

Segmentation of Pneumothorax on Chest CTs Using Deep Learning Based on Unet-Resnet-50 Convolutional Neural Network Structure

Adem Gencer ¹ , Yasin İlder Toker ² ¹ Department of Thoracic Surgery, Afyonkarahisar Health Sciences University, Faculty of Medicine, Afyonkarahisar, Turkey² Department of Thoracic Surgery, Afyonkarahisar State Hospital, Afyonkarahisar, Turkey

Received: 2024-01-22 / Accepted: 2024-02-25 / Published Online: 2024-02-25

Correspondence

Adem Gencer, Assistant Professor,
Department of Thoracic Surgery,
Afyonkarahisar Health Sciences
University, Faculty of Medicine,
Afyonkarahisar, Turkey
Address: Zafer Sağlık Külliyesi,
Dörtüol Mah. 2078 Sok. No:3 A Blok
Afyonkarahisar, Turkey
E-mail: dr.ademgencer@gmail.com

© 2024, European Journal of Therapeutics,
Gaziantep University School of Medicine.



This work is licensed under a Creative
Commons Attribution-NonCommercial 4.0
International License.

ABSTRACT

Objective: Pneumothorax refers to an abnormal accumulation of air in the pleural cavity. This condition is significant in terms of health and can provide a life-threatening risk, particularly when it is extensive or occurs alongside other medical conditions. Nevertheless, the scarcity of work on chest CT segmentation arises from the challenge of acquiring pixel-level annotations for chest X-rays. This paper presents and assesses a deep learning approach utilizing the Unet-Resnet-50 convolutional neural network architecture for accurately segmenting pneumothoraces on chest computed tomography (CT) images.

Methods: We employed a private dataset including 2627 manually annotated slices obtained from 16 patients. We assessed the model's performance by measuring the dice similarity coefficient (DSC or F1 score), accuracy, area under the curve (AUC), precision, and recall on both the validation and test sets.

Results: The binary accuracy of the test set was 0.9990; the precision was 0.9681; and the DSC was 0.9644. Although it contains less data (16 patients), we found that our deep learning-based artificial intelligence model has effective and compatible results with the literature.

Conclusion: Deep learning models that will be used to detect common pathologies in thoracic surgery practice, such as pneumothorax, to determine their localization and size, will provide faster diagnosis and treatment to patients, and especially improve radiology workflow.

Keywords: Pneumothorax segmentation, Deep learning, Convolutional neural networks, Medical imaging, Artificial intelligence

INTRODUCTION

A pneumothorax, characterized by an unusual accumulation of air in the pleural cavity, is a significant medical issue that can be life-threatening when it is extensive or occurs alongside other medical conditions [1].

Chest X-rays are typically used to diagnose patients with pneumothorax, as they can offer a rough estimate of the volume

of pneumothorax. Nevertheless, the assessment of volume in a single X-ray image is inaccurate and lacks consistency. Moreover, distinguishing pneumothorax from similar lung conditions as bullae and emphysema using X-rays is challenging. Consequently, chest computed tomography (CT) scans are now frequently employed to resolve the previously described concerns [2].

In the field of medical image processing, convolutional neural networks (CNNs) have demonstrated significant promise as deep learning-based technologies [3]. Fully convolutional networks, along with their subsequent extensions such as U-Net, have demonstrated remarkable efficacy in the segmentation of medical images, even when trained with a limited quantity of labeled data [3,4]. U-Net and U-Net-like models have shown effective in segmenting many anatomical structures such as the lungs, pulmonary nodules, clavicles, brain, heart, and prostate [1–3,5]

Various U-Net-inspired convolutional neural network (CNN) techniques have been suggested for pneumothorax identification using image-level annotation (classification) on chest X-rays [1,3,6–16]. Lesion semantic segmentation in medical imaging is a crucial tool for facilitating lesion analysis and therapy planning. Nevertheless, the scarcity of work on pneumothorax segmentation on chest x-rays is attributed to the challenges associated with acquiring pixel-level annotations for this condition [1].

The speed of pneumothorax segmentation on a chest x-ray is relatively high, but its accuracy may be worse when compared to segmentation on a CT scan [17]. Performing manual volume assessment on CT scans is a time-consuming task that is possible to carry out in clinical practice [18]. While there are several publications that utilize deep learning algorithms to segment the pneumothorax area on chest x-rays [1,5,19,20], the number of papers that focus on chest CT and employ deep learning, particularly convolutional neural networks, for pneumothorax segmentation is significantly lower [2,21].

