Dynamic and Quantitative Radiomics Analysis in Interventional Radiology

LIWEI ZHAO

Sid:490076071
Supervisor: XIUYING WANG

A thesis submitted in fulfilment of the requirements for the degree of Master of Philosophy in the Faculty of Engineering & Information Technologies

School Of Computer Science
University Of Sydney

AUGUST 2022

Student Plagiarism: Compliance Statement

I certify that:

To the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Name: LIWEI ZHAO

Abstract

Interventional Radiology (IR) is a subspecialty of radiology that performs invasive procedures driven by diagnostic imaging for predictive and therapeutic purpose. The development of artificial intelligence (AI) has revolutionized the industry of IR. Researchers have created sophisticated models backed by machine learning algorithms and optimization methodologies for image registration, cellular structure detection and computer-aided disease diagnosis and prognosis predictions. However, due to the incapacity of the human eye to detect tiny structural characteristics and inter-radiologist heterogeneity, conventional experience-based IR visual evaluations may have drawbacks.

Radiomics, a technique that utilizes machine learning, offers a practical and quantifiable solution to this issue. This technology has been used to evaluate the heterogeneity of malignancies that are difficult to detect by the human eye by creating an automated pipeline for the extraction and analysis of high throughput computational imaging characteristics from radiological medical pictures. However, it is a demanding task to directly put radiomics into applications in IR because of the heterogeneity and complexity of medical imaging data. Furthermore, recent radiomics studies are based on static images, while many clinical applications (such as detecting the occurrence and development of tumors and assessing patient response to chemotherapy and immunotherapy) is a dynamic process. Merely incorporating static features cannot comprehensively reflect the metabolic characteristics and dynamic processes of tumors or soft tissues.

To address these issues, we proposed a robust feature selection framework to manage the high-dimensional small-size data. Apart from that, we explore and propose a descriptor in the view of computer vision and physiology by integrating static radiomics features with time-varying information in tumor dynamics. The major contributions to this study include:

Firstly, we construct a result-driven feature selection framework, which could efficiently reduce the dimension of the original feature set. The framework integrates different feature selection techniques to ensure the distinctiveness, uniqueness, and generalization ability of the output feature set. In the task of classification hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (ICC) in primary liver cancer, only three radiomics features (chosen from more than 1, 800 features of the proposed framework) can obtain an AUC of 0.83 in the independent dataset. Besides, we also analyze features' pattern and contributions to the results, enhancing clinical interpretability of radiomics biomarkers.

Secondly, we explore and build a pulmonary perfusion descriptor based on 18F-FDG whole-body dynamic PET images. Our major novelties include: 1) propose a physiology-and-computer-vision-interpretable descriptor construction framework by the decomposition of spatio-temporal information into three dimensions: shades of grey levels, textures, and dynamics. 2) The spatio-temporal comparison of pulmonary descriptor intra and inter patients is feasible, making it possible to be an auxiliary diagnostic tool in pulmonary function assessment. 3) Compared with traditional PET metabolic biomarker analysis, the proposed descriptor incorporates image's temporal information, which enables a better understanding of the time-various mechanisms and detection of visual perfusion abnormalities among different patients. 4) The proposed descriptor eliminates the impact of vascular branching structure and gravity effect by utilizing time warping algorithms. Our experimental results showed that our proposed framework and descriptor are promising tools to medical imaging analysis.

Publications

Journal Papers:

(1) Jiang, C., **Zhao, L**., Xin, B., Ma, G., Wang, X., & Song, S. (2022). 18F-FDG PET/CT radiomic analysis for classifying and predicting microvascular invasion in hepatocellular carcinoma and intrahepatic cholangiocarcinoma. Quantitative imaging in medicine and surgery, 12(8), 4135.

(2) **Zhao, L**., Zhen, C., Zhou, Y., Wang, X. Quantitative Dynamic Descriptor in Regional Pulmonary Perfusion Visualization and Anomaly Detection. Quantitative Imaging in Medicine and Surgery, 2022. (Under preparation for submission.)

Book Chapter:

(1) Xin, B., **Zhao, L**., Li, J., Wang, X. (2021). Chapter: Artificial Intelligence in Interventional Radiology. Michael, B., Ulas B., Omer A., Xu Z., Daljeet C., Nasim P., and Alexandra G. Artificial Intelligence in Clinical Medicine. (Prepare an original and previously unpublished contribution.)

Authorship Attribution Statement

Journal Papers:

Chapter 3 of this thesis is published as [1]. I am the co-first author of this journal paper, and my major contributions include:

- Co-design the study with physicians
- Algorithms and experiments design and implement
- Data extraction, data analysis and results interpretation
- Main part of the manuscript writing

Dr. Jiang Chunjuan, the co-first author and industry collaborator, provision of study material or patients, provision of professional opinions in data analysis and discussion, and participation in manuscript revision. Dr. Bowen Xin, the third author, provision of opinions on algorithms and experiments design and participation in manuscript revision. Dr. Guang Ma, the industry collaborator, provision of study material or patients and participation in manuscript revision. Dr. Shaoli Song, the corresponding author, conception and design, paper revision support and administrative support. A/Prof. Xiuying Wang, the corresponding author, conception and design, paper revision support and administrative support.

Chapter 4 of this thesis is under preparation for submission for Quantitative Imaging in Medicine and Surgery. I am the first author of this journal paper, and my major contributions include:

- Co-design the study with physicians
- Algorithms and experiments design and implement

- Data extraction, data analysis and results interpretation
- Main part of the manuscript writing

Dr. Chaojie Zheng, the second author and industry collaborator, conception and design, provision of study material or patients, provision of opinions on experiments design, participation in manuscript revision. Dr. Yun Zhou, the industry collaborator, conception and design, paper revision support and administrative support. A/Prof. Xiuying Wang, the corresponding author, conception and design, paper revision support and administrative support.

Book Chapter:

Chapter 1 and Chapter 2 are partly based on 'Chapter: Artificial Intelligence in Interventional Radiology' in *Artificial Intelligence in Clinical Medicine*. My original unpublished contribution include:

- Wrote the section: Pre-Operative Radiomics with Machine Learning for Diagnostic IR.
- Participated in revision

Dr Bowen Xin, wrote the sections include Abstract, Introduction and Multimodal data fusion and interpretation for IR, participated in revision. Jiaru Li, wrote the section of Data balancing with machine learning for prognostic IR, participated in revision. A/Prof. Xiuying Wang revised the manuscript and organized the structure.

Authors list: The authors who contributed most to the idea, algorithm and experimental design were listed as the first authors, and other authors were listed based on their contributions to the study.

These chapters only included my contributions to papers. In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Student Name: Liwei Zhao

Signature:

Date: 08/30/2022

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Supervisor Name: Xiuying Wang

Signature:

Date: 08/30/2022

Acknowledgements

I am very grateful to all people that support my Master of Philosophy study and thesis. The journey of my study is quite difficult. Due to Covid-19 travel restriction, I couldn't book ticket to Sydney and had to strand in the U.S. for a very long time. I experienced a tough time when I tried to work from home. Nevertheless, I achieved more than I thought. My supervisor and fellow students gave me a lot of attention and support. With all these people's encouragement, I continue my research even though it takes much longer than others.

First, I want to express my sincere thanks to my supervisor Associate Prof. Xiuying Wang, for her invaluable guidance, efforts, and supports. She set a great model by her rigorous attitude and enthusiasm towards research. She also provided us with a lot of research opportunities, allowing us to work with multiple hospitals.

I would like to thank Prof. Yun Zhou and Dr. Zheng in United Imaging Healthcare (Shanghai, China) for their professional insight and guidance in my academic research.

Besides, I want to thank Dr. Xin, Dr. Hao, Lin Zhang, Peter and Yanxiao for their encouragement and support.

Finally, I wish to say thank you to my parents and other family members for unconditionally supporting during my MPhil Study.

Table of Contents

List of Figures

List of Tables

CHAPTER 1. Introduction

1.1 Research Motivations

Future breakthroughs in medical decision-making procedures will demand an ever-closer relationship with information systems [2]. Artificial intelligence (AI) has the potential to advance science in radiology, and it has recently been applied to Interventional Radiology (IR), a medical sub-domain that performs a variety of invasive operations under the guidance of medical imaging for diagnostic and prognostic purposes.

The contributions of IR start with the initial diagnosis and planning of cancer, then extend to the management of malignancy (procedural targeting, monitoring and control) and postprocedural assessment [3]. Research shows that a growing number of IR-based histopathological diagnoses are made through biopsies using minimally invasive methods [4]. IR also plays an important role in intraprocedural: more than two million central access devices are implanted annually in the US, which traditionally implanted by surgeons are now often sited using IR methods [5]. Besides, malignancy-related issues, which may arise from the disease itself or as a side consequence of treatment, have also assumed an important position in IR care [6].

The visual analysis of medical images is an indispensable technical means in IR, which has many applications like lesion recognition in diabetic retinal fundus images [7]. However, due to the incapacity of the human eye to detect tiny structural characteristics and inter-radiologist heterogeneity, physicians cannot offer accurate diagnosis based on conventional experience-based IR visual evaluations.

To provide a feasible and quantitative solution to address similar issues, radiomics, supported by big data technology and computer-aided diagnosis, are developed for applications of IR,

which are not limited to tumor detection, differential diagnosis, pathological typing and grading, prediction and evaluation of chemotherapy efficacy [8].

The radiomics pipeline generally consists of:

(1) Medical image acquisition,

(2) Segmentation and feature extraction,

(3) Biomarker mining and interpretation.

Biomarker mining is a major component in these steps, which is actively being improved using machine learning techniques. However, there are two inevitable challenges of biomarker mining in medical imaging applications. One is high-dimensional small-size problem, the other is the integration of dynamic image information. We will give more details in the following sections.

1.2 Challenges in biomarker mining

1.1.1 High-dimensional small-size data problem

Quantitative imaging biomarkers are usually based on mathematical definitions, which require high-throughput analyses. Besides, it's commonplace to encounter high-dimension small-size medical imaging dataset because of the confidentiality agreement, law issue and various acquisition protocols [9], making it difficult to find the informative biomarkers.

Larger feature space is a strong guarantee for achieving an accurate output, but it will unavoidably result in the presence of several duplicate and inappropriate features. Besides, it is more likely to induce the Curse of Dimensionality [10], that is, as the dimensionality grows, the time complexity rises significantly while the efficiency of the algorithm plummets. Besides,

conventional machine learning methods may fail because of the limited sample size, and it's more likely to encounter the overfitting problem.

Chapter 3 of this study was motivated by limitations mentioned above. We proposed a result-driven feature selection framework for prognostic prediction of primary liver cancer. Specifically, the algorithms were designed by three criterions to address the high-dimensional small-size data problem in biomarker mining applications. The first criterion is correlation, which means the selected features should be relevant with clinical labels. Filter methods such as Random Forest tree importance and statistical analysis, are used here to exclude noisy and irrelated features. The second criterion is non-redundancy. Introducing highly correlated features could unnecessarily increase the dimensionality and cause overfitting problem. Dynamic feature selection algorithms are utilized to exclude highly correlated features to ensure the discovered feature set is discriminative. The last standard is robustness. Nested validation strategies were applied to ensure the reproducibility and generalization ability of discovered biomarkers through different imaging devices and institutions. The framework could reduce features dimensions in an effective way (only three features were selected from more than 1800 features) and make feature combinations informative (AUC>0.8).

1.1.2 Integrates dynamic medical image information

Discovering more radiomics features is important for evaluating prospective influencing elements that may be effective in diagnostic and prognostic predictions. Many efforts have been made to explore and extract new radiomics features. No matter the feature extraction approach used, the results provided by the current methodologies are mostly dependent on static medical images. However, the formation and development process of cancer are a dynamic process and cannot be comprehensively captured solely by static features.

Studies show that dynamic imaging techniques could provide quantitatively metabolic and pharmacokinetic analysis for different kinds of cancers. For example, dynamic contrast-enhanced magnetic resonance (DCE-MR) imaging [11] is a technique for acquiring images by the administration of contrast, which improve the sensitivity for tumors detection. Dynamic PET/CT [12] is a technique that acquiring data in a dynamic way, collecting tracer dynamic information that typical static methods do not capture. It can avoid the influence of uptake kinetics, drug injection imaging time and BMI factors, and can be used to quantitatively evaluate tumor metabolism. In some clinical scenarios, taking medical images of different therapeutic phases may also be useful to evaluate patient's response to immunotherapy or chemotherapy. However, conventional analysis of dynamic medical images is based on kinetic modelling and pharmacokinetics, demanding sophisticated parametric calculation and professional software [13].

Recently, Delta Radiomics [14] and Dynamic Radiomics [15] are proposed based on static images acquired from subsequent examinations, which provide potential applications in the diagnosis and prognosis evaluation of liver and lung diseases. However, compared to modelling techniques based on dynamic imaging, these static-image-based methods hardly capture the pharmacokinetics characteristics of tumours, lacking in interpretability of relationships between metabolic variations and pathological processes.

With the development of high-res dynamic whole-body PET/CT Scanner [16], dynamic-image-based radiomics is a promising research topic with the potential to provide a computational-efficient and pathologically explainable solution to precise diagnosis and treatment planning.

Chapter 4 in this study was motivated by the topic mentioned above. Inspired by dynamic descriptor in the computer vision area. We build radiomics-based dynamic descriptor on 18F-

FDG dynamic whole-body PET/CT images. The descriptor can reflect visually uneven perfusion of patients' lung in a quantitative way. It also has potential to provide a more comprehensive assessment of patients' pulmonary diseases.

1.2 Thesis Contributions

The aim of this thesis is to give solutions to radiomics-based medical image biomarker analysis. Our major contributions to this thesis can be summarized as below:

A result-driven feature selection framework. It's common to encounter high-dimensional small size problem in medical imaging dataset. We construct a result-driven feature selection framework, which could efficiently reduce the dimension of the feature space and simultaneously avoid overfitting. The framework integrates different feature selection techniques to ensure the distinctiveness, uniqueness, and generalization ability of the output feature set. In the task of phenotype classification in primary liver cancer, only three features (choose from more 1800 features) can obtain an AUC of 0.83 in the independent dataset. Besides, we also analyze features' pattern and contribution to the results, enhancing clinical interpretability of radiomics biomarkers.

A computer-vision-based dynamic lung perfusion descriptor. Considering that medical image biomarker scarcely integrates temporal information, we explore and build a perfusion index based on FDG dynamic PET in the view of computer vision descriptor. Novelties include 1) propose a physiology-and-computer-vision-interpretable descriptor construction framework by the decomposition of spatiotemporal information into three dimensions: shades of grey levels, textures, and dynamics. 2) The quantitative comparison of descriptor intra and inter patients is feasible, making it possible to be an auxiliary diagnostic tool in pulmonary function assessment. 3) Compared with traditional PET metabolic biomarker analysis, the proposed descriptor

incorporating image's temporal information, which enables a better understanding of the time-various mechanisms and a more sensitive detection method of visual perfusion abnormalities among different patients. 4) The proposed descriptor eliminates the impact of vascular branching structure and gravity effect by utilizing time warping algorithms.

1.3 Thesis Organization

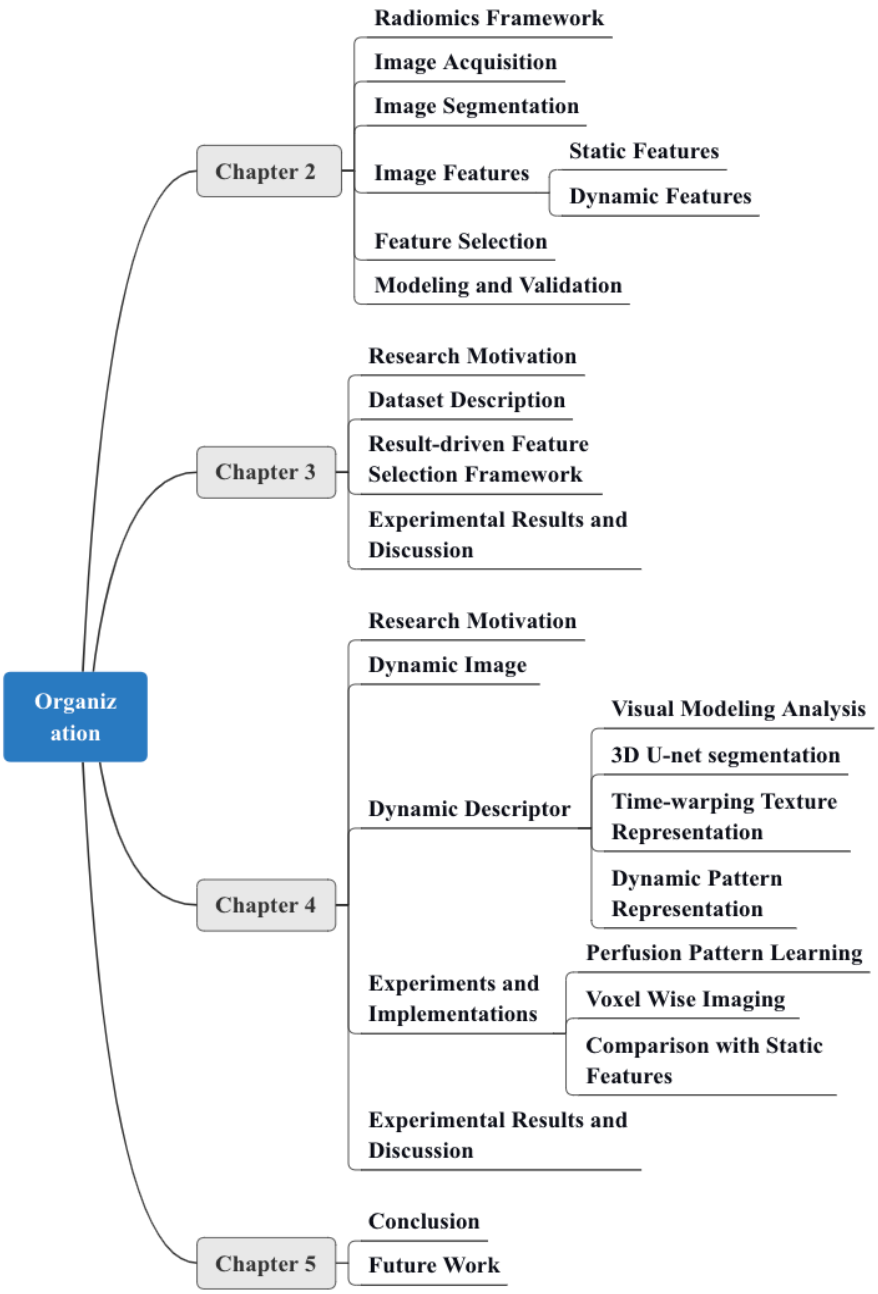

Figure 1-1 Thesis Organization

The remainder of this thesis is organized as follows: Chapter 2 provides a literature review of the recent progress in radiomics, feature selection method and the development in image

descriptors. Chapter 3 presents our radiomic analysis for classifying and predicting microvascular invasion in hepatocellular carcinoma and intrahepatic cholangiocarcinoma, which utilizing the feature selection framework we proposed. Chapter 4 presents the construction and analysis of computer-vision-based lung perfusion descriptor. Chapter 5 is conclusion and future work for medical image biomarker analysis.

CHAPTER 2. Literature Review

In this chapter, we first summarize the literature review on general framework of artificial-intelligence-based radiomics, which includes image acquisition, image segmentation, image features mining, feature selection, model construction and evaluation. Then, recent progress and challenges are discussed for each specific topic.

Image acquisition and segmentation are two important pre-processing steps deciding the inputs in biomarker mining algorithms. With the development of imaging technologies, more accurate and reliable information could be extracted, giving multi-dimensional training data for machine learning models. For medical image features, it can be categorized into static features and dynamic features according to whether the time variable is introduced. Image Biomarker Standardization Initiative[17] gives a comprehensive definition of static medical features, which major interest in conventional computer-vision-based descriptors. Traditional features well depict the areas of interest to physicians in many clinical scenarios, and furthermore provide analytical information that cannot be recognized by the naked human eyes. However, medical images of one single time point allow only limited information, especially for studies on pharmacokinetics of tumours. Lately, studies on dynamic medical image features mainly focus on Delta Radiomics [14] and Dynamic Radiomics [15], compensating the shortcomings of static features, strengthening the clinical interpretability of radiomics biomarkers. Besides, high-dimensional small-size data is easy to encounter in biomarker mining. Appropriate and interpretable feature selection methods are particularly important because it not only diminished impact of curse of dimensionality and algorithm's time complexity but retain the biomedical characteristics of original images. Methods on feature engineering are basically summarized as four major categorizes by algorithms' evaluation functions, which include filter method, wrapper

method, embedded method, and ensemble method. Modelling and evaluation are established to explore and analyze potential relationships between discovered biomarkers and its biological information. A trusted as well as robust machine learning framework is needed in this step.

This chapter is organized as follows: Radiomics Framework is summed up in section 2.1, Image Acquisition and Image Segmentation are introduced in section 2.2 and 2.3. Literatures in related to Image Features and Feature Selection are summarized in section 2.4 and 2.5. Finally, we give a presentation on Modelling and Evaluation on section 2.6.

2.1 Radiomics Framework

Radiomics studies are widely applied in survival prediction, tumor treatment response assessment, viral status analysis, etc. For instance, Microvascular Invasion (MVI), an important biomarker for assessing recurrence of liver cancer, is typically determined by biopsy, which may raise the patient's risk and may not reveal the whole tumor status [18]. Radiomics, on the other hand, might use non-invasive imaging features to aid in the detection of MVI. Figure 2-1 illustrates the radiomics analysis methodology for this specific task.

First, high throughput radiomics features are extracted from volume of interests (VOIs) on images. The machine learning classifier is then developed and trained based on the specified features. Third, pre-trained models predict the MVI status. Lastly, model interpretability

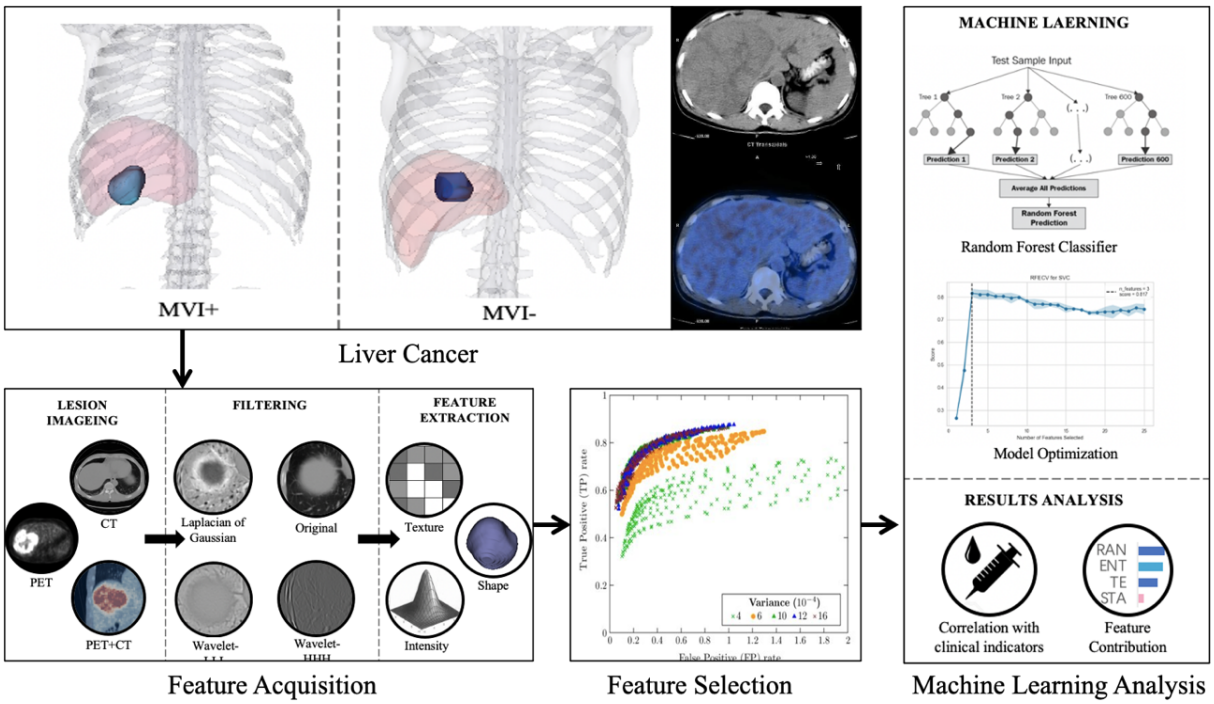

Figure 2-1 Workflow of Radiomics

strengthening via studying the connection between radiomics features and physiological indicators. We will cover the details of these steps in the following sections.

2.2 Image Acquisition

In most circumstances, medical imaging relates to a range of procedures for non-invasively creating pictures of the human body's anatomical structure, metabolic data, and malignant symptoms. Essentially, medical imaging might be viewed as the solution to mathematical invertible issues, which means that the cause (the metabolism of tumors) may be deduced from the result (the pixel value). Three widely used medical imaging techniques (Table 2-1) are described in the following paraphrase.

