

Automatika

Journal for Control, Measurement, Electronics, Computing and Communications



ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/taut20

PulmonU-Net: a semantic lung disease segmentation model leveraging the benefit of multiscale feature concatenation and leaky ReLU

H. Mary Shyni & Chitra E

To cite this article: H. Mary Shyni & Chitra E (2024) PulmonU-Net: a semantic lung disease segmentation model leveraging the benefit of multiscale feature concatenation and leaky ReLU, *Automatika*, 65:2, 641-651, DOI: [10.1080/00051144.2024.2314918](https://doi.org/10.1080/00051144.2024.2314918)

To link to this article: <https://doi.org/10.1080/00051144.2024.2314918>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 13 Feb 2024.



Submit your article to this journal [↗](#)



Article views: 510



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)



PulmonU-Net: a semantic lung disease segmentation model leveraging the benefit of multiscale feature concatenation and leaky ReLU

H. Mary Shyni and Chitra E

Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

ABSTRACT

Pulmonary diseases impact lung functionality and can cause health complications. X-ray imaging is an initial diagnostic approach for evaluating lung conditions. Manual segmentation of lung infections from X-rays is time-consuming and subjective. Automated segmentation has gained interest to reduce clinician workload. Semantic segmentation involves labelling individual pixels in X-rays to highlight infected regions. This article presents PulmonU-Net, an innovative semantic segmentation model using PulmonNet modules as the base network to highlight infected areas in chest X-rays. PulmonNet modules leverage global and local chest X-ray characteristics to create intricate feature maps. Incorporating leaky ReLU activation enables uninterrupted neuron functioning during learning. By adding PulmonNet modules in the encoder's deeper layers, the model addresses vanishing gradients and improves dice similarity coefficient to 94.25%. Real-time testing and prediction visualization demonstrate PulmonU-Net's effectiveness for automated lung infection segmentation from chest X-rays.

ARTICLE HISTORY

Received 12 December 2023
Accepted 31 January 2024

KEYWORDS

Deep learning; lung diseases; feature concatenation; semantic segmentation; U-Net; chest X-rays

1. Introduction

Pulmonary diseases cover an extensive range of ailments ranging from mild to potentially deadly infections that prevent the optimal functioning of the lungs and profoundly impact health and well-being [1]. It is essential to promptly identify and implement preventive measures to mitigate the consequences of these respiratory ailments as many of them tend to worsen over time [2]. X-ray imaging is often used as an initial diagnostic method for assessing lung diseases [3,4].

Healthcare practitioners can recognize abnormalities and make informed decisions regarding future diagnoses and treatments with the aid of segmentation, which highlights the diseased lung area in X-rays [5]. However, the manual annotation of the contaminated lung region is a demanding and labour-intensive process that requires skilled personnel and significant resources [6]. Hence, there is a substantial need for automated segmentation, due to its potential to alleviate the considerable workload associated with manual segmentation. Computerized segmentation employs deep learning algorithms to precisely identify and delineate the region of infection in the chest radiographs [7,8].

U-Net is a dedicated semantic segmentation framework specifically crafted for biomedical purposes which labels every individual pixel to localize the area of abnormality [9]. One advantage of U-Net is its ability to learn features from limited data samples, while the

skip connections help accurately locate features by preserving spatial information [10]. The U-Net encoder has been modified to incorporate richer feature representation patterns [11–13]. Some notable backbone enhancements include residual blocks [14] to train the network faster and DenseNet [15] to mitigate vanishing gradients. However, the inclusion of long skip connections increases the model complexity and dense connections introduce redundant parameters.

Therefore, the focus of this research is to introduce an innovative semantic segmentation model named PulmonU-Net that replaces the backbone of the conventional U-Net model with PulmonNet modules to generate dense predictions. PulmonNet modules incorporate filters of various sizes to handle features at multiple scales. The model integrates the local and global features of X-rays to alleviate challenges like vanishing gradient and overfitting, and it concurrently enhances GPU memory utilization due to its smaller parameter count in contrast to ResNet and DenseNet. The output generated by the model is a binary segmented mask that illustrates the anomaly and the background when the model is trained using chest radiographs comprising four distinct forms of lung illnesses.

The research presents several significant contributions:

1. The backbone of the U-Net segmentation model is replaced with PulmonNet modules to present

the PulmonU-Net semantic segmentation model which has produced promising outcomes in the categorization of pulmonary diseases.

2. Through the utilization of data extracted from X-rays at varying scales, the fusion process aids in segmenting infections with different sizes, efficiently addressing concerns related to overfitting and vanishing gradients.
3. The utilization of Leaky ReLU as the activation function makes sure that the neurons remain active throughout the feature learning process.
4. Real-time data is used to test the PulmonU-Net model to evaluate its ability to manage a wide variety of instances and produce reliable segmentations even under challenging circumstances.

The structure of the article unfolds as follows: The recent trends in the field concerning the localization of medical images are covered in Section 2. Section 3 clarifies the essential details regarding the dataset and delineates the framework of the PulmonU-Net model. In Section 4 the article focuses on the assessment of the proposed approach, encompassing both quantitative and qualitative performance metrics and Section 5 concludes the article by summarizing key findings and outlining future research considerations.

2. Related works

U-Net segmentation was presented by Olaf Ronneberger et al. and trained for three biological tasks by assigning a class label to each pixel. To achieve more accurate segmentations, elastic deformations have been applied to the training images. Skip connections concatenate the local details with the up-sampled feature maps improving the overall localization process [10]. To ensure semantic comparability between the encoder and the decoder feature maps prior to fusion, UNet++ redesigns the skip paths by including dense convolutions. The deep supervision of the model accurately segmented multi-scale lesions when evaluated using medical imaging datasets [16].

The DoubleU-Net introduced by Debesh Jha et al. employed two U-Nets in the analysis path to effectively capture the deep features and generate the output mask. With a greater number of model parameters, DoubleU-Net suppressed the performance of U-Net [17]. In DUNet, the retinal vessels are captured by stacking the encoder and the decoder with the deformable convolution blocks. This approach adjusts the receptive fields based on the shape and scale of the retinal vessels enabling the model to capture even the tiny and weakest vessels with a global accuracy of 96.97% [18].

Xiaomeng Li et al. improved the segmentation performance of the liver lesions by optimizing the integrated intra and inter-slice features. The H-DenseUNet model effectively tackled the limitations of 2D

convolutions in capturing volumetric contexts and the high memory usage of 3D convolutions [19]. Ozan Oktay et al. trained the U-Net model with attention gates for the dense prediction of the pancreas from the CT images. The attention coefficients suppress the irrelevant regions from the CT images thereby enhancing the capacity of the model without the need for multiple modules [20].

