Machine Learning in Medical Image Processing: Third Lumbar Detection for Rapid Visceral Fat Calculation

Luis F. Velasco, Archange G. Destiné, Anastasie Mvula, Japneet S. Kohli

RESEARCH REPORT

December 2020

GEORGE MASON UNIVERSITY

Table of Contents

Table of Figures

Abstract

Obesity has become a chronic medical condition internationally, deserving of increased medical attention and research. One recent development within obesity research is interest to improve efficiency and accuracy of the calculation of the Visceral Fat Index (VFI) obesity measure. VFI calculates the overall composition of the body by measuring the adipose tissue that lies between organs in the torso and is currently manually calculated by physicians through a visual analysis of CT scans. This manual process involves locating the slice image at the third lumbar vertebra (L3) level in the spine before computing the VFI in that location of the body. This is a tedious and time-consuming process, as the CT scan does not automatically identify the location of the L3 within the CT scan images. To address this issue, we have identified and used a previous implementation of a Fully Convolutional Neural Network model and optimized it to obtain a precision score of 100%. Our algorithm is also able to correctly process full-body CT scans, unlike the previous implementation, which was limited to only abdominal or chest CT scans. This tool significantly reduces the amount of time required by a physician to process a CT scan for VFI calculation purposes or any other study where L3 slice position is relevant.

1. Introduction

1.1 Background

Obesity is a complex disease caused by excessive amounts of body fat and is defined by the World Health Organization (WHO) as an “abnormal or excessive fat accumulation that presents a risk to health” [1]. Obesity is diagnosed by measuring the Body Mass Index (BMI), where the patient’s weight (in kilograms) is divided by their square height in meters (kg/m2). A person with a BMI of 30 or over is usually considered obese. Furthermore, a higher BMI can often increase a person's risk to develop diseases such as diabetes, hypertension, and certain types of cancer [1]. Figure 1: Obesity Rate by Country illustrates the obesity rate by country in 2020.

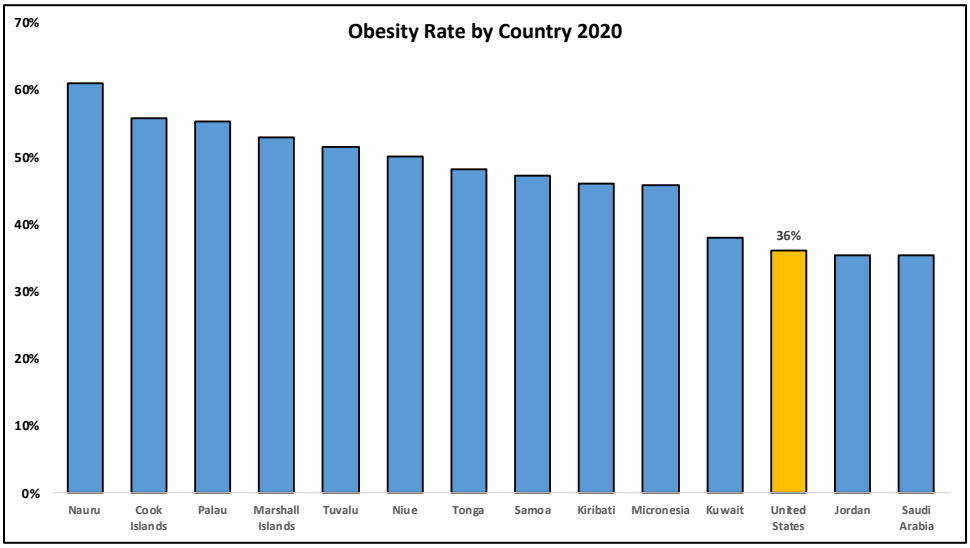

Figure 1: Obesity Rate by Country

Risk factors for obesity include genetics, problematic metabolism, and aspects of socio-behavioral habits. The United States is afflicted with one of the highest obesity rates among the industrial nations, which is primarily attributed to the nutritionally poor diets of the United States population at large. The severity of this health problem is a great cause of concern for the Center for Disease Control (CDC), which has declared obesity as a national epidemic.

The medical community in the United States views obesity as a leading factor for premature death because obesity is not only a major risk factor for metabolic disease, but also cardiovascular diseases. The medical community has invested time and resources to achieve a better understanding of this condition.

The adult population suffers tremendously under the linkage of obesity and respiratory diseases, which can lead to lethal cancer outcomes. It has been found that obese individuals are more prone to develop respiratory tumors. In this context, lung cancer is especially threatening, as it is the most common cancer in the world in women and men. Lung cancer accounted for over 12.3% of all American cancer diagnoses in the year 2018, the highest occurrence among all types of cancer [2].

Moreover, lung cancer is also one of the most dangerous and lethal cancers. There are many factors that can increase the risk of lung cancer. Some of these are avoidable, such as smoking. Other factors are not controllable, such as genetics. Lung cancer research is on the forefront of American cancer research, as medical providers and oncologists work tirelessly to research and understand the linkage between obesity and cancer. Identifying better diagnosis methods, treatments, and even cures are all long-term goals of these research teams that aim to reduce the number of deaths caused by the disease.

An outcome of the research studying obesity as an underlying health condition and its effect on lung cancer outcomes is controversy in the medical field about how to measure obesity. Some studies measure obesity in adults by calculating the BMI, a very common and reproducible calculation displayed in Figure 2. Other studies have found that higher BMI correlated with increased lung cancer survival. This theory is known in the medical field as the “obesity paradox”.

Body Mass Index Table

	Normal						Overweight					Obese										Extreme Obesity														
BMI	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
Height (inches)												Body Weight (pounds)																								
58	91	96	100	105	110	115	119	124	129	134	138	143	148	153	158	162	167	172	177	181	186	191	196	201	205	210	215	220	224	229	234	239	244	248	253	258
59	94	99	104	109	114	119	124	128	133	138	143	148	153	158	163	168	173	178	183	188	193	198	203	208	212	217	222	227	232	237	242	247	252	257	262	267
60	97	102	107	112	118	123	128	133	138	143	148	153	158	163	168	174	179	184	189	194	199	204	209	215	220	225	230	235	240	245	250	255	261	266	271	276
61	100	106	111	116	122	127	132	137	143	148	153	158	164	169	174	180	185	190	195	201	206	211	217	222	227	232	238	243	248	254	259	264	269	275	280	285
62	104	109	115	120	126	131	136	142	147	153	158	164	169	175	180	186	191	196	202	207	213	218	224	229	235	240	246	251	256	262	267	273	278	284	289	295
63	107	113	118	124	130	135	141	146	152	158	163	169	175	180	186	191	197	203	208	214	220	225	231	237	242	248	254	259	265	270	278	282	287	293	299	304
64	110	116	122	128	134	140	145	151	157	163	169	174	180	186	192	197	204	209	215	221	227	232	238	244	250	256	262	267	273	279	285	291	296	302	308	314
65	114	120	126	132	138	144	150	156	162	168	174	180	186	192	198	204	210	216	222	228	234	240	246	252	258	264	270	276	282	288	294	300	306	312	318	324
66	118	124	130	136	142	148	155	161	167	173	179	186	192	198	204	210	216	223	229	235	241	247	253	260	266	272	278	284	291	297	303	309	315	322	328	334
67	121	127	134	140	146	153	159	166	172	178	185	191	198	204	211	217	223	230	236	242	249	255	261	268	274	280	287	293	299	306	312	319	325	331	338	344
68	125	131	138	144	151	158	164	171	177	184	190	197	203	210	216	223	230	236	243	249	256	262	269	276	282	289	295	302	308	315	322	328	335	341	348	354
69	128	135	142	149	155	162	169	176	182	189	196	203	209	216	223	230	236	243	250	257	263	270	277	284	291	297	304	311	318	324	331	338	345	351	358	365
70	132	139	146	153	160	167	174	181	188	195	202	209	216	222	229	236	243	250	257	264	271	278	285	292	299	306	313	320	327	334	341	348	355	362	369	376
71	136	143	150	157	165	172	179	186	193	200	208	215	222	229	236	243	250	257	265	272	279	286	293	301	308	315	322	329	338	343	351	358	365	372	379	386
72	140	147	154	162	169	177	184	191	199	206	213	221	228	235	242	250	258	265	272	279	287	294	302	309	316	324	331	338	346	353	361	368	375	383	390	397
73	144	151	159	166	174	182	189	197	204	212	219	227	235	242	250	257	265	272	280	288	295	302	310	318	325	333	340	348	355	363	371	378	386	393	401	408
74	148	155	163	171	179	186	194	202	210	218	225	233	241	249	256	264	272	280	287	295	303	311	319	326	334	342	350	358	365	373	381	389	396	404	412	420
75	152	160	168	176	184	192	200	208	216	224	232	240	248	256	264	272	279	287	295	303	311	319	327	335	343	351	359	367	375	383	391	399	407	415	423	431
76	156	164	172	180	189	197	205	213	221	230	238	246	254	263	271	279	287	295	304	312	320	328	336	344	353	361	369	377	385	394	402	410	418	426	435	443

Figure 2: Body Mass Index (BMI) Table

1.2 Obesity Paradox

Healthcare community interest in pursuing VFI as a measure of body fat content instead of BMI has been further encouraged by the publication of the obesity paradox. According to the obesity paradox,

which is supported by BMI-based metrics, obese individuals have a protective mechanism that shields them against chronic disease, such as lung cancer [3]. This seemingly paradoxical statement claims that the alleged protective mechanism causes obese individuals to have a higher survival rate against certain diseases than other groups. As a response to these claims and to investigate the veracity of this theory, it has been suggested that alternative measures of body composition, such as the amount of fat and muscle in the abdominal segment, be considered. These may be a better indicator of obesity in the body and may better explain the correlation between chronic diseases, such as lung cancer, and obesity than BMI has accomplished traditionally. Although it is indisputable that obesity leads to major risk factors for metabolic and cardiovascular diseases in mature patients, BMI might not be the most accurate metric for obesity to compute the survival rate. As result, multiple studies have challenged the obesity paradox by proving that BMI is not the best indicator to determine obesity in adult population [3]

Physicians have shifted to a different and more accurate measurement of obesity that can incorporate the fat content more accurately. This alternative means of calculating obesity requires physicians to discern the amounts of adipose (fat) tissue and muscle mass in the body to determine obesity. There are two kinds of adipose tissues in the body: the visceral adipose tissue (VAT) and the subcutaneous adipose tissue (SAT). VAT represents the fat that surrounds the various organs in the body, whereas SAT refers to the fat that lies below the skin. To compute a metric for obesity using VAT and SAT, a ratio of VAT with total fat, which is a sum of VAT and SAT is taken. This metric is called the Visceral Fat Index (VFI) and is linked to age-related obesity caused by muscle mass loss, known as sarcopenia.

1.3 Sarcopenia & Lung Cancer

According to Pahor et al., sarcopenia is an age-related risk factor that leads to progressive loss of muscle mass and strength in aging individuals [4]. Furthermore, according to the Cambridge journal [3], sarcopenia is a malady which is not isolated from obesity but coexists with obesity in older patients. Medical researchers have found that, the decline of muscle mass begins with the age 30 in adults [5]. Sarcopenia is a severe illness, which intensifies with growing age by causing an increase in body fat and the decline in skeletal muscle. Consequently, middle-aged and older adults become more prone to develop respiratory diseases [6]. Although sarcopenic obesity is barely reflected in BMI, it can have a severe impact on lung cancer patient survival outcomes. Since the occurrence of morbidity in aging patients creates a diminishing cancer survival rate, the call for better understanding and treatment of this condition has become stronger. Further research in this field is required to improve understanding of the linkage between sarcopenia and lung cancer. Despite this known relationship between sarcopenia and cancer outcomes, there is still a dearth of knowledge about the connection between sarcopenia measures before and after cancer diagnoses. Currently, the medical community assesses sarcopenia by using Magnetic Resonance Imaging (MRI) or Computerized Tomography (CT) scans.

Medical researchers studying sarcopenia have determined that the best way to analyze sarcopenia is by analyzing segments of the lumbar vertebrae. Moreover, physicians have identified that muscle and fat areas are best represented in the third lumbar vertebra (L3) [7]. MRI and CT scans are the modality used to obtain L3 images. According to Pahor et al., MRI and CT are the golden standard for physicians to assess muscle and fat mass across sectional areas [4]. Both modalities provide accurate imaging

methods, which helps physicians to determine obesity rates in cancer patients. This project focuses on CT scan images as the sole modality for analysis considering that CT images are frequently obtained in cancer staging assessments.

1.4 CT and MRI Scans

Both the CT and MRI scan modalities provide a reliable and valid measure of muscle tissues. CT scans were first introduced in the year 1963 by Dr. William H. Oldendorf. Oldendorf's motivation came as he faced the difficulties of studying brain conditions without being able to perform medically invasive procedures. In order to make a better diagnostic tool, he was inspired by the engineers in his faculty who used X-rays to detect dehydrated regions in oranges. Consequently, based on this discovery, he developed a machine to mimic the approach. His invention was able to scan the head region through transmitted beams of X-rays. Moreover, it was also possible to reconstruct the radio density patterns of the head and brain using this invention [8]. Oldendorf's invention of the computerized axial tomography (CAT) scan, or CT scan, in short, has revolutionized the medical field and provided a giant leap in the technology available to future scientists to advance their research.

The first clinical CT scan was invented by Godfrey Hounsfield of the EMI Laboratories in England and by Allan Cormack of Tufts University in the United States in the year 1972. Soon after introducing the CT machine, Hounsfield and Cormack made the CT scanner available for commercial operations. Both scientists have massively improved the way medical images are taken. Consequently, these trailblazers were recognized with the Nobel prize in the year 1979 for their invention. According to the American Journal of Neuroradiology, over 72 million CT scans were performed in the United States in the year 2007 [9]. Overall, the invention of Dr William Oldendorf paved the way for more modern diagnostic imaging machines, like the modern CT scan and MRI. [8, 10, 11]

A CT scan measures different angles of the body and combines it in a series of X-ray images. Additionally, CT scans can also produce cross-sectional images, known as slices, of bones and soft tissues. The technology is capable of distinguishing different tissue types based on the image slices detected. In essence, the CT technology allows physicians and specialists to gain a view inside the body without having to perform surgery [12].

As shown in Figure 3, the CT scan machine uses multiple steps to obtain a full body image. Modern CT scans are equipped with a motorized table, which rotates the patient through a rounded opening of the CT scan machine. With every rotation, the X-ray beam passes through a section of the body. The images are created when the detectors on the opposite side of the X-ray source register the X-rays that are passing through the patient's body. During one complete rotation, the machine creates different "snapshots" of the beamed body part. For every rotation, the X-ray source and detector assembly collaborate to create a snapshot. This snapshot is rendered into an image, which is sent to a computer. Next, the computer reconstructs all individual "snapshots" into one or multiple cross-sectional images known as slices to obtain visibility of the internal tissues and organs, as can be seen in Figure 4 [10].

