UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias e Ingenierías

Application of the U-Net Architecture for

Lung Segmentation for Tuberculosis Diagnosis

Luis Alejandro Mosquera Berrazueta

Ingeniería en Ciencias de la Computación

Trabajo de fin de carrera presentado como requisito para la obtención del
título de Ingeniero en Ciencias de la Computación

Quito, 4 de enero de 2023

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias e Ingenierías

HOJA DE CALIFICACIÓN
DE TRABAJO DE FIN DE CARRERA

Application of the U-Net Architecture for

Lung Segmentation for Tuberculosis Diagnosis

Luis Alejandro Mosquera Berrazueta

Nombre del profesor, Título académico   Noel Pérez, Ph. D

Quito, 4 de enero de 2023

© DERECHOS DE AUTOR

Nombres y apellidos: Luis Alejandro Mosquera Berrazueta

Código: 00201404

Cédula de identidad: 1723558068

Lugar y fecha: Quito, 4 de enero de 2023

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en http://bit.ly/COPETheses.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on http://bit.ly/COPETheses.

RESUMEN

El pulmón es el órgano más comúnmente afectado en la infección tuberculosa, con estimaciones de afectación pulmonar en sujetos con tuberculosis activa del 79% al 87%. Se trata de una enfermedad infecciosa de transmisión aérea causada por el mycobacterium tuberculosis y es una de las principales causas de morbilidad y mortalidad, sobre todo en los países en desarrollo. Aunque los recientes avances de la medicina han sido la salvación para muchos pacientes, los médicos siguen estando sujetos a errores humanos a la hora de detectar y segmentar estas lesiones en los pulmones. En consecuencia, se han desarrollado modelos de aprendizaje profundo para ayudar a reducir los errores en la industria médica. Con el fin de detectar y segmentar con éxito las lesiones de tuberculosis en un conjunto de datos de imágenes de escáner, sugerimos una nueva técnica de aprendizaje profundo denominada Red-Unet basada en la arquitectura U-net convencional. Tanto la Red-Unet como la U-net se entrenaron utilizando un esquema de validación cruzada estratificada de diez veces bajo las mismas condiciones experimentales, y los resultados mostraron que la Red-Unet funciona tan bien como el modelo de referencia en menos tiempo y con menos requisitos computacionales, lo que es realmente importante. La media y la desviación estándar de las puntuaciones de los coeficientes IoU y DICE para la Red-Unet fueron de 0,91 (0,01) y 0,95 (0), mientras que para la U-net fueron de 0,92 (0,03) y 0,96. Además, el mejor modelo Red-Unet elegido con 100 épocas de entrenamiento se validó en un conjunto de pruebas externo, con una puntuación tan alta como los valores de entrenamiento. La calidad y la capacidad de generalización del modelo quedaron confirmadas por esta evaluación del rendimiento en el conjunto de datos de prueba, que demostró que era capaz de identificar y segmentar con éxito las anomalías pulmonares provocadas por la tuberculosis, independientemente de su tamaño.

Palabras clave: U-Net, Cross validation, Deep-learning, Tuberculosis, Pulmones anormales, Segmentación de imágenes

ABSTRACT

The lung is the most commonly affected organ in tuberculosis infection, with estimates of lung involvement in subjects with active tuberculosis of 79 to 87%. This type of disease is an airborne infectious disease caused by mycobacterium tuberculosis and is a major cause of morbidity and mortality, particularly in developing countries. Although recent advances in medicine have been the salvation for many patients, doctors are still subject to human errors when detecting and segmenting these lesions in the lungs. Deep-learning models have been developed as a result to aid in lowering errors in the medical industry. In order to successfully detect and segment tuberculosis lesions on a data set of scan pictures, we suggested a new deep-learning technique dubbed Red-Unet based on the conventional U-net architecture. Both the Red-Unet and U-net were trained using a stratified ten-fold cross-validation schema under the same experimental conditions, and the results showed that the Red-Unet performs as well as the baseline model in less time and with less computational requirements, which is really important. The mean and standard deviation of IoU and DICE coefficient scores for the Red-Unet were 0.91 (0.01) and 0.95 (0), while for the U-net they were 0.92 (0.03) and 0.96. Additionally, the top-chosen Red-Unet model with 100 training epochs was validated in an outside test set, scoring as highly as the training values. The model's quality and generalization ability were confirmed by this performance evaluation on the test data set, which demonstrated that it was successful in identifying and segmenting lung anomalies brought on by tuberculosis regardless of size.

