



HAL
open science

Classification and segmentation of vascular network in histological images : application to renal cell carcinoma

Rudan Xiao

► **To cite this version:**

Rudan Xiao. Classification and segmentation of vascular network in histological images : application to renal cell carcinoma. Image Processing [eess.IV]. Université Côte d'Azur, 2022. English. NNT : 2022COAZ4075 . tel-04472337

HAL Id: tel-04472337

<https://theses.hal.science/tel-04472337>

Submitted on 22 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT

Classification et Segmentation du Réseau Vasculaire dans les Images Histologiques : Application au Carcinome à Cellules Rénales

Rudan XIAO

Team Morpheme (Morphology & Images)
INRIA Sophia Antipolis-Méditerranée

Laboratoire d'Informatique, de Signaux et Systèmes de Sophia Antipolis (I3S)

UMR7271 UCA CNRS

Date of Defense: 18/10/2022

**Présentée en vue de l'obtention
du grade de docteur en Informatique
d'Université Côte d'Azur**

Dirigée par : Xavier DESCOMBES, Direc-
teur de recherche, Université Côte d'Azur,
INRIA, Sophia Antipolis

Devant le jury, composé de :

Nicolas PASSAT, Professeur des universités,
Université de Reims Champagne-Ardenne,
Reims

Thomas WALTER, Professeur, Mines Paris,
PSL Université, Paris

Damien AMBROSETTI, Praticien hospita-
lier, Université Côte d'Azur, IRCAN, CHU
de Nice, Hôpital Pasteur, Nice

Daniel RACOCEANU, Professeur des uni-
versités, Sorbonne Université, Paris

Cédric WEMMERT, Professeur des universi-
tés, Université de Strasbourg, Strasbourg

**CLASSIFICATION ET SEGMENTATION DU RÉSEAU VASCULAIRE
DANS LES IMAGES HISTOLOGIQUES : APPLICATION AU
CARCINOME À CELLULES RÉNALES**

*Classification and Segmentation of Vascular Network in Histological
Images: Application to Renal Cell Carcinoma*

Rudan XIAO



Jury :

Rapporteurs

Nicolas PASSAT, Professeur des universités, Université de Reims Champagne-Ardenne,
Reims

Thomas WALTER, Professeur, Mines Paris, PSL Université, Paris

Examineurs

Damien AMBROSETTI, Praticien hospitalier, Université Côte d'Azur, IRCAN, CHU de
Nice, Hôpital Pasteur, Nice

Daniel RACOCEANU, Professeur des universités, Sorbonne Université, Paris

Cédric WEMMERT, Professeur des universités, Université de Strasbourg, Strasbourg

Directeur de thèse

Xavier DESCOMBES, Directeur de recherche, Université Côte d'Azur, INRIA, Sophia
Antipolis

Rudan XIAO

Classification et Segmentation du Réseau Vasculaire dans les Images Histologiques : Application au Carcinome à Cellules Rénales

xiii+159 p.

Classification et Segmentation du Réseau Vasculaire dans les Images Histologiques : Application au Carcinome à Cellules Rénales

Résumé

Le carcinome à cellules rénales (RCC) est l'une des tumeurs malignes les plus courantes et le diagnostic pathologique postopératoire est la méthode étalon pour le diagnostic du RCC. La reconnaissance du type de tumeur RCC et la possibilité de migration cellulaire dépendent fortement des propriétés géométriques et topologiques du réseau vasculaire. Dans ce travail, nous explorons le réseau vasculaire visible dans les images histopathologiques du RCC et étudions s'il est suffisant pour caractériser le sous-type de RCC. Pour ce faire, nous construisons d'abord une nouvelle base de données d'images histopathologiques RCC, à savoir l'ensemble de données BigRCC, qui contient 63430 images de carcinome à cellules claires (ccRCC), de carcinome papillaire (pRCC), chromophile (chRCC) et oncocytome. Ensuite, nous construisons une seconde base de données d'images histopathologiques de RCC du réseau vasculaire de 31 patients, à savoir la base de données VRCC, contenant 424 annotations du réseau vasculaire, provenant d'images de patches tumoraux de ccRCC, pRCC et chRCC. Sur la base de ces réseaux vasculaires du jeu de données VRCC, nous proposons de nouvelles caractéristiques, à savoir des caractéristiques du squelette du réseau et des caractéristiques des zones entre les vaisseaux. Ces caractéristiques représentent bien les propriétés géométriques et topologiques des réseaux vasculaires des images histopathologiques. Ensuite, nous construisons des résultats de référence solides avec divers algorithmes (modèles d'apprentissage traditionnels et profonds) sur l'ensemble de données VRCC. Les résultats des caractéristiques de squelette et de réseau surpassent les modèles classiques par apprentissage profond. Malheureusement, il existe peu de grandes bases de données de réseaux vasculaires annotées manuellement. De plus, l'annotation du réseau vasculaire est très chronophage, ce qui limite son application potentielle. En outre, la segmentation automatique du réseau vasculaire à partir d'images histopathologiques reste un défi majeur en raison de la complexité de l'arrière-plan. Pour faciliter l'utilisation du réseau vasculaire nous proposons, dans cette thèse, une méthode qui réduit la dépendance aux données étiquetées grâce à l'apprentissage semi-supervisé (SSL). De plus, compte tenu de la corrélation entre la classification tumorale et la segmentation vasculaire, nous proposons un modèle d'apprentissage multi-tâches (MTL) qui peut simultanément segmenter le réseau vasculaire à l'aide du SSL et prédire la classe tumorale dans un contexte supervisé. Cette procédure d'apprentissage multi-tâches offre une solution d'apprentissage automatique de bout en bout pour segmenter conjointement le réseau vasculaire à partir d'images histopathologiques et classer les sous-types de tumeurs. Des expériences ont été menées sur l'ensemble de données VRCC, puis testées sur nos propres ensembles de données RCC et sur les données de la base TCGA. Les résultats montrent que le modèle MTL-SSL proposé surpasse l'approche conventionnelle de segmentation par apprentissage supervisé.

Mots-clés : Carcinome à cellules rénales, Base de Données d'images Histopathologiques RCC, Réseau Vasculaire, Apprentissage Semi-supervisé, Apprentissage Multi-tâches.

Classification and Segmentation of Vascular Network in Histological Images: Application to Renal Cell Carcinoma

Abstract

Renal Cell Carcinoma (RCC) is one of the most common malignancies and histopathological image analysis is a gold standard for RCC diagnostic method. Recognizing the type of RCC tumor and the possibility of cell migration highly depends on the geometric and topological properties of the vascular network. Motivated by the diagnosis pipeline, we explore whether the vascular network visible in RCC histopathological images is sufficient to characterize the RCC subtype. To achieve this, we firstly build a new RCC histopathological image dataset, namely the BigRCC dataset, which contains 63430 patch images of clear Cell RCC (ccRCC), papillary RCC (pRCC), Chromophobe (chRCC), and Oncocytoma. Then we build a vascular network-based RCC histopathological image dataset of 31 patients, namely the VRCC dataset, with 424 weakly-labeled vascular network annotations, coming from tumor patch images of ccRCC, pRCC, and chRCC. Based on these vascular networks of the VRCC dataset, we propose new handcrafted features, namely skeleton features and lattice features. These features well represent the geometric and topological properties of the vascular networks of RCC histopathological images. Then we build strong benchmark results with various algorithms (both traditional and deep learning models) on the VRCC dataset. The results of skeleton and lattice features outperform popular deep learning models. Unfortunately, there is a lack of large manually annotated vascular network databases for learning-based methods. Moreover, vascular network annotating is very time-consuming, which would limit its potential application. Besides, automatic vascular network segmentation from histopathological images is still a challenge due to the background complexity. To facilitate vascular network application, in this thesis, we propose a method that reduces reliance on labeled data through semi-supervised learning (SSL). Additionally, considering the correlation between tumor classification and vascular segmentation, we propose a multi-task learning (MTL) model that can simultaneously segment the vascular network using SSL and predict the tumor class in a supervised context. This multi-task learning procedure offers an end-to-end machine learning solution to jointly segment the vascular network from histopathological images and classify tumor subtypes. Experiments were carried out on the VRCC dataset and then tested on both own RCC and open-source TCGA datasets. The results show that the proposed MTL-SSL model outperforms the conventional supervised-learning segmentation approach.

Keywords: Renal Cell Carcinoma, RCC Histopathological Image Dataset, Vascular Network, Hand-crafted Features, Semi-Supervised Learning, Multi-Task Learning.

Acknowledgements

The three-year doctoral time has passed like a flash, and I have been exploring and devoting myself to making contributions and achievements in the field of human life sciences during my study career. In order to comprehensively and systematically learn the relevant knowledge and the application of extensive tools, I transitioned from medicine to genetics and then to computer science. The study and exploration of medical image processing at the doctoral stage strengthened my determination and enthusiasm to use computer knowledge to solve medical problems. I hope my enthusiasm will continue further. I am very fortunate to have met many good teachers and good friends who have influenced my life over these three years. I am really grateful for our encounter. It is your help and encouragement that allowed me to successfully complete my thesis and defense.

First of all, I would like to thank Professor Xavier Descombes for providing me with the opportunity for my doctoral study, and also for his patience, trust, support and guidance all the time. Professor Xavier Descombes inspired me with a lot of mathematical ideas, suggestions, and my Matlab code support. Your serious and professional attitude toward scientific research has also greatly influenced me. I have benefited from your creativity and openness to ideas during this thesis. And I should learn more from you along with my scientific research career in the future. Thanks to Professor Damien Ambrosetti for providing medical guidance and data support. The histopathological slides used in my project were all selected by your observation under a microscope. Thank you for your professional medical advice, which is in our interdisciplinary vital in cooperation. Thanks to Eric Debreuve for your help with my project. Thanks also to our project collaborators Francesco Ponzio and Mohammed Lamine Benomar, our cooperation has promoted international academic exchanges and made our research more valuable. We have interpreted the significance of artificial intelligence in the medical field from different perspectives and using different methods. Thanks to the efforts of my jury members who agreed to take the time to provide valuable insights during our interactions. I would like to especially thank the president of my jury, Professor Daniel Racoceanu, for your encouragement and valuable comments during my defense, and two reviewers, Nicolas Passat and Thomas Walter, for their insightful comments, for taking the time to review the manuscript and giving precious ideas to improve my thesis. Also thanks to Luca Calatroni, you are always very kind, organizing meetings to improve and interact with our work, and looking for learning opportunities and cooperation for us. I would also like to thank the secretaries Nadia Belfegas-Fagues and Isabelle Strobant for their patient help and careful guidance on complex administrative issues throughout my PhD and even after my contract ended. Thanks to the Université Côte d'Azur, I3S laboratory, INRIA and CHU Hospital for providing research funding, data support, learning resources, office, and leisure activities.

I am most grateful to the team that I have worked with over these years, the joint team of Morpheme, INRIA, and I3S. Thanks guys on our team. First of all, I would like to thank my first office roommate Sarah and the PhDs Asma, Marta, Vasilina, Xuchun, Cédric, Bastien, Alex, Imane, and also the interns Yueqiao, Fatima, Sai, Assia, Oumayma, Apolline, Kevin, Mayeul, Youness, really miss our coffee time, academic and outdoor activities we participated in together. Also thank you for your friendliness, Somia, Arne, Anca, and Ana, although we have not spent a long time.

Thanks also to my great friends, Yingyu, Enlin, Buntheng, Zhongwei, Yao, Fernando, Tong, Flora, Rémi, Dingge, Mulin, Zhuoer, Ruochen, Yangdi, Yangli, Yuqing, Yuqi, Dandan, Rui, etc., thank you for the care, companionship and support given to me over the years. We've been through a lot together, cooking, traveling, hiking, camping, etc. We share the joys of life with each other, which makes my life outside of scientific research colorful and joyful! I am so happy to have such a precious group of friends in my life!

I would also thank my family. Due to the coronavirus (COVID-19) pandemic, I did not go home even once during my doctoral period. Thank you for your online support and understanding. You always unconditionally support all my decisions on the road of scientific research. Your love is my solid backing in the face of all difficulties, and also my motivation and courage to keep moving forward!

Finally, I am grateful to all the people I met, and I am sorry for not being able to list everyone's names here. You are the light on my life path, and I am very lucky to meet all of you!

Contents

1	Introduction	3
1.1	Preface	3
1.2	General Context of Our Research	4
1.2.1	Brief introduction of medical aspects	4
1.2.2	Objective and motivations	4
1.3	Thesis Organization	5
1.3.1	Main contributions	5
1.3.2	Manuscript organization	6
1.3.3	List of publications	6
Background and Histopathological Images Dataset Building		
2	Machine Learning in Histopathological Images	11
2.1	Histopathological Images	13
2.1.1	Brief Introduction	13
2.1.2	Sample Preparation	14
2.1.3	Microscope Features and Scanning	18
2.1.4	WSIs and Image Preprocessing	19
2.2	Histopathological Images Analysis	23
2.2.1	Features Extraction from Histopathological Images	23
2.2.2	Deep Learning Approach	25
2.3	Computer-Aided Diagnosis and Prognosis (CAD & CAP)	34
2.3.1	CAD	34
2.3.2	CAP	35
2.3.3	Limitations and Improvements of Computer-Aided Technology	35
2.3.4	The Prospect of Computer-Aided Technology in the Diagnosis of Renal Cancer	36
2.3.5	The Future of Computer-Aided Technology	37
2.4	Conclusion	37
3	Histopathological Image Datasets Building	39
3.1	Epidemiology and Main Subtypes of RCC	41
3.1.1	The Phenotype and Histomorphology of RCC Subtypes	42
3.1.2	Cellular and Vascular Characteristics in RCC Subtypes	44
3.2	Histopathological Images Dataset	46
3.2.1	BigRCC Dataset Building	46
3.2.2	TCGA Open-access Dataset	51
3.2.3	VRCC (Vascular RCC) Dataset Building	53
3.3	Conclusion	55

RCC subtypes Classification of Histopathological Images

4	Vascular Network-based Features	61
4.1	Hand-crafted Features	63
4.1.1	Skeleton Features	63
4.1.2	Lattice Features	63
4.2	Graph-based Features	66
4.3	Deep Learning Features	66
4.4	Conclusion	67
 5	 Vascular-based RCC Classification	 71
5.1	Vascular-Based RCC Classification Benchmark Experiments	73
5.1.1	Classification Pipeline from Vascular Network of RCC	73
5.1.2	Experimental Environment	74
5.2	Deep Features Based Classification with Convolutional Neural Networks	75
5.2.1	Inputs of Raw Images	75
5.2.2	Inputs of Weak Vascular Network Mask	75
5.2.3	LeNet	75
5.2.4	AlexNet	75
5.2.5	VGGNet	78
5.2.6	ResNet	80
5.3	Graph-based Classification with Graph Convolutional Networks	81
5.3.1	A Variant of Graph Convolution Networks	81
5.3.2	Pooling Layer Based on Self-Attention Graph	82
5.4	Hand-crafted Features Based Classification with Traditional Machine Learning Algorithm	86
5.4.1	Hand-crafted Features Analysis	86
5.4.2	Adaptive Boosting	89
5.4.3	Decision Tree	90
5.4.4	Gradient Boosting Tree	91
5.4.5	K-Nearest Neighbor	92
5.4.6	Logistic Regression	93
5.4.7	Random Forest	95
5.4.8	SVM RBF & SVM Sigmoid	95
5.4.9	ROC Plotting and AUC Calculation	99
5.5	Results	101
5.5.1	Results of Hand-crafted Features Analysis	101
5.5.2	Performance of Hand-crafted Features on Traditional Algorithms under Non-parameter Test	101
5.5.3	Results of Vascular-Based RCC Classification Benchmark	102
5.6	Conclusion	103

Automatic Segmentation of Vascular Network

6	Coupling Vascular Network Segmentation and RCC Classification	107
6.1	Multi-Task Learning Pipeline	109
6.1.1	HRNet Backbone	109
6.1.2	Mean Teacher	109
6.1.3	Proposed MTL-SSL model	110
6.2	Proposed Vessel Loss Function	114
6.2.1	Definition of Loss Functions	114
6.2.2	Main Idea and Realization of Our Proposed Loss Function.	117
6.3	Vascular Network Segmentation Benchmark of RCC	119
6.3.1	Experimental Environment	119
6.3.2	Backbone Choice	119
6.3.3	Segmentation Benchmarks of Vascular Network	120
6.4	Test on New Subtype of RCC and Other Type of Cancers	122
6.4.1	New Subtype of RCC and Other Cancers	122
6.4.2	Test on New Subtype of RCC and Other Cancers Dataset	125
6.5	Conclusion	130

Conclusion and Perspectives

7	Conclusion and Perspectives	135
7.1	Conclusion	135
7.2	Perspectives	136

	Bibliography	139
--	---------------------	------------

	List of Figures	157
--	------------------------	------------

	List of Tables	159
--	-----------------------	------------

Abbreviations

CAD	Computer-aided diagnosis
WSI	whole slide image
AI	artificial intelligence
RCC	renal cell carcinoma
ccRCC	clear cell renal cell carcinoma
pRCC	papillary renal cell carcinoma
GCN	Graph Convolutional Network
CNN	Convolutional Neural Network
SSL	semi-supervised learning
MTL	multi-task learning
H&E staining	Hematoxylin and eosin staining
IHC	Immunohistochemistry
H&DAB staining	Hematoxylin and diaminobenzidine staining
PCA	Principal Component Analysis
TCGA	The Cancer Genome Atlas
CT	computed tomography
EMA	Exponential Moving Average
CE	cross-entropy
MSE	Mean Square Error
CPS	Cross Pseudo Supervision
CCT	Cross Consistency Training
EM	Entropy Minimization
DCT	Deep Co-Training
MT	Mean Teacher
MVD	Microvessel density
ROC	Receiver operating characteristic curve
AUC	Area under the ROC Curve
FPR	False positive rate
TPR	True positive rate
KNN	K-Nearest Neighbor
RF	Random Forest
SVM	The support vector machine
SMO	Sequential Minimal Optimization
GNN	Graph Neural Network
SAGPool	Self-Attention Graph Pooling
MLP	Multilayer Perceptron
LRN	Local Response Normalization
FIFO	first-in-first-out
MSE	Mean Squared Error
CE	Cross Entropy

CHAPTER 1

Introduction

1.1 Preface

Computer-aided diagnosis (CAD) is an interdisciplinary technology that combines elements of artificial intelligence and computer vision with radiology and medical imaging. This technology has the advantages of accurate quantitative analysis, repeatability, and a reduced workload for radiologists.

In recent years, artificial intelligence and smart medicine have become hot topics in the medical field and even globally in society. Conventional pathology consists of tissue sections after fixation, dehydration, wax immersion, embedding and other processes. After staining, the pathologist analyzes the characteristics of the lesions by reading microscopy images to determine the diagnosis result. Pathology has always been the "gold standard" in the medical diagnosis field, and pathological diagnosis is the final step in disease determination. CAD also has potential future applications in digital pathology with the advent of whole slide imaging (WSI) and machine learning algorithms. It is particularly investigated for the standard H&E stain [Shiraishi et al., 2011]. The emergence of digital pathology has effectively saved the manpower and time costs of the hospital pathology department, improved the quality and efficiency of pathological diagnosis, and even reduced the risk of misdiagnosis, alleviating the development dilemma of the pathology to a certain extent.

In digital pathology, digitizing slices is the most basic and critical step. The WSI is scanned and collected by an automatic microscope or optical magnification system to obtain high-resolution digital images. Then, a computer is used to automatically perform high-precision multi-scaled and thus obtain and process high-quality data for application in various fields of pathology. The development of digital pathology is based on high-throughput and fast WSI technology. The pathology department makes all routine slices into digital slices and integrates them into the daily workflow to realize the whole process of pathology with digital quality control and management. Combined with computer science technology, it establishes a pathological diagnosis platform. Based on the all-digital pathology, the archived digitized slices form rich datasets. When these datasets are combined with emerging computer algorithms such as artificial intelligence (AI) and the powerful computing power of computers, the computer can automatically detect the lesion area in the digital slices and quantitatively evaluate various indicators, helping pathologists to make a fast, accurate and highly repeatable pathological diagnosis. At present, CAD software will gradually become an indispensable tool in the daily diagnosis of pathologists.

1.2 General Context of Our Research

This project is launched by CHU de Nice - Hôpital Pasteur and our Morpheme team, a joint team between INRIA, CNRS and Université Côte d'Azur.

1.2.1 Brief introduction of medical aspects

Renal cell carcinoma (RCC) is a highly malignant tumor in the urinary system. 90% of kidney cancers are RCC [Hsieh et al., 2017], which is mainly divided into clear cell RCC [Lopez-Beltran et al., 2006] (ccRCC) accounting for 75% of RCC, papillary RCC (pRCC) accounting for 10% of RCC and Chromophobe (chRCC) accounting for 5% of RCC. RCC classification is a challenging task. Cell morphology, tumor architecture, phenotype and genetics data are mainly used to define the tumor subtype. Most of the current classification research is focused on the search for biological biomarkers, aiming to define the RCC subtype and also predict the behavior of the tumor [Cheville et al., 2003, Gao et al., 2020, Cheng et al., 2020]. However, vasculature also plays an important role [Loukas and Linney, 2004a] in histopathological diagnosis [Prasad et al., 2006]. The vascular structure of RCC subtypes is different and may lead to an accurate diagnosis. For example, ccRCC is characterized by a fishnet-like vascular architecture, the pRCC has a tree-like structure [Zubiolo et al., 2016] while the chRCC has a linear structure [Nagashima, 2000]. As shown in Figure 3.15 (see chapter 3) the ccRCC vascular networks are denser and contain more junctions but fewer end branches than pRCC, the density of chRCC vascular network is the lowest among subtypes. In this project, we explore the importance of the vascular network in the RCC diagnosis. We not only extracted hand-crafted features of the vascular networks to classify subtypes of RCC but also proposed an end-to-end semi-supervised learning (SSL) model to segment vascular networks automatically, which can outperform the fully supervised method.

1.2.2 Objective and motivations

Since there is no public RCC dataset with vascular network annotations, we build the **VRCC** dataset with vascular annotations. VRCC is extracted from our larger RCC histopathological image dataset (BigRCC) randomly, which is labeled with RCC categories. The BigRCC dataset contains the data of 167 WSI, coming from 74 patients, which are cropped into 39986 image patches of ccRCC, 18254 image patches of pRCC and 4420 image patches of chRCC respectively. VRCC has 424 vascular network segment images of ccRCC, pRCC and chRCC, coming from 31 patients (13 for ccRCC, 14 for pRCC and 4 for chRCC).

To explore the potential of traditional algorithms, we propose two sets of features, called "**skeleton features**" and "**lattice features**", which are extracted from the vascular network. Specifically, we compute the skeleton of the vascular network. The skeleton is a structure that embeds the topological and geometrical properties of the vascular network, as shown in Figure 3.15 (see chapter 3). It is composed of three types of elements : junction, non-end branch and end branch. Firstly, we define some meaningful features from the skeleton to form the skeleton features, containing the number of end branches (NE), small NE, long NE, and density, etc. Then, we perform a series of operations on the skeleton to obtain the lattice spatial map. The lattice features that represent the regions between vessels are extracted from this spatial map. Lattice features include a set of features, such as the mean area, median area, etc. in Table 4.1 (see chapter 4).

Then we build solid benchmark results of traditional and deep learning methods on the VRCC dataset. The results of traditional algorithms with our skeleton and lattice features outperform the results of popular deep learning models (Graph Convolutional Network (GCN) [Zhang et al., 2018a], Convolutional Neural Network (CNN) [He et al., 2016, Simonyan and Zisserman, 2014]).

In order to enhance the application value of the vascular network, we planned to build an automatic vascular network segmentation model. However, data labeling is often the most challenging task for the segmentation problem when considering deep learning approaches. Labeling large-scale images is laborious, time-consuming and exhibits low repeatability. This encouraged us to find out if it was possible to improve the vascular network segmentation performance using unlabeled datasets. This is indeed the paradigm of SSL models. Compared with the difficulty of obtaining manually a vascular network mask for the segmentation task, the labeling for the classification task is easy to obtain. We conjectured that joint supervised classification and SSL for vascular network segmentation, both embedded in a multi-task learning (MTL) model, may improve the performance of vascular network segmentation in RCC histopathological images. Additionally, we annotated 12 vascular network masks of another subtype of RCC, named Oncocytoma, only for test.

Finally, we conducted benchmark experiments of supervised learning, SSL, both single and multi-tasks, on RCC histopathological images. The proposed MTL-SSL model performs best, outperforming and being more robust than the supervised learning models which only use a few labeled data. Moreover, compared with the single-task SSL, our model indeed improves the segmentation efficiency of the vascular network while also performing tumor classification.

1.3 Thesis Organization

1.3.1 Main contributions

The main contributions of the work can be summarized as follows :

- We present the first work to investigate the importance of geometric and topological properties of the vascular network for RCCs classification.
- We build a new database and vascular annotated datasets, "BigRCC" and "VRCC", for RCCs classification and vascular network segmentation.
- We proposed two sets of hand-crafted features, "skeleton features" and "lattice features", to represent the vascular network, which is extracted from the vascular network segmentation images.
- We build benchmark results on the VRCC dataset, showing that our proposed features-based traditional classifier can outperform deep learning models.
- We propose an MTL-SSL model performing joint SSL segmentation and classification tasks to segment the vascular network using both labeled and unlabeled data.
- We apply an automatic, end-to-end machine learning vascular network segmentation method in histopathological images, which is robust and outperforms other popular SSL and supervised learning methods.
- The proposed MTL-SSL model forms a foundation for future developments in multi-task learning dealing with vascular segmentation and classification from histopathological images.

1.3.2 Manuscript organization

Part I (Chapter 1) introduces the background, main context, and objectives of this thesis focused on the segmentation and classification of the vascular network from the RCC histopathological images. Main contributions, as well as publications, are illustrated here.

Part II (Chapter 2) focuses on the background of the application of machine learning in histopathological images. Meanwhile, we present the state of the art and the current status and prospects of machine learning-assisted pathological diagnosis.

In Part III (Chapter 3), we present the medical background of the development and subtype character of RCC with cell nucleus and vascular network and the pipeline of the annotated histopathological images dataset building. The BigRCC dataset includes the category information of RCC, which divides into tumor patch images and non-tumor (fiber, necrosis, and normal) patch images. The VRCC dataset is a part of the tumor patch images from the BigRCC dataset, which contain the vascular network mask information.

Part IV (Chapter 4) describes the hand-crafted features (skeleton features and lattice features) extraction from the vascular network of ccRCC, pRCC, and chRCC, these features contain the topological and geometrical characterization of the vascular network. Then, we calculate the statistical significance between the features of 3 RCC subtypes. Finally, we choose the features that are significant in all significant tests to classify RCC subtypes and compare their performance with former work.

In Part V (Chapter 5), we compare the classification performance of our proposed hand-crafted features with the graph-based representation feature, which only contains the topological information, and with the deep features, which are represented by raw tumor patch images and the binary patch images of the vascular network masks.

Part VI (Chapter 6) mainly focuses on automatic vascular network segmentation. We propose an MTL-SSL model, in which joint tumor subtypes classify and SSL segment the vascular network to improve the performance of vascular network segmentation in RCC histopathological images. Both labeled and unlabeled patch images are used for training, which reduces the reliance on manually vascular network annotation. Then we conduct benchmark experiments of supervised learning, SSL, both single and multi-tasks, on RCC histopathological images and compare them with our former work as well. Moreover, we involve the new subtype of RCC, named Oncocytoma, and other cancer patch images obtained from another open-access database TCGA, which are added only to the test set for the model robustness without training.

Part VII (Conclusion and Perspectives) is the discussion, conclusion, and perspective of the project. In summary, we outperform the state of the art of RCC subtypes classification by hand-crafted features, and the vascular segmentation by the MTL-SSL model. Meanwhile, we illustrate the value of the vascular network for RCC subtyping and the prospect of this project in application to the pathological diagnosis.

1.3.3 List of publications

- **Rudan Xiao**, Damien Ambrosetti, Eric Debreuve, Xavier Descombes. "Renal Cell Carcinoma Classification from Vascular Morphology," *MICCAI 2021*. DOI : 10.1007/978-3-030-87231-1_59.

- Mohammed Lamine Benomar, Nesma Settouti, **Rudan Xiao**, Damien Ambrosetti, Xavier Descombes. "Convolutional Neuronal Networks for Tumor Regions Detection in Histopathology Images," *ICDTA 2021*. DOI : 10.1007/978-3-030-73882-2_2.
- **Rudan Xiao**, Damien Ambrosetti, Xavier Descombes. "Semi-Supervised Multi-Task Learning for Vascular Network Segmentation and Classification of Renal Cell Carcinoma," *MICCAI Workshop 2022 (Best Poster Award)*. DOI : 10.1007/978-3-031-16876-5_1.

Dissemination (presentation)

- Poster presentation, AI in medical data analysis, December 9-10th, 2019, Hôpital Pasteur, CHU Nice, France.
- Presentation in the conference, Imaging & machine learning meeting, January 29-31, 2020, Sophia-Antipolis, France.
- Poster presentation, MICCAI 2021, Strasbourg, France.
- Poster presentation, REMIA 2022, Resort World Convention Centre Singapore, Singapore.

Background and Histopathological Images Dataset Building

CHAPTER 2

Machine Learning in Histopathological Images

Dans ce chapitre, nous présentons le contexte médical de l'imagerie histopathologique et son analyse par la méthode traditionnelle d'extraction de caractéristiques et l'approche par apprentissage profond. Nous décrivons également la situation du développement du diagnostic/pronostic assisté par ordinateur et son application dans le cas du carcinome rénal à cellules claires.

In this chapter, we present the medical background of the histopathological image and its analysis with the traditional feature extraction method and deep learning approach. We also discuss the development situation of computer-aided diagnosis/prognosis and the application in RCC.

2.1	Histopathological Images	13
2.1.1	Brief Introduction	13
2.1.2	Sample Preparation	14
2.1.3	Microscope Features and Scanning	18
2.1.4	WSIs and Image Preprocessing	19
2.2	Histopathological Images Analysis	23
2.2.1	Features Extraction from Histopathological Images	23
2.2.2	Deep Learning Approach	25
2.2.2.1	Back Propagation Neural Network	25
2.2.2.2	Deep Learning Model Architecture	25
2.2.2.3	Deep Learning Development Platform	29
2.2.2.4	Data Augmentation	29
2.2.2.5	Histopathological Images Classification	29
2.2.2.6	Histopathological Images Segmentation	31
2.3	Computer-Aided Diagnosis and Prognosis (CAD & CAP)	34
2.3.1	CAD	34
2.3.2	CAP	35
2.3.3	Limitations and Improvements of Computer-Aided Technology	35
2.3.4	The Prospect of Computer-Aided Technology in the Diagnosis of Renal Cancer	36
2.3.5	The Future of Computer-Aided Technology	37
2.4	Conclusion	37

2.1 Histopathological Images

2.1.1 Brief Introduction

Histology is made up of the Greek words *histo* (organization) and *logos* (science). Histology is the science that studies the fine structure of the normal human body and its related functions. Microstructure refers to the structure seen under a microscope [Ross and Pawlina, 2006].

The first optical microscope was invented in 1650 by Jansen and Sons in the Netherlands. In 1665, Hooke observed thin slices of plant tissue with a self-made microscope and named cells for the first time. At present, there is not only an optical microscope but also a laser confocal microscope (mainly detecting fluorescence signal), electron microscope (divided into the scanning electron microscope, observe the surface of the sample, resolution 6-10 nm ; transmission electron microscope, transmission sample observation, resolution 0.1-0.2 nm) are widely used.

In 1801, Marie François Bichat (1771-1802) first proposed the word tissue and divided the human body into 21 kinds of tissues [Maulitz, 1973]. August Franz Mayer (1787-1865), a German, became Bern's former surgeon in 1813 and became the professor of anatomy, pathological anatomy and physiology of town in 1815. He divided tissues into 8 types and coined the term histology.

In the middle and late 19th century, scientists invented many histological techniques, which led to the discovery of many types of cells and the fine structure of tissues. Italian scientist Millo Golgi (1843-1926) and Spanish scientist Santiago Ramon y Cajal (1852-1934), who won the Nobel Prize in 1909, founded the silver staining technique and used it in the study of nerve cells and nerve tissue. They are considered the founder of Neuroscience.

From the end of the 19th century up to now, modern histology has developed into a golden age, and several Nobel Prizes have been awarded in this discipline. For example, in 1908, the relationship between phagocytes and body defense has been discovered. The functional junction between neurons is named synapse, which puts forward the theory of neural reflex and the concept of the synapse. In 1986, Charles Scott Sherrington (1857-1952) won the Nobel Prize for researching the important functions of nerve growth factor and epidermal growth factor in nerve growth and embryonic development [Perea et al., 2009].

Nowadays, histochemistry continues to advance, and in addition to basic research, it has also been fully applied in medical diagnosis. It is the main research method in the discipline of pathology. Autopsy, biopsy, and cell smears all require histological techniques for analysis.

In tumor diagnosis, apart from clinical features, biochemical diagnosis, and imaging data, only histopathology is the "gold standard" for diagnosing tumors : source, malignancy, the extent of tumor spread, etc.

A pathological section is a certain size of diseased tissue. Usually, the diseased tissue is embedded in a paraffin block, sliced into thin slices with a microtome, and then the staining method is selected according to the purpose. Further examination is carried out with a microscope. The occurrence and development of lesions are analyzed, and finally, a pathological diagnosis is done.

There are three ways to detect molecular pathology :

- The antigen-antibody reaction is used to specifically recognize the target protein, and the luciferase-labeled antibody is called immunofluorescence (IF). The corresponding sections need to be observed under a fluorescence microscope.
- The second is also based on the antigen-antibody reaction, but it is labeled with enzymatic chemistry and is called immunohistochemistry (IHC). After color development, the shades

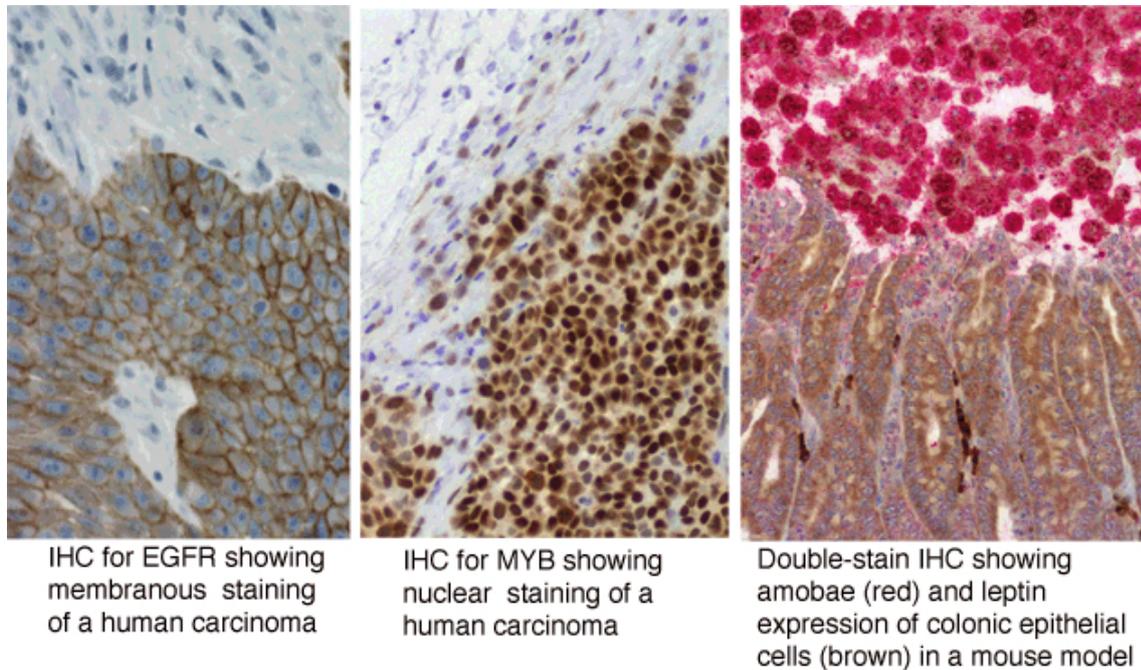


Figure 2.1 – Example of IHC staining.

Display specific tissue structures by applying different antibodies (source : Biorepository and Tissue Research Facility, School of Medicine, University of Virginia).

are different and can be observed under an optical microscope, an example of IHC staining is shown in Figure 2.1.

- For direct detection of target DNA or RNA, nucleic acid probes with labels are usually used for hybridization, which is called in situ hybridization (ISH). The probe markers are divided into radioisotopes labeled ^3H , ^{35}S , ^{125}I , etc., or non-radioactive. The radioactive requires autoradiography, and the non-radioactive can be observed by IHC.

Traditional immunohistochemical staining, such as H&E staining, uses Hematoxylin and Eosin staining. Hematoxylin is alkaline, which can make nuclear chromatin and ribosomes blue-violet, and eosin is acidic, marking cytoplasmic, thus the extracellular matrix becomes pink. Benign and malignant can be judged by qualitative and quantitative analysis, as shown in Figure 2.2 [Kausar et al., 2019]. In addition, there are also various staining methods shown in Figure 2.3, such as Sirius red staining for collagen components [Li et al., 2016a], Golgi staining for marking nerves [Du, 2019], Masson trichrome staining for connective tissue [Foot, 1933], and oil red O for fat [Stanton et al., 2011] staining. Labeling a wide variety of tissue results in a colorful diagnostic map.

2.1.2 Sample Preparation

Histopathological sections are divided into paraffin section, frozen section and vibrating section according to the different preparation methods. The sections used in our project are from paraffin.

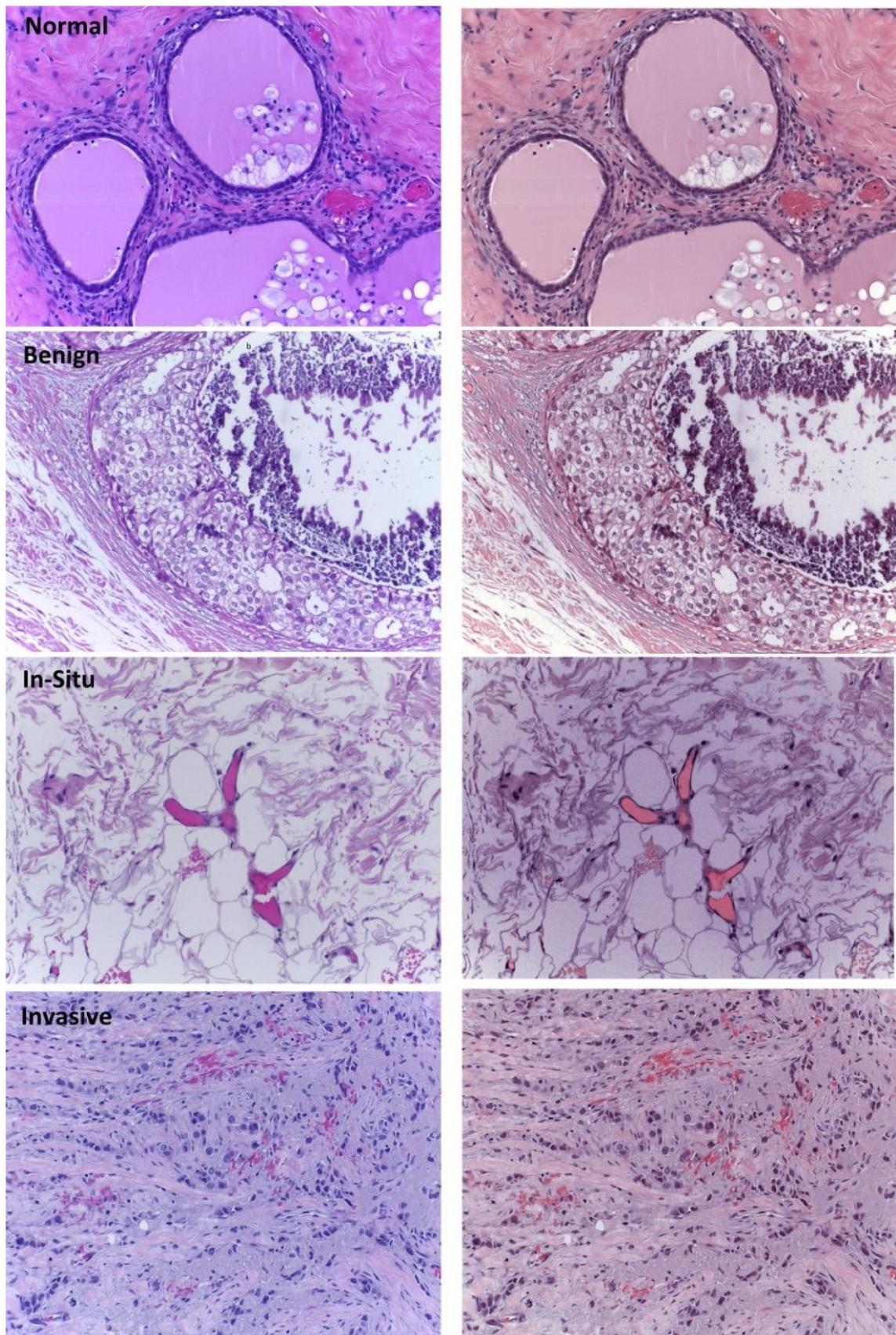


Figure 2.2 – Benign and malignant grading under H&E staining. Judging by the morphological features, geometric features, texture features and shape features of cells and nuclei. The left side is original raw H&E stained microscopy images and the right side is corresponding stain color normalized images. Microscopy images are taken from the ICIAR 2018 dataset (Kausar, 2019).

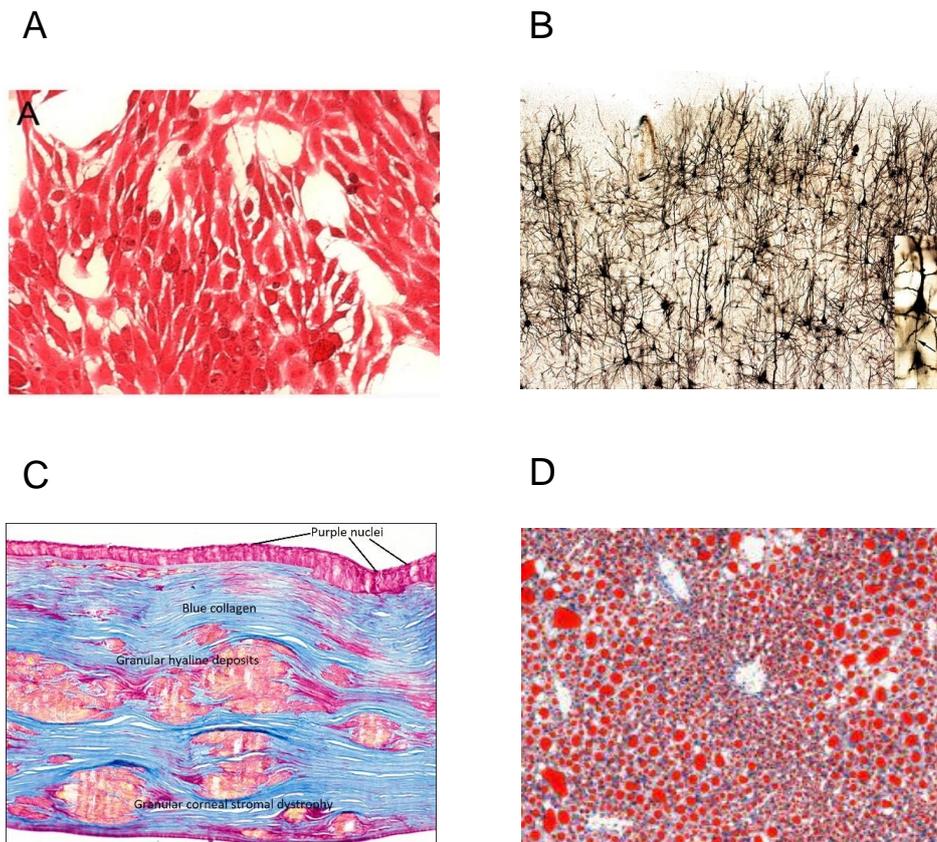


Figure 2.3 – Special stains to reveal specific tissue structures.

A : Sirius red staining (Li, 2016), B : Golgi staining (Du, 2019), C : Masson trichrome staining (source : Moran Center, University of Utah Health Care), D : Oil red O staining (Stanton, 2011).

The preparation process of paraffin sections includes the steps of sampling, fixing, washing and dehydration, clearing, dipping in wax, embedding, sectioning and sticking, dewaxing, staining, dehydration, clearing, and sealing. Generally, it takes several days from the sampling and fixing to the cover slip to make the slide sample.

— Fixing

Impregnate fresh materials with appropriate 4% paraformaldehyde to solidify or precipitate the material components in cells and tissues, stop all metabolic processes of cells, prevent cell autolysis or tissue changes, harden tissues, and maintain the structure of the living body as much as possible. Fixed time from 1 hour to several days, usually several hours to 24 hours.

— Washing and dehydration

After the tissue is fixed, the fixative infiltrated into it should be washed away, otherwise, the fixative left in the tissue will hinder the staining, and some will cause precipitation or crystallization.

The tissue after fixation or washing is full of water. If the water is not removed, subsequent clearing, wax dipping and embedding cannot be carried out, because most clearing agents are benzene, and neither benzene nor paraffin can be mixed with water. Therefore, dehydration must be carried out before transparent and wax dipping. Dehydration is to use of a dehydrating agent to replace the water in the tissue to facilitate the infiltration of organic solvents. Commonly used dehydrating agents are a series of different concentrations of ethanol.

The dehydration steps are : 80%, 90%, 95%, 100% ethanol of various concentrations for 2 hours. This process can be completed automatically by a dehydrator.

— Clearing

After tissue dehydration, because ethanol is insoluble in paraffin, for paraffin to be immersed in tissue blocks, it must go through a substitution process of a solvent that can be mixed with alcohol and dissolve paraffin, for alcohol to achieve the immersion of tissue blocks in paraffin. The material block is immersed in this kind of medium immersion liquid and appears transparent. This liquid is called the transparent agent.

Currently, the best clearing agent is xylene. Usually, the tissue is first immersed in a mixture of pure alcohol and half of the clearing agent for 1 to 2 hours and then transferred to the pure clearing agent for immersion.

— Wax dipping and embedding

After the tissue is transparent, it is immersed in molten paraffin, so that the paraffin penetrates into the tissue and replaces the xylene in the tissue. This process is called paraffin immersion.

The melting point of paraffin wax used for dipping wax is 56 °C, and the dipping wax should be carried out in an incubator about 3 °C higher than the melting point of paraffin wax.

The melted paraffin is poured into the embedding frame, and the paraffin-impregnated tissue block is placed with the cut side down with warmed tweezers. When the surface layer of the wax liquid solidifies, it will be rapidly cooled.

— Slicing

The embedded wax block is trimmed into a square or rectangle with a blade, cut into a 4 ~ 7 μ m wax strip on a microtome, and gently hold up with a brush and placed on the paper.

- Spread, fish and bake

Exhibit :

Ophthalmic tweezers are used to pick up the wax tape which is gently spread on the water surface of 40-45°C. The tension and the temperature of the water are used to naturally flatten the slightly wrinkled wax tape.

Fishing pieces :

After the section is fully flattened on the constant temperature water, the wax section is moved to the middle of the glass slide, and the remaining water on the glass slide is eliminated.

Baked slices :

Usually baked in an oven at 60°C

About 0.5 to 1 hour is necessary to remove the paraffin that dissolves the interstitial space.

- Section Dewaxing and Hydration

Dried sections need to be dewaxed and hydrated before staining in water-soluble dyes. If the dewaxing is not clean, and the sections are not easily stained or unevenly colored, the xylene used for dewaxed should be replaced frequently. Therefore, before H&E staining, it needs to be dewaxed with xylene, and then de benzene by graded alcohol until rehydrated with distilled water.

- Dyeing

As mentioned in the above section, we can carry out different staining according to different needs. In this project, the slices we use are based on H&E staining. H&E staining is routine staining of histological and pathological specimens.

First, the nuclei were stained with hematoxylin, and the sections were stained with Harris hematoxylin for 3-8 minutes, washed with tap water, differentiated with 1% hydrochloric acid alcohol for a few seconds, rinsed with tap water, returned to blue with 0.6% ammonia water, and rinsed with running water again. Then it is stained in eosin for 1-3 minutes to stain the cytoplasm.

Sections need to be dehydrated, transparent and mounted after H&E staining.

- Section dehydration, clearing and coverslip

The stained sections cannot be observed under the microscope, and need to be dehydrated with a gradient of alcohol from low concentration to the high concentration and transparent to xylene. Put the slices into 95% alcohol I for 5 minutes, 95% alcohol II for 5 minutes, absolute ethanol I for 5 minutes, absolute ethanol II for 5 minutes, xylene I for 5 minutes, and xylene II for 5 minutes, in turn, to dehydrate and transparent.

Take out the section, wipe off excess xylene, air dry, drop a drop of neutral gum, and cover the section with a coverslip for sealing.

2.1.3 Microscope Features and Scanning

Observing the H&E staining sections under a microscope, we can find that the nuclei are stained bright blue by hematoxylin, the cartilage matrix and calcium salt granules are dark blue, and the mucus is gray-blue. The cytoplasm was stained with eosin in different shades of pink to peach, and the eosinophilic granules in the cytoplasm were bright red with strong light reflection. Collagen fibers are pale pink, elastic fibers are bright pink, red blood cells are orange-red, and proteinaceous fluid is pink.

We take the Leica SCN400F slide scanner in our project. The Leica SCN400F scanner is the ideal solution for digital pathology laboratories. With its specially designed optics, scanning method and Z-Stack function, the Leica SCN400F slide scanner captures clear images of histopathology samples.

The Leica SCN400F slide scanner captures multiple labeled fluorescent samples through a combination of tri-frequency cubes and excitation filters, allowing sharp multi-channel images. This comprehensive slide scanning system incorporates the latest technology to provide truly high-quality images. The detailed technical specifications and parameters are as follows :

- Slide loading : four slides are placed in one holder, and automatic scanning.
- Scanning speed : 20x magnification 15×15mm, 100 seconds; 40x magnification 15×15mm, 220 seconds.
- Glass slide size : standard size 26×76mm
- Scanning range (for high resolution) : standard size 24×62mm
- Objective lens : Optical system with Leica lens and optical elements designed for dedicated digital sensor scanning
- Available magnifications : 5x, 10x, 20x, 40x
- Spatial resolution : 0.25 μ m/pixel (40x magnification)
- Scanning method : high-speed, high-sensitivity linear CCD equipment, user-selectable (automatic, semi-automatic, manual scanning)
- Focusing method : dynamic focus/auto focus
- Layer scanning : user-selectable thickness, optional layers from 1 to 100 layers
- Barcode reader : one-dimensional (code 128), two-dimensional (data matrix)
- Storage : 5TB to 20TB
- Image export : .scn, .svs
- Dimensions (W×H×L) : 723×595×582mm

The scanning process consists of a digital microscope or a magnification system to scan the glass slices one by one under a low-magnification objective lens to acquire an image. A frame of reference is defined for localizing slide images using a slide-based coordinate system. It specifies a particular corner of the slide as a nominal reference origin and a right-handed (X, Y, Z) coordinate system for positioning from that origin. The microscope scanning platform automatically scans and moves according to the XY axis direction of the slice, and automatically focuses in the Z axis direction, as shown in Figure 2.4. Then, the scanning control software collects high-resolution digital images employing program-controlled scanning based on the effective amplification of the optical magnifying device, and the image compression and storage software automatically splice the images seamlessly to produce a digital slice with a full field of view. The slide-based (X, Y) coordinate system is rotated 180 degrees from the original image matrix (row, column) orientation of the image frames with the left label (see Figure 2.4) [Committee et al., 2010].

2.1.4 WSIs and Image Preprocessing

In 1999, a whole slide image (WSI) is invented, making the preservation and transmission of pathological sections more convenient and safe [Weinstein et al., 2009]. This digitization of pathological sections is considered to be an important turning point in the pathological development process, which effectively saves the manpower, time and cost of hospital pathology departments, improves the quality and efficiency of pathological diagnosis, and alleviates the development dilemma of pathology to a certain extent [Fine, 2014].

