Imperial College London

Imperial College London
Derpartment of Mathematics

Segmentation of CT scans into Atrium/non Atrium

Thomas Vogel

CID:

Supervised by Prof Giovanni Montana

12th August 2015

Submitted in partial fulfilment of the requirements for the MSc in Statistics of Imperial College London

The work contained in this thesis is my own work unless otherwise stated.

Signed: Date:

Abstract

This is a LaTeX template to be used for project theses by the students of the Imperial College MSc in Statistics. Please have a look at the `*.tex` source and the code comments therein. You may use this template with only minor changes, or make any major style changes you desire (e.g. redesigning the title page, adding headers,...), or you may use a different template, or you may write your own approach from scratch, or you may just use some bits and pieces from the template's LaTeX code. It's up to you. Note that resources on how to use LaTeX are available on Blackboard under 'R&LaTeX intro'.

This template will quote relevant sections from the MSc in Statistics student handbook throughout; e.g.

> " The abstract should be a brief statement of the aims and outcomes of the project, to summarise/advise even for a casual reader!"

Acknowledgements

Thank you supervisor/friends/family/pet.

“ Include an acknowledgement.”

Table of Contents

1 Background Material: on convolutional neural networks

A convolutional neural network (CNN) is a specialised kind of feed-forward Neural Network that replaces standard matrix multiplication for a convolution. A convolutional layer has three components: a convolution part, an activation part, and a subsampling/pooling part.

1.1 Convolutional Neural Networks

1.1.1 Convolution

The input is processed by several kernels with learnable weights, each producing a set of outputs called feature maps. This leverages three ideas:

- sparse interaction: every output node is connected to a local sub- set of inputs.
- parameter sharing: the same kernel is used for every output node in a given feature map.
- translational equivariance: shifting the input results in an equiv- alent shift in the output.

1.1.2 Activation

Every node in the feature map is passed to an activation function, usually a Tanh unit or a Rectified Linear unit.

1.1.3 Subsampling/pooling

Reduces the output with a local summary statistic, e.g. maximum or average. This reduces the layer size and adds local translational invariance.

1.1.4 Typical Architecture

The architecture considered consists in an input layer which are square patches centred at the voxel of interest, two convolutional layers, a fully connected layer, and an output layer. Each layer in turn represents more abstract features the deeper its location in the architecture. Compared to a deep multiple layer perceptron, this architecture has the ad- vantage of being memory and computationally efficient.

2 Setting up the problem

2.1 Data

The data used is a set of 27 3D CT images that are 480*480*50 in dimension. They come in 2D arrays in DICOM files. The labelling of all the voxels come in a 3D matrix in a NRRD file.

2.2 The Tri-Planar Method

Classifying the voxels require building an input set containing local and global information to allow the Neural Network to extract relevant information and aid in its choice. One of the cheap ways of doing so is to use the so-called tri-planar method. This consists in generating 3 perpendicular square patches of a given size for each voxel to classify. Each patch is then fed into a different input channel of the Convolutional Neural Network, and are then connected to the classifying layer using a fully connected layer.

2.3 Technological Details

2.3.1 Libraries

We use an open-source library called Torch in Lua to train the Neural Networks and Python and a number of its libraries to handle all the logistics from generating the datasets to producing plots of segmentation results.

2.3.2 Computer Power

The code for training Neural Networks has been written to work on multi-GPU platforms that increase the speed of training by around 100 times. The graphic cards consist in two NVIDIA Tesla K40m and two Tesla K20Xm.