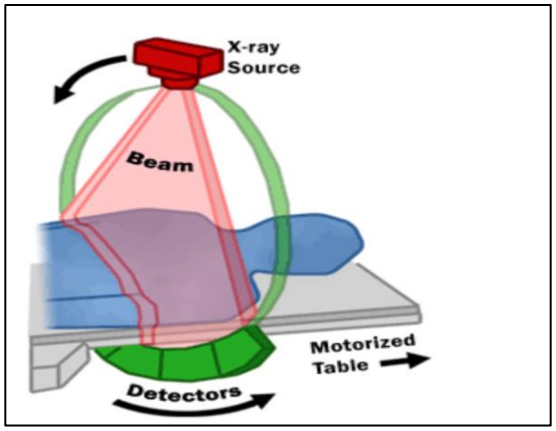

Figure 3: CT Scan Procedure

Figure 4: CT Scan Slice Image

CT images illustrate the scanned three-dimensional (3D) section of a patient's body over a two-dimensional (2D) map of pixels. Each CT image slice is represented by a row or rows of pixels on the 2D representation referenced above. Each of these slices contains 2D pixels and numbers, which correspond to locations on the overall body scan. CT scan images are standardized to have a pixel matrix of 512 × 512 on the scale. [5]

CT scans are expressed in Hounsfield units (HU), which is a relative quantitative measurement of radio density. The scale of the HU is used by radiologists and other specialists to interpret CT images [13, 14]. The scale was named after Sir Godfrey Newbold Hounsfield and was based on the oldest medical imaging machine, X-ray.

X-ray imaging uses radiation to take 2D full body images and the HU scale provides a way to read the images. The intention of the HU scale is to show radio density in X-ray images. CT scanning uses the same approach as X-ray images by using the HU scale values [15]. Overall, a CT scan is an X-ray machine that is connected to the computer, which uses HU values to read the scan.

The gray CT image shading, as codified in the HU scale, enables specialists reviewing CT images to differentiate fat from muscle or bone areas. Since body tissues and bones have different chemical characteristics, the HU scale applies numerical values from the range of −1000 to +1200 to the pixels representing these bodily structures. The measure of the HU density uses a linear scale density approach. The standard and reference points are defined by using water, which is set to have an HU value of 0 (0 HU). Adipose tissue, or fat, has values in the negative range from -190 to -30 HU. Muscle falls in the positive range from 0 to 100 HU. [5]

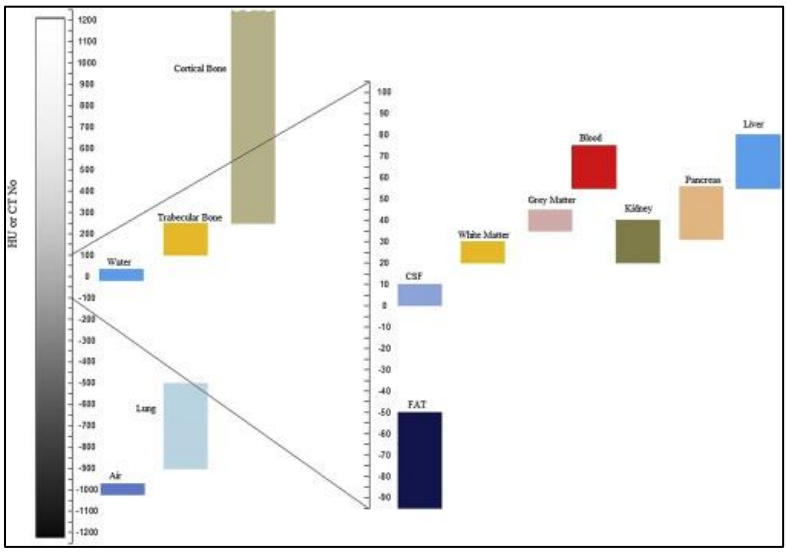

Figure 5: Hounsfield Scale

The images are shaded with a specific level of gray pixels, which enables the machine to distinguish and segment tissues. Adipose tissue is indicated by darker areas and muscle appears lighter in the images. The brightest contrast in the CT image is on bones.

The clinical analysis of this project is to automate L3 slice detections, which can assist physicians to compute obesity based on adipose tissues. Moreover, the adipose tissues are biomarkers that have shown strong associations with measures of sarcopenia. This means that a lower muscle density would indicate a higher fat content in elderly patients. Since the combination of sarcopenia and obesity in cancer patients is a critical measure of survival, this project aims to provide a tool that can detect from a CT scan the location of the L3 vertebra, the critical area needed to calculate obesity from the perspective of adipose content. The overall goal is to automate the process of L3 slice detections to help physicians and researchers in any field that require assessment of adipose tissue in the body, including the ongoing research to better understand sarcopenic obesity and the research to refute the “obesity paradox” by assessing the correlation between VFI and lung cancer outcomes.

Apart from CT scans, MRI scans are also a popular noninvasive modality that captures the composition of various tissues and organs in the body. An MRI is a tube-shaped machine, deploying strong magnetic fields and radio waves (radiofrequency energy) to create images of the whole body. The machine has the ability to temporarily rearrange water molecules in the patient’s body. The MRI technology works by distinguishing the different protons of fat and water molecules in the body. An MRI machine can send and receive radio waves by a transmitter or receiver. For each signal sent, a digital image of the scanned area of the body is created. Cross sectional images are created when radio waves signals are sent, which causes the alignment of atoms. [5]

While both CT scans and MRI scans are great tools to capture and analyze internal structures of the body, CT scans are more affordable and therefore more accessible to the general population. For this reason, this project will focus only on analyzing CT scans to provide a tool that can assist as many future patients as possible.

1.5 Problem Space

Pioneering studies have discovered that adverse health outcomes for lung cancer are correlated with worsening obesity. This correlation may be normal or inverse, depending on the risk factors for lung cancer included in the study as well as how obesity is measured, which is traditionally a variation of BMI. While certain risk factors of lung cancer and obesity as defined by BMI are inversely related, there is a dearth of research on whether lung cancer is correlated with obesity as defined by VFI. [16, 17, 18, 19, 20]

VFI is best calculated using adipose fat measurements observed at the abdominal region. The body composition at this region is best captured through images of the L3 vertebra. However, the connection between the L3 adipose and negative health outcome in lung cancer is poorly understood, in part due to a dearth of precise body composition measurements as collected from CT scan examinations. [16, 17, 18, 19, 20]

This dearth of data gleaned from body composition measurements as observed in CT scans is partially due to the time consuming and tedious nature of the manual work that is required to process the CT scan images to extract the body composition data. One way through which this data collection process can be improved is by employing Machine Learning (ML) techniques to automate the tedious manual labor. This can be achieved especially through deep learning techniques, which can help to identify the correct L3 region in the body scans where adipose tissue is well-defined.

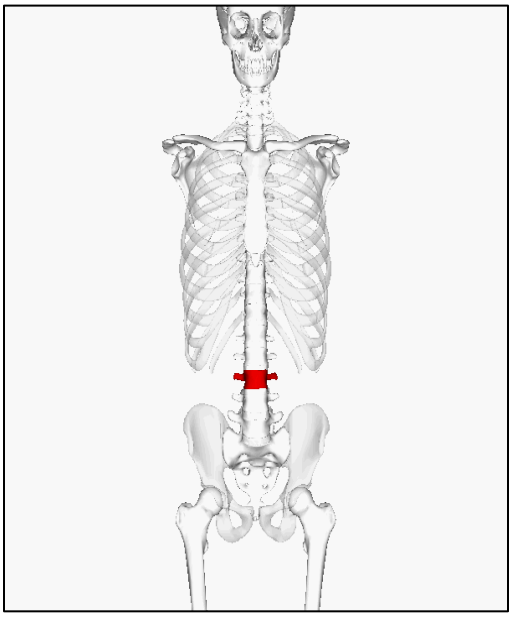

Figure 6: Lumbar Vertebra at the 3rd Position

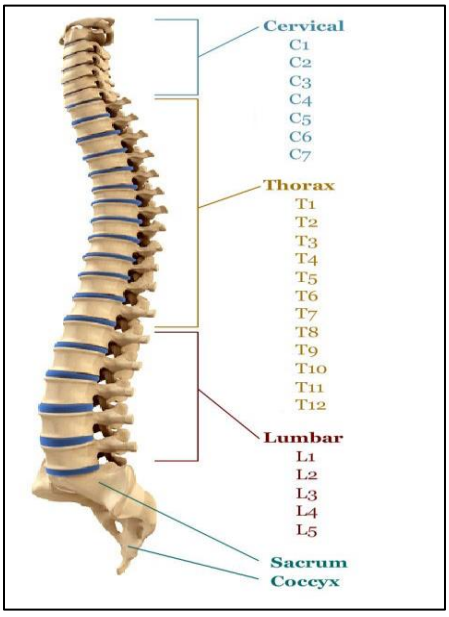

Figure 7: Spine Showing Vertebrae Structure

In summary, the relationship between metrics of obesity that are based on BMI and lung cancer is poorly understood, and research that demonstrates any correlation between VFI and lung cancer is sparse and so is the data that can be used to support any hypothesis regarding VFI and lung cancer [16, 17, 18, 19, 20]. This project will endeavor to provide valuable tools for physicians and scientists studying the relationship between obesity and lung cancer using imaging data from various cohorts of clinical

data derived from studies such as The Cancer Genome Atlas (TCGA) program and the Antibody directed against Programmed cell Death-1 ligand 1 (Anti-PD) immunotherapy study.

This study's focus is to utilize ML to read CT scans and to identify the location of the L3 vertebra at a cross sectional level. Utilizing the techniques proposed in this project, future research projects can build upon the results of the project to develop a second algorithm that will automatically calculate the VFI from the L3 vertebra image scans. The overall aim from such efforts is to research the relationship between adverse health outcomes for lung cancer and obesity, focusing on VFI and not BMI.

1.6 Research

There do exist studies that have attempted to automate the abdominal segmentation using CT scans. For example, the approach in Kanavati's 2018 research outlines the automation of L3 vertebra image slice detection using CT scans. However, this study has limitations with respect to identifying images from any CT scans that have body parts beyond just the abdominal section. [7]

In Kanavati's research, the semi-automatic approaches to detect L3 vertebra slice images function according to the sliding window approach, where convolutions from overlapping image areas are repeatedly recomputed. [7]

The training of the system is based on an elimination procedure, where the image is cropped by a fixed size. The output of this training results in a single number that indicates the predicted slice location, which represents the L3 slice for the training set. One drawback of this software that it is still time consuming since it requires refinement to obtain the L3 segment. [7]

The manual extraction of the L3 vertebra image slice can be accomplished by scrolling through the 3D image representing the frontal part of the human body scanned and identifying the L3 slice when it is reached. Semi-automated segmentation software such as Slice-O-Matic or ImageJ are used for this purpose and to perform further refinement, such as the identification of the VAT and SAT.

In Kanavati's report, the researchers have focused on the importance of context to distinguish the appearance of important features to get to the L3 vertebra image slice. The algorithm detects the L3 vertebra image slice using a system that allows Fully Convolutional Neural Networks (FCNN). [7]

In the past, the calculation of VAT and SAT has been difficult to perform accurately using only non-invasive procedures. However, several approaches have been tried and suggested over the last few years. Sarjoghian et al. developed a system using radio frequency (RF) to measure the SAT thickness in the abdominal region of humans, which they built upon to propose a method to estimate the visceral fat. [21]

Multifrequency bioelectrical impedance analysis (BIA) can assess the measure of body fat using electric signals to check for water content, which is a good indicator of fat in the body. This technique requires the usage of electrodes on the person's right hand and foot, through which a low-level imperceptible electrical current is sent across the body. This signal can be impeded by the different types of tissue in the body, including the amount of water, which is measured by the BIA device. The device determines that tissues containing large amounts of fluid and electrolytes, such as blood, have high conductivity,

whereas other structures, such as bone and fat, have low conductivity, which slows the electrical signal measured. This change in signal is measured as the resistance to the flow of current as it passes through the body, thereby providing estimates of body water from which body fat is then calculated. [14]

Azpiroz-Leehan et al. compared the BIA technique to estimate VAT with the measurement of VAT using MRI. They concluded that these two techniques are not comparable with each other because the BIA technique was found to be susceptible to the influence of SAT and was not a good indicator of VAT. [22]

1.7 Solution Space

An ML algorithm can be used to accelerate the process of the L3 vertebra image slice detection using a provided chest or abdominal or full-body CT scan. Considering that the detection is extremely critical for making diagnoses by physicians, it is imperative that the algorithm perform at close to 100% precision and correctly identify the L3 vertebra slice image every time it makes a positive detection decision. These detected images can be used by the physicians and researchers alike to perform VFI calculation.

Once the L3 vertebra image slice detection algorithm is tested and validated against the provided annotated CT scans, it can be implemented against larger datasets to which a physician or researcher might have access, thereby processing potentially thousands of available images in a batch mode. The algorithm can also be used to process data for just one patient at a time.

In summary, the proposed algorithm should be able to:

- Process any CT scans of the human body, such as chest CT, abdominal CT, and full-body CT scans;
- Accurately identify the L3 slice from the provided CT scan;
- Maintain a close to 100% precision for high-confidence detections by:
    - Providing an opportunity to review results for any low-confidence detections; and
    - Allowing the user to set confidence thresholds for detections.

The project aims to utilize a deep learning approach to achieve the above algorithm goals. However, there are a few challenges that can occur during the analysis of the CT scans. The quality of the provided image is a big factor that affects the quality of the analysis. Bad image quality could have a negative impact on the ability of the algorithm to correctly detect the L3 vertebra image slice and result in low-confidence detections that would need to be manually reviewed. Differing image quality can arise from a plethora of factors, including the settings of the CT machine used by the lab technician to perform a CT scan of a patient, which can differ from one medical institution to the next. The medical diagnosis which required a physician to prescribe a CT scan for a patient may also affect the protocols followed for obtaining the CT scan images and can lead to CT images with varying quality.

Another factor that could affect the analysis is the natural variation in the objects detected in CT scans. For example, the composition of skeletal muscles differs from one patient to the next. There are other factors that can lead to differing skeletal muscle composition as well, such as age, diseases or degree of obesity in the patient. All of these factors can negatively affect the analysis.

Additionally, there are other issues that can prove to be impediments for a robust analysis which might occur with the data processing. These may include factors such as the body parts that are included in a CT scan or any extraneous non-human objects that may show up in the CT scans that can cause the algorithm to return low-confidence outputs. Despite these challenges, the proposed algorithm for L3 vertebra image slice detection should be able to bypass all irrelevant objects and return detections with high precision.

1.8 Project Objectives

L3 vertebra image slices can be detected from CT scans by manual segmentation. This is a very tedious and time-consuming process. The level of effort required to perform this task means that it is not a feasible approach for evaluating bulk quantities of CT scans. Therefore, a different approach must be undertaken to make this task more feasible.

With this in mind, the project will attempt to automate the process of detecting L3 vertebra image slices from CT scans. The algorithm thus created will be able to reduce the detection time for L3 vertebra image slices and free up manual resources for other tasks. The algorithm should also be able to return the detections with a very level of precision, ideally 100% precision, and be able to accommodate variations in the input CT scans, such as in terms of quality or body parts included in the CT scans.

In summary, through this project, the project team hopes to successfully implement an image-recognition ML model that can identify and return the L3 vertebra image slice with high precision.

The team hopes to gain a better understanding of the kind of problems that data science can help tackle in the medical domain. Cancer research is an umbrella term that encompasses all the various efforts being made to alleviate some of the challenges that medical professionals face. This project also falls within this bracket as the VFI computed at the L3 vertebra is an important marker for obesity and can be used to study certain lung cancer outcomes. There are other health conditions as well for which VAT is an important factor. The project team hopes to gain a better understanding of the spheres of the medical field for which the automation of L3 vertebra image slice detection could be useful. In overall terms, this project affords an opportunity for the project team to experience real-world implementation of data science concepts and practices learnt in classroom settings.

Another objective of this project is to learn to work in an agile development environment. Agile has been gaining traction over the past several years and is now the preferred methodology among development teams in various industries.

