AITA
ARTIFICIAL INTELLIGENCE
THEORY AND APPLICATIONS

RESEARCH ARTICLE

Deep Learning Models for the Detection and Classification of COVID-19 and Associated Lung Diseases Using X-Ray Images

Osman Dikmen [a †]

[a] Department of Electric-Electronics, Düzce University, Düzce, Türkiye
[†] osmandikmen@duzce.edu.tr

RECEIVED SEPTEMBER 17, 2024
ACCEPTED SEPTEMBER 24, 2024

CITATION Dikmen. O. (2024). Deep learning models for the detection and classification of COVID-19 and associated lung diseases using X-ray images. *Artificial Intelligence Theory and Applications, 4(2)*, 121-142.

Abstract

The COVID-19 pandemic has introduced exceptional challenges to healthcare systems worldwide, underscoring the urgent need for swift and precise diagnostic solutions. In this research, we investigate the performance of various deep learning models, including VGG19, ResNet18, and a ResNet18-based U-Net, as well as a custom Convolutional Neural Network (CNN) developed in MATLAB, for the classification and segmentation of lung X-ray images. The dataset includes X-ray images from individuals diagnosed with COVID-19, viral pneumonia, lung opacity, and healthy individuals. The dataset was divided into 80% for training and 20% for testing, with data augmentation techniques implemented to enhance the model's effectiveness. The VGG19 model, utilizing transfer learning, demonstrated strong diagnostic capabilities, achieving high accuracy rates for COVID-19, lung opacity, healthy lungs, and viral pneumonia classification, with a test accuracy of 97.5%. ResNet18 was employed for both classification and as part of a hybrid model incorporating a U-Net-inspired decoder for lung disease segmentation. The ResNet18 model achieved competitive accuracy and loss metrics, while the ResNet18-based U-Net model excelled in image segmentation tasks, demonstrating its potential in biomedical image analysis. Additionally, a customized CNN model was developed using MATLAB for the classification of the four lung conditions. This model produced visual outputs, including training-validation loss/accuracy graphs and confusion matrices. Our results indicate that deep learning models, especially when combined with transfer learning and customized architectures, offer a powerful approach to diagnosing COVID-19 and related lung conditions. Future work will focus on refining these models with larger datasets and further experimentation to enhance diagnostic performance across diverse clinical settings.

Keywords: COVID-19, deep learning, medical image classification, x-ray imaging, convolutional neural networks

1. Introduction

The COVID-19 pandemic emerged as a result of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) virus, which triggered a profound global health emergency. Initially identified in the city of Wuhan in late 2019, the disease swiftly

Artificial Intelligence Theory and Applications, ISSN: 2757-9778. ISBN: 978-605-69730-2-4 © 2024 İzmir Bakırçay University

evolved into a global health crisis. COVID-19 presents a diverse array of clinical symptoms, ranging from mild to severe respiratory distress. Swift and precise diagnosis for this illness is essential for controlling its spread and determining appropriate treatment strategies [1].

The virus rapidly spread from Wuhan, reaching numerous countries around the globe. Its impact was notably felt in major regions, including North America, parts of South Asia, South America, Western Europe, and Eastern Europe In March 2020, the World Health Organization (WHO) in a formal manner announced COVID-19 as a pandemic, designating it as a worldwide health emergency [1]. During this time, billions of people were required to remain indoors, and many countries enforced lockdowns. By May 19, 2022, there had been approximately 525,080,438 confirmed cases of COVID-19 across more than 219 countries, with 484,920,117 recoveries and 6,294,856 fatalities reported [1].

Diagnosing COVID-19 usually involves several methods: Serology (Antibody) for detecting antibodies, Genetic Real-Time Reverse Transcription Polymerase Chain Reaction (RT-PCR) for genetic material analysis, and Antigen testing [2]. RT-PCR is widely regarded as the gold standard for detecting the coronavirus, delivering reliable results, especially during the initial phases of the infection [3], [4]. Nonetheless, this method has several drawbacks, including issues with false sampling, accessibility, specificity, cost, and extended turnaround times [5]. Additionally, many countries struggle with providing sufficient RT-PCR test kits.

The antibody test detects the presence of IgG and/or IgM antibodies through samples obtained from blood, serum, or plasma [6]. This test is designed to reveal whether an individual has been previously infected and how their immune system responded, but it cannot confirm a current infection. Antibodies typically begin to form between one to three weeks following the onset of symptoms [7]. This method is also costly and requires significant time.

The antigen test attempts to detect the presence of coronavirus infection by collecting samples from nasal swabs. It is less time-consuming and relatively inexpensive compared to other tests [8]. Given these limitations, there is a growing need to explore different diagnostic methods which are both accurate and accessible.

Medical imaging, including X-rays and Computed Tomography (CT), has proven essential in the rapid detection of lung abnormalities associated with COVID-19. X-rays, in particular, are widely accessible and commonly used in clinical settings. Recent advancements in deep learning have further enhanced the utility of medical imaging for disease diagnosis. Convolutional Neural Networks (CNNs), an essential element of deep learning, have shown remarkable accuracy in analyzing medical images. However, it may miss active infections when compared to RT-PCR tests.

Although RT-PCR testing is crucial among traditional diagnostic approaches, the demand for quick and precise alternative methods is growing due to the lengthy duration required for these tests, the need for laboratory facilities, and occasional false-negative results. In this regard, medical imaging techniques become essential. Imaging techniques, including X-rays, are valued for their swift accessibility and extensive

application in clinical practice. Similarly, Computed Tomography (CT) imaging is preferred for its expedited and comprehensive diagnostic capabilities.

In this study, we explore several deep learning models for the classification and segmentation of lung diseases, including COVID-19. The VGG19 architecture, renowned for its deep and wide network, is employed through transfer learning to classify X-ray images of COVID-19, lung opacity, viral pneumonia, and healthy lungs. Additionally, a ResNet18-based encoder is combined with a U-Net-inspired decoder to leverage both effective feature extraction and high-resolution segmentation for lung disease analysis. Furthermore, a customized Convolutional Neural Network (CNN) model was developed in MATLAB to classify the four lung conditions, offering a tailored approach to disease detection

Through these models, we aim to address the growing need for quick, precise, and accessible diagnostic tools for COVID-19 and related lung diseases. Our work demonstrates the capability of deep learning techniques in revolutionizing medical diagnostics, offering both high accuracy and clinical applicability in a time of global health crisis.

2. Related Studies

Various deep learning architectures have been introduced to diagnose COVID-19. In Nayak et al. [9], the researchers presented a compact CNN approach called LW-CORONet. This method incorporates convolutional layers, pooling layers, two fully connected (FC) layers, and a rectified linear unit (ReLU) activation function. With its five learnable layers, this architecture aids in extracting crucial features from CXR images.

Gupta and Bajaj [10] created a robust model for automatically detecting COVID-19 by utilizing deep learning (DL) techniques and chest CT scans. They incorporated two pre-trained DL models, DarkNet19 and MobileNetV2, alongside a lightweight DL approach, using publicly available CT scan visual data for concerning automated recognition of COVID-19.

