DOI: 10.15849/IJASCA.250330.11

Int. J. Advance Soft Compu. Appl, Vol. 17, No. 1, March 2025
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Multi-class Lung Disease Classification Using Convolutional Neural Networks

Ola Refai Momani[1]* and Mohammad Belal Al-Zoubi[2] and Marwan Al Tawil[3]

[1,2,3] Department of Computer Information System
University of Jordan
Amman-Jordan
Emails: Olamojor@gmail.com,[1], mba@ju.edu.jo[2], m.alTawil@ju.edu.jo [3]

Abstract

Lung diseases place a significant and growing burden on global healthcare systems. However, any technological tool that enables rapid and highly accurate screening of lung diseases can be crucially beneficial to healthcare professionals. Deep learning algorithms are critical in disease identification and medical decision-making, analyzing vast datasets, and uncovering patterns beyond human perception. This study aims to comprehensively evaluate artificial intelligence-based segmentation algorithms and their efficacy in segmenting X-ray images, three public Convolution Neural Networks (CNN)-based semantic segmentation models: Fully Convolutional Network (FCN) using pre-trained VGG16, (U-Net) trained from scratch, and (LinkNet) provided by Keras, the lung area was segmented and then fed to Convolutional Neural Networks to classify the CRX into four classes of lung diseases. Five different pre-trained CNN architectures (Xception, VGG16, ResNet50V2, MobileNetV2, and DenseNet201) were investigated on the plain and segmented lung CXR images. An image pre-processing pipeline was used in this research to enhance the classification process. The results of our experiments show that LinkNet by Keras outperformed the other segmentation networks in terms of accuracy, Jaccard, Dice, and Mean IoU of 98.13, 92.64, 96.09, and 96.49 respectively. The classification results show that classification performance from plain Chest X-ray images is better than the segmented; theDenesNet201 obtained the best classification results of non-segmented images with accuracy, F1-Score 91.32, 91.30, respectively.

Keywords: *Image Segmentation, Convolution Neural Networks, Deep Learning, Transfer Learning, X-Ray Images.*

1. Introduction

Chest X-rays (CXR), ultrasounds, and magnetic resonance imaging (MRI) are used to detect lung diseases and diagnose respiratory infections such as COVID-19, viral pneumonia, and lung opacity. Medical experts often face challenges in accurately diagnosing these diseases due to the similarity of their pathological symptoms. In recent years, computer-aided diagnosis

Received 4 January 2025; Accepted 25 February 2025

(CAD) systems and artificial intelligence (AI) algorithms have been employed to identify diseases and support medical decisions before treatment begins. Physicians and radiologists can leverage CAD systems, AI, and data mining to gain new insights into medical data, thereby complementing and enhancing their expertise. One of the most significant features of AI and machine learning is their ability to process vast amounts of data and learn from it, enabling the extraction of subtle features that might go unnoticed by the human eye. [1], [2].

In recent years, extensive research has focused on applying neural networks and deep learning technologies to X-ray images for lung disease detection. Convolutional Neural Networks (CNNs) are designed to analyze multi-dimensional inputs such as images for classification and recognition [1],[3],[4],[5].

These networks have demonstrated significant success and notable advancements in both accuracy and computational efficiency, as highlighted in studies[1],[6],[7]. In medical applications, CNNs have proven effective in classifying conditions such as brain tumors, breast cancer, pneumonia, arterial disorders, and throat conditions, as well as in detecting COVID-19 [8],[9],[10],[11].

During the COVID-19 pandemic, many researchers applied CNNs to detect the disease. Some focused on segmentation techniques, while others relied solely on classification methods. The high performance and accuracy of both segmentation and classification approaches have paved the way for enhanced automated clinical systems, aiding doctors and radiologists in making more informed patient care decisions [12].

Segmentation approaches using CNN networks are mainly based on FCN, which consists of an encoder and decoder [6], [13]. Various deep learning (DL) networks have been utilized for segmentation, including U-Net, which is widely used in medical image segmentation due to its skip connections between the encoder and decoder [10], [14]. LinkNet, now available by Keras, is recognized for its efficiency in real-time applications and has been mainly used in self-driving and augmented reality [15]. Several researchers have achieved high accuracy in detecting COVID-19 using segmentation techniques[16],[17],[18],[19],[20]. However, others have found that segmentation does not provide a competitive advantage in classifying COVID-19. Some studies have even reported that classification without segmentation yielded higher accuracy than using segmentation [6],[14].

A significant challenge in most of these studies is the limited availability of X-ray images of lung diseases, resulting in insufficient data for training and testing. This limitation affects the ability to properly train deep learning models and assess the reliability of machine learning approaches for use as assistive tools in the medical field [21],[22].

One solution to the limited availability of annotated medical images is transfer learning, which leverages pre-trained CNN models that have been trained on large datasets. However, many researchers have reported that ImageNet pre-trained networks may not accurately extract reliable feature representations from medical images, as they were originally trained on natural images rather than medical data [14],[22],[23],[24],[24]. Another effective approach to address this issue is data augmentation, which significantly increases data diversity during training and helps reduce overfitting caused by limited datasets [6],[7],[25],[14].

Medical images often suffer from noise, artifacts, low contrast, or poor quality, which can lead to incorrect diagnoses. Low-quality images are a major cause of inaccurate predictions. Another challenge in lung disease datasets is low resolution**,** which can hinder the extraction of essential features using neural networks [6], [7], [26], [13], [14], [22], [23], [27],

[28], [29], [30]. Therefore, a pre-processing step is crucial [31], [32], [33], [34]. Various techniques have been applied to enhance the quality of medical images, including edge detection [12], cropping images [25], and removing the majority of the diaphragm [35]. Additionally, contrast enhancement such as CLAHE, N-CLAHE, Gamma correction, Complement, and BCET filter, with different parameter values and settings have been used to improve image clarity and feature extraction[6],[19],[26].

This paper presents a pre-processing pipeline to improve the quality of X-ray images. It then classifies the COVID-19 CHEST X-RAY DATABASE into 4 classes by comparing and selecting the best pre-trained CNN architecture. This comparison includes using a segmentation approach and a non-segmentation approach.

The primary contributions of this research are:

1. Comparison of three different Neural Networks (U-Net, FCN, LinkNet) for segmentation to determine the most reliable approach for detecting lung diseases.
2. Proposal of a pre-processing pipeline to enhance chest X-ray images for accurate classification by five pre-trained CNNs (Xception, VGG16, MobileNetV2, ResNet50V2, and DenseNet201). The classification encompasses four classes (COVID-19, Viral Pneumonia, Normal, and Lung Opacity).
3. Comparison of the classification process with and without segmentation and subsequent analysis of the reasons behind these results.

The work is organized as follows. Section two summarizes the literature about COVID-19 detection and classification with and without segmentation. The methodology and the dataset description are clarified in section three. Experimental work, comparisons, results, and discussions of our experiments are clarified in sections Four and Five. Finally, the conclusions and future works are summarized in section Six.

2-Literature Review

Many studies have used machine learning to develop COVID-19 classification systems, primarily by employing Convolutional Neural Networks (CNN). This section explores two approaches for detecting COVID-19: those that utilize segmentation and those that do not.

2.1. COVID-19 detection without segmentation

Many researchers have proposed different ML approaches to detect COVID-19 using raw X-ray images. Some researchers have used transfer learning of well-known networks, and others have modified these architectures, or designed new ones to detect COVID-19.

The COVINet model was proposed by [12]. It used two different datasets. The first contains 79 images for each COVID-19 and bacterial pneumonia, and another dataset contains 78 X-ray images of COVID-19 and 28 images for normal people. The model applied edge detection using a value-based filter ([0,-1,0],[-1,6,-1],[0,-1,0]) through the proposed pre-processing pipeline, the results showed an accuracy of 0.85 and AUC 0.59480.

COVID-Net was introduced by [36], an open-source network architecture designed to detect COVID-19 cases from chest X-ray images, they got an accuracy of 93.3 and a sensitivity of 91.0. The CoroNet model was proposed by [37], the proposed CoroNet model which is based on Xception architecture, for 4 classes. The researchers achieved precision and recall rates of 93% and 98.2% respectively. HCN model was proposed by [24] which uses the first convolution layer from COVIDNet (since it was trained using X-ray images) then followed by

the convolutional layers from well-known pre-trained networks to extract the features, they got accuracy, sensitivity, specificity, and precision of 96.67, 96.08, 96.97 and 98.00 respectively.

To address the challenge of a small dataset, [23] introduced a Convolutional Neural Network (CNN) architecture based on class decomposition with pre-trained models. They employed Principal Component Analysis (PCA) to mitigate high dimensionality, creating more uniform classes and lowering memory demands. Additionally, k-means clustering was applied during the class decomposition phase. Their approach yielded impressive results, accomplishing an accuracy of 97.35%, a sensitivity of 98.23%, and a specificity of 96.34%.

The CoroDet is a deep learning-based classification for COVID-19 detection proposed by [38], they used the COVID-R dataset that contains 7390 X-ray images of 4 classes. The study trained their model using 5-fold cross-validation, the accuracy of the proposed model was 99.1% for 2 classes, 94.2% for 3 class cases, and 91.2% for 4 class classifications.

