

# **MSc. Thesis VIBOT**

Emmanouil Koutoulakis

Department of Computer Architecture and Technology  
University of Burgundy

A Thesis Submitted for the Degree of  
MSc in Vision and Robotics (VIBOT)

· 2022 ·

## Abstract

Magnetic resonance (MR) images play a vital role in rapid diagnosis and treatment planning. Physicians analyze images to determine the size of the organ, to plan the therapy and then to observe the evolution of the disease. Unfortunately, this process is deemed a time-consuming and laborious work for the physicians, creating impediments in the fast treatment planning. Following this incentive, the objective of this thesis is to evaluate and propose an efficient learning-based method for automatic segmentation of the organs at risk (OARs) in pelvic region in the context of radiotherapy planning, using MR images. Primarily, we centralized on organ-per-organ and multi-organ segmentation using Convolutional Neural Networks (CNN) trained on MR images which are manually segmented by experts. We tested 3 state of the art architectures to select the optimal one. The experiments were both 2D and 3D CNNs. Then, we concluded to one segmentation model having the best results and being efficient in terms of computational complexity. Due to the lack of enough patients, we implemented a data-augmentation methods to increase the size of the training set.

*Research is what I'm doing when I don't know what I'm doing. . . .*

Werner von Braun

# Contents

<b>Acknowledgments</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Preparing your dissertation . . . . .	1
1.1.1 Paper . . . . .	1
1.1.2 Font and margins . . . . .	1
<b>2 Background</b>	<b>2</b>
2.1 Introduction . . . . .	2
2.1.1 Interpolation . . . . .	3
<b>3 Limitations</b>	<b>5</b>
3.1 Class Imbalance . . . . .	5
3.2 Over-fitting . . . . .	5
<b>A The first appendix</b>	<b>7</b>
<b>Bibliography</b>	<b>7</b>

# List of Figures

2.1	Interpolation of corresponding coordinates . . . . .	4
-----	--	---

# List of Tables

# Acknowledgments

Any acknowledgments???

# Chapter 1

## Introduction

The objective of this chapter is to intuitively explain the thesis from its clinical and technical point of view. To begin with, we discuss aspects that are crucial to understanding the clinical incentives of the proposed approach. We start by defining the pelvic region and organs at risk (OARs), which are the main focus of this thesis. Then, we present the commonly used types of medical images, and we discuss the role of segmentation tasks in treatment planning. Second, we discuss important aspects related to deep learning, atlas-based and image processing techniques, which form the basis of most of the current state-of-the-art methods for image segmentation. In particular, we discuss its main advantages and drawbacks in the segmentation of organs at risk. In the last phase, we establish our main contributions and the structure of the thesis

### 1.1 Preparing your dissertation

You are strongly encouraged to use the Latex templates provided.

#### 1.1.1 Paper

The manuscript should be in A4 size, and the printed paper should be of at least 70 gsm.

#### 1.1.2 Font and margins

Thesis should be printed on both sides of the paper. Use no less than 1.5 spacing, with quotations and notes single-spaced. Regarding **Character size**, not less than 2.0mm for capitals and 1.5mm for x-height (the height of a lower-case x). Use a serif font (i.e. Times) between 10 and 12 points. Use consistent and clear fonts through all the document.

## Chapter 2

# Background

### 2.1 Introduction

Many scientific subject areas (i.e. medicine, remote sensing) have to deal with the problem of image correspondence (also known as image registration or image matching). This problem arises when there is a need to extract information about an object by comparing a set of different images in which this object appears. Most of the time this comparison cannot be performed easily because of differences between images such as viewpoint changes, use of different sensors, images taken at different times, and changes in general imaging conditions. Given such differences, a way to align different images of the same object is needed. In medical applications, and specifically in mammography, radiologists have to deal with all the above problems. For example, a way to detect abnormal structures in the breast is to compare temporal (images of the same breast taken at different times) or contralateral (using left-right breast) mammograms.

As images are taken under different conditions (i.e. film sensitivity, radiation exposure, breast compression and patient movement) and/or time intervals (e.g. screening interval of three years) the structure of the breast is likely to suffer some changes, although an overall similarity will be maintained. A way to detect those changes is to align the images and compare the results using, for instance, simple image subtraction. In addition, as described in the previous chapter, different imaging modalities are being used to provide a better understanding of a region of interest. The first step to incorporate information from those modalities is to align them to be able to establish areas of correspondence.