Key words: U-Net, Cross validation, Deep-learning, Tuberculosis, Abnormal lungs, Image segmentation

TABLE OF CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

INTRODUCTION

The lung is the most commonly affected organ in tuberculosis infection, with estimates of lung involvement in subjects with active tuberculosis of 79 to 87%[1]. Tuberculosis is an airborne infectious disease caused by Mycobacterium tuberculosis and is a major cause of morbidity and mortality, particularly in developing countries, most cases occur in Southeast Asia and Africa[2]. Tuberculosis is an ancient scourge. It has plagued humankind throughout known history and human prehistory. It has surged in great epidemics and then receded, thus behaving like other infectious diseases, but with a time scale that challenges accepted explanations for epidemic cycles. Mycobacterium tuberculosis may have killed more persons than any other microbial pathogen[3].

Thus, the first infection with tuberculosis frequently is clinically insignificant and unrecognized. In the majority of patients, the disease stays dormant either indefinitely or for many years and when a breakdown occurs, it may be secondary to a decrease in body immunity[1].

Different types of diagnosis methods can be applied but not all of them are accurate since the beginning of the disease. Physical examination of the chest is ordinarily of minimal help early in the disease. Routine laboratory examinations are rarely helpful in establishing or suggesting the diagnosis. The chest radiograph is the single most useful study for suggesting the diagnosis of tuberculosis [1].

If TB is detected early and fully treated, people with the disease quickly become noninfectious and eventually cured. Therefore, the imaging diagnosis would provide an appropriate therapy for infected patients before the definitive diagnosis by the bacteriology[2].

In [4] , a fast segmentation for lungs is proposed using two Deep learning models based on the U-Net architecture, the first maintaining the original architecture and the second called E-Net which has modifications for enhanced the performance. The models

were trained with images from their own dataset composed of 42 studies with data augmentation techniques such as rotation, translating in x and y directions and zooming. The best model was E-Net that achieved results of 95.90% Dice similarity coefficient.

In [5], an automated segmentation of Lungs is proposed using a robust model based on the U-Net architecture. The model was trained with images of two datasets that include the Montgomery County, Maryland, and Shenzhen No. 3 People's Hospital in China with data augmentation techniques such as zooming, cropping, horizontal flipping. The best model achieved a 98.6% Dice score.

In [6], a lung CT image segmentation is proposed using the U-Net architecture. The model was trained with LIDC image dataset which contains 1018 lung CTs images. The best model achieved a 95.02% Dice-Coefficient index.

In [7], a deep learning pipeline for segment and classify lung x-rays is proposed. The models used are U-Net and DCNN architectures and transfer learning which are trained by Montgomery Country and Shenzhen China datasets, techniques of data augmentation and image preprocessing were applied. The best U-Net model achieve 98.90% Dice-Coefficient score while the DCNN model achieve 97.1% of accuracy with sensitivity of 97.9%.

In [8], lung segmentation of x-ray images with U-Net + + is proposed. The model used is U-Net + + which looks to replace U-Net because the accuracy and mean_iou are grater, also minimize data leakage. The model was trained with Montgomery County and Shenzhen Hospital datasets, preprocessing techniques of the images were applied in order to use the dataset correctly. The best results were 98% of accuracy and 0.95 mean_iou.