A comparison study was made by [28], they compared two pre-trained CNNs using transfer learning: VGG19 and MobileNetV2 using two different datasets. The study concluded that DL may extract COVID-19 disease, but showed that more research is needed to evaluate the X-ray images with the corporation with the medical community. According to the study, the best accuracy, sensitivity, and specificity obtained were 96.78%, 98.66%, and 96.46% respectively. They concluded that MobileNet-v2 outperforms VGG19 in terms of specificity. The evaluations were made using 10-fold cross-validation. Another comparison between nine pre-trained convolutional neural networks was made by [22]. They used 5-fold cross-validation. The pre-trained model Se-ResNeXt-50 achieved the highest classification accuracy among the other architectures, the accuracy was 99.32% for the binary class and 97.55% for multi-class.

Eight different popular CNNs were compared by [39]. The researchers used transfer learning using 5-fold cross-validation. the results indicated the superiority of DenseNet201 which is a deep network. They concluded that deep networks perform better than shallow networks. The experiment accuracy was 97.94 sensitivity, 97.94 precision, 97.95 Specificity,98.80 F1-score,97.94 for 3 classes.

To put the results in perspective, Table 1 illustrates a summary of DL approaches for detecting COVID-19 without using segmentation.

TABLE 1
DL-based methods for detection of COVID-19 images without using segmentation approaches

Refers	Strategy	Dataset	Classification	Acc. (%)	Sens. (%)	Spec. (%)	Pre. (%)	F1-score (%)
[40]	RAM-Net	COVIDx dataset of 13675 X-ray	3 classes	95.33	92	-	99	-
[8]	CNN-LSTM	Dataset of 4575 X-ray images	3 classes	99.4	99.3	99.2	-	98.9
[12]	COVI Net	Dataset1: 158 images Dataset2: 106 X-ray images	4 classes	0.85	0.9899	0.9219	-	-
[36]	COVI D-Net	COVIDx of 13,975 CXR	3 classes	93.3	91.0	-	-	-
[41]	DarkCovidNet + YOLO	1125 X-ray images	2 class	98.08 87.02	95.13 85.35	95.30 92.18	-	96.51 87.37

			3 class					
[37]	CoroNet	1300 X-ray images	2 class	99	99.3	98.6	98.3	98.5
				95	96.9	97.5	95	95.6
			3 class	89.6	89.92	96.4	90	89.8
			4 class					
[35]	CAD system (removes diaphragm) /pre-trained VGG16.	Dataset of 8474 X-ray images	3 classes	94.5	98.4	98.0	-	-
[23]	DeTraC	Dataset1: JSRT Dataset2: 116 (with 4248 × 3480 pixels).	3 classes	97.35	98.23	96.34	-	-
[24]	HCN-DML model	COVIDx of 13,975 CXR	3 classes	96.67	96.08	96.97	98.00	
[42]	ResNet 101 plus J48	100 X-ray images	2 classes	98.54	99.32	97.4	99.32	98.53
[38]	CoroDet	COVID-R of 7390 x-ray images	2 class	99.12	95.36	97.36	97.64	96.88
				94.2	92.76	94.56	94.04	91.32
			3 class	91.2	91.76	93.48	92.04	90.04
			4 class.					
[27]	Faster R–CNN	X-Ray of 13800 images	2 classes	97.36	97.65	-	99.28	-
[43]	Fractal CovNet	Dataset1-6078 Xray images Dataset2-6278 Xray images	2-classes	0.98	0.94	-	0.88	0.92
[28]	Mobile Net v2 With transfer learning	Dataset1: 1427 X-ray images	2 class	96.78	98.66	96.46	-	-
		Dataset2: 1442 X-ray images	3 class	94.72	-	-	-	-
[22]	Transfer learning /Se-ResNeXt-50.	1- Dataset of 1428 X-ray images	2 class	98.36	99.11	98.02	95.76	97.4
		2- dataset of 1442 X-ray images		96.99	94.67	97.43	87.36	90.86
		3- 5232 X-ray images	4 class.					
[39]	DenseNet201 /Transfer learning	Dataset of 3487 X-ray images	2 class	99.70	99.70	99.55	99.70	99.69
				97.94	97.94	98.80	97.95	98.80
			3 class.					

Most of the studies in Table 1 used a transfer learning approach [22],[23],[28],[35],[39]. Transfer learning is known as the process of using a pre-trained model on a new problem. These models were trained using millions of available labeled training data, then instead of starting the learning process from scratch, the weights of the pre-trained model are used to perform a new task. Other researchers have trained their models from scratch using small datasets, such as researchers in [8] who used 4575 X-ray images to train their proposed model CNN-LSTM. Also, DarkCovidNet, designed by [41], was trained using 1125 X-ray images. In our research, we will use transfer learning to classify a dataset of 21,165 CXR images.

2.2. COVID-19 detection using Lung segmentation

Image segmentation has become increasingly important in radiology research and clinical practice [11],[17]. Many researchers have used different segmentation techniques. The following part discusses the state-of-the-art segmentation approaches used to detect COVID-19 using Neural networks.

The U-Net has been used in many researches, either by making modifications on U-Net, or by using U-Net embedded with the proposed model. A modified version of U-net architecture was proposed by [6] for lung segmentation from the X-ray images that have outperformed the U-net model then they made a classification using DenseNet201 with Gamma enhancement, they applied 5-fold cross-validation to have the accuracy, precision, and recall of 95.11%, 94.55%, and 94.56% respectively, they used the COVQU Dataset of 18479 CXR for classification and the same dataset with corresponding masks to train the proposed model. Attention U-Net (CXAU-Net) proposed by [19] for multi-class segmentation, that can be used to detect COVID-19 abnormalities from the chest X-ray images such as Ground-glass opacity (GGOs) and cardiomegaly. They applied spatial and channel attention, they also proposed a hybrid loss function for the model. The resulting accuracy was 95.15% using 840 X-ray images of COVID-19.

Another research has proposed by [44], they applied DenseNet103 based U-Net for segmentation of both the encoder and decoder frame, then the classification process was made using an ensemble model (EfficientNetB0, DenseNet121, VGG-19), they used 6581 X-ray images, the average IoU score was 0.90% and Dice coefficient of 0.92% finally the classification accuracy of COVID-19 was 99.2%.

Some researchers used the U-Net embedded with their proposed model [13],[14],[18],[45]. A channel-shuffled dual-branched architecture called CSDB with Distinctive Filter Learning (DFL) was proposed by [13]. The model contained a pre-trained UNet2D_4L for segmentation. They got accuracy, sensitivity, specificity, precision, f1-score, and AUC 97.94%, 97.54%, 99.25%, 96.34%, 96.90%, and 98.39% respectively. [14] proposed a High-Resolution Network (HRNet) for feature extraction embedding with the U-Net for segmentation, where different types of convolutional resolutions are linked in a parallel manner, as the first step of their work they trained the UNet model with the MC dataset. They used the generated model weights to get segmented lung images from the COVID dataset. They applied a region-based threshold to distinguish the pixels that would give the best segmentation lung area. They also utilized COVID and non-COVID-19 X-Ray datasets. Their results confirmed 99.26% accuracy, 98.53% sensitivity, and 98.82% specificity. On the other hand, on applying the experiments on plain X-ray images using HRNet. they achieved a higher accuracy of 100%.

In addition, segmentation research in [18] was made using the CT-Scan images. It has proposed a mechanism for the automatic segmentation of COVID-19 regions in CT images using a SegNet-based network with an attention gate (AG), the dataset contains 473 CT scan images and 100 CT with its ground truth, and the segmentation procedure was embedded in the proposed model, they obtained sensitivity, specificity, and dice scores of 92.73%, 99.51%, and 89.61%, respectively.

A Dual-branch combination network (DCN) based on U-net was proposed by [45]. The framework is a lung segmentation network to extract accurate lung regions and then classify the CT-scan images into Covid or Non-COVID-19 patients. A total of 1918 CT scans were used in their study. The findings indicated a slice-level accuracy of 95.99% an individual-level accuracy of 96.74%, a slice-level AUC of 0.9755, and an individual-level AUC of 0.9864.

Table 2. illustrates a summary of approaches for detecting COVID-19 using Lung segmentation models. Most of the previous studies have focused on classifying their datasets into two classes: COVID-19 and non-COVID-19. While other research classified the dataset into three and four classes. The work in [6] classified their dataset into three classes Covid, Non-COVID-19, and Normal. Another study has classified the dataset into four classes: COVID-19, Normal, Bacterial Pneumonia, and Viral Pneumonia [13]. This current research will classify the dataset into four classes: COVID-19, Viral Pneumonia, Normal, and lung Opacity.

TABLE 2
DL-based methods for detection of COVID-19 images using Lung segmentation models.

