Bachelor Thesis

Segmentation of three dimensional MRI data

Paul-Louis Pröve

mail@plpp.de

supervised by

Prof. Dr. Dennis Säring

dsg@fh-wedel.de

Hamburg,

July 11, 2017

Contents

1 Abstract

Password attacks are at the edge of accessing someones secrets. By learning to judge the strength of a password and by understanding how hackers execute attacks, users can make better estimations on how safe they are.

The entropy is widely used to measure how safe a password is, but many sources draw inaccurate conclusions between the entropy of a random password and the strength of a password that was chosen by a person. It is important to understand how these two differ and why realistic password strength is often hard to determine.

Todays hardware gives hackers incredibly powerful machines to launch different types of password attacks. Common password patterns lower possible permutations by such a magnitude that even seemingly safe passwords can be successfully attacked. In combination with frequently used passwords and personal information, hackers can further increase the effectiveness of their attacks.

By explaining common terminologies and analysing different datasets we will look at password attacks from the perspective of users, system administrators and hackers. All three benefit by understanding how the others operate in practice.

2 Introduction

Recent advances in Artificial Intelligence have led to workflows that not only run fully automated but also often exceed human performance. State of the art neural networks can classify images into thousands of categories more accurate and magnitudes faster than humans. They translate text between hundreds of languages, navigate cars autonomously through cities and detect intruders in computer networks. In most of these cases they have been trained on tens of thousands or even millions of data samples.

Neural networks have also found great success in the medical field,where new types of problems were introduced based on datasets that are much smaller. Medical imaging data includes MRI, Ultrasound and CT of which the latter is an invasive methods because it uses electromagnetic radiation.

3 Literature Review

3.1 Magnetic Resonance Imaging

text

3.2 Epiphyseal Plates

text

3.3 Neural Networks

text

3.4 Segmentation

texty text

4 Methods

4.1 Data Analysis

3d, mri, german males, 14-21, different resolutions, coronal/sagittal, segmentation maps for 76, mhd files, distribution of age.

4.2 Preprocessing

The number of parameters in a Neural Network commonly range from hundreds of thousands to hundreds of millions. This complexity allows the model to learn on its own what features of an image are relevant for any given task. It works in conjunction with the fact that high volumes of data are available for the training.

Because of the small dataset that was available for this study, several types of data preprocessing were applied to the images. These techniques reduce the amount of information per sample or the amount of variance between multiple samples. This results in a complexity reduction of the problem the network is supposed to solve. Other preprocessing methods experimented with the difference between 2D and 3D data as well as the influence of separate segmentation channels on the output.

4.2.1 Cropping

The framing of the raw images included large parts of the thigh and shin to be visible in the picture. Since these weren't relevant for the purpose of the study, they were cropped out. An algorithm was used to detect the center where Tibia and Femur meet and only use a square window around this point. There was no cropping applied on the z-axis.

4.2.2 Resizing

The images were also resized to a resolution that is common for Convolutional Neural Networks. 224x224 Pixels for width and height also allowed the use of Transfer Learning based on popular pre-trained models. This resolution still delivered enough detail for the segmentation maps.

On the z-axis everything was scaled to 36 slices, which meant a 1.5 upscale for the images provided by Jopp et al. and a minor downscale from the 41 slices that were taken specifically for this study.

Every image now had unified dimensions of 36x224x224 voxels.

4.2.3 Normalization

The normalization procedure of this dataset was executed in two steps. First the N4 Bias Correction was applied to the images, which tries to balance irregularities than happen when the MRIs were recorded. In the second step all intensity values were normalized to range from 0 to 1 so that every input is on the same scale.

4.2.4 2D and 3D data

Every three dimensional image can be converted to n two dimensional slices, where n is the resolution of the sliced axis. Since the z-axis shows a much lower resolution than x and y, each image was sliced in 36 224x224 2D images. This resulted in 36 times more samples, but reduced the information per image by the same factor.

Using three dimensional convolutions turned out to be helpful for the segmentation, because the network was able to draw conclusions from the order of the slices within one image.

4.2.5 Separate Bone Maps

The initial segmentation maps came with three separate channels for the Femur, Tibia and Fibula. With this information it was possible to train a model that would segment the three bones while still differentiating between them. This helped the accuracy of the prediction opposed to using just a single channel for all of the bones. In places where the Femur and Tibia were very close to one another, the separate channels prevented the closing of this region by the network.

For another experiment the three channels were treated as one to create a network that would segment any type of bone in the image. This resulted in better performance when applied to sagittal images of the knee provided by Maas et al. The network was able to generalize on a situation it wasn't trained on.

4.3 Architectures

In search for a network that would perform well on the segmentation, different architectures were looked at and multiple settings were tried.

4.3.1 Patch and Image based

4.3.2 Channels

growth and initial size

4.3.3 Dropout

Dropout is a popular regularization technique that randomly zeros out a fraction of the weights during training. It is understood that this helps the model to generalize better and reduce overfitting on a given dataset. Well known image classification architectures like VGG16, SqueezeNet or AlexNet use Dropout near the end of the network. Similarly U-Net uses Dropout at the end of the contracting path of the architecture to implicitly add image augmentation to the data.

Since overfitting was a big problem of this study, several other uses of Dropout were investigated. These experiments included increasing the dropout rate, as well as applying it in multiple places of the network. Adding a Dropout of 0.2 between the convolutional layers on the contracting side yielded the best results. While it didn't hurt the performance when low amounts of dropout were also added to the expanding side, it had a negative effect on the training speed.

4.3.4 Batch Size

The batch size describes the number of samples that will be used for one forward- and back-propagation step.

4.4 Training

IoU, Adam, early stopping, LR policy

5 Results

best model performance, parameter count, size

6 Discussion

age prediction, size of data, augmentation