Deep Learning Networks for Automatic Brain Tumour Segmentation from MRI Data

Keerati Kaewrak

Department of Electronic and Electrical Engineering

University of Strathclyde

Glasgow, United Kingdom

A thesis submitted for the degree of

Doctor of Philosophy

2023

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Keerati Kaewrak

2023

Acknowledgements

I would like to express my sincere gratitude to all the people who have supported me during my Ph.D journey. Without every guidance, help, support, and feedback, this PhD thesis would not be possible to accomplished.

Firstly, I would like to express the deepest gratitude to my primary supervisor, Professor John Soraghan, for giving me the opportunity to pursue my PhD. His knowledge and expertise in the field have been incredibly valuable, and I am grateful for his encouragement and kindness. His belief in my ability has boosted my confidence. He has always been there for me throughout the ups and downs of my PhD journey. He have always patiently answered my questions and provided me with necessary tools. His precious feedbacks has helped shape my work and expand my understanding. I truly appreciate his guidance and the positive impact he has had on my academic and personal growth. Thank you for being an amazing supervisor.

Secondly, I would like to express my sincere gratitude to my second supervisor, Dr. Gaetano Di-Caterina, for his exceptional support throughout my PhD. His deep understanding and expertise in technical matters have been remarkable. When Professor John Soraghan retired, he stepped in seamlessly, and I am grateful for his commitment. The computational resource he provided has been crucial for my research, and I am appreciate his efforts in making it possible. Additionally, his assistance with the necessary paperwork has been immensely helpful, allowing me to focus on my thesis writing up. I am also thankful for his insightful feedbacks of my writing. Without all his support, I would not be able to complete my experiments and writing up.

I would like to express my appreciation to Dr. Derek Grose from Beatson West of

Scotland Cancer Centre for his contribution to this research. He has provided helpful feedback and advice on clinical aspects for this research.

I would like to thank my friend, Dr. Amlan Basu, for his excellent work. I appreciate all the collaborative projects we have worked on and the research outputs we have achieved together. I am grateful to my friends and colleagues at the Centre of Signal and Image Processing, especially Dr. Huiyi Wu, Dr. Baixiang Zhao, Dr. Weijie Ke, Dr. Paul Kirkland, Dr. Yannan Xing, Yinyong Zhang, and Ming Gong, for their friendship and knowledge exchange. I also thank the Thai student community in Glasgow for their friendship and supportive network.

I would like to thank my family for their unconditional love and support. I appreciate everything my parents have done for me so I can pursue my studies. I am also grateful to Mr. Adam McGhie for his love and support during my time in Glasgow.

Lastly, I would like to express my sincere gratitude to my sponsor, The Development and Promotion of Science and Technology Talents Project (DPST), the Royal Thai Government for the opportunity and financial support to study abroad.

Abstract

Early diagnosis and appropriate treatment planning are the keys to brain tumour patients' survial rate. Radiotherapy (RT) is a common treatment for brain tumour. RT planning requires segmentation of a gross tumour volume (GTV). Manual segmentation of the brain tumour done by experts oncologists or clinicians is time-consuming and subject to intra- and inter-observer variability. This research presents novel image processing and deep learning methods for automatic brain tumour regions segmentation from MRI data. The MRI data of brain tumour patients from Brain Tumour Segmentation or BraTS dataset from 2018-2021 are used in this study.

2D deep neural networks for semantic segmentation of brain tumour regions from 2D axial multimodal (T1, T1Gd, T2, and FLAIR) MRI slices is presented. This proposed network is trained and tested on manual consensus labels by experts from BraTS 2018 dataset. The network has similar architecture to U-Net, which consists of a stream of down-sampling blocks for feature extraction and a reduction in the image resolution, then a stream of up-sampling blocks to recover image's resolution, integrate features, and classify pixels. The proposed network improved feature extraction by introducing two-pathways feature extraction in the first block of the down-sampling to extract local and global features directly from the input images. Transposed convolution was employed in up-sampling path. The proposed network was evaluated for the segmentation of five tumour regions: whole tumour (WT), tumour core(TC), necrotic and non-enhancing tumour (NCR/NET), edema (ED), and enhancing tumour (ET). The results obtained from the modified U-Net achieved mean Dice Similarity Coefficient (DSC) of 0.83, 0.62, 0.45, 0.69, and 0.70 for WT, TC, NCR/NET, ED, and ET, respectively.

These results show a 9% improvement compared to the original U-Net's performance. 2D predicted segmentation obtained from the proposed network are stacked to visualise the tumour volume.

A novel deep neural network called 2D TwoPath U-Net for multi-class segmentation of brain tumour region is described. The proposed network has improved two-pathways feature extraction to provided cascaded local and global features from 2D multimodal MRI input. The proposed networks was trained using MRI data from BraTS 2019 dataset and test using MRI data from BraTS 2020 dataset. Data augmentation and different training strategies including the use of full-size images and patches were employed to improve the predicted segmentation. The results obtained from the proposed network feature all intra-structure (NCR/NET, ED, ET) of tumour to form the segmentation of WT and TC regions, and achieved mean DSC of 0.72 and 0.66 for WT and TC, respectively.

A novel 3D deep neural network for brain tumour regions segmentation from MRI data called 3D TwoPath U-Net is described. The network has a similar structure to the 2D TwoPath U-Net, and uses two-pathways feature extraction to capture local and global features from volumetric MRI data from BraTS 2021 dataset. The volumetric data were created using T1Gd and FLAIR modalities. To construct a 3D deep neural network with significantly high computational parameters, cropped voxels from volumetric MRI were used to reduce the input resolution. Furthermore, high-performance GPUs were employed to implement the network. The proposed network achieved the mean DSC of 0.87, 0.70, and 0.58 for WT, TC, and ET segmentation, respectively, which represents a 25% improvement compared to the previous segmentation results obtained using the 2D approach. Moreover, the 3D smooth tumour volume generated from the proposed network output provide a more visually representative depiction of the tumour.

Contents

Declaration **i**

Acknowledgements **ii**

Abstract **iv**

List of figures **ix**

List of tables **xiii**

List of acronyms **xv**

List of symbols **xvii**

1 Introduction **1**

1.1 Preface 1

1.2 Research motivation 2

1.3 Research hypothesis and objectives 3

1.4 Summary of original contributions 3

1.5 Author's publications 4

1.6 Thesis organisation 5

2 Brain tumour and its imaging **7**

2.1 Introduction 7

2.2 The human brain 7

2.3 Brain tumour 9

2.3.1 Brain tumours treatment options 10
2.3.2 Medical imaging for target volumes delineation 11
2.4 Magnetic resonance imaging (MRI) 14
2.4.1 The basics of MRI 14
2.4.2 The imaging process 17
2.4.3 Image artefacts 20
2.4.4 MRI of brain tumour 21
2.5 Summary 23

3 Brain tumour segmentation from MRI data 24
3.1 Introduction 24
3.2 MRI based brain tumour segmentation 25
3.2.1 Research challenges 25
3.2.2 Automatic brain tumour segmentation 26
3.3 Conventional methods for brain tumour segmentation 28
3.3.1 Unsupervised methods 29
3.3.2 Supervised methods 32
3.4 Deep learning based segmentation methods 35
3.4.1 Convolutional Neural Networks 35
3.5 U-Net: Convolutional networks for biomedical image segmentation . . . 37
3.6 Implementation resources for automatic brain tumour segmentation . . 41
3.7 Summary 42

4 BraTS dataset and evaluation metrics 44
4.1 Introduction 44
4.2 BraTS dataset 45
4.2.1 BraTS labels annotation protocol 46
4.3 Labels distribution of BraTS dataset 50
4.4 Loss functions of imbalanced segmentation 52
4.5 Evaluation metrics 53
4.6 Summary 57

5 Automatic brain tumour segmentation using modified U-Net 59

5.1 Introduction . . . 59

5.2 Deep learning based segmentation framework . . . 60

5.3 Modified U-Net . . . 61

5.4 Details of modified U-Net . . . 63

5.4.1 Local and global feature extraction . . . 63

5.4.2 Up-sampling: Transposed convolution . . . 65

5.4.3 Network implementation . . . 66

5.5 Experimental results . . . 66

5.6 Summary . . . 72

6 TwoPath U-Net for multi-class 2D segmentation 73

6.1 Introduction . . . 73

6.2 TwoPath U-Net for multi-class segmentation . . . 74

6.2.1 Cascaded two-pathway feature extraction . . . 75

6.3 Experiments . . . 76

6.3.1 Data preparation . . . 76

6.3.2 Network implementation . . . 79

6.4 Experimental results . . . 79

6.5 Brain tumour segmentation (BraTS) 2020 Challenge . . . 80

6.6 Summary . . . 82

7 3D TwoPath U-Net for brain tumour regions segmentation 84

7.1 Introduction . . . 84

7.2 3D brain tumour segmentation . . . 85

7.2.1 3D brain tumour segmentation framework . . . 85

7.2.2 Impact of GPUs on 3D deep CNNs implementation . . . 86

7.3 3D TwoPath U-Net . . . 87

7.4 Experiments . . . 88

7.4.1 3D Data-preparation . . . 88

7.4.2 Network implementation . . . 89

7.5 Experimental results 89
7.5.1 Segmentation results 89
7.5.2 Segmentation time 92
7.6 Conclusion 94

8 Conclusions and future work **96**
8.1 Conclusions 96
8.2 Future work 99

A **101**
A.1 Visualisation of feature maps obtained from 2D TwoPath U-Net 101

B ARCHIE-WeST **106**
B.1 ARCHIE-WeST access 106
B.2 Example of single GPU job script 107

C **109**
C.1 Automatic segmentation results from 3D TwoPath U-Net 109

Bibliography **111**

List of Figures

2.1 The three anatomical planes of the brain 8
2.2 Nerve cells and surrounding glial cells. 9
2.3 An illustration of radiotherapy planning volumes. 11
2.4 CT image with contrast-enhancing of a patient with WHO grade IV glioma and its 3 different tumour target volumes. 12
2.5 Different types of medical imaging modalities from high-grade glioma patients. 13
2.6 An illustration of a hydrogen atom with a single proton and a single electron. 14
2.7 The different type of spins. 15
2.8 Relaxation. 16
2.9 Relationship between TR and T1 contrast. 17
2.10 Relationship between TE and T2 contrast. 18
2.11 An illustration of MRI imaging process. 19
2.12 (a) Gibbs phenomena artefact, (b) motion artefact, and (c) chemical shift artefact. 20
2.13 (a) MRI image of the brain before inhomogeneity field or bias field correction, (b) inhomogeneity field or bias field, and (c) MRI image of the brain after bias field correction. 21
2.14 Multimodal MRI of high-grade glioma patient. 22

3.1 Conventional methods for brain tumour segmentation. 28

3.2 Various operators (a) Roberts operator, (b) Prewitt operator, and (c) Sobel operator. 31
3.3 An illustration of a biological neuron (left) and artificial neural network. 34
3.4 Mathematical model of an artificial neuron. 35
3.5 TwoPathCNNs architecture. 36
3.6 The U-Net network architecture. 38

4.1 Manual annotation of the tumour structures in different MRI images. (a) The whole tumour appear bright in FLAIR image. (b) The tumour core is visible in T2 image. (c) The enhancing tumour structure visible in T1Gd surrounding the necrotic components of the core. (d) All tumour structures combined to generate the final labels : edema (yellow), non-enhancing tumour (brown), necrotic (green), and enhancing tumour (blue). 47
4.2 Examples of the BraTS training dataset. Each row shows FLAIR, T1, T1Gd, T2, and FLAIR with ground truth labels; NCR/NET (red), ED (yellow), and ET (blue) from HGG patients (rows 1-3) and LGG patients (row 4-6). 49
4.3 Labels distribution from BraTS 2020 dataset. (a) Labels distribution included label 0 for everything else and background, label 1 for NCR/NET, label 2 for ED, and label 3 for ET. (b) Labels distribution included all labels except 0. 51
4.4 An illustration of regions used for calculating evaluation metrics. 53
4.5 (a) Confusion matrix. (b) The overlapping of the ground truth (blue contour) and predicted segmentation (red contour). 54

5.1 Deep learning based segmentation framework from 2D MRI data. 61
5.2 Modified U-Net network architecture. 62
5.3 Two-pathways feature extraction. 63
5.4 Feature maps obtained from the first layer of modified U-Net using (a) 3x3 kernels, and (b) 9x9 kernels. 64

5.5 The transposed convolution of a 3x3 kernel over a 3x3 input with 1 zero inserted between inputs, padded with 1 zeros padding and using a stride of 1. 65
5.6 Learning curve from (a) WT segmentation, and (b) ET segmentation. . 68
5.7 Visualisation of the predicted WT segmentation (red) and corresponding ground truth (blue) contours from patient ID Brats18_TCIA08_469_1, (a) tumour contours obtained from the modified U-Net on 2D axial FLAIR images, (b) stacked 2D contours of the ground truth, and (c) stacked 2D contours of the predicted segmentation. 69
5.8 Visualisation of original T1Gd image, ground truth, and predicted segmentation of brain tumour regions. The images are from various patient of BraTS 2018 dataset. 71

6.1 TwoPath U-Net network architecture for multiclass segmentation. . . . 75
6.2 Cascaded local and global pathways. 75
6.3 Feature maps obtained from cascaded TwoPath feature extraction using 3x3, 9x9, and 12x12 kernels. From (a)-(c) shows feature maps obtained from the first to third down-sampling block of the proposed TwoPath U-Net. 76
6.4 Examples of 2D axial view (left), frontal view (middle), and sagittal view (right) from training data with overlaid ground truth labels; NCR/NET (red), ED (yellow), and ET (blue). 77
6.5 Examples of (a) middle cropped images, and (b) overlapping cropped images. 78
6.6 Segmentation results from 2D TwoPath U-Net for multiclass segmentation. 82

7.1 3D TwoPath U-Net network architecture. 85
7.2 3D TwoPath U-Net network architecture. 88
7.3 Visualisation of multi-view original FLAIR, T1Gd, ground truth and predicted segmentation of WT. 90

7.4 Visualisation of ground truth (blue) and predicted segmentation (red) tumour volumes from two patients. 91
7.5 Example of training time for TC segmentation task. 92
7.6 Example of WT segmentation with prediction time. 93

