INVESTIGATION OF ROBUSTNESS OF ^{18}F-FDG PET/CT RADIOMICS FEATURES FOR THE DEVELOPMENT OF IMAGE BASED BIOMARKERS FOR HEAD AND NECK CANCER

NOUSHIN ANAN

UNIVERSITI SAINS MALAYSIA

2023

INVESTIGATION OF ROBUSTNESS OF ^{18}F-FDG PET/CT RADIOMICS FEATURES FOR THE DEVELOPMENT OF IMAGE BASED BIOMARKERS FOR HEAD AND NECK CANCER

by

NOUSHIN ANAN

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

December 2023

ACKNOWLEDGEMENT

First and foremost, all praises and thanks to Almighty Allah for all blesses and strength, chances and endurance during my study and completion of this thesis.

I owe my earnest thankfulness to my mentor and main supervisor Associate Professor Dr. Rafidah Zainon for introducing me to the intriguing field of Medical Physics, and for her valuable guidance and insightful feedback to produce quality research. Her knowledge and wisdom motivated me, constantly. I am grateful for her great mentorship which has helped me become a better researcher. This work would not have been possible without clinical collaborators who identified clinical problems that we tried to address with computational solutions. My deepest gratitude to my co-supervisor, Dr. Mahabubunnabi Tamal for his dedication and guidance throughout every step of this project. The troubleshooting sessions, discussions and lines and lines of exchanged emails with Dr. Tamal helped me perform my thesis work in detail. He constantly motivated me and guided me throughout this journey.

My sincere thanks to Universiti Sains Malaysia and Advanced Medical and Dental Institute for supporting me with proper guidelines and facilities throughout this study.

I would also like to thank my family, who believed in me, inspired me, and supported me throughout the hardest times of this journey. I thank my fellow M.Sc. and Ph.D. colleagues in whom I found friendship and well-needed support through research activities, especially during the COVID-19 pandemic. Special thanks to my colleague Siti Zaleha Mat Isa and my main supervisor Associate Professor Dr. Rafidah Zainon for their contribution in translating Bahasa Malay abstract in my thesis.

Finally, I would like to acknowledge the financial support international collaboration fund (304.CIPPT.6501089.S158) from Ministry of Education, Saudi Arabia via

international research collaboration between Advanced Medical and Dental Institute, Universiti Sains Malaysia, Malaysia, and Imam Abdulrahman Bin Faisal University, Saudi Arabia.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

Θ	Angle
i	Directional vector along x axis
j	Directional vector along y axis
P	Matrix
d	Distance
μ	Linear attenuation constant of a specific substance
p	Proton
n	Neutron
e^+	Positron
v	Neutrino
e	Electron
keV	Kiloelectron volt

LIST OF ABBREVIATIONS

TCIA	The Cancer Imaging Achieve
IBSI	Imaging Biomarker Standardisation Initiative
CNR	Contrast to noise ratio
CT	Computed Tomography
PET	Positron Emission Tomography
GLCM	Grey-Level Co-occurrence Matrix
GLRLM	Grey-Level Run-Length Matrix
GLSZM	Grey-Level Size-Zone Matrix
NGTDM	Neighbouring Grey-Tone Difference Matrix
ROI	Region of Interest
SNR	Signal to Noise Ratio
VOI	Volume of Interest
2D	Two-Dimensional
3D	Three-Dimensional
FOV	Field of view
^{18}F-FDG	^{18}Flourine Fluorodeoxyglucose
SUV	Standardised Uptake Value
MTV	Metabolic tumour volume
DSC	Dice's Similarity Coefficient
GT	Ground Truth
ANOVA	Analysis of Variance
OSEM	Ordered Subset Expectation Maximization
FWHM	Full width at half maximum

SD	Standard Deviation
COV	Coefficient of variation
ROI_{true}	True segmentation

LIST OF APPENDICES

KAJIAN KETEGUHAN CIRI RADIOMIK ^{18}F-FDG PET/CT UNTUK PEMBANGUNAN BIOPENANDA BERASASKAN IMEJ UNTUK KANSER KEPALA DAN LEHER

ABSTRAK

Secara konvensional, ^{18}F-FDG PET/CT telah dinilai melalui matrik biopenanda seperti isipadu tumor metabolisme (MTV), nilai ambilan piawai (SUVs)untuk ramalan kanser. Walaubagaimanapun, ia tidak mengukur kepelbagaian tumor yang mencetuskan tentangan terhadap rawatan. Radiomik boeh mengukur kepelbagaian tumor dengan mengenal ciri dan corak yang ada pada ^{18}F-FDG PET/CT. Ciri radiomik ^{18}F-FDG PET/CT diperlukan untuk ketepatan atau dalam kata lain teguh untuk tetapan pensegmenan berbeza. Tujuan kajian penyelidikan ini adalah untuk mengkaji keteguhan ciri radiomik ^{18}F-FDG PET/CT untuk pembangunan imej berasaskan biopenanda untuk kanser kepala dan leher. Kajian penyelidikan ini mengkaji keteguhan 44 ciri radiomik terhadap variasi pensegmenan imej, dan hubungan dengan parameter kualiti imej- nisbah kontras kepada hingar (CNR), nisbah isyarat kepada hingar (SNR) dan MTV. Kajian ini melibatkan 59 data pesakit kanser kepala dan leher daripada arkib pengimejan kanser dengan kebenaran. Manual dan juga semi-automatik pensegmenan bahagian yang dikehendaki (ROI) telah dilakukan pada dataset. Kepelbagaian antara pemerhati dan dalam pemerhati telah dibuang dengan melakukan pensegmenan ROI dua belas kali sesi menggunakan kedua-dua kaedah oleh tiga pakar perubatan nuklear. 44 ciri radiomik daripada 6 kluster ciri radiomik telah dikeluarkan daripada ROI. 44 ciri ini dikategorikan kepada lima kumpulan termasuklah – 3 ciri Global, 9 Ciri Matrik Kejadian Bersama Skala Kelabu

(GLCM), 13 Ciri Marik Panjang-Lari Skala -Kelabu (GLRLM), 13 Ciri Matrik Saiz Zon Skala-Kelabu (GLSZM), 5 Ciri Matrik Perbezaan Ton-Kelabu Kejiranan (NGTDM) dan satu ciri berasaskan saiz dan bentuk. Analisis kepelbagaian pekali (COV) telah dilakukan untuk mengkategorikan keteguhan ciri radiomik. Analisis regresi linear telah menentukan hubung kait antara ciri radiomik dan parameter imej kualiti. 18% ciri adalah teguh (COV < 5%) berkenaan dengan kepelbagaian pensegmenan. Semua ciri Global, ciri bentuk dan saiz, ciri NGTDM, 5 ciri GLRLM, 7 ciri GLCM, 7 ciri GLSZM adalah kurang teguh (COV > 20%) berkenaan dengan kepelbagaian pensegmenan. 4 ciri teguh GLRLM bersama dengan kekasaran NGTDM dan LZE GLSZM tidak menunjukkan hubungan dengan MTV, CNR, dan SNR (nilai $p > 0.05$). Keenam-enam ciri ini adalah calon yang baik untuk pembangunan biopenanda. Perbandingan pensegmenan manual dan semi-automatik telah menunjukkan kebolehulangan kaedah pensegmenan semi-automatik adalah lebih besar daripada kaedah pensegmenan manual. Kepelbagaian penyempadanan ROI telah berlaku lebih banyak untuk MTV lebih kecil, CNR dan SNR yang rendah. Dalam kajan ini, pensegmenan semi-automatik lebih digemari berbanding pensegmenan manual untuk penyempadanan ROI bagi tumor kecil. Keteguhan ciri radiomik dengan kaedah konvensional matrik ^{18}F-FDG PET/CT boleh digunakan untuk melaporkan kepelbagaian maklumat untuk pencirian tumor yang lebih baik. Pencirian tumor yang baik mempromosikan ketepatan rawatan dan diagnostik yang lebih tinggi dengan menyelesaikan maslaah berkaitan kepelbagaian, perbezaan tisu sihat daripada tumor dan jangkaan ramalan maklumbalas.

