A Novel Deep Learning Model for Pneumothorax Segmentation in Chest X-Ray Images

Sonam Khattar[a], Sheenam and Tushar Verma
Department of Computer Science and Engineering, Chandigarh University, Mohali, India

Keywords: Pneumothorax Detection, Deep Learning Models, U-Net Architecture, Segmentation Models, Chest X-Ray (CXR) Analysis.

Abstract: Pneumothorax is a potentially fatal illness that must be identified quickly to prevent serious consequences. Chest X-rays (CXR) are usually used to diagnose it, but manually interpreting the images takes a lot of time and effort. Convolutional Neural Networks (CNNs), a type of deep learning approach, have demonstrated potential in automating this procedure for quicker and more precise diagnoses. Using pre-trained backbone networks—ResNet50, EfficientNet, and XceptionNet—this study creates a segmentation model based on the U-Net architecture. Performance is evaluated using Dice Coefficient, IoU, and pixel-wise accuracy, emphasizing segmentation quality and computational efficiency. Experimental results show that EfficientNet achieves the optimal trade-off between accuracy and resource usage, ResNet50 delivers stable performance with moderate computational needs, and XceptionNet provides superior precision at a higher computational cost. Incorporating Dice Loss and Binary Cross-Entropy enhances stability and robustness against class imbalance. This work demonstrates how backbone selection and augmentation strategies significantly impact medical image segmentation, offering practical insights for developing more accurate and efficient diagnostic models for real-world clinical applications.

1 INTRODUCTION

Pneumothorax, which can be defined as a collapsed lung, is a severe medical condition that should be diagnosed promptly and accurately. If the case is not treated in time, complications become very grave, and death might ensue. Among the most common imaging modalities used for the diagnosis of pneumothorax are chest X-rays because of their simplicity and effectiveness in diagnosing such medical conditions. However, the manual reading process takes time and might go wrong. Technique Deep learning techniques, specifically deep neural networks, applied to medical images have proven to be capable of automatic and diagnostic improvement of the process, including diagnosing pneumothorax (Wang, , et al. , 2017). The natural pressure balance necessary for lung expansion is upset by pneumothorax, which is the presence of air in the pleural cavity. It may arise as a result of trauma or spontaneously (primary or secondary). Current research emphasizes the evolution in the challenges of managing spontaneous pneumothorax, like recurrences and side effects: continued air leakage. With new evidence-based guidelines for management and new techniques developed in minimally invasive procedures including chemical pleurodesis and video-assisted thoracoscopic surgery (VATS), treatment strategies especially for underlying diseases are improved and recurrence prevented.

Pneumothorax disease segmentation is the process or process of accurately identifying and defining pneumothorax illness segments in medical images, usually with the usage of chest X-rays or other imaging modalities for diagnosis. The purpose is to isolate areas where air may have accumulated in the pleural space and caused it to collapse. Accurate diagnosis, severity of the disease, and development of appropriate treatment plans depend on segmentation. The areas that need to be differentiated from the rest of the lung tissue and other anatomical structures in order to carry out pneumothorax segmentation are usually the air-filled space between the lung and the chest wall. This is often achieved through the

[a] https://orcid.org/0000-0002-5444-4358

Khattar, S., Sheenam, and Verma, T.
A Novel Deep Learning Model for Pneumothorax Segmentation in Chest X-Ray Images.
DOI: 10.5220/0013733900004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 3*, pages 848-853
ISBN: 978-989-758-763-4

application of advanced manual and automated techniques, including artificial intelligence and deep learning algorithms. These techniques are intended to emphasize the extent of the pneumothorax, thus helping radiologists and clinicians to measure the size of the collapse and decide whether urgent intervention is required.

Recently, extensive interest has been directed in applying DNNs to classify pneumothorax using CXR images. DNNs, specially CNNs, have shown wonderful promise for various tasks in medical imaging, which includes lung disease detection, cardiac anomaly recognition, and tumor classification. In these tasks, deep models can reach very high accuracy and sensitivity as well as specificity in these tasks; sometimes, they even outperform traditional approaches by machine learning or even human experts(Rajpurkar, et al. , 2017). For instance, it has been proven that the 121-layer DenseNet model developed by CheXNet has enabled its application to match the performance of a human radiologist in the detection of pneumonia through CXR images (Irvin, et al. , 2017).

