

QUANTITATIVE ANALYSIS WITH MACHINE LEARNING MODELS FOR MULTI-PARAMETRIC BRAIN IMAGING DATA

Faculty of Engineering

A thesis submitted in fulfilment of the requirements for the
Degree of Master of Philosophy in the School of Computer Science at

The UNIVERSITY OF SYDNEY

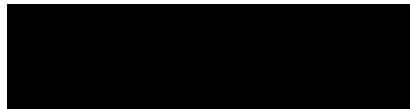
Dingqian Wang

September 2019

I declare that this thesis is my own work and has not been submitted in any other form for another degree at any university.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged. Information derived from the published or unpublished work has been acknowledged in the text and a reference list is given.

SIGNED:

A solid black rectangular box used to redact the signature of the author.

DATE: 30 Sep 2019

© Copyright by DINGQIAN WANG 2019

All Rights Reserved

Abstract

Gliomas are considered to be the most common primary adult malignant brain tumor. With the dramatic increases in computational power and improvements in image analysis algorithms, computer-aided medical image analysis has been introduced into clinical applications. Precision tumor grading and genotyping play an indispensable role in clinical diagnosis, treatment and prognosis. Gliomas diagnostic procedures include histopathological imaging tests, molecular imaging scans and tumor grading.

Pathologic review of tumor morphology in histologic sections is the traditional method for cancer classification and grading, yet human study has limitations that can result in low reproducibility and inter-observer agreement. Compared with histopathological images, Magnetic resonance (MR) imaging present the different structure and functional features, which might serve as noninvasive surrogates for tumor genotypes.

Therefore, computer-aided image analysis has been adopted in clinical application, which might partially overcome these shortcomings due to its capacity to quantitatively and reproducibly measure multilevel features on multi-parametric medical information.

Imaging features obtained from a single modal image do not fully represent the disease, so quantitative imaging features, including morphological, structural, cellular and molecular level features, derived from multi-modality medical images should be integrated into computer-aided medical image analysis. The image quality differentiation between multi-modality images is a challenge in the field of computer-aided medical image analysis.

In this thesis, we aim to integrate the quantitative imaging data obtained from multiple modalities into mathematical models of tumor prediction response to achieve additional insights into practical predictive value. Our major contributions in this thesis are:

1. Firstly, to resolve the imaging quality difference and observer-dependent in histological image diagnosis, we proposed an automated machine-learning brain tumor-grading platform to investigate contributions of multi-parameters from multimodal data including imaging parameters or features from Whole Slide Images (WSI) and the proliferation marker KI-67. For each WSI, we extract both visual parameters such as morphology parameters and sub-visual parameters including first-order and second-order features. A quantitative interpretable machine learning approach (Local Interpretable Model-Agnostic Explanations) was followed to measure the contribution of features for single case. Most grading systems based on machine learning models are considered “black boxes,” whereas with this system the clinically trusted reasoning could be revealed. The quantitative analysis and explanation may assist clinicians to better understand the disease and accordingly to choose optimal treatments for improving clinical outcomes.

2. Based on the automated brain tumor-grading platform we propose, multimodal Magnetic Resonance Images (MRIs) have been introduced in our research. A new imaging–tissue correlation based approach called RA-PA-Thomics was proposed to predict the IDH genotype. Inspired by the concept of image fusion, we integrate multimodal MRIs and the scans of histopathological images for indirect, fast, and cost saving IDH genotyping.

The proposed model has been verified by multiple evaluation criteria for the integrated data set and compared to the results in the prior art. The experimental data set includes public data sets and image information from two hospitals. Experimental results indicate that the model provided improves the accuracy of glioma grading and genotyping.

Publications

1. Wang, X., **Wang, D.**, Yao, Z., Xin, B., Wang, B., Lan, C., Qin, Y., Xu, S., He, D., Liu, Y. (2019). Machine Learning Models for Multiparametric Glioma Grading with Quantitative Result Interpretations. *Frontiers in Neuroscience*, 12(January 2019)
2. Xue, Z., Xin, B., **Wang, D.**, Wang, X. (2019). Radiomics-Enhanced Multi-Task Neural Network for Non-invasive Glioma Subtyping and Segmentation. In MICCAI workshop on Radiomics and Radiogenomics in Neuro-oncology using AI 2019

Journal Paper under Review

3. RA-PA-Thomics: A new approach to glioma genotype prediction

Attribution statements

Chapter 3 of this thesis is published as [1] and Chapter 4 is under review. The basic idea and algorithm design were from my supervisor who was the first author. My contributions include algorithm implementation, validation, and comparison. Therefore, the corresponding author positions are co-first authors.

Acknowledgement

Firstly, I would like to express appreciation to my supervisor Associate Professor Xiuying Wang. Thanks for the important opportunity to start my research journey and guide me through my challenging research. Dr. Wang's research enthusiasm, diligence and scientific attitudes set a role model for us and always motivate and encourage me to actively improve my research capacity and strive to achieve my personal best.

I would like to thank my friend Bowen Xin and Chongrui Xu for their support. In the research, thank you for knowledge sharing and exchange. In life, thank you for your encouragement, supports and friendship.

I am grateful to the Shandong Provincial Hospital for the valuable collaborations. Thanks to Dr. Yingchao Liu and the team for providing professional medical knowledge and inputs.

Finally, I would like to thank my family for supporting me over the years and thank my parents for their unconditional supports and encouragement. Especially thanks to my sister for her companionship and support. I am grateful for all the love and happiness from the family.

Contents

Abstract	1
Publications	4
Attribution statements	5
Acknowledgement	6
Contents	7
List of Figures	10
List of Tables	12
CHAPTER 1. Introduction	13
1.1. Motivation	13
1.2. Challenges	15
1.2.1. Multimodal Medical Image Fusion	15
1.2.2. Heterogeneity of medical information	16
1.3. Contributions	16
1.4. Thesis Organizations	17
CHAPTER 2. Background	19
2.1. Applications of Multi-parametric Medical Image Analysis	19
2.2. Multi-parametric Biomedical Images	21
2.2.1. Whole Slide images	21
2.2.2. Magnetic resonance imaging (MRI)	23
2.3. Multi-modality Image Analysis Procedure	24
2.4. Segmentation	25
2.5. Feature Extraction	28
2.5.1. Features used in Histopathological Image Analysis	29
2.5.2. Features used in Magnetic Resonance Imaging Analysis	31
2.6. Feature Selection	34
2.7. Machine-Learning Models Establishment and Prediction	37
CHAPTER 3. Machine Learning Models for Multi-parametric Glioma Grading With Quantitative Result Interpretations	41
3.1. Abstract	41

3.2.	Introduction.....	42
3.3.	Materials and Methods.....	45
3.4.	Automated interpretable Glioma grading with Machine-Learning Models	46
3.4.1.	Automated ROI Identification	48
3.4.2.	Multi-Parameter Extraction and Important Parameter Selection	48
3.4.3.	Machine Learning Model for Automated Grading	49
3.4.4.	Validation and Evaluation.....	51
3.5.	Results.....	51
3.5.1.	Multi-Parameter Extraction and Selection	51
3.5.2.	Automated Grading Performance with Different Models	52
3.5.3.	Further Investigation of SVM Performance	53
3.5.4.	Quantitative Grading Results Interpretation	55
3.6.	Discussion	60
3.7.	Chapter summary	64
CHAPTER 4. RA-PA-Thomics: A new approach to glioma genotype prediction.....		65
4.1.	Abstract	65
4.2.	Introduction.....	66
4.3.	Materials and Dataset.....	68
4.3.1.	Patient Enrolment.....	68
4.3.2.	Dataset.....	69
4.4.	Computer Analysis.....	71
4.4.1.	ROI Extraction	73
4.4.2.	Feature Extraction	74
4.4.3.	Feature Selection.....	75
4.4.4.	Modelling and Validation	75
4.5.	Results.....	76
4.5.1.	Feature Extraction and selection	76

4.5.2. Comparison of Performances Using Different Modalities of Biomedical Images and Feature Types	80
4.5.3. Univariate Feature Analysis.....	81
4.5.4. Quantitative IDH1 Status Prediction Results Interpretation	86
4.6. Discussion	88
4.7. Chapter Summary	90
CHAPTER 5. Conclusions and Future Work	91
5.1. Conclusions.....	91
5.2. Future Work.....	92
Bibliography	94

List of Figures

Figure 1 Outline of Thesis	18
Figure 2 Huge size of Histopathological Images	22
Figure 3 Multi-resolution of the same whole slide images.....	22
Figure 4 Multimodal MR Images. Brain MR Images using different imaging modalities: T1, T1 contrast, T2 and T2/FLAIR.....	24
Figure 5 The procedure of multi-modality biomedical images analysis.....	25
Figure 6 Presentation of different features in the histopathological image.	30
Figure 7 Prediction Model Establishment via machine learning algorithms.	37
Figure 8 Schematic flowchart of the automated grading framework.	48
Figure 9 Machine learning model with hyper-parameter tuning.	50
Figure 10 Predictive capability of selected important features and models.	52
Figure 11 Assessment of grading results with the support vector machine (SVM) model.	54
Figure 12 Grading result explanation of representative cases with the LIME algorithm for the SVM model (part 1)	57
Figure 13 Grading result explanation of representative cases with the Local Interpretable Model-Agnostic Explanations (LIME) algorithm for the SVM model (part 2)	58
Figure 14 Grading result explanation of representative cases with the Local Interpretable Model-Agnostic Explanations (LIME) algorithm for the SVM model (part 3)	59

Figure 15 Machine learning framework for automated prediction of IDH glioblastoma genotypes.	72
Figure 16 Feature extraction from histopathological images.	77
Figure 17 Measurements obtained from different regions of individual MR images.	79
Figure 18 Random forest classifier scores for IDH1 genotype prediction.	80
Figure 20 IDH1 status prediction result explanation of representative cases with the LIME algorithm for the RF model (IDH1-Wild type case)	86
Figure 21 IDH1 status prediction result explanation of representative cases with the LIME algorithm for the RF model (IDH1-Mutation case)	87

List of Tables

Table 1 F1, accuracy, precision, and recall for different machine learning models.	53
Table 2 Patient Characterizes.....	69
Table 3 TOP-performing features in IDH1 Status Prediction by means of Univariate Analysis.....	81
Table 4 Prediction of IDH1 Genotype Based in High Grade Gliomas.....	85

CHAPTER 1. Introduction

1.1. Motivation

The widespread use of computer-aided medical image analysis dates back to the emergence of digital breast cancer in the early 1990's [2]. Today, computer-aided design (CAD) has become an integral part of medical imaging and diagnostics [3, 4]. According to recent research, the possibilities of combining multimodal medical images with computer-aided image analysis is very broad [5, 6]. This technology can improve the accuracy of disease classification and can reduce the workload of clinicians.

Histopathological images contain a wealth of structural information, and pathologists can not only diagnose tumors but also the extent of cancer cells, the degree of malignancy, and the extent of tumor spread [7, 8]. At present, how to build a platform for automatically extracting useful information from medical images for diagnosis is an area of great interest to many researchers. Compared to magnetic resonance images, the extraction of whole slide images (WSIs) is damaging to the human body. Therefore, with the development of medical imaging technology, non-invasive magnetic resonance images have also begun to be included in the method of tumor diagnosis. For clinical medicine, accurate glioma grading is important. It is usually obtained by a pathologist manually grading histopathological images. In order to improve the accuracy and repeatability, and reduce the empirical differences between different observers, computer-aided imaging analysis has been introduced into this area. Accurate tumor grading and genotyping are an integral part of clinical cancer research and may affect the patient's clinical diagnosis and prognosis [9]. Gliomas diagnostic

procedures include imaging tests and tumor grade determination. Although the cause of glioma is still unclear, there are still some factors that may increase the risk of glioma, including age, radiation exposure, and family genetic history.

With the continuous development and advancement of medical imaging technology and computer technology, medical image analysis has become an indispensable tool and technical means in medical research, clinical disease diagnosis, and treatment. In recent years, machine learning has rapidly developed into a research hotspot in medical image analysis, which can automatically distinguish the characteristics of disease diagnosis from the medical image big data. This thesis first briefly describes the characteristics of medical image analysis; secondly, the basic principles of machine learning, then introduces the main machine learning algorithms mainly used for medical image analysis; and finally discusses their research status in the field of medical image applications.

Accurate tumor grading and genotyping play a significant role in oncology [9]. For instance, computational-aid tumor histopathological grading may help the junior pathologists improve their diagnose accuracy. Tumor grading and genotyping are the basis for the treatment of clinical diagnosis. As in clinical diagnose procedure, the whole process are depend on the experience of clinicians, which may contain subjective and not easy for junior doctors. Compared with traditional methods predicted by pathologists, computer-aided grading and genotyping have the advantages of repeatability and low labor consumption.

1.2. Challenges

Multi-parametric medical image analysis faces some serious challenges. Comprehensive diagnostic information cannot be obtained from a single medical alone. In computer-aided medical image analysis, different modal medical information with different qualities make it difficult to accurately integrated and fully utilized.

1.2.1. Multimodal Medical Image Fusion

The digital processing of medical images has become a practical clinical technique, which is widely used in various medical imaging disciplines, such as histopathological images, MRI and PET. Different imaging modalities are highly specific and can be divided into anatomical images (CT, MRI) and functional images (PET) from a medical perspective [10, 11]. Anatomical images can clearly show the anatomy of internal organs and bones with high spatial resolution and geometric characteristics, but does not have the function information display; while functional images can display the transformation of the human body's metabolic function, providing the body function information, but the image is blurred. It is impossible to clearly reflect the morphological structure, it is difficult to obtain accurate anatomical structure and stereoscopic positioning, and it is difficult to distinguish the boundaries of tissues and organs. Therefore, comprehensive diagnostic information cannot be obtained from a single image alone. For example, through image registration, images of different modalities are mapped into the same coordinate system, which reflects the corresponding tumor regions are in the same position in space. This can reflect the comprehensive information of various forms and functions of the human body, so that the lesion or ROI has clear visibility, which has important application value in the

positioning of the lesion, design of the radiotherapy plan, guiding the operation and checking the therapeutic effect.

Common multimodal medical image fusion methods found in recent researches including intensity-hue-saturation (IHS) and principal component analysis (PCA) method [12], fuzzy set method (type-2 fuzzy logical method [13] and fuzzy transform method [14]), Neural network method (non-subsampled contour-let transform and pulse-coupled neural network [15]), and wavelet transform method [11, 16, 17]. Heterogeneity of medical information

In conventional clinical image acquisition, differences in various imaging parameters, such as pixel size, matrix size, slice thickness, tumor position, and reconstruction algorithms, need to be considered. When meaningful data is extracted digitally to analyse the image phenotypes, discrepancies in these imaging parameters may lead to some changes that are not due to potential biological effects. It is therefore necessary to use standardized imaging protocols to eliminate the variability of unnecessary confounding factors.

1.3. Contributions

In this thesis, we target at integrating the quantitative imaging data obtained from multiple modalities into mathematical models of tumor prediction response to achieve additional insights into practical predictive value. Our major contribution in this thesis are summarized below:

1. We proposed an automated machine-learning brain tumor-grading platform to investigate contributions of multi-parameters from multimodal data including

imaging parameters or features from the Whole Slide Images (WSI) and the proliferation marker KI-67. For each WSI, we extract both visual parameters such as morphology parameters and sub-visual parameters including first-order and second-order features. On the basis of machine learning models, our platform classifies gliomas into grades II, III, and IV.

2. A quantitative interpretable machine learning approach (Local Interpretable Model-Agnostic Explanations) was proposed to measure the contribution of features for single case. Most of the grading systems based on machine learning models are considered “black boxes,” and it would be valuable for patient management if clinically trusted reasoning could be revealed. The quantitative analysis and explanation may assist clinicians to better understand the disease and accordingly to choose optimal treatments for improving clinical outcomes.
3. Based on the automated brain tumor-grading platform we proposed above, multimodal MRIs have been introduced in our research. A new imaging –tissue correlation based approached called RA-PA-Thomics was proposed to predict the IDH genotype. Inspired by the concept of image fusion, we integrate multimodal MRIs and the scans of histopathological images for indirect, fast, and cost saving IDH genotyping.

1.4. Thesis Organizations

This thesis has been organized to follow the normal procedures of the multi-modality biomedical imagery. As shown in **Figure 1**, the overview plan of this thesis has been listed. Computer-aided biomedical images analysis is reviewed and discussed in Chapter 2. Also, in this chapter, we introduced the machine-learning application in the

computer-aided biomedical imaging analysis field. As for Chapters 3 and 4 are present the proposed models we figure out. In the end, Chapter 5 summarise our conclusions and forecasts the research trend of multi-parametric imaging data processing analysis for brain tumour diagnoses.

Chapter 2 Background

- Application of Multi-Parametric Medical Image Analysis
- Multi-parametric Biomedical Images
- Multi-modality Image Analysis Procedure

Chapter 3 Machine Learning Models for Multi-parametric Glioma Grading With Quantitative Result Interpretations

- Automated Interpretable Glioma Grading with Machine-learning Models
- Quantitative Grading Interpretation

Chapter 4 RA-PA-Thomics: A new approach to glioma genotype prediction

- Multi-modality Biomedical Information Integration
- Optimal Feature Analysis
- Quantitative Grading Interpretation

Chapter 5 Conclusions and Future Work

- Conclusions
- Feature Work

Figure 1 Outline of Thesis

CHAPTER 2. **Background**

In this chapter, we have summarized related work on multi-parametric medical image analysis. We first briefly introduce the applications of computer-aided medical image analysis, followed by the image description (histopathological images and MRIs) techniques that we will typically utilize for model validation. Finally, we provide the literature review of computer-aided multi-parametric medical image analysis procedure.

2.1. Applications of Multi-parametric Medical Image Analysis

Computational Analysis for Clinical Diagnose: At present, the most widely used machine learning application in biomedical imaging field is the use of computer-aided biomedical imaging analysis and diagnosis. For example:

- 1) Targeting in the field of radiotherapy, the first half of radiotherapy should first outline the boundary between the surrounding organs and the tumor. In terms of computer-aided medical analysis, machine-learning based segmentation can better to figure out the edge of the tumor.
- 2) Automated analysis of pathological images, pathological diagnosis is the gold standard, but usually, the picture pixels are very high, where AI can effectively prevent missed diagnosis and accelerate diagnosis;
- 3) Early screening for disease, such as early screening for pulmonary nodules.

Image retrieval applied in medical applications: Machine-learning algorithms in computer-aided biomedical imaging fields have been widely used.

Content Based Image Retrieval (CBIR) is an image retrieval algorithm that is widely used in the medical field. CBIR is also known as Query by Image Content (QBIC), which provides technology that allows digital images to be organized through their visual features. The implementation of this technology depends on two aspects:

- 1) Computer vision technology,
- 2) An image library with a large amount of data.

CBIR is implemented by matching similar images to a fixed image in a massive database. CBIR is widely used in medical imaging, including segmentation and diagnosis of region of interest. For example, pathologists and researchers can use CBIR for tumor grade prediction by retrieving ROI in pathological images and extracting relevant features [18-24]. CBIR can be considered as unsupervised learning [23]. Due to the insufficient number of labelled data, this unsupervised learning algorithm is more suitable for biomedical field [21, 22].

Discovering the Relationships between Multimodalities Biomedical Images and Clinical Information: According to research, pathologists and researchers obtain important findings in the current clinical diagnosis of diseases such as invasive tumors by observing histopathological images. At the same time, along with the advancement of medical information digital and genomic analysis techniques, and the development of missiology, multi-modalities of medical imaging (MRI, WSIs, CT), genomic information, and clinical information are integrated into the final diagnosis [25]. Researchers and clinicians attempted to analyze the relationship between these data and discover the relationships between medical images of different modalities and the relationships between images in the same modality [26, 27]. However, a single whole

slide image may contain millions of nuclei information. It is unrealistic for researchers and clinicians to observe, identify and analyze such a large amount of data. At the same time, the analysis is considered subjective and non-repeatable. Machine learning has the potential to help with analyses that are instead both objective and repeatable. This is already being used in several fields, for example, prognostic analysis of current common cancer patients [28-31], genotype prediction [31, 32], and tumor grading [33-35].

2.2. Multi-parametric Biomedical Images

Whole Slide images

As a gold standard for pathological diagnosis, histopathological images have important applications in clinical and scientific research [36-38]. The pathologist performs a pathological diagnosis and a prognostic evaluation by microscopic examination of the pathological section, but this process is usually time consuming and laborious. The advent of digitization of pathological sections is considered an important turning point in the pathological development process [39, 40]. The production of digital pathology sections first requires tissue staining and subsequent digitization into digital pathology sections by a microscopy camera [41]. However, due to the low quality of the original digital slice, its popularity was limited. In 1999, the whole digital image appeared, making the preservation and transmission of pathological slices more convenient and safe [42].

Whole slide images (**Figure 2**) refer to the cutting of a certain size of the diseased tissue, using hematoxylin and bluish (H & E) and other staining methods to make the sliced tissue into a pathological slide, and then microscopic imaging technology to microscopic cells and glands. By analyzing the pathological images, the causes,

pathogenesis, and pathogenesis of the lesions can be explored to make a pathological diagnosis.

