TelLungNet - Enabling Telemedicine Utilizing an Improved U-Net Lung Image Segmentation

Rifat Al Mamun Rudro
Department of Computer Science
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
23-93113-3@student.aiub.edu

Api Alam
Department of Computer Science
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
19-40880-2@student.aiub.edu

Shafin Talukder
Department of Computer Science
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
18-36588-1@student.aiub.edu

Tanvir Ahmed
Department of Computer Science
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
20-42013-1@student.aiub.edu

Nayma Islam
Department of Computer Science
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
18-36607-1@student.aiub.edu

Kamruddin Nur
Department of Computer Science
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
kamruddin@aiub.edu

***Abstract*—This study presents a novel approach to web-based telemedicine services by utilizing and improving the U-Net deep learning architecture. Here we present a user-friendly web application designed for medical professionals to diagnose chest X-ray images easily using the service provided by our proposed TelLung-Net architecture. By using the web application a user can upload their X-ray images, the proposed architecture then segments the image, and instant segmentation results provide visual aids during remote consultation for both patients and medical professionals. In this research, we use 1228 X-Ray images from Mendeley data for training and testing our proposed architecture. Our experiments demonstrate that the proposed TelLung-Net achieves 96.26% accuracy with a 0.95 f1 score in chest X-ray segmentation. Incorporating image segmentation into U-Net deep learning architecture significantly improves the precision and accuracy of identifying issues in X-ray images. It also improves system reliability, and reduces time by identifying chest abnormalities.**

***Index Terms*—Image Segmentation, U-Net architecture, Biomedical Image Processing, Telehealth**

I. Introduction

The integration of telemedicine and artificial intelligence (AI) is revolutionizing healthcare by making it more accessible and dynamic. A major factor propelling the changing healthcare industry forward is incorporating technologies such as telemedicine and artificial intelligence (AI) [24]. This research aims to develop a novel process that combines Python, HTML, CSS, and JavaScript frameworks to build a TensorFlow.js web application. The web application effectively provides segmented chest X-ray pictures in medical healthcare. It also improves the diagnostic process through supervised segmented lung pictures over the original chest X-rays, providing essential visual context for healthcare professionals in one place. The current healthcare landscape has seen a growing need for telemedicine services that are both efficient and easily accessible [3] for both patients and doctors. The need for prompt and precise diagnosis is important and requires modern health technology, Where significant obstacle arises from latency concerns inherent to online implementation [2]. This is particularly crucial when patients demand immediate replies from the telemedicine systems.

The research intends to develop a specialized healthcare model for analyzing chest X-ray images and managing patients' medical documents in one place solution. The proposed model has been effectively incorporated into a web application using the Keras model, specifically offering a user-friendly interface for online deployment. In addition to addressing the technical intricacies of the model, the research lays considerable importance on enhancing the speed of inference to satisfy the expectations of online users for efficient telemedicine services [1].

Figure 1 depicts the first stages of web-based telemedicine services, including image processing techniques designed for individuals suffering from asthma and chest diseases [9]. Patients use the online application and utilize artificial intelligence-enabled filtering to authenticate medical information and its accompanying attachments. The U-Net architecture is used for image processing [13], specifically for analyzing

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00247

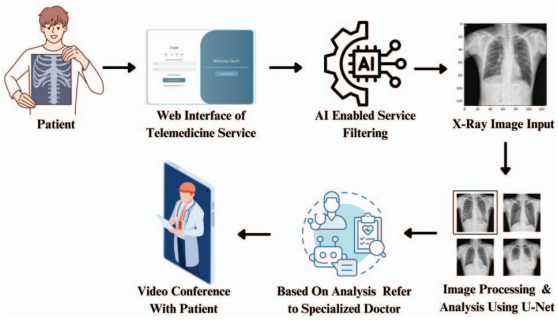

Fig. 1. Web-Enabled Chest X-Ray Image Segmentation using U-Net

and processing pictures linked to the chest. After the analysis, a specialized professional is effectively designated to the patient. This paper also improves the effectiveness and availability of telemedicine services for individuals suffering from chest disease. In today's world, the dynamic healthcare environment has necessitated the development of solutions that enable patients to access and use advanced medical services conveniently.The objectives of the research are:

- **System Implementation**: To create and execute an advanced web-enabled telemedicine platform that provides a smooth and secure virtual space for online consultations designed for individuals with chest diseases.
- **Image Processing with U-Net**: This research utilizes the proposed U-Net architecture to perform image processing tasks, primarily analyzing chest-related pictures by processing medical pictures for initial diagnostic purposes within the domain of chest disease.
- **Specialist Assignment**: Develop a method within the healthcare system to effectively allocate medical specialists to patients by leveraging the results obtained from image processing analysis. This approach expresses timely and focused virtual consultations customized to meet the needs of patients with chest diseases.