Computerized Tomography (CT). Using a revolving X-ray tube and a series of detectors, a CT scanner analyzes X-ray attenuations by various tissues inside the body. The computer then applies tomographic reconstruction methods to the numerous X-ray observations to generate tomographic pictures of a body. CT images are suitable for anatomical imaging for its high resolution in human body's structure [19]. CT also has several limitations, including its high radiation as well as the low contrast in tissue with high metabolism.

Magnetic Resonance Imaging (MRI). MRI employs a powerful magnetic field to create images that represent the anatomy and physiology of the body. MRI does not employ X-rays or any other kind of ionizing radiation, which makes it a painless and safer solution for medical analysis. Besides, MRI gives superior contrast for imaging soft tissue, e.g. in the brain or cardiac [20] [21]. However, the drawbacks of MRI include its high cost and patients may perceive it uncomfortable due to the long acquisition time, claustrophobic environment, and loud noise.

Table 2-1 Comparison for different medical images

Modality	**CT**	**MRI**	**PET**
Mechanism	X-ray attenuation	Strong magnetic field	Photon emission
Spatial Resolution	<1 mm	1-2mm	<1 mm
Acquisition Time	20-30 min	30-60 min	30-90 min
Advantages	1) Good structure imaging 2) Short acquisition time	1) Safer and more comfortable 2) Better contrast in soft tissue than CT	1) High sensitivity in metabolic activities
Disadvantages	1) Low contrast in tissue 2) High radiation dose exposure	1) Expensive 2) Uncomfortable acquisition method	1) Expensive 2) High radiation dose exposure 3) Low resolution
Major Applications	1) Bone imaging	1) Brain imaging 2) Cardiac imaging	1) Tumor imaging

Positron Emission Tomography. Different from CT and MRI, PET [22] utilizes radioactive materials to identify changes in metabolic activity and other physiological processes, including blood flow, absorption, and regional chemical properties. It is possible for PET imaging to detect metabolic processes taking place in the target body regions through the changes in biochemical substances, enabling the detection of disease progression, which may not be visible with anatomical imaging. PET may be used in a wider range of clinical and scientific situations than CT and MRI. The approach is widely used in the fields of oncology, neurology, and biodistribution studies to scan for tumors and the existence of metastases [23]. However, the high initial cost and low resolution of a PET scanner are its main disadvantages.

PET/CT & PET/MR. A composite PET/CT system, which combines the diagnostic information of two modalities, was introduced in 2000, replacing traditional PET scanners. By combining structural and functional tomographic imaging modalities, hybrid imaging technology provides superior information.

Dynamic whole body Positron Emission Tomography (dPET). Dynamic whole-body positron emission tomography (dPET) method is often utilized in scientific research projects because it requires special software and takes a long time to perform. dPET/CT has shown improvements over conventional static PET/CT by using methodologies of quantitatively extracting biological and kinetics knowledge from a tracer in tissue [24]. The major improvements are:

(1) With a dynamic image (Figure 2-2), one can capture the kinetics of the tracer over time instead of just recording it at a specific point after the tracer has been injected. In addition to being able to visualize specific tracer uptake compared to background uptake, it may be able to produce more accurate quantitative measures of disease. [25].

(2) The total body coverage. PET scanners with short axial fields of view (AFOV), which are often used in clinical settings, have some limitations, such as long acquisition times, and a high radiation exposure. Conventional scanners still need more than six bed positions to acquire whole body PET images. The total body system has an AFOV of 2m, which has a higher spatial resolution by covering the entire body with single bed [26].

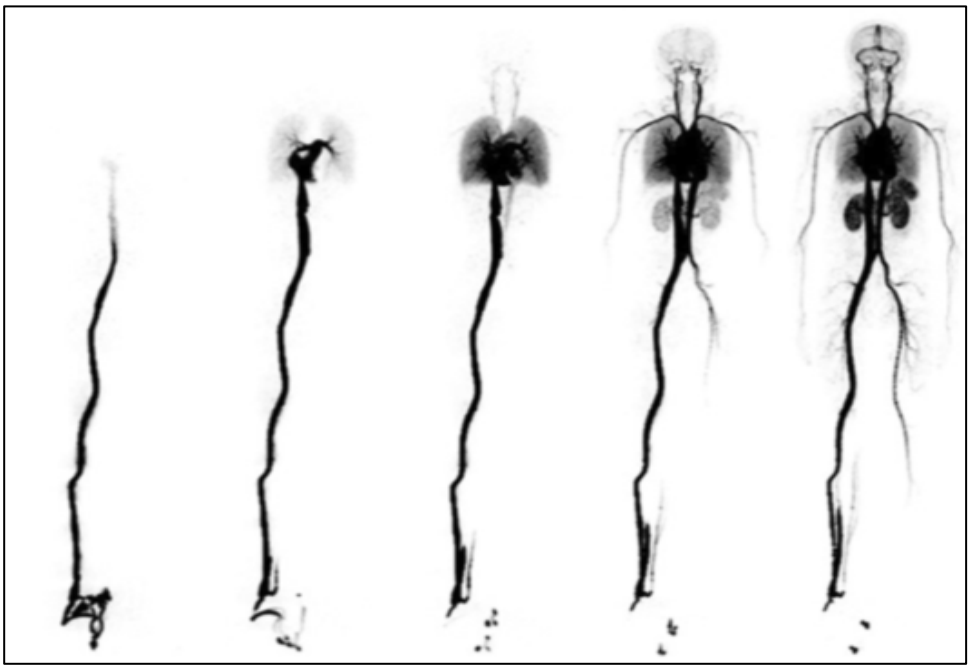

Figure 2-2 The initial three minutes of maximum-intensity projections of dynamic whole body PET images

2.3 Image Segmentation

The first step in the analysis and understanding of an image is usually to extract the image features of the object or object component from the image, such as extracting the boundary of the object component, or dividing the area where each component of the object is located. This process is called image segmentation. The aim of image segmentation is to divide different

parts of objects in an image to facilitate subsequent classification, identification, and interpretation of image objects.

Published medical image segmentation methods can be summarized into four categories [27]:

Spatial and geometric prior knowledge. Statistical Shape Models (SSMs) [28] are conventional statistical-shape-analysis-based method for medical image segmentation. Statistical shape analysis refers to the analysis of some set of shapes using statistical methods to determine its geometrical properties. Its major aspects include estimation of shape consistency within samples and mean shapes from (possibly random) samples.

Local image features with context layout. Graph Cut (GC) methods is a technique that utilize image's local information. Generally, GC segments the image into background and target area, and the voxels in the image are all represented by a set of vertex points. Probabilistic atlas (PA) is a classical model that analyzes images in a statistical way. The probability maps of PA are anatomical atlases containing statistically weighted fusions of many specimens[29].

Local image features with voxel-wise classification. This kind of method utilize machine learning classifiers to identify location and boundaries of target area and intuitive deformation for segmentation optimization. Zheng et al. [30] applied AdaBoost and random walk algorithms to do voxel-wise liver segmentation.

Neural networks. The heavy reliance on prior knowledge limited generalization capabilities of conventional algorithms, making it challenging to get acceptable results. Deep learning technologies then propelled advancements in biomedical segmentation. The traditional neural network model used histogram features for segmentation, depending extensively on preprocessing to eliminate unnecessary regions beforehand and guarantee clean boundaries by subsequently morphological operations [31]. Today, Convolutional Neural Networks (CNN) [32] has accomplished pixel-level classification by acquiring classification information for each pixel and

is frequently employed in medical imaging for automatic semantic segmentation tasks that do not rely on hand-crafted features. Ronneberger et al. [33] firstly proposed U-Net structure, which employs skip connections technology by merging the respective low-level detailed feature and the high-level feature. Recent studies showed that U-Net architecture is widely used for medical image segmentation and has already demonstrated its efficacy and effectiveness in many clinical applications [34].

2.4 Image Features

An image feature is a vector that contains several image-specific characteristics. It is a succinct representation that can be used to distinguish between images. We classify image features into two categories (Figure 2-3): static and dynamic, for different image modalities.

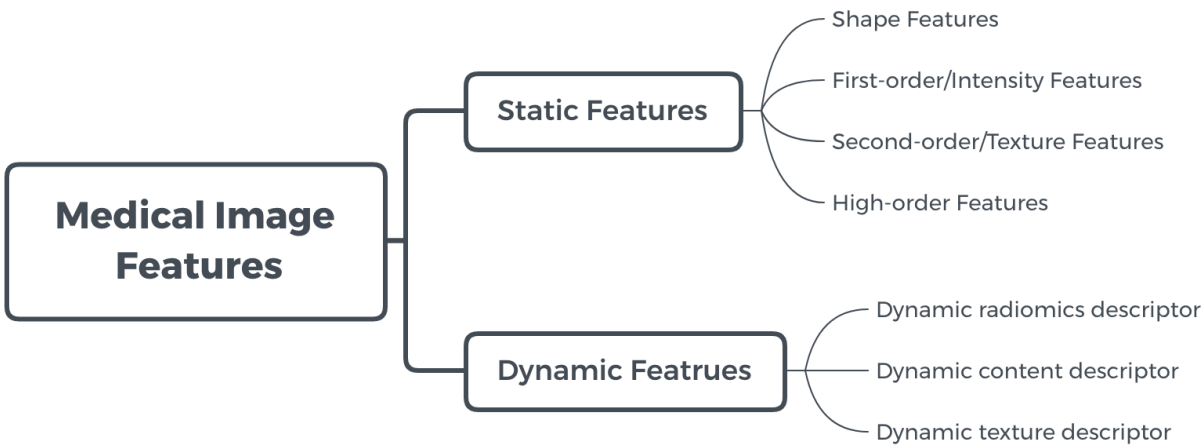

Figure 2-3 Categories of image features

2.4.1 Static features

In this section we mainly discuss static image features extracted from static medical images. According to Image Biomarker Standardization Initiative ‘s definition [17], image features can be categorized as morphological features, intensity features, texture features, and high-order

features. The static feature extraction methods will be automatically used when the area of interests (ROIs) within the images are defined. It is highlighted that the proper picture pre-processing methods, such as image intensity discretization [35], should be used to reduce noise and make it easier for feature computation. For improved consistency of the features, the most popular methods for image discretization employ a fixed bin count or fixed bin width [36].

2.4.1.1 Shape features

Morphological features represent the formation and geometry-related properties of the concentrated area of interest (ROI). Such as perimeter, maximum diameter, centroid, etc. For example, Sphericity (equation 2.1) is a measurement of the sphereness of the ROI. The value range is 0 to 1, and value 1 represents an absolute sphere.

$$\text{Sphericity} = \frac{\sqrt[3]{36\pi V^2}}{A} \quad (2.1)$$

where V is the volume of the mesh in cubic millimeter, and A is the surface area of the mesh in square millimeter.

2.4.1.2 First order/Intensity features

First-order statistics/Intensity features, including the mean, median, skewness, and kurtosis, are computed from the distribution of voxel intensities without considering voxels' spatial layouts [36]. Intensity features do not require discretization and can be used to depict continuous distributions. Variance is a first order feature (equation 2.2), which describes the average value of the squared distances of each grey level value from mean. Variance represents the homogeneity of ROI, and a higher value represents a higher homogeneity.

$$\text{Variance } = \frac{1}{N_p}\sum_{i=1}^{N_p}(\boldsymbol{X}(i) - \bar{X})^2 \quad (2.2)$$

where **X** is a group of Np in the ROI.

2.4.1.3 Second-order features

Second-order feature/texture feature is frequently employed to represent heterogeneity by exploiting the pixels' grey levels distribution of a co-occurrence matrix [37], i.e., Gray Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GLRLM), and Gray Level Dependence Matrix (GLDM).

Grey Level Co-occurrence Matrix (GLCM). GLCM is a matrix that describes the joint distribution of discretized intensities (grey levels) of neighboring pixels, or voxels in a 3D volume [37]. The value of (i, j) th element of GLCM is the count of combination of i and j that exists in distance m along the angle θ. Figure 2-4 gives an instance of GLCM calculation with four discrete grey levels (set distance m = 1 and θ = 0). The element (2,1) of the GLCM is 1 because there exists only one pair of voxels with intensity values of 2 and 1 in the direction of θ = 0. GLCM features include Autocorrelation, Joint Average and Cluster Prominence etc.

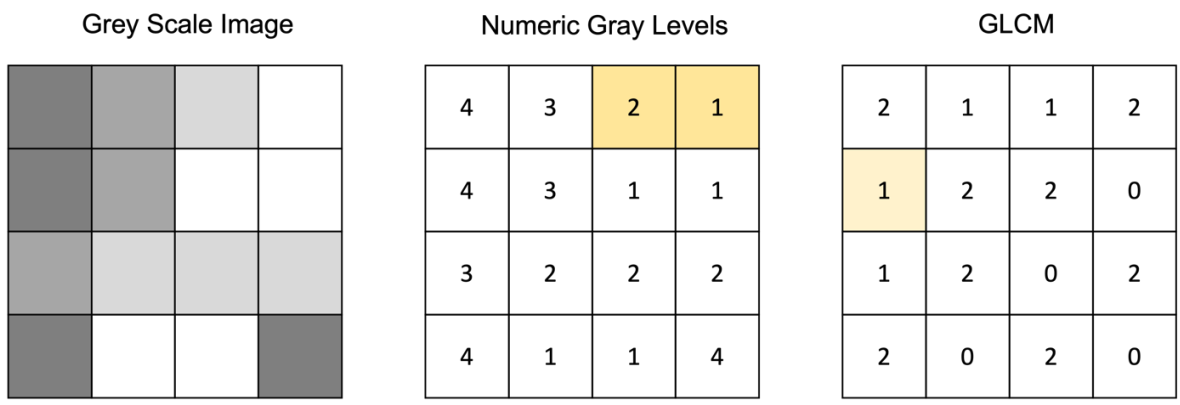

Figure 2-4 An example of GLCM computation

Grey Level Dependence Matrix (GLDM). The GLDM describes the voxel-wise dependencies in a grey level image [37]. The grey level dependency is that, within a distance of m, the counting of connected voxels which are dependent on the center voxel. Two adjacent voxels are dependent[38] (with grey level i and j, respectively) if $|i-j| \leq \alpha$ (α is a given scalar). The value of (i, j) th element of GLDM is the count that the center voxel with grey level i has j dependent adjacent voxels. Figure 2-5 gives an instance of GLDM calculation with four discrete grey levels (set m = 1 and α = 0). The element (2, 3) of the GLDM is 1 because there exists only one center voxel with value 2 and two dependencies. GLDM features include Small Dependence Emphasis and so on.

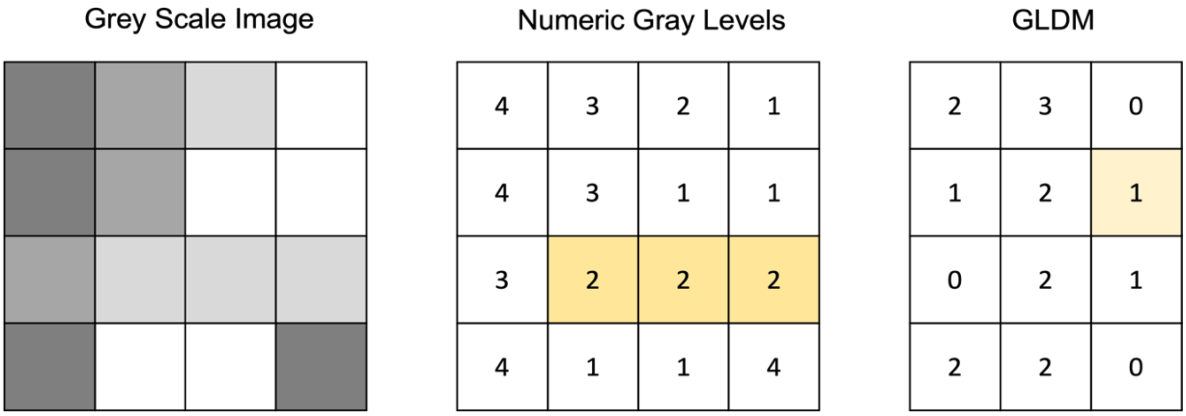

Figure 2-5 An example of GLDM computation

Grey Level Run Length Matrix (GLRLM). GLRLM is the quantification format of grey level runs, which represents the number of pixels comprising a sequence of pixels with the

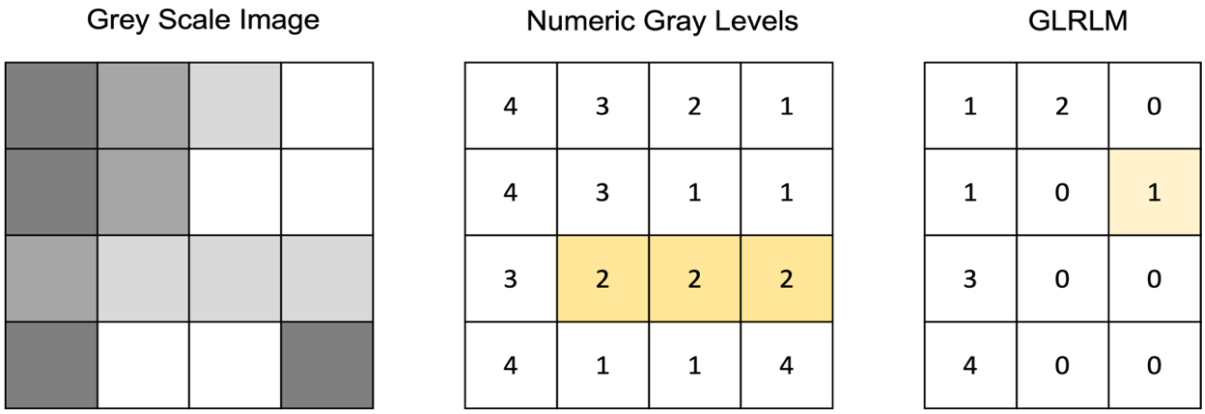

Figure 2-6 An example of GLRLM calculation

same gray level value [39]. The value of (i, j) th element of GLRLM is the counting of runs with grey level i and run length j in the image of an angle θ. Figure 2-6 gives an instance of GLRLM calculation with four discrete grey levels (set m = 1 and θ = 0).

The element (2, 3) of the GLRLM is 1 because there is only one run with the grey level 2 and length 3 in the θ = 0 direction. GLRLM features include Short Run Emphasis [40], Long Run Emphasis (LRE) and Run Percentage (RP), etc.

Grey Level Size Zone Matrix (GLSZM). Similar to GLRLM, GLSZM is the quantification format of an image's gray level zones, which is defined as the counting of voxel combinations that have identical gray level values [41]. The (i, j) th element of GLRLM shows the number of zones with the size j and intensity value i. Figure 2-7 gives an instance of GLSZM calculation with four discrete grey levels (set m = 1 and θ = 0). The element (3, 3) of the GLSZM is 1 because there exists only one zone with grey level three and size three. GLSZM features include Zone Variance (ZV), Zone Entropy [1] and Low Gray Level Zone Emphasis (LGLZE), etc.

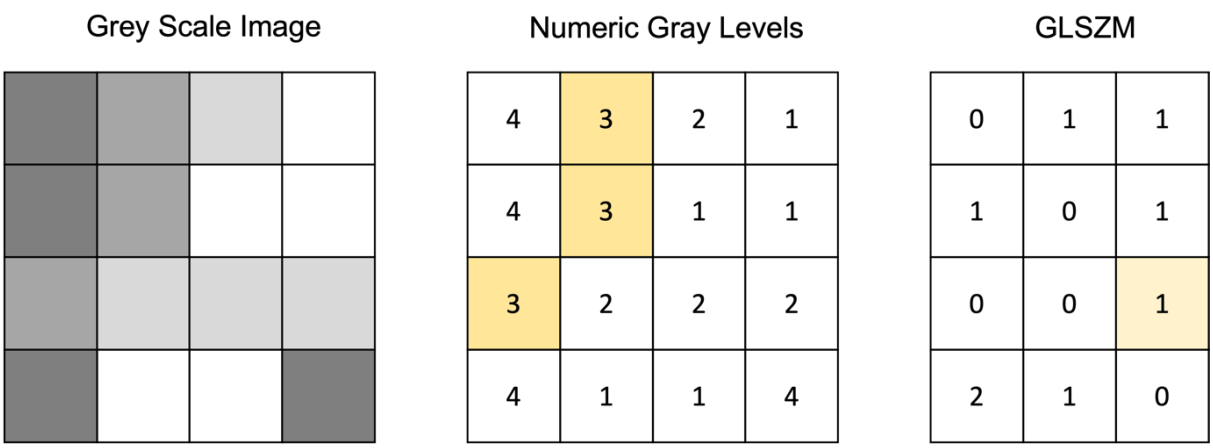

Figure 2-7 An example of GLSZM computation of an image with four grey levels.

Neighboring Grey Tone Difference Matrix (NGTDM). NGTDM is the quantification format of variance of a grey value and the mean grey value of its pixel combinations within the distance m [42]. The matrix stores the total value of absolute differences for gray level. Figure 2-8 gives an instance of NGTDM calculation with four discrete grey levels (set m = 1 and θ = 0). GLCM features include Coarseness, Contrast and Busyness.

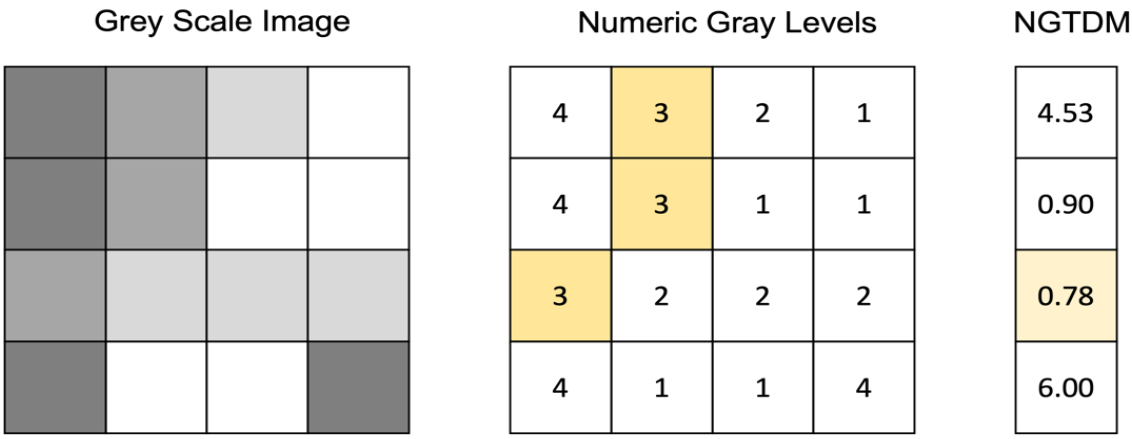

Figure 2-8 An example of a 2D NGTD matrix calculation

2.4.1.4 High-order features

High-order feature consists of first-order features, second-order features, and texture features from Laplacian of Gaussian (LoG) and wavelet pictures, which are designed to perform noise reduction and capture subtle information from images at different frequency domains. More specifically, LoG transformations exhibit remarkable performance for regions with hazy borders, metric descriptions, and capturing texture information at various coarse levels, while wavelet features use different frequency bands to disclose more useful subtle traits based on the original images.

2.4.2 Dynamic features

Dynamic images are common in medical scenarios, such as minimally invasive surgery [5] videos, dynamic PET/CT imaging, and time-related tomographic images in multi-stage chemotherapy evaluation [43]. Compared to static image features, there is no generally accepted definition of dynamic features based on medical images. We summarize the state-of-art dynamic feature (descriptor) techniques in computer vision field and classify dynamic features into three categories: dynamic radiomics features, dynamic content-based descriptors and dynamic texture-based descriptors.

2.4.2.1 Dynamic radiomics features

Most of the published studies on dynamic radiomics features are based on subsequent static medical images obtained from different time points. Delta Radiomics [17] [44] DR is a technique to assess the initial differentiation of FDG-PET radiomics features to predict therapeutic response for patients with non-small cell lung cancer (NSCLC). Dynamic radiomics features are calculated in the format of percentage change between two static scans. This method can offer more details to recognize and predict treatment-induced changes throughout therapy[14, 45, 46]. Recently, Qu et al. [15] extend DR with a more sophisticated dynamic radiomics framework, which can convert the static imaging features from different acquisition periods into dynamic features. The workflow and mathematical paradigm of this method could represent static radiomics features with time-varying characteristics and has been shown to achieve higher accuracy in prediction of gene mutation status and axillary lymph node.

2.4.2.2 Dynamic content-based descriptor

Computer aided surgery allowing surgeons to do sophisticated minimally invasive surgery[47], has been popular nowadays. Dynamic content-based descriptors are usually used in surgical video retrieval tasks, which need to describe motion and temporal information of a series of images.