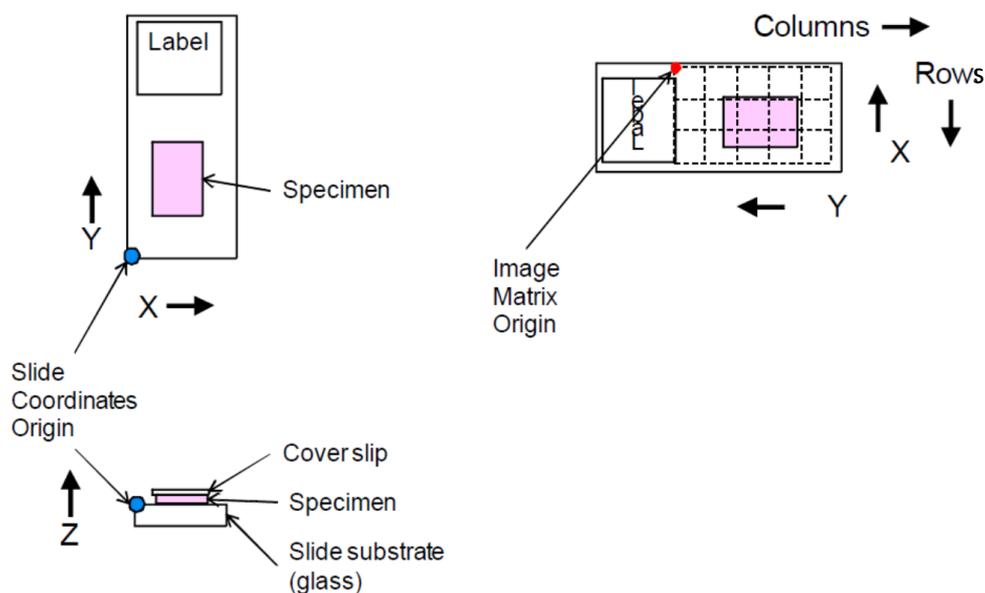


Figure 2.4 – Slide Coordinates Origin and (X, Y, Z) vs. Image Matrix Origin (Rows, Columns). (DICOM Standards Committee, 2010)

Digital Pathology (DP) uses the automatic microscope scanning system, combined with the virtual slicing software system, to scan and seamlessly splice traditional glass slides to generate a WSI, which is about the order of $10^6 \times 10^6$ pixels. DP has all the features of traditional slicing with the advantage of not being limited by space and time. The DP is a dynamic scale picture, it contains all the lesion information on the glass slide. On the computer, like under the microscope, it can be observed at different magnifications (4X, 10X, 20X, 40X, 100X, etc.), and within a specific range (1X ~ 100X), to achieve continuous zoom viewing slices. The average size of each slice is more than 1 Giga Byte (GB), even 3-5 GB, and each pixel is 160 nanometers. WSIs are usually stored in a multi-resolution pyramid structure [Bejnordi and van der Laak, 2016]. An image file contains multiple downsampled versions of the original image. Each image in the pyramid is stored as a series of image tiles to facilitate the fast loading of sub-regions of the image. The storage method of digital pathological slides is shown in Figure 2.5 [Wetteland et al., 2020]. It solves the problems that traditional glass slides that are easily damaged, easy to fade, easy to lose, and difficult to retrieve.

WSI can capture disease processes because they are developed at high resolution, each containing billions of pixels and up to one million or more microscopic objects with an important prognostic profile. Computational image analysis enables the mining of large-scale WSI datasets to extract quantitative morphological features that describe the visual quality of a patient's tissue.

WSI has several advantages :

- Glass slides are no longer unique.
- Access to cases is possible from any location.
- Digital image analysis can be applied.
- Archived WSI can be easily accessed.

From this point, several secondary advantages arise :

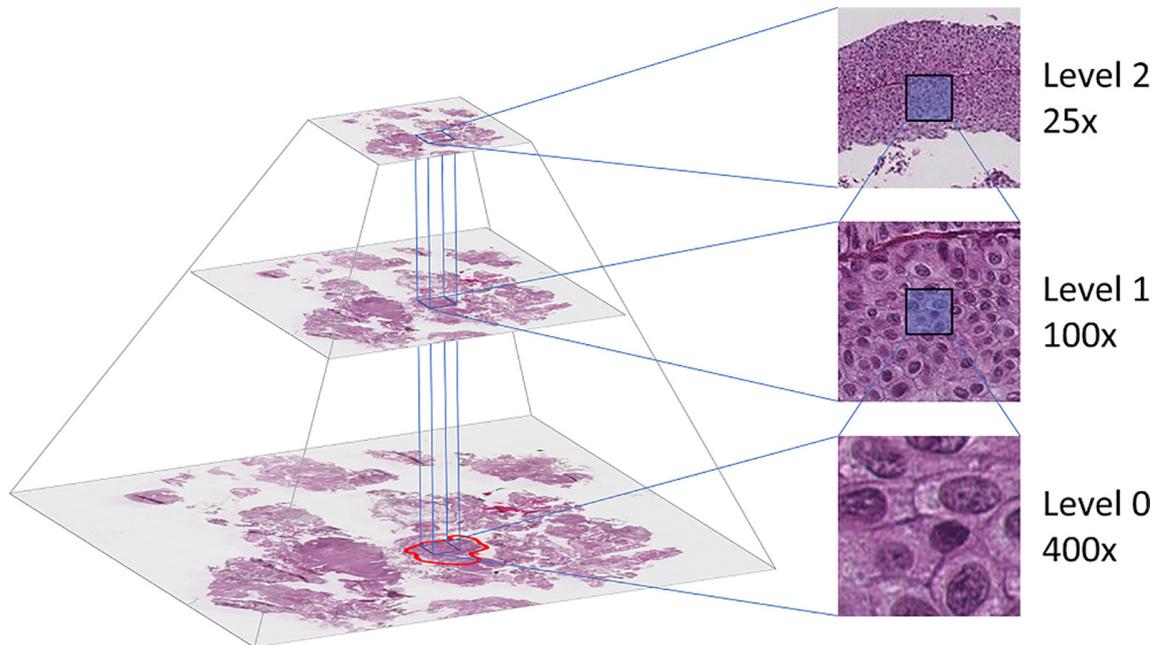


Figure 2.5 – The WSI is stored in a pyramidal file format, including several down-sampled versions of the base image.

The annotated region (marked with red at level 0) determines which tiles to extract. Tiles are then extracted at the desired location from all three levels. (Wetteland, 2020)

- The slide compilation of the case and the case assignment are fast and safe.
- Carrying the slices of cases to the pathologist for diagnosis is unnecessary and paperless work is possible.
- WSI can be used for a second opinion and be accessible in remote locations.
- WSI of referred cases are still accessible after returning the slides.
- Histological images can easily be provided on tumor boards ;
- The office desk is clean.
- A "home office" is possible [Grobholz, 2018].

At the same time, the WSI system can help segment the required regions and objects, determine cancer classification and prognosis, and quantify immunohistochemistry [Gandomkar et al., 2016].

The advancement of WSI quality has laid the foundation for the clinical application of digital images in anatomical pathology, thereby improving the possibility of computer-aided pathology diagnosis, and further optimizing the clinical interpretation of histopathology and medical care. With the popularity of WSI, the role of computer-aided technology in medical imaging becomes efficient for disease detection, diagnosis and prognosis.

From the image preprocessing perspective, it is not suitable to process so high-resolution huge images with conventional graphics software packages. OpenSlide is an open-source function library that provides interfaces to various programming languages such as C and Python [Goode et al., 2013]. It provides a unified set of application programming interfaces to read WSI scanned by manufacturers including Philips, Leica, and MIRAX. The automated slide analysis

platform (ASAP) is an open-source platform based on OpenSlide, Qt, and OpenCV for visualization, annotation, and automated analysis of histopathological images [Litjens, 2017].

We use Python programming languages in this project. Python is an ideal image processing language, there are many useful and convenient libraries we can use, such as OpenSlide to read WSI files and get patch images, Pillow is used for basic image processing in Python, NumPy is used for fast, concise, and powerful processing of images, Scikit-image is used for various image functions such as morphology, thresholding, and edge detection.

WSI preprocessing typically includes cutting patches, background removal, and staining normalization :

- patches cutting

The main method of cutting patches is patch sampling [Cruz-Roa et al., 2014], which treats the WSI as a grid of small-sized image patches that can be fed into an algorithm to extract features. By aggregating these features from small-sized blocks, the features on the whole WSI can be obtained, and then the classification results can be obtained using this information. [Xu et al., 2017] cropped the WSI into $151872\text{nm} \times 151872\text{nm}$ overlapping square blocks for the histopathological classification of brain tumors. These patches are sampled to 224×224 pixels and fed into pre-trained AlexNet [Krizhevsky et al., 2012a] for classification. During prediction, the 4096-dimensional vector output by the second fully connected (fc2) layer of the network is used as the block feature. The features of WSI are obtained by p-norm pooling and feature screening based on inter-class distance from all block features, and SVM is performed for glioblastoma multiforme (GBM) and low-grade glioma (LGG) discrimination. [Wang et al., 2016] randomly cropped breast cancer slices of the Camelyon dataset into 256×256 pixel patches, each patch has a benign and malignant label, indicating whether it is a lymph node metastasis. After the block is randomly cropped and horizontally flipped, the data are amplified and sent to GoogLeNet [Szegedy et al., 2015] for training to obtain a model for judging the malignant degree of the block.

- background removal

WSI usually contains a large number of white background areas, which brings meaningless calculation and storage overhead. Usually, when processing WSI, the histologists manually mark Region Of Interest (ROI) areas or use threshold-based Otsu binarization technology [Otsu, 1979] to remove these backgrounds. The experiments of [Wang et al., 2016] show that an average of 82% of the area of large full-section pathological images is background, so that, cropping can greatly improve the efficiency of diagnosis and analysis.

- staining normalization

The low repeatability of the manual production process and the difference in dyes and scanners cause color differences between WSIs, which leads to errors in subsequent analysis work. Dye normalization techniques are generally used to eliminate color differences between WSIs. [Reinhard et al., 2001] consider that the three channels of the RGB color space have an obvious correlation, and the adjustment of pixel color must be linked to the three channels simultaneously, which increases the complexity. While [Ruderman et al., 1998] proposed the $L\alpha\beta$ color space to reduce the correlation between channels. [Reinhard et al., 2001] proposed a color transfer method, which first transforms the image from RGB space to $L\alpha\beta$ space, then adjusts the image with the mean and standard deviation of the three channels of the $L\alpha\beta$ space, and finally transforms back to RGB space. Methods relying on statistical features imply the assumption that different sections

have the same staining composition, but the actual sections contain different types of tissue, and there is no guarantee that the staining composition will always be the same. The background also has color after color migration. In the field of digital pathology, researchers [Macenko et al., 2009, Vahadane et al., 2015] used the staining separation method to normalize the staining by mapping the image in the RGB space to the color space of the stain through a color deconvolution technique. Figure 2.2 displays the color normalized images corresponding to the original H&E stained images.

2.2 Histopathological Images Analysis

2.2.1 Features Extraction from Histopathological Images

Features extraction is the primary step for histopathological image analysis using machine learning. The typical steps for machine learning in digital pathological image analysis are shown in Figure 2.6 [Komura and Ishikawa, 2018].

Hand-crafted features perform an irreplaceable role in medical classification tasks due to low data requirements and ease of training. Li et al. [Li et al., 2016b] use support vector machines (HC-SVM) to obtain hand-crafted features and train a second-level window classifier on the posterior probabilities - as an output refinement - that can substantially improve the segmentation performance. The final performance of HC-SVM with refinement is comparable to CNN in colon histology images. And it also shows that combining and refining the posterior probability outputs of CNN and HC-SVM can further boost performance. Jeena et al. [Jeena et al., 2021] proposed hand-crafted texture features for stroke diagnosis. They trained the hand-crafted texture features by the custom CNN model and other five pre-trained models from ImageNet. The custom CNN model achieves the best performance with an accuracy of 95.8%. Our team's previous work [Zubiolo et al., 2016] uses hand-crafted features (number of end branches (NE) & junctions (NJ), the length of end branches (LE) & non-end branches (LJ) and their ratios) to represent the vascular network of RCC. In this project, we not only proposed a much bigger dataset, but also new hand-crafted features (skeleton and lattice features) of RCC and perform a series of classification tasks as opposed to the former study [Zubiolo et al., 2016].

Among feature extraction algorithms we can find :

- LBP (local binary patterns)

LBP is a simple and effective feature extraction algorithm for texture classification, proposed by [Ojala et al., 2002]. From the perspective of texture analysis, the texture feature of a pixel on an image refers to the relationship of this pixel with its neighbors.

- CLBP (completely local binary Pattern)

CLBP is the latest form of LBP. The traditional LBP operator loses some information when extracting texture features. Therefore, [Guo et al., 2010] propose a complete LBP operator, which consists of two parts, namely the LBP operator and the local part made of the central pixel.

- OTSU

The Otsu method [Otsu, 1979] was proposed by Nobuyuki Otsu. Among all the segmentation methods, the Otsu method is one of the most popular methods for image thresholding because of its simple calculation. Otsu is an automatic threshold selection region-based segmentation method.

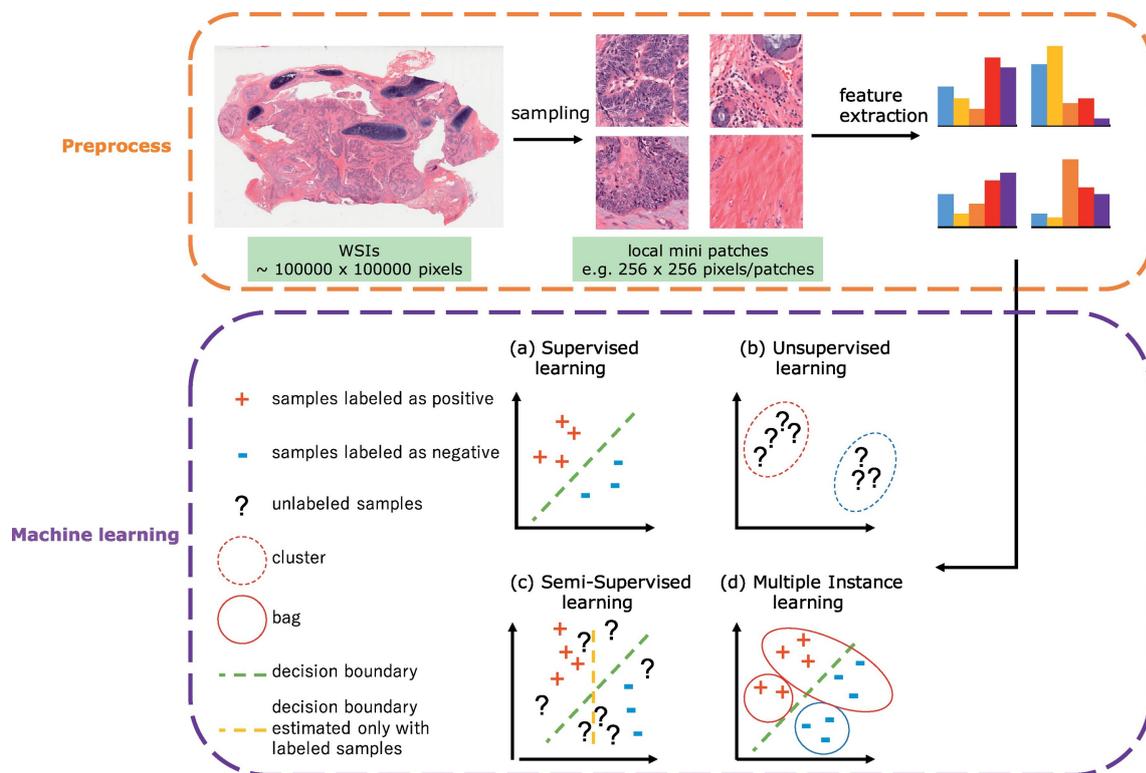


Figure 2.6 – Typical steps for machine learning in digital pathological image analysis. After preprocessing whole slide images, various types of machine learning algorithms could be applied including (a) supervised learning (b) unsupervised learning (c) semi-supervised learning, and (d) multiple instance learning. (Komura, 2018)

- LPQ (Local Phase Quantization)
LPQ is an effective texture analysis tool [Ojansivu and Heikkilä, 2008]. The method considers that the observed image is an image obtained after the original image has undergone a blurring process. To effectively reflect local information in the LPQ method, the Bayesian transformation [Digalakis and Neumeyer, 1996] of the partial image is usually calculated.
- GLCM (Gray-Level Co-occurrence Matrices)
The gray level co-occurrence matrix of an image can reflect the comprehensive information of the image grayscale concerning direction, adjacent interval, and variation amplitude. It is the basis for analyzing the local patterns of images and their arrangement rules [Saito et al., 2016].
- ORB (Oriented FAST and Rotated BRIEF)
The ORB feature was published by Ethan Rublee [Rublee et al., 2011]. Alleviating the high computational cost of the Scale Invariant Feature Transform (SIFT) feature and the Binary Robust Independent Elementary (BRIEF)-specific feature, the authors propose a binary feature (0,1 string) based on the Features From Accelerated Segment Test (FAST) and BRIEF features that are 100 times faster than SIFT, 10 times faster than Speeded Up Robust Features (SURF), with excellent matching effect as well.

2.2.2 Deep Learning Approach

2.2.2.1 Back Propagation Neural Network

A neural network usually consists of an input layer, hidden layer and output layer, as shown in Figure 2.7 [Bui, 2021]. Such a neural network model has a wider range of applications, covering both linearly and nonlinearly separable scenarios. It can be solved by the back-propagation algorithm. In addition, increasing the number of hidden layers can also improve the ability of data representation and function fitting. Adding more hidden layers based on three layers becomes a deep neural network. Once a hidden layer is added, the number of parameters of the model increases sharply, so deep learning requires particularly high computing resources, and the training time is also very long. Although it consumes computing resources, the advantages of deep learning are also very prominent. Compared with machine learning, the model automatically completes feature extraction and does not require manual feature engineering, which is particularly important for processing high-dimensional data. The common model structures of the neural networks are shown in Figure 2.8

2.2.2.2 Deep Learning Model Architecture

The deep learning method can learn the scale invariance and deformation-insensitive feature expression of the image from a large amount of training data according to automatic feature learning, and can more fully express the image features without being restricted by professional factors [Janowczyk and Madabhushi, 2016]. The deep learning algorithm can not only learn the characteristic expression of the data but also can be used as a classifier. It has stronger self-learning and self-adaptive ability and is more suitable for pathological analysis of large data volume.

In the field of image processing and computer vision, convolutional networks (CovNet) based on convolution, Pooling and ReLU functions are the most commonly used network structures for deep learning. Figure 2.9 is a Deep Convolutional Neural Network (DCNN) architecture [Izadyazdanabadi et al., 2018b].

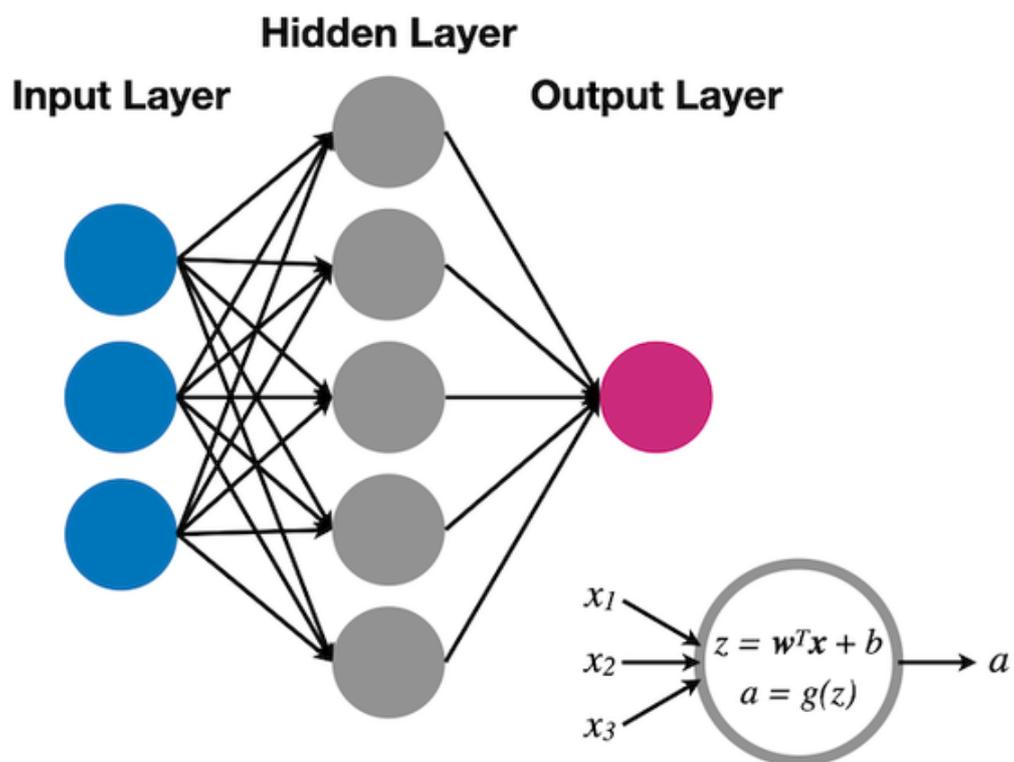


Figure 2.7 – The structure of a 3-layer neural network. (Christian, 2020)

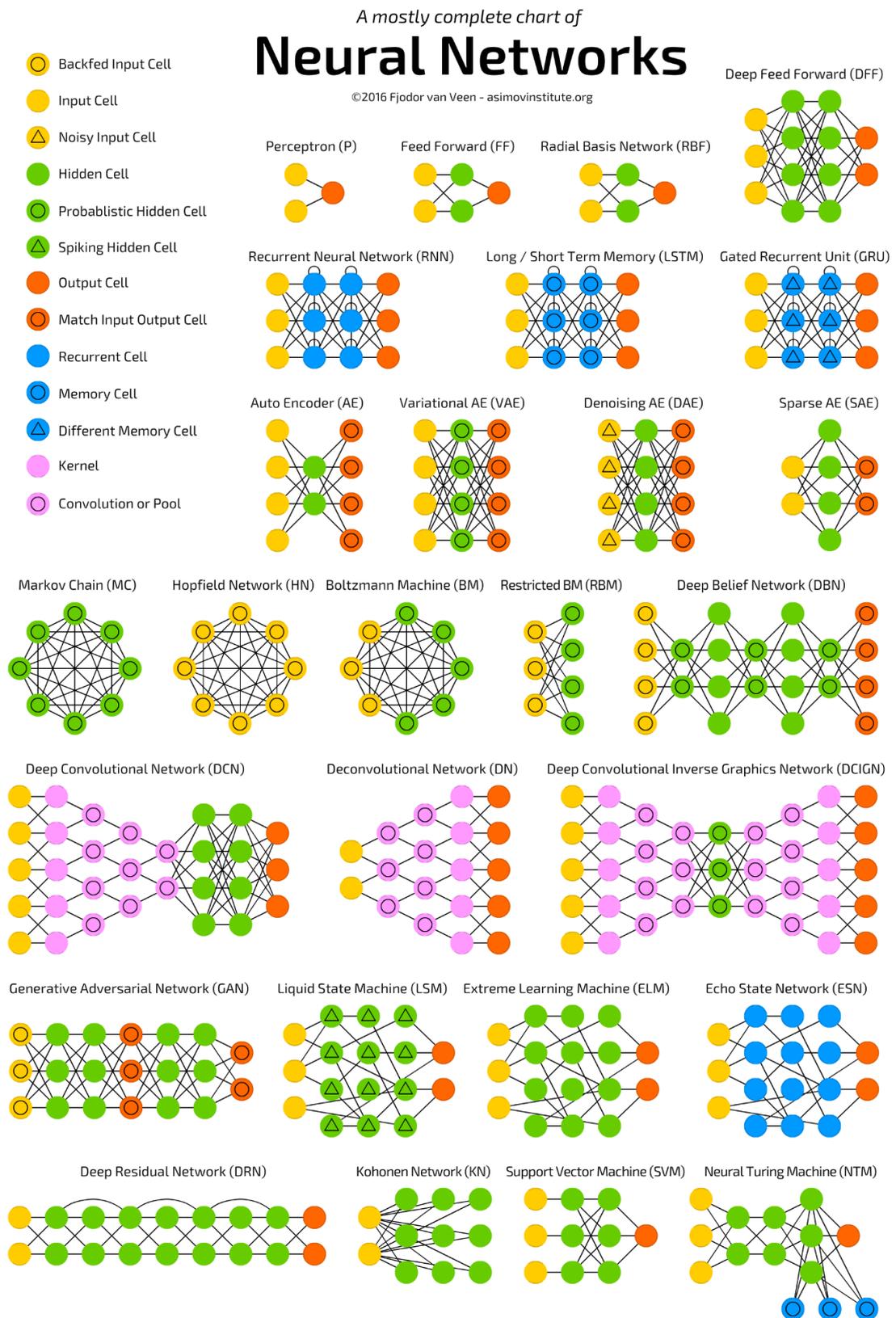


Figure 2.8 – General architectures of neural network. (source : Asimov Institute)

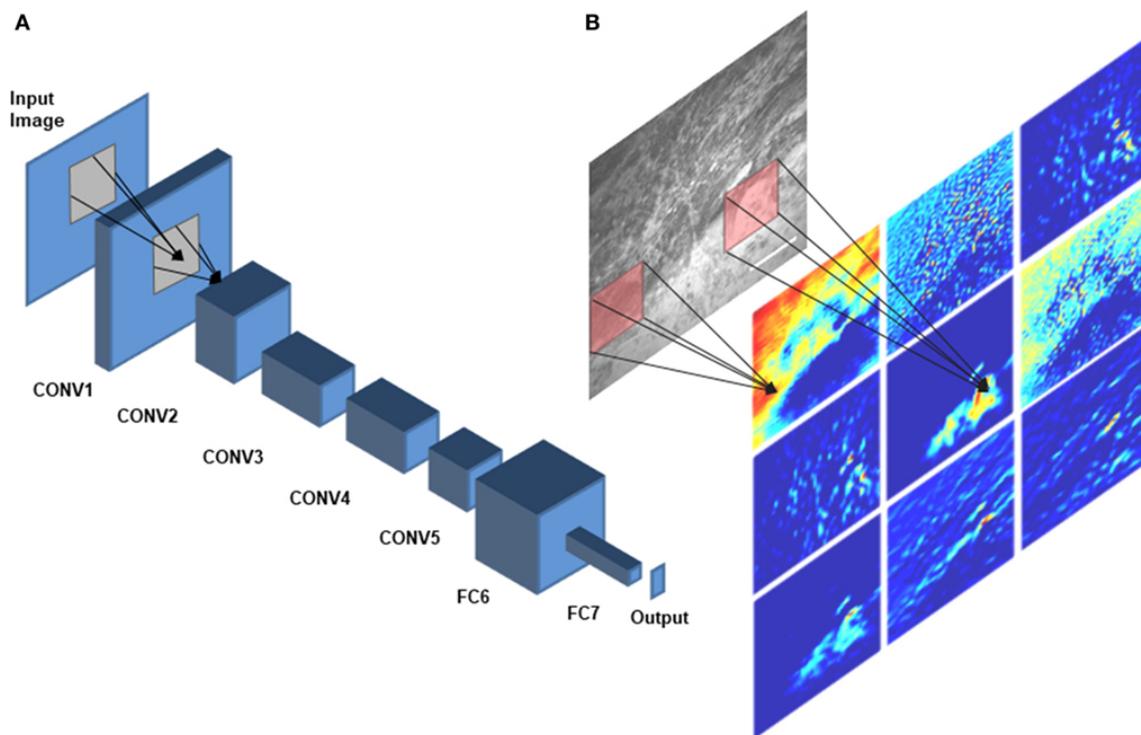


Figure 2.9 – Deep convolutional neural network (DCNN) architecture.

A schematic diagram of AlexNet [Krizhevsky et al., 2012b], a DCNN architecture that was trained on CLE images for diagnostic classification by Izadyyazanabadi et al. [Izadyyazanabadi et al., 2018a] is shown in panel (A). CONV1—CONV5 are the first five convolutional layers and FC6 and FC7 are the fully connected layers 6 and 7. Different feature maps of the first convolutional layer (color images) were calculated by convolving different filters (red squares) with the corresponding regions of the input image [illustrated in panel (B)]. (Izadyyazanabadi, 2018)

H&E stained sections, is appearance variability. To some extent, this can be explained by disease heterogeneity, but a large part is the result of the tissue preparation and staining process. In the histopathological image classification task, deep learning-based methods are currently the most widely used methods. Wang et al. [Wang et al., 2017] developed a bilinear CNN (BCNN) model to classify breast cancer histopathological images. The BCNN is a new CNN model for fine-grained classification. BCNN consists of two CNNs, whose convolutional-layer outputs are multiplied with an outer product at each spatial location. To evaluate the proposed BCNN-based method, the experimental results on eight classes of colorectal cancer histopathology image datasets show that the BCNN-based algorithm outperforms traditional CNNs. Wang et al. proposed a comprehensive recognition method with a newly proposed class structure-based deep convolutional neural network (CSDCNN) [Wang et al., 2022a], the CSDCNN employs an end-to-end training method, which can automatically learn semantic and discriminative hierarchical features from low-to-high levels. It provides an accurate and reliable solution for breast cancer multi-classification. Meanwhile, they also proposed LPMF2Net based on the deep manifold fusion of multilayer features [Wang et al., 2022b]. It can capture the complementarity between features at different levels, then cross-cascaded and fuse the multilayer features to improve the ability to characterize cell structure. When testing the BreakHis dataset with the proposed LPMF2Net model, it achieves the best and most consistent performance at $40\times$, $100\times$, $200\times$ and $400\times$ magnification levels, with accuracies of 94.91%, 96.12%, 95.51%, and 95.42%, respectively. Moreover, the performance of CNNs can be improved by deep high-order statistic models [Hou et al., 2022], which can outperform corresponding first-order counterparts in vision tasks. They integrated asymmetric convolution into the second-order network and proposed a novel second-order asymmetric convolution network (SoACNet). SoACNet uses a series of asymmetric convolutional blocks to replace each stand-square-kernel convolutional layer of the backbone architecture, and then computes second-order statistics of deep features by global covariance pooling. The classification result of SoACNet on the BreakHis dataset shows its competitive performance with the state-of-the-art.

However, CNNs do not explicitly extract the complex features of the spatial arrangement of cells from histopathological images. Graph Neural Networks (GCNs) can address this problem. [Gadiya et al., 2019] use GCNs by modeling a tissue section as a multi-attributed spatial graph of its constituent cells to classify cancers. Each cell appearance is captured as a multi-attributed high-dimensional vertex feature and the spatial relations between neighboring cells are captured as edge features based on their distances in a graph. The result demonstrated that using GCNs on histograms gives a competitive performance for detecting malignancy and invasiveness in BACH dataset [Aresta et al., 2019a] as compared to conventional patch-based CNN.

Almost all human solid malignant tumors require angiogenesis to sustain their aggressive growth. One way to observe angiogenesis is through histological images. Microvessel density (MVD) and vascularity have been used to describe and compare vascularization in different tumors or tumor subtypes. For example, breast cancer and sarcoma differ in their vascularization patterns [Tomlinson et al., 1999]. Likewise, renal cell carcinoma and breast cancer differ in terms of MVD [Eberhard et al., 2000]. Different vascular patterns in different tumors or tumor subtypes may lead to biologically and clinically relevant consequences. However, there is currently a lack of machine learning models for classifying and subtyping H&E staining histopathological images of cancers based on vascular morphology. In this project, we classify topological information of RCC vascular networks by GCNs with two different pooling methods : SAGPoolg [Zhang et al., 2018a] and SAGPoolh. [Cangea et al., 2018]. SAGPoolg is a novel Sort-Pooling layer that sorts graph vertices in a consistent order so that traditional neural networks

can be trained on the graphs. The experiments on some benchmark graph classification datasets (COLLAB [Gallicchio and Micheli, 2020a], D&D, IMDb-B [Cai and Wang, 2018], IMDb-M [Zhao and Wang, 2019], MUTAG [Zhang et al., 2021], NCI1 [Gallicchio and Micheli, 2020b]) demonstrate that the proposed SAGPoolg achieves highly competitive performance compared to graph kernels and other graph neural network methods. SAGPoolh combines several advances in graph neural network design to demonstrate that competitive hierarchical graph classification results are possible without sacrificing sparsity. The results on several established graph classification benchmarks show its state-of-the-art and can be scalable to large graphs.

Although the deep learning methods have achieved great progress for medical imaging tasks, traditional machine learning algorithms are still crucial due to interpretability and small datasets applicability. Saxena et al. [Saxena et al., 2021] proposed a hybrid ML model to solve the class imbalance problem. The proposed model employs a pre-trained ResNet50 and the kernelized weighted extreme learning machine to classify breast cancer histopathology images. The kernelized weighted extreme model achieved a reasonable performance for the classification of the minority as well as the majority class instances on the BreakHis and BisQue histopathological images datasets. It also can outperform the same training-testing folds of the BreakHis dataset. Moreover, a novel CoMHisP framework based on a fuzzy support vector machine with within-class density information (FSVM-WD) for histopathological image classification was proposed in [Kumar et al., 2020]. It defines a novel feature extraction technique by optimizing the block size to extract image micropatterns and compute the center of mass (CoM) for each pixel to extract feature vectors. The performances of the proposed model are evaluated on the CMTHis dataset comprising histopathological images of canine mammary tumors (CMT). The results demonstrated that the CoMHisP framework can be influenced by stain normalization and magnification. It works best with staining normalization and lower magnification, achieving a classification accuracy of 97.25% ($\pm 1.80\%$), outperforming traditional ML and deep FE-VGGNET16-based feature descriptors.

2.2.2.6 Histopathological Images Segmentation

The development of new blood vessels within a tumor is due to the increased demand for blood by cancerous cells [Ohgaki and Kleihues, 2005, Rodriguez et al., 2012]. Vasculature from histological images plays a key role in cancer development, subtyping and radiotherapy assessment [Loukas and Linney, 2004b]. Vascular segmentation is crucial in medical imaging tasks, for example, to obtain a complete liver vascular network from patient abdominal computed tomography (CT) images based on graph cut, thinning, and vascular combination [Guo et al., 2020]. Coarse-to-fine vessel segmentation network for OCTA images (OCTA-Net) also shows its ability to detect thick and thin vessels separately [Ma et al., 2020].

However, the current automatic vascular segmentation for histopathological images is limited to Immunohistochemistry (IHC) stained histology images. In [Bukanya et al., 2020], the authors first segment and quantify blood vessels from histopathological images of the brain in case of Alzheimer disease (AD). They use hematoxylin and diaminobenzidine (H&DAB) stained histology images. [Kather et al., 2015] obtained vascular hotspot probability maps of WSI by scanning whole CD34 immunostained histological images of colon cancer samples and then segmenting blood vessels. The comparison between the H&DAB staining image and the H&E staining image is shown in Figure 2.11. H&DAB staining for special coloration of blood vessels, the background is clean and easy to segment, but the background of the H&E image is more complex and has

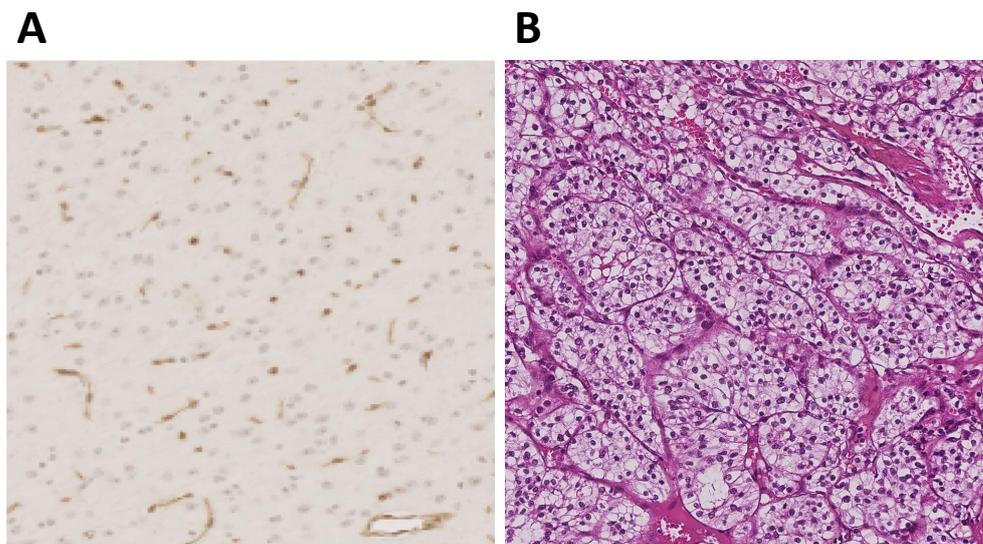


Figure 2.11 – Example of H&DAB staining and H&E staining histology images.
A : H&DAB staining, B : H&E staining.

some similar linear structures, such as cell membranes and fibers, etc., which makes the task of vascular segmentation from H&E images more challenging. Since H&E stained images contain more information, such as cell morphology, inflammatory cell infiltration, lymph node metastasis, etc., it is essential for pathological diagnosis. If we can obtain sufficient vascular information from the H&E images alone, additional operations such as IHC could be avoided, which can save time and resource costs, and make the diagnosis more accurate.

As far as we know, there is no vascular network segmentation research in H&E staining histopathological images. Our classification part of the project shows that the vascular network architecture can be used alone to classify RCC histopathological images [Xiao et al., 2021], but its application is limited due to its reliance on manual vascular masks. In this project, we propose an MTL-SSL model which can segment vascular networks automatically while predicting the tumor class.

SSL [Chapelle et al., 2009] plays a key role in segmentation tasks since it allows to reduce the reliance on large annotated datasets. It can provide an effective way of leveraging unlabeled data to improve model performance. Several approaches have been proposed for SSL, such as Deep Adversarial Networks [Zhang et al., 2017], Cross Pseudo Supervision [Chen et al., 2021] and Cross Consistency Training [Ouali et al., 2020]. In this project, we use the Mean Teacher approach [Tarvainen and Valpola, 2017] for SSL. This model was developed from the Temporal Ensembling Model (TEM). It uses the consistency regularization method to effectively reduce over-adaptation in SSL. However, only a few studies have investigated if SSL can be applied to achieve satisfactory results in H&E staining histopathological images, such as NAS-SGAN [Das et al., 2021] for atypia scoring of breast cancer; it consists of the discriminator and generator models that are trained in an adversarial manner using both labeled and unlabeled samples, the discriminator model is an unsupervised model and the generator model is trained over a stable fea-

ture matching objective function following a composite Generative Adversarial Networks (GAN) architecture. The results revealed that NAS-SGAN can better discriminate different cancer grades, and achieve good robustness and accuracy of the system along with limited labeled data. OSE-SSL is proposed in [Sparks and Madabhushi, 2016] for content-based image retrieval of prostate cancer. OSE-SSL incorporates semantic information, and partial class labels, into a machine learning (ML) scheme such that the low dimensional representation co-localizes semantically similar images. The yielded area under the precision-recall curve (AUPRC) of CBIR with OSE-SSL for prostate histology is 0.53 ± 0.03 . It outperforms a CBIR with Principal Component Analysis (PCA) to learn a low dimensional space yielded an AUPRC of 0.44 ± 0.01 . Moreover, breast cancer classification on the BreakHis dataset with Self-Paced Learning together [Asare et al., 2020] demonstrates the effectiveness of SSL, which integrates self-training and self-paced learning to generate and select pseudo-labeled samples. This model can retrain the model with the generated pseudo labels on unlabeled-target data, and also can prevent the model from ignoring samples from less represented classes (hard-to-learn samples) together with a class balancing framework, hence effectively handling the issue of data imbalance. In this project, we apply SSL to RCC histopathological images to provide the benchmarks for vascular network segmentation and propose to improve the segmentation accuracy by jointly solving the tumor classification task.

MTL [Caruana, 1997] aims at improving the performance of multiple related learning tasks by leveraging comprehensive information among them. MTL achieves better generalization properties than single-task learning. Classification and segmentation are both key tasks in medical image processing. Joint segmentation and classification of tumors in 3D automated breast ultrasound images show that learning these two tasks simultaneously improves the outcomes of both of them [Zhou et al., 2021]. The proposed framework consists of two sub-networks : an encoder-decoder network for segmentation and a lightweight multi-scale network for classification. The result of the 3D ABUS dataset indicates that segmentation and classification task can promote each other during the training process, and the performance improved compared to the single-task learning counterparts. Other joint tasks using MTL in histopathological imaging include cell detection and segmentation on PMS2-stained colon rectal cancer and tonsil tissue images [Chamanzar and Nie, 2020]. This model is an end-to-end deep learning algorithm to perform both single cell detection and segmentation using only point labels and shows significant improvement in cell detection and segmentation without increasing the annotation efforts. Mitotic cell detection and counting are one of the strongest prognostic markers for breast cancer diagnosis, MitosisNet [Alom et al., 2020] for mitosis detection from pathological images consists of segmentation, detection, and classification models. It fast enough to meet the requirements of clinical practice. MitosisNet is an end-to-end multi-task learning system, where the segmentation and detection models are used for mitosis reference region detection and the classification model is applied for further confirmation of the mitosis regions. The results evaluated on "MITOS 2012" and "MITOS 2014" [Ludovic et al., 2013], and Case Western Reserve University (CWRU) datasets show their reach state-of-the-art performance compared to the existing methods. The problem of poor interpretability of models becomes critical in medical applications : models that doctors do not understand and trust are less likely to be used in everyday clinical practice. MTL of a deep k-nearest neighbor network for histopathological image classification and retrieval [Peng et al., 2019] was proposed due to that consideration. It can leverage the classic concept of k-nearest neighbors to improve model interpretability. The original framework is built on top of any existing classification network and with an advanced loss function strategy. The results evaluated on the colorectal cancer histology slides show that confidence estimates are strongly correlated with model performance

and the nearest neighbors are intuitive and useful for expert evaluation. In this project, we combine the classification task for which labels are easy to obtain and the vascular network segmentation task for which images have complex backgrounds and manual delineation is cumbersome. Our MTL aims to improve the performance of the segmentation task on RCC histopathological images compared to the supervised learning and single segmentation tasks.

2.3 Computer-Aided Diagnosis and Prognosis (CAD & CAP)

2.3.1 CAD

CAD has proven to be an effective tool for assisting histopathological image processing analysis and can be used in traditional pathology. Especially for cancer patients, decision-making needs to be as accurate as possible to increase the likelihood of an optimal treatment plan. [Kostopoulos et al., 2017] proposed a new image acquisition library (HICL-histology image acquisition library), including 3831 histological images of three different diseases, freely open to the scientific community.

Color is the most important feature in quantitative immunohistochemistry (IHC) image analysis. IHC can be used to provide information related to the cause and to identify malignancies. [Shu et al., 2016] developed a color detection statistical model suitable for detecting staining in digital IHC images.

[Mousavi et al., 2015] extract discriminative information from the histopathology image slides for diagnosis and glioma tumor grading. [Momeni-Boroujeni et al., 2017] designed a computer model to help Fine-needle aspiration (FNA) biopsy to diagnose pancreatic cancer. The results of the study indicate that computer models can be successfully used to distinguish between benign and malignant pancreatic cytology. The classification accuracy rate is 77%.

To assess the potential impact of digital aids on the interpretation of digital slides, [Steiner et al., 2018] used a deep learning algorithm to detect breast cancer metastasis in lymph nodes. Six pathologists examined 70 digital slides with or without a deep learning algorithm-assisted reading mode. Deep learning algorithm assistance significantly increased the sensitivity of detection for micrometastases (91% vs 83%, $P = 0.02$). The image review of micrometastases is easier ($P = 0.0005$). This study demonstrates that deep learning algorithms improve the accuracy and efficiency of diagnosis in digital pathology workflows. Artificial intelligence (AI) is expected to fundamentally change the way we detect and treat breast cancer soon [Robertson et al., 2018, Feng et al., 2018]. [Fondón et al., 2018] also proposed a CAD tool for automatically assessing the degree of malignancy in breast tissue samples, this method is based on the calculation of three sets of features related to nuclei, color regions, and textures considering local characteristics and global image properties, accuracy levels range from 75.8% when the 5-fold cross-validation was performed to 75% with the external set of new images and 61.11% when the extremely difficult images were added to the classification experiment. In addition, [Maeda et al., 2017] developed a CADx for identifying breast epithelial proliferative lesions in invasive breast cancer. This system can be used for the pathological diagnosis of breast preparative core needle biopsy (CNB) in routine research. [Veta et al., 2014] outline the method of performing cost-effective and high-throughput histopathology slide digitization using WSI.

[Fatima et al., 2017] also developed a computer-aided system for eye diseases that aims to automatically detect papilledema from fundus images.

[Nayak et al., 2013] evaluates a method based on a variant of Restricted Boltzmann Machine (RBM) that learns the intrinsic features of image signatures in an unsupervised way. Patches from a curated library of images are then classified using computational code from the learned representation. The system has been evaluated on a small 1k-by-1k patch dataset of glioblastoma multiforme (GBM) and clear cell renal carcinoma (KIRC) extracted from The Cancer Genome Atlas (TCGA) open-access database. Evaluations on GBM and KIRC on 1400 and 2500 samples show the performance of 84% and 81%, respectively.

By combining the latest image analysis technology and medical image information, CADs provide a new non-invasive detection and analysis method for clinical diagnosis, surgical plan selection, and overall efficacy evaluation. With the rapid development of medical imaging, not only the image quality and speed have been improved but also the diagnostic and therapeutic technology has undergone revolutionary changes.

2.3.2 CAP

Fourier transform infrared (FTIR) spectroscopic imaging can be based on biochemical correlation with normal and diseased states and modern computer-aided mode to identify cancer or other physiologic conditions. It can also integrate infrared spectral data with patient information, predicting cancer risk, providing potential pathways for precision medicine and personalized care in cancer treatment [Tiwari and Bhargava, 2015].

The tumor-matrix ratio (TSR) reflected on the histological images of H&E staining is a potential prognostic factor for survival, and the use of automated image processing techniques to allow high throughput and precise differentiation of tumor epithelium and stroma can enhance the prognostic significance of TSR. The use of deep convolutional neural networks (CNN) to learn natural image features can alleviate the large sample size requirements for dealing with biomedical classification problems. [Du et al., 2018] used GoogLeNet [Szegedy et al., 2015] to achieve 90.2 accuracy (*vs* 91.1 for fine-tuning). This indicates the feasibility of using it to assist the pathological-based binary classification problem.

Stage II colon cancer (CC) patients lack a standardized analytic technique. [Eriksen et al., 2017] compared manual stereological estimates of tumor-infiltrating lymphocytes (TILs) with automatic counts obtained by image analysis, and at the same time investigated the heterogeneity of TILs. The Spearman's correlation coefficients for density estimates ranged from 0.9457 to 0.9638 ($p < 0.0001$), while those for area fraction estimates ranged from 0.9400 to 0.9603 ($p < 0.0001$). Regarding heterogeneity, intra-class correlation coefficients (ICC) for CD3+ TILs varied from 0.615 to 0.746 in the central area, and from 0.686 to 0.746 in the invasive area. ICC for CD8+ TILs varied from 0.724 to 0.775 in the central area, and from 0.746 to 0.765 in the invasive area.

2.3.3 Limitations and Improvements of Computer-Aided Technology

[Xiao et al., 2017] argued that due to the scarcity of publicly available repositories, it is difficult to validate advances in medical image processing techniques with real clinical data. Therefore data modeling is needed to establish data that can provide high-quality multimodal clinical data to validate and compare image registration algorithms, which may be beneficial for the accuracy and efficiency of brain tumor resection. In addition, the database can be used to test other image processing methods and neuro-navigation software platforms.

One of the main obstacles to the quantitative analysis of H&E images is the high variability observed between different samples and different laboratories. [Tam et al., 2016] used a decision tree-based support vector machine learning algorithm to classify H&E staining slide images of 44 resected breast tumors based on the existing histological structure. For example, a quantitative description of H&E variability for each structure is obtained by comparing the structural centroids of different images, researchers also use intensity centering and histogram equalization (ICHE) to reduce batch effects.

Multicolor immunofluorescence staining, multiplexed immunohistochemistry (mIHC) or cyclic immunofluorescence (cycIF) combined with standard histological staining such as H&E can provide an in-depth assessment of in situ cell complexity, helping to reveal relationships and spatial interactions in complex molecular disease states. However, these multiple imaging methods are expensive and can degrade tissue quality and antigenicity. [Burlingame et al., 2018] developed a new method using the conditional generation confrontation network (cGAN), speedy histopathological-to-immunofluorescent translation (SHIFT) of whole slide images (WSIs), with an accuracy of up to 94.5%. This approach not only enhances our understanding of the mapping of histology and morphological maps to protein expression profiles but also greatly improves the efficiency of diagnostic and prognostic decisions.

2.3.4 The Prospect of Computer-Aided Technology in the Diagnosis of Renal Cancer

Currently, computer-aided technology applied to a variety of cancers has made breakthroughs in the experimental stage. Although most of them are still far from clinical, they provide a scientific basis for informational pathology diagnosis in the future. On this basis, the diagnosis, grading, and prognosis of renal cancer can fully draw on these emerging computer-aided technologies to provide accurate diagnosis, personalized medical advice and nursing measures for each patient to optimize cancer treatment.

The review article [Suarez-Ibarrola et al., 2020] pointed out that the main area of research involving Computer-Aided Diagnosis in RCC concerns the differentiation between benign and malignant small renal masses, Fuhrman nuclear grade prediction, and gene expression-based molecular signatures. Computer-Aided Detection System (CADE) and Computer-Aided Diagnosis System (CADx) works on CT images can help pathological detection or identification of lesions and further diagnosis of renal cancer in time [Kaur et al., 2019]. [Coy et al., 2017] proposed a quantitative computer-aided diagnostic (CAD) algorithm on four-phase multidetector computed tomography (MDCT) to distinguish ccRCC from chRCC, pRCC, oncocytoma, and fat-poor angiomyolipoma (fp-AML). [Sim et al., 1999] used computer-aided tissue echo quantification to differentiate small hyperechoic renal cell carcinoma from angiomyolipoma by studying ultrasonographic images. [Raza et al., 2010] proposed the bag-of-features method, which according to scale-invariant features automatically classifies renal cell carcinoma subtypes. [Lee et al., 2017] developed a quantitative feature classification method to differentiate benign fat-poor fp-AML from ccRCC based on histogram and texture patterns from contrast-enhanced multidetector computer tomography (CE MDCT) images.

The application of Computer-Aided Techniques in histopathological images of renal cell carcinoma has great value and space for improvement. [Waheed et al., 2007] used a simple Bayesian classifier based on 8 extracted features that work on pathological imaging data to diagnose renal cell carcinoma as well as other types of cancer. Users can apply their expertise to the valida-

tion of feature extraction and quantification with this system, and they can also select from a list of features they think are most important and best suited for the classification at hand. The overall accuracy of the CNN model proposed by [Fenstermaker et al., 2020] can reach 99.1% in distinguishing normal and RCC (sensitivity 100%, specificity 97.1%), and the accuracy in distinguishing ccRCC, pRCC and chRCC cell type is 97.5%, predicting Fuhrman grade is as high as 98.4%. [Gao et al., 2021] proposed a composite high-resolution network named W-Net for nuclear grading of ccRCC in histopathology images. The W-Net can separate the clustered nuclei to automatically identify the rank of each core in ccRCC.

2.3.5 The Future of Computer-Aided Technology

Precision Medicine is a health care model designed to integrate large amounts of patients information to identify and classify disease progression and provide tailored treatment solutions for individual patients. With the increasing use of innovative IT technologies such as big data and cloud computing in hospital clinical business, the increasing computing power has greatly improved the data analysis capabilities and diagnostic levels in the medical field. Combined with algorithm optimization and machine learning, computer-aided technology will greatly enhance the speed and accuracy of medical diagnosis, and vigorously promote innovation and progress in the field of precision medicine. Artificial Intelligence & Precision Medicine is the development trend in the future!

From the doctor's experience and microscopic observation, we know that specific histological features and structures of different subtypes of RCC, such as red blood cells, blood vessels, lipid structures, papillary bodies, etc., may be involved. However, at present, the work of computer-aided techniques for automatic RCC classification of histopathological images from blood vessels is still in the blank stage. Our research has created a precedent and has important reference significance for future research.