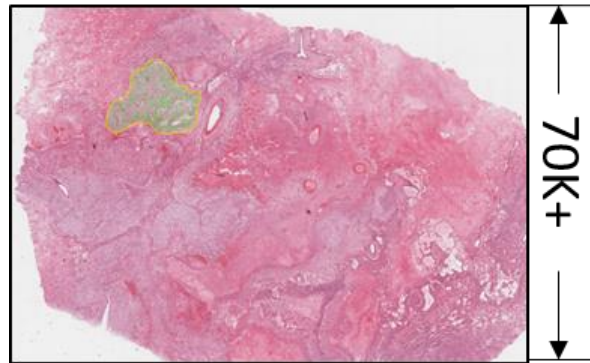


Figure 2 Huge size of Histopathological Images

Histopathological images are usually scanned at high resolution and the images we collected may exceed 1GB in size (**Figure 3**) [43, 44]. This kind of large tissue contains a large amount of redundant information.

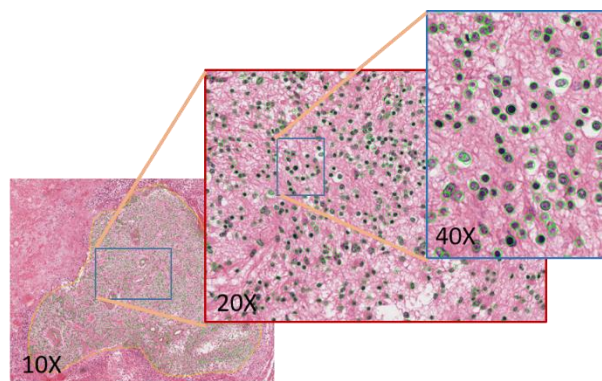


Figure 3 Multi-resolution of the same whole slide images. The left image was scanned at 10X resolution; it is difficult for pathologists to extract enough information directly from them. The image on the right is the field of view under 40x resolution.

2.2.2. Magnetic resonance imaging (MRI)

Human glioma is a common primary malignant tumor of the brain. The current treatment is still mainly surgical resection. After surgery, it is supplemented with radiotherapy and chemotherapy. However, different pathological grades have different malignant degrees and different responses to treatment. Recent studies have shown a relationship between MRI features and pathological grades of human gliomas and cell proliferation, genotype, and prognosis, which could be used to improve these outcomes.

Compared with WSIs, MRI of the brain is a safe and non-invasive medical imaging method that uses electromagnetic waves to draw detailed images of the brain. MRI [45], which is a measure of the magnitude of a magnetic resonance signal generated by a hydrogen atomic nucleus in human tissues and organs under the action of an strong external magnetic field, and is performed by the computer on the information data received by the external nuclear magnetic resonance signal detector. The concentration of atoms in the body, most commonly hydrogen, can be used to map the density of tissue in the body. Nuclear magnetic resonance can be used to identify where these atoms are concentrated, and the results can be mapped into an image of the body by a computer. 3D image reconstruction. It provides very clear images of human soft tissue anatomy and lesions.

Multimodal MRI imaging, a method in which multiple functional magnetic resonance imaging modes are fused together to diagnose and examine patients, is a rapidly evolving technique in brain imaging (**Figure 4**), including T1, T1 contrast, T2 and Fluid attenuation inversion recovery (FLAIR) [46, 47]. It provides information on tissue anatomy, functional metabolism, etc. [48] It is characterized by high resolution, high

precision, and non-invasiveness, which is helpful for a more accurate depiction of the edge of the lesion before surgery.

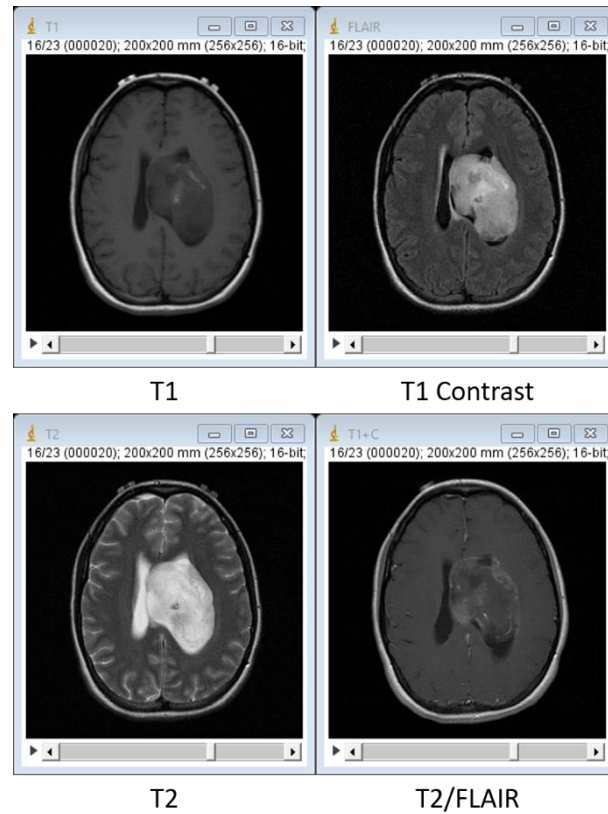


Figure 4 Multimodal MR Images. Brain MR Images using different imaging modalities: T1, T1 contrast, T2 and T2/FLAIR.

2.3. Multi-modality Image Analysis Procedure

The process of multi-modality image analysis can be divided into five steps as follows (**Figure 5**): image acquisition, the region of interest (ROI) extraction, features extraction, feature selection, consequence interpretation and analysis.

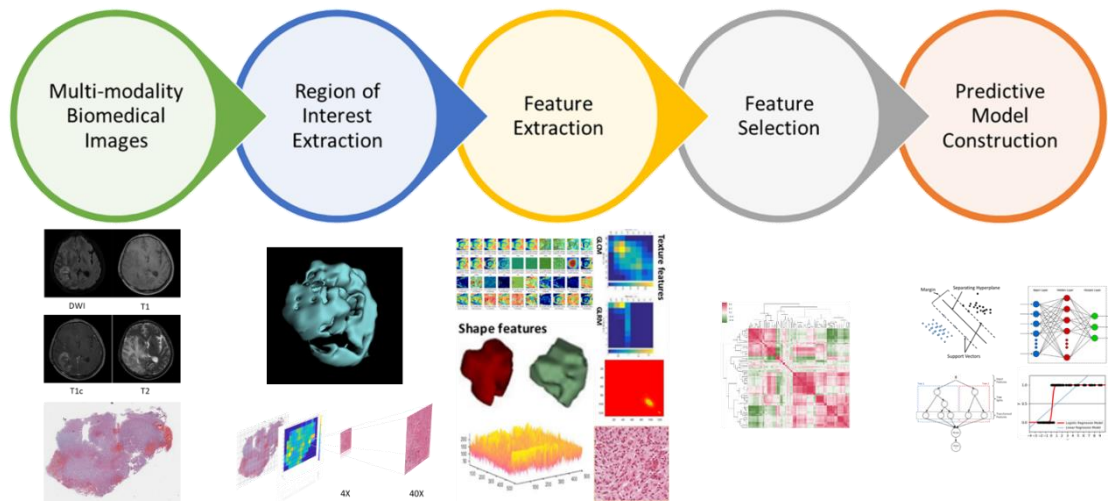


Figure 5 The procedure of multi-modality biomedical images analysis. The first step is to collect multi-modality biomedical images, including the scans of histopathological images, multi-modal MR images. Next, applying these images into segmentation algorithms. Furthermore, based on the ROIs, we automatically and quantitatively extract multi-parametric features. Then, through feature selection algorithms, we select optimal features and implement into machine-learning algorithms to predict the outcomes.

2.4. Segmentation

Medical image segmentation is the process of dividing an image into regions based on similarities or differences between regions. At present, images of various cells, tissues and organs are mainly used as objects of treatment. The traditional image segmentation technology has a region-based segmentation method and a boundary-based segmentation method. The former relies on the spatial local features of the image, such as the uniformity of grayscale, texture and other pixel statistical properties. The latter mainly uses gradient information. The latter mainly uses gradient information to

determine the boundaries of the target. With the combination of specific theoretical tools, image segmentation technology has been further developed.

In recent years, with the development of other emerging disciplines, some new image segmentation techniques have emerged such as statistical methods [49-51], fuzzy theory-based methods [52, 53], neural network-based methods [54, 55], wavelet analysis-based methods [56], model-based snake models (dynamic contour models [57, 58]), combined optimization models and other methods. Although new methods of segmentation have been proposed, the results are not satisfactory. The current research hotspot is a knowledge-based segmentation method, which introduces some transcendental knowledge into the segmentation process by some means, thus constraining the segmentation process of the computer so that the segmentation result is controlled within the range we can recognize. For example, when the intrahepatic mass and the normal liver gray value are very different, the mass and the normal liver are not regarded as two independent tissues.

The research of medical image segmentation method has the following remarkable features: any single image segmentation algorithm has difficulties obtaining satisfactory results for general images – more attention should be paid to the effective combination of multiple segmentation algorithms. Although there have been studies to distinguish the required organs, tissues or methods to find the lesions by automatic segmentation of medical images, currently available software packages generally still cannot complete fully automatic segmentation and still require anatomical aspects manual intervention [59]. In the current situation that the image segmentation task cannot be completely completed by a computer, the human-computer interactive segmentation method has gradually become the research focus; the new segmentation

methods have several goals: automatic, precise, fast, adaptive and robust segmentation. The comprehensive utilization of classical segmentation technology and modern segmentation technology (integration technology) is an appropriate goal for the future development of medical image segmentation technology [49, 60].

Compared to the 2-dimensional histopathological image, the Volume of Interest (VOI) outlined on the MRI is 3-dimensional. The specific steps are as follows: First, two experienced radiologists will outline the VOI of the image. When the two opinions are inconsistent, they are judged by another senior title physician. Secondly, using the MATLAB (R2017b)-based boundary sketching program, based on the enhanced T1WI image and other modal sequences as a reference, manually extract the maximum extent of the diseased tissue at each level of the visible lesion as the ROI of the layer. Stacking multiple levels of ROI together generates a VOI on the patient's enhanced T1WI image. Finally, the VOI boundary on the enhanced T1WI image is automatically copied to several other parameter maps that have been registered, and finally the multimodal VOI of each patient image is obtained.

Segmentation algorithms are usually divided into three categories; manual, semiautomatic, and automatic segmentation. There is no absolutely perfect algorithm for image segmentation as each category faces different opportunities and challenges.

Firstly, manual segmentation relies on physician experience; they evaluate each slice of the radiologic imaging sequence or pathological cell subtype information to delineate the ROIs. Although it is time-consuming and tiresomeness, it is very effective in reproducibility and reliability for tumors with large morphological variations and partially blurred volumes.

Secondly, in the face of the large number of data that is needed to delineate the ROIs, it is often necessary to rely on the assistance of semi-automatic algorithms to reduce the number of human interactions as much as possible and improve the segmentation efficiency.

Thirdly, with the development of segmentation algorithms based on artificial networks, many fully automated algorithms overcome the challenges of unclear boundaries and different shapes.

For clinical diagnosis, accurate histopathological grading and subtyping are critical, which may further influence the surgical treatment and prognosis analysis. From Figure 4, we can see that a single histopathological image may contain different kinds of features.

2.5. Feature Extraction

At present, there are many feature extraction methods, and the feature extraction methods used according to different purposes and different targets are also different. Commonly used image feature types can be divided into:

- 1) Grayscale features of the image [61, 62], such as the mean or variance of the image or local;
- 2) Texture features of the image [58, 63], such as co-occurrence matrix, equal grayscale run length, Fourier spectrum, random field model, etc.;
- 3) Spectral features of the image [64] Gradient features of image grayscale transformation;

- 4) Image object shape features [65, 66] such as area, perimeter, circularity, aspect ratio, moment, edge moment, Fourier descriptor, eccentricity, compactness, etc.;
- 5) Image information description and Signal to noise ratio and so on.

2.5.1. Features used in Histopathological Image Analysis

The WSI pathology image contains massive structural information [67, 68]. Therefore, converting images into mineable feature data is a necessary step in the pathological analysis [69, 70]. The method of feature extraction can be divided into artificial feature design and automatic feature learning. Artificial feature design refers to the selection and simplification of low-dimensional vectors that can best express image content from images. These features include gray histograms [71], shape features [72, 73], texture features, and relationships with surrounding tissues. Digital pathology slice analysis of traditional machine learning algorithms requires manual extraction of features. However, the application of artificial features has the following disadvantages: First, feature selection relies heavily on professional experience, objectivity is poor, and it is impossible to represent comprehensive information of pictures; second, there is a lack of principle criteria for integrating multiple artificial features.

As shown in **Figure 6**, we can figure out that a single histopathological image may contain different kind of features.

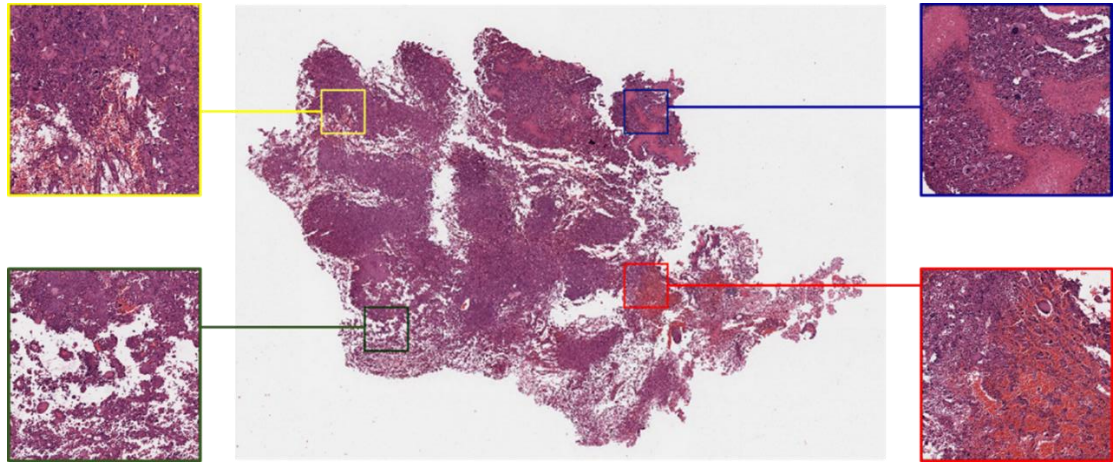


Figure 6 Presentation of different features in the histopathological image.

2.5.1.1 **Pixel-level Features** [68, 72]: Pixel-level image features refer to features that are unexplained in biology, which lied in the bottom of the information hierarchy. Pixel-level image features usually refer to some small details in the image, such as edge, corner, color, pixels, gradients, etc., which can be obtained by filter, SIFT or HOG.

2.5.1.2 **Object-level Features** [72, 74]: Compared with pixel-level characteristics, characterized in object-level information of the higher level hierarchy. This layer is characterized mainly described the histopathological characteristics of the image in the cell structure, for example, nuclei and glands. Generally, object-level features can be divided into the following categories: size and shape, radiometric and densitometry, texture, and chromatin-specific [43].

2.5.1.3 **Spatially Related Features** [75]: Spatially related features refer to the topological features of image structure information. This kind of feature was introduced as early as the 1990s, but it has only recently appeared in medical image analysis [76]. Such features provide a wealth of structure information to describe histopathological images, including cell distribution and image globalist.

2.5.2. Features used in Magnetic Resonance Imaging Analysis

In the field of biomedical imaging analysis, radiomics is a method of extracting massive features from emission images using different data feature algorithms. Radiomics features can be divided into shape-based features, first-order statistics features, texture features and features extracted by using different filters [77].

2.5.2.1 **Shape-based Features:** Measures include the description of the volume of tumor areas, such as volume, surface area, maximum diameter in two and three dimensions, and effective diameter, and features describing the degree of similarity between the ROI and the sphere, such as Surface volume ratio, density, eccentricity, sphericity, etc. In short, it is a feature of geometry, which has the same meaning as the volume we calculate. For example, sphericity describes the measure of the roundness of the tumor area relative to the sphere. The higher value ($0 < Sphericity \leq 1$) implies the higher the roundness of the ROI. The value is calculated as:

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

Where V is the volume of ROI and A is the surface of ROI.

2.5.2.2 First-order Statistics Features: The first-order statistics features, also named intensity features, which measures the intensity distribution characteristics in ROIs (it is a feature that can be calculated statistically on the grey value of an image), do not include spatial features between pixels. These intensity features are obtained by histogram analysis, including the mean, median Number, minimum, maximum, standard deviation, skewness, and kurtosis. These features reflect the symmetry, uniformity, and local intensity distribution of the measured voxels. For example, energy describes the magnitude of voxel values in the ROI [78]. A larger value implies a greater sum of the squares of these values. The function is defined as below:

$$energy = \sum_{i=1}^{N_p} (X(i) + c)^2 \quad (2.2)$$

2.5.2.3 Texture Features: The texture features describe the intensity of the voxel spatial distribution, which can also be called as the second-order statistics features. Image texture refers to a spatial change that is perceptible or measurable at the intensity level. It is treated as a gray level and is a combination of local features of visual perception. The second-order features include Gray Level Co-occurrence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GLRLM), Neighboring Gray Tone Difference Matrix (NGTDM) and Gray Level Dependence Matrix [78, 79]. These kinds of statistical features reflect the inhibition between different parts of the tumor region by quantifying the difference in grayscale regions in the image. For example, the GLDM features quantify the grey level dependences in an image. The function as below:

$$Dependence\ Non - Uniformity\ Normalized = \frac{\sum_{j=1}^{N_d} \left(\sum_{i=1}^{N_g} P(i, j) \right)^2}{N_z^2} \quad (2.2)$$

Where N_g be the number of discrete intensity values in the image, N_d be the number of discrete dependency sizes in the image, N_z be the number of dependency zones in the image, which is equal to $\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i, j)$, and $P(i, j)$ be the dependence matrix.

2.6. Feature Selection

In the practical application of machine learning, massive features are usually extracted, and there may be irrelevant features. There may also be interdependencies between features, which may lead to the following consequences: 1) The more features extracted, the longer it takes to analyze features and train the model; 2) The more features are extracted, the more complex the model will be and the lower the promotion ability.

Feature selection can reduce redundant information by removing invalid features, thereby improving model prediction capabilities. On the other hand, the model is simplified by selecting the optimal features to facilitate understanding and visualization of the data.

Feature selection has two main functions: 1) Reduce feature quantity and dimension reduction, make the model generalization ability stronger, and reduce over-fitting; 2) Enhance understanding between features and eigenvalues.

In this section, I will simply introduce several feature selection algorithms:

Removing Features with Low Variance: This should be the simplest feature selection method: suppose the eigenvalues of a feature are only 0 and 1, and in all input samples, 95% of the instances have a value of one, then it can be considered that this feature does not work. Big. If 100% is one, then this feature is meaningless. This method can be used when the eigenvalues are discrete variables. If it is a continuous variable, the continuous variables need to be discretized before they can be used. In practice, generally, 95% or more will take a certain value. The characteristics exist, so this method is simple but not very useful. It can be used as a pre-processing for feature selection. First, the features with small changes in value are removed, and then the appropriate feature selection is selected from the feature selection methods mentioned below for further feature selection.

Univariate Feature Selection: Univariate feature selection tests each feature measures the relationship between the feature and the response variable and discards bad features based on the score. For regression and classification problems, features can be tested by means of the chi-square test. This method is relatively simple, easy to run,

easy to understand, and usually has a good effect on understanding the data (but not necessarily effective for feature optimization and generalization); there are many improved versions and variants of this method.

Radom Forest: First, it will shuffle the original features and add them to the dataset as new features. Then, the random forest model is trained based on all features and the importance of each feature is evaluated. In each iteration, the method detects whether the real feature is more important than its shadow feature and removes features with the least difference in importance. Finally, the algorithm stops after all features are judged to be important or useless (or stops after a given number of iterations).

Recursive Feature Elimination (RFE): The `coef_` attribute returned by the learner or the `feature_importances_` attribute is used to get the importance of each feature. Then, remove the least important feature from the current feature set. This step of recursively is repeated over the set of features until the desired number of features is finally reached. This algorithm optimizes the model by continuously recursively removing features and re-establishing the model on the features. The stability of an RFE depends largely on which model is used at the bottom of the iteration. For example, if the ordinary regression used by RFE is unstable without regularization, then RFE is unstable; if Ridge is used, and the regression with Ridge regularization is stable, then RFE is stable.

2.7. Machine-Learning Models Establishment and Prediction

As shown in **Figure 7**, we can directly see that the normal procedure in biomedical imaging analysis field via machine-learning algorithms. As mentioned before, several procedures have been adopted include multi-modality biomedical images acquisition, the selection of region of interests, features extraction and optimal feature selection before establish predictive machine-learning models.