A novel technique has been proposed to advance the classification and screening of COVID-19 patients through chest X-ray (CXR) imaging. This approach integrates cutting-edge deep learning models with refined image analysis methods, aiming to significantly boost diagnostic precision and speed. The method described combines standard data augmentation strategies with generative adversarial networks (GANs) to address data limitations. Additionally, it incorporates various filter banks, including Gabor filters, Sobel filters, and the Laplacian of Gaussian (LoG), to achieve more comprehensive feature extraction [11].

Basu et al. [12] showcased a deep learning approach designed to ascertain the presence of COVID-19 through the assessment of chest X-ray images. This solution utilizes a model pre-trained on a small chest X-ray dataset and domain extension transfer learning. Specifically focusing on identifying cases of COVID-19, this method identifies the specific segments analyzed in classification using Gradient Class Activation Map, thus ensuring transparency in the detection process. The authors reported an overall accuracy rate of

90.1%. However, the proposed system can detect COVID-19 based on a restricted dataset.

Conversely, the ResNet-101 CNN model utilized by Azemin et al. [13] is notable within deep learning strategies. This approach utilized a vast number of images in the pre-training phase to extract crucial features, subsequently re-training the model to identify anomalies in chest X-ray images. Despite these efforts, the method achieved a reported accuracy rate of only 71.9%.

In another study, the MobileNetV2 deep learning model and k-nearest neighbor (k-NN) algorithm were used to detect brain tumors from MRI images, achieving an accuracy rate of 96.44% [14].

For lung cancer detection using CT images, a Convolutional Neural Network (CNN) model was outlined, demonstrating superior performance, achieving a high accuracy rate and fewer layers compared to existing deep learning models [15].

To tackle two distinct multi-class classification challenges, the Xception transfer learning method was employed. The first challenge involved distinguishing among control cases, COVID-19, and various forms of pneumonia, including viral and bacterial types. The second challenge focused on differentiating between control cases, COVID-19, and pneumonia overall. An undersampling technique was implemented to mitigate dataset imbalance by randomly removing samples from the more numerous classes. The dataset comprised 290 chest X-ray images for COVID-19, 310 for control cases, 330 for bacterial pneumonia, and 327 for viral pneumonia. The study reported an accuracy of 89% for classifying all four conditions and 94% for distinguishing between the three categories. For the dataset focusing on three categories, the accuracy was noted as 90% [16].

X-ray imaging is commonly employed to investigate conditions such as fractures, bone misalignments, pneumonia, and tumors. This method has a long history of use, providing a rapid assessment of the lungs and proving beneficial for identifying infections, including COVID-19 [17], [18].

X-rays can produce images showing lung damage, such as pneumonia brought about by the SARS-CoV-2 virus [19]. Due to their speed and low cost, X-rays can help prioritize patients in areas where the healthcare system is strained or access to complex technologies is limited. Furthermore, portable X-ray devices, which are easily transportable with ease to the needed location, are available [19].

CT scans, on the other hand, use X-ray principles to examine soft tissues in the body and are ideal for providing detailed images of organs and soft tissues [20]. Moreover, as they use less radiation, X-rays are faster, less harmful, and more economical than CT scans [21].

Narin et al. [17] introduced a method for the automatic uncovering of COVID-19 through the use of chest X-rays and convolutional neural networks (CNNs). Similarly, Apostolopoulos et al. [18] developed an automatic detection system for COVID-19, which involved analyzing and classifying three different categories: COVID-19, typical pneumonia, and normal conditions.

Transfer learning involves utilizing models that have been previously trained on one task to enhance performance on a new, related task on large and diverse datasets are re-trained on smaller and specific datasets. Using this technique, the model is able to learn faster and more accurately, particularly in fields with limited datasets like medical imaging. The VGG19 model holds significant importance in deep learning research. With its deep and wide layer architecture, VGG19 can classify complex images successfully. Therefore, in this study, the VGG19 model was used to identify a total of 2117 COVID-19 and related lung images.

3. Model Architecture and Methodology

This section outlines the methods and design of three different models used to classify COVID-19 and other lung diseases: the VGG19-based classifier model, the ResNet18-based encoder with a U-Net inspired decoder, and the MATLAB-based customized CNN model. Each model has been developed and evaluated on the dataset that consists of chest X-ray images across four categories: COVID-19, Normal, Lung Opacity, and Viral Pneumonia.

3.1. VGG19-Based Classifier Model

3.1.1 Model Architecture and Transfer Learning

The VGG19 model was employed for classifying the four classes (COVID-19, Normal, Lung Opacity, Viral Pneumonia). It leveraged transfer learning, with pre-trained weights on the ImageNet dataset. For classification purposes, the final layers of VGG19 were fine-tuned to accommodate four output classes. This enabled the model to adapt quickly by utilizing the generalized features learned from large-scale datasets.

3.1.2 Dataset and Preprocessing

The dataset utilized in this study includes chest X-ray images from COVID-19 patients, those with viral pneumonia, and healthy individuals, sourced from multiple databases. These images are available at: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database.

A series of preprocessing steps were undertaken to process and prepare the data for the training phase. The dataset consists of 21,165 chest X-ray images, which were classified into four categories:

- COVID-19: 3,616 images
- Normal: 10,192 images
- Lung Opacity: 6,012 images
- Viral Pneumonia: 1,345 images

Table 1. Organization of the Data Set.

Class	Training Set	Test Set
COVID-19	2893	723
Viral Pneumonia	1076	269
Healthy	8154	2038
Lung Opacity	4810	1202
Total	16933	4232

As shown in Figure 1, the dataset captures a wide variety of cases across different lung conditions. The dataset was split into 80% for training (16,932 images) and 20% for testing (4,233 images). All images were resized to 224x224 pixels to meet the input size requirements of the VGG19 model. Pixel values were normalized to the range of 0-1 for consistency in model processing.

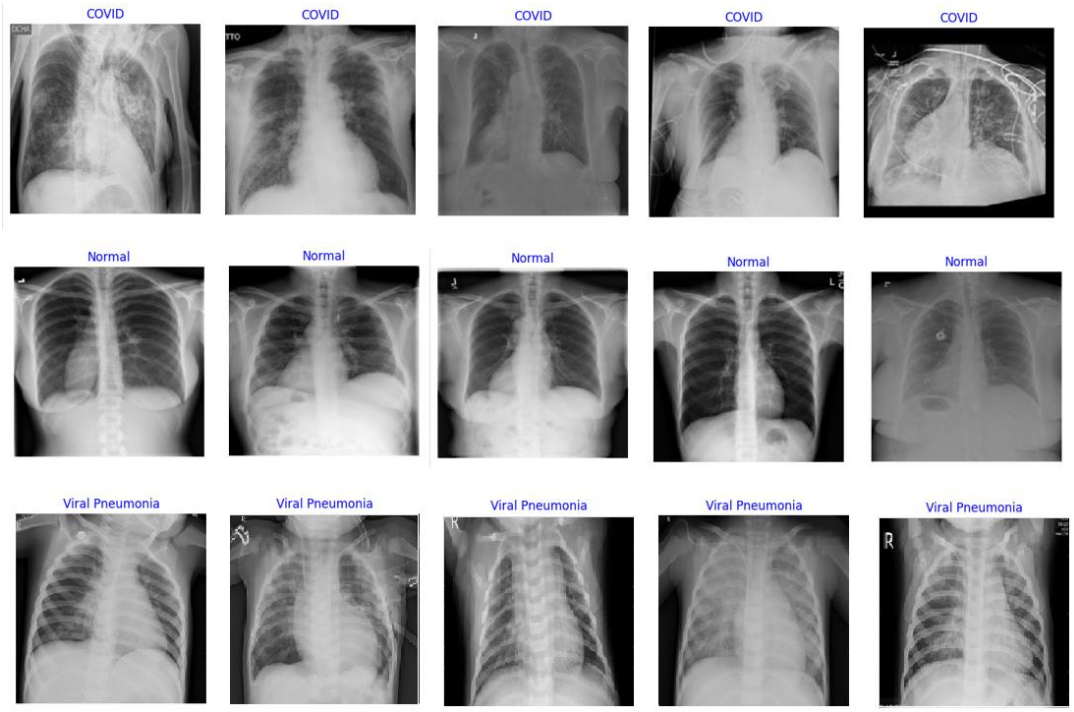

Figure 1. Example images utilized in the experimental analysis of this study.