2.4 Conclusion

In this chapter, we first discussed the history and development of histopathology. Pathological diagnosis remains the current gold standard for cancer diagnosis. However, manual diagnosis of histopathology images is a great challenge. There are several main reasons: (1) the professional background and rich experience of pathologists are difficult to inherit, the lack of skilled pathologists in small hospitals and clinics, (2) tedious work, expensive and time-consuming and poorly repeatable, and (3) pathologist fatigue may lead to misdiagnosis. Therefore, the use of computer-aided cancer diagnosis (CAD) is very urgent and important, which can reduce the heavy workload of pathologists and help to avoid misdiagnosis. Histopathological image processing is the most important part of CAD. Next, we mainly focus on the acquisition and processing of WSIs in this chapter and enumerate some major image analysis methods, including traditional machine learning methods and deep learning methods.

We elaborated on some feature extraction methods and demonstrated their interpretability. The interpretability of the model has always been a concern, so feature extraction based on histopathological images combined with interpretable traditional machine learning methods is still an important research direction. The hand-crafted feature is an artificially designed feature. According to the characteristics of human vision, what kind of features are sensitive to and what kind

of features are insensitive to extract the distinguishing features of the image? Therefore, each dimension of the extracted features often has specific characteristics. physical meaning. At present, advanced hand-crafted features include SIFT features based on directional gradients reflecting texture features, SURF features (SIFT improvement), HOG features, etc., as well as shape context reflecting contour shapes, etc. The hand-crafted feature also has some disadvantages : although it is based on the optic nerve theory, it is inevitably inappropriate because it is designed by humans. Moreover, this method depends on the database and needs to be designed according to the characteristics of the data, that is to say, the designed features are not applicable to all databases. When the data source changes, for example, the features designed for RGB data are replaced by Kinect depth images, these feature points will not be used. It must have adapted, so it had to be redesigned.

Then we also show the power of deep learning methods in WSIs processing tasks. Deep Convolutional Neural Networks (DCNNs) are one of the most powerful and successful deep learning methods particularly when the data are images. DCNNs have already delivered excellent performance in different medical image processing tasks including classification, segmentation, and detection. Tasks such as classification and segmentation from histopathological images play a key role in computer-aided cancer diagnosis or prognosis. However, deep learning also has certain limitations, such as the well-known poor interpretability, and we cannot judge which features play a role in deep learning. Fully supervised deep learning also typically requires a large amount of annotated data, which increases the workload. Semi-supervised learning may have better application prospects when annotation data is scarce.

At last, we discussed the content, application and development of CAD and CAP. Meanwhile, we also give talked about the limitations and how to improve their application well, especially in renal cell carcinoma research, which is related to our project. We also investigated their future opportunities and prospects.

Histopathological Image Datasets Building

Dans ce chapitre, nous illustrons le contexte médical de l'épidémiologie des sous-types de RCC et les différentes caractéristiques histopathologiques qui y sont associées. Puis nous présentons la méthodologie employée pour construire les bases de données BigRCC et VRCC et décrivons ces bases de données.

In this chapter, we first illustrate the medical background of epidemiology and the different histopathology characteristics of RCC subtypes, then present the methodology employed to build the BigRCC and VRCC databases and describe these databases.

3.1	Epidemiology and Main Subtypes of RCC	41
3.1.1	The Phenotype and Histomorphology of RCC Subtypes	42
3.1.2	Cellular and Vascular Characteristics in RCC Subtypes	44
3.2	Histopathological Images Dataset	46
3.2.1	BigRCC Dataset Building	46
3.2.1.1	Data Collection	46
3.2.1.2	Categories Annotation	46
3.2.1.3	Patch Images	49
3.2.2	TCGA Open-access Dataset	51
3.2.2.1	Downloaded Data	51
3.2.2.2	Sample Naming	52
3.2.2.3	Image Patches of TCGA	52
3.2.3	VRCC (Vascular RCC) Dataset Building	53
3.2.3.1	Data Source of VRCC	53
3.2.3.2	Weakly Vascular Network Annotation and Binary Mask	53
3.3	Conclusion	55

3.1 Epidemiology and Main Subtypes of RCC

Worldwide, kidney cancer is the thirteenth most common cancer [Ferlay et al., 2010]. Renal cell carcinoma (RCC) originates in the renal cortex, it accounts for 80% to 85% of malignant kidney cancer and 2% of all cancers [Kosary and McLaughlin, 1993]. RCC has the highest mortality rate among cancers of the genitourinary system. The incidence of RCC is on the rise; they are nearly 210000 new cases annually worldwide and approximately 102000 deaths. Men are almost twice as likely to have RCC malignancies compared to women. The main incidence age at diagnosis is between 55 and 75 years old. North America, Scandinavian countries, Australia, and several Western, Eastern, and Northern European countries have the highest incidence of renal malignancies. Africa and Asia have the lowest incidence [Jemal et al., 2005, Parkin et al., 2005, Franklin et al., 1996]. In 27 European Union (EU) countries, the estimated age-standardized incidence rate (the summary rate that would have been observed) of kidney cancer per 100000 Europeans was 15.8 in men and 7.1 in women. The Czech Republic, Lithuania, Latvia, Estonia and Iceland had the highest RCC incidence rates in Europe, while Romania, Cyprus and Portugal had the lowest rates [Levi et al., 2008]. The cause of RCC is not clear, but there is a certain genetic predisposition. Risk factors include smoking, chemical exposure, and diseases such as high blood pressure and obesity [Ross et al., 1989, Yu et al., 1986].

RCC has a certain relationship with age, and the incidence in Europe and the United States continues to increase with age, reaching a plateau around 70-75 years of age [Chow and Devesa, 2008]. The relationship between RCC and gender may be related to smoking rates, etc., and men have a higher risk [Ferlay et al., 2010]. Smoking is an established risk factor for RCC. A meta-analysis of nineteen case-control studies and five cohort studies [Hunt et al., 2005] confirmed that smoking or ever smoking was associated with a higher risk of RCC than non-smoking. Although the association between smoking and RCC was shown to be relatively weak, there was a clear dose-response relationship, with higher risk estimates associated with heavier smoking [Hunt et al., 2005]. Being overweight has also been identified as a risk factor for RCC. A meta-analysis of prospective studies [Renehan et al., 2008] found an association between body mass index (BMI) and the risk of RCC, with an estimated risk (per 5 kg/m² increase in BMI) of 1.24 for men and 1.34 for women. There are also studies investigating the distribution of body fat, demonstrating that the risk of RCC increases with an increasing waist-to-hip ratio [Adams et al., 2008, Pischon et al., 2006]. Other drug uses or diseases are also risk factors for RCC. Several large prospective cohort studies have demonstrated that hypertension or its treatment is associated with the risk of RCC [Chow et al., 2000, Weikert et al., 2008], and they found that the risk of RCC was associated with increased blood pressure with a clear dose-response relationship, and anti-stress drugs may have no causal risk for the development of RCC. In patients with acquired renal cystic disease (ARCD), the incidence of RCC is three to six times higher than in the general population [Port et al., 1989], which is a clear risk factor for RCC, and studies have also suggested that dialysis may be associated with the development of RCC [Nouh et al., 2010, Denton et al., 2002]. A history of urinary tract infection (UTI) may also be a risk for the development of RCC, and a population-based case-control study from the United States reported that a history of UTI was positively associated with RCC [Parker et al., 2004].

RCC is the most frequently occurring solid lesion in the kidney and includes different RCC types with specific histopathological and genetic features [Kovacs et al., 1997]. The 2002 TNM staging classification system is generally recommended for the utilization of clinical and scientific applications [Sobin, 2009]. The prognostic factors can be divided into anatomi-

cal, histological, clinical, and molecular. Anatomical factors include tumor size, venous invasion, renal capsule invasion, adrenal gland involvement, lymph node, and distant metastases. These factors are usually included in the commonly used 2002 TNM staging classification system. Histological factors affecting prognosis include Fuhrman grade, histological subtype, presence of sarcoid-like features, microscopic venous invasion, tumor necrosis, and invasion of the collecting system. The Fuhrman nuclear grade is the most widely accepted histological grading system in RCC [Ljungberg et al., 2007]. The TNM classification system, Furman classification system, and RCC subtype classification are recommended for clinical examination because of their impact on prognosis and treatment. According to the World Health Organization (WHO) classification [Ball, 2005], RCC has three major histologic subtypes : clear cell RCC [Lopez-Beltran et al., 2006] (ccRCC) accounting for 80 ~ 90%, papillary RCC (pRCC) accounting for 10 ~ 15% and Chromophobe RCC (chRCC) for 4 ~ 5% of RCC.

3.1.1 The Phenotype and Histomorphology of RCC Subtypes

ccRCC is the most common and aggressive RCC subtype, and approximately 2 ~ 3% of tumors in ccRCC are familial/hereditary cases associated with certain autosomal dominant syndromes, characterized by high metastasis, resistance to chemotherapy and radiotherapy, and poor prognosis. Concerning clinical aspects, ccRCC occurs with few symptoms and/or laboratory abnormalities, and approximately 20 ~ 30% of patients present advanced disease at diagnosis. Furthermore, approximately 30% of patients with locally diseased ccRCC will experience recurrence or metastasis after surgical resection of the tumor [Moch et al., 2016, Srigley et al., 2013]. So the early detection of ccRCC and monitoring of disease progression are critical steps to improve patient survival.

Although current imaging techniques can detect most renal masses (called small renal masses or SRMs, lower than 4.0 cm in size) when they are still small, 95% percent of SRM cases show slow growth kinetics, usually, only a few millimeters per year [Uzosike et al., 2018]. Among SRM, up to 20% of cases are benign, and only about 20% are found to be high-grade in surgically resected samples [Uzosike et al., 2018, Ball et al., 2015, Campbell et al., 2017]. The pathology of these small tumors cannot be accurately characterized without a percutaneous biopsy. ccRCC is defined histologically as malignant epithelial cells with clear cytoplasm, due to the vast accumulation of lipids and glycogen that are removed in standard histological preparations [Ericsson et al., 1966], as shown in Figure 3.1.

pRCC is the second most common type of kidney cancer after ccRCC. Chromophilic tumors, which mostly show a papillary growth pattern, are thus commonly referred to as pRCC with a relatively early age of onset and a lower male-female preponderance as compared to other renal cell cancers [Tonk et al., 1995, Dijkhuizen et al., 1998]. pRCC is the most heterogenous RCC; histological and molecular characterization studies have highlighted the heterogeneity of pRCC tumors [Wang et al., 2022c]. pRCC is usually histologically divided into 2 types, namely, type 1 and type 2. An example of the histopathological image of pRCC is shown in Figure 3.2.

chRCC is the third most common subtype of all RCCs, originating from progenitor cells (intercalated cells of the collecting tubule) [Gündüz et al., 2022]. Although rare chromophobe tumors have more aggressive behavior, especially those with adverse features such as vascular invasion, sarcomatoid change, or necrosis, most classic chromophobe tumors and almost all eosinophilic chromophobe tumors behave favorably after removal [Przybycin et al., 2011]. A histopathological image of chRCC is shown in Figure 3.3. Immunohistochemical staining for cytokeratin 7 recognize

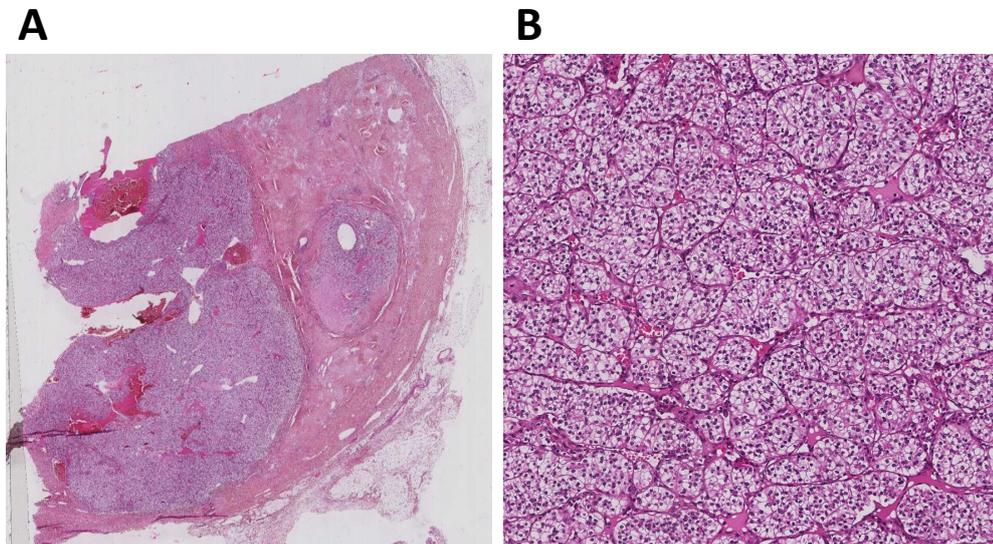


Figure 3.1 – Hematoxylin and eosin staining for ccRCC at different magnification. A : whole slice image, $\sim 60000 \times 60000$ Pixels. B : patch image, 2000×2000 Pixels

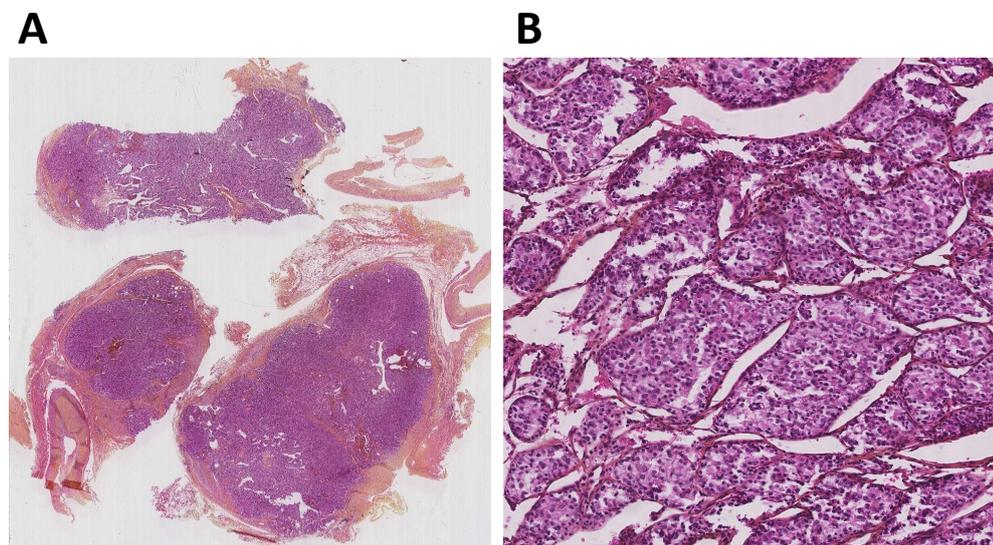


Figure 3.2 – Hematoxylin and eosin staining for pRCC at different magnification. A : whole slice image, $\sim 60000 \times 60000$ Pixels. B : patch image, 2000×2000 Pixels

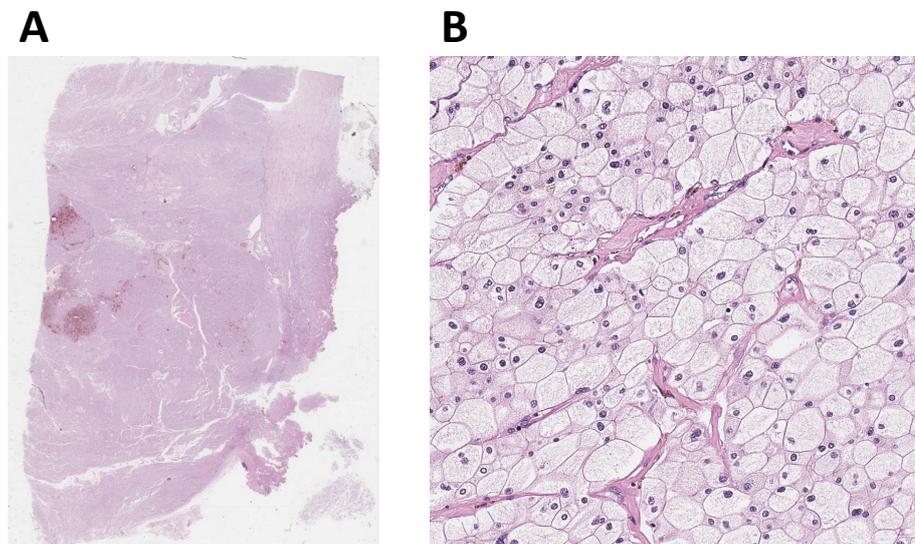


Figure 3.3 – Hematoxylin and eosin staining for chRCC at different magnification.
 A : whole slice image, $\sim 60000 \times 60000$ Pixels. B : patch image, 2000×2000 Pixels

to be a useful addition agent in the diagnosis of chromophobe renal cell carcinoma and in distinguishing this tumor from oncocytoma and conventional renal cell carcinoma [Mathers et al., 2002].

3.1.2 Cellular and Vascular Characteristics in RCC Subtypes

ccRCC has rich, sinusoid-like vascularity frequently used as a diagnostic criterion [Kryvenko et al., 2013]. Microscopically, in addition to the characteristic clear cell morphology determined by the accumulation of cytoplasmic lipids and glycogen, the tumor displayed a rich vascular network separated by fine sinus-like capillaries separating tumorigenic cells nests. The abundant vascular supply of ccRCC determines its typical contrast-enhanced radiographic appearance in contrast medium infusion studies [Grignon and Che, 2005, Tickoo et al., 2010, Grignon and Che, 2005].

pRCC has a lower vascular density than ccRCC. Type 1 consists of well-defined papillae lined by a monolayer of small cells with basophilic cytoplasm and low Fuhrman-grade nuclei [Fuhrman et al., 1982]. Type 2 is characterized by papillary lining with large pseudostratified cells, eosinophilic cytoplasm, enlarged nuclei, and prominent nucleoli, an example of Type 1 and Type 2 pRCC is shown in Figure 3.4 [Magers et al., 2019].

There are some key features [Williamson, 2022] distinguishing chRCC from potential mimics, such as prominent cell borders, wrinkled nuclei, variation in cell size, trabecular growth, and incomplete vascular network. The histological features most frequently associated with a diagnosis of chRCC were accentuated cell borders and a combination of hyperchromatic wrinkled nuclei and perinuclear halos. The most sensitive and specific marker for chRCC is parvalbumin [Abrahams et al., 2004].

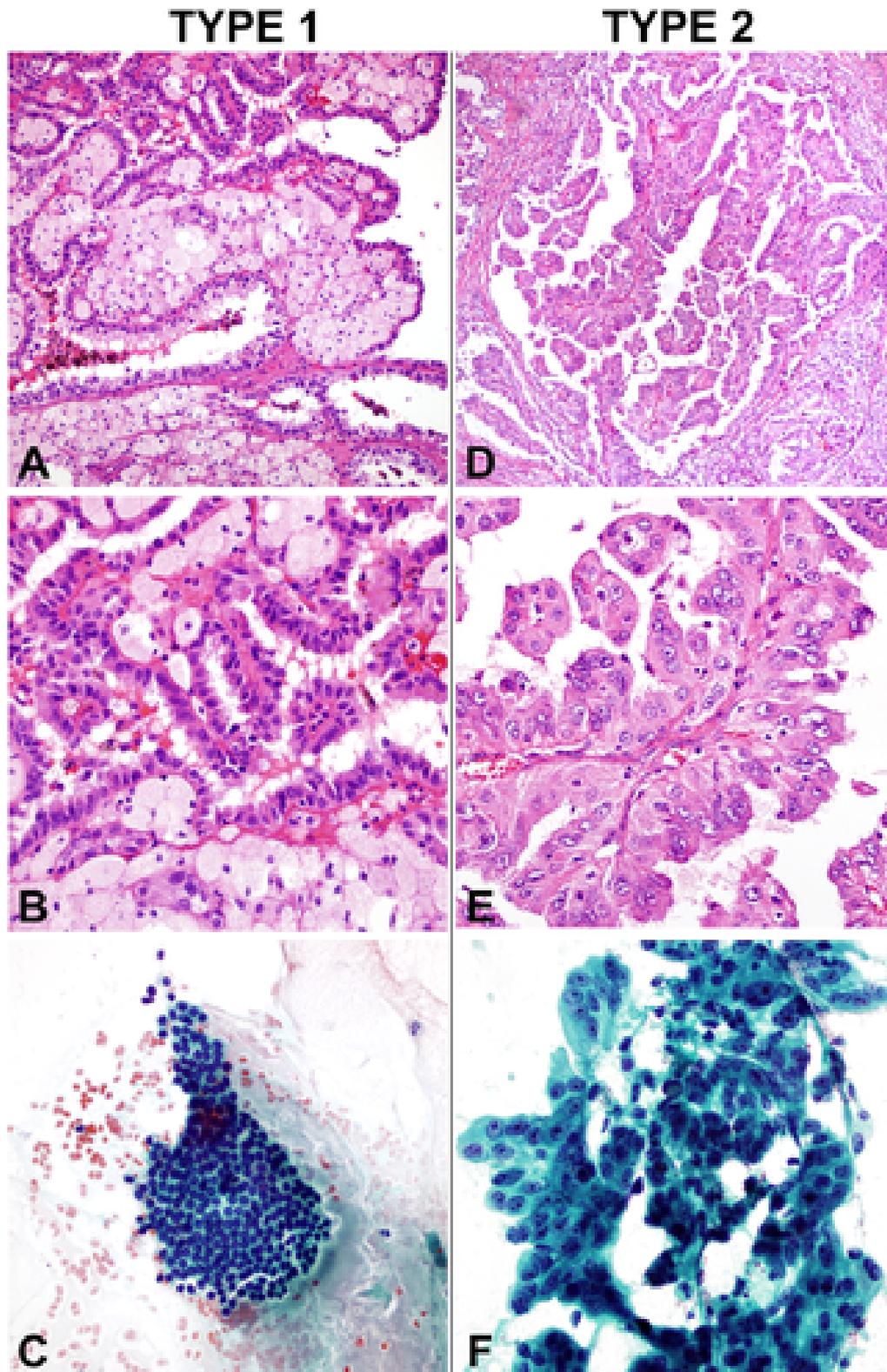


Figure 3.4 – The example of Type 1 and Type 2 of pRCC in histopathological image. Caption (A,B) Type 1 pRCC is characterized by papillae with fibrovascular cores lined by a single layer of neoplastic epithelial cells (H&E, $\times 200$ [panel A] or $\times 400$ [panel B]), and (C) type 1 pRCC usually exhibits a relatively low nuclear/nucleolar grade (eg, grade 1 or 2. Papanicolaou-stained direct smear, $\times 400$). (D,E) Type 2 pRCC is characterized by pseudostratified neoplastic epithelium lining fibrovascular cores (H&E, $\times 200$ [panel D] or $\times 400$ [panel E]), and (F) the nuclear grade is typically higher (eg, often grade 3. Papanicolaou-stained direct smear, $\times 400$). pRCC indicates papillary renal cell carcinoma. (Magers, 2019)

3.2 Histopathological Images Dataset

Classification and segmentation from histopathological images play a key role in computer-aided diagnosis or prognosis. There are some public histopathological images datasets for the classification tasks, such as BreaKHis [Spanhol et al., 2015] and BACH [Aresta et al., 2019b] for breast cancer, LC25000 [Borkowski et al., 2019] with five classes of lung tissue for lung cancer, and DigestPath [Li et al., 2019] for colon cancer, etc. However, RCC histopathological image datasets are rare, and there is no public database with vascular annotation. As far as we know, we are the first to perform RCC classification on the vascular network annotated histopathological image dataset.

3.2.1 BigRCC Dataset Building

3.2.1.1 Data Collection

The paraffin slices technique is very commonly used in the pathological examination. H&E staining is the most commonly used in the procedure of dyes. The hematoxylin staining solution is alkaline, and can stain the basophilic structure of the tissue (such as the ribosome, nucleus, and ribonucleic acid in the cytoplasm) into blue-violet; The eosin staining solution is an acid dye, which can stain the eosinophilic structure of the tissue (such as intracellular and intercellular proteins, including Lewy bodies, alcohol bodies, and most of the cytoplasm) stained pink, making the morphology of the entire cell organization visible.

We collected all the H&E staining slice samples of RCC subtypes from Hôpital Pasteur, Centre Hospitalier Universitaire (CHU) de Nice, France. They come from 46 patients of ccRCC, 22 patients of pRCC and 6 patients of chRCC. We have 107 slices of ccRCC, 51 slices of pRCC and 9 slices of chRCC total. Thus, we got the original WSIs, which are H&E-stained, scanned by Leica SCN400 scanner at 40x magnification ($\sim 60000 \times 60000$ pixels), the WSIs have a resolution of $0.25 \mu\text{m}$.

3.2.1.2 Categories Annotation

We use the open-access software ASAP to do the categories annotation, Figure 3.5 shows the display windows of ASAP at the 4X magnification of original WSIs. We consider four classes : tumor, necrosis, fiber and normal, the example of these four categories is shown in Figure 3.6.

The features of these four categories from H&E staining histopathological images are shown as follows :

- Tumor

Different subtypes of RCC have different tumor characteristics. ccRCCs include clear cells, acinar-like, sheet-like, or tubular growth patterns, and a rich vascular network. Some ccRCCs also have eosinophilic granular cytoplasm, inconspicuous interstitial blood vessels, and sarcomatoid differentiation. There are two subtypes of pRCC, type 1 papillary surface covered with smaller cuboidal cells, less or moderate amount of cytoplasm. Type 2 tumors have larger cells, higher nuclear grade, eosinophilic cytoplasm, and pseudostratified nuclei. While cells of chRCC have slightly opaque or finely reticular cytoplasm, they are usually large polygonal cells, showing a fine reticular cytoplasm and central round to oval nucleus, and are arranged in compact broad alveoli with or without tubular or tubulocystic areas [Yusenko, 2010].

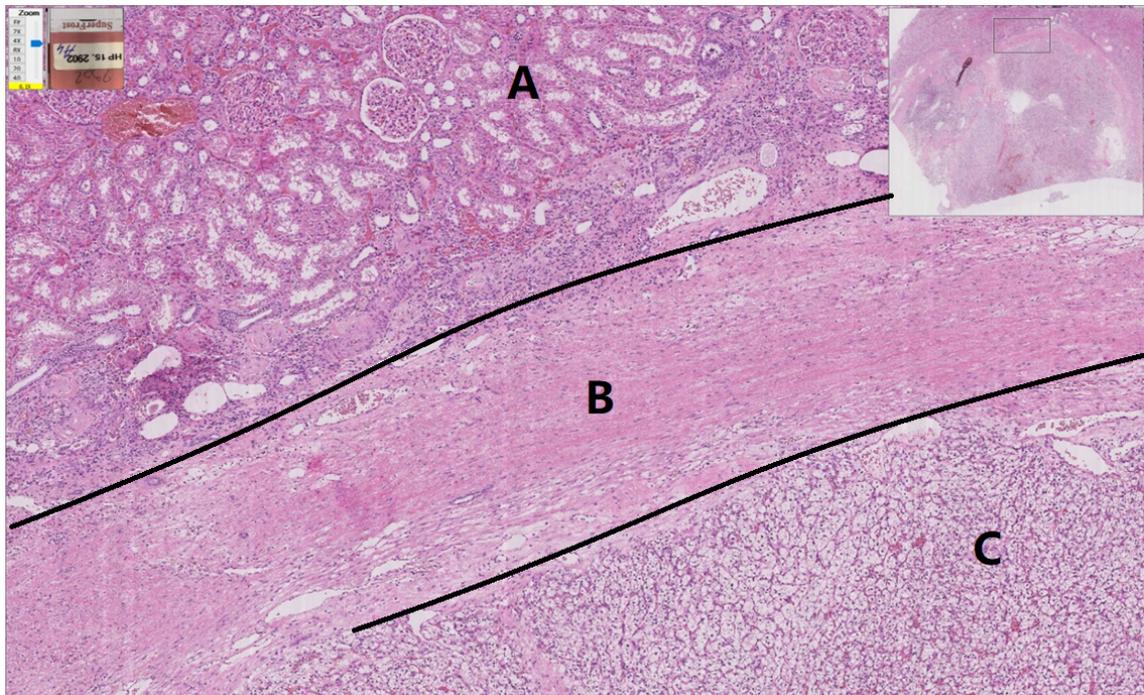


Figure 3.5 – The view of the histopathological image in ASAP.
The top left A is the normal area, middle B is the fiber normal, bottom right C is the tumor area.

— Necrosis

Changes in the nucleus are the main morphological markers of the necrosis area, which are manifested as : (1) pyknosis, the staining becomes darker, and the size of the nucleus becomes smaller. (2) karyorrhexis, the nuclear membrane ruptures. (3) Karyolysis, only the outline of the nucleus can be seen.

Changes in the cytoplasm are also a feature of necrosis, the cytoplasm of necrotic cells is red-stained due to the reduction or loss of the cytoplasmic basophilic substance nucleosome. At the same time, due to the disintegration of the cytoplasmic structure, the cytoplasm was granular.

— Fiber

The cells of the fibers are fusiform, arranged in bundles, and distributed in the collagen stroma.

— Normal

The normal area includes normal glomeruli with thin, clear vascular loops and normal surrounding tubules.

Then the different areas were annotated, and their coordinates were stored in an XML format file, shown in Figure 3.7. A grouped annotation example is shown in Figure 3.8.

Each slice has a unique identification. Let take the name of the slice in Figure 3.8 as an example, "HP17.7980.A2.pRCC.scn" means that this sample was collected in 2017, "7980" is the unique identification number of the patient, "A2" means the second H&E stained slice, "pRCC" is the pathologically diagnosed subtype of the patient, and ".scn" is the suffix for WSI image files.

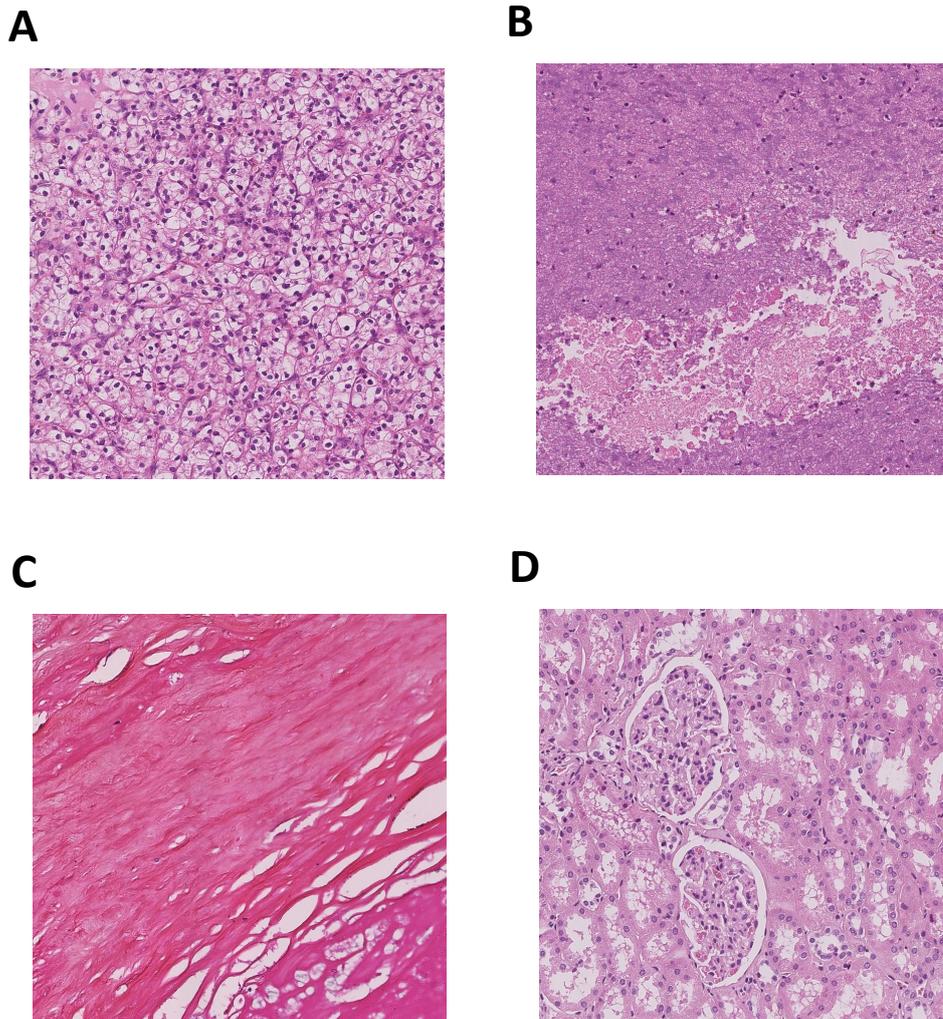


Figure 3.6 – The example of four categories of RCC histopathological images. A is the tumor patch, B is the necrosis patch, C is the fiber patch, D is the normal patch.

```

<?xml version="1.0"?>
<ASAP_Annotations>
  <Annotations>
    <Annotation Name="Annotation 0" Type="Spline" PartOfGroup="necrosis" Color="#F4FA58">
      <Coordinates>
        <Coordinate Order="0" X="43322.0586" Y="44901.5078" />
        <Coordinate Order="1" X="43434.875" Y="45747.6406" />
        <Coordinate Order="2" X="43491.2852" Y="46706.5938" />
        <Coordinate Order="3" X="44155.5234" Y="47633.9531" />
        <Coordinate Order="4" X="44154.4102" Y="48903.6953" />
        <Coordinate Order="5" X="44825.6797" Y="49968.7188" />
        <Coordinate Order="6" X="45572.5703" Y="51024.1094" />
        <Coordinate Order="7" X="46297.5" Y="52006.4844" />
        <Coordinate Order="8" X="47120.4023" Y="52775.6133" />
        <Coordinate Order="9" X="48005.1602" Y="53543.3203" />
        <Coordinate Order="10" X="48850.1328" Y="54378.207" />
        <Coordinate Order="11" X="49921.9023" Y="54885.8867" />
        <Coordinate Order="12" X="50993.6719" Y="55562.7969" />
        <Coordinate Order="13" X="52573.1211" Y="56070.4766" />
        <Coordinate Order="14" X="54671.0156" Y="55814.9141" />
        <Coordinate Order="15" X="54152.5703" Y="55449.9766" />
        <Coordinate Order="16" X="53525.3008" Y="54950.2539" />
        <Coordinate Order="17" X="53262.4375" Y="54373.7578" />
        <Coordinate Order="18" X="52379.3242" Y="53876.8867" />
        <Coordinate Order="19" X="51708.0195" Y="53019.957" />
        <Coordinate Order="20" X="51399.7734" Y="52109.2148" />
        <Coordinate Order="21" X="50937.2617" Y="50824.4453" />
        <Coordinate Order="22" X="49865.4922" Y="49470.6289" />
        <Coordinate Order="23" X="48286.043" Y="48568.0859" />
        <Coordinate Order="24" X="47721.9531" Y="48060.4063" />
        <Coordinate Order="25" X="46480.957" Y="46875.8203" />
        <Coordinate Order="26" X="45522.0039" Y="46198.9141" />
        <Coordinate Order="27" X="44450.2344" Y="45578.4141" />
      </Coordinates>
    </Annotation>
  </Annotations>
</ASAP_Annotations>

```

Figure 3.7 – The XML file format of grouped annotated areas.

3.2.1.3 Patch Images

Finally, we cut the annotated areas of WSIs into smaller patch images to form the BigRCC database. First, we combined the original image and XML file to find the area borderline according to the coordinates from an XML file. Then we use open-slide open source software (win64-20171122) to iterate every region of interest (ROI) of grouped annotations to generate 2000×2000 Pixels patch images from the original image, shown in Figure 3.9.

The pipeline of the dataset building is shown in Figure 3.10. The number of collected samples for the BigRCC dataset is shown in Table 3.1.

There are very few open-access histopathological image databases. TCGA database is the only one that contains the WSI data of RCC, which has two subtypes, ccRCC and pRCC. Our BigRCC includes three subtypes and has detailed annotations, which is particularly valuable and has great research contributions. Although in this project, we only use the tumor patch images for RCC subtyping research based on vascular networks, the large database BigRCC provides a strong data foundation for RCC-related research, which makes it possible to develop some deep learning models that require a large amount of input, such as our collaborator from Politecnico di Torino, Turin, Italy, using BigRCC data to classify RCC tumors and non-tumors with CNN-based models (publication under review). Another of our collaborators in the Abou Bakr Belkaid University of Tlemcen, Algeria uses BigRCC data for Tumor Regions Detection in Histopathology Images based on the CNN model [Benomar et al., 2021].

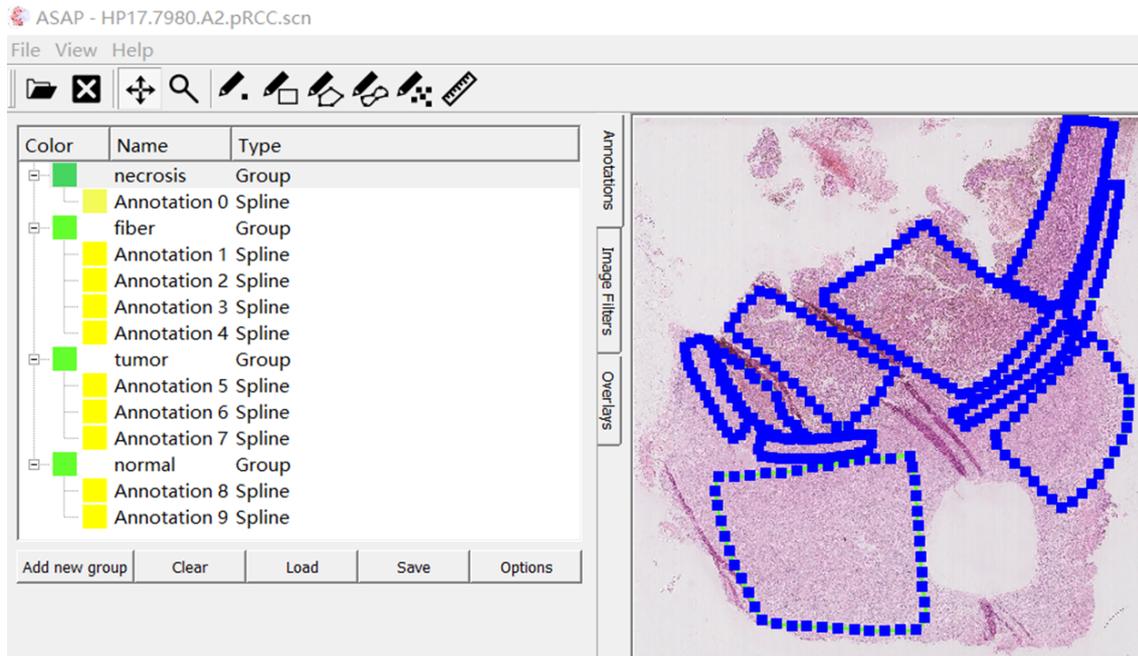


Figure 3.8 – The example of grouped annotation, the group include necrosis, fiber, tumor and normal.

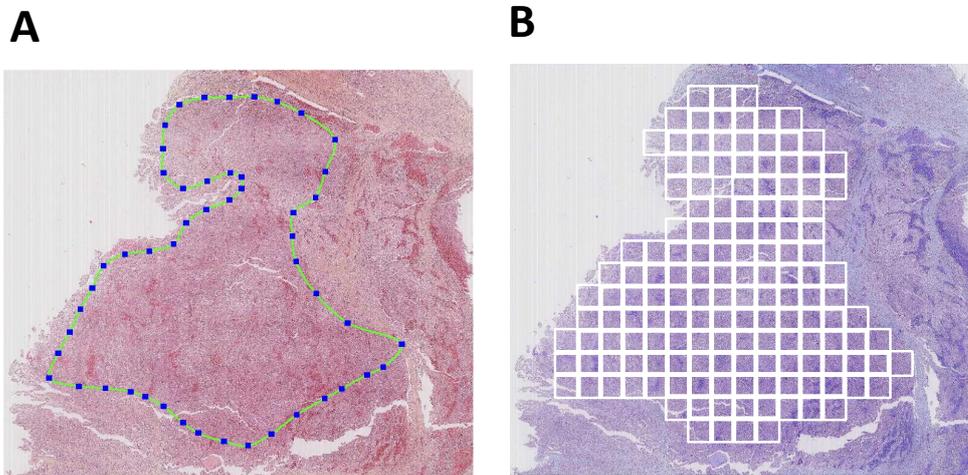


Figure 3.9 – Get patch images (2000×2000 pixels) from tumor annotation. A : WSI with tumor annotation. B : Display of patch images.

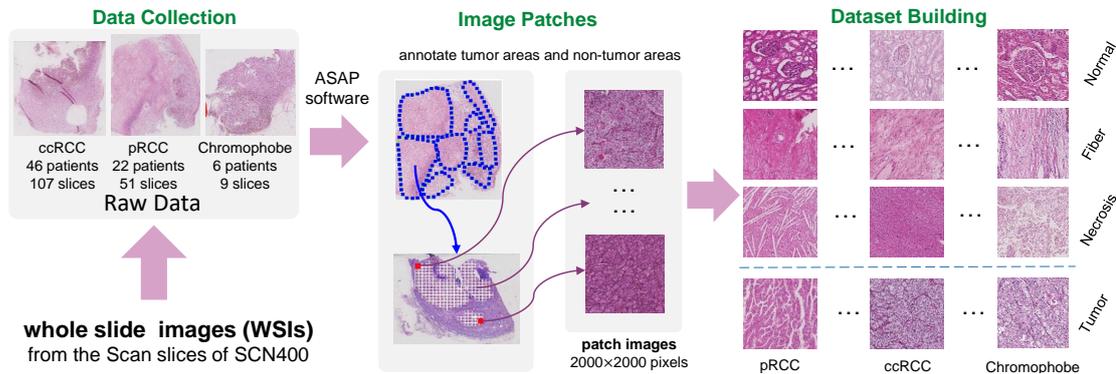


Figure 3.10 – BigRCC RCC histopathological images dataset building.

TABLE 3.1 – The number of image patches of each category in BigRCC.

BigRCC	Necrosis	Fiber	Normal	Tumor	Total
ccRCC	3324	1941	7459	27287	39986
pRCC	1602	920	2105	13637	18254
chRCC	79	170	1037	3134	4420

3.2.2 TCGA Open-access Dataset

The Cancer Genome Atlas (TCGA) is a large open-access database (<https://portal.gdc.cancer.gov/>), which was launched in 2006 by the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI). This database contains information about cancer patients, which includes data on gene expression, mutation, copy number, methylation, and clinical information of the same sample. As shown in Figure 3.11, it currently contains data from 20,000 patients and 36 types of cancer. We mainly use the pathological diagnosis slide data of H&E staining in this project.

3.2.2.1 Downloaded Data

TCGA officially provides the `gdc_client` download tool (<https://gdc.cancer.gov/access-data/gdc-data-transfer-tool>), which can download the data published by TCGA. This download tool application requires OS Linux (Ubuntu 16.x or later), OS X (10.9 Mavericks or later), or Windows (8 or later). Some data are open and can be downloaded without limit, while others are controlled and cannot be downloaded without a specific application. All the data we downloaded for my project are open and validated. The downloaded data of RCC in TCGA includes two subtypes, ccRCC and pRCC, and their project IDs are TCGA-KIRC and TCGA-KIRP respectively. Meanwhile, we also downloaded other types of cancers to test our vascular network segmentation model, which includes breast cancer, lung cancer, liver cancer and esophagus cancer.

Next, we downloaded the data through the Manifest option to get a .txt file containing the downloaded information of samples.

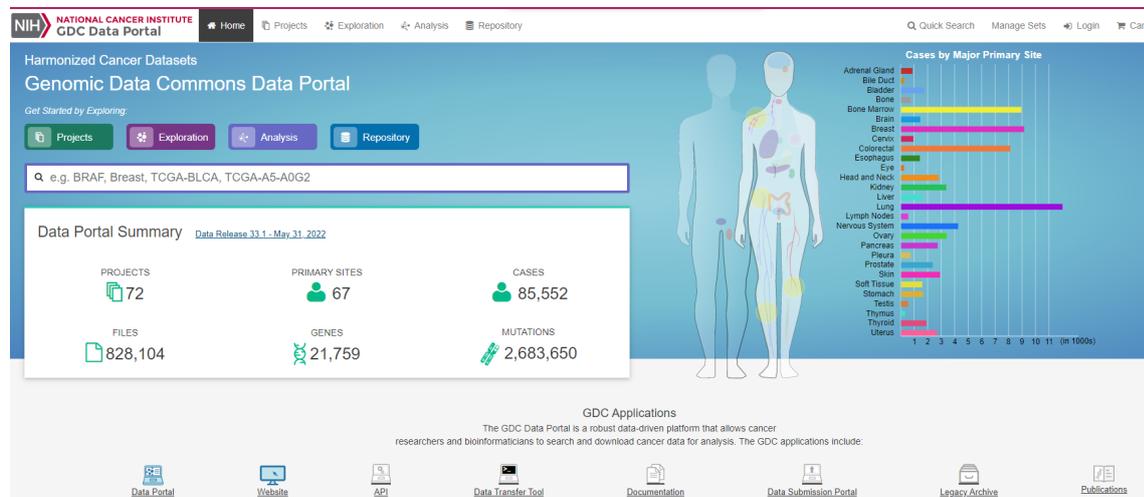


Figure 3.11 – The homepage of TCGA database.

3.2.2.2 Sample Naming

TCGA has a separate ID for each patient, such as 'TCGA-A6-6650-01B-02R'. This ID is common in the TCGA database. According to this ID, we can find the same patient in different types of databases, including clinical follow-up information. Let split this example ID with "-" and disassemble it for a detailed interpretation :

- TCGA : Project, all TCGA sample names start with this.
- A6 : Tissue source site, tissue source code, such as A6 means colon cancer tissue from Christiana Healthcare Center. For more information about the meaning of codes, can refer to : <https://gdc.cancer.gov/resources-tcga-users/tcga-code-tables/tissue-source-site-codes>.
- 6650 : Participant ID.
- 01 : Sample, where the numbers 01 09 represent tumors, and 10 19 represent normal controls.
- B : Vial, the order in a series of patient tissues, Vial indicates the order of the tissue, A, B, C, D. For example, B indicates paraffin-embedded tissue.
- 02 : Portion, the serial number of different parts belonging to the same patient tissue, the same tissue will be divided into 100-120 mg parts and used separately.
- R : Analyte, the type of molecule analyzed, the corresponding information is shown in this link : <https://gdc.cancer.gov/resources-tcga-users/tcga-code-tables/portion-analyte-codes>.

3.2.2.3 Image Patches of TCGA

Following the same methods for BigRCC dataset building, we get patch images of 5 cancers from the TCGA database, the number of slices and patch images for each cancer are shown in Table 3.2. We downloaded 100 slides and got 1029 image patches from TCGA totally.

TABLE 3.2 – The number of slices and image patches of each cancer in TCGA.

cancer	Project	slices	image patches
RCC	TCGA-KIRC & TCGA-KIRP	20	433
breast	TCGA-BRCA	20	60
liver	TCGA-LIHC	20	246
lung	TCGA-LUSC	20	120
esophagus	TCGA-ESCA	20	170

3.2.3 VRCC (Vascular RCC) Dataset Building

3.2.3.1 Data Source of VRCC

We randomly choose 424 tumor image patches from BigRCC and 90 from TCGA to construct the VRCC dataset (129 for ccRCC, 129 for pRCC, 166 for chRCC, 20 for RCC_TCGA, 15 for breast cancer, 20 for liver cancer, 20 for lung cancer and 15 of esophagus cancer). The data of the VRCC dataset comes from 31 patients of the BigRCC dataset (13 ccRCC patients, 14 pRCC patients and 4 chRCC patients) and 100 patients of TCGA (20 patients for each cancer).

3.2.3.2 Weakly Vascular Network Annotation and Binary Mask

We use the "ImageJ" software to annotate the vascular networks, as shown in Figure 3.12, and checked them by medical specialists. We choose the "Enable selection brush" with the size of 25 pixels to make the vascular network annotation, then save the annotations as .roi files from a series of operations (Analyze → Tools → ROI Manager). Then we got the binary mask image (Edit → Clear Outside → Fill). At last, we save the binary mask image (File → Save As → PNG with its name), as shown in Figure 3.13. The pixel value of vascular is 255 with white color, and the value of the background is 0 with black color.

The weakly label of the vascular network has been made with constant width bands and thus does not take into account the width information and local geometry. We used this annotation method to speed up the process. We plan to classify RCC subtypes only by extracting relevant features from the skeleton of the vascular network, so the length of the vessels and the network connectivity are more important than the width. So to save time and improve efficiency, we adopt this weakly labeled method to explore the importance of the vascular network in RCC typing. The comparison between the real vascular network mask and weakly label is shown in Figure 3.14. The width of the real vascular network mask is different in thickness and is thicker at some nodes than at the end parts. However, the weakly mask roughly keeps consistent in length with the real vascular network mask.

As shown in Figure 3.15, the vascular structure of ccRCC is like a "fishnet", pRCC looks like a "tree" while chRCC has a "linear" structure. To describe these vascular networks, we consider three elements of *junctions*, *end branches* and *non-end branches* (branches between two junctions) to represent it.

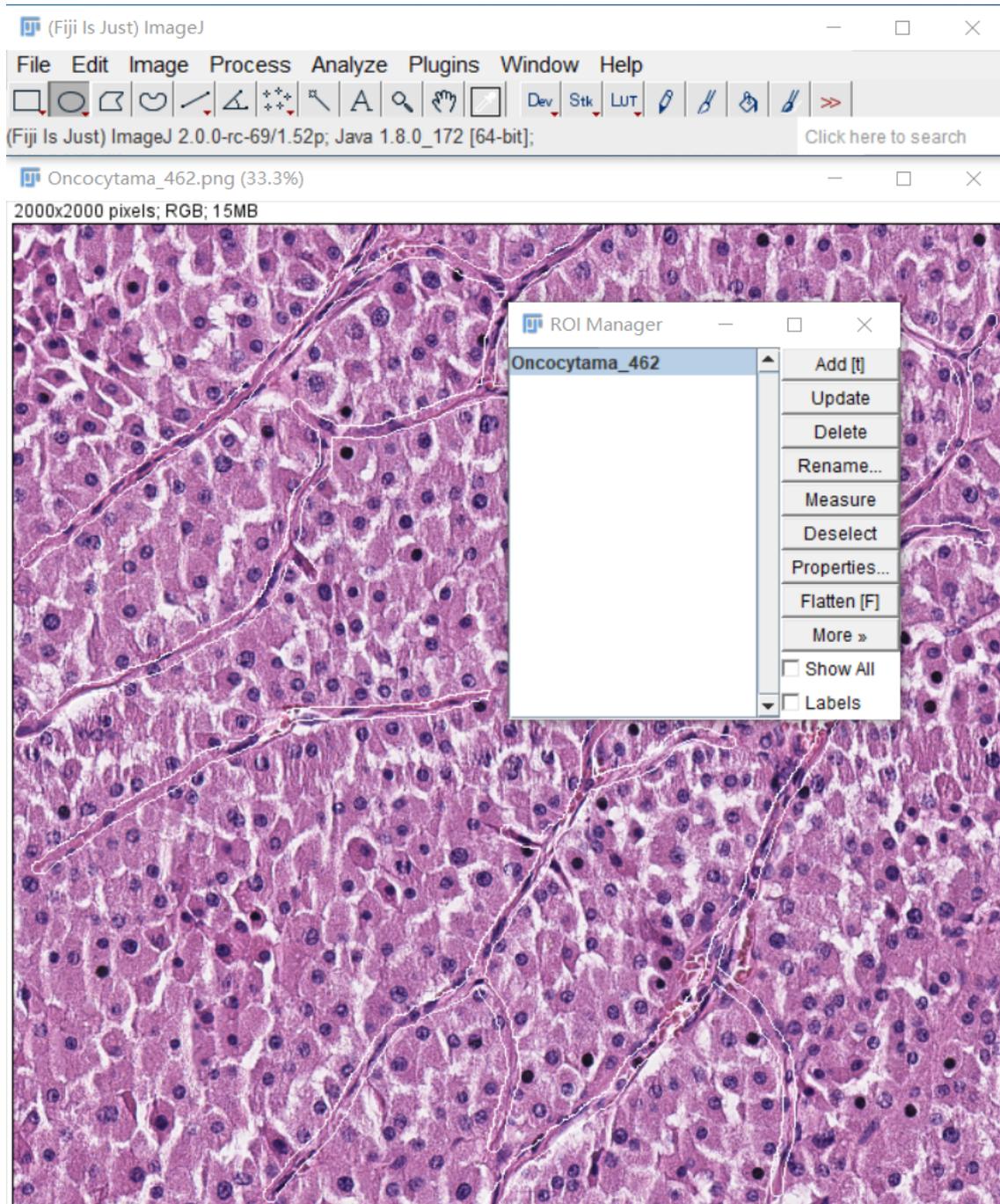


Figure 3.12 – The view of the histopathological image in ImageJ.