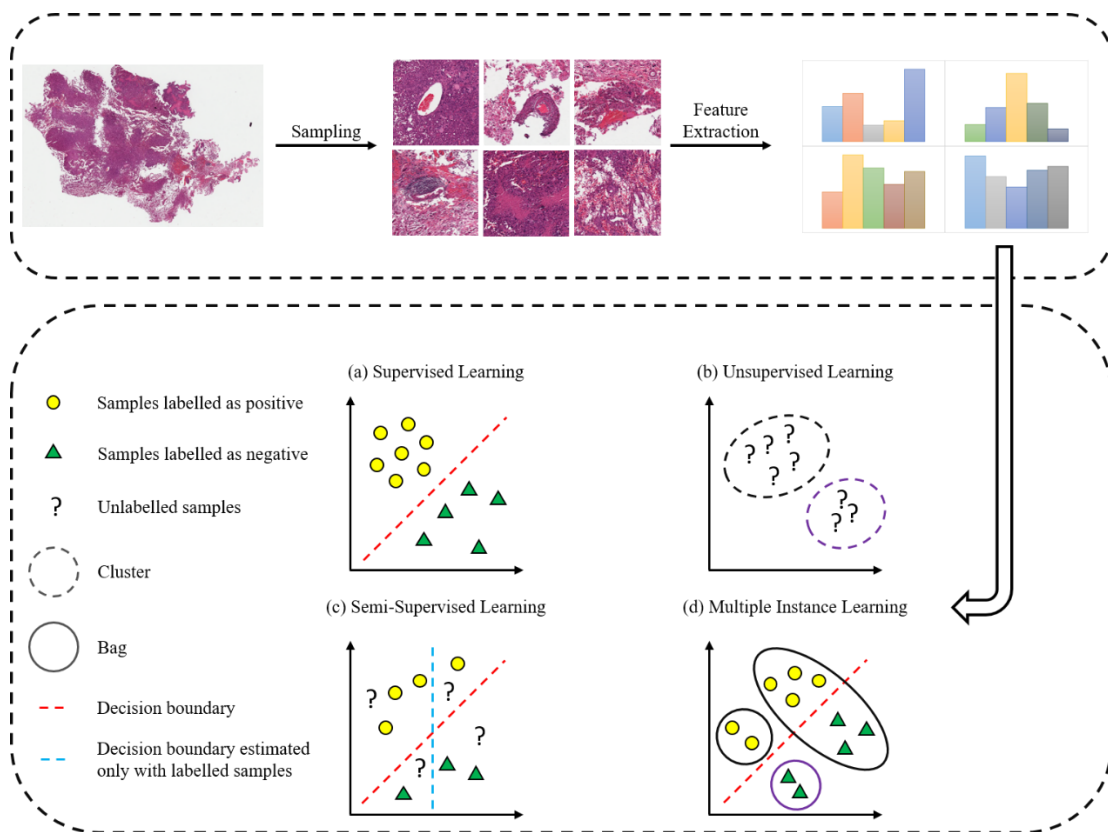


Figure 7 Prediction Model Establishment via machine learning algorithms. After the biomedical imaging pretreatment process, the commonly used machine learning algorithms in the biomedical imaging analysis field can be divided into

the following categories: (a) supervised learning, (b) unsupervised learning, (c) semi-supervised learning, and (d) multiple instance learning

As figured out in recent researches, machine-learning algorithms implemented into computer-aided biomedical imaging analysis field can be simply divided into the following categories: supervised learning, unsupervised learning, semi-supervised learning and multiple instance learning.

Supervised Learning: Supervised learning is a machine learning method in which a training model can be learned or established and a new instance can be inferred from this model. Training data consists of input objects (vectors) and expected outputs. The output of a function can be a continuous value (regression analysis) or a prediction of a classification label (classification).

The task of a supervised learner, after observing some training paradigms (inputs and expected outputs), predicts the output of this function for any possible input values. To achieve this goal, learners must generalize from existing sources to non-observed situations in a "reasonable" (see induction bias) manner. In human and animal perception, it is often referred to as concept learning.

Supervised learning has two models of morphology. Most generally, supervised learning produces a global model that maps input objects to expected output. The other is to implement this correspondence in a regional model. (such as case inference and nearest neighbor law).

The most widely used classifiers are artificial neural networks [80, 81], support vector machines [82], nearest neighbor [83], Gaussian mixture models [62], naive Bayesian methods [84], decision trees [85], and radial basis function classification [86].

Unsupervised Learning: No samples were labelled. In fact, unsupervised learning can't carry out targeted classification, fitting, and can only do some analytical tasks, such as clustering [87, 88], PCA [89], dictionary learning [90, 91].

Unsupervised Learning is an algorithm of artificial intelligence networks. Its purpose is to classify the original data in order to understand the internal structure of the data. Unlike the supervised learning network, the unsupervised learning network does not know whether the classification result is correct when learning, that is, it is not supervised (it tells it which learning is correct). It is characterized by providing input examples for only one type of network, and it automatically finds its potential category rules from these examples. After learning and testing, you can also apply it to new cases.

A typical example of unsupervised learning is clustering [87]. The purpose of clustering is to bring together similar things, and we do not care what this class is. Therefore, a clustering algorithm usually only need to know how to calculate the similarity and can start working.

Semi-supervised Learning: One part has a mark, and the other part only has data not marked. At this time, it is necessary to combine unmarked data with marked data for machine learning. This is semi-supervised learning.

The basic idea of semi-supervised learning is to use the model assumptions on the data distribution to create a learner to label unlabeled samples.

From the perspective of the sample, the semi-supervised learning problem is machine learning using a small number of labeled samples and a large number of unlabeled samples. From the perspective of probability learning, it can be understood as a study of how to use the correlation between the input edge probability and the conditional output probability of the training sample. Classifier. The existence of this connection is based on certain assumptions, namely the cluster assumption and the manifold assumption.

Multiple Instance Learning: The data is tagged, but the target of the tag is not a sample, but a packet. When the mark of a bag is negative, the mark of all samples in the bag is negative. When a bag is marked positive, at least one of the samples in the bag is marked as positive. Our goal is to learn to get a classifier so that the positive and negative signs can be given to the newly entered sample. One such type of problem is the multiple example problem. For example, to do detection problems, when marking the training picture sample, you need to give a rectangular box to indicate the position of the target. It is possible that the target is not accurate enough, resulting in different samples being misaligned. At this time, the marked rectangular frame can be partially disturbed. Get some new rectangles and think of them together as a bag, where there is always one that is the best positive sample, which is marked as positive. And take a picture without a target, as a negative sample package: no matter how the picture is intercepted inside, it is a negative sample.

CHAPTER 3. Machine Learning

Models for Multi-parametric

Glioma Grading With Quantitative

Result Interpretations

3.1. Abstract

Gliomas are the most common primary malignant brain tumors in adults. Accurate grading is crucial as therapeutic strategies are often disparate for different grades and may influence patient prognosis. This study aims to provide an automated glioma grading platform on the basis of machine learning models. In this paper, we investigate contributions of multi-parameters from multimodal data including imaging parameters or features from the Whole Slide images (WSI) and the proliferation marker Ki-67 for automated brain tumor grading. For each WSI, we extract both visual parameters such as morphology parameters and sub-visual parameters including first-order and second-order features. On the basis of machine learning models, our platform classifies gliomas into grades II, III, and IV. Furthermore, we quantitatively interpret and reveal the important parameters contributing to grading with the Local Interpretable Model-Agnostic Explanations (LIME) algorithm. The quantitative analysis and explanation may assist clinicians to better understand the disease and accordingly to choose optimal treatments for improving clinical outcomes. The performance of our grading model was

evaluated with cross-validation, which randomly divided the patients into non-overlapping training and testing sets and repeatedly validated the model on the different testing sets. The primary results indicated that this modular platform approach achieved the highest grading accuracy of 0.90 ± 0.04 with support vector machine (SVM) algorithm, with grading accuracies of 0.91 ± 0.08 , 0.90 ± 0.08 , and 0.90 ± 0.07 for grade II, III, and IV gliomas, respectively.

3.2. Introduction

Gliomas are the most common primary malignant brain tumours in adults, accounting for 30% of all primary central nervous system (CNS) tumours and 80% of all malignant brain tumours [92]. According to their histopathological characteristics, such as cellularity, pleomorphism, nuclear atypia, necrosis, and endothelial proliferation, gliomas can be classified into WHO grades I–IV [93]. Correctly differentiating tumor grades is critical because they are widely used to predict patient outcomes and determine the use of adjuvant therapy protocols including aggressive radiotherapy and concomitant chemotherapy [94]. A precise prediction of the prognosis depends on an accurate pathology diagnosis. Compared to glioblastoma (GBM), low-grade gliomas (LGGs; WHO II) have greater chemotherapy sensitivity and a better post-therapy prognosis [95-97].

Considered as one of the “gold standards” and conventionally used by pathologists for tumor grading, haematoxylin and eosin (H&E) stained images provide specific histopathological characteristics and patterns to differentiate gliomas grades [98]. The parameters extracted by clinicians from WSIs mostly focused on the cellular level, including the number and size of nuclei, which reflects the proliferation and

differentiation of the nuclei per unit area and heterogeneity of different levels of the tumor cells. For instance, compared to grade II gliomas, grade III, and grade IV both exhibit microvascular proliferation (MVP), indicating the presence of proliferation of enlarged blood vessels in the tissue [99]. Grade IV gliomas can be further distinguished from grade III by examining H&E stained images for the presence of highly pleomorphic cells with hyperchromatic, irregular nuclei and brisk mitotic activity. Grade IV gliomas are often more mitotically active, necrosis prone, and generally associated with neovascularity, infiltration of surrounding tissue and a rapid postoperative progression [9]. However, the conventional grading procedure based on H&E pathology images is often subjective and operator-dependent and of low reproducibility due to the inter-observer variability.

Apart from H&E stained images, immunocytochemical staining with Ki-67 antibodies has been widely accepted as an alternative reference for assessing the proliferative potential in tumor cells. Strictly associated with cell proliferation, Ki-67 nuclear antigen is present during all active phases of the cell cycle (Gap 1, Synthesis, and Gap 2 phases of the cell cycle) and mitosis but absent in resting (quiescent) cells (Gap 0) [100]. The proliferative index (PI), determined by Ki-67 immunohistochemistry (IHC), correlates well with the histopathological malignancy grade of gliomas [101]. PI values may assist in differentiating grade II from III or grade II from IV; however, due to the overlapping PI values of grades III and IV, PI itself would not be considered as sufficient evidence to adequately determine these two malignancy grades [101].

Computer-assisted grading systems are capable of mining a large number of quantitative features from digital pathology slide images to sort out the most important patterns for grading. This ability provides the opportunity for better quantitative

modeling of the disease appearance and hence possibly improves the prediction accuracy of tumor grading. In addition, it also provides a more reproducible, less labor-intensive and more efficient mechanism than manual grading by pathologists. However, the compelling opportunities offered by big digital pathology data come with optimal computational algorithm challenges [102]. For example, image analysis and computer-assisted detection models inadequately address the data density in high-resolution digitized whole slide images (WSI). Mousavi et al. (2015) [103] proposed to distinguish LGGs from GBMs using image features extracted from regions of interest (ROIs) in WSIs. However, the manual selection of ROI in this method may introduce a potential risk of inter-observer variance. Kong et al. (2009) [104] tiled the WSI into non-overlapping pieces and analyzed each image tile to grade LGGs from GBMs. While fully processing all the image tiles, the method might be less computationally efficient, in particular, for high-resolution H&E stained images. To improve computational efficiency of the automated prognosis of neuroblastoma from H&E images, Sertel et al. (2009) [105] proposed a multiresolution approach to extract local binary pattern and texture features of different scales.

While the above mentioned automated methods process and analyze WSIs for grading, the method proposed by Ertosun and Rubin (2015) [80] classified LGGs and high-grade glioma (HGGs) on the basis of automated segmentation and analysis of cell nuclei and morphological features [106] but neglected the tumor patterns important for manual grading. To systematically compare the performance of classic machine learning models on grading, Huang and Lee (2009) [107] implemented three machine learning models, Bayesian, k-nearest neighbors (KNN) and support vector machine (SVM), and concluded that KNN and SVM both achieved the highest accuracy. Apart from

conventional machine learning, Reza and Iftekharuddin (2016) [108] proposed a deep learning-based grading system, characterizing each tile type with convolutional neural network (CNN). However, these machine learning and deep learning models are considered as “black boxes” without transparent interpretation of either the models themselves or the grading results.

To improve reproducibility and avoid inter-operator variance in conventional manual grading, in this study, we design and implement a platform for automated grading gliomas into grades II, III, and IV from digital pathology images. In our approach, discriminative visual, sub-visual and IHC parameters are identified, and a reliable machine learning model is selected. With the machine learning based models, we integrate information from both histopathological morphology images and proliferation biomarkers into a single unified framework to predict the glioma grade in 116 patients, which surpass the current clinical paradigm for patients diagnosed with glioma. Our platform provides interpretation on the grading outcomes to disclose the contributions of multiparametric features to individual cases and presents an alternative for objective, accurate, and interpretable prediction of glioma grading in the clinic.

3.3. Materials and Methods

The dataset used in this study involves 146 cases of glioma grading from grades II to IV: 49 grade II, 45 grade III, and 52 grade IV images. All cases were from Shandong Provincial Hospital affiliated to Shandong University, and the pathology diagnoses of the cases were based on the 2007 WHO standards [93]. Paraffin-embedded samples were cut into three μm thick sections and stained with H&E stain. All H&E images in this study were obtained from WSIs scanned by a Leica SCN400 slide scanner (Leica

Biosystems, Nussloch, Germany) with multiresolution varying from 20× to 40×. Ki-67 immunohistochemical staining was prepared using an automated staining instrument (Ventana, Benchmark Ultra). As defined in Eq. 1, Ki-67 PI is the percentage of the number of immunoreactive tumor cells in relation to the total number of cells. At least 1000 tumor cells or alternatively, three high-power fields (HPFs) were examined by two independent experienced observers. The mean of Ki-67 PI is the average of the values calculated by different observers.

$$PI = \frac{\text{Positive Cells}}{\text{Total Cells}} \times 100$$

3.4. Automated interpretable Glioma grading with Machine-Learning Models

As illustrated in **Figure 8**, our automated grading framework is composed of five major components including automated ROI identification, feature extraction, important feature selection, automated grading, and result interpretation.



Figure 8 Schematic flowchart of the automated grading framework. We first automatically selected the representative regions of interest (ROIs) from the H&E images. Based on these ROIs, we extracted and selected important visual, sub-visual, and immunohistochemically features. We established automated machine learning models with these features for glioma grading. The grading results output from the model were further explained with the LIME algorithm.

3.4.1. Automated ROI Identification

Our proposed computational method automatically identifies the ROIs that reflect cell proliferation and cellular density based on the number of nuclei in the regions. H&E images were first partitioned as tiles, each with a resolution of 5120*5120 pixels to enable the process of high-resolution imaging [104, 109]. Then, we detected nuclei using watershed nuclei detection algorithm in each tile [110, 111] and based on the density of the detected nuclei, the five tiles with the highest densities of nuclear were identified as the ROIs.

3.4.2. Multi-Parameter Extraction and Important Parameter Selection

From the identified ROIs, we extracted multi-parameters including visual features and sub-visual features as the inputs for automated grading. The visual features, including seven nuclear morphological features, five nuclear staining features, and nuclei clusters or patterns, were extracted to reflect a basis for observation when pathologists make diagnostic decisions [106]. For instance, nuclear morphological characteristics such as shape, size, and circularity that reflect cellular atypia have been commonly used by pathologists to distinguish different grades of glioma.

In addition to the visual features, sub-visual features also contribute to accurate glioma grading. The sub-visual features are the computerized high-throughput intensity and texture image features that have been proved to have diagnostic, predictive, and prognostic power, although these features are somehow beyond human perception capabilities [112]. In our method, for instance, intensity features describe the first-order statistical information of the image intensity distribution, while the second-order grey-level co-occurrence matrix (GLCM) features capture both statistical intensity and relationship between neighbourhood pixels, revealing information such as homogeneity, contrast and entropy [113]. As illustrated in Figure 6, we used a random forest (RF)-based feature selection method combined with backward feature elimination to choose the most representative and informative features [114].

3.4.3. Machine Learning Model for Automated Grading

To select the best machine learning model for grading, we compared the performance of four types of classic machine learning models including RF [115], gradient boosting decision tree (GBDT) [116], SVM [117] and neural network (NN) [72, 118, 119]. To achieve the best performance for each grading model, the hyper-parameters of the model need to be optimized, for example, the tree number in RF model. The framework for automatically tuning the hyper-parameters is provided in **Figure 9**.

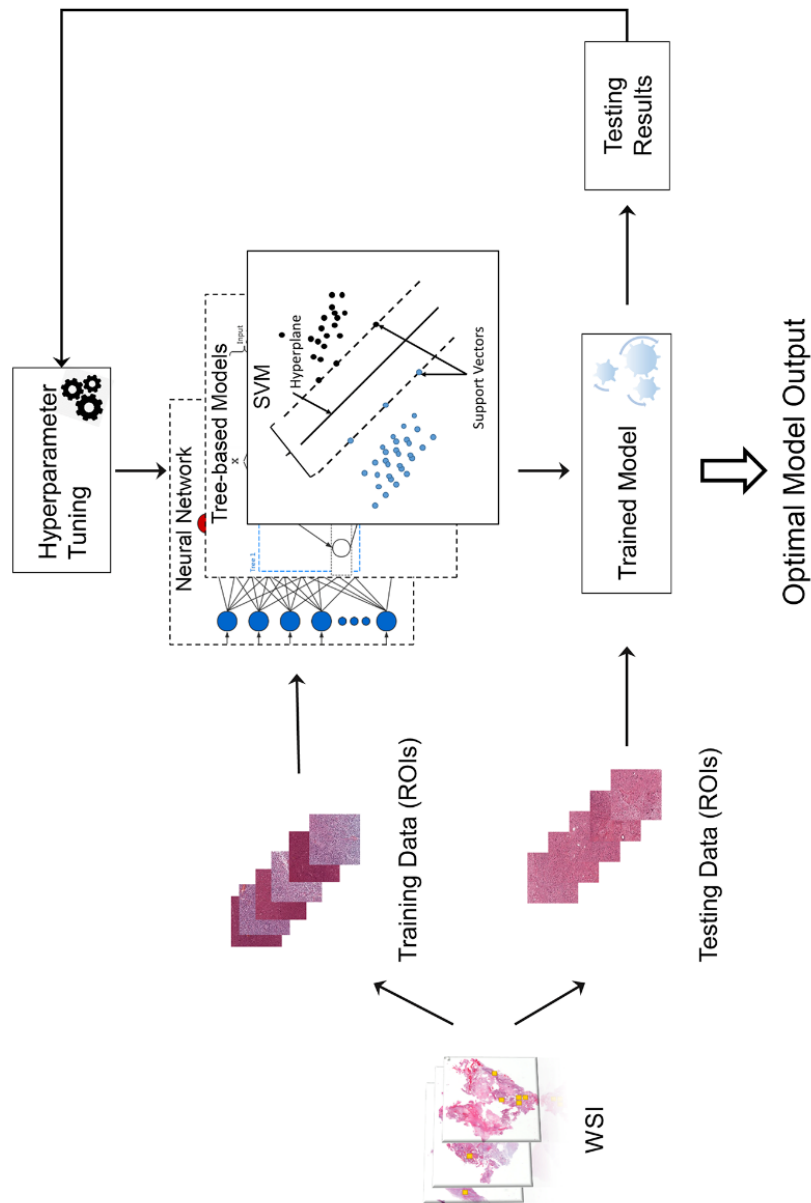


Figure 9 Machine learning model with hyper-parameter tuning. The original data were randomly separated into training data and testing data. Training data were used to train the machine-learning model, while testing data were used to evaluate the model performance. The hyper-parameters of the model can be continuously tuned with information acquired from testing results until the optimal model is obtained.

3.4.4. Validation and Evaluation

The 116 patients out of 146 was used to train the model while the other 30 patients was used for further validate the model. The 116 patients were randomly divided into training datasets (75%) and testing datasets (25%) to test performance and validate the models [120]. The corresponding training set was used to train the classifier, and the test set was used to verify the model performance. This procedure was repeated 30 times by changing random states while maintaining the same train-test split ratio. Apart from accuracy, other measurements including precision, recall, F1 score and confusion matrix were calculated to more completely validate grading performance. In addition, extra 30 cases were used as validation dataset with 30 times of cross-validation to further test the performance of our proposed model.

3.5. Results

3.5.1. Multi-Parameter Extraction and Selection

In the multi-parameter extraction phase, we extracted 24 visual parameters and 171 sub-visual parameters. Since the large size of the sub-visual features may suppress important morphological features, we performed the feature selection process on the sub-visual features twice to eventually choose 15 features from this category. The predictive performance of the selected 11 visual features, 15 sub-visual features and one IHC characteristic – Ki-67 in this study – were assessed with separate RF models. The box-plot in **Figure 10A** shows that visual features had the highest predictive power with 0.76 accuracy, while the accuracy for sub-visual features and Ki-67 only reached 0.62 and 0.53, respectively.

