Identify Pneumothorax disease in chest X-rays

Seyed Ehsan Seyed Bolouri

University of Alberta, Edmonton, Canada

Abstract. A significant challenge in pneumothorax treatment is the early diagnosis of pneumothoraces. An accurate AI algorithm to detect pneumothorax would be useful in a lot of clinical scenarios, although a perfect segmentation methodology has not yet been possible using conventional scientific methods. In this report, I hope to show how subject classification and collapsed lung segmentation tasks by using deep learning techniques might be achievable. The first step would be of distinguishing the subjects which experiencing collapsed lung phenomenon based on their chest radiographic images. The previously mentioned task has been completed with the accuracy of 79% by taking the concept of transfer learning and linear support vector machine into account. Moreover, segmentation of the pneumothorax mask with a Dice score of 45% has been achieved utilizing convolutional neural networks trained on the dataset consisting of 1000 X-ray images. The network architecture can be outlined as a 2D U-Net with a ResNet-34 encoder.

1 Introduction

The Pneumothorax phenomenon occurs when air traps in the space between the lung and chest wall. This air compresses the lung and correspondingly decreases the typical volume of the air that can be inhaled. This disease can be caused by any chest injury and also ruptured air blisters. Air blisters are blebs which can develop on the surface of the lungs and they are likely to occur in people between 20 and 40 years old, especially if they are tall and underweight. Moreover, the likelihood of having pneumothorax is far higher in men compared to women [1]. It is notable that we can divide the previously mentioned disorder into two different stages including primary spontaneous pneumothorax which occurs in persons without clinically apparent lung disease or history of respiratory illness and secondary spontaneous pneumothorax which is a complication of preexisting lung disease [2].

2 Background and literature survey

Mortality of 1.26 per million a year for men and 0.62 per million a year for women has been reported for pneumothorax disorder. Additionally, it can cause some serious damage to respiratory system that would lead to shortness of breath, sharp pain while inhaling and even loss of consciousness. Consequently, the main objective of this research paper is designing a pipeline that can be used as an axillary tool for medical specialists and therapist to diagnose pneumothorax more accurately in the early stages and initiate therapy with appropriate details.

Deep learning and specifically convolutional neural networks are the frameworks that have been proven their efficiency and superior performance over machine learning frameworks in similar medical image analysis projects. Subsequently, Using the previously mentioned tools would be a baseline to analyze the dataset.

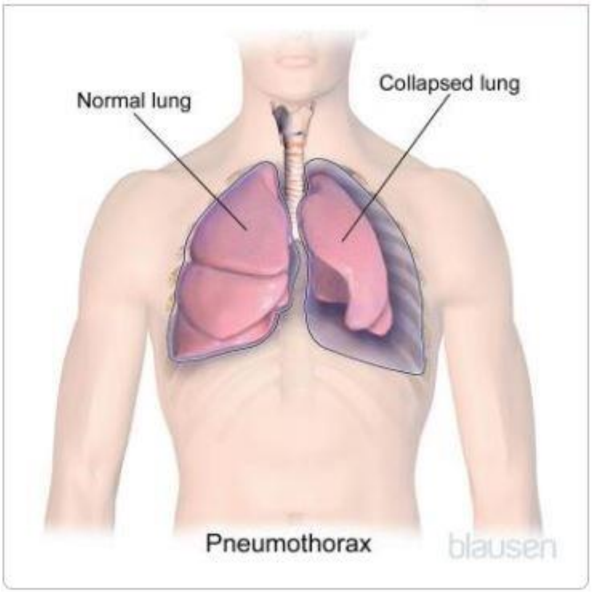


Fig. 1. Schematic comparison between the healthy lung (left) and the collapsed lung (right) [3]

3 Dataset description and data preprocessing

Data set is consisting of over 11500 X-Ray images with graphics display resolutions of 1024*1024 including the healthy subjects in addition to subjects with pneumothorax. 20 percent of images contain the masks that were labeled by a medical specialist. Run-Length encoding format has been used to specify the mask for each image.

