasizing the critical nature of accurate diagnosis and timely intervention. This paper presents convolutional neural network (CNN) models for the detection of pneumonia from chest X-rays images. Among 20 different CNN models, we identified EfficientNet-B0 as the most accurate and efficient, boasting an impressive accuracy rate of 94.13%. Furthermore, the precision, recall, and F-score metrics for this model stand at 93.50%, 92.99%, and 93.14%, respectively. This research underscores the potential of CNNs to revolutionize pneumonia diagnosis.

Keywords: pneumonia detection, CNN models, chest X-ray, medical imaging.

1. Introduction

Pneumonia, an infection targeting the delicate alveoli nestled within the lungs, triggers inflammation and swelling, complicating breathing. The emergence of pneumonia heralds a distressing escalation in respiratory challenges. Its presence is characterized by heightened sputum production amidst coughing bouts, fever, breathlessness, chest discomfort, and diminished appetite. Individuals with pre-existing respiratory ailments face a greater vulnerability to severe illness,

*Corresponding author.

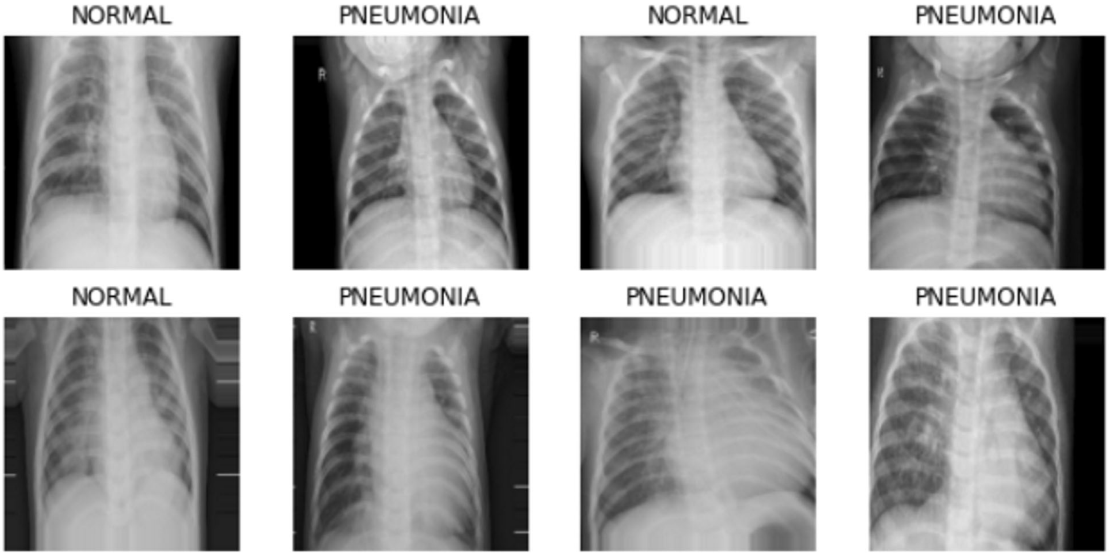

Fig. 1. Sample of normal and pneumonia images.

particularly accentuated during the winter months, when pneumonia emerges as the leading cause of mortality. Alarmingly, this condition claims the lives of over 15% of children under the age of five, underscoring its significant impact on public health (Rao *et al*., 2022).

Pneumococcal illness significantly affects populations in developing and underdeveloped nations, primarily due to inadequate healthcare resources and a burgeoning patient demographic (Regunath and Oba, 2021). The prevalence of community-acquired pneumonia (CAP) worldwide varies, with estimates ranging from 1.5% to 14% cases per 1000 person-years. Geographic location, seasonal fluctuations, and demographer yearly prevalence rate in the United States are documented as 24.8% scenarios per 10,000 individuals, and there is an apparent association between the rates of occurrence and advancing characteristics that influence this variability.

The WHO recognizes pneumonia as the leading infectious illness and ranks the seventh most common cause of death worldwide. Figure 1 is a sample and standard pneumonia images. It should be noted that mortality rates among patients who require admission to the intensive care unit can exceed 23%, highlighting the grave consequences associated with this disease (Yang *et al*., 2022). Hence, prompt treatment of pneumonia is imperative.

Radiologists rely on chest radiographs for the clinical diagnosis of lung disorders (Akbar *et al.*, 2023a). Chest X-rays are widely employed in medical practice due to their swift turnaround, cost-effectiveness, and satisfactory image clarity. However, despite these advantages, radiological findings alone may not definitively differentiate pneumonia from other pulmonary conditions. Moreover, the challenge extends beyond diagnosis. Hospitals and medical institutions generate a vast volume of medical images daily, necessitating radiologists to meticulously review numerous images manually (Shah and Shah, 2022).

Deep learning algorithms present a promising solution for helping doctors identify areas of the lungs affected by pneumonia. Integrating deep learning into medical imaging automation has garnered considerable attention recently (Akbar *et al.*, 2023b). Deep learning models have consistently performed better than human professionals in several IoT medical imaging scenarios, including IoT healthcare privacy and security (Singh and Tripathi, 2022; Hussain *et al*., 2024; Cheng *et al*., 2024). Since the introduction of AlexNet in 2012, deep learning models have made significant progress in classifying images (Heravi *et al*., 2016; Kang *et al*., 2024). ResNet and its modifications provide a robust foundation for precise object recognition and localization. The YOLO algorithm performs very quickly and accurately in detecting objects (Redmon *et al*., 2016), while RetinaNet is suitable for real-time applications of various types. GANs, as noted Iqbal and Ali (2018), play a critical role in unsupervised learning of the efficacy of CNN models when there is little training data or when they are poor quality.

Both issues are successfully resolved by incorporating deep learning automation. The study reported in this paper describes a potential method for reliably identifying pneumonia using CNN models trained on chest X-ray data. The key contributions of the proposed method are as follows:

- The most effective CNN architecture for pneumonia detection is identified. Among the models evaluated, the Efficient-B0 architecture emerged as the most accurate and efficient, achieving an impressive accuracy rate of 94.13%.

- The EfficientNet-B0 model also exhibited strong performance in key metrics, with a precision of 93.50% a recall of 92.99%, and an F-score of 93.14%.

- Various CNN architectures for pneumonia detection are explored and analyzed by training and evaluating 20 distinct CNN models on chest X-ray images to classify cases as pneumonia or normal.

- Challenges and limitations in identifying pneumonia are identified, as well as its significance within the broader context of medical research and healthcare.

2. Related work

Deep learning algorithms have significantly enhanced image processing techniques across various domains, including pathology, radiology, and mammography. In numerous biomedical applications, these algorithms have demonstrated capabilities that surpass human visual assessment (Litjens *et al.*, 2017). Recent technological breakthroughs have enabled the automation of pneumonia identification by analyzing chest X-rays. Several research initiatives have used deep CNN models for accurate and efficient pneumonia diagnosis. In the work of (Račić *et al.*, 2021), for pneumonia diagnosis, a deep learning model was developed to analyze chest X-ray images, yielding a 90% success rate. This model was assembled without data augmentation, image filtering, or segmentation tactics. Instead, data consistency was achieved by deliberately using images of low quality. Alqudah *et al.* (2021) developed an artificial intelligence approach to identify and categorize pneumonia in chest X-ray images. They proposed an AI strategy with 94% accuracy, using a CNN for improved feature extraction. The model underwent training with subpar images without using detection and filtering techniques to enhance image quality.

In the works of Manickam *et al.* (2021), Chen *et al.* (2024) and Shoaib *et al.* (2023), deep learning, transfer learning and several optimized algorithms were employed to detect pneumonia from chest X-rays. The images were utilized to detect the infection and then categorized as healthy or pneumonia-infected using a U-Net architecture. The first models were trained using selected ImageNet data. Evaluation techniques for the CNN models VGG16, RetinaNet + Mask RCNN, DenseNet169 + SVM, and Xception were modified from previous iterations. The ResNet50 model, renowned for its exceptional performance, attained 93% accuracy, 96% recall, 88% precision, and an F-measure of 92%. The model's accuracy was worse without image filtering and detection approaches, resulting in false positive diagnoses and inappropriate treatment decisions.

In the 121-layer CheXNet model proposed by Rajpurkar *et al.* (2017), 100,000 chest X-rays of 14 illnesses were used for training. The CheXNet model was tested on 420 chest X-ray images and compared with decisions made by professional radiologists. CheXNet and other deep learning-based CNN models surpassed radiologists in pneumonia detection.

Rajaraman *et al.* (2018) explored the effectiveness of designing specific customized CNN architectures for respective bacterial and viral pneumonia in 5232 community acquired paediatric chest radiographs. The return on investment (ROI) for the sequential CNN, residual CNN, VGG16, and inception CNN models was evaluated using a unique visualization approach. The modified VGG16 design performed more effectively than earlier versions, yielding 96.2% accuracy in pneumonia diagnosis and 93.6% in separating infections caused by viruses and bacteria.

In order to categorise various forms of pneumonia, Rahman *et al.* (2020) looked at 5247 chest X-ray images from the Kaggle pneumonia dataset. The transfer learning models used were AlexNet, ResNet18, Dense Net201, and Squeeze Net. DenseNet201 drove other models by comprehending distinctive pneumonia groups and accurately determining the two etiological variations with 98% accuracy.

Alqudah *et al.* (2021) demonstrated that modified CNNs could distinguish chest X-rays of patients with bacterial pneumonia from those without it. Afterwards, the support vector machine (SVM) and K-nearest neighbors (KNN) algorithms were applied. Ten-fold cross-validation generated a hybrid model that combined CNN-KNN and CNN-SVM. Earlier hybrid models achieved 94.03% accuracy, while the latter reached 93.9% accuracy. Additionally, chest X-rays were analyzed using deep learning to identify pneumonia. Through the use of pre-processing techniques before using the ResNet50 v2 deep learning framework, pneumonia identification accuracy was improved up to 96% according to Alsharif *et al.* (2021).