References	Strategy	Dataset	Classification	Acc. (%)	Sens. (%)	Spec. (%)	Pre. (%)	F1-score (%)
[6]	Modified U-Net. Classifier is DenseNet201	COVQU Dataset of 18,479 CXR	3-classes	95.11	94.56	95.59	94.55	94.53
[13]	UNet2D_4L within CSDB-DFLCNN model	Collection of datasets of 12,393 X-ray images.	4-classes	97.94	97.54	99.25	96.34	96.90
[18]	SegNet-based network with an attention gate (AG), U-Net embedded with the model.	473 CT images. and 100 CT with its GT	2-classes	-	92.73	99.51	-	89.61
[44]	DenseNet103 based U-Net	6581 X-ray images	2-classes	99.2	-	-	-	-
[45]	DCN- with U-Net	1918 CT scans	2-classes	92.87	92.86	92.91	-	-

[46]	HRNet-with U-Net	Segmentation Dataset: MC Classification: 2472 X-ray images	2-classes	99.26	98.53	98.82	-	99.25

In both previous methods -whether with or without segmentation- the results lack generalizability due to the small dataset sizes and the limited number of lung disease classes, typically fewer than four. The highest accuracy obtained when classifying plain medical images was 97.94 with an F1-score of 98.80, However, this was based on a dataset of only 3,487 CXR images classified into three categories [39]. In contrast, using the segmentation approach, the highest accuracy reached 99.26 with an F1-score of 99.25, but this dataset contained just 2,472 CXR images classified into only two categories [46]. In this research, we aim to compare both methods using a significantly larger dataset of 21,165 CXR images spanning four distinct classes, providing a more comprehensive evaluation.

3. Methods

3.1. Approach

Fig. 1. illustrates the proposed approach, which involves the use of two different datasets: the Lung segmentation dataset and the Lung classification dataset. The major experiments that are carried out in this study are:

1- Comparing the performance of three different segmentation neural networks (FCN, U-Net, and LinkNet) to choose the best architecture to make segmentation of the dataset.

2- Proposing a preprocessing pipeline to enhance the quality of the X-ray images.

3- Evaluating five pre-trained deep-learning networks (Xception, VGG16, DenseNet201, ResNet50V2, and MobileNetV2) for lung disease classification using raw X-ray images that are not segmented, as well as using images that have been segmented using the best segmentation model along with the calculation of different performance metrics to evaluate the performance of the networks.

Pseudo-codes are illustrated in Fig. 2. And 3.

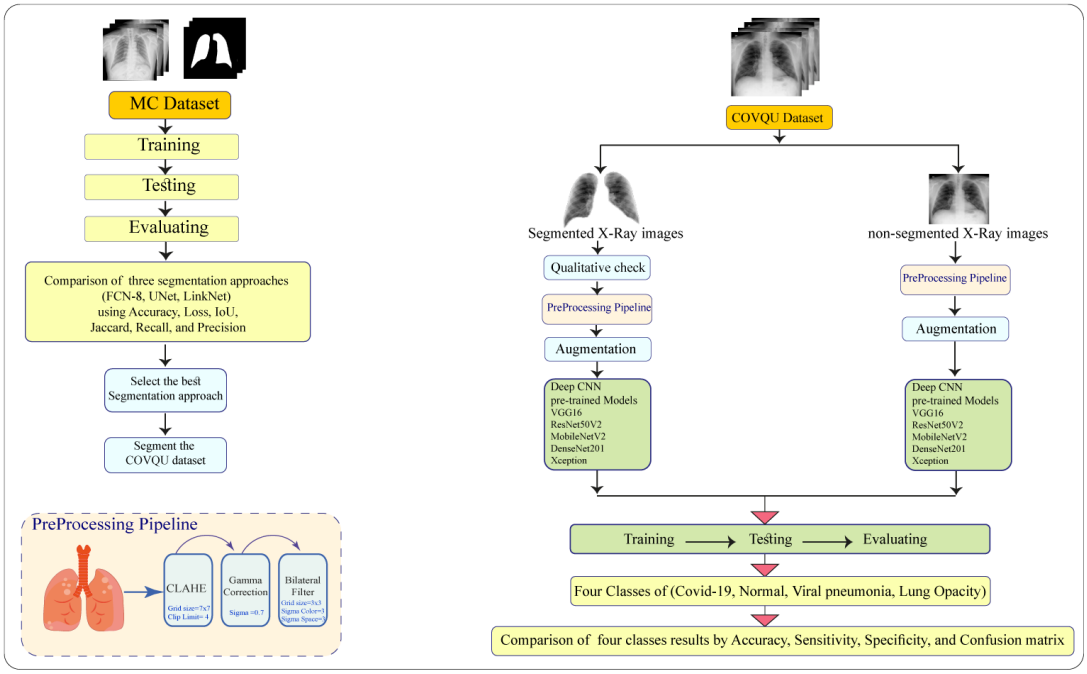

Fig.1: Diagram of Research Methodology

Pseudo-Code #1: Non-Segmentation Approach

1- Input: COVQU 2nd update dataset, contains 21,165 CXR images.

2- Preprocessing:

2.1. Resize images to fit CNN models (Xception, VGG16, DenseNet201, ResNet50V2, and MobileNetV2).

2.2. Preprocessing pipeline.

2.3. Data augmentation.

3- Split train data 80%, test 10%, and validation 10%.

4- Fine-tuning to the pre-trained CNN models.

5- Classify images.

Fig.2: Pseudo-Code for Non-Segmentation Approach

Pseudo-Code #2: Segmentation Approach
1- Input: (MC) CXR Datasets of 704 images with the corresponding ground truth.
2- Split train data 80%, test 10%, and validation 10%.
3- Train each model as follows:
3.1. U-Net from scratch.
3.2. Fine-tuning to the models: (FCN) using pre-trained VGG16, and (LinkNet) provided by Keras
4- Compare between previous models by performance metrics.
5- Select the best segmentation approach.
6- Input: COVQU 2nd update dataset, contains 21,165 CXR images.
7- Segment COVQU 2nd update dataset.
8- Qualitative check of the resulting images.
9- Go to step 2 in pseudo-code#1 to continue.

Fig.3: Pseudo-Code for Segmentation Approach

3.2. Dataset's description

Two datasets are used. The first dataset is used to train, evaluate, and compare the segmentation approaches, and the second dataset is used for COVID-19 detection.

3.2.1. Dataset for Lung Segmentation

To investigate lung segmentation models the study has used NLM-The Shenzhen and Montgomery County (MC) CXR Datasets available by Kaggle [47] which consists of 704 CXR images with the corresponding lung masks. The Dataset has been used in many articles [17],[19],[48],[49],[50] for training and testing different segmentation models, where annotation was made under the supervision of expert radiologists; the dataset contains both normal and abnormal X-ray lung images that are affected by tuberculosis. A sum of 360 images used are normal and 344 are infected lung images. Table 3 summarizes this information.

Table 3. Datasets -lung segmentation

NLM-The Shenzhen and Montgomery County (MC) CXR Datasets	Normal Lungs	Abnormal Lungs (tuberculosis disease)
704 X-ray images and masks	360	344

3.2.2. Dataset for image classification

This research uses the COVID-19 CHEST X-RAY DATABASE [6],[39] which is known as COVQU 2nd update dataset, it is one of the largest public COVID-19 positive cases datasets. It contains 21,165 CXR images of four classes; Normal, COVID-19, Pneumonia, and Lung Opacity disease, The dataset is a combination of the Radiological Society of North America (RSNA) CXR dataset and COVID-19 dataset. The details of this dataset are in Table 4.

Table 4. Datasets -image classification

Lung Conditions	Number of images
COVID-19	3616 CXR images
Lung- Opacity	6012 CXR images
Normal lung	10192 CXR images
Pneumonia	1345 CXR images
Summation	***21,165 CXR images***

3.3. Pre-processing and data augmentation

3.3.1. Pre-processing

Images in the classification dataset (COVQU) were resized to fit the input size of the pre-trained models. For the Xception model, images are resized into 299x299, for other models' images are resized into 244x244. Then the proposed pre-processing pipeline was used to enhance the quality of X-ray images. All original CXR images are in Portable Network Graphics (PNG) format of size 299x299x3. Fig. 4. shows the proposed pre-processing pipeline. First, CLAHE is applied for contrast enhancements followed by Gamma correction for further enhancement of low contrast in X-ray images. Then noise reduction is applied by using a Bilateral filter. An example is presented in Fig. 5, showing an X-ray image before and after the application of each process, along with the corresponding histogram.

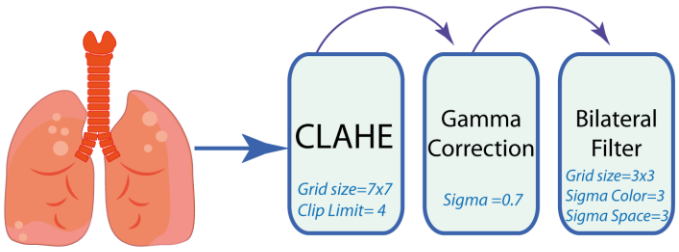

Fig. 4: pipeline of the pre-processing phase.

For the segmentation dataset, images were resized to 256x256 to fit the input size of U-Net and LinkNet architectures. For the FCN model images were resized to 128x128. The normalization step is made by a division of 255.0 to all models. The corresponding ground truth is a black-and-white image, where black is given the value of zero, representing the background, and white has the value of one, which represents the region of interest.