An image correspondence method is based on finding a mapping function ( $f(x, y)$ ) that maps each coordinate of one image ( $A$ ) into another image ( $B$ ).



$$B(x', y') = A(f(x, y)) \quad (2.1)$$

which maps spatial coordinates  $(x, y)$  in the reference image  $A$  to coordinates  $(x', y')$  in the warped (also referred to as target) image  $B$ . Usually the function  $f$  is expressed as two (or three in three dimensions) separate functions  $f_x$  and  $f_y$  (and  $f_z$ ). This review describes image registration in two dimensions. However extrapolation to 3D is often straightforward and is described in detail where appropriate.

A general methodology for image alignment typically follows those steps:

1. *Selection and extraction of features.* A registration method can be based on matching different image primitives. For instance points, edges, ridges, surfaces, or a whole image. The selection of a suitable primitive is a trade off between the information it provides and its complexity. For instance, raw pixels with only intensity information provide a straightforward comparison. On the other hand, taking the whole image provides more (and more complex) information. Primitives will be described by a set of features extracted from them. That is the case, for instance, of shape features extracted from regions or curvature values extracted from linear structures. There is an extensive literature about feature extraction, the reader is referred to Chapter 4 where a review on that subject is given.
2. *Similarity metric.* A way to measure the similarity between the extracted features is needed. Many different similarity metrics have been proposed and their suitability depends on the chosen feature. The reader is referred to Chapter 5 where similarity measures are reviewed and a novel measure is proposed.
3. *Selection of the mapping function and estimation of its parameters.* The complexity of a mapping function should be determined depending on the application field and type of misalignment we are dealing with. Once a mapping function is selected its parameters need to be estimated. This requires a definition of an optimum search space and search strategy which is often related to optimisation of a cost function related to a similarity measure previously mentioned.
4. *Alignment of images using the mapping function.* When the function parameters are known we are able to transform an image in order to minimise the initial misalignment.

### 2.1.1 Interpolation

Usually transformed coordinates  $(x', y')$  in the warped image do not match a given coordinate grid (i.e. they are non-integer values) and interpolation is needed. Different interpolation

methods have been used in image registration: nearest neighbour, tri-linear and partial volume distribution. The interpolation problem is graphically represented in Figure 2.1.

Figure 2.1: Interpolation of corresponding coordinates

Nearest neighbour interpolation assigns the intensity value of  $m = (x', y')$  to its nearest point:

$$B(m) = B(\underset{i}{\operatorname{Argmin}}(d_E[m, n_i])) \quad (2.2)$$

where  $d_E$  indicates Euclidean distance and  $n_i$  are the nearest neighbours. The nearest neighbour approach provides good accuracy and it is easy to compute.

Trilinear interpolation assigns to  $m$  the value of the weighted ( $w_i$ ) sum of the neighbouring pixels.

$$B(m) = \sum_i w_i B(n_i) \quad (2.3)$$

where each weight  $w_i$  is related to the distance from  $m$  (see Figure 2.1)

$$w_1 = (1 - dx)(1 - dy) \quad w_2 = (1 - dx)dy \quad w_3 = dx(1 - dy) \quad w_4 = dxdy \quad (2.4)$$

This interpolation creates a new intensity value, which could introduce undesirable effects in the intensity distribution. This fact is regarded as its main drawback, and can be solved using a partial volume distribution. This method adds the weight of each neighbour ( $w_i$  values) at the intensity value in the histogram distribution, without creating an additional intensity value.

## Chapter 3

# Limitations

### 3.1 Class Imbalance

One limitation in multi-organ segmentation is class imbalance. Class imbalance occurs when there are one or more classes that are more frequent occurring than the other classes. Simply put, there is a skewness towards the majority class. For instance, for segmentation of pelvic region, rectum and bladder are often different size or position from the femoral heads. Hence, training a network with class imbalanced data is causing unstable segmentation model, which is biased towards the classes of different organs. To alleviate this problem, a right use of loss functions is crucial for these tasks.

Proposed loss functions that tackle this problem

- Combo loss
- Focal Tversky loss
- 

### 3.2 Over-fitting

Another crucial challenge in multi-organ segmentation is data scarcity. In other words, the model has significantly good performance on the training set but weak performance on the validation and test set. This could be caused due to the lack of available dataset, on the grounds that manual segmentation of the organs is deemed a laborious and time-consuming process. Furthermore, the wrong (or partially wrong) delineated dataset produces similar consequences,

---

provoking misunderstandings on the trained model. A dominant solution to tackle this issue is data augmentation, however, a large dataset is essential for multiorgan segmentation.

## Appendix A

### The first appendix

If you need to add any appendix, do it here... Etc.