A.1 Example of feature maps obtained from the first down-sampling block of the TwoPath U-Net using 3x3 kernels. 101
A.2 Example of feature maps obtained from the first down-sampling block of the 2D TwoPath U-Net using 9x9 kernels. 102
A.3 Example of feature maps obtained from the first down-sampling block of the 2D TwoPath U-Net using 12x12 kernels. 102
A.4 Example of feature maps obtained from the second down-sampling block of the 2D TwoPath U-Net using 3x3 kernels. 103
A.5 Example of feature maps obtained from the second down-sampling block of the 2D TwoPath U-Net using 9x9 kernels. 103
A.6 Example of feature maps obtained from the second down-sampling block of the 2D TwoPath U-Net using 12x12 kernels. 104
A.7 Example of feature maps obtained from the third down-sampling block of the 2D TwoPath U-Net using 3x3 kernels. 104
A.8 Example of feature maps obtained from the third down-sampling block of the 2D TwoPath U-Net using 9x9 kernels. 105
A.9 Example of feature maps obtained from the third down-sampling block of the 2D TwoPath U-Net using 12x12 kernels. 105

B.1 ThinLinc remote desktop access. 106
B.2 The login node and the example of job script submission terminal. . . . 107

C.1 Visualisation of multi-view original FLAIR, T1Gd, ground truth and predicted segmentation of TC. 109
C.2 Visualisation of multi-view original FLAIR, T1Gd, ground truth and predicted segmentation of ET. 110

List of Tables

2.1 Signal intensities of difference tissues on T1 and T2-weighted images. . . 19

3.1 Summary of brain tumour segmentation techniques based on human interaction. 27
3.2 Summary of unsupervised brain tumour segmentation techniques. 29
3.3 Summary of supervised brain tumour segmentation techniques. 32
3.4 CNNs based brain tumour segmentation techniques. 37
3.5 U-Net and its variants on biomedical image applications. 40

4.1 Example of name mapping for BraTS training dataset 46
4.2 Summary of tumour region structure and corresponding annotated labels from BraTS dataset. 50

5.1 Segmentation results of the proposed method using BraTS 2018 dataset. 67
5.2 Brain tumour regions segmentation performance comparison. 68

6.1 Comparative segmentation results using BraTS 2019 validation dataset. 80
6.2 Brain tumour regions segmentation performance comparison. 81

7.1 Segmentation results of the proposed method using validation data from BraTS 2021 dataset. 90

List of acronyms

2D Two dimensional

3D Three dimensional

AI Artificial intelligence

BBB Blood-brain barrier

C-MET C-methionine

CNS Central nervous system

CT Computed tomography

CTV Clinical target volume

DSC Dice Similarity Coefficient

ED Peritumoral edema

EORTC European Organisation for Research and Treatment of Cancer

ET Enhancing tumour

FET F-Fluorethyl-tyrosine

FID Free induction decay

FLAIR Fluid attenuated inversion recovery image

IIH Intensity inhomogeneity

IoU Intersectio over Union

GBM Glioblastoma multiforme

GTV Gross tumour volume

HGG High-grade glioma

LGG Low-grade glioma

ML Machine learning

MRI Magnetic resonance imaging

NCR/NET Necrotic and non-enhancing tumour

PET Positron emission tomography

PTV Planning target volume

RF Radiofrequency

ROI Regions of interest

RT Radiotherapy

SPECT Single-photon emission computed tomography

T1/T1w T1-weighted image

T2/T2w T2-weighted image

T1Gd, T1w+Gd Gadolinium enhanced T1-weighted image

TC Tumour core

TE Echo time

WHO World Health Organization

WT Whole tumour

List of symbols

Section 2.4

B_0	External magnetic field
^{1}H	Hydrogen atom
MH$_z$	Megahertz
T	Tesla
ω_0	Larmor frequency
γ	Gyromagnetic ratio
M_{xy}	Transverse magnetization
M_z	Longitudinal magnetization

Section 3.3

f	Greyscale image (section 3.3.1)
f	Activation function (section 3.3.2)
h	Binary image
k	Number of nearest neighbors
k	Number of clusters
T_F	Threshold value
w_i	weights
x_i	inputs

Section 4.4

P	Prediction
P_0	Predicted non-tumour region
P_1	Predicted tumour region

T	Ground truth
T_0	Actual non-tumour region
T_1	Actual tumour region

Section 5.3

DL_2	Dice loss
i	Convolution input
i'	Transposed convolution input
$\widetilde{i'}$	Size of stretched input
k	Convolution kernel
k'	Transposed convolution kernel
N	Number of element
o	Convolution output
o'	Transposed convolution output
p	Padding of convolution
p'	Padding of transposed convolution
p_n	Pixel value of predicted probability map
r_n	Pixel value of ground truth
s	Stride of convolution
s'	Stride of tansposed convolution
β_1	Exponential decay rate for the first moment estimates
β_2	Exponential decay rate for the second moment estimates
ε	Small positive value

Section 6.3

c	Classes
i	Pixels
L_{CCE}	Categorical cross entropy
p	Matrix of predicted values

y	Ground truth
β_1	Exponential decay rate for the first moment estimates
β_2	Exponential decay rate for the second moment estimates

Section 7.4

D	Dice loss
G	Ground truth
N	Number of voxels
P	Prediction
β_1	Exponential decay rate for the first moment estimates
β_2	Exponential decay rate for the second moment estimates

Chapter 1

Introduction

1.1 Preface

According to the Cancer Research UK statistics, brain tumour is the 9^{th} most common cause of cancer death in the United Kingdom and only 11% of patients survived brain tumour for 10 or more years [1]. Over 12,000 of new cases diagnosed with brain tumour each year between 2016-2018 in the UK, and around 3% are preventable [2]. Early diagnosis and treatment planning are important keys to increase brain tumour patients survival rate. In 2016, the World Health Organization [3] classified brain tumour into more than hundred types, and glioma is the most common type of the brain tumour that found in both adults and children. Tumour region segmentation gives necessary reliable information to perform the treatment planning such as observation, surgery, radiotherapy and chemotherapy.

Radiotherapy is a common treatment option for pre-surgery and post-surgery patients. Delineation of clinical target volumes are required in order to deliver the appropriate radiotherapy dose to the patient [4]. Traditional estimation of the region of interest (ROI) that covers the whole tumour required anatomical expertise, and it is done manually by the expert oncologists. The process involves manually drawing of the tumour border on each MRI slide of the patients and it is a time-consuming task [5]. Difference medical imaging modalities were employed to improve the accuracy of tumour assessment and MRI imaging has become the modality of choice because

of its superior in soft-tissues contrast. Semi-automatic and automatic brain tumour segmentation were proposed to assist the delineation process [6], [7], [8]. Deep learning has become an effective technique in image classification and segmentation tasks [9], and it is also employed in medical imaging segmentation [10], [11].

This research aims to use image processing techniques and deep learning technology to create novel automatic segmentation frameworks in order to extract appropriate region of interest which covers the whole tumour and avoid non-tumour regions. The results of this research will lead to advance tools for clinical brain tumour assessment using MRI data and be a great value to help improving the early diagnosis and the treatment planning.

1.2 Research motivation

Radiotherapy (RT) is common treatment option for brain tumour patients by using high energy x-ray energy to destroy the tumour cells. Radiation dose is usually delivered to the tumour plus a margin of 2 cm following the European Organization for Research and Treatment of Cancer (EORTC) guideline [4, 12]. The RT planning and planning target volume (PTV) require the delineation of the gross tumour volume (GTV) to differentiate tumour from the healthy brain tissues. Traditionally, the delineation of the GTV is done manually and relied on the expertise of the oncologists. It is a time-consuming, subjective and impractical for large-scale study [13]. There is a need for automated segmentation algorithm that gives precise location of the GTV and time-efficient to assist the brain tumour assessment.

Development of segmentation framework combining image processing techniques and deep learning based model that can automatically segment brain tumour could greatly benefit tumour assessment, treatment planning, and treatment follow-up for brain tumour patients.

1.3 Research hypothesis and objectives

The hypothesis of this thesis is that, an automatic brain tumour regions segmentation framework, developed using images processing techniques and deep learning method can obtain similar results when compared to current manual segmentation done by expert oncologists but less time consuming. To assess the stated hypothesis, the objectives of this research were to:

1. Determination of the appropriate methodology for the automatic brain tumour regions segmentation.

2. Developed a novel 2D image processing techniques and deep learning algorithm for automatic segmentation and visualisation of the brain tumour regions from multimodal MRI data.

3. Developed a novel 3D image processing techniques and deep learning algorithm for automatic segmentation and visualisation of the brain tumour region from multimodal MRI data.

4. To achieve fast and accurate result of the segmentation results comparing to manually segmentation ground truth done by the experts.

1.4 Summary of original contributions

The main contributions of this thesis are described below:

1. A novel 2D deep learning segmentation algorithm for whole tumour and intra-structures of tumour region is developed. The algorithm makes use of local and global feature extraction paths to the first block of a modified version of classical U-Net model. The proposed algorithm achieves 2D segmentation of all tumour regions and improves the segmentation accuracy compared to that from the original U-Net.

2. A novel 2D deep learning algorithm for multi-class brain tumour regions segmentation is developed. The model is called TwoPath U-Net which is the im-

proved two-pathways feature extractions deep learning algorithm. The proposed algorithm achieves multi-class 2D segmentation of the whole tumour, enhancing tumour, and necrotic core regions.

3. A novel 3D deep learning algorithm for brain tumour region segmentation is developed. The algorithm introduced 3D two-pathways feature extraction to the 3D convolutional neural network. The algorithm used 3D voxel from multimodal MRI data and achieves the 3D segmentation of the brain tumour volume of the whole tumour, tumour core, and enhancing tumour regions.

1.5 Author's publications

This thesis is mainly based on the works that have been published in technical conferences as follows:

Technical conference papers

1. **Kaewrak, K.**, Soraghan, J., Caterina, G.D. and Grose, D., 2020, October. TwoPath U-Net for automatic brain tumor segmentation from multimodal MRI data. In International MICCAI Brainlesion Workshop (pp. 300-309). Springer, Cham.

2. **Kaewrak, K.**, Soraghan, J., Di Caterina, G. and Grose, D., 2019, September. Modified U-Net for automatic brain tumor regions segmentation. In 2019 27th European Signal Processing Conference (EUSIPCO) (pp. 1-5). IEEE.

Conference abstracts

1. **Kaewrak, K.**, Soraghan, J., Di Caterina, G. and Grose, D., 2020. Automatic brain tumour regions segmentation using modified U-Net. Academic Journal for Thai Researchers in Europe, 1(1), pp.45-48.

Here follows an extended lists of publications published during the course of this research, which are not presented in detail in this thesis but have contributed in related technical challenges and applications:

Conference papers

1. Basu, A., **Kaewrak, K.**, Petropoulakis, L., Di Caterina, G. and Soraghan, J.J., 2022, June. Indoor home scene recognition through instance segmentation using a combination of neural networks. In 2022 IEEE World Conference on Applied Intelligence and Computing (AIC) (pp. 167-173). IEEE.

2. Basu, A., **Kaewrak, K.**, Petropoulakis, L., Di Caterina, G. and Soraghan, J., 2022, May. 3-Dimensional object recognition using 1-dimensional capsule neural networks. In IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI).

3. Basu, A., **Kaewrak, K.**, Petropoulakis, L., Di Caterina, G. and Soraghan, J.J., 2020, July. Modified capsule neural network (Mod-CapsNet) for indoor home scene recognition. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-6). IEEE.

1.6 Thesis organisation

The remainder of this thesis is organised as follows:

Chapter 2 provides the fundamental of the human brain structure and its function. The cause of the brain tumour and the need for medical target volumes delineation are then discussed. The brain tumour classification and treatment options are reviewed. The basic concept of the magnetic resonance imaging (MRI) and it acquisition are presented. Further, the advantages and limitations of different MRI modalities for brain tumour assessment are discussed.

Chapter 3 describes tumour structure and the corresponding label annotation protocol. The challenges associated to the automatic segmentation are discussed. Various automatic segmentation techniques including conventional and deep learning based methods are explored. Furthermore, the development of U-Net and its variants including biomedical image applications are discussed. Finally, implementation resources for deep learning based model are presented.

Chapter 4 introduces the real (clinical) MRI data of brain tumour patient from dataset that are used in this thesis. Examples of MRI image, intra-tumour structures,

and their corresponding labels are presented. Furthermore, the distribution of labels in the dataset and its issue of class imbalance are discussed. The evaluation metrics that can be used to measure the performance of segmentation algorithms are presented.

Chapter 5 describes the novel novel deep neural networks called modified U-Net for automatic tumour regions segmentation using 2D slices from multimodal MRI data available in BraTS 2018 dataset. The modification with two-pathways called local and global feature extractions is presented. Experimental results including the visualisation of the segmentation obtained from the proposed method implementation are presented and discussed.

Chapter 6 describes the novel deep neural networks called TwoPath U-Net for multi-class tumour segmentation using 2D slices from multimodal MRI data available in BraTS 2019 and 2020. The improved cascaded two-pathways feature extraction is presented. Experimental results including the visualisation of the segmentation obtained from the proposed method implementation and BraTS 2020 challenge participation are presented and discussed.

Chapter 7 describes the novel 3D deep neural networks called 3D TwoPath U-Net for tumour regions segmentation using 3D volumetric from multimodal MRI data available in BraTS 2021. 3D segmentation framework is presented and the impact of GPUs accelerate on 3D deep neural networks implementation is discussed. Experimental results including visualisation of the segmentation obtained from the proposed method are presented and discussed.

Chapter 8 presents discussion of the thesis reflecting on research objectives, key contributions, and limitation of the research. Overall conclusions for this thesis and suggestion for future work are presented in this chapter.

Chapter 2

Brain tumour and its imaging

2.1 Introduction

Brain tumours are abnormal growth of cells in the brain that can be either benign (non-cancerous) or malignant (cancerous). They can vary in size, location, and aggressiveness. The proper diagnosis and treatment planning of the brain tumour relies on a thorough understanding of their imaging characteristics. In this chapter, a brief summary of the human brain structure and its function is presented. The cause of the brain tumour, tumour classification, and treatment options are reviewed. The basic concept of magnetic resonance imaging (MRI), and the images acquisition are presented. Finally, the advantages and limitations of different MRI modalities for brain tumour assessment are discussed.

2.2 The human brain

Human central nervous system (CNS) [14] consists of the brain and spinal cord. The CNS is capable of taking in information and providing related response. The CNS can be distinguished by its colour called grey matter and white matter. The brain is the centre of the CNS, and it is protected inside of skull. The brain contains three main parts: cerebrum, cerebellum, and brainstem. The cerebrum is the largest part of the brain, located in the front area of the skull and consisting of right and left hemispheres.