The Sparse-MLP incorporated by Jiaming Luo et al. into the U-shaped architecture is an efficient computer vision framework for mass delineation in mammograms. The encoder and decoder in the convolutional stage are connected through skip connections, allowing for the fusion of their respective features. The use of MLP blocks enhances the extraction of supplementary information from the feature maps [21]. In [22] MSD-Net was proposed to segment three different categories of COVID-19 infection from the CT images with a better specificity rate. This model utilizes ResNet-101 as the encoder backbone and is capable of segmenting infections of varying sizes while refining the feature maps.

Short skip connections were added to U-Net by Michal Drozdal et al. to enhance longer ones which ensures an uninterrupted flow of gradients in deep networks. With the two connections, the model speeds up the training process and increases the convergence of biomedical delineation [14]. To address the inadequate sensitivity of existing models in segmenting tiny blood vessels from retinal fundus images, DR-VNet was proposed in [23]. The model employs three encoders and decoders, utilizing DenseNet and ResNet backbone patterns to extract spatial information from input images.

3. Materials and methods

PulmonU-Net aims to precisely detect and segregate the impacted regions of four specific pulmonary abnormalities namely COVID-19, lung opacity, tuberculosis and pneumonia, from radiographs by generating binary masks that represent the abnormality and the background. The proposed approach follows a workflow consisting of four essential stages: Data collection, Data pre-processing, Model training, and Model evaluation, as illustrated in Figure 1.

3.1. Data collection

The PulmonU-Net segmentation model has been trained on two heterogeneous datasets from the Kaggle database: the COVID-19 Radiography Database [24,25] and the Tuberculosis (TB) Chest X-ray Database [26]. The combined data collection consists of three classes namely COVID-19, lung opacity and pneumonia from the COVID-19 Radiography Database and Tuberculosis class from the Tuberculosis

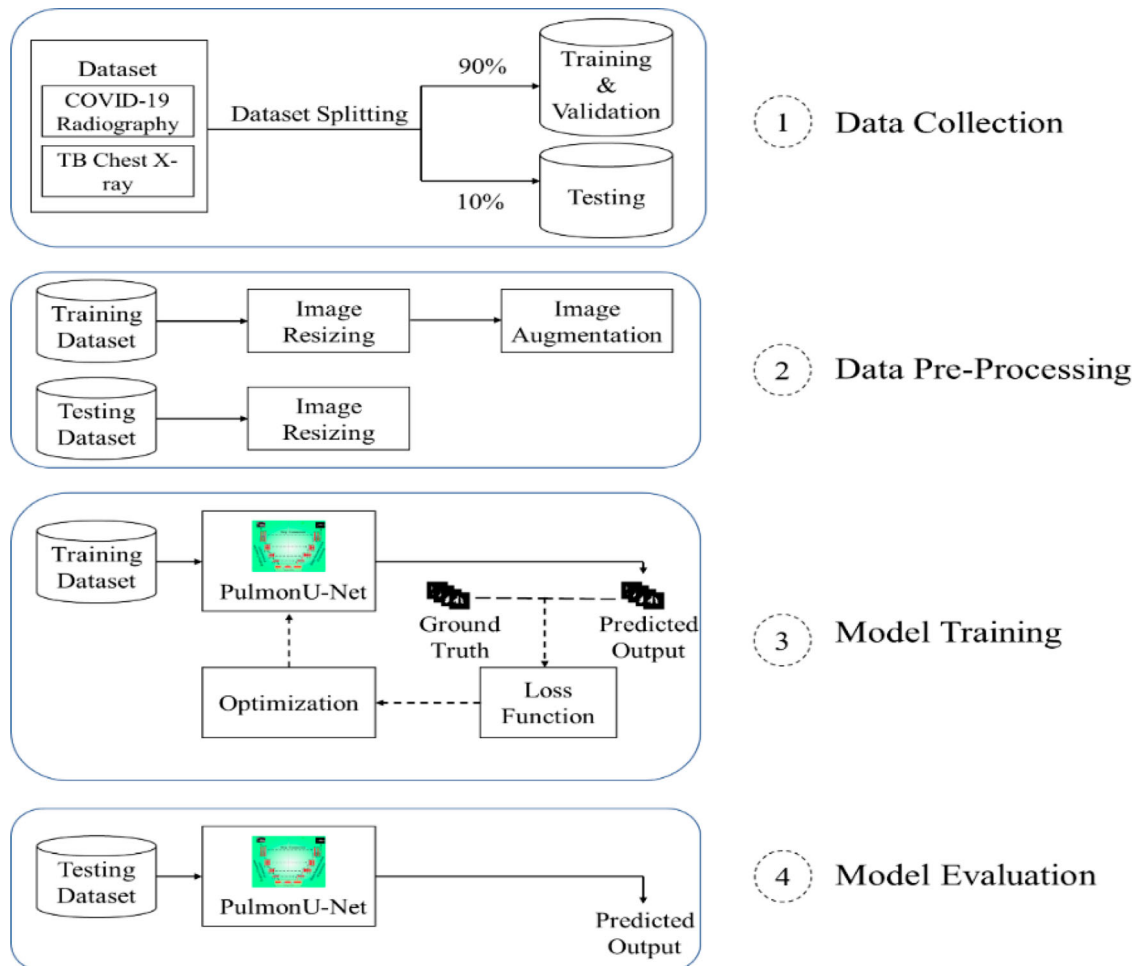


Figure 1. Procedural sequence of PulmonU-Net.

(TB) Chest X-ray Database. 50 infected chest radiographs from each class are manually annotated by a professional radiologist, resulting in the generation of corresponding binary masks. Out of the 200 annotated images, 90% (180 images) are used for training and validation, while 10% (20 images) are allocated for testing purposes. The study was granted ethical authorization by the SRM Medical College Hospital and Research Centre located in Kattankulathur, India. The hospital also contributed 25 real-time chest X-ray images that are utilized to assess the model.

3.2. Data pre-processing

The chest radiographs acquired from three distinct sources are resized to a standardized dimension of 224×224 . This resizing was performed to strike a balance between image resolution and computational efficiency [27]. The utilization of data augmentation has been necessitated by the sparse amount of annotated samples, in order to improve the generalization of the PulmonU-Net model [28,29]. In the pre-processing stage, six augmentation techniques were employed, including image rotation, vertical and horizontal shifting, shearing, zooming in or out, and horizontal flipping of the radiographs [30,31].