Primus et al. [48] employ multiple key point detection methods and Support Vector Machines (SVM) methodology to segment video information. DeMenthon et al. [49] propose a descriptor that makes use of the position, color, and dynamics of independently moving patches over consecutive frames. Deep learning algorithms have made tremendous developments over the past few years. The Similarity-Adaptive Deep Hashing (SADH) approach [50] uses training algorithms to learn similarity-preserving binary patterns from original images photos. Chitajallu et al. [51] used pre-trained 3D CNN models to learn features in dynamic frames and interact with users to iteratively enhance the model outputs.

2.4.2.3 Dynamic texture-based descriptor

Most statistical-based dynamic texture description methodologies extend the analysis of conventional spatial texture methods to the space-time domain by adapting standard spatial texture methods. In both static and dynamic texture analysis, Local Binary Pattern (LBP) has been extensively studied and refined because of its computational simplicity, invariance, and high performance. Ojala et al. (1996) [40] presented the fundamental principle of LBP. Since then, further development been suggested. Several research have suggested extending LBP to capture 3D textures and patterns for 3D pictures. The Local Phase Quantization on Three Orthogonal Planes (LPQ-TOP) [52] is an approach that measures the locally periodic properties of the Fourier Transform (FT) and is robust to blurry images.

Using system identification theory, some model-based texture presentation approaches characterize the layout properties and dynamics of a scene by estimating the parameters of a linear dynamical system (LDS) [53]. Recently, LBP descriptor and LDS techniques showed promising results by integrating together to examine the texture layout, spatial organization, and dynamics of image sequences.

2.5 Feature Selection

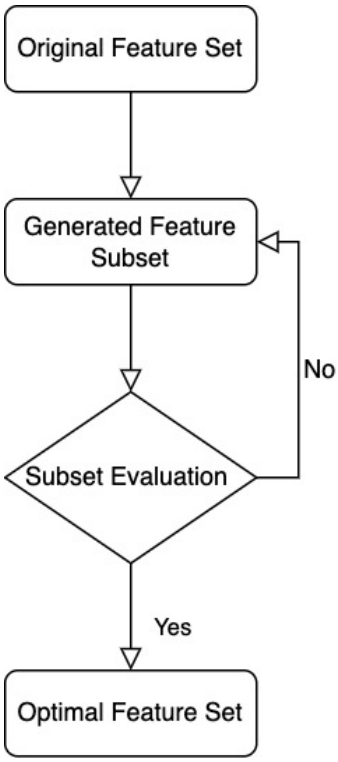

Figure 2-9 The general procedure of feature selection framework

From an algorithmic point of view, feature selection from high-dimensional small size data is a transformation-based method. Feature selection does not alter the original distribution of features. It simply picks certain suitable features to create a new low-dimensional space that retains the majority of the original feature space's attributes. A solid feature selection strategy may eliminate unnecessary and redundant features, limit the influence of noisy data on the

performance of the classifier, and make the chosen features more interpretable for high-dimensional, small-size data, which is prevalent in medical picture analysis.

The general procedure of feature selection framework (Figure 2-9) has four elements:

1. Formation of Subset
2. Assessment of Subset
3. Stopping Criteria
4. Confirmation of Output

According to different evaluation functions, the feature selection methods consist of four categories: filter method, wrapper method, embedded method, and ensemble method.

2.5.1 Filter Method

Filter methods (Figure 2-10) evaluate the categorization ability of a feature by assessing its internal properties, which is irrespective of whatever classifier is ultimately employed. Typically, such systems require a scoring index or a threshold. It can be separated further into the feature ranking approach (Table 2-2) and the spatial search method based on the various ways of formatting feature subsets. Generally, filtering algorithms may overlook features that are not valuable on their own but can be quite beneficial when paired with other feature selectors. Figure 2-10 depicts the graphical depiction of the filter model.

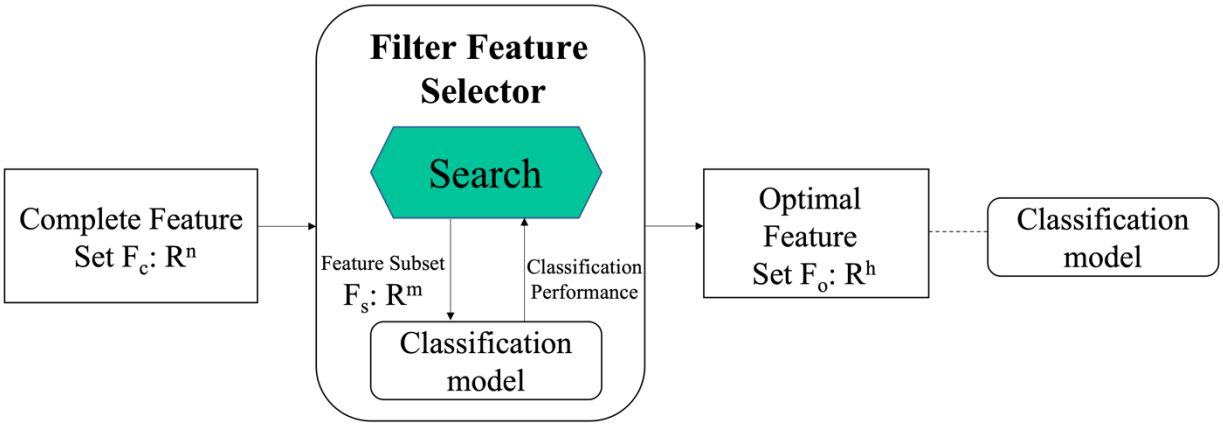

Figure 2-10 The feature filter model

Table 2-2 Filter method by ranking index

Filter Evaluation Criteria	*Description*	*Examples*
Statistics method	Measure the distribution difference between samples of class by statistical methods	t-test [54] Fold-change Ratio [55] Bayesian framework [56] Mutual information [57]
Information Theory method	Measure the information contained in the target feature through information entropy	Information gain [58] Log Likelihood ratio [59] Quasi-Poisson coefficient [60]
Correlation Method	Identify the relationship between desired attributes and categories	Kendall rank [61] Linear Discriminant Analysis[62]

In order to choose the feature subset from the whole feature set that has the greatest information and the least amount of redundancy, the spatial search approach primarily employs an optimization strategy, such as Correlation-based Feature Selection (CFS) [63] and Maximum Relevance Minimum Redundancy (MRMR) [64].

The majority of ranking techniques are univariate. The spatial search approach, on the other hand, is a multivariate method whose algorithm considers not only the connection between feature subsets and class labels, but also the connection between feature subsets. The precision of spatial search methods is usually high but the computational consumption of locating optimum subsets under high-dimensional settings is intensive.

2.5.2 Wrapper Method

The feature selection technique and classification model that make up the wrapper method, which encapsulates several categorization models. According to the results of the classifier on the feature subset, the wrapper method evaluates the selected features and adjusts

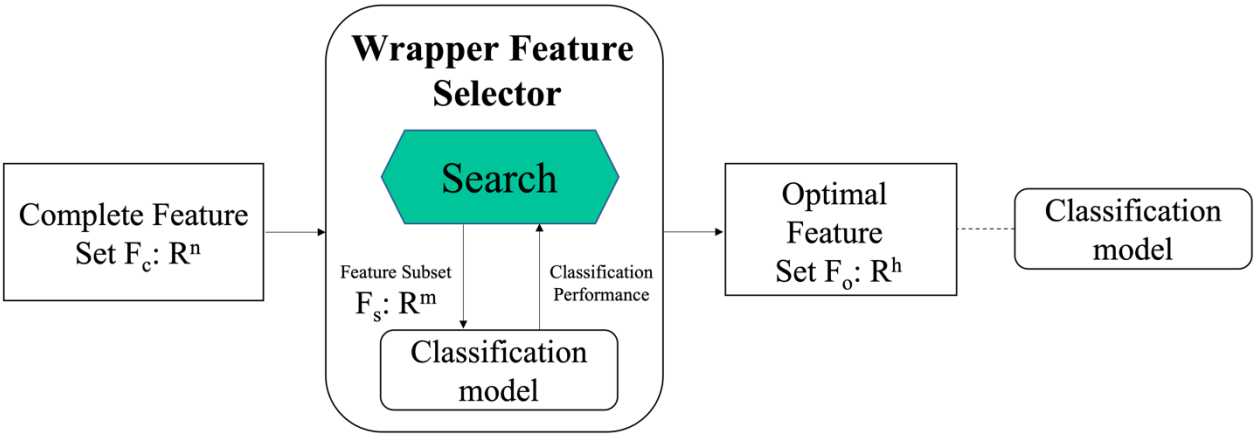

Figure 2-11 The wrapper model

the subset through some optimized search strategies, and finally obtains the approximate optimal subset. Genetic Algorithm (GA) [65] is often used to build classification models. The

number of potential feature subsets is 2^N for a dataset with N features. Finding the optimal subset of features has been shown to be NP-Hard problem [66]. The wrapper algorithm is shown in Figure 2-11. Approximately, sequential search and heuristic search are the two categories that the wrapper technique falls under:

(1) Sequential search algorithm.

The sequential search method begins with an empty feature subset and adds (or removes) features until the feature subset provides the optimal performance for the evaluation function. Some conditions for halting the search will be provided to expedite the selection of feature subsets, ensuring that the smallest number of feature subsets are chosen as the evaluation function continues to improve and reaches the best performance. Meanwhile, to improve temporal performance, Nakariyakul et al. [67] presented a novel feature selection approach known as the recursive feature elimination (RFE) method: adding features from the candidate feature set each time, examining the difference between the chosen features, and then eliminating the features least important to the class labels substantially improves the algorithm's time performance.

(2) Heuristic search algorithm.

From the whole collection of feature candidates, a random feature subset is generated as the starting point for the heuristic search algorithm, which then uses heuristic criteria to progressively move toward the ideal answer. This approach has a high level of search uncertainty, but the quality of the resulting feature subset is acceptable. Emmanouilidis et al. [68] employ an evolving genetic algorithm to overcome the challenge of feature selection in picture recognition. By adjusting the iteration threshold, the number of feature subsets that need to be iterated are sometimes decreased.

Due to the nature of the sequential search method, a secondary selection cannot be performed on the rejected features, nor can the picked features be discarded. It is simple to get trapped in local optima, often known as nesting effects. These issues may be effectively addressed by heuristic search methods. Additionally, it has been shown that the computational cost of parallel heuristic search methods is much lower than that of sequential selection algorithms. [69].

2.5.3 Embedded Method

The Embedded approach was created primarily to solve the time-consuming issue of wrapper method in processing large datasets. Embedded method has gradually become a hot spot for feature selection due to its efficient spatiotemporal performance and better accuracy. Figure. 2-12 depicts the graphical depiction of the embedded model.

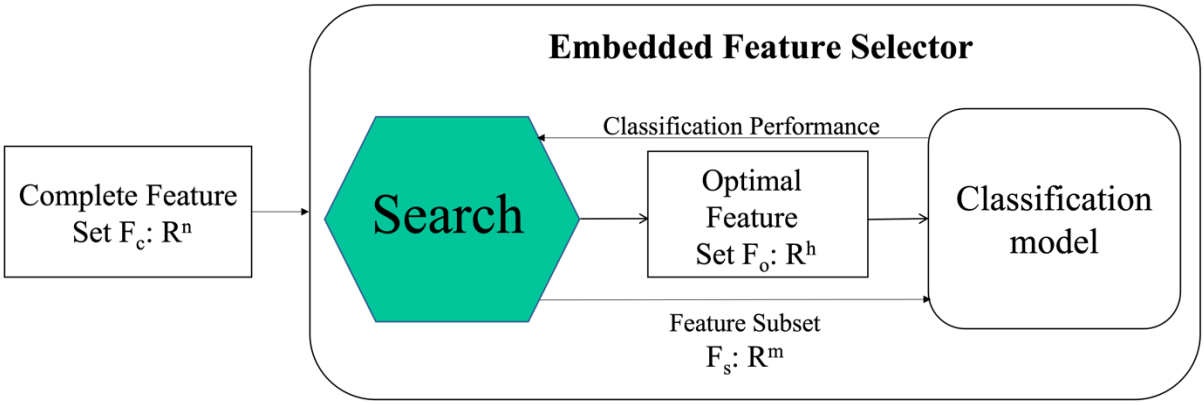

Figure 2-12 The embedded model

There are two popular embedded methods: SVM-based model [70] and Least Absolute Shrinkage Selection Operator (LASSO) [71]. SVM is widely used in handling high-dimensional small size feature selection problem. One explanation is that SVM may strike a compromise between model complexity and learning capacity on the basis of a small number of examples. Besides,

it can also effectively eliminate redundant features. Regularized sparse models can remove many redundant or noisy features and select a subset of features with good interpretability [72]. The representative algorithm of regularized sparse model is LASSO, which was proposed by Tibshirani in 1996 [73]. The main concept is to add coefficient absolute value as a punishment term to the least squares estimate such that the total absolute value of the coefficient is below or equal to a predetermined limit.

2.5.4 Ensemble Method

Ensemble learning is a technique that uses a number of feature selection strategies for learning and combines the learning outcomes according to a set of rules in order to achieve superior learning effects compared to single feature selection methods. This method has a lower computational complexity than the wrapper method since it interacts directly with the learning process. It considers not just the relationship between an input feature and an output feature, but also looks locally for traits that allow for more local discrimination. The ensemble approach has been employed in certain instances to increase the stability of feature selection algorithms. Li et al. [74] introduced a new method for selecting features that employs methods for resampling to disturb the data. It creates many training sets and test sets, continually invokes recursive decision tree, and picks features using classification error rate as the evaluation index. Dutkowski et al. [75] utilized multiple feature selection techniques for gene selection and combined them via an optimization approach, with the outputs of each algorithm composing the final subset of features. Sais et al. [76], Abil et al. [77] proposed an ensemble feature selection framework by bagging method. Saeys et al. [78] fuses the outcomes of numerous methods to finish the feature subset integration. Using sample resampling technology, Abeel et al. [79]

generates several feature subsets and achieves excellent performance on various high-dimensional small sample data sets.

2.6 Modeling and Evaluation

2.6.1 Model building

As a branch of AI, machine learning has advanced impressively quickly in medical imaging field [80]. In fields of research, machine learning demonstrates its distinctive talents. It functions as a crucial link between computer science studies and medical research [81]. Machine learning techniques used for medical imaging data analysis can help us better understand diseases and therapies, as well as produce individualized medicines and successful treatments [82]. A variety of difficult clinical tasks have applied machine learning algorithms, including brain tumor segmentation [83], detection Alzheimer disease with MR imaging [84], differentiation of liver tumor phenotypes [85], breast cancer detection and diagnosis [86], etc. There are two primary categories for ML algorithms: supervised learning and unsupervised learning.

2.6.1.1 Supervised Learning

Supervised learning algorithms learn or create a pattern from training data and infer output based on the learned pattern. A supervised learner is tasked with predicting the output with pre-labeled training data, which consists of classification problem and regression problem in data mining.

Support Vector Machines (SVM) Classifier. SVM [87] is a binary classifier. One advantage of SVM is that it doesn't need to calculate all samples and uses less computer memory when dealing with high-dimensional data. However, because SVM classifier needs to map low-

dimensional disordered data into high-dimensional feature space through a kernel function, and separate it through a hyperplane, the computational cost is relatively high.

K Nearest Neighbors (KNN) Classifier KNN [88] is a frequently used supervised classification algorithm. An object's categorization is decided by the majority vote of its neighbors, and it is assigned to the class with the highest percentage of members among its k nearest neighbors. KNN is a parameter-less training model which is simple to use and requires very little processing time. However, the calculation time and storage space will rise exponentially as the data size increases.

Logistic Regression Classifier Logistic Regression [89] is a model for binary classification with clarity and interpretability, which is also frequently utilized in the medical industry. The fundamental idea of logistic regression is to utilize maximum likelihood estimation to estimate parameters while supposing that the data follow a certain distribution. Logistic regression can be solved in two ways: the exact analytical solution and the Stochastic Gradient Descent (SGD) [90] algorithm estimation. Typically, we use analytical solutions when accuracy is required, and SGD iterations when time efficiency is required.

Decision Tree Classifier (DT) Decision Tree [91] is a tree-structure-based decision-making model It classifies data sets using several conditional discrimination methods and then achieves the desired outcomes. It classifies data sets using various conditional discriminating techniques before arriving at the desired outcomes. The decision tree's root node serves as its structural starting point, internal nodes represent intermediate decision-making steps, and leaf nodes represent the outcomes of categorization. Decision Tree has low data quality requirements and good interpretability but is very time-consuming.

Random Forest Classifier Random Forest [92] is an algorithm that integrates multiple decision trees, and it essentially belongs to a major branch of machine learning called Ensemble

Learning [93]. With the ability to handle input samples with high dimensional features and produce an internal, unbiased assessment of the generalization error as the forest building process moves along, RF classifiers perform well in huge datasets. However, similar to other ensemble learning algorithms, RF's outcomes are unpredictable and sensitive to probability.

Linear Regression In the analytical method of linear regression [94] the connection between one or more independent and dependent variables is modeled using least squares functional. Linear regression assumes that the data subject a linear distribution, limiting the generalization ability of the model because few data are strictly subject to the linear distribution in the real world.

2.6.1.2 Unsupervised Learning

Unsupervised learning [95] is a type of machine learning that automatically categorizes incoming data without given pre-labeled training examples. This learning technique helps identify the common characteristics inside a dataset that cannot be validated by specialists. In order for the algorithm to be able to model the input-output relationships, the final classes and labels are required. In order to extract useful insights and enhance data interpretability for users, these algorithms use approaches to mine the potential pattern input data, identify underlying information and describe or group comparable data elements. The main applications of unsupervised learning include cluster analysis, association rule [96], and dimensionality reduce [97].

K-means Clustering. K-means clustering [98] is a centroid-based technique, in which each cluster is described by a central vector. Many K-means algorithms need the hyper parameter, K, to be predetermined, which is regarded as one of their most significant disadvantages. Moreover, since they always assign an element to the closest centroid, the algorithms select clusters

with nearly comparable sizes. Therefore, the clusters edges are frequently erroneously chopped.

Hierarchical clustering. The major assumption of hierarchical clustering [99] is that things are more closely connected to one another than to distant objects. Based on the distance between elements, these algorithms link them to create clusters. The maximum distance required to connect cluster members can be used to characterize the cluster to a large extent. These methods will provide a hierarchy rather than a unique division of the data set. Hierarchical clustering is more sensitive to outliers and may result in new clusters or even the merging of previous clusters.

Evaluation metrics There exist many evaluation methods for clustering results. For internal assessment, the clustering is summarized as an evaluation score, such as Davies–Bouldin index and Silhouette coefficient [100]. External evaluation, in which the clustering is compared to an existing ground truth includes measurements like Purity [40] and confusion matrix [101].

2.6.2 Model Evaluation

Model Evaluation. Several factors may be taken into consideration when evaluating how successful classification models are. In medical studies, it is essential to distinguish between false positive (FP) and false negative (FN) misclassification. A sample's correct categorization is indicated by the terms of true positive (TP) and true negative [64] [48] in classification prediction. On the contrary, false positive (FP) and false negative (FN) are index representing misclassification. The receiver operating characteristic curve (ROC) is an exhaustive, objective indicator that represents the ongoing variations in sensitivity and specificity. It uses the composition technique to illustrate the connection between sensitivity and specificity. Multiple critical values for continuous variables are used in the ROC calculation, and the resulting series of

sensitivity and specificity are plotted on a curve with the ordinate representing the sensitivity and the abscissa representing the specificity. The accuracy of the diagnosis increases as the Area Under the Curve (AUC) increases.

Cross validation. The primary principle behind cross-validation is to group the original data (dataset) under specified circumstances, with one half serving as the training set and the other as the validation set or test set. Cross-validation techniques are performed to boost the generalization ability of the machine learning model. There are various cross-validation strategies such as S-Folde Cross Validation, Leave-One-Out Cross Validation etc., and nested cross-validation methods are prone to be effective in small and unbalanced datasets to avoid overfitting.

Bootstrap Replacement sampling is usually utilized in the bootstrap approach. First, a sample set as large as the original size is obtained for training, and then the unsampled data is used for testing. The probability that a sample is never taken in m samples is:

$$\lim_{m\to\infty}\left(1-\frac{1}{m}\right)^{m}=\frac{1}{e}\approx 0.368.\ (Equation\ 2.3)$$

Hence, bootstrap method may change the data distribution and will bring errors. On the other hand, the S-Folder Cross Validation, Leave-One-Out Cross Validation method etc. are commonly used when the sample size is sufficient.

CHAPTER 3. 18F-FDG PET/CT Radiomic Analysis for Classifying and Predicting Microvascular Invasion in Hepatocellular Carcinoma and Intrahepatic Cholangiocarcinoma

3.1 Introduction

Liver cancer is a high-risk malignancy which 5-year survival rate is only 10%[102]. Hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (ICC)account for over 95% of primary liver cancer and have significant differences in clinical treatment and prognosis[103-105]. Even in those cases when a radical resection is feasible, the probability of intrahepatic recurrent cancer and extrahepatic metastases is still very high[106]. HCC and ICC tend to invade vascular structures. Macrovascular invasion [104] refers to tumor invasion of larger vessels, and the most common is the portal vein tumor thrombus (PVTT), while microvascular invasion (MVI) refers to the presence of tumor cells within the portal or hepatic venous system[107]. The diagnostic gold standard of MVI positive in histopathology is defined as 5 or more tumor clusters visible within the peritumoral vascular (usually covered by endothelium) only on microscopy [108]. Previous studies have identified MVI as a major risk factor for early recurrence after liver resection [109-111]. However, lacking effective early diagnostic strategies, and with highly heterogeneous in clinical features and histological morphology [112, 113], liver cancer is difficult to distinguish from HCC and ICC as well as identify MVI status before surgery. Biopsy is a solution with invasive examination, yet it increases the risk of metastasis and cannot provide the whole status of tumors [114]. Therefore, there is an urgent need for a non-invasive quantitative evaluation method in vivo clinically, which can accurately

distinguish pathological subtype and reflect the biological characteristics of the whole tumor before surgery.

Radiomics, served as a quantitative high-throughput analysis method for mining medical images with high dimensional extractable data, has attracted increasing attention in recent years [115-117]. PET/CT (Positron Emission Tomography/ Computed Tomography)-based radiomics combined with medical imaging and molecular imaging could potentially be used as predictive or prognostic biomarkers for tumor diagnosis, treatment, efficacy evaluation and prognosis prediction before surgery[7, 118]. Some studies[119-121] have shown that PET/CT radiomics applications have obtained encouraging results, for instance, in differentiating benign and malignant tumors, identifying tumor stages.

Contrast-enhanced ultrasound (CEUS) is commonly used in HCC and ICC differentiation[122]. Besides, researchers have been making great efforts to find more precise ways to predict MVI status before surgery. MRI is widely used for detection the presence of MVI in ICC and HCC[123, 124]. Compared to CEUS and MRI, PET/CT scans noninvasively reflect tumor metabolism and molecular level changes in vivo and monitor tumor biological characteristics [125]. Cassim et al. stated that most HCC tumor cells were hypermetabolic activity stemming from an increased metabolic plasticity, which can be identified by PET/CT [126]. Lei et al. recently reported that FDG accumulation correlated with the degree of ICC differentiation [127]. Hence, PET/CT-based radiomics are expected to have great potential for predicting HCC and ICC type and MVI status with the advantages of high sensitivity, high specificity, repeatability.

Encouraged by the aforementioned promising applications, we attempted to explore 18F-FDG PET/CT imaging's potential capability in auxiliary diagnosis of its additional application in HCC and ICC classification, as well as detection in MVI presence before surgery. It's efficient

and convenient for patients to obtain a comprehensive quantification assessment of liver tumors after a single preoperative 18F-FDG PET/CT examination. Our aim of this chapter is to build a feasible and robust machine learning model with radiomics biomarkers and clinical characteristics that may provide preoperative prediction of HCC and ICC classification and MVI status in patients with primary liver cancer by using 18F-FDG PET / CT images.

3.2 Material and Methods

3.2.1 Dataset Description

The study was carried out in compliance with the International Guidelines for Human Research Protection of the Declaration of Helsinki (as revised in 2013) and International Conference on Harmonization in Good Clinical Practical (ICH-GCP). This retrospective study was approved by the Ethics Committee of Fudan University Shanghai Cancer Center and individual consent for this retrospective analysis was waived. We collected clinicopathological indicators and PET/CT images of 112 patients (58 females and 54 males) with liver cancer who underwent 18F-FDG PET/CT scan between January 2016 and December 2019 at Fudan University Shanghai Cancer Centre (Shanghai, China).

Inclusion criteria were as follows:(1) pathological diagnosis of either HCC or ICC confirmed by partial hepatectomy of primary liver lesion; (2) validation of 18F-FDG PET/CT scan images within two weeks before surgery; (3) with normal hematologic, renal, and hepatic function; (4) complete clinical characteristics and pathology immunohistochemistry results. Exclusion criteria included: (1) metastatic liver tumor; (2) preoperative PET/CT showed portal vein tumor thrombosis (PVTT); (3) incomplete clinical characteristics and pathology

immunohistochemistry results, including only performed liver biopsy; (4) blood glucose levels over 7.78 mmol/L or with abnormal laboratory indexes.