Alquran *et al.* (2021) categorized chest X-rays into three groups: pneumonia, COVID-19, and typical chest X-ray films, with a precision of 93.1% attained by the use of textural cues and conventional ML methods. According to Rajasenbagam *et al.* (2021), deep learning approaches effectively identified pneumonia infection. In image testing, the suggested CNN yielded 99.34% accuracy on a training dataset of 12,000 chest X-rays. The suggested CNN outperformed AlexNet, VGG16Net, and InceptionNet.

Table 1. Comparing our version of EfficientNet-B0 with other methods.

Approach	Models	Accuracy
Neuro-heuristic methodology (Ke *et al.*, 2019)	A new neuro-heuristic methodology is proposed for accurately identifying and categorizing the lung illnesses using X-ray pictures.	79.06%
Various pre-trained (CNNs) (Rahman *et al.*, 2020)	Pneumonia was correctly diagnosed from chest X-ray images using deep CNNs and transfer learning algorithms.	93.3%
VGG16 (Jain *et al.*, 2020)	CNNs and transfer learning techniques for detecting pneumonia in chest X-rays.	87.18%
VGG19 (Jain *et al.*, 2020)	Using CNNs and transfer learning techniques to detect pneumonia in chest X-ray images	88.46%
CNN architectures, including SVM and DenseNet-169 (Varshni *et al.*, 2019)	Detection of pneumonia using CNNs for feature extraction	80.02%
VGG16 and CNN (Liang and Zheng, 2020)	A paediatric pneumonia detection technique that combines deep residual networks with transfer learning	74.2%
RetinaNet + Mask RCNN (Sirazitdinov *et al.*, 2019)	A huge library of chest X-rays is used to locate pneumonia using a set of deep neural networks.	75.8%
VGG16 and Xception (Ayan and Ünver, 2019)	Chest X-ray images and deep learning are used to diagnose pneumonia.	87%
Fully connected RCNN (Rahmat *et al.*, 2019)	A faster R-CNN for categorizing chest X-ray images	62%
MobileNet + AEO (Sahlol *et al.*, 2020)	A unique technique for identifying tuberculosis in chest radiography involves optimizing deep neural network properties using artificial ecosystems.	90.20%
CNN+SVM (Alqudah *et al.*, 2021)	An artificial intelligence-based approach for accurately identifying and classifying pneumonia from chest radiography images	94%
DCNN (Rahimzadeh and Attar, 2020)	Xception and ResNet50V2 models are used in a modified deep CNN to identify COVID-19 and pneumonia cases from chest X-ray images.	91.4%
Transfer learning with deep learning (Manickam *et al.*, 2021)	The Xception and ResNet50V2 features are used in a complicated, deep CNN to accurately detect COVID-19 and pneumonia patients from chest X-ray images.	93.06%
EfficientNet-B0 (Frederich *et al.*, 2024)	Exploring different CNNs architectures to automate pneumonia detection on chest X-ray images.	94.13%

3. Methodology

This section covers recognizing and classifying pneumonia in chest X-rays. For the purpose of categorizing and detecting pneumonia, CNN architectures were used in this study. A CNN is implemented for finding irregularities in the gained chest X-ray data that signal pneumonia. A block diagram for the suggested methodology for various deep learning models is shown in Fig. 2.

3.1. Dataset. This dataset contains 5,863 X-ray images (JPEG), split into two groups: normal and with pneumonia. Images were collected from radiography exams on pediatric patients at the Guangzhou Medical Centre, which takes approximately 5,863 radiographs annually as part of standard clinical care. This research utilizes the pneumonia dataset (Kermany *et al.*, 2018) for training, validating, and testing as detailed in Table 2.

3.2. Data preprocessing. In this step, we applied augmentation using an ImageDataGenerator with

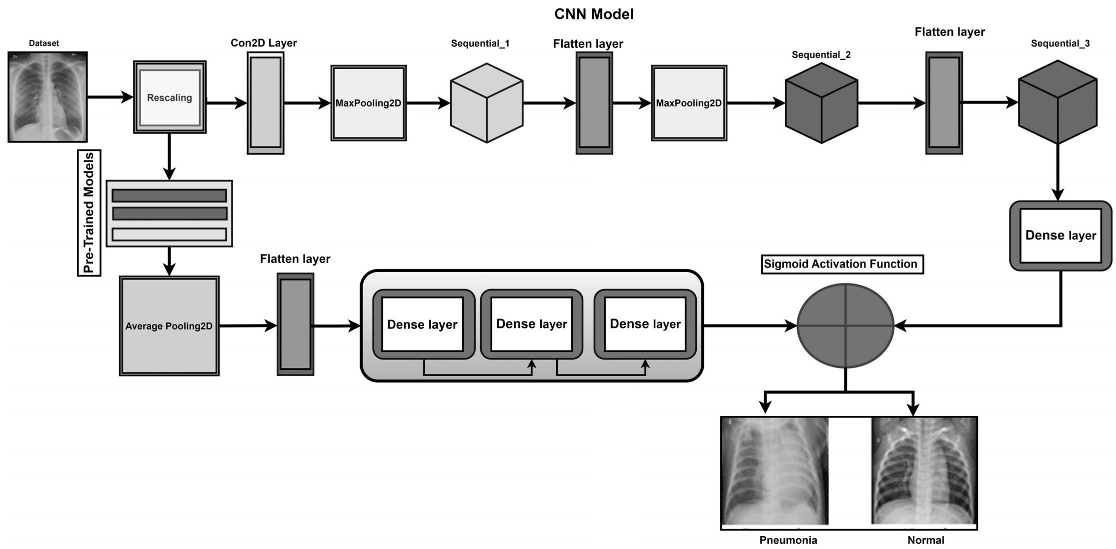

Fig. 2. Block diagram of the proposed model.

Table 2. Dataset splitting for models used.

Train	Pneumonia 3300	Normal 1147
Validation	Pneumonia 583	Normal 202
Test	Pneumonia 390	Normal 234

zooming, rotation, and shifting to avoid class imbalance in the data. This enriches the dataset and aids model generalization. Images are standardized to 224 $\times$ 224 pixels for consistency via rescaling. The augmented training data are used to train CNN models for classification and evaluation of the images. We normalized the image intensities before supplying the data as input for the models to train. Validation data tune the model, and the test set is used to determine its accuracy. This approach ensures effective training and generalization for pneumonia classification from chest X-rays.

3.3. Data augmentation. Machine learning and computer vision use data augmentation to enhance training datasets with modified versions of existing data. By applying the model to a wider variety of data instances, this procedure can help reduce overfitting, enhance model resilience, and boost prediction accuracy (Chlap *et al.*, 2021). Standard data augmentation methods include horizontal or vertical image flipping, image rotation, noise addition, and adjustment of brightness or contrast levels. The augmentation is applied to the training data set as shown in Table 4.

3.4. Rectified linear unit (ReLU) layer. The ReLU function belongs to the class of linear functions commonly utilized as an activation function within convolutional layers. The output is 1 for positive inputs and 0 for all other cases. ReLU has demonstrated superior performance in neural network architectures compared to alternative activation functions like sigmoid or hyperbolic tangent, primarily due to its ability to alleviate the vanishing gradient issue (Hara *et al.*, 2015). The ReLU function is

$$\begin{aligned} \mathrm{ReLU}(x) &= \max(0, x), \\ \text{Validation Loss} &= L(\theta_{\text{model}}). \end{aligned} \tag{1}$$

3.5. Early stopping. Cross-validation is a method used during training to measure the difference between training and validation errors, known as the generalization gap, which begins to widen instead of narrowing. This divergence typically indicates overfitting and can be mitigated by reducing model complexity, augmenting training data, applying regularization techniques, or implementing dropout (Prechelt, 2002). However, a pragmatic and practical approach to combat this issue is to terminate training prematurely once the generalization gap starts deteriorating (Prechelt, 2002). Specifically,

$$\text{Early Stopping:} \quad \begin{cases} \text{Stop training if} \quad L(\theta_{\text{model}}) \geq L_{\text{best}}, \\ \text{Update} \quad L_{\text{best}} \leftarrow L(\theta_{\text{model}}). \end{cases} \tag{2}$$

3.6. Sigmoid activation function. In binary classification tasks, neural networks often utilize the sigmoid activation function. This function effectively

Table 3. Comparison of the proposed and existing approaches.

Work	Type of nets	Dataset	Metrics	Optimizer	Problem type
Haloi *et al.* (2018)	Own implementation based on inception with 211 layers	ChestXray14, OCT and chest X-Ray (Mendeley), M-SC-Xray	AUC, ROC, sensitivity and specificity	Nesterov	Multiclass classification and multi-label
Liang and Zheng (2020)	Own implementation based on ResNet 51 layers with pretrained weights	OCT and chest X-Ray (Mendeley), M-SC-Xray	Confusion matrix, recall, F1-score, precision, accuracy, ROC, AUC	Adam	Binary classification
Stephen *et al.* (2019)	Own implementation	OCT and chest X-Ray (Mendeley)	Arithmetic mean of the accuracy of the network application with datasets of 5 different resolutions, accuracy, recall	Not specified	Binary classification
da Silva *et al.* (2020)	Networks without pretraining: VGG-16, InceptionResNetV2, ResNeXt50, InceptionV3 and ResNet50.	OCT and chest X-Ray (Mendeley)	Precision, confusion matrix and F1-score	Adam	Binary classification
Proposed approach	CNN + 19 pretrained models	Chest X-ray images (pneumonia)	Confusion matrix, recall, F1-score, precision, accuracy,	Adam	Binary classification

Table 4. Augmentation of the training dataset.

Rotation range	0.2
Zoom range	0.2
Height shift range	0.1
Width shift range	0.1

maps input values to a range between 0 and 1, enabling the calculation of probabilities from the sigmoid curve. Consequently, it is advantageous for binary prediction applications (Lau and Lim, 2018). Moreover, the smoothness and differentiability of the function enhance optimization tactics during neural network training. However, despite its widespread usage, the sigmoid function does have limitations (Sharma *et al.*, 2017).