3.3. PulmonU-Net

Localizing the anomaly in biomedical applications is made easier with the help of the semantic segmentation model U-Net, which is excellent at extracting complex characteristics from small datasets. However, recent enhancements to the U-Net algorithm aimed at addressing overfitting and vanishing gradient issues come at the cost of introducing redundant parameters and increasing model complexity [32]. The PulmonU-Net technique uses Leaky ReLU and multiscale feature concatenation to provide better segmentation to handle these barriers.

3.3.1. Architectural details

The PulmonU-Net, specifically crafted for segmenting lung diseases is built upon the U-Net architecture, featuring convolutional layers enhanced by PulmonNet modules. Figure 2 presents the layout of the model which consists of a compressing path rooted with encoders, an expanding path rooted with decoders, and the incorporation of skip connections.

The encoder layers responsible for feature extraction use PulmonNet modules as feature extractors which has produced promising outcomes in the classification of pulmonary diseases. Batch normalization and

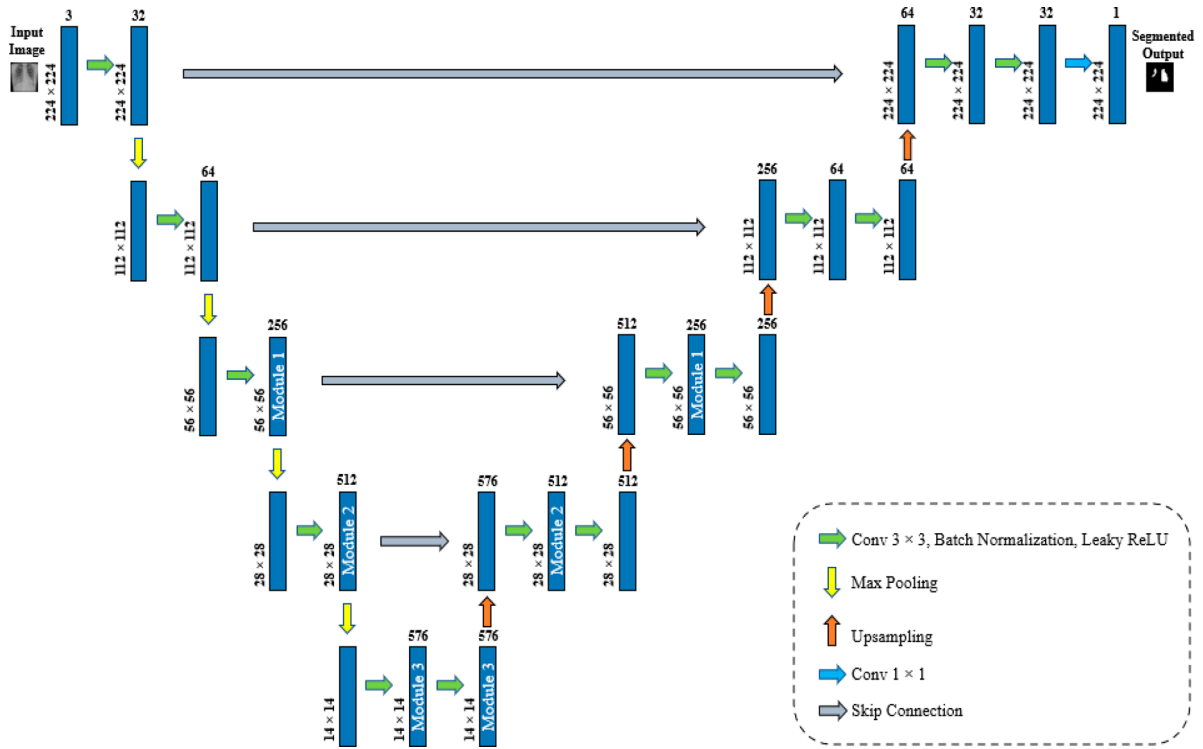


Figure 2. PulmonU-Net architecture with feature fusion approach.

a leaky ReLU activation function are added to each encoder layer to maintain training stability and guarantee increased neuron activity throughout the training process. As the encoder layers progress towards the bottleneck, the spatial resolution of the feature maps is decreased while their depth is increased by the convolutional and pooling operations carried out by the encoder layers. The initial encoder layers comprise convolutional layers that extract the shallow features of the input chest radiographs. Many existing algorithms encounter challenges in seamlessly merging the deep and shallow features to retain fine details, resulting in constraints on segmentation accuracy [33].

To enhance the feature fusion ability, the PulmonU-Net model uses PulmonNet modules at the deeper layers of the encoder and mitigates the effect of vanishing gradient and overfitting. These PulmonNet modules in the deeper layers merge the global features with the local features of the chest radiographs to produce more intricate feature maps. Figure 3 presents the three PulmonNet modules involved in the feature fusion process.

Incorporating filters of different sizes at the same level enables the model to capture a wide range of features from the input images [34]. The PulmonNet module 1 incorporates three filters of sizes 1×1 , 3×3 and 5×5 for deep feature extraction to capture multi-scale features. The concatenated multi-scale deep features of dimension $56 \times 56 \times 192$ are fused with the low-level features of dimension

$56 \times 56 \times 64$. The PulmonNet module 1 output features with dimensions $56 \times 56 \times 256$. Subsequently, a max pool layer follows module 1 to reduce the spatial dimension to $28 \times 28 \times 256$ which is fused with the down-sampled low-level features of dimension $28 \times 28 \times 64$, resulting in dimensions of $28 \times 28 \times 320$. The down-sampled high-level features are passed through 512 convolutions to produce feature maps of size $14 \times 14 \times 512$. These feature maps are fused with the down-sampled low-level features of size $14 \times 14 \times 64$ and outputs $14 \times 14 \times 576$ at PulmonNet module 3. Down-sampling of low-level features is done to align their spatial aspects with the high-level features for feature concatenation.

The decoders in the extended path reverse the encoding process by enhancing the spatial resolution of the feature maps, leading to a segmented output at the pixel level. The feature maps produced at each encoder layer are merged with the feature maps of their corresponding decoder layer through skip connections, maintaining detailed spatial information from the input images. The incorporation of a 1×1 convolution in the concluding decoder layer is essential for channel limitation in the feature maps. This process plays a vital role in aligning the feature maps with the desired number of output classes, ultimately leading to the intended result. The utilization of a sigmoid function in this layer facilitates the discrimination between the infected area and the background. Dropout layers are utilized in the PulmonU-Net model to promote the learning of more generalized features and prevent overfitting [3].