In [9], a method for automatic segmentation of lungs in CXR that addresses this problem by reconstructing the lung regions "lost" due to pulmonary abnormalities is proposed. The models proposed are two deep convolutional neural network that has four main steps, the model is trained with 138 chest x-ray images from Montgomery County's

Tuberculosis Control Program. The best results obtained are sensitivity of 97.54%, specificity of 96.79%, accuracy of 96.97%, Dice coefficient of 94%, and Jaccard index of 88.07%.

Despite the developments in deep learning architectures applying different training techniques there several factors to consider to have a successful performance, so further research is required to meet these standards. In this context, the purpose of this research is a deep learning model derived from the U-Net architecture for the identification and segmentation of lungs with tuberculosis affection in x-rays images scans. The main contributions are relation to modifications in the number of filters in each layer, also dropout layers were added to improve the performance. The research is based on judgement and knowledge, applying the ISO/IEC TR 24372:2021 standard [10] and the development of this research considered the principles of design engineering [11].

MATERIALS AND METHODS

A. Dataset

We employed a lung CT images database from Kaggle, which is a combination of two datasets, Montgomery County X-ray Set and the Shenzhen Chest X-ray Set. This database includes 394 cases with manifestation of tuberculosis and 406 normal cases. Moreover, all the images are in PNG file format and the image size varies for each X-ray, it is approximately 4000x4000 pixels[12].

Database	Tuberculosis cases	Normal cases	Total cases
Montgomery	58	80	138
Shenzhen	336	326	662
Total	394	406	800

Table 1. Database description
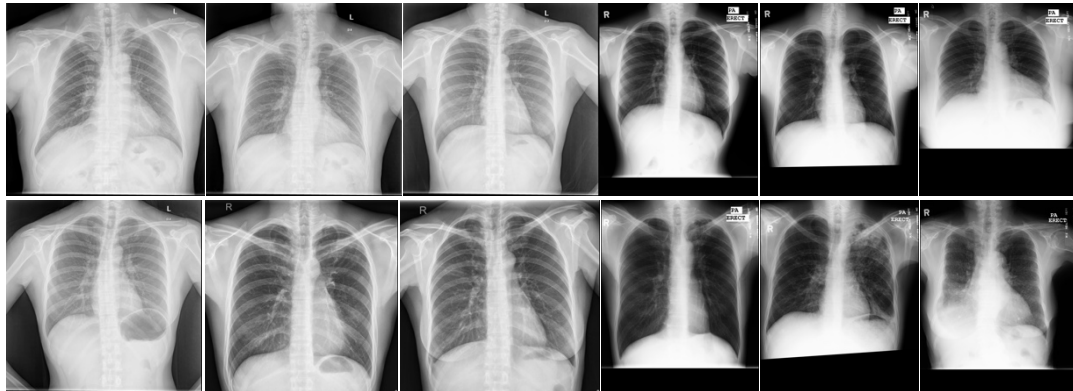
Figure 1. Samples of images of our dataset, left-side shows Shenzhen and right-side shows Montgomery, first row are normal cases, second row Tuberculosis cases[13].

B. Deep learning models

Deep learning is a field derived of machine learning and artificial intelligence (AI) concerned with algorithms based on the function of the brain called artificial neural networks, so it's going to be composed of several neuronal layers that the main

objective is to have a high complexity architecture, in addition to an iterative task and a large quantity and quality data will lead to a progressive learning[14].

Regarding to lesion analysis, the U-Net model is one of the most convenient methods to apply, as is a convolutional network architecture for fast and precise segmentation of images. It consists of down-sampling which is a contracting path and up-sampling which is an expansive path, these gives the appearance as u-shaped architecture. An image is given as an input and it goes though the architecture that consist in the twice application of a 3x3 kernel size convolutional layer followed by a rectified linear unit (ReLU), this is followed by a 2x2 max pooling operation, that consist in the down-sampling of the input in order to reduce it dimensions. During this down-sampling process the objective is to double the number of feature channels. On the other hand, for the up-sampling process we do an 2x2 convolution that consist in reducing to half the number of feature channels, then two 3x3 convolutions are applied followed by a ReLU. At the final layer a 1x1 kernel size convolutional operation is applied to map each vector to the desired number of classes[15].