INVESTIGATION OF ROBUSTNESS OF ^{18}F-FDG PET/CT RADIOMICS FEATURES FOR THE DEVELOPMENT OF IMAGE BASED BIOMARKERS FOR HEAD AND NECK CANCER

ABSTRACT

Conventionally, ^{18}F-FDG PET/CT is evaluated through biomarker matrices for example metabolic tumour volume (MTV) and standardised uptake values (SUVs) for cancer prognosis. However, they do not quantify tumour heterogeneity that induces resistance to treatment. Radiomics may quantify tumour heterogeneity by recognising features and patterns present in the ^{18}F-FDG PET/CT. The radiomics feature of ^{18}F-FDG PET/CT need to be accurate in other words robust for different segmentation settings. This study aimed to investigate the robustness of ^{18}F-FDG PET/CT radiomics features for image-based biomarker development for head and neck cancer diagnosis. The robustness of 44 radiomics features was examined over image segmentation variation, and their relationship with MTV and image quality parameters - signal to noise ratio (SNR), contrast to noise ratio (CNR). This study involved head and neck cancer image dataset in ^{18}F-FDG-PET/CT modality of 59 patients from the cancer imaging archive with permission. Manual and semi-automated segmentations were performed on the adopted dataset to delineate the region of interest (ROI). Twelve sessions for each ROI segmentation was performed using both methods by three nuclear medicine specialists to remove inter and intra observer variation. 44 radiomics features from the 6 radiomics feature clusters were obtained from the ROI. These 44 features were categorised into 6 groups including – 3 Global features, 9 Grey-Level Co-occurrence Matrix (GLCM) features; 13 Grey-Level Run-Length Matrix (GLRLM) features; 13 Grey-Level Size Zone Matrix (GLSZM) features; 5

Neighbourhood Grey-Tone Difference Matrix (NGTDM) features, and 1 shape and size-based feature. The coefficient of variation (COV) analysis was performed to categorise the robustness of the radiomics features. Linear regression analysis determined the correlation between radiomic features and image quality parameters. 18% of the features were robust (COV < 5%) with respect to segmentation variation. All the Global features, shape and size-based features, NGTDM features, 5 GLRLM features, 7 GLCM features, 7 GLSZM features were least robust (COV > 20%) with respect to segmentation variation. 4 GLRLM robust features along with coarseness of NGTDM and Large Zone Emphasis (LZE) of GLSZM showed no relation with MTV, CNR and SNR ($p > 0.05$). These six features are good candidate for biomarker development. Comparison of manual and semi-automated segmentation showed that repeatability of semi-automated segmentation method is greater than manual segmentation method. The variation in ROI delineation arises more with smaller MTV, low CNR and low SNR. In this study, semi-automated segmentation is preferred over manual segmentation for delineating ROI of small tumours. The robust radiomics features along with the conventional ^{18}F-FDG PET/CT matrices can be used to report tumour heterogeneity information for better tumour characterisation. Proper characterisation of tumour would promote higher diagnostic and treatment accuracy by solving the problem of heterogeneity, distinguishing healthy tissue from tumour and predicting prognostic response.

CHAPTER 1

INTRODUCTION

1.1 Background

Head and neck Cancer (HNC) claims 450,000 deaths annually and this number is predicted to rise by 30% by 2030 (Sung *et al.*, 2021, Johnson *et al.*, 2020). Globally. head and neck squamous cell carcinoma (HNSCC) ranked as the seventh predominant cancer with the highest incidence number in Asia (Cheong *et al.*, 2017). Patients with HNC diagnosis has doubled since 1982 and this trend is expected to continue (Koh *et al.*, 2019). In Malaysia, HNC is the third most dominant cancer with a total of 4,075 cases (Bray *et al.*, 2020, Husmeela *et al.*, 2021). The 5-year survival rates of only about 10 - 40% confer a poor prognosis of advanced HNC (Wong *et al.*, 2015). The incidence of HNC in Malaysia was reported to be 8.5 per 100,000 which is higher than the average global incidence in developed regions (Wong *et al.*, 2015). The inter and intra tumoral heterogeneity along with complexity present in HNC challenges the effective diagnosis of HNC (López *et al.*, 2021).

The qualitative diagnosis - invasive biopsy does not reveal the entire tumour characterisation as it includes extracting part of the tumour lesion. ^{18}F-fluoro-2-deoxy-D-Glucose positron emission tomography and computed tomography (^{18}F-FDG PET/CT) is broadly used for prognosis, observation and diagnosis of head and neck cancer as a means of quantitative diagnosis. The common quantitative measures of tracer uptake for quantification of ^{18}F-FDG PET/CT are metabolic tumour volume (MTV), standardised uptake value (SUV) and SUV derivatives. However, they have limited potential of reflecting the spatial distribution of ^{18}F-FDG. Integration of radiomics features into ^{18}F-FDG PET/CT head and neck cancer imaging quantifies tumour heterogeneity by providing precise information about intensity, shape, size,

volume, and texture of cell phenotype that is distinct or complementary to that provided by clinical reports and proteomic assays (Tixier *et al.*, 2011, Chicklore *et al.*, 2013, Bailly *et al.*, 2019).

The main goal of radiomics is the extraction of quantitative imaging features in an automated method and the development of prediction models for non-invasive diagnosis of lesion phenotypes. Radiomics enables extraction, collection and evaluation of higher order and statistical datasets through radiographic information conversion into large-scale and mineable entities (Rizzo *et al.*, 2018, Lambin *et al.*, 2012). Feature can be understood as an image-derived descriptor of intensity, shape, texture, or any other visually assessable or quantitatively measurable characteristics of image appearance. Several previous studies have described a true correlation between radiomics features and tumour biological characteristics such as cellularity, heterogeneity and necrosis, which are often directly involved in other diagnostic or outcome variables (Cook *et al.*, 2018, Sanduleanu *et al.*, 2018). Imaging features acknowledged as biomarker have diagnostic standard that characterises the biological and functional activity of the body (Boellaard, 2017). ^{18}F-FDG PET/CT radiomics features can be titled as biomarker only when the features become robust and standard.

Robustness is defined as the ability of a given methodology to generate accurate segmented volumes under varying acquisition and image reconstruction conditions. Standardisation includes precise, feasible and accurate radiomics feature quantification. Radiomics feature analysis includes several steps starting from image acquisition and ending in statistical analysis. So, ^{18}F-FDG PET/CT radiomics feature biomarker can be achieved by standardising the complete radiomics analysis process. Figure 1.1 represents the radiomics analysis workflow and steps toward imaging biomarker discovery for ^{18}F-FDG PET/CT for head and neck cancer diagnosis. Image

acquisition is the first step of radiomics feature analysis. The second step involves reconstruction of the acquired image using different software platforms. Sharpening and smoothing filter processes are applied during image reconstruction. Afterwards, tumour contour is defined using region of interest (ROI) delineation. Extraction of textural features from the ROI is performed and statistical model or machine learning algorithm is developed to attain biomarkers.

