Segmentic Segmentation of Chest X-Ray Images using InvertedNet

Mathilde Aubret, Osheen Sharma, Blanca Cabrera Gil

May 19, 2019

Abstract

In the past few years, Deep Learning has contributed in the clinical applications making it easier for early diagnosis and treatment of the diseases. The success of deep convolutional neural network relies on big data and large sets of training examples. This project focuses on the semantic segmentation of chest X-ray images using InvertedNet architecture which relies on the use of data augmentation technique [4]. The dataset used for the project comprises of 246 chest X-ray images acquired from JSRT along with ground truth segmentation masks available on SCR. The network is trained with AdaDelta optimizer, and the loss is computed as the sum of the categorical cross entropy and the Dice coefficient. The network acquired an accuracy of 96% on training set, 93% on validation set and 95% on testing set.
Keywords: diagnosis, segmentation, InvertedNet, data augmentation, accuracy.

The link for the GitHub Repository along with the code can be found with the link: `https://github.com/bcabgil/DD2424-Semantic-Segmentation-For-XRay-Images`

1 Introduction

Medical imaging field is going through large improvements thanks to computer aided tools that allow a better diagnosis to be performed. Among medical images, chest radiography are the most common images and represent an important tool in diagnosis and treatment. Segmentation of anatomical structures in chest radiography plays a key role for computer aided diagnosis. It allows to detect specific anatomical features to achieve accurate diagnosis and measure the performance of ongoing treatments.

The breakthrough of deep learning using convolutional neural networks was given by Krizhevsky et al. [1]. Their winning network for the ImageNet contest consisted of convolutional, max pooling and fully connected layers. In their architecture dropout was also used as regularizer. Following, Ronneberger et al. [3] presented the U-Net, a neural network architecture based on a contracting path to caputre the image context, followed by a symmetric expanding path to enable precise localization. This network relies on the power of data augmentation to use few images to train efficiently a network. Novikov, et al. [2] proposed a deep model called InvertedNet which outperformed the U-Net results on the segmentation of chest radiography by the usage of delayed sub-sampling, exponential linear units, highly restrictive regularization and a large number of high resolution low level abstract features.

The main aim of this project is to perform a multi class segmentation of chest images focusing on three areas of interest: the lungs, the clavicles and the heart. InvertedNet architecture performance is evaluated by means of the test accuracy score. The loss function is defined as the sum of the Categorical Crossentropy and the Dice coefficient.

This report is structured in three main sections: Methods, Results & Discussion and Conclusion. In the first section the data set structure, pre-processing of data and model implementation are described. The obtained results and performance are discussed in Results & Discussions section. Finally, the overall lessons learned are stated in Conclusions.

2 Methods

2.1 Data Set

The dataset used in the project comprises of 246 chest X-ray images and their corresponding masks taken from SCR Database and JSRT database respectively. For each chest radiography images, a total 5 masks are available: 2 for lungs, 2 for clavicle and 1 for heart.

2.2 Data pre-processing

The images are initially of size (2048,2048). According to the results obtained in the research [2] the images were resampled to have a size of (128,128). The five masks are merged and processed to label the two clavicles with label 1, the heart with label 2 and the right and left lung with label 3. The overlapping regions are dealt with in a way such as the clavicle stands out from the heart and the lungs, and the heart stands out from the lung. The figure below depicts a chest radiography image and a mask example after pre-processing of the data.

The data is then randomly split into the training, validation and testing set as follow:

- Training data: 198 images and masks
- Validation data : 23 images and masks
- Test data : 25 images and masks

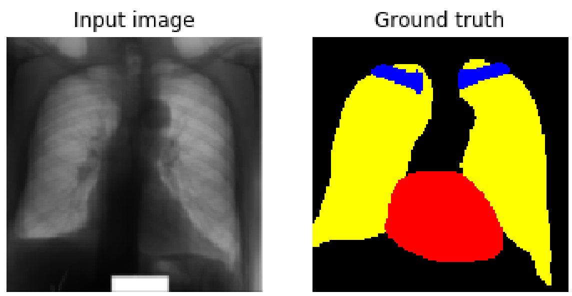

Figure 1: Input image and corresponding labelled ground truth mask.

2.3 Network Architecture

InvertedNet architecture was chosen in this project. It is a modified version of the original U-Net architecture. This neural network is based on a contracting path to capture the image context, followed by a symmetric expanding path to enable precise localization [3]. The aforementioned relies on the power of data augmentation to use few images to train efficiently a network.

The changes introduced to the U-Net model provided a better performance when predicting semantic segmentation labels for chest X-ray images. InvertedNet network is characterized by being a fully-convolutional network with fewer parameters compared to U-Net.