The sigmoid function can be defined as

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \tag{3}$$

3.7. Optimization techniques. The choice of optimization algorithm significantly impacts the efficiency of training an ML model by minimizing error rates. Typically, the effectiveness of an optimizer is assessed based on its convergence speed and ability to generalize well (Venter, 2010). Adam is an optimizer that offers advantages across various metrics, building on the strengths of both the adaptive gradient algorithm (AdaGrad) and the root mean square propagation (RMSProp), which are popular variations of stochastic gradient descent (SGD) (Foulds, 2012). Adam aims to enhance convergence speed and generalization performance by computing distinctive adaptive learning rates for various factors. However, recent researches has unveiled scenarios where Adam may falter in converging to an optimal solution under specific configurations (Foulds, 2012).

The Adam optimization update is given by

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \tag{4}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \tag{5}$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \tag{6}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}, \tag{7}$$

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t. \tag{8}$$

In the Adam optimization algorithm, the moving average of the gradients is updated according to (4). By contrast, the moving average of the squared gradients is updated as specified in (5). The bias-corrected estimates for these moving averages are calculated using (6) and (7), respectively. Finally, the parameter update is performed according to (8).

3.8. Data regularization. Dropout is a commonly used regularising method that assists in minimizing overfitting in neural networks. Unlike other methods, such as L1 and L2 regularization, which introduce penalty terms to the cost function, dropout dynamically alters the network's architecture during each training iteration. This involves randomly deactivating neurons in the network, effectively creating an ensemble of neural networks for each iteration. As a result, these diverse networks tend to overfit in distinct manners, ultimately minimizing the overall risk of overfitting. We implemented dropout with a rate of 0.5 in all the models (Tian and Zhang, 2022). Mathematical equations for dropout layers are described in the following subsections.

3.8.1. L1 regularization (lasso). The L1 appropriate regularization term, also known as the lasso penalty, is defined as

$$\text{L1 regularization} = \lambda \sum_{j=1}^{p} |\beta_j|. \tag{9}$$

Here, λ represents the regularization parameter, p denotes the number of features, and β_j indicates the coefficient for the j-th feature.

3.8.2. L2 regularization (ridge). The appropriate L2 regularization term, also known as the ridge penalty, is defined as

$$\text{L2 regularization} = \lambda \sum_{j=1}^{p} \beta_j^2. \tag{10}$$

In this case, β_j stands for the coefficient of the j-th feature, while λ represents the regularization parameter and the number of features.

3.9. Binary cross-entropy. Binary cross-entropy is a common loss function in deep learning. The difference between the true label and the expected probability distribution is measured as 0 or 1 (Ruby and Yendapalli, 2020). Mathematically, the binary cross-entropy loss is stated as follows:

$$\text{Loss} = -y \log(p) + (1 - y) \log(1 - p). \tag{11}$$

In binary classification models, cross-entropy losses should be minimized. The model penalizes false positives and false negatives, meaning it penalizes forecasts of positive labels when the actual labels are negative. In situations where the costs of the two types of errors are comparable, it makes sense to use this method (Ramos *et al.*, 2018). By minimizing the binary cross-entropy loss, the model learns to make predictions that closely match the true labels, leading to accurate binary classification outcomes.

3.10. Accuracy. The accuracy metric is commonly used to evaluate binary classification tasks. Data availability and valid predictions are compared to evaluate the overall accuracy of predictions (Eusebi, 2013; Liao *et al*., 2023). Mathematically, accuracy is defined as

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TN} + \text{TP} + \text{FP} + \text{FN}}. \tag{12}$$

3.11. Recall. The recall metric quantifies the capability of a model to point out all potential matches in a dataset. It is the ability of correctly identifying all the positive instances in relation to the overall positive population through a particular model (Yacouby and Axman, 2020). Mathematically, recall is defined as

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{13}$$

3.12. Precision. An indicator of a model's accuracy and reliability is the precision of a statistical parameter (Yacouby and Axman, 2020). Mathematically, it is defined as

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{14}$$

3.13. F1-score. F1-scores are mathematical measures of accuracy and recall in classification models (Yacouby and Axman, 2020). The F1-score is expressed mathematically as

$$\text{F1} - \text{score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \tag{15}$$

3.14. Confusion matrix. Confusion matrices compare a model's predictions with the actual data patterns. These comparisons are organized in a matrix. This matrix shows the model's performance (Townsend, 1971). Figure 3 illustrates an instance of a confusion matrix of our different models through a dataset in which the problem domain is addressed in this paper.

4. Model description

4.1. CNN architecture. A CNN model consists of four layers: convolution, flattening, pooling, and fully connected (Alzubaidi *et al*., 2021; Jeczmionek and Kowalski, 2023). These four layers the discussed in the following subsections.

4.1.1. Convolutional layer (CONV). The CONV layer is a fundamental building component of CNNs, translating images into matrix representations from which convolution operations are carried out using a collection of adjustable parameters that can be learned. This study used a 3 $\times$ 3 filter, kernel, and mask to extract feature maps from the input matrix, leading to a significant decrease in the image size for further processing. Despite the possibility of an information loss during convolution procedures, the CONV layer preserves vital components of image regions within feature maps. The input matrix undergoes many filters, resulting in a hierarchical arrangement of feature maps. The feature map layer is subsequently subjected to grouping, flattening, and fully linked layers, resulting in the complete CNN architecture (Albawi *et al*., 2017).

4.1.2. Pooling layer. The pooling layer plays a vital role in downsizing representations by employing subsampling techniques over each activation map, thus reducing dimensions. This reduction not only maintains object integrity, but also alleviates computational complexity. This study utilizes two subsampling techniques: average pooling and maximal pooling. Max-pooling is a discrete approach that involves scanning the feature map using a 2 $\times$ 2 window and selecting the maximum value within the frame to emphasize important features. In contrast, average pooling, akin to max-pooling, computes the mean of the data inside the window. Compared with maximum-pooling, this strategy preserves more information from the image (Sun *et al*., 2017).

4.1.3. Flattening and fully-connected layer. The procedure included in the flattening and a completely linked layer is as follows: The aggregated feature map, obtained from previous layers, transforms the features into a single column. This step prepares the features to be fed into the neural network. Afterwards, the compressed feature map is inputted into the fully linked layer. Within the fully connected layer, inputs are forward propagated through the network, weight calculations are made, and predictions are generated. Based on these predictions, a cost function is computed to evaluate the network's performance (Basha *et al*., 2020).

4.2. Convolutional neural networks (CNNs). A CNN employs deep learning techniques for recognising images and videos. CNNs classify input data based on meaningful characteristics. Layers are typically convolutional, activation, pooling, and completely linked in CNNs. Convolutional layers use filters to extract features from input data, activation layers add non-linearity, pooling layers minimize spatial dimensions, and fully connected layers produce final classification predictions. CNNs are trained using big labeled datasets. Optimization approaches such as stochastic gradient descent modify the network weights to minimize the discrepancy between the expected and actual labels. Once trained, CNNs can make predictions based on previously

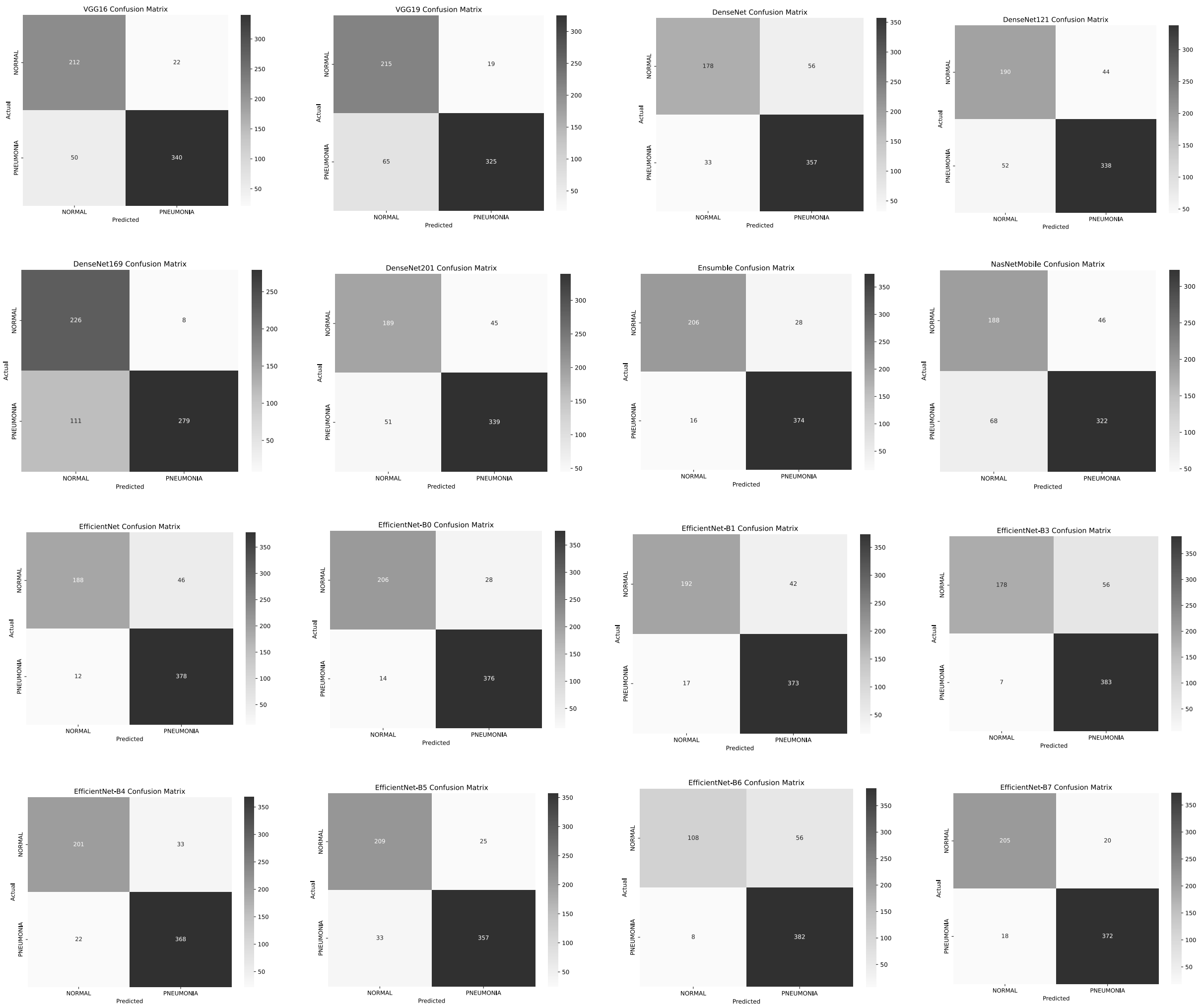

Fig. 3. Confusion matrices of our CNN models.

unidentified data. This study uses a CNN with twelve layers to achieve an accuracy of 87.82%, a precision of 89.31%, a recall of 84.87%, and an F-score of 86.33% (Albawi *et al.*, 2017). Furthermore, Table 5 presents a comparison with other models.