C. Proposed methods

We decided to modify the original U-Net architecture to develop an improved model that can over perform the baseline model in terms of lung segmentation. Image segmentation essentially aims to extract the silhouette of an interest item from an image. We would eliminate everything that is not important and would be adding noise to the system by concentrating just on the desired object. The lead modifications are the inclusion of batch normalization, dropout layers and one less layer in the architecture which results in the reduction of the number of channels.

Red-Unet. This proposed architecture describes a U-Net-like model. The down-sampling starts with the given image, going through two 3x3 convolutions of 32 feature channels, then a batch normalization layer is applied with the ReLU activation function. The max pooling operation with 2x2 kernel size is applied for

this contracting path, so the result is a reduced dimensions image. The dropout layer improvement with 0.5 probability leads to an overfitting reduction by amplifying the variability at the learning process.
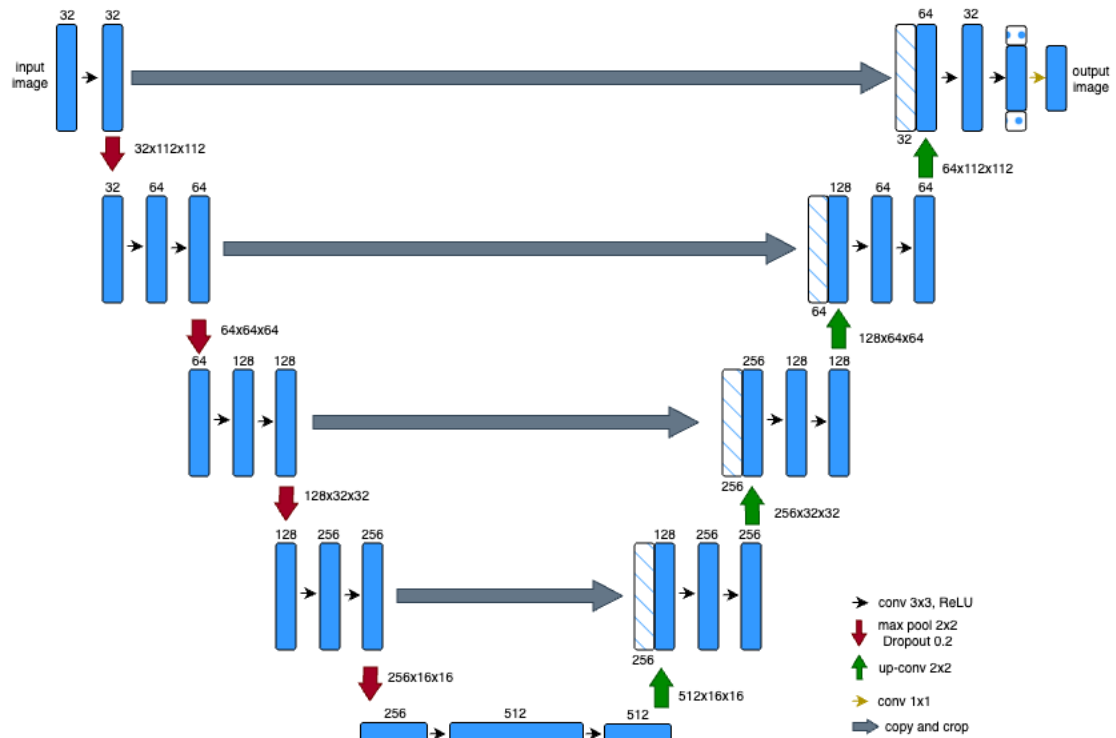

Figure 2. Red-Unet architecture[15].

D. Experimental setup

This section describes the experimental methodology employed to train, validate, and test the baseline U-net and the proposed Red-Unet architectures. The implementation developed to perform the experiments of this research is included in a Github repository[1].