One of the biggest challenges that fully-convolutional networks face is overfitting. This architecture deals with this problem by the addition of dropout layers after each convolutional layer to reduce the expressivity of the function. Delayed subsampling of the pooling layers is also introduced. A stride of 1 is used for the first convolutional and pooling layer. On the subsequent pooling layers a stride of 1 is maintained, but for the following convolutional layers a stride of 2 is set. Lastly, the solution space of the network is reduced by reordering the number of feature maps in the convolutional layers. The network starts with a large number of feature maps (128) which is divided by a factor of two after every pooling layer. This reduction takes place in the contraction part of the network. On the other hand, the number of feature maps is increased by a factor of two after every upsampling layer of the expansion part [2]. The final network architecture is displayed in figure 2.

2.4 Training strategies

All the experiments are run with Keras in Python, and the following training strategies are used with inspiration from [2].

Firstly, *ELU* activation function is chosen for all the convolution layers, and

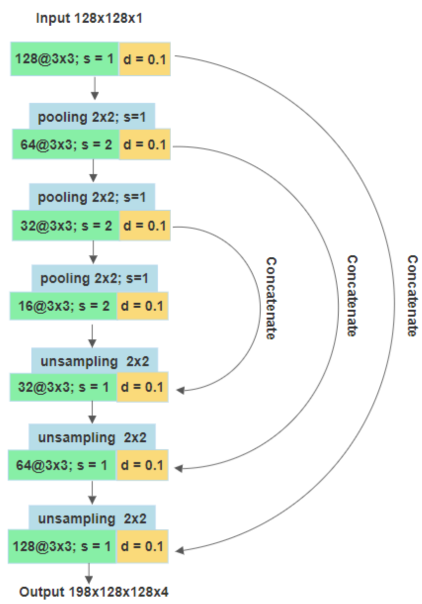

Figure 2: Overview of the network architecture

Softmax activation function is chosen for the output layer. This latter is used in the loss computation, which is defined as the the sum of the Categorical Crossentropy and the Dice coefficients. The network is trained using *AdaDelta* optimization, with the default parameters, namely initial learning rate of 1, a decay factor of 0.95, and an initial learning rate decay of 0.

Some improvements are implemented in order to achieve better results in the configuration of our problem. First, data augmentation is implemented, using Keras data generator with a zoom range of 0.01, a rotation range of 10 degrees and horizontal and vertical flip activated. Additionally, early stopping is added to the network in order to have an optimal number of epochs and avoid under-fitting or over-fitting issues. Lastly, in order to perform efficient experiment, Google Cloud is configured to handle our experiments.

3 Results and Discussion

The network is trained for 1000 epoch, but the training stops after 491 epoch due to early stopping with a patience of 100 epoch. Figure 3 depicts the performance of the model on training and validation set.

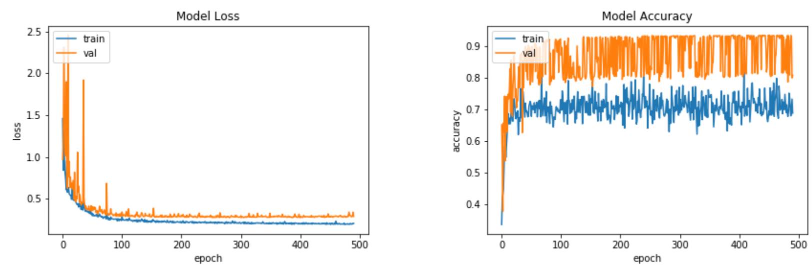

Figure 3: Performance of the model on traning and validation set

Figure 4 shows the segmentation results after the model have been trained. Despite the unbalanced data representation, and especially the under representation of the clavicles, the network managed to segment the biological structures with precision and coherence.

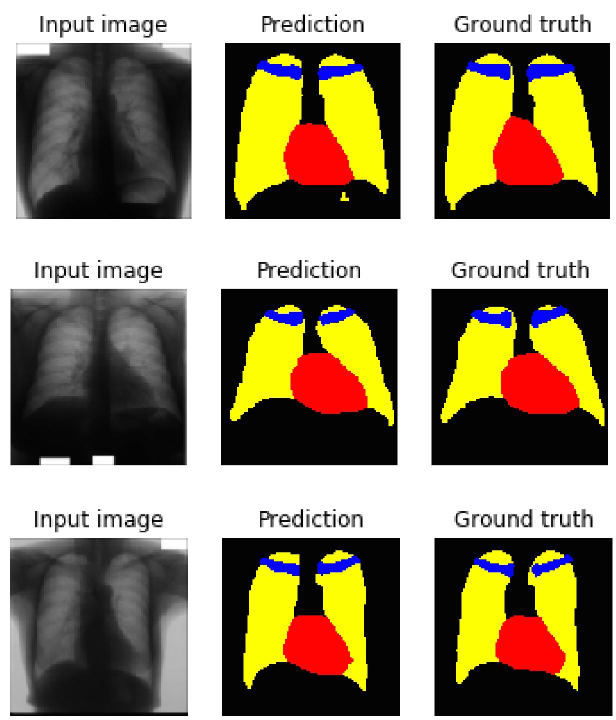

Figure 4: Segmentation results corresponding to the input test image and ground truth

The following table reveals the final accuracy obtained after training over

491 epoch. The accuracy obtained with the network built is similar to the one obtained in the reference paper [2].