The key contributions of this research are:

1) Proposed U-Net architecture for real-time lung X-ray image classification
2) Development of a telemedicine web application for both patients and doctors
3) Integration of U-Net classification engine with the telemedicine web application

The subsequent sections of the paper are structured in the following manner: Section II provides the preceding related works. With numerous figures and facts, Section III provides a thorough description of the methodology. The proposed model is explained in Section IV with some figures and equations. The proposed model outcome is discussed in Section V, and Section VI is the conclusion.

II. Literature Review

Cindy et al. [5] explored the effectiveness of U-Net-based semantic segmentation approaches for lung nodule detection, a critical aspect in combating the precarious nature of lung cancer in humans. The multi-step process involved dataset preprocessing, employing architectures such as U-Net2D, R2U-Net2D, U-Net++, and Attention U-Net, and training the models with rigorous evaluation metrics [16]. U-Net2D performed exceptionally well with 99.38% accuracy, a mean IOU of 74.34%, and a low binary cross-entropy loss of 0.01. Cui and Yanning [6] introduced CFUN+, a two-stage multi-modality segmentation strategy for detailed substructure segmentation of the whole heart. Using original CT and MRI data, this approach combined Faster R-CNN and a 3D U-Net. CFUN+ introduced a novel CIoU loss function, achieving a Dice score of 91.1% on the MM-WHS 2017 challenge CT dataset, surpassing the baseline CFUN model by 5.2%. The method significantly improved segmentation speed, reducing the time to less than 6 seconds for a single heart [17]. The limitations include sensitivity to image quality, difficulty in handling diverse datasets, high computational needs, and the requirement for extensive labeled data.

Naseer and Arfan [20] presented early lung cancer detection using a modified U-Net, extracting nodules, and classifying them as cancerous or not using Alex Net-SVM. Tested on the LUAN16 dataset, their method showed high accuracy (97.98%) and effectiveness in distinguishing lung cancer. Salama and Aly [22] introduced a novel framework for precise lung CT image segmentation, aiming to improve lung cancer analysis. They used the U-Net algorithm with data augmentation and various preprocessing steps. The algorithm achieved impressive results, including a 98.78% Jaccard Index and a 98.75% dice-coefficient index in lung cancer segmentation, with a computational time of 2.2145 seconds. Ferdinandus et al. contributed [8] to use deep learning, specifically the MultiResUNet architecture, to improve lung segmentation and COVID-19 infection analysis in CT images. The Mean-IoU evaluation metric demonstrated a 1.33% improvement in segmentation accuracy with MultiResUNet (93.05%) compared to U-Net (91.83%) in the utilized dataset. These findings suggest promising advancements in lung segmentation for early detection of COVID-19 infection.

Xiao et al. [26] introduced an innovative solution to address the limitations of existing lung segmentation algorithms by proposing a novel UNet-like model termed In-UNet. Additionally, integrating the Mish activation function contributed to the overall segmentation effectiveness [18]. The In-UNet showed superior lung parenchyma region segmentation compared to other networks in experimental validation on a lung CT dataset. Yang and Tian [27] proposed an advanced deep learning-based semantic segmentation network for efficient whole heart segmentation in cardiovascular disease diagnosis. By enhancing the attention mechanism and optimizing data preprocessing based on TransUNet, the model achieved a Dice score of 0.929 on the MM-WHS 2017 dataset, surpassing previous methods. While promising, further validation across diverse datasets and considerations for computational efficiency in real-time applications. Qiao and Song [21] enhanced U-Net for better whole heart segmentation in CT images by

introducing the ASPP module and attention mechanism to overcome complexity, low accuracy, tissue adhesion, and incomplete segmentation issues. The algorithm achieved 89.74% segmentation accuracy on the MMWHS dataset, showcasing improvement in whole heart structure segmentation. The potential limitations include generalization across datasets and computational efficiency in real-time applications.

