POLITECNICO DI TORINO

Master's Degree in Biomedical Engineering

Development of an automatic algorithm for the detection of the third lumbar vertebra in CT scan

Supervisor

Prof. Patricia Sánchez González

Prof. Filippo Molinari

Candidate

Constantin Bodnarescu

March 2021

Acknowledgements

Foremost, I would like to express my gratitude to Alexander Peter Seiffert, who always helped me to carry out this work at Universidad Politècnica de Madrid, and prof. Sanchez for giving me this opportunity.

Finally, I would thank all the persons who stayed close to me during this difficult period, especially my family and my friends. Without them, it would have been harder, but I am very luck to have such person in my life.

Summary

Muscle loss and tissue degradation are two conditions which manifest approximately in 40% of cancer patients. These conditions can also become a disease like sarcopenia or cachexia. The causes of cancer-related weight loss are different, and this condition reduces the response for both medical and surgical cancer treatments, as well as it increases the risk of mortality. For these reasons, weight loss should not be underestimated but, on the contrary, it has to be monitored in order to improve the benefits of cancer treatments.

There are different techniques to evaluate body composition and diagnose cancer-related muscle loss, both non-imaging techniques and imaging techniques. Several studies have shown that it is no more necessary to segment muscle tissue in the whole body 3D image, because techniques like magnetic resonance imaging (MRI) or computer tomography (CT) provide the possibility to study the third lumbar vertebra level (L3) where muscle area is correlated to whole body muscle mass. Thus, spotting L3 slice is an important task and the radiologist needs, at least, 5 minutes to detect and segment it.

Up until now, only few automatic methods based on computational neural network (CNN) have been proposed in order to detect L3 vertebra. However, they require a training part and, moreover, a big dataset and this can be a problem for the ones who do not have enough data. Therefore, the aim of the study in to create an automatic algorithm for L3 slice detection without using neural network architectures, so that a small database would not be a problem. To this effect, 24 patients are analysed in this study and, considering that every patient provides 3 or 4 CT scans, the final database is made by 80 acquisitions. For the purpose of the study, the technique is based on human body spinal cord characteristics at thoracic and lumbar level. Every CT image is converted into a binary mask to highlight bone presence, then the there is a focus on the number of white pixels (NoP) which denote a sort of minimum at lumbar vertebrae level. This consideration is used to analyse pixels presence just in a certain area of the masks, a region corresponding to vertebrae positions, called vertebra area mask (VAM). Thanks to the VAM it is possible to examine just the pixels referred to vertebrae in lumbar interval and, sliding down from the thoracic to the sacral ones, there is a sort of pending/rising trend of the NoP. Every maximum corresponds to a vertebra position.

As a result, the mean errors in slices is under 5, and for 50 out of 80 CT scans the error is under 3. In addition, for some patients the error is between 10 and 20 and it has been proved that in such cases the algorithm gives as output vertebra adjacent to L3. Finally, despite the limitations to be solved, it is concluded that the algorithm can be used especially for the ones who do not have the possibility to obtain a big database.

Keywords

Weight loss, Sarcopenia, Body Composition (BC), Third lumbar vertebra (L3), Computer Tomography.

Resumen

La pérdida muscular y la degradación del tejido son dos afecciones que se manifiestan aproximadamente en el 40% de los pacientes de cáncer. Estas afecciones también pueden convertirse en una enfermedad como la sarcopenia o la caquexia. Las causas de la pérdida de peso relacionada con el cáncer son diferentes, y esta condición reduce la respuesta para los tratamientos médicos y quirúrgicos del cáncer, así como aumenta el riesgo de mortalidad. Por estas razones, la pérdida de peso no debe subestimarse, sino que, por el contrario, debe vigilarse para mejorar los beneficios de los tratamientos del cáncer.

Hay diferentes técnicas para evaluar la composición corporal y diagnosticar la pérdida muscular relacionada con el cáncer, tanto técnicas no de imagen y técnicas de imagen. Varios estudios han demostrado que ya no es necesario segmentar el tejido muscular en la imagen 3D de todo el cuerpo, porque técnicas como la resonancia magnética (RMN) o la tomografía computarizada (TC) ofrecen la posibilidad de estudiar el tercer nivel de vértebras lumbares (L3) donde el área muscular se correlaciona con la masa muscular de todo el cuerpo. Por lo tanto, detectar la rebanada L3 es una tarea importante y el radiólogo necesita, al menos, 5 minutos para detectarla y segmentarla.

Hasta ahora, sólo se han propuesto pocos métodos automáticos basados en la red neuronal computacional (CNN) para detectar vértebras L3. Sin embargo, requieren una parte de entrenamiento y, además, un gran conjunto de datos y esto puede ser un problema para aquellos que no tienen suficientes datos.

Por lo tanto, el objetivo del estudio es crear un algoritmo automático para la detección de cortes L3 sin utilizar arquitecturas de red neuronal, de modo que una pequeña base de datos no sea un problema. A tal efecto, se analizan 24 pacientes en este estudio y, considerando que cada paciente proporciona 3 o 4 tomografías computarizadas, la base de datos final se realiza mediante 80 adquisiciones. A efectos del estudio, la técnica se basa en las características de la médula espinal del cuerpo humano a nivel torácico y lumbar. Cada imagen de TC se convierte en una máscara binaria para resaltar la presencia o sea, entonces hay un enfoque en el número de píxeles blancos (NoP) que denotan una especie de mínimo a nivel de las vértebras lumbares. Esta consideración se utiliza para analizar la presencia de píxeles sólo en un área determinada de las máscaras, una región que corresponde a las posiciones de las vértebras, llamada máscara de área de vértebras (VAM). Gracias al VAM es posible examinar sólo los píxeles referidos a las vértebras en el intervalo lumbar y, deslizándose desde las torácicas a las sacras, hay una especie de tendencia pendiente/ascendente de la NoP. Cada máximo corresponde a una posición de vértebras. Como resultado, los errores medios en los cortes están por debajo de 5, y para 50 de 80 tomografías computarizadas el error está por debajo de 3. Además, para algunos pacientes

el error está entre 10 y 20 y se ha demostrado que en tales casos el algoritmo da como vértebra de salida adyacente a L3.

Finalmente, a pesar de las limitaciones a resolver, se concluye que el algoritmo puede ser utilizado especialmente para aquellos que no tienen la posibilidad de obtener una gran base de datos.

Palabras clave

Pérdida de peso, sarcopenia, composición corporal (BC), tercera vértebra lumbar (L3), tomografía computarizada.

Contents

1 Introduction 1

1.1 Cancer and weight loss . . . 1

1.2 Sarcopenia and Cachexia: definition and epidemiology . . . 2

1.3 Diagnosis . . . 4

1.3.1 Non-imaging techniques . . . 5

1.3.2 Imaging techniques . . . 6

1.4 Treatments . . . 11

1.5 Context . . . 12

1.6 Chapter structure . . . 12

2 Objectives 14

3 State of art 15

3.1 Definition of dataset and pre-processing . . . 15

3.2 Spotting L3 slice using deep learning algorithm . . . 16

3.2.1 Dataset and pre-processing . . . 16

3.2.2 Model Architecture . . . 16

3.2.3 Data processing . . . 17

3.2.4 Decision process . . . 18

3.2.5 Results . . . 19

3.3 Automatic L3 slice detection in 3D CT images using fully-convolutional networks (L3 UNet-1D) . . . 20

3.3.1 Dataset and pre-processing . . . 20

3.3.2 Model Architecture . . . 22

3.3.3 Results . . . 24

3.4 Automatic Vertebra Annotator . . . 25

4 Materials and methods 27
4.1 Materials . . . 27
4.1.1 Database . . . 27
4.1.2 Software . . . 27
4.2 Methods . . . 28
4.2.1 Data preparation . . . 28
4.2.2 Data processing . . . 29
4.2.3 Anatomical features of spinal cord . . . 31
4.2.4 Validation . . . 34
5 Results 35
5.1 Pre-processing . . . 35
5.1.1 Interval of Interest . . . 36
5.1.2 Vertebrae Area Mask . . . 39
5.1.3 L3 detection . . . 41
5.2 Validation . . . 42
6 Discussions 45
7 Conclusion 49
A Ethical, economic, social and environmental aspects 55
A.1 Introduction . . . 55
A.2 Most relevant impacts of the project . . . 55
A.3 Conclusion . . . 56
B Economical budget 58

List of Figures

1.1 Most prevalent cancer types in 2018 in the world (both sexes, all ages) [2]. 1

1.2 Relationship between muscle strength and age [8]. 3

1.3 Survival rate of patients stratified by the presence of sacropenia [12]. 4

1.4 Representation of muscle-lipid system, with explanation of how muscle and lipid tissue are composed [4]. 5

1.5 Calculating BC with a DEXA image. (a): bone-fat-lean tissue mass map. (b): pixel-level fat percentage map. (c): fat percentage threshold map. In (d) t1 = 25% and t2 = 60% are the threshold for DXA system by default. (e): intensity thresholds [14]. 6

1.6 Main differences and locations of human vertebras [17]. 7

1.7 Example of two images at the same level: (a): T1 weighted MRI and it is possible to see in a better way soft tissue and adipose tissue. (b): T2 weighetd MRI image in which liquids and pathological conditions (cancer, inflammation, trauma) are seen in a better way [22]. 8

1.8 Example of different tissues at L3 level [6]. 10

3.1 Illustration of the main steps used to estimate L3 slice [30]. . . . 17

3.2 Example of normalized MIP images with L3 position highlighted by the green line [30]. 18

3.3 Example of CNN output sequence with the typical straight like of slope (theoretical one in green. In the graphyc of the right there is illustrated the correlation between CNN output and slope [30]. 19

3.4 Slice thickness distribution and image heights of the dataset [18]. 20

3.5 Examples of extractions of different MIP images from a stack of CT slices. From the starting stack of CT slices, it is possible to extract a frontal MIP image (on the left), a sagittal MIP image (on the center) and a restricted sagittal MIP image (on the right) [18]. 21

3.6 Network architecture of the 1D and 2D versions [18]. 22

3.7 Different images with ground-truth confidence map (in red) [18]. 24

3.8 Prediction outputs: four different predicted confidence maps. (a) and (b) refer to a L3UNet-2D frontal view, (c) refers to a L3UNet-1D frontal view and (d) refers to a L3UNet-2D sagittal view [18]. 25

3.9 Outlier cases: incorrect identification of L3 vertebra by L3UNet-2D (on the left) and by L3UNet-1D (on the right) [18]. 25

3.10 Example of AVA interface [33]. 26

4.1 Pipeline of the methods adopted in this study 28

4.2 Figure (a): schematic representation of the IOI detection containing the L3 slice. Figure (b): theoretical intersection between the threshold and the spinal cord [35]. 30

4.3 Main differences between thoracic vertebrae and lumbar vertebrae [17] 31

4.4 Example of CT slices from T12 vertebra to L4 vertebra. The most relevant features are explained in the figure [17]. 32

4.5 Focus on L4 vertebra which usually has shorter transverse processes and at its level can be presence of iliac crest [17]. 32

5.1 Figure in the left shows the result after image thresholding. The enhanced mask on the right is obtained thanks to the post-processing 35

5.2 Two different examples of mask creation. The left column shows the original CT slices taken from different patients, whereas the right column shows the respective binary masks. 36

5.3 Examples of the typical "downhill" trend in correspondence with the lumbar vertebrae. The green line corresponds to the threshold used to find the IOI. The orange curve corresponds to the smoothed NoP. The ellipses highlight the IOI 37

5.4 Red lines approximate the slope of the descending part, while black approximate the slope of the rising trend. These examples show how usually the rising slope is bigger (in absolute value) than the descending slope 38

5.5 Some graphs of NoP can produce an IOI that is too large. In cases like the one in figure (a), the IOI is automatically reduced to avoid creation of VAMs as the one in (b) 39

5.6 (a) (b) (c): three examples of VAM obtained from patients without relevant problems at the spinal cord. (d): example of an VAM probably obtained from a patient suffering from scoliosis, and this reflection is made by considering the non-symmetry of the VAM with respect to its Y axis 40

5.7 (a): example of region growing process: the user chooses the seeds with the mouse and the process starts, giving as output the VAM made by the elements connected to seeds. (b): Examples of the automatic method, which chooses the element with the largest area for the creation of the VAM. 41

5.8 Peak positions corresponding to lumbar vertebrae positions . . . 42

5.9 Group of CT slices with respective masks. These images show how, thanks to peaks detection, also the others lumbar vertebra can be estimated 42

5.10 (a) Graph of the errors calculated as the difference between the prediction and the correct L3 position. (b) boxplot of the errors, to show the shape of the distribution, its central value and its variability. 43

6.1 Histogram of the prediction errors. As it is possible to notice, a good percentage of prediction has an error under 5 slices, whereas the errors in the range [10-20] (in absolute values) usually refers to a vertebra adjacent to L3 (L2 or L4) 47

6.2 (a) (b) (c) are the NoP graphs of the outliers. For (a) and (b) the prediction become correct if the number of CT slices is reduced to 400, starting from the top (head). Regarding (c) there is still no prediction even with the reduction of the CT scan. Its shape is weird in respect to the graph in (d), which is the NoP of a CT scan with correct prediction. 48

List of Tables

1 Main characteristics of MRI, CT, DEXA and BIA [5]. 10

2 Comparison between errors given by CNN and by three different radiologists [30]. 20

3 Errors between annotators A and B, reported in mm and in number of slices. The error in slices was calculated by dividing the error in mm by the slice thickness (without rounding) [18]. 21

4 Cross validation results for L3UNet-2D and L3UNet-1D (both frontal and sagittal view). The last column refers to the number of outliers with error > 10 [18]. 24

5 Mean and SD of the computational time calculated for both the automatic (AM) and the semi-automatic methods (SAM). 44

6 Main ethical, economic and social aspects 56

7 Human Resources Cost 58

8 Material Resources Cost 58

9 Total cost. 59

Glossary

AM	Automatic Method
AVA	Automatic Vertebra Annotator
BC	Body Composition
BIA	Bioelectrical Impedance Analysis
CNN	Convolutional Neural Network
CT	Computer Tomography
DEXA	Dual Energy X-Ray Absorptiometry
DICOM	Digital Imaging and Communications in Medicine
FCNN	Fully Convolutional Neural Network
FFM	Fat-Free Mass
GH	Growth Hormone
HU	Hounsfield Value
IOI	Interval Of Interest
L1/L2/L3/L4/L5	First/second/third/fourth/fifth Lumbar vertebra
LBM	Lean Body Mass
MIP	Maximum Intensity Projection
MRI	Magnetic Resonance Imaging
NoP	Number of white Pixels
SAM	Semi-Automatic Method
SD	Standard deviation
T12	Twelfth Thoracic vertebra
TBW	Total Body Water
TCIA	Cancer Imaging Archive
TL	Transfer learning
VAM	Vertebrae Area Mask
VitD	Vitamin D

1 Introduction

In this chapter, the clinical context of the master's thesis is presented. Muscle loss in cancer patients, which can also become a disease (for example sarcopenia or cachexia), is explained and the methods of diagnosis as well as the possible treatments are examined. Tissue degradation is present in approximately 40% of cancer patients and this condition negatively affects the prognosis and survival rate of cancer patients [1].

1.1 Cancer and weight loss

Cancer disease is one of the biggest problems in the medical world, considering that in 2018 there were more than 18 million cases worldwide and this number is constantly increasing year by year [2]. This disease is not uniform but can affect different organs of the body in different ways and this makes it difficult to reduce the level of mortality. As it is possible to see in Figure 1.1, in 2018 the most affected organs were colon (13.7%), breast (12.1%), prostate (11.7%) and lung (10.1%).

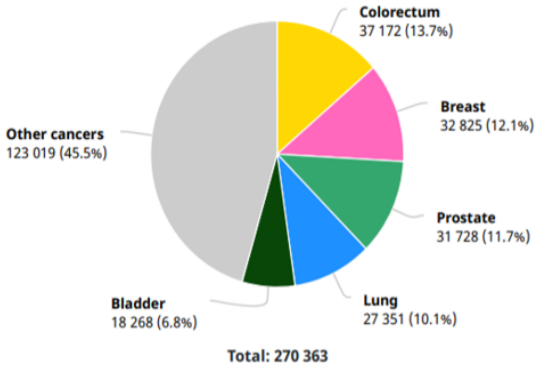


Figure 1.1: Most prevalent cancer types in 2018 in the world (both sexes, all ages) [2].

Cancer disease is very often associated with a progressive decline in the state of nutrition. In simple words, this means that the various elements that make up the body (organs, muscles, blood) are no longer in the state of balance

that characterizes the healthy individual (good state of nutrition). The most obvious sign of this condition is typically weight loss [3].

The causes of cancer-related weight loss are different. The most frequent are changes in the metabolism due to the cancer itself, but cancer treatments also have important side effects on the organism. The growth of the cancer induces changes that influence not only the organ or the affected tissue, but the entire organism. In fact, the cancer is perceived by the organism as a foreign body which triggers an immune system response, resulting in the production of substances that result in major changes in metabolism [3].

Cancer therapies (chemotherapy, radiotherapy, molecular target therapies) are effective, but they can cause appetite reduction, muscle weakness, nausea, vomiting, mouth ulcers and gastrointestinal mucous membranes alteration, especially in the period immediately following the therapeutic cycle. As a result, the patient also loses weight due to therapy. In addition, if the cancer required extensive surgery with removal of important parts of the digestive system, this can lead to alterations in digestion and nutrient absorption [4].

Weight loss reduces the response to both medical and surgical cancer treatments and increases the risk of mortality. For this reason, weight loss should not be ignored, nor underestimated, neither by the oncologist nor by the patient who must ask his doctor to be weighed regularly.