4.3. Dense convolutional network (DenseNet). A DenseNet enhances deep learning networks by including shorter connections across layers, thereby reducing training time, and increasing depth layers interconnecting CNN. There are links between adjacent layers, such as the first and second layers, second and third layers, and so on. As a consequence, most of the information is sent across network layers. Each layer gets information from the previous layer to retain the feed-forward nature. Unlike ResNet it concatenates features, rather than sums them. Thus, the feature maps from all preceding convolutional blocks have been incorporated into the i-th layer. The next layers, denoted as I-i, get the feature maps generated by this layer. This network differs from conventional deep learning designs in that it has a total of $I(I+1)/2$ connections. The proposed method requires fewer parameters than the CNN, eliminating the need to learn redundant feature mappings (Zhou *et al.*, 2022).

DenseNet needs fewer parameters than standard CNNs since it does not train duplicate feature maps (Rahman *et al.*, 2020). The layers of DenseNet are characterized by their narrow width, consisting of just 12 filters, which results in a reduced yet enriched collection of feature maps. We investigated pneumonia classification using DenseNet variations: DenseNet121, DenseNet169 (Akbar *et al.*, 2023c), and DenseNet201. DenseNet121 demonstrated the highest accuracy of 87.34%, with a precision of 87.03%, a recall of 85.68%, and an F-score of 86.26%. Following it closely, DenseNet achieved an accuracy of 85.74%, with a precision, a recall, and an F-score of 85.40%, 83.80%, and 84.46%, respectively. DenseNet201 also performed well, achieving an accuracy of 85.58%, with a precision, a recall, and an F-score of 84.58%, 84.70%, and 84.64%, respectively. However, DenseNet169 exhibited slightly lower performance, with 80.13% accuracy, 81.06% precision, 82.99% recall, and 79.96% F-score. Overall, DenseNet121 emerged as the best approach for pneumonia classification, closely followed by DenseNet and DenseNet201, while DenseNet169 showed slightly inferior performance. The performance of different models is shown in Table 5.

4.4. Visual geometry group VGG/VGG16/19. VGG, which is the abbreviation for the visual geometry group, is a CNN structure characterized by several layers. In this sense, "deep" refers to the level of complexity shown by the models, which is determined by the number of layers they possess. VGG-16 and VGG-19 are composed of 16 and 19 convolutional layers, respectively. They serve as the foundation for revolutionary VGG-based object identification models. In addition to outperforming baselines on various tasks and datasets beyond ImageNet, the VGG Net is being developed as a deep neural network. In addition, it remains one of the most prevalent patterns for image recognition today (Pérez-Pérez *et al.*, 2021). The name "VGG16" refers to 16 layers that make up the deep neural network architecture having over 138 million parameters, demonstrating its massive size.

Its impressive design remains noteworthy even in contemporary contexts due to its scale. However, the simplicity of the VGGNet16 surface makes it appealing. Just by looking at it, all its architecture becomes apparent. After a series of convolution layers, a pooling layer reduces the image's height and width dimensions. We have roughly 64 options regarding the filter numbers we can use when conducting our research, which we can expand to around 128 and 256. The VGG16 architecture shows 512 filters in the final stages. The VGG19 model, often known as VGGNet-19, comprises 19 layers. The model's weight layers are denoted by the numbers 16 and 19, which correspond to the convolutional layers. VGG-19 contains three more convolutional layers than VGG-16.

The study used cutting-edge deep-learning models, including VGG, VGG16, and VGG19, to diagnose pneumonia cases from radiographs with promising results. The VGG model outperformed all other models in this family, with an accuracy of 90.22%, a high precision of 89.23%, a recall of 90.64%, and an F-score of 89.77%. The results of different models are shown in Table 5. These results highlight the potential of deep learning methods to support clinical decision making, especially when using medical imaging analysis to diagnose pneumonia (Benaissa *et al.*, 2022).

4.5. MobileNetV2 . Google developed MobileNet-V2 (Sandler *et al.*, 2018) as the second version of MobileNet-V1. In MobileNet-V2, an inverted residual structure is the backbone for feature extraction, making it a better module. As such, MobileNetV2 exhibits state-of-the-art performance in tasks involving item detection and semantic segmentation. Specifically engineered for utilization in mobile and embedded devices constrained by computational resources, MobileNetV2 represents a significant advancement in efficient model design. It relies on "factorized convolutions," which reduce the amount of parameters and computations in the network. This makes MobileNetV2 faster and more efficient than many other CNNs, while still performing well on image classification and object identification tasks (Adedoja *et al.*, 2022). The research identified pneumonia using variants of MobileNet and MobileNetV2. MobileNet achieved 90.22% accuracy, with precision, recall, and F-score at 89.23%, 90.64%, and 89.77%, respectively. However, MobileNetV2 showed slightly worse performance, with an accuracy of 84.46%, a precision of 83.93%, a recall of 86.11%, and an F-score of 84.10%, as depicted in Table 5. These findings suggest that MobileNet outperformed MobileNetV2 in terms of efficiency and overall classification metrics for pneumonia detection.

4.6. Ensemble. An ensemble model is an ML model that is made up of several smaller models. Ensemble modeling is based on the independent training of multiple smaller models, which are then combined to produce a final prediction. By decreasing overfitting and enhancing generalization, the model's overall performance is intended to be improved. Using various modeling algorithms or training data sets, ensemble modeling entails building a variety of algorithms for predicting outputs. One final prediction is generated for the unseen data once the ensemble model has combined the predictions of each base model. The ensemble model's goal is to lower forecast generalisation inaccuracy. When the ensemble approach is applied, the model's prediction error is reduced as long as the basis models are diverse and independent. A prediction is made using the wisdom of crowds. The ensemble model functions and behaves like a single model even though it comprises several foundation models. Ensemble modeling is widely used in practical data mining solutions (Sagi and Rokach, 2018). An ensemble model was successfully applied to a dataset comprising radiographic images from pediatric patients at the Guangzhou Medical Centre with 92.95% accuracy, 92.91% precision, 91.97% recall, and an F-score of 92.40%, as shown in Table 5. The model demonstrates robust performance in identifying patterns and characteristics within these medical images. Such advancements promise to enhance medical decision making and ultimately improve patient outcomes in healthcare settings.

Table 5. Performance of various pre-trained and CNN models.

Sr. no.	Models	Accuracy	Precision	Recall	F-score
1	CNN	87.82%	89.31%	84.87%	86.33%
2	Mobile Net	90.22%	89.23%	90.64%	89.77%
3	Mobile NetV2	84.46%	83.93%	86.11%	84.10%
4	VGG	90.22%	89.23%	90.64%	89.77%
5	VGG16	88.46%	87.42%	88.89%	87.95%
6	VGG19	87.34%	86.60%	86.28%	86.44%
7	Dense Net	85.74%	85.40%	83.80%	84.46%
8	Densenet121	87.34%	87.03%	85.68%	86.26%
9	Densenet169	80.13%	81.06%	82.99%	79.96%
10	Desnet201	85.58%	84.58%	84.70%	84.64%
11	Ensemble	92.95%	92.91%	91.97%	92.40%
12	NASNet Mobile	81.73%	80.47%	81.45%	80.85%
13	Efficient Net	90.71%	91.58%	88.63%	89.76%
14	**EfficientNet-B0**	**94.13**%	**93.50**%	**92.99**%	**93.14**%
15	Efficient NetB1	91.03%	90.05%	91.62%	90.63%
16	Efficient NetB3	87.66%	90.63%	83.97%	85.85%
17	Efficient NetB4	91.19%	90.95%	90.13%	90.51%
18	Efficient NetB5	91.71%	89.91%	90.43%	91.15%
19	Efficient NetB6	91.35%	90.58%	91.11%	90.83%
20	Efficient Net B7	87.34%	87.03%	85.68%	86.26%

4.7. NASNet-Mobile. NasNet Mobile is a neural network architecture created by the Google Brain team for image categorization. It is meant to be lightweight and efficient, making it ideal for mobile and embedded applications. NasNet Mobile achieves high accuracy on the ImageNet dataset with few parameters and a low computational cost. It is also amenable to running on mobile devices in real-time with a high degree of accuracy. The NasNet Mobile model was used to detect pneumonia cases within the dataset (Tan *et al.*, 2019). The resulting accuracy was 81.73%, indicating the model's great performance in classify pneumonia and normal cases in chest X-ray. Moreover, the NasNet Mobile model yielded 80.47% precision, 81.45% recall, and an F-score of 80.85%, as shown in Table 5 suggesting its effectiveness in accurately identifying pneumonia in pediatric patients.