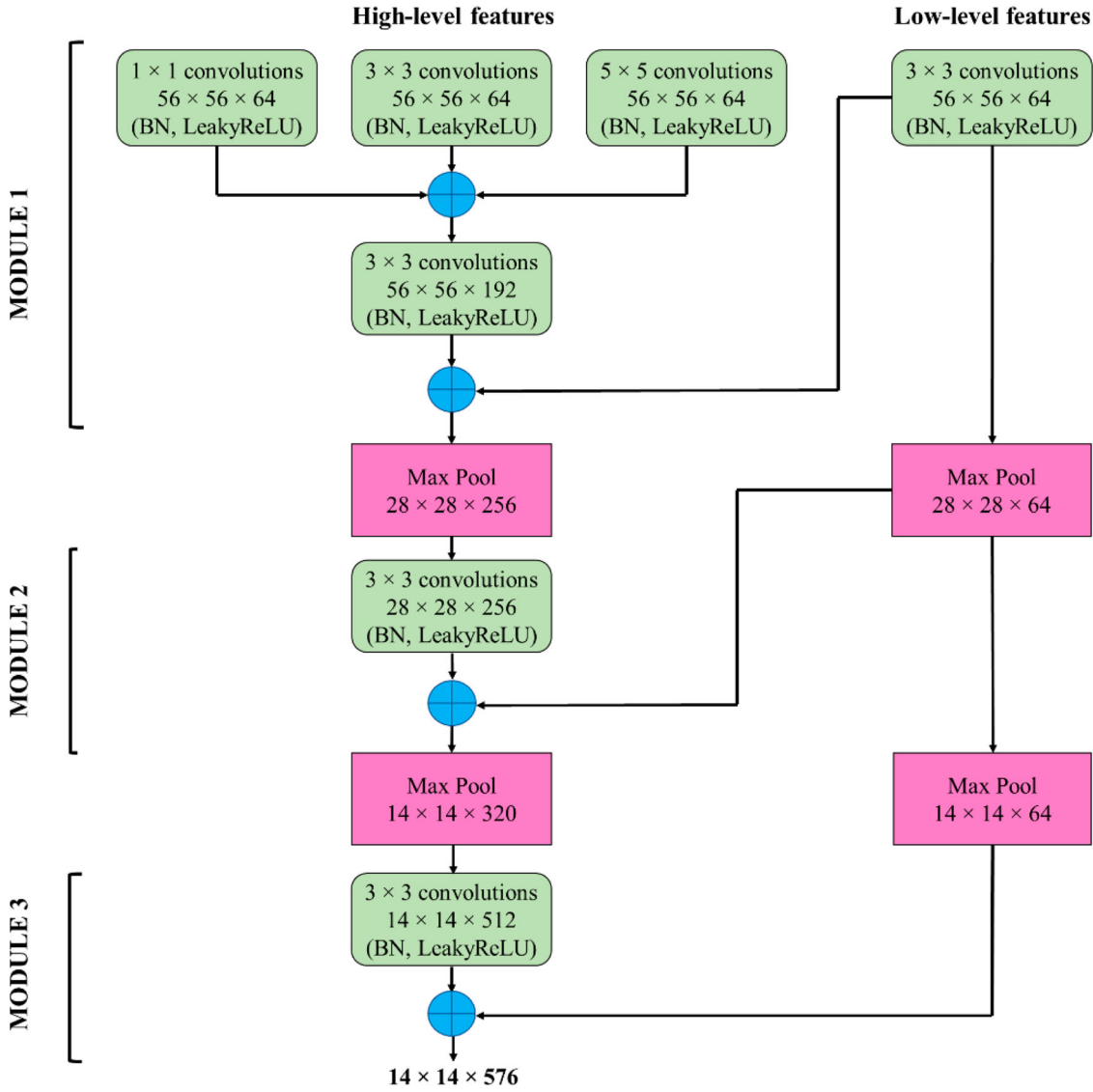


Figure 3. PulmonNet modules involved in the feature fusion process.

3.3.2. Feature concatenation

The PulmonNet modules employed at the deeper layers of PulmonU-Net undergo feature concatenation. Module 1 involves multi-scale deep and local feature concatenation, while the subsequent modules perform concatenation with the extracted deep and local features.

The combined features of module 1:

$$f_1 = [f_{d1}, f_{l1}] \quad (1)$$

The deep multi-scale features of module 1:

$$f_{d1} = [f_{1(1 \times 1)}, f_{1(3 \times 3)}, f_{1(5 \times 5)}] \quad (2)$$

where $f_{1(1 \times 1)} \in R^{h \times w \times c_1} \rightarrow 56 \times 56 \times 64$ denotes the feature map of 1×1 filter. h , w and c represent the height, width, and number of channels respectively. $f_{1(3 \times 3)} \in R^{h \times w \times c_2} \rightarrow 56 \times 56 \times 64$ denotes the feature map of 3×3 filter and $f_{1(5 \times 5)} \in R^{h \times w \times c_3} \rightarrow 56 \times 56 \times 64$ denotes the feature map of 5×5 filter.

$f_{d1} \in R^{h \times w \times (c_1 + c_2 + c_3)} \rightarrow f_{d1} \in R^{h \times w \times c_4} \rightarrow 56 \times 56 \times 192$ represents the deep multi-scale features of module 1 and $f_{l1} \in R^{h \times w \times c_5} \rightarrow 56 \times 56 \times 64$ denotes the low-level feature maps.

$f_1 \in R^{h \times w \times (c_4 + c_5)} \rightarrow f_1 \in R^{h \times w \times c_6} \rightarrow 56 \times 56 \times 256$ represents the combined deep multi-scale and low-level features of module 1.

The combined features of module 2:

$$f_2 = [f_{d2}, f_{l2}] \quad (3)$$

where $f_{d2} \in R^{h \times w \times c_6} \rightarrow 28 \times 28 \times 256$ represents the down-sampled deep feature maps and $f_{l2} \in R^{h \times w \times c_5} \rightarrow 28 \times 28 \times 64$ represents the down-sampled low-level feature maps.

$f_2 \in R^{h \times w \times (c_6 + c_5)} \rightarrow f_2 \in R^{h \times w \times c_7} \rightarrow 28 \times 28 \times 320$ represents the combined deep and low-level features of module 2.

The combined features of module 3:

$$f_3 = [f_{d3}, f_{l3}] \quad (4)$$

Table 1. PulmonU-Net hyper-parameters.

Hyper-Parameter	Value
Optimizer	Adam
Epochs	200
Batch size	32
Learning Rate	0.001
Momentum	$\beta_1 = 0.9$ and $\beta_2 = 0.999$
Weight Decay	0
Epsilon	$1e^{-7}$
Leaky ReLU Slope (α)	0.001

Table 2. Evaluation metrics comparison of PulmonU-Net with U-Net.