III. Methodology

At first, we split the dataset into training, validation and testing in 60:30:10 proportion. Then we apply the proposed TelLung-Net model to the selected dataset. Then validate and test the data on the training data. After validation, we evaluate the model performance on F1-score, precision, and recall. The detailed method is presented in the following subsections.

A. Dataset

The research utilizes two open-source Mendeley Data datasets to examine pulmonary chest X-ray anomalies. The dataset known as Pulmonary Chest X-Ray Abnormalities consists of a collection of pictures (as shown in Figure 2) obtained from two sources: Shenzhen and Montgomery in Table I. The Montgomery subset includes left and right lung masks for each of its 138 images. This dataset provides diverse chest X-ray data for training and validating models.

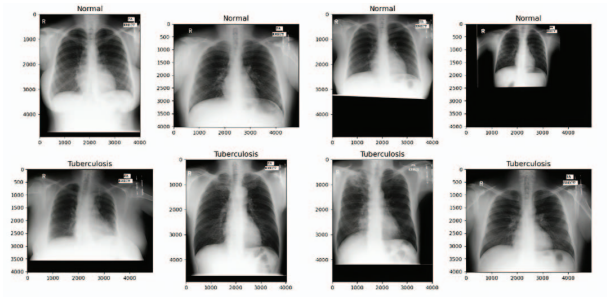

Fig. 2. Used Dataset Sample

TABLE I
Dataset Image Description

Dataset name	Source	Image
Chest X-Ray Abnormalities	Shenzhen and Montgomery [11]	662
Montgomery Subset	Shenzhen Hospital [10]	566

B. Dataset Preparation and Training

We created a complete data set containing all the photos to evaluate our image processing. We also set aside a separate test set for later assessment. After that, we divided the dataset into two parts for training and validation. We ensured an equal number of tuberculosis (TB) and normal patient images in both sets, maintaining a 60:30:10 ratio. The training dataset contained 362 samples, with 185 labeled as normal and 177 as tuberculosis, and 100 images for test cases. The validation dataset had 242 samples, of which 124 were labeled as normal, and 118 were tuberculosis in the Table II. The process of resizing each image and their related masks to a defined size of 128x128 pixels meets the purpose of establishing consistency within the dataset. This consistency, in turn, facilitates consistent processing of the data. Implementing the TelLung-Net model in the Keras framework involves the selection of the Adam optimizer [12], responsible for effective weight updates throughout the training process. To improve the convergence and stability of the model, a normalization technique is used, where, the pixel values of the images are divided by 255 [25].

TABLE II
Dataset Details at Different Stages

Stage	Total Samples	Normal Cases	Tuberculosis Cases
Training	362	185	177
Validation	242	124	118
Testing	100	50	50

IV. Proposed TelLung-Net Architecture

The proposed architecture consists of three main components, which are presented in Figure 3. The components are the Contracting Path, Bottleneck, and Expansive Path. The working processes of the following components are given below.

A. TelLung-Net Architecture

The U-Net architecture is a convolutional neural network (CNN) specifically developed for semantic segmentation tasks [7]. The proposed architecture Components are listed below:

- **Contracting Path**: The architecture comprises a series of convolutional layers, followed by max-pooling layers. The process of iteratively reducing the resolution of data to collect and retain contextual information. At every stage, the number of feature channels is doubled. [14]
- **Bottleneck**: The architecture comprises convolutional layers designed to retain contextual information.
- **Expansive Path**: The inverse of the contracting route is formed by using up-convolutions and combining them with high-resolution features obtained from the contracting path. At each stage, the number of feature channels is reduced by half (0.5).

The TelLung-Net model integrates conventional convolutional and pooling procedures. The fundamental operation used is the convolution.

$$Convolution, Y = \sigma(W * X + b) \quad (1)$$

In the Equation1, W represents the weights, X is the input, b is the bias term and sigma is the activation function (commonly ReLU) [23] [4]. The pooling operation is typically max-pooling:

$$MaxPooling, Y = maxpool(X) \quad (2)$$

where X is the input. For up-convolution, is commonly implemented using transposed convolutions or up sampling [15]

