



THE UNIVERSITY OF
SYDNEY

Deep Domain Adaptation Learning Framework for Associating Image Features to Tumour Gene Profile

Tian Xia

SID: 311150683

Supervisor: Jinman Kim
Associate Supervisor: Ashnil Kumar

This thesis is submitted in partial fulfilment of
the requirements for the degree of
Master of Philosophy (Engineering and IT)

School of Information Technologies
The University of Sydney
Australia

12th/July/2018

Student Plagiarism: Compliance Statement

I certify that:

I have read and understood the University of Sydney Student Plagiarism: Coursework Policy and Procedure;

I understand that failure to comply with the Student Plagiarism: Academic Board Policy: Academic Dishonesty and Plagiarism can lead to the University commencing proceedings against me for potential student misconduct under the [2012 Academic Dishonesty and Plagiarism in Coursework Policy](#).

This Work is substantially my own, and to the extent that any part of this Work is not my own I have indicated that it is not my own by Acknowledging the Source of that part or those parts of the Work.

Name: Tian Xia

Signature:

Date: 12th/July/2018

Abstract

Establishment of accurate cancer diagnosis involves a series of complex medical procedures and is rarely straightforward. These diagnoses are reliant upon the results of multiple medical examinations comprising of non-invasive imaging for disease localisation, histopathology analysis of tissue specimens and genetic sequencing to differentiate mutation subtypes. Although medical imaging and general pathology are becoming more routine, genetic sequencing is not always practical as human cancers exhibit strong phenotypic and genetic heterogeneity, which requires multiple site sampling in addition to its high cost and accessibility.

Medical imaging provides non-invasive assessments of phenotypic diversities in a number of cancer types. Medical images have the potential to offer insights into the patterns of disease spread and treatment response. Furthermore, advances in computerised medical image analysis enable the extraction of tumours' visual representation (features) to facilitate characterising tumour phenotypes. The association between tumour imaging biomarkers and existing genetic foreknowledge lead to the emergence of image-genomics.

Image-genomics leverages medical image analysis and genetic information to provide a complementary approach that contributes to a non-invasive and accurate cancer diagnosis. This is achieved through associating tumour imaging traits with clinical data (e.g., the underlying tumour genetic information), where recent image-genomics research has demonstrated its potential to identify imaging surrogates for tumour biomarkers.

However, the majority of image-genomics research has focused on the employment of tumour visual traits that are derived from high-level image object descriptors according to human understanding. In particular, image features that were designed to represent a range of image objects may not be ideal to describe tumour phenotypes. The challenge of image-genomics hence lies in the extraction of imaging representations to quantify tumour phenotypic characteristics. A secondary challenge lies within the image feature extraction technique, as it requires a large amount of disease-specific medical image datasets; such datasets are difficult to acquire, in addition to the need for precise annotations.

In this thesis we propose a deep domain adaptation learning framework for associating image features to tumour genetic information. Our approach exploits the potential of domain adaptation technique for image-genomics to quantify image features based on similar knowledge domains. This is accomplished by facilitating the learning of tumour image representations with larger datasets from similar domains to reduce the reliance on large volumes of disease-specific datasets for image-genomics research.

In addition, our proposed framework enables the extraction of additional tumour visual descriptors to provide abstract image representations for associating

with gene expressions. It leverages the current state-of-the-art in image object recognition to provide image features which encode subtle variations of tumour phenotypic characteristics. The quantification of such features is facilitated by the employment of domain adaptation techniques.

We evaluated our proposed deep domain adaptation learning framework by comparing with current state-of-the-art in: (i) tumour histopathology image classification and; (ii) the degree of image-genomics associations, compare with human-crafted tumour image descriptors.

Our results demonstrated that the proposed deep domain adaptation learning framework facilitates the learning of image descriptors and improves the accuracy of tumour classification. Our framework has also shown to provide additional image features that encode abstract representations of tumour phenotypic characteristics which exhibit stronger associations to patient-specific genetic information, compared to current state-of-the-art image-genomics analysis. This thesis advances the state-of-the-art in image-genomics by proposing a novel deep domain adaptation learning framework to offer additional image features for image-genomics analysis. The work in this thesis indicates the potential of image-genomics research to reveal additional imaging surrogates to genetic biomarker, which can be used to facilitate cancer diagnosis.

Acknowledgement

It's time to write another acknowledgement! Different to the honours thesis, I had less mental pressure this time. For this, I would like to express my sincerest appreciation towards my supervisor, Associate Professor Jinman Kim for the mental support and exceptional supervision for the duration of the candidature. I'd say Master of Philosophy is a little bit tougher than I thought, and I credit this to my glorious, supreme leader Kim Jinman as well.

I also would like to show my gratitude to my co-supervisor Ashnil Kumar, thanks for the caring and support in every way possible. I'll try my best to produce a nature comm paper during my PhD to repay you!

To all of my friends and mentors from the Biomedical & Multimedia Information Technology (BMIT) research group, I am grateful for the company, especially to Hoijoon Jung, Euijoon Ahn, Dr Younhyun Jung, Dr Lei Bi and the captain of team slack off: Dr Michael de Ridder.

The special place for special people: Jingshu Zhao I love you! And my parents: I love you too!

In the near future I'll get to write another thesis for the PhD. This thesis is surely the beginning of something greater!

Contents

ABSTRACT.....	III
ACKNOWLEDGEMENT	VI
CONTENTS.....	VII
LIST OF FIGURES	XI
LIST OF TABLES	XIII
THESIS PUBLICATIONS	XIV
CHAPTER 1	1
INTRODUCTION	1
1.1 MOTIVATION.....	3
1.2 AIMS AND OBJECTIVES.....	6
1.3 CONTRIBUTIONS OF THIS THESIS.....	7
IMAGE CLASSIFICATION AND DOMAIN ADAPTATION.....	7
IMAGE-GENOMICS FRAMEWORK	7

1.4 THESIS STRUCTURE	8
CHAPTER 2	9
BACKGROUND AND RELATED CONCEPTS	9
2.1 MEDICAL IMAGING AND IMAGE PROCESSING.....	9
2.1.1 DEFINITIONS.....	9
2.1.2 MEDICAL IMAGING MODALITIES	11
2.1.3 MEDICAL IMAGE SEGMENTATION	14
2.2 GENE EXPRESSION PROFILING	16
2.3 MACHINE LEARNING, DEEP LEARNING AND CONVOLUTIONAL NEURAL NETWORKS.....	17
2.3.1 MACHINE LEARNING.....	18
2.3.2 ARTIFICIAL NEURAL NETWORK.....	19
2.3.3 DEEP LEARNING	21
2.3.4 CONVOLUTIONAL NEURAL NETWORKS	22
2.4 DOMAIN ADAPTATION FOR TRANSFER LEARNING	28
CHAPTER 3	32
IMAGE-GENOMICS	32
3.1 OVERVIEW OF IMAGE-GENOMICS.....	32
3.2 IMAGE FEATURE EXTRACTION.....	36
3.3 SUMMARY OF GAPS	39
3.3.1 MEDICAL IMAGE FEATURE EXTRACTION	40
3.3.2 DOMAIN ADAPTATION	40
CHAPTER 4	42

DOMAIN ADAPTATION FACILITATED MEDICAL IMAGE ANALYSIS	42
4.1 OVERVIEW	42
4.1.1 FRAMEWORK OVERVIEW	44
4.2 DATASETS.....	46
4.3 IMAGE PRE-PROCESSING AND PATCH EXTRACTION.....	46
4.4 CNN TRAINING AND FINE-TUNING	48
4.5 EVALUATION STRATEGY.....	49
4.6 RESULTS	50
4.7 DISCUSSIONS AND SUMMARY.....	52
 CHAPTER 5	 54
DEEP IMAGE-GENOMICS	54
5.1 OVERVIEW	55
5.2 FRAMEWORK	55
5.2.1 DATASET	56
5.2.2 TUMOUR IMAGE FEATURE EXTRACTION	58
5.3.3 IMAGE-GENOMIC ASSOCIATION AND ENRICHMENT ANALYSIS	60
5.3.4 EVALUATION STRATEGY.....	61
5.3 RESULTS	62
5.3.1 DEEP FEATURES AS TUMOUR PHENOTYPE DESCRIPTOR	62
5.3.2 DEEP FEATURES DERIVED IMAGE-GENOMICS SIGNATURES	63
5.4 DISCUSSIONS AND SUMMARY.....	65
 CHAPTER 6	 71
CONCLUSION AND FUTURE WORK	71
6.1 CONCLUSION.....	71

6.2 FUTURE RESEARCH DIRECTIONS	73
REFERENCES.....	74
APPENDIX.....	86
HUMAN-CRAFTED IMAGE FEATURES.....	87
FIRST ORDER STATISTICS	87
SHAPE AND SIZE BASED FEATURES	89
TEXTURAL FEATURES	91
WAVELET FEATURES: WAVELET DECOMPOSITIONS.....	97

List of Figures

2.1: Pixels, voxels and spatial resolutions of medical images.....	10
2.2: The CT image of the lung field from a NSCLC patient.....	12
2.3: The overview and zoom-in view of WSI of sentinel lymph node with presence of breast cancer metastasis.....	13
2.4: The segmented region of interests in different medical image modalities.....	14
2.5: Representation of a single model neuron.....	20
2.6: An example 2D convolution.....	23
2.7: An example CNN with many convolution layers.....	24
2.8: The GoogLeNet architecture.....	26
2.9: The building block of ResNet.....	28
2.10: Traditional and transfer learning-based approaches for CNN models.....	29
2.11: Overview of shallow tuning-based transfer learning approach.....	30
3.1: The current image-genomic approach for associating image features with gene expression profiles.....	35
4.1: An overview of the proposed framework for patch-level tumour detection.....	45
4.2: Histopathology slides with background information removed.....	47

4.3: Accuracy of CNNs with various prostate cancer training size from fine-tuned and scratch-trained strategies.....	50
5.1: The proposed deep image-genomics framework.....	57
5.2: The 2.5D representation of a tumour physical centroid with axial, sagittal and coronal slices of the image volume.....	60
5.3: Association heat map between deep learning derived image features and tumour phenotype.....	69
5.4: Image-genomics association between deep image features and oncology gene sets using gene-set enrichment analysis (GSEA).....	70

List of Tables

4.1: Test accuracy of CNNs from both training approaches for cancer classification task with different training size.....	51
4.2: Area under receiver operation characteristics (AUC)	51
5.1: Statistical analysis of radiogenomic signatures from human-crafted and deep learning derived image features.....	64
5.2: Right-tailed two-sample t-test result summarisation of deep-learnt and human-crafted features.....	65
A1. Patients' characteristics of the NSCLC data set.....	73

Thesis Publications

The following publications have been produced over the course of the candidature:

Accepted paper:

- T. Xia, A. Kumar, D. Feng, J. Kim, “Patch-level tumour classification in digital histopathology images with domain adapted deep learning,” submitted to *the 40th International Conference of the IEEE Engineering in Medicine and Biology Society*, accepted, submitted in February, 2018

Paper under preparation

- T. Xia, A. Kumar, MJ. Fulham, D. Feng, J. Kim, “Deep radiogenomics: Quantifying Associations between Deep Learning-Based Medical Image Features with Tumour Gene Expression.” For submission to *Nature Scientific Reports*.

Chapter 1

Introduction

The current practices for the treatment of human cancers rely upon the establishment of accurate diagnosis, which may involve a series of complex medical procedures comprising of patient screening for symptom evaluation, non-invasive imaging for disease localisation and histopathology analysis of tissue specimens. In recent years, genetic sequencing is becoming an increasingly important addition to the existing diagnostic pipeline. Cancer diagnosis may be the result of analysis of data generated through each medical examinations; the integration of such data contributes towards the understanding of the disease and hence exhibit potentials to offer optimal cancer therapy at individual patient level [1].

Although medical imaging and general pathology are becoming more routine in current cancer diagnosis, genetic sequencing is not always practical. This is due to human cancers exhibits strong phenotypic and genetic heterogeneity, where disease develops at multiple sites with genetic differences [2]. It is not practical or feasible to obtain tissue samples from all sites of disease as invasive biopsies have the potential

to induce tumour haemorrhage/proliferation. Further, the cost of genetic examination is high in addition to its limited accessibility.

As an alternative, medical images provide non-invasive assessments of phenotypic diversities in a number of cancer types and contribute to clinical decision-making, e.g., tumour detection, subtype characterisation and treatment responses [3, 4]. Advances in computerised medical image analysis allow the extraction of tumour's visual representations (features) to facilitate the characterisation of tumour phenotypic differences, which offers insights into the patterns of disease spread and prognosis. The association between tumour imaging features and genetic foreknowledge allows the derivation of imaging surrogates to genetic biomarkers, and ultimately lead to the emerging field of image-genomics [5].

Image-genomics aims to associate the tumour imaging trait and clinical data (e.g., the underlying tumour gene expression) to provide an alternative approach that contributes to a non-invasive and accurate cancer diagnosis [6]. As such, the discovery and extraction of optimised tumour imaging descriptors represent a major challenge in the current image-genomics research.

This thesis address this challenge in the domain of medical image analysis, where our approach offers additional image features with domain adaptation technique to encode extra tumour phenotype representation. This thesis describes research to address the two key hypotheses: *can domain adaptation facilitate medical image analysis and processing applications?* If so, *is domain adaptation capable of deriving tumour visual descriptors that produce stronger associations with genetic information?*

1.1 Motivation

Precision medicine refers to the approach of disease prevention, treatment and care that takes into account individual variability in genes, environment, and lifestyle [1]. Precision medicine allows the classification of individuals into cohorts who are susceptible to particular diseases and forecasts response to a specific treatment, thereby improving disease prognosis [7]. The combined knowledge of patient's genetics, bioinformatics, imaging and proteomics ultimately contribute towards the optimisation of therapeutic strategies [8]. One of the near-term focus of precision medicine is with cancer, a disease that is among the leading causes of death incidence worldwide. Traditional cancer therapeutics involve the utilisation of existing knowledge of the clinical outcomes from previous patient populations to draft treatment plans for future patients. Precision medicine has the potential to improve clinical oncology practices by utilising the molecular profile of cancer genetic behaviour, together with patient's biological information, to optimise treatment plans at individual patient level [9].

However, pathways to cancer diagnosis are complicated and rarely straightforward. Cancer diagnosis may be the result of analysis of data generated by a series of medical procedures including non-invasive medical imaging, invasive tumour biopsies, genetic sequencing or from other evaluation of symptoms [10, 11]. Such procedures are prone to diagnostic errors, time-consuming and expensive in labour. In particular, obtaining tissue samples are not always feasible or practical as human cancers display phenotypic and genetic heterogeneity with multiple sites and genetic differences of the disease.

Medical imaging is a fundamental aspect of modern healthcare and plays an important role in cancer diagnosis and post-treatment monitoring. Medical images encode visual features which represent cancer phenotypic characteristics such as cancer location, size, texture and shape, which are useful for disease diagnosis and oncologic research [12]. Furthermore, the growing quantity and quality of medical images allow the creation of big data-driven approaches to computer-aided diagnosis (CAD, where clinicians routinely utilise CAD tools such as segmentation and registration to aid in their interpretation and decision making, as well as automated detection and classification of diseases as the second opinion to improve diagnostic accuracy [13, 14].

Although medical imaging is now a routine for cancer diagnosis, patient-specific genetic information can be valuable for its ability to reveal the underlying tumour biology and mutation types [15, 16]. Unfortunately, most human cancers exhibit strong genotypic heterogeneity, implying that tumours may have different gene expression at various sites and thus limits the usage of surgical operations such as tumour biopsies. In addition, tumour heterogeneity exhibits the potential requirement for multiple tumour biopsies during the diagnosis, the process of which further escalate the expenses along with patients discomfort, and risks the occurrence of undesired side effects such as tumour haemorrhage [17-19].

Imaging surrogates that reflect tumour genetic information have been explored as an alternative approach to invasive biopsies for genetic sequencing. This is achieved by associating tumour imaging descriptor with the tumours' underlying genetic information which exhibits strong potential to the discovery of tumour biomarkers, and has been a key factor for the emergence of "image-genomics" domain [20, 21].

Image-genomics aims to convert tumour imaging data into high-dimensional image features to derive associations with tumour genetic information through the use of quantitative feature extraction techniques [20].

In the medical domain, image-genomics has potential applications in medical image analysis and biomedical research. Recent studies have demonstrated the capability of associating tumour image features with genetic information [22]; and the feasibility for potential non-invasive imaging surrogates to specific tumour prognosis biomarkers through the use of image-genomics [23]. Image-genomics approaches have also been shown to associate image features to a number of biological molecules and processes such as the transcriptome [24]. Furthermore, studies suggested that image-genomics associations are capable of predicting clinical outcomes in hepatocellular carcinoma [25, 26].

Despite the advances in image-genomics for identifying imaging surrogate to genetic biomarkers, there is an inherent reliance on tumour imaging representations for use in the association. The current state-of-the-art image-genomics approach employs image features that are defined based upon the human understanding of a range of objects (we denote this as human-crafted image features). The direct employment of such image object descriptors (by extracting human-crafted image features from medical images) is restricted by its capacity to fully quantify tumour imaging descriptors and therefore unable to reflect specific cancer types, nor variants among the patient cohorts. An alternative approach is to extract additional image representations that describe the subtle variations in tumour images, which are potentially important to specific tumour types [27]. The quantification of such abstract image features offer complementary information to existing image-genomics approach

and have the opportunities to: (i) derive stronger associations to gene expressions and; (ii) explore and discover additional surrogate non-invasive imaging biomarkers.

Another challenge that exists in the field of image-genomics is the lack of annotated medical images for rare diseases or different image modalities, where large volumes of annotated datasets are required for the learning of image representations of particular diseases [22, 28]. The demand for such datasets limits the image-genomic applications to specific disease types or image modalities. This restricts the ability to derive associations between image features with tumour underlying gene expressions.

1.2 Aims and Objectives

In this thesis, we address the identified challenges through: (i) the extraction of additional image representation to specific cancer types and; (ii) overcome the reliance of large volumes of annotated datasets with a novel image-genomics framework. Our framework leverages deep learning image analysis and domain adaptation techniques to extract disease-specific abstract image representations for associating tumour gene expressions.

Our framework expands upon traditional approaches in medical image feature extraction by adapting the state-of-the-art deep learning techniques for image object recognition, and hence offering additional tumour image descriptors for genetic associations. The thesis address the following two aims:

1. The development of a medical image analysis scheme with domain adaptation approach. The proposed scheme emphasis on the learning of tumour image

representations which leads to the enhanced performance of medical image analysis applications, e.g., tumour image classification.

2. The derivation of an image feature extraction scheme based on deep learning techniques to quantify additional image representation from tumour image through the employment of domain adaptation technique to associate with tumour genetic information.

1.3 Contributions of this thesis

Our contributions can be summarised in the following two major areas:

Image Classification and Domain Adaptation

Chapter 4 presents our contribution towards the domain adaptation facilitated medical image analysis scheme with limited volumes of datasets. In our thesis, the tumour classification system employs the domain adaptation technique from different clinical histopathology images to facilitate the prediction of the presence of abnormalities. Our system overcomes the reliance of large volumes of annotated datasets and offers improved tumour classification accuracy through the utilisation of domain adaptation technique compared to the traditional approach.

Image-Genomics Framework

Detailed in Chapter 5, we present a new image-genomics framework to associate tumour imaging representations to the underlying tumour gene expressions. Our framework utilises the current state-of-the-art in image object recognition to learn the abstract image representations that describe the subtle variations in tumour

images through a domain adaptation facilitated data-driven approach. Our framework utilises different mathematical and genetic analysis tools to associate extracted image features to cellular pathways as well as cancer prognostic biomarkers.

1.4 Thesis Structure

The remainder of this thesis is structured as follows:

Chapter 2 and 3 orients the readers by providing theoretical background knowledge of the thesis. Chapter 2 provides an overview of medical imaging and the image processing techniques within this domain. Chapter 3 presents the overview of the state-of-the-art in the image-genomics analysis. In particular, it provides a summary of gaps from recent image-genomics research.

Chapter 4 and 5 provide the detailed contribution of this thesis. We describe the domain adaptation facilitated medical image analysis system in chapter 4. In chapter 5 we highlight the deep domain adaptation learning framework for associating image features to tumour genetic information.

Chapter 6 summarises the contribution of this thesis and indicates directions for future research in the domain of image-genomics analysis.

Chapter 2

Background and Related Concepts

2.1 Medical Imaging and Image Processing

Medical imaging is a fundamental component of the modern healthcare system and is essential for the accurate diagnosis and staging of cancers. Medical images provide a fast and non-invasive assessment of phenotypic diversity in many cancers [12] through a variety of different imaging modalities (or techniques). Medical image processing, in turn, allows the extraction of meaningful information which offers insights of different aspects of patients' conditions. In this chapter, we provide an overview of medical imaging and the theoretical background in image processing techniques that are crucial for cancer diagnosis.