Preoperative tumor staging followed the Barcelona Clinic Liver Cancer (BCLC) criteria revised by American Association for the Study of Liver Diseases (AASLD) in 2010. Postoperative pathological classification and the presence of MVI and number of satellite node were confirmed by two experienced pathologists. MVI positive is defined as 5 or more tumor clusters visible within the peritumoral vascular (usually covered by endothelium) only on microscopy. The selected serum AFP and CA-199 levels were measured within one week before surgery. The threshold value of serum AFP and CA-199 level was 20 ng/mL and 37u/ml respectively.

3.2.2 PET/CT imaging acquisition and reconstruction parameters

18F-FDG was produced by an RDS Eclipse ST medical cyclotron (Siemens Healthiness, Knoxville, TN, USA) and an Explore FDG4 synthesis module.18F-FDG radiochemical purity was > 95%. Blood glucose levels of all patients were less than 7.78 mmol/L. Patients fasted for at least 6 hours prior to injection. After intravenous administration of 18F-FDG (3.7 MBq/kg), all patients laid in a bed for one hour and imaged by a Biograph 16HR PET/CT scanner (Siemens Medical Systems, Erlangen, Germany). First, an unenhanced low-dose whole-body CT scan was performed from head to the top of the thighs (tube voltage, 120 kV; tube current, 80~250 mA; rotation time, 0.5 s; helical pitch 3.6; slice thickness,5 mm; matrix, 512 × 512). Images were performed for attenuation correction. Then, whole-body PET scan was acquired over the same extent at three minutes per bed position for a total of 6~7 bed positions. PET data were reconstructed using iterative protocols with gaussian-filter iterative method

(iterations, 4; subsets, 8). The PET and CT images were imported into the Siemens workstation for analysis.

3.2.3 Volume of interest (VOI) segmentation

To provide an accurate segmentation, the VOI of primary liver tumors was first semi-automatically delineated using the Grow Cut algorithm [128] implemented on 3D Slicer (https://www.slicer.org) based on PET standardized-uptake-value (SUV) data, which shows high reproducibility. For the instances where SUV data of tumors were similar with adjacent structures, LLC model [129] and an improved edge detector were used to separate the tumor from the background and highlight the regions with weak boundaries. All results were corrected by manual adjustment and validated independently by two senior nuclear medical physicians to ensure reproducibility and reliability. All masks were reshaped to the same pixel spacing as original PET-CT images and checked based on PET-CT fusion image on 3D Slicer. Besides, pixel value of PET image was replaced by SUVbw to eliminate the effects of patients' absorption differences. Conventional PET metrics were also considered as radiomics features. On the 18F-FDG-PET, the SUVmax (standardized uptake value of the highest-uptake voxel within a VOI) and MTV (metabolic tumor volume) were automatically calculated on the Siemens workstation. TLG was calculated as follows: TLG= MTV × SUVmean.

3.2.4 Radiomics feature selection and machine learning

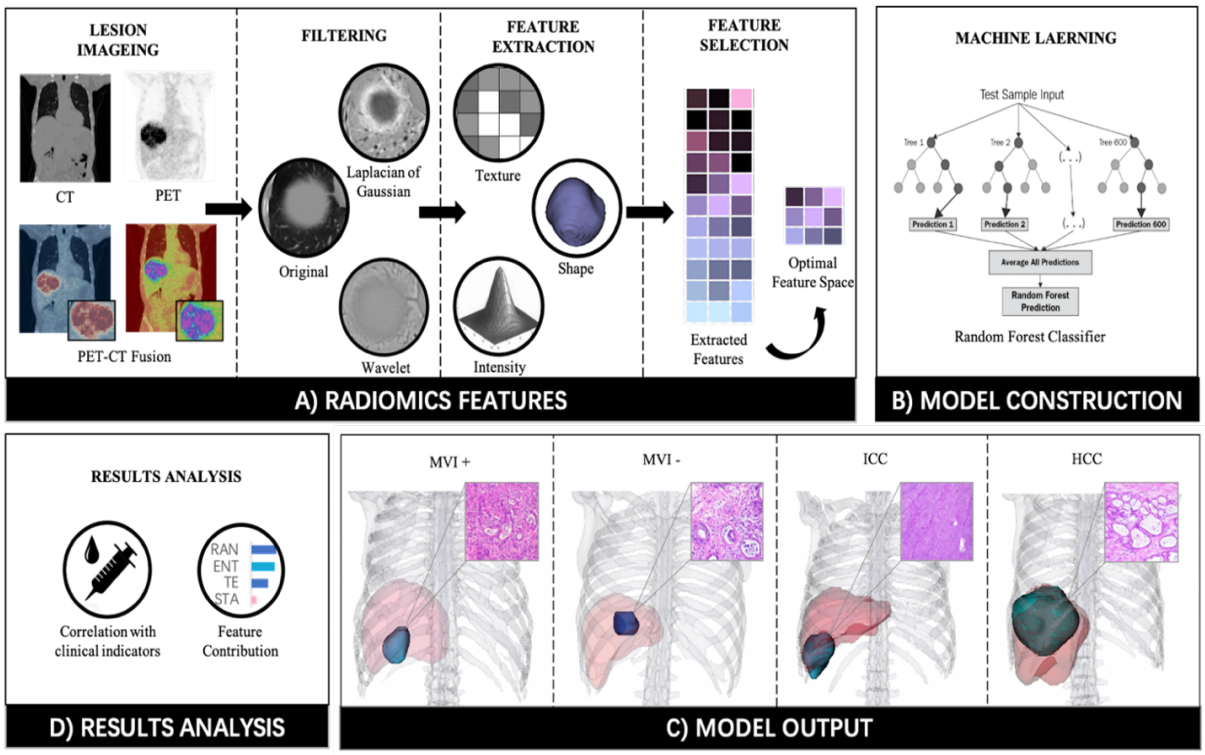

Figure 3-1 Outline of the workflow from feature acquisition, model construction, model output and results analysis.

The workflow of radiomic analysis by machine learning method is shown in Figure. 3-1, which consists of four key steps. At first, we obtain discriminative radiomics features from VOIs by using reproducible feature selection method; then the supervised machine learning classifier was constructed by random forest algorithm, which contributed to two tasks: HCC and ICC classification and MVI prediction; besides, we analyze potential correlations between radiomics and clinical features as well as each feature's contribution to model's results. Especially, for MVI prediction task, we divided all patients into HCC group and ICC group before feature selection step, then trained the MVIs identification models for each group separately.

Totally 1815 radiomics features including 918 CT features and 897 PET features were extracted for each patient. PyRadiomics (version 3.0), an open-source Python software package

was used to pre-process image and extract features, which is compliant with the Imaging biomarker standardization initiative as well. The matrix size of CT was 512 × 512 with the voxel size 1.0 × 1.0 × 3.3 mm3. The matrix size of PET was 128 × 128 with the voxel size 5.5 × 5.5 × 3.3 mm3. The images were discretized with a fixed bin size of 40 HU and 30 of SUV, and the mask images were resampled to the same pixel spacing as PET and CT images. From this package, we applied three filters for each PET and CT image before extraction: original channel, Laplacian of Gaussian (LoG) channel and wavelet channel. The extracted features reflected tumors' traits including intensity distribution, morphological characteristic, and texture pattern. The intensity feature is a first-order feature, which includes the maximum, mean, and average absolute deviation of the voxel values. The shape feature includes tumors' geometry properties such as edges and angles. Texture feature is a second-order feature and is used to express tumor's heterogeneity by the distribution of some common matrix, i.e., Gray Level Co-occurrence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GLRLM), and Gray Level Dependence Matrix (GLDM). High-order feature includes first-order features, second-order features and texture features from LoG and wavelet images, which aimed to reduce noises and obtain the subtle information from image at different frequency domains[130].

At feature selection part (Figure. 3-2), we aim to build a reproducible feature set. Take features' high dimensionality into consideration, we first eliminated statistical insignificant by Wilcoxon test (p-value < 0.05 was considered significant). Then we use uni-variable random forest feature selection to choose relevant features. Key features with tree importance greater than 0.001 were included. To further keep feature set discriminative, we use partwise Pearson Correlation matrix. We first identified pairs of related features ($|r| \geq 0.7$ for PET and CT features), then the feature with higher prediction ability (higher AUC using random forest classifier) will be included. Next, we utilize Sequential Forward Floating Algorithm[131-133] to recursively find

optimal feature combinations and avoid overfitting. We perform same steps for both HCC and ICC classification task and MVI prediction task for feature selection.

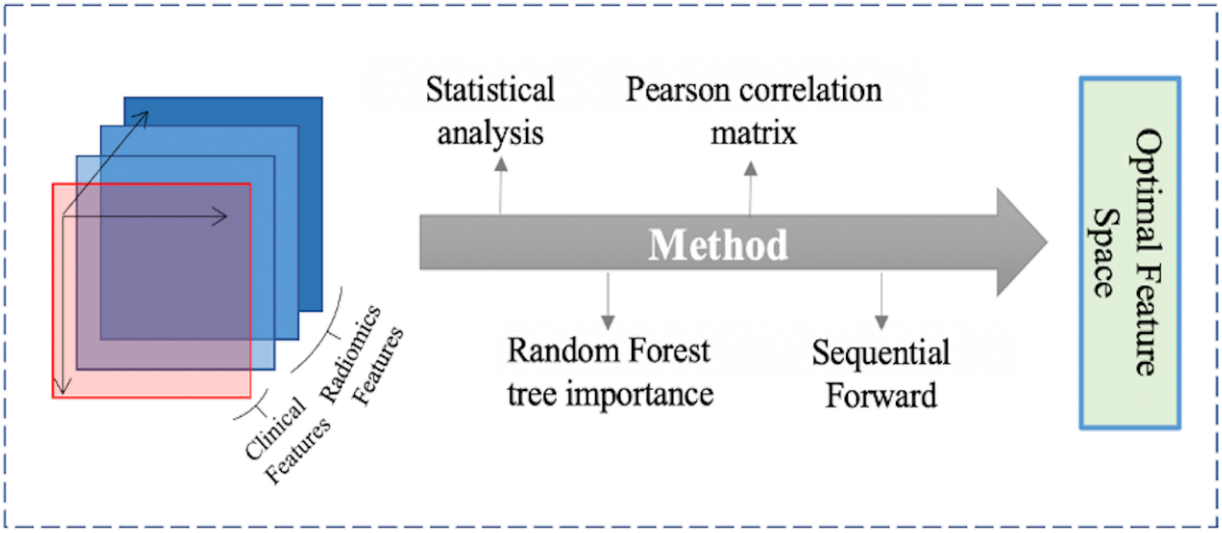

Figure 3-2 Outline of feature selection process

3.2.5 Modeling and Validation

The model was evaluated with cross validation and independent validation to achieve robustness and stability (Figure. 3-3). For HCC and ICC classification task (127 patients in total), they were randomly split into training (100 out of 127) and validation (27 out of 127) cohort. For MVI prediction in HCC (76 patients in total), there were 60 patients for training and 16 patients for validation; for MVI prediction in ICC (51 patients in total), there were 40 patients for training and 11 patients for validation. The proportion of positive and negative samples in the training and test sets is roughly the same as the proportion in the original dataset.

The feature selection procedure and random forest classifier were built on the entire training cohort. In the processes of model establishment, 20 times of stratified 10-fold (9-fold for training and 1-fold for testing) cross-validations were performed on training cohort for hyper

parameter searching. The performance analysis of machine learning models applied receiver operating characteristic curve, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) in the independent validation cohort.Statistical analyses were performed with 'Scipy 1.3.0', and 'math' packages in Python 3.6.8 programming language and environment.

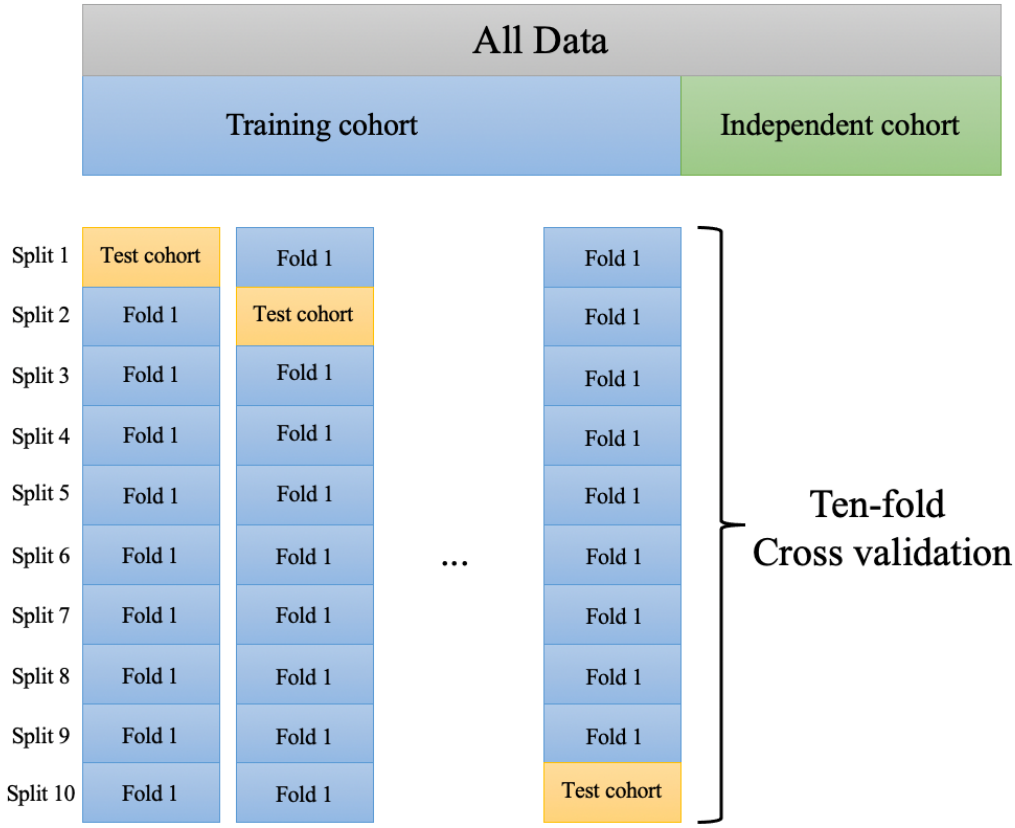

Figure 3-3 Flowchart for cohort divisions

3.3 Results

3.3.1 Clinical characteristics of patients

Patients' characters with split details are shown in Table 3-1 and 3-2. Totally 112 (55.5±28 years old) patients were included. The clinical indicators include Alpha-Fetoprotein (AFP), Carbohydrate Antigen 19-9 (CA19-9), age, tumor size, stage, tumor amount, and number of satellite nodes.

Demographic and clinical characteristics of patients

	Total	HCC/ICC	p-value
Number of patients	112	70/52	-
Age(yrs), median(range)	55 (±28)	54 (±28)/61 (±25)	0.211
Gender			0.194
male	79	58/21	
female	33	11/22	
AFP (ng/mL)			0.029
>=20	39	37/2	
<20	73	33/40	
CA19-9 (u/ml)			0.802
<=37	74	52/22	
>37	38	18/20	
Tumor Size [134]			0.001
3	21	16/5	
5	34	23/11	
10	38	17/21	
>10	21	14/7	
Tumor Stage			0.001
A	33	25/8	
B	46	32/14	
C	33	13/20	
Tumor Amount			0.112
multiple	12	10/2	
single	100	60/40	
Satellite Node			0.154
None	81	44/37	
1～3	41	26/15	

Table 3-1 Demographic & Clinical Characteristics of 112 paients for HCC and ICC task
*Abbreviations: **AFP**, Alpha-Fetoprotein. **CA19-9**, Carbohydrate antigen 19-9. **HCC**, Hepatocellular carcinoma. **ICC**, intrahepatic cholangiocarcinoma (ICC).

Demographic and clinical characteristics of patients

	Total	MVI present/MVI absent	p-value
Number of patients	112	64/48	-
Age(yrs), median(range)	55 (±28)	55 (±28)/56(±27)	0.486
Gender			0.347
male	79	48/31	
female	33	16/17	
AFP (ng/mL)			0.1
>=20	39	21/18	
<20	73	43/30	
CA19-9 (u/ml)			0.399
<=37	74	43/31	
>37	38	21/17	
Tumor Size [134]			0.0009
3	21	9/12	
5	34	16/18	
10	38	25/13	
>10	21	14/7	
Tumor Stage			0.090
A	33	11/22	
B	46	30/16	
C	33	23/10	
Tumor Amount			0.031
multiple	12	10/2	
single	100	54/46	
Satellite Node			0.359
None	81	32/49	
1～3	41	32/9	

Table 3-2 Demographic & Clinical Characteristics of 112 patients for MVI task
*Abbreviations: **AFP**, Alpha-Fetoprotein. **CA19-9**, Carbohydrate antigen 19-9. **MVI**, Microvascular invasion.

3.3.2 Radiomics features and performance of predictions

The most predictive feature combinations selected by feature engineering and corresponding explanation are shown in Table 3-3.

Task	*Feature Name*	*Feature Explanation*
HCC and ICC	**wavelet-LHL_Median_ct**	The median gray level intensity within the VOI. A higher value means higher density in the image.
	wavelet-HHL_Variance_pet	Variance is the mean of the squared distances of each intensity value from the Mean value. This is a measure of the spread of the distribution about the mean. Formula: $$variance = \frac{1}{Np}\sum_{i=1}^{Np}(\boldsymbol{X}(i) - \bar{X})^2$$
	log-sigma-3-0-mm-3D_ShortRun-HighGrayLevelEmphasis_pet	Short run high gray level emphasis (SRHGLE) measures the joint distribution of shorter run lengths with higher gray-level values. Formula: $$SRHGLE = \frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}\frac{P(i,j\|\theta)i^2}{j^2}}{Nr(\theta)}$$
MVI: HCC	**log-sigma-3-0-mm-3D_Range_pet**	The range of gray values in the VOI. Formula: $$range = \max(\boldsymbol{X}) - \min(\boldsymbol{X})$$
	wavelet-HHH_TotalEnergy_pet	Total Energy is the value of Energy feature scaled by the volume of the voxel in cubic mm. Formula: $$total\ energy = V_{voxel}\sum_{i=1}^{Np}(X(i) + c)^2$$
	wavelet-LLH_Entropy_pet	Here, ϵ is an arbitrarily small positive number (≈2.2×10−16≈2.2×10^{-16}). Entropy specifies the uncertainty/randomness in the image values. It measures the average amount of information required to encode the image values. Formula: $$entropy = -\sum_{i=1}^{Ng}p(i)log_2(p(i) + \epsilon)$$
MVI: ICC	**wavelet-HLL_Minimum_pet**	**X** be a set of Np voxels included in the VOI. Formula: $$minmum = \min(\boldsymbol{X})$$
	wavelet-HLL_TotalEnergy_pet	The same as wavelet-HHH_TotalEnergy_pet Formula: $$total\ energy = V_{voxel}\sum_{i=1}^{Np}(X(i) + c)^2$$

Table 3-3 Feature selection results

*Abbreviations: **VOI**, volume of interest. **HCC**, Hepatocellular carcinoma. **ICC**, intrahepatic cholangiocarcinoma (ICC). **MVI**, microvascular invasion. **LLH**, low, low, and high frequency. **HLL**,

high, low, and low frequency. Np is voxels included in ROI, Ng is the discrete intensity levels, and c is the voxel array shift.

We compare five categories of features: (1) clinical characteristic only; (2) optimal CT features only; (3) optimal PET features only; (4) optimal PET and CT features combination; (5) best PET, CT, and clinical characteristic combination that selected by feature engineering. Figure 3-4 gives the results of model performances in testing cohort in five categories. There were 2 PET features and one CT feature that gave the most prognostic value when working with HCC and ICC classification task (AUC = 0.86). As to MVI prediction tasks, three PET features and tumor stage shown great ability in HCC group (AUC = 0.88), meanwhile two PET features and CA19-9 performed well in ICC group (AUC = 0.90).

Task	*AUC*	*Accuracy*	*Sensitivity*	*Specificity*	*NPV*	*PPV*
HCC and ICC classification	0.86	0.82	0.78	0.88	0.91	0.88
MVI (HCC)	0.88	0.78	0.88	0.60	0.80	0.60
MVI (ICC)	0.90	0.77	0.75	0.80	0.75	0.80

Table 3-4 Model performances of optimal category of features

For HCC and ICC classification task, PET features show an AUC of 0.83. CT features also gave valuable information (AUC = 0.81), enhancing the results of PET-CT features combination (AUC = 0.86). But clinical features fail to give useful information (AUC = 0.56), which worse the results of combination as well (AUC = 0.80).

On the other hand, the results of MVI prediction tasks show that PET and clinical features outperform than CT features. For HCC in MVI task, three PET features plus one clinical feature (tumor stage) gave model highest AUC at 0.88. Only PET features can achieve AUC of 0.84, but CT features have AUC of 0.61. Since CT features' AUC is much lower than PET's, they

not only fail to give useful information, but worsening the results of PET-CT features (AUC = 0.71).

For ICC in MVI task, two PET features and CA19-9 can achieve AUC of 0.90. PET features also gave impressive AUC of 0.88. Meanwhile clinical features and CT features have AUC of 0.67 and 0.66 respectively, which are unable to give results in high accuracy, worsening the results of combinations. Table 4 shows detailed performances in optimal feature category for three tasks.

We also analyze the radiomics features' category. As it illustrated in Figure 3-5 and Figure 3-6, for all three tasks, PET features outperformed than CT features because only one CT feature was included. Besides, due to the spatial resolution of PET/CT is relatively low, and it has less advantages in defining the tumor boundaries all selected radiomics features were intensity as well as texture feature, which means shape features failed to give predictive information for both tasks. In Figure. 7, we show four representative patient examples.

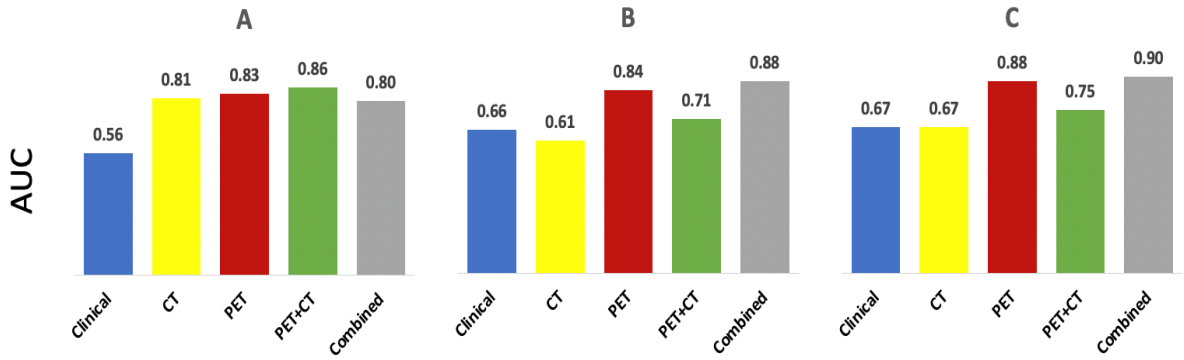

Figure 3-4 Model performance in training cohort (feature combination's category as x-axis, and feature combination's AUC (0~1) value as y-axis). [1] HCC and ICC classification task. (B) MVI prediction for HCC. [1] MVI prediction for ICC. Combined is the best feature combination of radiomics features and clinical features.*Abbreviations: HCC, Hepatocellular carcinoma. ICC, intrahepatic cholangiocarcinoma (ICC). MVI, microvascular invasion.