The cerebrum can be divided into four lobes: frontal lobe, parietal lobe, temporal lobe, and occipital lobe. Each of them is normally associated with different human functions. According to [15] and [16], the frontal lobe controls higher functions including judgement, reasoning, problem solving. The frontal lobe also responsible for planning and executing movement. It is the last brain region to fully develop and not completing development until individuals reach their 20s. Parietal lobe is associated with interpretation of language and words, sense of touch, interpretation of signals. Temporal lobe plays a role in sensory processing, hearing, smell, taste, and higher-level visual processing. The temporal lobe also controls speech and memory. Occipital lobe's main function is processing of visual information, image recognition and image perception.

Because the brain is three dimensional structure, there are planes or axes that can be used to examine the nervous system. As shown in Fig. 2.1, the frontal or coronal plane is a vertical plane that dividing the brain into front and back pieces. The sagittal plane is also a vertical plane but dividing the brain into left and right pieces. Finally, the transverse or axial plane divides the brain into top and bottom regions. These anatomical planes divide the brain to be able to view the internal regions and structures of the brain.

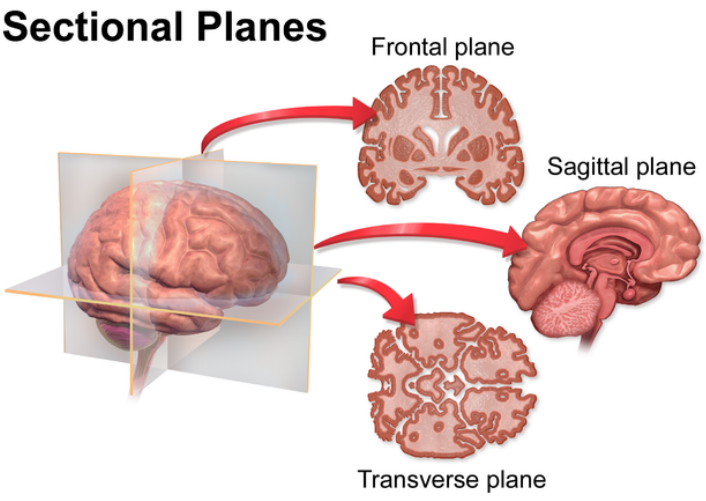


Figure 2.1: The three anatomical planes of the brain [17].

The brain is made of two types of cells; nerve cells or neurons and glial cells. The nerve cells transmit signals to and from the brain, and they consist of a cell body, dendrite, and axon. Axon conducts the nerve signal and transmits the signals to other neurons across the small gap called synapse. Dendrites act like arms to picking up the signals as messages. Neurons are surrounded by glial cells to support, protect and provide them oxygen. There are four types of glial cells in the CNS; astroglia or astrocytes, oligodendrocyte, ependymal, and microglia as shown in Fig. 2.2. The glial cells are 10 to 50 times more than nerve cells and they are the most common type of cells that involved brain tumours [18].

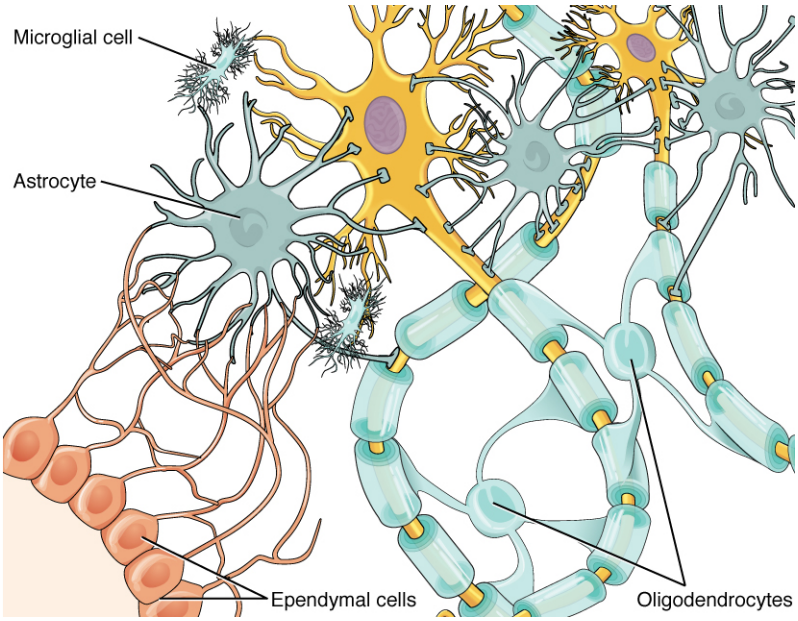


Figure 2.2: Nerve cells and surrounding glial cells [19].

2.3 Brain tumour

A brain tumour is an uncontrollable growth of abnormal cells of the brain tissues. Glioma [20] is the type of brain tumour that grows from glial cells. According to Cancer Research UK [21], gliomas are the most common type of the brain tumour that can be found in both adults and children. About 33 percent of all brain tumours are gliomas. They are graded based on the type of glial cells from which they grow. The

World Health Organization (WHO) [3] classification system categorized glioma from grade I to IV based on their histopathologic characterustics and their behaviours (how fast they are growing).

Grade I - Pilocytic astrocytoma typically occurs in the cerebellum or brainstem and found in children. Grade I tumours are slow growing.

Grade II - Low grade glioma includes astrocytoma, oligodendroglioma and mixed oligoastrocytoma typically occurs in the cerebral hemispheres and mostly found in young adults (20s-50s).

Grade III - Malignant glioma includes anaplastic astrocytoma, anaplastic oligodendroglioma, and anaplastic mixed oligoastrocytoma. They grow faster and more aggressive than grade II gliomas. They can spread to other part of the brain which is difficult for completely removal.

Grade IV - Glioblastoma multiforme (GBM) is malignant glioma and it is the most aggressive brain tumour. GBM spread to another part of the brain quickly with tentacle-like projections which are more difficult for completely removal.

Grade I and II are considered as low-grade glioma (LGG) [22], and are sometimes called benign. LGG are slow-growing and have less chance to recur after getting completely removed. Grade III and IV are high-grade glioma (HGG) [23], and are sometimes called malignant. They are fast-growing and more aggressive tumours.

2.3.1 Brain tumours treatment options

Treatment options for glioma patients vary depending on the size, location, and grade of the tumour. The goal of the treatment can also differ from patient to patient. It may be curative or it may focus on relieving symptoms. The treatment options are also often used in combination with one another. Following the guideline in [20], the treatment options for gliomas can be summarised as follows:

1. Observation is for low-grade gliomas where tumours are small and located in the areas that are not candidates for surgery or having high risk to cause damage of function loss after surgery. Some tumours may never grow.

2. Surgery is the most recommended choice for low-grade gliomas if the tumour can

be removed without causing damage of function loss.

3. Radiation can be a choice for either following surgery or in case that there is no surgery. Radiation treatment is recommended after surgery to slow the residual tumour growth, or recurrent tumours.

4. Chemotherapy is usually used for high-grade gliomas. For grade III gliomas, chemotherapy is given after radiation for 6-12 months. For grade IV gliomas, chemotherapy is given during and after radiation for 6-12 months.

2.3.2 Medical imaging for target volumes delineation

Radiotherapy (RT) is an important treatment option of brain tumour and it has been significantly improved over the past decade. For radiation oncologists, medical imaging is an important tool which gives reliable clinical information for the assessment of the brain tumours. Successful radiotherapy required the best possible characterisation of the location and extent of the tumour. Fig.2.3 illustrates the spatial relation of the three main volumes in radiotherapy target volumes [24]; the gross tumour volume (GTV), the clinical target volume (CTV), and the planning target volume (PTV). The GTV is the position and extent of the primary tumour. The CTV is the extent of tumour spread surrounds the GTV. After the GTV and CTV are established, the PTV must be added to allows uncertainties in treatment planning. The PTV is normal tissues surrounding the tumour.

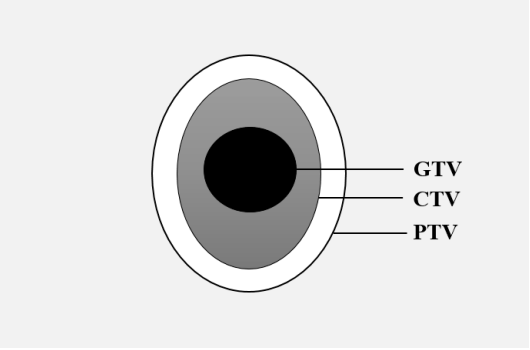


Figure 2.3: An illustration of radiotherapy planning volumes.

Fig. 2.4 shows an example of the three target volumes on Computed Tomography (CT) images of a patient with WHO grade IV glioma [24]. The first image (from the left) shows original CT with contrast-enhancing tumour, i.e. contrast agent is injected into the patient body before taking the scan. The second image shows the drawing of GTV (cyan line) which is the position and extent of the tumour. The third image shows the drawing of CTV (cyan dash-line) which covers the area of suspected tumour spread to the adjusted tissue, lymph vessels, or the draining lymph node stations. The radiation oncologist need expertise in anatomy to generate a rational and individual CTV. Finally, the PTV (magenta line) is introduced in the last image. It is the CTV with added margin, allows for uncertainties in radiotherapy planning and to ensure that the radiotherapy dose is delivered to the CTV. According to the European Organisation for Research and Treatment of Cancer (EORTC) [12], CTV is GTV plus a margin of 2 cm, and PTV is CTV plus a margin of 3-5 mm. In clinical practice, the edges of GTV are not always clear, better imaging to delineate gross tumour volume would help reduce the margin between CTV and PTV.

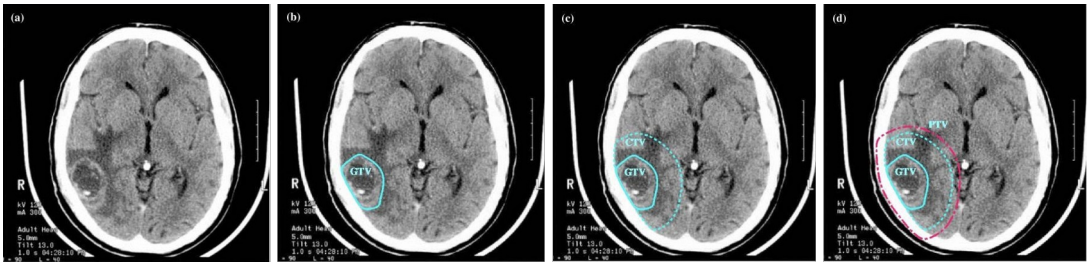


Figure 2.4: CT image with contrast-enhancing of a patient with WHO grade IV glioma and its three different tumour target volumes; GTV, CTV, and PTV [24].

Different medical imaging modalities are employed in different aspects of tumour volumes localisation and treatment planning. Computed Tomography or CT is a non-invasive medical procedure that uses specialized X-ray equipment to produce cross-sectional images of the body. Each cross-sectional image represents a slice of the person being imaged. As reported in modern brain tumour imaging [25], CT may be the first image modality employed in brain tumour patients. A CT scan of the brain can be

used to assess the brain tumour, lesions, and other injuries [26]. The CT scan process can be completed in under 30 minutes. However, the patient is exposed to radiation during the CT scan, which increases the potential risk of cancer [27].

Magnetic resonance imaging or MRI is current imaging modality of choice for brain tumour because of its excellent soft-tissue contrast and ability to image directly on multiple planes as described in [4]. The disruption of blood-brain barrier (BBB) cause the contrast enhancement on MRI. However, [28] shows that in case of large tumour parts that BBB is not yet affected, MRI sequences could underestimate the tumours mass. Subsequently, metabolic image modalities like Positron Emission Tomography (PET) and single-photon emission computed tomography (SPECT) are also employed to localise brain tumours rather than MRI alone. C-methionine (C-MET) and F-Fluor-ethyl-tyrosine (FET) are most two common amino acid tracers used in PET. C-MET and FET are superior in detecting infiltrative tumours cells that not detected in MRI. Fig. 2.5 shows an example of different medical imaging modalities from high-grade glioma patients. We can see that, the tumours and its extent are not easy to localise on CT and MRI images in this case as shown in Fig. 2.5(a) and Fig. 2.5(b). On the other hand, the tumour and the surrounding area are easier to localise in FET and C-MET as shown in Fig. 2.5(c) and Fig. 2.5(d).

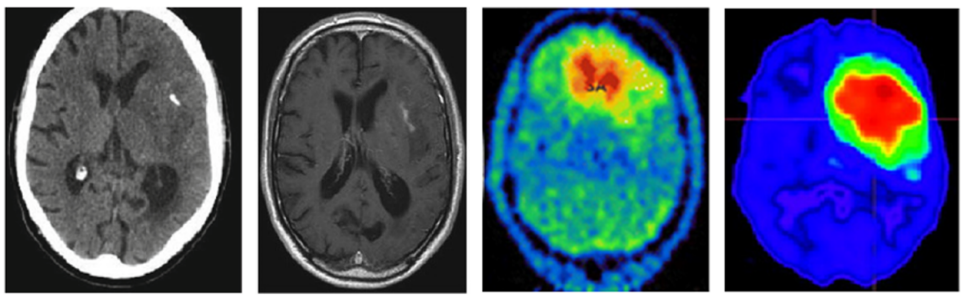


Figure 2.5: Different types of medical imaging modalities. (a) CT (b) MRI (c) FET and (d) C-MET from WHO high-grade glioma patients [28, 29].

2.4 Magnetic resonance imaging (MRI)

2.4.1 The basics of MRI

The basic concept of MRI is the spinning nuclear charge that produces small magnetic field called a magnetic moment. The medical MRI uses the signal from the nucleus of hydrogen atom (^{1}H) which naturally abundant found in human body for the image generation. As seen in Fig. 2.6, a hydrogen atom consists of a single proton and a single electron orbiting around the atom's nucleus. The proton has a positive charge and the electron has a negative charge which make hydrogen atom electrically neutral.

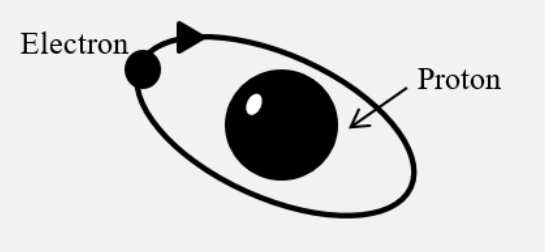


Figure 2.6: An illustration of a hydrogen atom with a single proton and a single electron.