Classification Metrics (%)	U-Net	PulmonU-Net
Precision	0.8312	0.9130
Recall	0.8580	0.9230
F1 Score	0.8670	0.9587
Specificity	0.8704	0.9480
Accuracy	0.8400	0.9708
Segmentation Metrics (%)	U-Net	PulmonU-Net
DSC	0.8687	0.9425
IoU	0.8232	0.8830

where $f_{d3} \in R^{h \times w \times c_8} \rightarrow 14 \times 14 \times 512$ represents the down-sampled deep feature maps passed through 512 convolutions and $f_{l3} \in R^{h \times w \times c_5} \rightarrow 14 \times 14 \times 64$ represents the down-sampled low-level feature maps.

$f_3 \in R^{h \times w \times (c_8 + c_5)} \rightarrow f_3 \in R^{h \times w \times c_9} \rightarrow 14 \times 14 \times 576$ represents the combined deep and low-level features of module 3.

3.3.3. Model training

The PulmonU-Net model undergoes training for 200 epochs, employing 5-fold cross-validation with the resized chest radiographs and their corresponding binary masks of size $224 \times 224 \times 3$. The mask layer obtained has a size of $224 \times 224 \times 1$, featuring the white representation of the affected portion of the lung region, while the rest of the background is depicted in black. The predicted binary mask is compared with the ground truth using the binary cross-entropy loss function. During backpropagation, the weights and biases are adjusted in every epoch using the default hyper-parameters of the Adam optimizer. Through weight updates, the model progressively captures more complex features from the training set, which helps it produce accurate segmentations on unobserved data. Table 1 reports the hyper-parameters used to train the PulmonU-Net model.

3.4. Model evaluation

After being trained, the PulmonU-Net model is loaded and undergoes testing. The predicted binary masks are compared with their ground truth and the model is evaluated using the segmentation metrics.

4. Experiments and results

A thorough comparative analysis has been conducted to systematically assess the PulmonU-Net model in relation to the baseline U-Net segmentation model. Furthermore, a visualization of the real-time predictions made by both models on data samples has been provided.

4.1. Experimental parameters

The PulmonU-Net model, which is specifically developed for segmenting pulmonary diseases, undergoes training on the Google Colaboratory platform for a total of 200 epochs. Through extensive experimentation, it has been determined that the most effective value for the slope of the Leaky ReLU activation function is 0.001. To enhance the model's ability to capture intricate features, the Adam optimizer is utilized along with the binary cross-entropy loss function to adjust the weights and biases. The output layer of the model employs the sigmoid activation function, enabling pixel-wise classification and generating a binary segmented output.

4.2. Performance metrics

The effectiveness of the PulmonU-Net model in precisely identifying regions of interest in a chest radiograph is measured through the application of multiple assessment criteria. To calculate these metrics, four measures are necessary: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). TP represents the number of infected pixels correctly identified, TN denotes the number of background pixels correctly identified, FP represents the number of pixels incorrectly identified as infected, and FN represents the number of pixels incorrectly identified as background. In terms of infection detection, the five evaluation metrics used are precision, recall, F1 score, specificity, and accuracy. On the other hand, infection segmentation is assessed using the Dice Similarity Coefficient (DSC) and Intersection over Union (IoU).

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

$$\text{Sensitivity / Recall} = \frac{TP}{TP + FN} \quad (6)$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (8)$$

$$\text{Accuracy} = \frac{TP + TN}{TN + TP + FN + FP} \quad (9)$$

To assess the agreement between the ground truth and predicted masks, two segmentation parameters are

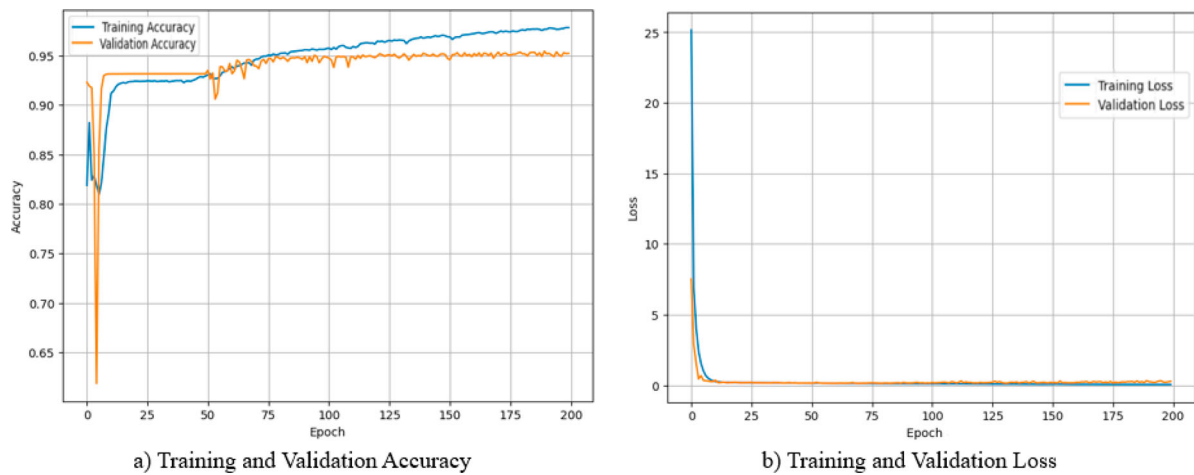


Figure 4. Performance curves of PulmonU-Net.

commonly used. The first one is the Dice Similarity Coefficient [35], which determines the fraction of overlapping pixels between the two masks relative to the total number of pixels in both masks. The second parameter is the Intersection over Union (IoU), which measures the percentage of common pixels between the ground truth and predicted masks with respect to the pixels that belong to either of the masks [36].

$$\text{Dice Coefficient} = \frac{2 * |N_G \cap N_P|}{|N_G| + |N_P|} \quad (10)$$

$$\text{Intersection over Union} = \frac{|N_G \cap N_P|}{|N_G \cup N_P|} \quad (11)$$

where N_G signifies the number of pixels in the ground truth mask and N_P signifies the number of pixels in the predicted mask.

4.3. Results and discussion

To evaluate the effectiveness of PulmonU-Net, a set of 20 data samples is utilized, with each class (COVID-19, Lung opacity, Pneumonia, and Tuberculosis) contributing 5 images. The testing process resulted in a binary segmented mask that accurately distinguishes the infected lung region from the background. The PulmonU-Net model is trained on a 4-class segmentation dataset for 200 epochs, achieving an impressive training accuracy rate of 98.25%. By analyzing the performance graphs presented in Figure 4, it becomes apparent that the accuracy of the training curve gradually increases until it reaches 175 epochs. Subsequently, between epochs 176 to 200, the accuracy remains relatively consistent, staying around 98%. This signifies that the PulmonU-Net model has achieved optimal feature learning from the 4-class segmentation dataset. Conversely, the validation accuracy has reached its peak at 95.08%. The minimal disparity between the training and validation accuracies serves as evidence that the model has successfully avoided both underfitting and overfitting. The accuracy and loss graphs depicted in

Figure 4 showcase the exceptional performance of the PulmonU-Net model.