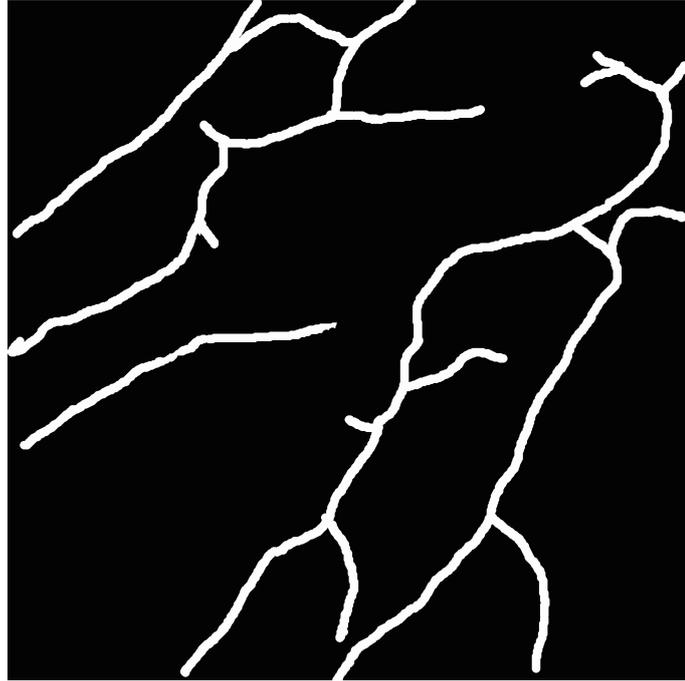


Figure 3.13 – The binary mask image of the vascular network.

3.3 Conclusion

Pathological diagnosis is still an important part of the diagnosis of RCC, and computer-aided diagnosis provides a bright vision for promoting medical development. Correctly classifying RCC subtypes through the WSIs using machine learning models is also an ongoing goal of researchers. In histopathological images, cellular and vascular morphologies showed different appearances in different subtypes of RCC. Based on a large number of cell-related characterization studies, in this project, we focused on exploring the importance of the vascular network in distinguishing RCC subtypes, to investigate whether the vascular network can be used alone for the RCC subtypes classification task. We collected the three most common RCC subtypes : ccRCC, pRCC, and chRCC. We annotated tumor and non-tumor regions by ASAP and then cut the original WSIs into patch images with 2000*2000 pixels according to the annotated regions to facilitate subsequent analysis. All patch images containing category information constitute the BigRCC dataset. Subsequently, for the study of the vascular network, we annotated the binary masks of the vascular network of 514 tumor patch images with ImageJ, thereby constructing the database VRCC with vascular network annotations.

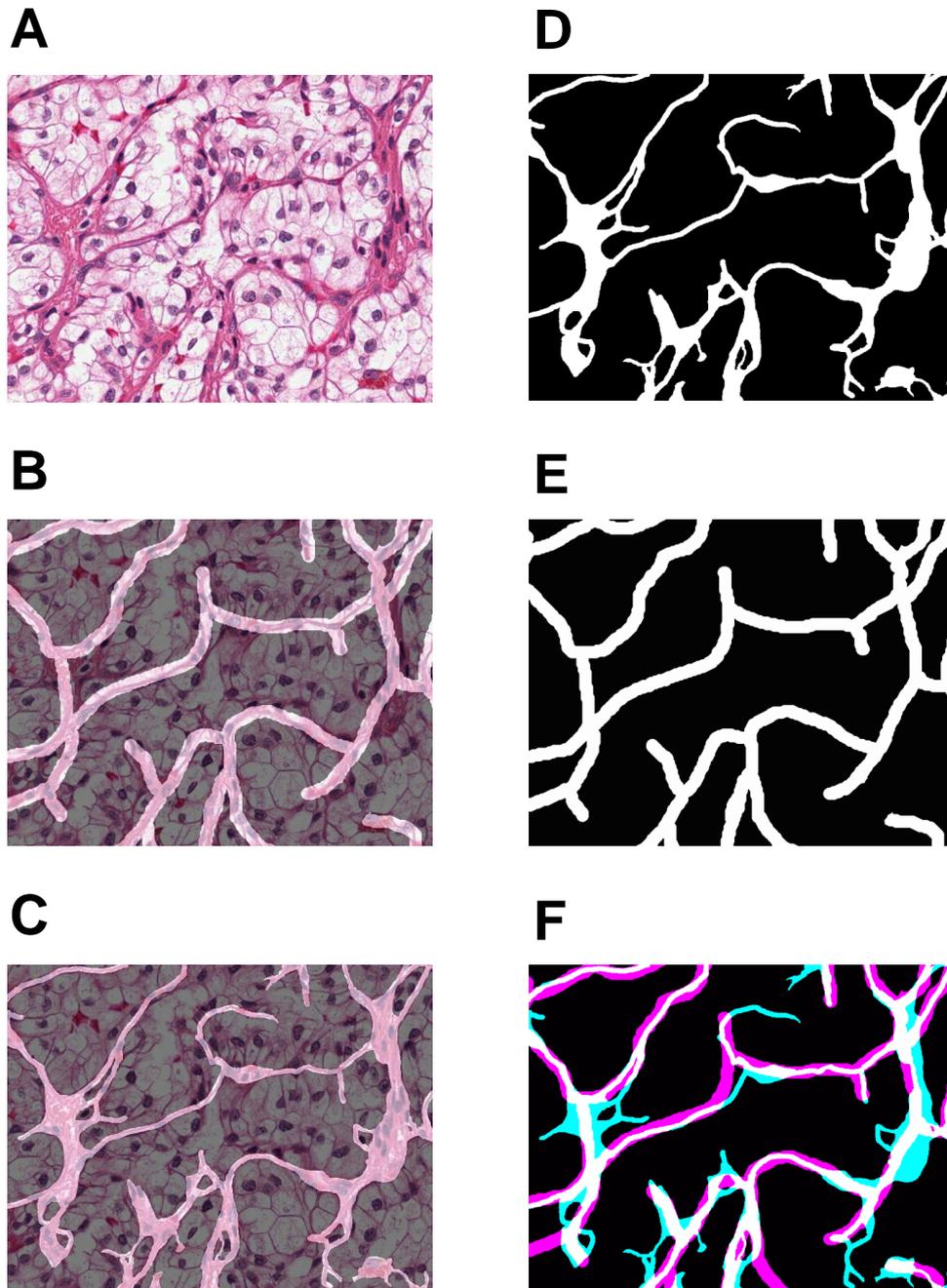


Figure 3.14 – The view of the histopathological image in ImageJ.

A : original image. B : overlay weakly mask on original image. C : overlay real mask on original image. D : real vascular mask. E : weakly vascular mask. F : overlay weakly mask on the real mask, the cyan color represents the real mask, the magenta color represent the weakly mask, the white color represents the overlap part.

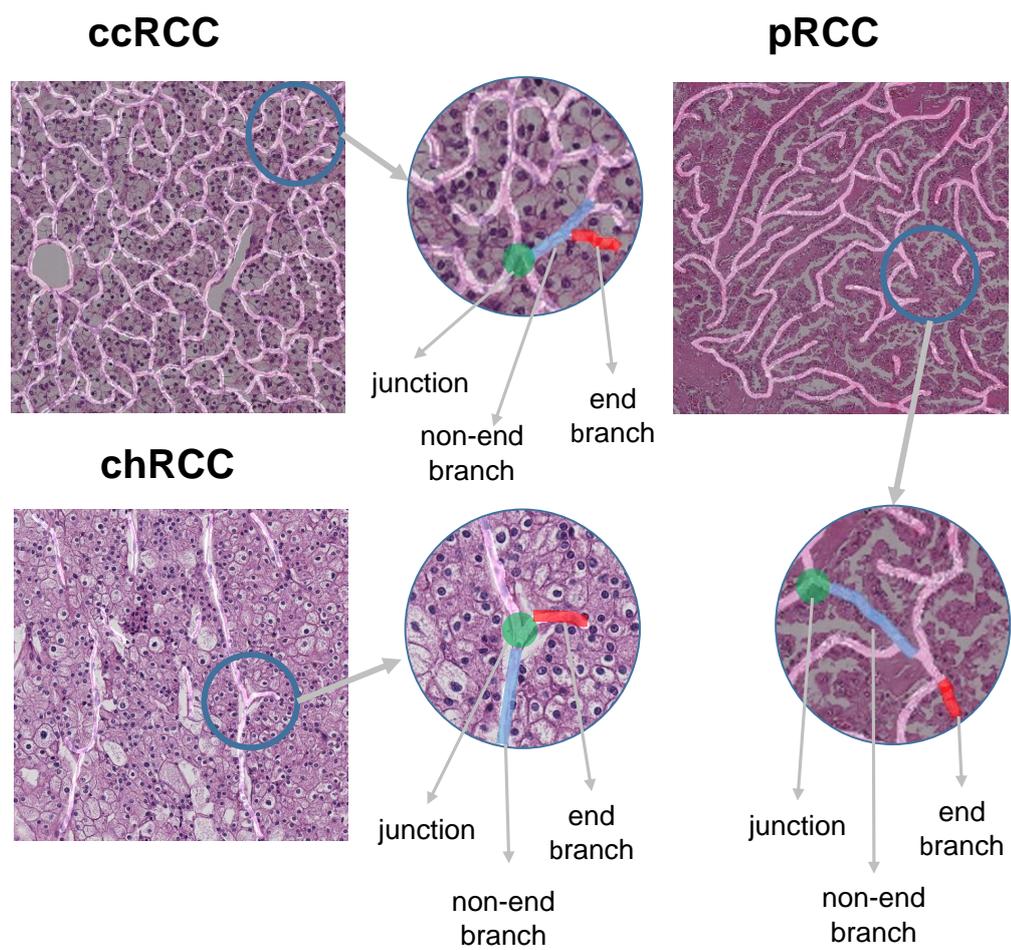


Figure 3.15 – The vascular network examples of ccRCC, pRCC and chRCC images.

RCC subtypes Classification of Histopathological Images

CHAPTER 4

Vascular Network-based Features

Dans ce chapitre, nous introduisons plusieurs types de caractéristiques liées aux réseaux vasculaires, telles que les caractéristiques ad-hoc extraites des réseaux, les propriétés des graphes modélisant le réseau vasculaire et les paramètres issus des approches par apprentissage profond.

In this chapter, we introduce several types of features related to vascular networks, such as hand-crafted features, graph features, and deep learning features.

4.1 Hand-crafted Features	63
4.1.1 Skeleton Features	63
4.1.2 Lattice Features	63
4.2 Graph-based Features	66
4.3 Deep Learning Features	66
4.4 Conclusion	67

4.1 Hand-crafted Features

4.1.1 Skeleton Features

We use the `bwskel` function of MatLab (version=R2018b) to convert the manually weak vascular network masks into a skeleton, shown in Figure 4.1. The vascular network mask mainly has three elements : junctions, end branches and non-end branches, as shown in Figure 3.15. The fishnet-like ccRCC has more junctions and more non-end branches, the tree-like pRCC has more end branches, while the linear-like chRCC has the lowest density and the least number of junctions. The graph visualization in Figure 4.4 also supports our assumption. The skeleton features are extracted from the skeleton directly, we use the Number of End branches (NE), small NE (less than nuclear size $\times 10$), long NE (more than nuclear size $\times 10$), average Length of the End branches (LE), Number of Junctions (NJ), average Length of the non-End branches (LJ), the density of the vascular network as the basic features of the skeleton. Moreover, we also consider NE/NJ , $NE/(LJ+LE)$, $NJ/(LJ+LE)$, LE/LJ , and $LJ/(LJ+LE)$ to make the skeleton features robust and comprehensive. The details are shown in Table 4.1.

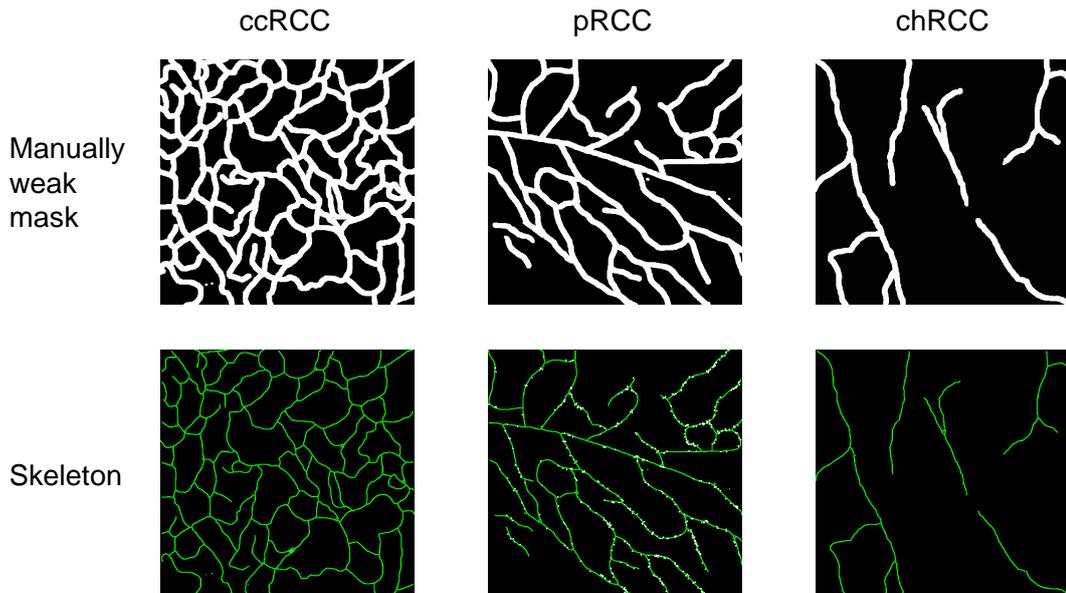


Figure 4.1 – The example of the skeleton of RCC subtypes.

The skeleton maintains the shape of the mask and keeps the width of the skeleton at 1 pixel, which facilitates subsequent graphics calculation operations.

4.1.2 Lattice Features

Based on the different structures of vascular networks of ccRCC, pRCC and chRCC, we further propose a new set of features "lattice features". But they are not extracted from the skeleton directly. As shown in Figure 4.3, we first compute the watershed [Soille and Vincent, 1990] on the vascular skeleton to obtain closed areas. The watershed calculation process is an iterative

labeling process. The classical calculation method of the watershed is proposed by L. Vincent [Soille and Vincent, 1990].

The watershed algorithm is a mathematical morphological segmentation method based on topology theory. The main idea is to treat the image as a topological landscape in geodesy. The water collection basin is the local minimum and its affected area, and the boundary of the water collection basin forms a watershed. The concept and formation of the watershed can be explained by simulating the immersion process. On the surface of each local minimum, a small hole is pierced, and then the entire model is slowly immersed in water. As the immersion deepens, the influence domain of each local minimum slowly expands outward. A dam is constructed at the confluence of the basins, forming a watershed. the procedure of the watershed is shown in Figure 4.2 [Chen et al., 2022].

The whole process of the watershed algorithm is as follows :

- All pixels in the gradient image are classified according to their gray values, and a geodesic distance threshold is set.
- Find the pixel with the smallest gray value (marked as the lowest gray value by default), let the threshold increase from the minimum value, and these points are the starting points.
- In the process of growing the horizontal plane, it will encounter the surrounding neighborhood pixels, and measure the geodesic distance from these pixels to the starting point (the lowest point of gray value). If it is less than the set threshold, these pixels are flooded, otherwise, dams are set on these pixels, so that these neighborhood pixels are classified. In the process of submerging from low to high, the influence domain of each local minimum at the h-order height is judged and marked by a first-in-first-out (FIFO) structure.
- As the water level gets higher and higher, more and higher dams will be set up, until the maximum value of the gray value, all areas meet on the watershed line, and these dams partition the entire image pixels.

The watershed algorithm responds well to weak edges and is guaranteed to form closed and continuous edges. We first calculate the Euclidean distance transform of the binary vascular skeleton image (Legend A of Figure 4.3), as shown in the Legend B of Figure 4.3. However, noise in the image and subtle grayscale changes on the surface of the object will cause over-segmentation, when we use the watershed algorithm directly on the inverse distance transform of the vascular skeleton (Legend C of Figure 4.3), this results in over-segmentation as shown in the Legend D of Figure 4.3.

Over-segmentation is due to many local minima. In this case, the general trick is to first use the "imextendedmin" function of MatLab in watershed-based image segmentation to filter out some particularly small (referring to small regions) local minima. Then we use the minima imposition method [Beucher, 1990, Meyer and Beucher, 1990] to solve the problem of local minimum and modify the distance transformation, as shown in Legend E of Figure 4.3. We modify the result of the distance transformation is to make sure there is no local minimum in the filtered region, which is the region after "imextendedmin" operation. This "minima imposition" operation can be implemented with the function "imimposemin" of MatLab.

Then, we remove surrounding lattice cells which do not include complete vascular information. The final obtained lattice spatial map is shown in Legend F of Figure 4.3.

Finally, the closed catchment basin obtained by the watershed algorithm supplies the probability to analyze the regional characteristics of the image. We define and extract six features from the lattice map, including mean area, median area, mean perimeter, median perimeter, mean eccentricity, and median eccentricity. The details are shown in Table 4.1.

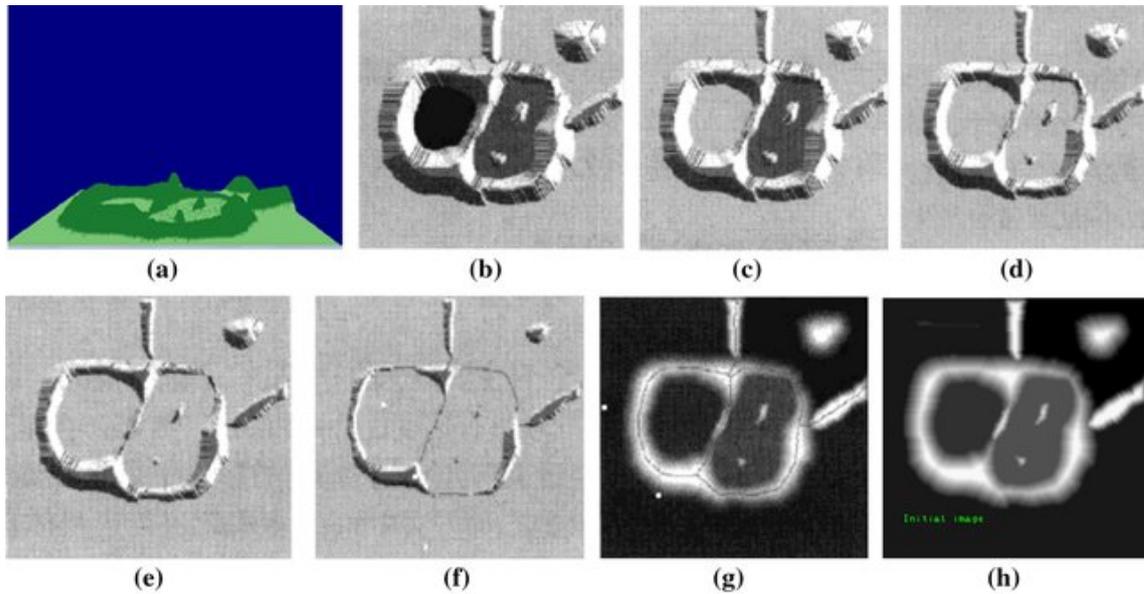


Figure 4.2 – Schematic diagram of the watershed algorithm.

a : Original image; b : Topographic view; c-d : Two stages of flooding; e : Result of further flooding; f : Beginning of merging of water from two catchment basins (a short dam was built between them); g : Longer dams; h : Final watershed (segmentation) lines. (Chen, 2022)

TABLE 4.1 – The explanation of each skeleton and lattice feature.

Skeleton Features	note
NE	the Number of End branches
LE	average Length of the End branches
small NE	NE that LE less than nuclear size $\times 10$
long NE	NE that LE more than nuclear size $\times 10$
NJ	Number of Junctions
LJ	average Length of the non-End branches
density	the sum of skeleton pixels
NE/NJ	NE/NJ Ratio
LE/LJ	LE/LJ Ratio
NE/(LJ+LE)	NE/(LJ+LE) Ratio
NJ/(LJ+LE)	NJ/(LJ+LE) Ratio
LJ/(LJ+LE)	LJ/(LJ+LE) Ratio
Lattice Features	note
mean area	mean of all lattice areas
median area	median of all lattice areas
mean perimeter	mean of all lattice perimeters
median perimeter	median of all lattice perimeters
mean eccentricity	mean of all lattice eccentricities
median eccentricity	median of all lattice eccentricities

4.2 Graph-based Features

Because the vascular network is a graph-like structure, we transform the vascular network skeleton into a graph by SKL-Graph [Debreuve, 2020]. SKL-Graph is a computer program whose purpose is to detect objects in a signal, where an object is defined as a piece of the signal represented by a set of parameters. A signal is typically a two-dimensional or three-dimensional image. SKL-Graph is programmed by Python language that can be used to create, manipulate, and learn the structure, dynamics, and functions of complex graph networks.

Any graph is composed of nodes and edges. Usually, any graph is represented by two sets : denoted as $G = (V, E)$, where V is a finite set of nodes, and E is a finite set of edges. As shown in Figure 4.4, in our case the nodes are represented by junctions and end-points, and the edges connect the nodes according to the shape of the vascular network.

We convert high-dimensional sparse matrices into compressed sparse row matrices (CSR matrices). The procedure is shown in Figure 4.5. While our original matrix stores data in 2D arrays, the transformed CSR matrices store them in three 1D arrays (Value array, Column index array, and Row index array). The interpretation is shown as follows :

- Value array : It stores all elements without 0 in the original matrix. The length of the array is equal to the number of nonzero entries in the original matrix. In this example, there are 7 non-zero elements. So the value array has a length of 7.
- Column index array : This array stores the column indices of the elements in the value array. (zero-based indexing is used here).
- Row index array : This array stores the cumulative count of all non-zero values in the current and previous rows. $row_index_array[j]$ encodes the total number of nonzeros above the j th row. The last element represents the number of non-zero elements in the original array. The length is $m + 1$; the number of rows in the original matrix is m .

Then we generated 4 files as input, which include sparse adjacency matrices, node labels, graph indicators, and graph labels, then feed them into GCN [Kipf and Welling, 2016]. The hand-crafted features have both topological and geometrical information on the vascular network, but graph-based features only contain the topological information, we compared these two sets of features' performance to assess the importance of geometrical information in the vascular network-based RCC subtype classification task. Most deep learning models are data-driven black-box models, although we consider only topological information as the branch length is not coded in the adjacency matrix, this feature is more explainable than the deep features systematically learned during network optimization.

4.3 Deep Learning Features

Deep Learning Features include two kinds of features. We consider both raw images and vascular segmentation images as inputs to further evaluate if the vascular network can be used alone for RCCs classification. Raw images contain all the information, not only the vascular network but also the complex background, which includes the character of cells and tissues. Vascular segmentation images include all the information about the vascular network, but their interpretability is harder than for hand-crafted features. We learn these deep features with deep learning models as baseline experiments, such as LeNet [Wright, 1995], AlexNet [Krizhevsky et al., 2012a], VggNet [Simonyan and Zisserman, 2014] and ResNet [He et al., 2016]. We compare the performance of

these features to evaluate whether our proposed hand-crafted features are representative and sufficient to distinguish different subtypes of RCC.

4.4 Conclusion

The features play a very important role in describing how vascular networks classify RCC subtypes. These three sets of features all represent the vascular network to a certain extent. The vascular network can be represented by the shape directly (Skeleton Features), by the area properties after processed (Lattice Features), by the whole mask images (Deep Learning Features), or by the adjacency matrices after conversion to graph (Graph-based features). These sets of features facilitate our next classification experiments.

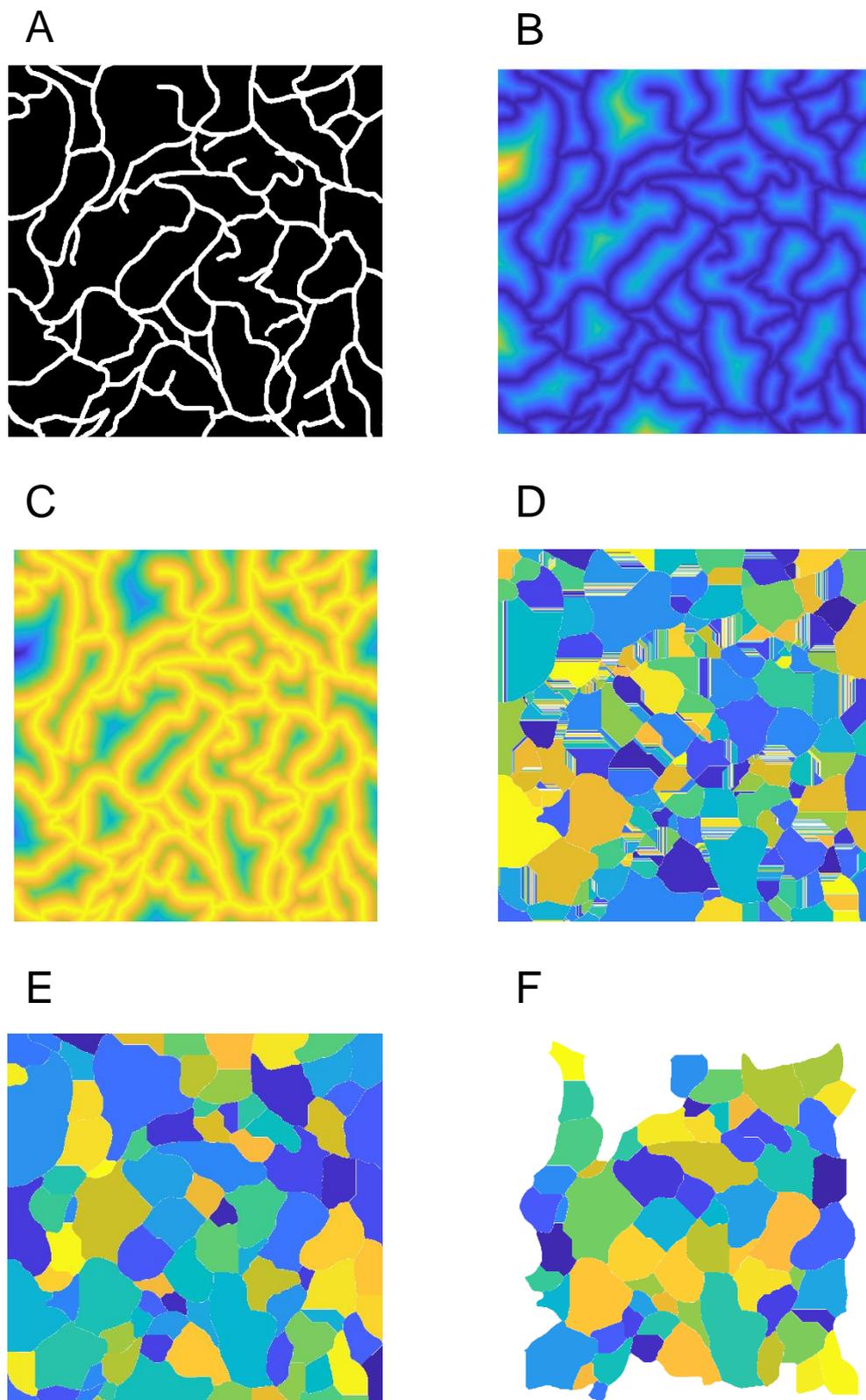


Figure 4.3 – The pipeline of obtaining the lattice features.

A : The original vascular skeleton ; B :The Euclidean distance transform of the binary vascular skeleton image ; C : The inverse distance transform of vascular skeleton ; D : The over-segmentation result after watershed work directly on the skeleton ; E : The distance transform after minima imposition modified ; F : The final lattice spatial map after surrounding remove.

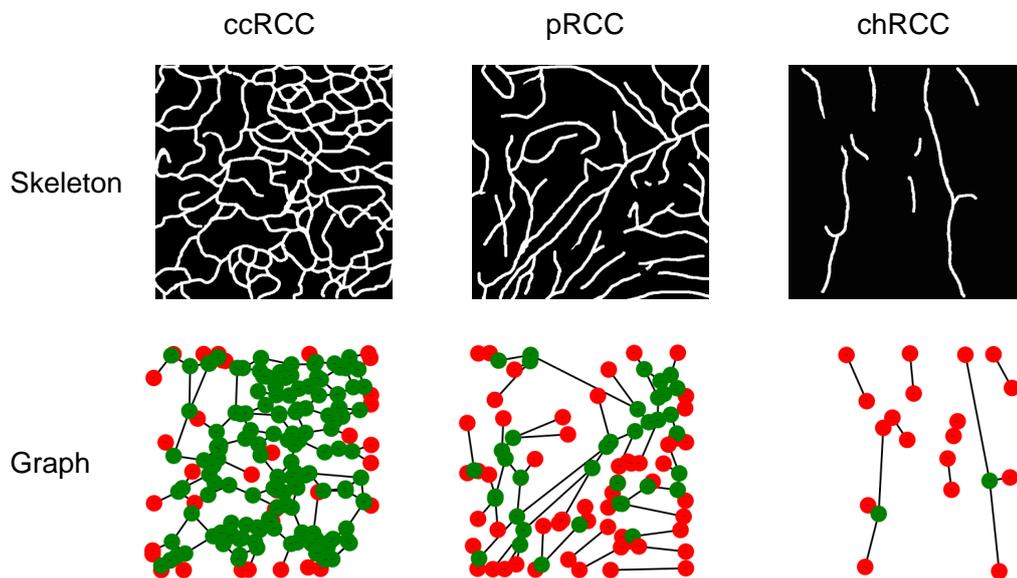


Figure 4.4 – The graph features of ccRCC, pRCC and chRCC. The red points represent end points of the vascular network, green points represent junctions of the vascular network.

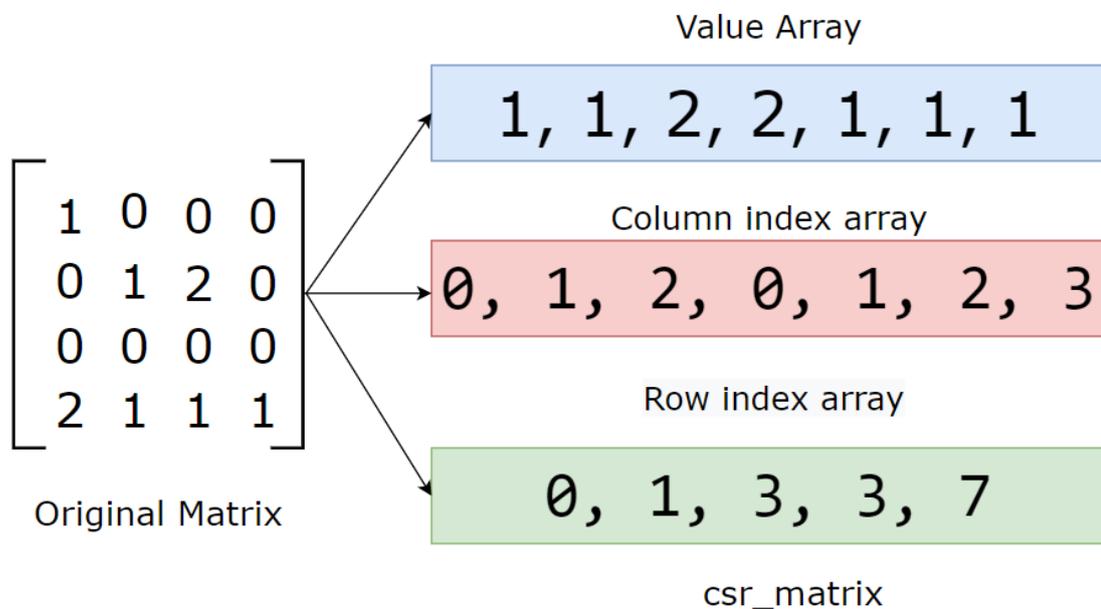


Figure 4.5 – The procedure of convert 2D Array Matrix to CSR Matrix. The CSR Matrix includes three arrays (Value array, Column index array, and Row index array). (Ransaka, 2021)

Vascular-based RCC Classification

Dans ce chapitre, nous avons entraîné les trois approches mentionnées dans le chapitre précédent (fondées sur des caractéristiques ad-hoc, fondées sur une modélisation par graphes et par apprentissage profond) en utilisant les méthodes traditionnelles d'apprentissage et d'apprentissage profond. Nous effectuons une analyse statistique des caractéristiques choisies et montrons leur pertinence pour la classification des réseaux vasculaires associés aux différents types de carcinome rénal à cellules claires. Les résultats montrent que les caractéristiques ad-hoc associées à des algorithmes de classification supervisée classiques donnent de meilleurs résultats que les approches par apprentissage profond sur l'ensemble des expériences réalisées.

In this chapter, we trained the three features mentioned in the previous chapter (our hand-crafted features, graph features and deep features) using traditional machine learning and deep learning methods. We analyze the hand-crafted features and give a solid benchmark of vascular-based RCC classification. The results show that our proposed hand-crafted features trained on the traditional machine learning approaches perform best among all the experiments.

5.1	Vascular-Based RCC Classification Benchmark Experiments . . .	73
5.1.1	Classification Pipeline from Vascular Network of RCC . . .	73
5.1.2	Experimental Environment	74
5.2	Deep Features Based Classification with Convolutional Neural Networks	75
5.2.1	Inputs of Raw Images	75
5.2.2	Inputs of Weak Vascular Network Mask	75
5.2.3	LeNet	75
5.2.4	AlexNet	75
5.2.5	VGGNet	78
5.2.6	ResNet	80
5.3	Graph-based Classification with Graph Convolutional Networks	81
5.3.1	A Variant of Graph Convolution Networks	81
5.3.2	Pooling Layer Based on Self-Attention Graph	82
5.4	Hand-crafted Features Based Classification with Traditional Machine Learning Algorithm	86
5.4.1	Hand-crafted Features Analysis	86
5.4.1.1	Evaluation Index	86
5.4.1.2	"Leave One Out" Cross-Validation	86
5.4.1.3	Non-parametric Tests	87
5.4.1.4	Kruskal-Wallis Test	87
5.4.1.5	Median Test	87
5.4.1.6	Jonckheere-Terpstra Test	89
5.4.2	Adaptive Boosting	89
5.4.3	Decision Tree	90
5.4.4	Gradient Boosting Tree	91
5.4.5	K-Nearest Neighbor	92
5.4.6	Logistic Regression	93
5.4.7	Random Forest	95
5.4.8	SVM RBF & SVM Sigmoid	95
5.4.9	ROC Plotting and AUC Calculation	99
5.5	Results	101
5.5.1	Results of Hand-crafted Features Analysis	101
5.5.2	Performance of Hand-crafted Features on Traditional Algorithms under Non-parameter Test	101
5.5.3	Results of Vascular-Based RCC Classification Benchmark	102
5.6	Conclusion	103

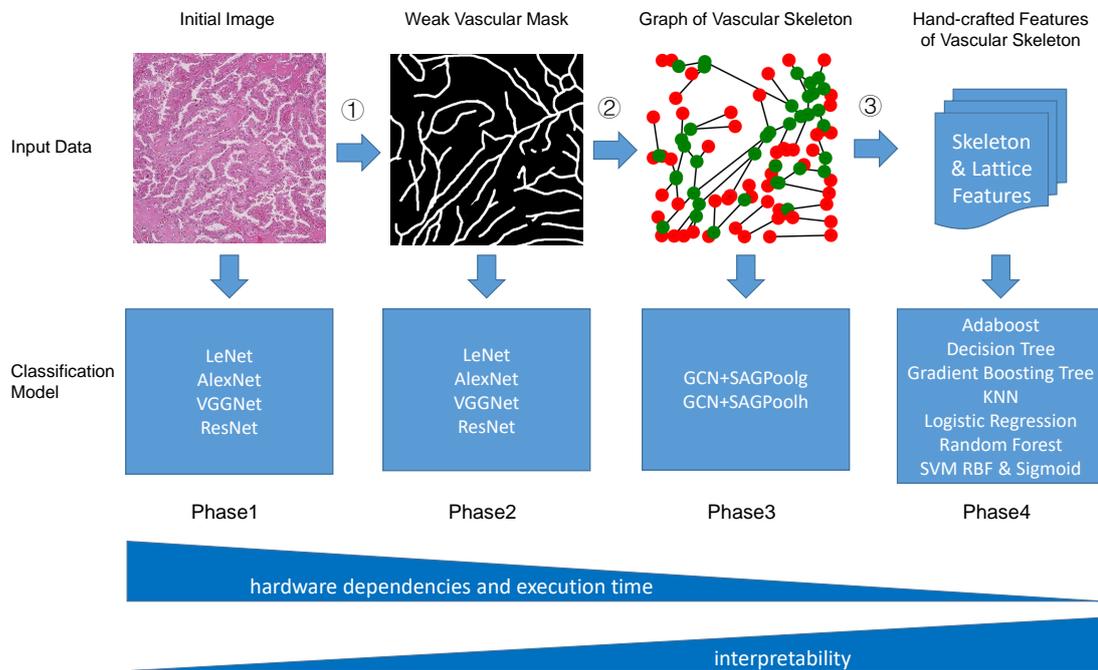


Figure 5.1 – The flow diagram of classification.

① : Manually vascular annotate; ② : Skeletonization weak vascular network mask and graph represent; ③ : Extract hand-crafted features from vascular network skeleton.

5.1 Vascular-Based RCC Classification Benchmark Experiments

5.1.1 Classification Pipeline from Vascular Network of RCC

Based on the introduction of input data in Chapter 4, we classify these four types of input data in this chapter through different classification models. The main process is shown in Figure 5.1. We mainly divide the experiments into four phases. In the first phase, we directly use the initial histopathology patch images as input to classify based on four popular deep learning models. In the second phase, we manually annotated vascular networks and also classified them based on four mainstream deep learning models according to manually annotated weak vascular network masks. Compared with the first phase, the second phase proved whether using blood vessels alone can be effective for RCC subtype classification. In the third phase, first, we transform the weak masks into vascular network skeletons, as shown in Figure 4.1 in Chapter 4; we represent the vascular network skeleton with a graph, where nodes are the set of junctions and end-points, and edges are the set of non-end points, which are then classified using two Graph Convolutional Networks (GCN) [Zhang et al., 2018a] deep learning models. In the fourth phase, we extract hand-crafted features from the vascular network skeleton, including Skeleton Features directly based on the skeleton and Lattice Features computed after the watershed; then we classify these hand-crafted features by eight common traditional machine learning methods.

Compared with deep learning, traditional machine learning methods have the following advantages :

- Reduce reliance on large amounts of data

The most important difference between traditional machine learning and deep learning is the performance when the scale of data is scaled up. Deep learning algorithms do not work well when the data is small because they need big data to recognize and understand it perfectly. However, machine learning algorithms can work in this small data situation.

- Reduce the dependence on hardware

Deep learning algorithms rely on high-computation machines because deep learning includes GPUs, which are an integral part of their work. Since deep learning operates through a large number of matrix multiplications, these operations can be efficiently optimized by using GPUs built for this purpose. In contrast, traditional machine learning algorithms can run on low-computing machines.

- Shorter execution time

Typically, training a deep learning algorithm takes a long time. This is because there are many parameters in deep learning algorithms. The training of machine learning will consume relatively little time, only a few seconds to a few hours.

- More interpretability

Explainability is also one of the factors that need to be considered in the application of deep learning. For example, if we use deep learning method for image classification, it will not reveal the reasons for the good classification. While it is possible to mathematically find out the nodes of a deep neural network that activate when classification is optimal, like a black box, we will not know the specific work of these neurons. This leaves us unable to interpret the results of deep learning models. However, machine learning algorithms provide us with a clear set of rules according to which features are selected, and the importance of individual features can also be calculated so that explaining the work behind traditional machine learning models becomes easy.

5.1.2 Experimental Environment

For the traditional machine learning experiment, the programming language used is Python, and the version is Python3.7.3. based on Windows 64-bit operating system. The hardware platform is Intel with CPU memory 64 GB (R) Core(TM) i7-8850H CPU @ 2.60GHz with 2 adapters : Intel with GPU Memory 32 GB (R) UHD Graphics 630 and NVIDIA with GPU Memory 32 GB Quadro P2000 Laptop.

For the deep learning experiment, the software used in this project is built with the deep learning framework of Pytorch, the programming language is Python, and the version is Python3.8.5. The operating system is CentOS Linux release 7.4.1708 (Core). The hardware platform is Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz with 2 physical CPUs and 40 logical CPUs, 10 cores per physical CPU, and 2 threads per core.

5.2 Deep Features Based Classification with Convolutional Neural Networks

We consider both raw images and vascular segmentation images as deep features, then learn these deep features with four popular Convolutional Neural Networks (CNN) models as the baseline experiments, such as LeNet [Wright, 1995], AlexNet [Krizhevsky et al., 2012a], VggNet [Simonyan and Zisserman, 2014] and ResNet [He et al., 2016]. The accuracy results of every model were got by Leave One Out cross-validation methods. All the models are based on the PyTorch frame (version=1.7.1) with Multi-GPU support. The programming language used is Python (3.8.5).

5.2.1 Inputs of Raw Images

The raw image of the tumor patch is a 2000×2000 pixel RGB image. As input to the deep learning model, it maintains an aspect ratio, resized to 256×256 pixels, 224×224 pixels, or 64×64 pixels depending on the particular model. We also augment the training set using data augmentation operations, such as cropping, rotation, flip and different normalization (mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225]).

5.2.2 Inputs of Weak Vascular Network Mask

A weak vascular network mask is obtained by annotating the vascular network with the constant width from the raw image. It is a 2000×2000 pixels binary image. In order to do a fair comparison with the raw image, it has also been resized and we considered the same augmentation and normalize operations to generate more training data as the raw image.

5.2.3 LeNet

LeNet is the collective name for a series of networks, including Lenet 1 - Lenet 5, proposed by Yann LeCun et al. in "Handwritten Digit Recognition with a Back-Propagation Network" [LeCun et al., 1989]. LeNet is a seven-layer neural network with three convolutional layers, two pooling layers, and one fully connected layer. All convolution kernels of all convolutional layers are 5×5 , stride=1, pooling method is global pooling, and activation function is Sigmoid. Compared with Multilayer Perceptron (MLP), LeNet uses relatively fewer parameters and achieves better results.

The network structure is shown in Figure 5.2 [LeCun et al., 1989].

Some setup used in our project :

- input size : 64×64 pixels
- loss_function : Cross Entropy Loss
- optimizer : Adam
- learning rate : 0.001

5.2.4 AlexNet

AlexNet was designed by 2012 ImageNet competition winner Hinton and his student Alex Krizhevsky [Krizhevsky et al., 2012a]. Compared with LeNet, AlexNet has more layers. LeNet has seven layers without the input layer. AlexNet has eight weighted layers, the first five are

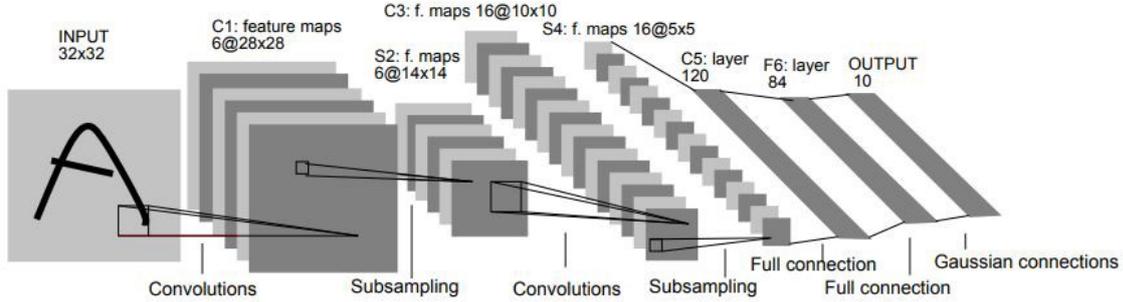


Figure 5.2 – The LeNet-5 architecture.

It consists of two sets of convolutional and average pooling layers, followed by a flattening convolutional layer, then two fully-connected layers and finally a softmax classifier. (LeCun, 1989)

convolutional layers, and the remaining three are fully connected layers. The first convolutional layer convolves an input image of size $224 \times 224 \times 3$ pixels with 96 kernels of size $11 \times 11 \times 3$ and stride=4. The second convolutional layer takes the output of the first convolutional layer as input and filters it with a $5 \times 5 \times 48$ kernel. The third, fourth, and fifth convolutional layers are connected without a pooling layer in between. The kernels of the second, fourth, and fifth convolutional layers are only connected to those kernel maps where the previous convolutional layer is also in the same GPU. The kernels of the third convolutional layer are connected to all kernel maps in the second convolutional layer. Neurons in a fully connected layer are connected to all neurons in the previous layer. A response normalization layer follows the first and second convolutional layers. The max pooling layer follows the response normalization layer and the fifth convolutional layer. ReLU nonlinearity is applied to the output of each convolutional and fully connected layer.

The network structure is shown in Figure 5.3 [Krizhevsky et al., 2012a].

The highlights of the network are :

- The first use of GPU for network acceleration training put half the cores on each GPU.
- The ReLU activation function is used instead of the traditional Sigmoid activation function and Tanh activation function. (The sigmoid derivation is more complex and the gradient disappears when the network is deep.)
- Local Response Normalization (LRN) is used.

LRN helps generalization, $a_{x,y}^i$ represents neuron activation, calculated by applying kernel i at (x, y) position, and then applying ReLU nonlinearity. The normalized activation of the Response Normalization $b_{x,y}^i$ is given by

$$b_{x,y}^i = a_{x,y}^i / \left(k + \alpha \sum_{j=\max(0, i-n/2)}^{\min(N-1, i+n/2)} (a_{x,y}^j)^2 \right)^\beta$$

where N is the number of convolution kernels, that is, the number of Feature Maps generated; k, α, β, n are hyperparameters, and the values used in the paper are $k = 2, n = 5, \alpha = 10^{-4}, \beta = 0.75$. The superscript of the output $b_{x,y}^i$ and input $a_{x,y}^j$ indicates the channel where the current value is located, that is, the direction of the superposition is along the channel, and the value to be normalized $a_{x,y}^i$ is in the same position as the nearby channels

The squares of the values are accumulated, $\sum_{j=\max(0, i-n/2)}^{\min(N-1, i+n/2)} (a_{x,y}^j)^2$

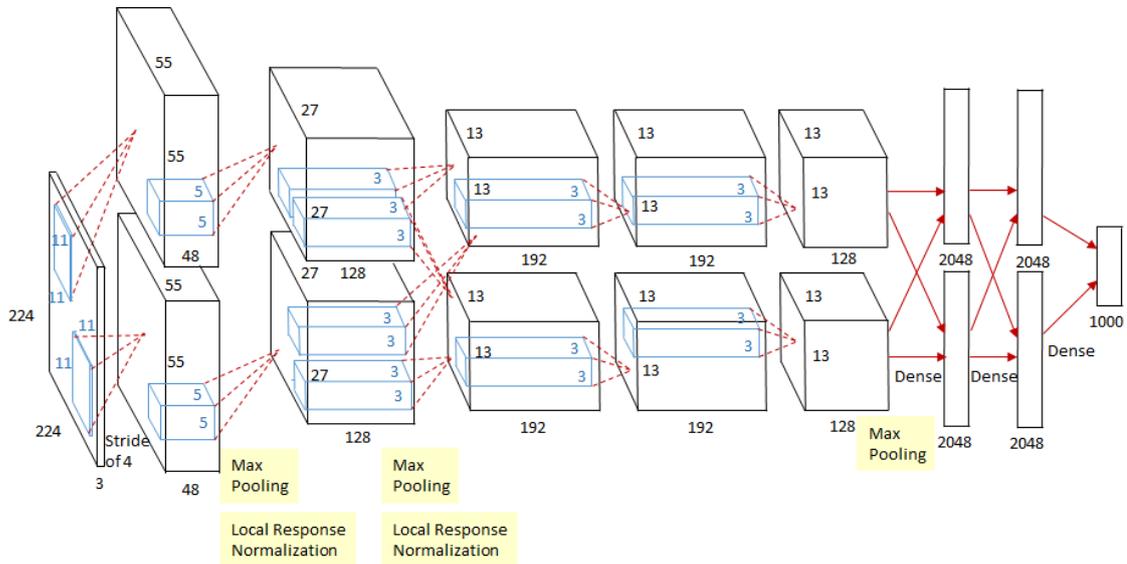


Figure 5.3 – An illustration of the architecture of AlexNet.

It explicitly shows the delineation of responsibilities between the two GPUs. One GPU runs the layer parts at the top of the figure while the other runs the layer parts at the bottom. The GPUs communicate only at certain layers. The network's input is 150,528-dimensional, and the number of neurons in the network's remaining layers is given by 253,440–186,624–64,896–64,896–43,264–4096–4096–1000. (Krizhevsky, 2012)

- The first two layers of the fully connected layer use Dropout to randomly deactivate neurons to reduce overfitting.
- Use overlapping max pooling in CNN. Previously, average pooling was commonly used in CNN, and AlexNet all used maximum pooling to avoid the blurring effect of average pooling. It is also proposed to make the stride smaller than the size of the pooling kernel so that there will be overlap between the outputs of the pooling layer, which improves the abundance of features.
- Data augmentation, randomly intercepting a 224×224 size area (and a horizontally flipped mirror) from the 256×256 pixels original image, which is equivalent to increasing the amount of data by $2 \cdot (256 - 224)^2 = 2048$ times. If there is no data augmentation, CNN with many parameters will fall into overfitting only with the original data volume. After data augmentation is used, overfitting can be greatly reduced and the generalization ability can be improved. When making predictions, take the four corners of the picture plus the middle five positions, and flip them left and right to obtain a total of ten pictures, make predictions on them, and average the ten results. At the same time, the AlexNet paper mentioned that the RGB data of the image will be processed by PCA, and a Gaussian perturbation with a standard deviation of 0.1 is performed on the principal component, this trick can reduce the error rate by another 1% by adding some noise.
- Cascade pooling.
Pooling is non-overlapping in LeNet, i.e. the pooled window size and stride are equal. While The pooling used in AlexNet is overlapping, that is, the step size of each move is smaller than the pooling window length when pooling. The AlexNet pools are 3×3

squares, and each pool moves with the stride=2, so there is overlap. Overfitting can be avoided by this overlapping pooling way, which contributes 0.3% to the Top-5 error rate. Compared with the non-overlapping scheme $s=2$, $z=2$, the dimensions of the outputs are equal, and overfitting can be suppressed to a certain extent.

Some setup used in our project :

- input size : 224×224 pixels
- loss_function : Cross Entropy Loss
- optimizer : Adam
- learning rate : 0.0002

5.2.5 VGGNet

VGGNet [Simonyan and Zisserman, 2014] was jointly developed by researchers from Oxford University's Visual Geometry Group and Google's DeepMind team, and won second place in ILSVRC 2014 (2014 ImageNet Image Classification Competition), reducing the Top-5 error rate. To 7.3%, it achieved 92.3% accuracy in Top-5.

The increased depth of the convolutional neural network and the use of small convolution kernels have a great impact on the final classification and recognition effect of the network. VGGNet explores the relationship between the depth and performance of a CNN. By constantly stacking 33 small convolution kernels and 22 max-pooling layers, VGGNet successfully builds a deep CNN with 16-19 layers.

VGG-16 has 13 convolutional layers with a kernel size of 3×3 , 5 max-pooling layers and 3 fully connected layers. The convolution layers are all convolution operations with a size of 3×3 , a stride of 1, and a padding of 1 (the height and width of the feature matrix will not be changed after convolution). The maximum pooling downsampling layers are all pooling operations with a pooling kernel size of 2 and a stride of 2. After each downsampling by maximum pooling, the height and width of the feature matrix will be reduced to half of the original.