Apart from its positive charge, the proton rotates about its axis and we called this process *spin* [30]. Fig. 2.7 illustrates the different types of the spins. In normal environment, millions of these spins are in random orientation and do not produce magnetic field as shown in Fig. 2.7(a). However, when these spins are put in a strong external magnetic field, the spins or the magnetic moments now align themselves in given directions and cause these spins to experience precession at a curtain frequency. This characteristic frequency is called Larmor frequency [31] and is given by the equation:

$$\omega_0 = \gamma B_0 \tag{2.1}$$

where ω_0 is Larmor frequency in megahertz (MHz), γ is a specific constant to a particular nucleus called gyromagnetic ratio and B_0 is the strong external magnetic field in tesla (T).

We can see from Fig. 2.7(b) that while the spins settle into steady state, the spins align parallel or anti-parallel to the magnetic field but the parallel alignment is slightly larger fraction. The longitudinal magnetization M_z is building up in z-direction by adding individual magnetic moment together. In this steady state, we can introduce energy to the spin system. A radiofrequency (RF) pulse is applied to tip the magnetization by exactly 90° and the result is that the longitudinal magnetization flips over and rotates to transverse magnetization M_{xy} on the xy-plane. The illustration of this process is shown in Fig. 2.7(c) and Fig. 2.7(d). This transverse magnetization (M_{xy}) can be detected as MR signal by a receiver coil placed in the xy-plane.

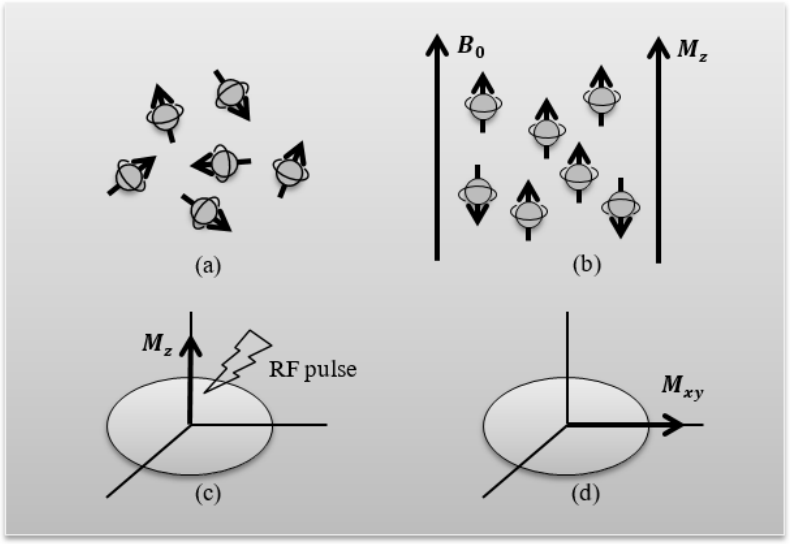


Figure 2.7: (a) the spins align randomly, (b) the spins align parallel to the strong external magnetic field B_0 and produce longitudinal magnetization M_z, (c) An RF pulse tips the magnetization vector by exactly 90°, (d) the entire longitudinal magnetization flips over after tipped by RF pulse and rotates into transverse magnetization M_{xy}.

After the RF pulse has been switched off, the M_{xy} rapidly fades and returns to the stable-state before the excitation. Fig. 2.8 illustrates the decreasing of M_{xy} which is called relaxation [32]. It involves two independent processes called spin-lattice interaction and spin-spin interaction. These two processes cause T1 relaxation and T2 relaxation, respectively.

T1 relaxation or longitudinal relaxation illustrated in Fig. 2.8(a) is the process that M_{xy} decays and slowly restores in z-direction. The time constant for this recovery is T1 and it depends on the strength of the external magnetic field B_0 and the internal motion of the molecules.

T2 relaxation or transverse relaxation illustrated in Fig. 2.8(b) is the decay of M_{xy} because the spins lose phase coherence. The phase coherence is when the spins have the 0° angle so the individual magnetization vector add up together and result with the M_{xy}. Losing phase coherence makes the M_{xy} becomes smaller and finally disappears. The loss of the phase coherence happens in two ways. Firstly, energy transfer between spins which causes a local change in the magnetic field and increase the loss of phase. This process occurs with the time constant T2. Secondly, time-independent of the external magnetic field B_0 which caused by the inhomogeneities in the magnetic field generator or by the person being imaged. It occurs in time constant T2* and normally shorter or equal to time T2. The loss of the MR signal due to T2* is called free induction decay (FID) and can be avoided by using spin echo sequences.

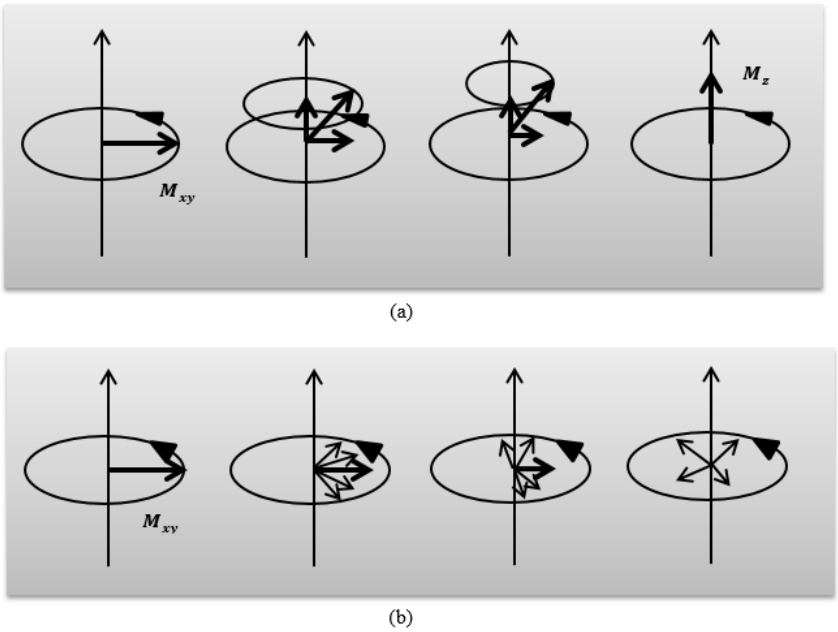


Figure 2.8: An illustration of (a) T1 relaxation. The transverse magnetization decays and restores from xy-plane to z-direction. (b) T2 and T2* relaxation. It is the process of transverse magnetization losing phase coherence and finally disappears.

2.4.2 The imaging process

The spins are excited many times and the result MR signal is recorded to generate MR slide. Each tissue has its own relaxation time which give us different image contrast [33]. Repetition time (TR) is the length of the relaxation period between two excitation pulses in millisecond (msec). Fig. 2.9 illustrates the relationship between TR and T1 contrast. When TR is long, the spins have time to regrow the longitudinal magnetization in z-direction before the next RF pulse is applied, which result with larger MR signal can be collected. If short TR (less than about 600 msec) is selected, tissues with a short T1 recover fast and give a large MR signal after next RF pulse and these tissues appear bright on MRI image. On the other hand, tissues with a long T1 which require more time to recover the longitudinal magnetization give less MR signal and appear dark on MRI image. An image acquired with short TR is called T1-weighted image or T1w. If a fairly long TR is selected (more than 1500 msec), all tissues have enough time to recover, the effect of T1 on image contrast will be small.

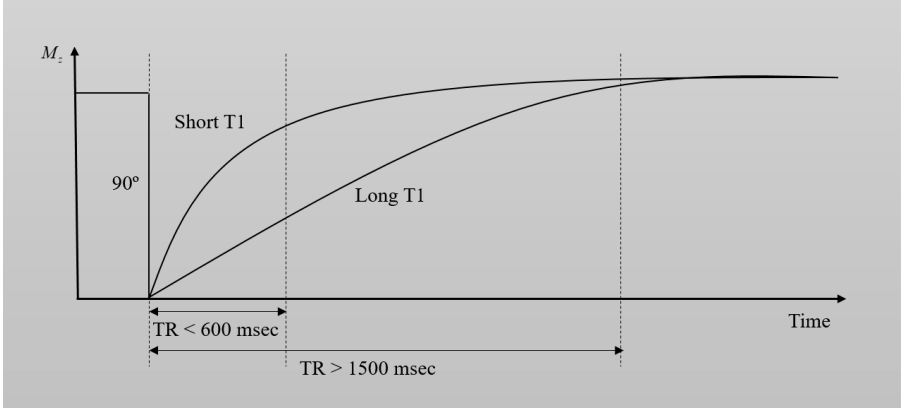


Figure 2.9: Relationship between TR and T1 contrast.

The echo time (TE) is the interval between application of the excitation pulse and collection of MR signal in millisecond (msec). TE determines the effect of T2 on image contrast. Fig. 2.10 illustrates the relationship between TE and T2 contrast. If the short TE (less than 30 msec) is selected, the signal difference between tissues are small because the relaxation is just started. If a longer TE (over 60 msec) is selected, tissues appear on MR image with different intensities due to their T2 relaxation times. Tissues

with short T2 appear dark because they lost most of their signal and tissues with long T2 still give strong signal so they appear bright. An image acquired with long TE is called T2-weighted image or T2w.

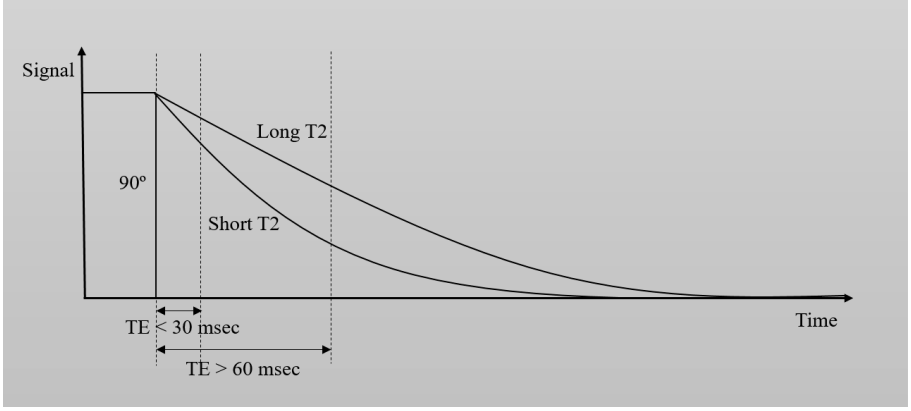


Figure 2.10: Relationship between TE and T2 contrast.

Table. 2.1 shows an example of different tissue and its contrast appearance on T1 and T2-weighted images. For T1-weighted image, tissues with a short T1 (e.g. fat) appear bright because they are able to regrow most of their longitudinal magnetization during the TR interval and produce strong MR signal. While, tissues with a long T1 (e.g. aqueous liquid, tumour, muscle, and compact bone) appear dark because they require more time to regrow their longitudinal magnetization. For T2-weighted image with long TE setting, tissues with a short T2 (e.g. muscle, connective tissues, and compact bone) appear dark because they lost most of their signal. While tissues with a long T2 (e.g. fat, aqueous liquid, and inflammatory tissue) still produce a strong signal and appear bright.

Fig. 2.11 illustrates the process of creating MRI images from the MR signals described in [34]. The process is that a patients body is divided into a set of slides and then each slide is cut into rows and columns to form a matrix of individual tissue voxel. The process consists of signal acquisition and image reconstruction from the acquired signals. During the acquisition process, the imaging cycle is repeated many times.

Table 2.1: Signal intensities of difference tissues on T1 and T2-weighted images [33].

Tissue	**T1-weighted image**	**T2-weighted image**
Fat	Bright	Bright
Aqueous liquid	Dark	Bright
Tumour	Dark	Bright
Inflammatory tissue	Dark	Bright
Muscle	Dark	Dark
Connective tissue	Dark	Dark
Compact bone	Dark	Dark
Air	No signal	No signal

The signals that emitted by tissues and received by the receiver coils are collected, digitized and stored in computer memory called k-space. After the signals data are collected, the data are sorted and delivered to the appropriate image pixels. Fourier transform [31] is the mathematical process using for this image reconstruction. The image is also divided into rows and columns to represents the corresponding voxel.

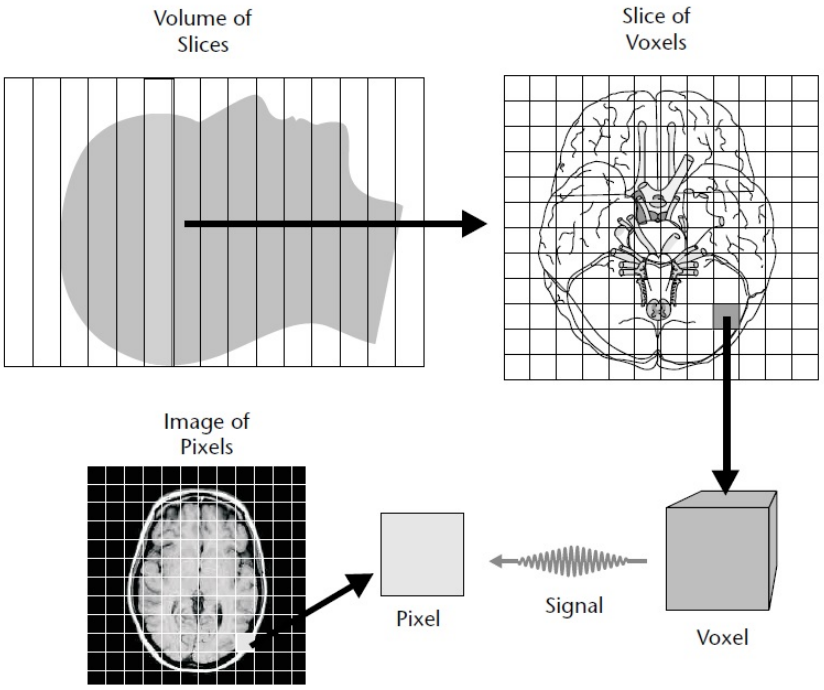


Figure 2.11: An illustration of MRI imaging process [34].