If natural nutrition is still possible, it is recommended to develop a personalized diet plan in the context of a nutritional program, taking into account the food preferences of the patient, if possible. If the amounts of food taken are insufficient to meet the nutritional needs of the patient due to anorexia or the gastrointestinal consequences of therapies, natural nutrition should be supplemented with specific products called oral nutritional supplements. If this approach proves insufficient, it is necessary to consider, on the advice of the doctor, the use of artificial nutrition, enteral or parenteral, in hospital or at home [4].

1.2 Sarcopenia and Cachexia: definition and epidemiology

The clinical definition of sarcopenia states that it "is a condition characterized by loss of skeletal muscle mass and function" [5]. In the last decade, the relation between diseases like sarcopenia (or cachexia) and cancer patients has been analysed [5]. Regarding cachexia, this pathology is considered as a side effect in cancer patients and it consists in muscle degradation[6].

These pathologies do not have a general definition valid for all the cases but they both have muscle tissue as target. In particular, sarcopenia involves a continuous and uniform loss of skeletal muscle mass, whereas cachexia can be more aggressive, causing an extreme muscle deterioration [7]. Since chronic

patients (especially cancer patients) often suffer from muscle loss, it is important to monitor the muscle mass. Studying muscle loss can help physicians to improve the benefits of cancer treatments. The causes of that can be due to different aspects like malnutrition, drug-related side effects, sickness and nausea. For example, 40% of the new diagnosticated cancer patients say that they experienced unexpected loss of weight, with other side effects being fatigue and weakness [1]. Figure 1.2 illustrates the correlations between a decrease of muscle strength and age, in correspondence of presence of a disease [8].

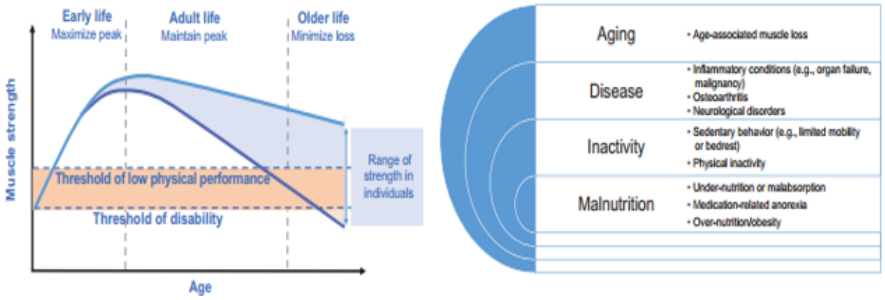


Figure 1.2: Relationship between muscle strength and age [8].

Different studies have shown that a prevalence of sarcopenia in advanced cancer patients exists, considering different types of cancer, different stages of the illness and different ways to analysing it [9] [10] [11]. Usually, muscle mass loss starts to appear during the third decade of life, which could be seen as a turning point where age can be associated with skeletal mass.

Scientific research has shown on more than one occasion that proper nutrition delays the onset of sarcopenia and prevents its worst consequences [9]. Figure 1.3 shows how the presence of sarcopenia decreases the survival rate of cancer patients [12].

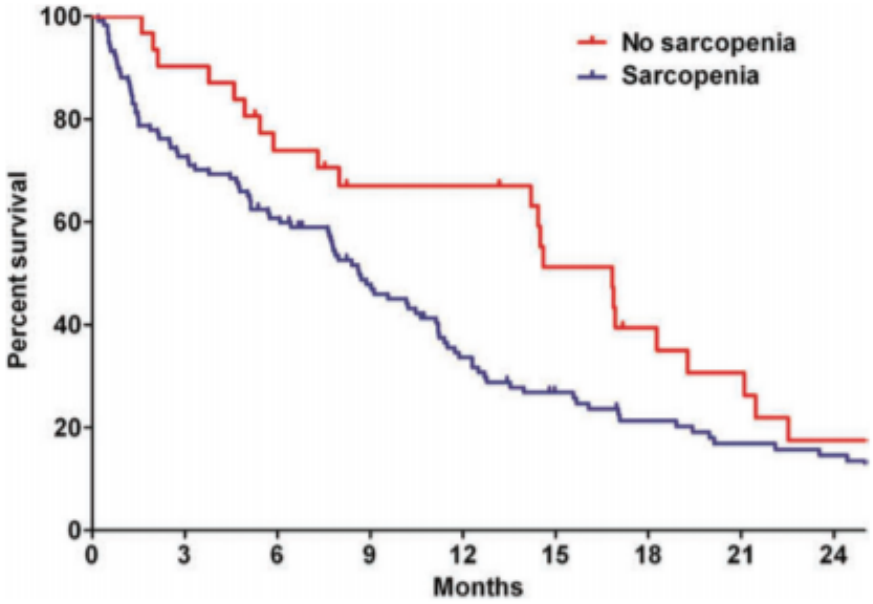


Figure 1.3: Survival rate of patients stratified by the presence of sacropenia [12].

1.3 Diagnosis

The cellular level of body composition (BC) consists of body cells (body cell mass) and their surrounding extracellular water, plus the skeleton and connective tissue. Although there is some lipid in the form of cell membranes, this compartment is largely fat-free and these components are sometimes termed the fat-free mass (FFM) or in older terminology the lean body mass (LBM) [13]. The body cell mass is responsible for almost all of the basal energy expenditure of the body, since that is where cellular metabolic and respiration processes take place. Together with the adipose tissue compartment (which consists mostly of fat), this level is often referred to as a two-compartment model, i.e., FFM and fat mass (FM). In the healthy individual, the FFM has a relatively constant composition, with a water content of 72–74%, an average density of 1.1 g cm-3 at 37 °C, a potassium content of 60-70 mmol kg-1 in men and 50–60 mmol kg-1 in women, and a protein content of 20% [13]. BC is an important factor to be studied, in patients with chronic diseases. This factor can be used also to enhance aspects like nutrition care and to personalize chemotherapy dose calculation taking in consideration muscle and adipose percentages. Figure 1.4 illustrates the muscle-lipid system.

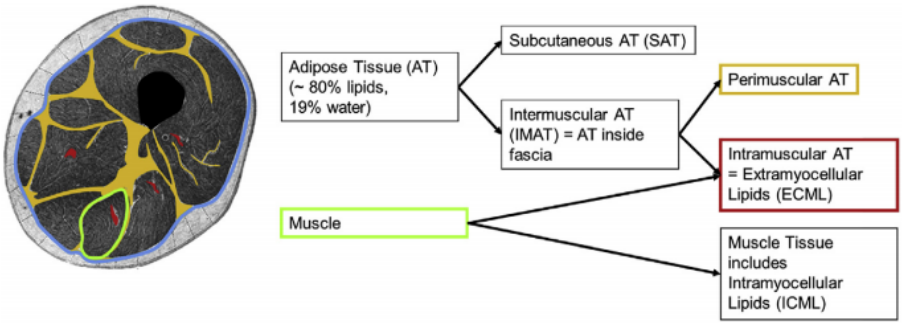


Figure 1.4: Representation of muscle-lipid system, with explanation of how muscle and lipid tissue are composed [4].

There are different techniques to evaluate BC and to diagnose cancer-related muscle loss, both imaging techniques and non-imaging techniques. For example, regarding the non-imaging techniques the most famous are Anthropometry (which measures subcutaneous fat) and Bioelectrical impedance analysis (BIA), which consists of a small AC electrical current injected through the body. Thanks to the inverse relationship between impedance and total body water, BIA can estimate the amount of muscle tissue.

Instead, imaging techniques are usually used mores widely because they offer better accuracy and resolution, which implicitly lead to a more accurate diagnosis. The following techniques are the landmarks for diagnosis of sarcopenia through imaging: Dual Energy X-Ray Absorptiometry (DEXA), computer tomography (CT) and Magnetic resonance imaging (MRI), echography [8].

1.3.1 Non-imaging techniques

BIA is currently recognized as a routine test to study body composition. It is a non-invasive technique, cost effective and it is also a reliable method for analysing BC in clinical routine. BC has a relevant impact on the transit of time of an electric current through the tissues. This physical aspect can be used to measure the electrical characteristics of the tissues and after that estimate different parameters as total body water (TBW) and fat-free mass (FFM). The biggest limitation of this technique is the variability in chemical composition of FFM, which depends on different aspects like age, sex and disease.

1.3.2 Imaging techniques

Dexa

It is based on the principle of differential attenuation of an X-ray beam, at two energy levels, when passing through the tissues. This attenuation is recordable and related to the body composition of the subject being examined. The device uses a matching X-ray beam with no dispersion in the environment.

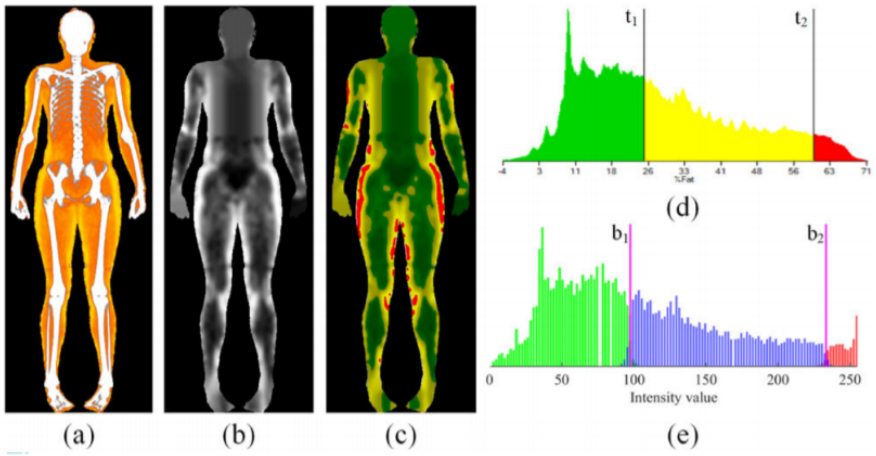


Figure 1.5: Calculating BC with a DEXA image. (a): bone-fat-lean tissue mass map. (b): pixel-level fat percentage map. (c): fat percentage threshold map. In (d) t1 = 25% and t2 = 60% are the threshold for DXA system by default. (e): intensity thresholds [14].

Compared to other imaging techniques, DEXA is inexpensive and has a shorter scan times and also radiation exposure. The device is made by a C-arm containing the x-rays source (below the patient in supine position) and there are two different energy levels that are specific for soft tissue and for cortical bone. Moreover, thanks to a collimator between the patient and the source, scatter is minimized. The two different energy level (a low one and a high energy photon emission) are detected and combined, creating a planar image. Figure 1.5 illustrates an example of a measurement of BC throw DEXA imaging [14].

Importance of L3 slice

To study body composition theoretically it is necessary to segment muscle tissue in the whole body 3D image to obtain the extend of it. However, studies have shown that muscle area at third lumbar vertebra (L3) is correlated to

whole body muscle mass [15] [16]. This parameter can be obtained with structural imaging techniques like MRI or CT, which provide cross sectional images to study the L3 level. Figure 1.6 illustrates an example of the whole spinal cord, underlining L3 position and explaining some differences between thoracic vertebrae and lumbar ones. For example, usually thoracic vertebras have long, sharp spinous processes and articular facets for ribs, while lumbar vertebras have shorter spinous processes and larger body [17]. To measure the patient's BC, radiologists needs at least 5 minutes [18] because this process is made by two steps: first of all, the radiologist has to spot the slice corresponding to L3 vertebra and then, to segment this image manually. Moreover, it is important to remember that these two steps are very operator-dependent, which causes a non-significant variability in this process (radiologists are prone to mistakes). Despite the variety of techniques, MRI and CT are still considered the standard techniques for body composition evaluation, since the cross-sectional images can be segmented and used to measure muscle and fat mass [19].

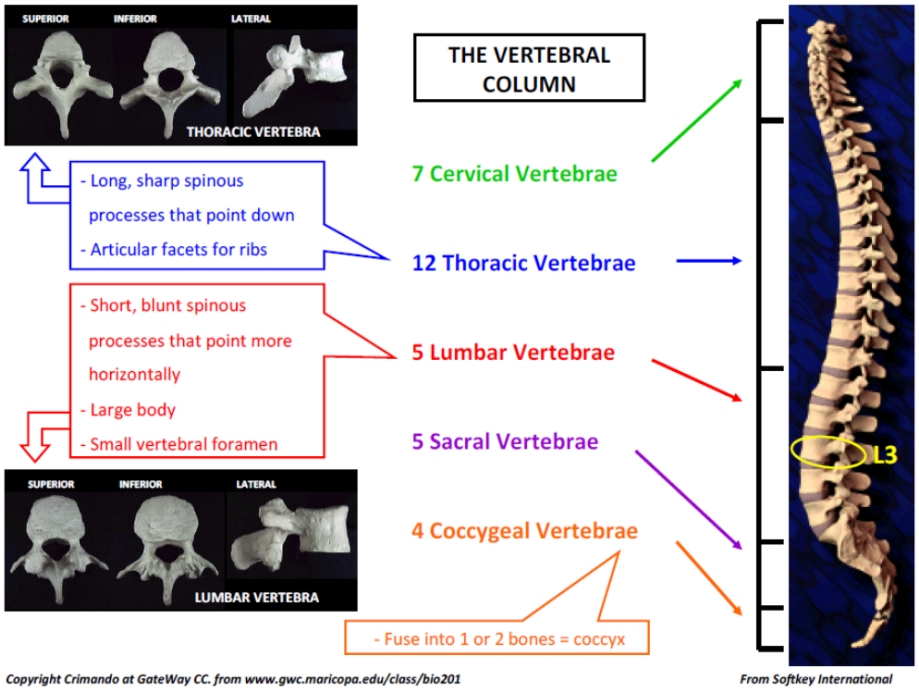


Figure 1.6: Main differences and locations of human vertebras [17].

MRI

MRI uses magnetic fields and radio waves to produce thin-layer tissue images (tomographic images). Normally, protons inside the tissues rotate, generating small magnetic fields that are randomly aligned. When they are surrounded by the strong magnetic field of the MRI, their magnetic axis aligns along that

field. Subsequently, the application of a radio frequency pulse causes the axis of many protons to align momentarily in the opposite direction of the field, in a condition of high energy. After the pulse, the protons relax and resume their original alignment in the magnetic field of the MRI. The magnitude and speed of the energy release that occurs with the return to basal alignment of protons (relaxation T1) and by their oscillation (precession) during the release of energy (relaxation T2) are recorded as localized signal strength spatially by a coil (antenna) present inside the apparatus RM. Computer algorithms analyse these signals and produce detailed anatomical images. By controlling radio frequency pulses and gradient oscillations, some computer programs produce specific pulse sequences that determine how the image is obtained (weighed) and how different tissues appear [20]. The MRI images can be:

- T1 weighted
- T2 weighted
- Proton density weighted

This technique manages to generate high resolution images and thanks to the differences between the molecular properties in different tissues it is possible to have a good representation of different soft tissue like water, fat and muscle. Moreover, it is possible to evaluate muscle structure by using different acquisition protocols. For example, with DIXON MRI it is possible to obtain water-fat separation and consequently reach a more precise measurements of muscle volumes and the percentage of fat infiltration [21].

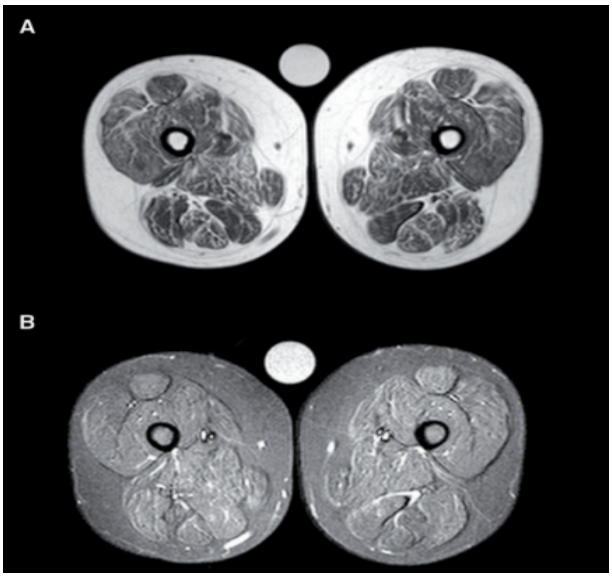


Figure 1.7: Example of two images at the same level: (a): T1 weighted MRI and it is possible to see in a better way soft tissue and adipose tissue. (b): T2 weighetd MRI image in which liquids and pathological conditions (cancer, inflammation, trauma) are seen in a better way [22].

As Figure 1.7 shows, there are two different MRI images: in the one on the left there is an example of T1 weighted MRI, hypointense image, whereas in the right there is an example of T2 weighted MRI image in which liquids are emphasized [22]. Moreover, with more advanced techniques there is also the possibility to detect fatty infiltration and percentage of intracellular fat. This can be an important aspect especially in cancer conditions, where usually this value tends to increase.

MRI does not produce side effects (there is no ion dose like in CT exams), and for all these reasons it is considered a reliable method for studying BC. The limits of MRI are high costs and limited availability, which are the reasons for which this technique is mainly limited to research .

CT

CT is a computerized X-ray imaging technique in which X-rays are beamed at a patient and rotated around the body very quickly, so that it is possible to obtain an image of a slice of the patient's body. These slices are named tomographic images and, compared to a simple X-ray test, contain much more detailed information, also because thanks to the creation of these slices it is possible to obtain a 3D image of the patient, which can be very useful for the identification of different problems like cancer or abnormalities. As said, CT, like MRI, provides cross-sectional images and, thanks to their different characteristics, it is possible to discriminate between different kinds of tissue.