The objective of this project was to create a deep learning

Main Points:

- Pneumothorax, an abnormal air accumulation in the pleural cavity, is a significant health problem and can be life-threatening.
- Deep learning, particularly convolutional neural networks (CNNs) like U-Net, shows promise in medical image analysis.
- This research successfully proposes a UNet CNN based on ResNet-50 for pneumothorax segmentation on chest CTs.

method using a convolutional neural network to accurately perform semantic segmentation on chest CT scans. Maybe this will provide pneumothorax patients with a faster diagnosis and an earlier treatment chance.

MATERIALS AND METHODS

The local committee has granted approval for the retrospective analysis of the imaging data for the study (Code: 2011-KAEK-2 No: 2023/2).

Dataset

Routine chest CTs of patients diagnosed with pneumothorax in the thoracic surgery department of the university hospital between 2021 and 2022 were obtained from the hospital's PACS (Picture Archiving Communication Systems). Following the anonymization process of DICOM (Digital Imaging and Communications in Medicine) files, a random selection of chest CTs from 16 patients was made from the acquired CTs. Two skilled thoracic surgeons utilized the 3D-Slicer, an open-source software, to carefully identify and delineate pneumothorax regions at the pixel level for each axial slice, creating ground truth annotations.

A total of 3737 slices were obtained from 16 individuals, resulting in an average of 164 slices having pneumothorax per patient. The distribution of CT slices is shown in Figure 1. The dimensions of all slices were 512x512 pixels, and the thickness of each slice was 2.5 millimeters.

Pre-processing

The DICOM images in the dataset undergo preprocessing through normalization, which involves adjusting the Hounsfield unit (HU) values to a range of 0 to 1. This is done to enhance the performance of the model and avoid the influence of excessively high values.

Creation of Training and Test Sets

The dataset was randomly divided into three sets using the `train_test_split` method from the Sklearn package. The training set consisted of 2391 slices, which accounted for 64% of the dataset. The validation set contained 598 slices, representing 16% of the dataset. The remaining 20% of the dataset was allocated to the test set, which consisted of 748 slices. The training set was used for model training. Hyper-parameter optimization was conducted on the validation set. The test set was only employed once, solely for the ultimate assessment of the current model's

precision. The table provided, Table 1, displays the data and pixel distribution throughout the training, validation, and test sets.

Model Architecture

The model architecture was constructed utilizing a Python library that is open-source and relies on Keras and TensorFlow. This library, known as segmentation models, specializes in picture segmentation using neural networks [22]. A convolutional neural network (CNN) model was constructed using the UNET architecture and the Resnet-50 encoder backbone. Figure 2 displays the structure of a Resnet-50 neural network model. The model comprises a grand total of 32,554,836 parameters, out of which 32,507,282 are eligible for training. The Adam optimizer was utilized with a learning rate of 1e-4. For model training, a hybrid of Binary Cross-Entropy (BCE) and Jaccard loss was employed. The training process was terminated prematurely using the validation set loss after 8 epochs, when patience was set to that value. The model training process was initiated with

1000 epochs and a batch size of 64.

Experimental Setup

Training was carried out on Google-Colaboratory® using Python 3.8.16 and Keras 2.9.0 with Tensorflow 2.9.2 as a backend on Google Compute Engine with Graphical Processing Unit (GPU).

Evaluation Metrics

In order to assess the model’s performance, we employed various metrics including the dice similarity coefficient (DSC or F1 score), accuracy, area under the curve (AUC), precision, and recall. These metrics were applied to both the validation and test sets. The outcomes of each pixel were tracked, distinguishing between true positives, false positives, true negatives, and false negatives. The dice similarity coefficient (DSC) is the primary parameter used to validate and analyze the performance of medical image segmentation [23].