To compare its performance, the same dataset was used to train and test the U-Net model for an equal number of epochs. The U-Net model achieved a training accuracy of 91.44%, which is 6.41% lower than that of the PulmonU-Net model. The performance curves of both the U-Net and PulmonU-Net models are visually represented in Figure 5.

Table 2 presents a quantitative comparison of the PulmonU-Net model and the U-Net model in terms of their performance. The proposed PulmonU-Net model, which integrates the features, achieves a dice coefficient of 94.25%, surpassing the baseline U-Net model by 7.38%. Furthermore, it attains an IoU of 88.3%, which is 5.98% higher than the U-Net model. These findings consistently demonstrate the superior performance of the PulmonU-Net model over the U-Net model, with a testing accuracy rate of 97.08%.

Figure 6 illustrates the comparison chart, which displays the classification metrics and segmentation metrics of both the PulmonU-Net model and the baseline U-Net model. The PulmonU-Net model exhibits an impressive capability to accurately detect infected pixels, while also effectively minimizing the occurrence of false positives. This is evident from its exceptional precision and recall metrics, which highlight its efficacy in accurately recognizing and segmenting the desired regions. The F1 score of 95.87% and specificity of 94.80% of the PulmonU-Net model prove its excellence in performance.

The SRM Medical College Hospital and Research Centre in Kattankulathur, India granted ethical approval for this study. A total of 25 real-time chest X-ray images, representing four different classes, are collected from the hospital to assess the performance of the model. Figure 7 depicts a sample of the visual predictions of real-time data made by the PulmonU-Net model and the baseline U-Net.

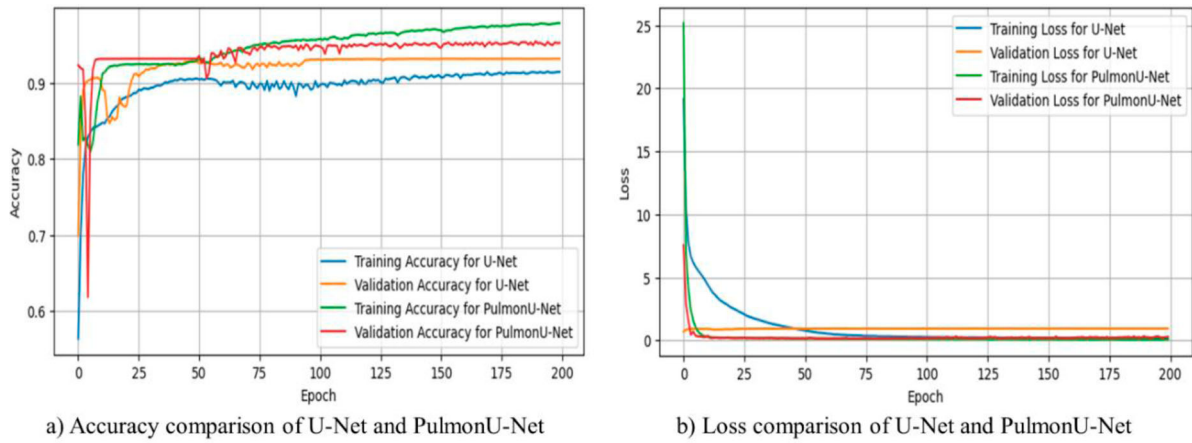


Figure 5. Performance comparison of U-Net and PulmonU-Net.

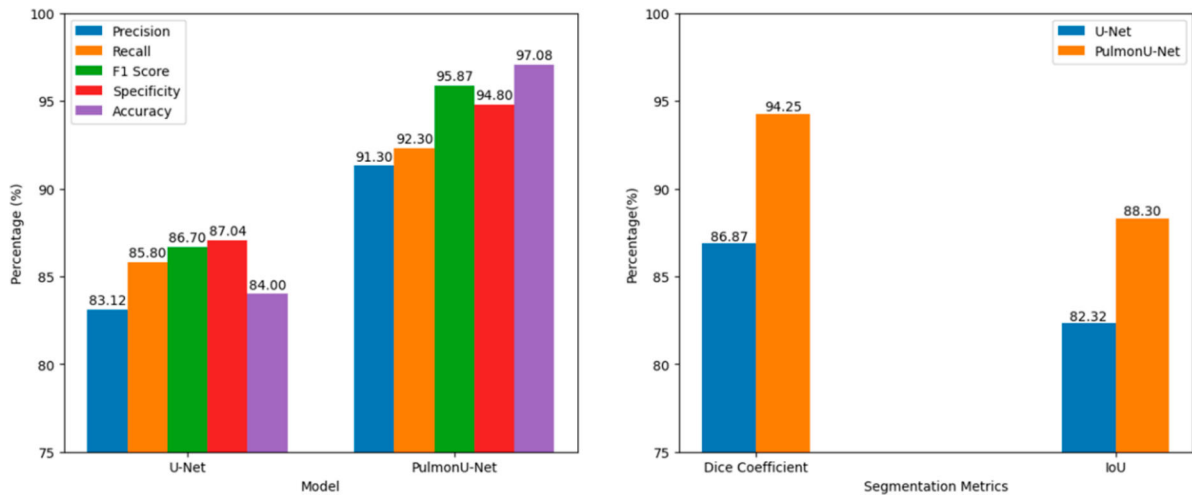


Figure 6. Evaluation metrics comparison chart of PulmonU-Net with U-Net.

The first row of Figure 7 presents the infected sample X-ray images from each class. The second and third rows represent the ground truth masks and the predicted masks respectively. When segmenting the infected region, two scenarios can occur: under-segmentation and over-segmentation. Under-segmentation can lead to loss of information which makes the model challenging to generate a robust segmentation boundary. However, in the case of over-segmentation, it is still possible to reconstruct the infected region since no useful information is lost [37]. From the visualized results it can be observed that the PulmonU-Net model highlights the infection effectively close to the ground truth with negligible over-segmentation in comparison with the baseline U-Net.