2.1.1 Definitions

Digital images consist of a collection of numerical picture elements, called *pixels*. The term *pixel resolution* refers to the number of pixels in an image, which can be

represented as a single number or by the number of pixels in each dimension. For example, an image that contains 8,294,400 pixels can also be referred to as having a resolution of 8.3 megapixels or 3840×2160 (width \times height) pixels.

Many medical imaging techniques consist of sampling 2D images along a third spatial axis to form 3D images, known as *volumetric images* or *image volumes*. These 3D volumetric images encode the spatial relationships between 3D pixels, called *voxels*, which both 2D and 3D images allow the extraction and interpretation of the encoded information through the use of image processing techniques. Both pixels and

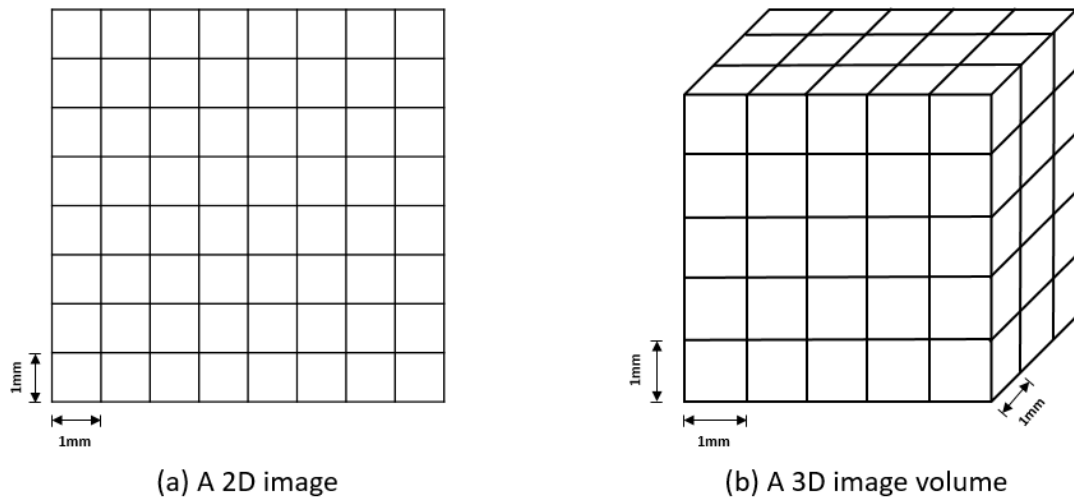


Figure 2.1: Pixels, voxels and spatial resolutions of medical images: (a) illustrate a 2D image with pixel resolution of 8×8 and a spatial resolution of $1\text{mm} \times 1\text{mm}$. (b) illustrate a 3D image volume with voxel resolution of $5 \times 5 \times 3$ where its spatial resolution is $1\text{mm} \times 1\text{mm} \times 1\text{mm}$.

voxels exhibit *spatial resolutions*, which describe the size of the details captured by individual pixels or voxels. For instance, a spatial resolution of $10.00\text{mm} \times 10.00\text{mm} \times 5.00\text{mm}$ means that a voxel depicts a region with a volume of 500.00mm^3 .

Contrast resolution refers to the range of distinct intensity or a set of intensities for red, green and blue (RGB) channels that can be distinguished in greyscale and coloured images respectively. A relatively low contrast resolution can be interpreted as pixel/voxel intensities that are similar in an image and are difficult to distinguish.

Figure 2.1 illustrates an example of a 2D image and a 3D volumetric image. Pixels and voxels are visualised in forms of 2D and 3D arrays of grids respectively.

In medical image processing, a region of interest (ROI) consists collections of pixels in 2D images that represent an area which encodes important information for particular domain or application in cancer diagnosis. The corresponding term volume of interest (VOI) refers to the emphasised anatomical structures that encode important knowledge for clinical uses.

2.1.2 Medical Imaging Modalities

Modern healthcare utilises different medical imaging modalities to capture different aspects of the human body for cancer diagnostic and treatment [29, 30]. Medical image modalities can be categorised into two classes based on their technique and process in visualising different aspects of the disease [31]: (i) Anatomical and (ii) Physiological (functional) imaging. Anatomical medical images capture and visualise the anatomical structures of the ROI in the form of 2D or 3D images. Anatomical medical images allow physicians to interpret and evaluate the disease conditions for diagnosis purposes, and can also be used for the monitoring of treatment responses [4]. Functional imaging, on the other hand, captures the metabolic status of the ROI, which allows physicians to assess the physiological status of patients, and to identify structures with abnormalities, such as tumours [32].

Common anatomical medical image modalities include X-ray, computed tomography (CT), magnetic resonance imaging (MRI). Common functional imaging includes single-photon emission computed tomography (SPECT) and positron

emission tomography (PET). Such image techniques produce a single type of image or image volumes that are referred as single-modality medical imaging.

Beyond the use of single-modality medical images for cancer diagnosis, the employment of multi-modality imaging types such as PET/CT and SPECT/CT has become one of the fundamental diagnosis protocol in clinical practices and basic medical research [33]. Despite multi-modality medical imaging exhibits strong potential to offer both tumour anatomical and functional properties, the evaluation of our proposed image-genomic framework did not directly involve datasets with multiple modalities. In this subsection, we emphasise on several single modality imaging techniques that are most relevant to our contribution.

2.1.2.1 Computed Tomography (CT)

Computed tomography is one of the most utilised imaging modalities in cancer detection, diagnosis and prognosis in modern clinical practices [34, 35]. CT scan makes use of the combination of multiple X-ray measurements to produce tomographic images. Figure 2.2 illustrates an example of CT image of the lung field from a Non-small Cell Lung Cancer Patient (NSCLC). In the coming chapters, we will explore the use of CT image volume with medical image analysis techniques,

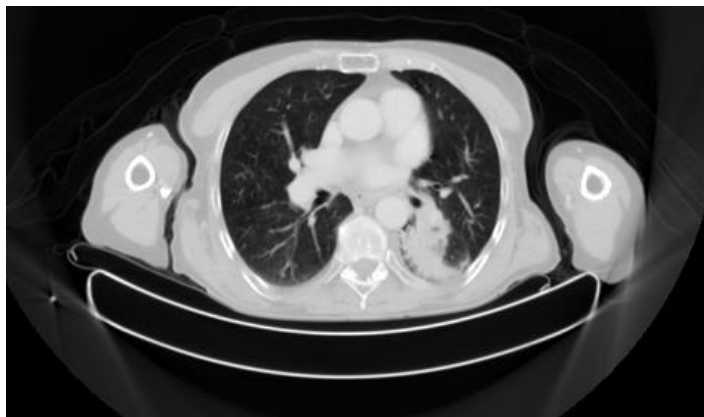


Figure 2.2: The CT image of the lung field from a NSCLC patient (data description detailed in Chapter 5).

including image feature extraction and its association to tumour phenotypic characteristics, which is the most relevant to our thesis contribution.

2.1.2.2 Digital Histopathology Images

Tumour histopathology is considered as the golden standard for cancer diagnosis. It is produced as the result of tumour tissue biopsies, where specimens of tumour tissue are stained and mounted onto glass slides for visual inspection [36]. Tumour histopathology offers the visualisation of cellular structures of tumour-containing ROIs, providing diagnostic information on mutated cell types, and hence offer therapeutic insights about the disease [36, 37].

The digitisation of histopathology was introduced through high-resolution whole-slide-imaging (WSI), which have enabled the quantitative analysis of tumour histomorphometry as well as its association with clinical practices for cancer diagnosis

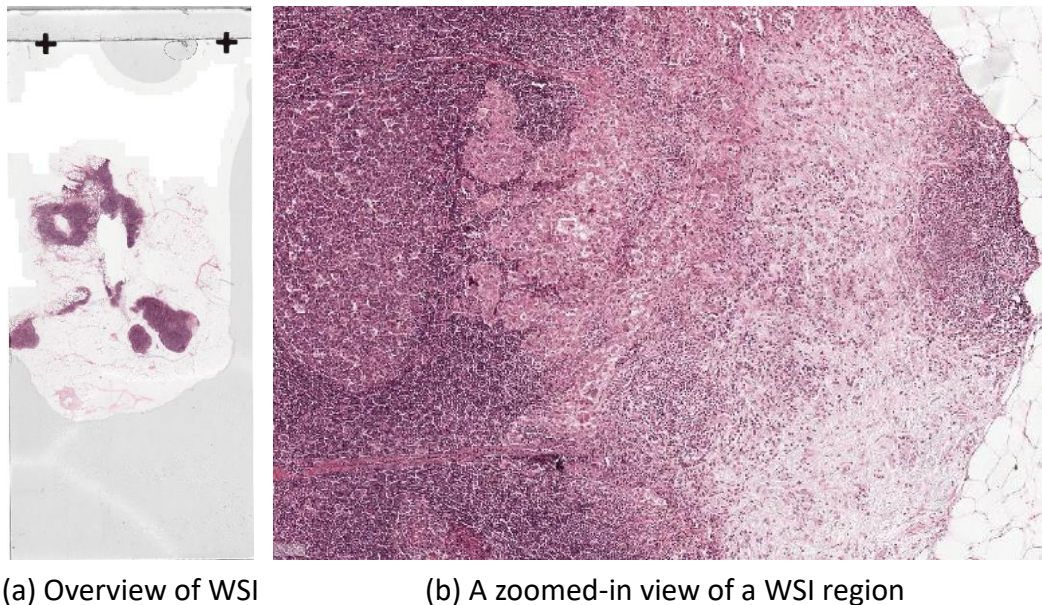


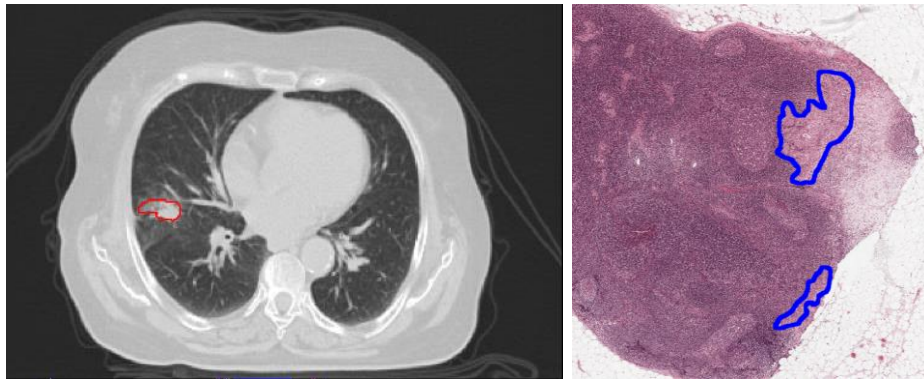
Figure 2.3: The overview and zoom-in view of WSI of sentinel lymph node with presence of breast cancer metastases. (a) The highest-level view of the entire WSI. (b) The zoom-in view of a region from the same WSI image (data description detailed in Chapter 4)

[21, 38]. Similar to CT images, digital WSI enables medical image analysis techniques to be applied to extract image features of abnormal cells. Image traits can then be exploited to correlate to genetic profiles of the tumour. Figure 2.3 illustrates an example of sentinel lymph node WSI with the presence of breast cancer metastases.

2.1.3 Medical Image Segmentation

This subsection describes several approaches for medical image segmentation that are most relevant to our contribution. We describe the current automated image segmentation approaches and compare with manual image segmentation by experienced physicians to reflect their distinct advantages.

Medical image segmentation refers to the process of image partitioning where each image partition contains a collection of pixels that together represent an ROI [39]. Medical image segmentation is commonly used to identify and emphasise ROIs, as it



(a) The region of interest (lung cancer) in a CT image (b) The region of interest (Breast cancer metastases) in a histopathology image

Figure 2.4: The segmented region of interests in different medical image modalities. (a) The delineated presence of lung cancer in an CT image. (b) The segmented regions in a WSI that shows the presence of breast cancer metastases in sentinel lymph node.

removes unrelated image regions to reduce the complexity of medical image analysis [40].

Medical image segmentation with manual delineation from experienced physicians is considered as the golden standard in most clinical applications. However, the current procedure of medical image segmentation for cancer patients relies upon the visual inspection of imaging data and the manual delineation of tumour ROI. Figure 2.4 provides examples of ROIs in different medical image modalities that were delineated by experienced physicians. The production of delineated tumour imaging data is time-consuming and prone to diagnostic errors, based on the expertise of the physician [41]. These challenges are exacerbated by the requirement of identifying quantitative morphological parameters of tumours such as shape and size [42].

To address this issue, computerised automated algorithms have been designed to facilitate medical image segmentation tasks. Commonly used medical image segmentation algorithm involves: pixel thresholding [43], region growing [44], edge detection [45] and fuzz clustering [46]. In the recent years, deep learning has also been utilised to perform classification and image segmentation tasks and achieved improved performance [47, 48]. Despite the performance improvement by deep learning-based approaches, few studies have shown superior accuracy to human perception in cancer diagnosis [49, 50]. The contribution of this thesis relies on accurate tumour delineations for domain adaptation purpose; we hence obtained tumour segmentation from experienced physicians as the foundation of our work. We will discuss the principle and applications of deep learning in later sections of this chapter.

2.2 Gene Expression Profiling

While the visual inspection of tumour histopathology is considered to provide a definitive diagnosis, it is prone to diagnostic errors in diseases where tumour classification is challenging due to visually indistinguishable morphological properties (e.g. diffuse large B-cell lymphoma and breast cancer) [51, 52]. Gene expression profiling refers to the technique of quantitative measurement of genetic activities that facilitate tumour classification and is also capable of predicting the clinical outcome for cancer patients [53]. In this subsection, we will describe the core concept of genes, gene expressions and its application in clinical applications that are relevant to our contribution.

A *gene* refers to a specific region of the deoxyribonucleic acid (DNA) strands, which exhibits genetic codes that produce ribonucleic acid (RNA) through **transcription**. RNAs are capable of synthesising proteins that encode biological functions through **translation**. However, human genetic codes are prone to mutation, permanent alteration of genetic elements that can be caused by many different factors such as DNA damage from environmental factors. Mutations can result in various types of changes in human DNA and may result in changes in behaviour as well as functions of genes. Clinical studies have shown evidence of genetic mutations in the prognosis of multiple human cancers [54-57]. Gene expression profiling identifies genes with abnormal expressions in tumour samples, thus offering insights into therapeutic options [16]. Gene expression profiling also provides patient-specific genetic information with normal expressions which encodes the genetic variables at the individual level that can ultimately contribute to the application of precision medicine [58].

Despite the significance of gene expression profiling in the diagnosis and prognosis of cancer patients, gene expression profiling is limited due to its reliance on invasive surgical procedures. Unfortunately, human cancers exhibit strong phenotypic and genetic heterogeneity within an individual with tumour manifestation at multiple sites [59]. Additionally, the differences in gene expressions across multiple sites may result in different responses to treatments [60]. The sampling and tissue acquisition of tumours from all disease sites is restricted due to patient discomfort and risk of undesired side-effects such as operation induced tumour proliferation. This thesis addresses this issue through the development of a novel image-genomic framework, discussed in detail in the following Chapter 4 and 5 of this thesis.

2.3 Machine Learning, Deep Learning and Convolutional Neural Networks

Machine learning techniques for medical image processing are a well-established field. The ability of machine learning to quantify the representation features of the input medical image empowers numerous automated medical image processing algorithms for different clinical applications. However, machine learning systems were limited due to its requirement of domain expertise with careful engineering to be able to learn and transform the input data [61]. Deep learning is a class of machine learning technique which allows learning of data representations with multiple layers of abstraction. Convolutional neural networks (CNNs) are a deep learning technique that extracts image features from imaging data to learn the sophisticated underlying representations with deep networks [62]. In this chapter, we emphasise on the principle of deep learning and CNNs as they represent the current state-of-the-art in image

object recognition. We will present the background and CNN architectures that are relevant to our study in the following sections.

2.3.1 Machine Learning

Machine learning is a major field of computer science that has been utilised to serve many aspects of modern healthcare systems. With carefully engineered mathematical models, machine-learning systems have been applied in pattern recognition [63], image classification [64], medical image retrieval [65] and tumour image segmentation[66]. Compare to the rule-based systems, the core of machine learning techniques are based on the development of models from statistical and artificial intelligence approaches. It is essential for the machine learning model to “learn” to recognise the distinguishing characteristics of the patterns within the data to produce meaningful outputs [67]. Generally, the learning approaches for machine learning models are divided into the following: supervised learning, unsupervised learning and reinforcement learning.

Although unsupervised and reinforcement learning exhibit strong potentials in multiple disciplines, our contribution rely on labelled medical imaging data to explore the region-specific genetic association. As the contributions of this thesis involve mainly supervised approaches, we will only cover those approaches here. Supervised training approaches require several different types of database, defined as follows:

- Training data: refers to a collection of data that are used to train the machine learning model. The machine learning model learns the representation of the training data and its predictive relationship to the output labels.

- Validation data: refers to a separate collection of data that are used in addition to the training data to adjust or guide the training process. This process involves the comparison between the predicted output with the training data labels. This provides an indication of the performance of the model on unseen data during the training process, and allows tuning of model parameters.
- Test data: refers to a collection of withheld data which is used to evaluate the performance of the model at the after the completion of the training process. Test data indicates the performance of the trained model with new examples. Test data are not involved in the training process.

2.3.1.1 Supervised Learning:

Supervised learning is one of the most common approaches for medical image processing and analysis. It refers to the approach where the model is trained with labelled data sets so that the model learns the internal representation of the input data to make predictions on the labels [68]. The resulting model from supervised learning is typically used to assign class labels with known predictive features for future data sets. Supervised learning is capable of performing classification or regression with training data set with discrete or continuous properties, respectively [69].

2.3.2 Artificial Neural Network

Artificial Neural Networks (ANNs) are a machine learning approach that was inspired by the biological neural networks that constitute human brains [70]. An ANN consists a collection of connected nodes or “artificial neurons” in a directed graph in the form of networks. ANNs are commonly used in machine learning to learn the complex non-

linear relationships from the dataset. This is achieved through the ANN's mechanisms where each neuron receives, processes and transmits a signal from one to another in a similar way to the biological synapse. To achieve the optimised learning outcome, ANN requires the design of an appropriate network structure and learning approach to tune the weights and biases of the network.

Neurons are the fundamental building blocks of many ANNs. Figure 2.5 illustrates the structure of a single neuron where multiple inputs values are processed in the neuron to produce a single output value, where X is a vector of inputs with n elements, W is the vector of the weights with a corresponding number of elements, which can be summarised in equation 2.3.1:

$$L(W, X) = \sum_{x=0}^n f(x_i, w_i) + \text{bias} \quad (2.3.1)$$

where $f(x_i, w_i)$ is the ANN function that generates the output value.

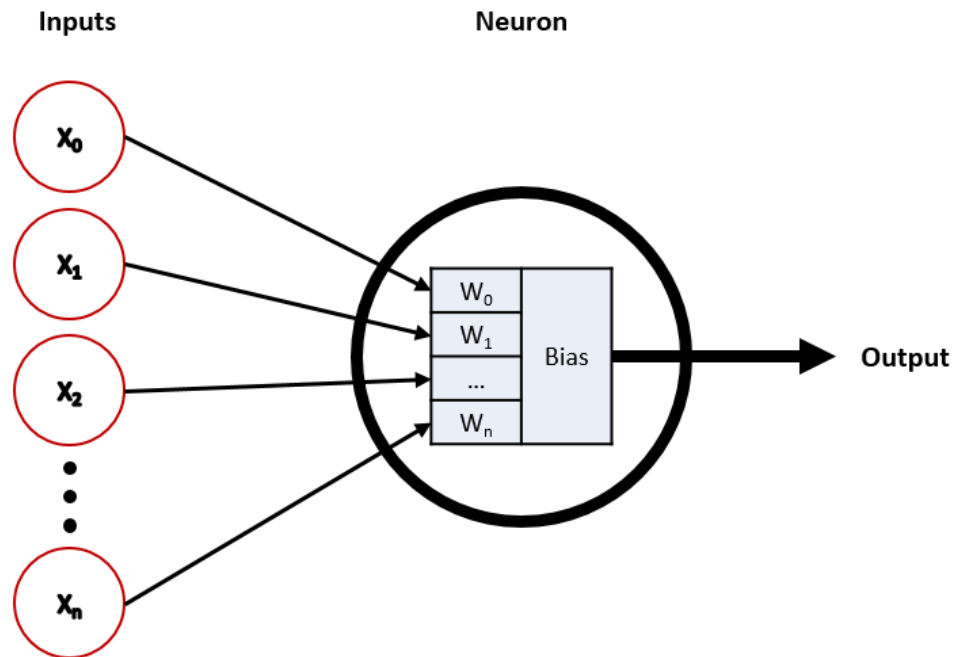


Figure 2.5: Representation of a single model neuron

The training of an ANN can be explained by illustrated single neuron model. For each element in the training dataset, its features are extracted and transmitted into the neuron in the form of a feature vector. Assuming that the classification task involves only two classes, the neuron processes the input feature vector by multiplying with its internal weights. A class prediction is based on whether the value of the product is above a threshold. If the predicted class does not match the corresponding label, each weight within the neuron is adjusted individually to refine the prediction process. Additionally, the training dataset is typically divided into collections of batches which allows the weight adjustment after training of each batch. An *epoch* refers to the full pass of the entire training dataset; robust weights for classification tasks are the outcome of hundreds of training epochs.