Due to the lack of computation resources, I selected approximately one-eleventh of the original dataset for my experiments. It is worth mentioning that the number of positive and negative cases in the selected dataset is very close. Dataset has not been normalized and each image has a different contrast ratio. The contrast limited adaptive histogram equalization (CLAHE) has been used to obtain a dataset with normalized histograms. This method is different from ordinary histogram equalization due to making use of the transformation matrix to transform each pixel derived from a neighborhood region. This method is a variant of adaptive histogram equalization in which the contrast amplification is limited, so the noise amplification phenomenon is reduced in most of the cases [4].

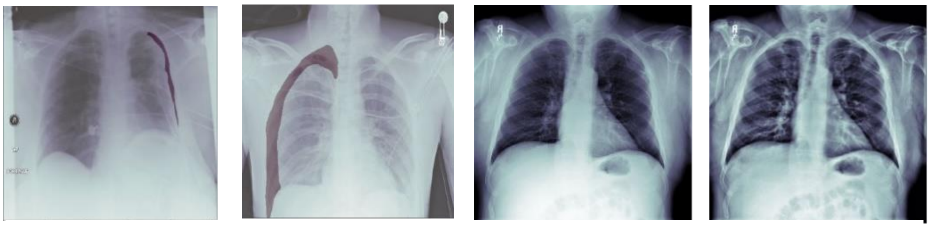


Fig. 2. From left, subject with slight pneumothorax, subject with severe pneumothorax, healthy subject before normalization, healthy subject after CLAHE normalization

4 Proposed methods for the classification task

The first step while analyzing the images would be classifying the dataset into two groups, X-Rays with pneumothoraces and healthy X-rays. Notably, due to the nature of collapsed lung disorder, it is even hard for medical specialists to decide whether the subject is sick or not.

The concept of transfer learning has been considered for the corresponding classification task. The main structure of the classification pipeline for this task consists of a pre-trained network named DenseNet121 and a classifier on top of the network. Fine-tuning the structure on the provided dataset would result in a classifier that can predict whether the subject has the pneumothorax or not.

4.1 SoftMax classifier

A neural network with a SoftMax function as the last layer is a very common structure for classifying the datasets due to the acceptable performance in prediction. The way how it works depends on the probabilities and distribution of the output from the layer before the SoftMax.

4.2 Support Vector Machine classifier (SVM)

Support Vector Machine is a shallow learning algorithm that does the prediction by finding the optimal hyperplane which divides the dataset into two contrasting classes. Hyperplane can be found by taking the concept of maximum margin into account. The common loss functions for SVMs are L1 and L2 hinge losses.
The idea of using SVMs on top of a deep learning framework has been borrowed from [5] which was published in 2013. They proved that DLSVM has shown superior performance over deep learning with the SoftMax function as the last layer. Their results are based on their experiment on MNIST and CIFAR-10 datasets.

L2 Hinge loss has been chosen over L1 for this task due to the differentiable nature of the function. Consequently, PyTorch would backpropagate through it without any further modifications. SVM constrained optimization and the corresponding derivatives can be written as below:

$$\min \ \frac{1}{2} w^T w + C \sum_{i=1}^{N} \max\,(1 - w^T x_n t_n, 0)^2 \tag{1}$$

$$\frac{\partial l(w)}{\partial h_n} = \ -2Ct_n w(\max(1 - w^T x_n t_n, 0)) \tag{2}$$

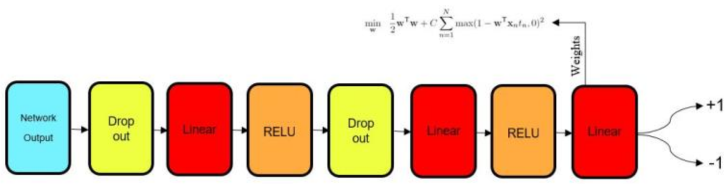


Fig. 3. SVM classifier with the L2 Hinge loss

DLSVM focuses on different types of features compared to SoftMax and it would result in better overall accuracy. After adjusting the hyperparameters for each method separately, it can be shown that DLSVM improved the accuracy of SoftMax from 75% to 79% for this specific classification task.