Despite the varying sizes of the infection in each sample, the model effectively detects them using the multi-scale filters of the PulmonNet modules, which retain contextual information at different scales. The PulmonU-Net model is specifically trained to segment infections in four different lung diseases. In Figure 7, each column represents a specific lung disease. Among the four real-time samples, the PulmonU-Net model successfully replicated the tuberculosis-infected sample

with a high dice coefficient of 95.02% and an IoU of 90.5%. In comparison, the U-Net model achieved a dice coefficient of 75.02% and an IoU of 60.03% for the same sample, which is 20% and 30% lower than the PulmonU-Net model, respectively.

5. Conclusion and future works

Segmentation of chest X-ray images helps in coming up with the apt medical treatment by diagnosing the specific lung condition. Automated segmentation is preferred over labour-intensive manual segmentation procedures. This research work has presented PulmonU-Net, an innovative semantic segmentation model that utilizes PulmonNet modules as the underlying network to highlight the infected region in chest radiographs. The PulmonNet modules have intricate feature maps that are created by taking into account the global and local characteristics of chest radiographs.

Leaky ReLU activation function is included in these modules to ensure the uninterrupted function of neurons throughout the learning process. The model effectively addresses the vanishing gradient issue by

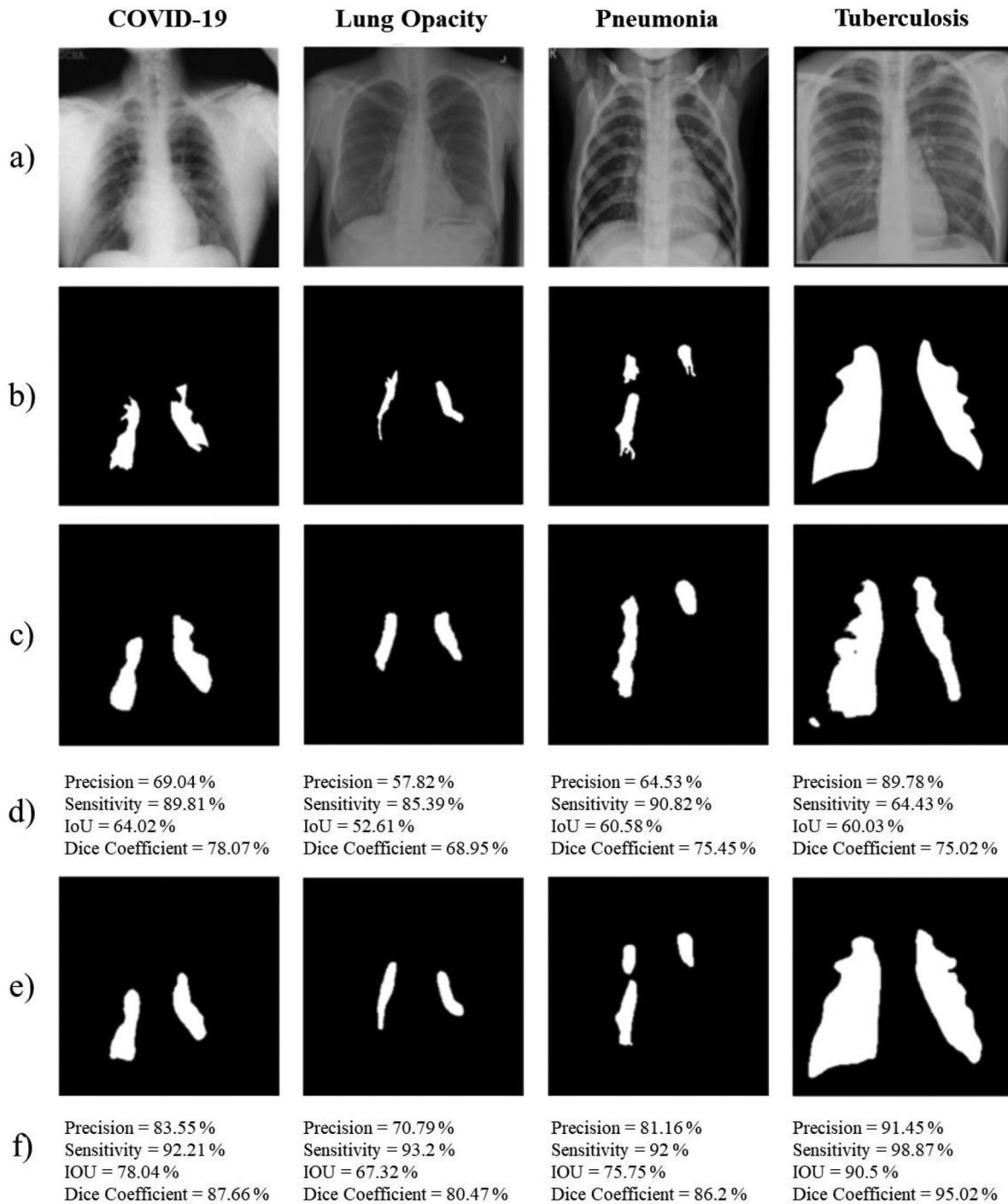


Figure 7. Simulation results of real-time sample prediction a) Original data samples b) Ground truth mask c) Predicted mask by U-Net d) Evaluation metrics of U-Net e) Predicted mask by PulmonU-Net d) Evaluation metrics of PulmonU-Net.

incorporating PulmonNet modules in the deeper layers of the encoder. This enhances.

The real-time testing of the semantic segmentation model PulmonU-Net reveals the impressive capabilities of the model in capturing fine-grained details and accurately delineating object boundaries across diverse data samples. The essence of PulmonU-Net lies in its ability to leverage the richness of multi-scale features and maintain active neurons throughout the training process.

The model achieved 94.25% dice coefficient and 88.3% IoU. This reflects its accuracy and reliability, making it well-suited for lung disease segmentation. Though the effectiveness of PulmonU-Net in highlighting infections across all four classes is notable, the inability to differentiate between different infection categories may impact its utility in providing detailed diagnostic information. A prospective direction for future research on the PulmonU-Net model is to encompass multi-class segmentation and provide a quantitative

assessment of infection severity within each class to broaden its utility.

Declarations

Competing Interests: The authors have no competing interests to declare that are relevant to the content of this article.

Ethical approval Ethical approval for the execution of this study was granted by the SRM Medical College Hospital and Research Centre, Kattankulathur, India.