3.1.3 Data Augmentation

To reduce overfitting and improve the model's ability to generalize, several data augmentation strategies were employed: Horizontal flipping, Random rotation, Zooming and shifting, Brightness adjustments. These augmentations allowed the model to become more invariant to orientation and lighting changes, improving its real-world applicability.

3.1.4 Training Process

The VGG19 model was optimized with the Adam algorithm, configured with a learning rate of 0.0001 to enhance its training efficiency. The training process was monitored with early stopping to prevent overfitting, and the model with the best validation accuracy was saved. Transfer learning enabled faster convergence by leveraging pre-trained weights. An evaluation was conducted on the model's performance using accuracy, precision, specificity, and F1 score metrics. These metrics were calculated using the following equations:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (3)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4)$$

Where TP is true positives, TN is true negatives, FP is false positives and FN is false negatives.

3.2. ResNet18-Based Encoder and U-Net Inspired Decoder Model

3.2.1 Encoder Architecture (ResNet18)

The encoder component of this model is based on ResNet18, a residual network that uses residual connections to facilitate learning in deep networks. The network begins with a 7x7 convolutional layer, followed by several 3x3 convolutional blocks and max-pooling layers. Residual connections ensure the gradient flows effectively through the network, improving convergence. To mitigate the learning difficulties in deep networks, residual connections are utilized. The residual block is defined as

$$\mathrm{Y} = \mathrm{F}(\mathrm{X}) + \mathrm{X} \quad (5)$$

where X is the input and F is the learned transformation.

To reduce feature dimensions, max pooling is applied:

$$\mathrm{X}_{\mathrm{pool}} = \mathrm{MaxPool}(\mathrm{X}, \text{kernel size} = 2) \quad (6)$$

3.2.2 Decoder Architecture (U-Net)

The decoder is inspired by the U-Net architecture, which is well-suited for segmentation tasks. The decoder uses transpose convolutional layers to upsample the feature maps to their original dimensions. These upsampled features are further refined by dual convolutional layers, and the final output is generated through a convolutional layer that produces a segmentation map of the lung regions, essential for differentiating between COVID-19 and other lung diseases. These layers upsample the feature maps to the original dimensions. Mathematically, this operation is represented as;

$$\mathrm{X}_{\mathrm{up}} = \mathrm{ConvTranspose2d}(\mathrm{X}_{\mathrm{in}}, \mathrm{W}_{\mathrm{trans}}, \text{stride} = 2) \quad (7)$$

These layers refine the upsampled features for detailed segmentation:

$$\mathrm{X}_{\mathrm{out}} = \mathrm{Conv2d}(\mathrm{X}_{\mathrm{up}}, \mathrm{W}_{\mathrm{conv}}, \text{kernel size} = 3) \quad (8)$$

This produces the final segmentation map:

$$\mathrm{Y_{seg}}=\mathrm{Conv2d}(\mathrm{X_{act}},\mathrm{W_{conv}},\mathrm{kernel\ size}=\mathbf{1}) \quad (9)$$

3.2.3 Training Process

The ResNet18 encoder and U-Net decoder model was trained to perform lung image segmentation. The model was optimized using the Adam optimizer, with careful tuning of the learning rate. The model was evaluated on its ability to identify lung regions affected by COVID-19 and other lung conditions, improving the overall diagnostic accuracy for the dataset. The loss function used is Binary Cross Entropy with Logits, defined as [22]:

$$Loss = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)] \quad (10)$$

where $\hat{y}_i$ represents the model's prediction, y_i denotes the true label, and N represents the total number of samples.

The model is trained using the Adam optimization algorithm, which updates weights according to:

$$\mathrm{W}_{t+1} = \mathrm{W}_t - \eta.\nabla Loss(\mathrm{W}_t) \quad (11)$$

where η is the learning rate, W_t denotes the model weights, and $.\nabla Loss(\mathrm{W}_t)$ represents the gradient of the loss function.

3.3. Matlab-Based Customized CNN Model

3.3.1 Model Architecture

A customized CNN model was designed using Matlab for the classification task. The input to the model consists of RGB images of size 256x256x3, and the architecture includes:

Input Layer: The model accepts RGB images of size 256x256x3.

Convolutional Layers: The initial convolutional layer has 32 filters with a 3x3 kernel size. This layer is succeeded by batch normalization and ReLU activation layers, which are employed to improve the effectiveness of feature extraction.

$$\mathrm{Output_i}=\mathrm{ReLU}(\mathrm{Conv}(\mathrm{Input},\mathrm{Filters_i},\mathrm{Stride_i})+\mathrm{Bias_i} \quad (12)$$

Here, Conv denotes the convolution operation, ReLU is the activation function, and Outputi is the output feature map.

Max Pooling Layer: Reduces spatial dimensions to increase computational efficiency. Max pooling is typically performed with 2×22 \times 22×2 filter sizes.

Fully Connected Layers: Enhance the learning capacity of the model and support the extraction of high-level features. The final layer includes a softmax activation function that provides output for four classes.

$$\text{Softmax}(z) = \frac{exp(z_i)}{\sum_j exp(z_j)} \quad (13)$$

Here, z denotes the output of the classification layer and Softmax is the activation function.

3.3.2 Dataset and Data Augmentation

To prepare the images for the model, several preprocessing techniques were applied. "imageDatastore" was used to split the labeled images into training and testing sets. Data augmentation was performed to enhance the model's generalization capability:

Rotation: Images are rotated at different angles to enable the model to identify and interpret images from various perspectives.

Horizontal and Vertical Flipping: Images are flipped both horizontally and vertically to aid the model in learning and recognizing symmetric patterns.

Zooming and Shifting: Images are zoomed and shifted to allow the model to learn objects at different scales.

3.3.3 Training and Evaluation

The model was trained using the "trainNetwork" function with the "adam" optimization algorithm. Training involved calculating loss and accuracy using the following equations:

Loss Function:

$$\text{Loss} = -\sum_i y_i \log \hat{y}_i \quad (14)$$

Where y_i are the true labels and $\hat{y}_i$ denotes the predicted probabilities.

Accuracy Calculation:

$$\text{Accuracy} = \frac{Correct\ Predictions}{Total\ Examples} \quad (15)$$

Loss and accuracy graphs, as well as confusion matrices, were presented. The loss and accuracy graphs illustrate the learning process of the model, while the confusion matrices detail the classification performance and the rates of correct and incorrect classifications for each class.