Radiomics has received much attention and interest in the field of ^{18}F-FDG PET/CT imaging. Nonetheless, reproducibility and validation of the published work are still a big challenge (Gillies *et al.*, 2016, Boellaard *et al.*, 2015, Berenguer *et al.*, 2018, Welch *et al.*, 2019, Meyer *et al.*, 2019). The absence of unanimously recognized reference values and definitions have hampered the clinical use of ^{18}F-FDG PET/CT image biomarker. Furthermore, well-established image processing platform required to extract, compute features is absent (Vallieres *et al.*, 2018, Hatt *et al.*, 2017, Bousabarah *et al.*, 2019). As a consequence, results published in one setting cannot be reproduced in different clinical settings. Manipulation and assessment of a single image set in two different software platforms result in dissimilar feature values (Foy *et al.*, 2018). Variation of imaging procedure, ^{18}F-FDG activities, image reconstruction, data comprehension and uptake time is significant (Messerli *et al.*, 2019, Beyer *et al.*, 2011, Graham *et al.*, 2011). Additionally, lack of detailed report of the reproducibility of the experiments and findings aggravates the situation (Traverso *et al.*, 2018).

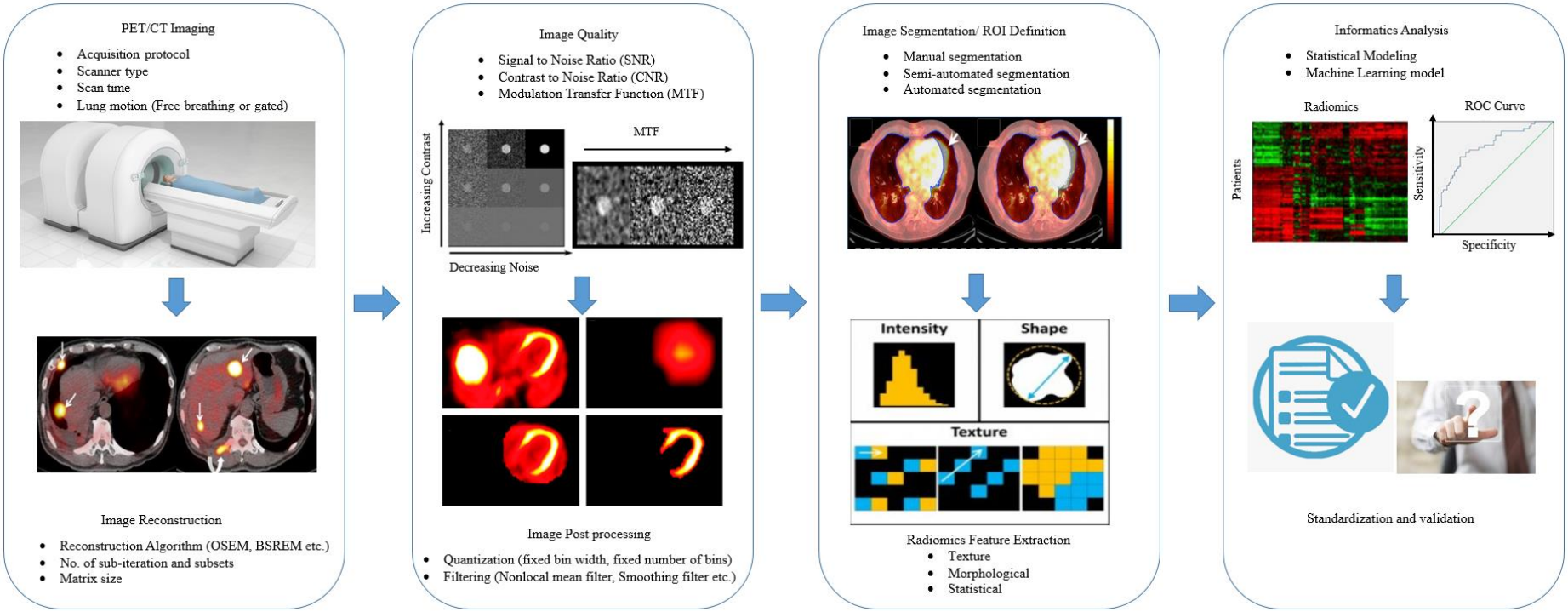


Figure 1.1 Overview of a radiomics workflow and imaging biomarker discovery.

The situation can be solved by standardisation of the radiomics features definition with supportable references and coherent execution of image assessment strategies for feature quantification (Uthoff *et al.*, 2019, Bogowicz *et al.*, 2017, Hatt *et al.*, 2017, Foy *et al.*, 2018). In the interest of strengthening the application of ^{18}F-FDG PET/CT as imaging biomarkers guidelines on tumour imaging using ^{18}F-FDG PET/CT have been published and revised (Boellaard, 2009, Schelbert *et al.*, 1998). Currently, it is well understood that harmonization of imaging modalities is vital alongside standardising imaging performance for realizing computation of ^{18}F-FDG PET/CT as biomarker (Boellaard, 2009).

1.2 Problem Statement

Imaging biomarkers, especially quantitative imaging biomarkers, are of great interest. They can provide a comprehensive view of the whole lesion while capturing clinically relevant biological predictors such as regional tumour intra-heterogeneity. Imaging biomarkers provides opportunities to tailor treatment decisions based on observed responses. Imaging-based quantification and characterisation of tumoural phenotypes has been the main goal of numerous efforts in recent years developing and integrating precision oncology in clinical practice (Creff *et al.*, 2020, O'Connor *et al.*, 2017, Gambhir, 2002). Identifying optimal quantitative image features for computer-aided diagnosis constitute crucial steps towards the development of robust, reproducible, standardised, and clinically relevant imaging biomarkers of head and neck cancer phenotypic characteristics (Hatt *et al.*, 2018, Boellaard *et al.*, 2015). In recent years, numerous quantitative imaging biomarkers based on different image features have been proposed. Clinical acceptance of novel imaging biomarkers is limited and translation into clinical practice generally takes years if not decades.

Currently, tumour response and tumour grading in head and neck cancer are essentially performed through qualitative measurements or using 1D or 2D descriptors of the size of lesions (Okada *et al.*, 2005, Mountain, 2000). Subjective visual evaluation of lesions on clinical medical images might not capture histopathological or genetic features of disease activity, including intra-tumoural heterogeneity, an important biomarker of cancer aggressiveness (Julesz *et al.*, 1973, Tixier *et al.*, 2014). Therefore, improved tumour treatment prescriptions could be achieved with comprehensive quantitative imaging biomarkers, overcoming the subjectivity of visual interpretation and over-simplistic assessment of shape markers of pathological structures on medical images. Thus, standardised and quantitative computational methods have the potential of improving radiology and oncology workflows in head and neck cancer patient screening, decision support, detection, and interpretation of findings to alleviate the current burden on radiologists and radio-oncologists. Image biomarkers cannot be subjective to segmentation settings rather they should be reproducible in any clinical settings. In this study, radiomics features were tested against segmentation, tumour size and image quality variation to identify the features that would be reproducible in different clinical settings.