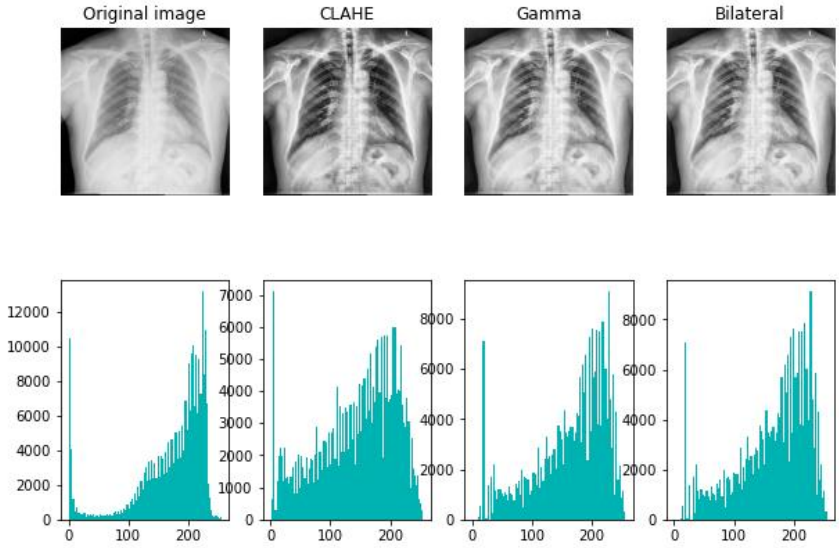

Fig.5: an example of an X-ray image going through the image processing pipeline

3.3.2. Data Augmentation

Deep learning techniques require huge amounts of images to be trained, lack of medical images will cause unbalanced data that will affect the algorithm's accuracy and loss [5],[6],[9],[12],[16],[22],[23],[26],[35],[40],[43]. Therefore, the Keras ImageDataGenerator class is used in this research to generate thousands of images during training time. For the UNet, FCN, and LinkNet, we applied the parameters: Rotation_range=25, Width_shift_range=0.2, Height_shift_range=0.2, Zoom=0.15, Horizontal_flip=True, Fill_mode= Reflect. For classification experiments, the same parameters except the Fill_mode= Nearest.

3.4. Experimental Setup

Datasets are split into the following proportions: 80% for training, 10% for validation, and 10% for testing. Tables 5 and 6 explain the details of training, validation, and test X-ray images for both datasets. All experiments conducted in this study are carried out using TensorFlow version 2.6.0, Python 3.8.8, Processor 11th Gen Intel(R) Core (TM) i7-11800H @ 2.30GHz, RAM 16.0 GB, Intel(R) UHD Graphics, NVIDIA GeForce RTX 3060 GPU. Each of our experiments will be discussed separately.

Table 5. Details of the segmentation dataset used for training, validation, and testing

Total Dataset	Training Set	Number of Augmented images	Resulted Training set	Validation set	Testing set
704 CXR images	563	55350	55913	70	71

Table 6: Details of the classification dataset used for training, validation, and testing

Total Dataset	Types	Training Set	Number of Augmented images	Resulted Training set	Validation set	Testing set
21,165 CXR images	COVID-19	2893	5261	8154	361	363
	Normal	8154	0	8154	1019	1020
	Viral pneumonia	1076	7078	8154	134	136
	Lung Opacity	4810	3344	8154	601	602

3.5. Experiments on Lung Segmentation

Three state-of-the-art semantic segmentation architectures are applied on the MC Dataset: (FCN) using pre-trained VGG16, (U-Net) trained from scratch, and (LinkNet) provided by Keras. FCN8, U-Net, and LinkNet models were trained for 100 epochs, no early stopping, using Adam optimizer, with the learning rate of 1e-4, batch size of 4, the loss function is Binary Focal Loss of gamma=2, Batch normalization was used in U-Net architecture.

3.6. Experiments on Lung Classification

The transfer learning approach is investigated for classification experiments for segmented and plain images. For this purpose, CNNs are applied (Xception, VGG16, ResNet50V2, MobileNetV2, and DenseNet201). The models are fine-tuned and trained for 50 epochs, with a batch size of 4, Stochastic Gradient Descent (SGD) optimizer with exponential decay for better convergence during training and validation, the learning rate starts with 1e-5, Momentum 0.9, Categorical-Cross-Entropy loss function, no early stopping was investigated.

3.7. Performance evaluation matrix

3.7.1. Evaluation matrix of lung segmentation

To compare the segmentation approaches, we used the same input images with the corresponding ground-truth masks. We used the same augmentation parameters for the training phase to be fed to our networks (FCN8, U-Net, LinkNet). The performance of each network was evaluated after the completion of the training and validation phase. In addition to the processing time, we used five performance metrics: accuracy, Intersection-Over-Union (IoU) or the Jaccard Index [17] [19] [29] [51], Dice Coefficient (or F1-score), Recall (or Sensitivity), Precision, and Mean IoU. All the metrics are shown in equations 1-5.

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

$$Jaccard\ Index\ (IoU) = \frac{(TP)}{(TP + FN + FP)} \quad (2)$$

$$Dice\ Coefficient\ (F1\text{-}score) = \frac{(2*TP)}{(2*TP + FN + FP)} \quad (3)$$

$$Recall\ (Sensitivity) = \frac{TP}{TP+FN} \quad (4)$$

$$Precision\ (Pre) = \frac{TP}{TP+FP} \quad (5)$$

Where TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives, respectively.

3.7.2. Evaluation matrix of COVID-19 Classification

The test sets are evaluated using several metrics, namely Total Accuracy, F1-score, Sensitivity, and Precision, as outlined in equations 1, 3, 4, and 5, respectively [8],[12],[22],[23],[28]. In these equations, TP corresponds to correctly predicting positively labeled images as COVID-19, and FP corresponds to cases where the model predicts COVID-19 when the image is labeled as Normal, Pneumonia, or Opacity. TN signifies the correct prediction of negative images labeled as Normal, Pneumonia, or Opacity. FN, on the other hand, is the scenario where the model wrongly predicts a positive label (COVID-19) as a non-COVID image. The confusion matrix is employed to assess the performance of Neural Networks classification approaches, considering both segmented and non-segmented scenarios.

4. Results

This section describes the performance comparisons of the lung segmentation models, then it will discuss the classification networks' performance on the segmented and non-segmented X-ray images.

4.1. Segmentation results

To compare the three neural networks during training, we plot the loss and accuracy curves for each model, the results are shown in Fig. 6. and Table 7. Fig. 7. shows a sample of the resulting segmentation for the three models.

4.2. Classification results

This section presents the experimental results evaluating the performance of the two classification methods, both with and without segmentation.

4.2.1. COVID-19 classification without segmentation

The confusion matrices for the classification without segmentation are illustrated in Fig. 8. Metrics are shown in Table 8.

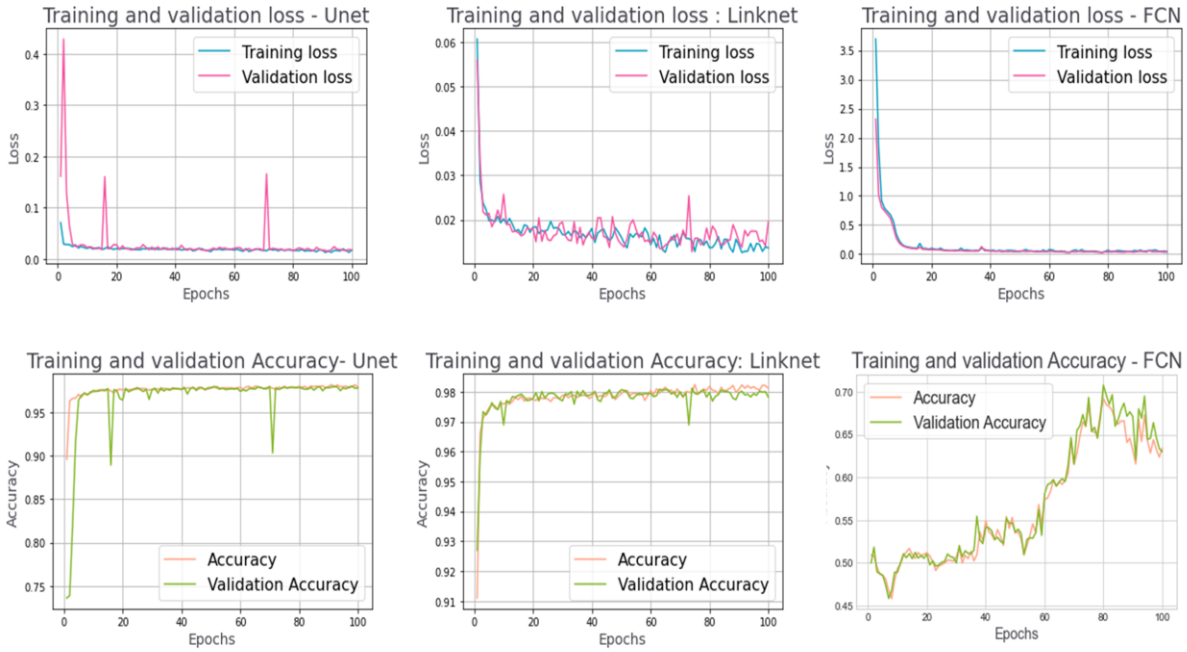

Fig. 6: Training and validation accuracies and losses for LinkNet, U-Net, and FCN8 neural networks