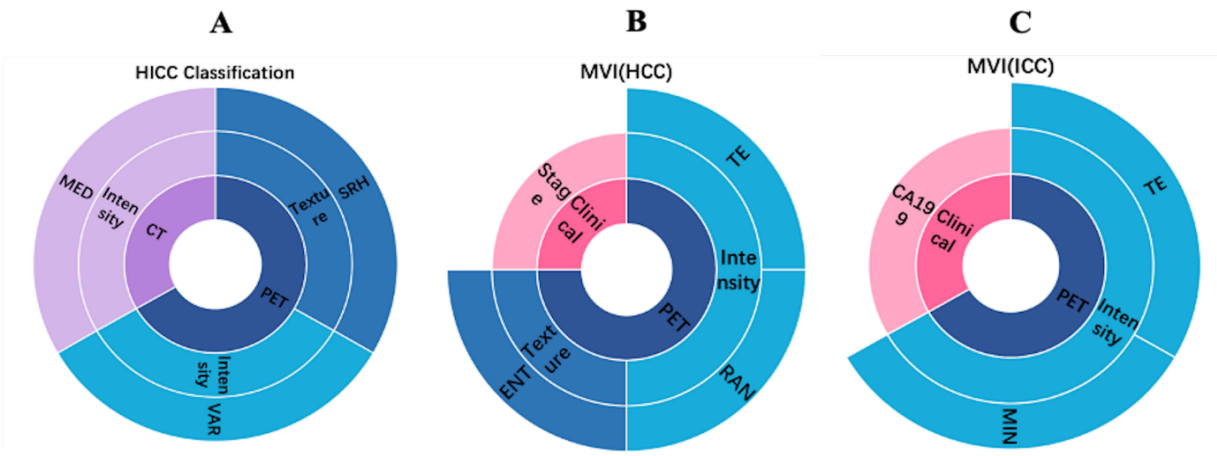

Figure 3-5 Selected feature's types three tasks. Blue are PET features; purple are CT features and pink are clinical features. *Abbreviation: MED, Median. SRH, Short Run High Gray Level Emphasis. VAR, Variance. TE, Total Energy. RAN, Range. ENT, Entropy. MIN, Minimum.**PET**, positron emission tomography. **CT**, computed tomography. **HCC**, Hepatocellular carcinoma. **ICC**, intrahepatic cholangiocarcinoma (ICC). **MVI**, microvascular invasion.

3.3.4 Correlation with Clinical and Conventional PET features

We also made use of Pearson correlation matrix to discovery potential relationships between radiomics features and clinical as well as conventional PET features. Figure. 8 shows the results of Pearson partwise correlations with four conventional PET features and six clinical features. p value <0.05 was considered significant. We found that for all three tasks, the selected feature combinations had significant relationships with metabolic indicators, tumor size and tumor stage.

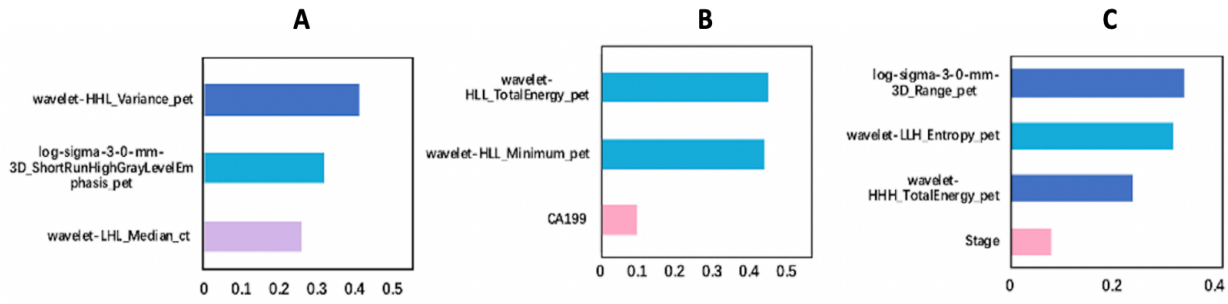

Figure 3-6 Feature importance in three tasks.

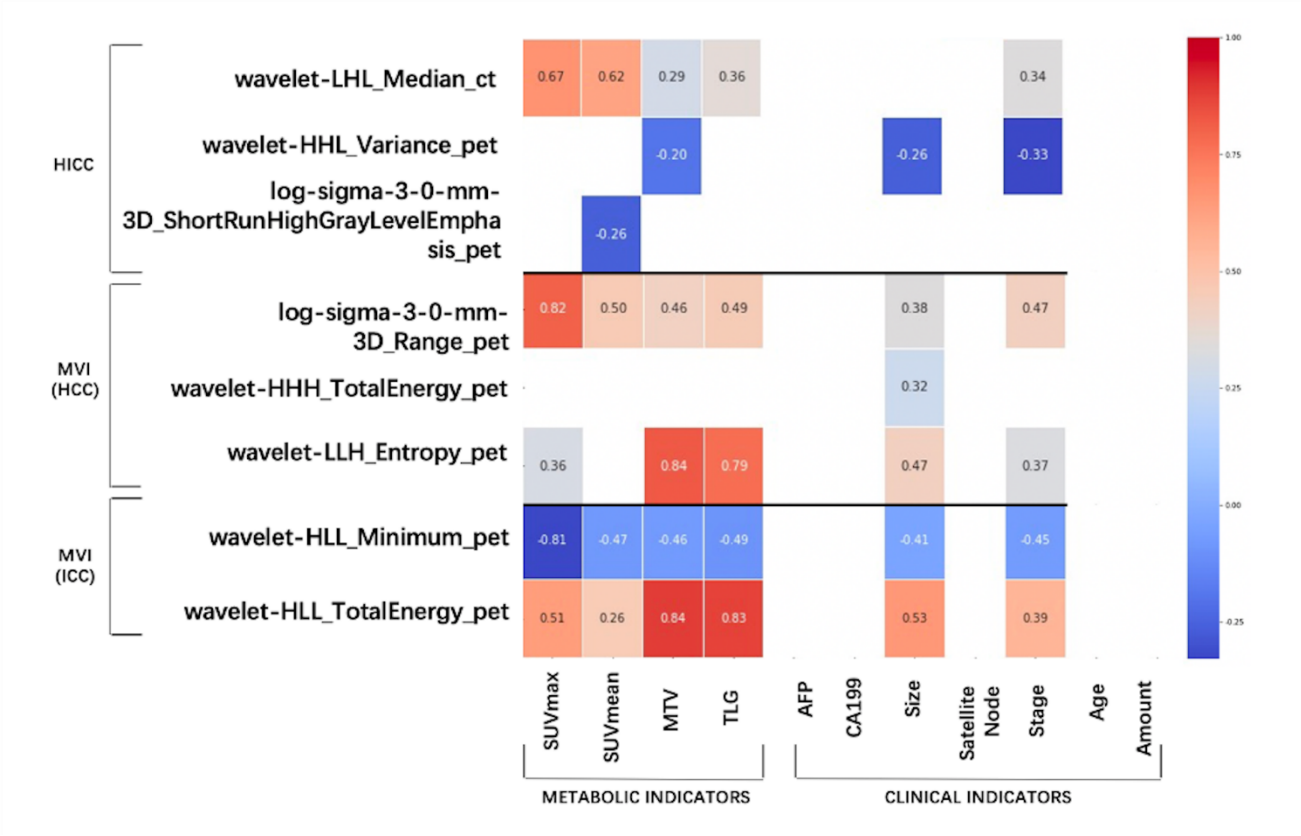

Figure 3-7 Correlation analysis of radiomic features with clinical features. *Strong correlation, $p \leq 0.05$; red color denotes positive correlation, blue denotes a negative correlation, and the shade of the color indicates the correlation intensity.

3.4 Discussion

HCC and ICC are two common subtypes in primary liver cancer with distinctive prognosis [135]. The metastasis and recurrence are the two major obstacles to improve the prognosis of liver cancer patients. More importantly, MVI status is an indicator of tumor's aggressiveness and an independent risk factor of metastasis and recurrence[115]. Hence, to provide precise information and appropriate treatment, the prediction of HCC and ICC classification and MVI

statues before surgery is crucial. CEUS is commonly used in HCC and ICC differentiation and achieves an AUC of 0.92[122], and MRI for MVI detection can achieve an AUC of 0.86 for HCC[123], an AUC of 0.81 for ICC[125]. Though the specialized medical imaging could give relatively high accuracy in detection, it's a great burden for patients to do many examinations. Encouraged by PET's promising applications, we aimed to explore whether 18F-FDG PET/CT imaging could provide a potential possibility for playing an auxiliary diagnosis and additional contribution for HCC and ICC classification and MVI before surgery, so that patients can obtain the comprehensive quantification of tumor phenotypes after a single preoperative 18F-FDG PET/CT examination for guiding oncologists or surgeons to establish a personalized therapeutic strategy. Generally, there are two findings in our study: we constructed a prediction model for HCC and ICC classification and MVI statues in primary liver cancer based on 18F-FDG PET/CT radiomics features and clinical factors. Moreover, we found that PET features had an impressive prediction capability in HCC and ICC classification and MVI, which outperformed than CT and clinical characteristics.

Raman et al. [136] described that radiomics features using computed tomography texture analysis can detect different liver lesion types and normal liver tissue. Our results showed that PET and CT radiomics features achieved achieved AUC of 0.86 (compared to 0.92 in CEUS) and a specificity of 0.88 for HCC and ICC classification, which has potential for informative reference. We also found that factors related to tumor intensity and texture were the most important components in predicting histological classification. This is partly in agreement with the findings of Minghui et al. [137]. Specifically, tumor intensity and texture features can reflect subtle information from PET/CT images. For instance, Median and Variance represented tumor area's degree of heterogeneity. Short Run High Gray Level Emphasis, as a texture feature, revealed joint distribution of dark small areas in VOIs. PET/CT radiomics provides molecular-based

image features and intratumoral heterogeneity [138], which could be an effective diagnostic tool in histological classification for primary liver cancers. This finding is contributed to the evaluation of differentiation between HCC and ICC, especially in cases that differentiation using conventional medical imaging methods is difficult.

Previous studies [139, 140] validated that MVI worsened the prognosis of liver cancer. Emerging studies have focused on the relationships of contrast-enhanced CT features and MVI status [138, 141]. But prediction model of MVI based PET/CT radiomics features has never been reported. In our study, three PET features integrating tumor clinical stage in HCC and two PET features integrating one clinical factor in ICC were selected for MVI prediction. The compound (PET, CT, and clinical characteristic) radiomics predictors can identify more than 0.77 of the MVI-positive cases with the AUC of 0.88~0.90 (compared to 0.86 and 0.81 in MRI). Our model exhibited better performance with the MRI model in both HCC and ICC. Besides, PET features were more important than clinical features, and intensity features seemed to perform better than texture features. One possible interpretation is that tumor intensity and texture features implied a range of discrete tumor activity and intratumor heterogeneity. Another is that PET/CT's spatial resolution is relatively low, and it has less advantages in defining the tumor boundaries. The greater values of these factors, the higher probability of MVI. This finding is consistent with the previous report [142] that the radiomics signature, nonsmoothed tumor margin, hypoattenuating halos and internal arteries were significantly associated with MVI status. We also found a positive association between some higher-order PET radiomics features (Range, Total Energy, Total Energy) and the 18F-FDG uptake activity (SUV max, SUV mean, TLG) of the lesion. This part of results is important since it may be an indicator of disease extent and tumor staging, especially in cases where evaluation using conventional clinical imaging methods might have been overlooked.

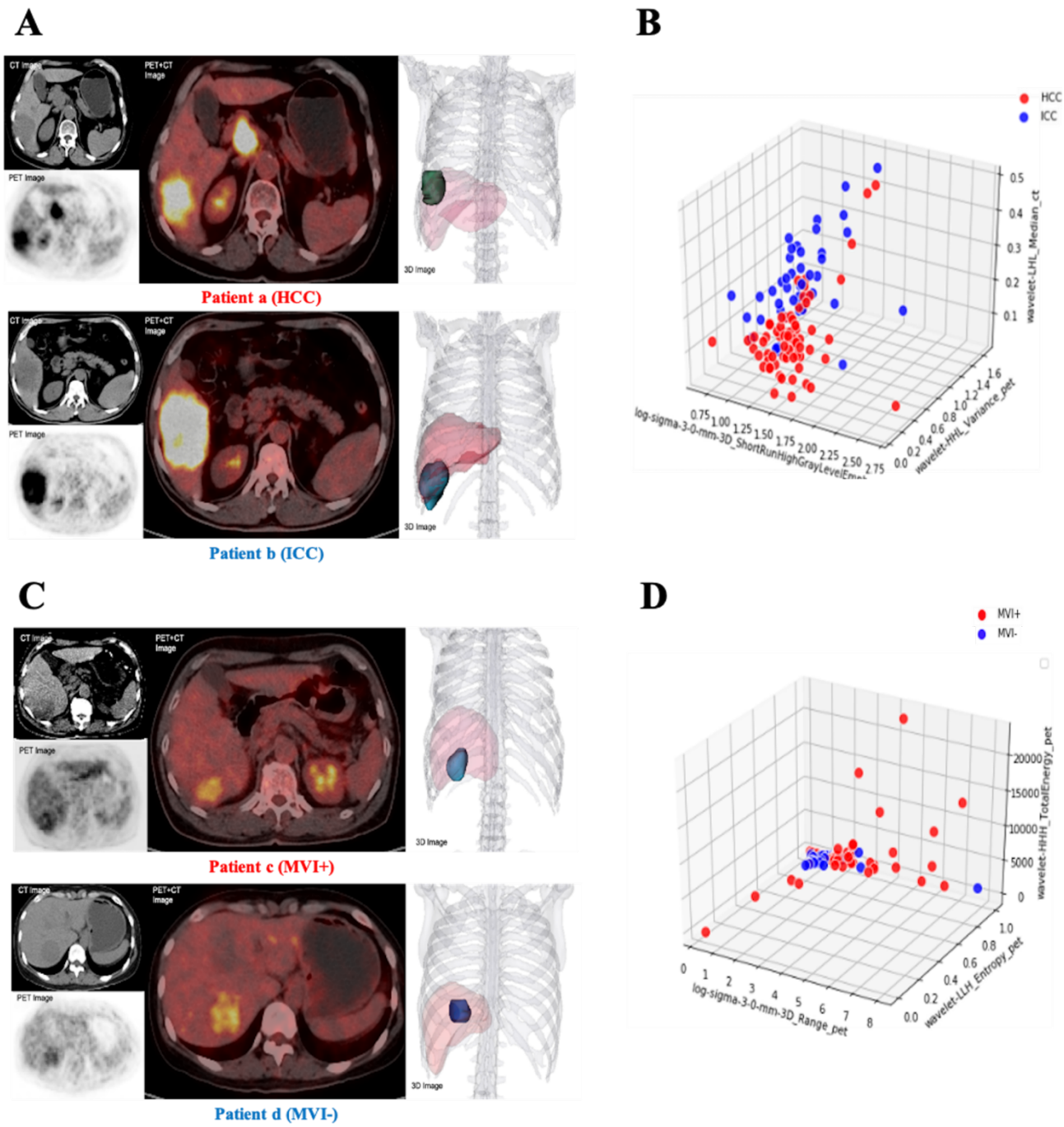

Figure 3-8 The lesions of all four patients are located in the right lobe of the liver. [1] patient a is a 43-year-old man with HCC and patient b is a 61-year-old woman with ICC. Both A and B showed high uptake and correctly predicted by machine learning model. (B) is a joint distribution of three selected radiomic features in a 3D space. There is a relatively clear distinction between HCC and ICC. [1] patient c is a 37-year-old man with MVI positive and patient d is a 69-year-old man with MVI negative. Both C and D showed moderate uptake and were correctly predicted by machine learning model. [1] is a joint distribution of three selected radiomic features in a 3D space. There is a relatively clear distinction between MVI+ and MVI-. *Abbreviation: **HCC**, Hepatocellular carcinoma. **ICC**, intrahepatic cholangiocarcinoma (ICC). **MVI**, microvascular invasion.

Notably, our model exhibited equivalent or superior performance with the CEUS and MRI model in HCC and ICC classification and MVI prediction. Therefore, 18F-FDG PET/CT imaging would contribute to its key role in evaluating and staging tumors and potential value in differentiating ICC and HCC and detecting MVI before surgery, which could help to provide an earlier indication of liver cancer to select a more appropriate treatment and relieve the medical burden of patients. Interestingly, our model revealed that PET features had dominant predictive power in HCC and ICC classification and MVI, which outperformed than CT and clinical characteristics. In this study only one CT feature was selected by in HCC and ICC classification task. The reasons may be as follows: CT scanning in PET/CT is unenhanced low-dose CT, which only provides limited information and is not capable for sufficient tumor detection or distinction. While PET reflects the metabolic activity of a whole tumor [134]. Though the importance of features in unenhanced low-dose CT has been validated in many studied[143, 144], such as HCC surveillance analysis[145], esophageal cancer[146] and lymphoma[147] , the value of CT features on HCC and ICC differentiation still requires a larger sample for further validation. Further, CT features were not included in the MVI model. One potential explanation is that MVI detection is relevant to find out the presence of tumor cells in inside portal or hepatic venous systems[107] , and most hepatic cancer cells were hypermetabolic activity stemming from an increased metabolic plasticity, which can be identified by PET/CTscan, especially in PET images. Hence unenhanced low-dose CT is insufficient in MVI detection in this study.

Our study has some limitations. Firstly, most clinical characteristics of patients cannot add the accuracy of the predictive model. It might be attributed to the small sample and a potential selection bias in this single-center retrospective study. Multicenter and larger clinical studies are necessary to be designed for validating our radiomics model. Nonetheless, our findings are

still reasonable and important. Cochran's formula[148] : if it's assumed 50% of patients are positive in 95% confidence level and 5% margin error, the ideal sample size for 382. Besides, machine learning models require around 50 patients for algorithm's training and validation to avoiding overfitting. Further, for PET/CT studies, Chalkidou et al.[149] found that for one radiomic feature, 10 to 15 patients are the minimum requirement. For three different tasks, though the idea sample size of 382 wasn't achieved, our feature selection model reduced the number of features to 3 (out of 127 patients), 4 (out of patients) and 2 (out of 51 patients), which suggests that our results are relatively valid with the minimum false detection rate. In addition, prognosis information with histologic MVI were not collected to investigate the predictive effectiveness of the model. The prediction model based on 18F-PET/CT radiomics features that widely used for liver cancer will be continually explored in future studies.

3.5 Chapter Summary

We construct a result-driven feature selection framework, which could efficiently reduce the dimension of the feature space and simultaneously avoid overfitting. The framework integrates different feature selection techniques to ensure the distinctiveness, uniqueness, and generalization ability of the output feature set.

CHAPTER 4. Quantitative Dynamic Descriptor in Regional Pulmonary Perfusion Visualization and Anomaly Detection

4.1 Introduction

Recent developments in lung imaging have a significant influence on the capacity to diagnose and quantify pulmonary diseases, giving both structural and functional data. Moreover, with the advent of digital chest radiography, AI-assisted diagnosis is becoming increasingly practical and significant[150]. Pulmonary metabolic process alterations, as well as those involving blood flow, regional chemical composition, and absorption, may be seen and measured by Positron Emission Tomography (PET) [22], a functional imaging method that employs radioactive chemicals called as radiotracers.

However, most PET studies are based on visual evaluation on static images (usually 60min after injection) and calculation of SUV values. In contrast, dynamic whole-body positron emission tomography/computed tomography (dPET/CT) scanner can provide more information by extracting physiological and pharmacokinetics properties from a tracer in tissue or tumours, allowing tracer kinetics to be recorded over time, not just at a certain point in time after the tracer is injected in a static image manner [151, 152].

Compared to conventional PET scanner with short axial field-of-view (AFOV), the total-body scanner with coverage of approximately 200 mm provides a higher sensitivity, a higher signal-to-noise ratio (SNR) and a lower radiation dose in human body imaging [153-155].

18F-FDG PET/CT has gained widespread clinical acceptability for its use in the care and evaluation of patients suffering from a range of pulmonary diseases. By analysing dynamic 18F-FDG PET images, we observed that the dynamic PET image series present discernible perfusion patterns among patients. It also has been validated that [156] lung perfusion indicators,

such as: permeable surface area product (PS), blood flow (BF), tumor micro vessel density (MVD) is significantly correlated with SUV in FDG PET ($p<0.05$). Hence, from patients' 18F-FDG dynamic PET images, a computer-vision based descriptor is expected to build for pulmonary perfusion analysis.

In this study, we explore and build a perfusion index based on 18F-FDG dynamic PET in the view of computer vision descriptor. The descriptor can reflect visually uneven perfusion of patients' lung in a quantitative way. It also has potential to provide a more comprehensive assessment of patients' pulmonary diseases.

4.2 Material and Method

4.2.1 Dataset description

Patients with Non-small cell lung cancer (NSCLC) were included in this study. Inclusion criteria were as follows:(1) untreated confirmed NSCLC; (2) between 18 and 75 years; (3) had unresectable stage IIIA-IV disease according to the 8th edition of the American Joint Committee on Cancer staging system; (4) expectancy of life ≥12 weeks.

4.2.2 Image Acquisition and Reconstruction Parameters

Patients were asked to avoid strenuous exercise 24 hours before each study and fast for up to 6 hours prior to PET/CT imaging. The image was acquired from head to toe between 0 ~ 60 minutes after the injection of 3.0 MBq/kg of FDG administered through the patient's feet. The methodology for total-body FDG-PET/CT imaging entails 60 minutes of dynamic capture right after the FDG injection and 10 minutes of delayed static acquisition two to three hours later. Prior to each PET acquisition, the Attenuation Correction Computed Tomography (ACCT) is

obtained, and a diagnostic CT is scanned before the dynamic PET scan. The acquisition of each PET scan occurs in list mode. In this study, we found perfusion-like patterns only exist in the initial image sequences, thus only the first 25 frames were selected for image analysis.

4.2.3 Quantitative Dynamic Descriptor Representation and Application

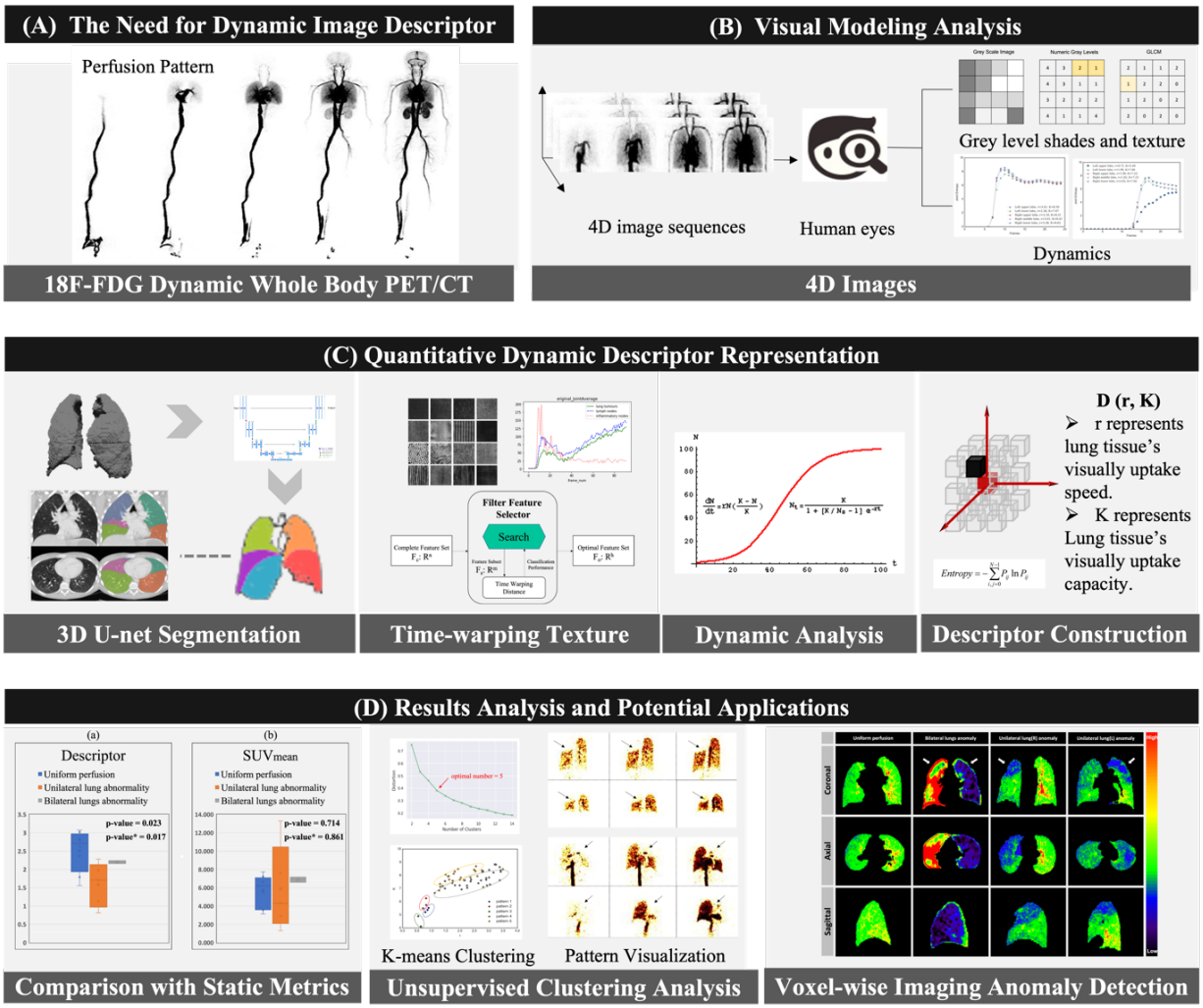

Figure 4-1 Workflow of descriptor construction

As shown in the Figure 4-1, the construction of descriptor includes some main steps. Based on patients' 18F-FDG PET images, since the visual perfusion pattern only appears in the initial uptake time, the first 3-min dynamic series of a 60-min data acquisition protocols are discussed.