2.3.3 Deep Learning

Deep learning can be categorised as a class of techniques of machine learning [61], which allows the computation models with multiple processing layers to learn the internal representation of the input data with multiple levels of abstraction [71]. Deep learning was proposed to address the limited performance of traditional machine learning approaches to process natural data in their raw form. Deep learning resolves this issue by utilising the multiple processing layers to learn and interpret the low level, abstract representations from the high-level understandings of the input dataset [61]. This is achieved by feeding the raw data through the successive multilayer architecture in a sequential manner, where deeper layers learn the abstract representation from the representation in the previous layers [67].

Deep learning technique employs backward propagation of error or “backpropagation” to train the multilayer model for supervised learning [61]. Backpropagation calculates the gradient of the error function of the ANN with respect to its weights and passes the gradient backwards through the neural network. Compared to the traditional approach where the gradient of the error function is calculated for each layer separately, the backflow of error gradient allows more efficient computation of gradient for ANNs [72].

The implementation of deep learning technique utilises specialised GPUs to improve the performance of training process by 10 to 20 times compared to the traditional training approach on standard CPUs [67]. Recent advances in deep learning have lead to the improved state-of-the-art in various domains such as visual object recognition [67], speech recognition [73] and also in medical image analysis [42].

2.3.4 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) is a particular type of ANN and designed to process input data that is in the form of multidimensional arrays, e.g., coloured 2D images which consist of 2D arrays for each RGB (colour) channel. Compared with traditional ANNs with fully connected adjacent layers, CNNs are much easier to train and are more generalised [67]. CNNs are structured in a series of stages where each stage consists of specialised layers with unique functions. The building blocks of a CNN consists three types of specialised layers: convolutional, pooling and activation layers, e.g., rectified linear unit (ReLU) layers [61].

Convolutional layers consist of organised units in the form of feature maps, where each unit is connected to local patches from the previous layer through a set of

weights which is referred as filter banks as shown in Figure 2.6. The resultant sum of the local filter bank is typically transmitted through a non-linear activation layer such as ReLU layer, e.g. $f(x) = \max(0, x)$. Only units within the same feature map share the same filter bank. This design allows the convolutional layers to detect local conjunctions of features from the previous layer, as local values are often highly correlated and are invariant to the location in an image input.

Pooling layers are designed to merge features in spatial proximity which share semantic similarities into one. The principle behind pooling layers is to detect the position of motifs that are typically formed by highly correlated features through a coarse-grained approach. An example pooling layer calculates the maximum of a local patch in one or more feature maps. Pooling layers act to reduce the dimensions of the representations and to create an invariance to small distortions and shifts.

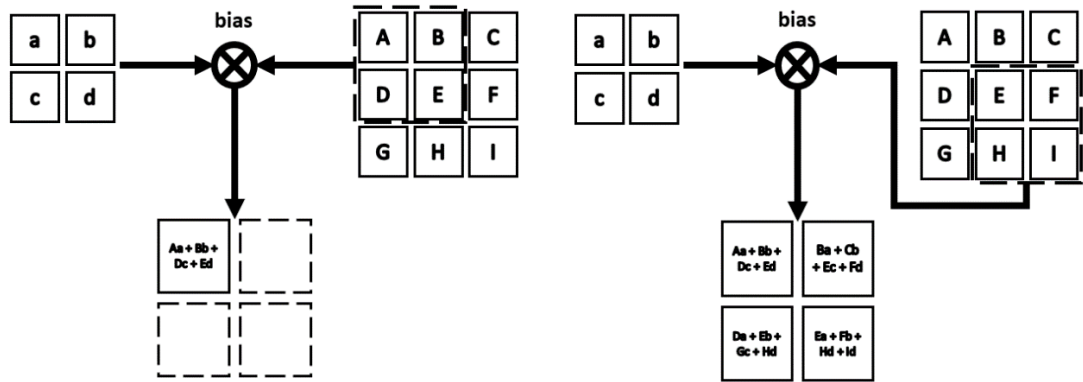


Figure 2.6: An example 2D convolution of a 3×3 input using a 2×2 filter with the 2×2 output feature map.

A standard CNN typically involves stacked layers in order of convolution, ReLU and pooling layers, and followed by fully connected layers as shown in Figure 2.7. CNNs also utilise backpropagation to train its filter banks in a similar way compare to regular ANNs.

CNNs are inspired by the biological signal transmission through cells with different complexity and functions in visual neuroscience [74]. CNNs exploit the compositional hierarchies in natural signal processing where high-level features are the result of the composition of low-level features, to apply onto images where objects are the results of the composition of edges, motifs and parts.

CNNs have demonstrated large improvement to the state-of-the-art in multiple disciplines. CNNs have been adopted in medical image analysis to perform a range of tasks including medical image segmentation [75], classification [76] and disease detection [50].

In the following subsections, we will explore the two CNN architectures that have shown outstanding performance in image object recognition and that are most relevant to our contributions.

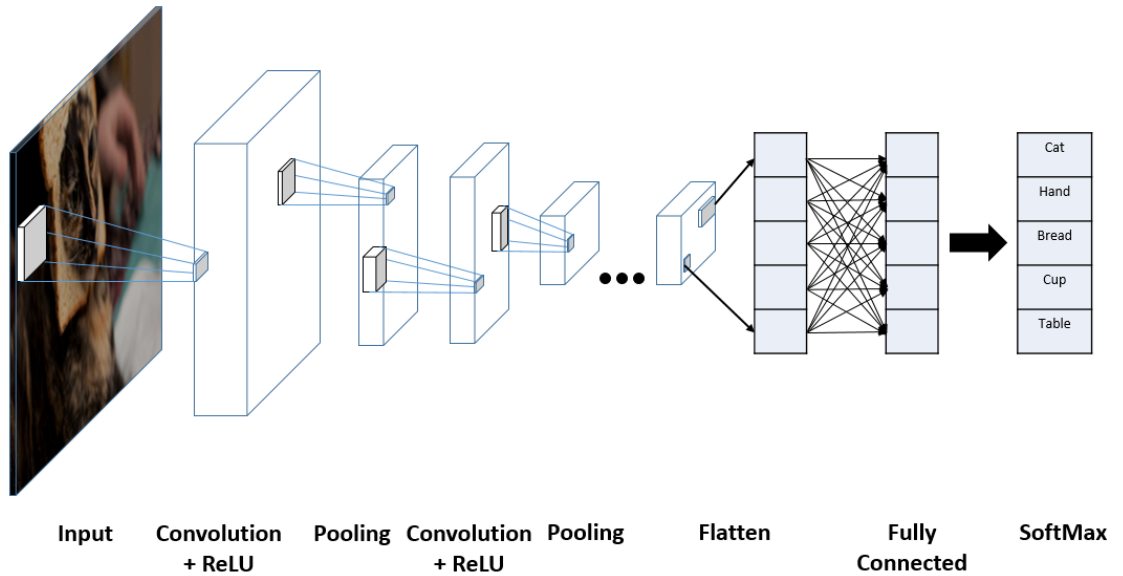


Figure 2.7: An example CNN with many convolution layers. Filters are applied to each input image and the output of each convolutional layer is used as the input to the next layer.

2.3.4.1 GoogLeNet (Inception)

GoogLeNet [77], introduced in 2014 as a submission to the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14) is a deep CNN architecture consists of 22 layers [78]. GoogLeNet was designed to act as a classifier for natural images with the underlying concept to increase the width and depth of CNN while keeping the computation cost constant. GoogLeNet demonstrated significant improvement to the state-of-the-art at the time by achieving a top-5 error of 6.67% for the classification challenge. A standard GoogLeNet architecture is shown in Figure 2.8.

GoogLeNet employs the “inception modules” concept to exploit the approximation and coverage of optimal local sparse structure by dense data structures. Within each inception model, the input is fed into separated convolution layers with dimensions of 1×1 , 3×3 , 5×5 and a 3×3 max pooling layer. Results from each layer are concatenated into a single output vector to serve as the input for the next stage.

Also, GoogLeNet exploits the “network-in-network” approach [79] by adding a 1×1 convolution layer followed by the standard stacked convolutional layers

structure for dimension reduction which removes the computation bottleneck and hence increases the representational power of the neural network.

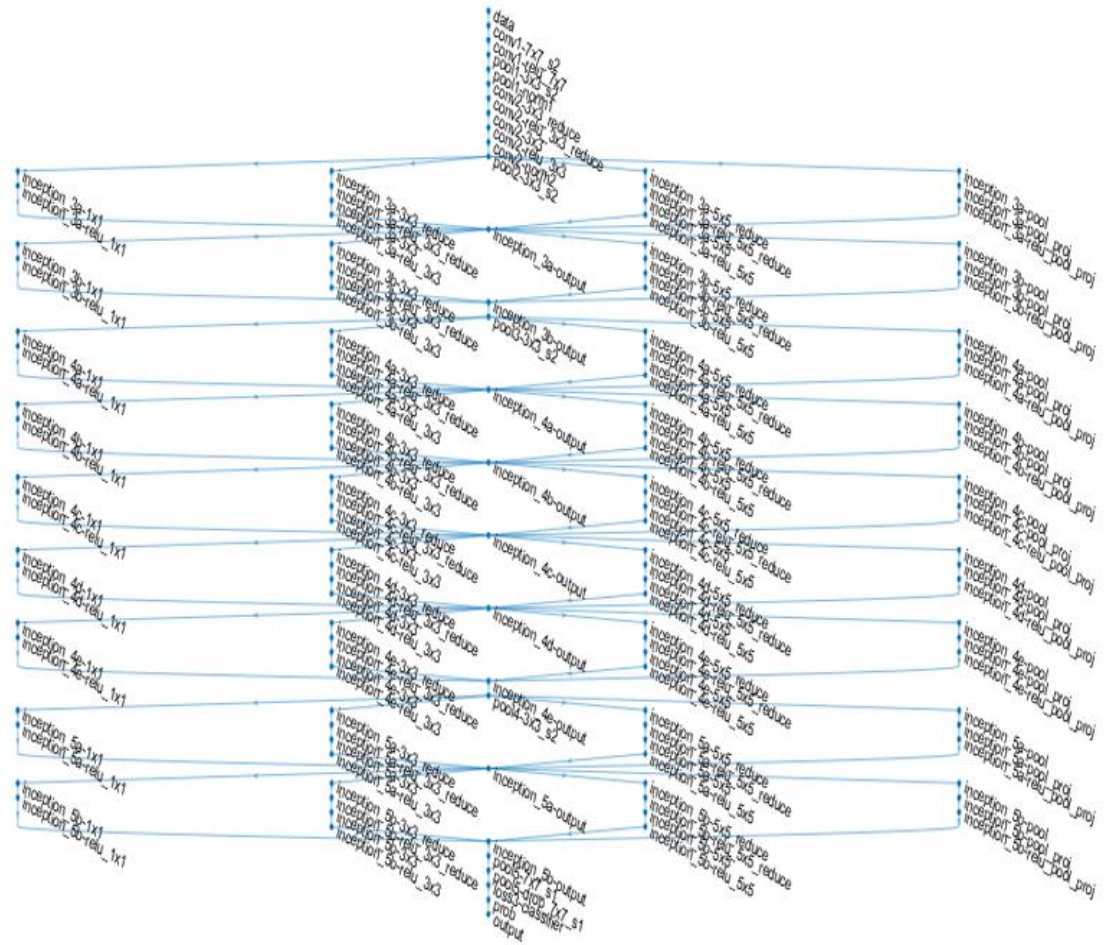


Figure 2.8: The GoogLeNet architecture. Layers labeled with “reduce” represents the additional 1×1 reduction layers before each of the 3×3 and 5×5 convolutions.

2.3.4.2 Deep Residual Networks (ResNet)

Several studies, including GoogLeNet, have demonstrated that the depth of CNNs is one of the crucial factors to improve the learning outcome for image object classification tasks with ImageNet dataset [80-82]. However, rapid degradation of learning accuracy was observed in CNN architectures that stack many layers together

in a traditional way. The rationale behind the training degradation was not due to overfitting, but an optimisation problem of the mapping of identities for the additional layers.

ResNet [83] was proposed to address the challenge of training accuracy degradation in deep CNN architectures through the use of residual mapping to fit the successive stacked layers. ResNet employed the concept of residual representations [84, 85] and created a shortcut connection [86] with closed gating functions [87, 88] for identity mapping. The residual function and identity mapping by shortcuts are according to:

$$f(x) = \mathcal{H}(x) - x \quad (2.3.2)$$

where $f(x)$ denotes the residual function, which is the result of removing the input x of the first layer of the stacked layers from the underlying mapping $\mathcal{H}(x)$. Residual learning drives the weights of the multiple non-linear layers towards zero to approach identity mapping, which is defined as:

$$y = f(x, W_i) + x \quad (2.3.3)$$

where y and x denote the output and the input of the stacked layers respectively and $f(x, W_i)$ is the function that represents the residual mapping that needs to be learned. The shortcut connections in ResNet does not increase the number of parameters of the network nor the computational complexity. A ResNet building block with residual learning and shortcut connection is illustrated in Figure 2.9.

ResNet further improved the state-of-the-art in image object recognition at a top-5 error of 4.49% with a 152-layer ResNet architecture, compared with 6.67% from GoogLeNet. ResNet also demonstrates a good generalisation performance for different

recognition tasks [83] and hence has been widely adopted in medical image analysis [89, 90].

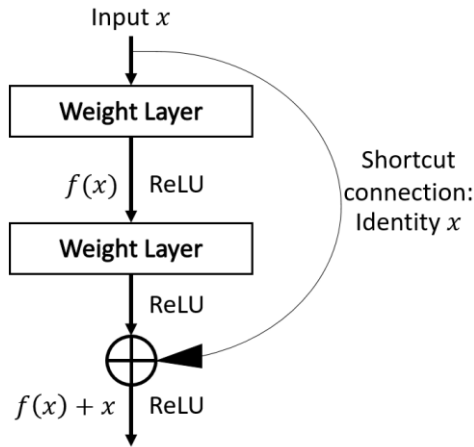


Figure 2.9: The building block of ResNet

2.4 Domain Adaptation for Transfer Learning

Common applications of machine learning-based techniques involve the training of the statistical model to quantify the internal representations of data from the target domain. Such an approach relies on the assumption that the training and real-world (used in a future application) datasets share the same feature space with a similar distribution [91]. The changes of the distribution of feature spaces in the datasets typically require that the trained statistical model is rebuilt to fit the newly collected data. However, the assumption does not always hold, as most machine learning methods require large annotated training datasets that are specific to the target domain. Machine learning applications are thus greatly limited in domains such as medical image analysis where the acquisition of labelled training data is expensive.

Transfer learning is the alternative approach to train the statistical model without the need for the large quantity of domain-specific datasets. Transfer learning refers to the process where the knowledge acquired from one large database is

transferred to facilitate learning applications on a related but smaller dataset. For neural network applications, transfer learning can be described as an approach to speed up the learning of a specific domain by using and transferring the weights obtained from a network that was trained for a related source task [92]. An illustration of the transfer learning approach is shown in Figure 2.10.

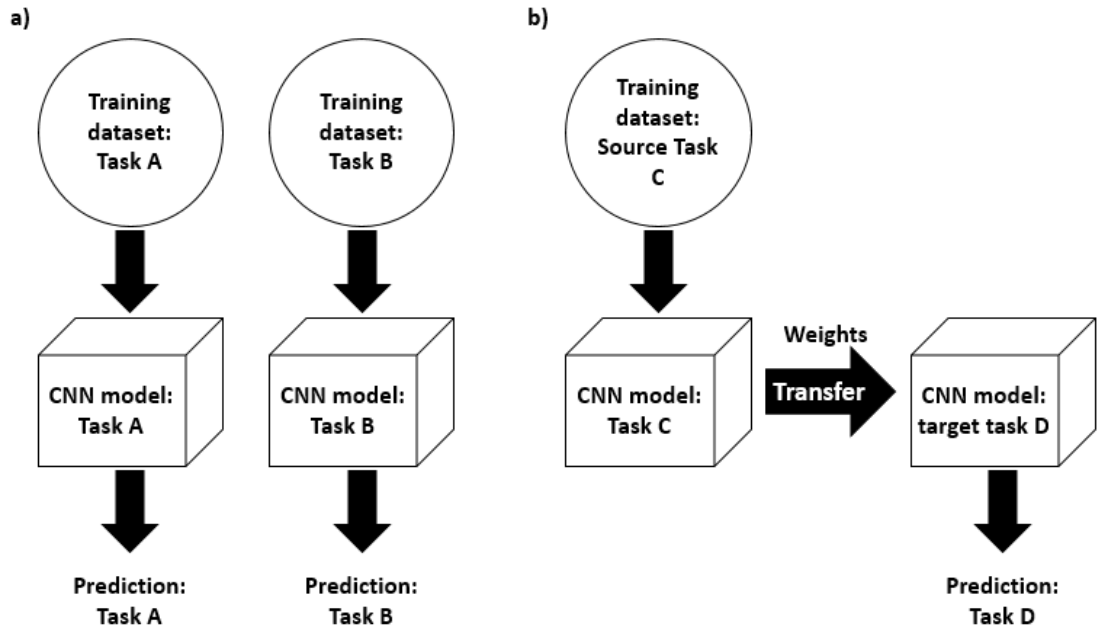


Figure 2.10: Traditional and transfer learning-based approaches for CNN models. a) Traditional machine learning. b) Transfer learning-based approach for CNN model.

Transfer learning technique allows CNNs to be trained and applied to datasets from different but related domains. The differences between domains can be categorised into two groups: (i) domains with different feature spaces, or (ii) domains with the same feature space but different marginal probability distributions [91]. To evaluate the representation and the similarity between two domains, techniques such as A-distance, have been adopted for transfer learning applications [93, 94].

There are two major approaches to employ domain adapted transfer learning consists for CNNs applications in medical image analysis: (i) using “off-the-shelf CNN” features; and (ii) using domain adaptation with fine-tuning techniques [95]. The

term “off-the-shelf CNN” refers to the procedure of utilising a CNN model which was trained on a larger dataset from scratch as a feature extractor directly on a new dataset with a smaller volume size. Parameters in the convolutional and fully connected layers are not changed. For classification purposes, the extracted features from the “off-the-shelf-CNN” can be used to train a separate classifier, such as support vector machines, random forest classifier or to train only the classification layer of the model to fit the number of classes of the new dataset [95-98].

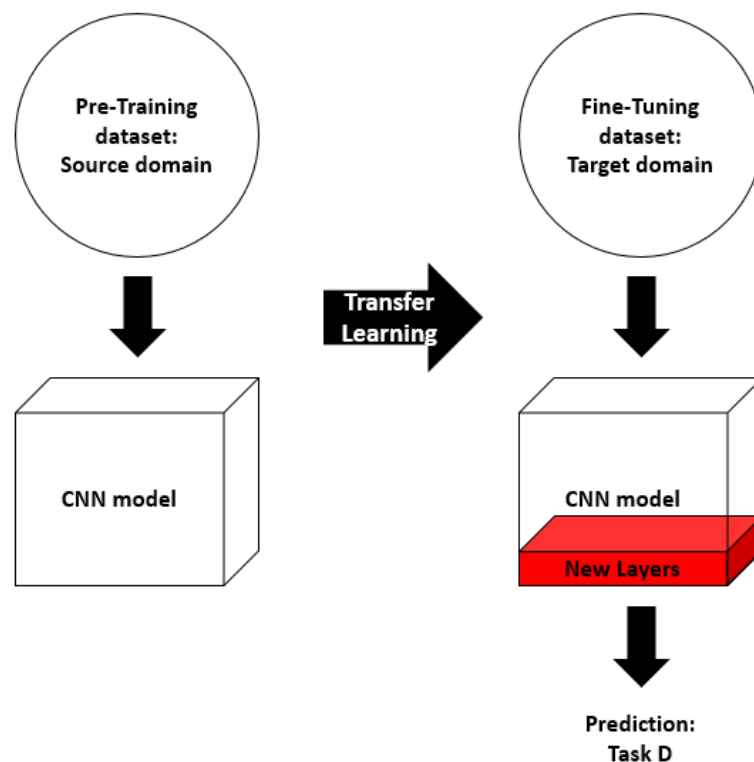


Figure 2.11: Overview of shallow tuning-based transfer learning approach.