1.3 Research Objectives

The main objective of this research was to investigate the robustness of ^{18}F-FDG PET/CT radiomics features for the development of image based biomarkers for head and neck cancer diagnosis. The specific objectives of this study were listed as follows:

i. To evaluate the consistency of the radiomics features for different segmentation methods

ii. To analyse the dependency of the radiomics features on MTV and image quality parameters including- CNR and SNR

iii. To evaluate the variation of manual and semi-automated segmentation methods for image quality parameters

1.4 Scope of Study

In this study, we focus on the importance of image segmentation for robust radiomics feature generation. We investigated the robustness of the radiomics features for obtaining stable and reliable features that reflect the biologic heterogeneity present in the ^{18}F-FDG PET/CT. 44 radiomics features were evaluated to determine the stable and sensitive radiomics features. The stability and sensitivity were examined against segmentation variation. Dependency of the features on image contrast, noise and tumour size was evaluated. We evaluated the head and neck tumour segmentation accuracy of two different segmentation methods. Variation of the segmentation for different contrast, noise and tumour size was also examined.

1.5 Thesis Organisation

The thesis contains five chapters. Chapter 1 comprises a general background of ^{18}F-FDG PET/CT radiomics and its potential as a biomarker. It also includes the problem statement that shows the need for this research, objectives, and scope of the research.

Chapter 2 presents the theoretical section related to the research area. It also consists of the literature review were some research done by the previous researchers on ^{18}F-FDG PET/CT radiomics in head and neck cancer is presented.

Chapter 3 includes the research methodology, where materials and the methodology are presented in detail. This chapter comprises open source QIN HEAD-NECK data collection, evaluation of ROI segmentation, radiomics feature extraction process, and stability analysis method of the extracted features.

The results and discussion of the results are presented in chapter 4. The chapter focus results obtained from stability analysis of the radiomics features. Selection of robust radiomics features is discussed in this chapter. This chapter includes relation among the image quality parameters- CNR, SNR, MTV and radiomics features. Results obtained from evaluation of segmentations and their relation with image contrast, noise and tumour are presented in this chapter with detailed discussion.

Finally, chapter 5 summarises and concludes the research work with some suggested future recommendations.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter comprises a detailed review of the available research work on ^{18}F-FDG PET/CT radiomics for head and neck cancer. In addition, concepts and theories related to this study were presented in this chapter. As diagnostic imaging is an important step towards radiomics analysis, the first section is dedicated towards diagnostic imaging. The main goal of radiomics is non-invasive quantification of tumour heterogeneity and so tumour heterogeneity is discussed in the following section. Finally, the radiomics definition, radiomics analysis process and relevant literature are presented in the later sections of this chapter.

2.2 Diagnostic imaging in head and neck cancer

The advances in clinical imaging play a centre part within the entirety of cancer management (Fass, 2008, Weissleder, 2006). It precisely identifies tumour area, ration, metastasis, and whether the treatment may include basic anatomical structures. Particularly, the integration of genomics and proteomics technologies with anatomical imaging conveys the molecular and physiological information with anatomical information, of the subject (Weissleder, 2006). The combination of molecular and anatomic imaging improves microlevel or macrolevel change distinction, survey and alter clinical planning in real-time, cancer drugs discovery simplification (Lambin *et al.*, 2012). More critically, this diagnostic imaging method visualizes tissue in non-invasive manner and avoids intrusive diagnostic tests. Conventional clinical imaging procedures are computed axial tomography (CT) imaging, positron emission tomography (PET) imaging, magnetic resonance imaging (MRI) and ultrasonography.

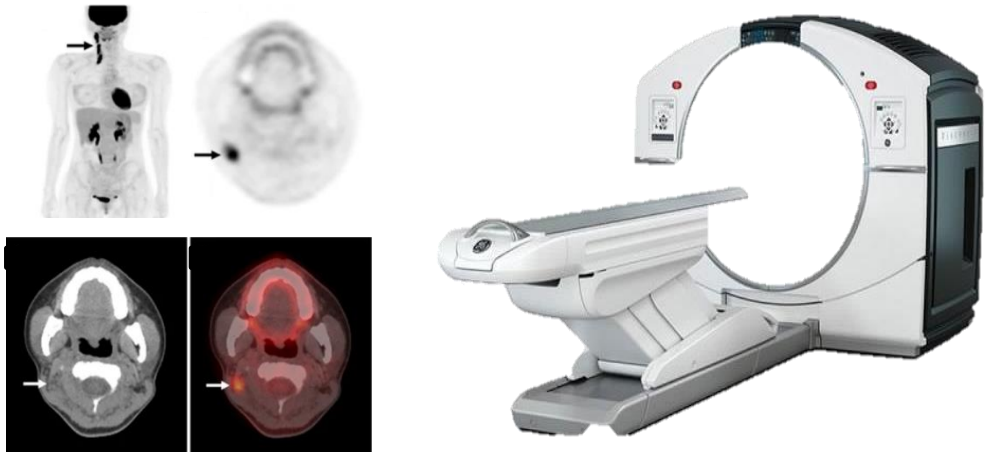
Our data set consisted PET/CT images only. So, the basic principles of hybrid PET/CT imaging is discussed in the next section.

2.2.1 Hybrid PET/CT Imaging

Positron emission tomography (PET) evaluates the metabolic and molecular features of a variety of malignancies, but its anatomical structure view is constrained. CT makes it easier to assess the physical features of tumours, but it cannot capture their metabolic and molecular characteristics (Seemann *et al.,* 2004). As a result, the combination of PET and CT enables the correct integration of metabolic and molecular features of the disorder with anatomical findings, providing additional information for the diagnosis and staging of tumours. Modern full-ring 3D PET and high-end 16-slice CT scanners are paired in the most updated design of high PET/CT scanners. Instead of employing ^{68}Ge sources for regular transmission scanning, PET/CT scanners attenuation-correct PET acquisition using a CT scan. As a result, the examination time is shortened. However, metallic objects and contrast agents that affect the quality of CT scans and quantitative measures of standardised uptake values (SUV) may cause artefacts in the PET images.

In comparison to PET or CT imaging alone, combining PET and CT imaging technology into a single scanner has a number of benefits. In integrated systems, the CT may be utilised to precisely localise where anatomical radiotracer uptake occurs, to adjust for attenuation, and to help accelerate the PET examination. A study of the uncorrected images may be required to distinguish between actual radiotracer uptake and tracer activity overestimation brought on by artefacts from the CT-based attenuation correction. In order to avoid "false" interpretations of infection, inflammation, or even cancer surrounding the body, only the absence of increased activity in the unfiltered

captures may actually indicate absent radiotracer activity in the area of the body. These methodological aspects must be taken into consideration when analysing changes in subjective or numerical terms (Seemann *et al.*, 2004).



(a) (b)

Figure 2.1 (a) PET/CT image; (b) PET/CT Scanner. (Retrieved from Kim *et al.*, 2013)

Typically, PET/CT imaging begins with a targeting beam computation radiograph, frequently referred as "topogram" that establishes PET scanning scope. At first, the patient undergoes CT acquisition equal to the length of topogram for attenuation rectification and even uptake area detection. Next, PET data acquisition is performed. The field of view determines the scan area in single acquisition. Typically, 15 cm (approximately) is the standard axial PET field of view. The present PET scanners have 22 cm, 26 cm field of view options. Only one field of view (FOV) is necessary for scanning the brain or the heart; however, whole-body imaging is implemented for investigating the disorder degree in oncology (Figure 2.1).