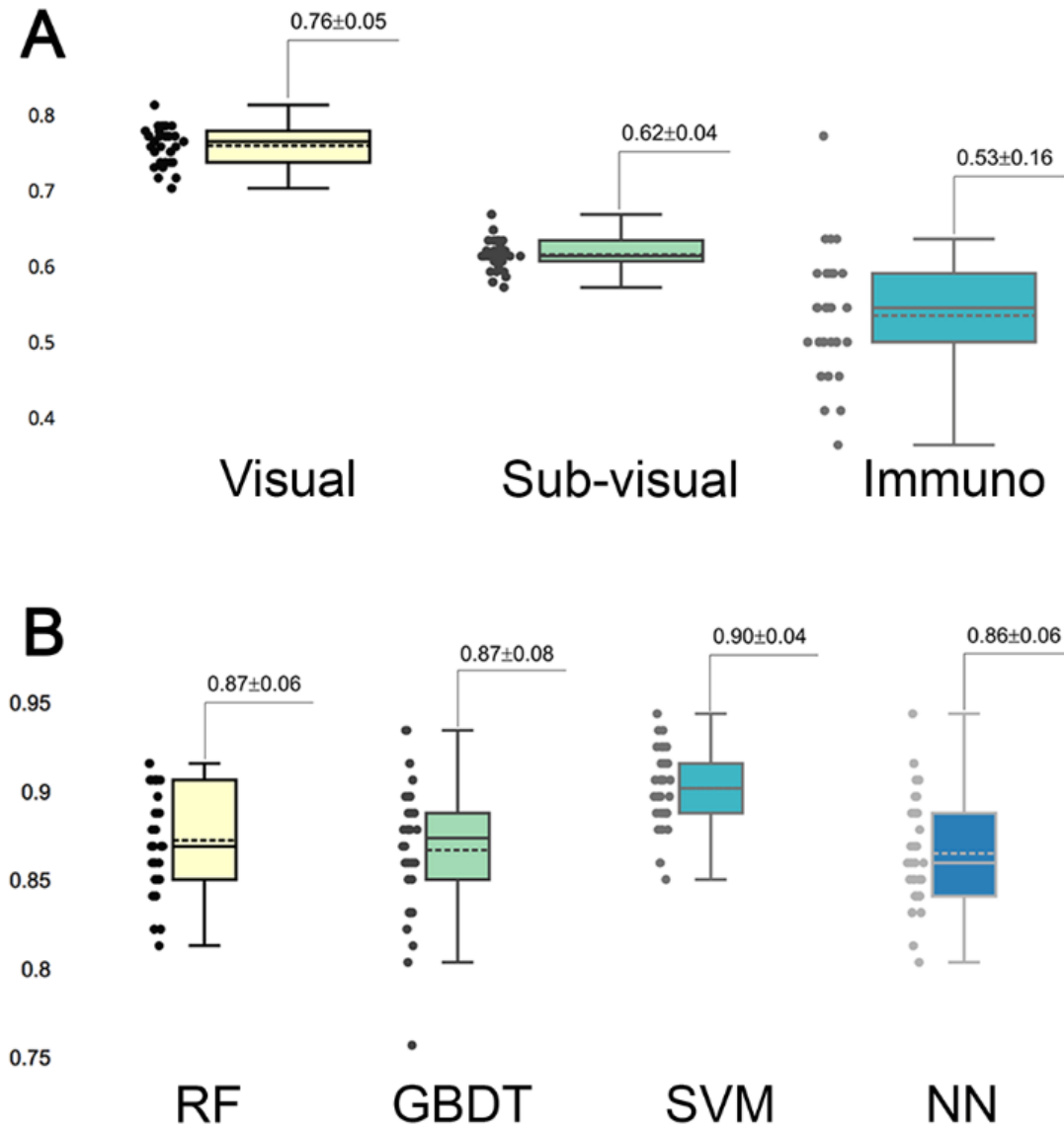


Figure 10 Predictive capability of selected important features and models. (A) Predictive accuracy of different categories of selected important features. In each category, features were assessed for accuracy separately. (B) Accuracy of different grading models.

3.5.2. Automated Grading Performance with Different Models

We fed these selected important parameters into the four machine learning models RF, GBDT, SVM, and NN to sort out the best grading model. Based on cross-validation,

SVM achieved the most accurate and stable grading performance, with the accuracy of 0.90 ± 0.04 , while the other three models had lower but nearly the same performance with accuracy stabilized at approximately 0.87 (**Figure 10B**). Table3 shows the results of other measurements including F1, precision and recall, among which the SVM model maintained the best performance.

Table 1 F1, accuracy, precision, and recall for different machine learning models.

Method	F1	Accuracy	Precision	Recall
Random Forest	0.86 ± 0.07	0.87 ± 0.06	0.86 ± 0.07	0.87 ± 0.06
GBDT	0.86 ± 0.08	0.87 ± 0.07	0.86 ± 0.07	0.86 ± 0.07
Neural Network	0.86 ± 0.05	0.86 ± 0.06	0.87 ± 0.06	0.87 ± 0.05
SVM	0.90 ± 0.07	0.90 ± 0.04	0.91 ± 0.04	0.91 ± 0.04

3.5.3. Further Investigation of SVM Performance

To better understand the performance of the SVM model for grading histologic grades of glioma II, III, and IV, we calculated a confusion matrix to uncover misclassification cases. As illustrated in **Figure 11A**, grade II had the highest accuracy (0.91) in the SVM model while the other two grades both had the accuracy of 0.90. As shown in the confusion matrix for the SVM model (**Figure 11B**), grade III gliomas may share common histological characteristics with both grade II and grade IV gliomas and thereby may lead to misclassification. In addition, we added 30 new patients of grade

II to IV for the further validation. The validation performance achieved 0.88 ± 0.14 with 30 times cross-validation.

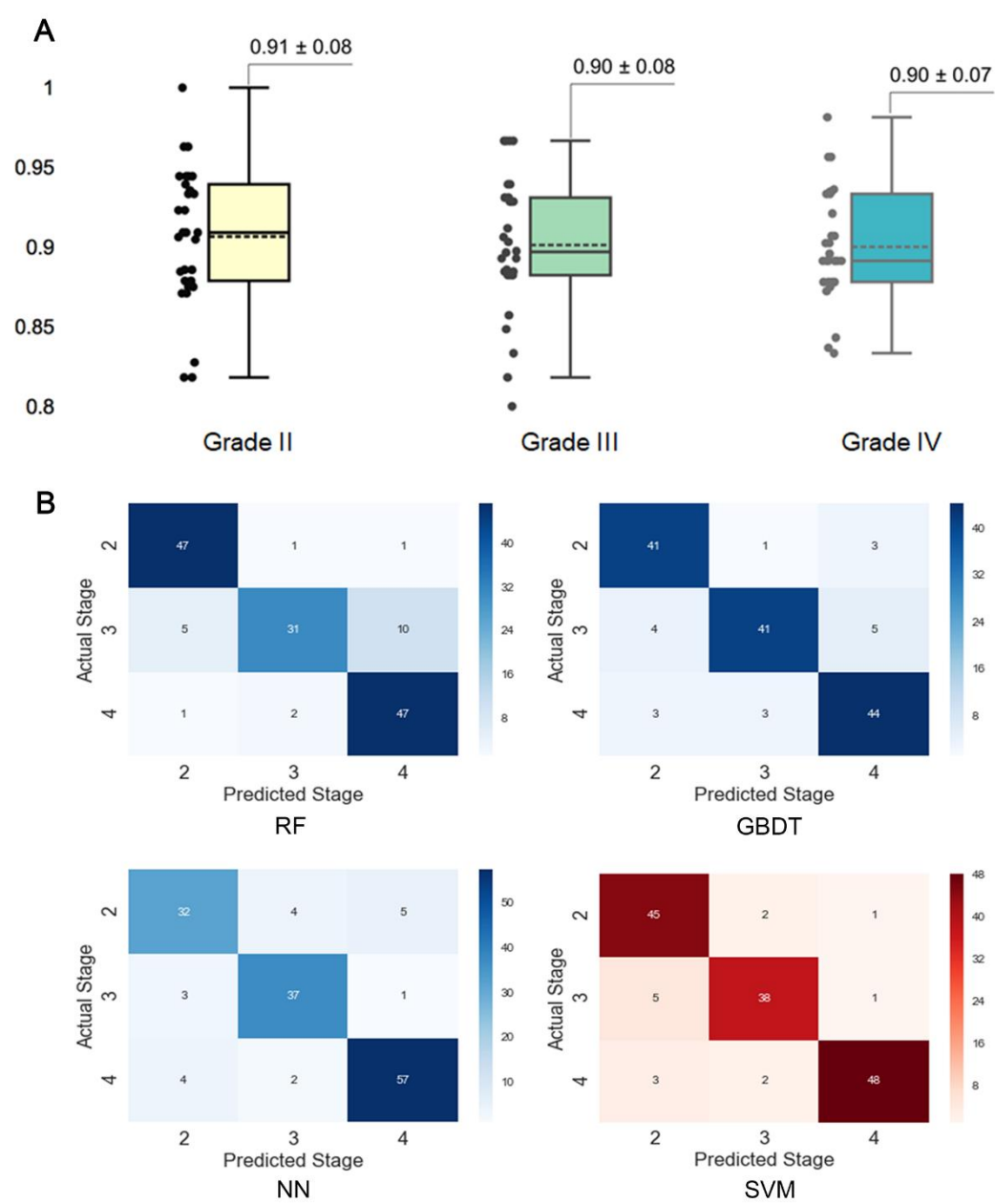


Figure 11 Assessment of grading results with the support vector machine (SVM) model. (A) Prediction accuracy for different histological grades in the SVM model. The accuracy for each grade was obtained by separating the results by grades in

the 30 times cross-validation. (B) Confusion matrixes for different models. The confusion matrixes reveal the number of misclassified cases for each grade.

3.5.4. Quantitative Grading Results Interpretation

The machine learning algorithm is usually complex and considered as a “black box” without explicit interpretation of the learning process or the outputs. To provide a better quantitative interpretable explanation for the grading results, we used Local Interpretable Model-Agnostic Explanations (LIME) algorithm [121] to reveal the importance of features and their underlying contribution to the grading decision. In this section, we illustrated three case studies to interpret the automated grading results and the underlying reasoning for our SVM model with the LIME algorithm.

As Ki-67 has been proved to be a clinically significant indicator for grading, we predicted the grading results by machine learning models while also taking the influence of Ki-67 into consideration. To highlight this strategy, we chose three cases to elucidate the workflow of SVM for precise grading. As shown in **Figure 12A**, Case 1 is a relatively easy case to classify, being grade II with low Ki-67 (0.05). However, Case 2 and Case 3 are more ambiguous, both with a value of 0.2 for Ki-67, but one is grade III while the other is grade IV. We established a result explainer with LIME algorithm for the SVM model and then calculated the feature contribution for each case.

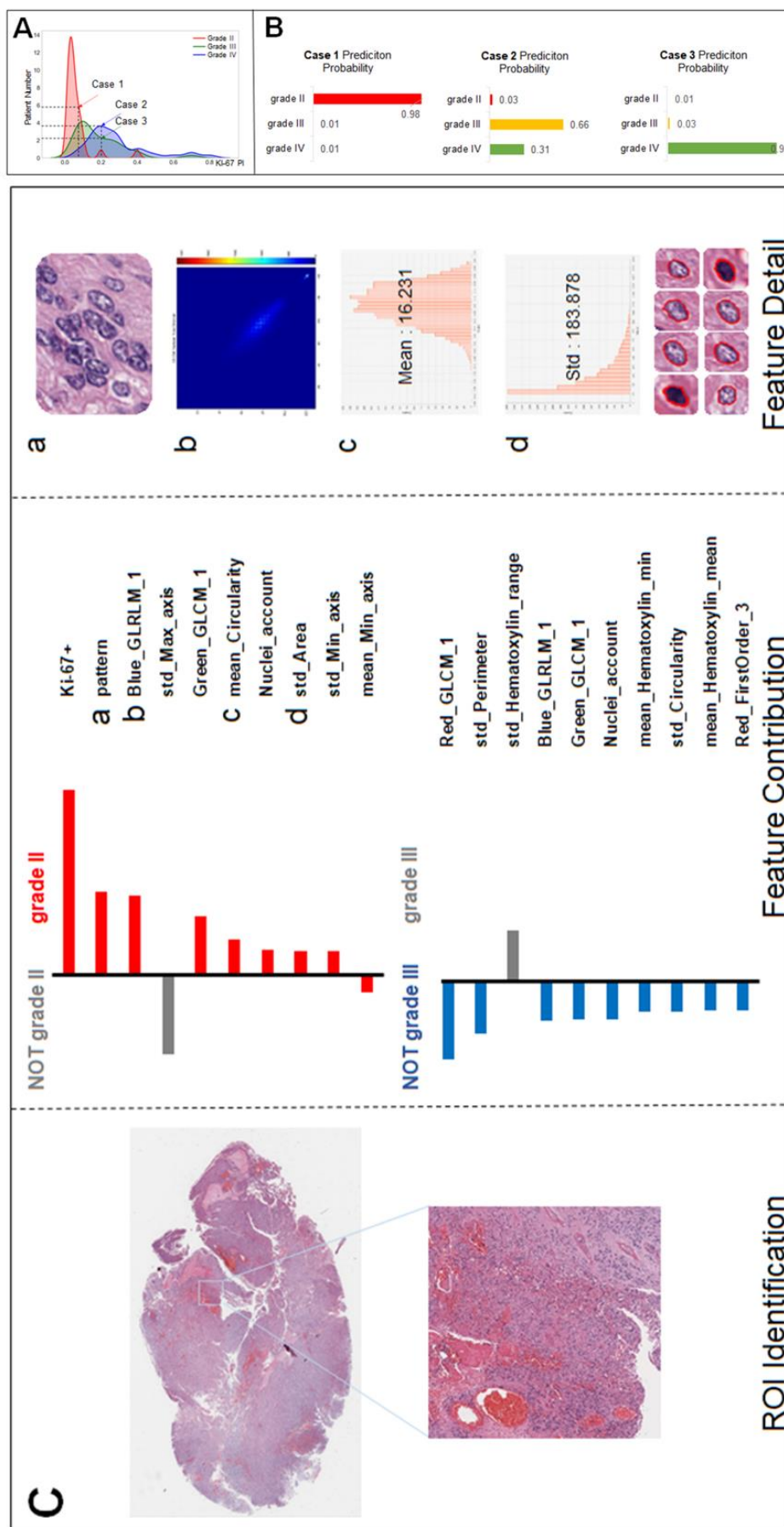


Figure 12 Grading result explanation of representative cases with the LIME algorithm for the SVM model (part 1) (A) Distribution of Ki-67 by different grades. Three selected cases had Ki-67 PI values of 0.08, 0.2, and 0.2. (B) Prediction probability for three cases.

Case 1 (**Figure 12C**) was correctly classified as grade II with a high probability. As Ki-67 has been clinically proved to have the discriminative capacity to distinguish between LGG and HGG, it is not surprising that Ki-67 has the dominant influence underlying this correct decision. Apart from Ki-67, visual features such as cell pattern and sub-visual features such as GLCM texture features also can correctly guide the grading. The result also shows that although some morphological features such as perimeter do not support the decision of grade II, they excluded grade III, which also contributed to accurate grading.

For the more ambiguous Case 2 (**Figure 13**) and Case 3 (**Figure 14**), with regard to Ki-67 itself, our classification model demonstrated its capability for an accurate decision. Case 2 (**Figure 13**) was the most challenging case among these three cases because in this case, Ki-67 had only a negative influence on the grading decision; however, with the strong support of GLCM texture features and the morphological nuclei count feature, our grading system made a correct decision. In Case 3, Ki-67 was the major contributor to the correct decision, while the morphological features such as standard deviation of cells' max axis and standard deviation of cells' perimeter were the second and third contributors.

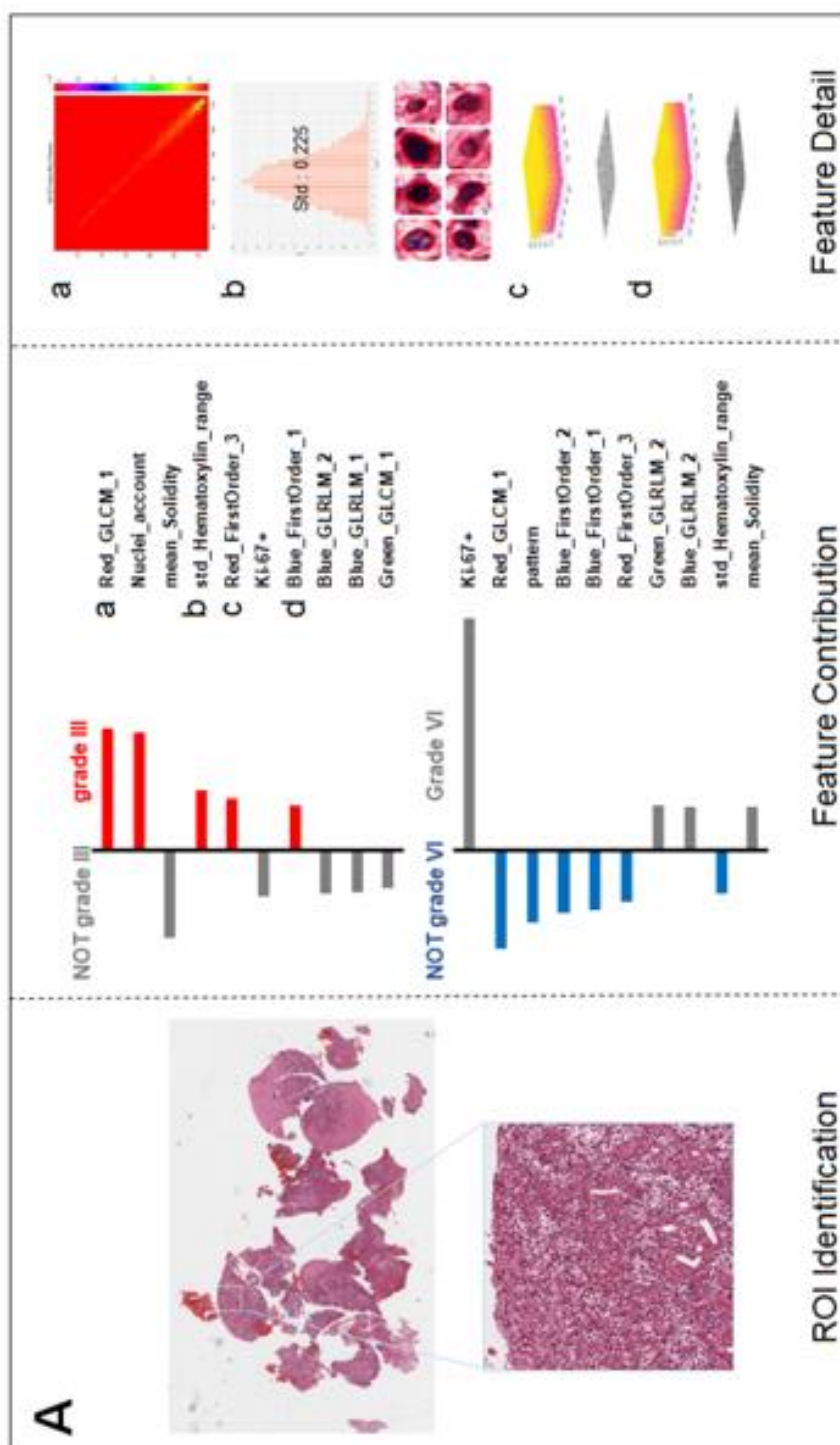


Figure 13 Grading result explanation of representative cases with the Local Interpretable Model-Agnostic Explanations (LIME) algorithm for the SVM model (part 2) (A) Glioma Case 2 for grade III with Ki-67 PI value of 0.2 while texture features and morphological nuclei count feature were the major contributors.

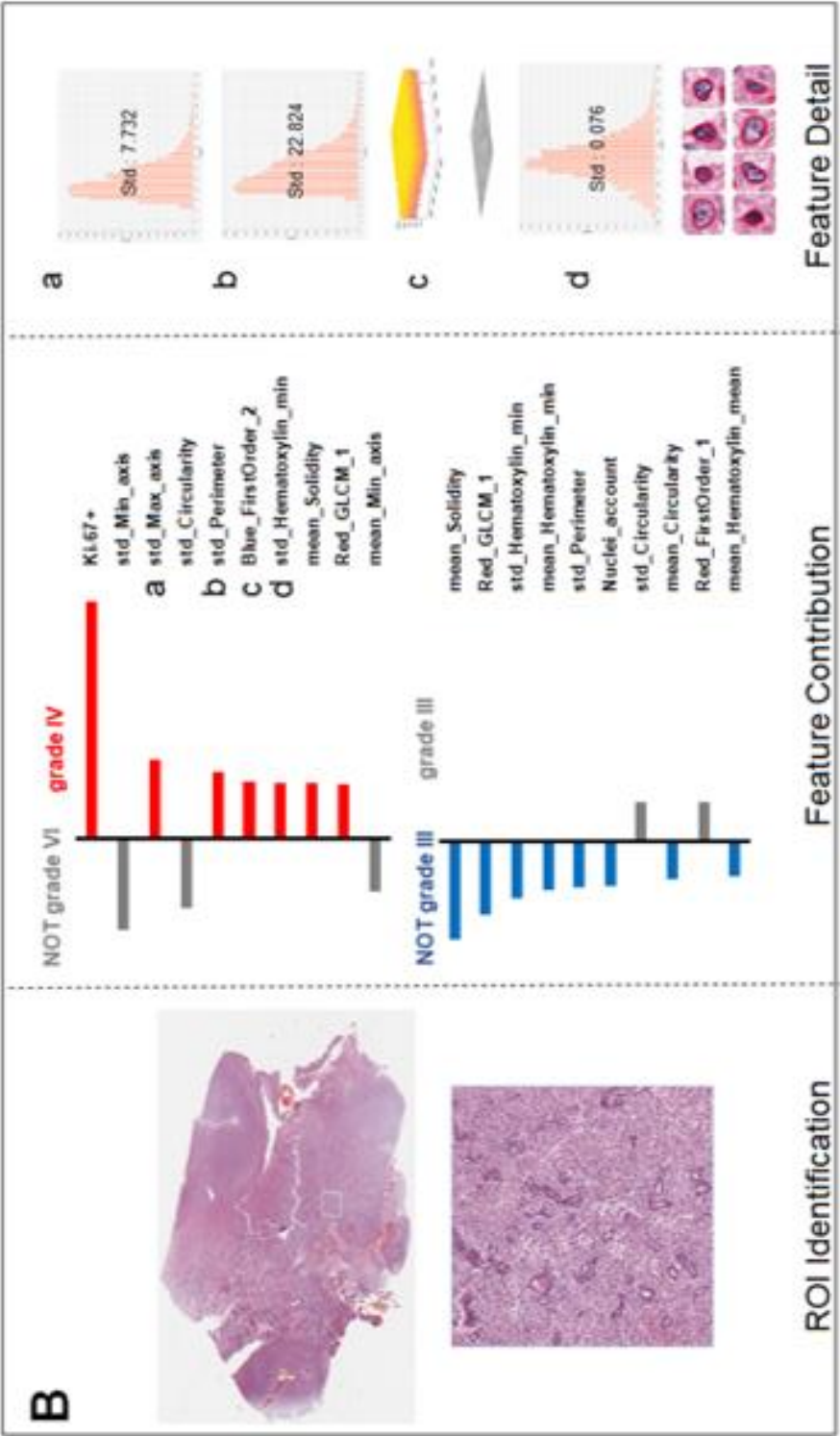


Figure 14 Grading result explanation of representative cases with the Local Interpretable Model-Agnostic Explanations (LIME) algorithm for the SVM model

(part 3) (B) Glioma Case 3 for grade IV also with Ki-67 PI value of 0.2 accounting for the dominant contribution.