This technique is based on different X-ray attenuations, which is quantitively measured in Hounsfield units (HU). The Hounsfield unit is a standard unidimensional form to express CT numbers. Reference values are 0 HU for water and -1000 for air. For muscle density, normal attenuation values are in the range between 40 to 100 HU but fatty infiltration can modify these in a huge way bringing them to negative ones (from -200 to -35 HU). Using those values it is possible, through post-processing analysis, to measure the extend of different tissues, like muscle, visceral adipose tissue or subcutaneous adipose tissue.

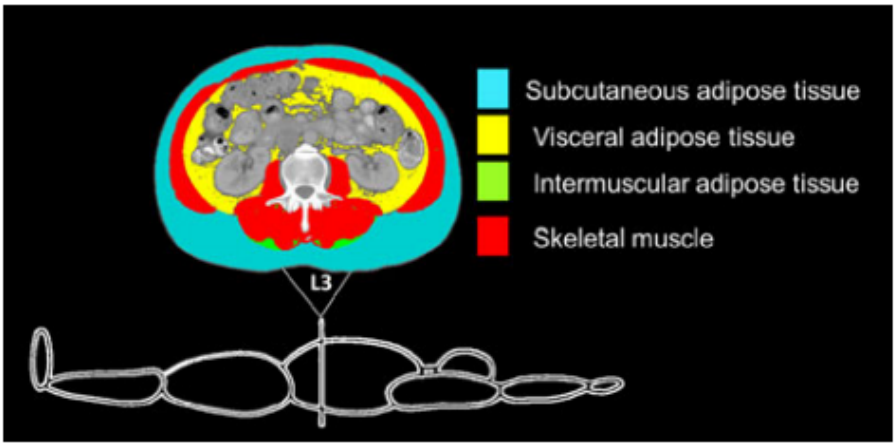


Figure 1.8: Example of different tissues at L3 level [6].

HU density thresholds can be used for muscle segmentation after removing visceral organs: in Figure 1.8 it is possible to see how images are chosen at a specific axial slice (usually L3 vertebra) and then having a muscle segmentation using a software, which includes HU values between -20 and 150.

However, this technique presents an important limitation: the radiation exposure must be considered also because each CT scan consist in 1 to 10 mSv, varying on the dose of radiation and the target. There is also a low-dose CT scan that are about 1.5 mSv, but, in any case, side effects due to X-rays must be considered [23].

Table 1 shows the main differences between some of the discussed techniques. As it is possible to notice from the table, all the techniques discussed are validated and the election between one of them depends on the specific patient and on the corresponding specialist.

	MRI	CT	DEXA	BIA
Costs	Very high	Very high	High	Relatively inexpensive
Sensitivity	Very high	Very high	High	Medium
Specificity	Very high	Very high	High	Medium
Personnel	Highly specialized	Highly specialized	Specialized	Non-specialized
Portable	No	No	No	Yes
Radiation	No	Considerable	Little	No
Time	15–20 min	15–20 min	15 min	5 min
Indication in diagnosis of sarcopenia	No, for research purposes	No, for research purposes	No, for research purposes. Indicated to assess BMD	Yes, fat mass, fat-free mass
Validated	Yes	Yes	Yes	Yes

Table 1: Main characteristics of MRI, CT, DEXA and BIA [5].

In order to provide a good spatial resolution and distribution of muscle and adipose tissue the best techniques are MRI and CT, with DEXA it is not possible to obtain the same results [5]. It is important to underline the

differences between MRI and CT, so that to have the correct information to understand when one is better to use than the other one. In particular, with MRI it is possible to obtain better soft tissue contrast in respect to CT but is a very low-cost test; on the other side, CT is much more accessible, cheaper and is faster.

1.4 Treatments

Currently, there is no approved medical treatment that can prevent sarcopenia. For several decades, however, there has been a number of pharmacological studies aimed at understanding whether the use of hormones, such as testosterone and growth hormone, allows the maintenance of muscle mass and strength, in spite of aging [24] [25].

While it is true that no medical treatment is capable of avoiding sarcopenia, it is also true that there are several natural remedies that can counteract the physiological decline of muscle mass and strength related to age and prevent the worst consequences [24]. These natural countermeasures consist in constant exercise and in a diet in line with the needs of muscle tissue and the human body in general, in old age.

As said, one possible treatment can be the change of diet habits. At the base of the ideal diet to fight sarcopenia there is the consumption of [26]:

- Foods rich in healthy protein. This is the most important point; foods rich in healthy proteins include: fish (e.g. trout or salmon), crustaceans, nuts, lentils, quinoa, beans, tofu, lean parts of poultry and lean cuts of beef
- 3 to 5 servings of fruit and vegetables
- To validate the implemented algorithm
- Foods low in sodium, fat and/or sugar. The advent of patient-centred care has increased attention to the fact that different molecular changes can result in the need to have different therapeutic approaches to similar conditions such as sarcopenia. Variety of molecular changes resulting in changes in myo-fibre metabolism and alterations in satellite cell properties

The practice of physical exercises has been shown to improve muscle mass and strength, but is not always feasible in elderly subjects, and it is not yet known how long its effects last after it is discontinued. Moreover, it has also been demonstrated that this kind of treatment it is not enough to contrast the loss of muscle mass in elderly people, for which it is necessary to add drug therapies [26].

Pharmacological treatment is a major area for research. A variety of drugs are being tested for their effects on muscle mass and strength, such as Ghrelin,

GH secretagogue and myostatin inhibitor, but at present few results have been achieved, some of which are controversial [26].

Testosterone is one of the most important steroid hormones and it stimulates development of secondary sexual characteristics in men, like muscle growth. The results of the different studies are not consistent. The administration of testosterone to young people is associated with an increase in muscle mass, but not in strength. In the elderly, high testosterone doses increase contraction force, but this treatment has been associated with severe complications, contraindicating its use. Growth hormone (GH) supplementation has created a multi-million market in anti-aging medicine programs. Administration of GH improves body composition by increasing muscle mass, reducing fat mass, and decreasing bone demineralization rate, but there is strong evidence that it does not improve contraction or functional capacity and does not induce positive metabolic changes. Moreover, long-term testosterone therapy could be also associated with side effects that can be physical (like prostate hypertrophy and high blood pressure) or behavioural (especially aggression) [27].

Monitoring vitamin D can be also useful in order to enhance problems with muscle loss and leak of strength. 25(OH)-vitamin D (vitD) quantity decreases with age. Many studies show that in elderly people there are extremely low vitD values. Low vitD is correlated with loss of muscle strength, this is why supplement of vitD produce a functional improvement. Many studies have reported extremely low vitD levels in elderly people. Without entering too much in details, normal values of vitD are considered those above 30 ng/mL, whereas there is vidD deficiency with values under 12 ng/mL. Usually cancer patients have vitD deficiency and dosing of vitamin D can help in not losing too much strength, as well as it can regulate the entire process of tumorigenesis, from initiation to metastasis process [28].

1.5 Context

The present Master's thesis is carried out in the Biomedical Engineering and Telemedicine Centre of the Escuela Técnica Superior de ingenieros de Telecomunicación, Centre for Biomedical Technology of the Universidad Politécnica de Madrid. Specifically, it is conducted in the "Medical images, surgical training and image guided surgery" research laboratory.

1.6 Chapter structure

This project of thesis contains 7 chapters, starting from the general problem and finishing with the future possibilities in this field. In particular, the first and current chapter is a summary of the clinical contest, giving a brief but intense idea of how important muscle loss is due to diseases like sarcopenia and cachexia in cancer patients. Moreover, this chapter contains also the reasons for which medical images are the best (thanks for MRI and CT cross-section

images) for analysing this loss. The second chapter contains the part regarding the objectives to achieve during this project. The third chapter is about the state of the art. After that, the fourth chapter presents the materials, i.e. the databases used for the training set and test set and the software used to write the algorithm. Chapter five contains the results obtained the validation of the algorithm. Chapter six and seven, finally, are related to the discussion and conclusion parts, respectively. After that there is the bibliography in which there are present all the articles from where the information where collected and used for the realization of this Master thesis.

2 Objectives

Time and interoperator variability are the two main factors that show the benefits of an automatic way to reach this goal is needed. Thus, the aim of this project is to develop an automatic algorithm to detect L3 level in whole body 3D CT scans. To perform this study, the following secondary objectives are defined:

- To define pre-processing steps to account for possible intensity-related differences and artifacts in different CT scans.
- To design and to develop the automatic L3 slice detection algorithm.
- To validate the implemented algorithm with a retrospective study cohort composed of CT scans acquired on the same scanner and same acquisition protocol.

3 State of art

As described in 1.3.2, CT is one of the gold standard methods to evaluate images and, thanks to its characteristics, it is possible to work with cross-sectional images (2D) or with 3D images. Normally, whole body CT scans are acquired and the L3 slice is selected manually. However, only two recent studies have been found that describe automatic L3 slice selection algorithms. There are different kind of possibilities to find L3 slice in a full CT scan, which usually includes more than 300 slices (a whole CT scan can contain also more than 700 images). In particular, in this subsection two techniques are presented that allow to find L3 slice in a different way.

These two techniques have the same characteristic: they are based on networks that need training. This is one of the challenges of this master thesis, trying to find a different method which does not need training part, saving computational time and power remaining aware of the fact that usually a network architecture can give better results.

3.1 Definition of dataset and pre-processing

DICOM[1] is the standard and universal format for the storage of CT images. Usually the images have a resolution of 512 by 512 pixels, which corresponds to a total of 264144 pixels per 2D image slice. To obtain better results and to make it easier to work with these images, pre-processing steps can be added to the image analysis methodology. Standard pre-processing steps can be, for example, the conversion of the DICOM files into other type of images (like png or jpeg), the normalization of the pixel values and the transformation of the images from RGB to grey-scale format.

[1]Digital Imaging and Communications in Medicine — international standard for medical images and related information. It defines the formats for medical images, with the data and quality necessary for clinical use. This format is commonly used in almost every cardiology, radiology imaging and radiotherapy device (X-ray, CT, MRI, ultrasound, etc.), and nowadays it starts to be used in different medical domains such as ophthalmology and dentistry. DICOM® is one of the most widely organised healthcare messaging Standards in the world. It is so used that it is possible to say that in this moment there are billions of DICOM® files in use for medical problems [29].

3.2 Spotting L3 slice using deep learning algorithm

In 2017, Soufiane Belharbi et al. [30] presented a whole automatic system for choosing a defined slide in a full 3D CT scan is illustrated. This study mainly relies on a machine learning regression method. It is possible to define 3 main steps:

- CT images converted into Maximum Intensity Projection (MIP) images
- Prediction of a window using a Convolutional Neural Network (CNN)
- Analysis of the prediction structure in order to predict the slice corresponding to L3 in the CT scan

3.2.1 Dataset and pre-processing

This approach is applied to the spot of the L3 vertebra and the starting database contained 642 CT scans, all from different patients. All the CT scans come from author's clinical centre and there is a heterogeneity between the patients in terms of age, sex, acquisition protocols (low dose acquisition (100–120 kV) and slice thickness (2-5 mm). the dataset is split into 5 folds, and this operation allows a cross-validation procedure. The split, applied at the patient level, avoid that a given CT scan specify windows in different sets.

3.2.2 Model Architecture

It is important to underline how using this kind of approach requires a huge amount of memory, as well as computational resources, especially for the training part, which is required in order to use a CNN architecture. In order to avoid these problems, the approach described in Figure 3.1 was adopted: first of all, the CT scan is converted into MIP images. This operation allows to convert a 3D input into a 2D input, drastically decreasing the computational and memory stores and avoiding an important leak of information.

After that, the MIP image is treated to create a sliding window which will be the input of a CNN, with fixed dimensions. Transfer learning[2] (TL) was used to train this CNN, creating a TL-CNN, which computes its predictions for every location of the sliding window, creating a prediction sequence. Finally, this sequence is processed and L3 position is estimated in the whole CT scan.

[2]Transfer learning: machine learning method where a model extracted for a task is reprocessed as the starting point for a model on a second task. It is a popular approach in deep learning: pre-trained models are used as the starting point on computer vision and natural language processing tasks [31].

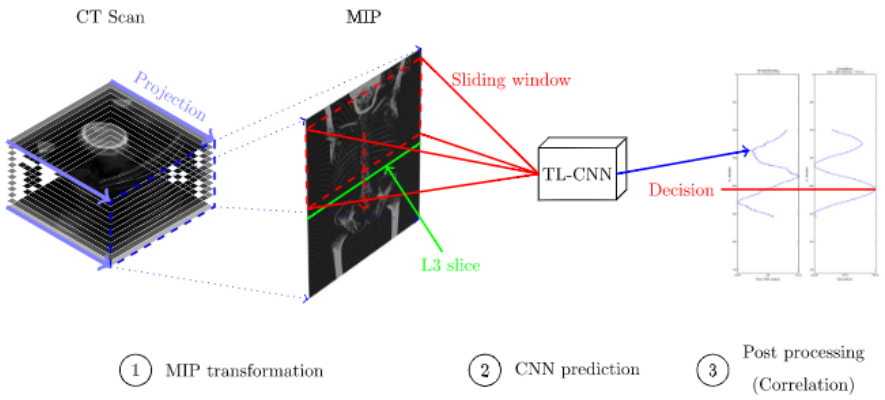


Figure 3.1: Illustration of the main steps used to estimate L3 slice [30].

3.2.3 Data processing

The main goal of the first step (MIP transformation) is to reduce the computational and storage weight in order to have a faster algorithm even losing some information. In this case, producing a MIP image brings to frontal view in which patient's skeleton is perfectly seen and the authors assume that this kind of image gives enough physical information so as to detect L3 slice. It was also tried to choose other views (as well as combination of different views) but no one of them could produce the same results as the one related to the frontal view alone.

The initial input size, for DICOM images with size 512 x 512, is given by 512 x 512 x N, where N is the number of slices. So, for example, if the whole CT scan contains 1000 slices, the final result would be 262 million inputs. Thanks to the conversion into MIP images the input size is drastically reduced to 512 x N, which corresponds to only 512 thousand inputs. In Figure 3.2 there is an example of detection of L3 slice for 3 different patients.

CNN is one of the neural network architectures where a convolutional layer, which performs a non-linear filtering process, is the main building block. This convolution could be seen as a feature extraction in which the values of the convolution kernel are the layer parameters. It is important to notice how every single layer can give a different set of features from the previous layer.

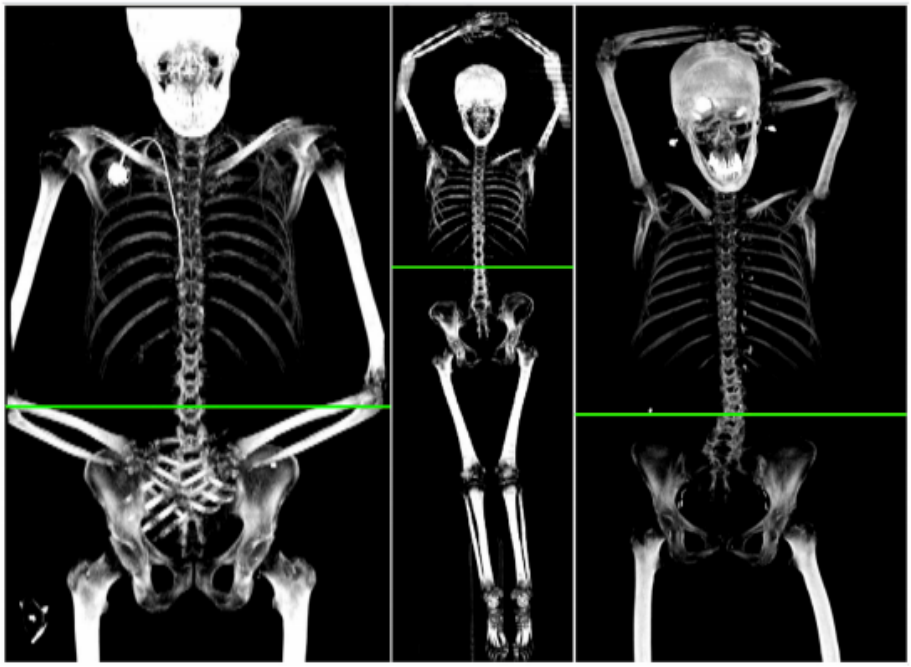


Figure 3.2: Example of normalized MIP images with L3 position highlighted by the green line [30].

During the last years, transfer learning methods are more and more used, also because these methods allow to overcome the loss of training data [30]. TS is based on adapting models, trained for several goals. The initial weights are set with weights taken from a different pre-trained CNN, and then adjusted in order to achieve the target.

3.2.4 Decision process

Sampling windows over a MIP image can produce two different sets of windows: one containing the L3 vertebra and one without the L3 vertebra. The authors propose a regression method, where adding the window without L3 to the dataset means that the CNN learns the offset of the L3. This offset is, unfortunately, very difficult to learn so that the decision is not include the window without L3 but only the one containing the target vertebra.

The decision process is made by the sliding window over the MIP images, which give as result the relative L3 position inside the window. The real position of the L3 slice should decrease one by one (in CNN outputs), as shown in Figure 3.3. CNN output should evolve linearly along the sequence of windows, leading to a noisy straight line with a slope of -1 (green segment represents the

theoretical slope). The L3 position can then be estimated as the central position of that segment.