Table 7. Comparison results of the segmentation

Network	*Acc %*	*Jaccard (IoU)%*	*Dice Coef (F1) %*	*Mean IoU %*	*Recall %*	*Pre %*	*Training time (H)*	*Testing time (sec)*
FCN8	96.75	87.72	93.38	91.72	94.57	92.51	**5:30:46.59**	**2.78**
U-Net	97.68	90.96	95.10	93.96	**96.27**	94.32	6:06:29.74	6.58
LinkNet	**98.13**	**92.64**	**96.09**	**96.49**	95.92	**95.09**	5:39:25.18	5.10

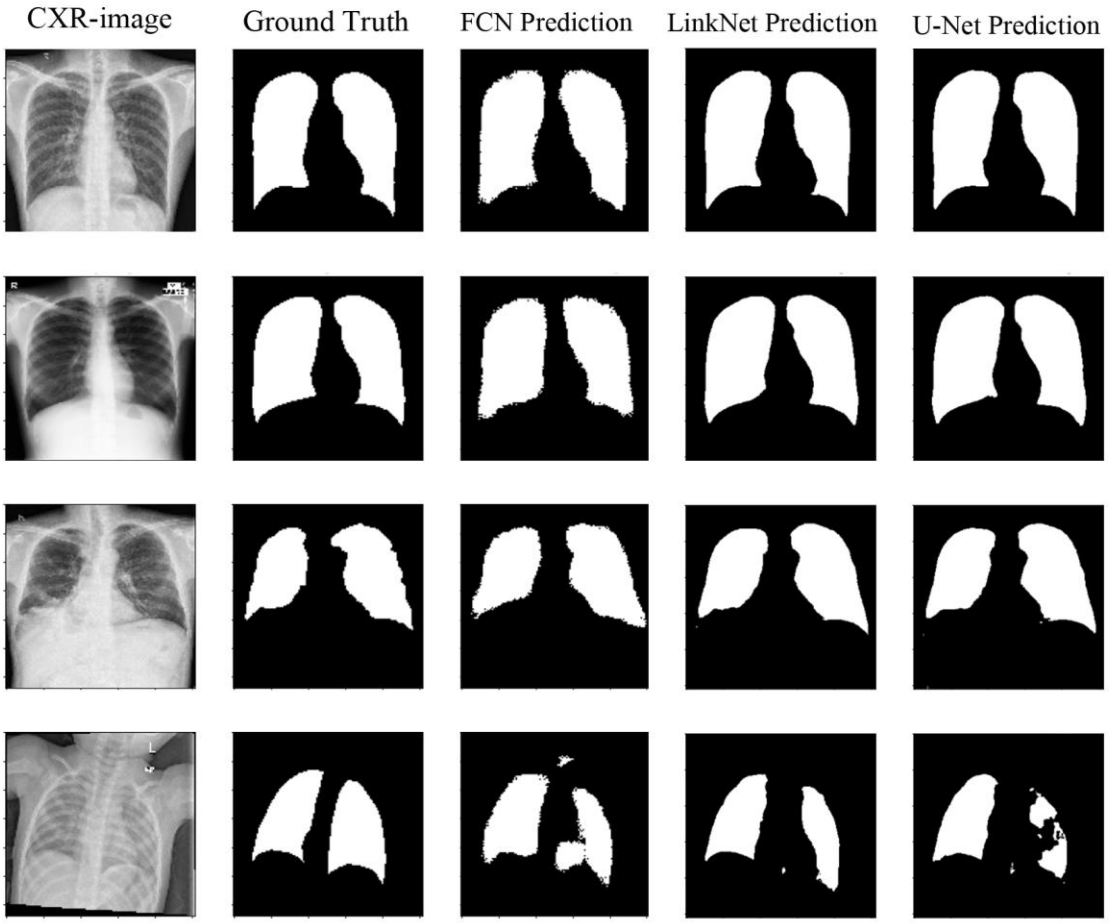

Fig. 7: X-ray image, then the ground truth of the image, and then the predicted result of FCN8, U-Net, and LinkNet respectively

4.2.2. COVID-19 classification using segmentation

Segmentation could play an important role in medical image classification since it captures the desired part of the image (region of interest). Consequently, the CNN model could learn from this part of the image [6]. In this part of the experiment, we will use the LinkNet segmentation model to segment the lungs of the COVQU dataset. To the best of our knowledge, this is the first study that utilizes LinkNet to classify lung diseases using a segmentation approach. Fig. 9. Shows samples of the resulting segmentations of the COVQU dataset using LinkNet.

For the classification tests of the segmented images, a preprocessing step was conducted utilizing our proposed pipeline. Subsequently, the segmented images were divided into training, validation, and testing datasets. Data augmentation was applied exclusively to the training set, followed by classification using pretrained models. (Xception, VGG16, ResNet50V2, MobileNetV2, and DenseNet201).

Fig. 10. Presents the confusion matrices generated from the classification of four classes using pretrained neural networks applied to the segmented X-ray images processed by

Table 8. The Total accuracy, F1 score, recall, and precision for the classification models on plain images

Model	Acc	F1 Score	Recall	Precision	Total Training Time (Hours)	Total Testing Time (Sec)
Xception	90.08	90.03	90.08	90.08	3:16:32.2	00:13.4
VGG16	90.18	90.24	90.18	90.41	2:24:49.6	00:09.8
ResNet50V2	88.2	88.18	88.2	88.18	2:00:48.4	00:07.9
MobileNetV2	89.76	89.78	89.76	89.81	**1:14:44.6**	**00:05.1**
DenseNet201	**91.32**	**91.30**	**91.32**	**91.32**	4:23:54.3	00:19.0

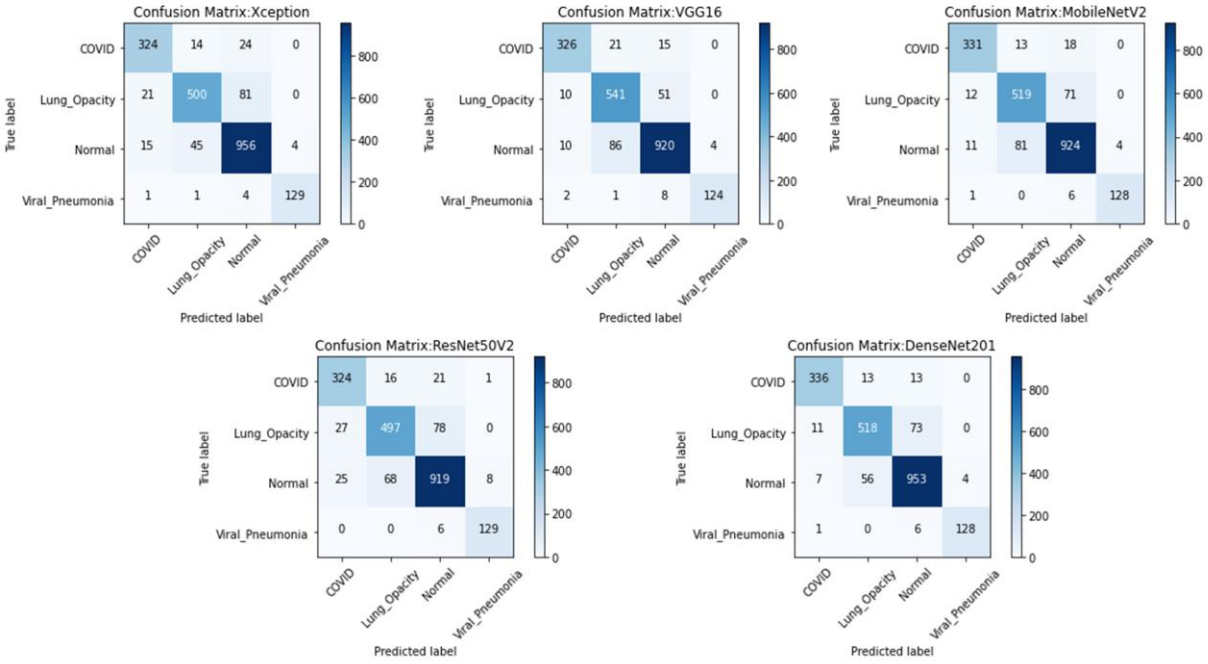

Fig.8: The confusion matrices for the classification without segmentation

LinkNet. Additionally, Table 9 summarizes the performance results of each pretrained model on the segmented X-ray images.

5. Discussion

The following discussion section aims to interpret the results of the previous experimental findings, situating them within the broader research landscape while identifying critical gaps in the field. It also addresses the study's limitations and proposes impactful directions for future research.