3.6. Discussion

Given the well-recognized difficulties associated with the grading of gliomas, in this study, we developed an automated and interpretable grading of gliomas with a machine-learning platform to provide an additional reference for diagnostic decision-making and patient management and prognosis, which is more efficient, effective, and objective than currently used methods. Our major findings were that (1) our method achieved a high grading accuracy of 90% for classifying gliomas into grades II, III, and IV; and (2) more importantly, for the first time, an interpretable insight into the grading outputs was provided for the histology grading system with the “black-box” machine learning models.

Computerized image perception in tissue histology is much more difficult due to the complexity of the cell patterns and inherent structural associations between different tissue components. Accurate histopathological diagnosis determines patient management, treatment, and follow-up, so any method that results in more objective glioma grading may be of great value. In the current literature, automated grading either was based on texture patterns extracted from the selected ROIs [103] to capture the histological structures in the histological images or utilized morphological features extracted from segmented tumor cells [80, 105]. Our method achieved high accuracy through fully utilizing the selected informative visual and sub-visual features from H&E images combined with the Ki-67 PI value. To take into account the visual cues used by pathologists for manual grading, visual features such as texture patterns and

cell morphological characteristics were extracted from selected ROIs. The sub-visual features that are beyond human perception capacities can help the precise grading through quantitative analyses that are more comprehensive. Furthermore, we innovatively integrated the IHC indicator Ki-67 PI into our grading platform, which contributed to boosting our grading accuracy with high discriminatory power.

Generally, the aim of this research was to objectify, standardize and quantify features that are already widely accepted as important by pathologists. Inter-observer discrepancies, which had long been a diagnostic problem in the past, could be partly overcome by employing the morphometric analysis used in the present study [122]. The ability of the system to objectively identify regions of tumor may additionally complement the pathologist's diagnosis and assist in tailored treatment. In this study, we attempted to automatically clarify the cellular composition and histopathological features of different grades of gliomas by utilizing morphometric- and immunohistochemically-based machine learning models. We assessed the importance of visual features, sub-visual features and Ki-67 separately, and the highest predictive power was achieved by visual features, achieving an accuracy of 0.76, which indicates that the histomorphology of glioma, such as nuclear morphological features and nuclear staining features and patterns play an important role in glioma grading.

The interpretability of the individuals' grading results provided quantitative insights into the feature contributions of each individual case. Most of the grading systems based on machine learning models are considered "black boxes," and it would be valuable for patient management if clinically trusted reasoning could be revealed. In our study, we used the LIME algorithm to provide an explanation for each individual grading result. Case 1 and Case 3 in **Figure 12C** and **Figure 14** endorse the important role that Ki-67

plays in grading decisions, which is in accordance with the clinical fact that Ki-67 PI is positively correlated with the histopathology grades [123]. However, Ki-67 sometimes has a detrimental influence on the grading process, as shown in Case 2, because Ki-67 PI of grade III gliomas largely overlaps with that of grade IV [101, 124-126]. As a result, Ki-67 itself cannot be the only determinant of grading. It has been accepted that the Ki-67 PI value is $3.0 \pm 2.1\%$ in grade II gliomas, $11.8 \pm 3.4\%$ in grade III anaplastic gliomas and $15.8 \pm 7.4\%$ in grade IV GBMs [126]. Furthermore, the Ki-67 PI values of grade III anaplastic glioma can overlap with values of grade II glioma at one end of the range and with those of GBM at the other in the diagnostic practice of pathologists. Apart from Ki-67, morphological features including the standard deviation of cells' max axis and perimeter were also found to be a significant contributor to glioma grading because HGG often exhibits strong heterogeneity with irregular cell shapes [9, 127]. The precise morphological features extracted from the most aggressive regions help to investigate intratumoral histological heterogeneity for precise histopathology grading [128]. The primary grading results still support the conclusion that the only sure way to determine the histopathological WHO grade remains the histopathology evaluation of the H&E stained tumor sample [129].

With three representative cases we could gain a deeper understanding of why misclassified grade III cases are often labelled as grade II, while grade IV cases could be wrongly labelled grade III as shown in the confusion matrix. Actually, the Ki-67 PI values of the GBM group can be as low as those for grade II tumours, indicating the limitation of Ki-67 values in the overlap region. In our work, we introduced a number of histological sub-visual features such as the intensity and GLCM texture features to distinguish these three grades of gliomas. These insights become easy once it is

understood what the algorithm models are actually doing, which in turn leads to models that generalize much better results. In many applications of machine learning, users are asked to trust a model to help them make decisions. There has always been a focus on “trust” in any type of modelling methodology, but with machine learning, many people feel that the black-box approach taken with these methods is not trustworthy. Through this machine learning-based approach, we could use the LIME algorithm to explain individual predictions to the decision-maker (the pathologist), and that understanding of the model’s predictions can be an additional useful tool when deciding whether a model is trustworthy or not for the final diagnosis from a pathologist.

Our study did, however, have its limitations. First, grade I gliomas are not included in our studies. They are accurately considered benign in clinical practice, in that complete surgical excision is considered curative. Therefore, grade I gliomas are different from grade II-IV gliomas in biological behaviour. The results of our discrimination of grade III and grade IV are just reasonable preliminary results but leave much room for improvement. Considering that necrosis is one of the remarkable features of GBM, we plan to use cell necrosis as an input feature to further train models to distinguish grade III from grade IV GBM. Actually, morphometric data research has indicated that the cellularity of oligodendroglioma’s type II was significantly higher than that of diffuse astrocytoma’s and that the conditional entropy of oligodendroglioma’s type III was significantly lower than that of diffuse astrocytoma’s [123] so further stratification of LGG (grade II) will lead to tailored glioma management according to their different biological behaviour. Hence, future work will focus on improvements utilizing larger datasets, including multi-centre cases. In addition, there is still much room to improve the grading performance and generalization of the algorithm model.

3.7. Chapter summary

In conclusion, our approach provides an objective alternative for quantitative pathology research and for the implementation of morphological data in routine diagnostic practice. The machine-learning model utilized multi-parameters including morphometric and sub-visual parameters as well as Ki-67 PI information to ensure high accuracy, efficiency and consistency in glioma grading. Interpretable grading platform has the potential to facilitate personalized medicine in the setting of malignant gliomas. With different important features identified for different patients, specific phenotypic tumor characteristics can be uncovered for optimal treatment selection. In addition, as our method is fully automated and quantitative with high reliability, it becomes easier for our platform to be introduced in routine clinical practice.

CHAPTER 4. RA-PA-Thomics: A new approach to glioma genotype prediction

4.1. Abstract

Glioma is the common primary brain tumor in adults. Preoperative identification of isocitrate dehydrogenase 1 (IDH1) mutations in gliomas may aid the classification of glioma and predict the clinical outcome. In this study, we aim to integrate multi-modal biomedical imaging information comprising both magnetic resonance image and digital histopathological scans, to predict IDH1 phenotypes and improve the accuracy. To deal with the multi-modal biomedical images, we have developed an automated framework for fast and cost saving IDH1 phenotypes prediction. Data from 217 histopathological diagnosed glioma patients were used. In contrast to most studies that solely relied on MRI radiomics features to predict glioma IDH1 phenotypes, we have also included digitized histopathological data to enhance the accuracy. Firstly, a random forest (RF)-based algorithm was used to identify and quantitatively analyze features in multi-modal biomedical images of radiological (RA) and pathological (PATH) origin. Then we integrated these RA- and PATH-omics data employing a machine-learning algorithm (Random Forest) to test the performance of IDH1 genotyping. We found that RF-based machine-learning model, which succeeds in using a combined pattern of RA and PATH images, is capable of predicting IDH1 phenotypes with high accuracy and reliability.

This combination-based prediction algorithm, RA-PATH-omics, is a supplement to the previous method. Accuracy achieved was 0.90 with multi-model medical images compared to 0.78 (brain tumor segmentation from the multi-slices FLAIR MR images), 0.68 (edema extraction from the FLAIR images), 0.72 (brain tumor segmentation from the T1CE MR images) and 0.73 (the edema area segmentation from the T1CE MR images), respectively, 0.82 for radiomics features alone and 0.86 for PATH-omics features. The integration of RA-PATH-omics in a single combined machine-learning model provides a new way of predicting glioma IDH1 phenotypes with great accuracy. The new method could thus be useful as a routine screening procedure for patients who do not have easy access to the lab IDH1 genotyping.

4.2. Introduction

Based on studies and clinical researches, the 2016 WHO classification has considered two more molecular anomalies included isocitrate dehydrogenase (IDH) and 1p/19q into adult glioma diagnostics [9, 130]. Competing with single histopathology character, genotype features in molecular level can achieve a higher accuracy of the classification and better prediction of clinical outcome [9]. According to recent researches, key molecular markers like IDH mutation plays a dominated character in the description of brain tumors. Because the molecular diagnosis is part of the diagnostic classification, clinical doctors suggest that gliomas with IDH mutation may have more value for diagnostic and prognostic [32, 131]. Therefore, accurate IDH genotype prediction paly a dominated role in the prognosis of glioma patients.

Histopathology images have been brought into clinical cancer research and conventionally used by pathologists in the area of tumor grading [132-134]. In addition,

pathologists can observe the cellular and sub-cellular level features from whole slide images [81]. However, for the IDH genotyping, recent researches have more focused on the analysis of the multimodality MRIs. Therefore, compared with most of the current experiments, we have included pathological images in this experiment.

Although current predictions are increasingly dependent on different genomic biomarkers, histology continues to play an important role in predicting disease progression in glioma patients. Because, compared to radiology, histology reflects a variety of information at the cellular level. However, competing with computational-aid imaging diagnoses, the traditional method by different pathologists are much more subjective and unreproducible. Computationally imaging analysis based on histology images can reduce the bias between different pathologists [31, 135, 136].

Machine learning algorithms are currently widely used in glioma research, and most are based on the analysis of features extracted from MRIs [135, 137-139]. Computationally biomedical imaging analysis could mine massive features from multimodality images for IDH genotyping. Zhou et al. [28] proposed to distinguish the IDH genotype with primary grade II and III glioma using clinical and pathological variables and textures features extracted from region of interest (ROI) in four sequences of MRI, including T1W, T2W, T1CE and T2F, which reached the accuracy of 0.86. Competing with the method mentioned by Zhou, Eichinger et al. [140] have designed local binary pattern texture features for IDH genotyping, which increase the accuracy to 92%. Apart from the traditional machine learning algorithm, the method proposed by Xing et al. [141] classified IDH mutation and IDH wild type (IDHwt) using conventional machine learning algorithm to extracted deep features from 4 sequence of MRI (T1W, T2W, T1CE, T2F).

As mentioned before, there are many researches have focused on lower-grade gliomas. Zhang et al. [31] aims to distinguish IDH mutation and IDHwt by the primary grade III and IV glioma with important features (age, intensity, texture and shape features) extracted from multimodality MRIs (T1, T1c, T2, T2/FLAIR and DWI), which the accuracy arrived at 0.8830 with random forest algorithm.

In this study, we predict the IDH genotype based the preoperative MRIs and whole slide images of patients diagnosed with low- and high-grade gliomas. In our research, the integration of WSIs, multimodality MRIs and clinical information via the machine-learning algorithm help us to improve the accuracy of IDH genotyping. Through the machine-learning algorithm, we combine the different modality of biomedical images into a single framework for IDH prediction.

4.3. Materials and Dataset

4.3.1. Patient Enrolment

The imaging data of 217 subjects with glioma were collected from two different sources. 126 cases were from Shandong Provincial Hospital Affiliated to Shandong University, comprising 41 histological grade III cases and 85 histological grade IV cases. The remaining 91 cases were retrieved from The Cancer Imaging Archive (TCIA), comprising 25 histological grade III cases and 66 histological grade IV cases (**Table 2**). The criteria for image acquisition in this study are as follows: (I) available histology, age at diagnosis and sex, and IDH1 status; (II)MR imaging data, including post-contrast T1-weighted images (T1CE), and T2-FLAIR, and (III) histopathological images.

This study was approved by the institutional review board of Shandong Provincial Hospital affiliated to Shandong First Medical University (NSFC NO: 2019-272), and the informed consent requirement was waived from each enrolled patient.

Table 2 Patient Characterizes

	Shandong provincial hospital	TCGA
Grade III (<i>n</i>; %)	41	25
IDH-mutated in Grade III (<i>n</i>; % column)	20, 48.9%	17, 68%
Grade IV (<i>n</i>; %)	85	66
IDH-mutated in Grade IV (<i>n</i>; % column)	20; 23.5%	12, 18.2%
Age (years; mean; range)	49; [5, 79]	53; [18, 81]
Sex (<i>n</i> male; % column)	55; 43.7%	55; 60.4%

4.3.2. Dataset

4.3.2.1 Imaging Data Acquisition and Pre-processing

Histopathological Images: Shangdong provincial hospital's cohort: Cases were diagnosed according to WHO criteria [142]. Paraffin-embedded samples were cut into three μm thick sections and stained with H&E stain for all patients in this cohort.

All H&E images were scanned by a Leica SCN400 slide scanner (Leica Biosystems, Nussloch, Germany) with multi-resolution varying from 20× to 40× for analysis.

Genomic DNA was isolated was performed to enrich for neoplastic cells from formalin-fixed paraffin-embedded glioma tissues. DNA was extracted using the QIAamp DNA micro kit (Qiagen GmbH, Hilden, Germany) as previously described [143].

Expression of IDH1-R132H mutant was firstly analyzed by IHC as previously described [144]. IDH1 and IDH2 gene mutations were identified by Sanger sequencing [145]. IDH status was defined according to the absence of IDH1-R132H immunopositivity and/or mutations in IDH1 and IDH2 genes.

The Cancer Genome Atlas cohort: Digital pathology slides diagnosed of diffuse glioma were downloaded from The Cancer Genome Atlas (TCGA) Data Portal (<http://cancergenome.nih.gov>.) with IDH status, and the corresponding MRI images were acquired from the The Cancer Imaging Archive (TCIA) Data Portal (<https://www.cancerimagingarchive.net>).

Multimodal MR Images: All patients were imaged in the supine position with a 3.0-T MRI machine (Magnetom, Skyra; Siemens Healthcare, Erlangen, Germany) using a transmit/receive quadrature 20-channel head-and-neck coil. The imaging protocol was the same for all patients.

T1-contrast: TR: 1820 ms, TE: 13 ms, slice number: 19, FOV: 230 mm, slice thickness: 5 mm, distance factor: 30%, FA: 150 deg, inversion time (TI): 825 ms, voxel size: 0.4×0.4×5.0 mm, accelerate factor: 2, bandwidth: 260 Hz/Px, echo spacing: 13 ms.

Fluid-attenuated inversion recovery imaging (T2-FLAIR): TR: 8000 ms, TE: 81 ms, slice number: 19, FOV: 220 mm, slice thickness: 5 mm, distance factor: 30%, FA: 150 deg, inversion time (TI): 2370 ms, voxel size: 0.7×0.7×5.0 mm, accelerate factor: 2, bandwidth: 289 Hz/Px, echo spacing: 9.02 ms.

All MRI sequences of each patient from our own datasets and from TCIA have the same imaging scale, position, slices and slice thickness.

4.4. Computer Analysis

An automated framework was designed to predict IDH1 genotype, consisting of the following steps, which were carried out in sequence: (I) automated image pre-processing to select the regions of interest (ROIs), (II) feature extraction, (III) feature selection, and (IV) automated IDH1 genotype prediction and results interpretation.

(Figure 15)

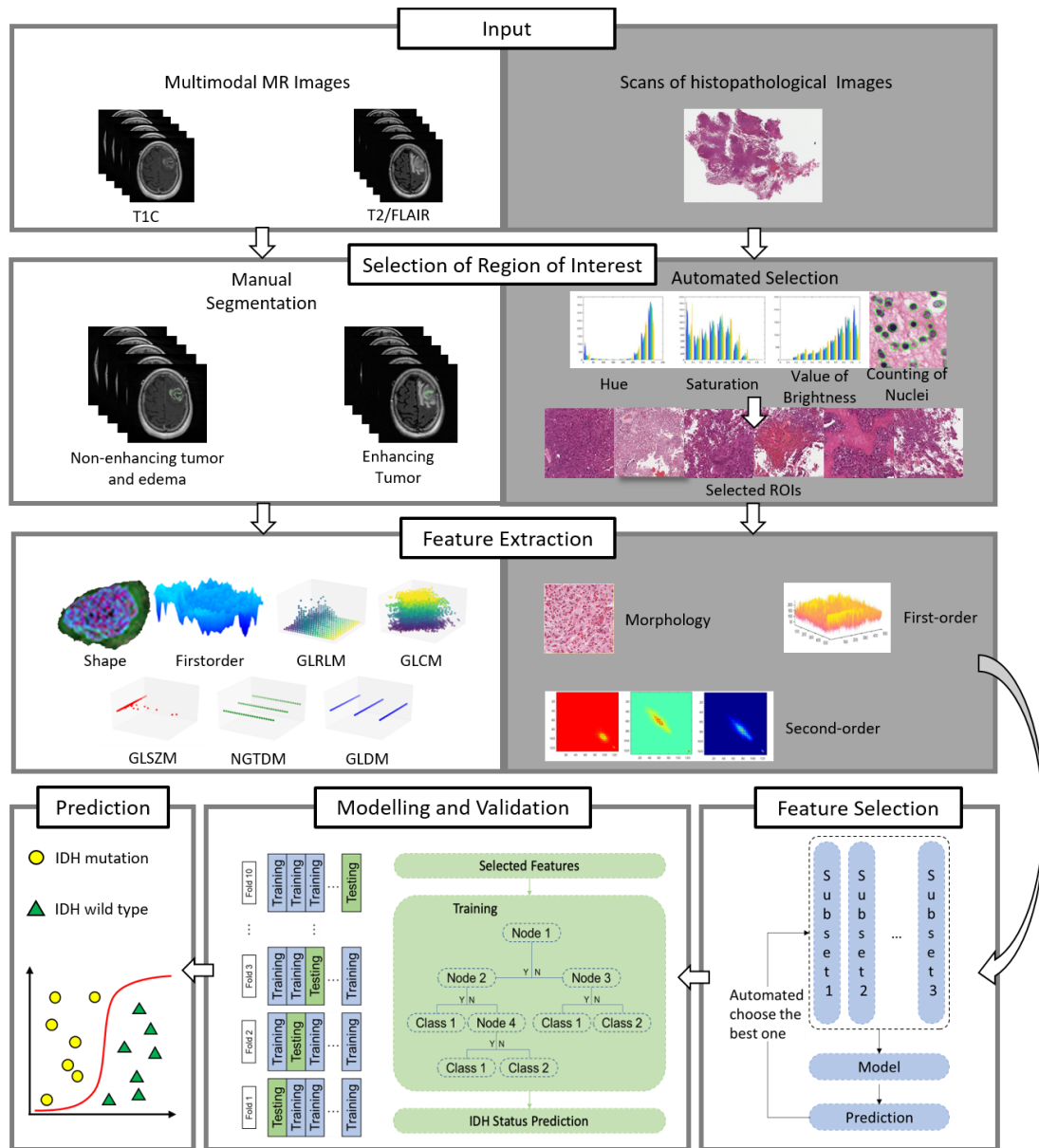


Figure 15 Machine learning framework for automated prediction of IDH glioblastoma genotypes. Representative regions of interest (ROIs) are selected on multimodal MR and histopathological images, respectively. After that, histopathological and radiomics features were derived from the ROIs. Then, in order to reduce the redundant features, the Random Forest model based Recursive Feature Elimination (RF-RFE) algorithm was adopt to select the relevant features. Following 10-fold cross-validation, the automated machine-

learning model with relevant features for glioma IDH genotype prediction is established. Abbreviations: GCLM, Grey Level Co-occurrence Matrix features; GLRLM, Grey Level Length Matrix features; GLSZM, Grey Level Size Zone Matrix features; NGTDM, Neighboring Gray Tone Difference Matrix features; GLDM, Grey level Dependence Matrix features. mass of features, which may have redundant information, are selected and processed improving the predictive power of the machine-learning model.