Fine-tuning refers to the process of updating the weights of a pre-trained CNN model new datasets from the desired domain through the use of backpropagation. Common approaches for fine-tuning involves the training of only the last few convolutional layers, referred as “shallow tuning” or the tuning of all convolutional layers of the network, called “deep tuning” [99]. Deep tuning is typically employed if

the distance between domains is significant. The tuning of early layers of CNNs involves the learning of low-level image features that are similar to Gabor filters and colour blobs, which are applicable to most vision tasks and hence are not applied to shallow tuning approach [100]. Studies have shown that fine-tuning is as effective as a CNN that is trained from scratch and is more robust to the size of training dataset [99]. Fine-tuning has been applied to a variety of medical image analysis tasks such as ultrasound anatomy identification and medical image modality classification [89, 101]. An illustration of “shallow tuning” is shown in Figure 2.11.

Chapter 3

Image-Genomics

This chapter presents the overview of image-genomics research in various clinical contexts. We provide a survey of the current state-of-the-art of image-genomics with a focus on the extraction of genetic-related image features, their application to medical images with different modalities, and identifying the gaps in the existing approaches.

3.1 Overview of Image-Genomics

Cancer treatment relies upon accurate diagnosis, which can benefit from accounting for individual variabilities in genetics, environment and lifestyle [1], but requires a complex combination of information gathered from clinical, histopathology, imaging, and genetic data. Medical imaging is one of the fundamental protocols in modern healthcare for the examination of human cancers [12]. It offers a fast and non-invasive assessment of tumour visual characteristics which can be used as oncologic diagnosis tool as well as providing treatment guidance. Briefly discussed in previous chapters, computerised analysis of medical images has been demonstrated that quantitative

image feature analysis is capable of capturing distinct tumour phenotypic characteristics, e.g. in soft tissue sarcomas (STSs) [102], lung adenocarcinoma [22] and colorectal cancers [103], thereby indicating the potential for image features analysis to discover and identify tumour imaging biomarkers. The patient-specific genetic information also offers insights into therapeutic options and hence can be used to optimise patients management and treatment plans, for example, the use of immune-modulating therapies in melanoma and lung cancer [104]. Also, patient-specific genetic data have been used in medical research to analyse high-dimensional mineable characteristics, which consists detailed characterisation of biological information such as DNA (genomics), RNA (transcriptomics), proteins (proteomics) and metabolites (metabolomics) [105].

Nevertheless, the clinical implementations are restricted by the challenges in acquiring tumour genetic profiles. As previously described in Chapter 1 and 2, human cancers exhibit strong phenotypic and genetic heterogeneity, where different gene behaviours are expressed in various tumour regions. Thus the potential need for multiple biopsies may introduce further risk of patient discomfort and the occurrence of undesired side effects such as tumour haemorrhage. As such, the consequence is that multiple biopsies are rarely performed [17-19].

The emerging field of image-genomics involves the integration of patient-specific clinical information from both imaging and genetics [106]. Image-genomics can be thought of as the identification of surrogate tumour biomarkers through correlating image features with gene expression and has been suggested as a non-invasive alternative to biopsies [5, 107]. It aims to covert clinical images into high dimensional mineable image features using quantitative, high-throughput feature

extraction algorithms for correlation to genetic information [12, 108]. Studies have shown promising result in identifying prognostic image-genomics biomarkers for several cancers including lung cancer [109], breast cancer [110], and high-grade primary brain tumours [111].

As such, current image-genomics research requires large volumes of annotated multi-dimensional data. The following clinical data are essential for image-genomics studies:

- Segmented medical images with precise annotations
- Annotated clinical metadata which provides patient-specific clinical parameters
- Labelled patient-specific gene profile or raw expression array data

At the current stage, most of the image-genomics research is limited to the publically available datasets that are published in The Cancer Imaging Archive (TCIA) [112] and The Cancer Genome Atlas [113], or data that is privately collected. The limited amount and quality of the dataset make image-genomic research more challenging, for example, the confined feature selection process due to the inadequate amount of clinical metadata and annotations.

In the following subsections, we will explore the use of image features and the techniques for image feature extraction in various contexts. Figure 3.1 outlines the crucial procedures of an image-genomic framework. We will focus on the extraction of image features from two of the widely adopted imaging modalities: CT and histopathology and their corresponding application in the field of image-genomics.

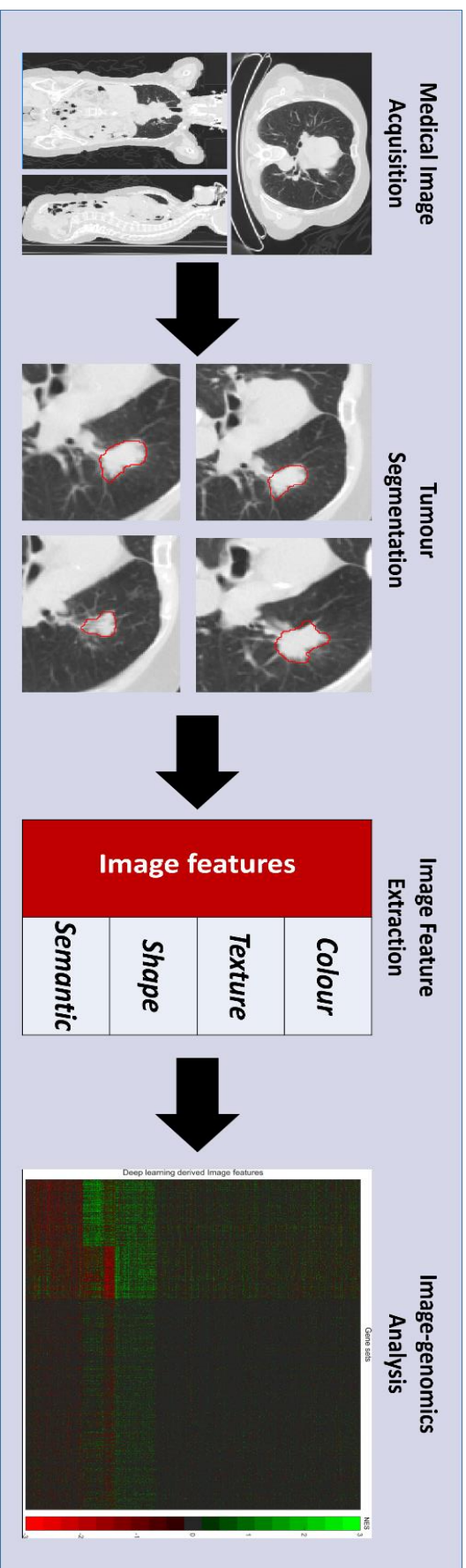


Figure 3.1: The current image-genomic approach for associating image features with gene expression profiles.

3.2 Image Feature Extraction

The famous quotation from Bartlett, “*A picture is worth ten thousand words*”¹, illustrates the amount of information encoded within a single image. Similar to the human behaviour, computerised image analysis techniques extract and select information from images to solve problems in various contexts. This is facilitated through the employment of models that were trained on specific domains [114].

Briefly mentioned in Chapter 2, image feature extraction is a process which finds and transforms image encoded information into suitable image representations to be used for various applications [115-117]. In the field of image analysis, extracted image features can be categorised into global, which are calculated from the entire image, or local features, which are extracted from specific ROIs. Image features that are applicable to many vision tasks consist colour, texture, shape and the structural representation of an image.

Texture-based features are one of the most important characteristics to identify objects and ROIs in an image. Texture is an innate property of surfaces in an image regardless of its modality [118] and contains crucial information about the structural arrangements of surfaces as well as its relationships to the surrounding environment of the ROIs. Texture-based features offer a precise description of the homogeneity of the local spatial variation of pixel intensity [119] and texture analysis serves computerised applications in image segmentation [120], classification [121] and pattern recognition [122]. Commonly adopted texture-based features include the use of Haralick and wavelet features [123, 124]. Haralick features are extracted from

¹ Original Chinese proverb, writing “畫意能達萬言”.

the distribution of co-occurring pixel values in the image within a specified spatial neighbourhood, hence providing a statistical summary of the relative distribution of gray tones in the image [118].

The extraction of texture features through the use of wavelets have also contributed to tasks such as content-based image retrieval [125], segmentation [126] and classification [127]. Wavelet features are extracted by decomposing the signal of images at various resolutions based on wavelet orthonormal [128]. The orthogonal multi-resolution representation provides a hierarchical framework to offer image characteristics in a coarse-to-fine strategy that extracts contextual and low-level image features [127, 128].

Colour-based features are one of the most adopted representations of images due to its robustness to background complications, as well as its independence to the size and orientation fo the image [129]. Colour features also offer both global and local image representations with various applications. Colour representations such as colour indexing, colour descriptors and compact colour moments have been adopted in applications such as image retrieval [130-132] and scene recognition [133]. Colour histograms provide insights on global colour distribution and ranges, and have been used in applications such as face detection [134] and bleeding detection in medical domain [135]. Additionally, the application of colour-based image representations to the three-channel domain offers complementary characteristics of the image compared to single-channel grayscale images [136].

Shape-based features offer descriptions to quantify the shape of ROIs in ways that agree with human perception for specific tasks [137]. Shape representations capture the geometric details within the ROIs of an image, and are invariant to

translation, rotation, scaling and resistant to noise [129]. Shape-based features have been extracted and applied through multiple approaches. For 2D images, shape-based features have been represented as point sets [138], outline curves [139] and shock graphs [140]. Such features have been adopted for image object classifications [141], recognition [142] and content-based image retrieval in various domains [143].

In 3D volumetric images, commonly adopted shape features use spherical harmonics that decompose the 3D object into orientation invariant and descriptive information [144]. Additional approaches for 3D shape-based features extraction involve the analysis of object surface curvature and the correlograms of the objects from various view points [145]. Such 3D shape-based features have been widely used in 3D model search engines [144], classification tasks [146] and content-based retrieval [147].

In addition to the image features that quantify the statistical information of pixels (hereafter referred to as “**agnostic features**”), yet another set of features have been commonly utilised to represent high-level tumour characteristics based upon human understanding (hereafter referred to as “**semantic features**”), e.g. tumour shape, size and necrosis [105]. Different from agnostic features, semantic features are commonly used by radiologists to describe tumours, with the foreknowledge of their prognostic value in cancer treatment [148]. The employment of semantic features in image-genomics studies has demonstrated their capability to identify prognostic imaging biomarkers [23] and to predict gene expression patterns in hepatocellular carcinoma [149]. As a collection, agnostic and semantic features quantify the image representation of tumour phenotypic traits that are based on established human knowledge in cancer physiology. To differentiate such a collection of image features

from artificial intelligence produced image descriptors, agnostic and semantic features are hereafter referred to as “**human-crafted**” features.

Human-crafted image features have been an integral component of image-based CAD systems [150] in many clinical applications such as disease detection and classification [151, 152], improving diagnostic performance [153] and ROI segmentations [154].

3.3 Summary of Gaps

A number of image-genomics approaches have been proposed and validated for different image modalities and clinical applications across various cancer types. The multiplicity of image-genomics research has encouraged studies to employ images with multiple modalities and to seek correlations with different genetic information.

Recent image-genomics research stimulated the extraction and quantification of image features that can be significantly correlated to genetic information. Other recent image-genomics studies attempted to derive image features from multi-modality images, e.g., PET/CT, to provide quantification of tumour phenotype as well as its underlying metabolic status to identify imaging surrogate of prognostic biomarkers [23].

Despite the progress and advances in image-genomics, a number of gaps remain unaddressed in the current state-of-the-art. In the following subsections, we provide a detailed explanation of aspects that should be pursued and improved to better seek image-genomics associations.

3.3.1 Medical Image Feature Extraction

Although current image-genomics approach has been validated in a number of clinical research, there is a gap in the quantification and selection of medical image features. Neither agnostic features nor semantic features were designed based on particular cancer or tissue type, thus may not be optimised to capture the subtle, low-level descriptions of specific disease or patients cohort [27]. As a result, human-crafted image features lack the ability to fully capture tumour phenotypic characteristics, where crucial biomarker may be encoded. Thus, the use of complementary image features that are optimised to represent specific disease and patient cohorts would be advantageous for future image-genomics studies.

Additionally, the curse of dimensionality is an issue to be addressed in current state-of-the-art. The process of feature extraction and selection should take into consideration of the available genetic biomarkers. Parameters from the clinical metadata have been adopted in studies to perform image feature stability test, hence offering features that are most stable in the patient cohort [22, 23].

The adoption of new imaging techniques in clinical practices also offer opportunities to derive image features with unique advantages. The investigation and extraction of new and more genetic related image features will also bring advantages for future image-genomic research.

3.3.2 Domain Adaptation

Discussed in the previous subsections, current image-genomics studies rely heavily on large volumes of multi-dimensional data. However, such datasets are difficult to acquire in addition to the need for precise annotations. As such, we suggest that the

employment of domain adaptation from our proposed framework exploits the data availability and hence reduce the reliance on large datasets for image-genomics research.

Domain adaptation has been widely adopted in image processing and analysis field and is a well-established tool to address the challenge of an insufficient amount of annotated medical image datasets. We have briefly discussed the fundamental concepts behind domain adaptation in Section 2.3.4. However, the current state-of-the-art in image-genomics relies solely on datasets that are either exclusively to the authors or publically available data. Hence studies that utilised small datasets may not be able to extract optimised tumour image representations for associating gene expressions.

Although the utilisation of domain adaptation technique offers opportunities to extract tumour descriptors from similar knowledge domains, its potential in image-genomics applications should be thoroughly examined. Such adaptations will benefit future image-genomic explorations where data for particular diseases or patient cohorts may be limited. Domain adaptation also has the potential to develop different approaches or techniques for image-genomics that are more efficient, practical and applicable.

Chapter 4

Domain Adaptation Facilitated

Medical Image Analysis

This chapter describes our proposed domain adaptation facilitated medical image analysis system. The domain adaptation scheme aims to overcome the reliance on large volumes of annotated datasets for learning tumour image representations and to enhance medical image analysis with regards to tumour image classification. We evaluate our method by comparing to the traditional scratch-trained CNNs for tumour classification task. Our results indicate that domain adaptation facilitates the training process of CNNs and show improved prediction accuracies for the tumour image classification tasks where training datasets are limited.

4.1 Overview

Tumour histopathology analysis is an important procedure which provides definitive cancer diagnosis [36]. However, the current practice to detect the presence of tumour

abnormalities relies upon visual inspection on histopathology slides, which is time-consuming and prone to diagnostic errors, based upon the expertise of the pathologist. These challenges are exacerbated when quantitative histomorphological parameters such as mitotic counts and surface areas are required to be identified and extracted.

The introduction of high-resolution whole-slide imaging (WSI) [155] has enabled the digitisation of tumour pathology slides. WSI has given rise to automated pathology image processing techniques that have demonstrated improvements in the efficiency and accuracy of pathology slide interpretations [156].

In this section of the thesis, we propose a domain adaptation facilitated medical image analysis system that aims to enhance tumour classification task in WSIs, especially in circumstances where labelled training data are limited. The technical contribution of this study addresses and overcomes a number of gaps in the medical image analysis research: (i) The learning of disease-specific image representations that encode tumour phenotypic characteristics and; (ii) to overcome the reliance of large volumes of annotated datasets for learning disease-specific image representations.

Detailed in Chapter 3, deep learning-based approaches have been widely adopted in medical image analysis, including applications in histopathology images. Recent deep learning methods for image classification and detection using CNNs have demonstrated marked improvements in the performance of a range of medical imaging analysis tasks, e.g., nuclei segmentation [76] and tissue classification [157]. However, most well-established deep learning methods require large annotated training datasets that are specific to the particular problem domain [61]; such datasets are difficult to acquire for histopathology data where visual characteristics differ between different tissue types, in addition to the need for precise annotations.

Our proposed system employs domain adaptation and deep learning-based techniques to learn the image representations of tumour phenotypic characteristics from WSIs with similar knowledge domain. Such image representations were employed to facilitate the tumour classification task with limited volumes of annotated medical image datasets.

4.1.1 Framework Overview

Figure 4.1 outlines the main procedures of the proposed domain adaptation facilitated tumour classification framework. The proposed framework initiates with an image pre-processing step where WSIs were processed to extract tumour-bearing and non-tumour patches for both datasets (please see section 4.2 for details of the datasets). Shown in Figure 4.1, a standard CNN model was trained from an annotated breast cancer dataset to learn abstract image representations of WSIs as part of the pre-training phase. In the fine-tuning phase, we optimise the initial weights from the pre-trained CNN to be more representative of the subtle variations in visual characteristics from a relatively smaller dataset that contains prostate cancer WSIs. Result evaluation calculates the 1) prediction accuracy with a subset of prostate cancer dataset; 2) The sensitivity and specificity of the classifier to plot an operational receiver characteristic graph. Section 4.5 and 4.6 provide detailed evaluation strategy and result interpretation, respectively.

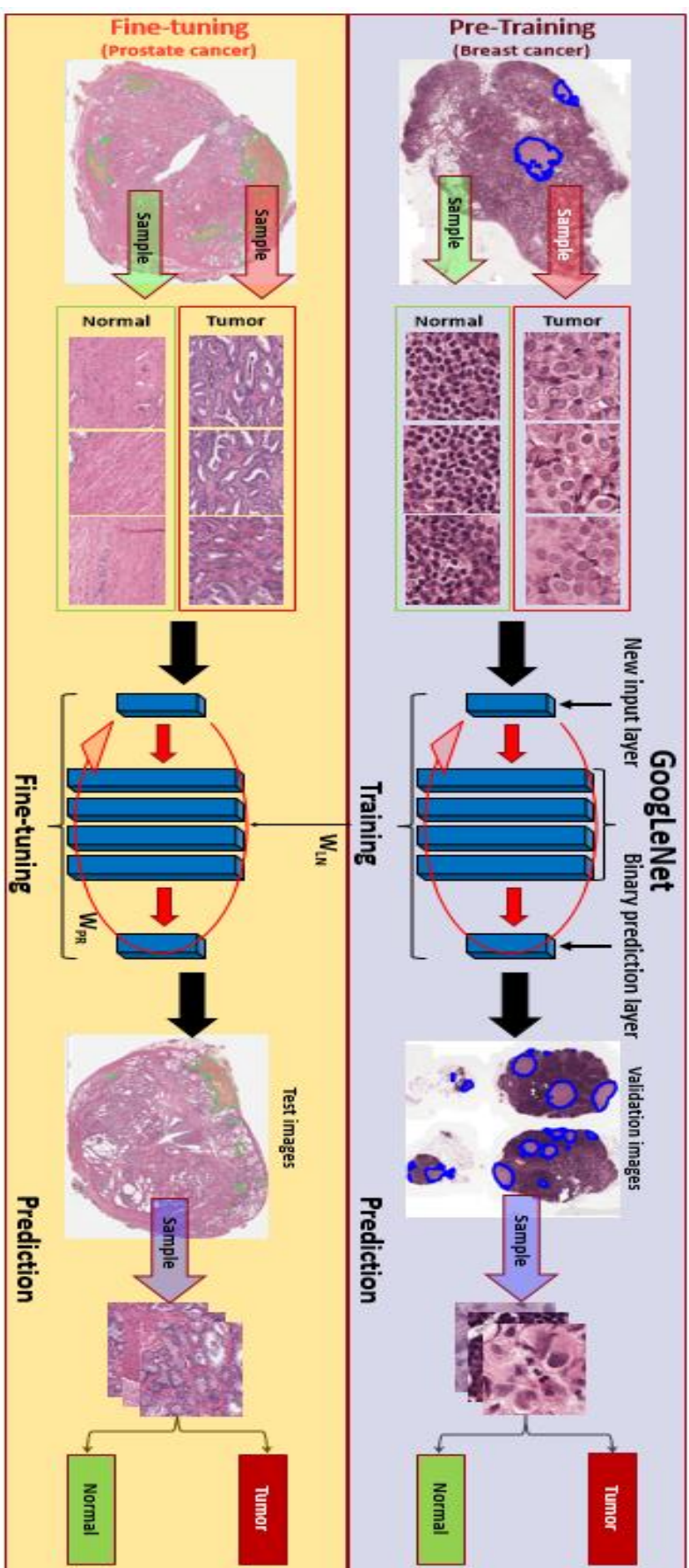


Figure 4.1: An overview of the proposed framework for patch-level tumour detection.

4.2 Datasets

We used high-resolution WSIs from two cancer types:

- Lymph node metastases dataset: We used 110 WSIs of metastases-containing sentinel lymph nodes digitised at a maximum of 40 \times magnification from the Camelyon-16 challenge dataset [22]. The ground truth data were provided in the form of WSI binary masks to indicate the regions of metastatic breast cancer. We randomly selected 100 WSIs for training and 10 WSIs for validation. This dataset was used as the larger set for CNN pre-training.
- Prostate cancer dataset: We used the publically available prostate fused-MRI-pathology dataset from the Cancer Imaging Archive (TCIA) [23, 24], comprising of 16 biopsies confirmed subjects of radical prostatectomy specimens. Each prostate specimen was digitised at 20 \times magnification and digitally stitched together from 4 separate slides to form a larger pseudo-mounted histopathology image. Ground truth annotations of cancer presence were made by an expert pathologist. We removed one subject from the study due to missing annotations. We randomly assigned the pseudo-mounted histopathology slides of the remaining subjects into two datasets: twelve were used for a training/fine-tuning dataset and three were used for the test dataset.