The structure diagram of VGG-16 is shown in Figure 5.4 [Simonyan and Zisserman, 2014].

Advantages of VGGNet-16 :

- Simple structure.

The structure of VGGNet is quite simple. It consists of five convolutional layers, three fully connected layers and one softmax layer. The maximum pooling connection is used between layers, and the activation functions used between hidden layers are all ReLU.

- Use small convolution kernels.

VGGNet uses convolutional layers with multiple small 3×3 convolution kernels instead of larger convolutional layers in AlexNet. The receptive field of stacking two 3×3 convolution kernels is equivalent to the receptive field of a 5×5 convolution kernel, and the receptive field of stacking three 3×3 convolution kernels is equivalent to the receptive field of a 7×7 convolution kernel. Therefore, using multiple small convolution kernels can not only reduce the number of parameters but also enhance the nonlinear mapping of the network to improve its expressiveness of the network. Using multiple 3×3 convolution stacks has two effects : one is to reduce the parameters without affecting the receptive field ; the other is to increase the nonlinearity of the network.

- Use small filters.

Compared with AlexNet, VGGNet all use 2×2 small filters in the pooling layer, and the stride is 2.

5.2.6 ResNet

The ResNet [He et al., 2016] was proposed by He Kaiming in Microsoft Labs and won first place in the classification task and first place in the target detection in the ImageNet competition that year. It won first place in object detection and image segmentation in the COCO dataset.

The ResNet network refers to the VGG19 network, modified on the basis, and added the residual unit through the short-circuit mechanism, as shown in Figure 5.5. The change is mainly reflected in the fact that ResNet directly uses the convolution of stride=2 for downsampling, and replaces the fully connected layer with the global average pool layer. An important design principle of ResNet is that when the feature map size is reduced by half, the number of feature maps is doubled, which maintains the complexity of the network layers.

For ResNet-18, 18 represents the depth of the network, specifying 18 layers with weights, including convolutional layers and fully connected layers, excluding pooling layers and BN layers.

The ResNet-18 network structure is shown in Figure 5.5 [Ramzan et al., 2020].

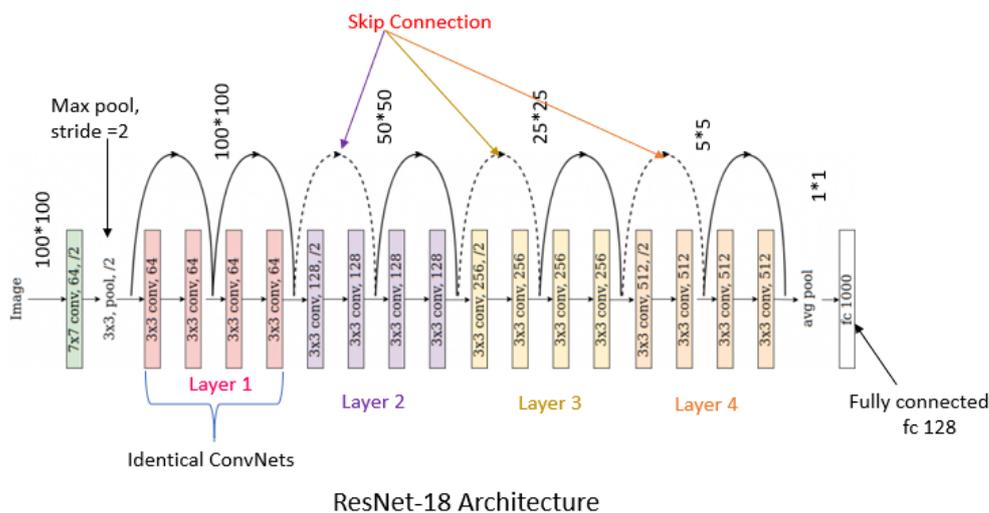


Figure 5.5 – The original network structure of ResNet-18. (Ramzan, 2020)

The highlights in the ResNet-18 Network :

- Super deep network structure (more than 1000 layers).
- Propose a residual (residual structure) module to alleviate the degradation problem in deep networks.

The residual structure uses a shortcut connection, which can also be understood as a short-cut. Let the feature matrix be added at intervals. The structure of the residual is shown in Figure 5.6. Note that $F(X)$ and X have the same shape. The addition is the addition of the numbers at the same position of the feature matrix. There are two different residuals in ResNet, BasicBlock and Bottleneck.

- Short cut for dimensionality reduction.

Looking at the ResNet18 layer network in figure 5.5, one can find that the shortcuts of some residual blocks are solid lines, while others are dotted lines. The short cuts of these dotted lines are dimensioned by a 1×1 convolution kernel (the feature matrix is downsampled in

the length and width directions, and the depth direction is adjusted to the channel required by the next layer of residual structure).

- Use Batch Normalization to speed up training (drop dropout) by solving the vanishing or exploding gradient problem. Batch Normalization satisfies the distribution law of data feature maps with a mean of 0 and a variance of 1.

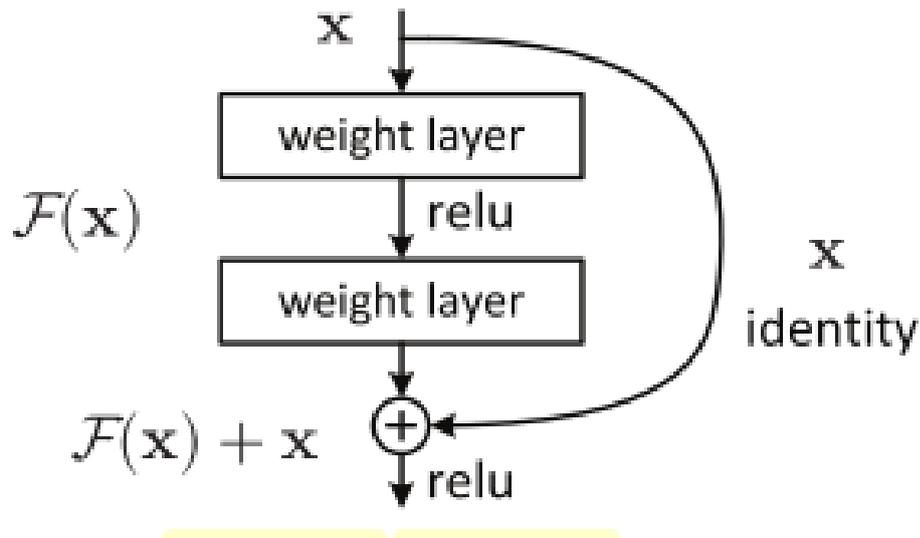


Figure 5.6 – The structure of residual. (He, 2016)

Some setup used in our project :

- input size : 256×256 pixels
- loss_function : Cross Entropy Loss
- optimizer : Adam
- learning rate : 0.0001

5.3 Graph-based Classification with Graph Convolutional Networks

In this part, we use the adjacency matrix of every graph, which represents the morphology of the vascular network from tumor patch images. The GCN [Zhang et al., 2018a] models were developed by PyTorch deep learning frame (version=1.6.0), and the programming language is Python (3.8.5).

5.3.1 A Variant of Graph Convolution Networks

GCN [Zhang et al., 2018a] designs a method to extract features from graph data so that we can use these features to perform node classification, graph classification and get the graph embedding by the way.

The graph convolution operation is defined as follows :

$$x_1 * x_2 = IGFT (GFT (x_1) \odot GFT (x_2))$$

The inverse Fourier transform IGFT is defined :

$$x_k = \sum_{i=1}^N V_{ki} \tilde{x}_i$$

Where $IGFT$ is The Inverse Graph Fourier Transform. \odot represents the Hadamard product of the Graph Fourier Transform (GFT). The GFT refers to the process of projecting the graph signal onto each feature vector (Fourier basis) to obtain a set of Fourier coefficients. Substituting $GFT(\mathbf{x}) = \tilde{\mathbf{x}} = V^T \mathbf{x}$ into the above formula, we get

$$\begin{aligned} x_1 * x_2 &= V \left(\left(V^T \mathbf{x}_1 \right) \odot \left(V^T \mathbf{x}_2 \right) \right) = V \left(\tilde{\mathbf{x}}_1 \odot \left(V^T \mathbf{x}_2 \right) \right) \\ &= V \left(\text{diag}(\tilde{\mathbf{x}}_1) \left(V^T \mathbf{x}_2 \right) \right) \\ &= \left(V \text{diag}(\tilde{\mathbf{x}}_1) V^T \right) \mathbf{x}_2 \end{aligned}$$

Then we can define the following neural network layers to parameterize the frequency response matrix :

$$\begin{aligned} X' &= \sigma \left(V \begin{bmatrix} \theta_1 & & & \\ & \theta_2 & & \\ & & \ddots & \\ & & & \theta_N \end{bmatrix} V^T X \right) \\ &= \sigma \left(V \text{diag}(\theta) V^T X \right) \\ &= \sigma(\Theta X) \end{aligned}$$

Where $\sigma(\cdot)$ is the activation function, $\theta = [\theta_1, \theta_2, \dots, \theta_N]$ is the parameter to be learned, Θ is the graph filter we need to learn, X is the input graph signal matrix, X' is the output graph signal matrix.

From a spatial perspective, the defined neural network layer introduces an adaptive graph displacement operator; from a frequency domain perspective, the neural network layer trains an adaptive graph filter between X and X' . The frequency response function can be used for supervised learning through the correspondence between tasks and data.

Define \tilde{L}_{sym} as a Laplace matrix in the renormalized form, the eigenvalues of \tilde{L}_{sym} range from $(-1, 1]$, $\tilde{L}_{sym} = \tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2}$, where $\tilde{A} = A + I$, $\tilde{D}_{ij} = \sum_j \tilde{A}_{ij}$. I is the identity matrix, which is added to A to make the diagonal elements 1. At the same time, a parameterized weight matrix W is defined to perform the affine transformation on the input image signal matrix. The neural network model with multiple layers as the main body is called GCN :

$$X' = \sigma \left(\tilde{L}_{sym} X W \right)$$

5.3.2 Pooling Layer Based on Self-Attention Graph

Self-Attention Graph Pooling (SAGPool) [Lee et al., 2019] uses a Graph Neural Network (GNN) to learn the importance of nodes. This GNN-based approach makes better use of the structural information of the graph to learn the importance of nodes. Based on the attention score calculated by GCN, the pooling operation is performed on each graph. To visualize the pooling process, we pool each graph separately and finally cascade them for the next calculation. In this project, we use the TopK-based pooling mechanism [Gao and Ji, 2019].

The pooling mechanism based on TopK is a process of continuously discarding nodes, which captures the information of different scales of the graph. Unlike the local sliding window-based pooling operation in CNN, TopK pooling places the scope on the full-graph nodes. Specifically, first set a hyperparameter $k, k \in (0, 1)$ that represents the pooling rate, then learn a value z representing the importance of nodes and sort them in descending order, and then put N nodes in the whole graph downsampled to kN nodes. The formula is expressed as follows :

$$\begin{aligned}\vec{i} &= \text{TopK}(z, kN) \\ X' &= X_{i,:} \\ A' &= A_{i,j}\end{aligned}$$

Where TopK is a sequence number that returns the first kN data, $\text{TopK}(z, kN)$ refers to sorting by the z index and selecting the first kN nodes. X_i indicates that the feature matrix is row-sliced according to the value of the vector \vec{i} , and $A_{i,j}$ indicates that the adjacency matrix is simultaneously row-sliced and column-sliced according to the value of the vector \vec{i} .

The pooling mechanism based on TopK only needs to discard $(1-k)N$ nodes from the original graph each time. The model structure of TopK given in the paper [Cangea et al., 2018] is shown in Figure 5.7 :

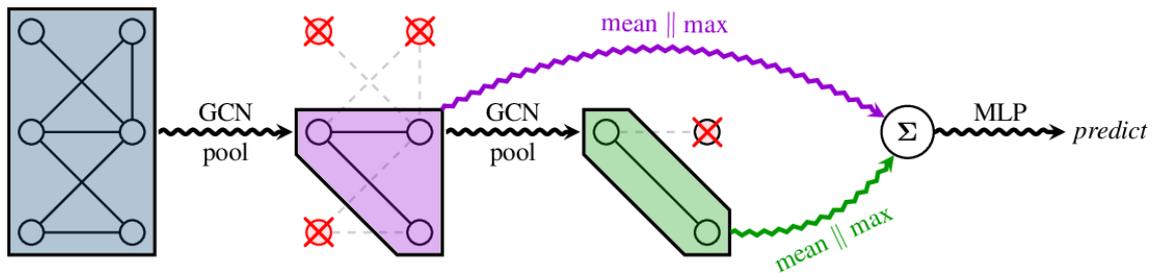


Figure 5.7 – The full pipeline of the TopK model (for $k = 0 : 5$), leveraging several stacks of interleaved convolutional/pooling layers. (Cangea, 2018)

Next, define two sets of SAGPool models, as shown in Figure 5.8. Among them, figure A uses only one pooling layer, called SAGPool_g [Zhang et al., 2018a], "g" stands for global; figure B uses multiple pooling layers, called SAGPool_h [Cangea et al., 2018], "h" stands for hierarchical. SAGPool_g has 3 layers of convolutional layers, after the pooling layer is aggregated by the readout layer, the vector of the entire graph can be obtained and connected to the MLP for the classification task. SAGPool_h has a multi-layer structure, which is shown on the right side of Figure 5.8. There are a total of 3 blocks, the output of each block is subjected to the summation operation after the readout layer and finally passed to the MLP for classification.

The gpool layer adopts the method of discarding nodes layer by layer to improve the fusion efficiency of distant nodes, but this method will make it lack the means of effective information fusion for all nodes. Therefore, in order to achieve effective fusion, a readout layer is added after the gpool layer to achieve a one-time aggregation of the global information of the graph at this scale. The specific implementation of the readout layer is to combine global average pooling and global max pooling, and the final global pooling is the sum of the pooling results of each layer.

The readout mechanism involved in the graph classification based on global pooling performs a one-time aggregation operation on all nodes after K rounds of iterations. Thus the global repre-

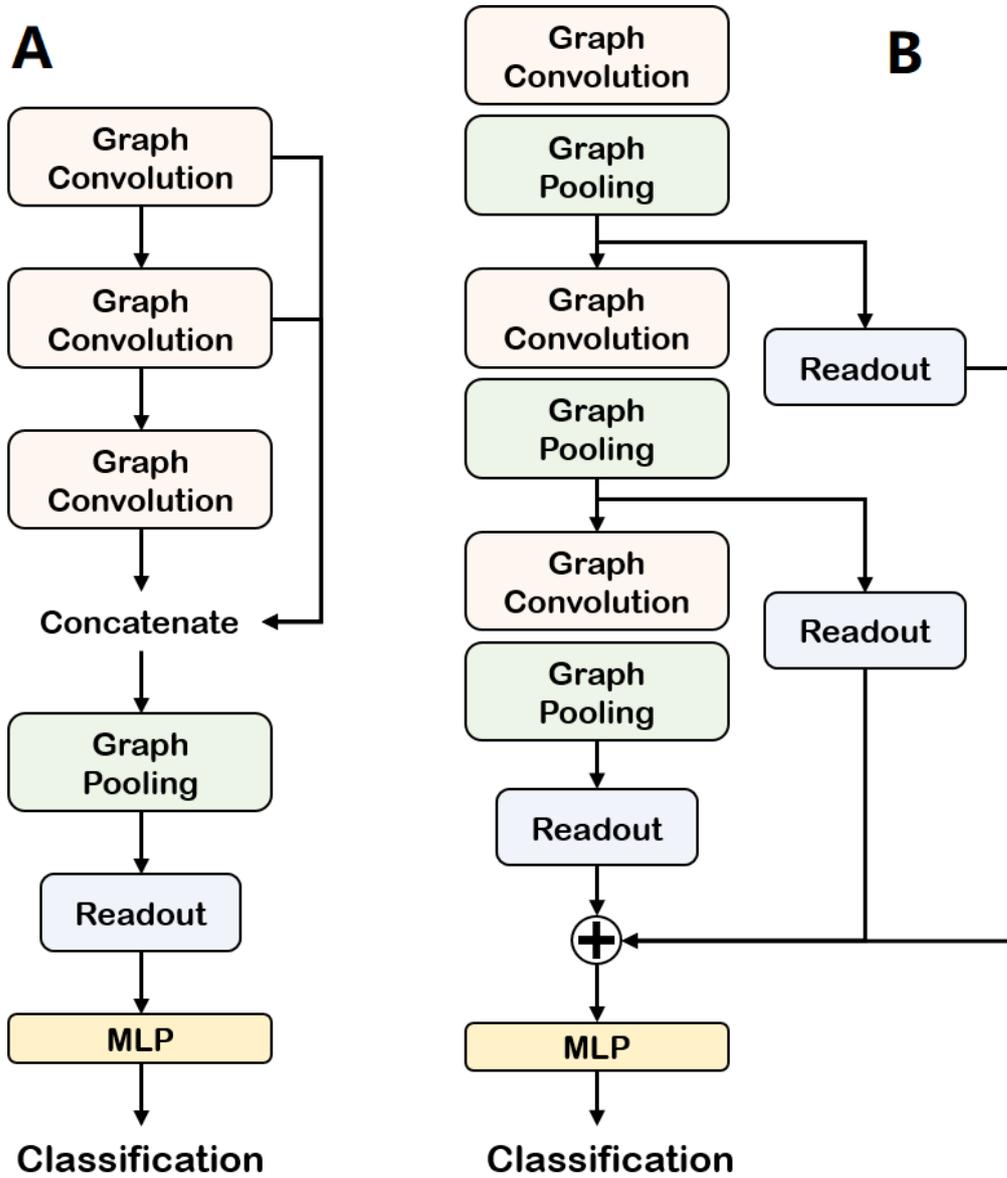


Figure 5.8 – The global pooling architecture (A) and the hierarchical pooling architecture (B). These architectures are applied to all the baselines and SAGPool for a fair comparison. In this paper, the architecture on the A side is referred to as POOLg and the architecture on the B side is referred to as POOLh with the POOL method (e.g. SAGPoolg, gPoolh). (Lee, 2019)

sentation of the output graph :

$$h_G = \mathbf{R} \left(h_v^{(k)} \mid v \in G \right)$$

Where the readout function \mathbf{R} is to perform graph-level predictions. $h_v^{(k)}$ is the feature representation of node v at the k -th iteration. The feature vector h_G is calculated for the entire graph (G) by the readout function \mathbf{R} .

The model summary of SAGPool_g and SAGPool_h in our project is shown in Figure 5.9, and the accuracy result of the two models are shown in Table 5.3.

A

```
print("Device:", DEVICE)
print(model)

Device: cuda
ModelA(
  (gcn1): GraphConvolution (89 -> 32)
  (gcn2): GraphConvolution (32 -> 32)
  (gcn3): GraphConvolution (32 -> 32)
  (pool): SelfAttentionPooling(
    (attn_gcn): GraphConvolution (96 -> 1)
  )
  (fc1): Linear(in_features=192, out_features=32, bias=True)
  (fc2): Linear(in_features=32, out_features=16, bias=True)
  (fc3): Linear(in_features=16, out_features=2, bias=True)
)
```

B

```
print("Device:", DEVICE)
print(model)

Device: cuda
ModelB(
  (gcn1): GraphConvolution (89 -> 32)
  (pool1): SelfAttentionPooling(
    (attn_gcn): GraphConvolution (32 -> 1)
  )
  (gcn2): GraphConvolution (32 -> 32)
  (pool2): SelfAttentionPooling(
    (attn_gcn): GraphConvolution (32 -> 1)
  )
  (gcn3): GraphConvolution (32 -> 32)
  (pool3): SelfAttentionPooling(
    (attn_gcn): GraphConvolution (32 -> 1)
  )
  (mlp): Sequential(
    (0): Linear(in_features=64, out_features=32, bias=True)
    (1): ReLU()
    (2): Linear(in_features=32, out_features=16, bias=True)
    (3): ReLU()
    (4): Linear(in_features=16, out_features=2, bias=True)
  )
)
```

Figure 5.9 – The model summary of SAGPool_g (A) and SAGPool_h (B) in our project.

The hyperparameter settings in our project are :

- INPUT_DIM = node_features.size(1) = 4 (int, the dimension of the input feature)
- NUM_CLASSES = 3 (int, number of categorical categories)
- EPOCHS = 30 (int, one epoch is equal to training once using all samples in the training set)
- HIDDEN_DIM = 32 (int, number of hidden layer units)
- LEARNING_RATE = 0.01 (Controls the learning progress of the model, determines whether the objective function can converge to the local minimum and when it converges to the minimum)
- WEIGHT_DECAY = 0.0001 (Weight decay is equivalent to $L2$ norm regularization. Regularization makes the weights smaller by adding a penalty term to the loss function of the model.)

5.4 Hand-crafted Features Based Classification with Traditional Machine Learning Algorithm

In this part, we input the sixteen feature data filtered by three nonparametric tests into eight traditional supervised machine learning models. The programming language used in this part is python (Version 3.7.3). First, we use the GridSearchCV function of 'sklearn.model_selection' to determine the best parameters for each model to fit our data. We train the model with the best parameters and test the model by leave-one-out cross-validation method to get the accuracy (as shown in Table 5.3), the area under the curve (AUC), and plot the Receiver Operating Characteristic (ROC) curve (as shown in Figure 5.12) of each model.

5.4.1 Hand-crafted Features Analysis

5.4.1.1 Evaluation Index

Our vascular network annotation data includes 424 patch image data from 31 patients, we consider a majority vote to translate the results at patch image level to patient level :

Predict labels at the patient level :

$$C = \operatorname{argmax}_j \{ \#_j \} \quad (5.1)$$

C : patient class (ccRCC, pRCC and chRCC)

$\#_j$: number of patch images that have been classified into class j

It means that if the predicted label of more than half of the patch images of the patient is class j , the class of the patient is considered to belong to class j , otherwise it belongs to class i .

Normalized confusion matrices at the patient level :

A confusion matrix A (also known as an error matrix [Townsend, 1971]) is a special type of contingency table with two dimensions (actual and predicted), and the set of classes in both dimensions.

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \quad (5.2)$$

Normalized confusion matrix B :

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} \quad (5.3)$$

$a_{i,j}$: number of patients of class i classified into class j

$b_{i,j}$: the probability of samples of class i to be classified into class j : $b_{i,j} = \frac{a_{i,j}}{\sum_k a_{i,k}}$

5.4.1.2 "Leave One Out" Cross-Validation

Leave-One-Out Cross-Validation (LOO-CV) is a common cross-validation method. First of all, the k-fold cross-validation is a very common machine learning method, that is, the data set is randomly divided into k parts, of which the training set has (k-1) parts, and the test set accounts for one part. For example, the first piece of data can be used to test the fit of the model based on the remaining (k-1) pieces of data, then the second piece of data can be used as the test set to test the fit

of the corresponding training set, and so on. After completion, each piece of data is used as a test set exactly once, and the fitness of the model can be obtained by averaging the fitness of the k tests. This approach is used to calibrate the parameters and has the advantage of avoiding over-fitting. In this project, there are 31 patients with the VRCC dataset, we take 30 patients for training and 1 for testing each time. In this project, we use the LeaveOneOut function of 'sklearn.model_selection' to implement the cross-validation method.

5.4.1.3 Non-parametric Tests

To study the relevance of the proposed features to discriminate the different cancer patch images, we apply statistical tests between the three populations : ccRCC, pRCC and chRCC.

Since the data of the proposed features are not normally distributed, as shown in Figure 5.10, we use three common non-parametric tests methods [Hollander et al., 2013] (Kruskal-Wallis Test, Median Test and Jonckheere-Terpstra Test) to calculate the statistical significance between the ccRCC, pRCC and chRCC features.

The statistical software SPSS, version 20.0, was used for all the statistical analyses. All the statistical tests were 2-sided. The level for statistical significance was set $\alpha = 0.05$.

5.4.1.4 Kruskal-Wallis Test

The idea of the Kruskal-Wallis test is to mix n groups of samples into a data set (that is, assuming they are from the same distribution), and then sort the merged sample values from small to large ; replace the sorted values with ranks, starting with the smallest value of 1. Then find the average rank of each group. If the n groups of data come from the same distribution, the rank of each group should be similar to the total average rank of the mixed data. If the difference is large, it means that each group is not from the same population.

The null hypothesis : The distributions of multiple populations from which multiple independent samples are drawn are not significantly different.

Inspection Steps :

- Suppose there are m independent simple random samples (X_1, \dots, X_{ni}) ($i=1, \dots, m$).
- Mix sets of sample numbers and sorts them in ascending order.
- Let R_i ($i=1, \dots, m$) represent the sum of the ranks of the n_i observations X_1, \dots, X_{ni} of the i -th sample in this arrangement.
- Calculate statistics $H = \frac{12}{N(N+1)} \sum_{i=1}^m \frac{R_i^2}{n_i} - 3(N+1)$
(When N is sufficiently large, H approximately obeys the chi-square distribution, with degrees of freedom $df = m - 1$)
- Find the upper quantile $\chi_{\alpha, df}^2$ of the chi-square distribution for a given significance level α and degrees of freedom $df = m - 1$.
- When $H \geq \chi_{\alpha, df}^2$, the m group samples are not all from the same population.

5.4.1.5 Median Test

The Mood's Median Test tests the null hypothesis that the population medians from which the samples come are not different. The alternative hypothesis suggests that there are differences in the medians of the populations from which the samples are drawn.

Once the median of the combined data is determined, the sample values are then put into one of two groups, one with all values above the median, and the other with all values at or below the

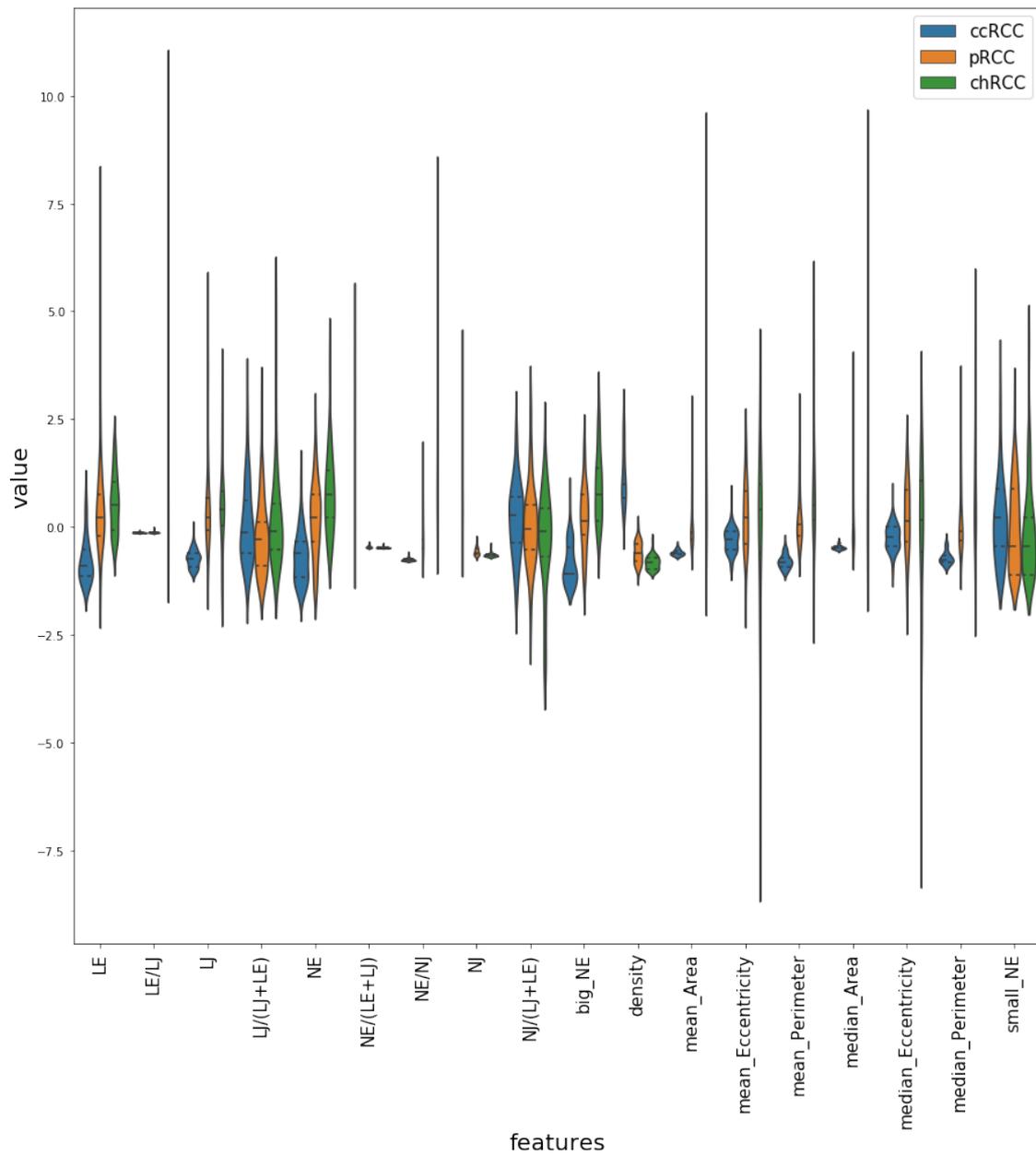


Figure 5.10 – The distribution of data from hand-crafted features.

median. A chi-square test was used to compare the actual frequencies of the two groups with the expected frequencies of no difference. The p-value of the test indicates whether the null hypothesis should be rejected.

While the Mood's Median test does not make any assumptions about the distribution of the data, there are still other assumptions that should be satisfied. They are :

- The sample values are independent both within and between samples
- The sample values come from populations with a continuous rather than discrete distribution
- The distributions of the populations from which the samples were drawn all have the same shape

5.4.1.6 Jonckheere-Terpstra Test

The Jonckheere-Terpstra test was independently proposed by Terpstra (1952) and Jonckheere (1954).

Assume that k independent samples (population) come from continuous distribution functions with the same shape, and their position parameters (such as median) are recorded as $\theta_1, \theta_2, \dots, \theta_k$. Assume that the sample size of k samples (population) is $n_i, i = 1, 2, \dots, k$. Let x_{ij} be the j -th independent observation from the i -th sample (population) ($i = 1, 2, \dots, k, j = 1, 2, \dots, n_i$). Then the observed value x_{ij} can be written as the following linear model : $x_{ij} = \mu + \theta_i + \varepsilon_{ij}, i = 1, 2, \dots, k, j = 1, 2, \dots, n_i$, where the errors are independent and identically distributed.

If the number of observations in one sample (population) is smaller than that of another sample (population), it can be considered that there is a size relationship between the locations of multiple samples (population).

Inspection Steps :

- calculate : $U_{ij} =$ The logarithm of the sample (population) i is smaller than the observations in the sample (population) $j = \#(X_{ik} < X_{jl}, k = 1, \dots, n_i, l = 1, \dots, n_j)$.
- Sum all U_{ij} in the range of $i < j$ to get the Jonckheere-Terpstra test statistic : $J = \sum_{i < j} U_{ij}$. When J is large, the null hypothesis should be rejected.
- The critical value c under the null hypothesis can be obtained from (n_1, n_2, n_3) and the test level α by looking up the critical value table, which satisfies $P(J \geq c) = \alpha$.

When encountering a large sample situation, the Jonckheere-Terpstra test statistic can be approximated normally as $Z = \frac{J - E(J)}{\sqrt{\text{Var}(J)}}$. That is, under the condition that the null hypothesis holds, when $\min_i \{n_i\} \rightarrow \infty$, and $\lim_{n_i \rightarrow \infty} \frac{n_i}{\sum_{i=1}^k n_i} = \lambda_i \in (0, 1)$, then $Z =$

$$\frac{J - \left(N^2 - \sum_{i=1}^k n_i^2 \right) / 4}{\sqrt{\left[N^2(2N+3) - \sum_{i=1}^k n_i^2(2n_i+3) \right] / 72}} \rightarrow N(0, 1).$$

5.4.2 Adaptive Boosting

Adaptive Boosting (Adaboost) was proposed by Yoav Freund and Robert Schapire [Freund and Schapire, 1997]. Its adaptation lies in that the samples misclassified by the previous basic classifier are strengthened (that is, get a higher weight), and the weighted total samples are used again to train the next basic classifier. At the same time, a new weak classifier is added in

each round until a predetermined small enough error rate or a pre-specified maximum number of iterations is reached, and the algorithm stops.

The Adaboost iterative algorithm consists of three steps :

- Initialize the weights for the training data.

If there are N samples, each training sample is initially given the same weight : $1/N$.

- Train weak classifiers.

In the specific training process, if a sample point has been accurately classified, its weight will be reduced when constructing the next training set; on the contrary, if a sample point has not been accurately classified, its weight will be reduced. Then, the next classifier is trained using the sample set of updated weights, and the whole training process is iterative.

- Combine the weak classifiers obtained from each training into a strong classifier.

After the training process of each weak classifier is completed, we increase the weight of the weak classifier with a small classification error rate, so that it plays a larger decisive role in the final classification function, and reduce the weight of the weak classifier with a large classification error rate. In other words, a weak classifier with a low error rate has a larger weight in the final classifier, otherwise, it is smaller.

The classifier with the best model parameter in our project is : AdaBoostClassifier(algorithm='SAMME.R', base_estimator=None, learning_rate=1.0, n_estimators=3).

5.4.3 Decision Tree

Decision Tree [Breiman et al., 1984] is a non-parametric supervised learning method that can summarize decision rules from a series of data with features and labels, and present these rules with a tree diagram structure to solve classification and regression problems. Decision trees can solve multi-classification problems and have very good interpretability.

The generation algorithms of decision trees include ID3, C4.5 and C5.0. A decision tree is a tree structure in which each internal node represents a judgment on an attribute, each branch represents the output of a judgment result, and finally, each leaf node represents a classification result.

ID3 : It is determined by the principle of Entropy which is the parent node and which node needs to be split. For a set of data, the smaller the entropy, the better the classification result. Entropy is defined as follows :

$$\text{Entropy} = -\sum [P(x_i) * \log_2(P(x_i))]$$

Where $P(x_i)$ is the probability of x_i appearing. Therefore, when Entropy tends to 1, it corresponds to the state with the worst classification effect, and when it goes to 0, it corresponds to a state of complete classification. The minimization of entropy is the process of improving classification accuracy.

C4.5 is developed by [Salzberg, 1994]. C4.5 is an extension of Quinlan's earlier ID3 algorithm by the gain ratio. The decision trees can be produced by C4.5, which are used for classification and are often referred to as statistical classifiers.

The C5.0 algorithm [Quinlan, 2004] is a commercial improved version proposed by Quinlan based on the C4.5 algorithm. The purpose is to analyze the data set containing a large amount of data. Compared with the C4.5 algorithm, the C5.0 algorithm has the following advantages :

- The construction time of the decision tree is faster than the C4.5 algorithm, and the resulting decision tree is also smaller in size and has fewer leaf nodes.

- The boosting method is used to combine multiple decision trees to make classification, which improves the accuracy greatly.
- Provide optional parameters, such as whether to consider the weight of the sample, the cost of sample misclassification, etc.

Therefore, the generation of the decision tree is mainly divided into the following two steps, which are usually realized by learning samples for which the classification results are known.

- Node splitting : Generally, when the attribute represented by a node cannot be judged, it is selected to divide the node into two sub-nodes (if it is not a binary tree, it will be divided into n sub-nodes)
- Determination of the threshold : select the appropriate threshold to minimize the classification error rate (Training Error).

The classifier with best model parameter in our project is : `DecisionTreeClassifier(class_weight=None, criterion='gini', max_depth=2, max_features=None, max_leaf_nodes=None, min_impurity_split=1e-07, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, presort=False, splitter='best')`.

5.4.4 Gradient Boosting Tree

The Gradient Boosting Tree algorithm [Friedman, 2001] was originally proposed by Friedman in 2001. It is an ensemble algorithm (or a combination algorithm). Its base learner is a decision tree, which can be used for both the training regression and classification model. The algorithm has three major advantages : it handles mixed types of data naturally, has a strong predictive ability, and is robust to outliers in the output space (through a powerful loss function).

The core of the Gradient Boosting Tree algorithm is that each tree learns from the residuals of all trees in the previous training, and then fits into a tree. During training, the residual is approximately equal to the negative gradient value of the loss function in the current model.

Each calculation of gradient boosting can reduce the previous residual, and then a new model can be established in the gradient direction of the residual reduction. Therefore, in the gradient boosting algorithm, the establishment of each new model is to reduce the residual of the previous model in the direction of the gradient, which is different from the traditional boosting algorithm that assigns different weights to positive and false samples.

The Gradient Boosting Tree algorithm consists of the following steps :

- Initialize the weak learner $f_0(x) = \arg \min_c \sum_{i=1}^N L(y_i, c)$
- For $m = 1, 2, \dots, M$ have :
 - Calculate the negative gradient for each sample $i = 1, 2, \dots, N$, that is, the residual

$$r_{mi} = - \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{m-1}(x)}$$

- Use the residual obtained in the previous step as the new real value of the sample, and use the data $(x_i, r_{mi}), i = 1, 2, \dots, N$ as the training data of the next tree to obtain a new regression tree $f_m(x)$, whose corresponding leaf node area is $R_{jm}, j = 1, 2, \dots, J$, where J is the number of leaf nodes of the regression tree.
- Calculate the best fit value for the leaf region $j = 1, 2, \dots, J$

$$\Upsilon_{jm} = \underbrace{\arg \min}_{\Upsilon} \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}(x_i) + \Upsilon)$$

- Update the strong learner

$$f_m(x) = f_{m-1}(x) + \sum_{j=1}^J \Upsilon_{jm} I(x \in R_{jm})$$

- Get the final learner

$$f(x) = f_M(x) = f_0(x) + \sum_{m=1}^M \sum_{j=1}^J \Upsilon_{jm} I(x \in R_{jm})$$

The classifier with best model parameter in our project is : GradientBoostingClassifier(criterion='friedman_mse', init=None, learning_rate=0.1, loss='deviance', max_depth=3, max_features=None, max_leaf_nodes=None, min_impurity_split=1e-07, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, n_estimators=8, presort='auto', subsample=1.0, verbose=0, warm_start=False).

5.4.5 K-Nearest Neighbor

K-Nearest Neighbor (KNN), originally proposed by Cover and Hart in 1967 [Cover and Hart, 1967], is a theoretically mature method and one of the commonest machine learning algorithms. The idea of this method is : if most of the K most similar samples in the feature space (that is, the closest neighbors in the feature space) belong to a certain category, then the sample also belongs to this category.

The KNN algorithm is a non-parametric model. The parametric model uses a fixed number of parameters or coefficients to define the model. The non-parametric model does not mean that no parameters are required, but the number of parameters is uncertain, it may increase with the number of training instances.

KNN can also be adapted to nonlinear cases.

The calculation process of the KNN classification algorithm is as follows :

- Calculate the distance between the point to be classified and the point of the known class. There are three popular used distances in the KNN algorithm, namely Manhattan distance, Euclidean distance and Minkowski distance. Assumed that we want to compare the differences between X individuals and Y individuals, both of which contain N -dimensional features. Their mathematical description is as follows

$$X = (x_1, x_2, x_3, \dots, x_n), \quad Y = (y_1, y_2, y_3, \dots, y_n)$$

- Minkowski distance

Minkowski distance is not a distance, but a definition of a class of distances. Its mathematical definition is as follows :

$$\left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p}$$

where p can take any value, it can be negative, positive, or infinite. When $p=1$, it is also called Manhattan distance, when $p=2$, it is also called Euclidean distance, and when $p = \infty$, it is also called Chebyshev distance.

- Euclidean distance

Euclidean distance ($L2$ norm) is the distance between two points in space, which is the most commonly used. Because the calculation is based on the square difference of each dimension feature, Euclidean measurement needs to ensure that each dimension index is at the same scale level. For example, using Euclidean distance for two indicators with different units of height (cm) and weight (kg) may invalidate the results. Therefore, when using Euclidean distance, each component should be normalized as much as possible to reduce the interference caused by the different scale levels of the values.

- Manhattan Distance

The actual driving distance from one intersection to another in Manhattan is called "Manhattan distance", which is not the simple straight-line distance between two points, also known as City Block distance.

In this project, we take the Minkowski distance.

- Sort the distances from smallest to largest in ascending order.
- Select the K points which have the smallest distance from the points to be classified.
- Count the number of appearances in the category to which the top K points belong.
- Return the category with the highest number of occurrences of the first K points as the predicted classification of the points to be classified.

The classifier with the best model parameter in our project is :

```
KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkowski', metric_params=None,
n_jobs=1, n_neighbors=5, p=2, weights='uniform')
```

5.4.6 Logistic Regression

Logistic Regression [Wright, 1995] is used to deal with regression problems where the dependent variable is categorical. The common problem is binary or binomial distribution, but it can also deal with multi-category problems. It is used as a classification method.

The logistic classifier classifies by probability. The algorithm predicts the probability that an individual belongs to a certain class according to the predictor variable and then assigns the individual to the class with the highest probability. When the response variable is binary, called binomial logistic regression, and multi-classification is called multinomial logistic regression.

For multinomial logistic regression problems :

Assuming that the value set of the discrete random variable Y is $\{1, 2, \dots, K\}$, and there are a total of k classes of the feature x , then the output probability of the multi-class logistic regression model is :

$$P(Y = k | x) = \frac{e^{w_k \cdot x}}{1 + \sum_{i=1}^{k-1} e^{w_i \cdot x}}, k = 1, 2, \dots, K - 1$$

$$P(Y = K | x) = \frac{1}{1 + \sum_{i=1}^{K-1} e^{w_i \cdot x}}, x \in R^{n+1}, w_i \in R^{n+1}$$

Among them, note that $Y = k$ is a category that takes 1 to $(K - 1)$ class, $Y = K$ refers to the K -th class, and $P(Y = K | x)$ is the K -th class probability.

The origin of multi-class LR is the softmax function,

$$\text{softmax}(x) = \frac{e^k}{\sum_{k=1}^C e^k}$$

Operation steps of the logistic regression :

- Write the maximum likelihood function and perform the logarithmic operation

$$\begin{aligned} L(W) &= \sum_{i=1}^n \ln p(y_i | X_i; W) \\ &= \sum_{i=1}^n \ln \left(\phi(z_i)^{y_i} (1 - \phi(z_i))^{1-y_i} \right) \\ &= \sum_{i=1}^n y_i \ln \phi(z_i) + (1 - y_i) \ln (1 - \phi(z_i)) \end{aligned}$$

$\phi(z)$ is the step function. $L(W)$ is the cost function $J(W)$,

$$\phi(z) = \begin{cases} 0, & z < 0 \\ 0.5, & z = 0 \\ 1, & z > 0 \end{cases}$$

$$J(W) = - \sum_{i=1}^n y_i \ln \phi(z_i) + (1 - y_i) \ln (1 - \phi(z_i))$$

- Find the minimum value of the cost function using gradient descent.

The negative direction of the gradient is the direction in which the cost function decreases the fastest. Therefore, we should gradually adjust the weight component w_j along the negative gradient direction until we get the minimum value, so the change of each weight component should be like this :

$$\Delta w_j = -\eta \frac{\partial J(W)}{\partial w_j}$$

where η is the learning rate, which controls the step size. $\frac{\partial J(W)}{\partial w_j}$ can be calculated as follows :

$$\begin{aligned} \frac{\partial J(w)}{\partial w_j} &= - \sum_{i=1}^n \left(y_i \frac{1}{\phi(z_i)} - (1 - y_i) \frac{1}{1 - \phi(z_i)} \right) \cdot \frac{\partial \phi(z_i)}{\partial w_j} \\ &= - \sum_{i=1}^n \left(y_i (1 - \phi(z_i)) - (1 - y_i) \frac{1}{1 - \phi(z_i)} \right) \cdot x_{ij} \\ &= - \sum_{i=1}^n (y_i - \phi(z_i)) \cdot x_{ij} \end{aligned}$$

A special property of the sigmoid function $\phi(z)$ is used in the derivation of the above formula :

$$\phi(z)' = \phi(z)(1 - \phi(z))$$

In this way, we get the variable that the gradient descent method updates the weights :

$$w_j := w_j + \eta \sum_{i=1}^n (y_i - \phi(z_i)) \cdot x_{ij}$$

The classifier with best model parameter in our project is : `LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True, intercept_scaling=1, max_iter=3, multi_class='ovr', n_jobs=1, penalty='l2', solver='liblinear', tol=0.0001, verbose=0, warm_start=False)`.

5.4.7 Random Forest

Random Forest (RF) [Breiman, 2001] is an extension of the bagging algorithm in ensemble learning. Based on building a bagging ensemble with a decision tree as the base learner, RF further adds random attribute selection in the training process of the decision tree. Specifically, the traditional decision tree selects an optimal attribute from all the candidate attributes of the current node (assuming there are d) when selecting the attributes for division; while in RF, for each node of the base decision tree, the first A subset containing k attributes is randomly selected from the set of candidate attributes of the node, and then an optimal attribute is selected from this subset for partitioning. The selection of the number of attributes k to be extracted is important, and $K = \log_2 d$ is generally recommended. As a result, the "diversity" of the base learner of the random forest comes not only from the perturbation of samples but also from the perturbation of attributes, which further enhances the generalization ability of the final ensemble.

Algorithm steps of RF :

- Assume that there is a dataset $D = \{x_{i1}, x_{i2}, \dots, x_{in}, y_i\}$ ($i \in [1, m]$), containing N features, a sampling with replacement process can generate a sampling space $(m * n)^{m*n}$
- Build a base learner (decision tree) : generate a decision tree for each sample $d_j = \{x_{i1}, x_{i2}, \dots, x_{ik}, y_i\}$ ($i \in [1, m]$) (where $K \ll M$), and record the result of each decision tree $h_j(x)$.
- Train T times to make $H(x) = \max \sum_{t=1}^T \phi(h_j(x) = y)$, where $\phi(x)$ is an algorithm (super majority voting, relative majority voting, weighted voting, etc.).

The classifier with best model parameter in our project is : `RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini', max_depth=None, max_features='auto', max_leaf_nodes=None, min_impurity_split=1e-07, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, n_estimators=31, n_jobs=1, oob_score=False, verbose=0, warm_start=False)`.

5.4.8 SVM RBF & SVM Sigmoid

The support vector machine (SVM) [Cao et al., 2008, Smola et al., 2000] estimates hyperplane to segment the samples. The principle of segmentation is to maximize the interval between the hyperplane and the nearest samples (support vectors), which is finally transformed into a convex quadratic programming problem to solve. It can solve both linearly separable and nonlinear problems using the kernel trick.

The main idea of SVM can be summarized in two points : (1) It analyzes the linearly separable case. For the linearly inseparable case, the non-linear mapping algorithm is used to convert the linearly inseparable samples in the low-dimensional input space into a high-dimensional feature space to make them linearly separable, therefore, it is possible to use a linear algorithm to linearly analyze the nonlinear characteristics of samples in high-dimensional feature space; (2) It constructs the optimal segmentation hyperplane in the feature space based on the structural risk

minimization theory, so that the learner is globally optimized, and the expected risk in the entire sample space satisfies a certain upper bound with a certain probability.

The key point in SVM is the kernel function. The main idea is to convert the samples that are linearly inseparable in the initial low-dimensional space into linearly separable samples in a high-dimensional space. Low-dimensional space vector sets are often difficult to divide, and the solution is to map them to high-dimensional spaces. But the difficulty brought by this method is the increase in computational complexity, and the kernel function just solves this problem through its interpretation as a scalar product.

In SVM theory, different kernel functions will lead to different SVM algorithms. Considering the generalization problem, two parameters, relaxation coefficient and penalty coefficient are introduced to smooth the frontiers between classes.

There are four commonly used kernel functions : Given two examples x_1 and x_2 :

— Linear kernel function

$$K(x_1, x_2) = x_1^\top x_2$$

— Polynomial kernel function

$$K(x_1, x_2) = (x_1^\top x_2 + c)^d$$

Where d is the degree of kernel, all interactions of order d or lower. c is a constant term greater than 0.

— Radial basis (RBF) kernel function (Gaussian kernel function)

$$K_{rbf}(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{c}\right)$$

Where $\|x_1 - x_2\|^2$ is euclidean distance between x_1 and x_2 .

— Sigmoid kernel function (two-layer neural network kernel function)

$$K(x_1, x_2) = \tanh(\alpha x_1^\top x_2 + c)$$

Where α and c are constant terms greater than 0.

We take the RBF kernel and sigmoid kernel in our project. RBF kernel function can map the original space to an infinite-dimensional space. For parameter σ , if it is selected very large, the weight on the high-order features would decay very quickly, which is equivalent to a low-dimensional subspace; conversely, if it is selected very small, we can map arbitrary data to be linearly separable and may lead to serious overfitting problem. However, in general, by adjusting the parameters, the Gaussian kernel is quite flexible and one of the most widely used kernel functions. When the sigmoid kernel uses the sigmoid function as the kernel function, the support vector machine implements a multi-layer perceptron neural network. The weights of nodes are automatically determined during the training process. Moreover, the theoretical basis of the SVM determines that it finally obtains the global optimal value instead of a local minimum value, and also ensures its good generalization ability for unknown samples.

Classification using SVM goes through three steps :

— Choose a suitable mathematical function as the kernel function.

— Using the kernel function for high-dimensional mapping, the data points change from linearly inseparable to linearly separable after mapping.

— The hyperplane that can maximize an interval is found, and the classification is finally completed.

This interval is determined by the sum of the distances from the point as the support vector to the hyperplane, and the distance here is the most common geometric distance. We use $wx + b$ to represent the hyperplane, and the formula for the distance from a point to a

three-dimensional plane is :

$$d = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}$$

Similarly, the distance r from a point to an N -dimensional hyperplane can be calculated by the following formula :

$$\gamma^{(i)} = \frac{(w^T x^{(i)} + b)}{\|w\|}$$

Where the dividend $w^T x^{(i)} + b$ is the expression of the hyperplane, and the divisor $\|w\|$ is the abbreviation of the $L2$ normal form. SVM uses this formula to calculate the distance of a point to the hyperplane.

The goal is to maximize the interval. 2 is a constant, so maximizing the distance between the two hyperplanes (margin) can be expressed as :

$$\min \frac{1}{2} \|w\|^2$$

The illustration for the support vectors, the margins and the decision boundary is shown in Figure 5.11. The margins are the space between the decision boundary and the support vectors furthest out.

Then use the Lagrangian multiplier method to transform it into the following Lagrangian function :

$$L(w, b, a) = \frac{1}{2} \|w\|^2 + \sum_{i=1}^m \alpha_i [1 - y_i (w^T x_i + b)]$$

where α is called the "Lagrange multiplier". Taking the derivative of the above formula concerning w and b respectively, and setting the derivative to 0, the right-hand formula can be transformed into the following formula :

$$\begin{aligned} & \sum_{i=1}^m \alpha_i - \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \max_{\alpha} & \sum_{i=1}^m \alpha_i - \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i^T x_j \end{aligned}$$

The constraints are :

$$\begin{aligned} \text{s.t.} & \sum_{i=1}^m \alpha_i y_i = 0 \\ & \alpha_i \geq 0 \end{aligned}$$

This last formula is usually solved by the quadratic programming algorithm Sequential Minimal Optimization (SMO) algorithm. The s.t. represents subject to, which means to be constrained.