$$UpConvolution, Y = transpose - convolution(X) \quad (3)$$

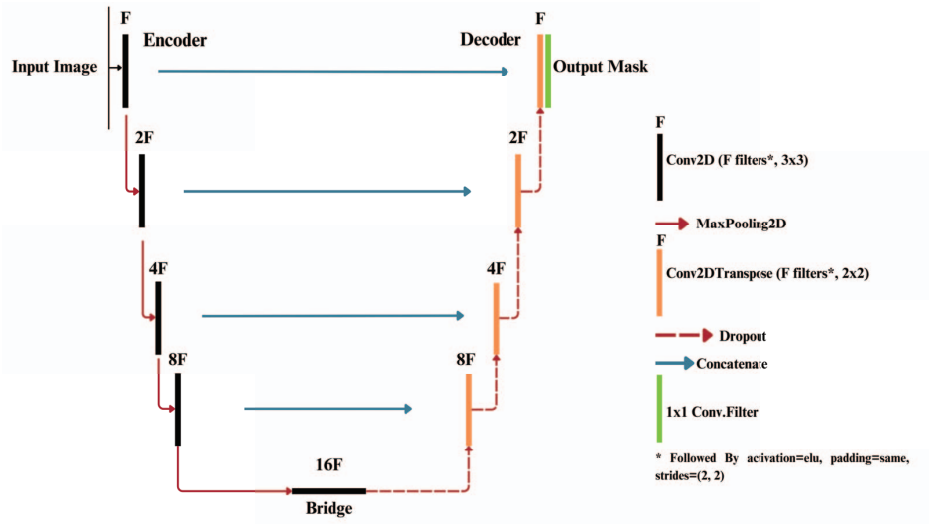

Fig. 3. Proposed U-Net architecture for Chest X-ray image segmentation.

This architecture captures context and helps pinpoint locations well. It's great for tasks like accurately finding lung boundaries in medical images [19]. This research examines the use of U-Net for lung segmentation in chest X-ray images. The architectural characteristics of U-Net, which form the foundation of its functionality, are outlined below:

Encoder (Contracting Path):

- The input layer (128x128x3) takes chest X-ray images as input. In Table III demonstrates the encoding portion of UNet architecture.
- Convolutional layers (Conv2D) with ReLU activation and dropout are applied to extract features.
- This process is repeated to create a deep encoding path that captures hierarchical features.

Decoder (Expanding Path):

- Convolutional layers with dropout are used to up sample and generate higher resolution feature maps. In Table IV describes the decoding of U-Net architecture.
- Transposed convolutional layers(Conv2DTranspose) are employed for up-sampling.
- Skip connections concatenate feature maps from the corresponding layers in the encoder, aiding in preserving spatial information.

Output Layer:

- The final layer utilizes a Conv2D layer with a sigmoid activation function, producing a binary mask representing the segmented lungs image.

Training:

- Binary cross entropy loss compares the predicted mask with the ground truth lung mask.
- Adam optimizer is employed for gradient descent during training.

Parameters: The total trainable parameters in the model are 19,41,105. The encoder captures key features, while the decoder generates segmented masks. The TelLung-Net architecture is trained using the fit_generator method, which undergoes several training cycles. The implementation of the proposed model architecture is accessible in the GitHub[1].

TABLE III
U-NET MODEL ENCODER

Layer Type	Filters	Size	Stride	Pad	Dropout
Input Layer	-	128x128x3	-	-	-
Conv2D	16	3x3	1	1	-
Dropout	-	-	-	-	0.25
...	...	...	...	...	...
Conv2D	128	3x3	1	1	-
Dropout	-	-	-	-	0.25
Conv2D	128	3x3	1	1	-

TABLE IV
U-NET MODEL DECODER

Layer	Filters	Size/Stride	Pad	Dropout
Conv2DTranspose	128	3x3/2	1	-
Concatenate	-	-	-	-
Conv2D	128	3x3/1	1	-
Dropout	-	-	-	0.25
...	...	...	...	...
Conv2D	16	3x3/1	1	-
Dropout	-	-	-	0.25
Conv2DTranspose	16	3x3/2	1	-
Concatenate	-	-	-	-
Conv2D	1	1x1/1	1	-

B. Image Masking with Generator Function

For the testing and training stages, we are able to implement a specialized generator function for image segmentation tasks. This generator function efficiently handled large numbers of datasets with complexity by providing batches of preprocessed

[1]GitHub https://github.com/RifatRudro/Image-Processing-Works

image-mask pairs during each training step. Our training dataset, referred to as X, consisted of 10 RGB images, each with a resolution of 128x128 pixels in Figure 4. The Y training dataset also included 10 binary masks with the same resolution. These datasets played a vital role in training a neural network to build accurate semantic segmentation models.