A successful implementation of this project would provide a solution for automating L3 scan identification by reviewing CT scan images. This solution will add a lot of value as it will speed-up the VFI calculations and will free-up resources in the medical domain that would otherwise be required to manually accomplish this task. The solution could also be useful for researchers to improve upon or adapt for their purposes. There may also be use cases for this research in other domains outside healthcare. For example, the performance of athletes and other professionals in sports could be analyzed in terms of visceral fat or overall fat using the VFI scores that can be computed from the L3 vertebra image slices returned by this algorithm.

1.9 Primary User Stories

This project will leverage the use of ML in the healthcare or medical domain. It will improve the accuracy of L3 slice vertebra image slice detections that physicians and/or researchers can rely upon for various purposes, such as the computation of VFI used for diagnosis or to forecast lung cancer treatment outcome or as input indicators for another research. Based on the user context and value proposition, the following user stories have been defined to guide the project:

- A Physician can use this solution to extract the L3 vertebra image slice from CT scans for their patients. The current followed by physicians is time-consuming and performed manually.
- A Researcher might use this solution to process a batch of CT scans and extract the L3 vertebra image slice detections as well as other accompanying details that inform the output, such as the detected L3 vertebra slice position, automatically.

1.10 Product Vision

The algorithm will benefit the medical field and healthcare industry, both directly through applications by physicians and indirectly through applications by researchers. The primary scenario in which our tool could be implemented is in a hospital or physician's office where the doctor or assistant needs to identify the L3 image slice from a CT scan quickly. By utilizing our tool, the healthcare provider will be able to spend less time looking for the L3 slice from the CT Scan and dedicate more time to perform other functions, such as providing medical diagnoses. It will also be possible for researchers and research labs to use the algorithm in batch mode to identify the L3 vertebra image slice from potentially large datasets, such as those including thousands of CT scans. In this application, researchers can prepare and catalogue datasets more efficiently to have the results ready for further study.

1.11 List of Terms: Acronyms

Table 1: List of Acronyms

Acronym	Full Form
AI	Artificial Intelligence
Anti PD 1	Antibody directed against Programmed cell Death-1 ligand 1
AI	Artificial Intelligence
BIA	Bioelectrical Impedance Analysis
BMI	Body Mass Index
CDAS	Cancer Data Access System
CDC	Center for Disease Control
CNN	Convolutional Neural Network

Acronym	Full Form
CT or CAT	Computerized Tomography
DICOM or DCM	Digital Imaging and Communications in Medicine
FCNN	Fully Convolutional Neural Networks
GPU	Graphical Processing Unit
HIPAA	Health Information Portability and Accountability Act
HU	Hounsfield Unit
L3	Third Lumbar
LUSC	Lung Squamous Cell Carcinoma
LUAD	Lung Adenocarcinoma
MIP	Maximal Intensity Projection
ML	Machine Learning
MRI	Magnetic Resonance Imaging
NBIA	National Biomedical Image Archive
NCI	National Cancer Institute
NIfTI	Neuroimaging Informatics Technology Initiative
NLST	National Lung Screening Trials
NSCLC	Non-Small Cell Lung Cancer
RF	Radio Frequency
PHI	Protected Health Information
SAT	Subcutaneous Adipose Tissue
SFA	Subcutaneous Fat Area
SOP	Study of Patient
TCGA	The Cancer Genome Atlas
TCIA	The Cancer Image Archive
TPU	Tensor Processing Unit
VAT	Visceral Adipose Tissue
VFA	Visceral Fat Area
VFI	Visceral Fat Index

Acronym	Full Form
WHO	World Health Organization

2. Data

2.1 Data Sources

The original NLST dataset is available through the Cancer Data Access System (CDAS). The CDAS is administered by the National Cancer Institute (NCI) [23]. The NCI has developed many collaborative efforts between researchers to promote shareable data exchange in hopes of a better understanding of the disease. In order to gain a better understanding of cancer, the NCI has developed a software environment that enables archiving of cancer imaging data. This software is hosted on the NCI website. It provides a publicly available open source that supports submission and distribution of cancer image data. [24]

This web site has gained a lot of traction and interest over the years, and, consequently, it has now expanded and hosts medical data beyond just the cancer images. Therefore, the website has been rebranded as the National Biomedical Imaging Archive (NBIA). The NBIA software is available through the NCI Center for Bioinformatics. Continuing to build upon their resources and driven by their need for a better understanding of genetic materials, the NCI contracted the Washington University in Saint Louis to create The Cancer Imaging Archive (TCIA), which is an open-source and open-access information resource, providing an advanced medical imaging portal of cancer patients. TCIA was built with the intent to support research and to promote educational initiatives to better understand and battle chronic diseases. [25]

When a user submits images to the TCIA database, TCIA checks the images for quality and privacy aspects. Privacy is a big concern when it comes to patient images. DICOM is widely used as the imaging format for the creation, transmission, and storage of digital medical image data. It is also the most common storage method for clinical imaging data.

The advantage of DICOM is that it offers a defined data dictionary and structure. DICOM compresses the data in a file format that enables a client server service. All images are checked by a “TagSniffer”, which provides a report about scanner manufacturer/model, and private tags attached to the DICOM images. Information that contains Protected Health Information (PHI) is removed by the tags and replaced by blanks. DICOM file is structured such that Study of Patient (SOP) instance, it contains a file header with metadata called the File Meta Information and the DICOM dataset for that particular SOP instance. [26]

As part of further processing by TCIA during the DICOM image intake process, before the DICOM images appear on the public server, a script runs that uniformly assigns a DICOM-header date by a day-interval assigned by TCIA.

This project primarily uses data obtained from the TCIA for all purposes. Various search filters were used to obtain the datasets used in this study. These search filters were used to obtain data from the Cancer

Genome Atlas Lung Squamous Cell Carcinoma study, the Cancer Genome Atlas Lung Adenocarcinoma study, and the Anti-PD 1 Immunotherapy study. For each selected study, the modality CT scan was selected to restrict the type of medical images processed to only CT scans. The anatomical areas included in the CT scans differed based on the datasets; they could cover either the chest or abdominal area specifically, or they could be full-body CT scans, covering other body parts as well, such as the legs and the skull.

The NBIA Data Retriever is a tool that is required to enable the acquisition of data from the TCIA website. For the datasets selected for download from TCIA, a manifest file is created by querying and filtering through the available datasets by using the filters defined above. The NBIA Data Retriever tool is able to read this manifest file as input and allows the user to download the selected medical image data. For this study, the data downloaded in this included CT scans in the DICOM format. The aggregate size of all the datasets thus acquired was 907 DCM files for the CT scans.

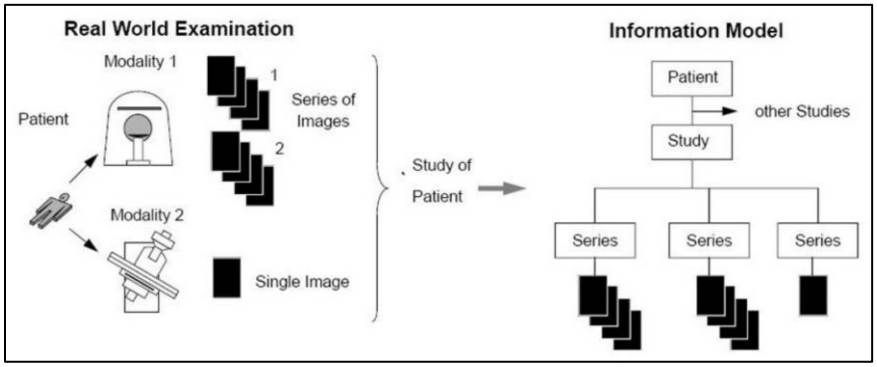

Figure 8: DICOM Model

2.2 DICOM Model

There is a particular hierarchy that is used to store and define the information model used by DICOM images. At the top of the hierarchy is the patient. Each patient can have multiple studies, which is also known in the medical world as exams or procedures. Next, each study may consist of multiple series of data. A series is equivalent to a specific type of data, and in the medical field it is also known as a modality of data acquired on a specific device. Therefore, in the context of this project, each series is one complete CT scan. This information model is described in Figure 8.

Each series of data may contain multiple DICOM object instances, which are images or reports of the patient at different positions of the body. Overall, all patient related information is stored in each DICOM object of a study. DICOM files enable processing and encoding of 2D and 3D axial DICOM slices.

The metadata of the DICOM files is a structured file, where the images are stores. The metadata provides information about the patient, the study, the series, and the general equipment settings from where the image was taken. It is important to note that the Health Information Portability and Accountability Act (HIPAA) regulates how medical data is managed and greatly restricts the kind of patient related data, especially PHI, that can be shared publicly. Therefore, the data provided by the

TCIA website scraps the PHI from their website and replaces the patient ID information with a generic Subject ID.

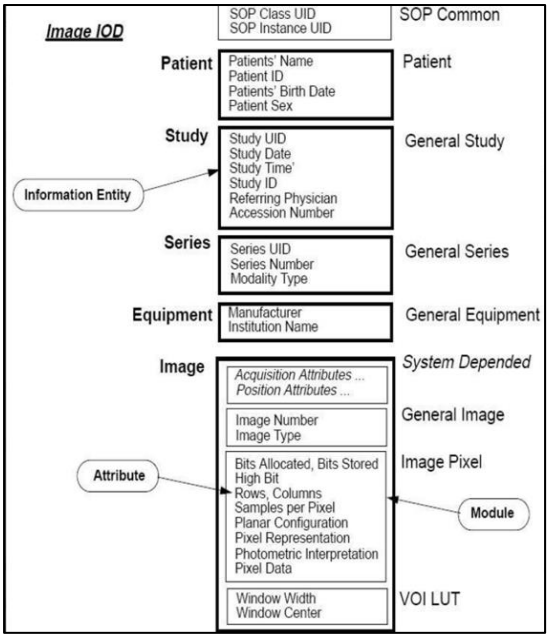

Figure 9: DICOM Metadata Structure

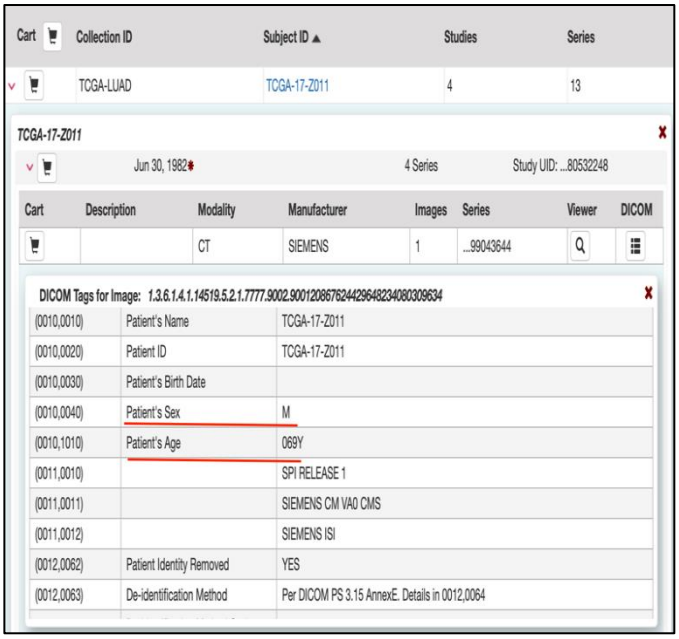

Figure 10: Metadata in TCIA Datasets

2.3 Data

The data used for this project came primarily from the publicly available datasets downloaded from the TCIA website. The datasets were used for different purposes as the project evolved. These details are listed in Table 2.

The datasets marked for training were used to run the algorithm to get initial results and make improvement. The validation dataset was used by the Project Sponsor Allwyn to run the algorithm at their end. The algorithm was optimized based on the results observed in both the training and validation datasets. Finally, after making all the optimization related changes to the algorithm, the final run of the algorithm was made against the test dataset, which was selected by Allwyn. The final results of the algorithm are reported against the test data.

Table 2: Data Sources

Data Source	Number of Series	Purpose
TCGA Lung Squamous Cell Carcinoma (LUSC) & Lung Adenocarcinoma (LUAD)	336	Train
Anti-PD 1 Lung	571	Train
Anti-PD 1 Lung	30	Validation
OmniSeq	49	Test
Anti-PD 1 Lung	3	Test
Anti-PD 1 Lung	30	Test

Out of all the datasets mentioned in Table 2, the Anti-PD 1 Lung dataset is the biggest. It has a size of 55 Giga Bytes (GB) and contains CT scan series for 46 patients. The studies in this dataset are divided over 571 series. Due to the large size of this dataset, only a sample dataset consisting of 30 CT scan images was selected as the validation data. The selection of these 30 CT scan series was made by the physician at the Roswell Park Cancer Center who collaborated with Allwyn for this project. The sample Anti-PD 1 Lung data has a size of 6 GB and includes data for 38 patients in the study.

2.4 Field Descriptions

The data is a collection of CT Scan images. These images are in the DICOM format. This format contains metadata about the images in the form of tags, where each tag represents an attribute or data element describing an aspect of the DICOM image. DICOM images support the storage over 4,100 different tags, indicating that the metadata captured for these images can be very diverse and comprehensive.

However, it is not necessary that each DICOM image would have a value for all the tags that can potentially be available. One of the reasons for why this may be is that many of these tags have now been retired and may not be available in newer datasets of DICOM images.

The data used for this project had a maximum of over 96 tags that could be available from the metadata. Table 3 below lists these tags.

Table 3: List of DICOM Metadata Tags

DICOM Metadata Fields/Tags Names	
AccessionNumber	ModifiedImageDate
AcquisitionDate	ModifiedImageID
AcquisitionNumber	ModifiedImageTime
AcquisitionTime	OriginalImageIdentification
BitsAllocated	OriginalImageIdentificationNomenclature
BitsStored	PatientAge
BodyPartExamined	PatientBirthDate
BurnedInAnnotation	PatientID
Columns	PatientIdentityRemoved
CompressionCode	PatientName
ContentDate	PatientPosition
ContentTime	PatientSex
ConvolutionKernel	PhotometricInterpretation
DataCollectionDiameter	PixelRepresentation
DataSetSubtype	PixelSpacing
DataSetType	PositionReferenceIndicator
DateOfLastCalibration	ReferringPhysicianName
DeidentificationMethod	RequestedProcedurePriority
DeviceSerialNumber	RescaleIntercept
DistanceSourceToDetector	RescaleSlope
DistanceSourceToPatient	RotationDirection
Exposure	Rows
ExposureTime	SOPClassUID

DICOM Metadata Fields/Tags Names	
FocalSpots	SOPInstanceUID
FrameOfReferenceUID	SamplesPerPixel
GantryDetectorTilt	ScanningSequence
GeneratorPower	SeriesDate
HighBit	SeriesInstanceUID
ImageDimensions	SeriesNumber
ImageFormat	SeriesTime
ImageGeometryType	SliceLocation
ImageLocation	SliceThickness
ImageOrientation	SoftwareVersions
ImageOrientationPatient	SourceImageIDs
ImagePosition	SpecificCharacterSet
ImagePositionPatient	StationName
ImageType	StudyComments
InstanceNumber	StudyDate
KVP	StudyID
Location	StudyInstanceUID
LongitudinalTemporalInformationModified	StudyTime
ManipulatedImage	TableHeight
Manufacturer	TimeOfLastCalibration
ManufacturerModelName	WindowCenter
MaskingImage	WindowWidth
Modality	XRayTubeCurrent

2.5 Data Context

Multiple DICOM objects are contained in each series of a CT scan. Each DICOM object represents one slice of the CT scan, representing a 2D image of a certain part of the body that has been scanned. To get a more complete understanding of how each individual CT scan slice is related to another, it is important to understand the hierarchy in which these scans are organized. Each complete series of CT scan slices is contained within a subfolder with each DICOM CT scan slice named in an ascending numbered order. This order contextualizes each CT scan image and helps the algorithm process each series separately by stacking all the individual CT scan slices one on top of the other in order to create a 3D representation of the body parts that were scanned.