4.3 Image Pre-processing and Patch Extraction

The white background space of WSIs contains no meaningful information and hence are removed in most classification methods [158]. We removed the background using a well-established threshold method based on the optical density of the RGB channels

of the WSI [42]. For each colour channel c , we calculated the optical density d_c across the set of all pixels \mathbf{P} , according to:

$$d_c(\mathbf{P}) = \log_{10} \frac{I_c(p)}{I_{max}} \quad (4.1)$$

where $I_c(p)$ is the intensity of pixel p in channel c and I_{max} is the maximum intensity of the channel, which is 255 with 8-bit quantisation. We removed the background by thresholding the original image where $d_c(\mathbf{P}) < 0.2$. Figure 4.2 illustrates the resulted binary masks which containing only the tissue regions of the WSI.

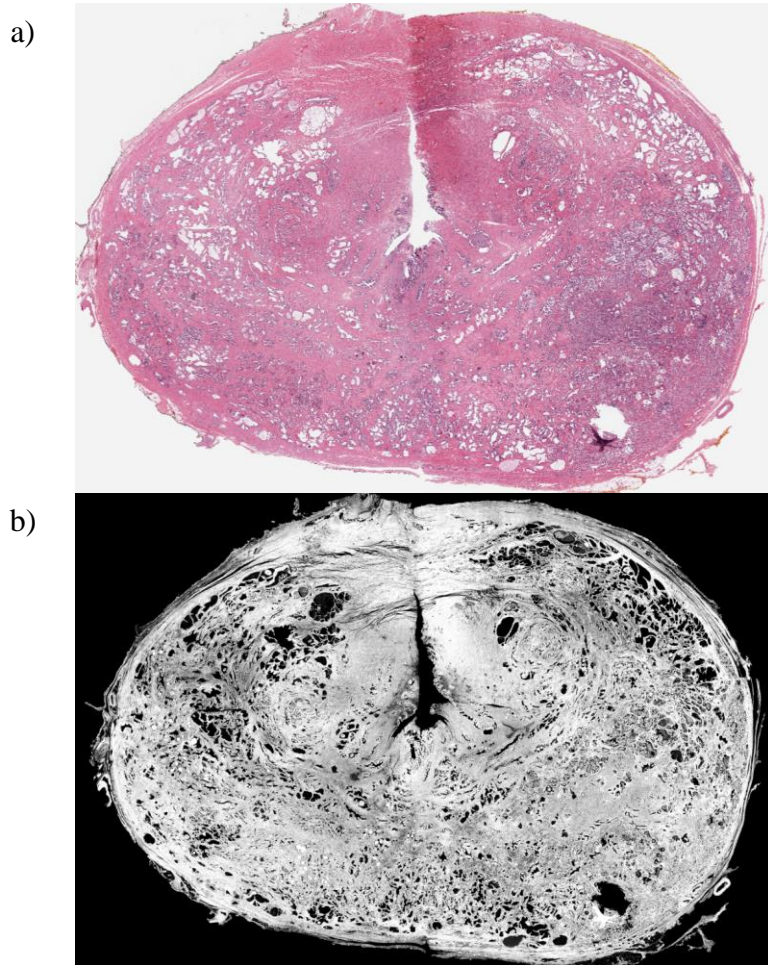


Figure 4.2: Histopathology slides with background information removed. a) The original tumour containing histopathology slide of prostate. b) The binary masks containing only tissue regions of the histopathology slide.

Due to the size of each WSI image (up to several gigapixels), we randomly extracted patches of size 256×256 pixels from WSIs with background removed. Tumour patches were those where $\geq 50\%$ of the total pixels belonged to the tumour regions while non-tumour patches were those where none of the pixels belonged to the tumour. To avoid bias in the training dataset, we selected an equal number of tumour and non-tumour patches from the entire set; this prevented oversampling predominant non-tumour patches. In total, we obtained 272540 patches from the lymph node metastases training set and 22086 patches from the validation set. For the prostate cancer dataset, we obtained a total of 4196 patches from the training slides and 274 patches from the test slides.

4.4 CNN Training and Fine-Tuning

We chose the GoogLeNet CNN for our method due to its state-of-the-art performance in WSI cancer detection in the Camelyon 16 challenge [159]. We modified the input layer to fit the size of the extracted image patches. Our pre-training on the lymph node patch dataset (denoted LN) involved finding the set of CNN layer weights W_{LN} that minimised the CNN's empirical loss:

$$L(W_{LN}, X_{LN}) = \frac{1}{n} \sum_{i=1}^n l(\text{GOOG}(\mathbf{x}_i, W_{LN}), C) \quad (4.2)$$

where \mathbf{x}_i is an image from the set of n images in X_{LN} , $\text{GOOG}()$ is a function that represents the GoogLeNet architecture to predict the class of \mathbf{x}_i , given the set of weights W_{LN} . C is the ground truth annotation for \mathbf{x}_i , and $l()$ is a penalty function for the class prediction, set to the logistic loss in our research. We determined the weights iteratively using mini-batch stochastic gradient descent for 172 epochs (the point at which the model converged) so that $W_{LN} = W_{172}$ given:

$$\mathbf{W}_{t+1} = \mathbf{W}_t + \eta \left[\alpha \Delta \mathbf{W}_t - \frac{\delta L(\mathbf{W}_t, \mathbf{X})}{\delta \mathbf{W}_t} - \lambda \mathbf{W}_t \right] \quad (4.3)$$

where $\Delta \mathbf{W}_t = \mathbf{W}_t - \mathbf{W}_{t-1}$ denotes the weight update from the previous iteration. The batch \mathbf{X} (a subset of \mathbf{X}_{LN}) contained 64 samples, coefficient $\eta = 1 \times 10^{-4}$ was the learning rate that controlled the size of the weight updates, and $\alpha = 0.9$ was the momentum that diminished irregular fluctuations in weight changes across consecutive iterations. In our formulation, \mathbf{W}_0 was randomly initialized from a uniform distribution with $\Delta \mathbf{W}_0 = 0$.

Our fine-tuning stage followed a similar process, substituting the prostate patch dataset (denoted PR) instead of the lymph node patch dataset, replacing \mathbf{W}_{LN} with \mathbf{W}_{PR} and \mathbf{X}_{LN} with \mathbf{X}_{PR} in Equation 4.2. Fine-tuning was iterated for 200 epochs (which is when the model converged) so that $\mathbf{W}_{PR} = \mathbf{W}_{200}$ given initial weights from the pre-trained model ($\mathbf{W}_0 = \mathbf{W}_{LN}$ and $\Delta \mathbf{W}_0 = 0$).

4.5 Evaluation Strategy

We compared the effectiveness of our system with CNNs which was trained from scratch using the prostate dataset. We trained the baseline CNN and our domain-adapted CNN with different numbers of prostate patches: 0, 1000, 2500 and 4196 patches.

The evaluation scheme comprises:

- Measure the performance of our system in predicting image patches (274 patches from prostate tumour WSIs) to contain tumour cells, compared with scratch-trained CNNs. This scheme reflects the ability of our system to extract image features that can differentiate tumour and non-tumour tissues in WSIs.

- Compute the area under receiver operating characteristics (AUC) [160] to compare the quality of the proposed system with traditional scratch-trained CNNs, which reflects the sensitivity (true-positive rate) and specificity (true-negative rate).

4.6 Results

Figure 4.3 illustrates the classification accuracy as it relates to the underlying size of the training/fine-tuning dataset; numerical results are provided in Table 4.1. Our domain adaptation facilitated system (highlighted in orange colour) achieved higher classification accuracy when compared with the traditional scratch-trained GoogLeNet architecture. As expected, better classification accuracies were obtained from our proposed domain adaptation facilitated training strategy compared to the traditional scratched-trained framework.

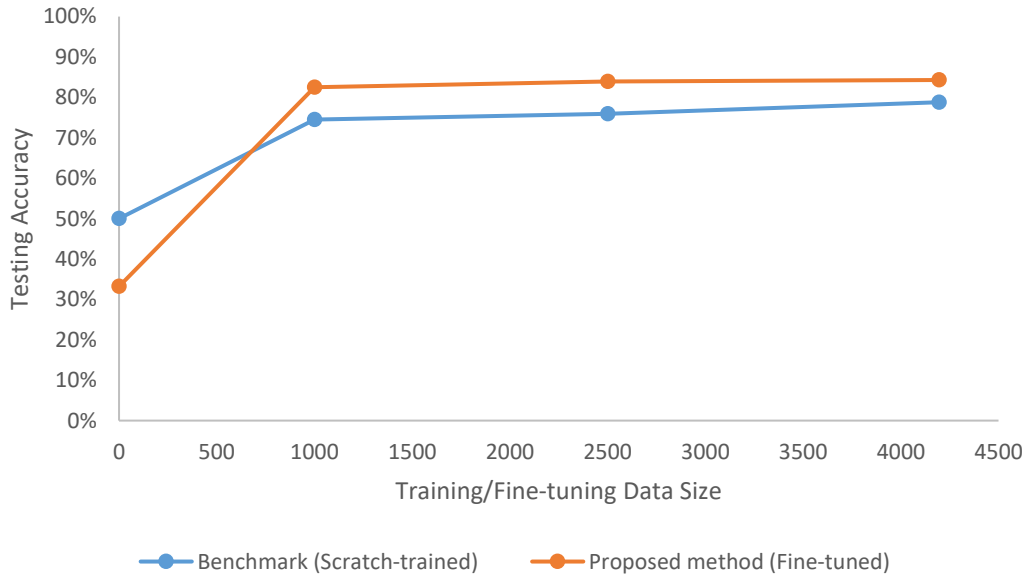


Figure 4.3: Accuracy of CNNs with various prostate cancer training size from fine-tuned and scratch-trained strategies.

Table 4.2 compares the AUC of the CNNs with different prostate cancer training dataset sizes. Our results show that our method has higher AUC values than the scratch-trained CNN. This demonstrates that our method enables GoogLeNet to perform cancer classification on WSIs with lower true/false positive rates across all three prostate cancer training sizes. It also suggests that the representation of WSIs derived from the fine-tuned CNN leverages the information from different tissue types to improve the specificity and sensitivity of cancer classification tasks.

Table 4.1: Test accuracy of CNNs from both training approaches for cancer classification task with different training size.

Prostate cancer training size	GoogLeNet: Scratch-trained (benchmark)	GoogLeNet: Fine-tuned (Proposed method)
0	50%	33.2%
1000	74.5%	82.5%
2500	75.9%	83.9%
4196	78.8%	84.3%

Table 4.2: Area under receiver operation characteristics (AUC) results of CNNs from both training approaches to perform cancer classification tasks with different training size.

Prostate cancer training size	GoogLeNet: Scratch-trained (benchmark)	GoogLeNet: Fine-tuned (Proposed method)
0	66.9%	27.3%
1000	79.9%	90.0%
2500	84.5%	91.9%
4196	86.7%	91.8%

4.7 Discussions and Summary

Our findings indicate that the proposed domain adaptation facilitated medical image analysis system improves the performance of CNNs for cancer classification tasks where training datasets are limited. At a dataset of 0 prostate images, the fine-tuned model (effectively trained only on lymph node metastases cancer) has an accuracy of 33.2%. This suggests that the lymph node feature space encodes some but not all of the information for classifying prostate images. In addition, the scratch-trained model at 0 prostate images (effectively untrained) appears to achieve an accuracy of 50%; this is equivalent to random guessing on a two-class dataset.

Our results also demonstrate that as the training dataset increases in size, the AUC and accuracy of the fine-tuned and scratch trained models to begin to converge; this is expected behavior from the employed CNNs, as given sufficiently larger training/fine-tuning datasets, each CNN would be able to learn image descriptors that are better representative of the specific data domain. As such, our findings suggest that our fine-tuning approach enables the classifier to reach its peak performance with considerably smaller training volume. This is attributed to its ability to learn subtle low-level differences between various tissue types [161], enhancing the ability to characterise and differentiate histopathology imaging data for different cancer types.

Our framework is not specific to particular CNN architecture or a particular cancer types. Hence we suggest that our approach has the potential to be extended with greater diversity in our training data and the adoption of various CNN models to support and improve the classification accuracy for multiple disease types.

The results indicate that our proposed domain adaptation facilitated medical image analysis system is capable of learning disease-specific image representations

that encode tumour phenotypic characteristics for tumour classification task; and overcame the reliance of large volumes of annotated datasets for the training process of CNN architectures. Our proposed framework allows automated image analysis to be applied to more disease types where large volumes of annotated medical image datasets are limited, and potentially contributes towards the advances in medical image analysis research.

Chapter 5

Deep Image-Genomics

This Chapter details our proposed deep domain adaptation learning framework for image-genomics analysis. Our work addresses the current gap in image-genomics research to extract additional, complementary image features that encode sophisticated abstract representations of tumour phenotypic characteristics for image-genomics association. We evaluate our proposed image-genomics framework with the current state-of-the-art in image-genomics analysis to compare the degrees of associations between image features and the tumour genetic information. Our results suggest that the proposed framework provides stable image features to encode tumour phenotypic characteristics. Results further indicate that the extracted image features offers stronger associations to tumour underlying genetic information and have the potential to identify imaging surrogates for tumour prognostic biomarkers.

5.1 Overview

Medical image feature extraction plays a crucial role in image-genomic research. Detailed in Chapter 3, one of the gaps in the current state-of-the-art in image-genomics research is the inherent reliance on traditional human-crafted features (agnostic and semantic) for associating to tumour genetic information.

In this thesis, we address this challenge by proposing a deep domain adaptation learning framework for associating image features to tumour genetic information. Our approach exploits the potential of domain adaptation and deep learning-based techniques to quantify additional image representations of tumour phenotypic characteristics from medical images with similar knowledge domain. Our framework also makes a number of modifications to the current state-of-the-art image-genomics approach to enable the extraction of abstract image representations, which encodes the subtle variations in tumour images.

5.2 Framework

Figure 5.1 outlines the crucial procedures of our proposed image-genomics framework, elements that are highlighted in the figure reflects our core contributions. The acquired medical image datasets and genetic information were processed with tumour sub-region delineations and gene expression normalisation [162] respectively. Our framework involves a widely adopted and computational efficient 2.5D representation of the tumours' physical centroids as the input for deep learning-based feature extraction [95]. Deep features were extracted through the employment of a pre-trained CNNs with domain adaptation techniques. Image-genomics associations were derived by calculating the Spearman's rank correlation between extracted image

features and the pre-defined gene sets, which are based on medical foreknowledge and are available from the Molecular Signatures Database C5 collection [163]. Image-genomics signatures were also visualised in the form of heat maps to reveal association patterns.

5.2.1 Dataset

We used the publicly available dataset from the Cancer Imaging Archive (TCIA) [112]. The dataset contains 89 histologically confirmed non-small cell lung cancer (NSCLC) patients: 29 women, 60 men, stages IA – IV with an age range 37- 85 years. All patients underwent a CT scan of the thorax/upper abdomen and a whole body 18F-FDG-PET-CT. The CT scans excluded tumour annotations. The CT scan slice thickness was between 1.5mm and 5mm. Genome-scale gene expression profiles were also acquired.

The three-dimensional tumour sub-regions of 69 patients (23 women, 46 men), were manually segmented by an experienced physician. The more detailed patient description can be found in Appendix Table A1. The remaining 20 patients from the dataset were excluded due to inconsistency of tumour locations that were specified between the specialist and the clinical metadata. The exclusion and segmentation were performed by a senior experienced medical imaging specialist who has read in excess of 80,000 PET-CT scans. The segmentation was done on transverse sections slice-by-slice, using lung windows, with an open source software tool (Medical Imaging Interaction Toolkit (MITK); version 2016.11) [164]. The segmented tumour sub-regions were used for feature extraction purposes. Gene expression was accessed via the Gene Expression Omnibus (GEO) [165].

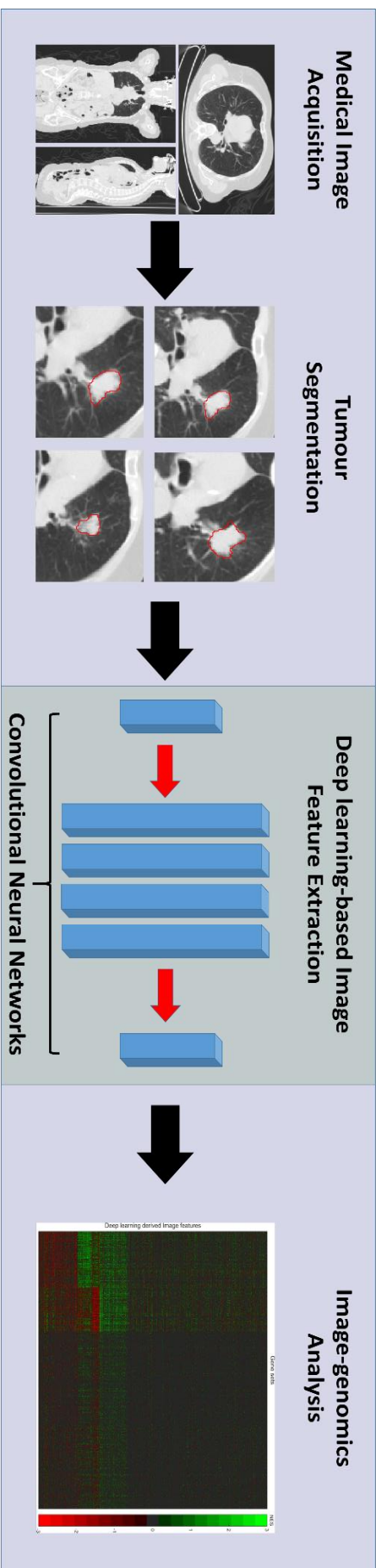


Figure 5.1: The proposed deep image-genomics framework. The deep learning-based image feature extraction is highlighted in gray to indicate the core contribution of our work.

5.2.2 Tumour Image Feature Extraction

In this study, we proposed a deep learning-based medical image feature extraction scheme to quantify tumour phenotypic descriptors that are complementary to human-crafted image features. Facilitated by the domain adaptation technique and the employment of CNNs, our framework extracted additional abstract image representations that were learned from images with similar domains. We also extracted human-crafted image features for the purpose of experimental comparisons. The following subsections describe the approaches and methods for the quantification of deep image features.

Deep Features: Deep features were extracted from a pre-trained CNN through the utilisation of domain adaptation techniques. We employed a 101-layer deep residual network (ResNet-101) that was trained on ImageNet ILSVRC challenge data. The pre-trained ResNet-101 was chosen as the feature extractor because of its performance in natural image classification with 3.57% error rate on the ImageNet test set [78, 83]. The top-5 classification error(s) on the validation dataset was 7.1%. The hyper-parameters of ResNet-101 comprise the following: input size of 224×224 , weight decay of 1×10^{-4} , the momentum of 0.9 with the initial learning rate of 0.1. During the training phase, the learning rate was decreased by a factor of 10 after the error-rate plateaued. Our implementation used the MatConvNet toolkit [166] and the pre-trained CNNs models that were provided in its model library.

We used a 2.5D representation of the CT data by resampling the axial, sagittal and coronal views at the physical centroid coordinates of the tumour; this was utilised as the input to the pre-trained ResNet-101 model, shown in Figure 4.2. For each of the views, the gray values were normalised to $[0, 255]$ using a linear transformation due

to 8-bit quantisation. All three input slices were then reduced to 224×224 with nearest-neighbour interpolation as well as zeros padding to preserve the aspect ratio of the tumour. Deep features were then derived through the forward propagation of the network with different input images, with final features extracted from the pooling layer named "pool5". We obtained a total of 6144 image features (2048 for each of the three planar views) from the CT volumes of each patient. We removed from the deep feature vector of all 'non-meaningful features', i.e., those with a median absolute deviation (MAD) of 0 across all patients (meaning that they did not vary across patients). After this removal, 4436 image features remained for each patient.

Human-defined Features: We extracted 431 human-crafted image features from CT volumes across all patients; their descriptions can be found in the Appendix. The human-crafted image features consist the following sub-categories, based on their properties: (a) first-order statistics that describe the distribution of voxel intensities, (b) shape and size descriptors of tumour geometric 3D characteristics such as compactness and surface area, (c) textural features which represent the spatial distribution of the voxel intensities, and (d) wavelets features, where the wavelet decomposition was applied to the raw image to derive first order statistics and textual features for all eight filter channels. Textural features were extracted using gray-level

co-occurrence matrix (GLCM) and gray-level run length matrix (GLRLM) toolboxes in Matlab 2016b.