4. Experimental Results

The experimental results for the three models—VGG19-based classifier, ResNet18-based encoder with U-Net inspired decoder, and the Customized Convolutional Neural Network (CNN)—are presented in this section. These results evaluate the performance of each model across the dataset. VGG19 focusing on key metrics such as accuracy,

precision, specificity, F1 score, and loss during both the training and testing phases. Additionally, confusion matrices and visualizations of accuracy and loss curves are presented to demonstrate the models' learning patterns and generalization capabilities.

4.1 VGG19-Based Classifier Model

4.1.1 Performance Metrics

The evaluation of the VGG19 model's performance was conducted using essential metrics such as accuracy, precision, specificity, and the F1 score. Table 2 offers a detailed overview of how the model performed for each class.

Table 2. VGG19 Model Performance Metrics.

Method	Precision (%)	Specificity (%)	F1 Score (%)	Accuracy (%)
COVID-19	98	98	98	97.5
Lung Opacity	95	91	93	91.02
Healthy	94	97	96	97.3
Viral Pneumonia	98	96	97	96.2

The COVID-19 class achieved the highest performance across most metrics, demonstrating the model’s effectiveness at distinguishing COVID-19 from other lung conditions.

4.1.2 Accuracy and Loss Curves

The training process for the VGG19 model is depicted in Figure 2 and Figure 3:

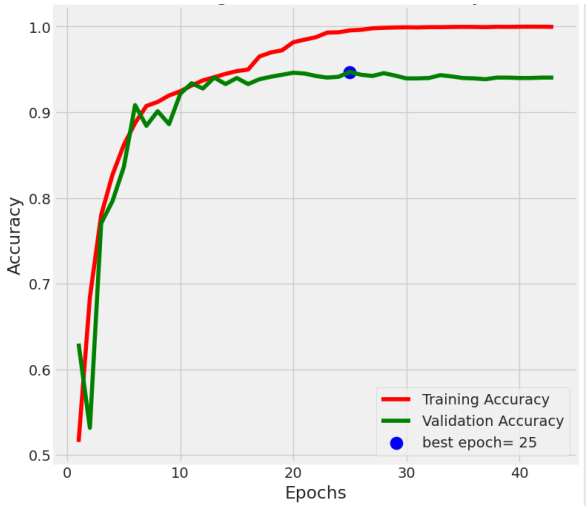

Figure 2. Training and Validation Accuracy Graph.

Figure 2 shows the accuracy progression over the training phase, where the training accuracy consistently increases, while validation accuracy stabilizes after reaching a peak.

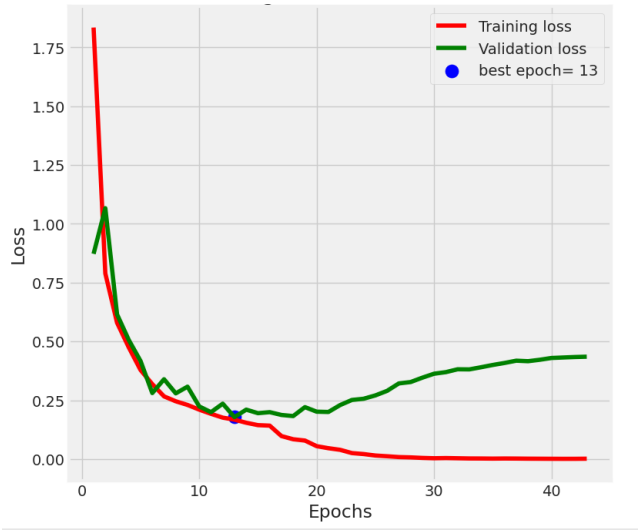

Figure 3. Training and validation loss graph.

Figure 3 illustrates the reduction in loss during training, with the training loss decreasing steadily, and validation loss stabilizing after a certain point. This suggests that the model has learned effectively and is able to generalize well to unseen data.

4.1.3 Confusion matrix

Figure 4 presents the confusion matrix, which provides insights into the classification performance for each class. The matrix shows high counts of true positives (TP) and true negatives (TN) across all classes, indicating strong model performance. The minimal instances of false positives (FP) and false negatives (FN) underscore the model's reliability in accurately distinguishing between COVID-19, lung opacity, healthy individuals, and viral pneumonia cases.

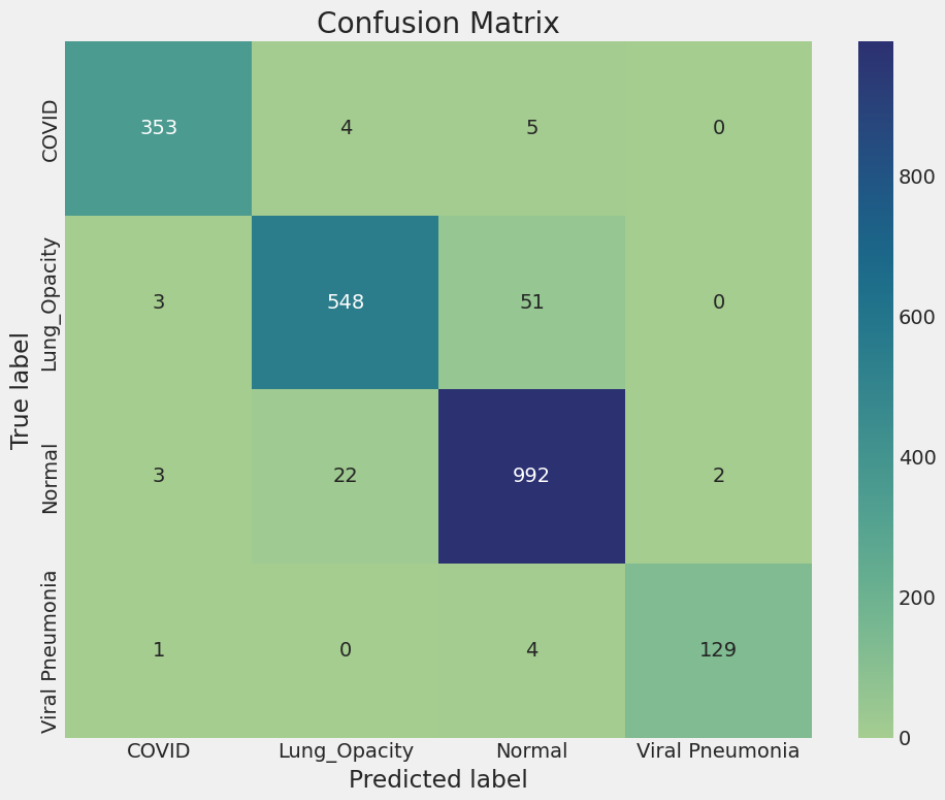

Figure 4. Confusion matrix.

The suggested approach for diagnosing COVID-19 demonstrates superior performance compared to other methods.

Table 4. Evaluating the Proposed COVID-19 Diagnostic Method Against Alternative Approaches.