4.4.1. ROI Extraction

Our computational algorithm used for the analysis of histological images approaches the region of interest at two different levels. First, one tile with the highest cell density ($5120 * 5120$) [105, 109] is extracted employing the watershed nuclei detection algorithm [111, 146, 147]. Then, based on Hue, Saturation, and Value of Brightness (HSV channel), five tiles representing the whole image [111, 146] at 40X resolution are identified. Third, based on the HSV channel, five tiles representing the entire image at 4X resolution are identified.

For the analysis of radiological images, we have segmented edema and non-enhanced tumors from T2-FLAIR image. In addition, T1CE images were used for enhancing tumor volume segmentation. The lesions were separated into three parts, enhancing tumor, tumorous necrosis and peritumoral edema. The process of tumor segmentation was performed manually using the ITK-SNAP software (version 3.6.0; www.itksnap.org). Firstly, all the MRI sequences were retrieved from the Picture Archiving and Communication System (PACS). Then we applied N4 bias field correction to remove the presence of low frequency intensity non-uniformity.

Intermodality co-registration with different 2D MRI sequences was achieved by means of ITK-SNAP. Using this method, ROIs of enhancing tumor were delineated on post-contrast T1WI images by a semi-automatic method, in which only the enhancing area was selected. Tumor necrosis was defined as the non-enhancing area within enhancing area on post-contrast T1WI. ROIs of peritumoral edema were delineated on T2-FLAIR, which was defined as the high-signal region beyond the enhancing area. The process was performed by a consultant neuro-radiologist. Finally, the ROIs were registered on each slice of each 2D MRI sequence.

4.4.2. Feature Extraction

Our computational algorithm used for the analysis of histological images approaches the region of interest at two different levels. First, one tile with the highest cell density ($5120 * 5120$) [105, 109] is extracted employing the watershed nuclei detection algorithm [111, 146]. Then, based on Hue, Saturation, and Value of Brightness (HSV channel), five tiles representing the whole image [1, 111, 146] at 40X resolution are identified. Third, based on the HSV channel, five tiles representing the entire image at 4X resolution are identified.

For the analysis of radiological images, we have segmented edema and non-enhanced tumors from T2-FLAIR image. In addition, T1CE images were used for enhancing tumor volume segmentation. The lesions were separated into three parts, enhancing tumor, tumorous necrosis and peritumoral edema. The process of tumor segmentation was performed manually using the ITK-SNAP software (version 3.6.0; www.itksnap.org). Firstly, all the MRI sequences were retrieved from the Picture Archiving and Communication System (PACS). Then we applied N4 bias field

correction to remove the presence of low frequency intensity non-uniformity. Intermodality co-registration with different 2D MRI sequences was achieved by means of ITK-SNAP. Using this method, ROIs of enhancing tumor were delineated on post-contrast T1WI images by a semi-automatic method, in which only the enhancing area was selected. Tumor necrosis was defined as the non-enhancing area within enhancing area on post-contrast T1WI. ROIs of peritumoral edema were delineated on T2-FLAIR, which was defined as the high-signal region beyond the enhancing area. The process was performed by a consultant neuro-radiologist. Finally, the ROIs were registered on each slice of each 2D MRI sequence.

4.4.3. Feature Selection

Although a large number of image features can be used to construct a model that fully reflects the characteristics of gliomas, removing redundant information can improve the efficacy of the model for glioma genotyping [148]. In order to reduce the amount of redundant information inherent to quantitative features, we built a random forest algorithm enhanced by a recursive feature elimination (RF-RFE) procedure in order to identify the relevant and important characteristics before implementation in a classification model all [114]. As shown in **Figure 15**, the feature with the lowest importance for classification calculated by the algorithm will be eliminated.

4.4.4. Modelling and Validation

We are proposing a binary classification model to differentiate patients with an IDH1 mutation from wild type HGGs based on clinical features (age and sex), digital histopathological image Pathomics features and MRI Radiomics features. In order to achieve this, the Random Forest algorithm was adopted, which is already widely used

in medical imaging analysis. The random forest model can deal with a huge feature set, especially the data involved in this study, which can obtain satisfactory results. All machine-learning methods were implemented with the Statistics and Machine Learning package on Python 3.6.

In addition, 10-fold cross-validation was utilized to evaluate the performance of the machine-learning model: cases were first divided into 10 cohorts, where one of the cohorts was used as the testing set and the other 9 cohorts were used for training. Specific, in the first iteration, the first cohort was used as the testing set and the rest cohorts were used as the training sets. In the second iteration, the second cohort was used for testing and the rest cohorts were used for training. This procedure was repeated until the each cohort of 10 cohorts have been used as testing set. The average performance of each iteration was used to evaluate the stability of the prediction model.

4.5. Results

4.5.1. Feature Extraction and selection

In this phase, we have extracted a total of 22 morphological features, which were identified in the glioma cases studied (Supplementary Materials I. Histopathologic Features Extraction), including nuclear shape and staining intensity (**Figure 16**), such as the mean of nuclei perimeter, area, and solidity. On the other hand, we have extracted 171 sub-visual features [1] from the high resolution of digital histopathology images, including intensity features, GLCM features, etc. As for results of IDH1 status prediction for HGGs, the histopathological features extracted from histopathology images, which reached an accuracy of 0.81 ± 0.03 with 10-fold cross validation.

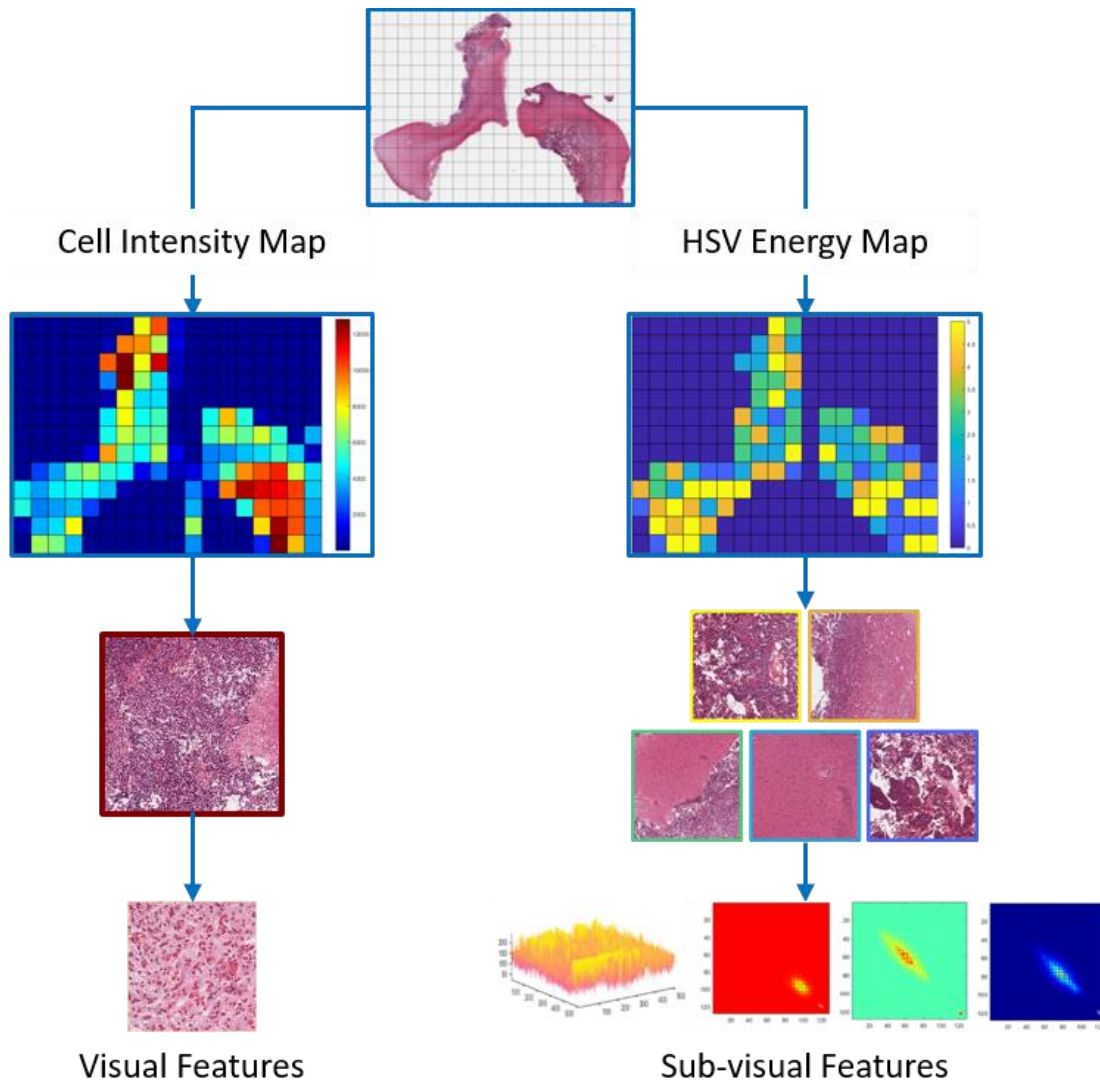


Figure 16 Feature extraction from histopathological images. Based on two different ROI extraction criteria, HSV (Hue, Saturation and Value of Brightness) and cell density, six ROIs in total per section presenting different resolutions are used. Subsequently, two types of features are extracted from these ROIs as shown, visual features and sub-visual features. The latter are derived from histopathology images at different resolutions. Visual features from the representative tile with the highest density of nuclei were obtained at 40X resolution, with density referring to the semantic level and features comprising both counts and sizes of

individual nuclei. In contrast, sub-visual features were derived from ROIs at 4X resolution and included high-throughput intensity and texture image features.

Turns to multimodal MRIs, 1132 features in all were obtained from each patients and each sequence, including 234 first order features, 14 shape-based features, 286 grey level co-occurrence matrix features, 208 grey level run length matrix features, 208 grey level size zone matrix features, 182 grey level dependence matrix features (**Figure 17**). For HGGs, after feature selection phase, for T1C and T2F, four and five features were remained. Based on 10-fold cross validation, we fed the selected features into random forest model to predict IDH genotype, which achieved the accuracy of 0.75 ± 0.02 and 0.71 ± 0.03 separately.

In this research, we did IDH genotyping prediction for HGGs. The AUC for features extracted from different histopathologic grade of biomedical images were 0.90 ± 0.09 .

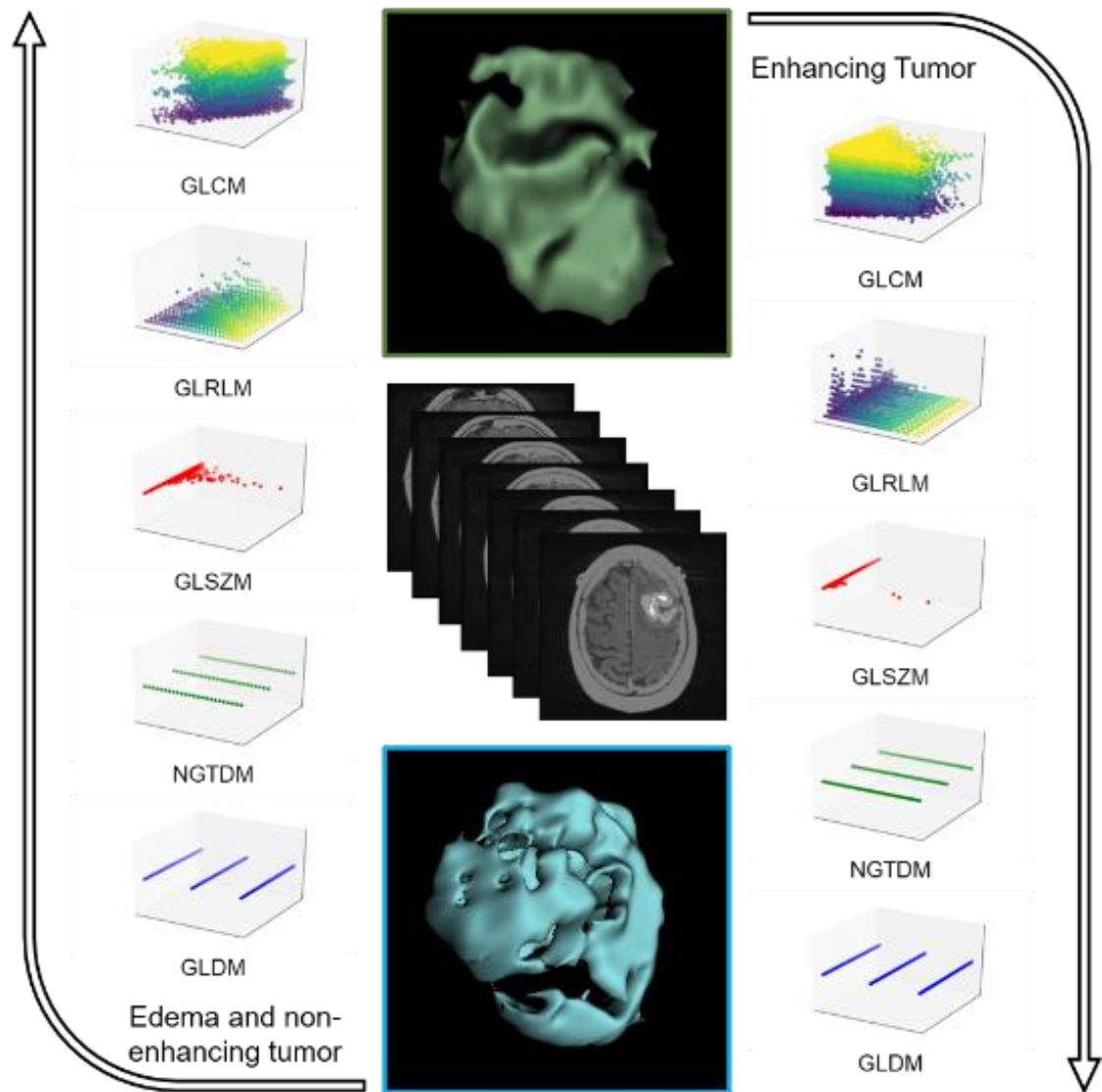


Figure 17 Measurements obtained from different regions of individual MR images. Briefly, the initial segmentation was performed by a clinicians who identified enhancing tumor, non-enhancing tumor and edema. Subsequently, quantitative radiomics features were derived from multiple volumes of interest (VOIs). The VOIs were divided into seven classes (see : first-order statistics features, shape-based features, GLCM, GLRLM, GLSZM, NGTDM, and GLDM features).

4.5.2. Comparison of Performances Using Different Modalities of Biomedical Images and Feature Types

In order to evaluate the role of different modalities of biomedical images (T1C, T2/FLAIR and WSIs) in the prediction of IDH genotype, random forest machine-learning model with 10-fold cross-validation was adopted in our research. In general, scans of histopathological images have better results in IDH genotype prediction than other images (**Figure 18C**). As for quantitative features obtained from different modalities of biomedical images, the Visual features show high accuracy and stability (**Figure 18 B**).

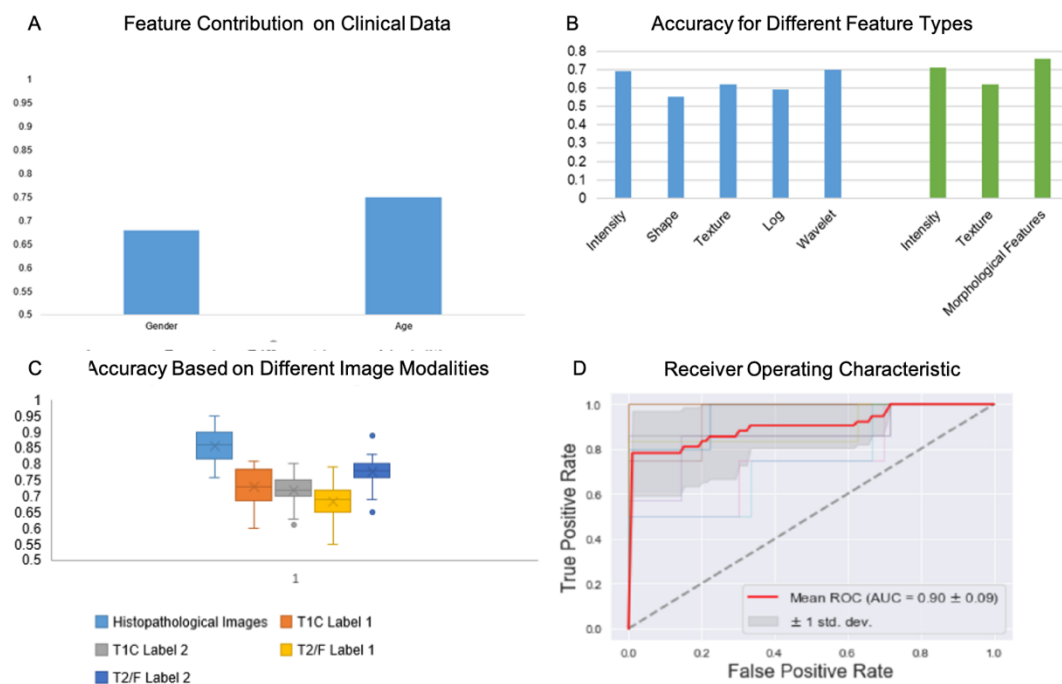


Figure 18 Random forest classifier scores for IDH1 genotype prediction. A, Prediction results based on Clinical data (age and gender); B, Prediction results based on different feature types; C, Prediction results based on different image

modalities; **D, Receiver Operating Characteristic (ROC) for IDH genotype prediction across multi-parametric medical images.**

4.5.3. Univariate Feature Analysis

For IDH1 genotype prediction, optimal features were selected from the different modalities of medical images, including 7 features from the digital histopathological images, 4 features from the T1CE images and 5 from the T2-FLAIR images. As shown in **Figure 18B**, GLRLM, Shape-based and GLCM features had the greatest power in predicting HGGs IDH1 status. Age, nuclei counts and first-order features were the most important factors that contributed to this result. As summarized in **Table 3**, top-performing features within different types of image features contributed to IDH1 status prediction as shown.

Table 3 TOP-performing features in IDH1 Status Prediction by means of Univariate Analysis

Types	Mask	Feature Name	Feature Description	Accuracy
Clinical	N/A	Age	Age at diagnose;	0.74
Intensity	T1C- edema	Uniformity	Formula: $F_u = \sum_{i=1}^{N_g} p(i)^2$	0.69

			<p>Where $p_{(i)}$ refers to the features calculated from N_g discrete pixel levels.</p> <p>Measuring the sum of the square of image VOI pixel value.</p>	
Shape	FLAIR- edema	Flatness	<p>Formula:</p> $F_{flatness} = \frac{\lambda_{least}}{\lambda_{major}}$ <p>Where λ_{major} and λ_{least} refer to the length of the maximum and minimum principal component axes, respectively.</p> <p>Measuring the relationship between the largest and smallest principal components in the VOI shape.</p>	0.67
Texture	T1C- tumor	wavelet- LLL_glrIm_ LRLGLE	Formula:	0.72

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

Where N_g refers to the gray level distribution within the VOI, N_r refers to the maximal length within the VOI, $P(i,j|\theta)$ refers to the run length matrix for an arbitrary direction θ , $N_r(\theta)$ is the number of runs in the image along θ .

This feature quantitative describes the joint distribution of long-run lengths with lower gray level values after a wavelet filter.

Wavelet	T1C-	wavelet-	Formula:	0.69
	tumor	HHH-glc-	$F_{MP} = \max (p(i,j))$	
		MP	Where $p(i,j)$ is the normalized co-occurrence matrix.	

<p>Quantify the occurrences of the most predominant pair of neighboring intensity values through a Gray Level Co-occurrence Matrix after an image filter by a high-frequency wavelet.</p>			
LoG	FLAIR-tumor	Log-glszm-SALGLE	<p>Formula: 0.65</p> $F_{SALGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i,j)}{i^2 j^2}}{N_z}$ <p>Where N_g refers to the distribution values within the VOI, N_s refers to the zone sizes quantity within the VOI., N_z refers to refers to the zones quantity within the VOI, and $P(i,j)$ is the size zone matrix.</p> <p>Quantify the proportion in the mask of VOI by quantify the Gray Level Size Zone joint distribution of smaller size zones with lower gray</p>

level values after the LoG filter.			
Morphology	Tile	Cell counts	Quantitative describes the cell intensity in the ROI. 0.78

The accuracy of IDH1 status prediction was as high as 0.88 ± 0.03 when multi-parametric features were extracted from different histopathological and radiomics images through our implementation of the random forest algorithm. **Table 4** reveals the features that play important roles in our classification model. As shown in this Table, clinical data (sex and age) and histopathological features had a dominant role in our classification procedure.

Table 4 Prediction of IDH1 Genotype Based in High Grade Gliomas.

Image Modalities	Accuracy
Digital Histopathological Images	0.86 (+/- 0.03)
T1CE (edema and non-enhanced tumor)	0.73 (+/- 0.06)
T1CE (enhanced tumor)	0.72 (+/- 0.04)
T2-FLAIR (edema and non-enhanced tumor)	0.68 (+/- 0.05)
T2-FLAIR (enhanced tumor)	0.78 (+/- 0.04)

Multi-modal Image Data

0.90 (+/- 0.05)

4.5.4. Quantitative IDH1 Status Prediction Results Interpretation

LIME (Local Interpretable Model-agnostic Explanations) is a tool for facilitating local model interpretability. The technique perturbs the input data to understand how the predictions are affected. Figure 3 and Figure 4 illustrate two representative cases from visual analysis and the machine-learning model. The first case is an IDH1-wt patient (**Figure 19**), who is 43 years old (age at diagnosis), female with a histopathological grade IV glioma. The second one is a patient with an IDH1 mutation (Figure 20), who is 22 years old (age at diagnosis), male with histopathological grade III glioma.