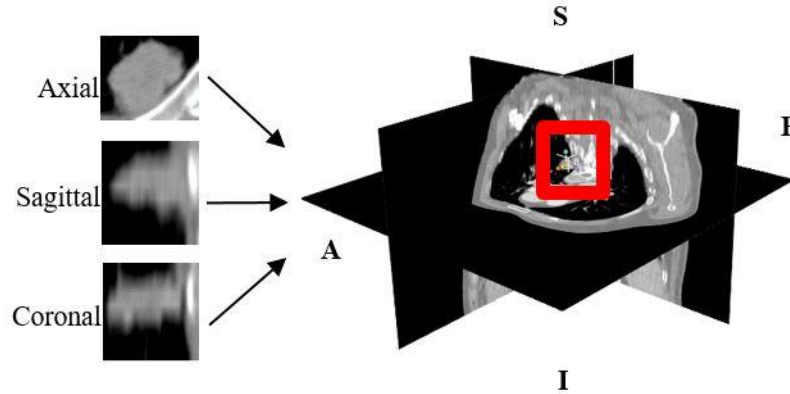


Figure 5.2: The 2.5D representation of a tumour physical centroid with axial, sagittal and coronal slices of the image volume. The orientation of the CT volume are illustrated on the figure: anterior (A), superior (S), posterior (P) and inferior (I).

5.3.3 Image-genomic Association and Enrichment Analysis

Our framework employed the following techniques: (i) Spearman’s rank correlation and; (ii) Gene Set Enrichment Analysis (GSEA) [167] to associate image features with patient-specific genetic information. We employed the pre-defined gene sets from the C5 collection of MsigDB that are based on specific gene ontology term (GO term), which denotes one of three categories: molecular function, cellular component or biological process in the experiment. Spearman’s rank correlation offered a robust estimation of the monotonic relationship between variables and was performed on image feature vectors and normalised gene expression to produce a pre-ranked gene list [168]. Genes from the pre-ranked gene list that had p-values < 0.001 according to the Spearman’s rank correlation were used to derive image-genomics associations. We then performed a pre-ranked version of GSEA to derive image-genomics signatures with gene sets. We followed the standard protocol by utilising gene sets that have

between 15 and 500 contributing genes to calculate the normalised enrichment score (*NES*) as follows:

$$P_{hit}(S, i) = \sum_{\substack{g_j \in S \\ j \leq i}} \frac{|r_j|}{N_R}, \text{ where } N_R = \sum_{g_j \in S} |r_j| \quad (5.1)$$

$$P_{miss}(S, i) = \sum_{\substack{g_j \notin S \\ j \leq i}} \frac{|r_j|}{N_R} \quad (5.2)$$

where the pre-ranked gene list contains genes g_1, \dots, g_N that were sorted based on the Spearman correlation values r , where $r(g_j) = r_j$. The foreknowledge defined gene set S contains N_H genes. The enrichment score (*ES*) reflects the degree to which a gene set S is over-represented at both extremes of the pre-ranked gene list, and is then defined as the maximum deviation from zero of $P_{hit} - P_{miss}$. The normalised *ES* (*NES*) is calculated through the rescaling of *ES* by dividing by the means of 1000 permutation of *ES*, where the permuted *ES* > 0 . Ultimately, *NES* is an indicator of the effectiveness of image features to represent a collection of genes in the pre-defined C5 gene sets.

5.3.4 Evaluation Strategy

The evaluation scheme for our experiment was designed to provide a direct comparison with the current state-of-the-art image-genomics approach. The proposed evaluation strategy consists of measuring of the peak and overall degrees of enrichment of image-genomics associations for deep features as follows:

- Quantification of the effectiveness of deep features in capturing tumour phenotypic representations. This is achieved by calculating the association between deep features and tumour clinical parameters such as primary tumour

stage with the χ^2 test. We compared our results with image-genomics associations that were derived using the 431 traditional human-crafted features.

- The measurement of the maximum, minimum and range of *NES* across all gene sets to evaluate the peak performance of image-genomics associations from both feature extraction approaches. We performed a two-sample *t*-test between deep learned and human-crafted image-genomics associations by comparing the absolute *NES* to evaluate the degree of enrichment across all gene set.
- The employment of χ^2 test to examine the association between deep features and the distribution of gene sets which contains key NSCLC prognostic biomarkers. The result indicates the strength of association between image-genomics signatures and the key prognostic biomarkers.

5.3 Results

5.3.1 Deep Features as Tumour Phenotype Descriptor

Our results demonstrate the associations between extracted deep features and tumour phenotypic characteristics in the form of heat map, shown in Figure 5.3. We employed unsupervised K-means clustering [169] to reveal three main clusters of patients with distinct association patterns. The effectiveness and stability of image features in capturing tumour phenotypic characteristics were measured by comparing the clusters with tumour clinical parameters (Fig. 5.3b), specifically with primary tumour (T-stage), lymph node (N-stage) and metastasis (M-stage), which are the standard indicator for lung cancer classification [170, 171]. Our results suggest a significant association to T-stage ($p < 6.8 \times 10^{-3}$, χ^2 test) where cluster I depicted greater

association with primary tumour stage pT4. However, image features that were extracted from our proposed framework did not show significant association with the N-stage and M-stage tumour phenotype ($p < 0.48$ and $p < 0.76$, respectively, χ^2 test). We found similar results from the 431 human-crafted image features, where primary tumour T-stage is significantly correlated to human-crafted image features ($p < 1.5 \times 10^{-3}$, χ^2 test) but correlates to neither tumour N-stage nor M-stage ($p < 0.68$ and $p < 0.82$ to N-stage and M-stage respectively, χ^2 test). Our findings indicate that both image feature extraction approaches share similar capabilities in capturing tumour phenotypic representations for our experimental dataset.

5.3.2 Deep Features Derived Image-genomics Signatures

The range of enrichment scores across image-genomics associations was used to investigate the extreme degrees of enrichment of the pair-wise association between extracted image features and pre-defined gene sets. Table 5.1 outlines the maximum, minimum and maximum range of *NES* from image-genomics associations that have a false discovery rate that is smaller than the standard threshold of 0.2. Image-genomics associations that were derived from deep features are shown to be better enriched to ontology gene sets compared to human-crafted image features. The range of *NES* for each image feature reflects the generalisability of the image representations to associate with different genetic information. Human-crafted and deep learning derived image features exhibits a maximum *NES* range of 8.22 and 8.46 respectively. Of the 1046 gene sets, 997 gene sets were shown to have a wider range of enrichment degrees when associated with deep features.

We also compared the overall performance of image-genomics association for image features through the employment of two-sample *t*-test on the absolute *NES*, as shown in Table 5.2. Image-genomics associations that were derived from deep features demonstrate improved degrees of enrichment in 518 of the total 1046 gene sets compare to the traditional human-crafted image features.

The association between gene expression and deep features are illustrated in Figure 5.4. Unsupervised K-means clustering reveals the three main gene set clusters and five image feature clusters with the distinct association to image features. The distribution of gene sets which contains key prognostic biomarker genes are shown in Figure 5.4b: (i) epidermal growth factor receptor (EGFR) (ii) ERCC1 and (iii) RRM1. Image-genomics signatures are shown to be strongly associated with key prognostic biomarkers ($p < 1.3 \times 10^{-5}$, χ^2 test). Gene set clusters I and III are observed to be more associated with ERCC1, while cluster II and III are associated with biomarker RRM1.

Table 5.1: Statistical analysis of radiogenomic signatures from human-crafted and deep learning derived image features.

Statistical analysis items	Human-crafted features	Deep features
Maximum NES	4.36	4.53
Minimum NES	-4.56	-4.78
Max NES Range (Single feature)	8.22	8.46

5.4 Discussions and Summary

Our results indicate that our proposed deep domain adaptation learning framework provides additional image features that capture tumour phenotypic characteristics, and is indicated by the significant association between deep features and tumour primary stage ($p < 6.8 \times 10^{-3}$, χ^2 test). However, deep features did not show a significant association to tumour N-stage nor M-stage. We found similar performance from human-crafted image features, which also showed significant association with primary tumour stage (T-stage) only. Neither of the image features to have a significant association with tumour N-stage nor M-stages, which is potentially attributed to the size of experiment dataset, where only a small portion of the patient cohort exhibits an N-stage or M-stage.

Table 5.2: Right-tailed two-sample t-test result summarisation of deep-learnt and human-crafted features

Deep features vs. Human-crafted features	Results
Signatures with p-values < 0.05	518
Total Valid Gene Set:	1046
Better Enriched Gene Sets (%)	49.6%
Two-sample t-test on <i>NES</i> range	
Gene sets with wider range	997
Gene sets with wider range (%)	95.3%

Our results also suggest that image-genomics associations were better enriched with deep features than to human-crafted image feature. Of the 1046 ontology gene sets, deep features have shown superior genetic representation across all gene sets compare to human-crafted image features. As shown in Table 5.1, the extreme degree of enrichment for both positive and negative correlations was improved through the utilisation of deep features. This finding suggests that the proposed framework extracts image features that are more capable of capturing the subtle visual variations from tumour medical images.

Further, our results also demonstrate that deep features exhibit stronger image-genomics associations to ontology gene set, as shown in Table 5.2. This was investigated by assessing the range and the absolute means of *NES* across all gene sets. Of the 1046 gene sets, more than 95% of the gene sets have a wider *NES* range when associated with deep features, compared to human-crafted features. Our results reflect that deep features-derived image-genomics associations are more representative to tumour ontology gene sets, compare to human-crafted image features. Our result hence suggests that the subtle visual representations that are encoded in deep features have the potential to better reflect tumour genetic information with the same experimental dataset, compare to the human-crafted image object descriptors.

The evaluation of the overall performance of image-genomics associations was shown in the form of *t*-test statistics. Our result indicates that 518 of the total 1046 gene set have improved absolute mean of *NES* with deep features. This observation illustrates that 49% of the gene sets are better associated with deep features compared to human-crafted image features. Our result further suggests that the deep features

have the potential to encode additional tumour phenotypic characteristics that exhibit strongly associations with tumour ontology gene sets.

Furthermore, the unsupervised K-means clustering of gene sets reveals distinct patterns with deep features in Figure 5.4. The distribution of gene sets that contain key prognostic biomarkers is shown to be associated with the clusters: cluster II is highly associated to prognostic biomarker EGFR, and cluster II and III contain prognostic biomarker gene RRM1 exclusively. This suggests that deep feature derived image-genomics associations comprise a distinctive association pattern to key prognostic biomarkers in the form of gene sets. Unfortunately, such findings require further validation from clinical contexts that needs to be explored for future studies.

Our proposed framework is not bound to specific CNN architecture or genetic information analysis protocols, and is sufficiently flexible to cater for other medical image modalities. As such, our framework can adapt to newer and improved CNNs architectures, thereby allowing for even better quantification of tumour visual characteristics from the medical image data.

Lastly, the deep domain adaptation learning framework for image-genomics analysis offers unique image-genomics associations when compared with the traditional human-crafted features. Our work showed improved performances in deriving image-genomics associations compared with the traditional human-crafted features. Our framework also demonstrated the potential for the extraction of imaging surrogates to tumour prognostic biomarkers. Our proposed framework allows a more comprehensive quantification of tumour phenotypic characteristics from medical images which exhibit potentials to be associated with tumour genetic information. Our work advances the current state-of-the-art and facilitates future research in identifying

image surrogates for tumour biomarkers, which has the potential to contribute towards accurate and non-invasive cancer diagnosis and precision medicine for patient-specific treatment.

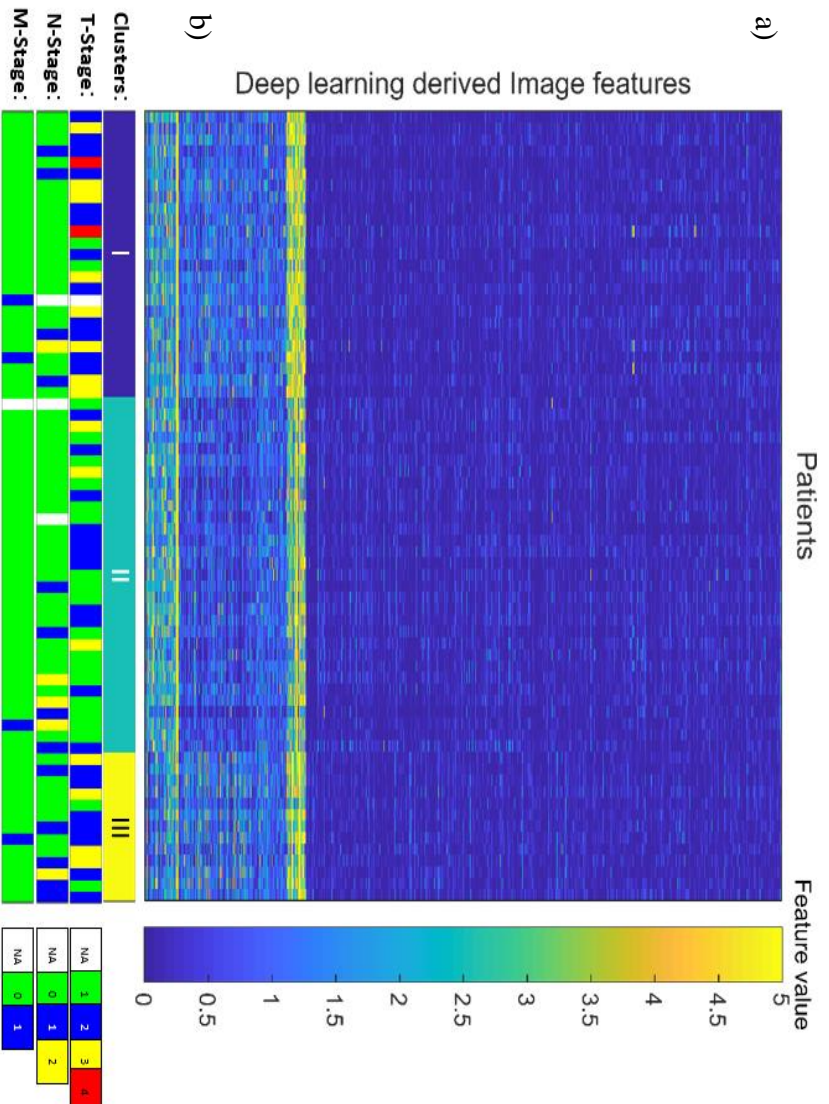


Figure 5.3: Association heat map between deep learning derived image features and tumour phenotype. (a) Unsupervised clustering of NSCLC patients (n=69) on the x axis and deep learning derived image features on the y axis. Clinical parameters demonstrating the association pattern between image features and tumour phenotypes. (b) Correspondence of clustered tumour phenotypes with association patterns.

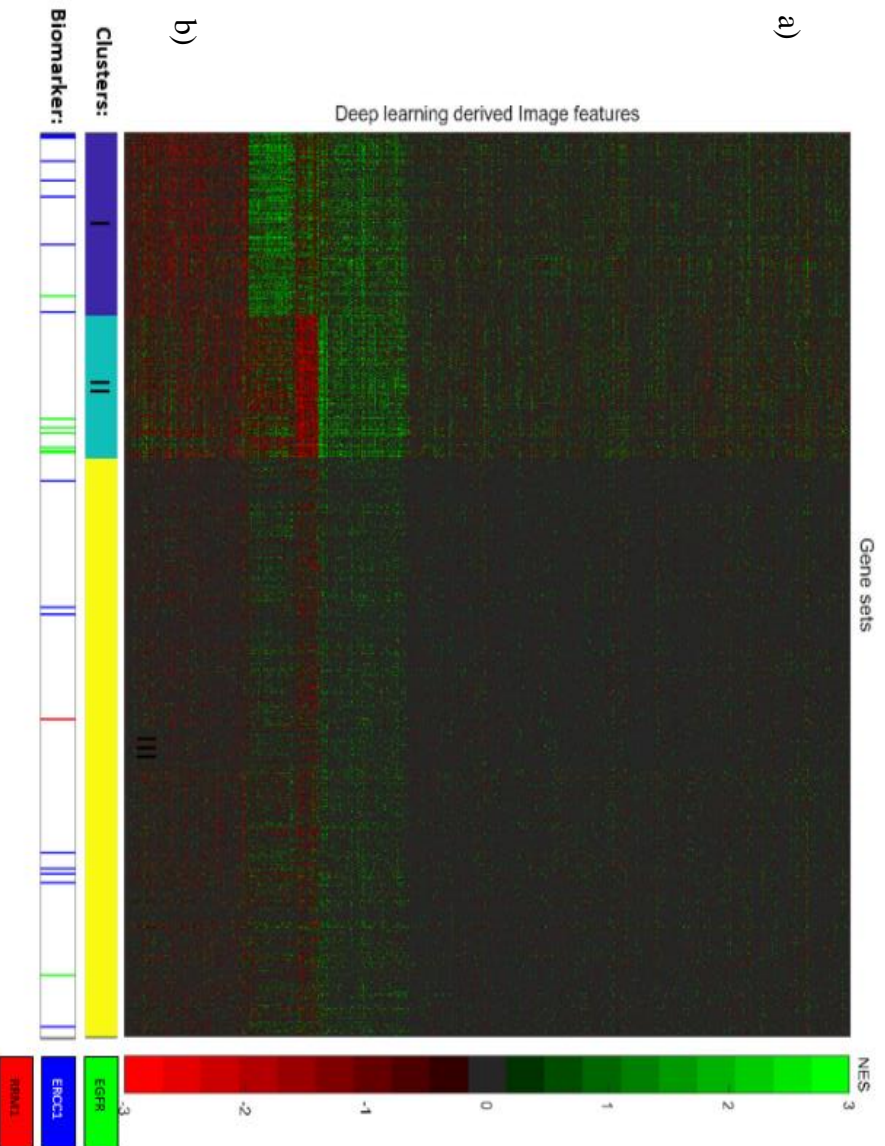


Figure 5.4: Image-genomics association between deep image features and oncology gene sets using gene-set enrichment analysis (GSEA). (a) Performance of image-genomics association is quantified and color-coded by the NES value. (b) Distribution of key NSCLC prognostic biomarker contributing oncology gene sets.

Chapter 6

Conclusion and Future work

This chapter highlights the capabilities of our proposed deep domain adaptation learning framework for image-genomics analysis, based on our experiments, together with the challenges that remain to be investigated. Further discussions on the future directions and limitations of our studies will also be addressed in the following subsections.

6.1 Conclusion

In this thesis, we presented novel deep domain adaptation learning framework for image-genomics analysis to improve image-genomics association. Our proposed framework emphasised the employment of domain adaptation of deep learning in image-genomics research to offer additional deep image features that encode abstract image representations of subtle variation in tumour medical images. Our domain

adaptation approach reduces the dependency on large volumes of annotated medical image dataset for deep learning models to learn the tumour image representation.

Our results demonstrated that domain adaptation technique facilitates the learning of image representations of abnormalities in medical images by extending or refining existing knowledge from similar domains. The quantification of such image features hence improves the accuracy of tumour image classification tasks, compared to traditional approaches where deep learning models were trained from scratch on limited volumes of domain-specific datasets. Our results also demonstrated that the proposed framework offers additional deep image features to encode abstract representation of tumour phenotypic characteristics which exhibit stronger associations to patient-specific genetic information, compared to human-crafted image features.

Our findings highlight the potential of deep domain adaptation learning framework in the field of image-genomics. With advances in imaging techniques and computation tools, the future employment of domain adaptation and deep learning-based approaches would offer more opportunities to reveal stronger image-genomics associations and identify imaging surrogates for tumour genetic biomarkers.

This thesis has potentially contributed toward the research field of image-genomics by allowing automated image analysis to be applied to more disease types where large volumes of annotated medical image datasets are limited. Our proposed deep domain adaptation learning framework also has the potential to contribute towards accurate and non-invasive cancer diagnosis and hence the clinical application of precision medicine for patient-specific treatment in the near future.

6.2 Future Research Directions

A number of new techniques and tools are now emerging to offer more opportunities in the field of image-genomics, e.g. advances in imaging agents prostate-specific membrane antigen (PSMA) PET/CT scans [172]. Image features that are extracted from our proposed domain adaptation scheme have the potential to be further fine-tuned. As described in Chapter 2 and 4, fine-tuning allows the image representation to be more optimised to the dataset, which may further improve image-genomics associations with gene expressions.

In addition, detailed clinical description with more characteristics of patients would also help future studies in the training of CNNs model, feature selection and to refine radiogenomics signatures to specific clinical parameters. Improved image feature selection could avoid over-fitting and hence reduce the number of false-positive radiogenomic signatures or vice versa.

Furthermore, patients can also be highly heterogeneous in many aspects: age, gender, hereditary, cell type and cancer stage. Those can be potentially reflected on their medical images and hence accounted for the image feature extraction process. With more clinical data made public in the future, future image-genomics framework has the potential to take account of patients' variability into the design of CNN architecture, image feature selection and to refine image-genomics associations. Such techniques allow more applications of image-genomics and open the door to greater clinical impact such as clinical trials.