Study	Type of Images	Number of Cases	Method Used	Accuracy (%)
[18]	X-ray	224 COVID-19 700 Pneumonia 504 Normal	Transfer learning+VGG19	93.48
[23]	X-ray	1300 COVID-19 1300 Pneumonia 1300 Normal	DTL+VGG-19	92.92
[24]	X-ray	219 COVID-19 1300 Pneumonia 1300 Normal	BND+VGG-19	95.48
[25]	CT	219 COVID-19 224 Pneumonia 758 Normal	ResNet+Location Attention	86.07
[26]	X-ray / CT	1493 COVID-19 2780 Pneumonia 1538 Normal	Inception Resnet V2	92.18
[27]	X-ray	2210 COVID-19 2340 Pneumonia 1480 Normal	CXRVN	93.07

[28]	X-ray	305 COVID-19 305 Pneumonia 305 Normal	CovXNet	89.6
[29]	CT	777 COVID-19 505 Pneumonia 708 Normal	ARENET	93.00
[30]	X-ray	260 COVID-19 300 Pneumonia 300 Normal	Transfer learning +VGG19	89.30
This Study	**X-ray**	**3616 COVID-19** **1345 Pneumonia** **10192 Normal** **6012 Lung Opacity**	**Transfer learning +VGG19**	**96.00**

4.2 ResNet18-Based Encoder and U-Net Inspired Decoder Model

4.2.1 Performance Evaluation

The ResNet18-based encoder with a U-Net inspired decoder was designed for lung segmentation and feature extraction from the X-ray images. The performance of the model was analyzed using both accuracy and loss metrics during training and validation.

The training and validation loss curves show that the model's training loss decreases consistently, while validation loss remains stable after a certain point, suggesting good generalization. Likewise, the accuracy curves show that training accuracy steadily improves with each epoch, and validation accuracy reaches a plateau, demonstrating that the model is not overfitting.

4.2.2 Segmentation Performance

The ResNet18-U-Net model was specifically evaluated for its capacity to identify and segment lung areas impacted by different conditions such as COVID-19 and viral pneumonia. The model attained strong performance in segmenting lung opacity and viral pneumonia areas, supporting its application in biomedical imaging.

The ResNet18 and ResNet18-Based Encoder with U-Net Inspired Decoder models were evaluated based on their training and validation accuracy, as well as their training and validation loss. The graphs below depict the models' performance during training. The training and validation accuracy graphs illustrate how the models progressively improved their accuracy over time, demonstrating their ability to learn effectively from the data. Likewise, the training and validation loss graphs show a consistent reduction in loss, indicating that both models successfully minimized errors during training while maintaining stable performance on the validation set. These results highlight the models' capability for generalization and effective learning throughout the training process.

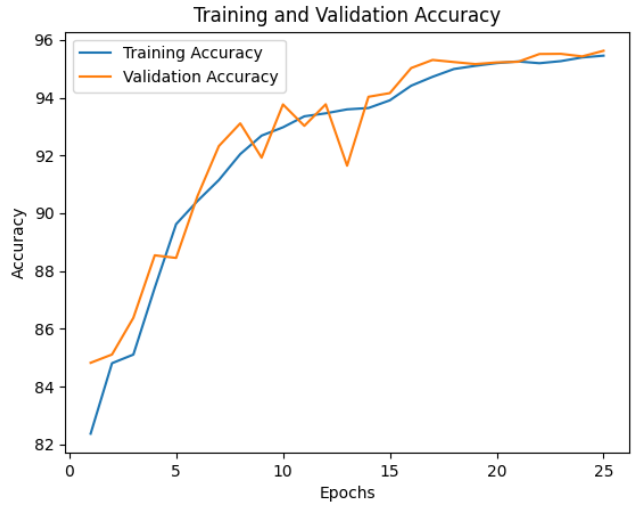

Figure 5. Training and validation accuracy graph for ResNet18.

Figure 6. Training and validation loss graph for ResNet18.

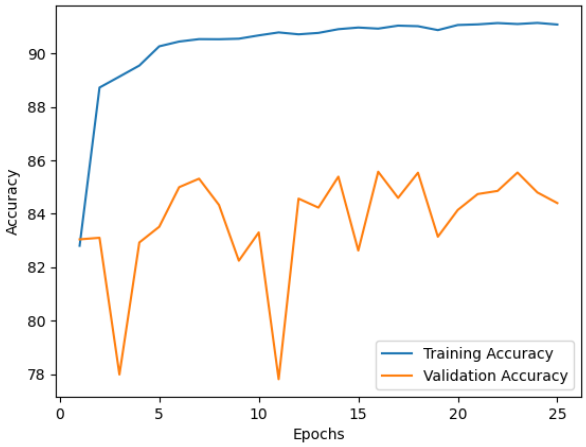

Figure 7. Training and validation accuracy graph for ResNet18-Based Encoder with U-Net Inspired Decoder.

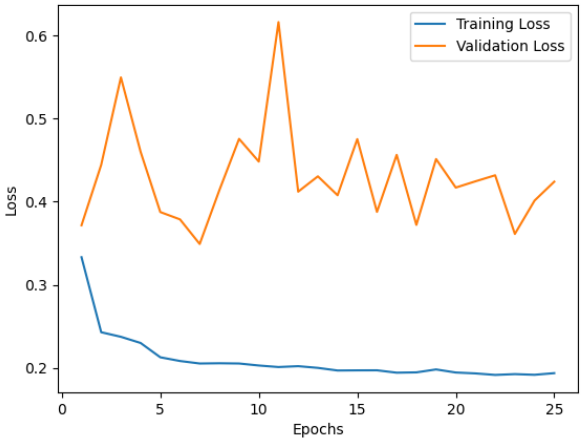

Figure 8. Training and validation loss graph for ResNet18-Based Encoder with U-Net Inspired Decoder.

4.3 Customized Convolutional Neural Network (CNN) Model (Matlab)

4.3.1 Model Performance

The customized CNN model developed in MATLAB was evaluated based on accuracy, with the model achieving an overall accuracy of 85.56% on the test set. The training process involved 2,645 iterations across 5 epochs, with each epoch consisting of 529 iterations. This setup illustrates the model's ability to accurately differentiate among COVID-19, normal, lung opacity, and viral pneumonia cases.

4.3.2 Accuracy and Loss Graphs

During the training process, the model's loss steadily decreased, indicating successful learning. The accuracy chart illustrates that the model attained high performance on both the training and validation datasets, demonstrating minimal signs of overfitting. The graphs depicting Training and Validation Accuracy and Loss are presented below, offering a visual summary of the model's performance across the training phase.