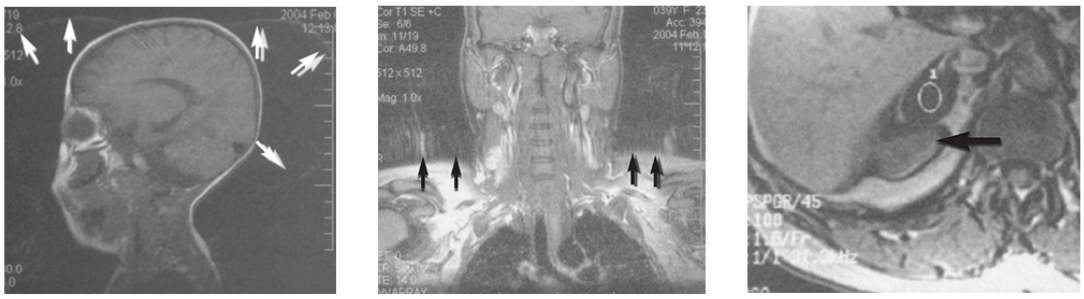
2.4.3 Image artefacts

Artefacts are undesirable objects that appear in images and do not represent a anatomical structure. There are many different factors that cause the artefacts in MRI images. In [35], the artefacts that appear on MRI images can be classified into groups based on their origin as follows;

1. Truncation artefacts or Gibbs phenomena [36] occur near sharp high contrast boundary and appear as multiple parallel bright and dark lines. They can be misinterpreted as a syrinx in the spinal cord. Fig. 2.12(a) illustrates Gibbs phenomena appears as multi parallel white and dark bands (white arrows) on T1 MRI of the brain [37]. Gibbs phenomenal is the consequence of using Fourier transform to reconstruct MR signals into images.

2. Motion artefacts or ghost artefacts caused by breathing, the heart beating, blood flow, cerebrospinal fluid pulsation and patients movement. These artefacts appear in the form of blurring or discrete set of lines or continuous smear as shown in Fig. 2.12(b).

3. Chemical shift artefacts are tissue-related artefacts which are the result of mismapped fat and water signal from the fluid-filled surrounding by fat structure. Chemical shift artefacts appear as dark or bright bands as shown in Fig. 2.12(c).



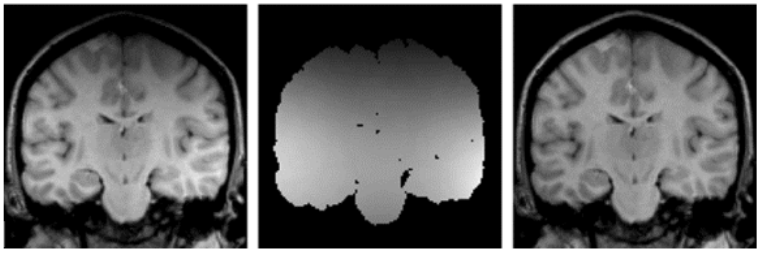
(a) (b) (c)

Figure 2.12: (a) Gibbs phenomena artefact, (b) motion artefact, and (c) chemical shift artefact [37].

4. Noise in MRI image are the results of a form of random, undesired RF energy [34] from the patient's body. The presence of noise in MRI image can reduce the image quality and limit the visibility of od low contrast object and difference among tissues.

5. Intensity inhomogeneity (IIH) or bias field [38] is a low frequency smooth undesirable signal that blurs the MRI images. IIH is the result of the inhomogeneities in the magnetic field of the MRI machine. Fig. 2.13 illustrates IIH in the MRI image of the brain. IIH reduces details of the image such as edges and contours as shown in Fig. 2.13(a). IIH affect the quality of medical image analysis methods such as segmentation and registration by creating the fake variations of image intensities.

The presence of noise and image artefacts can significantly reduce the ability of image analysis. Hence, image pre-processing is introduced to eliminate those undesired objects and improve the quality of the images. Morphological filtering [39] is a method that uses a structuring elements to remove small regions such as noise in the image. Gibbs phenomena can be reduced by applying post-processing algorithm developed by [40]. IIH that blurs the MRI image can be resolved by filtering methods [41], surface fitting [38], N3 or N4ITK methods [42]. The elimination of IIH (Fig. 2.13(b)) gives the image with better quality of edges and contours as shown in Fig. 2.13(c).



(a) (b) (c)

Figure 2.13: (a) MRI image of the brain before inhomogeneity field or bias field correction, (b) inhomogeneity field or bias field, and (c) MRI image of the brain after bias field correction. [41].

2.4.4 MRI of brain tumour

MRI has been an important tool for the oncologists and radiologists to diagnose and evaluate the brain tumour. With its superior in soft-tissue contrast, MRI provides an accurate determination of tumour location and its extent. T1-weighted, T2-weighted, fluid-attenuated inversion recovery (FLAIR), and postcontrast T1-weighted images are

typical modalities employed in tumour patient. Fig. 2.14 [43] shows an example of these four types of MRI image from a high-grade glioma patient. T1-weighted or T1w images are produced by using short TR and short TE times. T1w are most useful for providing anatomical detail and cerebrospinal fluid. As seen in Fig. 2.14(a) tumours appear as low signal intensity (dark) while the area of fat and haemorrhage appear as high intensity (bright) on T1w images. It is also allows for easy annotation of healthy tissues.

T2-weighted or T2w images are produced by using long TR and long TE times. T2w are most sensitive for lesion detection and cerebrospinal fluid which appear with high intensity as shown in Fig. 2.14(b). FLAIR images are produced by using very long TR an TE times. FLAIR images are T2w images with low signal of cerebrospinal fluid. Thus, FLAIR images can display lesion area including tumours and edema with higher signal intensity than T2w images which can be seen in Fig. 2.14(c).

However, the study in [43] shows that T2w or FLAIR images do not well present the distinction of tumour from surrounding edema. Gadolinium-enhanced T1-weighted or T1w+Gd images are current standard image modality for assessing glioma. Gd can be injected into patient during the scan and it changes signal intensities by shortening T1 time. In T1w+Gd images, tumour boarders appear brighter because of the contrast agent accumulates there due to the disruption of blood-brain barrier in the growing tumour region as shown in Fig 2.14(d). According to [44], T2 and T1w+Gd images are often used by radiologist to define radiological outer margins of the tumour.

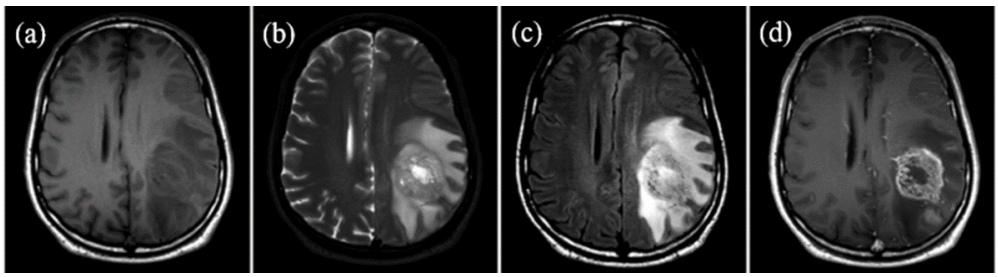


Figure 2.14: (a) T1w, (b) T2w, (c) FLAIR and (d) T1w+Gd images of patient with high-grade glioma [43].

2.5 Summary

The human brain anatomy and its functions were presented in this chapter. The cause of brain tumour, brain tumour classification and treatment options (observation, surgery, radiation, and chemotherapy) for different brain tumour grades were introduced. The medical image target volumes for tumour delineation were presented. We further discussed that MRI is the medical imaging of choice because of its superior in soft-tissue contrast. This chapter then presented the basic concept of MRI and image acquisitions to obtain different image modalities for brain tumour imaging. Finally, different types of MRI for brain tumour imaging are also discussed in this chapter. In the next chapter, machine learning and deep learning technology and its application for medical imaging will be presented.

Chapter 3

Brain tumour segmentation from MRI data

3.1 Introduction

The previous chapter introduced gliomas which are the most common type of brain tumour that can be found in both adults and children. Gliomas were categorised into different grades based on their characteristics and behaviour. Treatment options and treatment planning requires the segmentation of GTV and its location information. The segmentation process involves delineating the tumour regions from the surrounding healthy tissues, enabling accurate characterization and quantitative analysis. However, brain tumour segmentation is a challenging task due to the complex anatomical structures, and tumour appearance that varies from patient to patient. The medical imaging used in brain tumour image analysis was also discussed and MRI images have become the image modality of choice due to its superior in soft-tissues contrast.

In this chapter, MRI based brain tumour segmentation and the challenges associated with the task are discussed. Subsequently, we explore various conventional techniques employed for brain tumour segmentation. We then focus on the emerging role of deep learning, particularly U-Net based models, in achieving more accurate and robust brain tumour segmentation. Finally, we present the available implementation resources for

automatic segmentation algorithms.

3.2 MRI based brain tumour segmentation

MRI based brain tumour segmentation can be classified into 2D and 3D segmentation [6] based on dimensionality of the image domain. For 2D segmentation, tumour region is segmented separately from each MRI slice. For 3D segmentation, all sequential MRI slides are considered together as one MRI volume, then the segmentation of the tumour region is performed on the MRI volume to obtain the quantification and 3D visualisation of the tumour. In 2D images, the location od each image intensity is called a pixel, and in 3D MRI volume, it is called a voxel.

Manual segmentation of brain tumour is carried out by expert oncologists or clinicians to provides the ground truth information about the location of the brain tumour and its different tumour structures. Manual segmentation has been the traditional approach for brain tumour segmentation. However, it is a time-consuming task and subjective process. For example, all ground truth for BraTS 2012 challenge were segmented manually using the 3D slicer software [45] and it took about 60 minutes per subject, which making it impractical for large-scale studies and real-time applications. As a results, there is a need for automated and accurate segmentation techniques to overcome the limitations and enhance clinical decision-making.

3.2.1 Research challenges

Replacing the traditional manual segmentation with automated and accurate segmentation techniques could highly benefit the assessment of the brain tumour including diagnosis, treatment planning, and post-treatment follow-up for each patient. However, the brain tumour segmentation represents unique technical challenges and can be summarised as follows [5, 46]:

1. Lesion areas of the brain tumour are only defined based on the intensity changes relative to the surrounding normal tissues. This issue causes variations of manual segmentations done by experts especially when the intensity gradients between adjacent

structures are smooth or obscured by image artefacts.

2. Tumour structures vary from patient to patient in terms of size, extension, and localization. Glioma abnormally grow from glial cells which surround nerve cells as mentioned in previous chapter. Hence, both HGG or LGG could appear at any location inside the brain.

3. The mass effect induced by the growing lesion may displace normal brain tissues, and post-treatment resection cavities further limit the reliability of spatial prior knowledge for the healthy part of the brain.

4. Various imaging modalities, such as T1, T2, post-contrast T1Gd, FLAIR, perfusion and diffusion MRI, provides different types of biological information. Different type of image can be used to map changes in brain tissues caused by the tumour.

3.2.2 Automatic brain tumour segmentation

Table. 3.1 shows that brain tumour segmentation can also be classified into semi-automatic segmentation and automatic segmentation. Semi-automatic segmentation requires human-interaction to initialize a region of interest (ROI). In [47], users roughly draw around the suspected tumour on an MRI slide, then they applied the algorithm that combines the techniques such as region and edge-based active contours and level set method together to segment the tumour region. In Tumor-cut method [48], users draw the maximum diameter of the target GTV on the post-contrast T1 image before applied a cellular automata (CA) based seeded tumour segmentation algorithm.

Table 3.1: Summary of brain tumour segmentation techniques based on human interaction.

Techniques	**References**	**Description**
Manual segmentation	[5]	Tumour structures annotated by trained experts.
Semi-automatic segmentation		
- Guo et al.	[47]	Users give tumour initialization by roughly drawing the border.
- Tumor-cut method	[48]	Users draws maximum tumour diameter.
Automatic segmentation		
- Generative models	[49–51]	Model based on prior knowledge of brain anatomy and classify tumour as an outlier.
- Discriminative models	[10, 52–54]	Model extract low image features and learn the relationship between features and class label of the given input.

To eliminate the user interaction, fully automatic segmentation techniques are currently our main focused. We can broadly group the automatic MRI based brain tumour segmentation techniques into generative models and discriminative models as shown in Table. 3.1. Automated segmentations are obtained by combining explicit models of the anatomy and appearance in the generative model. In generative models, tumours are classified as an outlier from normal tissues. In [49, 50] Expectation Maximization (EM) algorithm [55] was employed to segment MRI volume into healthy tissue class and outlier class while [51] combined Standard hidden Markov field [56] with EM algorithm to classified five normal tissues classes (white matter, grey matter, ventricular cerebrospinal fluid, extraventricular cerebrospinal fluid, and other), then further classified the tumour into four tumour structures: edema, non-enhancing, enhancing and necrotic. On the other hand, discriminative models focus on learning the relationship between image intensities and the segmentation label. The models directly from the annotated ground truth labels to differentiate the appearance of normal tissues and the tumour. Classifiers such as decision tree and random forest in [52, 53] and deep neural networks in [10,54] extracted low level image features then find the relationship between

features and the label of the given voxel then classified each voxel into a certain class. These discriminative models have been giving outstanding accurate results according to the report in the multimodal brain tumour segmentation benchmark [5].

3.3 Conventional methods for brain tumour segmentation

Conventional techniques for automatic brain tumour segmentation have been extensively explored and reported [57–67]. Fig. 3.1 provides an overview conventional tumour segmentation techniques. These techniques include unsupervised and supervised methods. Unsupervised approaches include region-based methods, edge-based methods, thresholding, and clustering are used to define the boundary of the target tumour. On the other hand, supervised approaches apply training samples and learn from prior knowledge.

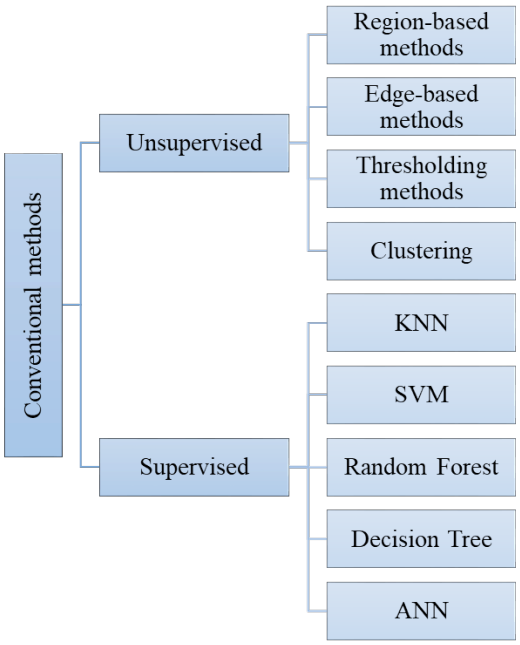


Figure 3.1: Conventional methods for brain tumour segmentation.

3.3.1 Unsupervised methods

Unsupervised methods for segmentation problem do not required prior anatomical knowledge nor training data to groups image pixels with homogeneous attributes together [66]. Table. 3.2 shows the summary of unsupervised techniques including edge-based methods, region-based methods, thresholding methods, and clustering methods.

Table 3.2: Summary of unsupervised brain tumour segmentation techniques.