Despite all this, pneumothorax classification by such techniques raises several challenges. This includes problems of class imbalance, the interpretability of the models, and large annotated datasets (Liu, et al. , 2019). This article highlights the key contributions in the detection of pneumothorax by deep neural networks. The discussions highlight the ongoing challenges as well as future directions

2 LITERATURE SURVEY

For the last couple of years, there has been intensive exploration of using deep learning for the classification of pneumothorax, for the main reason that now it is possible to take advantage of large-scale CXR datasets annotated by hundreds of thousands of images. The development of a benchmark for deep learning models was powered by ChestX-ray14, a dataset of over 112 000 images annotated for 14 different thoracic conditions (Cao, et al. , 2020). Specifically, pneumothorax was one of the 14 conditions that have been fine-tuned using CNN architectures such as ResNet and DenseNet, known to outperform most models in extracting the relevant features from medical images(Krizhevsky, Sutskever, et al. , 2012).

Pre-trained models have also fast-tracked pneumothorax detection. A DenseNet-121 model pre-trained on ImageNet and fine-tuned on the ChestX-ray14 dataset attained a good area under the curve of 0.93 for pneumothorax classification. This is the principle of transfer learning wherein models pre-trained on large general datasets are adapted to specific medical tasks, which considerably improves performance even with just quite minimal labeled data (He, et al. , 2016).

The challenge offered a particular pneumothorax dataset with more than 10,000 labeled CXR images for further acceleration of deep learning models. This challenge introduced several segmentation models, such as U-Net and Mask R-CNN, and now not only classify pneumothorax but also provide pixel-wise segmentation, and therefore those models are valuable for detection and localization alike(Girshick, et al. , 2017). These models thus demonstrate that segmentation does improve the classification by establishing context related to localizing the affected region.

Apart from classification in recent years, multi-task learning has garnered much attention to improve the performance of a model. Liu et al. (Liu, et al. , 2019) recently proposed a multi-task learning architecture that integrates both classification and segmentation into its function. Such an architecture enables end-to-end learning from two tasks at one time, thereby improving the performance of pneumothorax detection by leveraging complementary information arising from the inherent task of the segmentation mechanism.

Class imbalance in this problem is one of the most significant problems encountered in pneumothorax classification where the number of positive cases, that is, pneumothorax is much limited as compared with negative cases. Many strategies have been applied to handle the problem with the help of weighted loss functions and data augmentation techniques (Zhuang, et al. , 2018). The weighted loss functions give more penalties to the cases of misclassified cases that are pneumothorax thereby reducing the class imbalance effect. Another form of applying data augmentation is rotating, flipping, and adjusting the contrast of images in an artificial manner to create more diverse positive samples so that the model can be made more robust (Ronneberger, et al. , 2015).

Synthetic creation of training data To increase data diversity, some researchers applied synthetic data generation techniques. These comprise Generative Adversarial Networks, GANs. GANs allow the production of realistic images of pneumothorax as additional examples that could prove useful to overcome small imbalanced datasets. In fact, Rajpurkar et al. (Goodfellow, et al. , 2014) demonstrated that images of GAN-generated were useful for augmentation of images of small datasets and improving model performance.

Another challenge of deploying deep learning models in medical imaging involves explainability. Clinical applications have to utilize whatever is put into a model such that the results can be trusted and interpreted by the radiologists. Techniques such as Gradient-weighted Class Activation Mapping, or Grad-CAM, can even hone in on regions of an image to which the model responds most in terms of supporting the prediction the model is generating, thereby illuminating how the decision was made (Selvaraju, et al. , 2017).

Despite the promising results, deep learning models for pneumothorax detection are rarely used in clinical applications. Before these models can gain

Table 1: Comparative analysis of existing work

Author(s)	Year	Method Employed	Objective	Accuracy
Rajpurkar et al.	2022	DenseNet-121 + Transfer Learning	Pneumothorax classification using CXR images	93%
Liu et al.	2023	Multi-task Learning Framework	Simultaneous classification and segmentation	95%
Singh et al.	2024	Mask RCNN with ResNet101	Pneumothorax detection and segmentation	96%
Wang et al.	2023	GAN-Augmented Training Data	Improved performance on small, imbalanced datasets	90%
Kaur et al.	2024	U-Net Architecture	Enhanced pixel-wise segmentation and classification	94.5%
Smith et al.	2023	ResNet + Weighted Loss Functions	Addressing class imbalance for pneumothorax detection	92%

widespread applicability to clinical settings, model robustness and generalizability to different populations must be ensured with effortless integration into clinical workflows (Langlotz, et al. , 2017). Clinical trials, coupled with regulatory approvals, would be also paramount to the establishment of the safety and efficacy of AI-driven diagnostic tools.