1) Data processing and experimental data set creation. All images were transformed from PNG to JPG (Joint Photographic Expert Group) file format resulting in 800 images. The dataset employed were pretty clean and balanced between abnormal lung and normal lung, containing 634 images with and without Tuberculosis lesions. Also, each image was resized from a range between 3000 x 3000 to 4892 x 4020 to

---

[1] https://github.com/luismosquera99/tesis

256x256 to decrease the processing time in the training task while conserving the lesions' predicting property. For Data augmentation techniques such as horizontal flips, rotations, width and height shifts, zoom and cuts to the images to increase the training set size. Since these operations were performed in a way that no images were saved into the data set, our final created experimental data set consisted of a total of 634 lung images.

2) Training, validation, and test sets. We selected randomly 10% of images of both classes (with and without Tuberculosis) from the data set to form the set that will be used to test the models. The 90% was used for training and validation the model, also for feeding a stratified ten-fold cross. The use of the stratified cross-validation technique is an excellent way to avoid overfitting in the training phase [16].

3) Model configuration. The hyperparameter configuration consisted of a static batch size of 32 and 1000 epochs for both models, which are U-net and Red-Unet. Also, we employed a 10^4 learning rate and applied it to an Adam optimizer [17], mainly we applied this for the benefits in terms of memory requisites, so we reach a computationally efficient train, it's important to emphasize that this is optimal for data with large inputs, outputs, or parameters.

4) Assessment metrics. On each architecture's 10 folds, the mean and standard deviation of the Intersection Over Union (IoU) and Dice coefficient metrics were calculated. These metrics were selected on purpose since image segmentation problems frequently employ them. The IoU metric describes a ratio between the intersection of the predicted mask and the ground truth, and the union of these. This means the intersection of the real and expected masks is divided by their union to get at it. On the other hand, the Dice coefficient is the result of two times the intersection of the predicted[18].

5) Selection criterion. The mean IoU and Dice coefficients throughout the 10 folds were calculated by taking into account checkpoints every 100 epochs for a total of 10 inspected models per architecture (Unet and Red-Unet) during the training phase

(ensuring avoiding an exhaustive search in the whole learning space). The most effective model is the one that has the highest punctuation in both of the aforementioned measures.

RESULTS AND DISCUSSION

This section presents and discusses the evaluation of two different models on an experimental data set, containing 634 images with and without lung abnormalities in conjunction with a stratified ten-fold cross-validation schema. Table II contains the results that were derived using the mean of the IoU, Dice, and Dice loss metrics.

TABLE II: Performance results of deep learning models.

Architecture	Conv. layer (f)	Kernel size	Pool size per layer	Dropout factor	Batch size	Epochs (u)	Train IoU (u)	Train DICE (u)	Validation IoU(STD) (u)	Validation Dice(STD) (u)
U-Net	$(64, 128, 256, 512, 1024)$	(3×3)	(2×2)	NA		100	0.94	0.97	0.90(0.04)	0.95(0.02)
						200	**0.95**	**0.97**	**0.92(0.03)**	**0.96(0.01)**
						300	0.96	0.98	0.87(0.15)	0.92(0.11)
						400	0.97	0.99	0.54(0.37)	0.62(0.36)
					16	500	0.98	0.99	0.50(0.37)	0.58(0.37)
						600	0.98	0.99	0.73(0.25)	0.81(0.19)
						700	0.98	0.99	0.65(0.33)	0.73(0.29)
						800	0.98	0.99	0.63(0.28)	0.73(0.27)
						900	0.98	0.99	0.56(0.31)	0.67(0.30)
						1000	0.98	0.99	0.60(0.33)	0.70(0.28)
Red-UNet	$(32, 64, 128, 256, 512)$	(3×3)	(2×2)	0.2		**100**	**0.93**	**0.96**	**0.91(0.01)**	**0.95(0)**
						200	0.94	0.97	0.84(0.12)	0.91(0.07)
						300	0.94	0.97	0.53(0.18)	0.68(0.15)
						400	0.95	0.98	0.46(0.21)	0.61(0.17)
					16	500	0.96	0.98	0.41(0.19)	0.56(0.16)
						600	0.96	0.98	0.43(0.23)	0.58(0.19)
						700	0.96	0.98	0.35(0.11)	0.51(0.11)
						800	0.96	0.98	0.32(0.06)	0.48(0.07)
						900	0.97	0.98	0.36(0.14)	0.52(0.12)
						1000	0.97	0.98	0.33(0.06)	0.50(0.07)

Conv.- convolutional; f- number of filters per layer; u- units; Train IoU; Train DICE, Validation IoU, Validation DICE - mean of training and validation metrics DICE and IoU over ten folds; STD - standard deviations the respective metric.