The classifiers with the best model parameter in our project are :

```
'svm_rbf' : SVC(C=451, cache_size=200, class_weight='balanced', coef0=0.0, decision_function_shape=None, degree=3, gamma=0.002, kernel='rbf', max_iter=-1, probability=False, shrinking=True, tol=0.001, verbose=False)
```

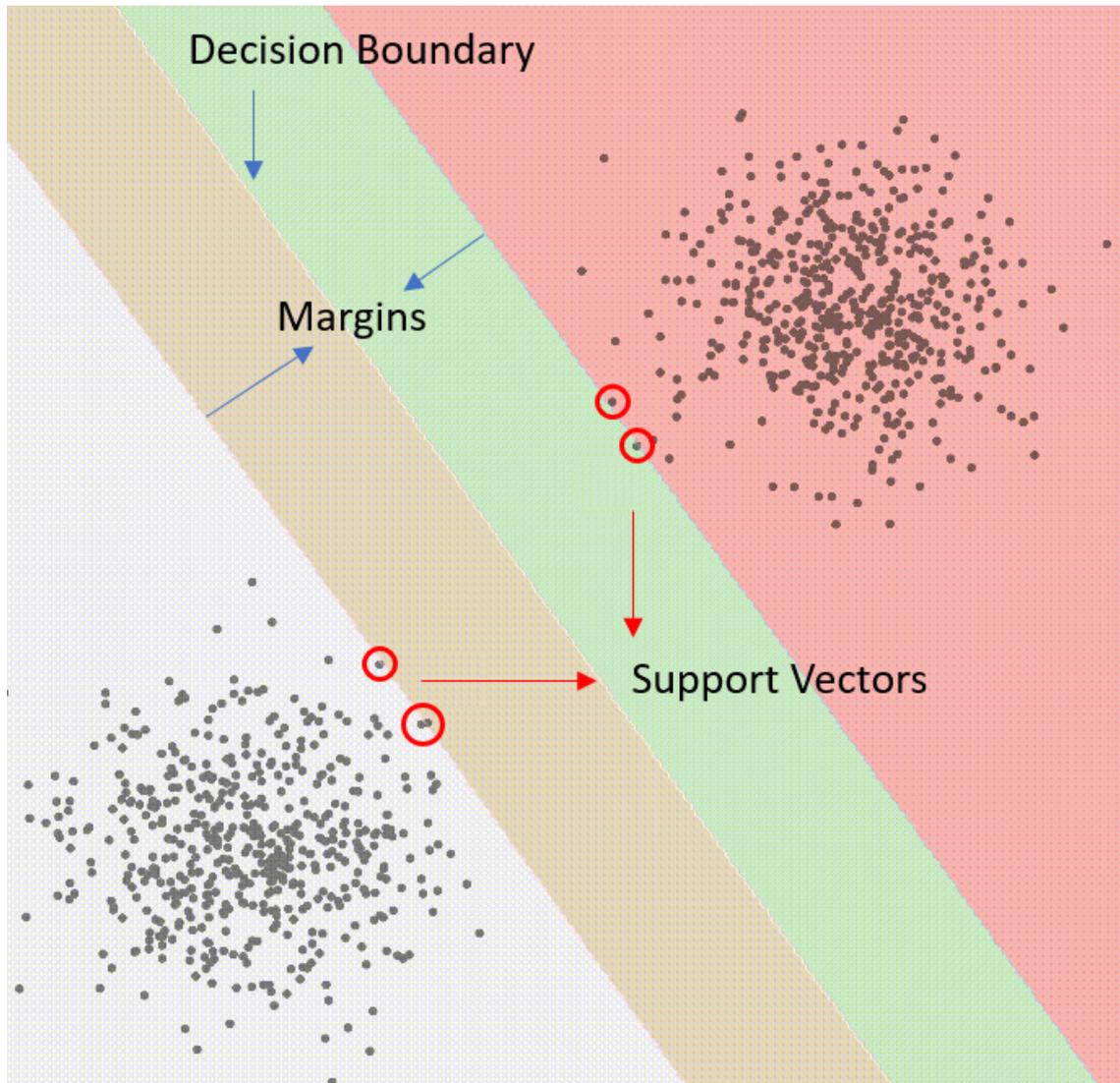


Figure 5.11 – The illustration of SVM.
It includes support vectors, the margins and the decision boundary. (Diego, 2022)

```
'svm_sigmoid' : SVC(C=1.0, cache_size=200, class_weight='balanced', coef0=0.33333, decision_function_shape=None, degree=3, gamma=0.012, kernel='sigmoid', max_iter=-1, probability=False, shrinking=True, tol=0.001, verbose=False)
```

5.4.9 ROC Plotting and AUC Calculation

Receiver Operating Characteristic (ROC) is a curve drawn on a two-dimensional plane, in which abscissa is the false positive rate (FPR) and the ordinate is the true positive rate (TPR). For a classifier, we can get a TPR and FPR point pair based on the performance of the test sample. In this way, this classifier can be mapped to points to get a ROC curve through (0, 0), (1, 1). In general, this curve should be above a random classifier, which is the line connecting (0, 0) and (1, 1). Area Under roc Curve (AUC) is a standard used to measure the quality of a classification model. The value of AUC is the size of the area below the ROC curve. Typically, the value of AUC is between 0.5 and 1.0, with a larger AUC representing better performance.

Assume that the number of test samples is m and the number of categories is n (assuming the category labels are : 0, 1, ..., $n-1$). After the training is completed, the probability or confidence of each test sample of each category is calculated, a matrix P in the shape of $[m, n]$ is obtained, and each row represents the probability value of a test sample under each category (by category label sort). Meanwhile, the label of each test sample is converted into a binary-like form, and each position is used to mark whether it belongs to the corresponding category (also sorted by the label, so that it corresponds to the previous one), thus a label matrix L of $[m, n]$ can also be obtained. For example, when n is 3, the label should be converted to :

$$\begin{aligned} 0 &\Rightarrow 100 \\ 1 &\Rightarrow 010 \\ 2 &\Rightarrow 001 \end{aligned}$$

Since the ROC curve is defined for binary classifications, for multi-class problems, there are two main ways to generalize the ROC curve :

Method 1 : Under each category, the probability that m number of test samples are of this category (columns in matrix P) can be obtained. Therefore, according to each column in the probability matrix P and the label matrix L , the FPR and the TPR under each threshold can be calculated to draw a ROC curve. In this way, a total of n ROC curves can be drawn. Finally, take the average of n ROC to obtain the final ROC curve.

Method 2 : At first, for a test sample, the label is only composed of 0 and 1, the position of 1 indicates its category (corresponding to "positive" in the two-class problem), and 0 indicates other categories ("negative"); Then if the test sample classifies correctly, the value of the position corresponding to 1 in the sample label from the probability matrix P is greater than the probability value of the position corresponding to the label 0. Based on these two points, the label matrix L and the probability matrix P are expanded in rows and transposed to form two columns, which results in a binary classification result. Therefore, this method can directly obtain the final ROC curve after calculation.

We take method 2 to calculate AUC in our project. The ROC curves and AUC values of each model are shown in Figure 5.12.

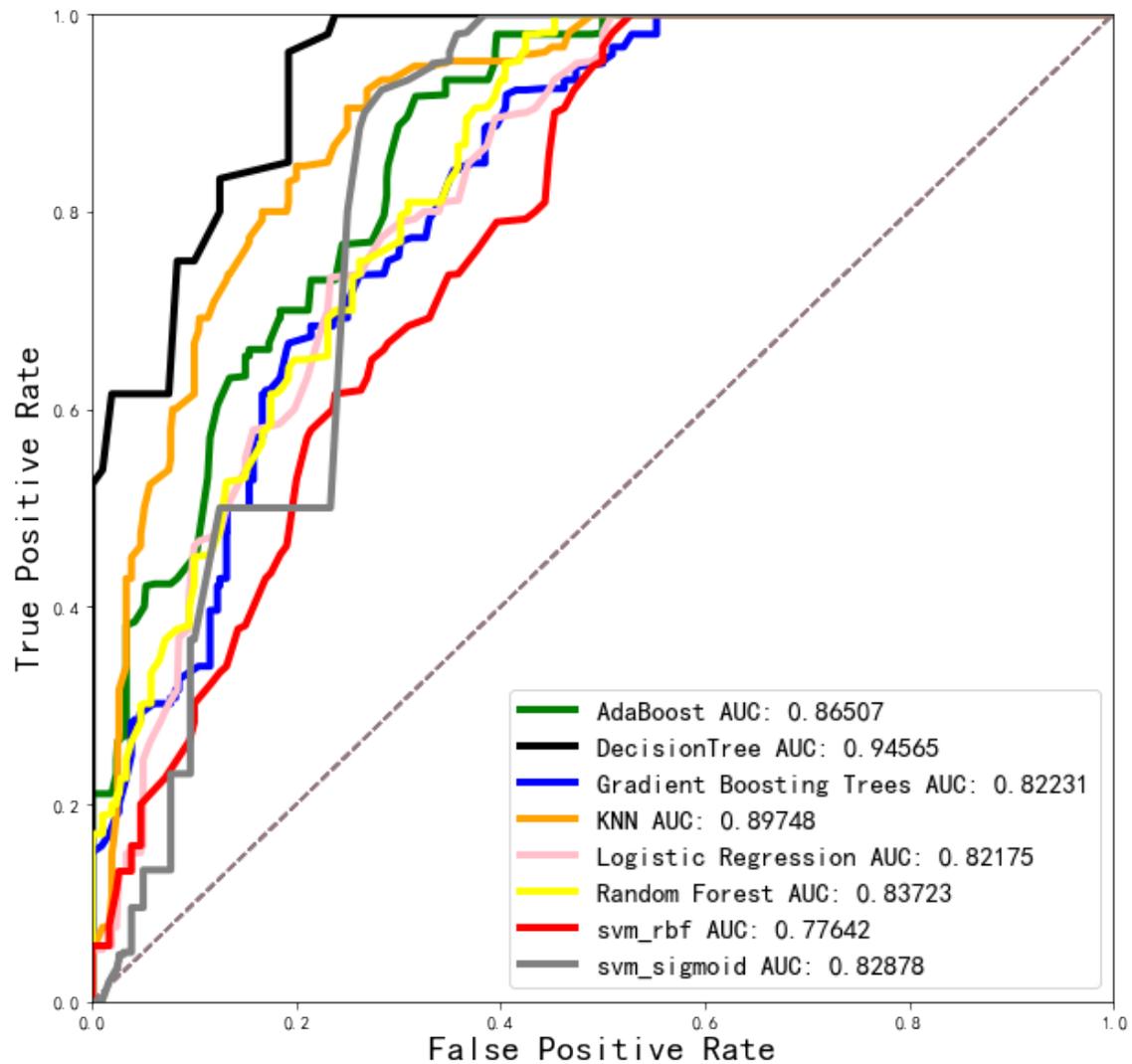


Figure 5.12 – The ROC curves and AUC values were obtained for each model with filtered hand-crafted features.

TABLE 5.1 – P-value of non-parametric tests of every feature.

Features	Kruskal-Wallis Test	Median Test	Jonckheere-Terpstra Test
NE	0.00E+00	0.00E+00	0.00E+00
small NE	7.44E-03	2.78E-02	2.42E-03
long NE	0.00E+00	0.00E+00	0.00E+00
NJ	0.00E+00	0.00E+00	0.00E+00
LE	0.00E+00	0.00E+00	0.00E+00
LJ	0.00E+00	0.00E+00	0.00E+00
density	0.00E+00	0.00E+00	0.00E+00
NE/NJ	0.00E+00	0.00E+00	0.00E+00
LE/LJ	1.61E-01	3.79E-01	9.56E-01
NE/(LJ+LE)	5.74E-02	4.30E-01	1.62E-02
NJ/(LJ+LE)	0.00E+00	0.00E+00	0.00E+00
LJ/(LJ+LE)	1.00E-06	2.35E-03	0.00E+00
mean area	0.00E+00	0.00E+00	0.00E+00
median area	0.00E+00	0.00E+00	0.00E+00
mean perimeter	0.00E+00	0.00E+00	0.00E+00
median perimeter	0.00E+00	0.00E+00	0.00E+00
mean eccentricity	0.00E+00	0.00E+00	1.00E-06
median eccentricity	1.80E-05	6.00E-06	5.40E-05

5.5 Results

In this section, we display the results of hand-crafted features analysis and benchmark experiments on the VRCC dataset with traditional classification schemes and deep learning models. The given accuracy results from the validation are the average of the "leave one out" cross-validation.

5.5.1 Results of Hand-crafted Features Analysis

As shown in Table 5.1, there are only two features (LE/LJ, NE/(LJ+LE)) whose P-value is larger than 0.05 for at least one statistical test. Therefore, we choose the 16 features (**font with blue color**) that are significant in all 3 tests. We assume that these 16 features are more suitable for classifying the ccRCC, pRCC and chRCC images. Subsequent experiments, as shown in Table 5.2 also proved this : features filtered by non-parametric tests perform best.

5.5.2 Performance of Hand-crafted Features on Traditional Algorithms under Non-parameter Test

As shown in Table 5.2, we compare the result of the skeleton and lattice features with the baseline features : NE, NJ, LE, LJ, NE/NJ and LE/LJ [Zubiolo et al., 2016]. All experiments used the leave-one-out cross-validation method and a majority vote to obtain patient-level classification results. Our proposed features, particularly filtered by non-parametric tests (filtered Skeleton & Lattice) achieve higher accuracy on almost all traditional machine learning algorithms (Adaboost [Freund and Schapire, 1997], DecisionTree [Breiman et al., 1984], Gradient Boosting Tree [Friedman, 2001], KNN [Cover and Hart, 1967], Logistic Regression [Wright, 1995], Ran-

dom Forest [Breiman, 2001] and SVM [Cao et al., 2008, Smola et al., 2000]) with two different kernels. This proved that the proposed features are more robust and efficient by adding lattice features.

TABLE 5.2 – The Accuracy obtained by 3 feature sets on different algorithms.

Methods	baseline Skeleton	full Skeleton & Lattice	filtered Skeleton & Lattice
Adaboost [Freund and Schapire, 1997]	0.6933	0.8533	0.9000
DecisionTree [Breiman et al., 1984]	0.6933	0.8700	0.9167
Gradient Boosting Tree [Friedman, 2001]	0.680	0.8733	0.9333
KNN [Cover and Hart, 1967]	0.7333	0.9100	0.9467
Logistic Regression [Wright, 1995]	0.7300	0.8967	0.9167
Random Forest [Breiman, 2001]	0.7100	0.9400	0.9367
SVM RBF [Cao et al., 2008]	0.6900	0.9067	0.9100
SVM Sigmoid [Smola et al., 2000]	0.6633	0.8433	0.9300

5.5.3 Results of Vascular-Based RCC Classification Benchmark

TABLE 5.3 – The benchmark results on the VRCC dataset.

Methods	Input	Feature	Accuracy (LeaveOneOut)
Traditional Algorithm			
Adaboost [Freund and Schapire, 1997]	Segment	Hand-crafted	0.9000
Decision Tree [Breiman et al., 1984]	Segment	Hand-crafted	0.9167
Gradient Boosting Tree [Friedman, 2001]	Segment	Hand-crafted	0.9333
KNN [Cover and Hart, 1967]	Segment	Hand-crafted	0.9467
Logistic Regression [Wright, 1995]	Segment	Hand-crafted	0.9167
Random Forest [Breiman, 2001]	Segment	Hand-crafted	0.9367
SVM RBF [Cao et al., 2008]	Segment	Hand-crafted	0.9100
SVM Sigmoid [Smola et al., 2000]	Segment	Hand-crafted	0.9300
Deep Learning			
LeNet [LeCun et al., 1989]	Raw	Deep	0.6616
AlexNet [Krizhevsky et al., 2012a]	Raw	Deep	0.6178
VGG-16 [Simonyan and Zisserman, 2014]	Raw	Deep	0.6440
ResNet-18 [He et al., 2016]	Raw	Deep	0.8091
LeNet [LeCun et al., 1989]	Segment	Deep	0.8905
AlexNet [Krizhevsky et al., 2012a]	Segment	Deep	0.8750
VGG-16 [Simonyan and Zisserman, 2014]	Segment	Deep	0.8590
ResNet-18 [He et al., 2016]	Segment	Deep	0.8850
GCN+SAGPoolg [Zhang et al., 2018a]	Segment	Graph	0.8596
GCN+SAGPoolh [Cangea et al., 2018]	Segment	Graph	0.8772

As shown in Table 5.3, the results of the traditional methods range from 90% to 94.67%, which is higher than most others. This demonstrates that our proposed skeleton and lattice features are efficient and robust.

For the results of deep learning baselines, the accuracy with vascular segmentation input is much better than raw image input. We argue that a bigger dataset is needed to train raw images, but vascular segmentation can work well on a small dataset and proved that we can just use the vascular network to do classification.

The results of the two GCN models are lower than some vascular segment input-based deep learning models (89.05% *vs* 87.72%) and traditional algorithm models (94.67% *vs* 87.72%). Maybe because the graph features only contain the topological information of vascular, whereas the other features contain both topological and geometrical information.

5.6 Conclusion

To demonstrate the importance of the vascular network structure of RCC, we build the VRCC datasets with accurate vascular annotations. In this way, we can characterize ccRCC, pRCC and chRCC from the topological and geometrical information of RCC vascular networks, by our proposed "skeleton and lattice features". Our Hand-crafted features trained with 8 popular traditional machine learning algorithms provide a better explanation for the RCC classification task and the model performance with our Hand-crafted features is much better than the "graph features" (only topological information) and the baseline deep learning features (both vascular segment masks and raw images) from the vascular network mask. And the accuracy of deep learning with vascular segment mask input is much better than with raw image input. This can be explained by the lack of training data. Segmenting the vascular network, and even more computing Hand-crafted features, somehow reduce the dimensionality of the problem and thus require less training data. All the results in this chapter proved that the vascular network can be used alone for RCC classification and is sufficient to define the tumor subtypes.

However, there are still some limitations of our classification research which is based on the manual vascular network annotation. Automatic subtyping application has been limited due to the time required for manual vascular labeling. It would be improved if we can segment the vascular network from the H&E stained histopathological images automatically. But supervised learning is challenging for the vascular segmentation task in our project due to lacking ground truth. This is the motivation for our next semi-supervised segmentation research.

Automatic Segmentation of Vascular Network

Coupling Vascular Network Segmentation and RCC Classification

Dans ce chapitre, nous proposons un modèle d'apprentissage multi-tâches (MTL) qui permet de segmenter le réseau vasculaire à l'aide d'un apprentissage semi-supervisé et de prédire la classe tumorale dans un contexte supervisé simultanément. Cette procédure d'apprentissage multitâche offre une solution d'apprentissage automatique de bout en bout pour la segmentation du réseau vasculaire et la classification des tumeurs. La validation a été effectuée sur une base de données d'images histopathologiques de carcinome à cellules claires (RCC), puis sur la base de données publique TCGA. Les résultats montrent que le modèle MTL-SSL proposé surpasse l'approche conventionnelle de segmentation par apprentissage supervisé.

In this chapter, we propose a multi-task learning (MTL) model that can simultaneously segment the vascular network using SSL and predict the tumor class in a supervised context. This multi-task learning procedure offers an end-to-end machine learning solution to joint vascular network segmentation and tumor classification. Experiments were carried out on a database of histopathological images of renal cell carcinoma (RCC) and then tested on both our own RCC and the open-access TCGA datasets. The results show that the proposed MTL-SSL model outperforms the conventional supervised-learning segmentation approach.

6.1	Multi-Task Learning Pipeline	109
6.1.1	HRNet Backbone	109
6.1.2	Mean Teacher	109
6.1.3	Proposed MTL-SSL model	110
6.2	Proposed Vessel Loss Function	114
6.2.1	Definition of Loss Functions	114
6.2.1.1	Cross Entropy Loss	114
6.2.1.2	Dice Loss	116
6.2.1.3	Mean Squared Error Loss	116
6.2.2	Main Idea and Realization of Our Proposed Loss Function.	117
6.3	Vascular Network Segmentation Benchmark of RCC	119
6.3.1	Experimental Environment	119
6.3.2	Backbone Choice	119
6.3.3	Segmentation Benchmarks of Vascular Network	120
6.4	Test on New Subtype of RCC and Other Type of Cancers	122
6.4.1	New Subtype of RCC and Other Cancers	122
6.4.1.1	Oncocytoma	122
6.4.1.2	Breast Cancer	122
6.4.1.3	Liver Cancer	123
6.4.1.4	Lung Cancer	124
6.4.1.5	Esophageal Cancer	124
6.4.2	Test on New Subtype of RCC and Other Cancers Dataset	125
6.5	Conclusion	130

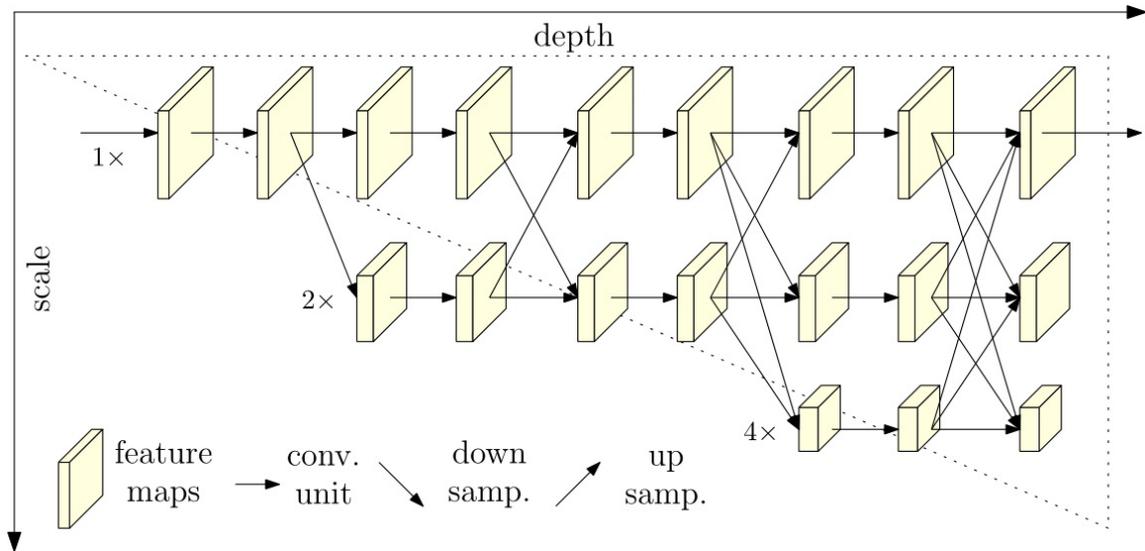


Figure 6.1 – The HRNet structure consists of parallel high-to-low resolution subnetworks. The parallel sub-networks repeatedly exchange information across resolutions : cross-resolution multi-scale fusion. The horizontal direction represents the network depth, and the vertical direction represents the feature map scale (resolution). (Sun, 2019)

6.1 Multi-Task Learning Pipeline

In this part, we chose the HRNet as our backbone of the model, and choose the Mean Teacher as the SSL method. These choices have been made according to experiments.

6.1.1 HRNet Backbone

HRNet was completed by a team led by Wang Jingdong at Microsoft Research Asia [Sun et al., 2019]. It has integrated image classification, image segmentation, target detection, face alignment, gesture recognition, style transfer, super-score, optical flow, depth estimation, edge detection, etc. HRNet backbone is sensitive to the target location and can output high-resolution feature maps. It starts with a high-resolution subnetwork as the first stage, and gradually adds high-to-low resolution subnetworks, forming more stages, and connecting the multi-resolution subnetworks in parallel. By performing multi-scale fusions, each high-to-low resolution repeatedly receives information from other parallel representations, resulting in rich high-resolution representations. The entire network structure is shown in Figure 6.1.

6.1.2 Mean Teacher

For the SSL structure, we chose the Mean Teacher [Tarvainen and Valpola, 2017] according to the backbone selection experiment, shown in Figure 6.2, which has two neural networks : a “student module” and a “teacher module”, these two modules shared the same architecture.

The core idea of Mean Teacher is as follows : The model acts as both student and teacher. The teacher generates the learning goals of the student ; the student uses the goals generated by the teacher model to learn.

Both the student and the teacher modules evaluate the input slightly perturbed with Gaussian noise (η and η') within their computation. The weights of the student module are updated using the Adam optimizer, whereas the weights of the teacher module are the Exponential Moving Average (EMA) of the student weights. EMA is a type of moving average (MA) that gives recent data points more weight and significance. The exponential moving average is also known as exponentially weighted moving averages.

Mean Teacher's training strategy :

- Training data : labeled samples x_1, y_1 , and unlabeled samples x_2 .
 - Input the labeled data x_1, y_1 into the student to calculate $loss_1$.
 - Enter the unlabeled data x_2 into the student to calculate the label1
 - Input the unlabeled data x_2 into the teacher to calculate the label2
 - We hope that the predicted labels of the two networks are as equal as possible, so the loss function $loss_2$ is obtained according to label1 and label2.
 - Update the student network according to $loss = loss_1 + loss_2$.
 - In each step, after updating the parameters of the student network, use the parameters of the student network to update the parameters of the teacher network,
- Therefore, the second term of the objective function of the Mean Teacher model is :

$$J(\theta) = \mathbb{E}_{x, \eta', \eta} [\|f(x, \theta', \eta') - f(x, \theta, \eta)\|^2]$$

where J is consistency cost, as the expected distance between the prediction of the student model (with weights θ and noise η) and the prediction of the teacher model (with weights θ' and noise η').

The update method of the model parameter θ is :

$$\theta'_t = \alpha \theta'_{t-1} + (1 - \alpha) \theta_t$$

where θ'_t is the EMA of successive θ weights at the training step t , α is a smoothing coefficient hyperparameter.

6.1.3 Proposed MTL-SSL model

Our proposed MTL-SSL model has a shared backbone encoder with task-specific heads shown in Figure 6.3. It consists of a classification task in a supervised learning context and a segmentation task using SSL, as shown in Figure 6.4. We chose HRNet [Sun et al., 2019] as the backbone after comparison with other models (see the experimental section). The entire backbone part of HRNet is divided into 4 stages, and each stage is divided into two parts : a blue box and an orange box. The blue box part is the basic structure of each stage, which is composed of multiple branches. The blue box of stage I in HRNet uses BottleNeck and the blue box of stages II&III&IV uses BasicBlock. The orange box part is the transition structure of each stage, the orange box of stage I in HRNet is a TransitionLayer, the orange box of stages II&III is a superposition of a FuseLayer and a TransitionLayer, and the orange box of stage IV is a FuseLayer. FuseLayer is used for information interaction between different branches. BasicBlock structure contains a residual branch and a short-cut branch, which is one more short-cut branch than the traditional convolution structure. BottleNeck structure can reduce the number of parameters and channels in the intermediate feature map, which can consume less memory and be conducive to building a network with more layers. TransitionLayer is used to generate an input feature map that downsamples twice the branch. The

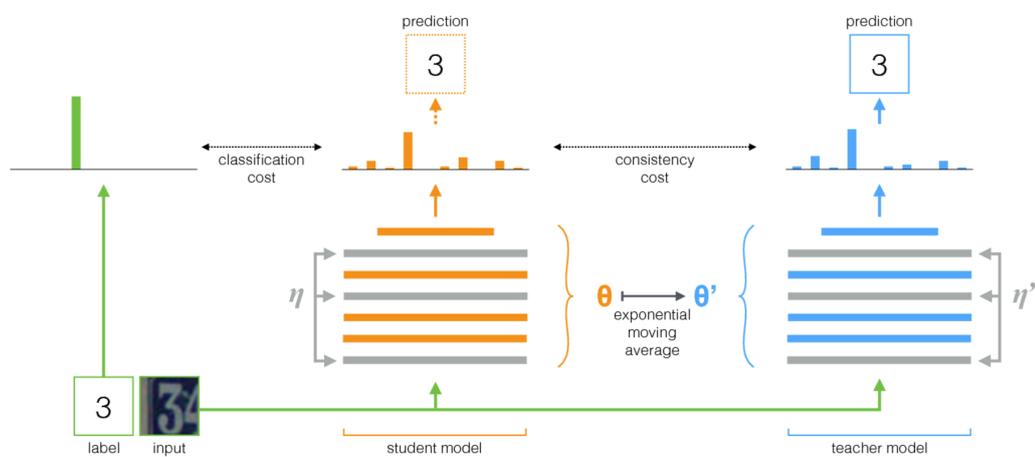


Figure 6.2 – The Mean Teacher method.

The figure depicts a training batch with a single labeled example. Both the student and the teacher model evaluate the input applying noise (η, η') within their computation. The softmax output of the student model is compared with the one-hot label using classification cost and with the teacher output using consistency cost. After the weights of the student model have been updated with gradient descent, the teacher model weights are updated as an exponential moving average of the student weights. Both model outputs can be used for prediction, but at the end of the training, the teacher's prediction is more likely to be correct. A training step with an unlabeled example would be similar, except no classification cost would be applied. (Tarvainen, 2017)

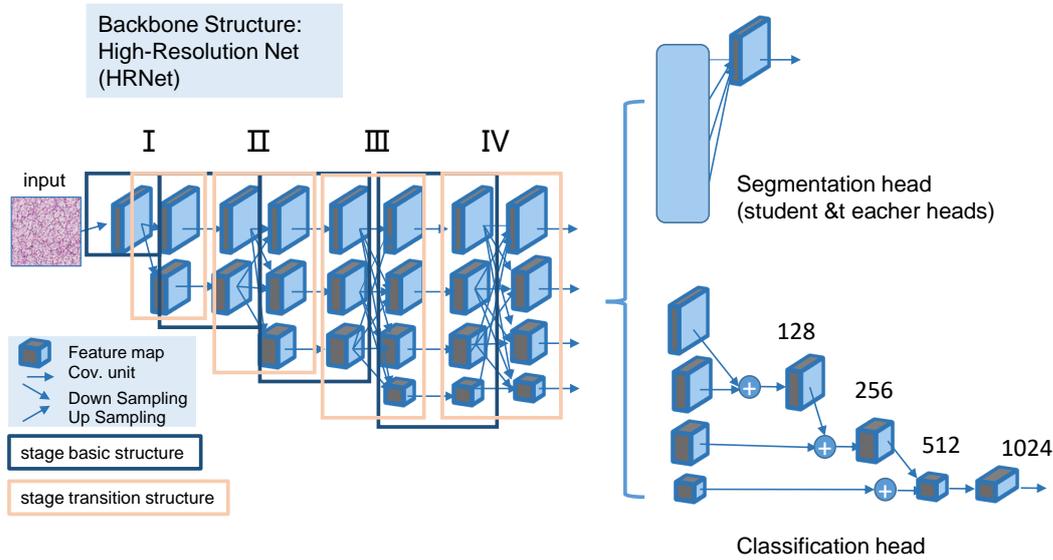


Figure 6.3 – The 4 stages of the HRNet backbone and its different heads.

orange box of stage I obviously cannot be used as FuseLayer, because the previous stage has only one branch and the orange of stage IV is followed by the neck and head, which also longer needed TransitionLayer.

HRNet segmentation heads (student head and teacher head) aggregate the output representations at four different resolutions, and then use 1×1 convolutions to fuse these representations. HRNet classification head fed the four-resolution feature maps into a bottleneck and the number of output channels is increased to 128, 256, 512, and 1024, respectively, then the high-resolution representations are downsampled by a 2-stride 3×3 convolution outputting 256 channels and added to the second-high-resolution representations. This scheme is repeated twice to get 1024 channels over the small resolution and finally, the 1024 channels are transformed to 2048 channels through a 1×1 convolution finally. The codes of Multi-task and HRNet backbone were developed according to the shared repository [Vandenhende et al., 2020] and [Sun et al., 2019]. The main hyperparameters used in this work are the same as [Vandenhende et al., 2020] and [Sun et al., 2019]. Ensuring fair comparison; all the models were trained using the same hyperparameters.

For the loss function, we use the cross-entropy (CE) and our proposed Vessel Loss functions between the student's predictions and the ground truth on the labeled dataset to get $loss_2$. The consistency cost, called $loss_3$ here, is computed from the student's prediction and the teacher's prediction by Mean Square Error (MSE) on the unlabeled dataset. The semi-supervised $loss_4$ is the sum of the supervised $loss_2$ and the consistency cost $loss_3$ by consistency weights, which were taken from [Tarvainen and Valpola, 2017]. Classification $loss_1$ is computed by the CE function on the class-labeled dataset. Final $loss_5$ of our MTL-SSL model is the weighted sum of semi-supervised $loss_4$ and classification $loss_1$, we defined 2 : 1 ratio between SSL task weight and classification task weight in our experiments.

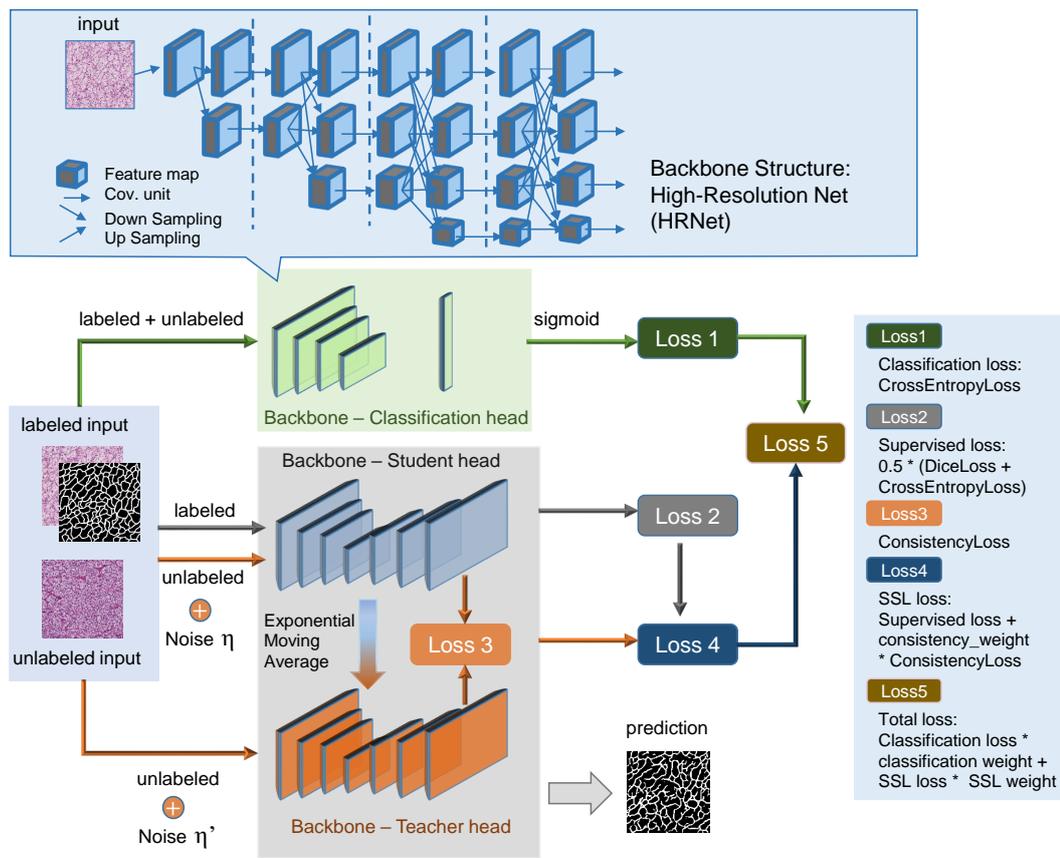


Figure 6.4 – Proposed MTL-SSL model architecture.

6.2 Proposed Vessel Loss Function

6.2.1 Definition of Loss Functions

Supervised learning in machine learning is essentially considered given a series of training samples (x_i, y_i) , trying to learn the mapping relationship of $x \rightarrow y$, so that given a x , even if this x is not in the training sample, it can output \hat{y} , as close as possible to the real y . The loss function is used to estimate the gap between the output \hat{y} of the model and the true value y , and guide the optimization of the model. The structural risk of the model includes empirical risk, and the loss function is the core part of the empirical risk function :

$$\hat{\theta} = \arg \min_{\theta} \frac{1}{N} \sum_{i=1}^N L(y_i, f(x_i; \theta)) + \lambda \Phi(\theta)$$

In the formula, the previous mean function is the empirical risk, $L(y_i, f(x_i; \theta))$ is the loss function, the latter item is the structural risk, and the $\Phi(\theta)$ measures the complexity of the model.

We used several classical loss functions in our models, such as CE, Dice and MSE loss functions.

6.2.1.1 Cross Entropy Loss

The most commonly used loss function for image semantic segmentation tasks is the pixel-level cross-entropy loss, which examines each pixel one by one, and the predictions (probability distribution vectors) for each pixel class are compared to our one-hot encoded vector. One-hot encoding is a sparse way of representing data in a binary string where only a single bit can be 1 and all other bits are 0. Our binary image segmentation task, which has two classes (vessel pixels and non-vessel pixels), can also be viewed as a binary classification task. In binary classification, we usually use the sigmoid function to compress the output of the model into the $(0,1)$ interval, and the $\hat{y}_i \in (0, 1)$ is used to represent the probability that the model is judged to be a positive class given the input x_i . Since there are only two classes (positive and negative classes), the probability of the negative class is also obtained :

$$\begin{aligned} p(y_i = 1 | x_i) &= \hat{y}_i \\ p(y_i = 0 | x_i) &= 1 - \hat{y}_i \end{aligned}$$

Combine the two equations :

$$p(y_i | x_i) = (\hat{y}_i)^{y_i} (1 - \hat{y}_i)^{1-y_i}$$

Assuming that the data points are independently and identically distributed, the likelihood can be expressed as :

$$L(x, y) = \prod_{i=1}^N (\hat{y}_i)^{y_i} (1 - \hat{y}_i)^{1-y_i}$$

Taking the logarithm of the likelihood, and then adding a negative sign to minimize the negative log-likelihood, gives the cross-entropy loss function :

$$NLL(x, y) = J_{CE} = - \sum_{i=1}^N y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)$$

However, the vessel pixels and non-vessel pixels are obviously imbalanced for one tumor patch image. If the loss function does not consider the sample balance problem, the learning process will tend to segment most of the non-vessel parts. To solve this imbalance problem, we adopt the class-balanced cross-entropy loss function [Xie and Tu, 2015], given as follows :

$$\begin{aligned} \ell_{\text{side}}^{(m)}(W, w^{(m)}) = & -\beta \sum_{j=Y_+} \log(\sigma(y_j^{(m)})) \\ & -\alpha(1-\beta) \times \sum_{j=Y_-} \log(1 - \sigma(y_j^{(m)})) \end{aligned}$$

Where $\ell_{\text{side}}^{(m)}(W, w^{(m)})$ is the loss of the m -th side output, W and $w^{(m)}$ denote the parameters of the convolutional layer and the m -th edge output, respectively. We use Y_+ and Y_- to denote vessel and non-vessel pixels. Weights β and α are used to balance classes. β is defined by the ratio of vessel pixels to all pixels. α is a hyperparameter, which we set to 2.5 in our experiments. $\sigma(y_j^{(m)})$ represents the probability map generated by applying a sigmoid on the activation value $y_j^{(m)}$ at pixel j .

In the multi-classification task, the derivation idea of the cross-entropy loss function is the same as that of the two-classification. The change is that the real value y_i is a one-hot vector, and the compression of the model output is changed from the original Sigmoid function to Softmax function. The Softmax function limits the output range of each dimension between (0,1), and the output sum of all dimensions is 1, which is used to represent a probability distribution.

$$p(y_i | x_i) = \prod_{k=1}^K (\hat{y}_i^k)^{y_i^k}$$

Among them, $k \in K$ represents one of the K categories, and the same assumption is made that the data points are independent and identically distributed, the negative log-likelihood can be obtained as :

$$NLL(x, y) = J_{CE} = - \sum_{i=1}^N \sum_{k=1}^K y_i^k \log(\hat{y}_i^k)$$

Since y_i is a one-hot vector, the output on all other categories except the target class is 0, so the above formula can also be written as :

$$J_{CE} = - \sum_{i=1}^N y_i^{c_i} \log(\hat{y}_i^{c_i})$$

Among them, c_i is the target class of x_i , usually, this cross-entropy loss function applied to multi-classification is also called Softmax Loss or Categorical Cross Entropy Loss. We use this Cross Entropy Loss for the classification task, which has 3 classes.

6.2.1.2 Dice Loss

Another loss function commonly used in semantic segmentation tasks is based on the Dice coefficient, which is essentially a measure of the overlap between two objects/shapes. This measure ranges from 0 to 1, where a Dice coefficient of 1 indicates complete overlap. Dice coefficients were originally intended for binary data and can be calculated as :

$$\text{Dice} = \frac{2|A \cap B|}{|A| + |B|}$$

$|A \cap B|$ represents the common element between sets A and B , and $|A|$ represents the number of elements in set A (same for set B).

For evaluating the Dice coefficients on the predicted segmentation mask, we can approximate $|A \cap B|$ by an element-wise multiplication between the predicted mask and the label mask, and then sum the resulting matrix.

To design a loss function that can be minimized, one can simply use $1 - \text{Dice}$. This loss function is called the soft Dice loss because we directly use the predicted probability instead of using a threshold to convert it into a binary mask.

Dice loss is proposed when the proportion of foreground is too small. The dice coefficient is derived from the binary classification, which is essential to measure the overlapping part of the two samples.

For the output of the neural network, the numerator is related to the predictions and labels, while the denominator is related to the number of activations in each mask, respectively, which has the effect of normalizing the loss according to the dimension of the label mask.

For each class of mask, a Dice loss is computed :

$$1 - \frac{2 \sum_{\text{pixels}} y_{\text{true}} y_{\text{pred}}}{\sum_{\text{pixels}} (y_{\text{true}}^2 + y_{\text{pred}}^2)}$$

The Dice loss for each class is summed and averaged to get the final Dice soft loss.

6.2.1.3 Mean Squared Error Loss

MSE loss is the most commonly used loss function in machine learning and deep learning regression tasks, also known as L2 Loss. Its basic form is as follows :

$$J_{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

In fact, under certain assumptions, we can use the maximum likelihood to get the form of the mean squared loss. Assuming that the error between the model prediction and the true value obeys the standard Gaussian distribution ($\mu = 0, \sigma = 1$), then given a x_i , the probability that the model will output the true value y_i is :

$$p(y_i | x_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_i - \hat{y}_i)^2}{2}\right)$$

Further, we assume that the N sample points in the data set are independent of each other, then given the probability of all x outputting all true values y , the likelihood is the cumulative product of all $p(y_i | x_i)$.

For computational convenience, we usually maximize the log-likelihood :

$$LL(x, y) = \log(L(x, y)) = -\frac{N}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Removing the first term irrelevant to \hat{y}_i , and then converting to minimize the negative log-likelihood, we obtain :

$$NLL(x, y) = \frac{1}{2} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

This loss is a form of mean squared loss. Therefore, under the assumption that the error between the model output and the true value follows a Gaussian distribution, minimizing the mean square error loss function is essentially consistent with the maximum likelihood estimation. Therefore, in the scenario where this assumption can be satisfied, the MSE loss is a good choice of the loss function.

In our project, we use MSE to calculate the consistency cost from the student's prediction and the teacher's prediction on the unlabeled dataset.

6.2.2 Main Idea and Realization of Our Proposed Loss Function.

The weak label of the vascular network has been made with constant width bands and thus does not take into account the local geometry [Xiao et al., 2021]. As shown in Figure 6.5, compared with the real mask, the weakly label is sometimes wider than the actual width of the vessel and sometimes narrower. The segmentation is the prediction mask of learning vessel features with weak labels, but its width is not always consistent. In order to figure out the accuracy of the segment model, we calculate the overlap ratio between the real mask and segmentation (0.712). It is even bigger than the overlap ratio between the real mask and weakly label (0.568). That means our predictions were successful with regard to vascular features. However, if the width information is considered, the accuracy is limited. Since our previous studies have shown that RCC subtypes can be well classified by only the information extracted from the vascular skeleton (without any width information), the classical loss functions like Dice loss are not adapted here due to false positive and negative pixels that may appear at the border of a vessel due to a lack of precision of the manually labeled image but not to the obtained segmentation. To overcome this imprecise ground truth we proposed a new loss function to restrict the results in terms of vessel detection, basically considering length but not the width of vessels. We dilated (with a disk of radius 3, according to experiment Table 6.2 the segmentation result from S to obtain DS and the ground truth GT to obtain DGT . We computed the ratio of miss-detected vessels as :

$$MV = \frac{|\{(i, j) : GT(i, j) = 1, DS(i, j) = 0\}|}{|\{(i, j) : GT(i, j) = 1\}|} \quad (6.1)$$

and the ratio of falsely detected vessels as :

$$VF = \frac{|\{(i, j) : S(i, j) = 1, DGT(i, j) = 0\}|}{|\{(i, j) : S(i, j) = 1\}|} \quad (6.2)$$

Finally, we defined the following global performance index :

$$IV = 1 - (MV + VF)/2 \quad (6.3)$$

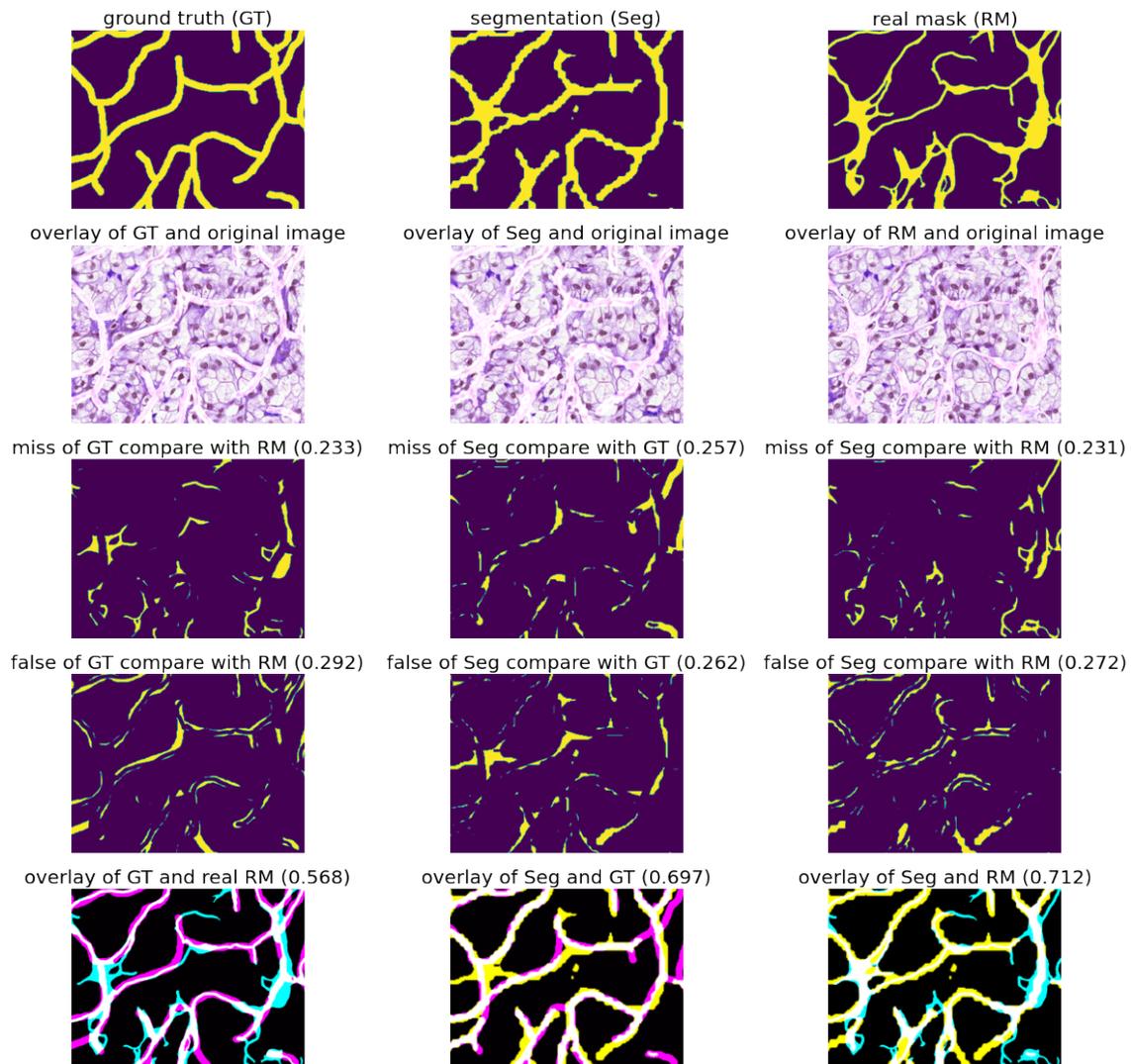


Figure 6.5 – Calculate the missing part and the false part between the generated segmentations, weakly labels, and real vascular masks by comparing them.

For the overlay images in the last line : the cyan color represent the real mask, the magenta color represent the ground truth (weakly mask), the yellow color represent the segmentation by our proposed MTL-SSL model, the white color represent the overlap part.

The loss value is $(MV + VF)/2$.

As shown in Figure 6.5, the generated vascular segmentation is even closer to the real vascular mask than the weakly ground truth. Generated segmentation has a smaller MV (0.231) than ground truth (0.233) and also a smaller FV (0.272) than ground truth (0.292) after overlap with the real mask. This loss function will be used in our CNN for segmentation but also as an evaluation method.

6.3 Vascular Network Segmentation Benchmark of RCC

6.3.1 Experimental Environment

The computing platform we used is Nef (https://wiki.inria.fr/ClustersSophia/Clusters_Home), which is Inria Sophia Antipolis Méditerranée research center multidomain scientific experimentation infrastructure for computation, storage and visualization. The platform relies on a cluster computing approach (several generations of high-performance servers are combined to provide heterogeneous parallel architecture), fast network interconnect, high capacity storage and inter-connection to Inria Sophia interactive visualization platform.

The experimental software in this project is built with the deep learning framework of Pytorch, the programming language is Python, and the version is Python3.8.5. The operating system is CentOS Linux release 7.4.1708 (Core). The hardware platform is Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz with 2 physical CPUs and 40 logical CPUs, 10 cores per physical CPU, and 2 threads per core.

6.3.2 Backbone Choice

We conducted experiments on different supervised classification models (GoogLeNet [Szegedy et al., 2015], ShuffleNet [Zhang et al., 2018b], VggNet [Simonyan and Zisserman, 2014], ResNet [He et al., 2016] and HRNet [Sun et al., 2019]) and competitive SSL models (Cross Pseudo Supervision (CPS) [Chen et al., 2021], Cross Consistency Training (CCT) [Ouali et al., 2020], Entropy Minimization (EM) [Vu et al., 2019] Deep Co-Training (DCT) [Qiao et al., 2018] and Mean Teacher (MT) [Tarvainen and Valpola, 2017]) to select the more efficient backbone. For the classification, to avoid using the data of the same patient for both training and testing, we split our database into train input with 18624 tumor patch images (8913 of ccRCC, 9079 of pRCC and 632 of chRCC), validation with 4843 tumor patch images (2049 of ccRCC, 2523 of pRCC and 271 of chRCC), and test with 6973 tumor patch images (4420 of ccRCC, 2250 of pRCC and 303 of chRCC). For training input of SSL segmentation, we used both the labeled vascular masks of our dataset with 335 tumor patch images (112 of ccRCC, 111 of pRCC and 112 of chRCC) and the 1005 unlabeled data from the RCC dataset randomly, which is 3 times the labeled data, so we had 1340 training input in total. Meanwhile, the validation of SSL segmentation with 32 tumor patch images (8 of ccRCC, 9 of pRCC and 15 of chRCC) and test with 57 tumor patch images (9 of ccRCC, 9 of pRCC and 39 of chRCC) were selected randomly. All the experiments have been repeated 5 times. The mean and standard deviation of the different model results are shown in Table 6.1. HRNet backbone performed best in both classification (0.9369) and SSL segmentation tasks (0.4614 *vs* 0.4537). We found that the results of the single SSL segmentation method are not ideal, and even the best segmentation result did not reach 0.5. So we try the multi-task approach to improve the segmentation results. The results of most SSL

methods have been greatly improved after the joint classification task. And MTL-SSL based on mean teacher reaches the best performance among all the segmentation methods (0.7979).

All the code and parameters of SSL models used were from the open-access repository SSL4MIS [Luo, 2020].

TABLE 6.1 – Performance of different backbones.