Techniques	References	Description
Thresholding		
- Ilhan et al.	[57]	Segmentation of skull, brain, and tumour obtained by combining morphological operations and intensities based threshold.
- Otsu's method	[58, 59]	Optimum value of a global threshold combined with morphological operations was used to detect brain tumour.
Edge-based		
- Canny operator	[60]	Brain tumours were segmented using watershed transform and edge detection algorithm on colour MRI images.
- Sobel operator	[68]	Tumours were extract using Sobel edge detection combined with thresholding method.
Region-based		
- Region growing	[61–63]	Morphological operators and seed region growing are combined to segment the tumour from the brain MRI.
Clustering		
- K-means	[69–71]	K-means clustered MRI image into segments, then a threshold separates the tumour from the non-tumour regions.
- Fuzzy C-Means	[72, 73]	FCM using the Radius Contraction and Expansion process
- Mean-shift	[64, 65]	Modified K-means clustering method and initial segmentation using MS

3.3.1.1 Thresholding methods

Thresholding [74] is one of the simplest techniques that converts a greyscale image $f(i,j)$ into a binary image $h(i,j)$. The intensity values of the image pixels are compared to the threshold value T_F. The pixels with the same or higher intensity value than threshold value are assigned the value 1, whereas the pixels with lower intensity values are assigned 0 and can be defined as:

$$h(i,j) = \begin{cases} 1 & \text{if} \quad f(i,j) \geq T_F \\ 0 & \text{if} \quad f(i,j) < T_F \end{cases} \tag{3.1}$$

Using global threshold when a single threshold value is chosen for the entire image is the common approach. Otsu's method [58] where optimum value of the global threshold was applied to segment the brain tumour in [59]. However, it is not always possible to divide an image into two regions by using a single threshold value.

3.3.1.2 Edge-based methods

Edge detection methods [75] detect object boundaries in the image by assume that pixel values rapidly change at the boundary between two regions. Robert operator [74] as shown in Fig. 3.2(a) is one of the first edge detection operator by applying 2×2 masks to calculate gradient of the image and give diagonal edges. 2×2 masks are simple computational but they are not as useful for computing edges that are symmetric about their centre and the smallest metric should be 3×3. The extension of 2×2 to 3×3 gives Prewitt operator [74] as shown in Fig. 3.2(b). Sobel operator [74] as shown in Fig. 3.2(c) is a modified version of Prewitt operator where the central column/row are multiplied by 2. Sobel operator is prefered as edge detection because its better noise-suppression or smoothing. [68] combined Sobel operator with thresholding method to extract tumour regions using intensity information from MRI images. Canny algorithm [76] is another edge detection used in combination with watershed method [74] to detect brain tumour from colour MRI image in [60]. However, the algorithm did not work well with greyscale images of the brain tumours.

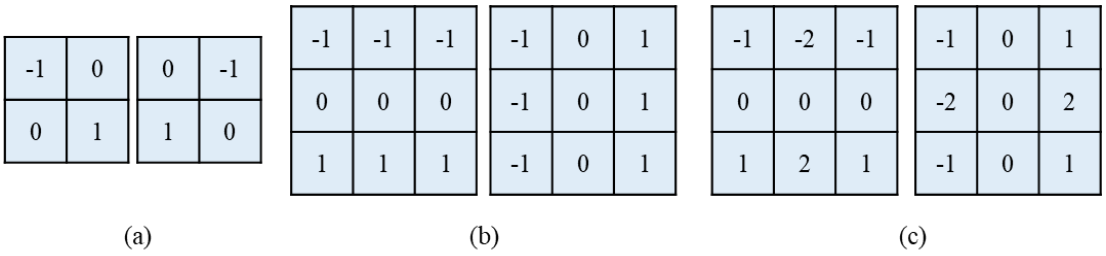


Figure 3.2: Various operators (a) Roberts operator, (b) Prewitt operator, and (c) Sobel operator.

3.3.1.3 Region-based methods

Region-based methods [75] is another intensity based techniques which assume that neighbouring pixels within one region have similar property. This leads to region growing [74] which is a process to group image pixels. The process begins with a seed point in each region of interest. Subsequently, the neighbouring pixels are connected to the seed point according to some predefined similarity criteria. The method is kept progressing till all the pixels are assigned in one of the regions. Region growing method was used to identify brain tumour from MRI in [62] while [61] combined region growing method based on the variances and gradients along and inside of the boundary curve to overcome the difficulty of threshold selection. Morphological operators such as dilation and erosion also applied to region growing method in [63] to enhance the edge detection of the tumour. Although, region-growing is effective intensity based segmentation method by incorporating local relationship between pixels, does not perform well for images with neighbouring regions sharing similar intensities [66].

3.3.1.4 Clustering

Clustering is segmentation technique that requires no training data to group image pixels with the same characteristics together. Common clustering techniques are K-means clustering, Fuzzy C-Means, and Mean-shift. K-means clustering algorithm [62] partitions image into distinct k clusters. The segmentation using K-means follows a few steps. Initially, k points are randomly selected as cluster centroids. Each image pixel then assigned to the nearest centroid. The centroid of each cluster is recalculated

by taking the mean of all image pixels assigned to that cluster. The steps are iterated until the new centroids no longer change. Although K-means algorithm is simple and computationally efficient, it may not optimised even after a large number of iterations.

Fuzzy C-Means (FCM) [77] is an extension of the K-means algorithm that allows soft or fuzzy clustering. Unlike K-means, which assign each image pixel to a single cluster, FCM assigns each image pixel a membership value indicating its degree of belongingness to each cluster. The main disadvantage of using FCM for image segmentation is that it is sensitive to noise [66]. Mean-shift (MS) [78] clustering is a non-parametric algorithm that identifies clusters by iteratively moving points towards the densest regions of the data distribution. One advantage of MS is that it can cluster data without requiring the predetermined number of clusters. However, it may struggle when the number of clusters changes abruptly, especially in high-dimensional data [67]. Brain tumour segmentation using clustering techniques can be found inn [64, 65, 69–73].

3.3.2 Supervised methods

Supervised methods are often used when training data with corresponding annotated labels is available. The supervised algorithms use labelled data to build a model that connects extracted features to classes during the training. Table. 3.3 shows the summary of supervised techniques including k-Nearest Neighbors, Support Vector Machine, Markov Random Field, and Artificial Neural Networks.

Table 3.3: Summary of supervised brain tumour segmentation techniques.

Techniques	**References**	**Description**
k-Nearest Neighbors	[79, 80]	Extracting different grey level pixel values to segment brain components
Support Vector Machine	[81, 82]	Using intensities and texture to differentiate tumour and non-tumour pixels
Markov Random Field	[83]	A three steps process that perform bias field correction and segmentation.
Artificial Neural Networks	[46, 66, 84],	Imitates the behaviour of biological neurons to learn classification

3.3.2.1 k-Nearest Neighbors

k-Nearest Neighbors (k-NN) [66] is a memory based algorithm that directly compared the new unlabelled data to the labelled data in the training set. The k in k-NN refers to the number of nearest neighbors to consider when making predictions. The classification process of k-NN occurs in two stages. First, the algorithm finds the k nearest neighbors of an unlabelled data in the training set. Second, the algorithm assigns the class to new unlabelled data using those neighbors. During the training phase, the algorithm stores the feature vectors and class labels of the training samples. Then the algorithm predicts the class labels of new unlabelled test sample by calculating the similarity of distance between the test sample and the training instances. It identifies the k nearest neighbors of the test sample from the training set and assigns the assigns their class label to the test sample based on majority voting. k-NN was employed in [79] to analyze the characteristics for each brain component including segmentation of light and dark abnormalities from FLAIR MRI, while [80] used inclusion of tissue type priors in features set to improve the performance of white matter lesion segmentation.

3.3.2.2 Support Vector Machine

Support Vector Machine (SVM) [66] is an algorithm that divides image into two classes by finding a hyperplane that best separate the data. The hyperplane is selected based on its maximum distance from the nearest data points on each side. To achieve this, SVM solves a particular optimization problem presented in [81]. [82] used SVM to differentiate tumour and non-tumour pixels on the basis of different features including intensities and texture. SVM can be used for binary segmentation of tumour region. However, there are more intra-tumour structures that can be extracted from the segmentation.

3.3.2.3 Markov Random Field

Markov Random Field (MRF) [6] is a statistics model that can be used within the segmentation method. MRF captures the spatial interactions between nearby pixels in an image, allowing for the modeling of various image properties. MRF are commonly employed in medical imaging because neighbor pixels of the image often belong to the

same class. This implies that anatomical structures consisting of only one pixel have a very low probability of occurring under the MRF assumption. MRF is usually used in combination with other techniques for segmentation such as K-means in [85], and Monte Carlo in [83]. MRF exhibits robustness to noise and artefacts in segmentation as presented in [6, 83].

3.3.2.4 Artificial Neural Networks

Artificial Neural Networks (ANNs) are inspired by biological neuron systems and modeled as collections of neurons. The biological neurons produce and transmit nerve signal through the gap called synapse, while ANN is connected by hidden layers as illustrated in Fig. 3.3. The output of some neurons can be an input of another neuron.

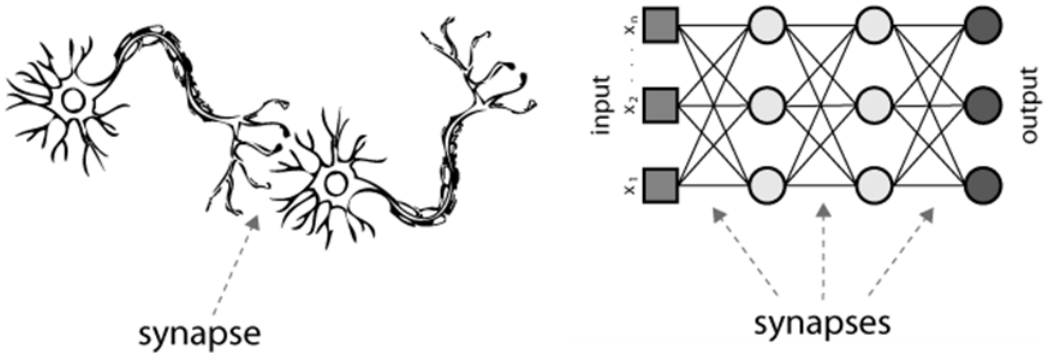


Figure 3.3: An illustration of a biological neuron (left) and artificial neural network [86].

Fig. 3.4 illustrate mathematical model of an artificial neuron. The neuron takes input $(x_1, x_2, ..., x_n)$, and each input is multiplied by its corresponding weights $(w_1, w_2, ..., w_n)$. The weighted inputs are subsequently summed to produce a net results. This net is then passed through an activation function f which gives the output of the neuron. Activation is used to introduce non-linearity to the neuron's output. Common activations for ANNs are Sigmoid, Tanh, Rectified Linear Unit (ReLU), and Leasky ReLU [87]. The number of layers in the networks defines how deep the ANN is. It could contain only 2 or 3 layers or even up to hundreds of layers [86]. Applications of ANNs in brain

tumour segmentation are reported in [46, 66, 84]. Details of these applications will be presented in the next section.

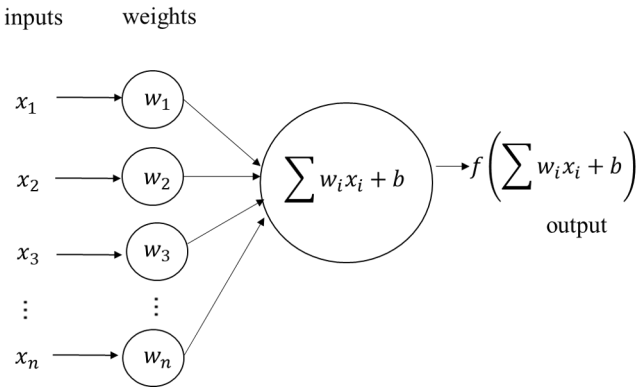


Figure 3.4: Mathematical model of an artificial neuron.

3.4 Deep learning based segmentation methods

3.4.1 Convolutional Neural Networks

Convolutional neural networks (CNNs or ConvNet) is the most popular deep learning model for image analysis including medical imaging applications. CNNs typically consists of convolutional layer, pooling layers and fully connected layer which are stacked together to form a complete CNNs architecture. They offer several advantages in medical image analysis, including modular interconnected operations, input modalities, input patch dimensions, time predictions, and contextual information [84]. CNNs have been developing over the past decades for image analysis. Initially, LeNet-5 [88] for handwritten digit recognition from MNIST dataset [89] was developed. It consists of 5 layers which makes it an excellent choice for grasping the basics of CNN due to its simplicity and fewer parameters. Series of 5x5 convolutions were used in LeNet-5. Alex-Net [90] revolutionized the computer vision by introducing the deep CNN architecture using 11x11, 5x5, and 3x3 convolutions, and won the ImageNet competition in 2012 [91]. AlexNet network architecture is much larger compared to the LeNet-5, which required significantly higher computational power. This results in the gap over 20 years between their development. Later VGG [92] introduced 16 and 19 layers CNNs

with series of small 3x3 filters and showed significant improvement over AlexNet. These CNNs methods classify entire images from ImageNet dataset into specific class. However, semantic segmentation problem requires each pixel of the image to be assigned to a certain class label.

Table. 3.4 shows the summary of CNNs based brain tumour segmentation techniques developed in the past years. [93] designed CNN for segmenting the infant brain tissues using various input patch sizes. The CNN classified central pixel of input patch into white matter, grey matter, or cerebrospinal fluid. Later, deep 11-layers CNNs using small kernel size (3x3) in convolution layers to extract features with the increasing depth for brain tumour segmentation was presented in [10]. TwoPathCNNs [54] which used both smaller kernel size (3x3) and larger kernel size (7x7 and 13x13) as illustrated in Fig. 3.5 to obtain local features and global features, then concatenated both type of features together and used the concatenation of feature to classify the class of central pixel. This work is one of the early multi-path networks. These methods performed segmentation from 2D input data. Deep 3D CNNs were proposed to extract brain from MRI volume in [94], while DeepMedic [95] used ensemble networks where each path of networks can be considered as sub-network to segment tumour regions. These applications of CNNs on brain imaging were implemented on available public data such as BraTS [10, 54], ISLES [95], IBSR, LPBA40 and OASIS [94].