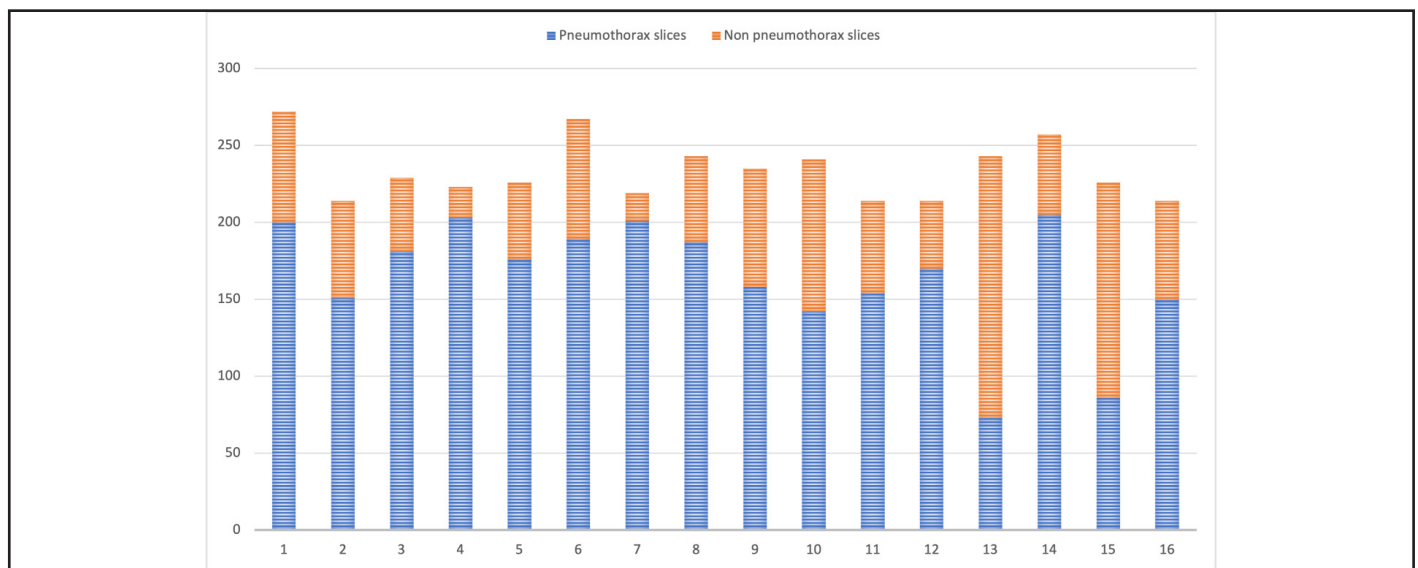


Figure 1. Distrubution of CT slices

Table 1. Data and pixel distrubution of dataset

	Training set	Validation set	Test set	Predicted	Total
Total slices	2391	598	748	748	3737
Non-pneumothorax slices	717	164	229	239	1110
Pneumothorax slices	1674	434	519	509	2627
Average pneumothorax pixels per slice	5244.37	5341.12	5405.43	5527.56	5292.17
Total pneumothorax pixels	8779077	2318050	2805419	2813530	13902546
Total pixels	626786304	156762112	196083712	2813530	979632128

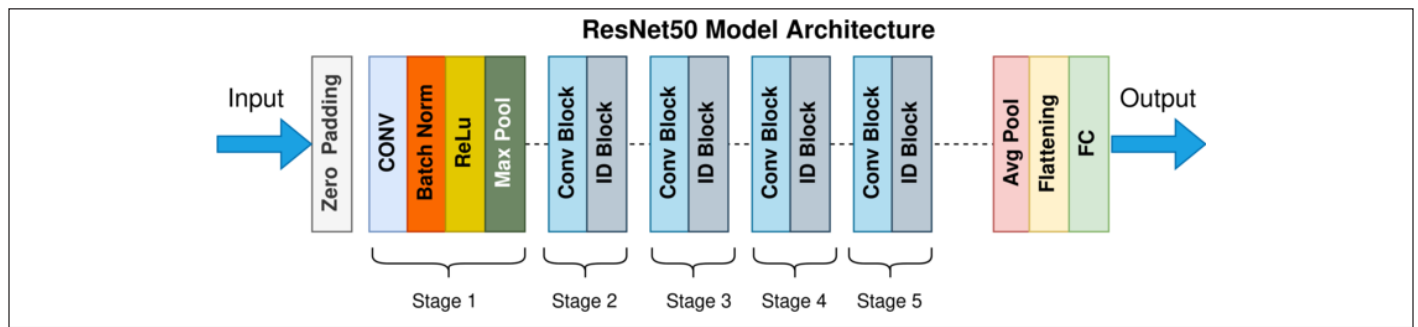


Figure 2. Sample ResNet neural network model structure

RESULTS

The training was completed at the 57th epoch as a result of early stopping. A total of 25.3 GB CPU Ram and 38.9 GB GPU Ram were used for training. The execution time for the training was approximately 46 minutes and the computations for the test set took 31 seconds on the same Google-Colaboratory® environment.

Following the completion of training and hyper-parameter tweaking, the model’s performance was assessed using a multitude of indicators. Table 2 displays the model’s segmentation performance on the validation and test sets. Figure 3 displays the incremental improvement in performance with each period of the training process.