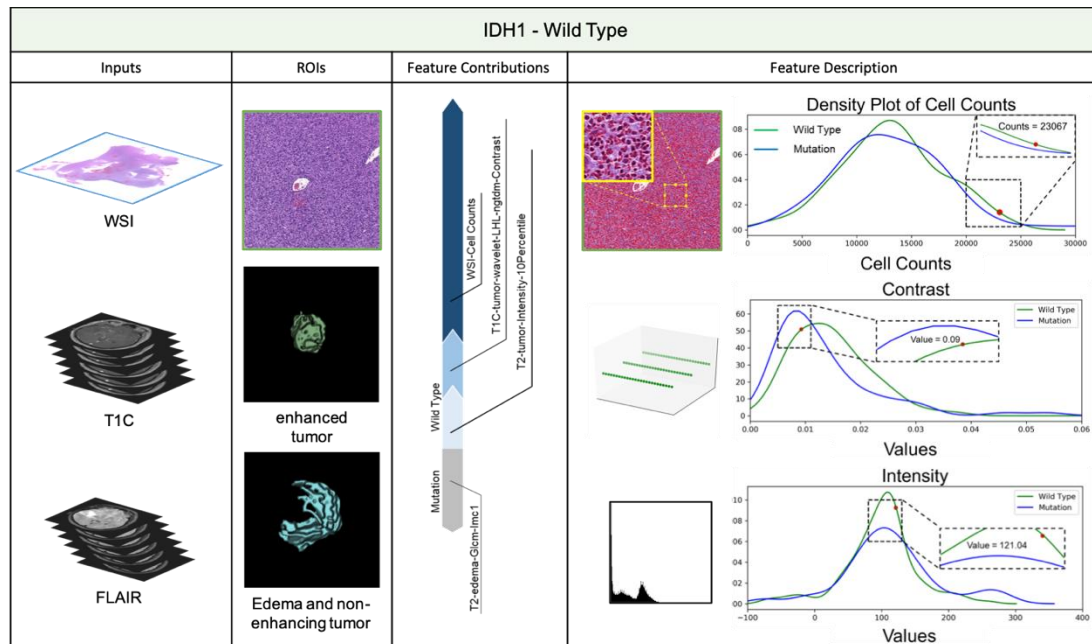


Figure 19 IDH1 status prediction result explanation of representative cases with the LIME algorithm for the RF model (IDH1-Wild type case)

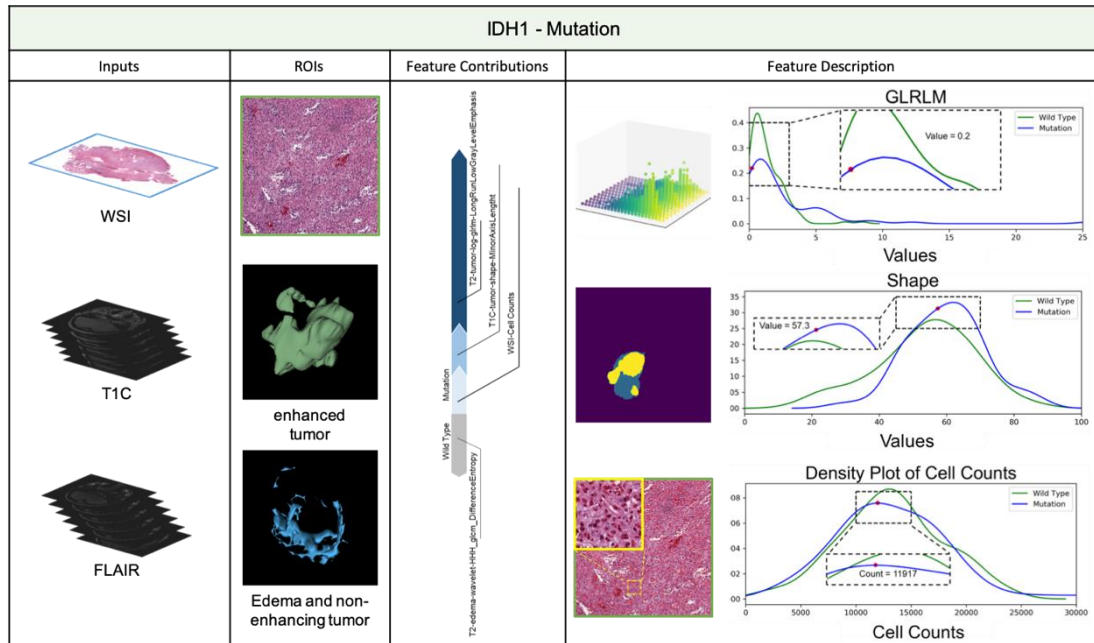


Figure 20 IDH1 status prediction result explanation of representative cases with the LIME algorithm for the RF model (IDH1-Mutation case)

Machine learning models taking into account their different contributions, then quantitatively combine these functions to predict the results. During model training, the LIME model can give weighted coefficients to explain the contributions made by different features. The LIME model perturbed the feature values and observed the resulted changes in prediction. The predictive ability of this algorithm made the most important contribution to achieve a higher weight value of the right. Positive weights reflect the increase in the corresponding features may make a positive contribution to the IDH1 status prediction. In contrast, negative weights would have negative predictive power. As shown in **Figure 19** and **Figure 20**, the feature contribution for IDH1 genotyping for two representative cases have been listed, which are derived from the LIME model to obtain the linear combination of feature values and weights

4.6. Discussion

The determination of isocitrate dehydrogenase (IDH1) status can help guide clinical decision making including post-operative treatment such as radio- and chemotherapy. We have developed this framework to integrate two cost saving imaging modalities to predict a clinically relevant somatic mutation. Thus, determination of IDH1 phenotypes as early as possible during the disease course may guide perioperative counselling, and the right choice of treatment option. In this study, we develop a random forest algorithm-based genotype classifier that allows the prediction of IDH mutation status in glioma patients making use of imaging features derived from pre-surgical MRI scans and post-operative digital data obtained from routine H&E slides. The primary results show this combined digital platforms will allow very large data sets to be created and mined to extract new information for high reliability genotype prediction.

Going beyond the identification of biomarkers, one of the greatest unmet challenges for the successful clinical implementation of AI is to build the bridges that enable multiscale imaging integration of radiomics, pathomics into readily available pipelines that can transform clinical routine. In this regard, it is very important to integrate two cost-effective imaging modalities to make precise key somatic mutation predictions. In order to improve the accuracy of IDH1 phenotype prediction before the post-treatment plan, visual and sub-visual features extracted from digital histopathological images and quantitative radiomics feature extracted from different multimodality MRIs were implemented into our RF-RFE feature selection model to select the optimal features for further analysis. However, the accuracy of the combined RA-PATH-omics within the independent testing set was lower than that of the testing set. The most likely reason for

this is the differences in MRI model and scan parameters and in IDH1 mutation rate among the different patient cohorts. In an ideal scenario, all patient scans would be collected with consistent acquisition parameters, and IDH1 mutation prevalence would be the same or similar. However, this would be challenging in practice, as MR scanner models and acquisition parameters, as well as the demographics of patient captured vary widely from institution to institution.

Our novel integrated approach, which we have termed RA-PATH-omics, was found to be a more accurate predictor of IDH1 genotype than either radiomics or histopathological feature recognition alone. RA-PATH-omics characterizes tumor properties at different biological scales and meet the need to understand correlations between image features, genomics, and clinical outcomes.

Furthermore, the IDH1 predictive performance of histopathological image analysis was found to be superior to T2-FLAIR and T1CE (0.86 vs 0.71, 0.75) in HGGs. Among the top histomorphometric features, the mean area of cells and the mean of cells axis were most significant, where statistically describe the central tendency of cell shape. (In addition, age was a good predictor (Accuracy=0.8)). The top identified features are in line with the fact that gliomas with an IDH1 mutation have a more coherent nuclear architecture, i.e. they are morphologically less atypical than IDH1 wild type, which is associated with a higher risk of recurrence. Our IDH1 phenotypes prediction achieved high accuracy for two reasons: First, we integrated MRIs, digital histopathological images and clinical information for IDH1 status prediction. Second, we used a relatively data to significantly reduce the number of parameters in the model to avoid overfitting while making our model more general. To the best of our knowledge, this study is the first to integrate MRI and digital pathology images in a computerized model for

predicting IDH1 genotype. It is worth noting that T1CE and T2-FLAIR images conferred a higher predictive value than other MR sequences. Our study makes an important first step towards achieving the goal of radio-genomic tools be used in a clinical setting.

4.7. Chapter Summary

Our work is a step towards more effective use of combined Radiomics and Pathomics data for genotype prediction by utilizing structural MRI sequences and digital pathology images for IDH1 phenotype prediction. In particular, it should be helpful to patients who did not receive molecular diagnosis or for retrospective studies. In a more general sense, our results (i) show that machine learning can robustly identify genomic features within structural MR imaging and histopathological datasets, (ii) suggest a complementary method for the genotyping of gliomas, which is suitable for glioma patient screening, and (iii) demonstrate the potential for algorithmic tools within the clinic to aid clinical decision-making. Taken together, it is expected that integration of multimodality omics in oncology research will become commonplace as technology evolves, with significant potential for prognosis prediction and treatment strategies for brain tumor management in the future.

CHAPTER 5. **Conclusions and**

Future Work

5.1. Conclusions

In this thesis, we target to integrate quantitative imaging data obtained from multiple modalities into mathematical models of tumor prediction response to achieve additional insights into practical predictive value.

Firstly, we proposed an automated machine-learning brain tumor-grading platform to investigate contributions of multi-parameters from multimodal data, including imaging parameters or features from Whole Slide Images (WSI) and the proliferation marker KI-67. For each WSI we extract both visual parameters, such as morphology parameters, and sub-visual parameters, including first-order and second-order features. On the basis of machine learning models, our platform classifies gliomas into grades II, III, and IV.

Secondly, a quantitative interpretable machine learning approach (Local Interpretable Model-Agnostic Explanations) was proposed to measure the contribution of features for single case. Most grading systems based on machine learning models are considered “black boxes,” and it would be valuable for patient management if clinically trusted reasoning could be revealed. The quantitative analysis and explanation may assist clinicians to better understand the disease and accordingly to choose optimal treatments for improving clinical outcomes.

Thirdly, based on the automated brain tumor-grading platform we proposed before, multimodal MRIs have been introduced in our research. A new imaging–tissue correlation based approach called RA-PA-Thomics was proposed to predict the IDH genotype. Inspired by the concept of image fusion, we integrate multimodal MRIs and scans of histopathological images for indirect, fast, and cost saving IDH genotyping.

5.2. Future Work

This chapter introduces the work that we would like to implement into the procedure of multimodal medical image analysis;

1. deep learning algorithms adopted for tumor segmentation and grading
2. multi-task learning framework established for multi-parametric image features obtained from multimodal medical images
3. correlation between histopathological images and molecular imaging.

Deep learning: Accurate brain tumor segmentation is based on digital image processing. At present, the manual segmentation of clinical experts is still a recognized segmentation method, but this work is time-consuming and observe-dependent. Tumor incoherence and differences in contrast uptake will affect the segmentation of gliomas. It is inevitable to employ the machine-learning algorithm into glioma segmentation. For example, the combination of unsupervised classification with FCM and Knowledge-based image processing [56] can be applied on gliomas segmentation to segment non-enhancing tumors. Recently, deep learning has been implemented into automatic gliomas segmentation. For example, a 3D convolutional network [57] can be applied on the multimodal MRI glioma segmentation. However, for tumor segmentation and staging, the proposed methods need further study.

Multi-task learning: Competing with traditionally machine-learning algorithms, multi-task learning refers to a method that learns multiple related tasks together based on shared representations. As mentioned before, deep learning has been introduced into the field of medical image analysis, quantitative deep features have been dig out. In the future, we would like to combine the deep features and radiomics features we used in chapter 4 to implement into multi-task learning and achieve the better result in tumor segmentation or grading.

Correlation between histopathological images and molecular imaging: On the account of vastly different image features of individual modalities, we would like to propose an information-theoretic approach to combine the multi-contribute features. In the end, we hope to identify patterns or features extracted from MRIs that will help interpret the features from histopathological images.

Bibliography

- [1] X. Wang *et al.*, "Machine learning models for multiparametric glioma grading with quantitative result interpretations," *Frontiers in neuroscience*, vol. 12, 2018.
- [2] A. J. Méndez, P. G. Tahoces, M. a. J. Lado, M. Souto, and J. J. Vidal, "Computer-aided diagnosis: Automatic detection of malignant masses in digitized mammograms," *Medical Physics*, vol. 25, no. 6, pp. 957-964, 1998.
- [3] G. D. Tourassi, "Journey toward computer-aided diagnosis: role of image texture analysis," *Radiology*, vol. 213, no. 2, pp. 317-320, 1999.
- [4] Q. Li and R. M. Nishikawa, *Computer-aided detection and diagnosis in medical imaging*. Taylor & Francis, 2015.
- [5] S. P. Constantinos, M. S. Pattichis, and E. Micheli-Tzanakou, "Medical imaging fusion applications: An overview," in *Conference Record of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers (Cat. No. 01CH37256)*, 2001, vol. 2: IEEE, pp. 1263-1267.
- [6] K. Horsch, M. L. Giger, C. J. Vyborny, L. Lan, E. B. Mendelson, and R. E. Hendrick, "Classification of breast lesions with multimodality computer-aided diagnosis: observer study results on an independent clinical data set," *Radiology*, vol. 240, no. 2, pp. 357-368, 2006.
- [7] C. Demir and B. Yener, "Automated cancer diagnosis based on histopathological images: a systematic survey," *Rensselaer Polytechnic Institute, Tech. Rep*, 2005.

- [8] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," *Neurocomputing*, vol. 191, pp. 214-223, 2016.
- [9] D. N. Louis *et al.*, "The 2016 World Health Organization classification of tumors of the central nervous system: a summary," *Acta neuropathologica*, vol. 131, no. 6, pp. 803-820, 2016.
- [10] N. Soomro and S. Bakhtiar, "Image-Based Modeling and Precision Medicine," in *Progress and Challenges in Precision Medicine*: Elsevier, 2017, pp. 141-170.
- [11] V. Bhateja, H. Patel, A. Krishn, A. Sahu, and A. Lay-Ekuakille, "Multimodal medical image sensor fusion framework using cascade of wavelet and contourlet transform domains," *IEEE Sensors Journal*, vol. 15, no. 12, pp. 6783-6790, 2015.
- [12] C. He, Q. Liu, H. Li, and H. Wang, "Multimodal medical image fusion based on IHS and PCA," *Procedia Engineering*, vol. 7, pp. 280-285, 2010.
- [13] Y. Yang, Y. Que, S. Huang, and P. Lin, "Multimodal sensor medical image fusion based on type-2 fuzzy logic in NSCT domain," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3735-3745, 2016.
- [14] M. Manchanda and R. Sharma, "A novel method of multimodal medical image fusion using fuzzy transform," *Journal of Visual Communication and Image Representation*, vol. 40, pp. 197-217, 2016.
- [15] S. Das and M. K. Kundu, "NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency," *Medical & biological engineering & computing*, vol. 50, no. 10, pp. 1105-1114, 2012.

- [16] G. Qu, D. Zhang, and P. Yan, "Medical image fusion by wavelet transform modulus maxima," *Optics Express*, vol. 9, no. 4, pp. 184-190, 2001.
- [17] R. Singh, M. Vatsa, and A. Noore, "Multimodal medical image fusion using redundant discrete wavelet transform," in *2009 Seventh International Conference on Advances in Pattern Recognition*, 2009: IEEE, pp. 232-235.
- [18] B. Marshall, "A Brief History of the Discovery of *Helicobacter pylori*," in *Helicobacter pylori*: Springer, 2016, pp. 3-15.
- [19] J. C. Caicedo, F. A. González, and E. Romero, "Content-based histopathology image retrieval using a kernel-based semantic annotation framework," *Journal of biomedical informatics*, vol. 44, no. 4, pp. 519-528, 2011.
- [20] N. Mehta, A. Raja'S, and V. Chaudhary, "Content based sub-image retrieval system for high resolution pathology images using salient interest points," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009: IEEE, pp. 3719-3722.
- [21] X. Qi *et al.*, "Content-based histopathology image retrieval using CometCloud," *BMC bioinformatics*, vol. 15, no. 1, p. 287, 2014.
- [22] A. Sridhar, S. Doyle, and A. Madabhushi, "Content-based image retrieval of digitized histopathology in boosted spectrally embedded spaces," *Journal of pathology informatics*, vol. 6, 2015.
- [23] J. A. Vanegas, J. Arevalo, and F. A. González, "Unsupervised feature learning for content-based histopathology image retrieval," in *2014 12th International Workshop on Content-Based Multimedia Indexing (CBMI)*, 2014: IEEE, pp. 1-6.

- [24] R. Sparks and A. Madabhushi, "Out-of-sample extrapolation utilizing semi-supervised manifold learning (ose-ssl): Content based image retrieval for histopathology images," *Scientific reports*, vol. 6, p. 27306, 2016.
- [25] J. N. Weinstein *et al.*, "The cancer genome atlas pan-cancer analysis project," *Nature genetics*, vol. 45, no. 10, p. 1113, 2013.
- [26] A. Yoshida *et al.*, "Comprehensive histologic analysis of ALK-rearranged lung carcinomas," *The American journal of surgical pathology*, vol. 35, no. 8, pp. 1226-1234, 2011.
- [27] M. Dal Molin *et al.*, "Clinicopathological correlates of activating GNAS mutations in intraductal papillary mucinous neoplasm (IPMN) of the pancreas," *Annals of surgical oncology*, vol. 20, no. 12, pp. 3802-3808, 2013.
- [28] H. Zhou *et al.*, "MRI features predict survival and molecular markers in diffuse lower-grade gliomas," *Neuro-oncology*, vol. 19, no. 6, pp. 862-870, 2017.
- [29] F. G. Davis, S. Freels, J. Grutsch, S. Barlas, and S. Brem, "Survival rates in patients with primary malignant brain tumors stratified by patient age and tumor histological type: an analysis based on Surveillance, Epidemiology, and End Results (SEER) data, 1973–1991," *Journal of neurosurgery*, vol. 88, no. 1, pp. 1-10, 1998.
- [30] A. D. Norden *et al.*, "An exploratory survival analysis of anti-angiogenic therapy for recurrent malignant glioma," *Journal of neuro-oncology*, vol. 92, no. 2, pp. 149-155, 2009.
- [31] B. Zhang *et al.*, "Multimodal MRI features predict isocitrate dehydrogenase genotype in high-grade gliomas," *Neuro-oncology*, vol. 19, no. 1, pp. 109-117, 2016.

- [32] H. Yan *et al.*, "IDH1 and IDH2 mutations in gliomas," *New England Journal of Medicine*, vol. 360, no. 8, pp. 765-773, 2009.
- [33] L. Lüdemann, W. Grieger, R. Wurm, M. Budzisch, B. Hamm, and C. Zimmer, "Comparison of dynamic contrast-enhanced MRI with WHO tumor grading for gliomas," *European radiology*, vol. 11, no. 7, pp. 1231-1241, 2001.
- [34] M. H. Lev *et al.*, "Glial tumor grading and outcome prediction using dynamic spin-echo MR susceptibility mapping compared with conventional contrast-enhanced MR: confounding effect of elevated rCBV of oligodendrogliomas," *American Journal of Neuroradiology*, vol. 25, no. 2, pp. 214-221, 2004.
- [35] A. von Deimling, T. Ono, M. Shirahata, and D. N. Louis, "Grading of diffuse astrocytic gliomas: a review of studies before and after the advent of IDH testing," in *Seminars in neurology*, 2018, vol. 38, no. 01: Thieme Medical Publishers, pp. 019-023.
- [36] B. F. Waller, C. A. Pinkerton, and J. Slack, "Intravascular ultrasound: a histological study of vessels during life. The new 'gold standard' for vascular imaging," *Circulation*, vol. 85, no. 6, pp. 2305-2310, 1992.
- [37] D. R. Genest, "Partial hydatidiform mole: clinicopathological features, differential diagnosis, ploidy and molecular studies, and gold standards for diagnosis," *International journal of gynecological pathology*, vol. 20, no. 4, pp. 315-322, 2001.
- [38] L. B. Rorke, "Pathologic diagnosis as the gold standard," *Cancer*, vol. 79, no. 4, pp. 665-667, 1997.
- [39] S. Beer, "The viable system model: Its provenance, development, methodology and pathology," *Journal of the operational research society*, vol. 35, no. 1, pp. 7-25, 1984.

- [40] S. Park *et al.*, "The history of pathology informatics: A global perspective," *Journal of pathology informatics*, vol. 4, 2013.
- [41] L. Pantanowitz *et al.*, "Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center," *Archives of Pathology and Laboratory Medicine*, vol. 137, no. 12, pp. 1710-1722, 2013.
- [42] M. D. Zarella *et al.*, "A practical guide to whole slide imaging: a white paper from the digital pathology association," *Archives of pathology & laboratory medicine*, vol. 143, no. 2, pp. 222-234, 2018.
- [43] M. N. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE reviews in biomedical engineering*, vol. 2, p. 147, 2009.
- [44] Y. Xu *et al.*, "Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features," *BMC bioinformatics*, vol. 18, no. 1, p. 281, 2017.
- [45] L. Landini, V. Positano, and M. Santarelli, *Advanced image processing in magnetic resonance imaging*. CRC press, 2018.
- [46] P. Korfiatis, T. L. Kline, and B. J. Erickson, "Automated segmentation of hyperintense regions in FLAIR MRI using deep learning," *Tomography*, vol. 2, no. 4, p. 334, 2016.
- [47] L. Chato and S. Latifi, "Machine learning and deep learning techniques to predict overall survival of brain tumor patients using MRI images," in *2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2017: IEEE, pp. 9-14.