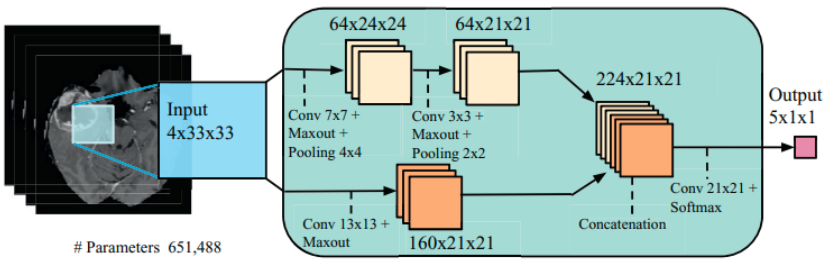


Figure 3.5: TwoPathCNNs architecture [54].

Table 3.4: CNNs based brain tumour segmentation techniques.

Method	Dataset	Description
Zhang et al. [93]	private dataset	Infant brain structures segmentation using various input patch sizes
Kleesiek et al. [94]	IBSR, LPBA40 and OASIS	3D CNNs for extracting brain from MRI
Pereira et al. [10]	BraTS 2013, BraTS 2015	Deep 11 layers CNNs with small kernel
TwoPathCNNs [54]	BraTS 2013	Deep CNNs with two-pathways feature extraction.
DeepMedic [95]	BraTS 2015, ISLES 2015	Ensemble 3D CNN with conditional random field

3.5 U-Net: Convolutional networks for biomedical image segmentation

The U-Net [96] is a fully convolutional network [97] that take arbitrary input size and produce correspondingly-size output. The network consists of down-sampling path, bottleneck, and up-sampling path forming the U shape architecture as shown in Fig. 3.6. The down-sampling path built of typical Convolutional networks. It consists of two 3x3 convolutions, each follows by ReLU as activation function and a 2x2 max pooling operation with stride of 2. The number of feature channels are doubled at each downsampling step, starting with 64 feature channels for the first block, 128 for the second, and so on. This downsampling path captures the context of the input image in order to perform segmentation. The bottleneck is the part between the downsampling path and upsampling path. It consists of 3x3 convolutions follows bu ReLu. The up-sampling path expand the features by a 2x2 up-convolution that half the number of feature channels, then concatenated with the cropped features from the corresponding down-sampling step, and two 3x3 convolutions, each followed by ReLU. The cropping is necessary due to the loss of border pixels in every convolution. The finally layer is a 1x1 convolution that map each 64-component feature vector to the desired number of classes. In total the network has 23 convolutional layers.

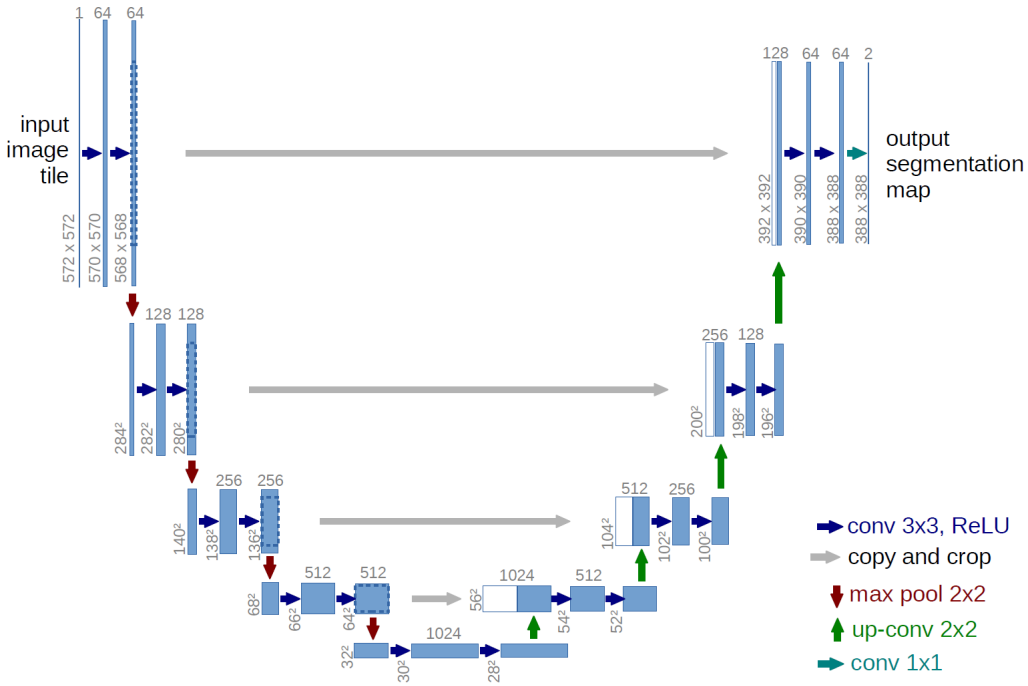


Figure 3.6: The U-Net network architecture [96].

U-Net model was designed to tackle the downside of typical CNNs that require large amount of training data to obtain accurate classification. U-Net requires less training data to achieve accurate classification and also includes the localisation as semantic segmentation that assign class label to each pixel. Hence, U-Net is suitable for medical image segmentation including brain tumour segmentation where there is limited training data. Applications of U-Net to biomedical imaging can be found in [96, 99–103]. They include both 2D and 3D segmentation of cell structures, organs, and lesion. The image modalities used ranges from microscopical images, CT, and MRI. Various modification of U-Net [104–106] also employed in different organs segmentation including blood vessels, lung, and hippocampus of the brain.

Table. 3.5 shows applications of U-Net and its variant on biomedical image segmentation problems. U-Net based methods for brain tumour segmentation can be found in [100, 101, 107–109]. Original U-Net was employed in [100] for 2D brain tumour segmentation and achieved 0.86 dice similarity coefficient (DSC) for whole tumour (WT) and tumour core (TC) segmentation. However, the method was unable to segmented

enhancing tumour (ET) from low-grade glioma (LGG) patients. Two-stage cascaded U-Net [101] was developed to segment the substructures of brain tumour from coarse to fine. The two-stage of the method connected two U-Net like models together to form cascaded network. The first stage returned rough or course segmentation map to the next stage U-Net to predict more accurate segmentation. Two-stage cascaded U-Net achieved 0.90 DSC for WT segmentation, over 0.8 DSC for TC and ET segmentation, and won first place in BraTS 2019 challenge. nnU-Net [107, 110] later won BraTS 2020 challenge with 0.89, 0.85, and 0.82 DSC for WT, TC, and ET segmentation, respectively. nnU-Net used a method that automatically configures itself including preprocessing, network architecture, training and post-processing for any new task. The nnU-Net set new state-of-the-art solution for brain tumour segmentation. Optimized U-Net [111] and NAVUTO [108] which are top rank teams from BraTS 2021 were developed based on nnU-Net achieved over 0.90 DSC in all tumour regions segmentation. These top performance methods from the BraTS 2019 to 2020 employed 3D volumetric MRI as input data.

Table 3.5: U-Net and its variants on biomedical image applications.

Method	**Application**	**Description**	**Performance**
U-Net [96]	neuronal structures from microscopical images	Introduced fully networks with features concatenation	0.92 IoU
V-Net [104]	Prostate from MRI	3D segmentation and proposed dice loss	0.87 DSC
U-Net [100]	Brain tumour from MRI	2D U-Net	0.86 DSC
3D U-Net [99]	Xenpus kidney from confocol microscopic data	3D segmentation form volumetric data	0.86 IoU
U-Seg-Net [105]	hippocampus from MRI	multi-view ensemble network	0.89 DSC
CE-Net [106]	Optic disc, blood vessel, lung, and cell	Context encoder network	0.95 Accuracy
Two-Stage Cascaded U-Net [101]	Brain tumour from MRI	3D segmentation and the winner of BraTS 2019	0.88 DSC
Channel-Unet [102]	Liver and tumour from CT images	Introduced spatial channel-wise convolution	0.94 DSC
SA-UNet [103]	Retinal vessel from fundus images	Proposed spatial attention module	0.97 Accuracy
nnU-Net [107]	Brain tumour from MRI	BraTS 2020 winner	0.88 DSC
Optimized U-Net [111]	Brain tumour from MRI	BraTS 2021 winner	0.91 DSC
NVAUTO [108]	Brain tumour from MRI	BraTS2021 rank 2	0.91 DSC

3.6 Implementation resources for automatic brain tumour segmentation

Deep Learning has been providing *state-of-the-art* solutions for computer vision problems including medical image analysis such as classification and segmentation. The availability of frameworks and tools in the programming languages such as MATLAB and python is one of the key to the success of deep learning in computer vision [112].

MATLAB is a matrix-based language software commonly used for mathematics computation and graphic visualization. MATLAB provided built-in tools for various data analysis problems including machine learning application. The most relevant toolboxes that provided by MATLAB for the machine learning application are Statistics and Machine Learning Toolbox, Neural Network Toolbox, Computer Vision System Toolbox, and System Identification Toolbox. These toolboxes are available to access in [122]. MATLAB also provide image processing applications e.g. Image Segmenter that can be used to generate binary mask or ROI label for medical image segmentation problems. These applications can be found in [123].

Python is one of the most popular programming languages among machine learning and deep learning researchers. Various powerful libraries and frameworks available in python make it easier to analyse data, develop algorithms, and create models. For deep learning programming task, the examples of important python libraries are NumPy, SciPy, Scikit-learn, and Matplotlib [113–117]. We can also choose to implement the model on different frameworks such as Keras [118], TensorFlow [119], Tensorlayer [120], and Pytorch [121].

Traditionally, the training phase of deep learning model takes the longest time to achieve. Using Graphical processing units or GPUs can significantly reduce the training time. GPUs allow us to run deep learning model with massive numbers of parameters quickly and efficiently by enable parallel training and tasks distributing. GPUs are also optimised to perform computations faster then non-specialised hardware.

Google Colaboratory or Colab is an online Jupyter notebook from Google which allows any user to write and execute python code through browser. Colab provides

Jupyter notebook with limited free GPUs access, memory, and runtime. Colab also provides paid subscription service called Colab Pro and Colab Pro+ for faster GPUs access, more memory, and longer runtime.

ARCHIE-WeSt [124] is the Research Computing Centre for the West of Scotland based at the University of Strathclyde. ARCHIE-WeSt provides computing facilities for multi-disciplinary research. The machine consists of different computational nodes including two GPU nodes (Lenovo SR670 servers) of eight NVIDIA A100 GPUs, and 192GB RAM (four GPUs per server). Storage of 825TB which capable of over 12GB/s data transfer is also available. Further information of ARCHIE-WeSt can be found from [124].

In this thesis, we used the combination of available implementation resources for different tasks. We used MATLAB and Colab notebook to initially visualise and study the data. 2D segmentation framework using Tensorflow and Keras libraries with local GPUs will be presented in chapter 5 and 6. 3D segmentation frameworks using high performance computing resource from ARCHIE-WeSt will be presented in chapter 7.

3.7 Summary

This chapter reviewed MRI based brain tumour segmentation techniques. Traditional segmentation of the brain tumour was done manually which is time-consuming and subjective to individual expertise. The need to replace traditional segmentation with automated segmentation framework was addressed. Due to the tumour appearance is unique in each patient, the challenges come with the automatic segmentation framework was also discussed.

This chapter also reviewed the conventional techniques for brain tumour segmentation including unsupervised and supervised approaches. Unsupervised approaches such as thresholding, edge-based, region-based, and clustering techniques explored the characteristics of the data without prior knowledge. While supervised approaches such as k-Nearest Neighbors, Support Vector Machine, Markov Random Field, and Artificial Neuron Networks utilized the labelled data to associate the extracted features and the

labels.

Further, the development of deep learning based techniques including convolutional neural networks or CNNs for computer vision problems were presented. Different CNNs architectures using different filter sizes for brain structures and brain tumour segmentation were explored. These CNNs used patches input of both 2D and 3D MRI data. However, these CNNs were only able to classify the centre pixel of the patches input which resulted in a need of post-processing to obtain the segmentation mask of the tumour.

Subsequently, this chapter reviewed the semantic segmentation network architecture called U-Net which has become the *state-of-the-art* solution for biomedical imaging application. U-Net and its variations of modification were employed in medical image segmentation including brain tumour segmentation. U-Net has been used as the backbone of the automatic brain tumour segmentation algorithms development since 2016 utilizing both 2D and 3D MRI data. However, the ET region remains a challenge of the segmentation task due to it small appearance or sometimes completely absent especially in LGG patients. Finally, available deep learning libraries and implementation resources for the development of automatic segmentation framework were presented.

In the next chapter, the real clinical MRI data of the brain tumour patients, and the manual segmentation for the labels annotation protocol will be presented. The labels distribution and its imbalanced problem will be discussed. Finally, the techniques used to mitigate the biases of the imbalanced data including loss functions and evaluation metrics will be introduced.

Chapter 4

BraTS dataset and evaluation metrics

4.1 Introduction

The previous chapter shown that deep learning based techniques have become the powerful tool for automatic brain tumour segmentation. The performance of deep learning models surpass the traditional approaches and enable precise and efficient tumour segmentation. However, the success of these techniques relied on the quality and diversity of the data used in model training. Real data capture the complexity and variety presented in clinical scenarios.

In this chapter, the real MRI data of the brain tumour patients that are available for segmentation task is introduced. Examples of MRI image, intra-tumour structures, and their corresponding labels annotation protocol are presented. Furthermore, the distribution of labels in the dataset and its issue of class imbalance are discussed. Finally, the loss function for imbalanced segmentation, and the evaluation metrics that can be used to train and measure the performance of segmentation algorithms are presented.

4.2 BraTS dataset

In this thesis, we used MRI data from the Brain Tumour Segmentation or BraTS from 2018 - 2021. The dataset contain multimodal MRI scans of glioma patients. The data are published mainly for the BraTS challenge. The challenge runs in conjunction with Medical Image Computing and Computer Assisted Interventions (MICCAI) conference [5, 125–129]. The clinical tasks focused in the challenge are the segmentation of the brain tumour, the prediction of patients overall survival, and the evaluation of the uncertainty in tumour segmentation. The MRI scans of the glioma patients are available in the Neuroimaging Informatics Technology Initiative or NIfTI format (.nii.gz). The dataset provides native (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR) MRI volumes. Each year, BraTS challenge publish training dataset, validation dataset and testing dataset for the competition. The CBICA's Image Processing Portal [130] and Synapse [131] are the platforms for the BraTS data request submission for BraTS 2018-2020 and BraTS 2021, respectively. These platforms are also used for the evaluation for the clinical tasks.

The number of glioma patients images data from the BraTS training dataset varies each year and includes both high-grade glioma (HGG) and low-grade glioma (LGG) patients. In BraTS 2018, there were MRI scans available from 210 HGG and 75 LGG patients. For BraTS 2019, there were MRI scans from 259 HGG and 76 LGG patients. In BraTS 2020, the dataset included MRI scans from 263 HGG and 76 LGG patients. Finally, BraTS 2021 provided 1251 MRI scans from glioma patients without the separation of HGG and LGG. The number of available patients' data has increased every year. All the samples available in the previous years of the BraTS dataset are included in the current year. For example, all the MRI data from BraTS 2018, 2019, and 2020 are included in the BraTS 2021 with addition of new patients.