Fig. 9: **LinkNet** segmentation results for the COVQU dataset, where A: belongs to COVID-19 X-ray images, B: belongs to Lung opacity, C: belongs to Normal, and D: belongs to Pneumonia X-ray images.

Table 9. The Total accuracy, F1 score, recall, and precision for the classification models using segmented COVQU dataset by **LinkNet**.

Model	Acc	F1 Score	Recall	Precision	Total Training Time (Hours)	Total Testing Time (Sec)
Xception	84.35	84.28	84.35	84.25	3:33:54.6	00:12.5
VGG16	84.16	83.96	84.16	83.92	2:11:02.6	00:08.6
ResNet50V2	83.26	83.22	83.26	83.22	1:58:34.5	00:07.1
MobileNetV2	83.45	83.51	83.58	83.45	**1:11:14.0**	**00:04.2**
DenseNet201	**85.11**	**84.78**	**85.11**	**84.94**	4:23:03.9	00:17.4

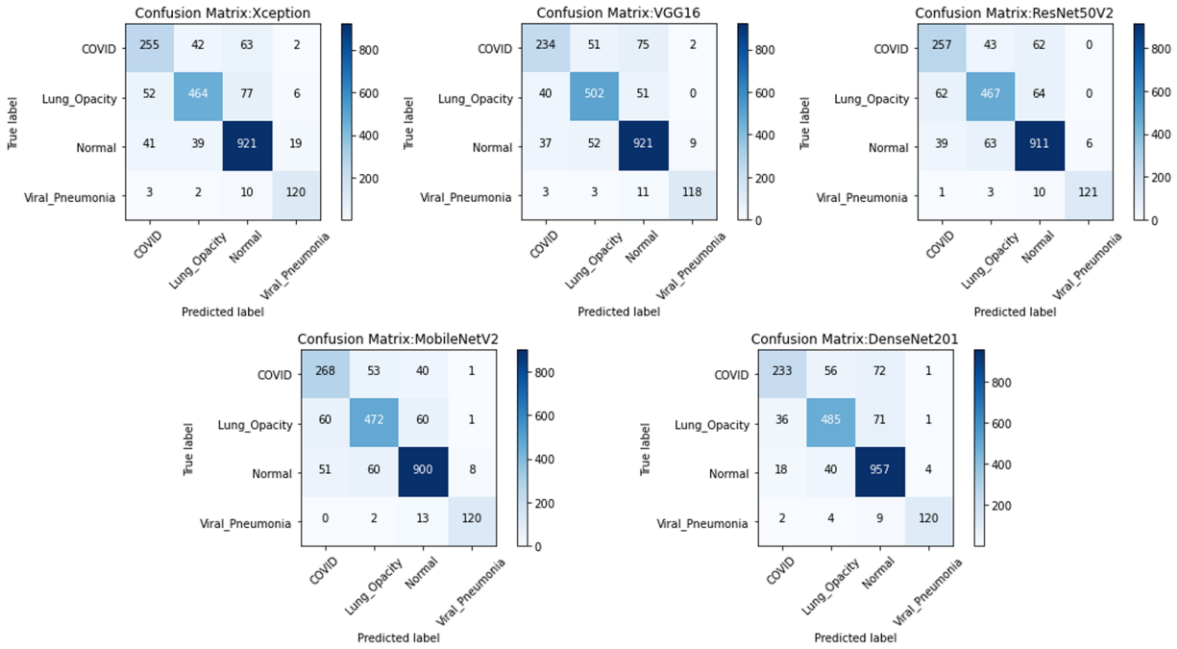

Fig. 10. Confusion matrices for the segmented COVQU dataset using the **LinkNet**, classification by pretrained neural networks.

5.1. Segmentation results Discussion

It is evident that LinkNet, developed by Keras, delivered superior performance compared to the other networks in terms of accuracy, Jaccard index, Dice coefficient, Mean IoU, and Precision of 98.13, 92.64, 96.09, 96.49, 95.09 respectively. On the other hand, U-Net exhibited the highest sensitivity at 96.27%, while FCN8 showed the shortest training and testing time. Both LinkNet and U-Net have better results. They are converging in both the training/validation accuracy and loss curve, while FCN8 has converged in the loss only, but its accuracy is highly swinging.

The children's X-ray images were not well segmented, as the process of taking these images is not an easy task for radiologists, which may affect the quality and clarity of the image.

A comparative studies have been conducted in the literature to determine the most robust segmentation model. Researchers in [14] compared FCN, U-Net, and DeepLabv3 for person segmentation using a top view dataset, the study revealed that FCN achieved an intersection over union (IoU) of 83%, mean IoU (mIoU) of 80%, and pixel accuracy of 91%, U-Net achieved an IoU of 84%, mIoU of 82%, and pixel accuracy of 92%. DeepLabv3 achieved an IoU of 86%, mIoU of 84%, and pixel accuracy of 93%, additionally, FCN was found to be faster than U-Net and DeepLabv3 in terms of computational performance for both CPU and GPU. A comparison study between UNet, SegNet, ENet, and ErfNet was made by [1] using a dataset containing 536 Laryngeal Endoscopic Images which contain 7 classes. Their results were UNet and ErfNet was the best with a mean IoU of 84.7 %.

A comparison of different encoders as the backbone of U-Net in segmenting histopathology Images was made by [10], the highest performance was to EfficientNet-B3 as a backbone, Jaccard Index Dice Coeff. Sensitivity and Specificity were 0.95%, 0.97%, 0.99%, and 0.99%

respectively. ENet was compared with Fully convolutional networks FCN8s, FCN16s, and FCN 32s [52], they implemented semantic segmentation to derive the contours of the tea rows and identify the obstacles in the field scene. The result of the experiments was that ENet outperformed other models with a mean IOU of 0.734 and an accuracy of 0.941. In our comparisons, we obtained the highest accuracy using the LinkNet model, so we adopted it in the segmentation process to segment the images of the COVQU dataset.

5.2. Classification results Discussion

This section includes the experimental discussion on the performance of the two methods of COVID-19 classification (with and without segmentation).

5.2.1. COVID-19 classification without segmentation

This research found that **DenseNet201** has the highest scores of 91.32%, 91.3%, 91.32%, and 91.32% for Accuracy, F1-Score, Recall, and Precision, respectively. The fastest overall training time and testing time was for **MobileNetV2**.

COVID-19 is a subset of Pneumonia diseases so X-ray images from both diseases may have similarities. This may lead to a decline in accuracy [12],[41]. Literature stipulates that the accuracy of four classes is less than the accuracy of two or three classes. Researchers [12] have got an accuracy of 0.85 to 4 classes with sensitivity, specificity, and AUC 98.994%, 92.190 %, and 59.480 % respectively. The four classes contain four different cases but with fewer datasets than we have. Researchers in [53] have made their experiments on four different classes, which are COVID-19, normal, bacterial pneumonia, and viral pneumonia images. They got results of accuracy 94.79 %. While it was 99.81% for binary class, their dataset contains only 6884 X-ray images.

In our research we have four different cases of the X-ray images; COVID-19, Pneumonia, Normal, and Lung Opacity using a dataset of 21,165 CXR images. The result achieved by DenseNet201 of F1-Score was the highest among the other research works.

Any ML approach must minimize false-negative predictions, especially in the medical context since it may lead to diagnostic and treatment delays that can be dangerous to patient health and will lose confidence relating to medical services. It can also lead to legal consequences. The high number of false negatives for experiments in ML to detect COVID-19 is due to the limited sample size and low image quality of the COVID-19 data sets used for the X-ray images [26]. The main reason behind the missing correct classification is the low quality and resolution of the X-ray images [41]. Good images must have good light focus on both the left and right lobes [39]. Many of these X-ray images were originally in DICOM format (Digital Imaging and Communications in Medicine) that had been converted to PNG or JPEG format and this could cause lower-quality images [17].

5.2.2. COVID-19 classification using segmentation

This research found that some X-ray images were not segmented properly. Some of them were almost black, and some were completely black. Table 10 displays the count of segmentation failures for each segmentation model; the LinkNet model has the lowest failures

among the other models. Instances of failures include cases where the resulting images turned completely black. Additionally, other failures were identified. For example, there were cases where only one lung side was captured in the segmentation result, or where the segmentation result consisted of tiny, isolated pieces.

The segmentation of the COVQU dataset of Normal and Pneumonia X-ray images was better than the segmentation of COVID-19 and Lung-Opacity X-ray images.

Table 10. Number of segmentation failures for each segmentation model to the COVQU dataset where the result was a completely black image.

Lung disease	Segmentation failures by FCN8	Segmentation failures by LinkNet	Segmentation failures by U-Net
COVID-19	13	4	22
Lung- Opacity	38	3	12
Normal lung	5	2	3
Pneumonia	2	2	2
Summation	**58**	**11**	**39**

The good-quality X-ray images that have well-balanced light in the COVQU dataset were segmented well, some of these good samples are shown in Fig. 11-A. On the other hand, we found some images in the COVQU dataset that were not segmented properly, we can summarize the reasons behind that to:

1- Poor quality images and improper lighting resulted in black or semi-black segmented images, as shown in Fig. 11-B.