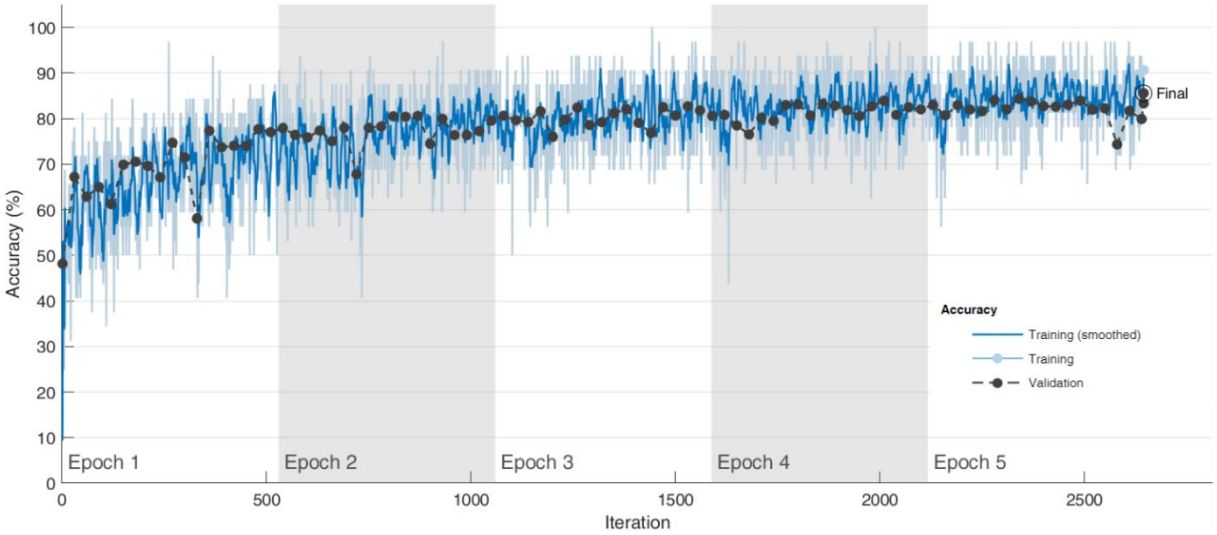

Figure 9. Accuracy vs. Iteration for Matlab.

This figure displays the model's accuracy throughout the training process. The x-axis represents the number of iterations, while the y-axis illustrates the accuracy achieved by the model at each iteration. The graph presents how the model's accuracy evolves as training progresses, showing the improvements made with each iteration and helping to assess the overall effectiveness of the training process. The plot highlights both the training accuracy and the validation accuracy, enabling a comparison of the model's performance on training data versus unseen validation data.

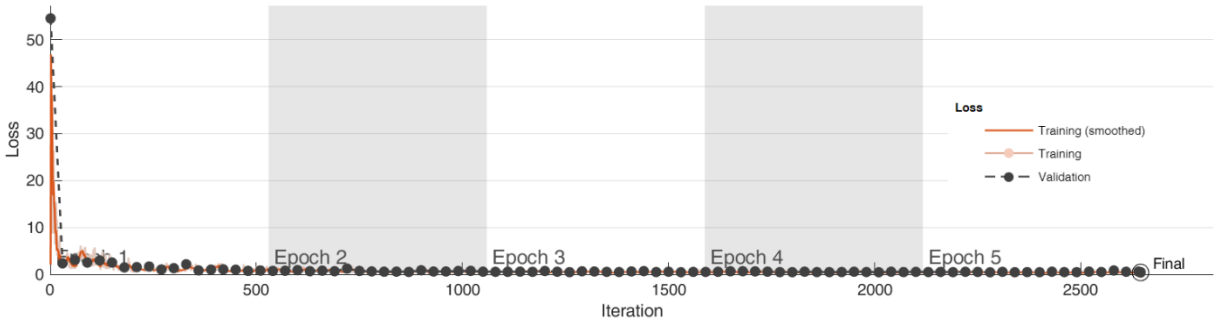

Figure 10. Loss vs. Iteration for Matlab.

This figure shows the model's loss during training. The x-axis signifies the number of iterations, while the y-axis reflects the loss value recorded for the model at each iteration. This plot helps visualize how the loss decreases as the model learns and improves over time. By comparing the loss values across iterations, we can evaluate the convergence of the model and the effectiveness of the training. The graph typically displays both the training loss and the validation loss, providing a clear view of how well the model is generalizing to unseen data as training progresses.

4.3.3 Confusion Matrix

The confusion matrix for the customized CNN model illustrates the performance of the model in classifying the four lung conditions: COVID-19, Normal, Lung Opacity, and Viral Pneumonia. The matrix offers a detailed account of the predictions, categorizing them into true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) for each class. Each cell in the matrix reflects the number of predictions made by the model for a particular class compared to the actual class.

The diagonal elements represent the correct predictions for each class, showing high precision and recall rates across all classes. This indicates the model's effectiveness in accurately distinguishing between different lung conditions. The off-diagonal elements represent misclassifications, providing insights into the types of errors the model makes. By analyzing these errors, we can assess areas for improvement and understand how well the model generalizes to various conditions.

This detailed confusion matrix is crucial for evaluating the model's overall classification performance and for identifying any biases or specific challenges the model may face in distinguishing between certain classes. The confusion matrix is shown in Figure 11.

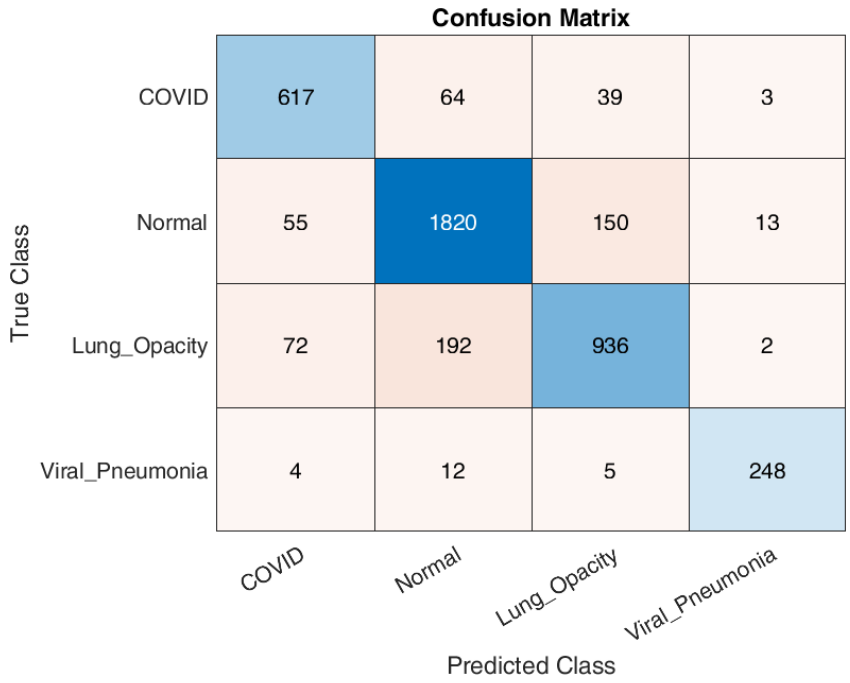

Figure 11. Confusion Matrix for Customized CNN Model (MATLAB).

4.4 Overall Comparison and Discussion

The performance of each model was compared, revealing unique strengths depending on the task:

- The VGG19-based model excelled in classification tasks, particularly in identifying COVID-19 cases with a precision of 98%.
- The ResNet18-based encoder with a U-Net decoder performed well in segmenting lung regions affected by various conditions, making it suitable for segmentation tasks.
- The customized CNN model achieved strong classification results, benefiting from various data augmentation techniques.

The VGG19-based model's superior performance in classifying COVID-19 can be attributed to its transfer learning strategy, while the ResNet18-U-Net architecture was optimized for biomedical image segmentation, excelling at feature extraction. The customized CNN model in Matlab offered flexibility in classification, achieving robust results across all metrics.