Table 1: Evaluation results

Data Set	Accuracy (%)	Loss
Training	96.5	0.20
Validation	93	0.20
Test	95	0.25

In terms of timing performance, the network depicts bad performance when the algorithm is run on a local PC. It takes around 8 hours to train for 100 epoch. On the other hand, when the network is trained on Google Cloud, it exposes a significantly better timing performance. Indeed, the network takes 20 minutes of computation on the Google Cloud for 100 epoch, so 1.5 hours approximately for 500 epochs.

4 Conclusions

This project proposes an approach to multi class segmentation of anatomical structures in X-ray images. An InvertedNet convolutional neural network was implemented to deal with the problem. This latter was improved by several means including random splitting of the data, data augmentation techniques and early stopping. Data augmentation techniques are necessary due to the size of the data set. Overall results show that even under represented organs can be segmented with the network implemented, even though class weights were not introduced. The timing efficiency of the network is improved with configuration of the Google Cloud account, and the overall training of the network takes approximately one hour and a half. An accuracy over 90 % is obtained for the three data set, namely training, validation and test.

On a broader scale, this project has proved the power of Deep Learning when applied to medical images. It was in this case applied to the segmentation problem, but it can show promising results on broader application of medical image processing for diagnosis application.

References

[1] Geoffrey E. Hinton Alex Krizhevsky, Ilya Sutskever. Imagenet classification with deep convolutional neural networks. 2015.

[2] David Major Jirı Hladuvka Maria Wimmer Alexey A. Novikov, Dimitrios Lenis and Katja Buhler. Fully convolutional architectures for multi-class segmentation in chest radiographs. 2018.

[3] Philipp Fischer Olaf Ronneberger and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. 2015.

[4] Philippand Brox Thomas Ronneberger, Olafand Fischer. U-net: Convolutional networks for biomedical image segmentation. 2015.

Appendix : Personal conclusions

Mathilde Aubret During the Deep Learning course, I acquired knowledge about deep learning in general, and I applied it during the project work.

I learned particularly during the project how to build Convolutional Neural Network from scratch with Deep Learning libraries, such as Keras. During the project work, I learned how to structure my ideas and implement a CNN to fill my stated goals, here inspired by a precise network architecture given in the reference paper, the InvertedNet. More precisely, I learned how get familiar with Keras Deep Learning libraries, with some practice and insights on common mistakes or trick for a good implementation. I also learned how to correctly justify the parameter of the network, as the library provides with a broad range of already written functions that you have to correctly implement in the network. Lastly, I learned some performance improvement that can be implemented in the network, such as data augmentation that we used due to the size of our dataset.

The project I chose was linked with my research interests, medical images segmentation, and I really realized the power of Deep Learning techniques by seeing the promising results achieved in this project for semantic segmentation of biological structures, in a reduced time. It was the first time for me building a convolutional neural network, and I think this method is really interesting when working with medical clinically relevant images, and with my previous knowledge in medical images segmentation, I was able to see the power of this method among others.

Osheen Sharma After the completion of the project I understand the importance of deep learning and its ability to segment medical images and making it easier for diagnosis and treatment. I learnt the implementation of Convolutional Neural Network using keras (the deep learning library) for segmentation of chest x-ray images. The segmentation techniques I learnt previously using matlab were graph cut, normalized cut etc which were very much complex and difficult for me to understand. This project makes me familiar with an interesting technique which helped me improve my skills and knowledge in the field of deep learning. The dataset we used in our project was not so huge, so I also learnt the technique of data augmentation so that the training of the model can be improved. I learnt how to implement InvertedNet architecture for segmentation of medical images. I also learnt about different activation layers and the differences we can get in our prediction on implementing those activation layers.

Blanca Cabrera Gil The realization of this project gave me the opportunity to gain three main skills. In the first place, get more familiar with the usage of Keras to test and work with state of the art network architectures. Also, to learn best practices of the library, its APIs and how to solve common mistakes when working with CNNs. Following, we were able to implement a deep learning solution for a very studied problem, semantic segmentation of biological structures. This method resulted to be more accurate than state of the art rule-based approaches for segmenting medical images. Giving us a tool for further research and work. Finally, we learned how to work with U-Net architecture, understanding the utility of autoencoder network architectures for semantic segmentation, find and evaluate latest research works and implement an improved version of the network which focused in medical image purposes.