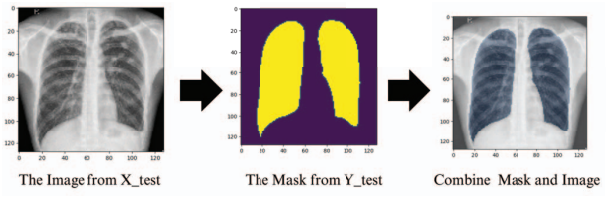

Fig. 4. Model Training Phase

C. Validation and Model Implementation

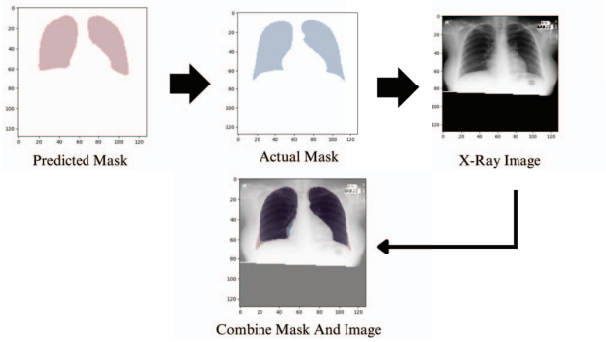

Fig. 5. Model Testing

In the training phase, we start by performing an initial validation. This involves overlaying each image with its associated mask to ensure accurate alignment, as shown in Figure 5. Once the training phase concludes, we apply the U-Net Model Architecture with specific parameter configurations, including a dropout rate of 0.1 and an initialization seed of 101. The model undergoes a training process, allowing us to predict the degree of overlap between the masked and actual photos. Following training, we randomly select photos and their corresponding masked counterparts to assess the accuracy of the dual image overlay.

D. AI Enabled Filtering

This paper includes optimizing healthcare procedures, improving patient accessibility to medical care, and introducing automation into specific aspects of the medical consultation process. Figure 6 illustrates the procedure for designing an artificial intelligence (AI) filter process. Datasets are compiled and reprocessed (resizing, normalization). Post-processing involves converting predictions to binary outcomes using a threshold operation (0.7). For configuring the model dropout rate of 0.1 prevents over-fitting during training. An initial

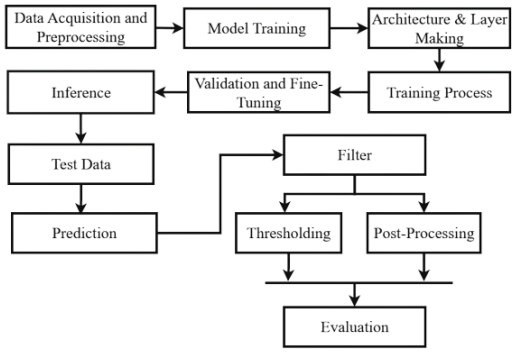

Fig. 6. AI Filtering Working Process

seed value (101) enhances reproducibility. Metrics include accuracy, precision, recall, F1 score, and a confusion matrix analysis. A feedback loop supports model iteration, potential data inclusion, or parameter adjustments for continuous improvement.

E. Wearable Device Integration

The patient's blood pressure, heart rate, and stress levels data are transmitted to the server using a smartwatch wearable to check up on the preliminary health condition, which can then be viewed on the doctor's web interface. Both wearable data and TelLung-Net-produced output are used for the patient's prescription. A prescription given by a doctor can be viewed, downloaded, or printed on the patient's web interface. The details of the web application interface are presented in Figure 9.

F. Web Interface

In our developed telemedicine platform's web interface, patients register and sign in using their credentials in Figure 7, grant authorization for telemedicine services, and activate chest-related medical services. They complete a medical questionnaire, provide wearable device data readings in Figure 8, and upload X-ray documents for analysis by an artificial intelligence filtering system, as shown in Figure 9. Based on this analysis, patients are referred to specialized chest disease doctors for further evaluation and treatment.

G. Model Evaluation

When evaluating the performance of a chest X-ray segmentation model for applications such as lung region detection, selecting appropriate metrics is crucial. The precision in Equation 4, recall in Equation 5, and F1 score in Equation 6 metrics are particularly suitable.

$$\text{Precision} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Positive (FP)}} \tag{4}$$

$$\text{Recall} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Negative (FN)}} \tag{5}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

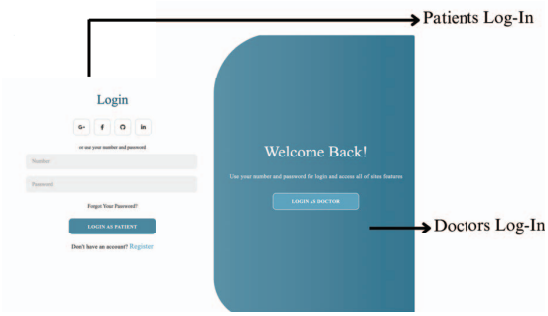

Fig. 7. A dual login portal for a healthcare platform with separate access points for patients and doctors.