The binary accuracy of both the validation and test sets was 0.9990, with a validation loss of 0.0729 and 0.0757, respectively. The area under the receiver operating characteristic curve (AUC)

was 0.9962 and the Dice similarity coefficient (DSC) was 0.9657 in the validation set. In our training and validation set, we assess the true positive, false positive, true negative, and false negative values for each pixel.

The evaluation of the segmentation performance was also conducted on a separate test set. The test set comprises data that the artificial intelligence model has not been exposed to during its training phase. The precision on the testing set was 0.9681, the recall was 0.9648, and the dice coefficient was 0.9644.

Figure 4 shows instances of segmentation outcomes for dicom pictures obtained from random slices. The first image represents the unmodified CT slice, the subsequent image shows the exact mask, and the last image demonstrates the predicted mask. Additionally, an image is included that displays both the accurate annotation of the ground truth and the predicted mask.

Table 2. Segmentation performance of the model

Metric	Training set	Validation test	Test set
Loss	0.0407053	0.0729687	0.0757
AUC	0.996446	0.996279	0.7265
Binary Accuracy	0.999513	0.999054	0.9990
Binary IoU	0.982674	0.982674	0.9669
Mean IoU	0.690134	0.707968	0.6883
Precision	0.982771	0.978086	0.9681
Recall	0.982476	0.984776	0.9648
True Positive	8625236	2242176	4943309
True Negative	617856000	154367216	347559360
False Positive	151207	76839	162979
False Negative	153841	75874	180160
F1 score (DSC)	0.980604	0.965768	0.9644
IoU score (Jaccard index)	0.961969	0.933909	0.9313