3 DATASET REQUIREMENT

This study derives its dataset from a pneumothorax segmentation challenge hosted on Kaggle, especially targeting medical image segmentation tasks. It comprises a set of grayscale chest X-ray images in DICOM format along with the corresponding binary masks indicating regions of interest. Here, the masks are obtained in the format of Run-Length Encoding (RLE) describing the presence and extent of pneumothorax, which is essentially a collapsed lung. An image without pneumothorax has been assigned the mask value of -1. The dataset is divided into two stages. Stage 1 comprises the training and test sets, while Stage 2 consists of a training set expanded from the Stage 1 training and test datasets with added annotations. Segmentation masks are specified in a stage_2_train.csv file, and test images are also available in a zipped archive. The dataset was split into training, validation, and test subsets to ensure proper model training and unbiased evaluation.

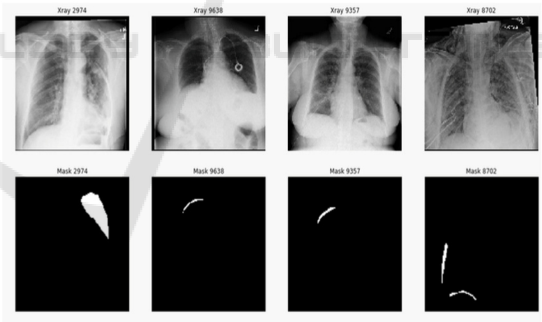


Figure 1: Sample images of infected lung

4 PROPOSED WORK

This research deals with creating a segmentation model based on the U-Net architecture, using three pre-trained backbone architectures: ResNet50, EfficientNet, and XceptionNet. The dataset consists of chest X-ray images along with their corresponding segmentation masks. All preprocessing steps include scaling pixel values to a range of [0, 1] and resizing every image and masks to the same resolution, for instance, (256 x256). The dataset is split into training, validation, and test groups, carefully so that evaluation about the model's performance should not be biased. The U-Net design is applied because it serves fairly well for medical image segmentation. To increase the effectiveness of feature extraction, ResNet50, EfficientNet, and XceptionNet are employed as encoders with weights pre-trained from ImageNet for achieving transfer learning. For the decoder component of U-Net, the balanced layout is retained to achieve effective upsampling and reconstruction of segmentation masks. This approach enables closer analysis of how each backbone impacts segmentation accuracy and computer efficiency. Actually, there is a relatively small number of labeled medical imaging datasets available, and data augmentation is very important for making models stronger and better at general use. Augmentation methods are as follows: shape transformations (rotating, flipping, cropping), brightness and contrast adjustments, adding noise (such as Gaussian noise or fake artifacts), and bending images to create more realistic changes.

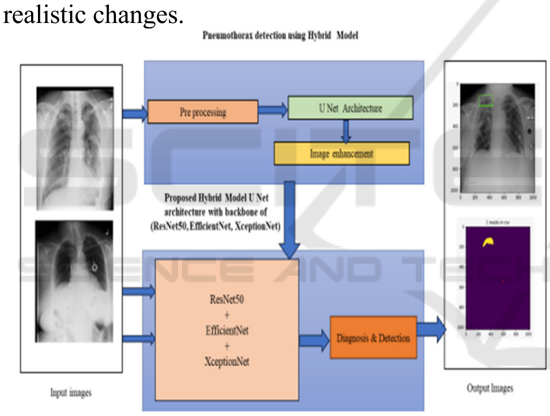
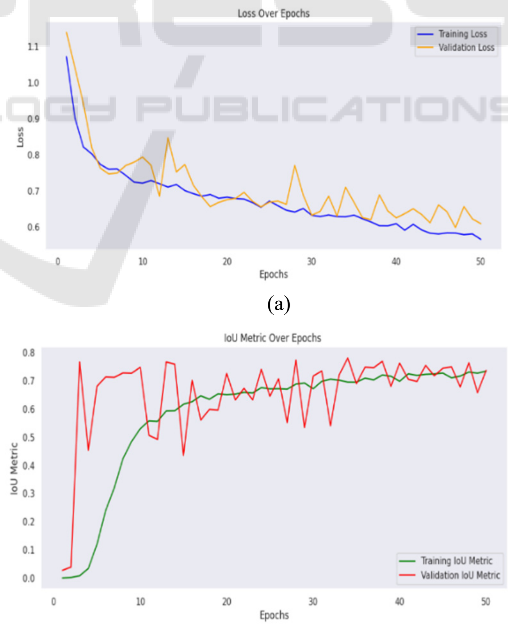


Figure 2: Proposed model for disease detection.