- [48] S. Napel, W. Mu, B. V. Jardim-Perassi, H. J. Aerts, and R. J. Gillies, "Quantitative imaging of cancer in the postgenomic era: Radio (geno) mics, deep learning, and habitats," *Cancer*, vol. 124, no. 24, pp. 4633-4649, 2018.
- [49] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annual review of biomedical engineering*, vol. 2, no. 1, pp. 315-337, 2000.
- [50] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: a review," *Medical image analysis*, vol. 13, no. 4, pp. 543-563, 2009.
- [51] K. H. Zou *et al.*, "Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports," *Academic radiology*, vol. 11, no. 2, pp. 178-189, 2004.
- [52] Y. Lin, J. Tian, and X.-p. Zhang, "Application of a new medical image segmentation method based on fuzzy connectedness and FCM," *Chinese journal of stereology and image analysis*, vol. 2, 2001.
- [53] M. M. Sharma, "Brain tumor segmentation techniques: A survey," *Brain*, vol. 4, no. 4, pp. 220-223, 2016.
- [54] H. Jia, Y. Xia, Y. Song, W. Cai, M. Fulham, and D. D. Feng, "Atlas registration and ensemble deep convolutional neural network-based prostate segmentation using magnetic resonance imaging," *Neurocomputing*, vol. 275, pp. 1358-1369, 2018.
- [55] B. Gaonkar, D. Hovda, N. Martin, and L. Macyszyn, "Deep learning in the small sample size setting: cascaded feed forward neural networks for medical image segmentation," in *Medical Imaging 2016: Computer-Aided Diagnosis*, 2016, vol. 9785: International Society for Optics and Photonics, p. 97852I.

- [56] Y.-h. Liu, M. Muftah, T. Das, L. Bai, K. Robson, and D. Auer, "Classification of MR tumor images based on Gabor wavelet analysis," *Journal of Medical and Biological Engineering*, vol. 32, no. 1, pp. 22-28, 2012.
- [57] S. Luo, R. Li, and S. Ourselin, "A new deformable model using dynamic gradient vector flow and adaptive balloon forces," in *APRS Workshop on Digital Image Computing, Brisbane, Australia*, 2003: Citeseer.
- [58] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "A novel content-based active contour model for brain tumor segmentation," *Magnetic resonance imaging*, vol. 30, no. 5, pp. 694-715, 2012.
- [59] F. C. Ghesu, B. Georgescu, T. Mansi, D. Neumann, J. Hornegger, and D. Comaniciu, "An artificial agent for anatomical landmark detection in medical images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2016: Springer, pp. 229-237.
- [60] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE transactions on medical imaging*, vol. 18, no. 11, pp. 1049-1075, 1999.
- [61] R. C. Patil and A. Bhalchandra, "Brain tumour extraction from MRI images using MATLAB," *International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSSE)*, vol. 2, no. 1, p. 1, 2012.
- [62] A. Chaddad, "Automated feature extraction in brain tumor by magnetic resonance imaging using gaussian mixture models," *Journal of Biomedical Imaging*, vol. 2015, p. 8, 2015.
- [63] A. Islam, S. M. Reza, and K. M. Iftekharuddin, "Multifractal texture estimation for detection and segmentation of brain tumors," *IEEE transactions on biomedical engineering*, vol. 60, no. 11, pp. 3204-3215, 2013.

- [64] J. Luts *et al.*, "Effect of feature extraction for brain tumor classification based on short echo time 1H MR spectra," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 60, no. 2, pp. 288-298, 2008.
- [65] A. Singh, M. K. Dutta, M. ParthaSarathi, V. Uher, and R. Burget, "Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image," *Computer methods and programs in biomedicine*, vol. 124, pp. 108-120, 2016.
- [66] V. Rathi and S. Palani, "Brain tumor MRI image classification with feature selection and extraction using linear discriminant analysis," *arXiv preprint arXiv:1208.2128*, 2012.
- [67] T. Kurc *et al.*, "Scalable analysis of Big pathology image data cohorts using efficient methods and high-performance computing strategies," *BMC bioinformatics*, vol. 16, no. 1, p. 399, 2015.
- [68] A. Cruz-Roa *et al.*, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," in *Medical Imaging 2014: Digital Pathology*, 2014, vol. 9041: International Society for Optics and Photonics, p. 904103.
- [69] B. Q. Huynh, H. Li, and M. L. Giger, "Digital mammographic tumor classification using transfer learning from deep convolutional neural networks," *Journal of Medical Imaging*, vol. 3, no. 3, p. 034501, 2016.
- [70] N. Tajbakhsh *et al.*, "Convolutional neural networks for medical image analysis: Full training or fine tuning?," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1299-1312, 2016.

- [71] W. M. Hafizah, E. Supriyanto, and J. Yunus, "Feature extraction of kidney ultrasound images based on intensity histogram and gray level co-occurrence matrix," in *2012 Sixth Asia Modelling Symposium*, 2012: IEEE, pp. 115-120.
- [72] 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, no. 6, pp. 1099-1108, 2013.
- [73] E. Mercan, S. Aksoy, L. G. Shapiro, D. L. Weaver, T. T. Brunyé, and J. G. Elmore, "Localization of diagnostically relevant regions of interest in whole slide images: A comparative study," *Journal of digital imaging*, vol. 29, no. 4, pp. 496-506, 2016.
- [74] T. Wan, J. Cao, J. Chen, and Z. Qin, "Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features," *Neurocomputing*, vol. 229, pp. 34-44, 2017.
- [75] J. Saltz *et al.*, "Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images," *Cell reports*, vol. 23, no. 1, pp. 181-193. e7, 2018.
- [76] R. Albert, T. Schindewolf, I. Baumann, and H. Harms, "Three-dimensional image processing for morphometric analysis of epithelium sections," *Cytometry: The Journal of the International Society for Analytical Cytology*, vol. 13, no. 7, pp. 759-765, 1992.
- [77] H. J. Aerts *et al.*, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature communications*, vol. 5, p. 4006, 2014.
- [78] A. Zwanenburg, S. Leger, M. Vallières, and S. Löck, "Image biomarker standardisation initiative," *arXiv preprint arXiv:1612.07003*, 2016.

- [79] J. J. Van Griethuysen *et al.*, "Computational radiomics system to decode the radiographic phenotype," *Cancer research*, vol. 77, no. 21, pp. e104-e107, 2017.
- [80] M. G. Ertosun and D. L. Rubin, "Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks," in *AMIA Annual Symposium Proceedings*, 2015, vol. 2015: American Medical Informatics Association, p. 1899.
- [81] 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.
- [82] A. Kharrat, K. Gasmi, M. B. Messaoud, N. Benamrane, and M. Abid, "A hybrid approach for automatic classification of brain MRI using genetic algorithm and support vector machine," *Leonardo journal of sciences*, vol. 17, no. 1, pp. 71-82, 2010.
- [83] C. Decaestecker *et al.*, "Nearest-neighbor classification for identification of aggressive versus nonaggressive low-grade astrocytic tumors by means of image cytometry-generated variables," *Journal of neurosurgery*, vol. 86, no. 3, pp. 532-537, 1997.
- [84] J. Kazmierska and J. Malicki, "Application of the Naïve Bayesian Classifier to optimize treatment decisions," *Radiotherapy and Oncology*, vol. 86, no. 2, pp. 211-216, 2008.
- [85] M. Artzi, G. Liberman, D. T. Blumenthal, O. Aizenstein, F. Bokstein, and D. Ben Bashat, "Differentiation between vasogenic edema and infiltrative tumor in patients with high-grade gliomas using texture patch-based analysis," *Journal of Magnetic Resonance Imaging*, vol. 48, no. 3, pp. 729-736, 2018.

- [86] D. Patel, F. Vankawala, and B. Bhatt, "A Survey on Identification of Glioblastoma Multiforme and Low-Grade Glioma Brain Tumor Type," in *2019 International Conference on Communication and Signal Processing (ICCSP)*, 2019: IEEE, pp. 0335-0339.
- [87] R. Shai *et al.*, "Gene expression profiling identifies molecular subtypes of gliomas," *Oncogene*, vol. 22, no. 31, p. 4918, 2003.
- [88] R. Inano *et al.*, "Voxel-based clustered imaging by multiparameter diffusion tensor images for glioma grading," *NeuroImage: Clinical*, vol. 5, pp. 396-407, 2014.
- [89] V. Kumar, J. Sachdeva, I. Gupta, N. Khandelwal, and C. K. Ahuja, "Classification of brain tumors using PCA-ANN," in *2011 World Congress on Information and Communication Technologies*, 2011: IEEE, pp. 1079-1083.
- [90] T. H. Vu, H. S. Mousavi, V. Monga, U. A. Rao, and G. Rao, "DFDL: Discriminative feature-oriented dictionary learning for histopathological image classification," in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015: IEEE, pp. 990-994.
- [91] T. H. Vu, H. S. Mousavi, V. Monga, G. Rao, and U. A. Rao, "Histopathological image classification using discriminative feature-oriented dictionary learning," *IEEE transactions on medical imaging*, vol. 35, no. 3, pp. 738-751, 2015.
- [92] M. L. Goodenberger and R. B. Jenkins, "Genetics of adult glioma," *Cancer genetics*, vol. 205, no. 12, pp. 613-621, 2012.
- [93] D. N. Louis *et al.*, "The 2007 WHO classification of tumours of the central nervous system," *Acta neuropathologica*, vol. 114, no. 2, pp. 97-109, 2007.
- [94] D. W. Hsu, D. N. Louis, J. T. Efird, and E. T. Hedley-Whyte, "Use of MIB-1 (Ki-67) immunoreactivity in differentiating grade II and grade III gliomas,"

- Journal of Neuropathology & Experimental Neurology*, vol. 56, no. 8, pp. 857-865, 1997.
- [95] J. Cillekens *et al.*, "A histopathological contribution to supratentorial glioma grading, definition of mixed gliomas and recognition of low grade glioma with Rosenthal fibers," *Journal of neuro-oncology*, vol. 46, no. 1, pp. 23-43, 2000.
 - [96] R. Cavaliere, M. B. S. Lopes, and D. Schiff, "Low-grade gliomas: an update on pathology and therapy," *The Lancet Neurology*, vol. 4, no. 11, pp. 760-770, 2005.
 - [97] Y. Wang and T. Jiang, "Understanding high grade glioma: molecular mechanism, therapy and comprehensive management," *Cancer letters*, vol. 331, no. 2, pp. 139-146, 2013.
 - [98] R. Rubin, D. S. Strayer, and E. Rubin, *Rubin's pathology: clinicopathologic foundations of medicine*. Lippincott Williams & Wilkins, 2008.
 - [99] P. Burger, F. Vogel, and L. Rubinstein, "Surgical Pathology of the Nervous System and its Coverings," *Journal of Neuropathology and Experimental Neurology*, vol. 42, no. 3, 1983.
 - [100] S. Bruno and Z. Darzynkiewicz, "Cell cycle dependent expression and stability of the nuclear protein detected by Ki-67 antibody in HL-60 cells," *Cell proliferation*, vol. 25, no. 1, pp. 31-40, 1992.
 - [101] A. J. Skjulsvik, J. N. Mørk, M. O. Torp, and S. H. Torp, "Ki-67/MIB-1 immunostaining in a cohort of human gliomas," *International journal of clinical and experimental pathology*, vol. 7, no. 12, p. 8905, 2014.
 - [102] A. Madabhushi and G. Lee, "Image analysis and machine learning in digital pathology: Challenges and opportunities," ed: Elsevier, 2016.

- [103] H. S. Mousavi, V. Monga, G. Rao, and A. U. Rao, "Automated discrimination of lower and higher grade gliomas based on histopathological image analysis," *Journal of pathology informatics*, vol. 6, 2015.
- [104] J. Kong, O. Sertel, H. Shimada, K. L. Boyer, J. H. Saltz, and M. N. Gurcan, "Computer-aided evaluation of neuroblastoma on whole-slide histology images: Classifying grade of neuroblastic differentiation," *Pattern Recognition*, vol. 42, no. 6, pp. 1080-1092, 2009.
- [105] O. Sertel, J. Kong, H. Shimada, U. Catalyurek, J. H. Saltz, and M. N. Gurcan, "Computer-aided prognosis of neuroblastoma on whole-slide images: Classification of stromal development," *Pattern recognition*, vol. 42, no. 6, pp. 1093-1103, 2009.
- [106] J. Kong *et al.*, "Machine-based morphologic analysis of glioblastoma using whole-slide pathology images uncovers clinically relevant molecular correlates," *PloS one*, vol. 8, no. 11, p. e81049, 2013.
- [107] P.-W. Huang and C.-H. Lee, "Automatic classification for pathological prostate images based on fractal analysis," *IEEE transactions on medical imaging*, vol. 28, no. 7, pp. 1037-1050, 2009.
- [108] S. M. Reza and K. M. Iftexharuddin, "Glioma grading using cell nuclei morphologic features in digital pathology images," in *Medical Imaging 2016: Computer-Aided Diagnosis*, 2016, vol. 9785: International Society for Optics and Photonics, p. 97852U.
- [109] P. Mobadersany *et al.*, "Predicting cancer outcomes from histology and genomics using convolutional networks," *Proceedings of the National Academy of Sciences*, p. 201717139, 2018.

- [110] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 841-852, 2010.
- [111] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, and A. Sethi, "A dataset and a technique for generalized nuclear segmentation for computational pathology," *IEEE transactions on medical imaging*, vol. 36, no. 7, pp. 1550-1560, 2017.
- [112] R. Thawani *et al.*, "Radiomics and Radiogenomics in Lung Cancer: A Review for the Clinician," *Lung Cancer*, 2017.
- [113] S. Selvarajah and S. Kodituwakku, "Analysis and comparison of texture features for content based image retrieval," *International Journal of Latest Trends in Computing*, vol. 2, no. 1, 2011.
- [114] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *bioinformatics*, vol. 23, no. 19, pp. 2507-2517, 2007.
- [115] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R news*, vol. 2, no. 3, pp. 18-22, 2002.
- [116] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367-378, 2002.
- [117] T. S. Furey, N. Cristianini, N. Duffy, D. W. Bednarski, M. Schummer, and D. Haussler, "Support vector machine classification and validation of cancer tissue samples using microarray expression data," *Bioinformatics*, vol. 16, no. 10, pp. 906-914, 2000.
- [118] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. De Jesús, *Neural network design*. Pws Pub. Boston, 1996.

- [119] E. I. Zacharaki *et al.*, "Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 62, no. 6, pp. 1609-1618, 2009.
- [120] J. Larsen and C. Goutte, "On optimal data split for generalization estimation and model selection," in *Neural Networks for Signal Processing IX, 1999. Proceedings of the 1999 IEEE Signal Processing Society Workshop.*, 1999: IEEE, pp. 225-234.
- [121] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you?: Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016: ACM, pp. 1135-1144.
- [122] S. Kinjo, J. Hirato, and Y. Nakazato, "Low grade diffuse gliomas: shared cellular composition and morphometric differences," *Neuropathology*, vol. 28, no. 5, pp. 455-465, 2008.
- [123] M. H. Faria, B. Gonçalves, R. M. Patrocínio, M. O. d. Moraes-Filho, and S. H. Rabenhorst, "Expression of Ki-67, Topoisomerase II α and c-MYC in astrocytic tumors: Correlation with the histopathological grade and proliferative status," *Neuropathology*, vol. 26, no. 6, pp. 519-527, 2006.
- [124] T. Shibata, P. Burger, and P. Kleihues, "Ki-67 immunoperoxidase stain as marker for the histological grading of nervous system tumours," in *Proceedings of the 8th European Congress of Neurosurgery, Barcelona, September 6–11, 1987, 1988*: Springer, pp. 103-106.

- [125] F. Kayaselçuk, S. Zorludemir, D. Gümürdülü, H. Zeren, and T. Erman, "PCNA and Ki-67 in central nervous system tumors: correlation with the histological type and grade," *Journal of neuro-oncology*, vol. 57, no. 2, pp. 115-121, 2002.
- [126] A. L. Johannessen and S. H. Torp, "The clinical value of Ki-67/MIB-1 labeling index in human astrocytomas," *Pathology & Oncology Research*, vol. 12, no. 3, p. 143, 2006.
- [127] S. Shuangshoti, Y. Navalitloha, V. Kasantikul, and A. Mutirangura, "Genetic heterogeneity and progression in different areas within high-grade diffuse astrocytoma," *Oncology reports*, vol. 7, no. 1, pp. 113-120, 2000.
- [128] C. Walker *et al.*, "Characterisation of molecular alterations in microdissected archival gliomas," *Acta neuropathologica*, vol. 101, no. 4, pp. 321-333, 2001.
- [129] G. S. Stoyanov, D. L. Dzhenkov, M. Kitanova, I. S. Donev, and P. Ghenev, "Correlation Between Ki-67 Index, World Health Organization Grade and Patient Survival in Glial Tumors With Astrocytic Differentiation," *Cureus*, vol. 9, no. 6, 2017.
- [130] M. B. S. Lopes, "The 2017 World Health Organization classification of tumors of the pituitary gland: a summary," *Acta Neuropathologica*, vol. 134, no. 4, pp. 521-535, 2017.
- [131] J. Balss, J. Meyer, W. Mueller, A. Korshunov, C. Hartmann, and A. von Deimling, "Analysis of the IDH1 codon 132 mutation in brain tumors," *Acta neuropathologica*, vol. 116, no. 6, pp. 597-602, 2008.
- [132] L. He, L. R. Long, S. Antani, and G. R. Thoma, "Histology image analysis for carcinoma detection and grading," *Computer methods and programs in biomedicine*, vol. 107, no. 3, pp. 538-556, 2012.

- [133] J. Torhorst *et al.*, "Tissue microarrays for rapid linking of molecular changes to clinical endpoints," *The American journal of pathology*, vol. 159, no. 6, pp. 2249-2256, 2001.
- [134] S. Robertson, H. Azizpour, K. Smith, and J. Hartman, "Digital image analysis in breast pathology—from image processing techniques to artificial intelligence," *Translational Research*, vol. 194, pp. 19-35, 2018.
- [135] K. E. Emblem *et al.*, "A generic support vector machine model for preoperative glioma survival associations," *Radiology*, vol. 275, no. 1, pp. 228-234, 2014.
- [136] K. S. Choi, S. H. Choi, and B. Jeong, "Prediction of IDH genotype in gliomas with dynamic susceptibility contrast perfusion MR imaging using an explainable recurrent neural network," *Neuro-oncology*, 2019.
- [137] L. Macyszyn *et al.*, "Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques," *Neuro-oncology*, vol. 18, no. 3, pp. 417-425, 2015.
- [138] E. I. Zacharaki, N. Morita, P. Bhatt, D. O'rourke, E. Melhem, and C. Davatzikos, "Survival analysis of patients with high-grade gliomas based on data mining of imaging variables," *American Journal of Neuroradiology*, vol. 33, no. 6, pp. 1065-1071, 2012.
- [139] B. M. Ellingson *et al.*, "Volumetric analysis of functional diffusion maps is a predictive imaging biomarker for cytotoxic and anti-angiogenic treatments in malignant gliomas," *Journal of neuro-oncology*, vol. 102, no. 1, pp. 95-103, 2011.
- [140] P. Eichinger *et al.*, "Diffusion tensor image features predict IDH genotype in newly diagnosed WHO grade II/III gliomas," *Scientific reports*, vol. 7, no. 1, p. 13396, 2017.

- [141] Z. Xing, X. Yang, D. She, Y. Lin, Y. Zhang, and D. Cao, "Noninvasive assessment of IDH mutational status in World Health Organization grade II and III astrocytomas using DWI and DSC-PWI combined with conventional MR imaging," *American Journal of Neuroradiology*, vol. 38, no. 6, pp. 1138-1144, 2017.
- [142] Louis DN, Ohgaki H, Wiestler OD, and C. WK, *World Health Organization Histological Classification of Tumours of the Central Nervous System*. Lyon, France: International Agency for Research on Cancer (IARC), 2016.
- [143] M. Perizzolo, B. Winkfein, S. Hui, W. Krulicki, J. A. Chan, and D. J. Demetrick, "IDH Mutation Detection in Formalin-Fixed Paraffin-Embedded Gliomas Using Multiplex PCR and Single-Base Extension," *Brain Pathology*, vol. 22, no. 5, pp. 619-624, 2012.
- [144] G. Reyes-Botero *et al.*, "Molecular analysis of diffuse intrinsic brainstem gliomas in adults," *Journal of neuro-oncology*, vol. 116, no. 2, pp. 405-411, 2014.
- [145] M. Sanson *et al.*, "Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas," *J Clin Oncol*, vol. 27, no. 25, pp. 4150-4154, 2009.
- [146] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 841-852, 2009.
- [147] X. Wang *et al.*, "Machine learning models for multiparametric glioma grading with quantitative result interpretations," *Frontiers in neuroscience*, vol. 12, p. 1046, 2019.

- [148] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157-1182, 2003.