A. Performance evaluation in the training set

From Table II, we can appreciate that both the Dice and IoU scores obtained by the proposed Red-Unet were the pretty good since the first epochs and maintain excellent values in all 1000 epochs, which can be described as a normal behavior since deep learning models requires several epochs with large data to improve and stabilize values, the epochs require several iterations to appropriately adequate their internal parameters. Additionally, we can determine that the training values for both metrics are higher for the U-net model than the proposed Red-Unet model. The best performances in terms of high mean Dice and IoU validation scores and low standard deviations are shown in bold in Table II (one for each architecture). In comparison to the actual data, the mean scores demonstrate the models' ability to segment tuberculosis accurately, and the standard deviations reflect how closely all the data from each fold was collected to the corresponding mean scores. The U-net architecture showed its best performance with 200 epochs, obtaining values of 0.95 for IoU and 0.97 for Dice. In the case of the Red-Unet model, the highest performance was obtained with 100 epochs, where the mean scores (0.93 and 0.96 for IoU and Dice respectively) were maximized with minimal standard deviations.

Despite Table II shows that the scores of the original U-net are better in an 0.01 in both metrics than the proposed method, we can see Red-Unet starts showing accurate predictions and its curve its excellent since the first epoch which means that less computational requirements and time to achieve excellent results thus predicts. This conveys that with less training iterations, Red-Unet has the tendency to enhance Tuberculosis predictions and better the performance of the detector in comparison to its baseline model. In this context and comparing the best models of both architectures previously mentioned,

we established that the best overall performance and the best- selected model was Red-Unet with 100 epochs (underlined in Table II) since it has almost the same performance but in less time and with less computational requirements. These excellent scores, together with the best picked model's low standard deviations, show us to have strong generalization and prediction abilities.

Figure 4. Performance of the best-chosen model regarding our Dice

loss function

In Figure 4, we can see the behavior of the best selected model through a training and validation loss graph. This graph indicates us that the training and validation loss values decreased in a similar manner and reached a value of less than 0.05. Even thought, the validation values start raising after a certain epoch, this will be discussed in the next section. This similarity of the curves shows us that overfitting was avoided successfully, during the first hundred epochs; the training process was aided by the use of dropout layers and the reduction of the model's complexity.