The other important context of note with the data is the study for which these images were created. The data belongs to the following studies: TCGA-LUSC, TCGA-LUAD, Anti-PD 1 Lung Immunotherapy, and OmniSeq.

Each of these studies is related to studying different aspects of lung cancer or studying the effectiveness of certain treatments for lung cancer. These studies were chosen because the likelihood of finding usable CT scan images of the L3 vertebra was determined to be high. Whenever a CT scan of the chest is taken, it is very likely that the abdominal area is also scanned due to its proximity with the chest. Furthermore, after analyzing the data obtained, it was found that the L3 vertebra was consistently included in the scans and that the scans extended to full body scans in certain cases as well.

2.6 Data Conditioning

For the purpose of this project, all the data was manually downloaded from The Cancer Imaging Archive (TCIA) website using the NBIA Data Retriever. The TCIA website allows a user to select the data required for download and creates a manifest file for the same. In computing, a manifest file is a file that contains metadata for a group of accompanying files that are part of a set or coherent unit. The NBIA Data Retriever tool takes the manifest file as the input and uses it to download the required data from the TCIA website. After downloading the data, it was placed on a Google Drive Cloud that has been assigned for this project.

The CT scan datasets used for this project are more than 60 GB in size. This size is difficult to accommodate on personal devices and non-enterprise cloud infrastructure. In order to accommodate these datasets, the Google Drive storage capacity was increased to 200 GB.

Furthermore, the analysis platform used for the project is a Google Collaboratory environment, which allows the usage of virtually enabled Graphic Processing Units (GPUs) and Tensor Processing Units (TPUs) on a Jupyter Notebook based platform. The GPU capabilities were necessary for this project due to the extensive size of the dataset as well as the requirement to run memory hungry algorithms.

2.7 Data Quality Assessment

After conditioning the data, it was possible to perform data quality analysis. The image quality was found to be good generally, with some exceptions. The TCGA-LUSC and TCGA-LUAD datasets were also analyzed to see the distribution of gender and age among the datasets. This is shared in Figures 11 and 12 below.

From Figure 11, we can see that there are more men than women participating in the study. The dataset has 162 females and 175 males. However, the difference in the numbers is not large enough to say that the dataset is skewed towards males. The overall distribution is quite even, and both the genders are represented well in the data.

From Figure 12, we can see that the age of the participants in the TCGA studies ranges from 45 to 85. The largest age group is between 57 and 73 in the dataset. Overall, the membership in the various age brackets seems to increase as the age brackets are observed in an ascending order. The highest frequencies are observed for the age brackets representing the ages between 57 and 72. On either side of 57 and 72, the age brackets have considerably lower membership, indicating that the study focuses on patients aged in their late 50s and early 70s.

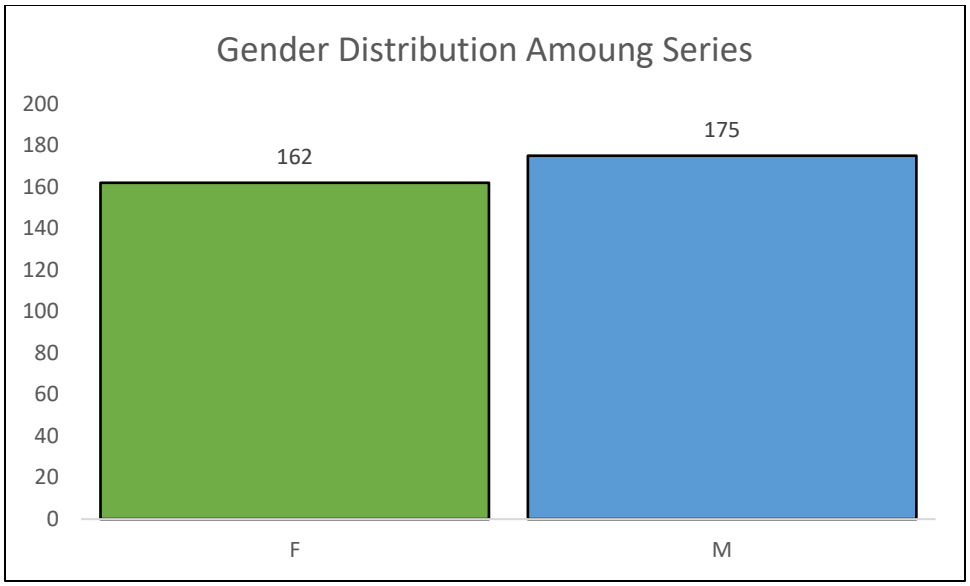

Figure 11: Gender Distribution in TCGA Datasets

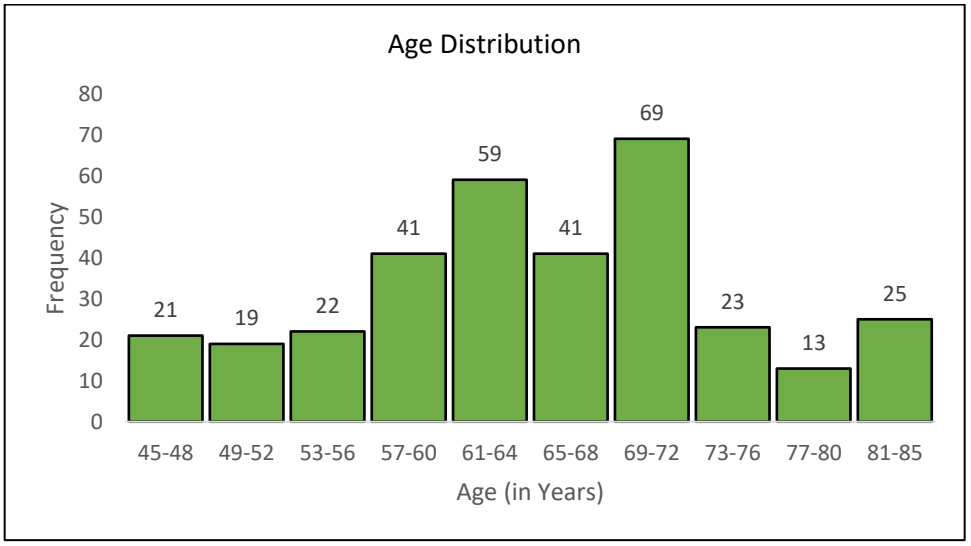

Figure 12: Age Distribution in TCGA Datasets

The files were converted to the Neuroimaging Informatics Technology Initiative (NIfTI) file format prior to processing them through the L3 vertebra image slice detection algorithm. The training results revealed several facts about the quality of the images processed.

Successful outputs produced a correct L3 detection and corresponding L3 vertebra slice image. Certain CT scan series were found to be of relatively inferior quality; for these series, the algorithm failed to produce an image output.

During the training process, the CT series in the TCGA datasets were converted to 98 NIfTI files, out of which 56 were successfully analyzed. Furthermore, 10 files produced the L3 slice with a confidence probability of over 10% (from a cumulative probability spread across all slices). 46 NIfTI files did not produce an L3 vertebra image slice because none of the slices breached the 10% confidence threshold.

The poor image quality observed might be caused by the fact that the TCGA CT scan datasets are dated; they belong to studies conducted in the mid-1980s and their quality is restricted to match the CT technology provided at the time.

The conversion of DICOM format images into NIfTI file format resulted in loss of data, as observed through the truncation of the metadata fields. However, the overall image quality and quantity remained the same. For analyzing the data, the metadata can be accessed through the original format.

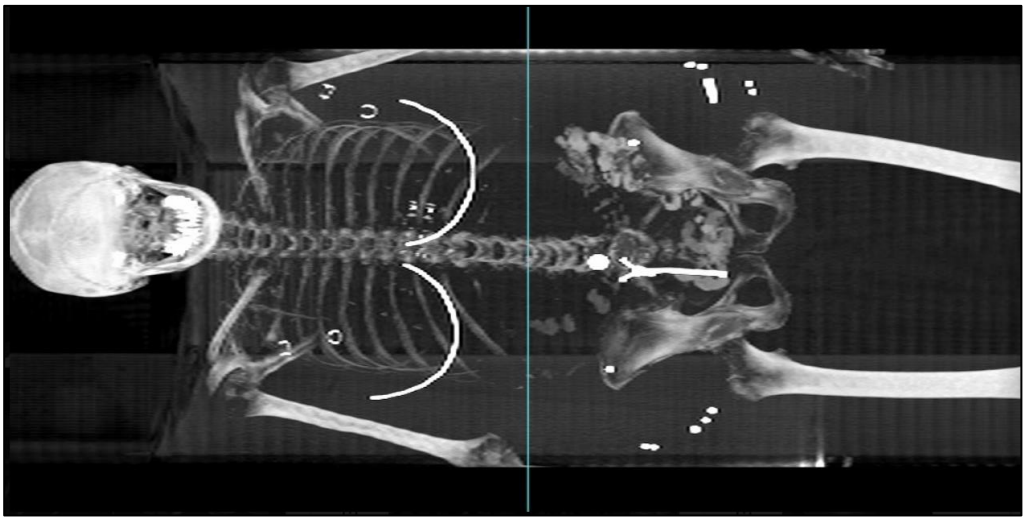

Figure 13: Metal Objects in CT Scan

As can be seen in Figure 13, image quality was also affected by extraneous objects that were included in a number of CT scans. In order not to affect the image quality, patients are usually asked to remove any metal objects, such as earrings, belts, and jewelry before undergoing the CT scan procedure. Metallic objects can interfere with the image quality of the CT scans as they show up as bright white spots and make it impossible to view the parts of the body over which they are placed. This is because metal objects are visibly distinguishable on X-ray images, which is what are used to create CT scans. Since metal objects are extremely dense, the image shines brighter on the foreign objects. Radiation does not penetrate these objects as well as it does the soft tissues. This explains why bones appear bright on an X-ray and subsequently the CT scans. Bones are denser than blood, cartilage, and soft organs. [27, 28]

This could be problematic for the algorithm not only when picking the L3 vertebra slice, but also for any further analysis that is performed using the detected L3 vertebra image slices, such as for computing VFI. For this reason, metal and other foreign objects can be problematic for L3 vertebra image slice detection. They can lower the confidence with which the algorithm makes a decision regarding the position of the detected L3 vertebra image slice and may end up requiring manual review.

It is also important to keep in mind that several series of CT scans present in the various datasets were not considered for processing due to their structure. These series either do not contain the L3 vertebra image slice or are too small in terms of the number of DICOM objects contained within each series to make them viable for analysis.

3. Analytics and Algorithms

3.1 Algorithm: L3 Vertebra Image Slice Detection

The goal for this algorithm is to detect the L3 vertebra image slice from a 3D CT scan. It should be able to accurately detect the L3 vertebra slice image from all the other content in the DICOM files that make up a series of CT scan for a particular patient or subject

The algorithm that has been used in this project is based on a trained model provided in the GitHub code for Kanavati's research [7]. The TCGA datasets were initially run on this algorithm without making any changes to set a benchmark. The accuracy as observed from the output returned was assessed, and subsequent training and validation of this algorithm was performed on iterations of optimizations applied over the algorithm to improve the evaluation metrics for the algorithm. As the project progressed, several rounds of validation and a final round of testing were performed to evaluate the algorithm after making all the modifications deemed necessary to get the best results.

The algorithm is expected to produce two outputs: the detected L3 vertebra image slice position and a confidence level for this detection. The L3 vertebra slice detections and associated confidence levels need to be highly accurate to be considered acceptable for use by a physician or researcher. Therefore, the algorithm must perform at an accuracy level with a precision score of 100% and an acceptable recall rate. What this means is that the algorithm should not include any false positives in its output. The algorithm should behave in such a way that whenever it assigns a "Yes" confidence flag to a detection, it must be correct.

The original model created by Kanavati et. al [7] could not guarantee such a strict accuracy and precision requirement. It has, therefore, been modified to adhere to the accuracy requirements. These modifications and improvements are presented later in the report.

3.2 Preprocessing of Data: DICOM & NIfTI

The CT scans included in the various datasets used throughout this project are composed of individual slice images in the DICOM format. Some of these images were not part of a viable series due to lack of a sustainable minimum number of slices that could be considered worthy of analysis. These images were, therefore, unusable for the project and were removed from analysis.

To remove such unusable CT scan series from consideration, a different imaging format has been used. The NIfTI format is an alternative medical imaging format, which is widespread in the neurological medical field. NIfTI is a simpler image format than DICOM, thus enabling a more rapid analysis process. Overall, NIfTI is an improved version to analyze medical imaging. Furthermore, the NIfTI file format is also simpler to use than the relatively older medical file format DICOM. This is because NIfTI enables users to store medical images in a single file instead of having a large tree of medical subfolders.

Overall, due to the differences between DICOM and NIfTI, it can be stated that the NIfTI format has several advantages over DICOM and it makes it easier and faster for users to load and process medical image data.

In summary, while this project makes use of DICOM images from the TCIA website as the source of data these images are not used in this form and are first converted into the NIfTI format for ease of processing. Consequently, NIfTI files are loaded into the model to perform L3 vertebra image slice detection. The output of the detection consists of 2 image files in the commonly used JPEG format, a frontal image showing the overall areas of the body included in the CT scan series and a slice image showing a 2D representation of the L3 vertebra slice detected by the algorithm.

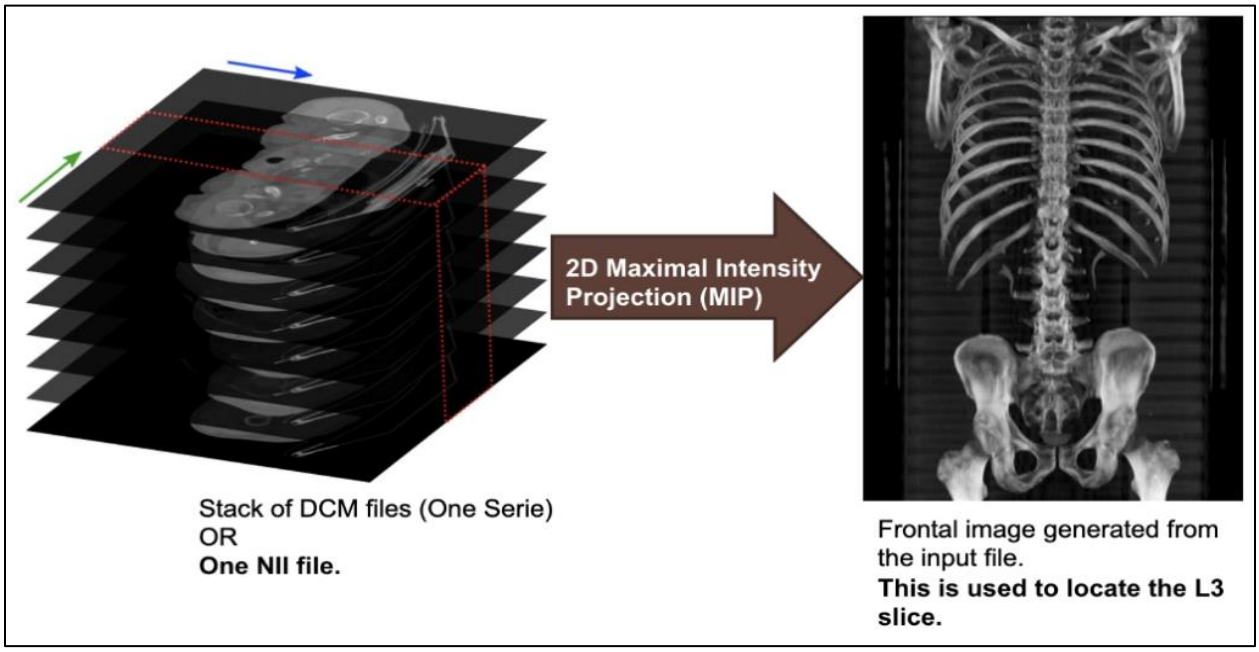

Figure 14: Generation of Frontal Image

The algorithm stacks the DCM slices used to create the NIfTI file together to create a frontal image of the CT scan series. This process is depicted in Figure 14. The frontal image thus produced is used to detect the location of the L3 vertebra image slice.