Firstly, visual modeling analysis of 4D image serious was performed. Secondly, the VOIs were defined by automatic deep-learning-based segmentation. Then the texture representation was defined by a time-warping algorithm and dynamic analysis, based on which the descriptor was constructed. At last step, we performed unsupervised clustering and voxel-wise imaging for pulmonary anomaly detection compared to static image features.

4.2.3 Visual Modelling Analysis for Dynamic PET/CT Images

While searching for effective descriptors to represent dynamic picture sequences, it is reasonable to consider the sorts of features that humans employ to understand visual information. Generally, grey level shades, textural, and dynamics are three fundamental dimensions for human to interpret images [157]. Spectral features indicate the average total changes in separate bands of the visible and infrared sections of an electromagnetic spectrum, while textural features describe the spatial distribution of tonal variations within a band [157].

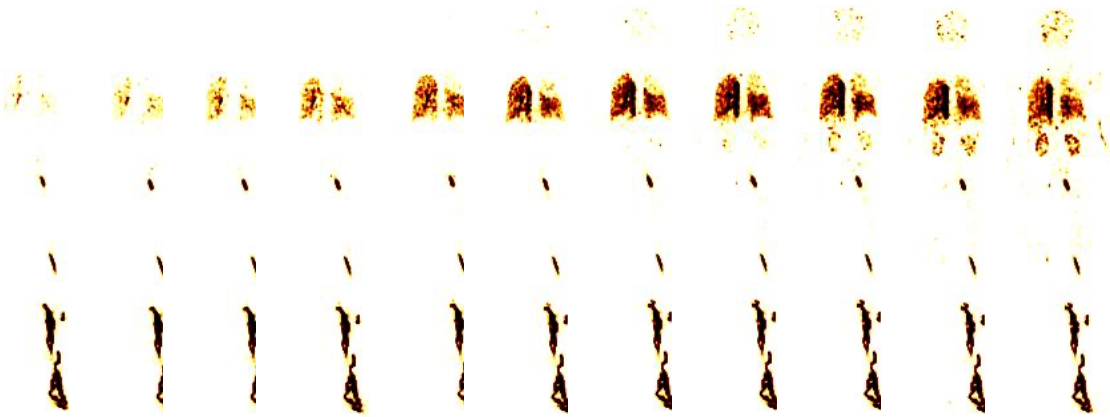

Figure 4-2 Hot MIP of total-body FDG PET images from 0min to 3min

Figure 4-2 demonstrated a hot MIP of total-body FDG PET images from 0min to 3min. The overall spatiotemporal pattern of pulmonary in dynamic PET/CT images perceived by human

eyes tends to be a coherent process: the uptake in lung area gradually accumulated and then diminishes. That's because, for soft tissue in lung, 18F-FDG will be initially absorbed and then metabolized. Besides, previous studies [158] have validated that the ventilation and vascularization of the lung parenchyma are correlated with the 3D textural features of the pulmonary soft tissue, and pulmonary embolism results in wedge-shaped pleura-based zones of heterogeneous enhanced attenuation. This study mainly focuses on pulmonary visually uptake pattern rather than the vanishing process.

Hence, in the view of computer vision and graphics, we have the assumption that the descriptor modeling capturing varying shades of grey levels, texture, and dynamics, is adequate for analysis.

4.2.4 3D U-net segmentation for VOI definition.

The first step to build a visual descriptor is to give the definition of its VOI. In this study, the descriptor needs to reflect visually uneven perfusion of patients' lung in a quantitative way. Hence, to ensure accuracy and robustness, we define the region (or VOI) of descriptor as a single lobe in the lung. Basically, the left lung contains the upper and lower lobes, and the right lung contains the upper, middle, and lower lobes. The descriptor calculation is based on one single lobe. A robust deep learning-based segmentation method (U-net, LTRCLobes_R231) [159] was performed on CT to generate lung masks. The model uses a single slice to extract the left and right lungs separately, including air pockets, tumors, and effusions. Trachea will be excluded in the segmentation.

4.2.4 Time-warping Texture pattern representation.

The second step is to give the representation of descriptor's region. Generally, we have two choices: describe an area in terms of its external attributes (its boundaries) or its internal attributes (such as the pixels that make up the area). In this study, the description of perfusion involves changes in VOIs' texture. From segmented VOIs of lungs, totally fifty texture features were extracted with the PyRadiomics package which is compliant with the Imaging Biomarker Standardization Initiative [17].

Before quantitative analysis, we need to select the feature that could best qualitatively describe visual changes in lobes' uptake. In this study, we utilized a general inductive framework and physiological analysis. The time series curves of all intensity and texture features extracted in the previous step are drawn. For every feature, there are five curves representing five different VOIs: left upper lobe, left lower lobe, right upper lobe, right middle lobe, and right lower lobe. According to the curves, we give feature' inclusion criteria:

(1) Based on lung tissue's physiological characteristics and visually analysis (slow uptake, fast uptake, and slow uptake to stable, see figure 4-2), the feature's curves are generally in the shape of a logistic growth (S-shape).
(2) The feature's curves can distinguish visually uneven perfusion between right and left lung.
(3) The feature's curves can distinguish visually uneven perfusion among upper, middle, lower lodes of unilateral lung.

To achieve criteria (1), we perform distribution detection for all features. The skewness-kurtosis plot proposed by Cullen and Frey [98] is an effective method to check the distribution. The skewness could reflect symmetry, while the kurtosis reveals tails' weight compared to normal

distribution. The skewness, sk and kurtosis, kr [99] and their unbiased estimator from $(X_i)_i \sim X$ with observations $(x_i)_i$ are:

$$sk(X) = \frac{E[(X-E(X))^3]}{Var(X)^{\frac{3}{2}}}, \widehat{sk} = \frac{\sqrt{n(n-1)}}{n-2} \times \frac{m_3}{{m_2}^{\frac{3}{2}}} \text{ (Equation 4.1)}$$

$$kr(X) = \frac{E[(X-E(X))^4]}{Var(X)^2}, \widehat{kr} = \frac{n-1}{(n-2)(n-3)}\left((n-1) \times \frac{m_4}{{m_2}^2} - 3(n-1)\right) + 3$$

(Equation 4.2)

where m_1, m_2, m_3, m_4 denote empirical moments.

Besides, a nonparametric bootstrap procedure is performed to ensure estimators' robustness [160]. The boot was set to 500. Based on Cullen and Frey graph, we can find the feature that fits the shape of Logistic Curve. Statistical analyses were performed with 'Scipy 1.3.0', and 'math' packages in Python 3.6.8 programming language and environment.

To achieve criteria (2) and (3), we need to calculate the similarity between different curves. Due to pulmonary vascular structure and gravity effects [161], the curves don't need to align. Therefore, we utilized Dynamic Time Warping (DTW) distance[162] , which does not give low similarity score for curves with similar shape and different phase, for all curves pairs of every feature to calculate curves' similarity. For a good feature, its' curves pairs with same perfusion pattern should have shorter DTW distance. Otherwise, the curves paired with different perfusion patterns should have longer DTW distances. The DTW distance is calculated by dynamic programming:

$Input$: $Time\ series\ Q\ and\ C\ with\ same\ length\ n\ where$

$$Q = q_1, q_2, \dots, q_n \quad C = c_1, c_2, \dots, c_n$$

$Output$: $DTW\ distance\ between\ Q\ and\ C$

$Step\ 1$: $Construct\ an\ n \times n\ matrix\ M, whose\ i, j^{th}\ element\ is\ the\ Euclidean\ distance\ between\ q_i\ and\ c_j$

$$M = m^1, m^2, \dots, m^k, where\ each\ element\ of\ M\ represents\ the\ distance\ between\ a\ point\ i\ in\ Q\ and\ j\ in\ C$$

$Step\ 2$: $Find\ the\ path\ with\ the\ minimum\ distance$: $M^* = argmin_m(\sqrt{\sum_{k=1}^{K} m_k})$

$Step\ 3$: $Solve\ the\ problem\ by\ solving\ recursive\ equation$:

$$\gamma(i,j) = d(q_{i,}c_j) + min\ (\gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1))$$

Algorithm.1 Pseudo Code for DTW distance calculation

For qualitative descriptor construction step, GLCM joint entropy is selected as the best representative feature. GLCM is a matrix (Figure 4-3) that describes the joint distribution of discretized intensities (grey levels) of neighboring pixels, or voxels in a 3D volume [37].

GLCM Joint entropy reflects the variability in neighbourhood intensity values. The definition:

$$joint\ entropy = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) log_2\ (p(i,j) + \epsilon) \quad \text{(Equation 4.3)}$$

The ROI is defined by $\boldsymbol{P}(i,j|\delta,\theta)$. The $(i,j)^{th}$ element in GLCM represents the value of times the levels i and j occur in two pixels, which separated by pixels' δ distance of θ angle. A greater value of GLCM-joint entropy implies more texture pattern heterogeneity.

Dynamics pattern representation. After we have determined the optimal features, we need to quantitatively reflect the pattern of perfusion procedure.

The logistic growth function with time t:

$$P(t) = \frac{KP_0e^{rt}}{K+P_0(e^{rt}-1)} \text{ (Equation 4.4)}$$

The two parameters r and K are calculated based on fitting the parameter Logistic Curve by least squares regression equation method.

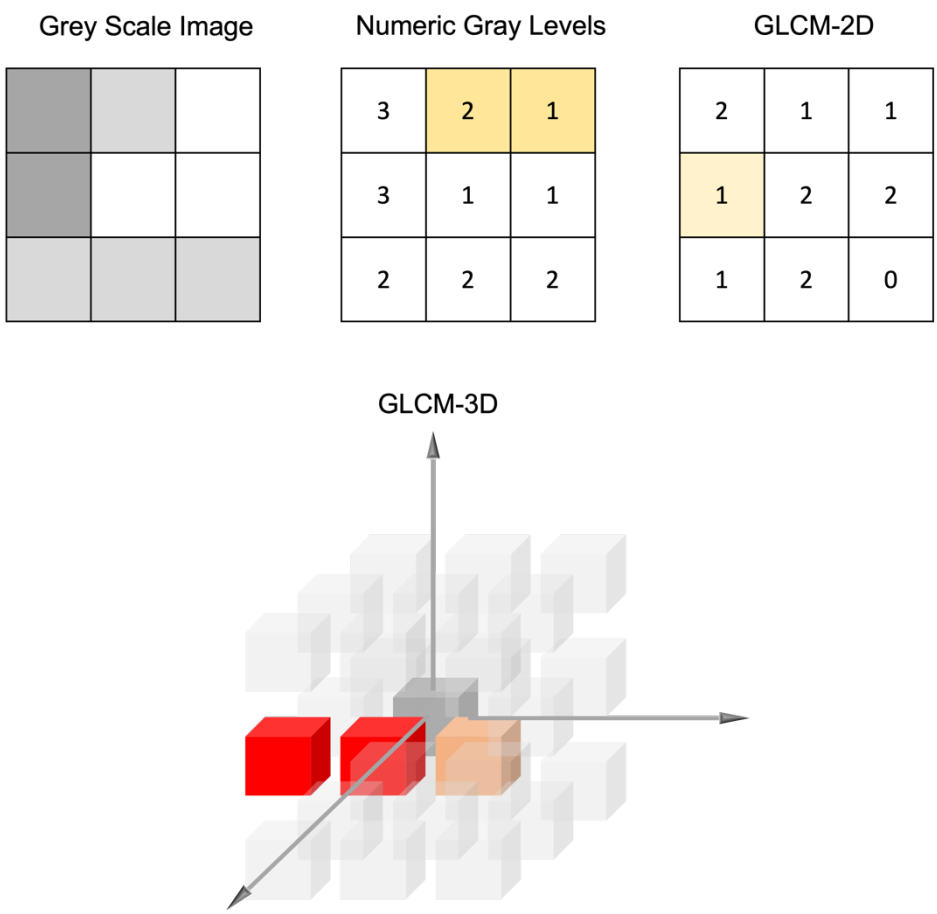

Figure 4-3 GLCM in 2D and 3D

A two-dimensional dynamic descriptor. Since the feature curve is generally consistent with Logistic Growth, we build the two-dimensional perfusion descriptor D (r, K) by GLCM-joint entropy, in which:

(1) r represents lung tissue's visually uptake speed for tracers.

(2) K represents lung tissue's visually capacity for tracers.

Unsupervised K-means Clustering for learning local lobe's patterns. Patients were classified into distinct classes using an unsupervised K-Means clustering approach. The algorithm was performed using scikit-learn package (https://scikit-learn.org). Descriptor value (r, K) extracted from VOIs were evaluated.

Voxel-wise imaging for perfusion anomaly detection. A voxel-wise descriptor was also performed for pulmonary anomaly detection. GLCM joint entropy, which characterize texture heterogeneity, has been validated as one of repeatable voxel-wise features in improvements in providing structurally different information, and correlates better with the tumor biology and clinical outcomes [163]. We represent each voxel with descriptor calculated in its closest eight neighborhood pixels. A distance between two descriptors is calculated by entropy weight [164] sum of squares for visualization.

4.3 Results

4.3.1 Demographics of Patients

A total of thirty newly diagnosed stage IIIA-IV NSCLC patients were prospectively enrolled between September 2020 and December 2020. The clinical characteristics of twenty patients were summarized in table 4-1. Of all patients, the median age at diagnosis was 57 years (range, 41-68). Twenty cases (87%) were males and three were females (13%). Twenty-one patients (91.3%) had stage III and two (8.7%) had stage IV diseases. There were ten cases (43.5%) with squamous cell carcinoma and eleven cases (47.8%) with adenocarcinoma. The median FEV1 for all the patients was 2.3L (range, 0.99-3.39).

Characteristics	Number (%)
Age, years	57 (41-68)
Gender	
Male	20 (87.0)
Female	3 (13.0)
ECOG	
0	5 (21.7)
1	18 (78.3)
Smoking	
Yes	12 (52.2)
No	11 (47.8)
Tumor location	
Left upper	9 (39.1)
Left lower	3 (13.0)
Right upper	7 (30.4)
Right middle	1 (4.3)
Right lower	2 (8.7)
Mediastinum	1 (4.3)
Histology	
Squamous	10 (43.5)
Adenocarcinoma	11 (47.8)
Lymphoepitheli-oma-like	1 (4.3)
NSCLC-NOS	1 (4.3)
Stage	
IIIA	5 (21.7)
IIIB	12 (52.2)
IIIC	4 (17.4)
IV	2 (8.7)
FEV1, L	2.3 (0.99-3.39)

Table 4-1 Clinical characteristics of patients

4.3.2 Unsupervised K-means clustering for descriptor patterns classification

A K-means cluster algorithm was performed by descriptor D (r, K) for categorizing lobes into different visual patterns. The Elbow Method was used by the Python Sklearn module to establish the ideal number of clusters into which the data may be divided. The distortion and inertia

values for each value of k in the specified range may be calculated by iterating over the values of k fro n and inertia start to l ne given data is five.

Figure 4-4 Elbow method for K-means clustering

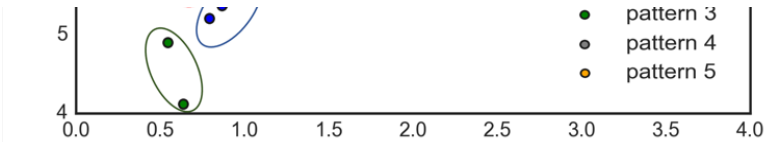

Figure 4-5 Visualization of clustering results

The preliminary grouping criterion is based on lobe distribution for each pattern. Since pattern 1, pattern 4 and pattern 5 all include five lobes, we classify it as whole lung patterns. Pattern 2 only includes left upper lobe and pattern 3 only includes right upper lobe, hence we classify them as single lobe patterns.

We performed a further grouping based on visual uptake speed (Figure 4-6). For whole lung patterns, pattern 1(r = 1.868), 4 (r = 1.973) and 5(r = 1.827) were defined as whole lung slow

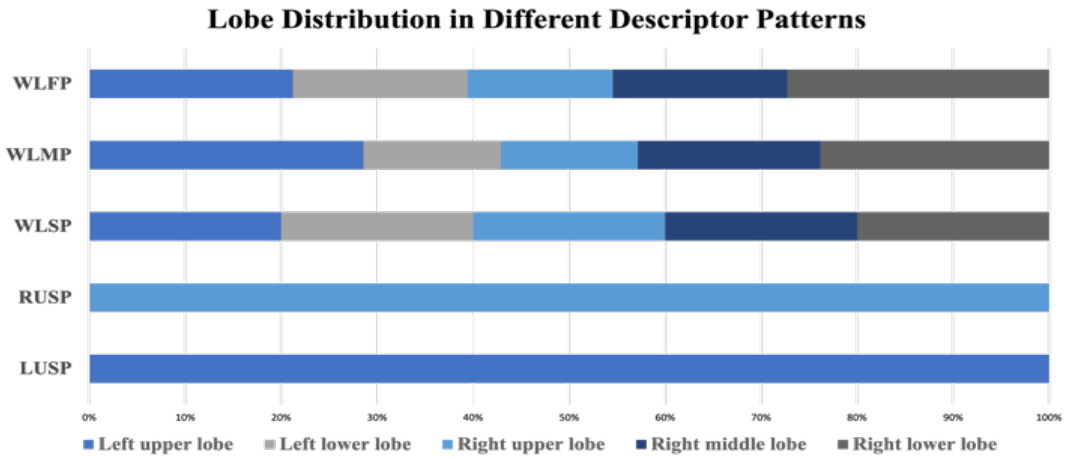

Figure 4-6 Lobe distribution in different descriptor patterns

pattern (WLSP), whole lung medium pattern (WLMP) and whole lung slow pattern (WLSP), respectively. For single lobe patterns, pattern 2 (r = 0.859) and pattern 3 (r = 0.843) was defined as right upper lobe slow pattern (RUSP) and left upper lobe slow pattern (LUSP).

Pattern	**R**		**K**	
	Mean	stdev	Mean	stdev
WLFP	1.97301018	0.97263815	6.95690717	1.14150983
WLMP	1.82744345	1.00800428	6.5324481	1.14861448
WLSP	1.60836739	0.00085102	5.37051834	0.00636746
RUSP	0.84326999	0.04347414	5.3136795	0.09855066
LUSP	0.85943405	0.04801155	5.36724245	0.13398052

Table 4-2 Mean and standard deviation for each pattern

Figure 4-6 shows a detailed lobe distribution in descriptor patterns. Table 4-2 shows mean and standard deviation of D (r, K) for each pattern. WLFP exists most in the right lower lobe, and it also has fast visual uptake speed (r = 1.973) and higher visual uptake capacity (K = 6.956). WLMP exists most in left upper lobe, while it has relatively moderate visual uptake speed (r =

1.827) and visual uptake capacity (K = 6.532) compared to WLFP. All five lobes have an average amount of WLSP, which has substantially slower visual absorption rates (r = 1.60) and poorer visual uptake capacities (K = 5.37). With a similar pattern of WLSP, RUSP and LUSP are found in a single upper lobe and have sluggish visual uptake capacities (K = 5.313 and K = 5.367) as well as slow visual uptake speed (r = 0.843 and 0.859).

4.3.3 Case Studies for Whole Lung Patterns and Single Lobe Patterns

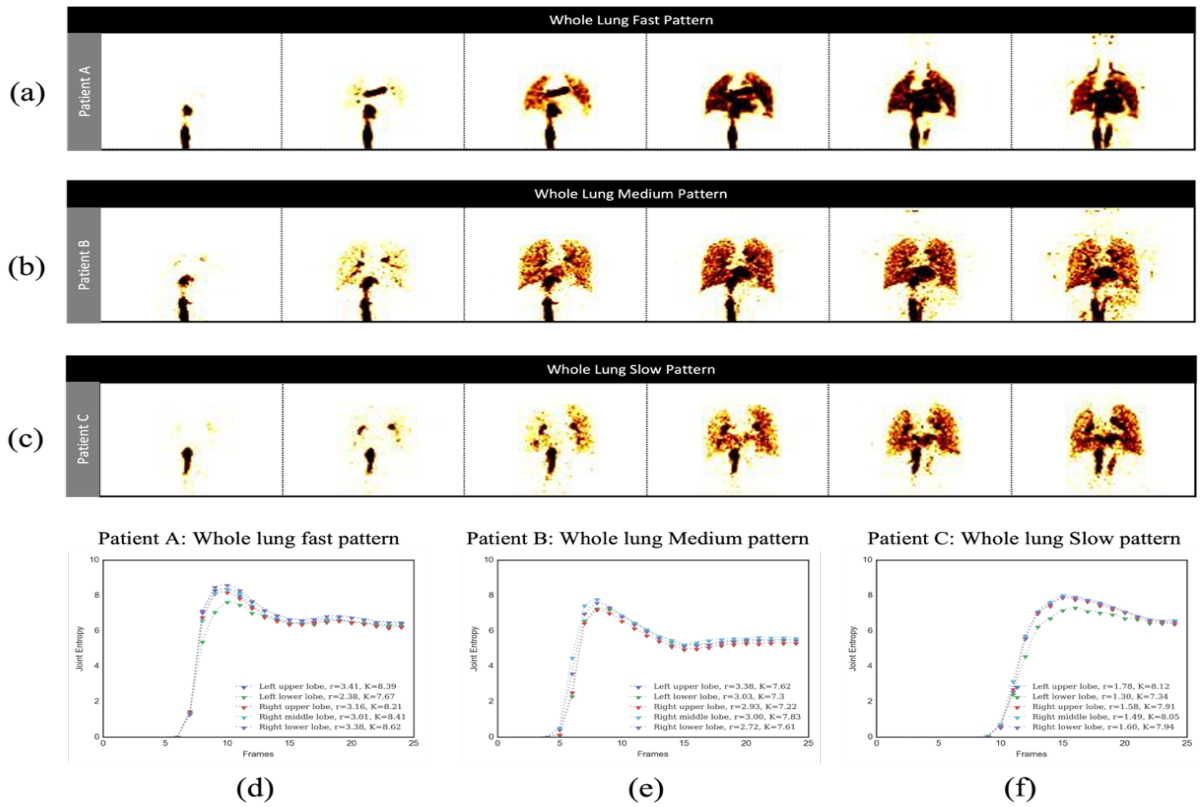

Figure 4-7 Whole lung perfusion patterns

As illustrated in Figure 4-7, descriptor **D (r, K)** could discriminate visually different perfusion pattern of whole lung in a quantitative way. For patient A with WLFP, the descriptor value for his left upper lobe, left lower lobe, right upper lobe, right middle lobe, right lower lobe is (3.41, 8.39), (2.38, 7.67), (3.16, 8.21), (3.01, 8.41), (3.38, 8.62), respectively. For patient B with WLMP, the descriptor value for his left upper lobe, left lower lobe, right upper lobe, right middle lobe, right lower lobe is (3.41, 8.39), (2.38, 7.67), (3.16, 8.21), (3.01, 8.41), (3.38, 8.62), respectively. For patient C with WLSP, the descriptor value for his left upper lobe, left lower lobe, right upper lobe, right middle lobe, right lower lobe is (1.78, 8.12), (1.30, 7.34), (1.58, 7.91), (1.49, 8.05), (1.60, 7.94), respectively. Figure 8a and Figure 8b show that WLFP and WLMP share similar uptake speed, but WLFP has larger capacity than WLMP. On the other hand, WLSP in Figure 8c represents slower uptake speed and smaller capacity than WLFP and WLMP.