4.8. EfficientNet models. EfficientNet is a CNN specifically created by Google Research for image categorization. It is designed to be more precise in terms of parameters, computing cost, and accuracy than prior designs like ResNet and Inception. To scale up CNN models, EfficientNet employs a compound scaling technique. The approach concurrently increases the model's resolution, depth, and width. The scaling approach employs a predetermined set of computational resources to achieve an optimal balance of accuracy and efficiency. EfficientNet produced superior accuracy on the ImageNet dataset while being more efficient than previous models with comparable accuracy. It offers many applications, including object detection, image segmentation, and video classification (Marques *et al.*, 2020). Notably, each variant demonstrated robust performance metrics, reflecting their efficacy in identifying pneumonia cases. Among these, EfficientNet-B0 showcased the highest accuracy (94.13%) and precision (93.50%), as shown in Table 5 underlining its capability in accurately discerning pneumonia instances from the radiographs. Additionally, other variants like EfficientNet-B1 to B6 also exhibited commendable accuracy, precision, recall, and F-score values, further emphasizing the versatility and effectiveness of the EfficientNet architecture in medical image analysis tasks.

5. Results and discussion

The results presented in Table 5 comprehensively compare the training and testing accuracies achieved by various CNNs employed in classification problems. Among the twenty different models for classification evaluated, EfficientNet-B0 consistently outperforms the other models, demonstrating superior performance in both the training and testing phases. Notably, EfficientNet-B0 achieves an impressive accuracy of 94.13%, with precision, recall, and F-score metrics further reinforcing its effectiveness, with scores of 93.50%, 92.99%, and 93.14%, respectively. These results underscore the efficacy of EfficientNet-B0 in accurately classifying input data, highlighting its potential as a robust and reliable model for pneumonia detection and classification tasks.

In applications such as medical image interpretation, pretrained models have shown excellent performance. CNNs can be used to learn from large data sets such efficiently. Researchers can improve the generalization efficacy of models for certain tasks like pneumonia diagnosis by fine-tuning them on smaller, domain-specific datasets. As shown in Table 5, EfficientNet-Bo attains superior accuracy, precision, and recall compared to models trained from scratch. Their success underscores the importance of transfer learning in medical image analysis tasks, where access to large-scale annotated datasets may be limited. The relationship between model complexity and performance is critical in machine learning, particularly in medical image analysis tasks like pneumonia classification. Table 3 compares the proposed approach and existing works.

Although increased model complexity often leads to improved performance, this comes at the cost of computational resources and interpretability. In our study, we explored this trade-off by comparing the performance of DenseNet and EfficientNet, two pretrained models known for their different levels of complexity. DenseNet, with its densely connected layers, offers a highly expressive model architecture capable of capturing intricate patterns in the data. On the other hand, EfficientNet uses an innovative compound scaling strategy to maximize model depth, breadth, and resolution, yielding outstanding results with fewer parameters. Our findings reveal that EfficientNet achieves comparable results with a significantly lower computational overhead. This highlights the importance of balancing model complexity and performance, considering factors like computational cost, interpretability, and scalability in real-world applications.

While our work shows promising results in pneumonia identification utilizing deep learning models and data augmentation strategies, certain limitations should be considered. Firstly, the generalizability of our findings may be constrained by the relatively small size and imbalanced nature of the dataset, which primarily comprises chest X-ray images from a single institution. Increasing the dataset's diversity in demographics and imaging methods may improve the model's resilience and suitability in various healthcare environments. Moreover, although data augmentation helps mitigate the limitations of limited data availability, it may introduce biases or distortions that affect the model's performance, highlighting the need for careful validation and monitoring of augmented datasets. The interpretability of deep learning models remains challenging, as they operate as complex black-box systems, making it challenging to comprehend the fundamental assumptions underlying their projections.

To improve model interpretability and promote provider confidence, future research should concentrate on creating explainable AI strategies and fostering trust among healthcare providers. Furthermore, deploying AI-based diagnostic tools in clinical practice requires rigorous evaluation of their real-world performance, including prospective validation studies and assessing their impact on clinical outcomes and workflow efficiency. Addressing these constraints and embracing continuous advances in AI and medical imaging technologies will open the path for the further development and deployment of creative solutions to improve patient care and pneumonia diagnosis in the future.

6. Comparative analysis of classification results

An assessment was conducted to compare the classification outcomes and determine the accuracy as shown in Table 1 of the suggested strategy in detecting and categorizing pneumonia. The findings are summarized in Table 5. The table shows that the suggested approach outperforms numerous current approaches in terms of performance. Figure 4 shows the line plots for training and validation accuracy and losses for various models. Notably, we intend to enhance all models' efficiency in future research by meticulously fine-tuning hyperparameters and variables. Table 3 compares the proposed approach and existing works. Moreover, this fine-tuning approach could be extended to facilitate early identification of pneumonia and COVID-19, thereby aiding physicians in more accurate disease diagnoses.

7. Conclusions

The automation of pneumonia detection using chest X-rays and CT scans has made remarkable progress in recent years, particularly with the advent of deep-learning algorithms. Base deep learning model architectures have changed significantly during the last four years. COVID-19 has emerged as the most crucial worldwide problem for saving human life. Several healthcare institutions are working hard to find appropriate solutions. However, artificial intelligence applications in computer-assisted diagnosis (CAD) has demonstrated its value and efficacy in resolving a range of medical problems. Because of several forms of pneumonia, such as viral, bacterial, tubercular, and COVID-19, a system for multi-class classification was required since present methods provide less reliable solutions.

In this study, we trained twenty alternative CNN models on the pneumonia dataset. The Efficient Net-B0 model surpassed all other models with an accuracy of 94.13%, a precision of 93.50%, a recall of 92.99%,

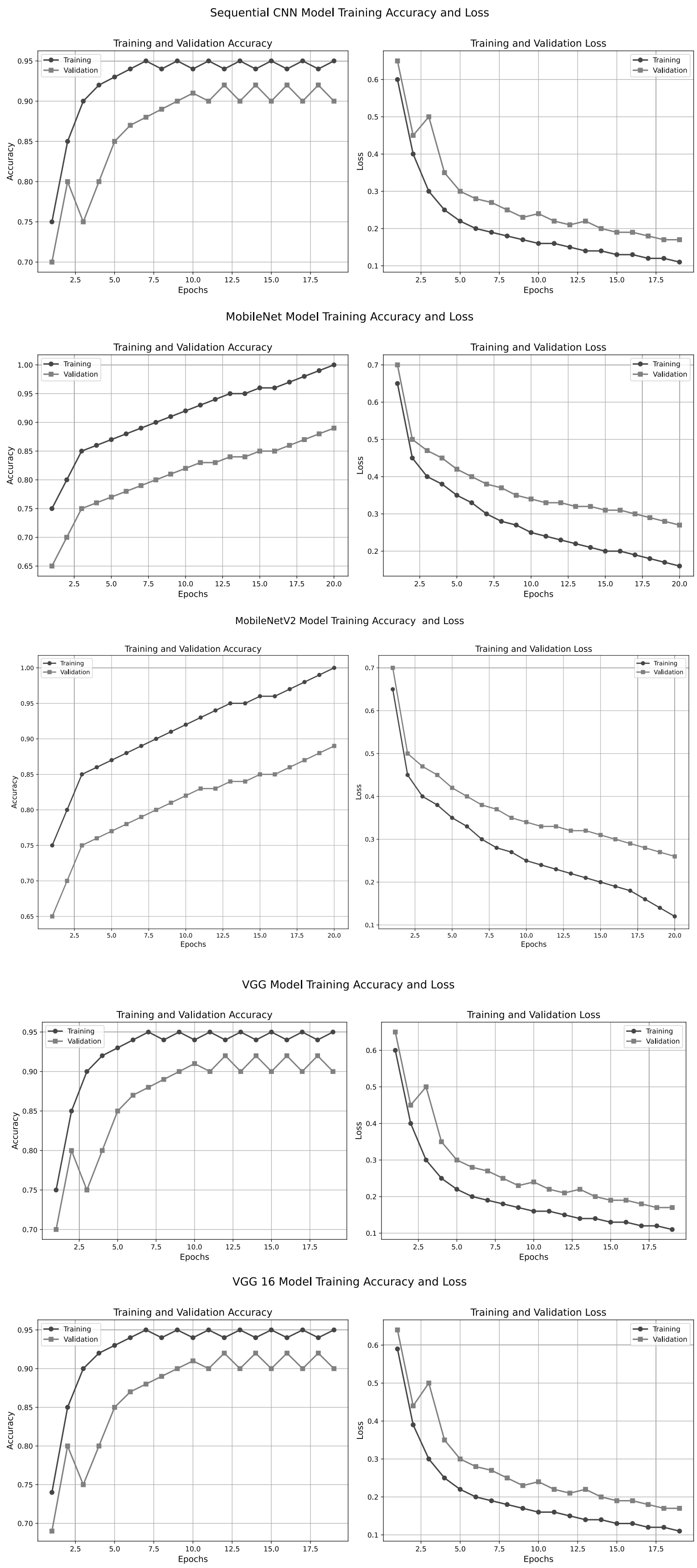

Fig. 4. Training and validation results (Part 1).