B. Performance evaluation in the test set

We used the Red-Unet trained with 100 epochs, which was the best model selected in the training step, to test its generalization power on an external test set, containing 72 images with and without Tuberculosis. The obtained results of 0.91 and 0.95 for the IoU and Dice coefficient highlighted the promising performance of the proposed method in segmenting lung abnormalities on unseen lung scans. An example of successful lung abnormalities segmentation cases of the test set is shown in Fig. 4.
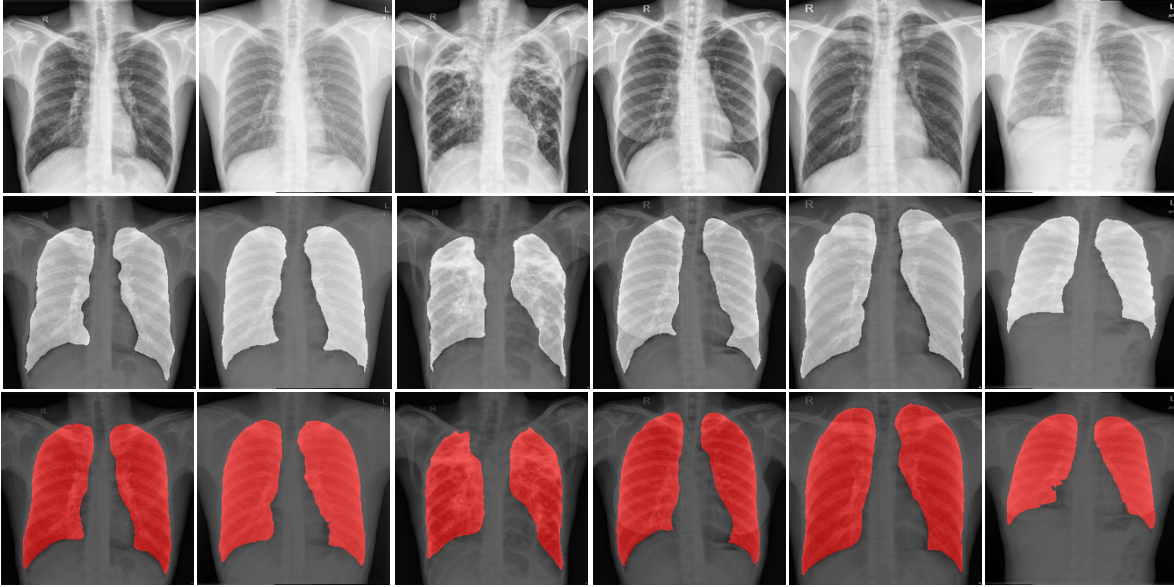

Figure 4. Images from our test set. The first row contains the original Tuberculosis images, the second row contains the same Tuberculosis images with their respective ground truth, and the third row contains the images with our predicted masks.

To clarify the doubts of why the metrics in test set of the U-Net and Red-Unet start to decrease in meaning of IoU and Dice coefficient after the epoch 250 in Red-Unet and 330 U-Net approximately it is because the model is so accurate, the dataset has a high quality images and the computer in which the models are being training is powerful this leads to both models obtain excellent results in training and validation set during the first hundred epochs, this change mainly because of the size of the test set, which contains really few images compared with the train set where we don't see an overfitting problem[16]. In spite of, we can determine that both are excellent because validation set achieves really high (over 0.95) metrics, we can solve this with more images in the validation set or applying techniques like early stopping to the models[19], but one objective of this investigation was achieve the 1000 epochs to see the behavior of the model, if we applied this technique the training would have stopped around the 300 epochs because the loss value stop improving after that, so we could conclude that the model already achieve the maximum best performance it can for this dataset [19], this is excellent because the model can be train and generalized successfully with not too much computational requirements and time, which are two main problems in the Deep Learning area.

CONCLUSIONS AND FUTURE WORK

In order to successfully identify and segment Tuberculosis anomalies on a data set of lung scan pictures, we suggested a new deep-learning technique dubbed Red-Unet based on the conventional U-net architecture. Both the Red-Unet and U-net were trained using a stratified ten-fold cross-validation schema under the same experimental conditions, and the results showed that the Red-Unet performs as well as the baseline model in less time and with less computational requirements, which is really important. The mean and standard deviation of IoU and DICE coefficient scores for the Red-Unet were 0.91 (STD 0.01) and 0.95 (STD 0), while for the U-net they were 0.92 (STD 0.03) and 0.96 (STD 0.01). Additionally, the top-chosen Red-Unet model with 100 training epochs was validated in an outside test set, scoring as highly as the training values. The model's quality and generalization ability were confirmed by this performance evaluation on the test data set, which demonstrated that it was successful in identifying and segmenting lung anomalies brought on by tuberculosis regardless of size.

We intend to implement additional deep-learning iterations of the suggested technique in the future, enhancing the architecture and parameter settings. We will also consider alternatives that involve combining different deep-learning models. In order to repeat the process and find more generalizable models, we wish to look for various and larger databases.