The whole body cannot be scanned in a FOV. Multiple FOVs are obtained to cover the total body. However, detectors sensitivity is very poor at FOV edges and so the fields of view comprise small amount of overlaps (Tout *et al.*, 2016). Figure 2.2 represents the multiple bed positions with small overlaps during the whole-body PET

scan. The term "bed position" refers to each of these fields of vision, and every bed position takes 1.5 to 5 minutes to complete. The scan time depends on the scanner's sensitivity and the radiopharmaceutical's affinity. Alignment between the PET and the CT is utilised to identify and correct attenuation as well as reduce movement artefacts. Balance between patient comfort and immobilisation is very important in order to preserve PET and CT alignment (Tout *et al.*, 2016).

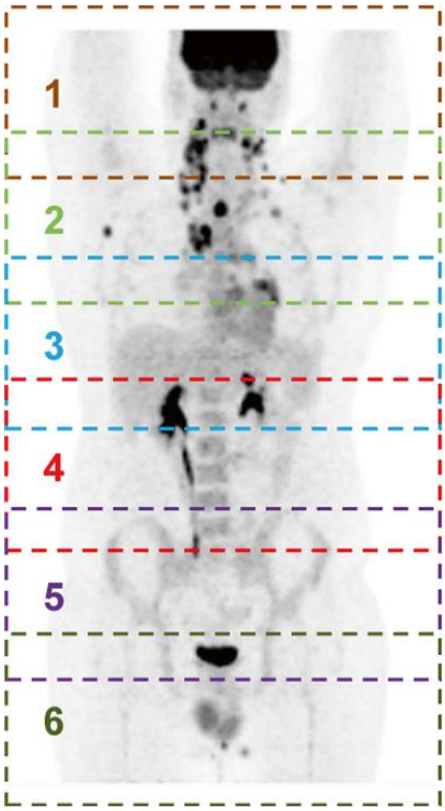


Figure 2.2 Multiple bed positions with small overlaps during PET scan (Retrieved from Tout *et al.*, 2016).

Positron emission tomography (PET) imaging begins with infusing radiopharmaceutical into patient body. The radiopharmaceutical is composed of a positron emitting radionuclide coupled to a chemical component known as the "tracer," which functions as a physiological analogue (Basu *et al.*, 2011b). The tracer is chosen to specifically target the metabolic activity of tumours. The radionuclide is employed to acquire images and serves as a source of radiation emission that is recorded by the

imaging scanner. The unstable isotopes known as radionuclides exhibit brief radioactive decay by positron emission. ^{18}F-Fluorodeoxyglucose (FDG) is the most common radiopharmaceutical that is employed in PET imaging. The hydroxyl group of glucose molecule present in FDG is replaced by positron-emitting radioisotope fluorine-18 (^{18}F) having 110 minutes half-life, to create the ^{18}F FDG tracer.

$$p \rightarrow n + e^{+} + \nu \qquad \text{(Eq. 2.1)}$$

Equation 2.1 represents the decay of a proton (p) into a neutron (n), a positron (e^{+}) and a neutrino (ν).

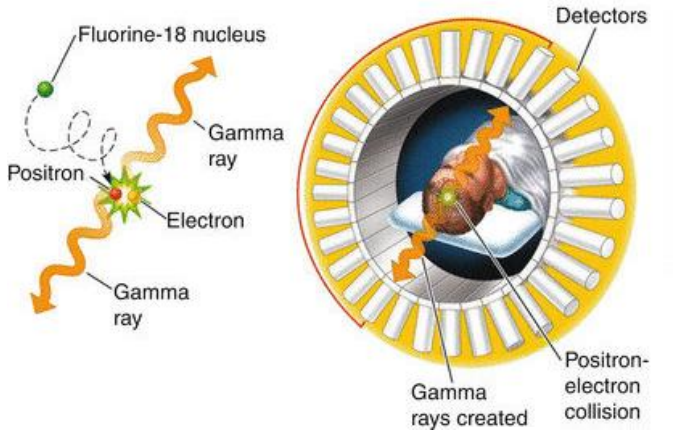
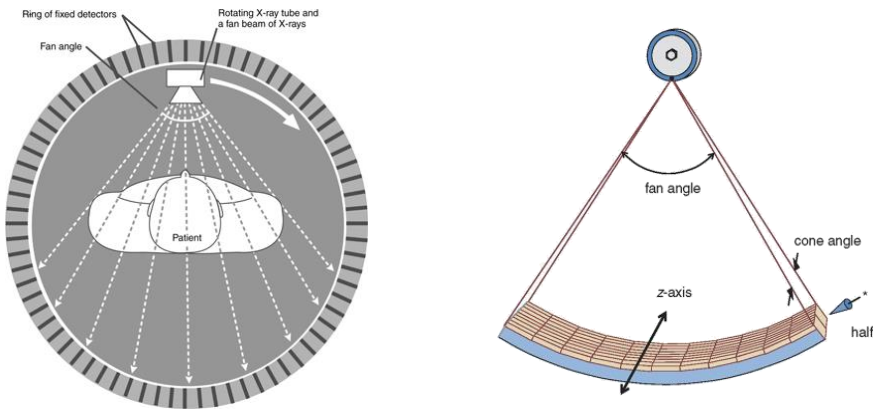


Figure 2.3 (a) The interaction between positron and electron during PET imaging; (b) Gamma ray collection by the detectors during PET imaging. (Retrieved from Tout *et al.*, 2016)

The positron (e^{+}) and a neutrino (ν), which are emitted from the nucleus with a continuous kinetic energy spectrum, receive the energy released during the conversion process. Depending on its energy, the positron travels a few millimetres through tissues after being released at a specific point in the body and encounters multiple scattering process. When a positron meets an electron (e) at the endpoint of its trajectory, they annihilate, and the remaining mass energy of the two particles is split into two photons that are each 511 keV in energy and approximately anti-parallel to one another. Figure

2.3 a represents this interaction electron and positron. In the PET scanner, several rings of radiation detectors are positioned circulating the body, exteriorly. These detectors note the detection timing of annihilating photons. Therefore, it detects photons escaping from the inside of patients.

Figure 2.3 b shows the basic structure of a cylindrical PET scanner, which comprises many rings of detectors mounted in an axial direction with a patient in the centre. The high-energy photons are converted into short pulses of visible light every time an annihilation photon strikes a single detector on a ring composed of a scintillating crystal. The optically connected crystal and photomultiplier tube (PMT), transform and multiplies the scintillation ray into an electrical pulse. Line of response (LOR) is line between two detectors that identifies each annihilation. Individual LOR of the two concurrent photons carry the information of determining the radiopharmaceutical position inside the body. During image reconstruction, image projection is formed from the lines of response. Accurate and precise PET image reconstruction involves adjustment of attenuation scatters, point spread function and non-uniform response for a uniform source.



(a) (b)

Figure 2.4 (a) Fan beam projection in CT imaging, (b) Multiple detector array in CT imaging (Retrieved from Bushberg and Boone, 2011).