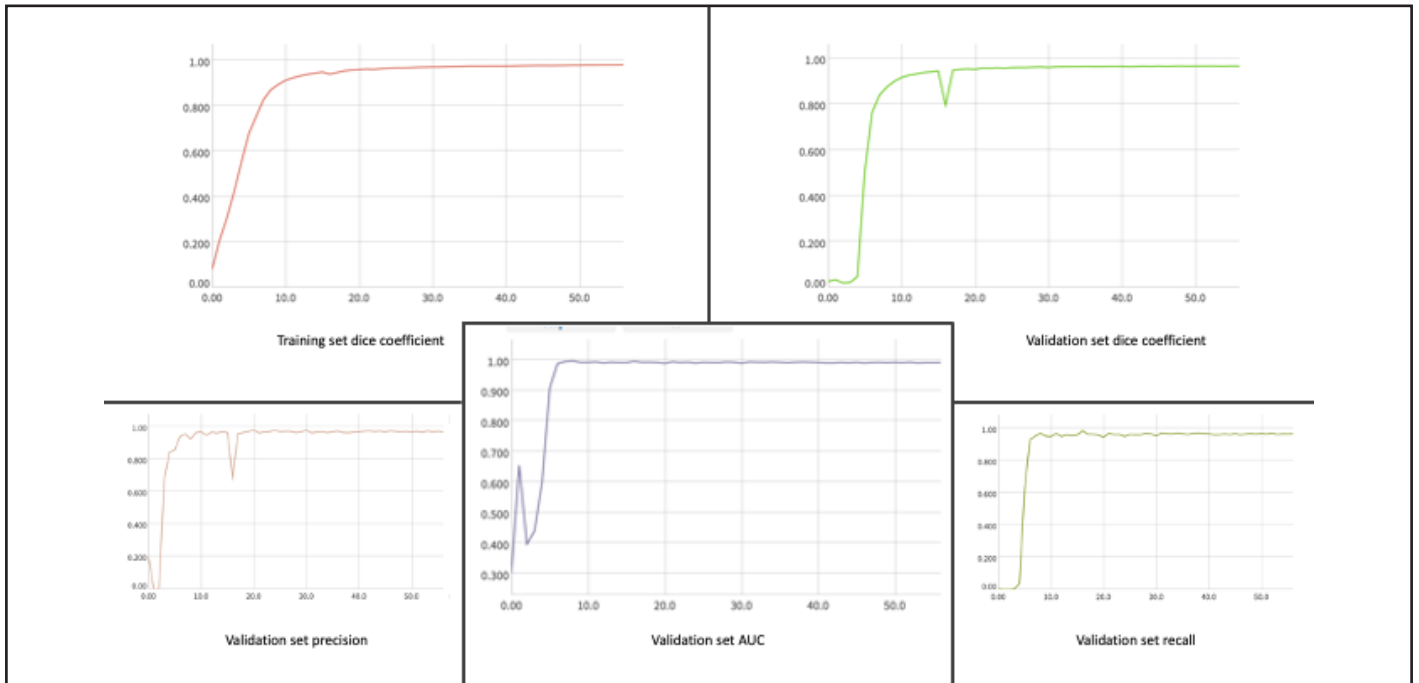


Figure 3. Performans metrics of deep learning model

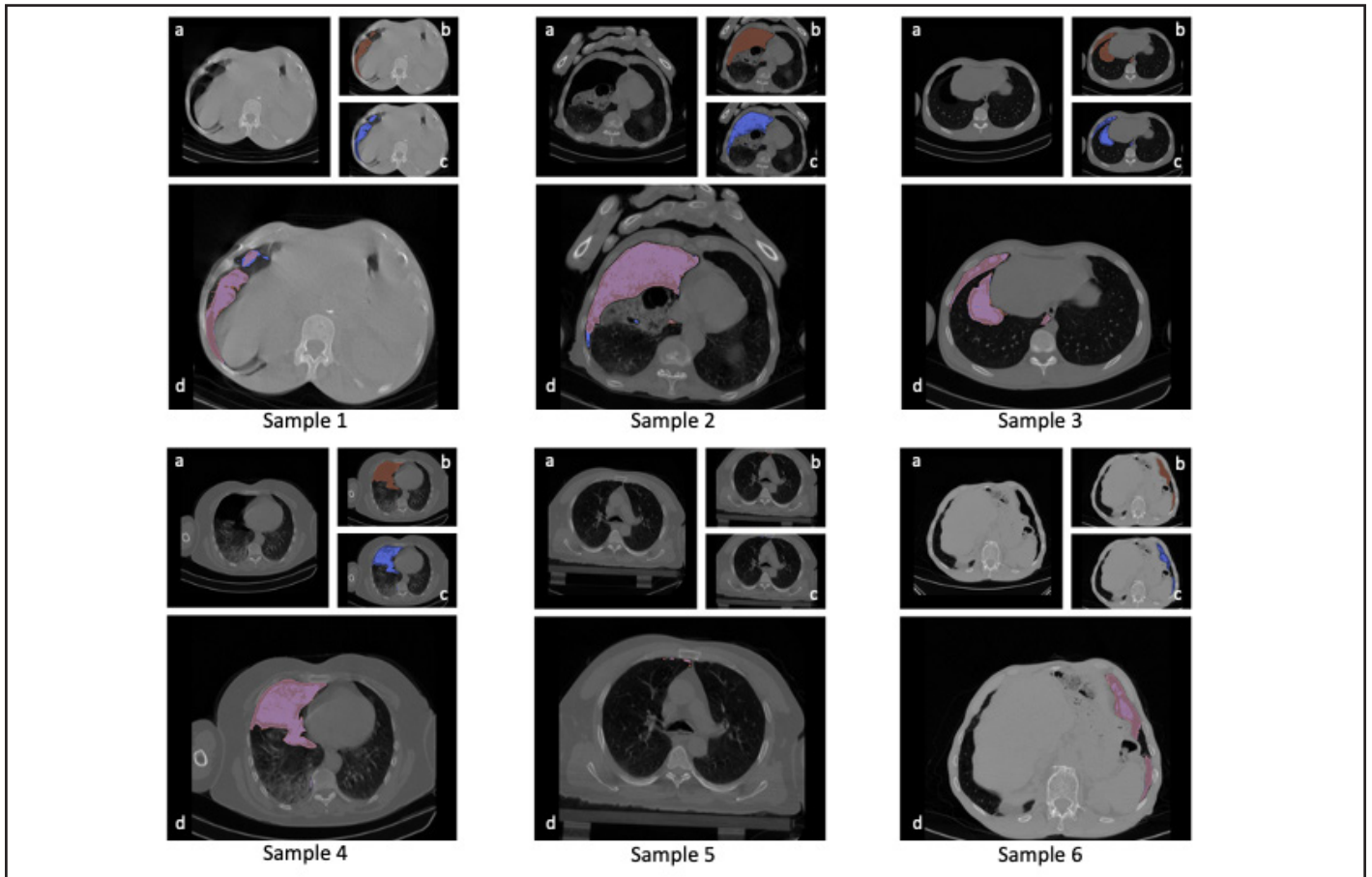


Figure 4. Example CT scans and corresponding segmentations. a. Initial CT image b. Accurate annotations c. Segmentation prediction d. Merged ground truth and predicted segmentations.