Data Availability Statement i) The dataset analyzed during the current study is available in the Kaggle repository, <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database> and <https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>. ii) The real-time data samples gathered from SRM Medical College Hospital and Research Centre are used under license for the current study and are not publicly available.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- [1] MacNee W. Pathogenesis of chronic obstructive pulmonary disease. *Clin Chest Med.* 2007;28(3):479–513. doi:10.1016/j.ccm.2007.06.008
- [2] Doyle TJ, Hunninghake GM, Rosas IO. Subclinical interstitial lung disease: why you should care. *Am J Respir Crit Care Med.* 2012;185(11):1147–1153. doi:10.1164/rccm.201108-1420PP
- [3] Kim S, Rim B, Choi S, et al. Deep learning in multi-class lung diseases' classification on chest X-ray images. *Diagnostics.* 2022;12(4):915. doi:10.3390/diagnostics12040915
- [4] Alshmrani G, Mufarah M, et al. A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images. *Alexandria Eng J.* 2023;64:923–935. doi:10.1016/j.aej.2022.10.053
- [5] Khan A, et al. A novel threshold-based segmentation method for quantification of COVID-19 lung abnormalities. *Signal Image Video Process.* 2023;17(4):907–914. doi:10.1007/s11760-022-02183-6
- [6] Karar ME, et al. Smart IoMT-based segmentation of coronavirus infections using lung CT scans. *Alexandria Eng J.* 2023;69:571–583. doi:10.1016/j.aej.2023.02.020
- [7] Souza JC, et al. An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks. *Comput Methods Programs Biomed.* 2019;177:285–296. doi:10.1016/j.cmpb.2019.06.005
- [8] Elharrouss O, Subramanian N, Al-Maadeed S. An encoder–decoder-based method for segmentation of COVID-19 lung infection in CT images. *SN Comput Sci.* 2022;3:1–12. doi:10.1007/s42979-021-00874-4
- [9] Guo Y, et al. A review of semantic segmentation using deep neural networks. *Int J Multimed Inf Retr.* 2018;7:87–93. doi:10.1007/s13735-017-0141-z
- [10] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18.* Springer International Publishing, 2015.
- [11] Liu W, et al. Automatic lung segmentation in chest X-ray images using improved U-Net. *Sci Rep.* 2022;12(1):8649. doi:10.1038/s41598-022-12743-y
- [12] Tiwari T, Saraswat M. A new modified-unet deep learning model for semantic segmentation. *Multimed Tools Appl.* 2023;82(3):3605–3625. doi:10.1007/s11042-022-13230-2
- [13] Azad R, et al. Medical image segmentation review: the success of u-net. *arXiv preprint arXiv:2211.14830.* 2022.
- [14] Drozdal M, et al. The importance of skip connections in biomedical image segmentation. *International Workshop on Deep Learning in Medical Image Analysis, International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis.* Springer, Cham, 2016.
- [15] Cai S, et al. Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network. *Quant Imaging Med Surg.* 2020;10(6):1275.
- [16] Zhou Z, et al. Unet++: A nested u-net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4.* Springer International Publishing, 2018.
- [17] Jha D, et al. Doubleu-net: A deep convolutional neural network for medical image segmentation. *2020 IEEE 33rd International symposium on computer-based medical systems (CBMS).* IEEE, 2020.
- [18] Jin Q, et al. DUNet: A deformable network for retinal vessel segmentation. *Knowl Based Syst.* 2019;178:149–162. doi:10.1016/j.knosys.2019.04.025
- [19] Li X, et al. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE Trans Med Imaging.* 2018;37(12):2663–2674. doi:10.1109/TMI.2018.2845918
- [20] Oktay O, et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999.* 2018.
- [21] Luo J, et al. USMLP: U-shaped Sparse-MLP network for mass segmentation in mammograms. *Image Vis Comput.* 2023;137:104761. doi:10.1016/j.imavis.2023.104761
- [22] Zheng B, et al. MSD-Net: Multi-scale discriminative network for COVID-19 lung infection segmentation on CT. *Ieee Access.* 2020;8:185786–185795. doi:10.1109/ACCESS.2020.3027738
- [23] Karaali A, Dahyot R, Sexton DJ. DR-VNet: retinal vessel segmentation via dense residual UNet. *International Conference on Pattern Recognition and Artificial Intelligence.* Cham: Springer International Publishing, 2022.
- [24] Chowdhury ME, et al. Can AI help in screening viral and COVID-19 pneumonia? *Ieee Access.* 2020;8:132665–132676. doi:10.1109/ACCESS.2020.3010287
- [25] Ahmed MR, Ashrafi AF, Ahmed RU, et al. DoubleUNetPlus: a novel attention and context-guided dual U-Net with multi-scale residual feature fusion network for

- semantic segmentation of medical images. *Neural Comput Appl.* 2023;35:14379–14401. doi:10.1007/s00521-023-08493-1
- [26] Rahman T, et al. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Comput Biol Med.* 2021;132:104319. doi:10.1016/j.combiomed.2021.104319
- [27] Rahman T, et al. Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. *IEEE Access.* 2020;8:191586–191601. doi:10.1109/ACCESS.2020.3031384
- [28] Sharma A, et al. Machine learning approach for detection of diabetic retinopathy with improved pre-processing. 2021 International conference on computing, communication, and intelligent systems (ICCCIS). IEEE, 2021.
- [29] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data.* 2019;6(1):1–48. doi:10.1186/s40537-018-0162-3
- [30] Salama WM, Aly MH. Deep learning in mammography images segmentation and classification: automated CNN approach. *Alexandria Eng J.* 2021;60(5):4701–4709. doi:10.1016/j.aej.2021.03.048
- [31] Nayak SR, et al. Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: a comprehensive study. *Biomed Signal Process Control.* 2021;64:102365. doi:10.1016/j.bsp.2020.102365
- [32] Ottoni ALC, et al. Tuning of data augmentation hyperparameters in deep learning to building construction image classification with small datasets. *Int J Mach Learn Cybern.* 2023;14(1):171–186. doi:10.1007/s13042-022-01555-1
- [33] Kugelman J, et al. A comparison of deep learning U-Net architectures for posterior segment OCT retinal layer segmentation. *Sci Rep.* 2022;12(1):14888. doi:10.1038/s41598-022-18646-2
- [34] Wang X, et al. A deep learning method for optimizing semantic segmentation accuracy of remote sensing images based on improved UNet. *Sci Rep.* 2023;13(1):7600. doi:10.1038/s41598-023-34379-2
- [35] Szegedy C, et al. Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2015.
- [36] Dhamija T, et al. Semantic segmentation in medical images through transfused convolution and transformer networks. *Appl Intell.* 2023;53(1):1132–1148. doi:10.1007/s10489-022-03642-w
- [37] Srivastava S, Vidyarthi A, Jain S. Analytical study of the encoder-decoder models for ultrasound image segmentation. *Serv Oriented Comput Appl.* 2023:1–20. doi:10.1007/s11761-023-00373-9