3.3 Setting Up Resources

The algorithm requires a computer resource environment of a Linux OS with a GPU enabler. GPU is preferred for this project since it meets all computing demands of the GitHub repository used as the starting point for this project [7]. Furthermore, GPU processors have become more common in the ML and artificial intelligence (AI) areas. GPUs offer enough computing power to break down multiple tasks and still ensure fast processing at once [27].

The project team is using the Google Colab platform to create the python environment and leverage the existing code available on GitHub from Kanavati's research [7]. The Google Drive platform is utilized for

data storage needs as it can serve the required storage capacity at a low cost for the entire duration of the project.

3.4 How the L3 Vertebra Image Slice Detection Algorithm Works

The CNN is based on a model of the U-Net that produces 1D confidence maps. The authors of the study decided to convert the 3D images into 2D to reduce the dimension and to obtain a clear view of the images. The image reduction was done via Maximal Intensity Projection (MIP). The 2D MIP images are used as input to the network. The end result of the network is the creation of 2D confidence map output. The obtained output in the confidence map provides possible confidence for the L3 location. Overall, the confidence map prediction is a good method for a fully automated L3 segment detection.

The measurement of visceral fat using CT scans is a common approach among practicing physicians and researchers. However, the automation of this task, which is the underlying goal of this project, is a novel idea.

In this paper, the algorithm framework utilizes CNN image classifications. CNN is the main tool used to perform image recognition and images classification. CNN is a deep learning model aids with analyzing images.

CNN emerged from the study of the brain's visual cortex, and it has been used in image recognition since the 1980s. In the last few years, thanks to the increase in computational power, the amount of available training data, software capabilities, CNNs have managed to achieve superhuman performance on some complex visual tasks. Moreover, CNNs are not restricted to visual perception: they are also successful at other tasks, such as voice recognition or natural language processing (NLP). [23]

CNN works in the following way as shown in Figures 15, 16, and 17. It uses an architecture of hidden layers to map the input from one layer, process it, and send an output to the next layer, which becomes its input, and so on.

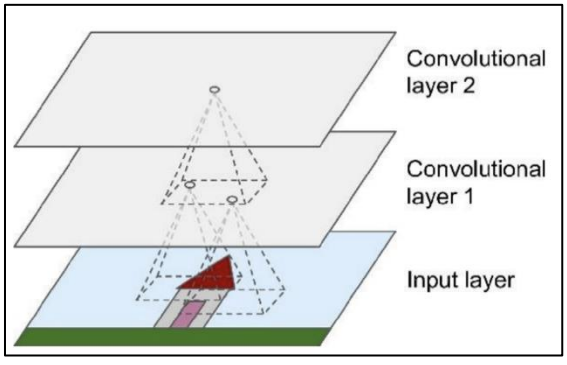

Figure 15: Layers in a Convolutional Neural Network

In the case of our project, each group of pixels is processed and produces an output for the following layer. This involves a significant amount of computation for images. In fact, each CT volume can have in average approximately 230 slice images. As discussed, each slice image has 512 pixels x 512 pixels. Parallel programming is what makes such computation possible in a reasonable time. Looking at the

convolution operation, we notice that they are independent at any layer level. So, it is possible and highly recommended that they are performed in a GPU environment instead of sequentially. For that purpose, we are using Google Colab which make it easier to have access to such material.

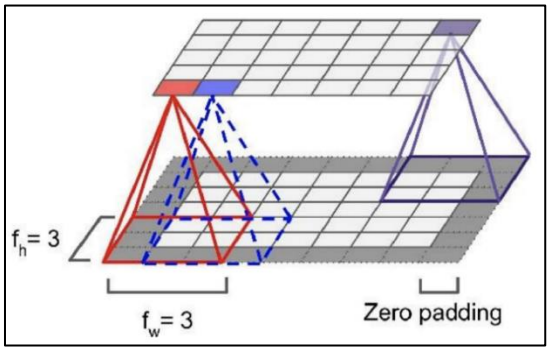

Figure 16: Mapping Output from Layer to Layer

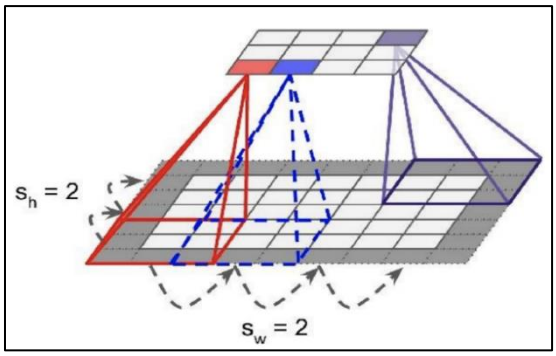

Figure 17: Mapping Output from Layer to Layer

3.5 Context of L3 Vertebra Image Slices Detection

To detect the L3 slice, the algorithm creates a 2D Maximal Intensity Projection (MIP) image like illustrated in Figure 14. So instead of trying to classify each slice as L3 or not, from the hundreds of slices in the 3D image, the algorithm creates a 2D Frontal image of the body. From that, a contextual detection is possible since the algorithm can detect the L3 vertebra based on other surrounding characteristics. A dataset of MIP has been used to train the model used in our solution.

3.6 Adjustments Made to the Original L3 Vertebra Detection Model

One major issue encountered during the validation phase was that the images used in validation set are full body images and therefore include the subject's skull and legs. This explained why the model initially did not perform as well as expected. To fix this issue, the team conducted an analysis of the correct and incorrect L3 slice detections. It was clear that the correct predictions fell within a predictable location range of the body. A solution, therefore, was put in place to influence the algorithm to run in such a fashion that whenever a position was predicted outside of the identified standard range, the algorithm would ignore it and consider it unlikely to be an L3 slice. The algorithm will rather consider the slice within that range that is most likely to have an L3.

Our purpose is to classify the CT Scans into two buckets based on the confidence of the solution in the result:

- Automatic Detection: Detect L3 accurately 100% of the times.
- Requires Manual Review: Suggest an L3 slice position but the output is flagged for human review.

The “Yes” or “Review” labels attributed by the algorithm to a particular CT scan series is done by the analysis of the probability density associated with all the slices in the CT scan series. The slice with the highest probability density is assigned the detection of L3 vertebra. If the probability density associated with the highest probability density slice breaches a particular probability threshold, then the algorithm designates this particular slice with a “Yes” label for L3 vertebra image slice detection. If the probability density threshold is not breached, then the algorithm assigns the “Review” to the detection so that manual inspection of the particular case could be performed.

To obtain detections with 100% precision, which is desirable for obtaining the most reliable results, the value of the probability threshold should be conservative. It can be influenced by the quality of the dataset that is run against the algorithm. To allow flexibility to the users, the probability density threshold value has been parameterized so that the users can set a value that is best according to their judgment. As will be shown in the Findings section, the threshold value can be adjusted to include more correctly identified L3 vertebra image slices within the “Yes” bucket. If a conservative number for this parameter is used, then it is likely that a lot of the correctly identified L3 vertebra image slices would be marked with the “Review” label. On the flip side, if a lower value for the probability threshold is used, then it is possible to include the L3 vertebra slice images that were labeled in the previous case as “Review” to be now correctly placed in the “Yes” bucket. However, if the probability threshold is lowed by a large number, then it is likely going to include misidentified L3 vertebra detections within the “Yes” bucket. This will cause the precision score to drop from the required 100% for most reliability. This is an undesirable situation as it would lead to the inclusion of False Positive outcomes in the outputs. Therefore, it is very important to strike the right balance to get the probability threshold values. Using conservative values are recommended to achieve 100% precision scores for the outcomes.

Another adjustment that was made to the algorithm was in terms of the area of a CT scan that can be searched by the algorithm to make a detection for the L3 vertebra slice image. As the L3 vertebra is typically found in the abdominal section of the body, which is in the center or core of the body, it is likely going to be represented spatially in the center of an MIP of a CT scan series, representing the frontal image of the body scanned. Therefore, to reduce the likelihood that the algorithm could pick a slice of CT scan in the extremities of the body as the L3 vertebra slice, the search area within which the detection could be made to restrict it only to the middle portion of the CT scans.

3.7 Algorithm: Manual L3 Vertebra Detection and VFI Computation

While the architecture and algorithm adjustments presented thus far incorporate the automated L3 vertebra slice image detection scheme, it is also important to understand how the physicians currently perform this procedure manually. After finding the L3 vertebra slice from a stack of CT scan series

images, the physicians typically continue processing for other purposes. One such process incudes the computation of the VFI using the VFA and SFA identified at the L3 vertebra level.

Physicians currently process CT scan series one-by-one through software that can read these types of images. One particular process involves using ImageJ, which is a Java based image processing program. The ImageJ software allows a user to scroll through the frontal image of the CT scan and manually identify the L3 vertebra image slice. Once this identification is made, the L3 vertebra image slice can be exported and opened by itself in ImageJ to allow further processing.

Such processing could be used to perform tasks such as VFI calculation. This is a multi-step sequential process, as defined in Figure 18 and as described below.

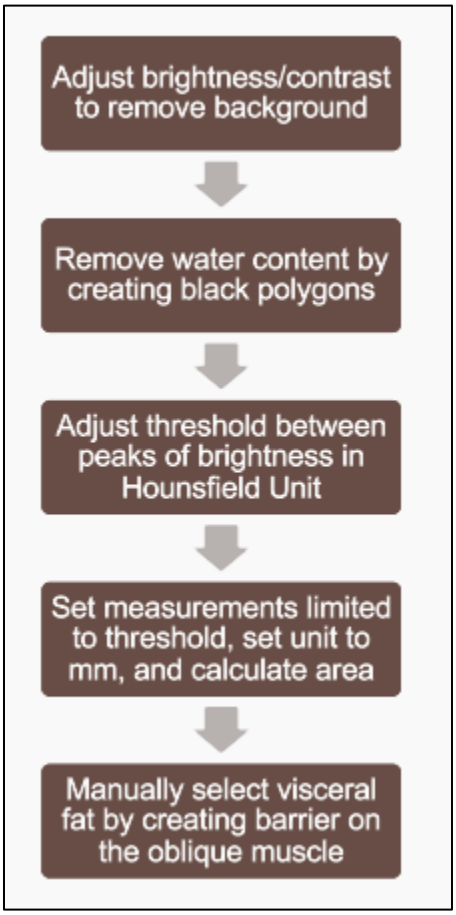

Figure 18: VFI Calculation Steps using ImageJ

- The raw L3 vertebra image slice has a lot of noise that is overlaid as bright gray/white colors throughout the image. These need to be removed from the image to allow the user to see the structure of the body parts that are captured within the L3 vertebra image slice. This removal of the noise is performed by adjusting the brightness and contrast settings in ImageJ. The before and after pictures of the L3 vertebra slice image are shown in Figures 19 and 20.

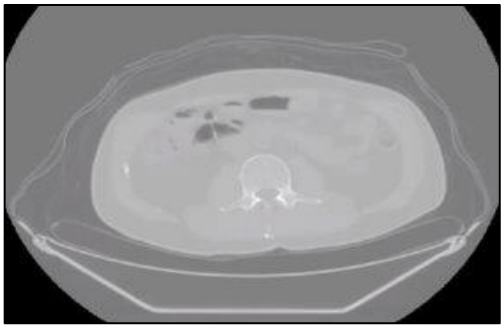

Figure 19: L3 Slice Before Adjusting Brightness/Contrast

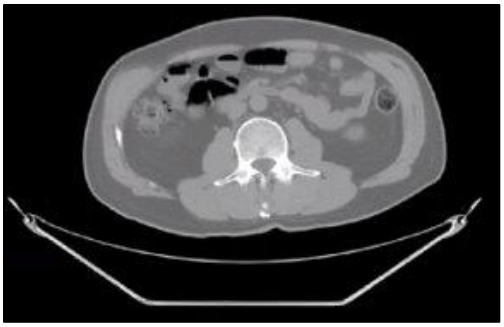

Figure 20: L3 Slice After Adjusting Brightness/Contrast

- Next, it is important to remove water content from the L3 vertebra slice image. This is because water content shows up in the same place as the fat content in the image and can be misinterpreted as fat because both water and fat show values within similar color spectra. The removal of the water content is accomplished in ImageJ by using a tool to create polygons within the image. The identification of the water content is done manually based on the expertise of the user. The “Polygon” tool changes the color spectrum for the area of the image that falls within the polygons to something that is completely different from the spectrum of colors used to represent fat. In Figure 21, this change can be seen. Notice that certain polygons are created in black color and how these are different from the observations made in Figure 20.

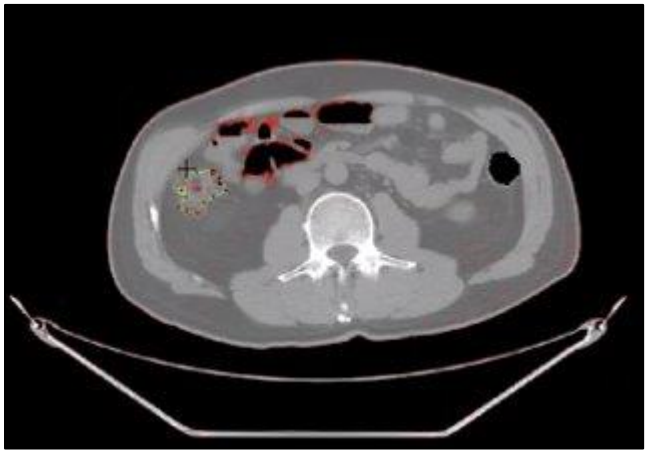

Figure 21: Creating Polygons in ImageJ to Remove Water Content

- After creating the polygons, the “Threshold” tool in ImageJ can be used to find and set the spectra of color that are represented in the image. This tool can be used to assign a starting and ending value to define a range of color. These values are in HU. For all the parts of the image that fall within this range of colors defined, ImageJ assigns a different color. In Figures 22 and 23, it can be observed that the color red was chosen to represent the areas of the image that fall within the defined range. The range is defined based on the expertise of the user to capture only those areas that represent adipose tissue.