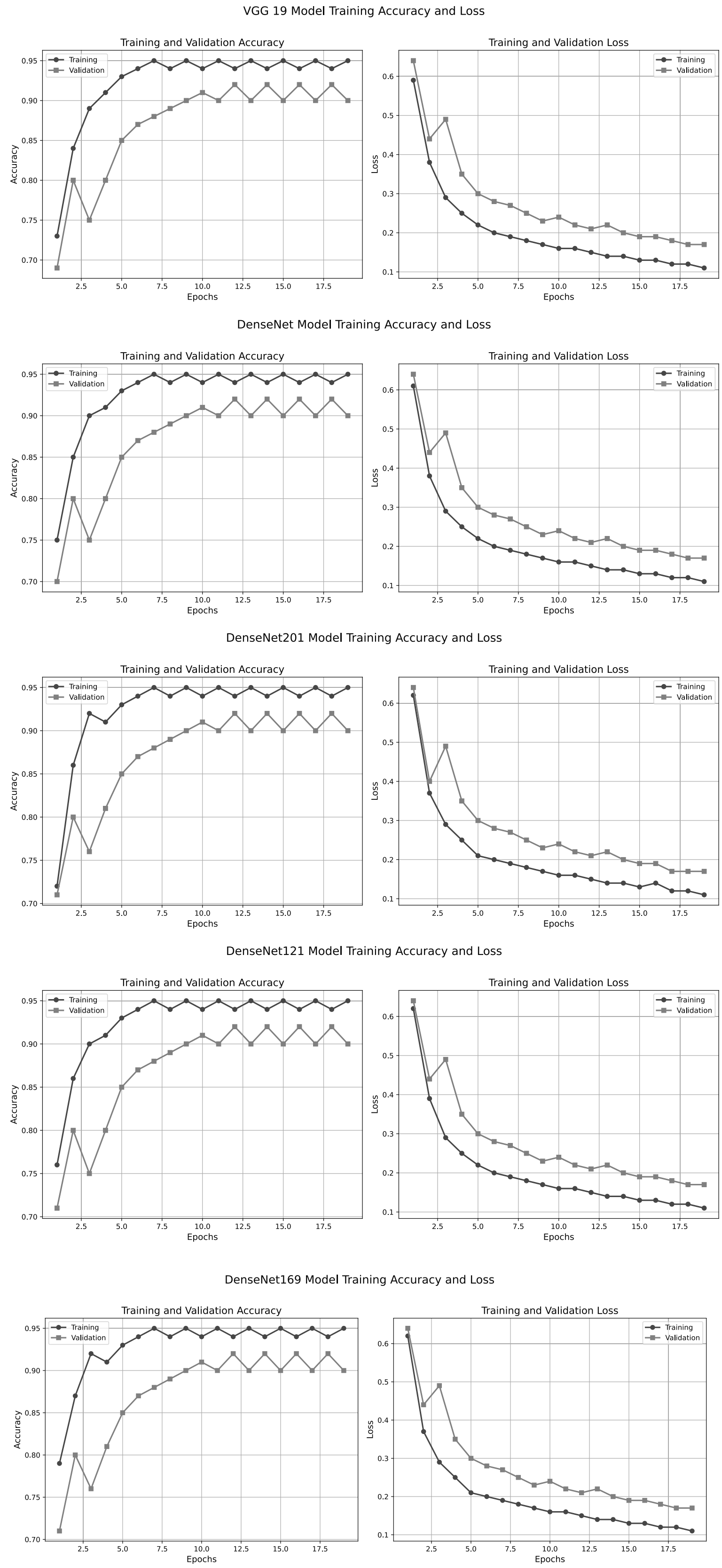

Fig. 5. Training and validation results (Part 2).

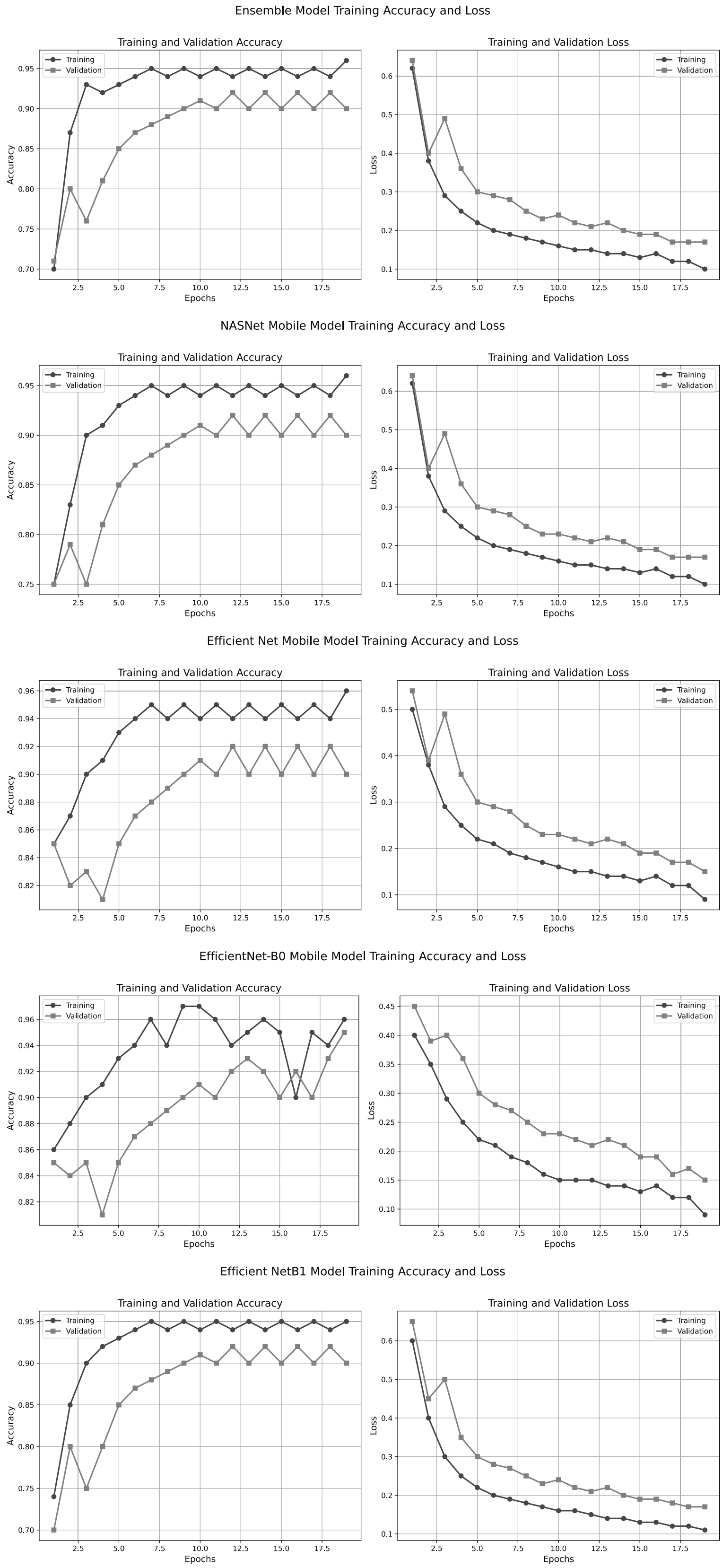

Fig. 6. Training and validation results (Part 3).

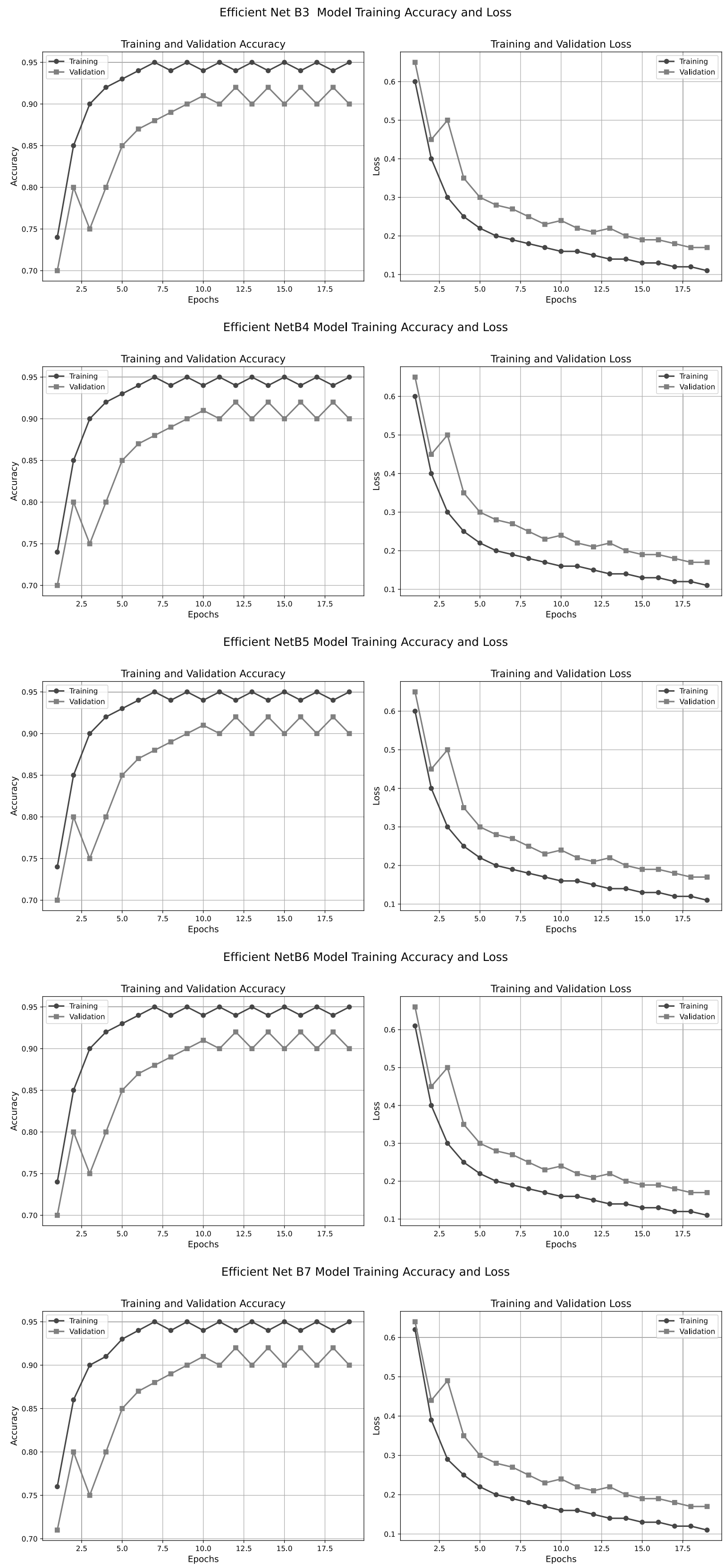

Fig. 7. Training and validation results (Part 4).

and an F-score of 93.14%. These results highlight the potential of advanced neural network models to revolutionize pneumonia diagnosis through enhanced accuracy and efficiency. Given the significant burden of pneumonia worldwide, particularly among vulnerable populations, this research offers a compelling solution to improve early detection and timely intervention. Furthermore, the findings illustrated that the outcomes of the proposed study surpassed those of other architectures in accuracy, indicating superior overall performance compared to existing research. The efficacy could be enhanced by increasing the dataset size and incorporating additional pretrained architectures. Consequently, deep learning approaches exhibited notably superior treatment quality and accuracy results compared with conventional methods.

