DeepMedic and U-Net neuronal network architectures for lung segmentation in CT scans

Lionel Rajaona
University of Western Ontario
Email: lrajaona@uwo.ca

Rajat Balhotra
University of Western Ontario
Email: rbalhotr@uwo.ca

Daniel Allen
University of Western Ontario
Email: dallen44@uwo.ca

Steffen Bleher
University of Western Ontario
Email: sbleher@uwo.ca

***Abstract*—This work examines the DeepMedic and the U-Net architectures of neuronal nets for segmentation of lung structures in computed tomography scans. The application of the algorithms will be evaluated on the LUNA16 dataset.**

I. Introduction

Since hardware specifications and computational power increased dramatically in the last years machine learning approaches can be applied in various disciplines today. On top of that the progress in research on convolutional neuronal networks (CNN) made it a very powerful tool for image processing where information is gained from image data.
One challenging application is medical image computing (MIC). The main goal of MIC is to extract clinically relevant information or knowledge from medical images. Furthermore Segmentation is the process of partitioning an image into different meaningful segments (e. g. organs, bones, ...).
In this project the goal is to segment the lung of a human body from computed tomography (CT) scans.
The used dataset is from the LUNA16 challenge [1] and each scan contains a number of slices which are 512 x 512 pixel greyscale images. The algorithm creates a 512 x 512 pixel label map for each slice marking every pixel that is part of the lung. An example of one scan and the corresponding labeling is shown in figure 1.

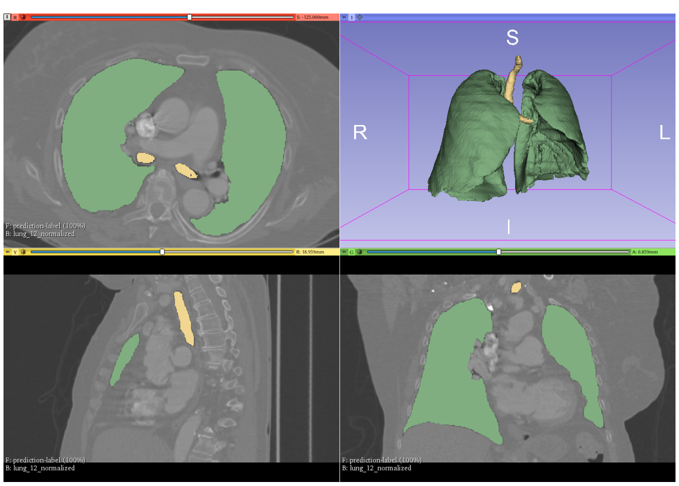

Fig. 1. CT scan of the lung and labeled parts. The first picture represents one slice of the image, the second a 3D representation of the label and the lower two pictures are visualizations of the side and front view.

The segmentation of the lung from the rest of the picture is the first step for further image processing. In the LUNA16 dataset the final goal for example is to detect nodules of the lung indicating cancer. Machine learning approaches can be a powerful support for the doctors who treat patients with suspected cancer. The algorithms can reduce human errors, and it might even have the potential to outperform human capabilities or could automatize the process of cancer detection. This could have a positive effect on health care quality and costs.
In this work two different neuronal networks will be trained and tested to segment the lung on the above described dataset. Furthermore they will be examined and compared under different metrics.
First a short overview on the DeepMedic Network and the U-Net will be given and will be related to other approaches for medical image segmentation. After that the process and implementation will be explained and then the results of the two algorithms will be examined and compared.

II. Background

A. DeepMedic architecture

DeepMedic is a 3D neural network. It has been initially used for segmentations in biomedical 3D scans, especially for detecting brain anomalies such as injuries, tumors and lesions. In this project we will try this algorithm for lung segmentations.
A DeepMedic model consists of detecting a particular pattern in an image. This is achieved by multi-layer convolution in the network. We distinguish two components in the model. First there is a three-dimension convolutional neural network (CNN) model used for dense segmentation, then a three-dimension fully connected conditional random field (CRF) model for deeper predictions. (Might we only use the first component?)
Each layer of the CNN model contains channels called feature maps, i.e. a group of neurons identifying a feature in the previous layer. Then a feature is defined by associated kernel weights. Number of inputs, outputs, feature-maps and kernels are tunable.