DISCUSSION

This paper introduces and assesses a deep learning approach utilizing the Unet-Resnet-50 convolutional neural network architecture to segment pneumothorax on chest CT scans. We utilized a private dataset including 2627 carefully annotated slices obtained from 16 patients diagnosed with pneumothorax. The dataset contained a total of 3737 slices.

The test set evaluation yielded a binary Intersection over Union (IoU) of 0.9669 and a dice score of 0.9644 for this model. Röchrich et al. [21] present their segmentation model based on UNet architecture, achieving an average precision of 0.97 and a dice similarity coefficient of 0.94. The model was evaluated using a private dataset consisting of 2487 annotated slices from 43 patients. They marked one slice out of every five to ten slices and automatically filled in the missing slices in between. In contrast, we carefully annotate each every slice in a manual manner to create accurate ground-truth annotations. The improvement in our dice coefficient may be attributed to this factor.

The study conducted by Wu et al. [2] focuses on the segmentation of pneumothorax on CT sections. In this study, they performed a comparison of 9 distinct convolutional neural networks that resemble UNet architecture. Their confidential dataset comprises 12,535 CT slices obtained from 60 patients diagnosed with pneumothorax. The researchers examine various neural networks, such as Unet, Resnet-Unet, MultiRes-Unet, MFP-Unet, Dilated-Unet, Attention-Unet, and PSP-Net. The Unet++ model yielded the highest dice coefficient scores, achieving 0.9938 for a big pneumothorax and 0.9592 on average. The results of our investigation indicate that the validation set achieved a dice coefficient score of 0.9657, while the test set had a slightly lower value of 0.9644. According to our perspective, the primary determinant influencing the model's performance is the precision with which the ground truth annotation is conducted. For ground truth annotation, each individual pixel must be determined independently. Identifying whether each pixel falls inside the pneumothorax region is challenging, particularly in CT scans that contain artifacts and at the corners.

Ibtehaz et al. [4] sought to enhance spatial resolution by incorporating a third dimension into 2D images. They established a 3D Unet model called MultiResUnet and evaluated its performance on five distinct heterogeneous datasets, including magnetic resonance images from the BraTS17 database. Wu et al. [2] conducted a study on pneumothorax CT scans using

the MultiResUnet model. They discovered that MultiResUnet achieved the highest level of accuracy in segmenting big pneumothorax. However, the improvement in the average dice coefficient was only small, with values of 0.9866 and 0.9592, respectively. Reiterating, we achieved nearly identical outcomes in the model we employed, with a dice coefficient of 0.9644 in the test set.

Deep learning models developed on medical images with artificial intelligence are basically divided into two groups. In classification problems, it is tried to determine whether the relevant pathology is present in the medical image given. In segmentation problems, besides the presence of pathology, the location of the pathology is tried to be determined. In order to perform the segmentation, the artificial intelligence model makes a separate prediction for each pixel in the image.

There are many studies in the literature about x-ray classification of pneumothorax [6–16]. Deep learning studies using DICOM CT images for pneumothorax segmentation are very rare due to the intense processing power required. Each thorax CT series contains approximately 200 tomography sections with 512*512 resolution (for 2.5 mm thickness). This means that approximately 52,428,800 pixels are calculated separately for a patient. Some studies on segmentation of other organs and pathologies are available in the literature [3]. However, apart from the studies mentioned in detail above [2,4,21], we could not find any research in the literature that tried to develop a CNN model that could segment pneumothorax using CT slices. It is likely that as the chance of researchers to access high processing power increases, artificial intelligence studies using medical images such as CT, MR and PET, which require many processing, will increase.

Several research have conducted pneumothorax segmentation on x-ray images [1,1,5,20]. Wang et al. [5] created a convolutional neural network named CheXLocNet, which is built upon Mask R-CNN. They applied this model to analyze 12047 X-ray pictures from the “SIIM-ACR Pneumothorax Segmentation Competition” database on Kaggle®. The metrics assessed include AUC, F1 score, sensitivity, specificity, and positive predictive value (0.86, 0.64, 0.82, 0.92, and 0.65, respectively).

In their publication, Wang et al. [1] present an article on ScSE-Dense Net, a convolutional network that is fully implemented. The model was tested using a 11051 chest x-ray, which consisted

of a 5566 case of pneumothorax and a 5485 case of non-pneumothorax. The segmentation analysis yielded a pixel-wise accuracy of 0.93 and a dice similarity coefficient (DSC) of 0.92. The obtained results exhibit higher performance with values of 0.9990 and 0.9644, respectively.