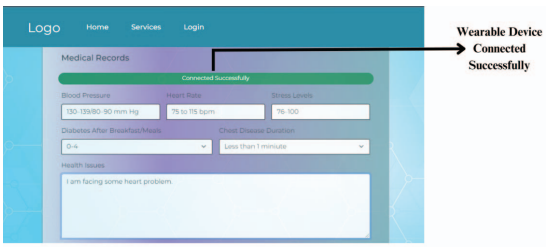

Fig. 8. A medical records web page user interface displaying vital signs and health status, indicating a successful connection to a wearable device.

V. Results and Discussion

A. Performance Analysis

The training in Table V and validation over the test dataset in Table VI analysis reveals a consistent improvement in accuracy and F1 score throughout 15 epochs. Starting with a training accuracy of 89.26% and validation accuracy of 89.76%, the model demonstrates steady progress, eventually achieving remarkable performance with a training accuracy of 97.26% and validation accuracy of 96.02% and validation loss is very less in Figure 10. This indicates the model's ability to learn and generalize from the training data. Additionally, the F1 scores reflect a balance between precision and recall, confirming the model's effectiveness for chest X-ray segmentation. The performance indicates the model's practical efficacy in medical diagnostics, making it a reliable tool for assisting medical professionals and minimizing the likelihood of false positives or negatives in chest disease diagnostics.

VI. Conclusion and Future Work

In this research, we have introduced an advanced telemedicine service with a web-based interface with sophisticated AI filtering systems designed to bridge the gap between patients and specialists in the chest medicine department. This novel telemedicine solution steams the diagnostic workflow and ensures timely access to expert care for a respiratory time frame. The service's advanced feature is its ability to revolutionize telemedicine practices and elevate the standard of patient care in diagnosing chest diseases. Our findings

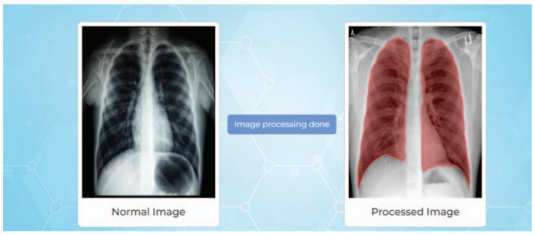

Fig. 9. Patient Attachment Processing For Doctor. Before-and-after comparison of a chest X-ray, with image processing highlighting lung areas.

TABLE V
Training Model Performance Metrics Over Chest X-Ray Abnormalities Dataset

Epoch	Loss	Training Accuracy	F1 Score	Precision	Recall
1	0.2444	89.26%	0.88	0.90	0.86
2	0.2411	90.12%	0.89	0.91	0.88
3	0.2388	90.74%	0.90	0.92	0.88
4	0.2365	91.18%	0.90	0.92	0.89
5	0.2343	91.56%	0.91	0.93	0.89
...	...	...	...	...	...
10	0.2239	92.81%	0.93	0.95	0.91
11	0.2219	92.95%	0.93	0.95	0.92
12	0.2200	94.06%	0.94	0.96	0.93
13	0.2181	95.14%	0.95	0.97	0.94
14	0.2163	96.21%	0.96	0.98	0.94
15	**0.2145**	**97.26%**	**0.97**	**0.98**	**0.96**

TABLE VI
Testing Model Performance Metrics Over Chest X-Ray Abnormalities Dataset