2- Unbalanced light for the left or right lung or bad postures of the patient will cause segmentation for one lung side or arbitrary parts of the other side. A sample is shown in Fig. 11-C.

3- Bad shooting positions or images that were not taken properly will not be segmented accurately. A sample is shown in Fig. 11-D.

4- Unclear and blurry images will not be segmented accurately. A sample is shown in Fig. 11-E.

5- The existence of some medical equipment or other signs or lines in the image will cause errors in the predicted segmented image. A sample is shown in Fig. 11-F.

6- Some children's X-ray images were not segmented properly, since children and adolescents need alerted imaging approaches. A sample is shown in Fig. 11-G.

For the classification tests of the segmented images, a preprocessing step was conducted utilizing our proposed pipeline. Subsequently, the segmented images were divided into training, validation, and testing datasets. Data augmentation was applied exclusively to the training set, followed by classification using pretrained models. (Xception, VGG16, ResNet50V2, MobileNetV2, and DenseNet201).

The higher accuracy, F1-Score, recall, and precision were obtained by DeneseNet201 of 85.11, 84.78, 85.11, and 84.94 respectively. MobileNetV2 has the lowest training and testing time in all our experiments.

Numerous researchers have indicated that classification conducted without segmentation yielded higher accuracy than approaches that incorporate segmentation [6],[14]. This discrepancy is believed to stem from suboptimal segmentation quality. As illustrated in Fig. 9. and 7, LinkNet occasionally failed to accurately predict the region of interest, resulting in diminished classification accuracy.

Our results without segmentation demonstrated superior performance, achieving an F1-score of 91.30, compared to 84.78 when employing LinkNet for segmentation and DenseNet201 as the classifier in both approaches.

Deep learning models trained on non-medical image datasets struggle with low-resolution X-rays, which hurt the accuracy. Collaborating with radiology experts will be crucial for enhancing and expanding the high-resolution dataset, enabling deeper insights and improving the performance of deep learning algorithms in the medical field.

Fig. 11: segmentation Failures for COVQU dataset, where A: Good quality X-ray images, B: Poor images, C: Unbalanced light, D: Bad shooting positions, E: Unclear and blurry images F: medical equipment or other signs, G: children X-ray images.

6. Conclusions

In this research work, different deep learning-based semantic segmentation models FCN8, U-Net, and LinkNet have been investigated for lung segmentation to find out the segmentation effect on multi-class classification for lung disease. U-Net architecture has been trained from scratch, while FCN8 has been trained from scratch using pretrained VGG16 as a backbone, LinkNet which is a pre-trained model provided by Keras was investigated in this comparison.

Results show that LinkNet by Keras outperformed the other segmentation networks in terms of accuracy, Jaccard, Dice, Mean IoU, and Precision of 98.13, 92.64, 96.09, 96.49, 95.09 respectively, while U-Net has the higher sensitivity of 96.27. FCN8 has the lowest training and testing time.

The research adopted the LinkNet to segment the COVQU 2nd update dataset which contains 21,165 CXR images then used five pretrained architectures (Xception, VGG16, ResNet50V2, MobileNetV2, and DenseNet201) to make a 4-class classification (COVID-19, Viral Pneumonia, Normal and lung Opacity). We found that the accuracy of non-segmented X-ray images outperforms the segmented. We have provided a detailed analysis to explain this result by studying the quality of images that were not segmented well. Therefore, we highly recommend that researchers select high-quality medical images whether for classification or segmentation experiments.

The best classification results of non-segmented images were obtained by DenesNet201 of accuracy, F1-Score, recall, and precision 91.32, 91.30, 91.32, 91.32. while MobilNetV2 has the lowest training and testing time for all our experiments in this research. The best classification accuracy for segmented X-ray images was achieved using the LinkNet segmentation model and the DenseNet201 classifier, with accuracy, F1-Score, recall, and precision values of 85.11%, 84.78%, 85.11%, and 84.94%, correspondingly.

Deep learning models leveraging transfer learning, which can achieve remarkable results, are typically trained on vast datasets of non-medical images. However, their performance encounters significant challenges, particularly when processing low-resolution or blurry X-ray images, which can severely affect accuracy. Consequently, it is crucial for radiology specialists to meticulously evaluate the datasets employed by researchers in artificial intelligence to ensure reliability and effectiveness.

One of the primary limitations of this research is the low resolution of many images in the COVQU dataset. This issue poses significant challenges in achieving accurate segmentation of the lung region. In the future, we aim to address this limitation by acquiring a high-resolution dataset and leveraging hybrid deep-learning models for training. Additionally, we plan to implement a cross-validation approach to rigorously evaluate the performance and robustness of the future model.

References

[1] M. H. Laves, J. Bicker, L. A. Kahrs, and T. Ortmaier, "A dataset of laryngeal endoscopic images with comparative study on convolution neural network-based semantic segmentation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 3, pp. 483–492, 2019,10.1007/s11548-018-01910-0

[2] G. Mufarah, M. Alshmrani, Q. Ni, R. Jiang, H. Pervaiz, and N. M. Elshennawy, "ORIGINAL ARTICLE A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images," *Alexandria Eng. J.*, vol. 64, pp. 923–935, 2023,10.1016/j.aej.2022.10.053

[3] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, "Medical image classification with convolutional neural network," *2014 13th Int. Conf. Control Autom. Robot. Vision, ICARCV 2014*, vol. 2014, no. December, pp. 844–848, 2014,10.1109/ICARCV.2014.7064414

[4] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," *Proc. 2017 Int. Conf. Eng. Technol. ICET 2017*, vol. 2018-Janua, pp. 1–6, 2018,10.1109/ICEngTechnol.2017.8308186

[5] A. Bria, C. Marrocco, and F. Tortorella, "Addressing class imbalance in deep learning for small lesion detection on medical images," *Comput. Biol. Med.*, vol. 120, no. March, p. 103735, 2020,10.1016/j.compbiomed.2020.103735

[6] T. Rahman *et al.*, “Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images,” *Comput. Biol. Med.*, vol. 132, no. March, p. 104319, 2021,10.1016/j.compbiomed.2021.104319

[7] W. Gómez-Flores and W. Coelho de Albuquerque Pereira, “A comparative study of pre-trained convolutional neural networks for semantic segmentation of breast tumors in ultrasound,” *Comput. Biol. Med.*, vol. 126, p. 104036, 2020,10.1016/j.compbiomed.2020.104036

[8] M. Z. Islam, M. M. Islam, and A. Asraf, “A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images,” *Informatics Med. Unlocked*, vol. 20, p. 100412, 2020,10.1016/j.imu.2020.100412

[9] A. Wagh *et al.*, “Semantic segmentation of smartphone wound images: Comparative analysis of AHRF and CNN-based approaches,” *IEEE Access*, vol. 8, pp. 181590–181604, 2020,10.1109/ACCESS.2020.3014175

[10] A. Riasatian, M. Rasoolijaberi, M. Babaei, and H. R. Tizhoosh, “A Comparative Study of U-Net Topologies for Background Removal in Histopathology Images,” *Proc. Int. Jt. Conf. Neural Networks*, 2020,10.1109/IJCNN48605.2020.9207018

[11] T. Sakinis *et al.*, “Interactive segmentation of medical images through fully convolutional neural networks,” no. v, pp. 1–10, 2019.

[12] M. Umer, I. Ashraf, S. Ullah, A. Mehmood, and G. S. Choi, “COVINet: a convolutional neural network approach for predicting COVID-19 from chest X-ray images,” *J. Ambient Intell. Humaniz. Comput.*, vol. 13, no. 1, pp. 535–547, 2022,10.1007/s12652-021-02917-3

[13] R. Karthik, R. Menaka, and M. Hariharan, “Learning distinctive filters for COVID-19 detection from chest X-ray using shuffled residual CNN,” *Appl. Soft Comput.*, vol. 99, no. xxxx, p. 106744, 2021,10.1016/j.asoc.2020.106744

[14] I. Ahmed, M. Ahmad, F. A. Khan, and M. Asif, “Comparison of deep-learning-based segmentation models: Using top view person images,” *IEEE Access*, vol. 8, pp. 136361–136373, 2020,10.1109/ACCESS.2020.3011406

[15] A. Chaurasia and E. Culurciello, “LinkNet: Exploiting encoder representations for efficient semantic segmentation,” *2017 IEEE Vis. Commun. Image Process. VCIP 2017*, vol. 2018-Janua, pp. 1–4, 2018,10.1109/VCIP.2017.8305148

[16] S. Walvekar and S. Shinde, “Efficient medical image segmentation of COVID-19 Chest CT images based on deep learning techniques,” *2021 Int. Conf. Emerg. Smart Comput. Informatics, ESCI 2021*, pp. 203–206, 2021,10.1109/ESCI50559.2021.9397043

[17] T. Rahman *et al.*, “Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization,” *IEEE Access*, vol. 8, no. July, pp. 191586–191601, 2020,10.1109/ACCESS.2020.3031384