Methods	Backbone	Accuracy			
Classification	GoogLeNet	0.9348(0.01)			
	ShuffleNet	0.7753(0.07)			
	VggNet	0.9114(0.01)			
	ResNet	0.8863(0.03)			
	DenseNet	0.762(0.001)			
	HRNet	0.9369(0.03)			
Method	Backbone	MV(↓)	FV(↓)	IV(↑)	
SSL Segmentation	DCT (ECCV 2018)	UNet	0.3770(0.02)	0.7453(0.05)	0.4388(0.02)
	EM (CVPR 2019)	UNet	0.3340(0.04)	0.7686(0.01)	0.4487(0.02)
	CCT (CVPR2020)	Single UNet	0.3644(0.03)	0.7474(0.03)	0.4417(0.004)
	CPS (CVPR 2021)	Task UNet	0.3459(0.01)	0.7467(0.01)	0.4537(0.01)
	MT (NIPS 2017)	UNet	0.3622(0.002)	0.7827(0.002)	0.4275(0.001)
	DCT (ECCV 2018)	HRNet	0.2926(0.02)	0.7846(0.01)	0.4614(0.004)
	EM (CVPR 2019)	HRNet	0.3049(0.01)	0.7844(0.01)	0.4554(0.01)
	CCT (CVPR2020)	Single HRNet	0.2842(0.02)	0.7951(0.01)	0.4604(0.01)
	CPS (CVPR 2021)	Task HRNet	0.3190(0.02)	0.7733(0.01)	0.4539(0.01)
	MT (NIPS 2017)	HRNet	0.2934(0.02)	0.7932(0.01)	0.4567(0.01)
	DCT (ECCV 2018)	HRNet	0.3032(0.01)	0.1073(0.02)	0.7948(0.01)
	EM (CVPR 2019)	HRNet	0.1307(0.04)	0.2893(0.06)	0.7900(0.01)
CCT (CVPR2020)	Multi HRNet	0.1562(0.05)	0.4189(0.05)	0.7209(0.05)	
CPS (CVPR 2021)	Task HRNet	0.2142(0.03)	0.8455(0.01)	0.4702(0.02)	
MT (NIPS 2017) Our	HRNet	0.2798(0.02)	0.1243(0.03)	0.7979(0.01)	

6.3.3 Segmentation Benchmarks of Vascular Network

We conducted benchmark experiments on supervised learning, SSL, single segmentation task and MTL. For SSL, the split ratios were the same as in the backbone choice experiment : 335 labeled and 1005 unlabeled data for training input, 32 for validation and 57 for the test. For supervised learning, we only used the labeled vascular masks of our dataset, 335 labeled data for training input, 32 for validation and 57 for the test. All the data were selected randomly according to the patient. For the common parameter setup, the input size was 512*512 pixels, the optimizer was Adam. We used batches of size 8, the epoch of 200 and a poly learning rate decay scheme. The initial learning rate was 0.002 and weight decay was 1×10^{-6} . For the parameter setup specific to SSL, the error moving average decay was 0.99, the consistency type was mean square error, the consistency was 0.1, and the consistency ramp-up was 50.

Table 6.3 gives a quantitative evaluation. The proposed MTL-SSL model, which combines a classification task and a semi-supervised segmentation task, reaches the best performance (0.7979) among all the experiments (0.4567 ~ 0.7925). The accuracy of classification also improved from

TABLE 6.2 – Evaluate with different Strel Radius.

Radius	MV(↓)	FV(↓)	IV(↑)
1	0.1353(0.02)	0.3514(0.06)	0.7566(0.02)
2	0.1208(0.02)	0.3361(0.06)	0.7716(0.02)
3 (Our)	0.2798(0.02)	0.1243(0.03)	0.7979(0.01)
4	0.1040(0.02)	0.3205(0.06)	0.7878(0.02)
5	0.0835(0.02)	0.3426(0.04)	0.7870(0.01)

TABLE 6.3 – Performance of different models.

	Supervised	MV(↓)	FV(↓)	IV(↑)
	single-task	0.2551(0.02)	0.1600(0.03)	0.7924(0.01)
	multi-task	0.3119(0.05)	0.1032(0.01)	0.7925(0.02)
Segmentation	Semi-Supervised	MV(↓)	FV(↓)	IV(↑)
	single-task	0.2934(0.02)	0.7932(0.01)	0.4567(0.01)*
	multi-task (Our)	0.2798(0.02)	0.1243(0.03)	0.7979(0.01)
		Accuracy		
Classification	single-task	0.8236 (0.06)		
	multi-task (Our)	0.8586 (0.05)		

TABLE 6.4 – Performance of differently labeled data in training.

		Labeled=200	Labeled=300	Labeled=all (335)
Supervised	MV(↓)	0.091(0.01)	0.1067(0.02)	0.2551(0.02)
Task	FV(↓)	0.4120(0.02)	0.3670(0.03)	0.1600(0.03)
	IV(↑)	0.7483(0.01)	0.7631(0.01)	0.7924(0.01)
	MV(↓)	0.1100(0.02)	0.1046(0.004)	0.2798(0.02)
Multi-task (Our)	FV(↓)	0.3580(0.04)	0.3302(0.03)	0.1243(0.03)
	IV(↑)	0.7660(0.01)	0.7826(0.02)	0.7979(0.01)

0.8236 with the single task to 0.8586 with the multi-tasks. And Table 6.4 shows that the smaller the scale of the labeled data, the more advantages of our model, the difference of IV between our proposed model and the fully supervised method is from 0.0055 to 0.0195.

6.4 Test on New Subtype of RCC and Other Type of Cancers

6.4.1 New Subtype of RCC and Other Cancers

In our project, to validate our models, we collected patch images of Oncocytoma, a new subtype of RCC, from our database and also download the diagnosis slice of breast cancer, liver cancer, lung cancer and esophagus cancer from TCGA (<https://portal.gdc.cancer.gov/>) to test our model, so here I briefly introduce the pathological features of these five cancers. Some examples of histopathological images of these cancers are shown in Figure 6.7.

6.4.1.1 Oncocytoma

Renal oncocytoma accounts for approximately 3 – 7% of RCC. Histologically, tumor cells with finely granular cytoplasm proliferate in edematous, myxomatous, or hyalinoid stroma with a nested, tubular-cystic, solid, or trabecular pattern [Kuroda et al., 2003]. Oncocytoma has been recognized as a distinct subtype of benign kidney tumor for decades [Klein and Valensi, 1976, Moch et al., 2016]. Imaging techniques currently used cannot reliably distinguish oncocytoma from malignant lesions; therefore, patients must undergo resection or biopsy for a definitive diagnosis. The adjunctive use of pathological features and immunostaining helps differentiate oncocytoma from other renal tumors characterized by granular, eosinophilic cytoplasm, especially chromophobe renal cell carcinoma [Schatz and Lieber, 2003]. Despite the evaluation of many biomarkers over the years, the pathological diagnosis of oncocytoma and differentiation from its mimics remains a challenge [Wobker and Williamson, 2017]. An example of the histopathological image of oncocytoma is shown in Figure 6.6.

Nuclei are characterized by roundness and regularity, however, it has also been demonstrated that oncocytomas may contain areas of "degenerative" cytological atypia, resulting in plaques of tumor cells with large nuclei, irregular nuclear outlines, and chromatin smears [Trpkov et al., 2010]. An unusual pattern is sometimes encountered in oncocytomas where areas of less cytoplasm are found, resulting in an appearance known as "small cell" oncocytoma, oncocytoma with pseudorosettes, or "oncoblastic" cells [Petersson et al., 2011].

Although oncocytomas are recognized as benign, vascular invasions have been reported for up to 5.4% [Davis, 1991]. However, it has low vascular density compared to other subtypes of RCC, as shown in Figure 6.6.

6.4.1.2 Breast Cancer

Breast cancer is the most common malignancy, with approximately 1.4 million women diagnosed with breast cancer in 2008, accounting for 23% of all new cancer cases [Jemal et al., 2011].

With the close integration of histopathology and clinical medicine, pathological types are gradually classified according to the degree of invasion of cancer cells to surrounding tissues and the possibility of distant metastasis roughly divided into non-invasive cancer, early invasive cancer and invasive cancer.

One of the hallmarks of cancer is the formation of new vascular, and vascular invasion is also considered to be an important step in breast cancer metastasis and is the main cause of morbidity, progression, and death of the disease. Determining the metastatic potential of primary tumors often requires testing vascular invasion. This process results in locally increased microvessel density (MVD) with dilated and tortuous abnormal vessels. The tumor vasculature was studied pathologically using microscopy on immunohistochemically (IHC)-stained tissue specimens [McDonald and Choyke, 2003], and quantification of microvessels is a prognostic parameter [Uzzan et al., 2004].

6.4.1.3 Liver Cancer

Hepatocellular carcinoma (HCC) is the most common subtype of liver cancer and is the fourth leading cause of cancer-related death and is currently the leading cause of liver-related death, killing more than 1 million people worldwide each year [Siegel et al., 2016, Miller et al., 2018, Bray et al., 2018]. The characteristics of HCC are influenced by various factors, such as tumor size, relevance to cirrhosis, tumor thrombus in the portal vein, intrahepatic metastasis, necrosis, and hemorrhage. Unlike cancers of other organs, HCC is often nodular with fibrous capsules and septa, and the presence of capsules and septa is considered a useful diagnostic clue for the imaging diagnosis of HCC [Kojiro, 2005].

In most cases, HCC can be differentiated by cytological atypia and structural abnormalities (eg, thickening of the liver plate), including endothelial lining, pseudo glandular configuration, tumor involvement of lymphatic vessels (more common in resected material), and portal veins and loss of hepatic lobules. Most HCCs show trabecular, nested, solid, or pseudoscalar growth patterns. Abnormal vessels can be identified within the lobules rather than the portal vein, which is often helpful in well-differentiated HCC. Bile production is seen in a considerable number of tumor cells and is characteristic of hepatocyte differentiation. Well-differentiated HCC has abundant fine-grained eosinophilic cytoplasm, round nuclei, scattered chromatin, and prominent nucleoli. If the biopsy is indeed taken from a liver mass and shows hepatocytes on H&E slides, the differential list includes HCC and its analogs such as focal nodular hyperplasia (FNH), hepatic adenoma (HA), and DN. Tumor cells in HCC can also exhibit fatty, clear cellular changes, or prominent nuclear inclusions [Jiang et al., 2018].

HCC is detected as a hypervascular tumor on angiography and/or contrast-enhanced CT. Meanwhile, most of the small HCCs of the inapparent nodular type were uniformly distributed in the well-differentiated cancer tissues, and different numbers of portal vein bundles were preserved in the nodules. At the tumor/non-tumor boundary, well-differentiated HCC cells proliferated in a replacement growth pattern as if they were replacing hepatocytes in the surrounding liver tissue, and there was no envelope formation. In these tumors, tumor cell invasion into the portal vein retained within the tumor was observed and was defined as "stromal invasion" [Kojiro, 2005]. [Tomizawa et al., 1995]. Well-differentiated early-stage HCC has the following four main histological features : (1) increased cell density, increased nuclear/cytoplasmic ratio, and an irregular trabecular pattern; (2) increased eosinophilic or basophilic staining intensity; (3) frequent acinar and/or pseudo glandular patterns; (4) frequent fatty changes and/or marked changes in cancer cells. Underdevelopment of arterial tumor vessels (unpaired arteries) is thought to be a major cause of fat changes.

6.4.1.4 Lung Cancer

Lung cancer has the highest mortality rate of all cancers [Howlader et al., 2016]. The prognosis for lung cancer is poor and varies significantly based on tumor stage at diagnosis. Adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC) are the most common subtypes of lung cancer.

Two-thirds of squamous cell carcinomas present a central lung tumor, while many of the remaining third are peripheral [Travis, 2011]. [Funai et al., 2003] suggest that morphological features of squamous differentiation include intercellular bridging, squamous pearl formation, and single-cell keratinization. These features are evident in well-differentiated tumors; however, they are difficult to find in poorly differentiated tumors. Squamous cell carcinoma occurs most often in segmental bronchi and involves lobes and main bronchi [Carlile and Edwards, 1986]. Squamous cell carcinoma can have papillary, clear cell, small cell, and basal cell subtypes [Melamed et al., 1977, Churg et al., 1980]. Papillary squamous cell carcinoma often exhibits a pattern of exophytic endobronchial growth [Dulmet-Brender et al., 1986, Sherwin et al., 1962]. The diagnosis of LUAD relies entirely on histopathological features. Important information related to prognosis and diagnosis can be obtained from the histopathological features of LUAD [Barletta et al., 2010, Borczuk et al., 2009]. The pathological subtypes of LUAD (squamous, acinar, papillary, solid, and micropapillary) can lead to different recurrence and survival rates [Sica et al., 2010]. Meanwhile, novel histomorphological features of LUAD have been discovered, such as tumor budding, lymphovascular invasion, and tumor spread through air spaces [Higgins et al., 2012, Yamaguchi et al., 2010, Lee et al., 2020].

The vascular invasion proved to be an independent significant recurrence risk factor for lung cancer [Shimada et al., 2012]. In the study, [Yoo et al., 2010], histological and immunofluorescence studies showed that primitive blood vessels had distinct features from similar-looking lymphatic vessels. They have multiple channels surrounded by a loose collagen matrix, in stark contrast to the single-channel structure of other vascular systems. The rod-shaped nuclei lined longitudinally along the channel are considered to be provascular endothelial cells, but do not express LYVE-1, a specific marker for lymphatic vessels, indicating a distinct difference from lymphatic endothelial cells. Taken together these findings and the characterization of novel linear vascular structures in cancer models may have important implications for cancer prognosis and therapy.

6.4.1.5 Esophageal Cancer

Esophageal cancer is one of the least studied and deadliest cancers worldwide, with the sixth highest mortality rate of all cancers [Enzinger and Mayer, 2003]. Esophageal cancer affects more than 450,000 people worldwide and the incidence is increasing rapidly [Pennathur et al., 2013]. More than 90% of esophageal cancers are squamous cell carcinomas or adenocarcinomas. About three-quarters of adenocarcinomas occur in the distal esophagus [Daly et al., 2000], whereas squamous cell carcinomas are more evenly distributed between the middle and lower thirds. The cervical esophagus is a rare disease site that is extremely aggressive and has poor survival prognostic [Mao et al., 2011, Enzinger and Mayer, 2003].

Squamous cell carcinoma arises from the stratified squamous epithelial lining of the organ, while adenocarcinoma affects the columnar gland cells that replace the squamous epithelium [Blot et al., 1993]. Sarcomas and small cell carcinomas typically account for less than 1% – 2% of all esophageal cancers [Young Jr et al., 1981, Kwatra et al., 2003]. In rare cases, other carcino-

mas, melanoma, leiomyosarcoma, carcinoid, and lymphoma may also develop in the esophagus [Blot et al., 1993].

Histopathological evaluation is based on the invasion depth (vertical tumor invasion depth), tumor size, lymphatic invasion, venous invasion, lymph node metastasis, tumor differentiation, growth pattern, degree of nuclear atypia and histological grade [Tajima et al., 2000]. The distance from the invasive front to the muscular mucosa is measured using microscopy with an ocular lens scale and defined as vertical tumor invasion depth. In specimens where the muscular mucosa is disrupted by ulceration or tumor invasion, the muscular mucosa level is estimated by drawing a line to connect the remaining muscular mucosa [Eguchi et al., 2006]. Tumor differentiation is evaluated according to the World Health Organization classification [Travis et al., 2004]. The degree of nuclear atypia is classified as low (sharp nuclei of uniform size with homogeneous chromatin : one point) or high (polymorphic nuclei with vesicular or dot-like chromatin : two points). The growth pattern is classified as expansive (large solid nests with a clear tumor margin : one point) or infiltrative (small nests and dissociated cells with an unclear tumor margin : two points). The degree of nuclear atypia and growth pattern are evaluated at the deepest area of the invading tumor. The histological grade of each tumor is calculated by adding the nuclear atypia and growth pattern scores to obtain Grade 1 : two points; Grade 2 : three points; or Grade 3 : four points [Eguchi et al., 2006]. Meanwhile, vessel permeation, vertical tumor invasion depth in the submucosal layer, and histological grade are important factors in determining patients who do not require additional surgical treatment after endoscopic mucosal resection [Tajima et al., 2000].

In squamous epithelium, the microvascular pattern of intracapillary loops is the only reliable indicator of tissue atypia [Inoue et al., 2015]. Histopathological features of superficial esophageal cancers can be diagnosed by evaluating microvascular patterns on magnifying endoscopy [Arima et al., 2005]. The size of avascular areas and associated type 4 vessels can be used to assess the extent and depth of tumor invasion. The size of avascular areas is related to the depth of tumor invasion. Four types of Microvascular patterns are defined as follows : Type 1 is characterized by thin, linear capillaries in the subepithelial papilla and is generally seen in normal mucosa. Type 2 is characterized by distended, dilated vessels, and the shape of capillaries in the subepithelial papilla is preserved. Type 2 is generally seen in inflammatory lesions. Type 3 is characterized by spiral vessels with an irregular caliber and crushed vessels with red spots, and the arrangement of the vessels is irregular. Type 4 is characterized by multilayered, irregularly branched, reticular vessels with an irregular caliber. Type 4 is generally seen in cancers with deeper invasion. Avascular areas and stretched type 4 vessels are seen in cancers with downward growth.

6.4.2 Test on New Subtype of RCC and Other Cancers Dataset

Figure 6.8 shows the segmentation results of the proposed MTL-SSL and fully supervised models on our own database, which include subtypes within the training set and new subtypes without training, it reveals that our MTL-SSL performs better and especially has smaller false detection than the fully supervised model. Figure 6.9 shows the segmentation results of MTL-SSL on the TCGA dataset. And Table 6.5 shows that our MTL-SSL model is more robust than fully supervised segmentation when considering with new subtype (0.7867 vs 0.7067) and other dataset tests (0.8786 vs 0.8691 in TCGA-RCC; 0.7418 vs 0.6965 in TCGA-BRCA; 0.8127 vs 0.7551 in TCGA-LIHC; 0.8080 vs 0.7644 in TCGA-LUSC; 0.7959 vs 0.7737 in TCGA-ESCA). When testing a new cancer type that is not involved in training, traditional supervised segmentation methods are more intended to over-segment, so they have lower MV scores but higher FV, thus

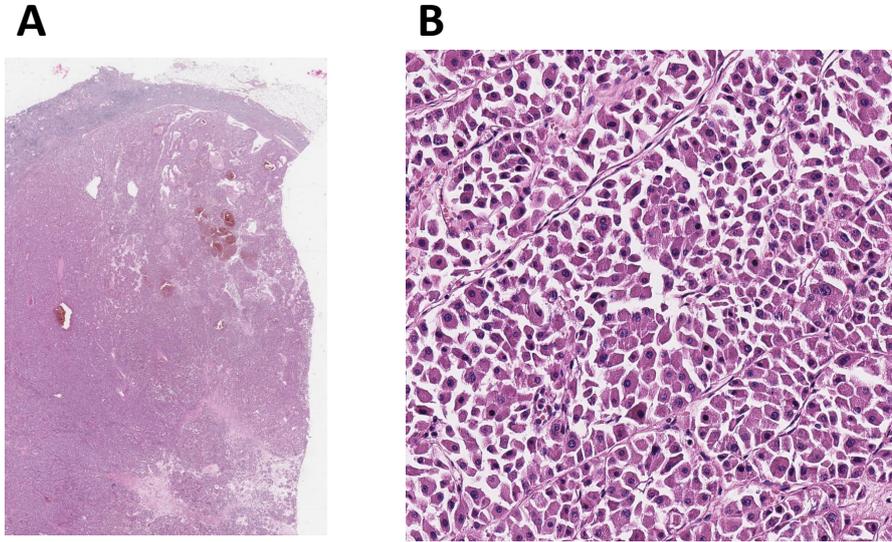


Figure 6.6 – Hematoxylin and eosin staining for oncocytomas at different magnification. A : whole slice image, $\sim 60000 \times 60000$ Pixels ; B : patch image, 2000×2000 Pixels.

TABLE 6.5 – Performance of new subtype of RCC and other cancers.

Task	Dataset	MV(\downarrow)	FV(\downarrow)	IV(\uparrow)
Supervised Task	Our-RCC (new subtype)	0.1366(0.03)	0.4501(0.04)	0.7067(0.01)
	TCGA-RCC (2 subtypes)	0.087(0.01)	0.1746(0.03)	0.8691(0.01)
	TCGA-BRCA (breast)	0.2179(0.05)	0.3892(0.21)	0.6965(0.10)
	TCGA-LIHC (liver)	0.1720(0.02)	0.3177(0.03)	0.7551(0.02)
	TCGA-LUSC (lung)	0.1865(0.02)	0.2847(0.04)	0.7644(0.02)
	TCGA-ESCA (esophagus)	0.1734(0.02)	0.2792(0.03)	0.7737(0.01)
Semi-Supervised Multi-Task (Our)	Our-RCC (new subtype)	0.1631(0.02)	0.2635(0.04)	0.7867(0.01)
	TCGA-RCC (2 subtypes)	0.068(0.01)	0.1746(0.02)	0.8786(0.01)
	TCGA-BRCA (breast)	0.1863(0.03)	0.3301(0.1)	0.7418(0.03)
	TCGA-LIHC (liver)	0.1483(0.03)	0.2262(0.08)	0.8127(0.03)
	TCGA-LUSC (lung)	0.1930(0.02)	0.1910(0.07)	0.8080(0.03)
	TCGA-ESCA (esophagus)	0.1663(0.02)	0.2419(0.08)	0.7959(0.03)

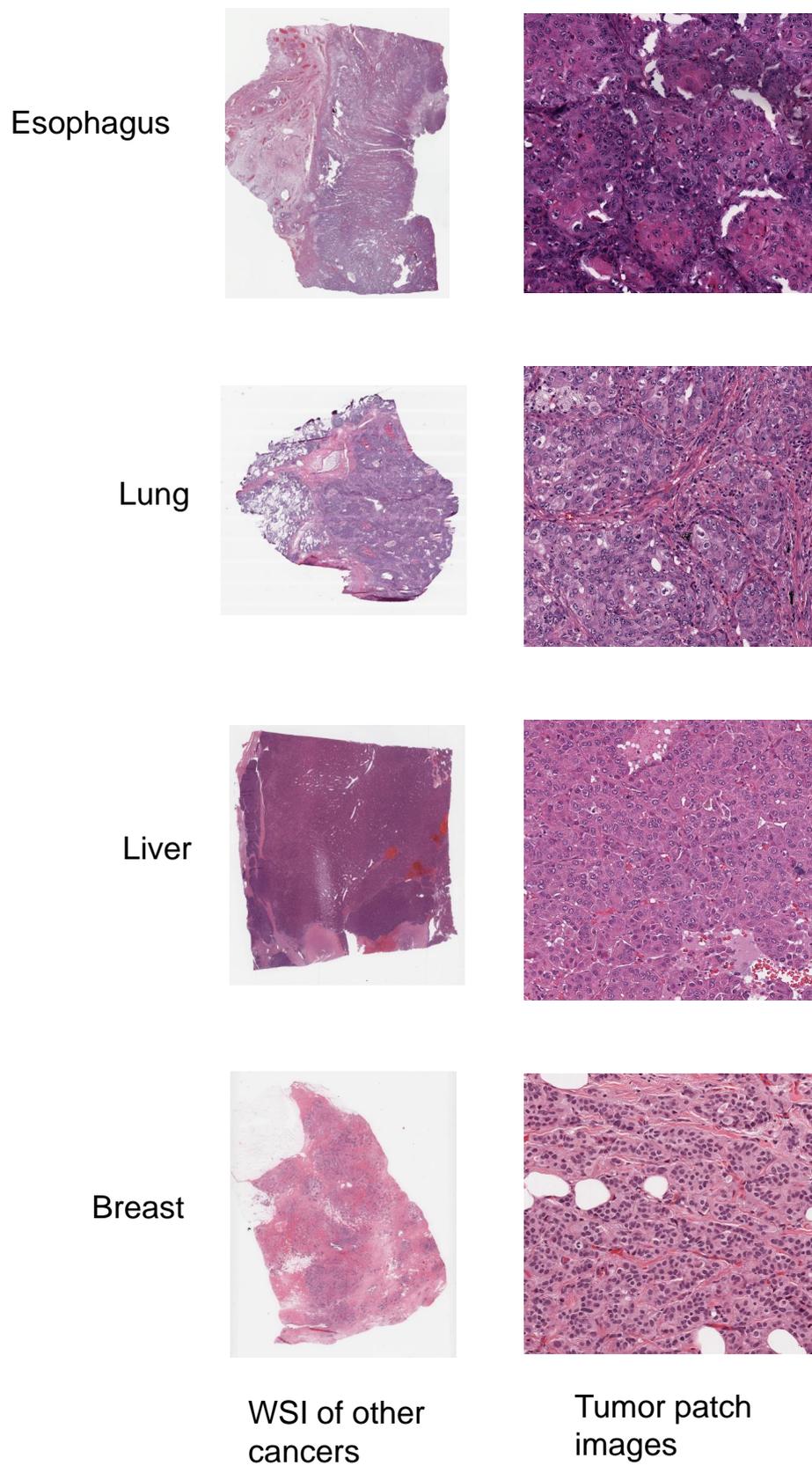


Figure 6.7 – H&E staining for other cancers at different magnification. Their vascular morphology is quite different.

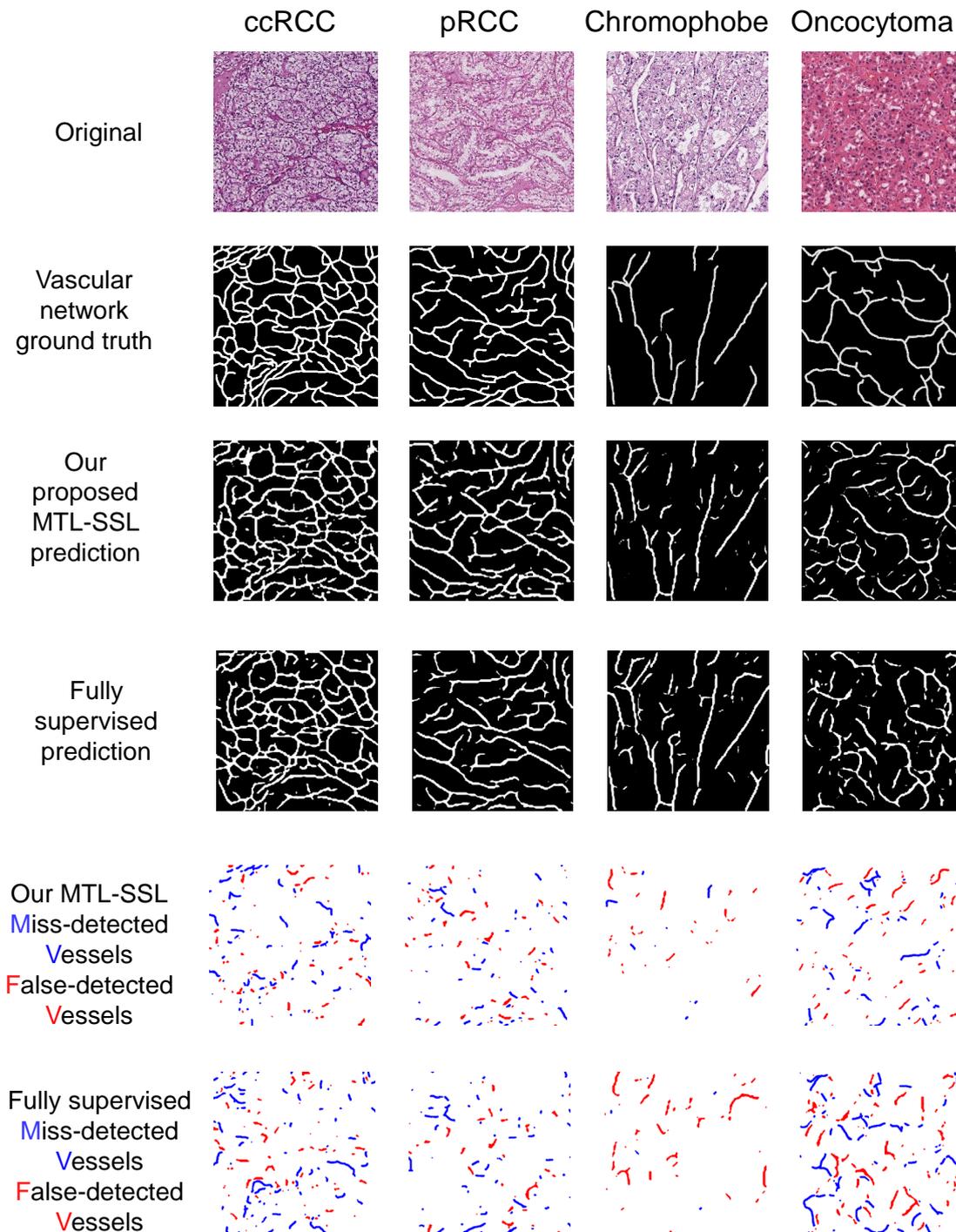


Figure 6.8 – Segmentation results of the proposed MTL-SSL and fully supervised models for the four subtypes of RCC.

The four subtypes of RCC include cRCC, pRCC, chRCC, and the new subtype Oncocytoma which is without training. (The blue color represents the miss detected part of the vessels and the red color represents the false detected part of the vessels.)

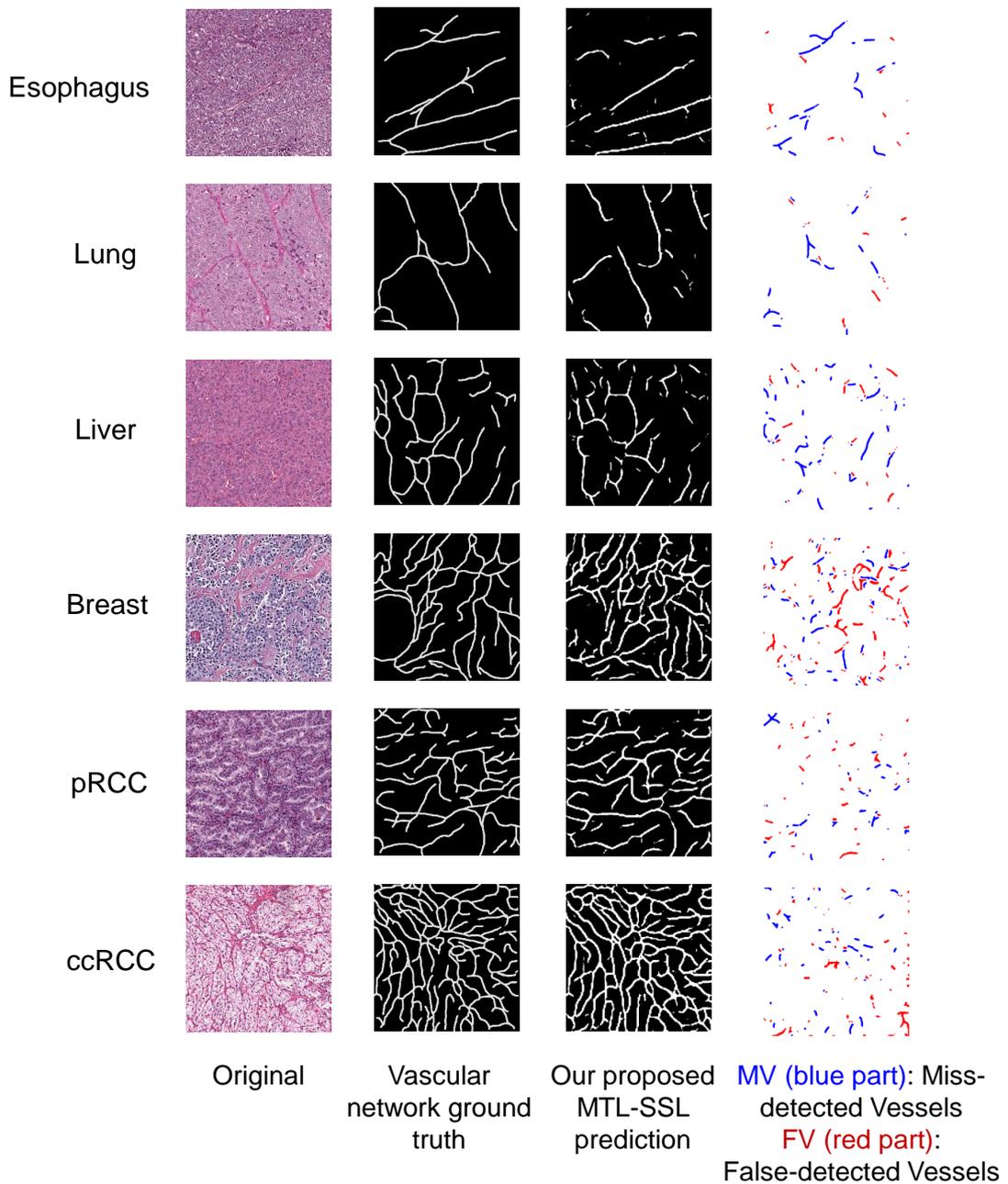


Figure 6.9 – Segmentation results of the proposed MTL-SSL model on the TCGA dataset.

affecting the final segmentation result. Our MTL-SSL model appears to be versatile with respect to vascular segmentation tasks, it has the potential to segment vascular from other subtypes of RCC and even other cancers without adding manual vascular masks for training. This provides a solid foundation for the study of the vascular network in H&E stained histopathological images of cancers, which is no longer limited to immunostaining and artificial labeling.

Furthermore, segmentation of vascular networks in histopathology images is very challenging, in this context, our MTL-SSL model has improved state of art by favorable HRNet backbone, loss strategy, and MTL with classification.

6.5 Conclusion

We propose an MTL-SSL model performing joint SSL segmentation and classification tasks to segment the vascular network using both labeled and unlabeled data. It can reduce the reliance on manually segmented vascular network masks to some extent and can achieve automatic segmentation.

We apply the first automatic, end-to-end vascular network segmentation method in challenging H&E staining histopathological images. And in our experiments, this MTL-SSL model obtained the best performances both in segmentation and classification tasks among fully supervised learning, SSL, multi, single segmentation tasks. It is also versatile on the new RCC subtype and other cancers testing. Meanwhile, the MTL-SSL model can outperform the fully supervised in both our own dataset and other TCGA datasets. That clarified applying the HRNet backbone-based multi-task model (jointly with an SSL principle segmentation and fully supervised classification) to the vascular network segmentation of histopathological images is valuable.

We conducted plentiful experiments for backbone selection, HRNet performs best among them, which may be because HRNet is the connection of multi-resolution sub-networks. HRNet can maintain high-resolution feature maps, so that more suitable for segmentation tasks with pixel-imbalanced binary images. For other networks, their strong semantic information is obtained by downsampling, and then upsampling to restore high-resolution recovery position information. However, this approach will cause a large amount of effective information to be lost in the continuous up-and-down sampling process. HRNet achieves the purpose of strongly semantic information and precise location information at the same time by paralleling multiple resolution branches and continuously performing information interaction between different branches, so that easier to capture the details of vascular networks by HRNet.

We choose the mean teacher method from a bunch of competitive semi-supervised segmentation networks. The advantage of the mean teacher depends on its average weight method, which can update the teacher model at each training step and guide the learning of the student model in a timely manner. Additionally, the model has a better intermediate representation because weight averaging improves the output of all layers, not just the top output. These aspects have two practical advantages in terms of temporal aggregation : First, more accurate target labels lead to a faster feedback loop between the student model and the teacher model, which improves test accuracy. Second, the method is scalable to large datasets and online learning.

To adapt our weak labeling method, we proposed a new evaluation method. We computed the ratio of miss-detected vessels(MV), the ratio of falsely detected vessels (FV), and the global performance index (IV), this evaluation method can help segment the vascular network, which more focus on the length of vessels rather than the width. That's because we have already proved

that extracting features only with the length of the vessels and the network connectivity can classify RCC subtypes effectively.

Our study lays the foundation for exploring the importance of vascular geometry and topological morphology in different cancers. However, our research still has limitations. We use the weak labeling method to save annotation time so that our method can't adapt to some tasks which need exact segmentation of the vascular network. We have only seen its application in H&E stained histopathological images so far. In other types of images, vascular networks play an important role, such as OCT images. In order to expand its automatic segmentation value, we should try to apply it to different kinds of images.

Conclusion and Perspectives

Conclusion and Perspectives

7.1 Conclusion

The classification of RCC subtypes is of great value for RCC diagnosis, treatment, and prognosis. Due to the many similarities in the tissue morphology of each subtype, the subtype classification task of RCC based on H&E staining has always been a challenging task in the medical field. In this project, instead of distinguishing subtypes, as usually done by cell morphology, we present the first work to investigate the importance of the vascular network characters (geometric and topological properties) for RCCs classification.

However, there are not enough open-access datasets of histopathological images of RCC. In order to overcome this problem and achieve our goal, we established a BigRCC database with 62220 patch images from ccRCC, pRCC and chRCC. Then we manually annotated the vascular network of the 424 tumor patch images from the BigRCC to form the VRCC dataset to facilitate the extraction of the hand-crafted features of the vascular network, which include skeleton and lattice features.

The results of feeding these hand-crafted features into a machine learning model outperform the results of graph features based on GCN models that only include the topological character of vascular, and also beat the deep features based on popular deep learning approaches. This demonstrates that we can distinguish RCC subtypes only by vascular networks. Since our model is based on the traditional machine learning models, the proposed hand-crafted features, which reach the accuracy of 0.9467 in our research, can also provide good results on a small dataset and reduce the dependence on a large amount of data.

Although our classification algorithm reaches excellent performance, it is time-consuming and based on manual vascular network annotations. Therefore, its application is limited to a certain extent. To solve this issue, we propose an end-to-end model to segment the vascular networks automatically. This model is an MTL-SSL model performing joint SSL segmentation and classification tasks using both labeled and unlabeled data, which even can reach considerable results based on only limited data.

Usually, the segmentation of blood vessels based on H&E staining is very complicated, because the background includes many structures similar to the vascular, such as linear structures of cell membranes and fibers. Most of the previous vascular segmentation studies were based on immunostaining, our research is the first vascular network segmentation study based on H&E staining.

To prove the robustness of our MTL-SSL model, we also labeled a new subtype of RCC (Oncocytoma) from our database and downloaded data on RCC subtypes and other cancers (breast cancer, lung cancer, liver cancer and esophagus cancer) from the open-access TCGA database.

Our proposed MTL-SSL model reaches the best performance among all the experiments, it can outperform the fully supervised method in our own dataset (0.7979 *vs* 0.7925), new subtype testing (0.7867 *vs* 0.7067), and other data from TCGA (0.8786 *vs* 0.8691 in RCC; 0.7418 *vs* 0.6965 in breast cancer; 0.8127 *vs* 0.7551 in liver cancer; 0.8080 *vs* 0.7644 in lung cancer; 0.7959 *vs* 0.7737 in esophagus cancer). This shows that compared to the fully supervised approach, our model can not only reduce the reliance on a huge labeled data required through the SSL method but also demonstrates its excellence and robustness in tests on new subtypes and other datasets. Meanwhile, it forms a foundation for future developments in multi-task learning dealing with vascular segmentation and classification from histopathological images.

7.2 Perspectives

Although computer-aided pathology has received extensive attention in the fields of medicine and oncology, there are still many deficiencies in its clinical practical application. First, differences in the baseline experiments of the study designs, the features extracted, and the machine learning methods employed can make it difficult to compare multiple methods; second, due to the lack of open-access databases of histopathological images, most of the machine learning algorithms used in these studies are based on the researchers' own data. If there is no external validation, the generalizability and reproducibility of the results cannot be applied to other datasets or populations; Meanwhile, the repeatability, reproducibility, sample size, statistical power, and standardization are also important factors to be considered in future research. Finally, the transition from experimental research tools to basic clinical applications will be a challenge if data heterogeneity is not considered in future studies.

There are many kinds of traditional hand-crafted features, and most of them are obtained by summarizing a large amount of prior knowledge. The design of features requires a certain basic knowledge of histological morphology. The image preprocessing process also has a great impact on the evaluation results of hand-crafted features, so how to preprocess the images and select appropriate features to classify RCC subtypes is a difficult task. Our proposed hand-crafted features not only meet the above requirements but are also validated by statistical methods. Their classification performance is even better than that of deep learning methods, which provides a solid benchmark result for classifying RCC subtypes by vascular networks, and also provides a reference for subsequent related studies.

Some conventional wisdom is that large amounts of data underpin the development of AI actually underestimates the potential of machine learning on small datasets. These new techniques do not require large labeled datasets. An artificial intelligence approach based on small data numbers is suitable for situations where the scale of data is small or when not enough labeled data are available, which reduces the reliance on large data collection. Our study demonstrates that effective methods can still achieve comparable or even better results than fully supervised deep learning methods on small datasets for classification and segmentation tasks on limited datasets. Furthermore, the research on multi-task and semi-supervised learning is progressing rapidly, and it would be more effective and more widely applied in the medical field in the future.

Even though our MTL-SSL has constructive results for vascular segmentation, especially in the RCC test of TCGA data (0.8786), it still has a big space to improve. Further investigations will concern the adjustment of the labeled and unlabeled data ratio, the regulation of the loss function and the improvement of the edge-preserving regularization model used as segmentation post-processing, or the involvement of an attention module in the proposed MTL-SSL framework.

From the medical perspective, to further improve our work, we can also consider more markers from other immunostaining, not only from H&E staining images; but study the tumor micro-environment as well, such as surrounding blood vessels, immune cells, fibroblasts, bone marrow-derived inflammatory cells, various signaling molecules and extracellular matrix; graph modeling with the nucleus, then do the interact analysis between the vascular skeleton graph and nucleus graph.

Bibliography

- [Abrahams et al., 2004] Abrahams, N., MacLennan, G., Khoury, J., Ormsby, A., Tamboli, P., Doglioni, C., Schumacher, B., and Tickoo, S. (2004). Chromophobe renal cell carcinoma : a comparative study of histological, immunohistochemical and ultrastructural features using high throughput tissue microarray. *Histopathology*, 45(6) :593–602.
- [Adams et al., 2008] Adams, K. F., Leitzmann, M. F., Albanes, D., Kipnis, V., Moore, S. C., Schatzkin, A., and Chow, W.-H. (2008). Body size and renal cell cancer incidence in a large us cohort study. *American journal of epidemiology*, 168(3) :268–277.
- [Alom et al., 2020] Alom, M. Z., Aspiras, T., Taha, T. M., Bowen, T., and Asari, V. K. (2020). Mitosisnet : end-to-end mitotic cell detection by multi-task learning. *IEEE Access*, 8 :68695–68710.
- [Alom et al., 2019] Alom, M. Z., Yakopcic, C., Nasrin, M., Taha, T. M., Asari, V. K., et al. (2019). Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. *Journal of digital imaging*, 32(4) :605–617.
- [Aresta et al., 2019a] Aresta, G., Araújo, T., Kwok, S., Chennamsetty, S. S., Safwan, M., Alex, V., Marami, B., Prastawa, M., Chan, M., Donovan, M., et al. (2019a). Bach : Grand challenge on breast cancer histology images. *Medical image analysis*, 56 :122–139.
- [Aresta et al., 2019b] Aresta, G., Araújo, T., Kwok, S., Chennamsetty, S. S., Safwan, M., Alex, V., Marami, B., Prastawa, M., Chan, M., Donovan, M., et al. (2019b). Bach : Grand challenge on breast cancer histology images. *Medical image analysis*, 56 :122–139.
- [Arima et al., 2005] Arima, M., Tada, M., and Arima, H. (2005). Evaluation of microvascular patterns of superficial esophageal cancers by magnifying endoscopy. *Esophagus*, 2(4) :191–197.
- [Asare et al., 2020] Asare, S. K., You, F., and Nartey, O. T. (2020). A semisupervised learning scheme with self-paced learning for classifying breast cancer histopathological images. *Computational Intelligence and Neuroscience*, 2020.
- [Ball et al., 2015] Ball, M. W., Bezerra, S. M., Gorin, M. A., Cowan, M., Pavlovich, C. P., Pierorazio, P. M., Netto, G. J., and Allaf, M. E. (2015). Grade heterogeneity in small renal masses : potential implications for renal mass biopsy. *The Journal of urology*, 193(1) :36–40.
- [Ball, 2005] Ball, R. (2005). Pathology and genetics of tumours of the urinary system and male genital organs. *Histopathology*, 46(5) :586–586.
- [Barletta et al., 2010] Barletta, J. A., Yeap, B. Y., and Chirieac, L. R. (2010). Prognostic significance of grading in lung adenocarcinoma. *Cancer : Interdisciplinary International Journal of the American Cancer Society*, 116(3) :659–669.
- [Bejnordi and van der Laak, 2016] Bejnordi, B. E. and van der Laak, J. (2016). Camelyon16 : Grand challenge on cancer metastasis detection in lymph nodes.
- [Benomar et al., 2021] Benomar, M. L., Settouti, N., Xiao, R., Ambrosetti, D., and Descombes, X. (2021). Convolutional neuronal networks for tumor regions detection in histopathology

- images. In *International Conference on Digital Technologies and Applications*, pages 13–23. Springer.
- [Beucher, 1990] Beucher, S. (1990). *Segmentation d'images et morphologie mathématique*. PhD thesis, Ecole Nationale Supérieure des Mines de Paris.
- [Blot et al., 1993] Blot, W. J., Devesa, S. S., and Fraumeni, J. F. (1993). Continuing climb in rates of esophageal adenocarcinoma : an update. *Jama*, 270(11) :1320–1320.
- [Borczuk et al., 2009] Borczuk, A. C., Qian, F., Kazeros, A., Eleazar, J., Assaad, A., Sonett, J. R., Ginsburg, M., Gorenstein, L., and Powell, C. A. (2009). Invasive size is an independent predictor of survival in pulmonary adenocarcinoma. *The American journal of surgical pathology*, 33(3) :462.
- [Borkowski et al., 2019] Borkowski, A. A., Bui, M. M., Thomas, L. B., Wilson, C. P., DeLand, L. A., and Mastorides, S. M. (2019). Lung and colon cancer histopathological image dataset (lc25000). *arXiv*.
- [Bray et al., 2018] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018 : Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA : a cancer journal for clinicians*, 68(6) :394–424.
- [Breiman, 2001] Breiman, L. (2001). Random forests. *Machine learning*, 45(1) :5–32.
- [Breiman et al., 1984] Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and regression trees*. CRC press.
- [Bui, 2021] Bui, D. T. (2021). Applications of machine learning in ekyc's identity document recognition.
- [Bukenya et al., 2020] Bukenya, F., Nerissa, C., Serres, S., Pardon, M.-C., and Bai, L. (2020). An automated method for segmentation and quantification of blood vessels in histology images. *Microvascular Research*, 128 :103928.
- [Burlingame et al., 2018] Burlingame, E. A., Margolin, A. A., Gray, J. W., and Chang, Y. H. (2018). Shift : speedy histopathological-to-immunofluorescent translation of whole slide images using conditional generative adversarial networks. In *Medical Imaging 2018 : Digital Pathology*, volume 10581, pages 29–35. SPIE.
- [Cai and Wang, 2018] Cai, C. and Wang, Y. (2018). A simple yet effective baseline for non-attributed graph classification. *arXiv preprint arXiv :1811.03508*.
- [Campbell et al., 2017] Campbell, S., Uzzo, R. G., Allaf, M. E., Bass, E. B., Cadeddu, J. A., Chang, A., Clark, P. E., Davis, B. J., Derweesh, I. H., Giambarresi, L., et al. (2017). Renal mass and localized renal cancer : Aua guideline. *The Journal of urology*, 198(3) :520–529.
- [Cangea et al., 2018] Cangea, C., Veličković, P., Jovanović, N., Kipf, T., and Liò, P. (2018). Towards sparse hierarchical graph classifiers. *arXiv*.
- [Cao et al., 2008] Cao, H., Naito, T., and Ninomiya, Y. (2008). Approximate rbf kernel svm and its applications in pedestrian classification. In *MLVMA Workshop*.
- [Carlile and Edwards, 1986] Carlile, A. and Edwards, C. (1986). Poorly differentiated squamous carcinoma of the bronchus : a light and electron microscopic study. *Journal of Clinical Pathology*, 39(3) :284–292.
- [Caruana, 1997] Caruana, R. (1997). Multitask learning. *Machine learning*, 28(1) :41–75.

- [Chamanzar and Nie, 2020] Chamanzar, A. and Nie, Y. (2020). Weakly supervised multi-task learning for cell detection and segmentation. In *ISBI*, pages 513–516. IEEE.
- [Chapelle et al., 2009] Chapelle, O., Scholkopf, B., and Zien, A. (2009). Semi-supervised learning. *IEEE Transactions on Neural Networks*, 20(3) :542–542.
- [Chen et al., 2022] Chen, D., Miao, D., Zhao, C., and Zhou, H. (2022). Candidate region acquisition optimization algorithm based on multi-granularity data enhancement. *International Journal of Machine Learning and Cybernetics*, 13(7) :1847–1860.
- [Chen et al., 2021] Chen, X., Yuan, Y., Zeng, G., and Wang, J. (2021). Semi-supervised semantic segmentation with cross pseudo supervision. In *CVPR*, pages 2613–2622.
- [Cheng et al., 2020] Cheng, J., Han, Z., Mehra, R., Shao, W., Cheng, M., Feng, Q., Ni, D., Huang, K., Cheng, L., and Zhang, J. (2020). Computational analysis of pathological images enables a better diagnosis of tfe3 xp11. 2 translocation renal cell carcinoma. *Nature communications*, 11(1) :1–9.
- [Cheville et al., 2003] Cheville, J. C., Lohse, C. M., Zincke, H., Weaver, A. L., and Blute, M. L. (2003). Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma. *The American journal of surgical pathology*, 27(5) :612–624.
- [Chow and Devesa, 2008] Chow, W.-H. and Devesa, S. S. (2008). Contemporary epidemiology of renal cell cancer. *The Cancer Journal*, 14(5) :288–301.
- [Chow et al., 2000] Chow, W.-H., Gridley, G., Fraumeni Jr, J. F., and Järholm, B. (2000). Obesity, hypertension, and the risk of kidney cancer in men. *New England Journal of Medicine*, 343(18) :1305–1311.
- [Churg et al., 1980] Churg, A., Johnston, W. H., and Stulbarg, M. (1980). Small cell squamous and mixed small cell squamous–small cell anaplastic carcinomas of the lung. *The American journal of surgical pathology*, 4(3) :255–263.
- [Committee et al., 2010] Committee, D. S. et al. (2010). Digital imaging and communications in medicine (dicom) : Supplement 145 : Whole slide microscopic image iod and sop classes. *DICOM Standards*, pages 1–59.
- [Cover and Hart, 1967] Cover, T. and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1) :21–27.
- [Coy et al., 2017] Coy, H., Young, J. R., Douek, M. L., Brown, M. S., Sayre, J., and Raman, S. S. (2017). Quantitative computer-aided diagnostic algorithm for automated detection of peak lesion attenuation in differentiating clear cell from papillary and chromophobe renal cell carcinoma, oncocytoma, and fat-poor angiomyolipoma on multiphasic multidetector computed tomography. *Abdominal Radiology*, 42(7) :1919–1928.
- [Cruz-Roa et al., 2014] Cruz-Roa, A., Basavanahally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., and Madabhushi, A. (2014). Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In *Medical Imaging 2014 : Digital Pathology*, volume 9041, page 904103. SPIE.
- [Daly et al., 2000] Daly, J. M., Fry, W. A., Little, A. G., Winchester, D. P., McKee, R. F., Stewart, A. K., and Fremgen, A. M. (2000). Esophageal cancer : results of an american college of surgeons patient care evaluation study. *Journal of the American College of Surgeons*, 190(5) :562–572.