Epoch	Loss	Validation	Test	F1 Score	Precision	Recall
1	0.2555	89.76%	89.90%	0.88	0.89	0.87
2	0.2511	90.12%	90.21%	0.89	0.90	0.88
3	0.2467	90.48%	90.70%	0.89	0.91	0.87
4	0.2424	90.84%	90.73%	0.90	0.91	0.88
5	0.2383	91.18%	91.20%	0.90	0.92	0.88
...	...	...	...	...	...	...
10	0.2194	92.40%	92.59%	0.93	0.94	0.92
11	0.2159	92.55%	93.89%	0.93	0.95	0.92
12	0.2125	93.68%	94.99%	0.94	0.95	0.93
13	0.2092	95.78%	95.98%	0.95	0.96	0.94
14	0.2060	95.87%	96.01%	0.96	0.97	0.95
15	**0.2029**	**96.02%**	**96.26%**	**0.97**	**0.98**	**0.96**

showcase the TelLung-Net semantic segmentation's impressive accuracy rate of 96.29% with 0.97 in F1 score in diagnosing chest conditions, a testament to the effectiveness of our comprehensive dataset management and the precision of our custom-built generator. The robust accuracy of our model not only proves its proficiency in classifying and segmenting chest images but also confirms its reliability for clinical deployment. Although, the system currently works on chest X-ray images, research on MRI images is outlined as the future work of this work.

References

[1] Mustafa Almahdi Algaet, Zul Azri Bin Muhamad Noh, Abdul Samad Shibghatullah, and Ali Ahmad Milad. Provisioning quality of service of wireless telemedicine for e-health services. In *2013 IEEE Conference on Information & Communication Technologies*, pages 199–202, 2013.

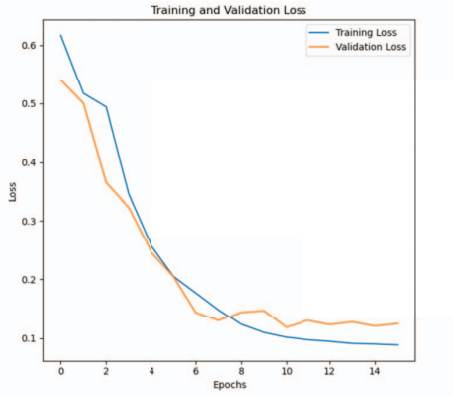

Fig. 10. The graph depicts the decrease in training and validation loss over epochs, the model's training process indicating learning over time.

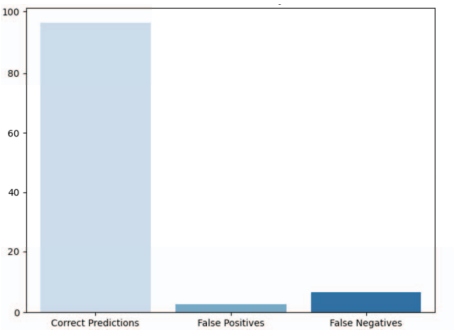

Fig. 11. The bar chart shows correct predictions substantially outnumbering both False Positives and False Negatives

[2] Madona Azar and Robert Gabbay. Web-based management of diabetes through glucose uploads: Has the time come for telemedicine? *Diabetes Research and Clinical Practice*, 83(1):9–17, 2009.

[3] R. Bellazzi, S. Montani, A. Riva, and M. Stefanelli. Web-based telemedicine systems for home-care: technical issues and experiences. *Computer Methods and Programs in Biomedicine*, 64(3):175–187, 2001.

[4] Yuming Cai, Haotian Li, Junyi Xin, and Guanqun Sun. Mlda-unet:multi level dual attention unet for polyp segmentation. In *2022 16th ICME International Conference on Complex Medical Engineering (CME)*, pages 372–376, 2022.

[5] P Stoila Cindy, Ananya Bhattacharjee, R Murugan, Ram Kumar Karsh, and Tripti Goel. Implementation of different u-net architectures for segmentation of lung cancer ct images. In *2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1)*, pages 1–6, 2023.

[6] Hengfei Cui, Yifan Wang, Yan Li, Di Xu, Lei Jiang, Yong Xia, and Yanning Zhang. An improved combination of faster r-cnn and u-net network for accurate multi-modality whole heart segmentation. *IEEE Journal of Biomedical and Health Informatics*, 27(7):3408–3419, 2023.

[7] Jing Di, Shuai Ma, Jing Lian, and Guodong Wang. A u-net network model for medical image segmentation based on improved skip connections. In *2022 14th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pages 298–302, 2022.

[8] Fx Ferdinandus, Esther Irawati Setiawan, Eko Mulyanto Yuniarno, and Mauridhi Hery Purnomo. Lung segmentation using multiresunet cnn based on computed tomography image. In *2022 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, pages 1–6, 2022.