8. Limitations and future work

Our research does not identify the specific location of pneumonia in chest X-rays or provide any information about the duration of the disease in the patient's chest. Increasing datasets for pneumonia diagnosis using CNNs will improve model robustness. Optimizing network architectures and hyperparameters can enhance accuracy and efficiency. Multi-class classification abilities will enable variation of pneumonia types and harshness stages. Furthermore, upcoming exploration aims to identify the specific location of pneumonia in chest X-rays and estimate the period of the disease in patients. We recommend further exploring and validating the EfficientNet-B0 model on larger, more diverse datasets to assess its generalizability and robustness. Additionally, investigating the model's interpretability and exploring ways to integrate it into clinical workflows would be valuable next steps. In short, this work presents a promising step forward in leveraging the power of deep learning to combat the global challenge of pneumonia.

Acknowledgment

This work was supported in part by the Princess Nourah bint Abdulrahman University Researchers Supporting Project PNURSP2024R 343. The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA, for funding this research work through the project NBU-FFR-2024-1092-10. This work was also supported by the Researchers Supporting Project RSP2024R395, King Saud University, Riyadh, Saudi Arabia.

References

Adedoja, A.O., Owolawi, P.A., Mapayi, T. and Tu, C. (2022). Intelligent mobile plant disease diagnostic system using NASNet-Mobile deep learning, *IAENG International Journal of Computer Science* **49**(1): 216–231.

Akbar, W., Inam Ul Haq, M., Soomro, A., Muhammad Daudpota, S., Shariq Imran, A. and Ullah, M. (2023a). Automated report generation: A GRU based method for chest X-rays, *2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan*, pp. 1–6.

Akbar, W., Soomro, A., Ghanghro, S.A., Ul Haq, M.I. and Ullah, M. (2023b). Performance evaluation of deep learning models for breast cancer classification, *2023 IEEE International Conference on Emerging Trends in Engineering, Sciences and Technology (ICES&T), Bahawalpur, Pakistan*, pp. 1–4.

Akbar, W., Soomro, A., Ullah, M., Inam Ul Haq, M., Ullah Khan, S. and Ali Shah, T. (2023c). Performance evaluation of deep learning models for leaf disease detection: A comparative study, *2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan*, pp. 01–05.

Albawi, S., Mohammed, T.A. and Al-Zawi, S. (2017). Understanding of a convolutional neural network, *2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey*, pp. 1–6, DOI: 10.1109/icengtechnol.2017.8308186.

Alqudah, A.M., Qazan, S. and Masad, I.S. (2021). Artificial intelligence framework for efficient detection and classification of pneumonia using chest radiography images, *Journal of Medical and Biological Engineering* **41**(5): 599–609.

Alquran, H., Alsleti, M., Alsharif, R., Qasmieh, I.A., Alqudah, A.M. and Harun, N.H.B. (2021). Employing texture features of chest X-ray images and machine learning in COVID-19 detection and classification, *Mendel* **27**(1): 9–17, DOI: 10.13164/mendel.2021.1.009.

Alsharif, R., Al-Issa, Y., Alqudah, A.M., Qasmieh, I.A., Mustafa, W.A. and Alquran, H. (2021). Pneumonianet: Automated detection and classification of pediatric pneumonia using chest X-ray images and CNN approach, *Electronics* **10**(23): 2949.

Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M. and Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions, *Journal of Big Data* **8**(53): 1–74.

Ayan, E. and Ünver, H.M. (2019). Diagnosis of pneumonia from chest X-ray images using deep learning, *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), Istanbul, Turkey*, pp. 1–5.

Basha, S.S., Dubey, S.R., Pulabaigari, V. and Mukherjee, S. (2020). Impact of fully connected layers on performance of convolutional neural networks for image classification, *Neurocomputing* **378**: 112–119.

Benaissa, A., Bahri, A., El Allaoui, A. and Bourass, Y. (2022). Character recognition using pre-trained models and performance variants based on datasets size: A survey, *ITM Web of Conferences* **43**: 01008.

Chen, C., Han, D. and Chang, C.-C. (2024). MPCCT: Multimodal vision-language learning paradigm with context-based compact transformer, *Pattern Recognition* **147**: 110084.

Cheng, H., Lu, T., Hao, R., Li, J. and Ai, Q. (2024). Incentive-based demand response optimization method based on federated learning with a focus on user privacy protection, *Applied Energy* **358**: 122570.

Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L. and Haworth, A. (2021). A review of medical image data augmentation techniques for deep learning applications, *Journal of Medical Imaging and Radiation Oncology* **65**(5): 545–563.

da Silva, L. O., dos Santos Araújo, L., Souza, V. F., de Barros Neto, R. M. and Santos, A. (2020). Comparative analysis of convolutional neural networks applied in the detection of pneumonia through X-ray images of children, *Learning & Nonlinear Models* **18**(2): 4–15.

Eusebi, P. (2013). Diagnostic accuracy measures, *Cerebrovascular Diseases* **36**(4): 267–272.

Foulds, L. R. (2012). *Optimization Techniques: An Introduction*, Springer, New York.

Frederich, J., Himawan, J. and Rizkinia, M. (2024). Skin lesion classification using EfficientNet B0 and B1 via transfer learning for computer aided diagnosis, *AIP Conference Proceedings* **3080**(1), Article no. 110002.

Haloi, M., Rajalakshmi, K.R. and Walia, P. (2018). Towards radiologist-level accurate deep learning system for pulmonary screening, *arXiv:* 1807.03120.

Hara, K., Saito, D. and Shouno, H. (2015). Analysis of function of rectified linear unit used in deep learning, *2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland*, pp. 1–8.

Heravi, E.J., Aghdam, H.H. and Puig, D. (2016). Classification of foods using spatial pyramid convolutional neural network, *in* À. Nebot *et al.* (Eds), *Artificial Intelligence Research and Development*, IoS Press, Amsterdam, pp. 163–168.

Hussain, A., Akbar, W., Hussain, T., Bashir, A.K., Al Dabel, M.M., Ali, F. and Yang, B. (2024). Ensuring zero trust IoT data privacy: Differential privacy in blockchain using federated learning, *IEEE Transactions on Consumer Electronics* **PP**(99): 1–1.

Iqbal, T. and Ali, H. (2018). Generative adversarial network for medical images (MI-GAN), *Journal of Medical Systems* **42**, Article no. 231, DOI: 10.1007/s10916-018-1072-9.

Jain, R., Nagrath, P., Kataria, G., Kaushik, V.S. and Hemanth, D.J. (2020). Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning, *Measurement* **165**: 108046.

Jeczmionek, E. and Kowalski, P.A. (2023). Choice of the-norm for high level classification features pruning in modern convolutional neural networks with local sensitivity analysis, *International Journal of Applied Mathematics and Computer Science* **33**(4): 663–672, DOI: 10.34768/amcs-2023-0047.

Kang, L., Tang, B., Huang, J. and Li, J. (2024). 3D-MRI super-resolution reconstruction using multi-modality based on multi-resolution CNN, *Computer Methods and Programs in Biomedicine* **248**: 108110.

Ke, Q., Zhang, J., Wei, W., Połap, D., Woźniak, M., Kośmider, L. and Damaševǐ (2019). A neuro-heuristic approach for recognition of lung diseases from X-ray images, *Expert Systems with Applications* **126**: 218–232.

Kermany, D., Zhang, K. and Goldbaum, M. (2018). Labeled optical coherence tomography (OCT) and chest X-ray images for classification, *Mendeley Data* **2**(2): 651.

Lau, M.M. and Lim, K.H. (2018). Review of adaptive activation function in deep neural network, *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Sarawak, Malaysia*, pp. 686–690.

Liang, G. and Zheng, L. (2020). A transfer learning method with deep residual network for pediatric pneumonia diagnosis, *Computer Methods and Programs in Biomedicine* **187**: 104964.

Liao, L., Guo, Z., Gao, Q., Wang, Y., Yu, F., Zhao, Q., Maybank, S.J., Liu, Z., Li, C. and Li, L. (2023). Color image recovery using generalized matrix completion over higher-order finite dimensional algebra, *Axioms* **12**(10): 954.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B. and Sánchez, C.I. (2017). A survey on deep learning in medical image analysis, *Medical Image Analysis* **42**: 60–88.

Manickam, A., Jiang, J., Zhou, Y., Sagar, A., Soundrapandiyan, R. and Samuel, R.D.J. (2021). Automated pneumonia detection on chest X-ray images: A deep learning approach with different optimizers and transfer learning architectures, *Measurement* **184**: 109953.

Marques, G., Agarwal, D. and de la Torre Díez, I. (2020). Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network, *Applied Soft Computing* **96**: 106691.

Pérez-Pérez, B.D., Garcia Vazquez, J.P. and Salomón-Torres, R. (2021). Evaluation of convolutional neural networks' hyperparameters with transfer learning to determine sorting of ripe Medjool dates, *Agriculture* **11**(2): 115.

Prechelt, L. (2002). Early stopping—But when?, *in* G.B. Orr and K.R. Müller (Eds), *Neural Networks: Tricks of the Trade*, Springer, Berlin/Heidelberg, pp. 55–69, DOI: 10.10007/3-540-49430-8_3.

Račić, L., Popović, T., Čakić, S. and Šandi, S. (2021). Pneumonia detection using deep learning based on convolutional neural network, *2021 25th International Conference on Information Technology (IT), Zabljak, Montenegro*, pp. 1–4.