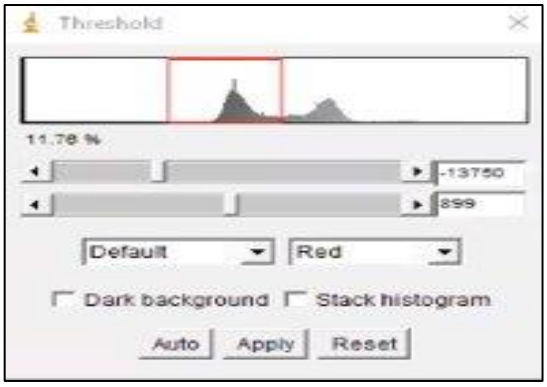

Figure 22: Setting Threshold to Mark Fat Areas

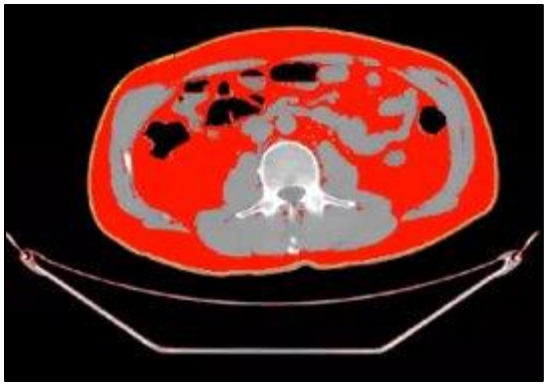

Figure 23: Resulting Image After Marking Fat Areas

- Once the threshold has been defined, ImageJ allows the user to compute the area covered by all the parts of the image that fall within the threshold. One of the units of measurement for calculating the area can be millimeter. This step is accomplished by using the "Set Measurements" tool in ImageJ, as shown in Figures 24 and 25.

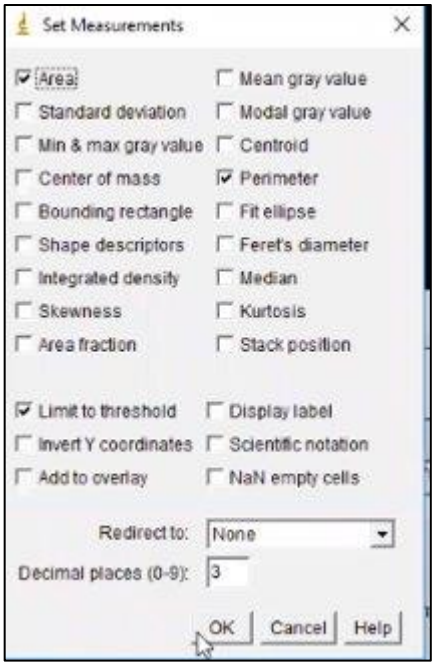

Figure 24: Set Measurements for Threshold Area Calculation

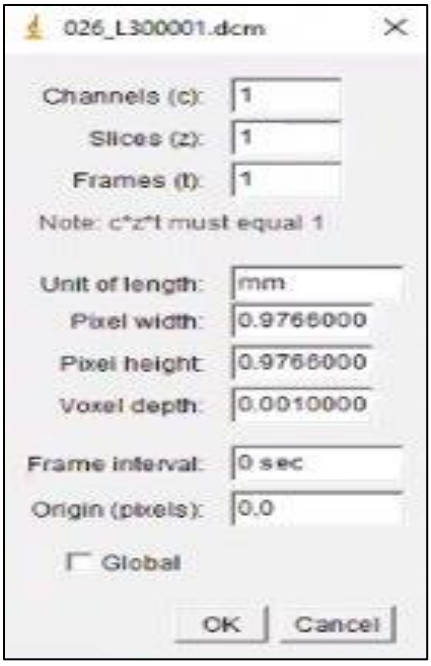

Figure 25: Set Unit of Measurement for Area

- Next, the user a "Line" tool in ImageJ to create two concentric oval structures that define the boundary of the abdomen itself and also defines the boundary of the oblique muscle shown in the abdomen. This is done because the adipose tissue identified outside the oblique muscle boundary but within the overall abdomen boundary is SAT, while the fat that lies inside the boundaries of the oblique muscle is VAT. The two boundaries can be observed in Figure 26. The lines defining these boundaries are shown in yellow color.

- By making this distinction using the “Line” tool, the area that is calculated based on the “Set Measurements” values defined above is shown for the perimeters defined by the line tool. In essence, it provides two values for the area, as can be seen in Figure 27. The smaller area calculated represents the VFA, while the larger area represents VFA and SFA combined, which is the value for total fata area in the body. By dividing the area value calculated for VFA by the combined area value calculated for VFA and SFA both, the VFI is obtained.

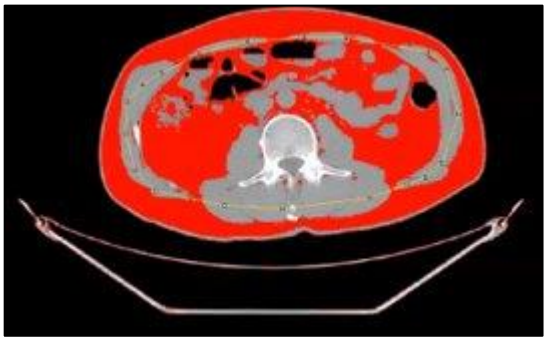

Figure 26: Using Line Tool to Create Two Perimeters to Segment VFA

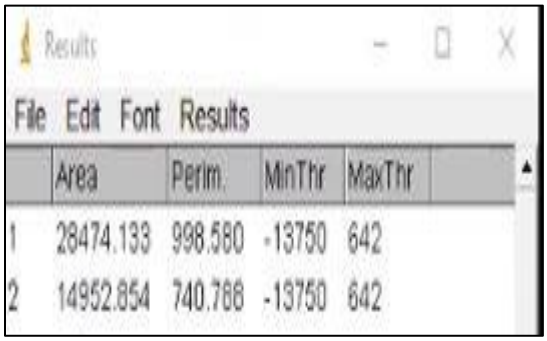

Figure 27: Calculated Area Values for Total Fat Area and VFA

As can be seen from the above manual process, it is tedious and time-consuming to calculate VFI manually. Automation of this process is possible and can be the subject of future studies and projects.

4. Findings

Several tests and analyses were performed on the algorithm to detect L3 slices automatically. Some of the key results are outlined below.

Figure 28 shows the distribution of probability density for series of CT Scans where no L3 vertebra image slice was detected. This analysis helped analyze the viability of the probability threshold as an indicator of the detections and helped come up with a third label of “Review” in addition to the “Yes/No” labels to manually inspect outcomes that were too close to the assigned probability threshold.

The “Review” label was put in place to allow human intervention to take place to supplement the decisions made by the algorithm. As the project progress, it was usually found to be case that viable CT scan series that were analyzed contained the L3 vertebra image slices. Therefore, there were practically no cases when the algorithm should return a value of “No” for L3 vertebra image slice detection. Either the algorithm should be able to identify the L3 vertebra image slice, or it should mark the closest detection for “Review”. Therefore, the algorithm was tweaked to return only two types of output cases: “Yes” and “Review”. The inclusion of the “Review” label was possible only due to the analysis of the “No” decisions as seen in Figure 28.

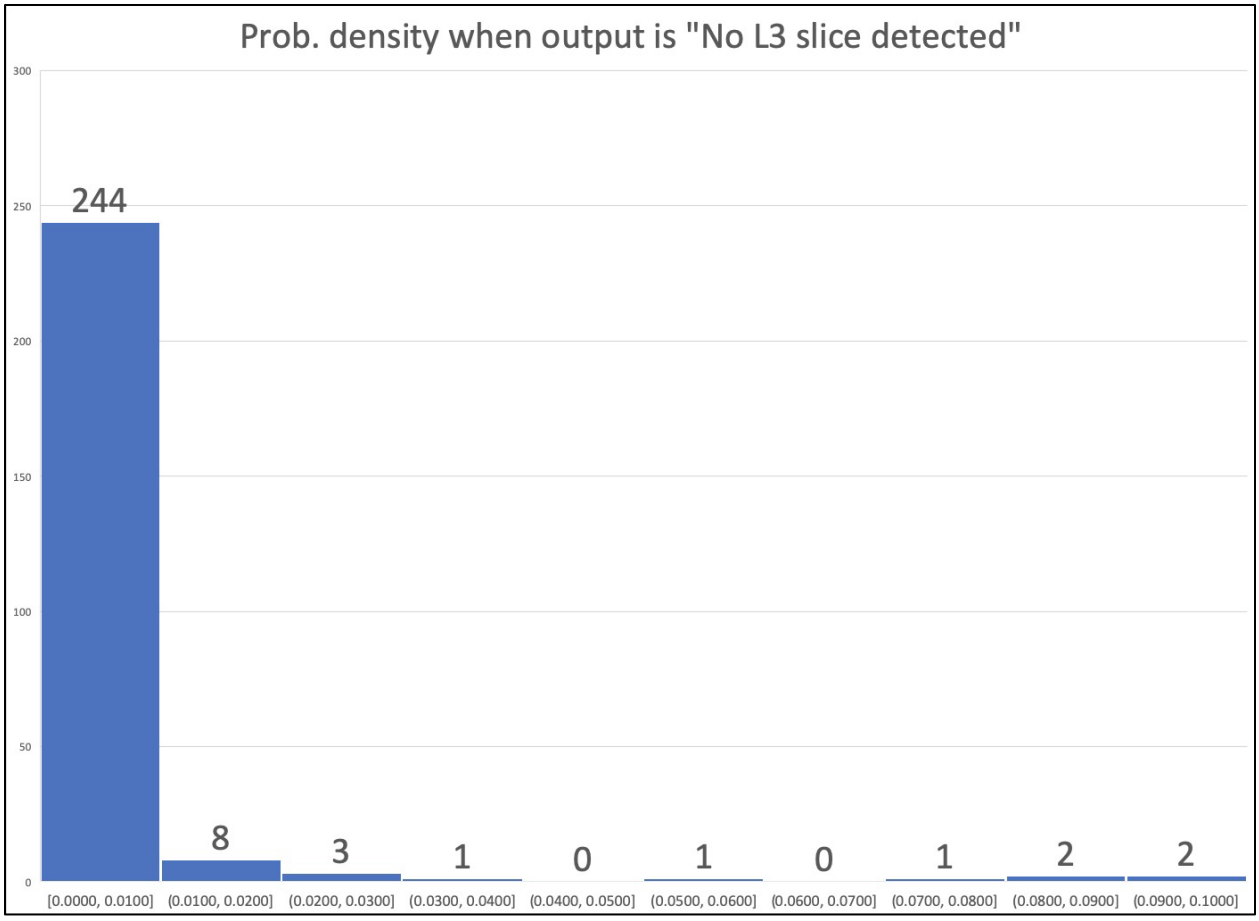

Figure 28: Probability Density for L3 Slice Detections

The original model for the algorithm incorrectly detected the L3 vertebra slice in certain cases, as shown in Figure 29. The detections made by the algorithm are represented by a green colored horizontal line that runs across the frontal image created from MIP of the CT scan series. These are shown in Figure 29. As can be seen, several detections were erroneously made in the legs or the skull region of the body. These erroneous detections were handled by restricting the search area in which the algorithm could make a prediction for the L3 vertebra image slice. This will be discussed later in the project.

The validation of the algorithm was performed on a small subset of the Anti-PD 1 dataset. This dataset had a membership of 30 series of CT scans. For these scans, manual testing or quality assurance validation testing of whether the correct slice was detected was also performed, with the results finding that 22 out of the 30 scans were correctly detected. This is shown in Figure 30 below.

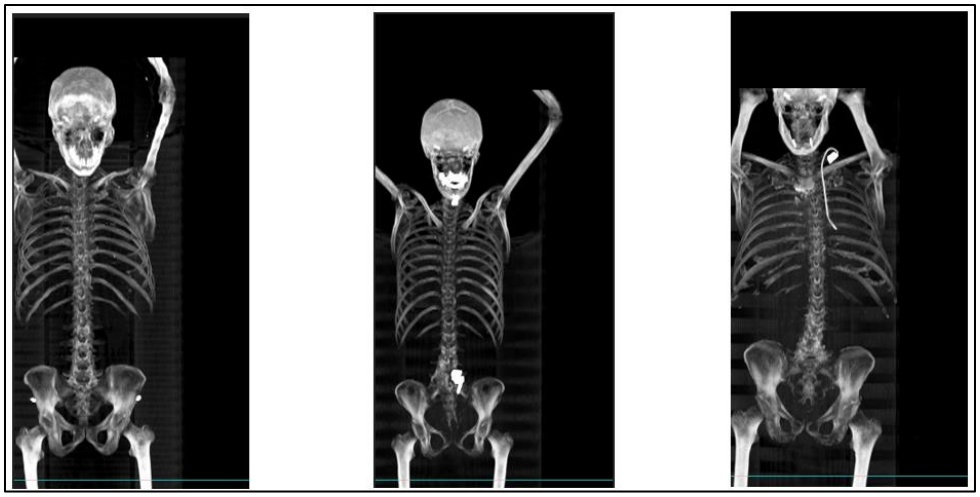

Figure 29: Incorrect L3 Slice Detections

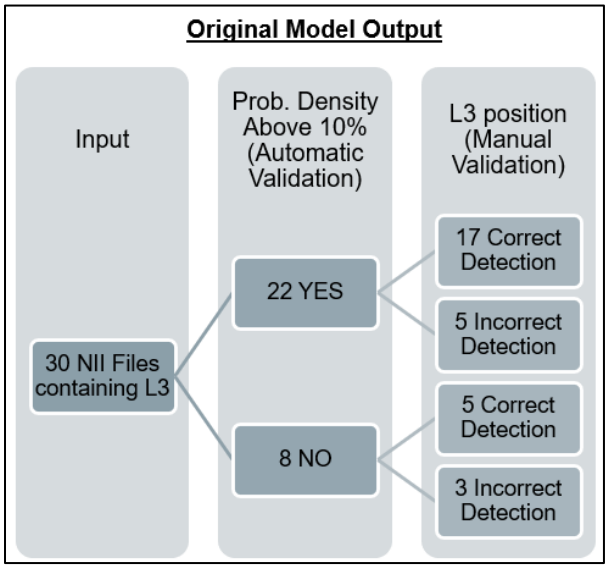

Figure 30: Binomial Tree Representation of Validation Outcomes

For the incorrectly identified L3 vertebra image slices as shown in Figure 29, the correct detections were made after an analysis of the range within which detections were made was performed. This analysis determined that the correct detections were typically in the middle portion of the CT scans. Figures 31 and 32 show the raw and scaled ranges for the correct L3 vertebra slices.

The search windows within which the L3 vertebra could be reasonably expected to be found works well for CT scan series that have extraneous bodily features, such as legs and skulls. This means that the search window technique works well on full-body scans, which the original CNN model used for the predictions was not trained with. This enhances the algorithm to allow to run a more diverse selection of imaging data to obtain dependable results.

The normalization of the slice position search window scales was an important step to get the most accurate positions in a CT scan series where the L3 vertebra could be found. This is because the total number of slices that a CT scan series can contain is not fixed and can vary.