References

- [1] F. S. Collins and H. Varmus, "A new initiative on precision medicine," *New England Journal of Medicine*, vol. 372, pp. 793-795, 2015.
- [2] P. L. Bedard, A. R. Hansen, M. J. Ratain, and L. L. Siu, "Tumour heterogeneity in the clinic," *Nature*, vol. 501, p. 355, 2013.
- [3] C. A. Karlo, P. L. Di Paolo, J. Chaim, A. A. Hakimi, I. Ostrovnaya, P. Russo, *et al.*, "Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations," *Radiology*, vol. 270, pp. 464-471, 2014.
- [4] P. M. Evans, "Anatomical imaging for radiotherapy," *Physics in Medicine & Biology*, vol. 53, p. R151, 2008.
- [5] A. M. Rutman and M. D. Kuo, "Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging," *European journal of radiology*, vol. 70, pp. 232-241, 2009.
- [6] H. Z. Sailem and C. Bakal, "Identification of clinically predictive metagenes that encode components of a network coupling cell shape to transcription by image-omics," *Genome research*, vol. 27, pp. 196-207, 2017.
- [7] A. A. Friedman, A. Letai, D. E. Fisher, and K. T. Flaherty, "Precision medicine for cancer with next-generation functional diagnostics," *Nature Reviews Cancer*, vol. 15, p. 747, 2015.
- [8] J. L. Jameson and D. L. Longo, "Precision medicine—personalized, problematic, and promising," *Obstetrical & Gynecological Survey*, vol. 70, pp. 612-614, 2015.
- [9] M. Schwaederle, M. Zhao, J. J. Lee, A. M. Eggermont, R. L. Schilsky, J. Mendelsohn, *et al.*, "Impact of precision medicine in diverse cancers: a meta-analysis of phase II clinical trials," *Journal of clinical oncology*, vol. 33, pp. 3817-3825, 2015.
- [10] D. A. Fisher, L. L. Zullig, S. C. Grambow, D. H. Abbott, R. S. Sandler, R. H. Fletcher, *et al.*, "Determinants of medical system delay in the diagnosis of colorectal cancer within the Veteran Affairs Health System," *Digestive diseases and sciences*, vol. 55, pp. 1434-1441, 2010.
- [11] A. L. Stanton and P. R. Snider, "Coping with a breast cancer diagnosis: A prospective study," *Health Psychology*, vol. 12, p. 16, 1993.

- [12] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R. G. van Stiphout, P. Granton, *et al.*, "Radiomics: extracting more information from medical images using advanced feature analysis," *European journal of cancer*, vol. 48, pp. 441-446, 2012.
- [13] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Physics in medicine & biology*, vol. 46, p. R1, 2001.
- [14] K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," *Computerized medical imaging and graphics*, vol. 31, pp. 198-211, 2007.
- [15] E.-J. Yeoh, M. E. Ross, S. A. Shurtleff, W. K. Williams, D. Patel, R. Mahfouz, *et al.*, "Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling," *Cancer cell*, vol. 1, pp. 133-143, 2002.
- [16] L. J. Van't Veer, H. Dai, M. J. Van De Vijver, Y. D. He, A. A. Hart, M. Mao, *et al.*, "Gene expression profiling predicts clinical outcome of breast cancer," *nature*, vol. 415, p. 530, 2002.
- [17] C. R. King, "Patterns of prostate cancer biopsy grading: trends and clinical implications," *International journal of cancer*, vol. 90, pp. 305-311, 2000.
- [18] B. L. Baisden, H. Kahane, and J. I. Epstein, "Perineural invasion, mucinous fibroplasia, and glomerulations: diagnostic features of limited cancer on prostate needle biopsy," *The American journal of surgical pathology*, vol. 23, p. 918, 1999.
- [19] B. Bain, "Bone marrow biopsy morbidity and mortality: 2002 data," *International Journal of Laboratory Hematology*, vol. 26, pp. 315-318, 2004.
- [20] H. Su, Y. Shen, F. Xing, X. Qi, K. M. Hirshfield, L. Yang, *et al.*, "Robust automatic breast cancer staging using a combination of functional genomics and image-omics," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 2015, pp. 7226-7229.
- [21] A. Madabhushi and G. Lee, "Image analysis and machine learning in digital pathology: Challenges and opportunities," *Medical image analysis*, vol. 33, pp. 170-175, 2016.
- [22] H. J. Aerts, E. R. Velazquez, R. T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, *et al.*, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature communications*, vol. 5, 2014.
- [23] O. Gevaert, J. Xu, C. D. Hoang, A. N. Leung, Y. Xu, A. Quon, *et al.*, "Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data—methods and preliminary results," *Radiology*, vol. 264, pp. 387-396, 2012.
- [24] P. Kickingereder, F. Sahm, A. Radbruch, W. Wick, S. Heiland, A. Von Deimling, *et al.*, "IDH mutation status is associated with a distinct hypoxia/angiogenesis transcriptome signature which is non-invasively predictable with rCBV imaging in human glioma," *Scientific reports*, vol. 5, p. 16238, 2015.
- [25] S. Banerjee, D. S. Wang, H. J. Kim, C. B. Sirlin, M. G. Chan, R. L. Korn, *et al.*, "A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma," *Hepatology*, vol. 62, pp. 792-800, 2015.

- [26] M. D. Kuo, J. Gollub, C. B. Sirlin, C. Ooi, and X. Chen, "Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma," *Journal of Vascular and Interventional Radiology*, vol. 18, pp. 821-830, 2007.
- [27] Y. Bengio, "Learning deep architectures for AI," *Foundations and trends® in Machine Learning*, vol. 2, pp. 1-127, 2009.
- [28] B. S. Rosenstein, C. M. West, S. M. Bentzen, J. Alsner, C. N. Andreassen, D. Azria, *et al.*, "Radiogenomics: radiobiology enters the era of big data and team science," *International Journal of Radiation Oncology• Biology• Physics*, vol. 89, pp. 709-713, 2014.
- [29] A. S. Takhar, P. Palaniappan, R. Dhingsa, and D. N. Lobo, "Recent developments in diagnosis of pancreatic cancer," *BMJ: British Medical Journal*, vol. 329, p. 668, 2004.
- [30] K. Horsthuis, P. C. Stokkers, and J. Stoker, "Detection of inflammatory bowel disease: diagnostic performance of cross-sectional imaging modalities," *Abdominal imaging*, vol. 33, pp. 407-416, 2008.
- [31] D. W. Townsend and S. R. Cherry, "Combining anatomy and function: the path to true image fusion," *European radiology*, vol. 11, pp. 1968-1974, 2001.
- [32] R. D. Pascual-Marqui, M. Esslen, K. Kochi, and D. Lehmann, "Functional imaging with low-resolution brain electromagnetic tomography (LORETA): a review," *Methods and findings in experimental and clinical pharmacology*, vol. 24, pp. 91-95, 2002.
- [33] S. R. Cherry, "Multimodality imaging: Beyond pet/ct and spect/ct," in *Seminars in nuclear medicine*, 2009, pp. 348-353.
- [34] S. Sone, S. Takashima, F. Li, Z. Yang, T. Honda, Y. Maruyama, *et al.*, "Mass screening for lung cancer with mobile spiral computed tomography scanner," *The Lancet*, vol. 351, pp. 1242-1245, 1998.
- [35] I. E. L. C. A. P. Investigators, "Survival of patients with stage I lung cancer detected on CT screening," *New England Journal of Medicine*, vol. 355, pp. 1763-1771, 2006.
- [36] A. H. Fischer, K. A. Jacobson, J. Rose, and R. Zeller, "Hematoxylin and eosin staining of tissue and cell sections," *Cold Spring Harbor Protocols*, vol. 2008, p. pdb. prot4986, 2008.
- [37] M. J. Reardon, J.-C. Walkes, and R. Benjamin, "Therapy insight: malignant primary cardiac tumors," *Nature Reviews Cardiology*, vol. 3, p. 548, 2006.
- [38] S. Kothari, J. H. Phan, T. H. Stokes, and M. D. Wang, "Pathology imaging informatics for quantitative analysis of whole-slide images," *Journal of the American Medical Informatics Association*, vol. 20, pp. 1099-1108, 2013.
- [39] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annual review of biomedical engineering*, vol. 2, pp. 315-337, 2000.
- [40] R. A. Poldrack, "Region of interest analysis for fMRI," *Social Cognitive and Affective Neuroscience*, vol. 2, pp. 67-70, March 1, 2007 2007.
- [41] S. S. Raab, D. M. Grzybicki, J. E. Janosky, R. J. Zarbo, F. A. Meier, C. Jensen, *et al.*, "Clinical impact and frequency of anatomic pathology errors in cancer diagnoses," *Cancer*, vol. 104, pp. 2205-2213, 2005.

- [42] G. Litjens, C. I. Sánchez, N. Timofeeva, M. Hermsen, I. Nagtegaal, I. Kovacs, *et al.*, "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis," *Scientific reports*, vol. 6, p. 26286, 2016.
- [43] S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," *IEEE transactions on medical imaging*, vol. 20, pp. 490-498, 2001.
- [44] J. Fan, D. K. Yau, A. K. Elmagarmid, and W. G. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," *IEEE transactions on image processing*, vol. 10, pp. 1454-1466, 2001.
- [45] J. Canny, "A computational approach to edge detection," in *Readings in Computer Vision*, ed: Elsevier, 1987, pp. 184-203.
- [46] K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, and T.-J. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," *computerized medical imaging and graphics*, vol. 30, pp. 9-15, 2006.
- [47] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, *et al.*, "Brain tumor segmentation with deep neural networks," *Medical image analysis*, vol. 35, pp. 18-31, 2017.
- [48] O. Ronneberger, 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, pp. 234-241.
- [49] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115-118, 2017.
- [50] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *Jama*, vol. 316, pp. 2402-2410, 2016.
- [51] J. O. Armitage, "A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma," *Blood*, vol. 89, pp. 3909-3918, 1997.
- [52] W. L. McGuire, "Breast cancer prognostic factors: evaluation guidelines," ed: Oxford University Press, 1991.
- [53] J. Khan, J. S. Wei, M. Ringner, L. H. Saal, M. Ladanyi, F. Westermann, *et al.*, "Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks," *Nature medicine*, vol. 7, p. 673, 2001.
- [54] A. M. Dunning, C. S. Healey, P. D. Pharoah, M. D. Teare, B. A. Ponder, and D. F. Easton, "A systematic review of genetic polymorphisms and breast cancer risk," *Cancer Epidemiology and Prevention Biomarkers*, vol. 8, pp. 843-854, 1999.
- [55] A. G. Knudson, "Mutation and cancer: statistical study of retinoblastoma," *Proceedings of the National Academy of Sciences*, vol. 68, pp. 820-823, 1971.
- [56] D. Ford, D. F. Easton, D. T. Bishop, S. A. Narod, and D. E. Goldgar, "Risks of cancer in BRCA1-mutation carriers," *The Lancet*, vol. 343, pp. 692-695, 1994.

- [57] S. Kobayashi, T. J. Boggon, T. Dayaram, P. A. Jänne, O. Kocher, M. Meyerson, *et al.*, "EGFR mutation and resistance of non-small-cell lung cancer to gefitinib," *New England Journal of Medicine*, vol. 352, pp. 786-792, 2005.
- [58] J. E. Celis, M. Kruhøffer, I. Gromova, C. Frederiksen, M. Østergaard, T. Thykjaer, *et al.*, "Gene expression profiling: monitoring transcription and translation products using DNA microarrays and proteomics," *FEBS letters*, vol. 480, pp. 2-16, 2000.
- [59] A. Marusyk, V. Almendro, and K. Polyak, "Intra-tumour heterogeneity: a looking glass for cancer?," *Nature Reviews Cancer*, vol. 12, p. 323, 2012.
- [60] M. R. Junttila and F. J. de Sauvage, "Influence of tumour micro-environment heterogeneity on therapeutic response," *Nature*, vol. 501, p. 346, 2013.
- [61] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, p. 436, 2015.
- [62] Y. Bengio, A. Courville, and G. Hinton, "Representation learning: a review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013.
- [63] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, *Machine learning: An artificial intelligence approach*: Springer Science & Business Media, 2013.
- [64] O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based image classification," *IEEE transactions on Neural Networks*, vol. 10, pp. 1055-1064, 1999.
- [65] C. Brodley, A. Kak, C. Shyu, J. Dy, L. Broderick, and A. M. Aisen, "Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision," in *AAAI/IAAI*, 1999, pp. 760-767.
- [66] J. Zhang, K.-K. Ma, M.-H. Er, and V. Chong, "Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine," in *International Workshop on Advanced Image Technology (IWAIT'04)*, 2004, pp. 207--211.
- [67] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [68] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerging artificial intelligence applications in computer engineering*, vol. 160, pp. 3-24, 2007.
- [69] G. Carneiro, A. B. Chan, P. J. Moreno, and N. Vasconcelos, "Supervised learning of semantic classes for image annotation and retrieval," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, pp. 394-410, 2007.
- [70] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, pp. 115-133, 1943.
- [71] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85-117, 2015.
- [72] L. Deng and D. Yu, "Deep learning: methods and applications," *Foundations and Trends® in Signal Processing*, vol. 7, pp. 197-387, 2014.

- [73] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-r. Mohamed, N. Jaitly, *et al.*, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Processing Magazine*, vol. 29, pp. 82-97, 2012.
- [74] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *The Journal of physiology*, vol. 160, pp. 106-154, 1962.
- [75] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *3D Vision (3DV), 2016 Fourth International Conference on*, 2016, pp. 565-571.
- [76] K. Sirinukunwattana, S. E. A. Raza, Y.-W. Tsang, D. R. Snead, I. A. Cree, and N. M. Rajpoot, "Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images," *IEEE transactions on medical imaging*, vol. 35, pp. 1196-1206, 2016.
- [77] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, *et al.*, "Going deeper with convolutions," 2015.
- [78] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, *et al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, pp. 211-252, 2015.
- [79] M. Lin, Q. Chen, and S. Yan, "Network in network. CoRR, abs/1312.4400 (2013)," ed.
- [80] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [81] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026-1034.
- [82] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*, 2015, pp. 448-456.
- [83] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
- [84] H. Jegou, F. Perronnin, M. Douze, J. Sánchez, P. Perez, and C. Schmid, "Aggregating local image descriptors into compact codes," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, pp. 1704-1716, 2012.
- [85] F. Perronnin and C. Dance, "Fisher kernels on visual vocabularies for image categorization," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, 2007, pp. 1-8.
- [86] C. M. Bishop, *Neural networks for pattern recognition*: Oxford university press, 1995.
- [87] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735-1780, 1997.
- [88] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway networks," *arXiv preprint arXiv:1505.00387*, 2015.
- [89] A. Kumar, J. Kim, D. Lyndon, M. Fulham, and D. Feng, "An ensemble of fine-tuned convolutional neural networks for medical image classification," *IEEE journal of biomedical and health informatics*, vol. 21, pp. 31-40, 2017.

- [90] P. V. Tran, "A fully convolutional neural network for cardiac segmentation in short-axis MRI," *arXiv preprint arXiv:1604.00494*, 2016.
- [91] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, pp. 1345-1359, 2010.
- [92] L. Y. Pratt, "Discriminability-based transfer between neural networks," in *Advances in neural information processing systems*, 1993, pp. 204-211.
- [93] S. Ben-David, J. Blitzer, K. Crammer, and F. Pereira, "Analysis of representations for domain adaptation," in *Advances in neural information processing systems*, 2007, pp. 137-144.
- [94] X. Glorot, A. Bordes, and Y. Bengio, "Domain adaptation for large-scale sentiment classification: A deep learning approach," in *Proceedings of the 28th international conference on machine learning (ICML-11)*, 2011, pp. 513-520.
- [95] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, *et al.*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, vol. 35, pp. 1285-1298, 2016.
- [96] B. van Ginneken, A. A. Setio, C. Jacobs, and F. Ciompi, "Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans," in *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, 2015, pp. 286-289.
- [97] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology detection using deep learning with non-medical training," in *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, 2015, pp. 294-297.
- [98] F. Ciompi, B. de Hoop, S. J. van Riel, K. Chung, E. T. Scholten, M. Oudkerk, *et al.*, "Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box," *Medical image analysis*, vol. 26, pp. 195-202, 2015.
- [99] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, *et al.*, "Convolutional neural networks for medical image analysis: Full training or fine tuning?," *IEEE transactions on medical imaging*, vol. 35, pp. 1299-1312, 2016.
- [100] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in *Advances in neural information processing systems*, 2014, pp. 3320-3328.
- [101] A. Kumar, P. Sridar, A. Quinton, R. K. Kumar, D. Feng, R. Nanani, *et al.*, "Plane identification in fetal ultrasound images using saliency maps and convolutional neural networks," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, 2016, pp. 791-794.
- [102] M. Vallières, C. R. Freeman, S. R. Skamene, and I. El Naqa, "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities," *Physics in Medicine & Biology*, vol. 60, p. 5471, 2015.
- [103] Y.-q. Huang, C.-h. Liang, L. He, J. Tian, C.-s. Liang, X. Chen, *et al.*, "Development and validation of a radiomics nomogram for preoperative

- prediction of lymph node metastasis in colorectal cancer," *Journal of Clinical Oncology*, vol. 34, pp. 2157-2164, 2016.
- [104] S. L. Topalian, C. G. Drake, and D. M. Pardoll, "Immune checkpoint blockade: a common denominator approach to cancer therapy," *Cancer cell*, vol. 27, pp. 450-461, 2015.
 - [105] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: images are more than pictures, they are data," *Radiology*, vol. 278, pp. 563-577, 2015.
 - [106] P. M. Thompson, N. G. Martin, and M. J. Wright, "Imaging genomics," *Current opinion in neurology*, vol. 23, p. 368, 2010.
 - [107] B. S. Rosenstein, C. M. West, S. M. Bentzen, J. Alsner, C. N. Andreassen, D. Azria, *et al.*, "Radiogenomics: radiobiology enters the era of big data and team science," *International journal of radiation oncology, biology, physics*, vol. 89, p. 709, 2014.
 - [108] V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, *et al.*, "Radiomics: The process and the challenges," *Magnetic Resonance Imaging*, vol. 30, pp. 1234-1248, 2012.
 - [109] V. S. Nair, O. Gevaert, G. Davidzon, S. K. Plevritis, and R. West, "NF- κ B protein expression associates with 18F-FDG PET tumor uptake in non-small cell lung cancer: A radiogenomics validation study to understand tumor metabolism," *Lung Cancer*, vol. 83, pp. 189-196, 2014.
 - [110] S. Yamamoto, D. D. Maki, R. L. Korn, and M. D. Kuo, "Radiogenomic analysis of breast cancer using MRI: a preliminary study to define the landscape," *American Journal of Roentgenology*, vol. 199, pp. 654-663, 2012.
 - [111] P. O. Zinn, B. Majadan, P. Sathyan, S. K. Singh, S. Majumder, F. A. Jolesz, *et al.*, "Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme," *PloS one*, vol. 6, p. e25451, 2011.
 - [112] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, *et al.*, "The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository," *Journal of digital imaging*, vol. 26, pp. 1045-1057, 2013.
 - [113] C. G. A. R. Network, "Comprehensive genomic characterization defines human glioblastoma genes and core pathways," *Nature*, vol. 455, p. 1061, 2008.
 - [114] J. H. Larkin and H. A. Simon, "Why a diagram is (sometimes) worth ten thousand words," *Cognitive science*, vol. 11, pp. 65-100, 1987.
 - [115] R. S. Choras, "Image feature extraction techniques and their applications for CBIR and biometrics systems," *International journal of biology and biomedical engineering*, vol. 1, pp. 6-16, 2007.
 - [116] A. Hyvarinen, E. Oja, P. Hoyer, and J. Hurri, "Image feature extraction by sparse coding and independent component analysis," in *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on*, 1998, pp. 1268-1273.
 - [117] J. Liu, Y. Y. Tang, and Y. Cao, "An evolutionary autonomous agents approach to image feature extraction," *IEEE Transactions on Evolutionary Computation*, vol. 1, pp. 141-158, 1997.