Rahimzadeh, M. and Attar, A. (2020). A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of xception and ResNet50v2, *Informatics in Medicine Unlocked* **19**: 100360.

Rahman, T., Chowdhury, M.E., Khandakar, A., Islam, K.R., Islam, K.F., Mahbub, Z.B., Kadir, M.A. and Kashem, S. (2020). Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray, *Applied Sciences* **10**(9): 3233.

Rahmat, T., Ismail, A. and Aliman, S. (2019). Chest X-ray image classification using faster R-CNN, *Malaysian Journal of Computing* **4**(1): 225–236.

Rajaraman, S., Candemir, S., Kim, I., Thoma, G. and Antani, S. (2018). Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs, *Applied Sciences* **8**(10): 1715.

Rajasenbagam, T., Jeyanthi, S. and Pandian, J.A. (2021). Detection of pneumonia infection in lungs from chest X-ray images using deep convolutional neural network and content-based image retrieval techniques, *Journal of Ambient Intelligence and Humanized Computing* **2021**, DOI: 10.10007/s12652-021-03075-2.

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M.P. and Ng, A.Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning, *arXiv:* 1711.05225.

Ramos, D., Franco-Pedroso, J., Lozano-Diez, A. and Gonzalez-Rodriguez, J. (2018). Deconstructing cross-entropy for probabilistic binary classifiers, *Entropy* **20**(3): 208.

Rao, D.S., Nair, S.A.H., Rao, T.V.N. and Kumar, K.S. (2022). Classification of pneumonia from chest x-ray image using machine learning models, *International Journal of Intelligent Systems and Applications in Engineering* **10**(1s): 399–408.

Regunath, H. and Oba, Y. (2021). *Community-Acquired Pneumonia*, StatPearls Publishing, Treasure Island, https://europepmc.org/article/nbk/nbk430749.

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016). You only look once: Unified, real-time object detection, *IEEE Conference on Compute Vision and Pattern Recognition, Las Vegas, USA*, pp. 779–788.

Ruby, U. and Yendapalli, V. (2020). Binary cross entropy with deep learning technique for image classification, *International Journal of Advanced Trends in Computer Science and Engineering* **9**(4), DOI: 10.30534/ijatcse/2020/175942020.

Sagi, O. and Rokach, L. (2018). Ensemble learning: A survey, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **8**(4): e1249.

Sahlol, A.T., Abd Elaziz, M., Tariq Jamal, A., Damaševičius, R. and Farouk Hassan, O. (2020). A novel method for detection of tuberculosis in chest radiographs using artificial ecosystem-based optimisation of deep neural network features, *Symmetry* **12**(7): 1146.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA*, pp. 4510–4520.

Shah, A. and Shah, M. (2022). Advancement of deep learning in pneumonia/COVID-19 classification and localization: A systematic review with qualitative and quantitative analysis, *Chronic Diseases and Translational Medicine* **8**(03): 154–171.

Sharma, S., Sharma, S. and Athaiya, A. (2017). Activation functions in neural networks, *Towards Data Science* **6**(12): 310–316.

Shoaib, M., Sayed, N., Shah, B., Hussain, T., AlZubi, A.A., AlZubi, S.A. and Ali, F. (2023). Exploring transfer learning in chest radiographic images within the interplay between COVID-19 and diabetes, *Frontiers in Public Health* **11**: 1297909, DOI: 10.3389/fpubh.2023.1297909.

Singh, S. and Tripathi, B. (2022). Pneumonia classification using quaternion deep learning, *Multimedia Tools and Applications* **81**(2): 1743–1764.

Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., Yixuan, Y., Kuleev, R. and Ibragimov, B. (2019). Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database, *Computers & Electrical Engineering* **78**: 388–399.

Stephen, O., Sain, M., Maduh, U.J., Jeong, D.-U. (2019). An efficient deep learning approach to pneumonia classification in healthcare, *Journal of Healthcare Engineering* **2019**, Article ID: 4180949, DOI: 10.1155/2019/4180949.

Sun, M., Song, Z., Jiang, X., Pan, J. and Pang, Y. (2017). Learning pooling for convolutional neural network, *Neurocomputing* **224**: 96–104.

Tan, M., Chen, B., Pang, R., Vasudevan, V., Sandler, M., Howard, A. and Le, Q.V. (2019). MnasNet: Platform-aware neural architecture search for mobile, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA*, pp. 2820–2828.

Tian, Y. and Zhang, Y. (2022). A comprehensive survey on regularization strategies in machine learning, *Information Fusion* **80**: 146–166.

Townsend, J.T. (1971). Theoretical analysis of an alphabetic confusion matrix, *Perception & Psychophysics* **9**: 40–50, DOI: 10.3758/BF03213026.

Varshni, D., Thakral, K., Agarwal, L., Nijhawan, R. and Mittal, A. (2019). Pneumonia detection using CNN based feature extraction, *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India*, pp. 1–7.

Venter, G. (2010). Review of optimization techniques, *in* R. Blockley and W. Shyy (Eds), *Encyclopedia of Aerospace Engineering*, DOI: 10.1002/9780470686652.eae495.

Yacouby, R. and Axman, D. (2020). Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models, *Proceedings of the 1st Workshop on Evaluation and Comparison of NLP Systems*, pp. 79–91, (online).

Yang, C.-P., Zhu, J.-Q., Yan, T., Su, Q.-L. and Zheng, L.-X. (2022). Auxiliary pneumonia classification algorithm based on pruning compression, *Computational and Mathematical Methods in Medicine* **2022**, Article ID: 8415187, DOI: 101155/2022/8415187.

Zhou, T., Ye, X., Lu, H., Zheng, X., Qiu, S. and Liu, Y. (2022). Dense convolutional network and its application in medical image analysis, *BioMed Research International* **2022**, Article ID: 2384830, DOI: 10.1155/2022/2384830.

Wajahat Akbar is a PhD student in the School of Electronic and Control Engineering at Chang'an University, Xi'an, China. He received his BS in computer science from Khushal Khan Khattak University Karak in 2019. He obtained his MS in computer science from the same university, specializing in artificial intelligence, in 2023. His research interests include artificial intelligence, deep learning, natural language processing, computer vision, computer networks, and network security, focusing on healthcare applications.

Abdullah Soomro is currently working as a lecturer in the Department of Computer Science, Faculty of Computing, The Islamia University of Bahawalpur, Pakistan. He received his MS in computer science with specialization in data knowledge engineering from Sukkur IBA University, Sukkur, Pakistan, in 2019. From 2017 to 2019, he worked as a teaching assistant in the Computer Science Department at Sukkur IBA University. His research interests include machine learning, deep learning, natural language processing, and computer vision, focusing on healthcare applications.

Altaf Hussain received his BS in computer science from the University of Agriculture, Peshawar (Pakistan) in 2016 and his MS in computer science from Khushal Khan Khattak University, Karak (Pakistan), in 2023. He specializes in data security through blockchain, differential privacy, and federated learning. His research interests are data security, network security, artificial intelligence, and computer networks.

Tariq Hussain holds BS as well as MS degrees in information technology from the University of Malakand, Pakistan (2015) and the Institute of Computer Sciences and Information Technology at the University of Agriculture Peshawar, Pakistan (2019). He has published numerous research papers in the area of computer networks. He is currently a doctoral candidate at the School of Computer Science and Technology and the School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou, China. His research interests are the Internet of things, big data, data analytics, 3D point clouds, and artificial intelligence.

Farman Ali is an assistant professor in the Department of Computer Science and Engineering, Sungkyunkwan University, South Korea. He received his BS degree in computer science from the University of Peshawar, Pakistan, in 2011, his MS degree in computer science from Gyeongsang National University, South Korea, in 2015, and his PhD degree in information and communication engineering from Inha University, South Korea, in 2018. He has registered over four patents and published more than 100 research articles in reputable international conference proceedings and journals. His research interests include sentiment analysis, data science, text mining, recommendation systems, healthcare monitoring systems, ontology, fuzzy logic, type-2 fuzzy logic, artificial intelligence, and deep learning.

Muhammad Inam Ul Haq holds an MS from the Institute of Management Sciences, University of Peshawar, Pakistan, and a PhD from Jean Monnet University, Saint-Etienne, France. He works as an assistant professor in the Department of Computer Science and Bioinformatics at Khushal Khan Khattak University, Karak, Pakistan. His research interests include computer vision, image processing, networks, optonumeric security, deep learning, and NLP.

Razaz Waheeb Attar is an associate professor in the Department of Management, Princess Nourah bint Abdul Rahman University, Riyadh, Saudi Arabia. She holds a PhD in business administration from Dublin City University Ireland. She also holds MS degrees in electronic commerce from Dalhousie University in Nova Scotia, Canada, and in business research from Kingston University in Kingston upon Thames, UK. Her research interests include social innovation, higher education, strategic management, contemporary issues in marketing and management, digital strategy, social commerce, digital ethics, sustainability, AI risks, and e-commerce.

Ahmed Alhomoud is an assistant professor in the Department of Computer Science at Northern Border University, Kingdom of Saudi Arabia. He holds a PhD in computer science from the University of Southampton, UK. His research interests include, but are not limited to, digital forensics, cyber security, the Internet of things and blockchains.

Ahmad Ali AlZubi is a professor at King Saud University. He obtained his PhD in computer networks engineering from the National Technical University of Ukraine in 1999. His current research interests include computer networks, grid computing, cloud computing, big data and data extracting, and cybersecurity.

Reem Alsagri is a lecturer at the University of Hafr Al Batin and a PhD candidate at the University of Southampton. She holds an MSc in computer science from the University of Southampton and a Bachelor's degree in information technology from the University of Solent. Alsagri's research interests include the IoT, biotechnology, wearable devices, artificial intelligence, and signal processing.

Received: 19 February 2024
Revised: 30 April 2024
Re-revised: 12 May 2024
Accepted: 23 August 2024