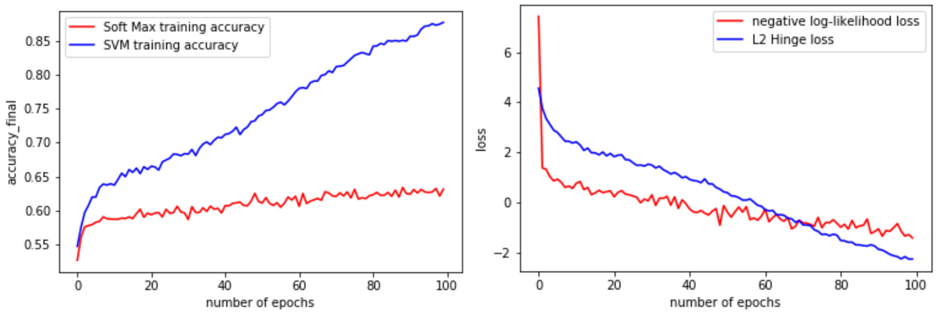


Fig. 4. Training Accuracy(left), Training loss(right)

5 Proposed methods for pneumothorax segmentation

5.1 U-Net

The U-NET architecture was developed by Olaf Ronneberger et al. [6] for Biomedical Image Segmentation. Due to the outstanding performance in medical image analysis, it can be considered as a baseline for every segmentation task. I used the Vanilla U-Net followed by DICE loss in my first attempt to segment the pneumothorax from the original image. This attempt was completely unsuccessful! The output of U-Net for each image was an empty mask and the network did not learn during the training process.

The possible reasons that U-Net could not predict the outcome correctly are:

1. I used positive and negative cases to train the network. Images with empty masks distract the network from focusing on the positive cases and force it to predict empty masks for all the images
2. Images with the resolution of 128*128 have been used to train the network and down sampling process omits many deterministic features in the prediction of masks
3. Although Dice loss is a good metric for most of the segmentation problems, it did not help the network capture the masks properly for this case due to the huge difference between the background and foreground

5.2 U-Net with a ResNet backbone

ResNet encoder would reinforce the previously mentioned network by adding extra layers and depth. Although the structure is deeper and more complicated compared to U-Net, it would not face the vanishing gradient problem because of the skip connections that embedded in the structure of ResNet.

The encoder section works as a feature extractor in this architecture and it extracts the features from the training set that contains only images with masks. It must be noted that I used the image with the original size (1024*1024) to train the network.

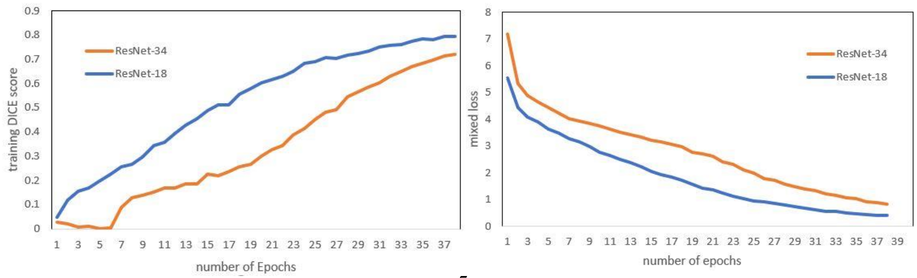
The loss function for this task is a combination of DICE and Focal loss. Focal loss has been added to improve the performance of the network by focusing more on the pixels that form the masks. It penalizes the loss function for pixels that are 1 in the ground truth image and network classifies them as zero with higher weights. The formula that has been used for the Focal loss function [7] is shown below:

$$FL(p_t) = -(1 - p_t)^{\gamma} \log p_t \tag{3}$$

$(1 - p_t)^{\gamma}$ is the modulating factor with a tunable focusing factor γ. By utilizing the Focal loss, class imbalance problem has been treated in this object detection task with the minimum error.

6 Experiments and results

ResNet-18 and ResNet-34 have been considered for the decoder section of the network. Experiments have shown that ResNet-34 has a better performance on the test set. ResNet-18 seems too shallow and ResNet-50 is too deep that may cause the overfitting problem. According to Fig. 5 the training loss is decreasing, and the training DICE score is increasing. Consequently, the network is working properly, and the weights are getting updated in the optimization process.



5.