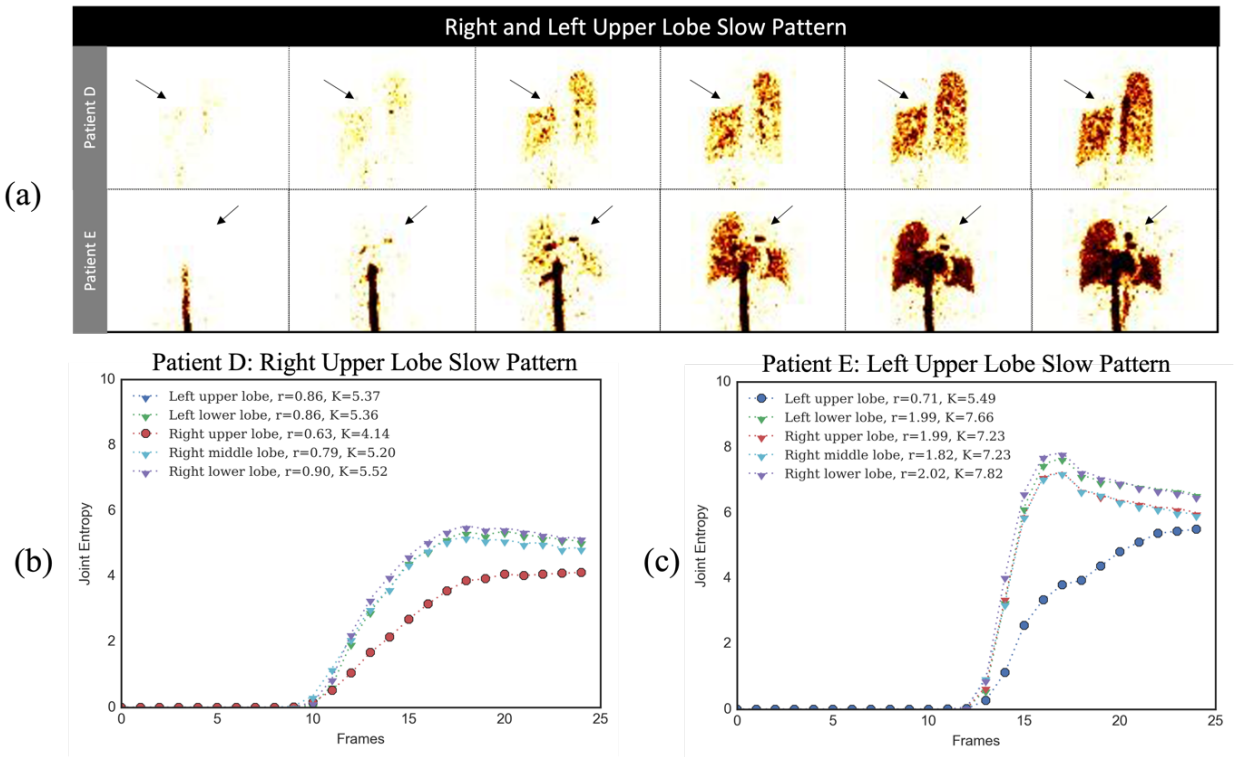

Figure 4-8 Right/left upper lobe slow pattern

Figure 4-8 shows descriptor **D (r, K)** could discriminate visually different perfusion pattern of single lobe in a quantitative way. For patient D with RUSP, the descriptor value for his left upper lobe, left lower lobe, right upper lobe, right middle lobe, right lower lobe is (0.86, 5.37), (0.86, 5.36), (0.63, 4.14), (0.79, 5.20), (0.90, 5.52), respectively. For patient E with LUSP, the descriptor value for his left upper lobe, left lower lobe, right upper lobe, right middle lobe, right lower lobe is (0.71, 5.49), (1.99, 7.66), (1.99, 7.63), (1.82, 7.63), (2.02, 7.82), respectively. Figure 9a shows that RUSP and LUSP both have slower uptake speed and smaller capacity than WLFP, WLMP and WLSP.

4.3.3 Voxel wise imaging for unilateral lung and bilateral lung perfusion anomaly

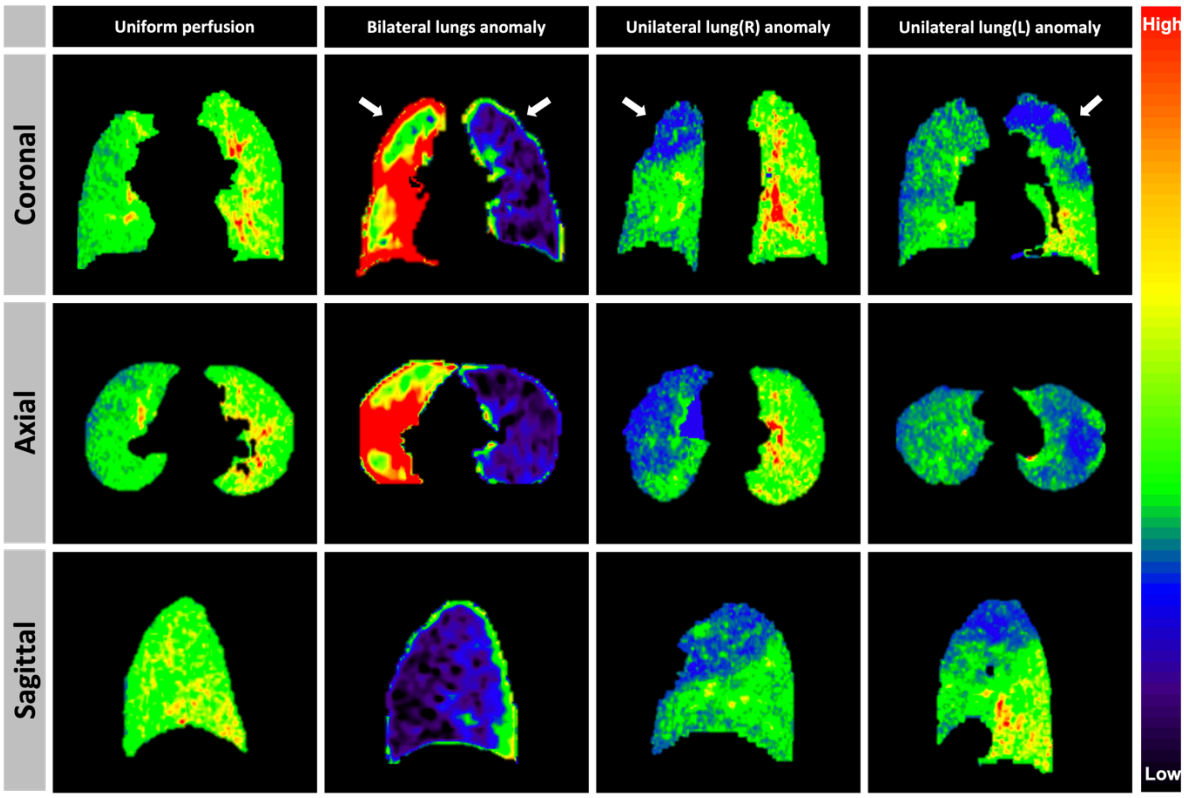

Figure 4-9 Voxel-wise imaging of pulmonary perfusion maps

Figure 4-9 demonstrates descriptor's voxel-wise imaging in axial, coronal, and sagittal views of uniform perfusion (first column), bilateral lungs anomaly (second column), unilateral lung(R) anomaly (third column) and unilateral lung(L) anomaly (fourth column). Table 4-3 shows descriptor value of single lobe for uniform and anomaly patterns. Compared to uniform perfusion, the bilateral lungs anomaly and unilateral lung anomalies all have abnormal lobes with descriptor value (r, K) lower than other lobes

	Uniform Perfusion	Bilateral lungs anomaly	Unilateral lung(R) anomaly	Unilateral lung(L) anomaly
Left upper lobe	(3.2, 7.6)	**(0.7, 5.4)**	(2.78,8.2)	**(0.81, 5.49)**
Left lower lobe	(3.0, 7.3)	**(0.5, 4.8)**	(3.2, 8.8)	(1.89, 7.6)
Right upper lobe	(2.9, 7.2)	(3.7, 8.3)	**(0.8, 5.7)**	(1.99, 7.2)
Right Middle lobe	(3.0, 7.8)	(3.5, 8.3)	(2.8, 8.4)	(1.82,7.2)
Right lower lobe	(2.8,7.6)	(3.5, 8.6)	(2.5, 7.1)	(2.02, 7.8)

Table 4-3 Descriptor value for uniform and anomaly patterns

Comparison with static PET metabolic metrics of visual perfusion abnormality detection. As shown in Figure 4-10, the descriptor value of uniform perfusion is significantly higher than unilateral lung abnormality (p-value = 0.023, $p<0.05$) and bilateral lung abnormality (p-value = 0.017, $p<0.05$). However, for traditional PET metabolic metrics, there is no significant difference among uniform perfusion, unilateral lung abnormality and bilateral lung abnormality.

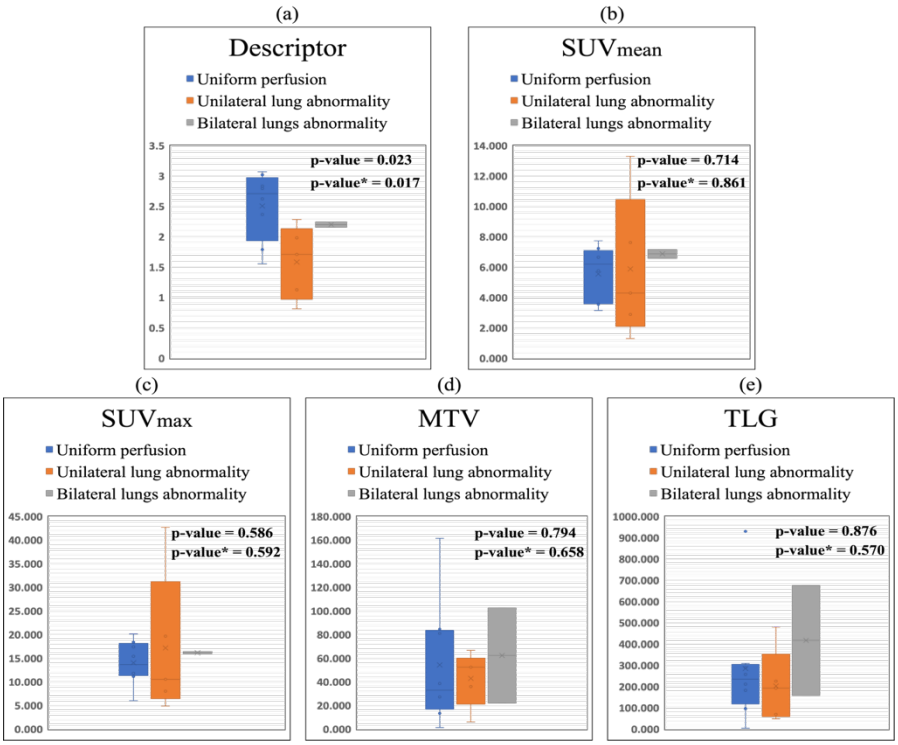

Figure 4-10 The comparisons of descriptor value and SUVmean, SUVmax, MTV and TLG among uniform perfusion, unilateral and bilateral abnormality

4.4 Discussion

In this prospective study, we explore and build a pulmonary perfusion descriptor based on 18F-FDG whole-body dynamic PET images in the view of physiology and computer vision. Our major finding includes: 1) We proposed a descriptor construction framework which is interpretable in the perspective of both physiology and computer vision. 2) The quantitative comparison of descriptor intra and inter patients is feasible, making it possible to be an auxiliary diagnostic tool in pulmonary function assessment. 3) Compared with traditional PET metabolic biomarker analysis, the proposed descriptor incorporating image's temporal information, which enables a better understanding of the time-various mechanisms and visual perfusion

abnormalities detection among different patients. 4) The proposed descriptor eliminates the impact of vascular branching structure and gravity effect by utilizing time warping algorithms.

The first finding of this study is that the proposed descriptor can reflect dynamic textural variations, as well as physiological changes in patient's pulmonary uptake. High imaging sensitivity is a substantial benefit of the whole-body PET/CT scanner [165]. By analysing dynamic image series from 0~3 minutes of patients with lung cancers, we model the descriptor by the observation that pulmonary uptake of different individuals exhibited various 'lung perfusion' patterns, which also represented glucose uptake of pulmonary tissue [166]. The trustworthiness of automated selected texture feature in descriptor, GLCM-Joint Entropy, could be further reinforced by its feasibility in previous texture analysis in lung tissue and soft tissue. For instance, the roughness of pulmonary texture exhibited a decreasing trend in entropy for normal lung, followed by embolism and then emphysema, indicating an association with the degree of perfusion[167], and GLCM also shown great performance in micro texture representation [168].

Secondly, the lobe's uptake phenotype might be stratified based on the proposed description into whole lung patterns and single lobe patterns, which may associate with the patient's pulmonary function. Three whole lung patterns exist in all five lobes and represented higher **D (r, K)** values than two single lobe patterns (table 2), which is also consistent with the visual illustration in figure 8 and figure 9. Similar to previous finding [166], the speed of perfusion may reveal patients' response to treatments. Patients with whole lung patterns shared analogous tumour shrinkage to varying degrees after chemotherapy and immunotherapy a few months after examination, while patients shared single lobe patterns showing no obvious changes in tumour size. One possible interpretation is that a faster perfusion pattern reflects better pulmonary vascular permeability, and the therapeutic result could be enhanced by increased capillary permeability.

Besides, compared with static SUVmean and other traditional PET metrics, the descriptor can significantly ($p<0.05$) distinguish visual perfusion abnormalities among different patients in a quantitative way. The voxel-wise imaging (figure 10) demonstrated descriptor's sensitivities in discriminating normal perfusion, bilateral lungs anomaly and unilateral lung anomaly. Recently published articles also pointed out that underlying perfusion heterogeneity may be illustrated by impaired hypoxic pulmonary vasoconstriction in infected lung regions [169]. Although computed tomography (CT) scans served as primary workhorse in pulmonary imaging [170] for its accurate assessment of morphological changes in the lung parenchyma, empowered by total-body coverage scanner, dynamic PET imaging achieve more than 40-fold gain in in effective sensitivity [171]. Our proposed descriptor explored additional capability in auxiliary diagnosis of dynamic whole-body PET in assessing patients' pulmonary diseases.

Further, parameters in descriptor are adjusted by utilizing time warping algorithm to eliminate the influence of gravity and other factors on perfusion. Passive mechanisms include vascular branching structure and the effect of gravity on ventilation and perfusion have effects on the distribution of pulmonary blood flow [161]. This mechanism causes two lobes with the same perfusion pattern to have different time-intensity curves of trajectories but similar curve shapes. Hausdorff distance [172] , Euclidean distance and other comparab methods are insensitive to temporal information. In contrast, we utilize time warping distance to address this issue by eliminating the influences of lag points on the curves, obtaining better discrimination between different modes of perfusion.

Our study has some limitations. First, only twenty patients were included in this study, and the results should be validated with a larger and multi-centre dataset. Secondly, the modelling of dynamics was not validated with kinetic characteristics. Third, the results were not validated in oncological patients.

4.6 Chapter Summary

We explore and build a pulmonary perfusion descriptor based on 18F-FDG whole-body dynamic PET images in the view of physiology and computer vision. Compared to traditional static radiomics features, the descriptor incorporating image's temporal information. Moreover, it can quantitatively and visually reflect pulmonary perfusion abnormality without the influence of passive mechanism. We suggest that the descriptor may serve as an auxiliary diagnostic tool for the comprehensive assessment of patients' pulmonary diseases.

CHAPTER 5. Conclusion and Future Work

5.1 Conclusion

Radiomics studies are widely used for biomarker mining in medical image analysis. Due to the multi-modality of medical data, there are two challenges in biomarker mining: one is dimensionality reduction of features when the number of features is significantly larger than the sample size, and the other is the integration of temporal information in dynamic medical images.

In this thesis, we proposed one results-driven feature selection framework for handling high-dimensional small-sample data. We applied this framework on two medical challenges when selecting representative radiomics features, including the preoperative prediction of Hepatocellular carcinoma (HCC) and Intrahepatic cholangiocarcinoma (ICC), and Microvascular Invasion (MVI) status identification for patients with primary liver cancer. We also build a feasible and robust machine learning model with selected radiomics features and clinical characteristics.

Besides, we explore and build a pulmonary perfusion descriptor for non-small-cell primary lung cancer (NSCLC) patients on 18F-FDG whole-body dynamic PET/CT images. Compared to traditional static radiomics features, the descriptor could reflect a dynamic computer-vision-based pattern within and between patients in a quantitative way. We also explored different slow/fast perfusion patterns by unsupervised machine learning algorithm. Furthermore, voxel-wise imaging by proposed descriptor in regional pulmonary perfusion visualization offers potentially application for anomaly detection.

5.2 Future Work

Future study should concentrate on improving the interpretability of retrieved space-temporal information and adding more complete information for accurate decision making. As a result, radiomics in the medical field might help physicians make judgments that are more accurate and reliable, which would also enhance patient care and the course of therapy.

Enhancement of radiomics research' generalizability and repeatability

The ineffectiveness of radiomic feature quantification quality control is a significant drawback of radiomics. Due to the sensitivity of radiomics characteristics to various picture modalities, normalization techniques, filter settings, and reconstruction techniques. The incorporation of fuzzy logics into the process of acquiring radiomic feature sets is another prospective future development in fuzzy radiomics, which aims to create a stable and repeatable radiomic feature combination.

More sophisticated tumor segmentation algorithms

VOIs were currently segregated manually, which took a lot of time and effort. Automatic segmentation may be strengthened and made more dependable by using a hybrid technique. By integrating the benefits of several segmentation techniques, many segmentation faults may be corrected. Defining the right normal boundary on normal scans may be quite challenging for radiologists. Due to the diverse looks of the tumor's perimeter, automated systems have an even more difficult time detecting it. By accepting that some scans may be too tough to automatically segment, error detection takes on a new significance. Indicating to the radiologist which scans need manual editing. The inclusion of this to current clinical segmentation techniques might be practically beneficial.

Involvement of deep learning in whole body dynamic radiomics analysis

The majority of early-learn descriptors were still created manually, and machine learning was only utilized to determine the ideal set of parameters. Recently, neural networks that can learn almost any descriptors, detect the probable response important to the final output, and display the complicated structures and information of high-dimensional spatial-temporal pictures were developed. Deep-learning-enhanced technology may improve the robustness and reproducibility of dynamic radiomics analysis.

Reference

1. Wei, J., et al., *Prediction of Microvascular Invasion in Hepatocellular Carcinoma via Deep Learning: A Multi-Center and Prospective Validation Study.* Cancers, 2021. **13**(10): p. 2368.
2. Maier-Hein, L., et al., *Surgical data science for next-generation interventions.* Nature Biomedical Engineering, 2017. **1**(9): p. 691-696.
3. Solomon, S.B. and S.G. Silverman, *Imaging in interventional oncology.* Radiology, 2010. **257**(3): p. 624-640.
4. Gupta, S., et al., *Quality improvement guidelines for percutaneous needle biopsy.* J Vasc Interv Radiol, 2010. **21**(7): p. 969-75.
5. Timsit, J.-F., et al., *Expert consensus-based clinical practice guidelines management of intravascular catheters in the intensive care unit.* Annals of Intensive Care, 2020. **10**(1): p. 118.
6. Lin, P.H., L.N. Ochoa, and P. Duffy, *Catheter-directed thrombectomy and thrombolysis for symptomatic lower-extremity deep vein thrombosis: review of current interventional treatment strategies.* Perspect Vasc Surg Endovasc Ther, 2010. **22**(3): p. 152-63.
7. Antunovic, L., et al., *PET/CT radiomics in breast cancer: promising tool for prediction of pathological response to neoadjuvant chemotherapy.* Eur J Nucl Med Mol Imaging, 2019. **46**(7): p. 1468-1477.
8. Ko, G.Y., K.B. Sung, and D.I. Gwon, *The Application of Interventional Radiology in Living-Donor Liver Transplantation.* Korean J Radiol, 2021. **22**(7): p. 1110-1123.
9. Asri, H., et al. *Big data in healthcare: challenges and opportunities*. in *2015 International Conference on Cloud Technologies and Applications (CloudTech)*. 2015. IEEE.
10. Verleysen, M. and D. François. *The curse of dimensionality in data mining and time series prediction*. in *International work-conference on artificial neural networks*. 2005. Springer.
11. Paldino, M.J. and D.P. Barboriak, *Fundamentals of quantitative dynamic contrast-enhanced MR imaging.* Magnetic resonance imaging clinics of North America, 2009. **17**(2): p. 277-289.
12. Dimitrakopoulou-Strauss, A., L. Pan, and C. Sachpekidis, *Kinetic modeling and parametric imaging with dynamic PET for oncological applications: general considerations, current clinical applications, and future perspectives.* European Journal of Nuclear Medicine and Molecular Imaging, 2021. **48**(1): p. 21-39.
13. Dimitrakopoulou-Strauss, A., L. Pan, and L.G. Strauss, *Quantitative approaches of dynamic FDG-PET and PET/CT studies (dPET/CT) for the evaluation of oncological patients.* Cancer Imaging, 2012. **12**(1): p. 283.
14. Yang, M., et al., *Treatment Response Prediction Using Ultrasound-Based Pre-, Post-Early, and Delta Radiomics in Neoadjuvant Chemotherapy in Breast Cancer.* Front Oncol, 2022. **12**: p. 748008.
15. Qu, H., et al., *Dynamic radiomics: A new methodology to extract quantitative time-related features from tomographic images.* Applied Intelligence, 2022. **52**(10): p. 11827-11845.

16. Wen, J., et al., *Determination of optimal 68 Ga-PSMA PET/CT imaging time in prostate cancers by total-body dynamic PET/CT.* European Journal of Nuclear Medicine and Molecular Imaging, 2022. **49**(6): p. 2086-2095.
17. Zwanenburg, A., et al., *The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping.* Radiology, 2020. **295**(2): p. 328-338.
18. Xu, X., et al., *Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma.* Journal of hepatology, 2019. **70**(6): p. 1133-1144.
19. Buckwalter, K.A., et al., *Musculoskeletal Imaging with Multislice CT.* American Journal of Roentgenology, 2001. **176**(4): p. 979-986.
20. Koretsky, A.P., *Early development of arterial spin labeling to measure regional brain blood flow by MRI.* Neuroimage, 2012. **62**(2): p. 602-7.
21. Chyou, J.Y. and L. Axel, *Considerations for Stress Perfusion Cardiac Magnetic Resonance Imaging in Patients with Cardiac Implantable Electronic Devices.* J Cardiovasc Electrophysiol, 2022.
22. Bailey, D.L., et al., *Positron emission tomography*. Vol. 2. 2005: Springer.
23. Laruelle, M., *Imaging synaptic neurotransmission with in vivo binding competition techniques: a critical review.* J Cereb Blood Flow Metab, 2000. **20**(3): p. 423-51.
24. Cherry, S.R., *In vivo molecular and genomic imaging: new challenges for imaging physics.* Phys Med Biol, 2004. **49**(3): p. R13-48.
25. Rahmim, A., et al., *Dynamic whole-body PET imaging: principles, potentials and applications.* European Journal of Nuclear Medicine and Molecular Imaging, 2019. **46**(2): p. 501-518.
26. Zhang, X., et al., *Subsecond total-body imaging using ultrasensitive positron emission tomography.* Proceedings of the National Academy of Sciences, 2020. **117**(5): p. 2265-2267.
27. Bilic, P., et al., *The liver tumor segmentation benchmark (lits).* arXiv preprint arXiv:1901.04056, 2019.
28. Ling, H., et al. *Hierarchical, learning-based automatic liver segmentation*. in *2008 IEEE Conference on Computer Vision and Pattern Recognition*. 2008. IEEE.
29. Okada, T., et al., *Automated segmentation of the liver from 3D CT images using probabilistic atlas and multilevel statistical shape model.* Academic radiology, 2008. **15**(11): p. 1390-1403.
30. Zheng, Y., et al., *Feature learning based random walk for liver segmentation.* PloS one, 2016. **11**(11): p. e0164098.
31. Fishwick, P.A. *Neural network models in simulation: A comparison with traditional modeling approaches*. in *Proceedings of the 21st conference on Winter simulation*. 1989.
32. Tajbakhsh, N., et al., *Convolutional neural networks for medical image analysis: Full training or fine tuning?* IEEE transactions on medical imaging, 2016. **35**(5): p. 1299-1312.
33. Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation*. in *International Conference on Medical image computing and computer-assisted intervention*. 2015. Springer.
34. Chen, K.-b., et al., *Lung computed tomography image segmentation based on U-Net network fused with dilated convolution.* Computer Methods and Programs in Biomedicine, 2021. **207**: p. 106170.

35. Scapicchio, C., et al., *A deep look into radiomics.* Radiol Med, 2021. **126**(10): p. 1296-1311.
36. van Griethuysen, J.J.M., et al., *Computational Radiomics System to Decode the Radiographic Phenotype.* Cancer Res, 2017. **77**(21): p. e104-e107.
37. Nanni, L., et al., *Different approaches for extracting information from the co-occurrence matrix.* PLoS One, 2013. **8**(12): p. e83554.
38. Stoecker, W.V., C.S. Chiang, and R.H. Moss, *Texture in skin images: comparison of three methods to determine smoothness.* Comput Med Imaging Graph, 1992. **16**(3): p. 179-90.
39. Galloway, M.M., *Texture analysis using grey level run lengths.* NASA STI/Recon Technical Report N, 1974. **75**: p. 18555.
40. Sripada, S.C. and M.S. Rao, *Comparison of purity and entropy of k-means clustering and fuzzy c means clustering.* Indian journal of computer science and engineering, 2011. **2**(03).
41. Thibault, G., et al., *Texture indexes and gray level size zone matrix.* Application to Cell Nuclei Classification. PRIP, 2009: p. 140-145.
42. Amadasun, M. and R. King, *Textural features corresponding to textural properties.* IEEE Transactions on systems, man, and Cybernetics, 1989. **19**(5): p. 1264-1274.
43. Rao, S.X., et al., *CT texture analysis in colorectal liver metastases: A better way than size and volume measurements to assess response to chemotherapy?* United European Gastroenterol J, 2016. **4**(2): p. 257-63.
44. Carvalho, S., et al., *Early variation of FDG-PET radiomics features in NSCLC is related to overall survival-the "delta radiomics" concept.* Radiother Oncol, 2016. **118**(suppl 1): p. S20-S21.
45. Corino, V.D.A., et al., *Refining Tumor Treatment in Sinonasal Cancer Using Delta Radiomics of Multi-Parametric MRI after the First Cycle of Induction Chemotherapy.* J Imaging, 2022. **8**(2).
46. Gong, J., et al., *A short-term follow-up CT based radiomics approach to predict response to immunotherapy in advanced non-small-cell lung cancer.* Oncoimmunology, 2022. **11**(1): p. 2028962.
47. Wang, Z., et al., *Unsupervised feature disentanglement for video retrieval in minimally invasive surgery.* Medical Image Analysis, 2022. **75**: p. 102296.
48. Primus, M.J., K. Schoeffmann, and L. Böszörmenyi. *Instrument classification in laparoscopic videos*. in *2015 13th international workshop on content-based multimedia indexing (CBMI)*. 2015. IEEE.
49. Nishigaki, M., C. Fermüller, and D. Dementhon. *The image torque operator: A new tool for mid-level vision*. in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2012. IEEE.
50. Shen, F., et al., *Unsupervised deep hashing with similarity-adaptive and discrete optimization.* IEEE transactions on pattern analysis and machine intelligence, 2018. **40**(12): p. 3034-3044.
51. Chittajallu, D.R., et al. *XAI-CBIR: Explainable AI system for content based retrieval of video frames from minimally invasive surgery videos*. in *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019)*. 2019. IEEE.
52. Päivärinta, J., E. Rahtu, and J. Heikkilä. *Volume local phase quantization for blur-insensitive dynamic texture classification*. in *Scandinavian conference on image analysis*. 2011. Springer.