In the CT imaging, the total body of medical interest is imaged as slices (cross-sections) for understanding of the physiological condition of any particular location. The determination of X- ray attenuation beam's coefficients in the investigated region of interest (ROI) provides the basis of CT. During CT acquisition of patient, X-ray attenuation is recorded in a plane perpendicular to the lateral axis of subject together with a number of lines in that direction. Afterwards, the attenuation coefficients (μ) map is reconstructed for the plane. When the photons of the X-ray travel through the body they interact with tissue or pass through the vacuum unaffected. In the case of interaction between tissue and X-ray photons, beam attenuation takes place by scattering or absorption (Council, 1996). The grey-scale values denote the calculated amount of attenuation. Hounsfield unit (HU) is the greyscale measurement in CT imaging (Goldman, 2007).

Generally, CT scanners employ fan-beam projection calculation by one else more arc positioned detector arrays with reference to the tube of x-ray (Bushberg and Boone, 2011). Figure 2.4 a represents the CT scan settings for implementing fan beam projection. The x-ray tube is at the top of the fan. As seen in Figure 2.4 a, the detectors are placed in a circle for covering a 360° view of the subject and the source moves in this circular path. Each detector measurement relates to a certain set of photons. In this geometry, the collection of rays is called a fan beam projection. The simultaneous acquisition of manifold slices is made possible by the employment of several X-ray detector arrays (Figure 2.4. b). For image reconstruction, cross-sectional measurements are treated as a slice of a complete ROI. When the slices are placed one after one the ROI becomes visible.

2.3 ^{18}F-FDG PET/CT radiomics in tumour heterogeneity quantification

Tumours in the head and neck region are very heterogeneous and therefore difficult to treat. This heterogeneity arises from multiple clonal sub-populations present in the unit composition of tissues within tumour having totally different properties (Padhani and Miles, 2010). The proliferation rate, expression of biosignature, ability to metastasize, and immunological traits are completely different in the multiple clonal sub-populations. The distinction in properties arises from the distinction in cell proliferation, blood vasculature, metabolic activity, pH, oxygenation level and necrotic areas present in sub-population at intervals tumour (Fouad and Aanei, 2017, Lin *et al.*, 2017). As a result, different spatial intensity patterns are formed from the intratumoural variations. The difference present in multiple clonal sub-populations within a tumour is called intratumoural heterogeneity (O'Connor *et al.*, 2015, Davnall *et al.*, 2012, Sala *et al.*, 2017). Figure 2.5 represents tumour heterogeneity.

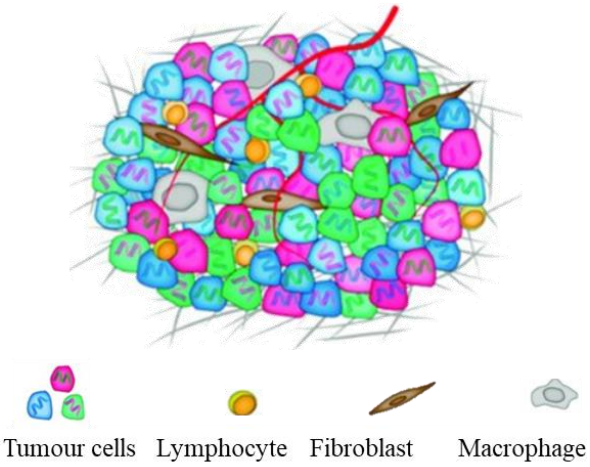


Figure 2.5 Conceptual illustration of tumour heterogeneity in head and neck cancer comprising tumour cells, lymphocytes, fibroblast and macrophages.

In solid cancers, tumour heterogeneousness creates resistance towards treatment leading to poor prognosis (Samanta and Semenza, 2018). However, the spatial and temporal variations are captured in ^{18}F-FDG PET/CT images in multi-level due to

underlying cellular microenvironments, tissue and anatomical landmarks within tumour (Basu *et al.*, 2011a, Yang and Knopp, 2011). Adequate information of tumour heterogeneity may lead towards precision medicine. Standard tumour heterogeneity study would provide patient specific molecular traits and these may accelerate tumour aggressiveness and sensitivity to therapeutic response identification prior to treatment. Investigation of tumour heterogeneity from histopathological samples (Biopsies) is challenging because inherent variation of sampling (Vaidyanathan *et al.*, 2019). Also, characteristic of a particular tumour region does not carry the information of the total tumour as tumours are mostly heterogenous (Dagogo-Jack and Shaw, 2018, Mroz *et al.*, 2013). An alternative to the invasive approach could be a deeper analysis of medical imaging.

Images contain more information than our eye can decipher (Gillies *et al.*, 2016). Researchers are actively investigating on biomarker measured from medical images as it holds the potential to quantify tumour heterogeneity. Particularly, morphology (shape, volume, eccentricity), histograms (variance, skewness, kurtosis) and texture traits hold information related to tumour heterogeneity (Willaime *et al.*, 2012). ^{18}F-FDG PET/CT radiomics analysis has the highest potential for characterizing tumour heterogeneity as it represents the spatial arrangement of grey-level intensities within a given volume of interest (VOI) as numerical descriptors. Presently, there are five leading radiomics metrices investigated by the clinical imaging researchers. These five texture units are grey level co-occurrence matrix (GLCM), grey level run length matrix (GLRLM), grey level size zone matrix (GLSZM) and neighbourhood grey tone distinction matrix (NGTDM) (Ang *et al.*, 2010, Cheng *et al.*, 2015, Vakkila and Lotze, 2004, Proskuryakov and Gabai, 2010, Ahn *et al.*, 2016). Section 2.4 is dedicated for the features definition and methodology of computation details.

2.4 Definition of radiomics

Advanced quantitative analysis of medical images holds the potential to capture the genomic heterogeneity of aggressive tumours that are reflected in the heterogenous tumour metabolism and anatomy. The term 'radiomics' represents the advanced computational analysis of diagnostic images. Radiomics is the study of tumour characteristics through the generation of higher order spatial data extracted from medical images (Hatt *et al.*, 2017, Yip and Aerts, 2016). Radiomics analysis is a "top to bottom" approach for understanding the underlying tumour biology. Substantial computational textural traits are mined from clinical images in the radiomics analysis process. These extracted features are associated with different tumour phenotypes. In the past years, the emerging field of radiomics experienced an exponential growth. Radiomics is in its early development stage needing standardisation and validation. However, the use of high-order imaging biomarkers dedicated to the quantification of intratumoural heterogeneity holds great promise for better tumour aggressiveness assessment and subsequent treatment personalization.

The workflow of radiomics analysis for its translation into clinical settings is illustrated in Figure 2.6. Feature extraction from the region of interest (ROI) in the first and foremost step (Figure 2.6 a) (Mayerhoefer *et al.*, 2020). A complete description of radiomic features mentioned in the thesis is listed in appendix A. Afterwards, Spearman's correlation coefficient, Pearson correlation coefficient, concordance correlation coefficient or interclass correlation are evaluated so that coefficient robustness and the reproducibility can be determined (Figure 2.6 b). Depending on the robustness and reproducibility results, the optimum features are nominated and redundant features are omitted (Figure 2.6 c) (Cutaia *et al.*, 2021). Artificial intelligence-based models are developed for disease prediction, prognosis and diagnosis

for incorporating automatization into the clinical practice (Figure 2.6 d) (Mayerhoefer *et al.*, 2020). The final outcome will be application of radiomics for achieving clinical outcomes such as survival prediction and prognosis prediction (Figure 2.6 e) (Li *et al.*, 2018).