Fig. 5. Comparison between ResNet-18 and ResNet-34 performance. Training DICE score(left) ,Training mixed loss(right)

The DICE score for the test set is 45%. It means that this model is capable of predicting the location and surface area corresponding to pneumothorax for positive cases up to a good extent. Some successful and unsuccessful samples of segmentation can be found below:

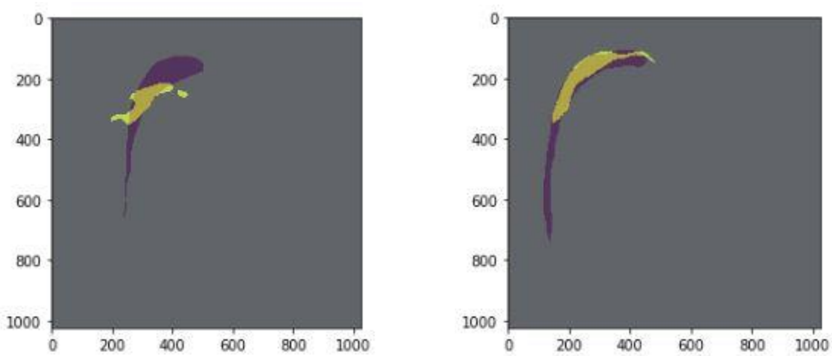


Fig. 6. Purple mask is the ground truth and yellow label is the output of network

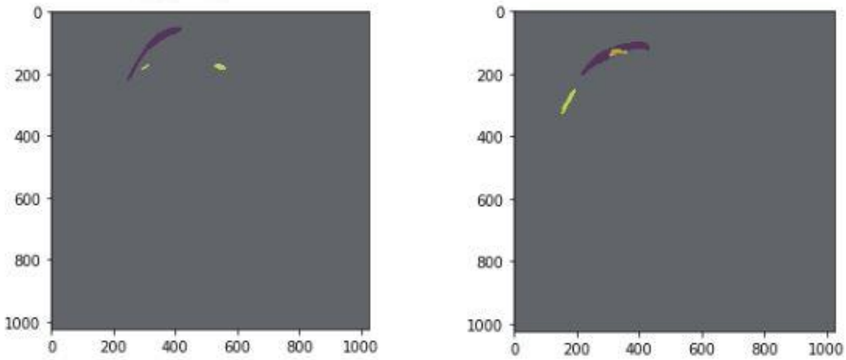


Fig. 7. Purple mask is the ground truth and yellow label is the output of network

7 Conclusions

Using the SVM classifier on top of a deep learning framework instead of the SoftMax function can be a more efficient approach in some cases. SoftMax uses the probability and distribution of features to determine the outcome of the network. However, Support Vector Machine focuses on relatively real features due to the nature of algorithm.

Utilizing deeper networks to extract more deterministic features in segmentation tasks such as pneumothorax segmentation that is even difficult for medical specialists to identify the intended area would increase the accuracy of segmentation. For imbalanced images, weights should be assigned in a way that keeps the focus of the network on the pixels that are less abundant.

8 Future improvements

Training the model has been done on the GPU provided by Google Collaboratory and because of the limited power of this cloud service, I selected one-eleventh of the original data set to complete the different tasks. Having enough computation resources to train the entire dataset would result in far better accuracy for each task.

Adding nonlinearity to the nature of SVM by using Gaussian or Polynomial kernels would increase the accuracy of the classifier.

Using the metadata provided for each X-Ray image such as age, sex, modality and view position and comparing the importance of these features. Finally, Dataset can be divided into different batches based on the common features and analyzing those batches by deep learning would increase the accuracy in the segmentation task.

Fig. 8. Chest X-Rays with metadata provided for each patient

References

1. Mark S. Allen, M.D .: Pneumothorax. Retrieved from www.mayoclinic.org (2019)
2. SA Sahn, JE Heffner.: Spontaneous pneumothorax. New England Journal of Medicine (2000)
3. Richard W. Light, Pneumothorax. Retrieved from www.merckmanuals.com
4. Stephan Saalfeld.: Enhance Local Contrast (CLAHE), ImageJ (2010)
5. Yichuan Tang.: Deep Learning using Linear Support Vector Machines. arXiv:1306.0239 (2013)
6. Olaf Ronneberger, Philipp Fischer, and Thomas Brox.: U-Net Convolutional Networks for Biomedical Image Segmentation. MICCAI (2015)
7. Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, Piotr Dollar.: Focal Loss for Dense Object Detection. arXiv:1708.02002 (2018)