ACKNOWLEDGMENT

REFERENCES

[1]   L. S. M and R. M. D, "Pulmonary Tuberculosis," *Microbiol Spectr*, vol. 5, no. 1, p. 5.1.24, Feb. 2017, doi: 10.1128/microbiolspec.TNMI7-0032-2016.

[2]   Y. Jeong and K. Lee, "Pulmonary Tuberculosis: Up-to-Date Imaging and Management," *AJR Am J Roentgenol*, vol. 191, pp. 834–844, Oct. 2008, doi: 10.2214/AJR.07.3896.

[3]   T. M. Daniel, "The history of tuberculosis," *Respir Med*, vol. 100, no. 11, pp. 1862–1870, 2006, doi: https://doi.org/10.1016/j.rmed.2006.08.006.

[4]   A. Comelli *et al.*, "Lung Segmentation on High-Resolution Computerized Tomography Images Using Deep Learning: A Preliminary Step for Radiomics Studies," *J Imaging*, vol. 6, no. 11, 2020, doi: 10.3390/jimaging6110125.

[5]   J. Islam and Y. Zhang, "Towards Robust Lung Segmentation in Chest Radiographs with Deep Learning," *ArXiv*, vol. abs/1811.12638, 2018.

[6]   B. Ait Skourt, A. el Hassani, and A. Majda, "Lung CT Image Segmentation Using Deep Neural Networks," *Procedia Comput Sci*, vol. 127, pp. 109–113, 2018, doi: https://doi.org/10.1016/j.procs.2018.01.104.

[7]   C. Dasanayaka and M. Dissanayake, *Deep Learning Methods for Screening Pulmonary Tuberculosis Using Chest X-rays*. 2020.

[8]   S. Gite, A. Mishra, and K. Kotecha, "Enhanced lung image segmentation using deep learning," *Neural Comput Appl*, 2022, doi: 10.1007/s00521-021-06719-8.

[9]   J. C. Souza, J. O. Bandeira Diniz, J. L. Ferreira, G. L. França da Silva, A. Corrêa Silva, and A. C. de Paiva, "An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks," *Comput Methods Programs Biomed*, vol. 177, pp. 285–296, 2019, doi: https://doi.org/10.1016/j.cmpb.2019.06.005.

[10]   ISO, "Iso/iec tr 24372:2021(en) - international organization for standardization," *https://www.iso.org/obp/ui/ #!iso:std:78508:en*, 2021.

[11]   Y. and S. S. and S. T. M. Haik, *Engineering design process*. Cengage Learning, 2015.

[12]   K. S. Mader, "Pulmonary chest X-ray abnormalities," *Pulmonary chest X-ray abnormalities*, Mar. 09, 2019.

[13]   S. Urooj, s Suchitra, L. Krishnasamy, N. Sharma, and N. Pathak, "Stochastic Learning-Based Artificial Neural Network Model for an Automatic Tuberculosis Detection System Using Chest X-Ray Images," *IEEE Access*, vol. PP, p. 1, Jan. 2022, doi: 10.1109/ACCESS.2022.3208882.

[14]   Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[15]   O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *CoRR*, vol. abs/1505.04597, 2015, [Online]. Available: http://arxiv.org/abs/1505.04597

[16]   B. Ghojogh and M. Crowley, *The Theory Behind Overfitting, Cross Validation, Regularization, Bagging, and Boosting: Tutorial*. 2019.

[17]   D. Yi, J. Ahn, and S. Ji, "An Effective Optimization Method for Machine Learning Based on ADAM," *Applied Sciences*, vol. 10, no. 3, 2020, doi: 10.3390/app10031073.

[18] Z. Wang, E. Wang, and Y. Zhu, "Image segmentation evaluation: a survey of methods," *Artif Intell Rev*, vol. 53, no. 8, pp. 5637–5674, 2020, doi: 10.1007/s10462-020-09830-9.

[19] X. Ying, "An Overview of Overfitting and its Solutions," *J Phys Conf Ser*, vol. 1168, p. 22022, Feb. 2019, doi: 10.1088/1742-6596/1168/2/022022.