- [118] R. M. Haralick and K. Shanmugam, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, pp. 610-621, 1973.
- [119] X. Zhang, J. Cui, W. Wang, and C. Lin, "A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm," *Sensors*, vol. 17, p. 1474, 2017.
- [120] J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation," *International journal of computer vision*, vol. 43, pp. 7-27, 2001.
- [121] D. Wang and S. N. Srihari, "Classification of newspaper image blocks using texture analysis," *Computer Vision, Graphics, and Image Processing*, vol. 47, pp. 327-352, 1989.
- [122] A. Leone and C. Distanto, "Shadow detection for moving objects based on texture analysis," *Pattern Recognition*, vol. 40, pp. 1222-1233, 2007.
- [123] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern recognition*, vol. 40, pp. 262-282, 2007.
- [124] J. H. du Buf, M. Kardan, and M. Spann, "Texture feature performance for image segmentation," *Pattern recognition*, vol. 23, pp. 291-309, 1990.
- [125] D. Zhang, A. Wong, M. Indrawan, and G. Lu, "Content-based image retrieval using Gabor texture features," *IEEE Transactions PAMI*, pp. 13-15, 2000.
- [126] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern recognition*, vol. 24, pp. 1167-1186, 1991.
- [127] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 15, pp. 1186-1191, 1993.
- [128] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 11, pp. 674-693, 1989.
- [129] Y. Rui, T. S. Huang, and S.-F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *Journal of visual communication and image representation*, vol. 10, pp. 39-62, 1999.
- [130] Y. Deng, B. Manjunath, C. Kenney, M. S. Moore, and H. Shin, "An efficient color representation for image retrieval," *IEEE Transactions on image processing*, vol. 10, pp. 140-147, 2001.
- [131] M. A. Stricker and M. Orengo, "Similarity of color images," in *Storage and Retrieval for Image and Video Databases III*, 1995, pp. 381-393.
- [132] M. J. Swain and D. H. Ballard, "Color indexing," *International journal of computer vision*, vol. 7, pp. 11-32, 1991.
- [133] K. Van De Sande, T. Gevers, and C. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, pp. 1582-1596, 2010.
- [134] J. Fang and G. Qiu, "A colour histogram based approach to human face detection," 2003.
- [135] M. W. Mackiewicz, M. Fisher, and C. Jamieson, "Bleeding detection in wireless capsule endoscopy using adaptive colour histogram model and

- support vector classification," in *Medical Imaging 2008: Image Processing*, 2008, p. 69140R.
- [136] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, pp. 1349-1380, 2000.
 - [137] M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," ed: In-Tech, 2008.
 - [138] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, pp. 509-522, 2002.
 - [139] R. Basri, L. Costa, D. Geiger, and D. Jacobs, "Determining the similarity of deformable shapes," *Vision Research*, vol. 38, pp. 2365-2385, 1998.
 - [140] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing their shock graphs," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 26, pp. 550-571, 2004.
 - [141] H. Soltanian-Zadeh and F. Rafiee-Rad, "Comparison of multiwavelet, wavelet, Haralick, and shape features for microcalcification classification in mammograms," *Pattern recognition*, vol. 37, pp. 1973-1986, 2004.
 - [142] X. W. Gao, L. Podladchikova, D. Shaposhnikov, K. Hong, and N. Shevtsova, "Recognition of traffic signs based on their colour and shape features extracted using human vision models," *Journal of Visual Communication and Image Representation*, vol. 17, pp. 675-685, 2006.
 - [143] S. Brandt, J. Laaksonen, and E. Oja, "Statistical shape features for content-based image retrieval," *Journal of Mathematical Imaging and Vision*, vol. 17, pp. 187-198, 2002.
 - [144] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, *et al.*, "A search engine for 3D models," *ACM Transactions on Graphics (TOG)*, vol. 22, pp. 83-105, 2003.
 - [145] S. Berretti, A. Del Bimbo, and P. Pala, "Description, matching and retrieval by content of 3D objects," in *Digital Libraries: Research and Development*, ed: Springer, 2007, pp. 288-297.
 - [146] M. Ankerst, G. Kastenmüller, H.-P. Kriegel, and T. Seidl, "3D shape histograms for similarity search and classification in spatial databases," in *International Symposium on Spatial Databases*, 1999, pp. 207-226.
 - [147] J. W. Tangelder and R. C. Veltkamp, "A survey of content based 3D shape retrieval methods," in *Shape Modeling Applications, 2004. Proceedings*, 2004, pp. 145-156.
 - [148] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: Images Are More than Pictures, They Are Data," *Radiology*, vol. 278, pp. 563-77, 2016 Feb (Epub 2015 Nov 2016).
 - [149] E. Segal, C. B. Sirlin, C. Ooi, A. S. Adler, J. Gollub, X. Chen, *et al.*, "Decoding global gene expression programs in liver cancer by noninvasive imaging," *Nature biotechnology*, vol. 25, p. 675, 2007.
 - [150] R. L. Engle Jr, "Attempts to use computers as diagnostic aids in medical decision making: a thirty-year experience," *Perspectives in biology and medicine*, vol. 35, pp. 207-219, 1992.
 - [151] H. P. Chan, S. Galhotra, C. J. Vyborny, H. MacMahon, and P. M. Jokich, "Image feature analysis and computer-aided diagnosis in digital radiography.

- I. Automated detection of microcalcifications in mammography," *Medical physics*, vol. 14, pp. 538-548, 1987.
- [152] T. W. Way, L. M. Hadjiiski, B. Sahiner, H. P. Chan, P. N. Cascade, E. A. Kazerooni, *et al.*, "Computer-aided diagnosis of pulmonary nodules on CT scans: Segmentation and classification using 3D active contours," *Medical physics*, vol. 33, pp. 2323-2337, 2006.
 - [153] Y. Jiang, R. M. Nishikawa, R. A. Schmidt, C. E. Metz, M. L. Giger, and K. Doi, "Improving breast cancer diagnosis with computer-aided diagnosis," *Academic radiology*, vol. 6, pp. 22-33, 1999.
 - [154] B. Sahiner, N. Petrick, H.-P. Chan, L. M. Hadjiiski, C. Paramagul, M. A. Helvie, *et al.*, "Computer-aided characterization of mammographic masses: accuracy of mass segmentation and its effects on characterization," *IEEE transactions on medical imaging*, vol. 20, pp. 1275-1284, 2001.
 - [155] J. Ho, A. V. Parwani, D. M. Jukic, Y. Yagi, L. Anthony, and J. R. Gilbertson, "Use of whole slide imaging in surgical pathology quality assurance: design and pilot validation studies," *Human pathology*, vol. 37, pp. 322-331, 2006.
 - [156] F. Ghaznavi, A. Evans, A. Madabhushi, and M. Feldman, "Digital imaging in pathology: whole-slide imaging and beyond," *Annual Review of Pathology: Mechanisms of Disease*, vol. 8, pp. 331-359, 2013.
 - [157] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz, "Patch-based convolutional neural network for whole slide tissue image classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2424-2433.
 - [158] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, "Deep learning for identifying metastatic breast cancer," *arXiv preprint arXiv:1606.05718*, 2016.
 - [159] "CAMELYON16 ISBI challenge on cancer metastasis detection in lymph node," ed, 2016.
 - [160] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, pp. 29-36, 1982.
 - [161] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," in *Advances in neural information processing systems*, 2014, pp. 1988-1996.
 - [162] R. A. Irizarry, B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, *et al.*, "Exploration, normalization, and summaries of high density oligonucleotide array probe level data," *Biostatistics*, vol. 4, pp. 249-264, 2003.
 - [163] A. Liberzon, A. Subramanian, R. Pinchback, H. Thorvaldsdóttir, P. Tamayo, and J. P. Mesirov, "Molecular signatures database (MSigDB) 3.0," *Bioinformatics*, vol. 27, pp. 1739-1740, 2011.
 - [164] I. Wolf, M. Vetter, I. Wegner, T. Böttger, M. Nolden, M. Schöbinger, *et al.*, "The medical imaging interaction toolkit," *Medical image analysis*, vol. 9, pp. 594-604, 2005.
 - [165] T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, *et al.*, "NCBI GEO: archive for functional genomics data sets—update," *Nucleic acids research*, vol. 41, pp. D991-D995, 2012.

- [166] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proceedings of the 23rd ACM international conference on Multimedia*, 2015, pp. 689-692.
- [167] A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, *et al.*, "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences*, vol. 102, pp. 15545-15550, 2005.
- [168] M. Hollander, D. A. Wolfe, and E. Chicken, *Nonparametric statistical methods*: John Wiley & Sons, 2013.
- [169] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory*, vol. 28, pp. 129-137, 1982.
- [170] D. S. Ettinger, W. Akerley, G. Bepler, M. G. Blum, A. Chang, R. T. Cheney, *et al.*, "Non-small cell lung cancer," *Journal of the national comprehensive cancer network*, vol. 8, pp. 740-801, 2010.
- [171] L. H. Sobin and I. D. Fleming, "TNM classification of malignant tumors, (1997)," *Cancer*, vol. 80, pp. 1803-1804, 1997.
- [172] A. Ghosh and W. D. Heston, "Tumor target prostate specific membrane antigen (PSMA) and its regulation in prostate cancer," *Journal of cellular biochemistry*, vol. 91, pp. 528-539, 2004.

Appendix

Table A1. Patients' characteristics of the NSCLC data set (n = 69).

Variable	Counts (Frequency %)
<i>Gender</i>	
Male	46 (66.7%)
Female	23 (33.3%)
<i>T-stage</i>	
T1	21 (30.4%)
T2	30 (43.4%)
T3	15 (21.7%)
T4	2 (2.9%)
Tx	1 (1.4%)
<i>N-stage</i>	
N0	48 (69.6%)
N1	13 (18.8%)
N2	5 (7.2%)
Nx	3 (4.3%)
<i>M-stage</i>	
M0	64 (92.8%)
M1	4 (5.8%)
Mx	1 (1.4%)

Human-crafted Image Features

We have extracted a total number of 431 image features from CT volumes from the NSCLC dataset. The extracted image features can be sub-divided into the following groups, based on their feature properties (see Chapter 3 and 4 for feature description):

1. First order statistics
2. Shape and size based image features
3. Textural features
4. Wavelet features

First order statistics

Let X denote the three dimensional CT volume with N voxels. The first order histogram P with N_l discrete intensity levels. The following fourteen features were extracted in our experiment:

- A1. Energy:

$$energy = \sum_i^N X(i)^2$$

- A2. Entropy:

$$entropy = \sum_i^{N_l} P(i) \log_2 P(i)$$

- A3. Kurtosis:

$$kurtosis = \frac{\frac{1}{N} \sum_{i=1}^N (X(i) - \bar{X})^4}{(\sqrt{\frac{1}{N} \sum_{i=1}^N (X(i) - \bar{X})^2})^2}$$

where \bar{X} denotes the mean of X .

A4. Maximum:

The maximum voxel intensity value in the CT volume X .

A5. Mean:

$$mean = \frac{1}{N} \sum_i^N X(i)$$

A6. Mean absolute deviation:

The mean of the absolute deviation of all voxel intensities around the mean intensity in the CT volume X .

A7. Median:

The median voxel intensity value in the CT volume X .

A8. Minimum:

The minimum voxel intensity value in the CT volume X .

A9. Range:

The range of voxel intensity values in the CT volume X .

A10. Root mean square (RMS):

$$RMS = \sqrt{\frac{\sum_{i=1}^N X(i)^2}{N}}$$

A11. Skewness:

$$skewness = \frac{\frac{1}{N} \sum_{i=1}^N (X(i) - \bar{X})^3}{(\sqrt{\frac{1}{N} \sum_{i=1}^N (X(i) - \bar{X})^2})^3}$$

where \bar{X} denotes the mean of X .

A12. Standard deviation:

$$standard\ deviation = (\frac{1}{N-1} \sum_{i=1}^N (X(i) - \bar{X})^2)^{\frac{1}{2}}$$

where \bar{X} denotes the mean of X .

A13. Uniformity:

$$uniformity = \sum_{i=1}^{N_l} P(i)^2$$

A14. Variance:

$$variance = \frac{1}{N-1} \sum_{i=1}^N (X(i) - \bar{X})^2$$

where \bar{X} denotes the mean of X .

The standard deviation, mean absolute deviation and variance measure the dispersion of the intensity histogram of the CT volume X . The skewness measures the degree of asymmetry around the mean of the intensity histogram, while kurtosis quantifies the sharpness of the intensity histogram. Uniformity and entropy are descriptors for the randomness of the intensity histogram.

Shape and size based features

The following image features represent the three-dimensional size and shape of the tumour VOI. Let V denotes the volume and A the surface area of the volume of the VOI:

A15. Compactness 1:

$$compactness\ 1 = \frac{V}{\sqrt{\pi} A^{\frac{2}{3}}}$$

A16. Compactness 2:

$$compactness\ 2 = 36\pi \frac{V^2}{A^3}$$

A17. Maximum 3D diameter:

The maximum three-dimensional tumour diameter is measured as the largest pairwise Euclidean distance between voxels on the surfaces of the tumour VOI.

A18. Spherical disproportion:

$$sphericity\ disproportion = \frac{A}{4\pi R^2}$$

A19. Sphericity:

$$sphericity = \frac{\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}}{A}$$

A20. Surface area:

The surface area is measured as the sum of the number of border voxels that are located on each slice of the tumour VOI.

A21. Surface to volume ratio:

$$surface\ to\ volume\ ratio = \frac{A}{V}$$

A22. Volume:

The volume of the tumour VOI is measured by the total count of voxels in the tumour region and multiplying by the voxel size.

The maximum 3D diameter, surface area and volume feature quantify the size of the tumour VOI. Compactness, spherical disproportion, sphericity and the surface to volume ratio offer information to describe the shape aspects of the tumour VOI.

Textural features

This group of image features describes the patterns and spatial distribution of voxel intensities in the tumour VOI. In this thesis, the extracted gray-level co-occurrence (GLCM) and gray-level run-length (GLRLM) texture matrices were employed to describe the texture features of the tumour VOIs.

Gray-Level Co-Occurrence Matrix based features

A GLCM is defined as $\mathbf{P}(i,j;\delta,\alpha)$, \mathbf{P} is a matrix with the size of $N_g \times N_g$, where N_g denotes the number of discrete gray-level intensities in the VOI. The element at (i,j) denotes the number of co-occurrence of intensity values i and j in two voxels that are separated by a distance of δ voxels in the direction α . In our experiment, the distance δ was set to 1 with 13 directions in three dimensions for α . In addition, we denote the following terms, let:

μ be the mean of $\mathbf{P}(i,j)$,

$p_x(i) = \sum_{j=1}^{N_g} \mathbf{P}(i,j)$ be the marginal row probabilities,

$p_y(i) = \sum_{i=1}^{N_g} \mathbf{P}(i,j)$ be the marginal column probabilities,

μ_x be the mean of p_x ,

μ_y be the mean of p_y ,

σ_x be the standard deviation of p_x ,

σ_y be the standard deviation of p_y ,

$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \mathbf{P}(i,j)$, where $k = i + j, k = 2, 3, \dots, 2N_g$,

$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \mathbf{P}(i,j)$, where $k = |i - j|, k = 0, 1, \dots, N_g - 1$,

$HX = -\sum_{i=1}^{N_g} p_x(i) \log_2[p_x(i)]$, be the entropy of p_x ,

$HY = -\sum_{i=1}^{N_g} p_y(i) \log_2[p_y(i)]$, be the entropy of p_y ,

$H = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \mathbf{P}(i,j) \log_2[\mathbf{P}(i,j)]$, be the entropy of $\mathbf{P}(i,j)$,

$HXY1 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \mathbf{P}(i,j) \log(p_x(i)p_y(j))$, and

$HXY2 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)p_y(j) \log(p_x(i)p_y(j))$.

A23. Autocorrelation:

$$autocorrelation = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ij \mathbf{P}(i,j)$$

A24. Cluster prominence:

$$cluster\ prominence = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [i + j - \mu_x(i) - \mu_y(j)]^4 \mathbf{P}(i,j)$$

A25. Cluster shade:

$$cluster\ shade = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [i + j - \mu_x(i) - \mu_y(j)]^3 \mathbf{P}(i,j)$$

A26. Cluster tendency:

$$cluster\ tendency = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [i + j - \mu_x(i) - \mu_y(j)]^2 \mathbf{P}(i,j)$$

A27. Contrast:

$$contrast = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |i - j|^2 \mathbf{P}(i,j)$$

A28. Correlation:

$$correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ij \mathbf{P}(i,j) - \mu_x(i)\mu_y(j)}{\sigma_x(i)\sigma_y(j)}$$

A29. Difference entropy:

$$difference\ entropy = \sum_{i=0}^{N_g-1} \mathbf{P}_{x-y}(i) \log_2[\mathbf{P}_{x-y}(i)]$$

A30. Dissimilarity:

$$dissimilarity = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |i-j| \mathbf{P}(i,j)$$

A31. Energy:

$$energy = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [\mathbf{P}(i,j)]^2$$

A32. Entropy (H):

$$entropy = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \mathbf{P}(i,j) \log_2[\mathbf{P}(i,j)]$$

A33. Homogeneity 1:

$$homogeneity\ 1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{\mathbf{P}(i,j)}{1 + |i-j|}$$

A34. Homogeneity 2:

$$homogeneity\ 2 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{\mathbf{P}(i,j)}{1 + |i-j|^2}$$

A35. Informational measure of correlation 1 (IMC1):

$$IMC1 = \frac{HXY - HXY1}{\max\{HX, HY\}}$$

A36. Informational measure of correlation 2 (IMC2):

$$IMC2 = \sqrt{1 - e^{-2(HXY2-HXY)}}$$

A37. Inverse Difference Moment Normalised (IDMN):

$$IDMN = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{\mathbf{P}(i,j)}{1 + (\frac{|i-j|^2}{N^2})}$$

A38. Inverse Difference Normalised (IDN):

$$IDN = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{\mathbf{P}(i,j)}{1 + (\frac{|i-j|}{N})}$$

A39. Inverse variance:

$$inverse\ variance = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{\mathbf{P}(i,j)}{|i-j|^2}, i \neq j$$

A40. Maximum probability:

$$maximum\ probability = \max\{\mathbf{P}(i,j)\}$$

A41. Sum average:

$$sum\ average = \sum_{j=2}^{2N_g} [i\mathbf{P}_{x+y}(i)]$$

A42. Sum entropy:

$$sum\ entropy = \sum_{i=2}^{2N_g} \mathbf{P}_{x+y}(i) \log_2[\mathbf{P}_{x+y}(i)]$$

A43. Sum variance:

$$sum\ variance = \sum_{i=2}^{2N_g} (i - SE)^2 \mathbf{P}_{x+y}(i)$$

A44. Variance:

$$variance = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 \mathbf{P}(i,j)$$

Gray-Level Run-Length matrix based features

The gray-level run-length matrix is employed to quantify the number of consecutive pixels that have the same gray level in an image. Given a gray-level run-length matrix $p(i, j|\theta)$, the element at (i, j) describes the number of times, denoted as j , of pixels with gray level of i appears consecutively in the direction of θ . N_g denotes the number of discrete gray-level intensities. In our research, we computed GLRL matrix with 13 directions in the three-dimensional CT image volumes. The following GLRL features were extracted in the experiment, let:

$p(i, j|\theta)$ be the (i, j) th entry in the gray-level run-length matrix p in direction θ .

N_r denotes the number of different run lengths,

N_p denotes the number of voxels in the image volume.

A45. Short Run Emphasis (SRE):

$$SRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} [\frac{p(i, j|\theta)}{j^2}]}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

A46. Long Run Emphasis (LRE):

$$LRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 p(i, j|\theta)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

A47. Gray Level Non-Uniformity (GLN):

$$GLN = \frac{\sum_{i=1}^{N_g} [\sum_{j=1}^{N_r} p(i, j|\theta)]^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

A48. Run Length Non-Uniformity (RLN):

$$RLN = \frac{\sum_{j=1}^{N_r} [\sum_{i=1}^{N_g} p(i, j|\theta)]^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

A49. Run Percentage (RP):

$$RP = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{p(i, j|\theta)}{N_p}$$

A50. Low Gray Level Run Emphasis (LGLRE):

$$LGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} [\frac{p(i, j|\theta)}{i^2}]}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

A51. High Gray Level Run Emphasis (HGLRE):

$$HGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} i^2 p(i, j|\theta)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

A52. Short Run Low Gray Level Emphasis (SRLGLE):

$$SRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} [\frac{p(i, j|\theta)}{i^2 j^2}]}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

A53. Short Run High Gray Level Emphasis (SRHGLE):

$$SRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} [\frac{p(i, j|\theta) i^2}{j^2}]}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

A54. Long Run Low Gray Level Emphasis (LRLGLE):

$$LRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} [\frac{p(i, j|\theta) j^2}{i^2}]}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

A55. Long Run High Gray Level Emphasis (LRHGLE):

$$LRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta) i^2 j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)}$$

Wavelet features: wavelet decompositions

Wavelet transformation allows the decomposition of images in low- and high-frequencies. Our experiment involves the application of a discrete, one-level and undecimated three-dimensional wavelet transformation to each of the CT image volumes. The original CT image volume X was decomposed into eight decompositions. Let L and H denote the low- and high-pass function respectively. The resulted decompositions of X are labelled as X_{LLL} , X_{LLH} , X_{LHL} , X_{LHH} , X_{HLL} , X_{HLH} , X_{HHL} and X_{HHH} . The decomposition of X_{LHL} represents the directional filtering of X with low-pass filters along the x- and z-directions, and a high-pass filter along the y-direction given Equation A66:

$$X_{LHL} = \sum_{p=1}^{N_L} \sum_{q=1}^{N_H} \sum_{r=1}^{N_L} L(p)H(q)L(r)X(i+p, j+q, k+r)$$

where N_L denotes the length of filter L while N_H denotes the length of filter H . In our experiment, wavelet “Coiflet 1” was applied to CT image volumes to derive all 8 decompositions. The first order statistics (group 1) and textural features (group 3) were computed for each of the decompositions.