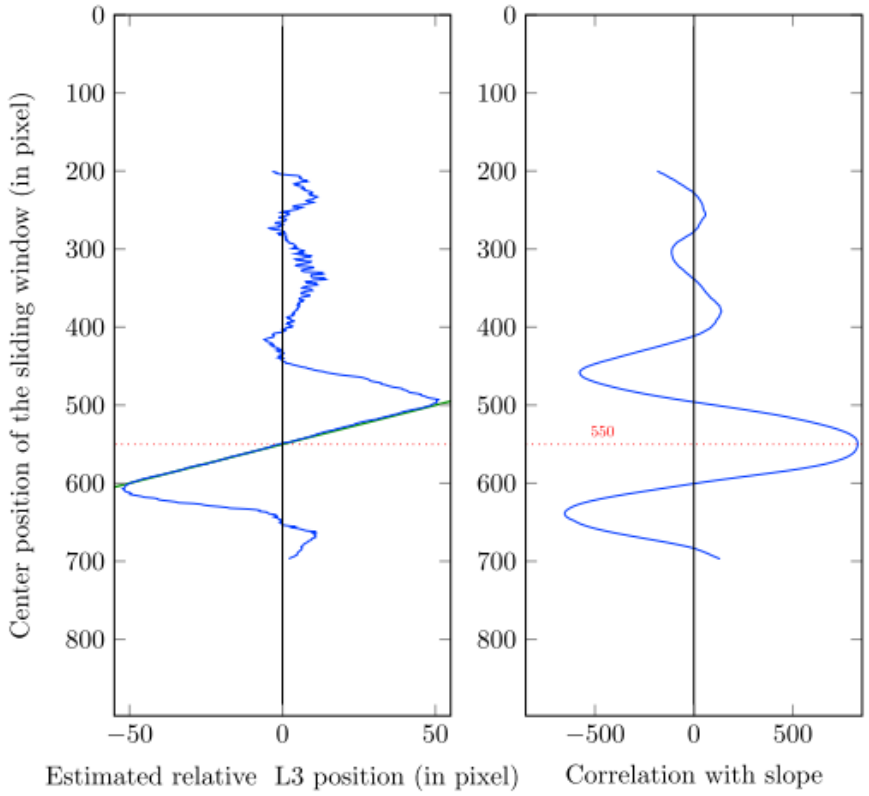


Figure 3.3: Example of CNN output sequence with the typical straight like of slope (theoretical one in green. In the graphyc of the right there is illustrated the correlation between CNN output and slope [30].

3.2.5 Results

In Table 2 it is possible to compare the errors given by CNN and the errors done by three different radiologists. In this case, it is important to underline how radiologists usually are more precise than automatic methods, but these results show also that there is some variabilities among different radiologists whereas the automatic methods give the same output. This prediction shows robustness by averaging several predictions.

Errors (slices)/ operator	CNN4	VGG16	Radiologist #1	Radiologist #2	Radiologist #3
Review1	2.37±2.30	1.70±1.65	0.81± 0.87	0.72±1.51	0.51±0.62
Review2	2.53±2.27	1.58±1.83	0.77±0.68	0.95±1.61	0.86±1.30

Table 2: Comparison between errors given by CNN and by three different radiologists [30].

3.3 Automatic L3 slice detection in 3D CT images using fully-convolutional networks (L3 UNet-1D)

In 2018, Kanavati Fahdi et al. [18] took inspiration from [30], trying to enhance further the performances of the previous study. In particular, there are two methods described in the article: the first one, as the previous one, converts a 3D CT scan into MIP image which is then used as input to a 2D convolutional network (in order to predict L3 slice position); on the other hand, the second method is a modified architecture which works with 1D output size. The second method, compared to the previous study, is faster thanks to the fact that there is one dimension less, and the results show the same prediction accuracy [18].

3.3.1 Dataset and pre-processing

The authors collected a dataset consisting of 1070 CT images, from different sources like Cancer Imaging Archive (TCIA) or Hammersmith Hospital (HH), London. Since the final dataset is a combination of different datasets, Figure 3.4 illustrates two parameters: 1) the distribution of the slice thicknesses and 2) the image heights, while Table 3 contains the errors between annotators A and B.

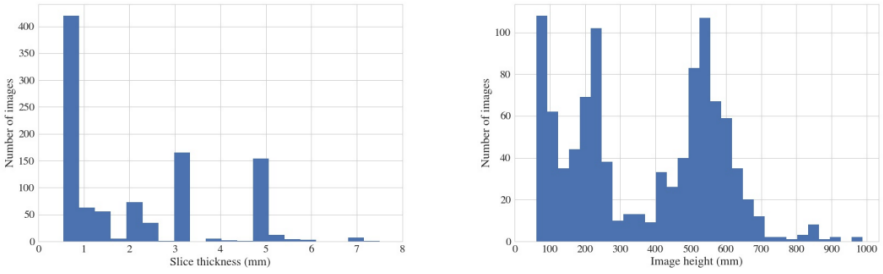
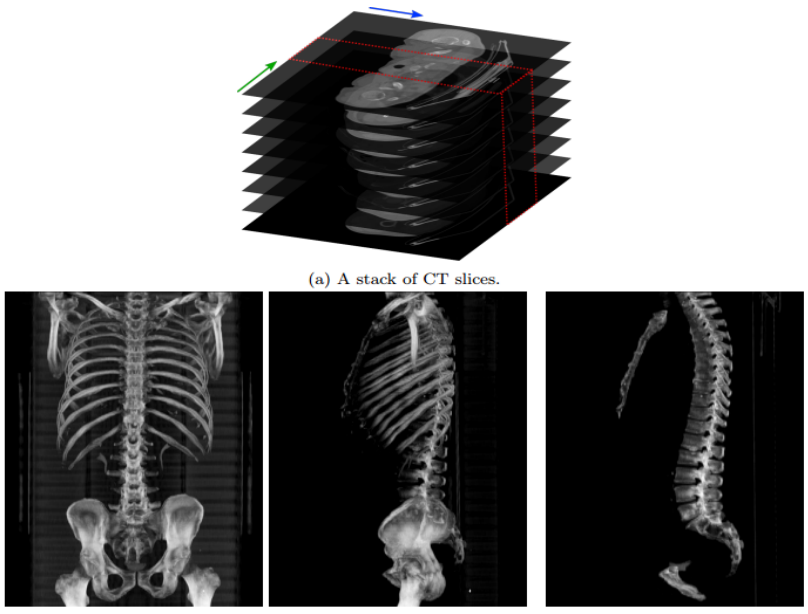


Figure 3.4: Slice thickness distribution and image heights of the dataset [18].

	mean	std	median	max
Error between A and B (mm)	1.90	1.76	1.00	9.00
Error between A and B (slice)	1.94	2.36	0.80	11.43
Error between A/B and mean of A and B (mm)	1.14	0.97	1.00	5.00
Error between A/B and mean of mean A and B (slice)	0.97	1.18	0.40	5.71

Table 3: Errors between annotators A and B, reported in mm and in number of slices. The error in slices was calculated by dividing the error in mm by the slice thickness (without rounding) [18].

Regarding pre-processing, as in the previous study the first step is a conversion of the 3D CT scan into MIP images, which returns an image of the sagittal plane. Differently from the method described in [30], the authors of this article choose to compute a smaller sagittal view, eliminating the outer edges of the pelvis and having in this way a better view of the sacrum vertebra. Sacrum vertebras position, if a bottom-up approach is used to find L3, is essential. It was considered that, in the majority of the patients, the spinal column is central in the image so MIP image was restricted to the range [-20, 20] starting from the centre of the MIP image. Moreover, considering that every CT slice has a different thickness, a pixel size normalization is adopted to have a reliable input to the algorithm. The last step consists in the elimination of majority of soft tissues: this is obtained thresholding the images in the range 100 – 1500 HU. Figure 3.5 shows a representation of a restricted MIP image.



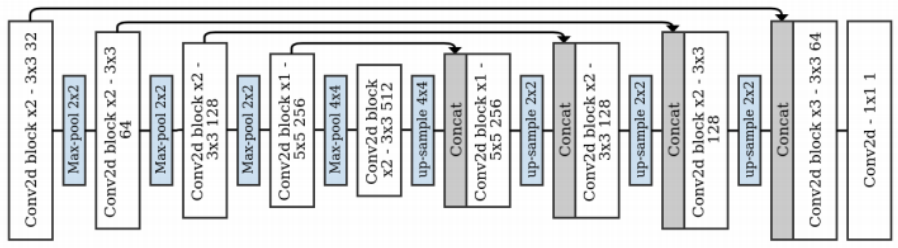
(a) A stack of CT slices.

Figure 3.5: Examples of extractions of different MIP images from a stack of CT slices. From the starting stack of CT slices, it is possible to extract a frontal MIP image (on the left), a sagittal MIP image (on the center) and a restricted sagittal MIP image (on the right) [18].

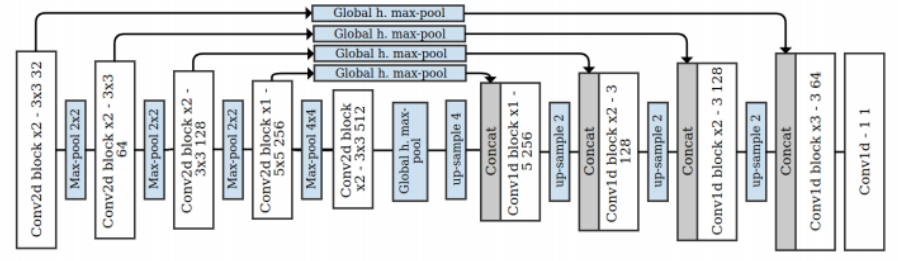
3.3.2 Model Architecture

The 2D architecture consists of a several down-sampling and up-sampling blocks, this is the reason it is called UNet-like architecture (with the latter copying the former). Each block is made by one or two convolutional units and every unit is a 3x3 convolution. The concatenation between the output of the up-sampling blocks and the outputs of the down-sampling blocks is called skip connection, used to associate the down-sampling path to the up-sampling path. The result is a 2D confidence map prediction output, which has the same dimensions of the input image. This network has 8,493,537 parameters and it is named L3UNet-2D.

Similarly, in the 1D version the FCNN is based on the UNet structure. In this case, the down-sampling path remains the same as the previous one but there are important differences in the part regarding the application global horizontal max-pooling along the up-sampling path. Here the result consists into a 1D convolutions for the up-sampling, obtaining a 1D output with same size as the height of the input image. The network in this case has 6,189,025 parameters and it is named L3UNet-1D. The authors use a fully convolutional network (FCNN) based on UNet architecture[3] with two different outputs: a 2D confidence map and a 1D confidence map output. Figure 3.6 shows the two network architectures.



(a) L3UNet-2D



(b) L3UNet-1D

Figure 3.6: Network architecture of the 1D and 2D versions [18].

[3]Unet is one of the existing network architectures, used for semantic segmentation. It is made by contrasting path and expansive path [32].

To improve generalisation execution, an augmentation approach is used. In order to achieve this, some transformations were applied to the images to simulate variants. The transformation applied are:

- Horizontal flipping
- Intensity offset
- Vertical image sub-sampling
- Scaling
- Piece-wise affine deformations

Point-wise annotations are converted into confidence maps for each MIP image. Position of L3 slice along the y-axis is the only accessible annotation. There is one assumption: the spine is mostly located in a fixed x-axis range. There is a generation of the ground truth confidence maps for the frontal MIP images. Defining y_i as the ground truth coordinates of L3 slice for image i, the value of the confidence map is expressed as:

$$\mathbf{H}_i(x, y) = A \times (f_i * g_\sigma)(x, y),$$

Where:

- f_i is the step function below:

$$f_i(x, y) = \begin{cases} 1 & x_0 - v \leq x \leq x_0 + v, y = y_i \\ 0 & \text{otherwise,} \end{cases}$$

- g_σ is a gaussian filter function
- ν is an offset
- A is the maximum normalization $||Hi(x, y)||_\infty$

In the 1D architecture fi is simply an indicator function:

$$f_i(y) = \begin{cases} 1 & y = y_i \\ 0 & \text{otherwise.} \end{cases}$$

Figure 3.7 illustrates example images with ground-truth confidence maps covering the images.

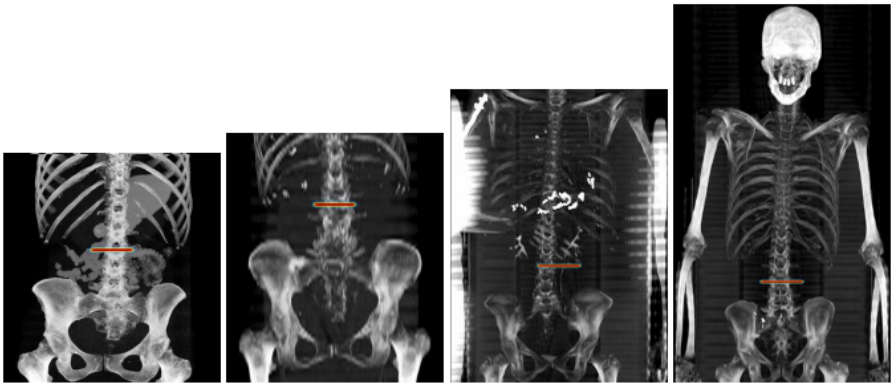


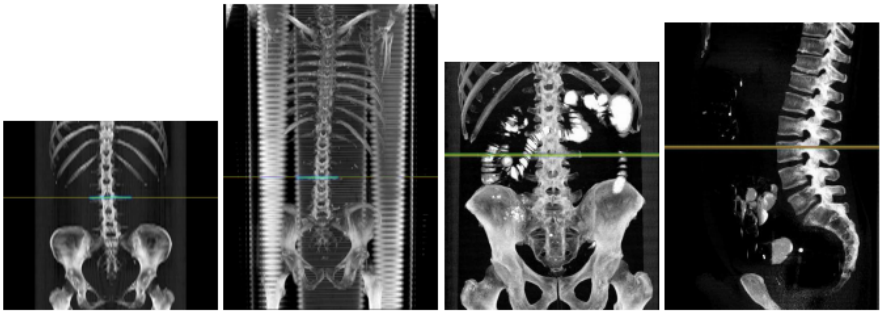
Figure 3.7: Different images with ground-truth confidence map (in red) [18].

3.3.3 Results

Performance obtained are comparable to the ones obtained in [30], with a median error of 1mm (almost equal to the human error) using the L3UNet models. Moreover, 1D architecture gives slightly better results than the 2D architecture. The slice location along the y-axis is the only accessible annotation, it is easy to apply the 1D architecture both to the frontal MIP and to the sagittal MIP. In Table 4 there are error values for 2D frontal architecture and for 1D architecture applied to both the frontal and the sagittal view and Figure 3.8 shows the different views.

		mean	std	median	max	> 10
L3UNet-2D - frontal	error (mm)	2.09	4.56	1.00	51.00	19
	error (slice)	1.43	3.52	0.80	51.00	12
L3UNet-1D - frontal	error (mm)	2.12	4.56	1.00	38.00	22
	error (slice)	1.53	4.22	0.67	45.71	15
L3UNet-1D - sagittal	error (mm)	1.99	5.41	1.00	52.00	28
	error (slice)	1.41	5.02	**0.50**	65.00	23
VGG16 Regression [BCH+17] - frontal	error (mm)	13.78	8.57	12.00	48.00	591
	error (slice)	10.26	9.92	6.12	60.00	360
VGG16 Regression with dual output - frontal	error (mm)	6.94	5.90	6.00	62.00	191
	error (slice)	5.54	6.29	3.20	40.00	180

Table 4: Cross validation results for L3UNet-2D and L3UNet-1D (both frontal and sagittal view). The last column refers to the number of outliers with error > 10 [18].



(a) L3UNet-2D frontal (b) L3UNet-2D frontal (c) L3UNet-1D frontal (d) L3UNet-1D sagittal

Figure 3.8: Prediction outputs: four different predicted confidence maps. (a) and (b) refer to a L3UNet-2D frontal view, (c) refers to a L3UNet-1D frontal view and (d) refers to a L3UNet-2D sagittal view [18].

In a small group of images, the prediction is incorrect and the difference to the L3 slice is usually one vertebra. Figure 3.9 illustrates some examples of outliers, one referring to an error occurred in 2D architecture and one referring to the 1D architecture [18]. In conclusion, it is possible to say that this method works, and the results are comparable to the ones obtained in [30].

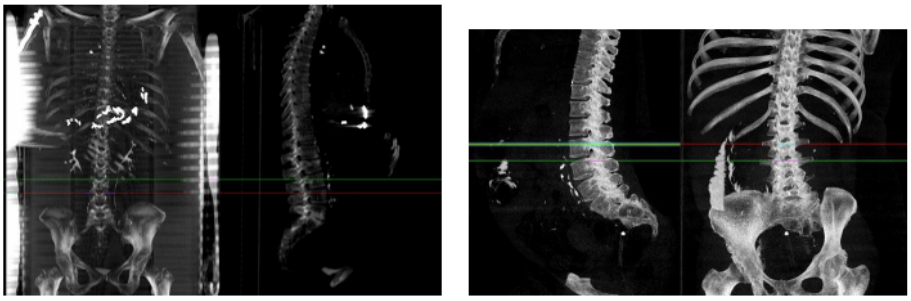


Figure 3.9: Outlier cases: incorrect identification of L3 vertebra by L3UNet-2D (on the left) and by L3UNet-1D (on the right) [18].

3.4 Automatic Vertebra Annotator

Voronoi Health Analytics is a medical software company that develops medical image analysis solutions for clinics, hospitals and research institutes. Among the several diagnostic and clinical support tools created by this company, a new software is under development which analyses and extracts patient's BC after predicting vertebrae positions in a whole body CT scan. The Automatic Vertebra Annotator (AVA) is an application to quickly select and extract individual

slices, starting at the first thoracic vertebra until the sacrum [33]. Figure 3.10 illustrates an example of the AVA interface.