53. Saisan, P., et al. *Dynamic texture recognition*. in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. 2001. IEEE.
54. Pauole, K., et al., *Reliability and validity of the T-test as a measure of agility, leg power, and leg speed in college-aged men and women.* The Journal of Strength & Conditioning Research, 2000. **14**(4): p. 443-450.
55. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome biology, 2014. **15**(12): p. 1-21.
56. Strens, M. *A Bayesian framework for reinforcement learning*. in *ICML*. 2000.
57. Kraskov, A., H. Stögbauer, and P. Grassberger, *Estimating mutual information.* Physical review E, 2004. **69**(6): p. 066138.
58. Azhagusundari, B. and A.S. Thanamani, *Feature selection based on information gain.* International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2013. **2**(2): p. 18-21.
59. Mishra, D. and B. Sahu, *Feature selection for cancer classification: a signal-to-noise ratio approach.* International Journal of Scientific & Engineering Research, 2011. **2**(4): p. 1-7.
60. Gao, Y. and H.-l. Wang. *A feature selection algorithm based on poisson estimates*. in *2009 Sixth international conference on fuzzy systems and knowledge discovery*. 2009. IEEE.
61. Auffarth, B., M. López, and J. Cerquides. *Comparison of redundancy and relevance measures for feature selection in tissue classification of CT images*. in *Industrial conference on data mining*. 2010. Springer.
62. Anderson, T.W., *RA Fisher and multivariate analysis.* Statistical Science, 1996: p. 20-34.
63. Hall, M.A., *Correlation-based feature selection for machine learning*. 1999, The University of Waikato.
64. Unler, A., A. Murat, and R.B. Chinnam, *mr2PSO: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification.* Information Sciences, 2011. **181**(20): p. 4625-4641.
65. Leardi, R., R. Boggia, and M. Terrile, *Genetic algorithms as a strategy for feature selection.* Journal of chemometrics, 1992. **6**(5): p. 267-281.
66. Tkatek, S., et al., *Artificial intelligence for improving the optimization of NP-hard problems: a review.* International Journal of Advanced Trends Computer Science and Applications, 2020. **9**(5).
67. Nakariyakul, S. and D.P. Casasent, *An improvement on floating search algorithms for feature subset selection.* Pattern Recognition, 2009. **42**(9): p. 1932-1940.
68. Emmanouilidis, C., et al., *A multi-objective genetic algorithm approach to feature selection in neural and fuzzy modeling.* Evolutionary Optimization, 2001. **3**(1): p. 1-26.
69. Mahafzah, B.A., *Performance evaluation of parallel multithreaded A* heuristic search algorithm.* journal of Information Science, 2014. **40**(3): p. 363-375.
70. Bradley, P.S. and O.L. Mangasarian. *Feature selection via concave minimization and support vector machines*. in *ICML*. 1998. Citeseer.
71. Kukreja, S.L., J. Löfberg, and M.J. Brenner, *A least absolute shrinkage and selection operator (LASSO) for nonlinear system identification.* IFAC proceedings volumes, 2006. **39**(1): p. 814-819.
72. Bi, J., et al., *Dimensionality reduction via sparse support vector machines.* Journal of Machine Learning Research, 2003. **3**(Mar): p. 1229-1243.

73. Tibshirani, R., *Regression shrinkage and selection via the lasso.* Journal of the Royal Statistical Society: Series B (Methodological), 1996. **58**(1): p. 267-288.
74. Li, X., et al., *Gene mining: a novel and powerful ensemble decision approach to hunting for disease genes using microarray expression profiling.* Nucleic Acids Research, 2004. **32**(9): p. 2685-2694.
75. Dutkowski, J. and A. Gambin, *On consensus biomarker selection.* BMC bioinformatics, 2007. **8**(5): p. 1-13.
76. Saeys, Y., T. Abeel, and Y.V.d. Peer. *Robust feature selection using ensemble feature selection techniques*. in *Joint European conference on machine learning and knowledge discovery in databases*. 2008. Springer.
77. Abeel, T., et al., *Robust biomarker identification for cancer diagnosis with ensemble feature selection methods.* Bioinformatics, 2010. **26**(3): p. 392-398.
78. Van Gassen, S., B. Callebaut, and Y. Saeys, *Using self-organizing maps for visualization and interpretation of cytometry data.* Cytometry A, 2014(87).
79. Galagan, J.E., et al., *The Mycobacterium tuberculosis regulatory network and hypoxia.* Nature, 2013. **499**(7457): p. 178-183.
80. Zhou, Z.-H., *Machine learning*. 2021: Springer Nature.
81. Erickson, B.J., et al., *Machine learning for medical imaging.* Radiographics, 2017. **37**(2): p. 505.
82. Langs, G., et al., *Machine learning: from radiomics to discovery and routine.* Der Radiologe, 2018. **58**(1): p. 1-6.
83. Havaei, M., et al., *Brain tumor segmentation with deep neural networks.* Medical image analysis, 2017. **35**: p. 18-31.
84. Teipel, S.J., et al., *Relevance of magnetic resonance imaging for early detection and diagnosis of Alzheimer disease.* Medical Clinics, 2013. **97**(3): p. 399-424.
85. Jiang, C., et al., *18F-FDG PET/CT radiomic analysis for classifying and predicting microvascular invasion in hepatocellular carcinoma and intrahepatic cholangiocarcinoma.* Quantitative imaging in medicine and surgery, 2022. **12**(8): p. 4135.
86. Bazazeh, D. and R. Shubair. *Comparative study of machine learning algorithms for breast cancer detection and diagnosis*. in *2016 5th international conference on electronic devices, systems and applications (ICEDSA)*. 2016. IEEE.
87. Moguerza, J.M. and A. Muñoz, *Support vector machines with applications.* Statistical Science, 2006. **21**(3): p. 322-336.
88. Peterson, L.E., *K-nearest neighbor.* Scholarpedia, 2009. **4**(2): p. 1883.
89. Liao, X., Y. Xue, and L. Carin. *Logistic regression with an auxiliary data source*. in *Proceedings of the 22nd international conference on Machine learning*. 2005.
90. Ketkar, N., *Stochastic gradient descent*, in *Deep learning with Python*. 2017, Springer. p. 113-132.
91. Pal, M. and P.M. Mather, *An assessment of the effectiveness of decision tree methods for land cover classification.* Remote sensing of environment, 2003. **86**(4): p. 554-565.
92. Petkovic, D., et al. *Improving the explainability of Random Forest classifier–user centered approach*. in *PACIFIC SYMPOSIUM ON BIOCOMPUTING 2018: Proceedings of the Pacific Symposium*. 2018. World Scientific.
93. Dong, X., et al., *A survey on ensemble learning.* Frontiers of Computer Science, 2020. **14**(2): p. 241-258.
94. Weisberg, S., *Applied linear regression*. Vol. 528. 2005: John Wiley & Sons.
95. Hinton, G. and T.J. Sejnowski, *Unsupervised learning: foundations of neural computation*. 1999: MIT press.

96. Agrawal, R., T. Imieliński, and A. Swami. *Mining association rules between sets of items in large databases*. in *Proceedings of the 1993 ACM SIGMOD international conference on Management of data*. 1993.
97. Long, B., P.S. Yu, and Z. Zhang. *A general model for multiple view unsupervised learning*. in *Proceedings of the 2008 SIAM international conference on data mining*. 2008. SIAM.
98. Pham, D.T., S.S. Dimov, and C.D. Nguyen, *Selection of K in K-means clustering.* Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2005. **219**(1): p. 103-119.
99. Murtagh, F. and P. Contreras, *Algorithms for hierarchical clustering: an overview.* Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2012. **2**(1): p. 86-97.
100. Petrovic, S. *A comparison between the silhouette index and the davies-bouldin index in labelling ids clusters*. in *Proceedings of the 11th Nordic workshop of secure IT systems*. 2006. Citeseer.
101. Visa, S., et al., *Confusion matrix-based feature selection.* MAICS, 2011. **710**(1): p. 120-127.
102. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2018. **68**(6): p. 394-424.
103. Rizvi, S. and G.J. Gores, *Emerging molecular therapeutic targets for cholangiocarcinoma.* J Hepatol, 2017. **67**(3): p. 632-644.
104. Mavros, M.N., et al., *Treatment and Prognosis for Patients With Intrahepatic Cholangiocarcinoma: Systematic Review and Meta-analysis.* JAMA Surg, 2014. **149**(6): p. 565-74.
105. Armengol, C., M.R. Sarrias, and M. Sala, *Hepatocellular carcinoma: Present and future.* Med Clin (Barc), 2018. **150**(10): p. 390-397.
106. Tabrizian, P., et al., *Recurrence of hepatocellular cancer after resection: patterns, treatments, and prognosis.* Ann Surg, 2015. **261**(5): p. 947-55.
107. Sumie, S., et al., *Microvascular invasion in patients with hepatocellular carcinoma and its predictable clinicopathological factors.* Ann Surg Oncol, 2008. **15**(5): p. 1375-82.
108. Rodriguez-Peralvarez, M., et al., *A systematic review of microvascular invasion in hepatocellular carcinoma: diagnostic and prognostic variability.* Ann Surg Oncol, 2013. **20**(1): p. 325-39.
109. Lim, K.C., et al., *Microvascular invasion is a better predictor of tumor recurrence and overall survival following surgical resection for hepatocellular carcinoma compared to the Milan criteria.* Ann Surg, 2011. **254**(1): p. 108-13.
110. Renzulli, M., et al., *Imaging features of microvascular invasion in hepatocellular carcinoma developed after direct-acting antiviral therapy in HCV-related cirrhosis.* Eur Radiol, 2018. **28**(2): p. 506-513.
111. Tang, Z., et al., *Prognostic Value and Predication Model of Microvascular Invasion in Patients with Intrahepatic Cholangiocarcinoma.* J Cancer, 2019. **10**(22): p. 5575-5584.
112. Hiley, C., et al., *Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine.* Genome Biol, 2014. **15**(8): p. 453.
113. Lu, L.C., et al., *Tumor Heterogeneity in Hepatocellular Carcinoma: Facing the Challenges.* Liver Cancer, 2016. **5**(2): p. 128-38.
114. Sherman, M. and J. Bruix, *Biopsy for liver cancer: how to balance research needs with evidence-based clinical practice.* Hepatology, 2015. **61**(2): p. 433-6.

115. Aerts, H.J., *The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review.* JAMA Oncol, 2016. **2**(12): p. 1636-1642.
116. Larue, R.T., et al., *Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures.* Br J Radiol, 2017. **90**(1070): p. 20160665.
117. Ochiai, T., et al., *Clinicopathologic features and risk factors for extrahepatic recurrences of hepatocellular carcinoma after curative resection.* World J Surg, 2012. **36**(1): p. 136-43.
118. Hatt, M., et al., *Radiomics in PET/CT: More Than Meets the Eye?* J Nucl Med, 2017. **58**(3): p. 365-366.
119. Liu, Q., et al., *Predicting EGFR mutation subtypes in lung adenocarcinoma using (18)F-FDG PET/CT radiomic features.* Transl Lung Cancer Res, 2020. **9**(3): p. 549-562.
120. Li, P., et al., *(18)F-FDG PET/CT radiomic predictors of pathologic complete response (pCR) to neoadjuvant chemotherapy in breast cancer patients.* Eur J Nucl Med Mol Imaging, 2020. **47**(5): p. 1116-1126.
121. Kirienko, M., et al., *Prediction of disease-free survival by the PET/CT radiomic signature in non-small cell lung cancer patients undergoing surgery.* Eur J Nucl Med Mol Imaging, 2018. **45**(2): p. 207-217.
122. Chen, L.-D., et al., *Differentiation of intrahepatic cholangiocarcinoma from hepatocellular carcinoma in high-risk patients: A predictive model using contrast-enhanced ultrasound.* World journal of gastroenterology, 2018. **24**(33): p. 3786-3798.
123. Qian, X., et al., *A Multi-Parametric Radiomics Nomogram for Preoperative Prediction of Microvascular Invasion Status in Intrahepatic Cholangiocarcinoma.* Front Oncol, 2022. **12**: p. 838701.
124. Nebbia, G., et al., *Pre-operative Microvascular Invasion Prediction Using Multi-parametric Liver MRI Radiomics.* Journal of digital imaging, 2020. **33**(6): p. 1376-1386.
125. Na, S.J., et al., *(18)F-FDG PET/CT Can Predict Survival of Advanced Hepatocellular Carcinoma Patients: A Multicenter Retrospective Cohort Study.* J Nucl Med, 2017. **58**(5): p. 730-736.
126. Cassim, S., et al., *Metabolic reprogramming enables hepatocarcinoma cells to efficiently adapt and survive to a nutrient-restricted microenvironment.* Cell Cycle, 2018. **17**(7): p. 903-916.
127. Findlay, J.M., et al., *Routinely staging gastric cancer with (18)F-FDG PET-CT detects additional metastases and predicts early recurrence and death after surgery.* Eur Radiol, 2019. **29**(5): p. 2490-2498.
128. Linagjia Zhu, I.K., Yi Gao,Ron Kikinis,Allen Tannenbaum, *An effective interactive medical image segmentation method using fast grow cut.* MICCAI workshop on interactive medical image computing, 2014.
129. Li, C., et al., *A likelihood and local constraint level set model for liver tumor segmentation from CT volumes.* IEEE Trans Biomed Eng, 2013. **60**(10): p. 2967-77.
130. Van Griethuysen, J.J., et al., *Computational radiomics system to decode the radiographic phenotype.* Cancer research, 2017. **77**(21): p. e104-e107.
131. Ferri, F.J., et al., *Comparative study of techniques for large-scale feature selection*, in *Machine Intelligence and Pattern Recognition*. 1994, Elsevier. p. 403-413.
132. Haralick, R.M. and L.G. Shapiro, *Computer and robot vision*. Vol. 1. 1992: Addison-wesley Reading.
133. Pudil, P., J. Novovicová, and J. Kittler, *Floating search methods in feature selection.* Pattern Recognit. Lett., 1994. **15**: p. 1119-1125.

134. Boellaard, R., et al., *FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0.* Eur J Nucl Med Mol Imaging, 2015. **42**(2): p. 328-54.
135. Zhou, M., et al., *Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017.* Lancet, 2019. **394**(10204): p. 1145-1158.
136. Raman, S.P., et al., *Preliminary data using computed tomography texture analysis for the classification of hypervascular liver lesions: generation of a predictive model on the basis of quantitative spatial frequency measurements--a work in progress.* J Comput Assist Tomogr, 2015. **39**(3): p. 383-95.
137. Wu, M., et al., *Predicting the grade of hepatocellular carcinoma based on non-contrast-enhanced MRI radiomics signature.* Eur Radiol, 2019. **29**(6): p. 2802-2811.
138. Banerjee, S., et al., *A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma.* Hepatology, 2015. **62**(3): p. 792-800.
139. Eguchi, S., et al., *Predictor for histological microvascular invasion of hepatocellular carcinoma: a lesson from 229 consecutive cases of curative liver resection.* World J Surg, 2010. **34**(5): p. 1034-8.
140. Imamura, H., et al., *Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy.* J Hepatol, 2003. **38**(2): p. 200-7.
141. Ma, X., et al., *Preoperative radiomics nomogram for microvascular invasion prediction in hepatocellular carcinoma using contrast-enhanced CT.* Eur Radiol, 2019. **29**(7): p. 3595-3605.
142. Peng, J., et al., *A radiomics nomogram for preoperative prediction of microvascular invasion risk in hepatitis B virus-related hepatocellular carcinoma.* Diagn Interv Radiol, 2018. **24**(3): p. 121-127.
143. Liu, J., et al., *Development and validation of qualitative and quantitative models to predict invasiveness of lung adenocarcinomas manifesting as pure ground-glass nodules based on low-dose computed tomography during lung cancer screening.* Quantitative imaging in medicine and surgery, 2022. **12**(5): p. 2917-2931.
144. Gao, Y., et al., *Reproducibility of radiomic features of pulmonary nodules between low-dose CT and conventional-dose CT.* Quantitative imaging in medicine and surgery, 2022. **12**(4): p. 2368-2377.
145. Yoon, J.H., et al., *A Comparison of Biannual Two-Phase Low-Dose Liver CT and US for HCC Surveillance in a Group at High Risk of HCC Development.* Liver cancer, 2020. **9**(5): p. 503-517.
146. Ganeshan, B., et al., *CT-based texture analysis potentially provides prognostic information complementary to interim fdg-pet for patients with hodgkin's and aggressive non-hodgkin's lymphomas.* European radiology, 2017. **27**(3): p. 1012-1020.
147. Ganeshan, B., et al., *Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival.* Clin Radiol, 2012. **67**(2): p. 157-64.
148. Woolson, R.F., J.A. Bean, and P.B. Rojas, *Sample size for case-control studies using Cochran's statistic.* Biometrics, 1986. **42**(4): p. 927-32.
149. Chalkidou, A., M.J. O'Doherty, and P.K. Marsden, *False Discovery Rates in PET and CT Studies with Texture Features: A Systematic Review.* PLoS One, 2015. **10**(5): p. e0124165.
150. Jang, Y.M., et al., *Quantitatively assessed dynamic contrast-enhanced magnetic resonance imaging in patients with chronic obstructive pulmonary disease: correlation*

of perfusion parameters with pulmonary function test and quantitative computed tomography. Invest Radiol, 2008. **43**(6): p. 403-10.
151. Wu, Y., et al., *Whole-Body Parametric Imaging of (18)F-FDG PET Using uEXPLORER with Reduced Scanning Time.* J Nucl Med, 2022. **63**(4): p. 622-628.
152. Zhang, X., et al., *Total-Body Dynamic Reconstruction and Parametric Imaging on the uEXPLORER.* J Nucl Med, 2020. **61**(2): p. 285-291.
153. Cherry, S.R., et al., *Total-body imaging: Transforming the role of positron emission tomography.* Sci Transl Med, 2017. **9**(381).
154. Surti, S. and J.S. Karp, *Impact of detector design on imaging performance of a long axial field-of-view, whole-body PET scanner.* Phys Med Biol, 2015. **60**(13): p. 5343-58.
155. Vandenberghe, S., P. Moskal, and J.S. Karp, *State of the art in total body PET.* EJNMMI Phys, 2020. **7**(1): p. 35.
156. Xing, N., et al., *The use of CT perfusion to determine microvessel density in lung cancer: comparison with FDG-PET and pathology.* Chinese Journal of Cancer Research, 2011. **23**(2): p. 118-122.
157. Haralick, R.M., K. Shanmugam, and I.H. Dinstein, *Textural features for image classification.* IEEE Transactions on systems, man, and cybernetics, 1973(6): p. 610-621.
158. Ganeshan, B., et al., *Three-dimensional selective-scale texture analysis of computed tomography pulmonary angiograms.* Investigative radiology, 2008. **43**(6): p. 382-394.
159. Hofmanninger, J., et al., *Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem.* European Radiology Experimental, 2020. **4**(1): p. 1-13.
160. Hesterberg, T., *Bootstrap.* Wiley Interdisciplinary Reviews: Computational Statistics, 2011. **3**(6): p. 497-526.
161. Henderson, A.C., et al., *The gravitational distribution of ventilation-perfusion ratio is more uniform in prone than supine posture in the normal human lung.* J Appl Physiol (1985), 2013. **115**(3): p. 313-24.
162. Chen, Y., et al., *Delineating urban functional areas with building-level social media data: A dynamic time warping (DTW) distance based k-medoids method.* Landscape and Urban Planning, 2017. **160**: p. 48-60.
163. Bernatowicz, K., et al., *Robust imaging habitat computation using voxel-wise radiomics features.* Scientific Reports, 2021. **11**(1): p. 20133.
164. Zhu, Y., D. Tian, and F. Yan, *Effectiveness of entropy weight method in decision-making.* Mathematical Problems in Engineering, 2020. **2020**.
165. Liu, G., et al., *Ultra-low-activity total-body dynamic PET imaging allows equal performance to full-activity PET imaging for investigating kinetic metrics of (18)F-FDG in healthy volunteers.* Eur J Nucl Med Mol Imaging, 2021. **48**(8): p. 2373-2383.
166. Wang, D., et al., *Assessing dynamic metabolic heterogeneity in non-small cell lung cancer patients via ultra-high sensitivity total-body [(18)F]FDG PET/CT imaging: quantitative analysis of [(18)F]FDG uptake in primary tumors and metastatic lymph nodes.* Eur J Nucl Med Mol Imaging, 2022.
167. Ganeshan, B., et al., *Three-dimensional selective-scale texture analysis of computed tomography pulmonary angiograms.* Invest Radiol, 2008. **43**(6): p. 382-94.
168. Roy, S.D., et al., *Computer Aided Breast Cancer Detection Using Ensembling of Texture and Statistical Image Features.* Sensors (Basel), 2021. **21**(11).
169. Herrmann, J., et al., *Modeling lung perfusion abnormalities to explain early COVID-19 hypoxemia.* Nat Commun, 2020. **11**(1): p. 4883.

170. Ley, S., *Lung imaging.* Eur Respir Rev, 2015. **24**(136): p. 240-5.
171. Zhang, X., et al., *Theoretical study of the benefit of long axial field-of-view PET on region of interest quantification.* Physics in Medicine & Biology, 2018. **63**(13): p. 135010.
172. Huttenlocher, D.P., G.A. Klanderman, and W.A. Rucklidge, *Comparing Images Using the Hausdorff Distance.* IEEE Trans. Pattern Anal. Mach. Intell., 1993. **15**(9): p. 850–863.

Glossary

Alpha-Fetoprotein
AFP, 56
American Association for the Study of Liver Diseases (AASLD)
AASLD, 50
area of interests
ROIs, 27
Area Under the Curve
AUC, 46
artificial intelligence
AI, 3
axial fields of view
AFOV, 24
Barcelona Clinic Liver Cancer
BCLC, 50
blood flow
BF, 70
Carbohydrate Antigen 19-9
CA1-99, 56
Computerized Tomography
CT, 22
Contrast-enhanced ultrasound
CEUS, 48
Convolutional Neural Networks
CNN, 26
Correlation-based Feature Selection
CFS, 38
Decision Tree Classifier
DT, 43
dynamic contrast-enhanced magnetic resonance
DCE/MR, 16
Dynamic whole body Positron Emission Tomography
dPET, 24
false negative
FN, 45
false positive
FP, 45
Genetic Algorithm
GA, 39
Graph Cut
GC, 26
Gray Level Dependence Matrix
GLDM, 29
Gray Level Run Length Matrix
GLRLM, 29
Gray Level Size Zone Matrix
GLSZM, 29
hepatocellular carcinoma
HCC, 4
International Conference on Harmonization in Good Clinical Practical
ICH-GCP, 49
Interventional Radiology
IR, 3
intrahepatic cholangiocarcinoma
ICC, 4
K Nearest Neighbors
KNN, 43
Laplacian of Gaussian
LoG, 32
Least Absolute Shrinkage Selection Operator
LASSO, 40
left upper lobe slow pattern
LUSP, 82
linear dynamical system
LDS, 35
Local Binary Pattern
LBP, 34
Local Phase Quantization on Three Orthogonal Planes
LPQ-TOP, 35
Magnetic Resonance Imaging
MRI, 22
Maximum Relevance Minimum Redundancy
MRMR, 38
Microvascular Invasion
MVI, 21
negative predictive value
NPV, 55
non-small cell lung cancer
NSCLC, 33
permeable surface area product
PS, 70
portal vein tumor thrombus
PVTT, 47