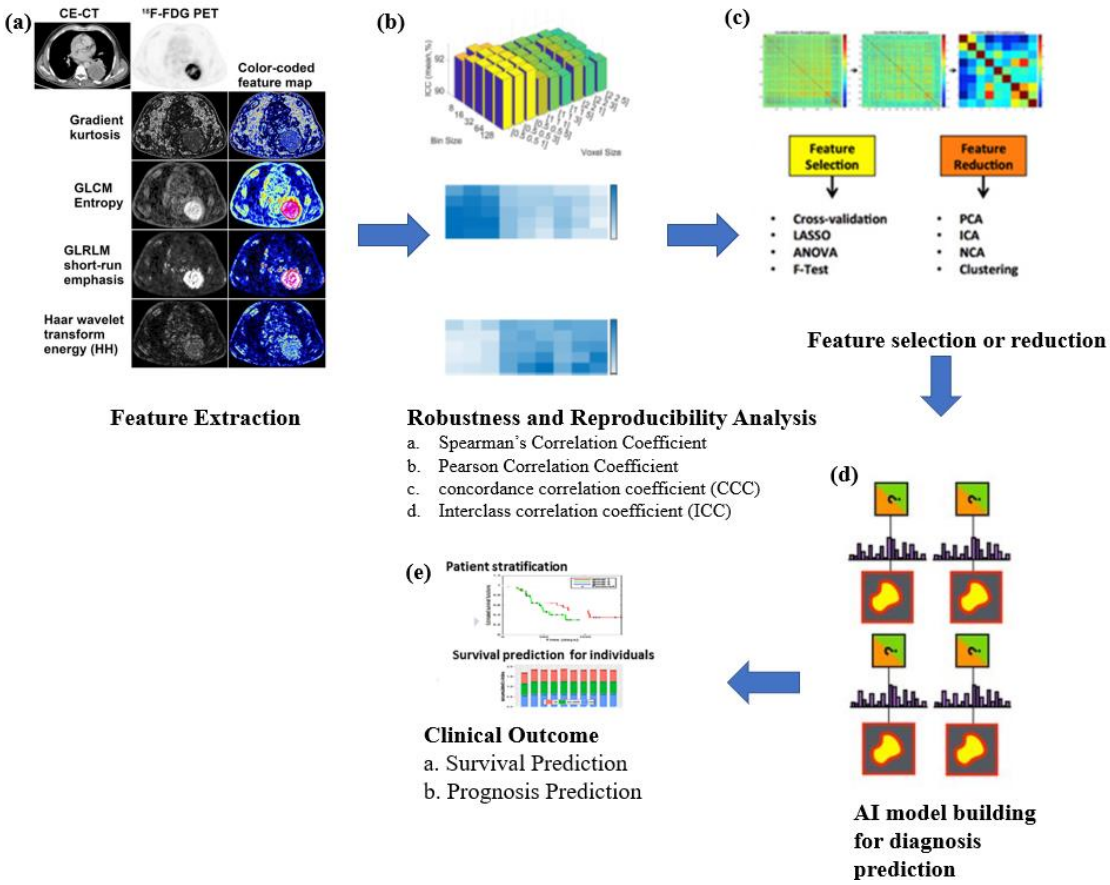


Figure 2.6 The workflow of radiomics analysis for its translation into clinical settings.

Radiomic characteristics are derived from ROI. Thousands of distinct features may presently be obtained by using various mathematical algorithms and operations, even using artificial intelligence (Figure 2.7). Manual features and automated features are the two main categories of radiomics features. Through the use of some appropriate mathematical functions, manual properties are achieved. Shape and texture characteristics are the most prevalent ones. Deep learning features are acquired

intuitively by training on massive image samples. The features employed in this thesis are briefly mentioned below without claiming to be all-inclusive. "Image biomarker standardisation initiative" (IBSI) document extensively explains each feature (Hatt *et al.*, 2018).

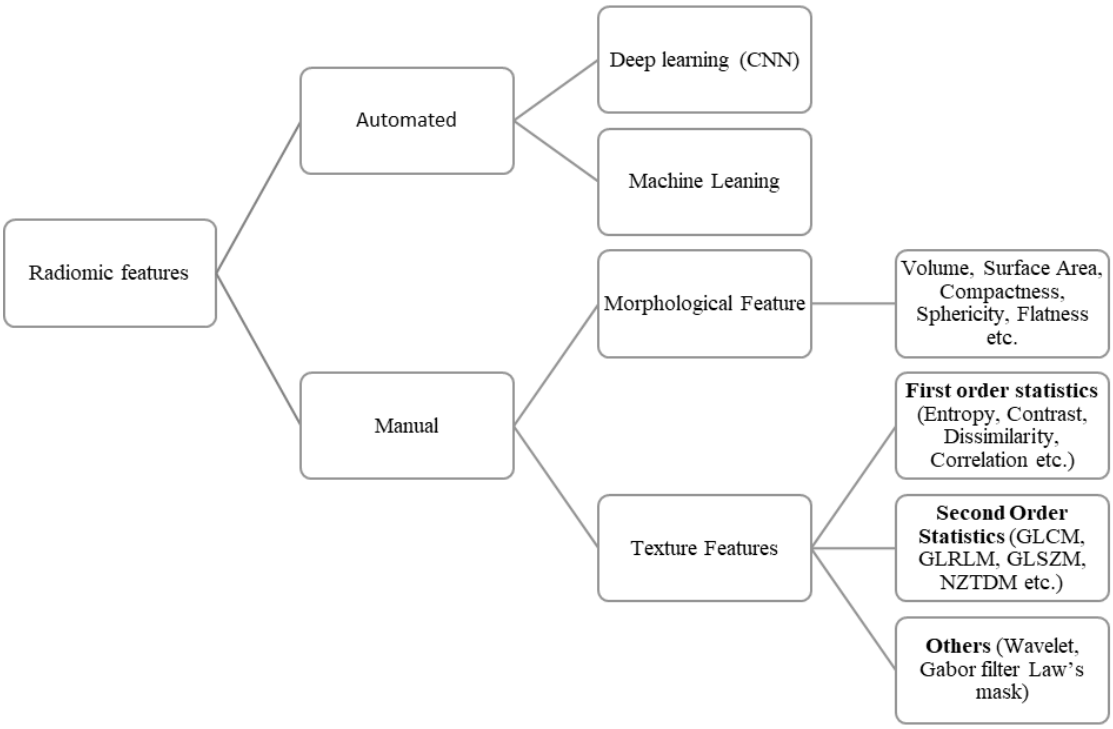


Figure 2.7 Categories of radiomic features.

In radiomics analysis, textures are a central type of features that can be extracted from a tumour ROI. Other types of features include morphological and histogram-based features (Appendix A). However, textures remain the core of radiomic feature computation given their higher-order characterisation of spatial patterns in imaging volumes. In this thesis, texture features from five major categories were extracted: a) Global; b) Grey-Level Co-occurrence Matrix (GLCM) features; c) Grey- Level Run-Length Matrix (GLRLM) features; d) Grey-Level Size Zone Matrix (GLSZM) features; and e) Neighbourhood Grey-Tone Difference Matrix (NGTDM) features. The primary and important step towards the computation of the different texture features from these

four categories is to calculate a matrix P summarizing the neighbourhood properties of interest (differently for each category). Thereafter, different mathematical operations can be applied to the different matrices to obtain the final texture features.