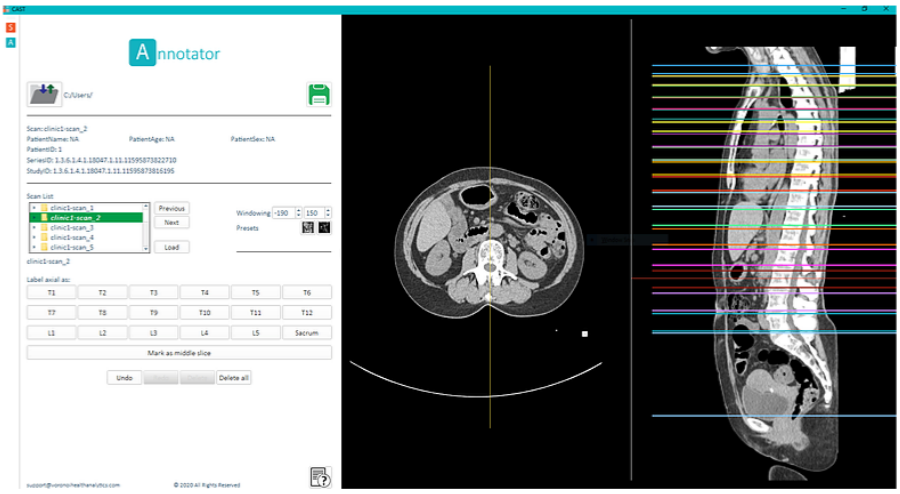


Figure 3.10: Example of AVA interface [33].

4 Materials and methods

4.1 Materials

4.1.1 Database

The database used for this project is composed by 24 patients treated for Hodgkin and non-Hodgkin lymphoma. 3 or 4 whole body CT scans were acquired for each patient, taken during different months in order to monitor the clinical process. This database is provided by the Department of Nuclear Medicine of the *Hospital Universitario 12 de Octubre*, Madrid, and every CT scan is composed by approximately 300-700 DICOM images. A total of 80 CT scans are available from 24 patients However, due to the time gap between image acquisitions of the same patient, in this project, each image, and more importantly L3 slice position, is considered sufficiently different for the validation process.

4.1.2 Software

Nowadays the panorama of coding software offers several possibilities: for this project, the choice fell on Python, which is a high-level programming software. This software is simple, it is easy to learn how to code in Python and, in addition, supports different modules and packages. Python needs an interpreter and, for this project, Anaconda is used, in particular Spyder. As said, Python supports different libraries and packages and, for this project, the used packages are the following:

- The Pydicom[4] package is a group of functions designed to manipulate DICOM files using Python code. This package is needed for loading the DICOM files.
- The Skimage[5] package contains several image processing routines like filtering or creating binary masks.
- Matplotlib[6] is a Python library through which different visualizations can be created. This plotting library is useful to visualize data.

[4]https://pydicom.github.io/ [Last access: Jenuary, 02 2021]
[5]https://scikit-image.org/docs/0.13.x/overview.html [Last access: January, 02 2021]
[6]https://matplotlib.org/contents.html [Last access: January, 02 2021]

- Scipy[7] is a library, which collects different mathematical algorithms and functions. These provide level commands and classes useful to manipulate and to visualize data.
- Numpy[8] access to a multidimensional array objects and many functions to obtain fast operations on arrays, like mathematical or logical operations, like shape manipulation, random simulation and a lot of other functionalities.
- OpenCV[9] (CV2) is a library which has important tools regarding computer vision and machine learning.
- The Os[10] package is principally used for reading and writing files. Through Os it is possible to create temporary files and directories.
- The time[11] package provides several time-related functions and in this project it is used in order to calculate the computational time of the algorithm.

4.2 Methods

As explained, the goal of the project is to find the slice corresponding to L3 vertebra in a whole CT scan, using an automatic algorithm. The following steps are the summary of the adopted approach used to reach that goal: first of all there is a data preparation part, in which DICOM files are loaded; then some pre-processing is done in order to increase the homogeneity of the different CT scans; after that, there is the prediction of L3 slice; finally, a validation of the algorithm is executed. Figure 4.1 illustrates the pipeline of the project:

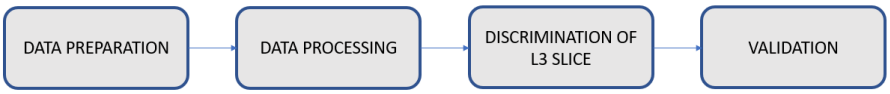


Figure 4.1: Pipeline of the methods adopted in this study

4.2.1 Data preparation

The first step is to load every CT scan and one by one the slices are inserted into an array. The dimensions are 512x512xN, where N is the number of slices inside a CT scan and, in this database, there are CT scans with different number of slices (from 300 to more than 700 slices).

[7]https://www.scipy.org/docs.html [Last access: January, 02 2021]
[8]https://numpy.org/devdocs/user/whatisnumpy.html [Last access: January, 02 2021]
[9]https://opencv-python-tutroals.readthedocs.io/en/latest/ [Last access: January, 02 2021]
[10]https://docs.python.org/3/library/os.html [Last access: January, 02 2021]
[11]https://docs.python.org/3/library/time.html Last access: January, 02 2021]

The arrays are created to be able to use cycling operations. Once the CT slices are loaded, the algorithm determines if the sliding sequence is bottom-up or top-down. This is a very important aspect because the script is written for top-down sequences, not for bottom-up ones. After determining if the direction of the slices is a top-down one or not, if the answer is positive the array containing the CT slices remains the same, otherwise it is simply inverted.

Having a homogenous dataset helps in the project, and since the CT scans are taken from different patients the first thing to do in order to enhance the homogeneity is an image intensity normalization. In this project, image normalization is done simply by dividing each pixel by the maximum value in the image, in order to have images where 1 corresponds to white colour and 0 corresponds to black.

4.2.2 Data processing

At this point, with a matrix composed by N normalized images, binary masks are created in order to isolate as much as possible the pixels corresponding to bone. Bone presents high intensity values (226-3071 HU) in CT scans [34] and, as result, is hyperintense in the images. But it is also important to consider attenuations and other side aspects that could still appear in masks, so a relatively high threshold is chosen. This decision is taken in order to account for intratissue variations of pixel intensity. After these considerations, a global bone segmentation threshold of normalized images is used, so that pixel values lie in the range [0,1]. Lastly, after threshold segmentation and mask creation, two more operations based on region connectivity are performed to clean the masks:

- Deletion of the element with the maximum axis length, which corresponds to the patient bed.
- Deletion of small elements that do not have a certain area to avoid random white pixels in the masks.

The idea of segmenting only bone areas arises from some considerations regarding human anatomy. Starting from the head area and sliding down through the spinal cord, first, a progressive decrease of bone presence due to the end of the ribs can be observed, reaching a minimum at the lumbar vertebrae levels. After that, an increase of bone presence can be observed, in this case due to the iliac crest. These considerations can be quantified examining the number of white pixels (NoP) in each slice of the previously created binary mask. To study its trend, the NoP of each slice is plotted. Moreover, in order to avoid problems linked to some discontinuities in the NoP, the plot is smoothed.

As said, a negative slope of the NoP can be observed that is due to the absence of ribs in the CT scan when moving downwards slice by slice. The NoP increases again once reaching the iliac creast. In this case, it is possible to fix an NoP threshold, which encompasses the slices of the least NoP and should

end with the L4 or L5 slice. Since the majority of CT scans present this trend, the mean of NoP is chosen as threshold. This way, the choice of the threshold is specific to every CT scan. The intersection of the threshold with the NoP function shows how there is a central interval in which the NoP is under the value, or interval of interest (IOI). Figure 4.2 [35] illustrates the detection of the slices containing the L3.

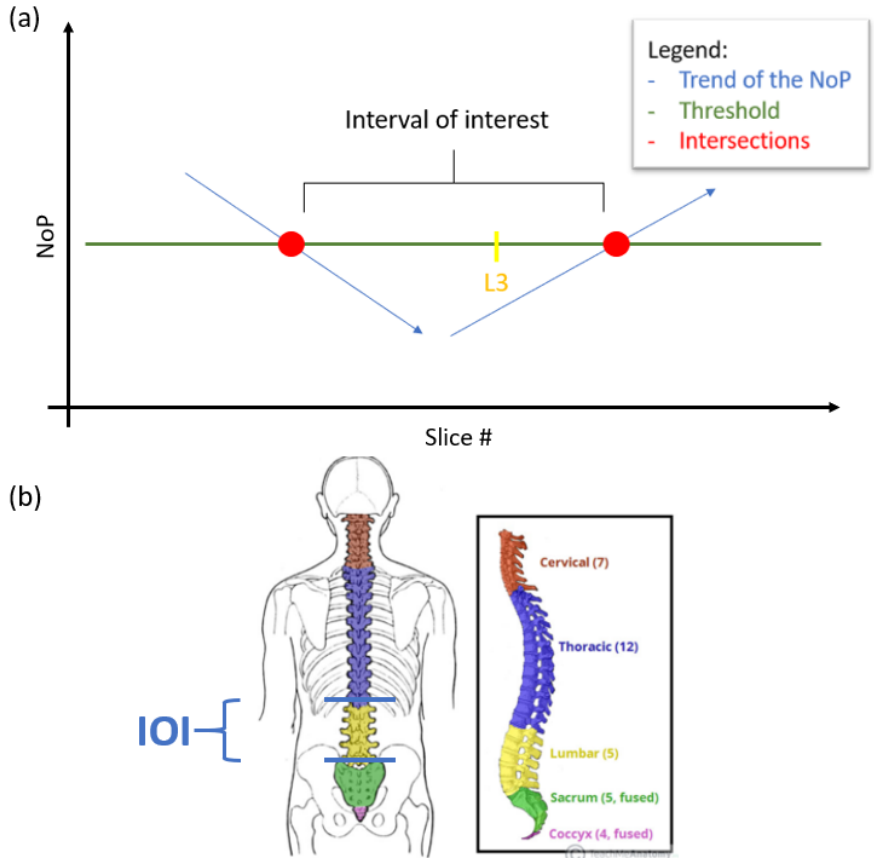


Figure 4.2: Figure (a): schematic representation of the IOI detection containing the L3 slice. Figure (b): theoretical intersection between the threshold and the spinal cord [35].

This part of the methodology is based on the visual identification of landmarks in axial CT slices when manually selecting the L3 slice following the Alberta protocol. An in-depth explanation is given below in section Physical features of spinal cord.

4.2.3 Anatomical features of spinal cord

The Alberta protocol is a standardized method to calculate BC starting from a CT scan, defined by TomoVision (Virtual Magic, Inc., Magog, Canada). A crucial step is the identification of the L3 slice, and its detection is described based on anatomical landmarks. As shown in Figure 1.6, there are some differences between the several kinds of vertebrae and these differences can be used in order to detect a specific one [17].

Figure 4.3 show the position of the lumbar vertebrae between the thoracic and sacral vertebrae. In particular, there are 12 thoracic vertebrae, 5 lumbar vertebrae and 5 sacral vertebrae and despite their closeness, there are important features to discriminate them.

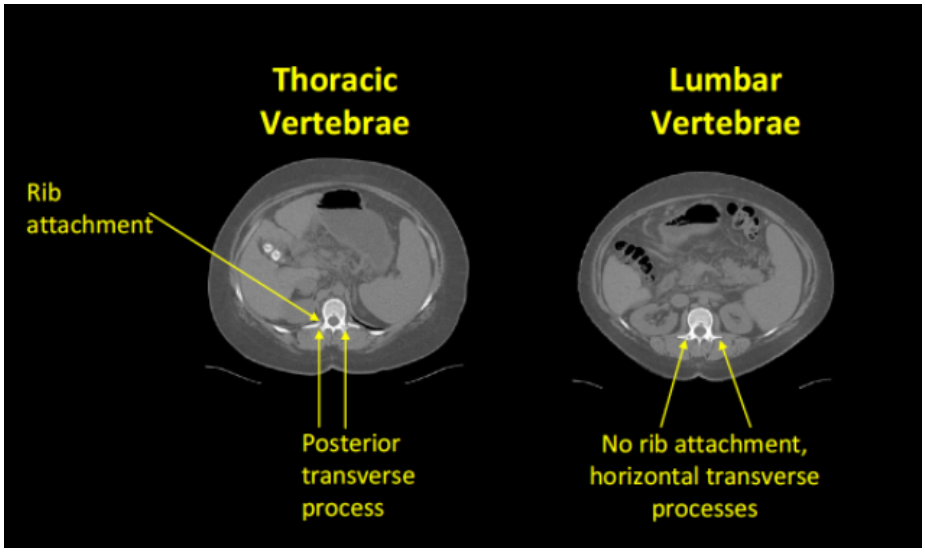


Figure 4.3: Main differences between thoracic vertebrae and lumbar vertebrae [17]

Thoracic vertebrae are characterized by a rib-attachment, producing the bony thorax. This feature is the key to discriminate between thoracic vertebrae and lumbar ones. Since lumbar vertebrae have no rib-attachment they have longer transverse processes that point more horizontally. The first lumbar vertebra is the first vertebra without rib attachment and with horizontal transverse processes. Now that L1 is found, to find the other vertebrae it is sufficient to count descending in the image volume, using the transverse processes as yardstick. Since there is space between one vertebra and the following one and considering that the distance between 2 CT or MRI slices is in the order of millimetres, during the sliding of the slices it is possible to notice how the transverse processes disappear and reappear. This disappear/reappear process can be used as a counter in order to define the vertebrae numbering.

Another possible feature is that, usually, not only lumbar vertebra does not have rib attachment but also, at L3 level, there is no rib presence. If ribs are present in the image it probably means that that slice is at L1 or L2 level. Figure 4.4 [17] shows the main features present from T12 to L4 vertebrae.

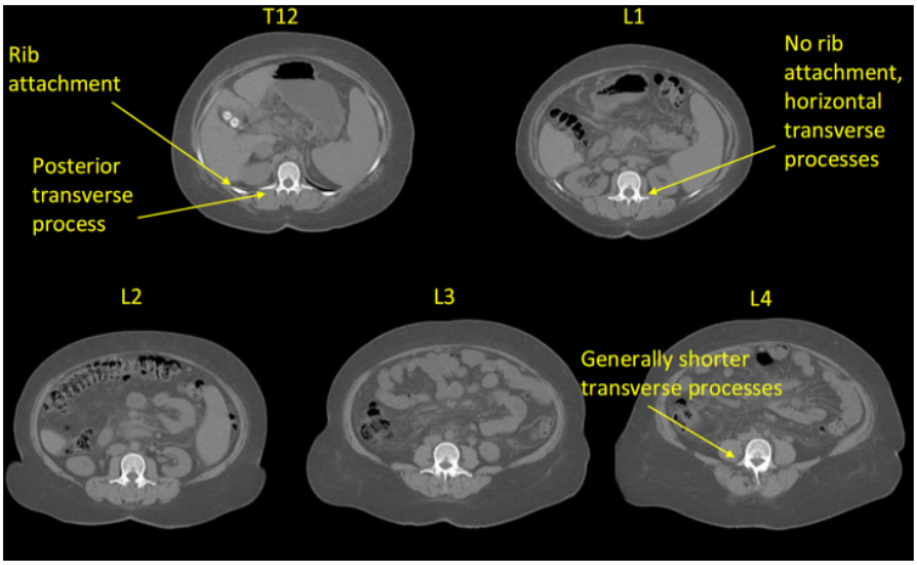


Figure 4.4: Example of CT slices from T12 vertebra to L4 vertebra. The most relevant features are explained in the figure [17].

Apart from the features regarding thoracic vertebrae, the iliac crest is present at the L4 level. However, at the superior L3 level it is not usually possible to see iliac crest. Moreover, L4 vertebra is also usually characterized by shorter transverse processes (Figure 4.5).

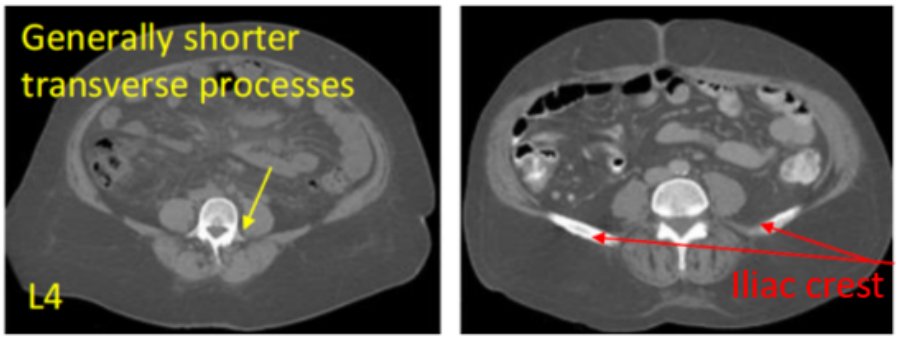


Figure 4.5: Focus on L4 vertebra which usually has shorter transverse processes and at its level can be presence of iliac crest [17].

Finally, sacral vertebrae are easy to be detected because, besides the presence of the pelvis, they are fused together. In any case, usually it is not necessary to consider sacral vertebrae in order to find L3 position because the other features give enough keys to find it.

After the definition of the IOI, the next step is to identify the L3 slice, which is located inside the IOI. In this case, the analysis is just focused on the vertebrae, and for this reason a specific 2D mask is created, which is used as a reference region of interest. The IOI is a 3D binary mask composed of approximately 50-100 slices and is reduced to a single 2D mask representing the vertebrae area. This mask is obtained performing slice by slice operations. Starting from a total black image (all pixel intensities equal to 0), an OR-logical operation is performed between this image and the first slice of the IOI. The result of this operation, along with the subsequent slice of the IOI, are then the inputs of a new OR-logical operation. This way, since both 1 OR 0 and 1 OR 1 are equal to 1, all pixels referring to bone are selected. This operation is performed subsequently until the last slice of the IOI, obtaining a binary mask in which the 'true' pixels correspond to the superimposed masks of the individual slices of the IOI.