- [Das et al., 2021] Das, A., Devarampati, V. K., and Nair, M. S. (2021). Nas-rgan : A semi-supervised generative adversarial network model for atypia scoring of breast cancer histopathological images. *IEEE Journal of Biomedical and Health Informatics*.
- [Davis, 1991] Davis, C. (1991). Renal oncocyoma. clinicopathological study of 166 patients. *J. Urogenital. Pathol.*, 1 :41–52.
- [Debreuve, 2020] Debreuve (2020). <https://gitlab.inria.fr/edebreu/sklgraph>.
- [Denton et al., 2002] Denton, M. D., Magee, C. C., Ovuworie, C., Mauiyyedi, S., Pascual, M., Colvin, R. B., Cosimi, A. B., and Tolkoff-Rubin, N. (2002). Prevalence of renal cell carcinoma in patients with esrd pre-transplantation : a pathologic analysis. *Kidney international*, 61(6) :2201–2209.
- [Digalakis and Neumeyer, 1996] Digalakis, V. V. and Neumeyer, L. G. (1996). Speaker adaptation using combined transformation and bayesian methods. *IEEE transactions on speech and audio processing*, 4(4) :294–300.
- [Dijkhuizen et al., 1998] Dijkhuizen, T., Van Den Berg, E., Störkel, S., Terpe, H., Bürger, H., and De Jong, B. (1998). Distinct features for chromophilic renal cell cancer with xp11. 2 breakpoints. *Cancer genetics and cytogenetics*, 104(1) :74–76.
- [Du, 2019] Du, F. (2019). Golgi-cox staining of neuronal dendrites and dendritic spines with fd rapid golgistain™ kit. *Current protocols in neuroscience*, 88(1) :e69.
- [Du et al., 2018] Du, Y., Zhang, R., Zargari, A., Thai, T. C., Gunderson, C. C., Moxley, K. M., Liu, H., Zheng, B., and Qiu, Y. (2018). Classification of tumor epithelium and stroma by exploiting image features learned by deep convolutional neural networks. *Annals of biomedical engineering*, 46(12) :1988–1999.
- [Dulmet-Brender et al., 1986] Dulmet-Brender, E., Jaubert, F., and Huchon, G. (1986). Exophytic endobronchial epidermoid carcinoma. *Cancer*, 57(7) :1358–1364.
- [Eberhard et al., 2000] Eberhard, A., Kahlert, S., Goede, V., Hemmerlein, B., Plate, K. H., and Augustin, H. G. (2000). Heterogeneity of angiogenesis and blood vessel maturation in human tumors : implications for antiangiogenic tumor therapies. *Cancer research*, 60(5) :1388–1393.
- [Eguchi et al., 2006] Eguchi, T., Nakanishi, Y., Shimoda, T., Iwasaki, M., Igaki, H., Tachimori, Y., Kato, H., Yamaguchi, H., Saito, D., and Umemura, S. (2006). Histopathological criteria for additional treatment after endoscopic mucosal resection for esophageal cancer : analysis of 464 surgically resected cases. *Modern Pathology*, 19(3) :475–480.
- [Enzinger and Mayer, 2003] Enzinger, P. C. and Mayer, R. J. (2003). Esophageal cancer. *New England Journal of Medicine*, 349(23) :2241–2252.
- [Ericsson et al., 1966] Ericsson, J. L., Seljelid, R., and Orrenius, S. (1966). Comparative light and electron microscopic observations of the cytoplasmic matrix in renal carcinomas. *Virchows Archiv für pathologische Anatomie und Physiologie und für klinische Medizin*, 341(3) :204–223.
- [Eriksen et al., 2017] Eriksen, A. C., Andersen, J. B., Kristensson, M., dePont Christensen, R., Hansen, T. F., Kjær-Frifeldt, S., and Sørensen, F. B. (2017). Computer-assisted stereology and automated image analysis for quantification of tumor infiltrating lymphocytes in colon cancer. *Diagnostic pathology*, 12(1) :1–14.

- [Fatima et al., 2017] Fatima, K. N., Hassan, T., Akram, M. U., Akhtar, M., and Butt, W. H. (2017). Fully automated diagnosis of papilledema through robust extraction of vascular patterns and ocular pathology from fundus photographs. *Biomedical optics express*, 8(2) :1005–1024.
- [Feng et al., 2018] Feng, Y., Zhang, L., and Mo, J. (2018). Deep manifold preserving autoencoder for classifying breast cancer histopathological images. *IEEE/ACM transactions on computational biology and bioinformatics*, 17(1) :91–101.
- [Fenstermaker et al., 2020] Fenstermaker, M., Tomlins, S. A., Singh, K., Wiens, J., and Morgan, T. M. (2020). Development and validation of a deep-learning model to assist with renal cell carcinoma histopathologic interpretation. *Urology*, 144 :152–157.
- [Ferlay et al., 2010] Ferlay, J., Shin, H.-R., Bray, F., Forman, D., Mathers, C., and Parkin, D. M. (2010). Estimates of worldwide burden of cancer in 2008 : Globocan 2008. *International journal of cancer*, 127(12) :2893–2917.
- [Fine, 2014] Fine, J. L. (2014). 21st century workflow : a proposal. *Journal of pathology informatics*, 5(1) :44.
- [Fondón et al., 2018] Fondón, I., Sarmiento, A., García, A. I., Silvestre, M., Eloy, C., Polónia, A., and Aguiar, P. (2018). Automatic classification of tissue malignancy for breast carcinoma diagnosis. *Computers in biology and medicine*, 96 :41–51.
- [Foot, 1933] Foot, N. C. (1933). The masson trichrome staining methods in routine laboratory use. *Stain technology*, 8(3) :101–110.
- [Franklin et al., 1996] Franklin, J., Figlin, R., and Belldegrun, A. (1996). Renal cell carcinoma : Basic biology and clinical behavior. In *Seminars in urologic oncology*, volume 14, pages 208–215.
- [Freund and Schapire, 1997] Freund, Y. and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1) :119–139.
- [Friedman, 2001] Friedman, J. H. (2001). Greedy function approximation : a gradient boosting machine. *Annals of statistics*, pages 1189–1232.
- [Fuhrman et al., 1982] Fuhrman, S. A., Lasky, L. C., and Limas, C. (1982). Prognostic significance of morphologic parameters in renal cell carcinoma. *The American journal of surgical pathology*, 6(7) :655–663.
- [Funai et al., 2003] Funai, K., Yokose, T., Ishii, G.-i., Araki, K., Yoshida, J., Nishimura, M., Nagai, K., Nishiwaki, Y., and Ochiai, A. (2003). Clinicopathologic characteristics of peripheral squamous cell carcinoma of the lung. *The American journal of surgical pathology*, 27(7) :978–984.
- [Gadiya et al., 2019] Gadiya, S., Anand, D., and Sethi, A. (2019). Histograms : Graphs in histopathology. *arXiv*.
- [Gallicchio and Micheli, 2020a] Gallicchio, C. and Micheli, A. (2020a). Fast and deep graph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 3898–3905.
- [Gallicchio and Micheli, 2020b] Gallicchio, C. and Micheli, A. (2020b). Ring reservoir neural networks for graphs. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE.

- [Gandomkar et al., 2016] Gandomkar, Z., Brennan, P. C., and Mello-Thoms, C. (2016). Computer-based image analysis in breast pathology. *Journal of pathology informatics*, 7(1) :43.
- [Gao and Ji, 2019] Gao, H. and Ji, S. (2019). Graph u-nets. In *international conference on machine learning*, pages 2083–2092. PMLR.
- [Gao et al., 2020] Gao, Z., Puttapirat, P., Shi, J., and Li, C. (2020). Renal cell carcinoma detection and subtyping with minimal point-based annotation in whole-slide images. In *MICCAI*, pages 439–448. Springer.
- [Gao et al., 2021] Gao, Z., Shi, J., Zhang, X., Li, Y., Zhang, H., Wu, J., Wang, C., Meng, D., and Li, C. (2021). Nuclei grading of clear cell renal cell carcinoma in histopathological image by composite high-resolution network. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 132–142. Springer.
- [Goode et al., 2013] Goode, A., Gilbert, B., Harkes, J., Jukic, D., and Satyanarayanan, M. (2013). Openslide : A vendor-neutral software foundation for digital pathology. *Journal of pathology informatics*, 4(1) :27.
- [Grignon and Che, 2005] Grignon, D. J. and Che, M. (2005). Clear cell renal cell carcinoma. *Clinics in laboratory medicine*, 25(2) :305–316.
- [Grobholz, 2018] Grobholz, R. (2018). Digitale pathologie. *Wiener klinisches Magazin*, 21(6) :270–279.
- [Gündüz et al., 2022] Gündüz, N., Eser, M., Yıldırım, A., and Kabaalioglu, A. (2022). Radiomics improves the utility of adc for differentiation between renal oncocytoma and chromophobe renal cell carcinoma : Preliminary findings. *Actas Urológicas Españolas (English Edition)*.
- [Guo et al., 2020] Guo, X., Xiao, R., Zhang, T., Chen, C., Wang, J., and Wang, Z. (2020). A novel method to model hepatic vascular network using vessel segmentation, thinning, and completion. *Medical & biological engineering & computing*, pages 1–16.
- [Guo et al., 2010] Guo, Z., Zhang, L., and Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE transactions on image processing*, 19(6) :1657–1663.
- [He et al., 2016] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *CVPR*, pages 770–778.
- [Higgins et al., 2012] Higgins, K. A., Chino, J. P., Ready, N., D’Amico, T. A., Berry, M. F., Sporn, T., Boyd, J., and Kelsey, C. R. (2012). Lymphovascular invasion in non–small-cell lung cancer : implications for staging and adjuvant therapy. *Journal of Thoracic Oncology*, 7(7) :1141–1147.
- [Hollander et al., 2013] Hollander, M., Wolfe, D. A., and Chicken, E. (2013). *Nonparametric statistical methods*, volume 751. John Wiley & Sons.
- [Hou et al., 2022] Hou, C., Li, J., Wang, W., Sun, L., and Zhang, J. (2022). Second-order asymmetric convolution network for breast cancer histopathology image classification. *Journal of Biophotonics*, 15(5) :e202100370.
- [Howlader et al., 2016] Howlader, N., Noone, A., Krapcho, M., Miller, D., Bishop, K., Altekruse, S., Kosary, C., Yu, M., Ruhl, J., Tatalovich, Z., et al. (2016). Seer cancer statistics review, 1975–2013. *National Cancer Institute*, pages 1992–2013.
- [Hsieh et al., 2017] Hsieh, J. J., Purdue, M. P., Signoretti, S., Swanton, C., Albiges, L., Schmidinger, M., Heng, D. Y., Larkin, J., and Ficarra, V. (2017). Renal cell carcinoma. *Nature reviews Disease primers*, 3(1) :1–19.

- [Hunt et al., 2005] Hunt, J. D., Van der Hel, O. L., McMillan, G. P., Boffetta, P., and Brennan, P. (2005). Renal cell carcinoma in relation to cigarette smoking : meta-analysis of 24 studies. *International journal of cancer*, 114(1) :101–108.
- [Inoue et al., 2015] Inoue, H., Kaga, M., Ikeda, H., Sato, C., Sato, H., Minami, H., Santi, E. G., Hayee, B., and Eleftheriadis, N. (2015). Magnification endoscopy in esophageal squamous cell carcinoma : a review of the intrapapillary capillary loop classification. *Annals of Gastroenterology : Quarterly Publication of the Hellenic Society of Gastroenterology*, 28(1) :41.
- [Izadyyazdanabadi et al., 2018a] Izadyyazdanabadi, M., Belykh, E., Mooney, M., Martirosyan, N., Eschbacher, J., Nakaji, P., Preul, M. C., and Yang, Y. (2018a). Convolutional neural networks : Ensemble modeling, fine-tuning and unsupervised semantic localization for neurosurgical cle images. *Journal of Visual Communication and Image Representation*, 54 :10–20.
- [Izadyyazdanabadi et al., 2018b] Izadyyazdanabadi, M., Belykh, E., Mooney, M. A., Eschbacher, J. M., Nakaji, P., Yang, Y., and Preul, M. C. (2018b). Prospects for theranostics in neurosurgical imaging : empowering confocal laser endomicroscopy diagnostics via deep learning. *Frontiers in oncology*, 8 :240.
- [Janowczyk and Madabhushi, 2016] Janowczyk, A. and Madabhushi, A. (2016). Deep learning for digital pathology image analysis : A comprehensive tutorial with selected use cases. *Journal of pathology informatics*, 7(1) :29.
- [Jeena et al., 2021] Jeena, R., Shiny, G., Sukesh Kumar, A., and Mahadevan, K. (2021). A comparative analysis of stroke diagnosis from retinal images using hand-crafted features and cnn. *Journal of Intelligent & Fuzzy Systems*, (Preprint) :1–9.
- [Jemal et al., 2011] Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA : a cancer journal for clinicians*, 61(2) :69–90.
- [Jemal et al., 2005] Jemal, A., Murray, T., Ward, E., Samuels, A., Tiwari, R. C., Ghafoor, A., Feuer, E. J., and Thun, M. J. (2005). Cancer statistics, 2005.
- [Jiang et al., 2018] Jiang, K., Al-Diffalha, S., and Centeno, B. A. (2018). Primary liver cancers—part 1 : Histopathology, differential diagnoses, and risk stratification. *Cancer Control*, 25(1) :1073274817744625.
- [Kather et al., 2015] Kather, J. N., Marx, A., Reyes-Aldasoro, C. C., Schad, L. R., Zöllner, F. G., and Weis, C.-A. (2015). Continuous representation of tumor microvessel density and detection of angiogenic hotspots in histological whole-slide images. *Oncotarget*, 6(22) :19163.
- [Kaur et al., 2019] Kaur, R., Juneja, M., and Mandal, A. K. (2019). Computer-aided diagnosis of renal lesions in ct images : a comprehensive survey and future prospects. *Computers & Electrical Engineering*, 77 :423–434.
- [Kausar et al., 2019] Kausar, T., Wang, M., Idrees, M., and Lu, Y. (2019). Hwdcnn : Multi-class recognition in breast histopathology with haar wavelet decomposed image based convolution neural network. *Biocybernetics and Biomedical Engineering*, 39(4) :967–982.
- [Kipf and Welling, 2016] Kipf, T. N. and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv :1609.02907*.
- [Klein and Valensi, 1976] Klein, M. J. and Valensi, Q. J. (1976). Proximal tubular adenomas of kidney with so-called oncocytic features. a clinicopathologic study of 13 cases of a rarely reported neoplasm. *Cancer*, 38(2) :906–914.

- [Kojiro, 2005] Kojiro, M. (2005). Histopathology of liver cancers. *Best Practice & Research Clinical Gastroenterology*, 19(1) :39–62.
- [Komura and Ishikawa, 2018] Komura, D. and Ishikawa, S. (2018). Machine learning methods for histopathological image analysis. *Computational and structural biotechnology journal*, 16 :34–42.
- [Kosary and McLaughlin, 1993] Kosary, C. and McLaughlin, J. (1993). Kidney and renal pelvis. seer cancer statistics review. 1973–1990. miller ba, ries lag, hankey bf. Technical report, et al : NIH publication.
- [Kostopoulos et al., 2017] Kostopoulos, S., Ravazoula, P., Asvestas, P., Kalatzis, I., Xenogiannopoulos, G., Cavouras, D., and Glotsos, D. (2017). Development of a reference image collection library for histopathology image processing, analysis and decision support systems research. *Journal of digital imaging*, 30(3) :287–295.
- [Kovacs et al., 1997] Kovacs, G., Akhtar, M., Beckwith, B. J., Bugert, P., Cooper, C. S., Delahunt, B., Eble, J. N., Fleming, S., Ljungberg, B., Medeiros, L. J., et al. (1997). The heidelberg classification of renal cell tumours. *The Journal of Pathology : A Journal of the Pathological Society of Great Britain and Ireland*, 183(2) :131–133.
- [Krizhevsky et al., 2012a] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012a). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25 :1097–1105.
- [Krizhevsky et al., 2012b] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012b). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- [Kryvenko et al., 2013] Kryvenko, O. N., Roquero, L., Gupta, N. S., Lee, M. W., and Epstein, J. I. (2013). Low-grade clear cell renal cell carcinoma mimicking hemangioma of the kidney : a series of 4 cases. *Archives of Pathology & Laboratory Medicine*, 137(2) :251–254.
- [Kumar et al., 2020] Kumar, A., Singh, S. K., Saxena, S., Singh, A. K., Shrivastava, S., Lakshmanan, K., Kumar, N., and Singh, R. K. (2020). Comhisp : A novel feature extractor for histopathological image classification based on fuzzy svm with within-class relative density. *IEEE Transactions on Fuzzy Systems*, 29(1) :103–117.
- [Kuroda et al., 2003] Kuroda, N., Toi, M., Hiroi, M., Shuin, T., and Enzan, H. (2003). Review of renal oncocytoma with focus on clinical and pathobiological aspects. *Histology and histopathology*.
- [Kwatra et al., 2003] Kwatra, K. S., Prabhakar, B., Jain, S., and Grewal, J. S. (2003). Sarcomatoid carcinoma (carcinosarcoma) of the esophagus with extensive areas of osseous differentiation : a case report. *Indian journal of pathology & microbiology*, 46(1) :49–51.
- [LeCun et al., 1989] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4) :541–551.
- [Lee et al., 2017] Lee, H. S., Hong, H., Jung, D. C., Park, S., and Kim, J. (2017). Differentiation of fat-poor angiomyolipoma from clear cell renal cell carcinoma in contrast-enhanced mdct images using quantitative feature classification. *Medical physics*, 44(7) :3604–3614.
- [Lee et al., 2019] Lee, J., Lee, I., and Kang, J. (2019). Self-attention graph pooling. In *International conference on machine learning*, pages 3734–3743. PMLR.

- [Lee et al., 2020] Lee, M. A., Kang, J., Lee, H. Y., Kim, W., Shon, I., Hwang, N. Y., Kim, H. K., Choi, Y. S., Kim, J., Zo, J. I., et al. (2020). Spread through air spaces (stas) in invasive mucinous adenocarcinoma of the lung : Incidence, prognostic impact, and prediction based on clinicoradiologic factors. *Thoracic cancer*, 11(11) :3145–3154.
- [Levi et al., 2008] Levi, F., Ferlay, J., Galeone, C., Lucchini, F., Negri, E., Boyle, P., and La Vecchia, C. (2008). The changing pattern of kidney cancer incidence and mortality in europe. *BJU international*, 101(8) :949–958.
- [Li et al., 2019] Li, J., Yang, S., Huang, X., Da, Q., Yang, X., Hu, Z., Duan, Q., Wang, C., and Li, H. (2019). Signet ring cell detection with a semi-supervised learning framework. In *IPMI*, pages 842–854. Springer.
- [Li et al., 2016a] Li, M., He, P., Wu, Y., Zhang, Y., Xia, H., Zheng, Y., and Han, Y. (2016a). Stimulatory effects of the degradation products from mg-ca-sr alloy on the osteogenesis through regulating erk signaling pathway. *Scientific reports*, 6(1) :1–13.
- [Li et al., 2016b] Li, W., Manivannan, S., Akbar, S., Zhang, J., Trucco, E., and McKenna, S. J. (2016b). Gland segmentation in colon histology images using hand-crafted features and convolutional neural networks. In *ISBI*, pages 1405–1408. IEEE.
- [Litjens, 2017] Litjens, G. (2017). Automated slide analysis platform (asap).
- [Ljungberg et al., 2007] Ljungberg, B., Hanbury, D. C., Kuczyk, M. A., Merseburger, A. S., Mulders, P. F., Patard, J.-J., and Sinescu, I. C. (2007). Renal cell carcinoma guideline. *European urology*, 51(6) :1502–1510.
- [Lopez-Beltran et al., 2006] Lopez-Beltran, A., Scarpelli, M., Montironi, R., and Kirkali, Z. (2006). 2004 who classification of the renal tumors of the adults. *European urology*, 49(5) :798–805.
- [Loukas and Linney, 2004a] Loukas, C. G. and Linney, A. (2004a). A survey on histological image analysis-based assessment of three major biological factors influencing radiotherapy : proliferation, hypoxia and vasculature. *Computer Methods and Programs in Biomedicine*, 74(3) :183–199.
- [Loukas and Linney, 2004b] Loukas, C. G. and Linney, A. (2004b). A survey on histological image analysis-based assessment of three major biological factors influencing radiotherapy : proliferation, hypoxia and vasculature. *Computer Methods and Programs in Biomedicine*, 74(3) :183–199.
- [Ludovic et al., 2013] Ludovic, R., Daniel, R., Nicolas, L., Maria, K., Humayun, I., Jacques, K., Frédérique, C., Catherine, G., et al. (2013). Mitosis detection in breast cancer histological images an icpr 2012 contest. *Journal of pathology informatics*, 4(1) :8.
- [Luo, 2020] Luo, X. (2020). SSL4MIS. <https://github.com/HiLab-git/SSL4MIS>.
- [Ma et al., 2020] Ma, Y., Hao, H., Xie, J., Fu, H., Zhang, J., Yang, J., Wang, Z., Liu, J., Zheng, Y., and Zhao, Y. (2020). Rose : a retinal oct-angiography vessel segmentation dataset and new model. *IEEE transactions on medical imaging*, 40(3) :928–939.
- [Macenko et al., 2009] Macenko, M., Niethammer, M., Marron, J. S., Borland, D., Woosley, J. T., Guan, X., Schmitt, C., and Thomas, N. E. (2009). A method for normalizing histology slides for quantitative analysis. In *2009 IEEE international symposium on biomedical imaging : from nano to macro*, pages 1107–1110. IEEE.

- [Maeda et al., 2017] Maeda, I., Kubota, M., Ohta, J., Shinno, K., Tajima, S., Ariizumi, Y., Doi, M., Oana, Y., Kanemaki, Y., Tsugawa, K., et al. (2017). Effectiveness of computer-aided diagnosis (cadx) of breast pathology using immunohistochemistry results of core needle biopsy samples for synaptophysin, oestrogen receptor and ck14/p63 for classification of epithelial proliferative lesions of the breast. *Journal of Clinical Pathology*, 70(12) :1057–1062.
- [Magers et al., 2019] Magers, M. J., Perrino, C. M., Cramer, H. M., and Wu, H. H. (2019). Cytomorphologic comparison of type 1 and type 2 papillary renal cell carcinoma : a retrospective analysis of 28 cases. *Cancer Cytopathology*, 127(6) :370–376.
- [Mao et al., 2011] Mao, W.-M., Zheng, W.-H., and Ling, Z.-Q. (2011). Epidemiologic risk factors for esophageal cancer development. *Asian Pac J Cancer Prev*, 12(10) :2461–2466.
- [Mathers et al., 2002] Mathers, M. E., Pollock, A., Marsh, C., and O’Donnell, M. (2002). Cytokeratin 7 : a useful adjunct in the diagnosis of chromophobe renal cell carcinoma. *Histopathology*, 40(6) :563–567.
- [Maulitz, 1973] Maulitz, R. C. (1973). *A TREATISE ON MEMBRANES : CONCEPTS OF TISSUE STRUCTURE, FUNCTION AND DYSFUNCTION, FROM XAVIER BICHAT TO JULIUS COHNHEIM*. Duke University.
- [McDonald and Choyke, 2003] McDonald, D. M. and Choyke, P. L. (2003). Imaging of angiogenesis : from microscope to clinic. *Nature medicine*, 9(6) :713–725.
- [Melamed et al., 1977] Melamed, M. R., Zaman, M. B., Flehinger, B. J., and Martini, N. (1977). Radiologically occult in situ and incipient invasive epidermoid lung cancer : detection by sputum cytology in a survey of asymptomatic cigarette smokers. *The American journal of surgical pathology*, 1(1) :5–16.
- [Meyer and Beucher, 1990] Meyer, F. and Beucher, S. (1990). Morphological segmentation. *Journal of visual communication and image representation*, 1(1) :21–46.
- [Miller et al., 2018] Miller, K. D., Goding Sauer, A., Ortiz, A. P., Fedewa, S. A., Pinheiro, P. S., Tortolero-Luna, G., Martinez-Tyson, D., Jemal, A., and Siegel, R. L. (2018). Cancer statistics for hispanics/latinos, 2018. *CA : a cancer journal for clinicians*, 68(6) :425–445.
- [Moch et al., 2016] Moch, H., Cubilla, A. L., Humphrey, P. A., Reuter, V. E., and Ulbright, T. M. (2016). The 2016 who classification of tumours of the urinary system and male genital organs—part a : renal, penile, and testicular tumours. *European urology*, 70(1) :93–105.
- [Momeni-Boroujeni et al., 2017] Momeni-Boroujeni, A., Yousefi, E., and Somma, J. (2017). Computer-assisted cytologic diagnosis in pancreatic fna : An application of neural networks to image analysis. *Cancer Cytopathology*, 125(12) :926–933.
- [Mousavi et al., 2015] Mousavi, H. S., Monga, V., Rao, G., and Rao, A. U. (2015). Automated discrimination of lower and higher grade gliomas based on histopathological image analysis. *Journal of pathology informatics*, 6(1) :15.
- [Nagashima, 2000] Nagashima, Y. (2000). Chromophobe renal cell carcinoma : clinical, pathological and molecular biological aspects. *Pathology international*, 50(11) :872–878.
- [Nayak et al., 2013] Nayak, N., Chang, H., Borowsky, A., Spellman, P., and Parvin, B. (2013). Classification of tumor histopathology via sparse feature learning. In *2013 IEEE 10th international symposium on biomedical imaging*, pages 410–413. IEEE.

- [Nouh et al., 2010] Nouh, M. A., Kuroda, N., Yamashita, M., Hayashida, Y., Yano, T., Minakuchi, J., Taniguchi, S., Nomura, I., Inui, M., Sugimoto, M., et al. (2010). Renal cell carcinoma in patients with end-stage renal disease : relationship between histological type and duration of dialysis. *BJU international*, 105(5) :620–627.
- [Ohgaki and Kleihues, 2005] Ohgaki, H. and Kleihues, P. (2005). Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas. *Journal of Neuropathology & Experimental Neurology*, 64(6) :479–489.
- [Ojala et al., 2002] Ojala, T., Pietikainen, M., and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7) :971–987.
- [Ojansivu and Heikkilä, 2008] Ojansivu, V. and Heikkilä, J. (2008). Blur insensitive texture classification using local phase quantization. In *International conference on image and signal processing*, pages 236–243. Springer.
- [Otsu, 1979] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1) :62–66.
- [Ouali et al., 2020] Ouali, Y., Hudelot, C., and Tami, M. (2020). Semi-supervised semantic segmentation with cross-consistency training. In *CVPR*, pages 12674–12684.
- [Parker et al., 2004] Parker, A. S., Cerhan, J. R., Lynch, C. F., Leibovich, B. C., and Cantor, K. P. (2004). History of urinary tract infection and risk of renal cell carcinoma. *American journal of epidemiology*, 159(1) :42–48.
- [Parkin et al., 2005] Parkin, D. M., Bray, F., Ferlay, J., and Pisani, P. (2005). Global cancer statistics, 2002. *CA : a cancer journal for clinicians*, 55(2) :74–108.
- [Peng et al., 2019] Peng, T., Boxberg, M., Weichert, W., Navab, N., and Marr, C. (2019). Multi-task learning of a deep k-nearest neighbour network for histopathological image classification and retrieval. In *MICCAI*, pages 676–684. Springer.
- [Pennathur et al., 2013] Pennathur, A., Gibson, M. K., Jobe, B. A., and Luketich, J. D. (2013). Oesophageal carcinoma. *The Lancet*, 381(9864) :400–412.
- [Perea et al., 2009] Perea, G., Navarrete, M., and Araque, A. (2009). Tripartite synapses : astrocytes process and control synaptic information. *Trends in neurosciences*, 32(8) :421–431.
- [Pettersson et al., 2011] Pettersson, F., Šíma, R., Grossmann, P., Michal, M., Kuroda, N., Hora, M., Yang, X., Kinkor, Z., Trivunic, S., Žalud, R., et al. (2011). Renal small cell oncocyoma with pseudorosettes : A histomorphologic, immunohistochemical, and molecular genetic study of 10 cases. *Human pathology*, 42(11) :1751–1760.
- [Pischon et al., 2006] Pischon, T., Lahmann, P. H., Boeing, H., Tjønneland, A., Halkjær, J., Overvad, K., Klipstein-Grobusch, K., Linseisen, J., Becker, N., Trichopoulou, A., et al. (2006). Body size and risk of renal cell carcinoma in the european prospective investigation into cancer and nutrition (epic). *International journal of cancer*, 118(3) :728–738.
- [Port et al., 1989] Port, F. K., Ragheb, N. E., Schwartz, A. G., and Hawthorne, V. M. (1989). Neoplasms in dialysis patients : a population-based study. *American journal of kidney diseases*, 14(2) :119–123.
- [Prasad et al., 2006] Prasad, S. R., Humphrey, P. A., Catena, J. R., Narra, V. R., Srigley, J. R., Cortez, A. D., Dalrymple, N. C., and Chintapalli, K. N. (2006). Common and uncommon

- histologic subtypes of renal cell carcinoma : imaging spectrum with pathologic correlation. *Radiographics*, 26(6) :1795–1806.
- [Przybycin et al., 2011] Przybycin, C. G., Cronin, A. M., Darvishian, F., Gopalan, A., Al-Ahmadie, H. A., Fine, S. W., Chen, Y.-b., Bernstein, M., Russo, P., Reuter, V. E., et al. (2011). Chromophobe renal cell carcinoma : a clinicopathologic study of 203 tumors in 200 patients with primary resection at a single institution. *The American journal of surgical pathology*, 35(7) :962–970.
- [Qiao et al., 2018] Qiao, S., Shen, W., Zhang, Z., Wang, B., and Yuille, A. (2018). Deep co-training for semi-supervised image recognition. In *ECCV*, pages 135–152.
- [Quinlan, 2004] Quinlan, J. R. (2004). Data mining tools see5 and c5.0. <http://www.rulequest.com/see5-info.html>.
- [Ramzan et al., 2020] Ramzan, F., Khan, M. U. G., Rehmat, A., Iqbal, S., Saba, T., Rehman, A., and Mehmood, Z. (2020). A deep learning approach for automated diagnosis and multi-class classification of alzheimer’s disease stages using resting-state fmri and residual neural networks. *Journal of medical systems*, 44(2) :1–16.
- [Raza et al., 2010] Raza, S. H., Parry, R. M., Sharma, Y., Chaudry, Q., Moffitt, R. A., Young, A., and Wang, M. D. (2010). Automated classification of renal cell carcinoma subtypes using bag-of-features. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 6749–6752. IEEE.
- [Reinhard et al., 2001] Reinhard, E., Adhikhmin, M., Gooch, B., and Shirley, P. (2001). Color transfer between images. *IEEE Computer graphics and applications*, 21(5) :34–41.
- [Renehan et al., 2008] Renehan, A. G., Tyson, M., Egger, M., Heller, R. F., and Zwahlen, M. (2008). Body-mass index and incidence of cancer : a systematic review and meta-analysis of prospective observational studies. *The lancet*, 371(9612) :569–578.
- [Robertson et al., 2018] Robertson, S., Azizpour, H., Smith, K., and Hartman, J. (2018). Digital image analysis in breast pathology—from image processing techniques to artificial intelligence. *Translational Research*, 194 :19–35.
- [Rodriguez et al., 2012] Rodriguez, F. J., Orr, B. A., Ligon, K. L., and Eberhart, C. G. (2012). Neoplastic cells are a rare component in human glioblastoma microvasculature. *Oncotarget*, 3(1) :98.
- [Ross and Pawlina, 2006] Ross, M. H. and Pawlina, W. (2006). *Histology*. Lippincott Williams & Wilkins.
- [Ross et al., 1989] Ross, R. K., Paganini-Hill, A., Landolph, J., Gerkins, V., and Henderson, B. E. (1989). Analgesics, cigarette smoking, and other risk factors for cancer of the renal pelvis and ureter. *Cancer research*, 49(4) :1045–1048.
- [Ruble et al., 2011] Rublee, E., Rabaud, V., Konolige, K., and Bradski, G. (2011). Orb : An efficient alternative to sift or surf. In *2011 International conference on computer vision*, pages 2564–2571. Ieee.
- [Ruderman et al., 1998] Ruderman, D. L., Cronin, T. W., and Chiao, C.-C. (1998). Statistics of cone responses to natural images : implications for visual coding. *JOSA A*, 15(8) :2036–2045.
- [Saito et al., 2016] Saito, A., Numata, Y., Hamada, T., Horisawa, T., Cosatto, E., Graf, H.-P., Kuroda, M., and Yamamoto, Y. (2016). A novel method for morphological pleomorphism and

- heterogeneity quantitative measurement : Named cell feature level co-occurrence matrix. *Journal of pathology informatics*, 7(1) :36.
- [Salzberg, 1994] Salzberg, S. L. (1994). C4. 5 : Programs for machine learning by j. ross quinlan. morgan kaufmann publishers, inc., 1993.
- [Saxena et al., 2021] Saxena, S., Shukla, S., and Gyanchandani, M. (2021). Breast cancer histopathology image classification using kernelized weighted extreme learning machine. *International Journal of Imaging Systems and Technology*, 31(1) :168–179.
- [Schatz and Lieber, 2003] Schatz, S. M. and Lieber, M. M. (2003). Update on oncocytoma. *Current urology reports*, 4(1) :30–35.
- [Sherwin et al., 1962] Sherwin, R. P., Laforet, E. G., and Strieder, J. W. (1962). Exophytic endobronchial carcinoma. *The Journal of Thoracic and Cardiovascular Surgery*, 43(6) :716–730.
- [Shimada et al., 2012] Shimada, Y., Saji, H., Yoshida, K., Kakihana, M., Honda, H., Nomura, M., Usuda, J., Kajiwarra, N., Ohira, T., and Ikeda, N. (2012). Pathological vascular invasion and tumor differentiation predict cancer recurrence in stage ia non–small-cell lung cancer after complete surgical resection. *Journal of Thoracic Oncology*, 7(8) :1263–1270.
- [Shiraishi et al., 2011] Shiraishi, J., Li, Q., Appelbaum, D., and Doi, K. (2011). Computer-aided diagnosis and artificial intelligence in clinical imaging. In *Seminars in nuclear medicine*, volume 41, pages 449–462. Elsevier.
- [Shu et al., 2016] Shu, J., Dolman, G., Duan, J., Qiu, G., and Ilyas, M. (2016). Statistical colour models : an automated digital image analysis method for quantification of histological biomarkers. *Biomedical engineering online*, 15(1) :1–16.
- [Sica et al., 2010] Sica, G., Yoshizawa, A., Sima, C. S., Azzoli, C. G., Downey, R. J., Rusch, V. W., Travis, W. D., and Moreira, A. L. (2010). A grading system of lung adenocarcinomas based on histologic pattern is predictive of disease recurrence in stage i tumors. *The American journal of surgical pathology*, 34(8) :1155–1162.
- [Siegel et al., 2016] Siegel, R. L., Miller, K. D., and Jemal, A. (2016). Cancer statistics, 2016. *CA : a cancer journal for clinicians*, 66(1) :7–30.
- [Sim et al., 1999] Sim, J. S., Seo, C. S., Kim, S. H., Kim, J. H., Kim, W. S., Han, J. K., and Choi, B. I. (1999). Differentiation of small hyperechoic renal cell carcinoma from angiomyolipoma : computer-aided tissue echo quantification. *Journal of ultrasound in medicine*, 18(4) :261–264.
- [Simonyan and Zisserman, 2014] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv*.
- [Smola et al., 2000] Smola, A. J., Bartlett, P., Schölkopf, B., and Schuurmans, D. (2000). Probabilities for sv machines.
- [Sobin, 2009] Sobin, L. (2009). International union against cancer (uicc) tnm classification of malignant tumours. *Oesophagus including Oesophagogastric Junction*, pages 66–72.
- [Soille and Vincent, 1990] Soille, P. and Vincent, L. M. (1990). Determining watersheds in digital pictures via flooding simulations. In *VCIP*, volume 1360, pages 240–250. International Society for Optics and Photonics.
- [Spanhol et al., 2015] Spanhol, F. A., Oliveira, L. S., Petitjean, C., and Heutte, L. (2015). A dataset for breast cancer histopathological image classification. *IEEE transactions on biomedical engineering*, 63(7) :1455–1462.

- [Sparks and Madabhushi, 2016] Sparks, R. and Madabhushi, A. (2016). Out-of-sample extrapolation utilizing semi-supervised manifold learning (ose-ssl) : Content based image retrieval for histopathology images. *Scientific reports*, 6(1) :1–15.
- [Strigley et al., 2013] Strigley, J. R., Delahunt, B., Eble, J. N., Egevad, L., Epstein, J. I., Grignon, D., Hes, O., Moch, H., Montironi, R., Tickoo, S. K., et al. (2013). The international society of urological pathology (isup) vancouver classification of renal neoplasia. *The American journal of surgical pathology*, 37(10) :1469–1489.
- [Stanton et al., 2011] Stanton, M. C., Chen, S.-C., Jackson, J. V., Rojas-Triana, A., Kinsley, D., Cui, L., Fine, J. S., Greenfeder, S., Bober, L. A., and Jenh, C.-H. (2011). Inflammatory signals shift from adipose to liver during high fat feeding and influence the development of steatohepatitis in mice. *Journal of inflammation*, 8(1) :1–14.
- [Steiner et al., 2018] Steiner, D. F., MacDonald, R., Liu, Y., Truszkowski, P., Hipp, J. D., Gammage, C., Thng, F., Peng, L., and Stumpe, M. C. (2018). Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. *The American journal of surgical pathology*, 42(12) :1636.
- [Suarez-Ibarrola et al., 2020] Suarez-Ibarrola, R., Hein, S., Reis, G., Gratzke, C., and Miernik, A. (2020). Current and future applications of machine and deep learning in urology : a review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer. *World journal of urology*, 38(10) :2329–2347.
- [Sun et al., 2019] Sun, K., Xiao, B., Liu, D., and Wang, J. (2019). Deep high-resolution representation learning for human pose estimation. In *CVPR*, pages 5693–5703.
- [Szegedy et al., 2015] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions. In *CVPR*, pages 1–9.
- [Tajima et al., 2000] Tajima, Y., Nakanishi, Y., Ochiai, A., Tachimori, Y., Kato, H., Watanabe, H., Yamaguchi, H., Yoshimura, K., Kusano, M., and Shimoda, T. (2000). Histopathologic findings predicting lymph node metastasis and prognosis of patients with superficial esophageal carcinoma : analysis of 240 surgically resected tumors. *Cancer*, 88(6) :1285–1293.
- [Tam et al., 2016] Tam, A., Barker, J., and Rubin, D. (2016). A method for normalizing pathology images to improve feature extraction for quantitative pathology. *Medical physics*, 43(1) :528–537.
- [Tarvainen and Valpola, 2017] Tarvainen, A. and Valpola, H. (2017). Mean teachers are better role models : Weight-averaged consistency targets improve semi-supervised deep learning results. *arXiv*.
- [Tickoo et al., 2010] Tickoo, S., Vankalakunti, M., and Reuter, V. (2010). Clear cell renal cell carcinoma. *Diagnostic Pathology : Genitourinary*. Altona, Manitoba, Canada : Amirsys, pages 1–34.
- [Tiwari and Bhargava, 2015] Tiwari, S. and Bhargava, R. (2015). Extracting knowledge from chemical imaging data using computational algorithms for digital cancer diagnosis. *The Yale journal of biology and medicine*, 88(2) :131–143.
- [Tomizawa et al., 1995] Tomizawa, M., Kondo, F., and Kondo, Y. (1995). Growth patterns and interstitial invasion of small hepatocellular carcinoma. *Pathology international*, 45(5) :352–358.

- [Tomlinson et al., 1999] Tomlinson, J., Barsky, S. H., Nelson, S., Singer, S., Pezeshki, B., Lee, M. C., Eilber, F., and Nguyen, M. (1999). Different patterns of angiogenesis in sarcomas and carcinomas. *Clinical cancer research*, 5(11) :3516–3522.
- [Tonk et al., 1995] Tonk, V., Wilson, K. S., Timmons, C. F., Schneider, N. R., and Tomlinson, G. E. (1995). Renal cell carcinoma with translocation (x ; 1) further evidence for a cytogenetically defined subtype. *Cancer genetics and cytogenetics*, 81(1) :72–75.
- [Townsend, 1971] Townsend, J. T. (1971). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, 9(1) :40–50.
- [Travis, 2011] Travis, W. D. (2011). Pathology of lung cancer. *Clinics in chest medicine*, 32(4) :669–692.
- [Travis et al., 2004] Travis, W. D., Brambilla, E., Muller-Hermelink, H. K., and Harris, C. C. (2004). World health organization classification of tumours. *Pathology and genetics of tumours of the lung, pleura, thymus and heart*, 10 :179–84.
- [Trpkov et al., 2010] Trpkov, K., Yilmaz, A., Uzer, D., Dishongh, K. M., Quick, C. M., Bismar, T. A., and Gokden, N. (2010). Renal oncocytoma revisited : a clinicopathological study of 109 cases with emphasis on problematic diagnostic features. *Histopathology*, 57(6) :893–906.
- [Uzosike et al., 2018] Uzosike, A. C., Patel, H. D., Alam, R., Schwen, Z. R., Gupta, M., Gorin, M. A., Johnson, M. H., Gausepohl, H., Riffon, M. F., Trock, B. J., et al. (2018). Growth kinetics of small renal masses on active surveillance : variability and results from the disrrm registry. *The Journal of urology*, 199(3) :641–648.
- [Uzzan et al., 2004] Uzzan, B., Nicolas, P., Cucherat, M., and Perret, G.-Y. (2004). Microvessel density as a prognostic factor in women with breast cancer : a systematic review of the literature and meta-analysis. *Cancer research*, 64(9) :2941–2955.
- [Vahadane et al., 2015] Vahadane, A., Peng, T., Albarqouni, S., Baust, M., Steiger, K., Schlitter, A. M., Sethi, A., Esposito, I., and Navab, N. (2015). Structure-preserved color normalization for histological images. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 1012–1015. IEEE.
- [Vandenhende et al., 2020] Vandenhende, S., Georgoulis, S., and Gool, L. V. (2020). Mti-net : Multi-scale task interaction networks for multi-task learning. In *ECCV*, pages 527–543. Springer.
- [Veta et al., 2014] Veta, M., Pluim, J. P., Van Diest, P. J., and Viergever, M. A. (2014). Breast cancer histopathology image analysis : A review. *IEEE transactions on biomedical engineering*, 61(5) :1400–1411.
- [Vu et al., 2019] Vu, T.-H., Jain, H., Bucher, M., Cord, M., and Pérez, P. (2019). Advent : Adversarial entropy minimization for domain adaptation in semantic segmentation. In *CVPR*, pages 2517–2526.
- [Waheed et al., 2007] Waheed, S., Moffitt, R. A., Chaudry, Q., Young, A. N., and Wang, M. D. (2007). Computer aided histopathological classification of cancer subtypes. In *2007 IEEE 7th International Symposium on BioInformatics and BioEngineering*, pages 503–508. IEEE.
- [Wang et al., 2022a] Wang, C., Gong, W., Cheng, J., and Qian, Y. (2022a). Dblcnn : Dependency-based lightweight convolutional neural network for multi-classification of breast histopathology images. *Biomedical Signal Processing and Control*, 73 :103451.

- [Wang et al., 2017] Wang, C., Shi, J., Zhang, Q., and Ying, S. (2017). Histopathological image classification with bilinear convolutional neural networks. In *EMBC*, pages 4050–4053. IEEE.
- [Wang et al., 2016] Wang, D., Khosla, A., Gargeya, R., Irshad, H., and Beck, A. H. (2016). Deep learning for identifying metastatic breast cancer. *arXiv preprint arXiv :1606.05718*.
- [Wang et al., 2022b] Wang, P., Li, P., Li, Y., Xu, J., Yan, F., and Jiang, M. (2022b). Deep manifold feature fusion for classification of breast histopathology images. *Digital Signal Processing*, page 103400.
- [Wang et al., 2022c] Wang, Q., Zhang, Y., Zhang, B., Fu, Y., Zhao, X., Zhang, J., Zuo, K., Xing, Y., Jiang, S., Qin, Z., et al. (2022c). Single-cell chromatin accessibility landscape in kidney identifies additional cell-of-origin in heterogenous papillary renal cell carcinoma. *Nature Communications*, 13(1) :1–12.
- [Weikert et al., 2008] Weikert, S., Boeing, H., Pischon, T., Weikert, C., Olsen, A., Tjønneland, A., Overvad, K., Becker, N., Linseisen, J., Trichopoulos, A., et al. (2008). Blood pressure and risk of renal cell carcinoma in the european prospective investigation into cancer and nutrition. *American journal of epidemiology*, 167(4) :438–446.
- [Weinstein et al., 2009] Weinstein, R. S., Graham, A. R., Richter, L. C., Barker, G. P., Krupinski, E. A., Lopez, A. M., Erps, K. A., Bhattacharyya, A. K., Yagi, Y., and Gilbertson, J. R. (2009). Overview of telepathology, virtual microscopy, and whole slide imaging : prospects for the future. *Human pathology*, 40(8) :1057–1069.
- [Wetteland et al., 2020] Wetteland, R., Engan, K., Eftestøl, T., Kvikstad, V., and Janssen, E. A. (2020). A multiscale approach for whole-slide image segmentation of five tissue classes in urothelial carcinoma slides. *Technology in Cancer Research & Treatment*, 19 :1533033820946787.
- [Williamson, 2022] Williamson, S. R. (2022). Renal cell carcinoma classification : what matters ? *Diagnostic Histopathology*.
- [Wobker and Williamson, 2017] Wobker, S. E. and Williamson, S. R. (2017). Modern pathologic diagnosis of renal oncocytoma. *Journal of Kidney Cancer and VHL*, 4(4) :1.
- [Wright, 1995] Wright, R. E. (1995). Logistic regression.
- [Xiao et al., 2021] Xiao, R., Debreuve, E., Ambrosetti, D., and Descombes, X. (2021). Renal cell carcinoma classification from vascular morphology. In *MICCAI*, pages 611–621. Springer.
- [Xiao et al., 2017] Xiao, Y., Fortin, M., Unsgård, G., Rivaz, H., and Reinertsen, I. (2017). Retrospective evaluation of cerebral tumors (resect) : A clinical database of pre-operative mri and intra-operative ultrasound in low-grade glioma surgeries. *Medical physics*, 44(7) :3875–3882.
- [Xie and Tu, 2015] Xie, S. and Tu, Z. (2015). Holistically-nested edge detection. In *Proceedings of the IEEE international conference on computer vision*, pages 1395–1403.
- [Xu et al., 2017] Xu, Y., Jia, Z., Wang, L.-B., Ai, Y., Zhang, F., Lai, M., Chang, E. I., et al. (2017). Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC bioinformatics*, 18(1) :1–17.
- [Yamaguchi et al., 2010] Yamaguchi, Y., Ishii, G., Kojima, M., Yoh, K., Otsuka, H., Otaki, Y., Aokage, K., Yanagi, S., Nagai, K., Nishiwaki, Y., et al. (2010). Histopathologic features of the tumor budding in adenocarcinoma of the lung : tumor budding as an index to predict the potential aggressiveness. *Journal of thoracic oncology*, 5(9) :1361–1368.

- [Yoo et al., 2010] Yoo, J. S., Ayati, M. H., Kim, H. B., Zhang, W.-b., and Soh, K.-S. (2010). Characterization of the primo-vascular system in the abdominal cavity of lung cancer mouse model and its differences from the lymphatic system. *Plos one*, 5(4) :e9940.
- [Young Jr et al., 1981] Young Jr, J., Percy, C., Asire, A., Berg, J., Cusano, M., Gloeckler, L., Horm, J., Lourie Jr, W., Pollack, E., and Shambaugh, E. (1981). Cancer incidence and mortality in the united states, 1973-77. *National Cancer Institute Monograph*, (57) :1–187.
- [Yu et al., 1986] Yu, M. C., Mack, T. M., Hanisch, R., Cicioni, C., and Henderson, B. E. (1986). Cigarette smoking, obesity, diuretic use, and coffee consumption as risk factors for renal cell carcinoma. *Journal of the National Cancer Institute*, 77(2) :351–356.
- [Yusenko, 2010] Yusenko, M. V. (2010). Molecular pathology of chromophobe renal cell carcinoma : a review. *International Journal of Urology*, 17(7) :592–600.
- [Zhang et al., 2018a] Zhang, M., Cui, Z., Neumann, M., and Chen, Y. (2018a). An end-to-end deep learning architecture for graph classification. In *AAAI*, volume 32.
- [Zhang et al., 2018b] Zhang, X., Zhou, X., Lin, M., and Sun, J. (2018b). Shufflenet : An extremely efficient convolutional neural network for mobile devices. In *CVPR*, pages 6848–6856.
- [Zhang et al., 2017] Zhang, Y., Yang, L., Chen, J., Fredericksen, M., Hughes, D. P., and Chen, D. Z. (2017). Deep adversarial networks for biomedical image segmentation utilizing unannotated images. In *MICCAI*, pages 408–416. Springer.
- [Zhang et al., 2021] Zhang, Z., Jia, J., Wang, B., and Gong, N. Z. (2021). Backdoor attacks to graph neural networks. In *Proceedings of the 26th ACM Symposium on Access Control Models and Technologies*, pages 15–26.
- [Zhao and Wang, 2019] Zhao, Q. and Wang, Y. (2019). Learning metrics for persistence-based summaries and applications for graph classification. *Advances in Neural Information Processing Systems*, 32.
- [Zhou et al., 2021] Zhou, Y., Chen, H., Li, Y., Liu, Q., Xu, X., Wang, S., Yap, P.-T., and Shen, D. (2021). Multi-task learning for segmentation and classification of tumors in 3d automated breast ultrasound images. *Medical Image Analysis*, 70 :101918.
- [Zubiolo et al., 2016] Zubiolo, A., Debreuve, E., Ambrosetti, D., Pognonec, P., and Descombes, X. (2016). Is the vascular network discriminant enough to classify renal cell carcinoma? In *CBMI*, pages 1–6. IEEE.

List of Figures

2.1	Example of IHC staining.	14
2.2	Benign and malignant grading under H&E staining.	15
2.3	Special stains to reveal specific tissue structures.	16
2.4	Slide Coordinates Origin and (X, Y, Z) vs. Image Matrix Origin (Rows, Columns).	20
2.5	The WSI is stored in a pyramidal file format, including several down-sampled versions of the base image.	21
2.6	Typical steps for machine learning in digital pathological image analysis.	24
2.7	The structure of a 3-layer neural network. (Christian, 2020)	26
2.8	General architectures of neural network. (source : Asimov Institute)	27
2.9	Deep convolutional neural network (DCNN) architecture.	28
2.10	Four example images with corresponding augmented images	29
2.11	Example of H&DAB staining and H&E staining histology images.	32
3.1	Hematoxylin and eosin staining for ccRCC at different magnification.	43
3.2	Hematoxylin and eosin staining for pRCC at different magnification.	43
3.3	Hematoxylin and eosin staining for chRCC at different magnification.	44
3.4	The example of Type 1 and Type 2 of pRCC in histopathological image.	45
3.5	The view of the histopathological image in ASAP.	47
3.6	The example of four categories of RCC histopathological images.	48
3.7	The XML file format of grouped annotated areas.	49
3.8	The example of grouped annotation, the group include necrosis, fiber, tumor and normal.	50
3.9	Get patch images (2000×2000 pixels) from tumor annotation.	50
3.10	BigRCC RCC histopathological images dataset building.	51
3.11	The homepage of TCGA database.	52
3.12	The view of the histopathological image in ImageJ.	54
3.13	The binary mask image of the vascular network.	55
3.14	The view of the histopathological image in ImageJ.	56
3.15	The vascular network examples of ccRCC, pRCC and chRCC images.	57
4.1	The example of the skeleton of RCC subtypes.	63
4.2	Schematic diagram of the watershed algorithm.	65
4.3	The pipeline of obtaining the lattice features.	68
4.4	The graph features of ccRCC, pRCC and chRCC.	69
4.5	The procedure of convert 2D Array Matrix to CSR Matrix	69
5.1	The flow diagram of classification.	73
5.2	The LeNet-5 architecture.	76
5.3	An illustration of the architecture of AlexNet.	77
5.4	The network structure of VGG-16.	79

5.5	The original network structure of ResNet-18. (Ramzan, 2020)	80
5.6	The structure of residual. (He, 2016)	81
5.7	The full pipeline of the TopK model (for $k = 0 :5$), leveraging several stacks of interleaved convolutional/pooling layers. (Cangea, 2018)	83
5.8	The global pooling architecture (A) and the hierarchical pooling architecture (B).	84
5.9	The model summary of SAGPool_g (A) and SAGPool_h (B) in our project.	85
5.10	The distribution of data from hand-crafted features.	88
5.11	The illustration of SVM.	98
5.12	The ROC curves and AUC values were obtained for each model with filtered hand-crafted features.	100
6.1	The HRNet structure consists of parallel high-to-low resolution subnetworks.	109
6.2	The Mean Teacher method.	111
6.3	The 4 stages of the HRNet backbone and its different heads.	112
6.4	Proposed MTL-SSL model architecture.	113
6.5	Calculate the missing part and the false part between the generated segmentations, weakly labels, and real vascular masks by comparing them.	118
6.6	Hematoxylin and eosin staining for oncocytomas at different magnification.	126
6.7	H&E staining for other cancers at different magnification.	127
6.8	Segmentation results of the proposed MTL-SSL and fully supervised models for the four subtypes of RCC.	128
6.9	Segmentation results of the proposed MTL-SSL model on the TCGA dataset.	129

List of Tables

3.1	The number of image patches of each category in BigRCC.	51
3.2	The number of slices and image patches of each cancer in TCGA.	53
4.1	The explanation of each skeleton and lattice feature.	65
5.1	P-value of non-parametric tests of every feature.	101
5.2	The Accuracy obtained by 3 feature sets on different algorithms.	102
5.3	The benchmark results on the VRCC dataset.	102
6.1	Performance of different backbones.	120
6.2	Evaluate with different Strel Radius.	121
6.3	Performance of different models.	121
6.4	Performance of differently labeled data in training.	121
6.5	Performance of new subtype of RCC and other cancers.	126