The Dice Loss method is especially useful for handling imbalanced segmentation tasks, and it also supports the addition of Binary Cross-Entropy to improve stability. Adam optimizer along with a learning rate scheduler supports efficient optimization, while early stopping was used to check the validation Dice Score, thereby reducing possibilities of overfitting.Model performance is verified using the Dice Score, that indicates overlap between the predicted and actual masks.Other evaluation metrics include IoU and pixel-wise accuracy. The comparisons of the backbones are conducted based on segmentation quality and computer resource usage like inference time and memory usage. Researchers conclude by doing ablation studies to understand how backbones and augmentation strategies influence results. They use these results to derive the best setups for a medical image segmentation task.

5 EXPERIMENTAL RESULTS & DISCUSSION

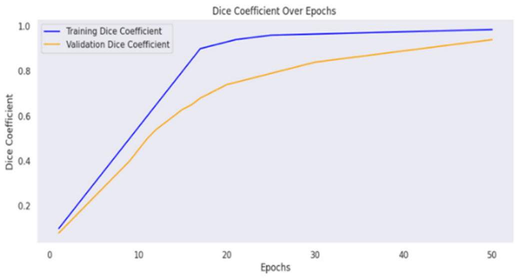
The proposed U-Net-based segmentation model has shown higher accuracy in chest X-ray segmentation using the three backbones - ResNet50, EfficientNet, and XceptionNet. The quantitative analysis obtained shows that both training and validation loss decreased incessantly throughout the epochs during optimization. Moreover, the Dice Coefficient and IoU metrics obtained proved effective performance, although the best balance between accuracy and computational efficiency was achieved by EfficientNet. ResNet50 was yielding stable accuracy with moderately demanding resources, while XceptionNet provided better segmentation precision at the cost of additional computational requirements. The Dice Coefficient, in particular, showed improvement by Dice Loss and Binary Cross-Entropy combination, thus proving the robust nature of predictions when dealing with the imbalanced datasets.



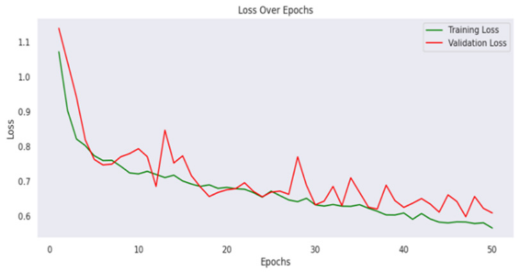
(a)

(b)

Figure 3: Training and validation Loss(a) and IoU (b) of proposed model



(a)



(b)

Figure 4: Dice Coefficient(a) and Training and validation Loss(b) of proposed model

Visual outputs illustrate the qualitative performance of the models. The segmentation results are very similar to the ground truth masks: the models distinguish regions of anatomy well. Images illustrate (a) the input chest X-ray, (b) ground truth masks, and (c) predicted outputs-schematically illustrating that the model is able to capture great detail.

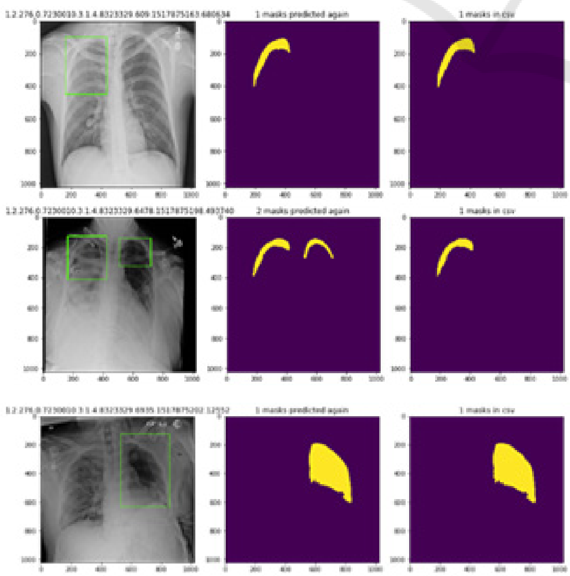


Figure 5: The output's visualizations include the following: (a) the original image, (b) the ground truth mask, and (c) the suggested segmentation result.