Feng et al. [19] performed a CNN-based artificial intelligence model on the CANDID-PTX dataset, which contains 19237 chest x-rays of the New Zealand population. DeepLabV3+ and Effientnet-B3, the models with the best segmentation performance they found, have lower dice coefficient than ours (0.91, 0.9644, respectively).

Limitations

In the field of artificial intelligence, the effectiveness of models is directly related to the magnitude of the dataset. Despite utilizing a dataset with a smaller amount of data compared to other studies [16], our achievement rates were on the same level as those of analogous studies documented in the literature. The reason why this gap is bridged is due to the presence of similar patterns in medical imaging.

CONCLUSIONS

This paper introduces a UNet convolutional neural network that utilizes a ResNet-50 backbone for the purpose of segmenting pneumothorax on chest CT scans. Experiments on our chest CT dataset show that our deep learning-based artificial intelligence model has results effective and compatible with the literature. Studies on the segmentation of medical images allow to better distinguish pathological events from normal anatomical tissues and focus on this area. In addition, deep learning models that will be used to detect common pathologies in thoracic surgery practice, such as pneumothorax, to determine their localization and size, will provide faster diagnosis and treatment to patients, and especially improve radiology workflow. Further studies on the subject are needed to develop better models especially on 3D images likes CT and MRI.

Funding: No funding was received for this work.

Data availability statement: Model structure and python codes are available publicly at Github Repository (***) . Due to its sensitive nature the dataset (DICOM images and pixel data) analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Authorship confirmation/contribution statement: G.A.: Conceptualization, Methodology, Software, Data Curation, Writing - Original Draft. T.Y.I.: Investigation, Resources, Data Curation, Writing - Review & Editing.

Conflict of Interest: We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

REFERENCES

- [1] Wang Q, Liu Q, Luo G, Liu Z, Huang J, Zhou Y, et al. Automated segmentation and diagnosis of pneumothorax on chest X-rays with fully convolutional multi-scale ScSE-DenseNet: a retrospective study. *BMC Med Inform Decis Mak.* 2020 Dec;20(S14):317. <https://doi.org/10.1186/s12911-020-01325-5>
- [2] Wu W, Liu G, Liang K, Zhou H. Pneumothorax Segmentation in Routine Computed Tomography Based on Deep Neural Networks. In: 2021 4th International Conference on Intelligent Autonomous Systems (ICoIAS) [Internet]. Wuhan, China: IEEE; 2021:78–83. <https://doi.org/10.1109/ICoIAS53694.2021.00022>
- [3] Malhotra P, Gupta S, Koundal D, Zaguia A, Enbeyle W. Deep Neural Networks for Medical Image Segmentation. Chakraborty C, editor. *Journal of Healthcare Engineering.* 2022 Mar 10;2022:1–15. <https://doi.org/10.1155/2022/9580991>
- [4] Ibtihaz N, Rahman MS. MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Networks.* 2020 Jan;121:74–87. <https://doi.org/10.1016/j.neunet.2019.08.025>
- [5] Wang H, Gu H, Qin P, Wang J. CheXLocNet: Automatic localization of pneumothorax in chest radiographs using deep convolutional neural networks. Xie H, editor. *PLoS ONE.* 2020 Nov 9;15(11):e0242013. <https://doi.org/10.1186/s12911-020-01325-5>
- [6] Taylor AG, Mielke C, Mongan J. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study. Saria S, editor. *PLoS Med.* 2018 Nov 20;15(11):e1002697. <https://doi.org/10.1371/journal>