However, given that the IOI is composed by 50-100 slices, ribs and probably part of the iliac crest might also be present, apart from the vertebrae. To solve this problem, 2 approaches are proposed. Both techniques generate a binary mask in which just the pixels referring to vertebrae are considered, called vertebrae area mask (VAM).

- A semi-automatic approach using region growing: Once the OR operations are finalised, , the user chooses the correct seeds and the algorithm creates a mask in which only the white pixels connected to the seeds are taken into consideration.
- An automatic approach using vertebrae characteristics: Once the 2D binary mask is created, the different non-connected elements correspond to vertebrae, ribs and/or other regions. It is probable that the element with the largest area is the one corresponding to the vertebrae. Looking at the vertebrae from an axial point of view it can be observed how they are not exactly in the same position due to the curvature of the spinal cord. In this case, this aspect is taken advantage of because it increases the area of the element corresponding to vertebrae locations. In this automatic approach, the element with the largest area is isolated and all other regions that are not connected are eliminated.

Now that the VAM is created the L3 slice can be detected in the IOI. Moreover, thanks to this algorithm, it is also possible to detect the other lumbar vertebras. Due to the spaces between adjacent vertebrae the function of the NoP of the IOI masks shows peaks and valleys. In a central slice of a vertebra the NoP is higher, whereas the space between two vertebras corresponds to a relative minimum. Lastly, the peaks of that specific signal are detected and the number corresponding to the L3 slice is identified. The assumption for L3 identification

is that the intersection between IOI and NoP is somewhere between L4 and L5 level, so that if the last peak is close (< 10 slices) to the end of the IOI it is considered to be L4. On the contrary, if the last peak has more than 10 slices distances from the end point of the IOI, this is defined as L4. In order to define the position of the others lumbar vertebrae, it is sufficient to count backwards the peaks. In this way L3 position is determined.

4.2.4 Validation

The correct L3 slice of all images of the available database is visually selected following the Alberta protocol described above. In order to validate the algorithm, the distance (in number of slices) from the predicted L3 slice and the one visually selected is calculated. This operation is done for both the semi-automatic method and the fully automatic one. Calculating the distance between the prediction and the correct L3 slice is considered the accuracy of the algorithm.

Moreover, the computational time of the two methods is calculated, which is measured for all the CT scan per each technique. All the time variables per each method are then averaged in order to obtain the mean values of the processes.

5 Results

5.1 Pre-processing

As explained in 4.2, the first binary masks are created from DICOM files by thresholding. Figure 5.1 shows the limitations of thresholding, because other tissues present the same pixel intensity as the bone, as well as part of the structure of the patient bed. The masks are then cleaned deleting automatically the elements of the patient bed and little parts of different tissues. Region props is used both for the measurement of the element with major axis length and for the deleting little elements (with area < 30 pixels). For both these operations, 8-connected neighbourhood pixels of the elements are considered as structural element. Figure 5.2 shows examples of the cleaned masks.

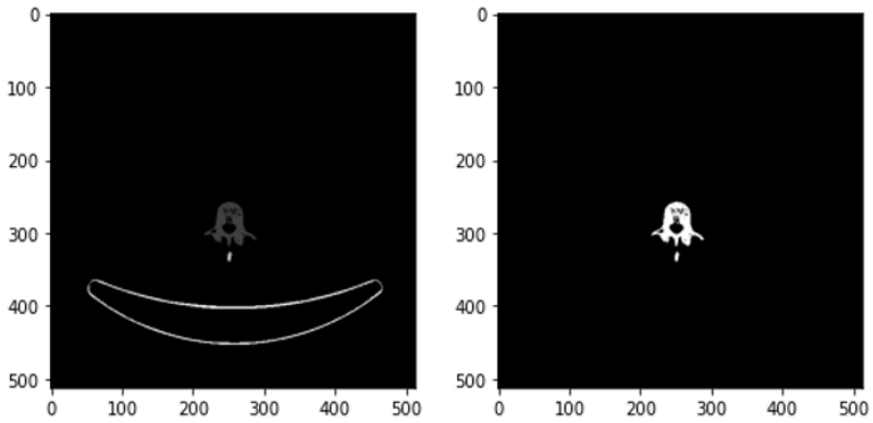


Figure 5.1: Figure in the left shows the result after image thresholding. The enhanced mask on the right is obtained thanks to the post-processing

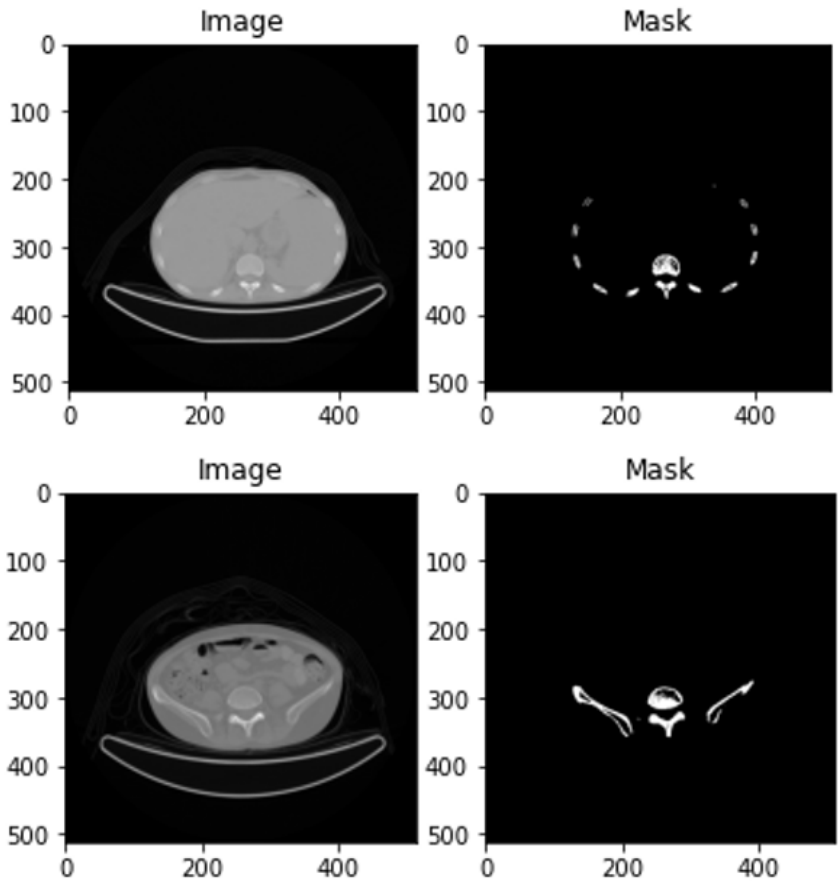


Figure 5.2: Two different examples of mask creation. The left column shows the original CT slices taken from different patients, whereas the right column shows the respective binary masks.

5.1.1 Interval of Interest

The detection of the IOI is done by analysing the regions of the binary masks. As explained in the previous chapter, at L3 level the only bone component is usually due to the vertebra and the expectation is a negative slope of the NoP in each slice until a global minimum when starting from the top of the 3D volume. Figure 5.3 shows some examples where the graphs represent the NoP (Y axis) of each slice of the binary masks (X axis).

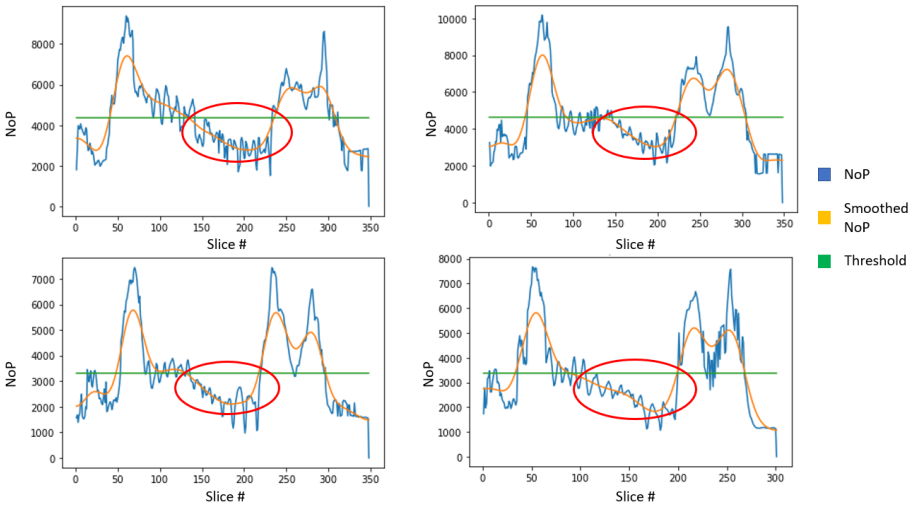


Figure 5.3: Examples of the typical "downhill" trend in correspondence with the lumbar vertebrae. The green line corresponds to the threshold used to find the IOI. The orange curve corresponds to the smoothed NoP. The ellipses highlight the IOI

The IOI is then used for the creation of the VAM, but it is necessary to precise some points:

- First, adjacent masks can have big differences in NoP, therefore the graph is smoothed. This operation is necessary to avoid problems linked to this sort of discontinuities.
- Second, as it is possible to notice in the examples of Figure 5.4, the slopes are usually steeper in the rising part. This is due to the iliac crest and it is possible to imagine that the intersection of the threshold and the graphic of white pixels is at L5 level.

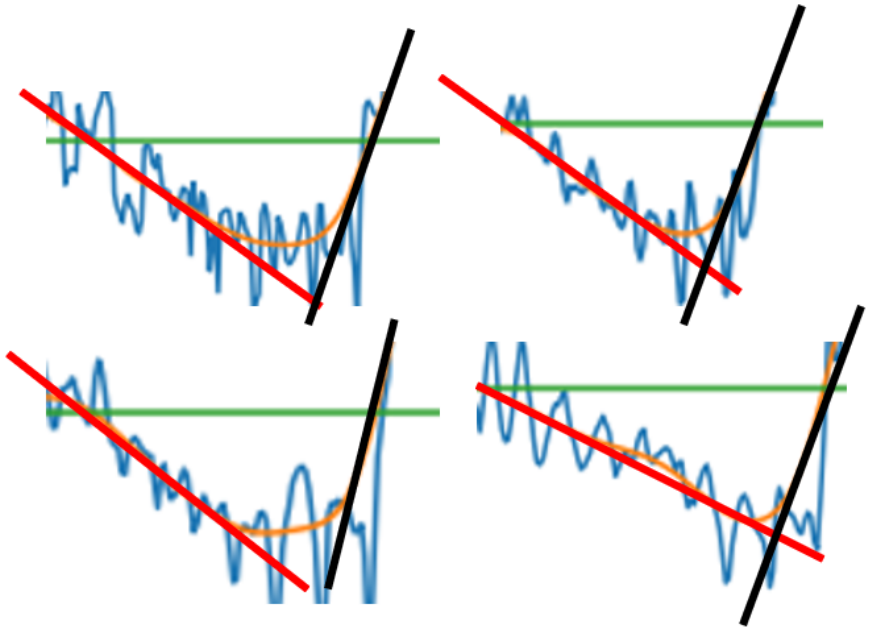


Figure 5.4: Red lines approximate the slope of the descending part, while black approximate the slope of the rising trend. These examples show how usually the rising slope is bigger (in absolute value) than the descending slope

- Lastly, there are CT scans in which the descending slope is flatter compared to the rising one, and result in a large IOI. Having a large IOI produces a VAM with more elements, because it also includes slices with ribs and maybe other tissue. To avoid this, a limit is fixed and, if the left NoP-threshold cut-point is under 140, it is automatically set to 140. Figure 5.5 shows an example of this situation.

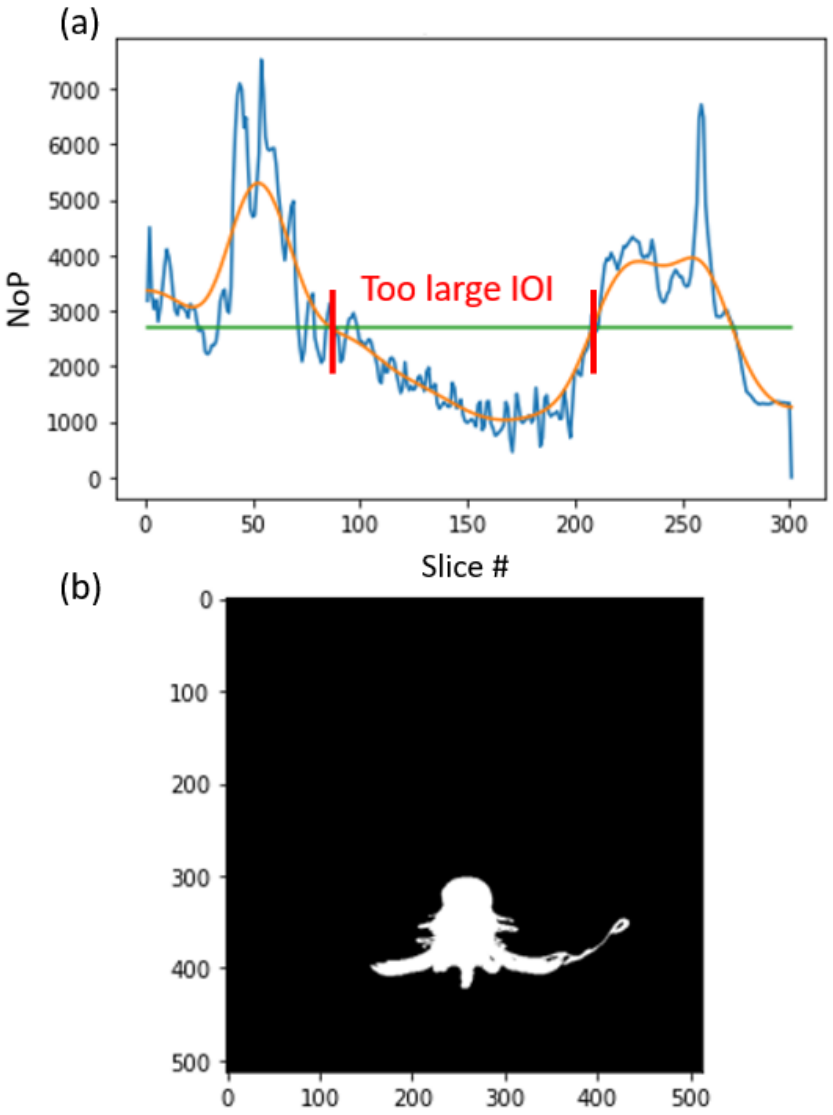


Figure 5.5: Some graphs of NoP can produce an IOI that is too large. In cases like the one in figure (a), the IOI is automatically reduced to avoid creation of VAMs as the one in (b)

5.1.2 Vertebrae Area Mask

With the creation of the VAM it is possible to monitor just the area related to vertebrae and the corresponding variation of the NoP. Figure 5.6 shows some examples: images (a), (b) and (c) show a lateral stability of the position of the vertebrae in patients without relevant problems in the spinal cord. Moreover,

due to the curvature, the shape of the VAM is similar to a bigger vertebra with many transversal processes. However, image (d) is a VAM of a patient which probably suffers for scoliosis (d), this is the reason for which the or mask is not symmetrical with the y axis.

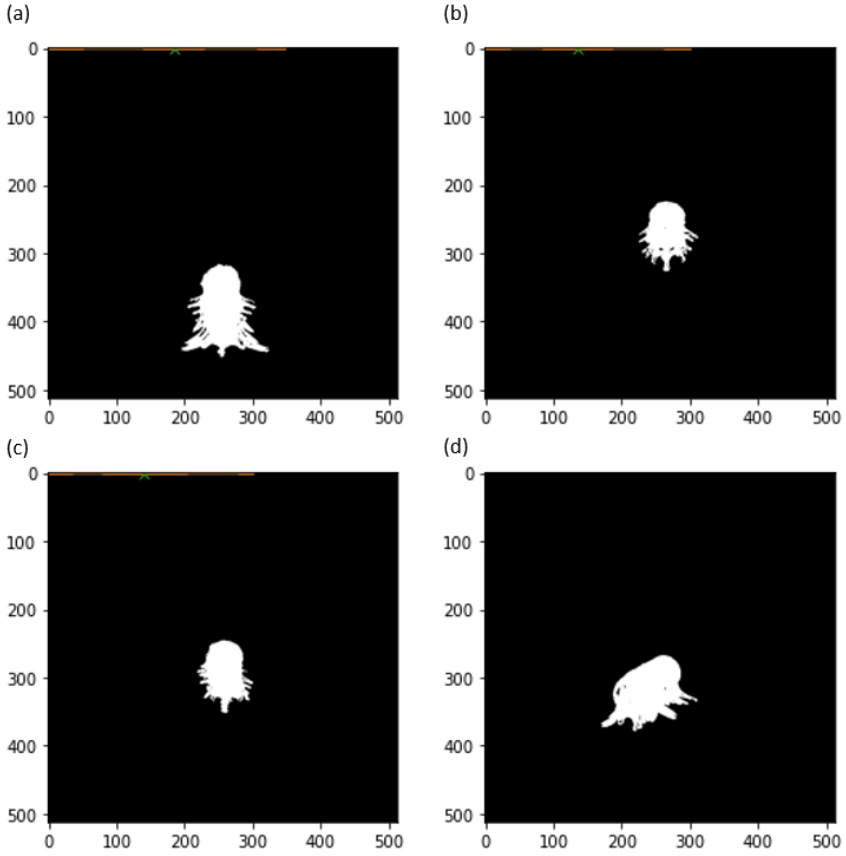


Figure 5.6: (a) (b) (c): three examples of VAM obtained from patients without relevant problems at the spinal cord. (d): example of an VAM probably obtained from a patient suffering from scoliosis, and this reflection is made by considering the non-symmetry of the VAM with respect to its Y axis

As said in 4.2.3, in order to obtain the VAM it is necessary to select the correct element of the mask created through OR-operations, and this is done in two ways. The first method is a semi-automatic one and the user chooses the correct element manually. This is done thanks to a region-growing algorithm which, after receiving the input seeds, starts to grow until all the 8-connected neighbourhood pixels of the element give origin to the final mask. Instead, the second method consists in the choice of the element with bigger area. Figure 5.7 shows an example for both the techniques.