First order grey level statistics are known as the Global texture features. In the ROI, the frequency distribution per voxel intensity is expressed by Global texture features. Histogram of intensity frequency are employed for Global texture features determination and include mean, skewness and kurtosis. Among the second order grey level statistics, grey level co-occurrence matrix (GLCM) represents the neighbourhood probability of pixel intensity i and pixel intensity j (Figure 2.8). Along a specific orientation and at a specific length, GLCM denotes how two voxels "co-occur" with relation to one another. GLCM has a neighbourhood of 26 connected and so 13 distinct direction vector are generated in 3D with neighbouring length of 1 (Zwanenburg *et al.*, 2020). In 2D image, 8 connected neighbourhood with 4 distinct direction vectors is present with neighbouring length of 1. As a result, a ROI contains 13 distinct GLCMs for each of the 13 directions in a three- dimensional framework in a neighbourhood length of one.

GLCM features include contrast, energy, entropy, correlation, homogeneity and dissimilarity, sum average. Higher value of sum average and homogeneity is correlated to enhancing lesion. Randomness of intensities in ROI in defined by entropy, where higher randomness results in higher value of entropy. Enhancing lesions tend to have low value of entropy. Value of correlation is higher for linear structures such as honeycomb patterns present in normal median nerves (Ardakani *et al.*, 2022). Based on the relationships between three or more voxels, various texture matrices are used to derive the higher-order texture characteristics. The grey level run length-based matrix (GLRLM) measures the pixels length of connected pixels with equal value to determine

grey level path length (Figure 2.8). The (i, j)th element of a grey level run length matrix P(i, j |Θ), indicates the path value with the grey level i and length j that are present in the picture at angle Θ (Zwanenburg *et al.*, 2020). GLRLM features includes short run emphasis (SRE), long run emphasis (LRE), run percentage (RP), grey level nonuniformity (GLN), run length nonuniformity (RLN) to name a few.

CT scans of COVID infection, present presented fine textures in the ROI. This fine textures are quantified by higher value of short run emphasis and low value of long run emphasis (Ardakani *et al.*, 2022). In the case of COVID-19 infection diagnosis, GLN is smaller and RLN is larger compared to non-COVID-19. Similar to correlation of GLCM, RP tends to have lower value while capturing linear structures. In an image, grey level regions are denoted using the grey level size zone-based matrix (GLSZM). Collection of linked voxels with the same intensity of grey is referred to grey level zone (Figure 2.8). The number of zones in the image that have grey levels i and length j is represented by the (i, j)th element of the grey level size zone matrix P(i, j) (Zwanenburg *et al.*, 2020). Nodes are composed of solid and cystic components. Cystic components are more heterogenous compared to solid components. As a result, small zone emphasis is greater and large zone emphasis are smaller of cyst compared to solid.

The neighbourhood grey tone difference matrix (NGTDM) measures the distinction in grey values among a given value and the mean of its neighbouring value within a certain distance, d. The matrix contains the total of the absolute distinctions of i grey level (Figure 2.8) (Zwanenburg *et al.*, 2020). For example, coarseness describes the inconsistency of grey value within ROI. Contrast represents spatial change of grey values and busyness represents rate of intensity shift within ROI. Healthy and entrapped median nerves can be distinguished utilizing NGTDM features. Coarseness is higher in healthy nerves compared to entrapped median nerves (Ardakani *et al.*, 2022).

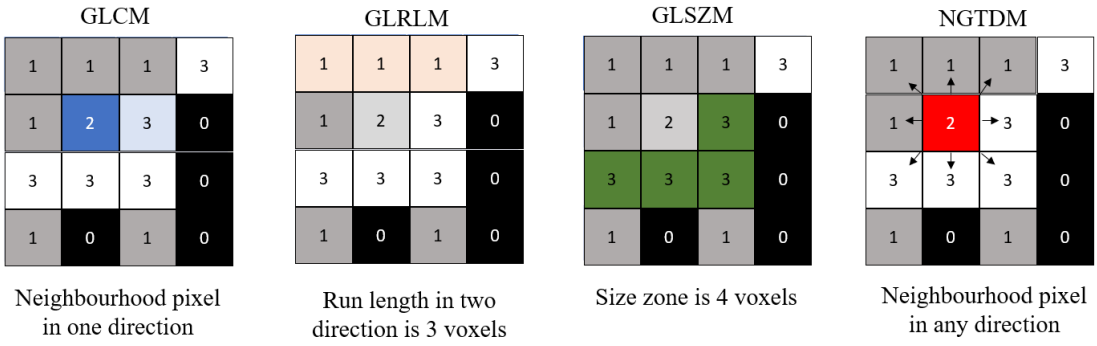


Figure 2.8 Schematic view of the radiomics feature.

2.5 ^{18}F-FDG PET/CT radiomics for head and neck cancer

Radiomics textures features hold the potential for head and neck tumour tissue characterization, prediction of therapeutic response and monitoring prognosis. Busyness, contrast and coarseness of the NGTDM feature group are found to have differentiation capability. These features recognised tumour tissue from the healthy tissue (Yu *et al.*, 2009). Tumour tissues presents higher contrast and lower busyness and contrast compared to healthy tissue in the PET/CT image. Texture features can also predict therapeutic outcome (El Naqa *et al.*, 2009). A study performed by El Naqa *et. al.* demonstrated that first and second order features characterises tumour uptake in the microenvironment and conveys information about treatment resistance. In another study on oesophageal cancer, researchers found that GLCM features can classify chemotherapy responders from non-responders (Tixier *et al.*, 2011). The heterogeneity detected in image is due to difference in tumour tissue component arrangement (Henriksson *et al.*, 2007).

^{18}F-FDG uptake is higher in tumour cells compared to stroma and necrosis. The association of texture feature with tumour characteristics is complex hence careful investigation is required to establish reliable and accurate relation between these two parameters. Researchers have focused on repeatability and reproducibility of radiomics

features. GLCM features homogeneity, correlation and GLRLM features long run emphasis and short run emphasis were concluded to be robust compared to SUV (Shiri *et al.*, 2017, Pfaehler *et al.*, 2019). The metrices used to report reproducibility were mainly inter-class correlation coefficient (ICC), concordance correlation coefficient (CCC) and coefficient of variation (COV). The clinical studies found in the literature mainly concentrated on the dependency of the radiomics features on voxel discretization and feature extraction parameters (Belli *et al.*, 2018, Lv *et al.*, 2018). All of the studies were coherent with a strong correlation between feature reproducibility and image quantization. According to one study, the features of GLSZM were the least reproducible. Impact of tumour delineation variation on features was also focused in two studies.

The researchers reported a strong correlation between tumour delineation variation and radiomics features. GLSZM feature zone percentage was sensitive to tumour delineation variation and GLCM features entropy, dissimilarity, and GLSZM feature high intensity large area emphasis were robust against tumour delineation variation. Drawing any conclusion in reproducibility and repeatability of the radiomics features is challenging due to the large variation of tumour types analysed and diagnostic settings. However, most studies found the first order features GLCM and GLRLM robust and GLSZM least robust. Following a standardised process might eliminate this challenge. The future studies should include detail information about the matrices and cut-offs used to classify the features into a degree of reproducibility that is absent in the studies found in the present literature.

Another group of researchers focused on the development of radiomics signatures to improve predictive models for specific cancers (Aerts *et al.*, 2014). The study included 440 features from 1,019 patients with either head and neck cancer