[9] Manh-Hung Ha, Do Anh Luyen, Oscal Tzyh-Chiang Chen, Truong Cong Doan, and Nguyen Van Tinh. Ndau: Non-local dual attention-gated unet for brain tumor segmentation. In *2023 International Conference on System Science and Engineering (ICSSE)*, pages 543–547, 2023.

[10] Shenzhen Hospital. Shenzhen hospital x-ray set. https://paperswithcode.com/dataset/shenzhen-hospital-x-ray-set.

[11] Stefan Jaeger, Sema Candemir, Sameer Antani, Yì-Xiáng J Wáng, Pu-Xuan Lu, and George Thoma. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. *Quant. Imaging Med. Surg.*, 4(6):475–477, December 2014.

[12] Devidas T Kushnure and Sanjay N Talbar. MS-UNet: A multi-scale UNet with feature recalibration approach for automatic liver and tumor segmentation in CT images. *Comput. Med. Imaging Graph.*, 89(101885):101885, April 2021.

[13] Dinh-Hung Le, Nhat-Minh Le, Khac-Hung Le, Van-Truong Pham, and Thi-Thao Tran. Dr-unet++: An approach for left ventricle segmentation from magnetic resonance images. In *2022 6th International Conference on Green Technology and Sustainable Development (GTSD)*, pages 1048–1052, 2022.

[14] Chen Li, Yusong Tan, Wei Chen, Xin Luo, Yuanming Gao, Xiaogang Jia, and Zhiying Wang. Attention unet++: A nested attention-aware u-net for liver ct image segmentation. In *2020 IEEE International Conference on Image Processing (ICIP)*, pages 345–349, 2020.

[15] Dongyun Lin, Yi Cheng, Yiqun Li, Shitala Prasad, and Aiyuan Guo. Mlsa-unet: End-to-end multi-level spatial attention guided unet for industrial defect segmentation. In *2022 IEEE International Conference on Image Processing (ICIP)*, pages 441–445, 2022.

[16] Konstantinos Loverdos, Andreas Fotiadis, Chrysoula Kontogianni, Marianthi Iliopoulou, and Mina Gaga. Lung nodules: A comprehensive review on current approach and management. *Ann. Thorac. Med.*, 14(4):226–238, October 2019.

[17] G S Masson, D R Mestre, and L S Stone. Speed tuning of motion segmentation and discrimination. *Vision Res.*, 39(26):4297–4308, October 1999.

[18] Diganta Misra. Mish: A self regularized non-monotonic activation function, 2020.

[19] Seyyed Ali Mortazavi-Zadeh, Alireza Amini, and Hamid Soltanian-Zadeh. Brain tumor segmentation using u-net and u-net++ networks. In *2022 30th International Conference on Electrical Engineering (ICEE)*, pages 841–845, 2022.

[20] Iftikhar Naseer, Sheeraz Akram, Tehreem Masood, Muhammad Rashid, and Arfan Jaffar. Lung cancer classification using modified u-net based lobe segmentation and nodule detection. *IEEE Access*, 11:60279–60291, 2023.

[21] Guang Xiao Qiao and Ji Hong Song. Cardiac image segmentation based on improved u-net. In *2022 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML)*, pages 133–137, 2022.

[22] Wessam M. Salama and Moustafa H. Aly. Lung ct image segmentation: A generalized framework based on u-net architecture and preprocessing models. In *2021 31st International Conference on Computer Theory and Applications (ICCTA)*, pages 141–146, 2021.

[23] Kashish D. Shah, Dhaval K. Patel, Minesh P. Thaker, Harsh A. Patel, Manob Jyoti Saikia, and Bryan J. Ranger. Emed-unet: An efficient multi-encoder-decoder based unet for medical image segmentation. *IEEE Access*, 11:95253–95266, 2023.

[24] Li Su. Application of telemedicine diagnosis assistant system for breast diseases patients. In *2020 13th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, pages 453–456, 2020.

[25] Yan Wang, Lingjia Gu, Tao Jiang, and Fang Gao. Mde-unet: A multitask deformable unet combined enhancement network for farmland boundary segmentation. *IEEE Geoscience and Remote Sensing Letters*, 20:1–5, 2023.

[26] Hongxin Xiao, Lingxi Peng, Shaohu Peng, and Yifan Zhang. Lung image segmentation based on involution unet model. In *2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*, pages 184–187, 2022.

[27] Xiaoniu Yang and Xiaolin Tian. Transmunet: Using enhanced attention mechanism for whole heart segmentation. In *2022 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, pages 566–569, 2022.