- [pmed.1002697](https://doi.org/10.1007/s10140-020-01767-4)
- [7] Yi PH, Kim TK, Yu AC, Bennett B, Eng J, Lin CT. Can AI outperform a junior resident? Comparison of deep neural network to first-year radiology residents for identification of pneumothorax. *Emerg Radiol*. 2020 Aug;27(4):367–75. <https://doi.org/10.1007/s10140-020-01767-4>
- [8] Thian YL, Ng D, Hallinan JTPD, Jagmohan P, Sia SY, Tan CH, et al. Deep Learning Systems for Pneumothorax Detection on Chest Radiographs: A Multicenter External Validation Study. *Radiology: Artificial Intelligence*. 2021 Jul 1;3(4):e200190. <https://doi.org/10.1148/ryai.2021200190>
- [9] Malhotra P, Gupta S, Koundal D, Zaguia A, Kaur M, Lee HN. Deep Learning-Based Computer-Aided Pneumothorax Detection Using Chest X-ray Images. *Sensors*. 2022 Mar 15;22(6):2278. <https://doi.org/10.3390/s22062278>
- [10] Tian Y, Wang J, Yang W, Wang J, Qian D. Deep multi-instance transfer learning for pneumothorax classification in chest X-ray images. *Medical Physics*. 2022 Jan;49(1):231–43. <https://doi.org/10.1002/mp.15328>
- [11] Hallinan JTPD, Feng M, Ng D, Sia SY, Tiong VTY, Jagmohan P, et al. Detection of Pneumothorax with Deep Learning Models: Learning From Radiologist Labels vs Natural Language Processing Model Generated Labels. *Academic Radiology*. 2022 Sep;29(9):1350–8. <https://doi.org/10.1016/j.acra.2021.09.013>
- [12] Seah J, Tang C, Buchlak QD, Milne MR, Holt X, Ahmad H, et al. Do comprehensive deep learning algorithms suffer from hidden stratification? A retrospective study on pneumothorax detection in chest radiography. *BMJ Open*. 2021 Dec;11(12):e053024. <https://doi.org/10.1136/bmjopen-2021-053024>
- [13] Hillis JM, Bizzo BC, Mercaldo S, Chin JK, Newbury-Chaet I, Digumarthy SR, et al. Evaluation of an Artificial Intelligence Model for Detection of Pneumothorax and Tension Pneumothorax in Chest Radiographs. *JAMA Netw Open*. 2022 Dec 15;5(12):e2247172. <https://doi.org/10.1001/jamanetworkopen.2022.47172>
- [14] Kitamura G, Deible C. Retraining an open-source pneumothorax detecting machine learning algorithm for improved performance to medical images. *Clinical Imaging*. 2020 May;61:15–9. <https://doi.org/10.1016/j.clinimag.2020.01.008>
- [15] Al-antari MA, Hua CH, Bang J, Lee S. “Fast deep learning computer-aided diagnosis of COVID-19 based on digital chest x-ray images.” *Appl Intell*. 2021 May;51(5):2890–907. <https://doi.org/10.1007/s10489-020-02076-6>
- [16] Li X, Thrall JH, Digumarthy SR, Kalra MK, Pandharipande PV, Zhang B, et al. Deep learning-enabled system for rapid pneumothorax screening on chest CT. *European Journal of Radiology*. 2019 Nov;120:108692. <https://doi.org/10.1016/j.ejrad.2019.108692>
- [17] Hoi K, Turchin B, Kelly AM. How accurate is the Light index for estimating pneumothorax size? *Australas Radiol*. 2007 Apr;51(2):196–8. <https://doi.org/10.1111/j.1440-1673.2007.01705.x>
- [18] Do S, Salvaggio K, Gupta S, Kalra M, Ali NU, Pien H. Automated Quantification of Pneumothorax in CT. *Computational and Mathematical Methods in Medicine*. 2012;2012:1–7. <https://doi.org/10.1155/2012/736320>
- [19] Feng S, Liu Q, Patel A, Bazai SU, Jin C, Kim JS, et al. Automated pneumothorax triaging in chest X-rays in the New Zealand population using deep-learning algorithms. *J Med Imag Rad Onc*. 2022 Dec;66(8):1035–43. <https://doi.org/10.1111/1754-9485.13393>
- [20] Abedalla A, Abdullah M, Al-Ayyoub M, Benkhelifa E. Chest X-ray pneumothorax segmentation using U-Net with EfficientNet and ResNet architectures. *PeerJ Computer Science*. 2021 Jun 29;7:e607. <https://doi.org/10.7717/peerj-cs.607>
- [21] Röhrich S, Schlegl T, Bardach C, Prosch H, Langs G. Deep learning detection and quantification of pneumothorax in heterogeneous routine chest computed tomography. *Eur Radiol Exp*. 2020 Dec;4(1):26. <https://doi.org/10.1186/s41747-020-00152-7>
- [22] Pavel Y. Segmentation Models [Internet]. GitHub; Available from: https://github.com/qubvel/segmentation_models.
- [23] Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Res Notes*. 2022 Dec;15(1):210. <https://doi.org/10.1186/s13104-022-06096-y>

How to Cite;

Gencer A, Toker YI (2024) Segmentation of Pneumothorax on Chest CTs Using Deep Learning Based on Unet-Resnet-50 Convolutional Neural Network Structure. Eur J Ther. 30(3):249-257. <https://doi.org/10.58600/eurjther2018>