[18] Ü. Budak, M. Çıbuk, Z. Cömert, and A. Şengür, “Efficient COVID-19 Segmentation from CT Slices Exploiting Semantic Segmentation with Integrated Attention Mechanism,” *J. Digit. Imaging*, vol. 34, no. 2, pp. 263–272, 2021,10.1007/s10278-021-00434-5

[19] R. Arora, I. Saini, and N. Sood, “Multi-label segmentation and detection of COVID-19 abnormalities from chest radiographs using deep learning,” *Optik (Stuttg).*, vol. 246, no. August, p. 167780, 2021,10.1016/j.ijleo.2021.167780

[20] P. Kalane, S. Patil, B. P. Patil, and D. P. Sharma, “Automatic detection of COVID-19 disease using U-Net architecture based fully convolutional network,” *Biomed. Signal Process. Control*, vol. 67, no. February, p. 102518, 2021,10.1016/j.bspc.2021.102518

[21] H. Reza Tizhoosh and L. Pantanowitz, “Artificial intelligence and digital pathology: Challenges and opportunities,” *J. Pathol. Inform.*, vol. 9, no. 1, 2018,10.4103/jpi.jpi_53_18

[22] S. Hira, A. Bai, and S. Hira, “An automatic approach based on CNN architecture to detect Covid-19 disease from chest X-ray images,” *Appl. Intell.*, vol. 51, no. 5, pp. 2864–2889, 2021,10.1007/s10489-020-02010-w

[23] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, “Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network,” *Appl. Intell.*, vol. 51, no. 2, pp. 854–864, 2021,10.1007/s10489-020-01829-7

[24] K. Dev, S. A. Khowaja, A. S. Bist, V. Saini, and S. Bhatia, “Triage of potential COVID-19 patients from chest X-ray images using hierarchical convolutional networks,” *Neural Comput. Appl.*, vol. 9, no. November 2020, 2021,10.1007/s00521-020-05641-9

[25] A. G. U. Juarez, H. A. W. M. Tiddens, and M. de Bruijne, "Automatic airway segmentation in chest CT using convolutional neural networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11040 LNCS, pp. 238–250, 2018,10.1007/978-3-030-00946-5_24

[26] M. J. Horry *et al.*, "COVID-19 Detection through Transfer Learning Using Multimodal Imaging Data," *IEEE Access*, vol. 8, pp. 149808–149824, 2020,10.1109/ACCESS.2020.3016780

[27] K. H. Shibly, S. K. Dey, M. T. U. Islam, and M. M. Rahman, "COVID faster R–CNN: A novel framework to Diagnose Novel Coronavirus Disease (COVID-19) in X-Ray images," *Informatics Med. Unlocked*, vol. 20, 2020,10.1016/j.imu.2020.100405

[28] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 635–640, 2020,10.1007/s13246-020-00865-4

[29] S. Hassantabar, M. Ahmadi, and A. Sharifi, "Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches," *Chaos, Solitons and Fractals*, vol. 140, p. 110170, 2020,10.1016/j.chaos.2020.110170

[30] Z. Nabizadeh-Shahre-Babak, N. Karimi, P. Khadivi, R. Roshandel, A. Emami, and S. Samavi, "Detection of COVID-19 in X-ray images by classification of bag of visual words using neural networks," *Biomed. Signal Process. Control*, vol. 68, no. September 2020, p. 102750, 2021,10.1016/j.bspc.2021.102750

[31] S. V. Mohd Sagheer and S. N. George, "A review on medical image denoising algorithms," *Biomed. Signal Process. Control*, vol. 61, p. 102036, 2020,10.1016/j.bspc.2020.102036

[32] M. Agarwal and R. Mahajan, "Medical Image Contrast Enhancement using Range Limited Weighted Histogram Equalization," *Procedia Comput. Sci.*, vol. 125, no. 2017, pp. 149–156, 2018,10.1016/j.procs.2017.12.021

[33] Z. Al-Ameen, "Contrast Enhancement of Medical Images Using Statistical Methods with Image Processing Concepts," *Proc. 6th Int. Eng. Conf. ' 'Sustainable Technol. Dev. IEC 2020*, pp. 169–173, 2020,10.1109/IEC49899.2020.9122925

[34] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *J. VLSI Signal Process. Syst. Signal Image. Video Technol.*, vol. 38, no. 1, pp. 35–44, 2004,10.1023/B:VLSI.0000028532.53893.82

[35] M. Heidari, S. Mirniaharikandehei, A. Z. Khuzani, G. Danala, Y. Qiu, and B. Zheng, "Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms," *Int. J. Med. Inform.*, vol. 144, no. September, p. 104284, 2020,10.1016/j.ijmedinf.2020.104284

[36] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Sci. Rep.*, vol. 10, no. 1, pp. 1–12, 2020,10.1038/s41598-020-76550-z

[37] A. I. Khan, J. L. Shah, and M. M. Bhat, "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images," *Comput. Methods Programs Biomed.*, vol. 196, p. 105581, 2020,10.1016/j.cmpb.2020.105581

[38] E. Hussain, M. Hasan, M. A. Rahman, I. Lee, T. Tamanna, and M. Z. Parvez, "CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images," *Chaos, Solitons and Fractals*, vol. 142, p. 110495, 2021,10.1016/j.chaos.2020.110495

[39] M. E. H. Chowdhury *et al.*, "Can AI Help in Screening Viral and COVID-19 Pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020,10.1109/ACCESS.2020.3010287

[40] M. A. Rab Ratul, M. Tavakol Elahi, K. Yuan, and W. S. Lee, "RAM-Net: A Residual Attention MobileNet to Detect COVID-19 Cases from Chest X-Ray Images," *Proc. - 19th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2020*, pp. 195–200, 2020,10.1109/ICMLA51294.2020.00040

[41] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, vol. 121, no. April, p. 103792, 2020,10.1016/j.compbiomed.2020.103792

[42] G. Dhiman, V. Chang, K. Kant Singh, and A. Shankar, "ADOPT: automatic deep learning and optimization-based approach for detection of novel coronavirus COVID-19 disease using X-ray images," *J. Biomol. Struct. Dyn.*, vol. 0, no. 0, pp. 1–13, 2021,10.1080/07391102.2021.1875049

[43] H. Munusamy, J. M. Karthikeyan, G. Shriram, S. Thanga Revathi, and S. Aravindkumar, "FractalCovNet architecture for COVID-19 Chest X-ray image Classification and CT-scan image Segmentation," *Biocybern. Biomed. Eng.*, vol. 41, no. 3, pp. 1025–1038, 2021,10.1016/j.bbe.2021.06.011

[44] M. Hasan Jahid, M. Alom Shahin, and M. Ali Shikhar, "Deep Learning based Detection and Segmentation of COVID-19 Pneumonia on Chest X-ray Image," *2021 Int. Conf. Inf. Commun. Technol. Sustain. Dev. ICICT4SD 2021 - Proc.*, pp. 210–214, 2021,10.1109/ICICT4SD50815.2021.9396878

[45] K. Gao *et al.*, "Dual-branch combination network (DCN): Towards accurate diagnosis and lesion segmentation of COVID-19 using CT images," *Med. Image Anal.*, vol. 67, p. 101836, 2021,10.1016/j.media.2020.101836

[46] S. Ahmed, T. Hossain, O. B. Hoque, S. Sarker, S. Rahman, and F. M. Shah, "Automated COVID-19 Detection from Chest X-Ray Images: A High-Resolution Network (HRNet) Approach," *SN Comput. Sci.*, vol. 2, no. 4, pp. 1–17, 2021,10.1007/s42979-021-00690-w

[47] "Tuberculosis Chest X-rays (Montgomery) | Kaggle." Accessed: Feb. 20, 2022. [Online]. Available: https://www.kaggle.com/raddar/tuberculosis-chest-xrays-montgomery?select=montgomery_metadata.csv

[48] S. Jaeger *et al.*, "Using Chest Radiographs," *Iee*, no. c, pp. 1–13, 2013.

[49] S. Candemir *et al.*, "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," *IEEE Trans. Med. Imaging*, vol. 33, no. 2, pp. 577–590, 2014,10.1109/TMI.2013.2290491

[50] L. Zhang, A. Liu, J. Xiao, and P. Taylor, "Dual Encoder Fusion U-Net (DEFU-Net) for Cross-manufacturer Chest X-ray Segmentation," *Proc. - Int. Conf. Pattern Recognit.*, pp. 9333–9339, 2020,10.1109/ICPR48806.2021.9412718

[51] B. Shuai, T. Liu, and G. Wang, "Improving Fully Convolution Network for Semantic Segmentation," 2016.

[52] Y. K. Lin and S. F. Chen, "Development of Navigation System for Tea Field Machine Using Semantic Segmentation," *IFAC-PapersOnLine*, vol. 52, no. 30, pp. 108–113, 2019,10.1016/j.ifacol.2019.12.506

[53] R. C. Joshi *et al.*, "A deep learning-based COVID-19 automatic diagnostic framework using chest X-ray images," *Biocybern. Biomed. Eng.*, vol. 41, no. 1, pp. 239–254, 2021,10.1016/j.bbe.2021.01.002