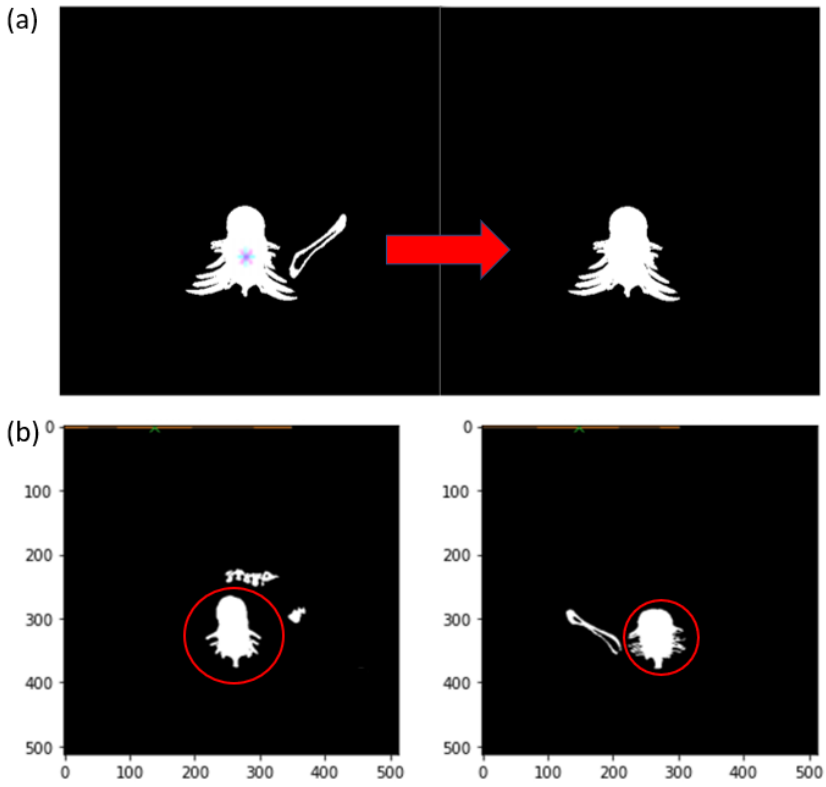


Figure 5.7: (a): example of region growing process: the user chooses the seeds with the mouse and the process starts, giving as output the VAM made by the elements connected to seeds. (b): Examples of the automatic method, which chooses the element with the largest area for the creation of the VAM.

5.1.3 L3 detection

Figure 5.8 shows the graph of the NoP in the IOI after clearing the binary masks. The relative maximums correspond to vertebrae positions, whereas the relative minimums correspond to the space between two adjacent vertebrae, where the presence of bone is lower. In this way, considering the positions of the local maximums, it is possible to define not only the L3 slice, but also the others lumbar vertebrae. Figure 5.8 illustrates how each maximum corresponds to a lumbar vertebra. Comparing the CT slices to the masks at the peak positions, the anatomical landmarks described previously can be identified (Figure 5.9). First of all, as explained in the Alberta protocol, at L1 and L2 level can still be rib presence, as figure (a) and (b) show. Instead, at L3 level (c) usually there is no rib presence. Finally, at L4 level the iliac crest can be seen, as figure (d) illustrates.

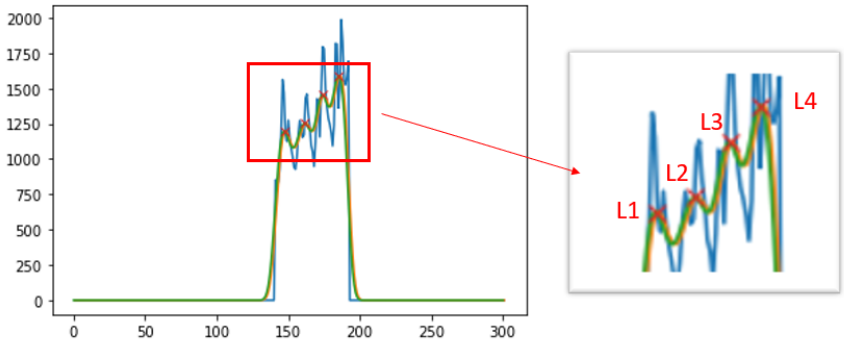


Figure 5.8: Peak positions corresponding to lumbar vertebrae positions

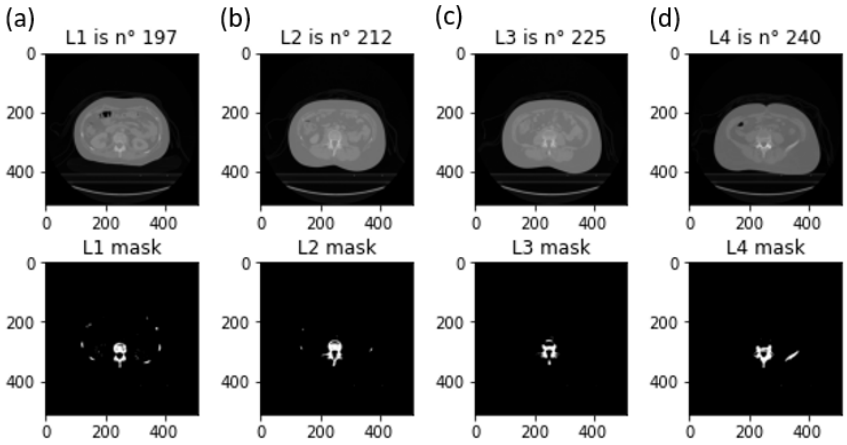


Figure 5.9: Group of CT slices with respective masks. These images show how, thanks to peaks detection, also the others lumbar vertebra can be estimated

5.2 Validation

The developed algorithm is validated with all available images of the database. The performance is determined by calculating the difference between the predicted position and the correct position defined manually following the Alberta protocol. Figure 5.10 shows a graph of the differences, not taking into consideration 3 outliers. A positive error indicates that the detected L3 slice is above/under the visually identified L3 slice. The mean (± SD) error is 5.12 ± 7,96 slices with the largest error being 27 slices.

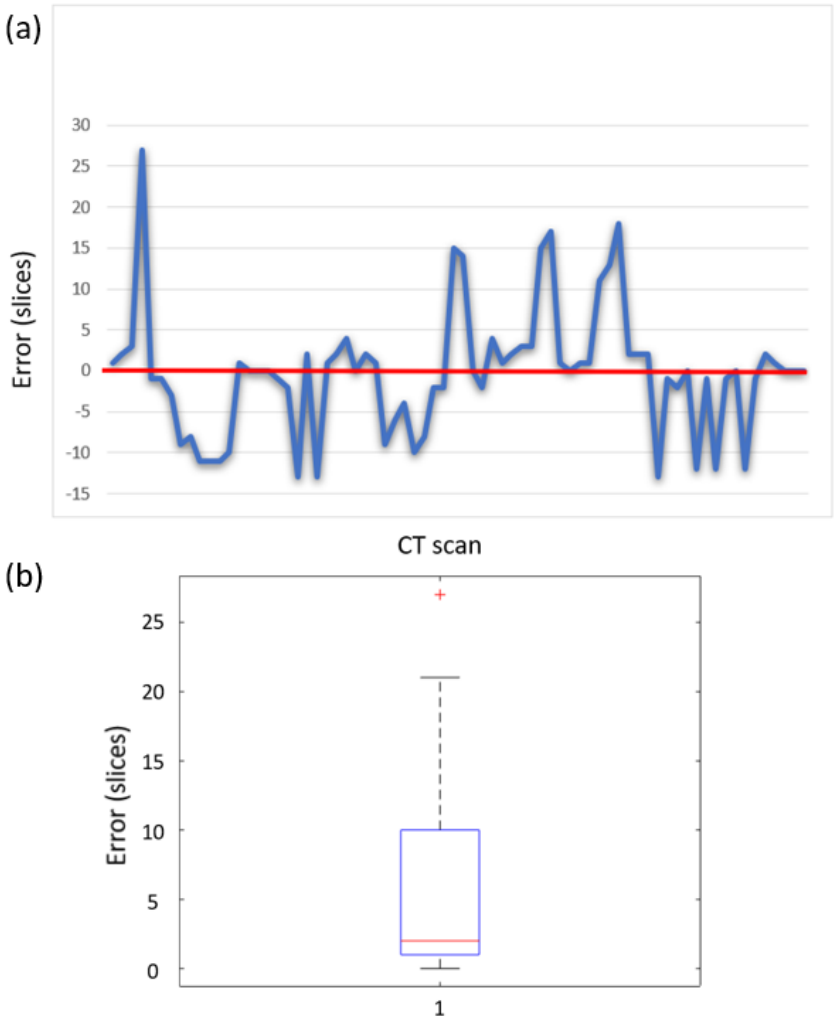


Figure 5.10: (a) Graph of the errors calculated as the difference between the prediction and the correct L3 position. (b) boxplot of the errors, to show the shape of the distribution, its central value and its variability.

The last aspect that is interesting to consider is the computational time of the algorithm. In order to give significant numbers, all the computational times of the automatic processes are averaged to obtain the mean (±SD). The same is done for the semi-automatic technique. Table 5 shows the average time needed to obtain the L3 slice position for both techniques.

	Time (s)
Mean AM ± SD	14.75±1.53
Mean SAM ± SD	49.01±16.59

Table 5: Mean and SD of the computational time calculated for both the automatic (AM) and the semi-automatic methods (SAM).

6 Discussions

Measurement of BC has a significant importance on the valuation of nutritional status in cancer patients. In particular, L3 level is considered the gold standard region of the human body to study this feature. BC is calculated thanks to an image segmentation after the detection of this slice, however, up until now the only ways to perform a detection of that particular level, except for the visual detection, are based on CNN architectures which requires a vast dataset and a training part. Therefore, the aim of the current study is to detect that level (and, potentially, also detect other vertebrae in a CT scan) without neither neural networks nor any type of training, but to base the algorithm on anatomical landmarks detectable in CT scans. A dataset of 80 CT scans provided by the *Hospital Universitario 12 de Octubre* in Madrid is analysed in order to evaluate the efficiency of the technique.

The masks creation shows good performance in highlighting just the bone components of the images. While some present other elements (like part of the structure of the patient table), the final prediction is not affected for one reasons: by focusing the study just in the VAM area, all the other elements in the mask become insignificant. Moreover, regarding the IOI, it is shown that the rising trend just after L3 position is due to the iliac crest, which usually appears at L4 or L5 level. For this reason, the intersection between the threshold and the positive slope of the NoP graph is assumed to be at L5 level, because the increment is very rapid, and this is due to the presence of the iliac crest in the image (and respectively in the masks).

Regarding patients with particular problems to the anatomy of spinal cord, the only case present in the database is analysed correctly. Even if there are no other cases of this kind, theoretically diseases like scoliosis should not be problems for the algorithm. However, they should be assessed on a case-by-case basis.

The analysis of the NoP demonstrates the theoretical assumption of the alternating relative maximums and minimums in correspondence to vertebrae positions. This consideration was to be expected because the spaces between vertebrae there is less bone presence (intervertebral discs are smaller that the vertebrae). Moreover, thanks to this trend not just the L3 level, but also other lumbar vertebrae can be detected. The only operation to do is count the peaks.

Regarding the differences between the automatic method and the semi-automatic one, for this database, the performances in terms of final prediction are the same. The semi-automatic method is proposed for the cases where the operator prefers a visual feedback. For example, it may be useful for patients with very particular problematics which have internal prothesis, which could influence a lot the prediction. This method can be considered more reliable

because there is a feedback from the operator but since the results are the same it is possible to say that the automatic method is better.

Apart from that, the only consideration to do about the difference between the two methods regards the computational time. As can be expected, the semi-automatic method is slower than the automatic one. This difference of time is due to the region growing process. The creation of the final element starting from the seeds takes more time than just choosing the element with the maximum area. Table 5 shows the values obtained for both the techniques and the conclusions to be made are:

- The mean of time for the automatic method is lower.
- The computational time for the automatic method is much more stable. Its standard deviation is of the order of 10%, whereas for the semi-automatic technique it is more than 30%.
- The region growing process is highly relevant in terms of time, and it is also highly variable as regards different CT scans. This is due to the dimension of the VAM, because the bigger it is and the more time it takes to be processed.

The results obtained can be considered acceptable, but the present algorithm does not reach the performances of the studies discussed in the state of art, where the errors on the final prediction are 1.53 ± 4.22 slices [18] and 2.37 ± 2.30 [30]. The results are worst for this algorithm, but it is necessary to take into consideration that this limitation can be avoided because one of the peaks is L3.

Regarding the wrong predictions, it is important to evaluate a recurrent condition: for all the cases in which L3 is not detected correctly the algorithm gives as output still a lumbar vertebra, adjacent to L3. In the majority of cases it is the L4 vertebra, but in few cases it is L2. The histogram of the error in Figure 6.1 shows that usually the error is under 5 slices, or over 10. These results can be seen as the fact that the algorithm gives as output a central position of the vertebra, but in the wrong case the output refers to a vertebra adjacent to L3. The distance between 2 vertebrae is different from people to people but, considering a mean, it can be considered as 3-4cm [36]. Since the slice thickness of the of the CT scans is 5mm, and considering that the distance between two vertebrae is about 10-20 slices, this can be seen as a good result because the mean error is less than the distance of two vertebrae. Moreover, there are only two cases where the error is more than 20 slices.

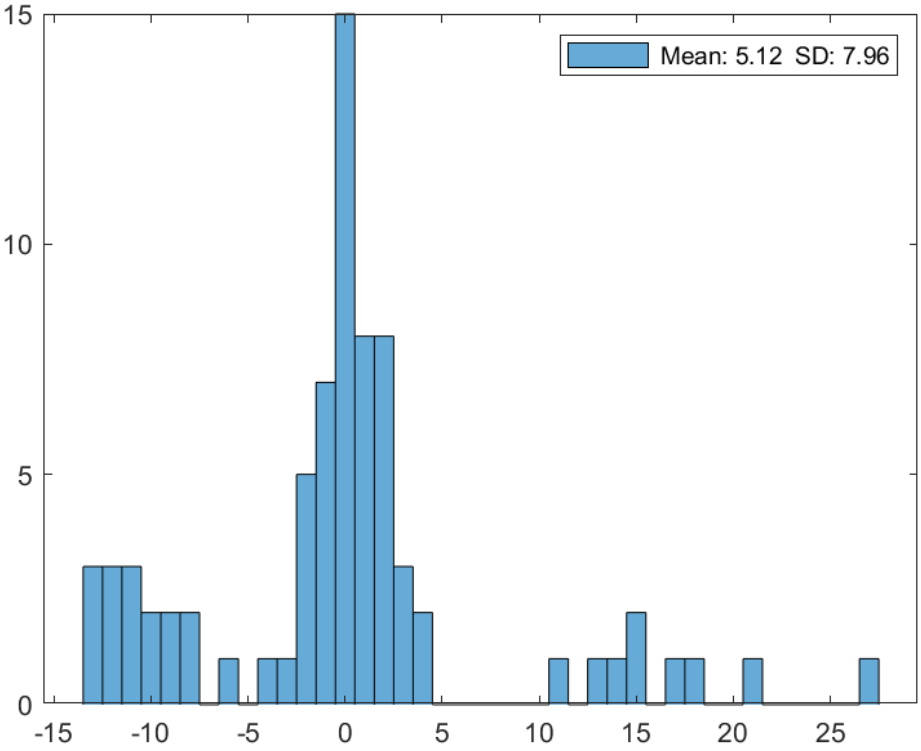


Figure 6.1: Histogram of the prediction errors. As it is possible to notice, a good percentage of prediction has an error under 5 slices, whereas the errors in the range [10-20] (in absolute values) usually refers to a vertebra adjacent to L3 (L2 or L4)

Concerning the outliers mentioned in the previous 5.2, Figure 6.2 shows the graphs of the NoP for the relative CT scans. As it is possible to notice, the graphs present large peaks and valleys along the whole CT volume. This trend makes it impossible for the algorithm to choose the correct IOI and this brings to a failure to predict . For (a) and (b) the prediction is solved just reducing the CT to 400 slices, starting the count from the top (head). Regarding figure (c), the pixel intensities are more homogeneous (less contrast between different tissues) than the one in (d), which is a NoP graph of a CT scan with correct prediction. This causes a NoP graph with less pronounced peaks and valleys in its smoothed version as more pixels are segmented by the selected bone tissue threshold.