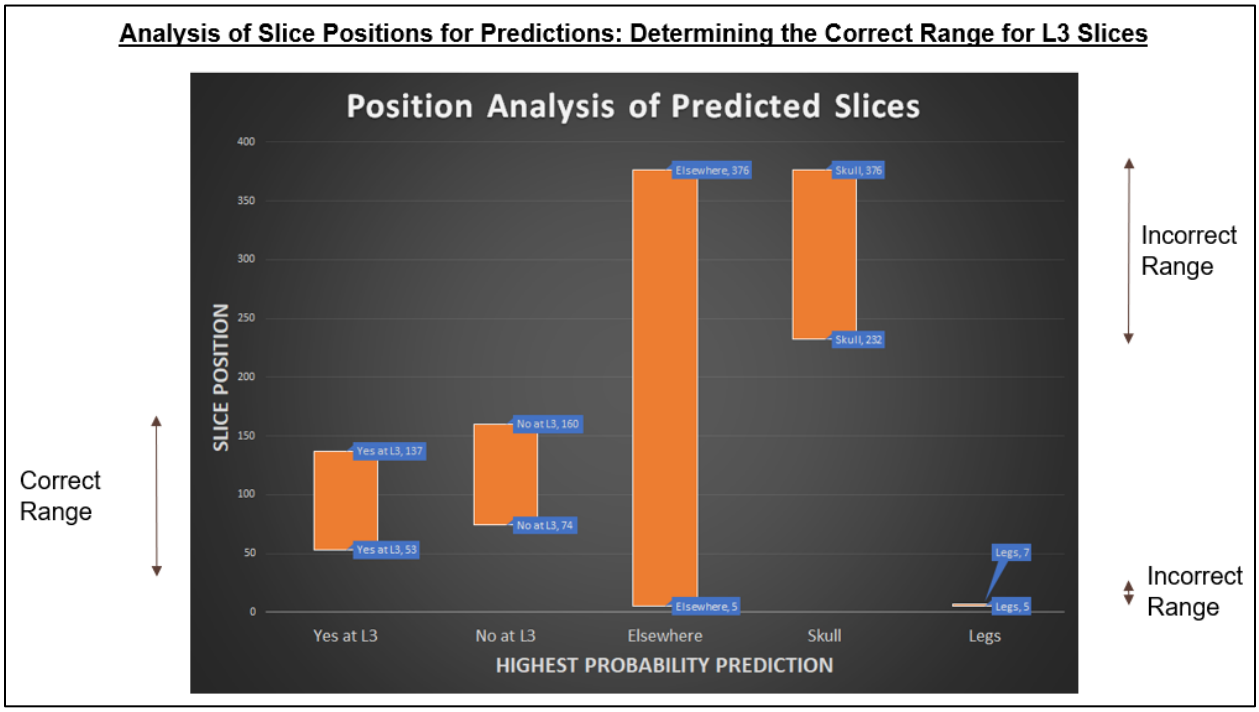

Figure 31: Analysis of Raw Slice Positions to Determine Correct Range

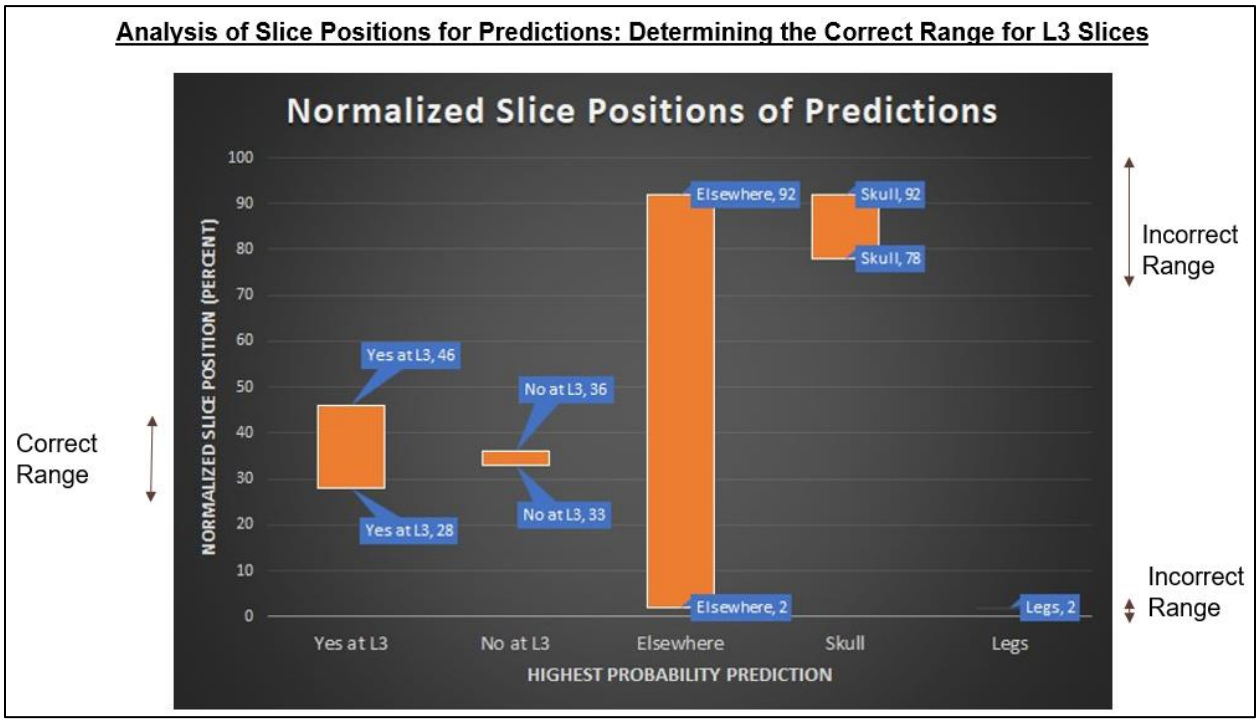

Figure 32: Analysis of Scaled Slice Positions to Determine Correct Range

Based on the slice position analysis, the algorithm was adjusted to reduce its search area to make determinations only in the middle portions of the CT scans. By excluding the areas away from the middle portion of the CT scans, any bodily structures found away from the middle portion of the body were excluded from the area within which a detection could be made. This is a good outcome for the algorithm as the L3 vertebra is found in the middle part of the body and can be reasonably expected to show up in the middle portion of the CT scans as well.

After making this adjustment, several previously incorrectly identified L3 vertebra image slices were correctly identified. This has been depicted in Figures 33 and 34, where it can be observed that previous detections were made in the leg or skull area, whereas the new detections are made in the correct L3 vertebra location.

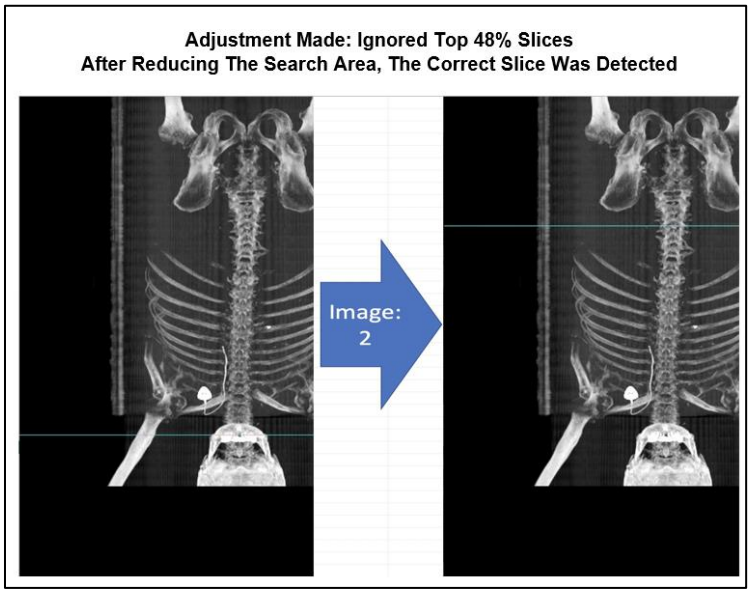

Figure 33: Improved Detection with Reduced Search Area (Restricted Skull Area)

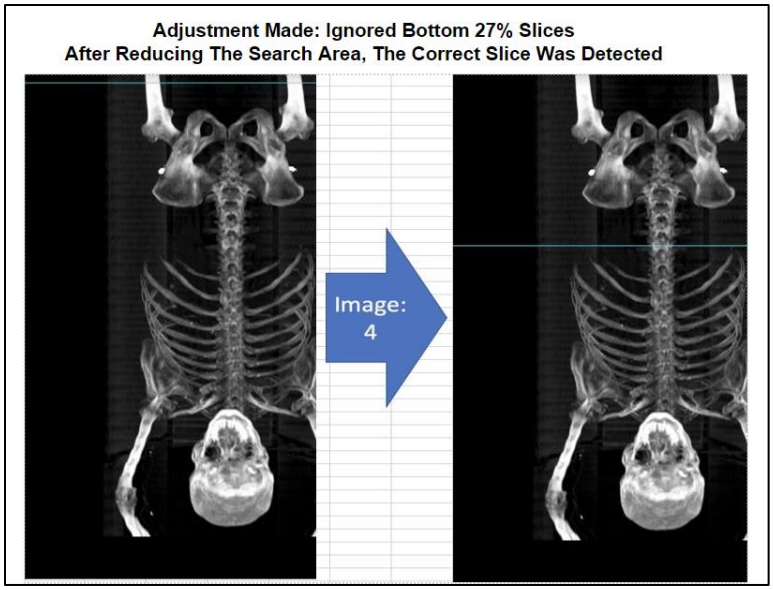

Figure 34: Improved Detection with Reduced Search Area (Restricted Legs Area)

For the 30 images in this validation dataset, the correct detections were performed for 27 images after making the adjustment in the algorithm. Figure 35 sheds light on this improvement.

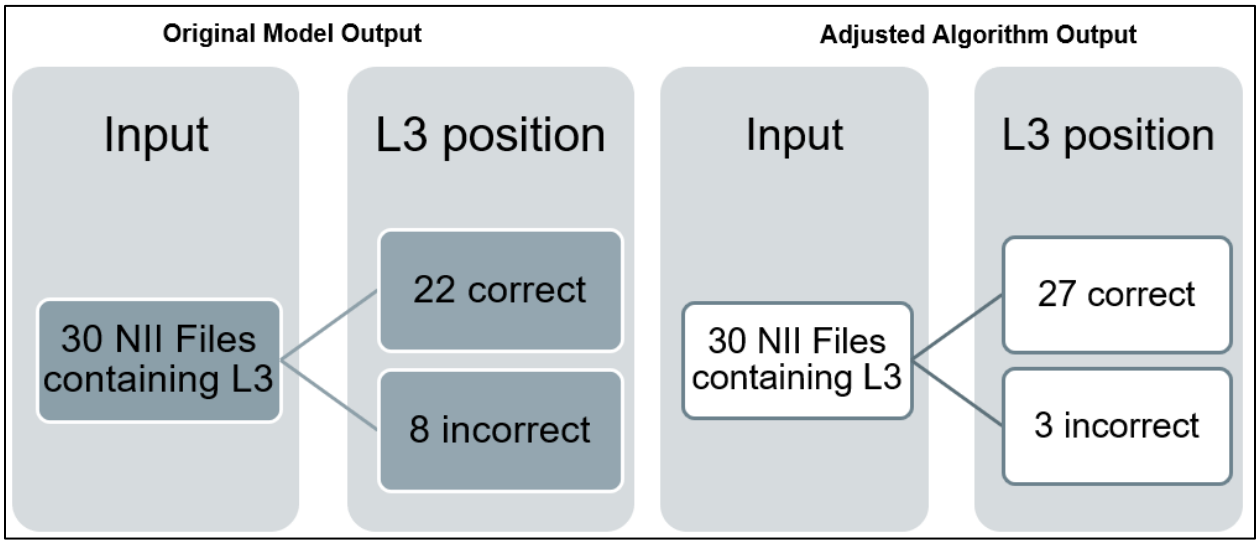

Figure 35: Detection Comparison Between Original and Adjusted L3 Slice Detection Models

The final round of testing was performed using the data provided by the Project Sponsor Allwyn. The results showed that the algorithm detects the L3 slice with high degree of confidence in a 100% of the cases. There were no False Positives included in the results, which implies that the algorithm is suitable to be used in situations where extreme reliability of systems is required, such as in healthcare facilities and laboratories. For the testing phase, overall, 82 CT scan series were tested from a variety of different datasets. The L3 vertebra image slice was correctly detected by the algorithm in 76 out of the 82 series. For the 6 series where this detection was missed, the algorithm assigned the outputs with a “Review” label, indicating that manual processing is required for these cases. This is a testament to the reliability

of the system created. The details of the various test databases and the results are shown in Figures 36 and 37.

	#1	#2	#3
Data Set Description	AntiPD Lung Dataset	AntiPD Lung Dataset	Omniseq Dataset
Total Number of Scans	3	30	49
Threshold used for classification	12%	12%	12%
# of scans classified as successful	**1**	**19**	**23**
# of successful L3 scans verified and found to be correct(QA)	**1**	**19**	**23**
# of scans classified as manual review required	2	11	26

Figure 36: Description of Test Results for Each Test Dataset

	ALGORITHM OUTPUT	
TRUTH	**L3 Position is Correct (YES)**	**Marked for Review**
L3 Position is Correct	43	33
L3 Position Missed (Review Necessary)	0	6

Figure 37: Classification Report Shows 100% Precision

The reliability of the algorithm is based on the fact that whenever it marks a CT series with a "Yes" label, it correctly identifies the L3 slice. Whenever the algorithm misidentifies the L3 slice, it marks the output with a "Review" label. This behavior is explained in Figure 38.

Output: YES, Algorithm is confident L3 Position is Correct

Precision 100%

Recall 57%

Interpretation

Whenever the Output is YES, the L3 Slice Detection is Correct.

From the Overall Correct Detections (76), the Algorithm is Confident only 57% of the time.

Output: REVIEW, Algorithm is not confident about L3 Slice Detected

Precision 15%

Recall 100%

Interpretation

From the overall Review recommendations (39), it is necessary only 15% of the times.

The Algorithm recognizes when detection is not at L3 Slice Position 100% of the times.

Figure 38: Reliability of L3 Detection Algorithm Based on Precision

While a lot of the outputs were marked with the "Review" label, only 6 cases required an actual human review. For all the other cases, the algorithm correctly detected the L3 slice. The only reason why it

didn't assign a "Yes" label is because the probability density threshold applied was a conservative 12%. This is shown in Figure 39.

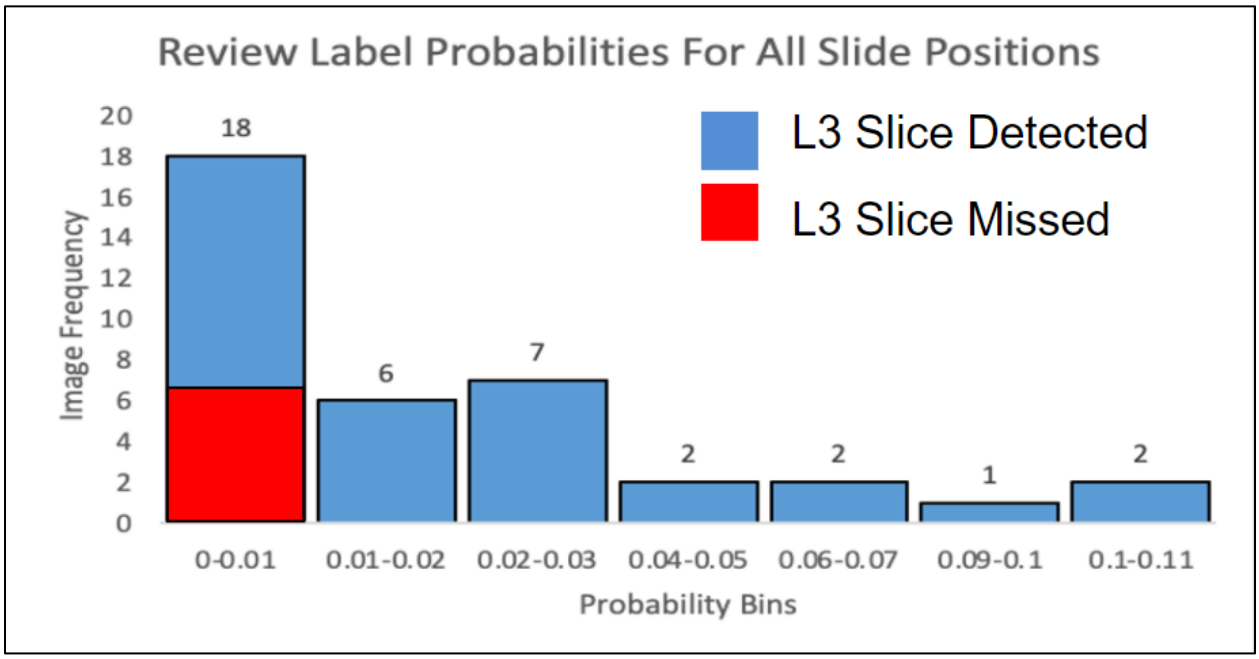

Figure 39: Review Label Probabilities

A high threshold is required to maintain 100% precision that is critical to the reliability of the algorithm. However, to allow users flexibility with the kind of threshold risks they are comfortable with, the probability density threshold values have been parameterized so users can set the values that they deem appropriate. As the quality and type of CT scans varies from dataset to dataset, the reliability of the threshold may be variable from one dataset to another as well. Therefore, with the parameterized threshold, users should be able to gauge the best fit for their needs. The threshold values that have been parameterized as well as how these values will affect the output decision is shown in Figure 40.