B. U-Net architecture

Semantic segmentation is partition of an image in coherent parts. U-Net is mostly used for biomedical image

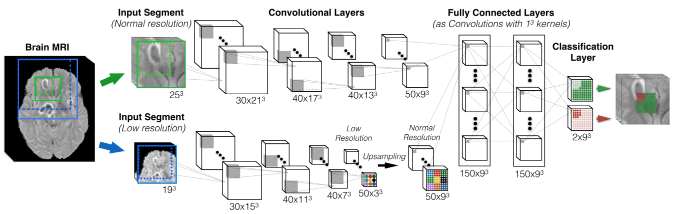

Fig. 2. Overview of DeepMedic architecture

segmentation. In figure 3 the structure of the U-Net is shown.

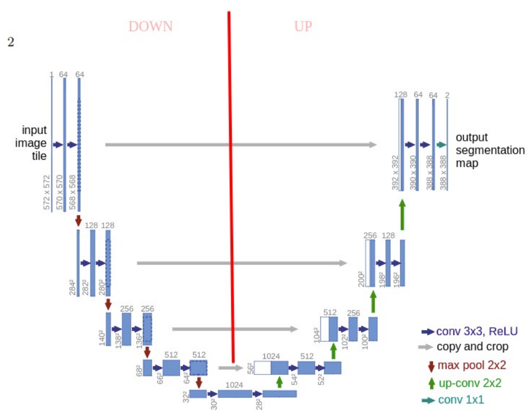

Fig. 3. Structure of U-Net architecture

Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. X-Y-size is provided at the lower left edge. White boxes represent copied feature maps. Arrows denote the different operations.
First part is called down or encoder part. Convolutional blocks followed by maxpool downsampling layers are applied to encode the input image into feature representations at multiple different levels. The second part of the network consists of upsampling and concatenation layers followed by regular convolution operations. The dimensions from left are expanded to meet the original image size. The grey and green arrows indicate where to concatenate future maps together.
In comparison to other fully convolutional segmentation networks, the main feature of U-Net is that while upsampling and going deeper in the networks, it also concatenates the higher resolution features from the down sampling part with upsampled features in order to better localize and learn representation.

C. Metrics

1) Dice Loss: formulas

2) BCE Dice Loss:

3) Hausdorff distance: The Hausdorff distance is a metric to calculate the maximum of the difference between two sets of coordinates.
For the set of coordinates A and B it is defined as following

$$d_{hausdorff}(A,B) = \max_{a\in A}(\min_{b\in B} d(a,b)) \tag{1}$$

where $d(a,b)$ represents the Euclid distance between the two coordinates a and b. It can be interpreted as the maximum deviation of set B to set A and is therefore a important metric for evaluating the difference between two label maps.

4) Mean distance: The mean distance is similar to the Hausdorff metric but is a measurement for the mean of the difference between two sets of coordinates.
It is therefore defined as following

$$d_{mean}(A,B) = \frac{1}{|A|}\sum_{a\in A}\min_{b\in B} d(a,b) \tag{2}$$

where again $d(a,b)$ represents the Euclid distance between the two coordinates a and b. It furthermore can be seen as the average difference between the two sets and will also be used for evaluation.

III. Related Work

text and subsections

IV. Process and occurring problems

overview

A. dataset

The Luna16 dataset consists of about 10 GB of mdh and raw files. The files contain a varying number of slices of 512 x 512 pixel images. For this work just a subset of the data was used for training and evaluation.

B. Preprocessing of image data

file conversion normalization on mean and sd picture of normalized loading data in numpy label transformation to just one/two labels

C. Training and validation

building of networks describe it (dice loss and bce dice loss) hold out evaluation threshold, export files to nii.gz

D. Handling huge dataset

- ram issue - generator solution - still problems with data import size

V. Evaluation

tensorflow keras, (implement unet on our own, used library for DeepMedic)

graph loss on training evaluation (dice loss, bce dice loss, batch size (correlates to gpu), epochs, steps per epoch, optimizer) prediction image on label and prediction evaluating and examining models with hausdorff and mean distance on pictures (graphs or statistics, table) comparison of unet and deepmedic

VI. Conclusion

- conclusion on network comparison - complex and specific networks especially for this kind of task - difficulties with huge data - training time and allocation of computational resources - draw conclusion on how good ml is for MIC, issues, ... - further steps (nodules, ...)

References

[1] "Lung nodule analysis 2016." [Online]. Available: https://luna16.grand-challenge.org/