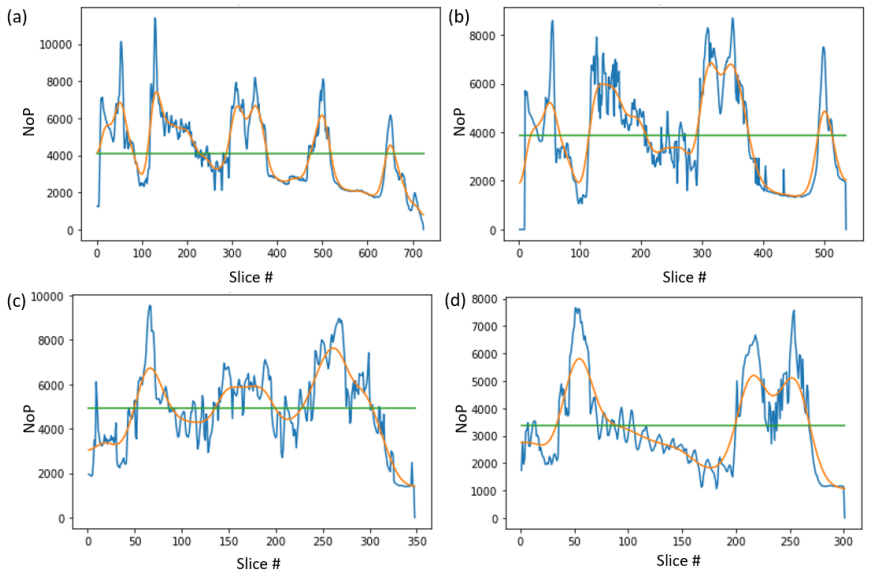


Figure 6.2: (a) (b) (c) are the NoP graphs of the outliers. For (a) and (b) the prediction become correct if the number of CT slices is reduced to 400, starting from the top (head). Regarding (c) there is still no prediction even with the reduction of the CT scan. Its shape is weird in respect to the graph in (d), which is the NoP of a CT scan with correct prediction.

Finally, since the only physical aspect present into the DICOM files is weight, the correlation between this parameter and the error in slice is evaluated. This is done by calculating Pearson's correlation coefficient between the patient weight and the error in slices. The obtained correlation coefficient is r = 0.09 (p=0.39), showing a weak and not statistically significant correlation between the two variables. Therefore, it can be concluded that patient weight does not influence the detection of the L3 slice. The main limitation of this study is the recurrent problem of predicting an adjacent vertebra, despite the fact that the correct L3 position is still present because it is one of the peaks. Another limitation is the scarcity of information of the database: sex, age and height are not present, and this makes difficult to evaluate the correlation between those factors and the wrong prediction. Despite those limitations, the results are coherent in the majority of the cases with the operator predictions with just 1 outlier.

The main limitation of this study is the recurrent problem of predicting an adjacent vertebra, despite the fact that the correct L3 position is still present because it is one of the peaks. Another limitation is the scarcity of information of the database: sex, age and height are not present, and this makes difficult to evaluate the correlation between those factors and the wrong prediction. Despite those limitations, the results are coherent in the majority of the cases with the operator predictions with just 1 outlier.

7 Conclusion

This master's thesis is based on the study of L3 slice detection with the aim of a segmentation executed on that level in order to study body composition in cancer patients. In particular the database is made by CT scans coming from patients treated for Hodgkin and non-Hodgkin lymphoma and it is known that cancer patients suffer for muscle loss. Monitoring BC can be important for the treatment and studying its composition at L3 level is one of the gold standards for that clinical background. Currently, the automatic methods to detect L3 slice are based on neural network architectures like CNN, and such techniques require huge datasets and training parts. Thus, the main objective of the present study is to provide a non-neural network architecture-based method to detect L3 slice automatically. Thus, taking into consideration the physical features of human body described by the Alberta protocol, this algorithm has been developed. Thanks to the VAM it is possible to analyse just the area of each slice in which vertebrae are present. In this way, it is theoretically possible to spot also other vertebrae. L3 detection does not reach the same performances as the studies in the state of art, but this algorithm is structured in order to give more vertebrae position and for this reason which the results are good. Lastly, considering the limitations of the study described in the previous chapter, it is possible to define several possible future works. The following future improvements and research lines are proposed:

- To enhance the L3 prediction: as explained, finding lumbar vertebrae positions is not a problem for the developed algorithm, but not always the choice is the correct one. To solve this, the L4 position can be focused on instead of L3, as it presents shorter transversal processes and at L4 level the iliac crest starts to appear.
- To detect of other vertebrae: Taking into consideration physical aspects of the body it may be possible to detect any vertebra in the spinal cord. Therefore, the detection of other vertebrae apart from L3 is proposed as a future addition to the algorithm.
- To consider particular conditions that affect the anatomy of the spinal cord like scoliosis or the presence of internal prothesis.
- To implement a tissue segmentation algorithm (vertebrae prediction + BC assessment): in this way the algorithm would be complete starting with the detection of the slice corresponding to L3 and finishing with a BC measurement.
- To study the correlation between a wrong prediction and the error, and features like height, sex age. Since the database used for this project has just the weight of the patients.

In conclusion , this master's thesis gives an innovative idea for vertebrae detection. This project can be the first part of a complete algorithm which give as final output the BC at any level of the spinal. Moreover, it can be useful both in that cases in which a large database is not available and also it can be used as a first tool before radiologist's prediction.

References

[1] K. C. Fearon, D. J. Glass, and D. C. Guttridge, "Cancer cachexia: Mediators, signaling, and metabolic pathways," *Cell Metabolism*, vol. 16, no. 2, pp. 153 – 166, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1550413112002483

[2] TheGlobalCancerObservatory, "Gco.iarc.fr, 2021," December 2020. [Online]. Available: https://gco.iarc.fr/today/data/factsheets/populations/724-spain-fact-sheets.pdf

[3] S. B. Heymsfield and C. B. Mcmanus, "Tissue components of weight loss in cancer patients. a new method of study and preliminary observations," *Cancer*, vol. 55, no. S1, p. 238–249, 1985.

[4] K. Engelke, O. Museyko, L. Wang, and J.-D. Laredo, "Quantitative analysis of skeletal muscle by computed tomography imaging—state of the art," *Journal of Orthopaedic Translation*, vol. 15, p. 91–103, 2018.

[5] F. Pamoukdjian, T. Bouillet, V. Lévy, M. Soussan, L. Zelek, and E. Paillaud, "Prevalence and predictive value of pre-therapeutic sarcopenia in cancer patients: A systematic review," *Clinical Nutrition*, vol. 37, no. 4, p. 1101–1113, 2018.

[6] C. M. Prado, S. J. Cushen, C. E. Orsso, and A. M. Ryan, "Sarcopenia and cachexia in the era of obesity: clinical and nutritional impact," *Proceedings of the Nutrition Society*, vol. 75, no. 2, p. 188–198, 2016.

[7] P. M. Graffy, J. Liu, P. J. Pickhardt, J. E. Burns, J. Yao, and R. M. Summers, "Deep learning-based muscle segmentation and quantification at abdominal ct: application to a longitudinal adult screening cohort for sarcopenia assessment," *The British Journal of Radiology*, vol. 92, no. 1100, p. 20190327, 2019.

[8] A. J. Cruz-Jentoft, G. Bahat, J. Bauer, Y. Boirie, O. Bruyère, T. Cederholm, C. Cooper, F. Landi, Y. Rolland, A. A. Sayer, and et al., "Sarcopenia: revised european consensus on definition and diagnosis," *Age and Ageing*, vol. 48, no. 4, p. 601–601, 2019.

[9] J. Chindapasirt, "Sarcopenia in cancer patients," *Asian Pacific Journal of Cancer Prevention*, vol. 16, no. 18, p. 8075–8077, 2016.

[10] R. D. Nipp, G. Fuchs, A. El-Jawahri, J. Mario, F. M. Troschel, J. A. Greer, E. R. Gallagher, V. A. Jackson, A. Kambadakone, T. S. Hong, and et al., "Sarcopenia is associated with quality of life and depression in patients with advanced cancer," *The Oncologist*, vol. 23, no. 1, p. 97–104, 2017.

[11] S. Onishi, M. Tajika, T. Tanaka, Y. Hirayama, K. Hara, N. Mizuno, T. Kuwahara, N. Okuno, Y. Inaba, T. Kodaira, and et al., "Prognostic significance of sarcopenia in patients with unresectable advanced esophageal cancer," *Journal of Clinical Medicine*, vol. 8, no. 10, p. 1647, 2019.

[12] N. Chargi, S. Bril, M. Emmelot-Vonk, and R. Bree, "Sarcopenia is a predictor for overall survival in elderly patients with head and neck cancer," *Radiotherapy and Oncology*, vol. 132, p. 94, 2019.

[13] S. J. Withrow and D. M. Vail, "Chapter 16 - supportive care for the cancer patient," May 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780721605586500198

[14] Y. Lu, J. K. Hahn, and X. Zhang, "3d shape-based body composition inference model using a bayesian network," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 1, p. 205–213, 2020.

[15] W. Shen, M. Punyanitya, Z. Wang, D. Gallagher, M.-P. St.-Onge, J. Albu, S. B. Heymsfield, and S. Heshka, "Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image," *Journal of Applied Physiology*, vol. 97, no. 6, p. 2333–2338, 2004.

[16] S. L. Gomez-Perez, J. M. Haus, P. Sheean, B. Patel, W. Mar, V. Chaudhry, L. Mckeever, and C. Braunschweig, "Measuring abdominal circumference and skeletal muscle from a single cross-sectional computed tomography image," *Journal of Parenteral and Enteral Nutrition*, vol. 40, no. 3, p. 308–318, 2015.

[17] Y. Martel, "Alberta protocol," February 2017. [Online]. Available: https://tomovision.com/Sarcopenia_Help/index.htm

[18] F. Kanavati, S. Islam, E. O. Aboagye, and A. Rockall, "Automatic l3 slice detection in 3d ct images using fully-convolutional networks," Nov 2018. [Online]. Available: https://arxiv.org/abs/1811.09244

[19] M. Marra, R. Sammarco, A. D. Lorenzo, F. Iellamo, M. Siervo, A. Pietrobelli, L. M. Donini, L. Santarpia, M. Cataldi, F. Pasanisi, and et al., "Assessment of body composition in health and disease using bioelectrical impedance analysis (bia) and dual energy x-ray absorptiometry (dxa): A critical overview," *Contrast Media amp; Molecular Imaging*, vol. 2019, p. 1–9, 2019.

[20] M. Kocak, "Risonanza magnetica - argomenti speciali." [Online]. Available: https://www.msdmanuals.com/it-it/professionale/argomenti-speciali/principi-di-imaging-radiologico/risonanza-magnetica

[21] S. Sammet, "Magnetic resonance safety," *Abdominal Radiology*, vol. 41, no. 3, p. 444–451, 2016.

[22] L. Yao, A. L. Yip, J. A. Shrader, S. Mesdaghinia, R. Volochayev, A. V. Jansen, F. W. Miller, and L. G. Rider, "Magnetic resonance measurement of muscle t2, fat-corrected t2 and fat fraction in the assessment of idiopathic inflammatory myopathies," *Rheumatology*, 2015.

[23] S. Felson, "Ct scan and cancer risk," Jan 2020. [Online]. Available: https://www.webmd.com/cancer/can-ct-scans-lead-to-cancer

[24] "Fda.vog," https://www.guitex.org/home/en/forum/5-tex-e-latex/105703-note-a-fondo-pagina-risolto, (Accessed on 01/10/2021).

[25] I. Liguori, G. Russo, L. Aran, G. Bulli, F. Curcio, D. Della-Morte, G. Gargiulo, G. Testa, F. Cacciatore, D. Bonaduce, and et al., "Sarcopenia: assessment of disease burden and strategies to improve outcomes," *Clinical Interventions in Aging*, vol. Volume 13, p. 913–927, 2018.

[26] J. E. Morley, "Treatment of sarcopenia: the road to the future," *Journal of Cachexia, Sarcopenia and Muscle*, vol. 9, no. 7, p. 1196–1199, 2018.

[27] "Testosterone replacement therapy: Side effects, dosages, treatment, interactions, warnings," Feb 2017. [Online]. Available: https://www.rxlist.com/consumer_testosterone-medication/drugs-condition.htm

[28] M. Fioranelli, "Vitamina d e cancro," Apr 2020. [Online]. Available: https://www.artoi.it/vitamina-d-e-cancro/

[29] "About dicom- overview." [Online]. Available: https://www.dicomstandard.org/about

[30] S. Belharbi, C. Chatelain, R. Hérault, S. Adam, S. Thureau, M. Chastan, and R. Modzelewski, "Spotting l3 slice in ct scans using deep convolutional network and transfer learning," *Computers in Biology and Medicine*, vol. 87, p. 95–103, 2017.

[31] J. Brownlee, "A gentle introduction to transfer learning for deep learning," Sep 2019. [Online]. Available: https://machinelearningmastery.com/transfer-learning-for-deep-learning/

[32] H. Lamba, "Understanding semantic segmentation with unet," Feb 2019. [Online]. Available: https://towardsdatascience.com/understanding-semantic-segmentation-with-unet-6be4f42d4b47

[33] "Dafs: Voronoi health analytics." [Online]. Available: https://www.voronoihealthanalytics.com/dafs3

[34] V. Chougule and B. Ahuja, "Clinical case study: Spine modeling for minimum invasive spine surgeries (miss) using rapid prototyping." 2017.

[35] "The vertebral column." [Online]. Available: https://teachmeanatomy.info/back/bones/vertebral-column/

[36] I. Busscher, J. J. W. Ploegmakers, G. J. Verkerke, and A. G. Veldhuizen, "Comparative anatomical dimensions of the complete human and porcine spine," *European Spine Journal*, vol. 19, no. 7, p. 1104–1114, 2010.

A Ethical, economic, social and environmental aspects

A.1 Introduction

The study of body composition can be of great importance for cancer patients because therapies and treatments can be customized to individual patient needs. Given the high percentage of cancer patients suffering from tissue degradation, the study of body composition has now a bigger importance. One of the methods to evaluate it without having to analyse the entire body of the patient is to study the composition at the level of the third lumbar vertebra.

Over the years, automatic methods have been developed for the detection of this level, but nowadays the available methods are based on CNN architectures, which require a large amount of samples in addition to a training process. For this reason, it would be desirable to have an algorithm still automatic, but that can be used even in those cases where you do not have a large amount of data available. An algorithm based on vertebrae characteristics is developed in this master's thesis. This algorithm could lead to a reduction in costs as well as a greater effectiveness of the times, since it would be possible to study more patients in the same unit of time. Finally, more importantly, in the first place reducing time would benefit patients, who would receive diagnosis/care in less time.

A.2 Most relevant impacts of the project

Primarily, this project has a social impact that affects the patient's health and would also reduce costs and time within a radiology department. To avoid a negative social aspect and considering that all the data analysed came from real hospital patients, all the resources were anonymized. In addition, the regulations followed are:

- Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de carácter personal (LOPD)
- Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamentode desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de carácter personal (RDLOPD)
- Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica (LAP)

- Reglamento (UE)2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento General de Protección de Datos (RGPD), aplicable a partir de mayo de 2018
- Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales

Furthermore, this project has no environmental aspects and does not produce waste that could damage the ecological system. Table 6 illustrates the main ethical, economic and social aspects described so far.

Relevant aspects	**Description**	**Affected group/ sectors**	**Regulations, laws standards and ethical codes of reference**
Social	Health and quality life of cancer patients with muscle loss	Cancer Patients which express weight loss	In relation to the data protection law
Ethical	Related to the commitee of ethical aspects	Medical personnel and patients	In relation to human rights and bioethics
Economic	Cost reduction due to time reduction	Medical sector	

Table 6: Main ethical, economic and social aspects

A.3 Conclusion

In conclusion, this study aims to improve the efficiency of what can be scientific research or can also, after implementing the necessary improvements, be used within a radiology department. In addition, as the software does not suffer operator-dependent errors, the implementation of such software (in addition to the automatic segmentation algorithm) would greatly reduce costs but especially the time, since the speed of the algorithm would allow many more patients to be evaluated in the same time frame. However, it is important to note that, at the moment, the algorithm is not tested with special cases of altered spinal cord anatomy, like scoliosis. The only case present in the database has had a correct prediction but not enough to determine the effectiveness of the algorithm for such specific cases.

Therefore, to conclude, the long-term objective of this project is a future hospitalization of the same, to be used as the first necessary step in the tissue segmentation process. Also, since the algorithm is able to recognize not only L3 but also other vertebrae, it could be interesting for the doctor so as to have a

first impression on the clinical state of the patient and then decide whether to deepen the exams or not.

B Economical budget

- Human Resources Costs: the salary of the people who are part of the project are considered. Table 7 illustrates the cost of the engineering student.

	Hours	**Cost/Hours**	**Total**
Engeering student	900	20€	18000€
Total			**18000€**

Table 7: Human Resources Cost

- Material Resources Costs: this part contains the costs of materials and resources needed for the whole project. Cost per unit, the month of use and the depreciation of the device are all considered and Table 8 shows the total cost.

	Purchase price	**Use (Months)**	**Deprecitation (years)**	**Total**
Personal computer	900€	5	8%	75€

Table 8: Material Resources Cost

- Total cost: the general cost and the industrial benefit are both considered in this estimation. It is important to underline that the general cost is considered as the 15% of the direct cost, whereas the industrial benefit is calculated as the 6% of the total amount of direct and indirect costs. Table 9 shows the total budget.

Human Resources Costs (Direct cost)				
		Hours	**Cost/Hours**	**Total**
Engeering student		900	20€	18000€
Total				**18000€**
Material and resources costs (Direct cost)				
	Purchase price	**Use (Months)**	**Deprecitation (years)**	**Total**
Personal computer	900€	5	8%	75€
Total				18075€
Overhead cost (Indirect cost)		15%DC		1575€
Industrial profit		6%(DC+AC)		724,5€
Subtotal Budget				20374,5€
IVA (21%)				4278,6€
TOTAL BUDGET				**24653,1€**

Table 9: Total cost.