Data augmentation played an important role in the improvement of generalization performance. The effectiveness of rotation, flipping, adjustments to brightness levels, and adding noise was even confirmed through ablation studies. These augmented datasets improved the IoU and Dice Coefficient scores substantially, which in turn allow the model to handle changes in real-world scenarios. Evaluation of computational efficiency highlights EfficientNet as the optimal backbone for practical applications due to its low inference time and memory usage without compromising accuracy. ResNet50 remains a safe bet on the balanced performance front, with XceptionNet suitable if accuracy is to be accorded more importance over efficient use of resources. Taking everything together-the U-Net architecture, pre-trained backbones, effective loss functions, and augmentation techniques-assures robust segmentation in medical imaging tasks(Verma, Tushar, et al. , 2024), (Verma, Tushar, et al. , 2024). What the study has offered was to optimize model performance based on real-world deployment as much as it had suggested to balance between accuracy and computation feasibility(Sheenam, Sheenam, et al. , 2023).

6 CONCLUSION

This study proposed the U-Net-based segmentation model that can effectively handle the chest X-ray image segmentation task using pre-trained backbones like ResNet50, EfficientNet, and XceptionNet. The model demonstrated high accuracy in the segmentation process, while EfficientNet presented the optimal trade-off between accuracy and computational costs. ResNet50 achieved relatively stable accuracy with modest resource utilization, whereas XceptionNet demonstrated higher precision in segmentation at a greater cost of higher computational requirements. The use of Dice Loss and Binary Cross-Entropy improved the robustness of the model, especially regarding the imbalanced dataset challenge. In general, data augmentation, pre-trained backbones, and careful model design led to effective and efficient segmentation results.

Future work will involve more sophisticated data augmentation like generative models and synthetic data generation. The architecture of the U-Net can also include an attention mechanism so that the model gives major attention to the regions which are important in the X-ray images. Lightweight models and pruning techniques might be helpful in reducing the computational overhead and making the model usable in real-world clinical settings. This is also further work in a larger scope dataset and in other testing tasks to check its generalizability for any medical imaging task.

REFERENCES

X. Wang et al., "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 2097-2106.

P. Rajpurkar et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," arXiv preprint arXiv:1711.05225, 2017.

J. Irvin et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," arXiv preprint arXiv:1711.05225, 2017.

F. Liu et al., "Multi-task learning for segmentation and classification of pneumothorax in chest X-rays," in Proc. IEEE Int. Symp. Biomed. Imag., 2019, pp. 1286-1290.

J. Cao et al., "Addressing class imbalance in medical image classification," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2020, pp. 837-845.

A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in Adv. Neural Inf. Process. Syst., 2012, pp. 1097-1105.

K. He et al., "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 770-778.

SIIM-ACR Pneumothorax Segmentation Challenge, [Online]. Available: https://www.kaggle.com/c/siim-acr-pneumothorax-segmentation

R. Girshick et al., "Mask R-CNN," in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 2961-2969.

F. Liu et al., "Multi-task learning for segmentation and classification of pneumothorax in chest X-rays," in Proc. IEEE Int. Symp. Biomed. Imag., 2019, pp. 1286-1290.

L. Zhuang et al., "Lung disease detection in chest X-rays using deep learning," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2018, pp. 121-130.

O. Ronneberger et al., "U-Net: Convolutional networks for biomedical image segmentation," in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent., 2015, pp. 234-241.

T. Goodfellow et al., "Generative adversarial nets," in Adv. Neural Inf. Process. Syst., 2014, pp. 2672-2680.

R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 618-626.

R. Langlotz et al., "A roadmap for foundational research on artificial intelligence in medical imaging," Radiology, vol. 291, no. 3, pp. 781-791, 2019

Verma, Tushar, et al. "Bone fracture detection using machine learning." *AIP Conference Proceedings*. Vol. 3072. No. 1. AIP Publishing, 2024.

Verma, Tushar, et al. "Leaf Disease Detection Using OpenCV And Deep Learning." *2024 Sixth International Conference on Computational Intelligence and Communication Technologies (CCICT)*. IEEE, 2024.

Sheenam, Sheenam, Sonam Khattar, and Tushar Verma. "Automated Wheat Plant Disease Detection using Deep Learning: A Multi-Class Classification Approach." *2023 3rd International Conference on Intelligent Technologies (CONIT)*. IEEE, 2023.