5. Discussion and Future Work

5.1. Discussion

This study explored the application of deep learning models, including VGG19, ResNet18-based encoder with a U-Net inspired decoder, and a Customized CNN model in the classification and segmentation of COVID-19 and related lung diseases using medical imaging. The performance of each model demonstrates the effectiveness of deep learning techniques in medical diagnostics, but also highlights certain challenges

and opportunities for improvement. Below, we present a more detailed discussion of the models' limitations and the challenges faced during development.

VGG19 Model

The VGG19 model achieved high validity in classifying COVID-19 and other lung conditions, demonstrating the potential of transfer learning in medical imaging. Its success is largely due to the ability to leverage pre-trained layers to capture complex features. However, a significant limitation observed was the model's tendency to overfit on the training data, which can be attributed to the restricted diversity of the dataset. This overfitting suggests that the model's generalization ability could be compromised when applied to unseen data. Additionally, VGG19's computational complexity can hinder real-time or large-scale deployment in clinical settings. Increasing the dataset size, especially by including more diverse samples, and enhancing data augmentation techniques (e.g., brightness adjustment, rotation, scaling) would help mitigate this issue. Furthermore, regularization techniques like L2 regularization, dropout, and batch normalization will prevent overfitting and improve generalization.

ResNet18-Based Encoder with U-Net Inspired Decoder

The ResNet18-based encoder with U-Net inspired decoder was highly effective for biomedical image segmentation tasks, particularly when identifying lung regions affected by diseases like COVID-19 and viral pneumonia. The combination of ResNet18 for feature extraction and U-Net for segmentation enhanced the model's ability to precisely capture both global and local structures within the images. However, a notable challenge was the model's high computational cost, which makes it less suitable for scenarios with limited resources or where real-time predictions are needed. To improve its applicability, future work should focus on optimizing computational efficiency and validating the model on additional biomedical datasets to ensure robustness across a broader range of medical conditions. Testing it on more complex or unseen medical images would further enhance its real-world performance.

Customized CNN Model (Matlab)

The Customized CNN model developed in Matlab also delivered promising results, but its limitations were more evident when compared to the transfer learning-based architectures. While the customized CNN performed well in classification tasks, the absence of pre-trained layers, as used in VGG19 and ResNet18, limits its capacity to generalize across more complex datasets. Additionally, the analysis of the training-validation accuracy/loss graphs and confusion matrices revealed areas for further optimization, particularly in the management of class imbalances and feature extraction. Improving the model's architecture and experimenting with deeper convolutional layers could lead to enhanced performance. Moreover, the model could benefit from more advanced data augmentation and regularization techniques to improve robustness across various datasets.

Overall, these deep learning models underscore the potential for AI-driven diagnostics in medical imaging. However, it is critical to address the challenges of generalization, overfitting, and computational cost to maximize the models' clinical applicability.

5.2. Challenges in Model Development

In the process of model development, several key challenges were identified:

1. **Dataset Limitations:** The relatively small and less diverse dataset posed challenges for generalization. Class imbalance was particularly noticeable, which could skew the model's performance toward overrepresented classes, such as healthy individuals or common lung diseases like pneumonia. More comprehensive datasets and synthetic data generation techniques would help improve performance on underrepresented classes like COVID-19.
2. **Overfitting Issues:** Overfitting, especially in the VGG19 model, was a significant concern. While data augmentation helped, more sophisticated techniques, such as adversarial training or contrastive learning, could offer more robust solutions. Additionally, cross-validation strategies were employed to better estimate the model's true performance, but further work is needed to explore how different regularization and early stopping mechanisms can be combined for optimal results.
3. **Computational Constraints:** The high computational cost of training deep learning models, especially for U-Net-inspired architectures, required the use of advanced hardware. Reducing the model's footprint via model pruning or quantization techniques could enhance its usability in real-time clinical environments.

5.3. Future Work

Based on the challenges and findings of this study, several directions for future research are proposed to enhance model performance and reliability:

VGG19 Model

Future research should prioritize expanding the dataset to improve the model’s generalizability. This could involve curating a larger set of X-ray images from a wider demographic and geographical spectrum to ensure that the model can handle more diverse cases. Additionally, exploring alternative deep learning architectures such as InceptionNet or EfficientNet could yield improvements in both accuracy and computational efficiency. Further development of data augmentation techniques that target underrepresented classes could also significantly enhance the model's performance in real-world clinical settings. Finally, additional regularization methods (e.g., label smoothing, dropout) and ensemble learning approaches should be tested to further prevent overfitting.

ResNet18-Based Encoder with U-Net Inspired Decoder

For the ResNet18-based U-Net model, future research should focus on optimizing the model's computational efficiency for real-time use. This could involve implementing model compression techniques like distillation or pruning to reduce resource consumption without sacrificing performance. Furthermore, applying the model to other types of medical imaging (e.g., CT scans, MRI) could validate its versatility across different domains of medical diagnostics. Collaborating with medical professionals to create user-friendly interfaces that allow for practical, real-time use of these models in clinical environments would be another beneficial step.

Customized CNN Model (Matlab)

Future work on the Customized CNN should focus on exploring more complex architectures, such as deeper networks or hybrid models that combine CNNs with attention mechanisms to improve feature detection and classification. Additionally,

improving the model's generalization by testing on larger datasets and applying more aggressive regularization techniques (e.g., early stopping, weight decay) will help address overfitting. Lastly, comparing the model against newer architectures could offer insights into further performance improvements.

In conclusion, this research highlights the potential of deep learning models in the field of medical imaging, particularly for diagnosing COVID-19 and associated lung diseases. While promising, these models still face limitations, particularly in terms of generalization, computational efficiency, and handling diverse clinical data. Addressing these challenges through larger datasets, more robust architectures, and computational optimizations will be crucial in refining these models for real-world use.

6. Conclusion

This study explored the application of various deep learning models, including VGG19, a ResNet18-based encoder with a U-Net inspired decoder, and a Customized CNN model, for the classification and segmentation of COVID-19 and related lung diseases using chest X-rays and biomedical images. The results from all three models demonstrate the significant potential of deep learning in medical imaging, particularly for tasks that require rapid and accurate diagnoses, such as COVID-19 detection.

The VGG19 model achieved high accuracy, precision, specificity, and F1 score, proving that transfer learning and data augmentation techniques are effective for medical image classification. Its success underscores the potential of deep learning models to assist in clinical settings, particularly for urgent health conditions where timely diagnosis is critical. Despite these achievements, future work should focus on expanding datasets and exploring alternative architectures to further improve the model’s performance and generalization ability.

Similarly, the ResNet18-based encoder and U-Net inspired decoder demonstrated its effectiveness in biomedical image segmentation tasks, outperforming traditional methods. This model's success in segmenting lung diseases highlights its potential application in diverse medical imaging tasks. Future research should focus on testing the model on a broader range of biomedical datasets, optimizing its computational efficiency, and developing user-friendly interfaces for real-world clinical deployment.

The Customized CNN model, developed using Matlab, also delivered promising results in classifying lung diseases, achieving high accuracy and demonstrating its suitability for clinical applications. However, future efforts should involve testing the model with larger and more diverse datasets, and exploring advanced data augmentation and regularization techniques to further enhance its performance.