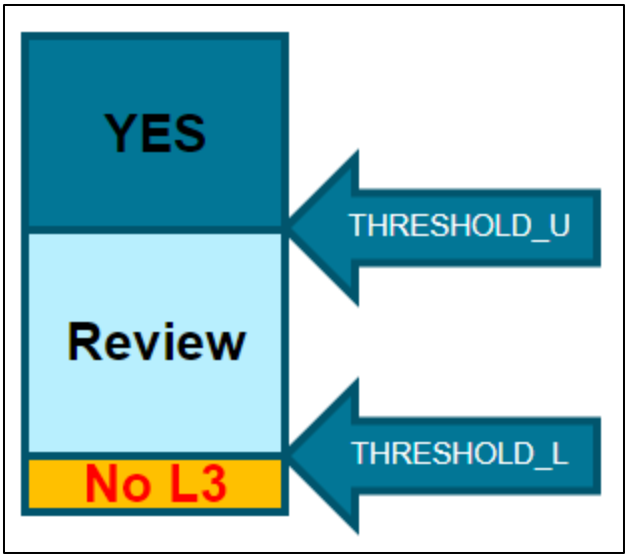

Figure 40: Parameterization of Probability Threshold Values

5. Conclusion

Since most diseases are triggered by the amount of CAT and SAT in the body, it has become increasingly important to calculate the amount of unhealthy adipose tissues. The best location to analyze the adipose tissues that are in the body is the abdominal section surrounding the L3 vertebra. This project was intended to improve upon the manual approach of how the L3 images are detected, creating a simpler scenario to support the calculation of the VFI. Future studies regarding the automation of VFI can focus on automating the next steps of VFI calculation by utilizing the HU scale to compute the distribution of adipose tissue, bone, and muscle within the L3 image slice.

The current clinical approach of reviewing the L3 slice consists of manually labeling and color marking of the different tissues. A complementary algorithm can be created to label the different components automatically. An algorithm for automated VFI calculation was discussed by the research paper of Masoudi et al. and could be leveraged as a starting point for customization. Their method describes the use of ML algorithm to segment fat based on MRI images, so it would need to be adjusted to reflect the requirements of CT scan images. Nonetheless, this study of the fat tissue algorithm could be used as a reference point to calculate the VFI automatically. [29]

The CNN algorithm adjustments presented through our project contribute a critical first step towards the significant goal of automating the VFI calculation.

Appendix A: Code References

A publicly available GitHub repository has been used to assist with the L3 Vertebra Detection algorithm. A link to this repository is mentioned below.

https://github.com/fk128/sarcopenia-ai

Modifications and improvements to the above code were made to complete the project. All of these updates are located in the repository mentioned below.

https://github.com/Team-Metallicus/L3-Slice-Detection

Appendix B: Risks

Risks were assessed throughout the project; however, to compartmentalize the risks, they were recognized over each sprint. Mitigation plans were put in place to assist in case of a high probability risk. These risks have been outlined for Sprint 2 (Datasets) and Sprint 3 (Analytics/Algorithms) in Tables 4 and 5 respectively, shown below.

Table 4: Risks and Mitigation Plan for Sprint 2 (Datasets)

Risk Name	Description	Probability	Impact	Mitigation
Team members not available temporarily (Demand and Capacity)	Member not able to contribute in project progress at some stage of the project (sickness or other reasons)	Low	High	Each team member has a specific role; However, everyone will contribute together in critical tasks and provide support to others.
Missing skills	Team member might not have the skills necessary to conduct the study.	Low	High	The team is confident that we shall be able to accomplish the tasks based on our collective skillsets.
Quality of Dataset for Algorithm 2	Not good enough data	Low	Medium	Have more information on the dataset as soon as possible. Perform quality analysis on data for Algorithm 1.

Risk Name	Description	Probability	Impact	Mitigation
Evaluation of Algorithm 1 Output: Ability to Detect L3 Slice Accurately	The Algorithm 1 code runs and does not produce an accurate output for the L3 Vertebra Slice.	Low	High	Awaiting validation feedback from sponsors. To mitigate this risk, we may have to work only with the “Painted Images” dataset provided by the sponsors. This may not be sufficient as it does not have any of the metadata from the DICOM images

Table 5: Risks and Mitigation Plan for Sprint 3 (Analytics/Algorithms)

Risk Name	Description	Probability	Impact	Mitigation
Team members not available temporarily (Demand and Capacity)	Member not able to contribute in project progress at some stage of the project (sickness or other reasons)	Low	High	Each team member has a specific role; However, everyone will contribute together in critical tasks and provide support to others.
Missing skills	Team member might not have the skills necessary to conduct the study.	Low	High	The team is confident that we shall be able to accomplish the tasks based on our collective skillsets.
Evaluation of Algorithm 1 Output: Ability to Detect L3 Slice Accurately	The Algorithm 1 code runs and does not produce an accurate output for the L3 Vertebra Slice.	Low	High	We have received validation feedback from sponsors. To mitigate this risk, we have to create two buckets. One will be high accuracy and other human review, when algorithm fails to predict the L3.

Appendix C: Agile Development Workflow

To following flowchart is used to track the workflow of the project.

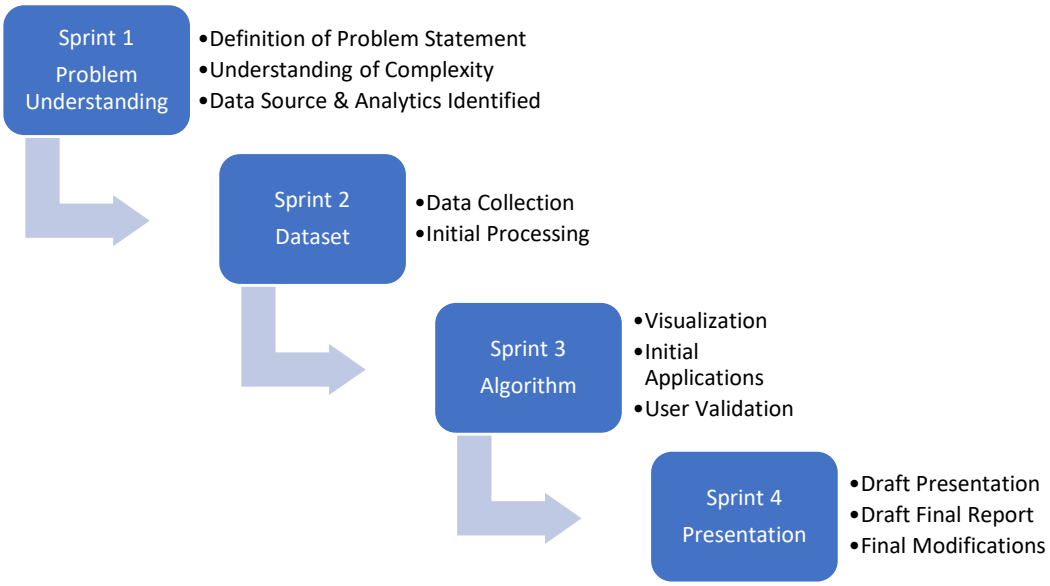

Figure 41: Project Workflow

Appendix D: References

[1] World Health Organization, "Obesity," [Online]. Available: https://www.who.int/topics/obesity/en/. [Accessed 30 September 2020].

[2] World Cancer Fund International, "Worldwide Cancer Data," August 2018. [Online]. Available: https://www.wcrf.org/dietandcancer/cancer-trends/worldwide-cancer-data. [Accessed 10 November 2020].

[3] S. G. Wannamethee and J. L. Atkins, "Muscle loss and obesity: the health implications of sarcopenia and sarcopenic obesity," *Proceedings of the Nutrition Society,* vol. 74, no. 4, pp. 405-412, 2015.

[4] M. Pahor, T. Manini and M. Cesari, "Sarcopenia: clinical evaluation, biological markers and other evaluation tools," *The Journal of Nutrition, Health & Aging,* vol. 13, no. 8, pp. 724-728, 2009.

[5] B. H. Goodpaster, F. L. Thaete and D. E. Kelley, "Composition of Skeletal Muscle Evaluated with Computed Tomography," *Annals of the New York Academy of Sciences,* vol. 904, pp. 18-24, 2000.

[6] K. M. Choi, "Sarcopenia and Sarcopenic Obesity," *Korean Journal of Internal Medicine,* vol. 31, no. 6, pp. 1054-1060, 2016.

[7] F. Kanavati, S. Islam, E. O. Aboagye and A. Rockall, "Automatic L3 slice detection in 3D CT images using fully-convolutional networks," 22 November 2018. [Online]. Available: https://arxiv.org/abs/1811.09244. [Accessed 01 September 2020].

[8] S. K. Mishra and P. Singh, "History of neuroimaging: the legacy of William Oldendorf," *Journal of Child Neurology,* vol. 25, no. 4, pp. 508-517, 2010.

[9] M. Castillo, "The Industry of CT Scanning," *American Journal of Neuroradiology,* vol. 33, pp. 583-585, 2012.

[10] US Food & Drug Administration, "Computed Tomography (CT)," 14 June 2019. [Online]. Available: https://www.fda.gov/radiation-emitting-products/medical-x-ray-imaging/computed-tomography-ct. [Accessed 2020 October 2020].

[11] US Food & Drug Administration, "MRI (Magnetic Resonance Imaging)," 29 August 2018. [Online]. Available: https://www.fda.gov/radiation-emitting-products/medical-imaging/mri-magnetic-resonance-imaging. [Accessed 01 November 2020].

[12] Mayo Clinic, "CT Scan," 28 February 2020. [Online]. Available: https://www.mayoclinic.org/tests-procedures/ct-scan/about/pac-20393675. [Accessed 30 September 2020].

[13] T. D. DenOtter and J. Schubert, "Hounsfield Unit," StatPearls, 11 May 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK547721/. [Accessed 01 November 2020].

[14] Doylestown Health, "BIO-ELECTRICAL IMPEDANCE ANALYSIS (BIA) - BODY MASS ANALYSIS," Doylestown Health, [Online]. Available: https://www.doylestownhealth.org/services/nutrition/bio-electrical-impedance-analysis-bia-body-mass-analysis. [Accessed 21 October 2020].

[15] A. Kalra, "Developing fe human models from medical images," in *Basic Finite Element Method as Applied to Injury Biomechanics*, Academic Press, 2018, pp. 389-415.

[16] K. Hidayat, X. Du, G. Chen, M. Shi and B. Shi, "Abdominal Obesity and Lung Cancer Risk: Systematic Review and Meta-Analysis of Prospective Studies," *Nutrients,* vol. 8, no. 12, p. 810, 2016.

[17] A. Gupta, K. Majumdar, N. Arora, H. Mayo, P. Singh, M. Beg, R. Hughes, S. Singh and D. Johnson, "Premorbid body mass index and mortality in patients with lung cancer: A systematic review and meta-analysis," *Lung Cancer,* vol. 102, pp. 49-59, 2016.

[18] N. Shen, P. Fu, B. Cui, C. Bu and J. Bi, "Associations between body mass index and the risk of mortality from lung cancer," *Medicine (Baltimore),* vol. 96, no. 34, 2017.

[19] J. Wang, H. Xu, S. Zhou, D. Wang, L. Zhu, J. Hou, J. Tang, J. Zhao and S. Zhong, "Body mass index and mortality in lung cancer patients: a systematic review and meta-analysis," *European Journal of Clinical Nutrition,* vol. 72, no. 1, pp. 4-17, 2018.

[20] D. Yu, W. Zheng, M. Johansson, Q. Lan, Y. Park, E. White, C. E. Matthews, N. Sawada, Y. Gao, K. Robien, R. Sinha, A. Langhammer, R. Kaaks, E. L. Giovannucci, L. M. Liao, Y. Xiang, D. Lazovich, U. Peters, X. Zhang, B. Bueno-de-Mesquita, W. C. Willett, S. Tsugane, Y. Takata, S. A. Smith-Warner, W. Blot and X. Shu, "Overall and Central Obesity and Risk of Lung Cancer: A Pooled Analysis," *Journal of the National Cancer Institute,* vol. 110, no. 8, pp. 831-842, 2018.

[21] S. Sarjoghian, Y. Alfadhl and X. Chen, "A novel Wide-Band reflection-based system for measuring abdominal fat in humans," in *2016 URSI International Symposium on Electromagnetic Theory (EMTS)*, Espoo, 2016.

[22] J. Aspiroz-Leehan, M. Cadena and N. Ramos-Ibañez, "Comparative statistical analysis between two methods for the measurement of visceral fat in humans," in *2013 Pan American Health Care Exchanges (PAHCE)*, Medellin, 2013.

[23] J. J. Morrison, J. Hostetter, K. Wang and E. L. Siegel, "Data-Driven Decision Support for Radiologists: Re-Using the National Lung Screening Trial Dataset for Pulmonary Nodule Management," *Journal of Digital Imaging,* vol. 28, no. 1, pp. 18-23, 2015.

[24] Cancer Data Access System, "NLST," National Cancer Institute, [Online]. Available: https://cdas.cancer.gov/nlst/. [Accessed 30 September 2020].

[25] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox and F. Prior, "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," *Journal of Digital Imaging,* vol. 26, no. 6, pp. 1045-1057, 2013.

[26] D. Damien, "A Very Basic DICOM Introduction," dcm4che, 18 November 2015. [Online]. Available: https://dcm4che.atlassian.net/wiki/spaces/d2/pages/1835038/A+Very+Basic+DICOM+Introduction. [Accessed 10 November 2020].

[27] Intel Corporation, "CPU vs. GPU: What's the Difference?," Intel Corporation, [Online]. Available: https://www.intel.com/content/www/us/en/products/docs/processors/cpu-vs-gpu.html. [Accessed 02 November 2020].

[28] A. L. Helmenstine, "What Happens If You X-Ray Metal?," ThoughtCo., 13 January 2020. [Online]. Available: https://www.thoughtco.com/x-rays-and-metal-interference-608418. [Accessed 15 November 2020].

[29] S. Masoudi, S. Anwar, S. Harmon, P. Choyke, B. Turkbey and U. Bagci, "Adipose Tissue Segmentation in Unlabeled Abdomen MRI using Cross Modality Domain Adaptation," in *IEEE Engineering in Medicine and Biology Conference 2020*, Online, 2020.