In conclusion, the findings of this study emphasize the practical benefits and potential of deep learning models in medical image analysis and diagnostics. With continued research focused on optimizing model architectures, expanding datasets, and improving generalization, these models can play a pivotal role in improving the quality of healthcare by enabling faster, more accurate, and reliable diagnoses in real-world clinical environments.

References

[1] Punn, N. S., & Agarwal, S. (2021). Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks. Applied Intelligence, 51(5), 2689-2702.

[2] Srinivas, K., Gagana Sri, R., Pravallika, K., Nishitha, K., & Polamuri, S. R. (2024). COVID-19 prediction based on hybrid Inception V3 with VGG16 using chest X-ray images. Multimedia Tools and Applications, 83(12), 36665-36682.

[3] Govindarajan, S., & Swaminathan, R. (2021). Differentiation of COVID-19 conditions in planar chest radiographs using optimized convolutional neural networks. *Applied Intelligence*, *51*(5), 2764-2775.

[4] Wang, Y., Kang, H., Liu, X., & Tong, Z. (2020). Combination of RT-qPCR testing and clinical features for diagnosis of COVID-19 facilitates management of SARS-CoV-2 outbreak. *Journal of medical virology*, *92*(6), 538.

[5] Adhikari, S. P., Meng, S., Wu, Y. J., Mao, Y. P., Ye, R. X., Wang, Q. Z., ... & Zhou, H. (2020). Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review. *Infectious diseases of poverty*, *9*, 1-12.

[6] Pan, Y., Li, X., Yang, G., Fan, J., Tang, Y., Zhao, J., ... & Li, Y. (2020). Serological immunochromatographic approach in diagnosis with SARS-CoV-2 infected COVID-19 patients. *Journal of Infection*, *81*(1), e28-e32.

[7] Spicuzza, L., Montineri, A., Manuele, R., Crimi, C., Pistorio, M. P., Campisi, R., ... & Crimi, N. (2020). Reliability and usefulness of a rapid IgM-IgG antibody test for the diagnosis of SARS-CoV-2 infection: A preliminary report. *The Journal of infection*, *81*(2), e53.

[8] Porte, L., Legarraga, P., Vollrath, V., Aguilera, X., Munita, J. M., Araos, R., ... & Weitzel, T. (2020). Evaluation of a novel antigen-based rapid detection test for the diagnosis of SARS-CoV-2 in respiratory samples. *International Journal of Infectious Diseases*, *99*, 328-333.

[9] Nayak, S. R., Nayak, D. R., Sinha, U., Arora, V., & Pachori, R. B. (2022). An efficient deep learning method for detection of COVID-19 infection using chest X-ray images. *Diagnostics*, *13*(1), 131.

[10] Gupta, K., & Bajaj, V. (2023). Deep learning models-based CT-scan image classification for automated screening of COVID-19. Biomedical Signal Processing and Control, 80, 104268.

[11] Barshooi, A. H., & Amirkhani, A. (2022). A novel data augmentation based on Gabor filter and convolutional deep learning for improving the classification of COVID-19 chest X-Ray images. *Biomedical Signal Processing and Control*, *72*, 103326.

[12] Basu, S., Mitra, S., & Saha, N. (2020, December). Deep learning for screening covid-19 using chest x-ray images. In *2020 IEEE symposium series on computational intelligence (SSCI)* (pp. 2521-2527). IEEE.

[13] Che Azemin, M. Z., Hassan, R., Mohd Tamrin, M. I., & Md Ali, M. A. (2020). COVID-19 deep learning prediction model using publicly available radiologist-adjudicated chest x-ray images as training data: preliminary findings. *International Journal of Biomedical Imaging*, *2020*(1), 8828855.

[14] Aslan, M. (2022). Derin Öğrenme Tabanlı Otomatik Beyin Tümör Tespiti. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, *34*(1), 399-407.

[15] Erdoğan, A. M. N., Öztürk, T., & Talo, M. (2022). Yeni bir Evrişimsel Sinir Ağı Modeli Kullanarak Bilgisayarlı Tomografi Görüntülerinden Akciğer Kanseri Tespiti. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, *34*(2), 795-802.

[16] Khan, A. I., Shah, J. L., & Bhat, M. M. (2020). CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Computer methods and programs in biomedicine*, *196*, 105581.

[17] Narin, A., Kaya, C., & Pamuk, Z. (2021). Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *Pattern Analysis and Applications*, *24*, 1207-1220.

[18] Apostolopoulos, I. D., & Mpesiana, T. A. (2020). Covid-19: automatic detection from x-ray images

utilizing transfer learning with convolutional neural networks. *Physical and engineering sciences in medicine*, *43*, 635-640.

[19] Wang, L., Lin, Z. Q., & Wong, A. (2020). Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific reports*, *10*(1), 19549.

[20] Kroft, L. J., van der Velden, L., Girón, I. H., Roelofs, J. J., de Roos, A., & Geleijns, J. (2019). Added value of ultra–low-dose computed tomography, dose equivalent to chest x-ray radiography, for diagnosing chest pathology. *Journal of thoracic imaging*, *34*(3), 179-186.

[21] Damilakis, J., Adams, J. E., Guglielmi, G., & Link, T. M. (2010). Radiation exposure in X-ray-based imaging techniques used in osteoporosis. *European radiology*, *20*, 2707-2714.

[22] Su, Q., Kloukinas, C., & Garcez, A. D. A. (2024, June). FocusLearn: Fully-Interpretable, High-Performance Modular Neural Networks for Time Series. In *2024 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.

[23] Fayemiwo, M. A., Olowookere, T. A., Arekete, S. A., Ogunde, A. O., Odim, M. O., Oguntunde, B. O., ... & Kayode, A. A. (2021). Modeling a deep transfer learning framework for the classification of COVID-19 radiology dataset. *PeerJ Computer Science*, *7*, e614.

[24] Cao, Z., Huang, J., He, X., & Zong, Z. (2022). BND-VGG-19: A deep learning algorithm for COVID-19 identification utilizing X-ray images. *Knowledge-Based Systems*, *258*, 110040.

[25] Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., ... & Li, L. (2020). A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering*, *6*(10), 1122-1129.

[26] El Asnaoui, K., & Chawki, Y. (2021). Using X-ray images and deep learning for automated detection of coronavirus disease. *Journal of Biomolecular Structure and Dynamics*, *39*(10), 3615-3626.

[27] Elzeki, O. M., Shams, M., Sarhan, S., Abd Elfattah, M., & Hassanien, A. E. (2021). COVID-19: a new deep learning computer-aided model for classification. *PeerJ Computer Science*, *7*, e358.

[28] Mahmud, T., Rahman, M. A., & Fattah, S. A. (2020). CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Computers in biology and medicine*, *122*, 103869.

[29] Song, Y., Zheng, S., Li, L., Zhang, X., Zhang, X., Huang, Z., ... & Yang, Y. (2021). Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM transactions on computational biology and bioinformatics*, *18*(6), 2775-2780..

[30] Rahaman, M. M., Li, C., Yao, Y., Kulwa, F., Rahman, M. A., Wang, Q., ... & Zhao, X. (2020). Identification of COVID-19 samples from chest X-Ray images using deep learning: A comparison of transfer learning approaches. *Journal of X-ray Science and Technology*, *28*(5), 821-839.