Table. 4.1 shows examples of the name mapping reference for glioma patients across the BraTS training dataset from 2018-2020. Each row of the table represents a specific grade (HGG or LGG) and the corresponding patient ID for each year. In BraTS 2018 and BraTS 2019, patient ID change between years but maintain the same format.

While the patient ID in BraTS 2020 are presented in sequential format. Note that some entries are labelled as "NA" indicating that specific patient data is not available for that particular grade and year combination. The training dataset provided the name mapping information sheet along with the ground truth segmentation of each patient. Since BraTS 2021 no longer separates HGG and LGG patients, and the patient IDs are in sequential format, we did not include BraTS 2021 name mapping in the table.

Table 4.1: Example of name mapping for BraTS training dataset

Grade	**BraTS 2018 subject ID**	**BraTS 2019 subject ID**	**BraTS 2020 subject ID**
HGG	Brats18_CBICA_AAB_1	BraTS19_CBICA_AAB_1	BraTS20_Training_001
HGG	Brats18_CBICA_AAG_1	BraTS19_CBICA_AAG_1	BraTS20_Training_002
HGG	Brats18_CBICA_BHK_1	BraTS19_CBICA_BHK_1	BraTS20_Training_087
HGG	Brats18_CBICA_BHM_1	BraTS19_CBICA_BHM_1	BraTS20_Training_088
HGG	NA	BraTS19_CBICA_ANV_1	BraTS20_Training_089
HGG	NA	BraTS19_CBICA_AOC_1	BraTS20_Training_090
HGG	Brats18_2013_10_1	BraTS19_2013_10_1	BraTS20_Training_130
HGG	NA	BraTS19_TMC_30014_1	BraTS20_Training_157
HGG	NA	NA	BraTS20_Training_336
HGG	NA	NA	BraTS20_Training_369
LGG	NA	BraTS19_TMC_09043_1	BraTS20_Training_270
LGG	Brats18_TCIA09_451_1	BraTS19_TCIA09_451_1	BraTS20_Training_271
LGG	Brats18_TCIA10_449_1	BraTS19_TCIA10_449_1	BraTS20_Training_287
LGG	Brats18_TCIA13_624_1	BraTS19_TCIA13_624_1	BraTS20_Training_334
LGG	Brats18_TCIA13_634_1	BraTS19_TCIA13_634_1	BraTS20_Training_335

4.2.1 BraTS labels annotation protocol

The multimodal MRI (T1, T1Gd, T2, and FLAIR) scans of the BraTS dataset were obtained from various MRI scans with different image acquisition settings. Therefore, these MRI were co-registered to a common anatomical template, and re-sampled into 1x1x1 mm^3 with the size of 240x240x155 where 155 is the number of MRI slices per patient [126]. The ground truth of the tumour structures were manually segmented by a team of expert radiologists. The annotations are label 1 denoted the necrotic and non-enhancing tumour (NCR/NET), label 2 denoted the peritumoral edema (ED), label 4 denoted the enhancing tumour (ET), and label 0 for everything else (non-tumour

region and background of the image). The manual segmentation protocol for defining different visual tumour structures of the BraTS dataset can be summarised as follows:

1. The edema was primarily segmented from T2 images. Then the extension of the edema was cross-check using FLAIR images. The initial edema contain the core structures as shown in Fig. 4.1(a). It is then relabelled in subsequence steps.

2. The gross tumour core including both enhancing and non-enhancing structures was first segmented by evaluating hyper-intensities in T1Gd in Fig. 4.1(c) together with the inhomogenous component of the hyper-intense lesion visible in T1 and the hypo-intense regions visible in T1 as shown in Fig. 4.1(b).

3. The enhancing core was subsequently segmented by thresholding T1Gd intensities within the resulting gross tumour core including the Gd enhancing tumour rim and excluding the necrotic centre as shown in Fig. 4.1(c). The appropriate intensity thresholding was determined on a case-by-case basis.

4. The necrotic core was defined as the low intensity necrotic structure within the enhancing rim visible in T1Gd as shown in Fig. 4.1(c).

5. Finally, the non-enhancing core structures were defined as the remaining part of the gross tumour core after subtraction of enhancing core and necrotic core structures as shown in Fig. 4.1(d).

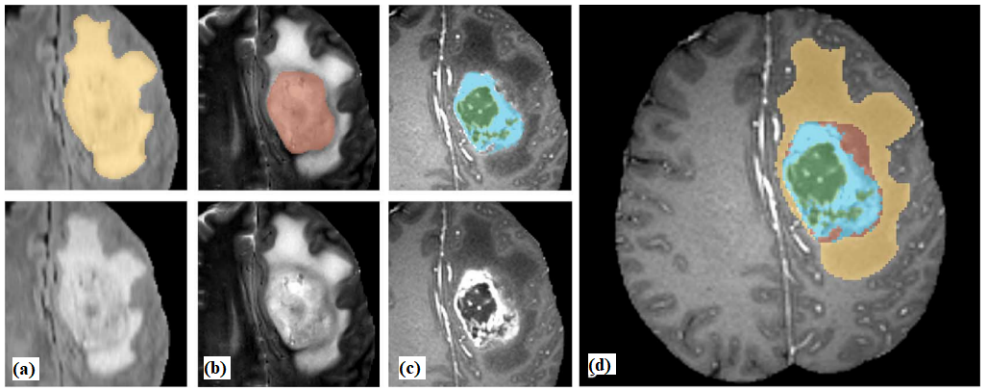


Figure 4.1: Manual annotation of the tumour structures in different MRI images. (a) The whole tumour appear bright in FLAIR image. (b) The tumour core is visible in T2 image. (c) The enhancing tumour structure visible in T1Gd surrounding the necrotic components of the core. (d) All tumour structures combined to generate the final labels : edema (yellow), non-enhancing tumour (brown), necrotic (green), and enhancing tumour (blue) [5].

Fig. 4.2 illustrates example of MRI images from the BraTS training dataset. Each row displays axial view images of FLAIR, T1, T1Gd, T2, and ground truth overlaid on FLAIR of each patient. The ground truth is visually represented by red, yellow, and blue contours, corresponding to the NCR/NET, ED, and ET tumor regions, respectively. The first three rows feature patients with HGG, while the fourth to sixth rows depict patients with LGG. We can see that all tumour regions (NCR/NET, ED, and ET) are presented in all three of the HGG patients. However, ENT region appears very small in the LGG patient shown in the fourth row. In the fifth row, the ET region is completely absent. These examples show that the presence and the size of tumour regions varies in different patients. Additionally, the sixth row of the Fig. 4.2 shows an example of images without any tumour regions presented.

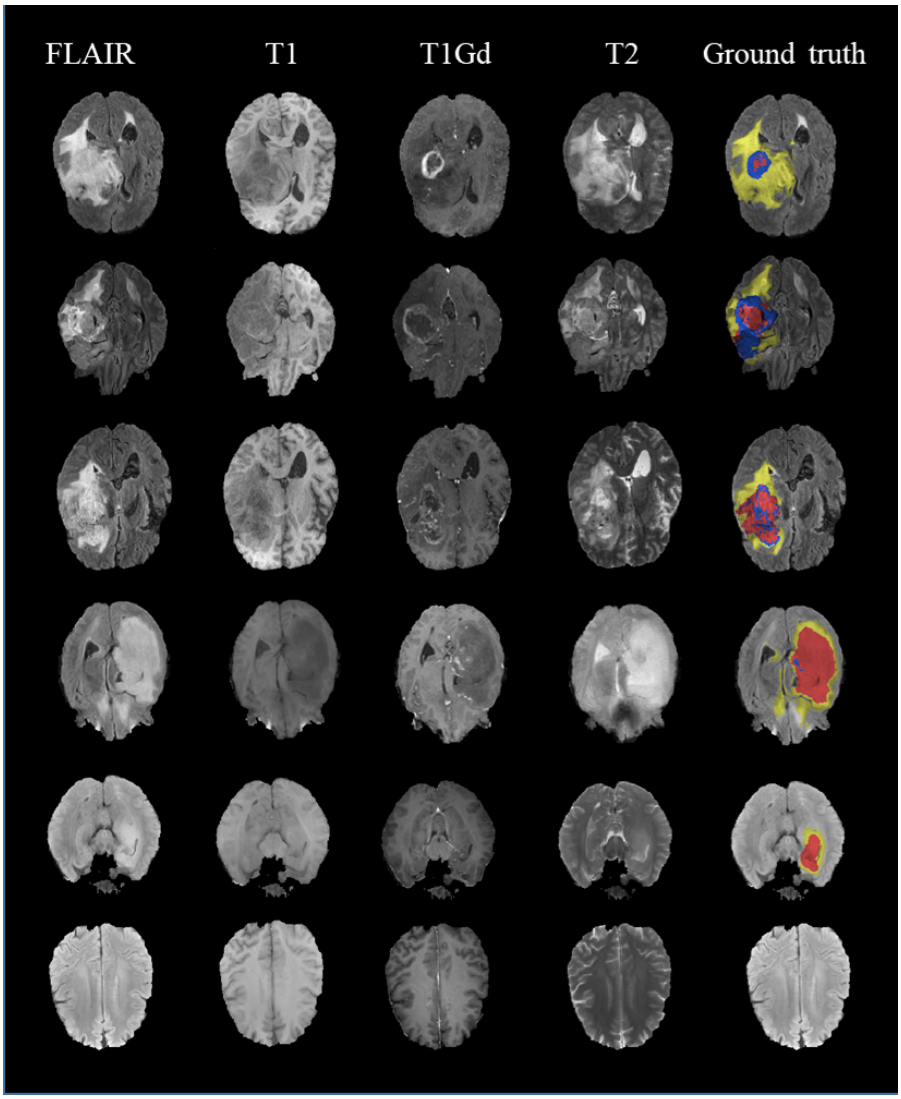


Figure 4.2: Examples of the BraTS training dataset. Each row shows FLAIR, T1, T1Gd, T2, and FLAIR with ground truth labels; NCR/NET (red), ED (yellow), and ET (blue) from HGG patients (rows 1-3) and LGG patients (row 4-6).

For the purpose of algorithm implementation in this thesis the original label 4 for ET was changed to label 3. Hence, we now use label 1 for NCR/NET, label 2 for ED, label 3 for ET, and label 0 for everything else including background. Table 4.2 shows that we can also group the intra-tumour structures into three tumour regions namely the whole tumour (WT) region, the tumour core (TC) region, and the enhancing tumour (TC). The WT region is the complete extent of the tumour including all intra-tumour structures (ED,NCR/NET,ET). The corresponding labels for WT are 1,2,and 3. The TC region includes all tumour structures except ED. The corresponding labels for the TC are 1 and 3. Finally, ET or active tumour region only contains the enhancing core of the tumour with label 3. These three tumour regions better represent in clinical tasks including segmentation task in BraTS challenge [5, 127, 128].

Table 4.2: Summary of tumour region structure and corresponding annotated labels from BraTS dataset.

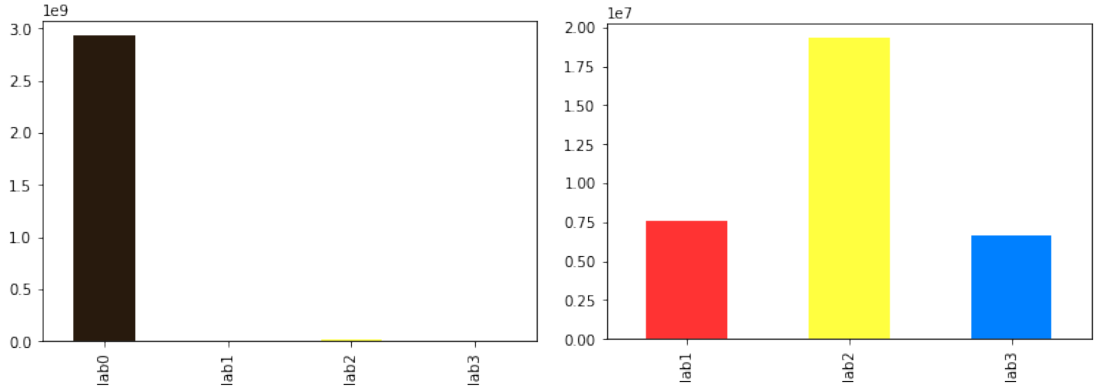
Tumour regions	**Intra-tumour structures**	**labels**
Whole tumour (WT)	ED, NCR/NET, ET	1,2,3
Tumour core (TC)	NCR/NET, ET	1,3
Enhancing tumour (TC)	ET	3

4.3 Labels distribution of BraTS dataset

The field of medical image segmentation faces significant challenges due to class imbalance problem in the data as presented in [132–134]. The ROI target presented in a medical image usually takes small percentage of pixels in the image comparing to the remaining of the image annotated as background. This situation creates an extreme inequality in class distribution, with the background class often prevailing over 90% or even close to 100% of the data. It affects all aspects of the segmentation pipeline including image preprocessing, model architecture, model training strategy, and performance evaluation.

Recall to Fig. 4.2 in the previous section, it can be seen that tumour regions and the size of the tumour vary in different patients. The tumour region can appear very small or completely absent in some patients especially LGG patients. The proportion

of the ROI can also be very small comparing to the non-tumour region and background of the image. Fig. 4.3 shows an example of labels distribution across the BraTS 2020 training dataset. We can see from Fig. 4.3(a) that the majority pixels of the MRI scans are everything else and background of the image as label 0 is the highest number across the dataset. It means that we will train the automatic segmentation algorithms on mostly background of the images. To be able to see the ROI that contain glioma sub-regions, Fig. 4.3(b) shows label distribution without label 0. We can see that, ED regions appear almost 2.5 times more than NCR/NET and ET tumour regions across the whole dataset.



(a) (b)

Figure 4.3: Labels distribution from BraTS 2020 dataset. (a) Labels distribution included label 0 for everything else and background, label 1 for NCR/NET, label 2 for ED, and label 3 for ET. (b) Labels distribution included all labels except 0.

In summary, the BraTS dataset that we used in this thesis represent the extreme inequality in class distribution. The majority of the image pixels are the background class over 90% of the data. This results in the segmentation algorithms of this strong class imbalance can be challenging to train and evaluate. The appropriate loss functions and evaluation metrics that are used to train and evaluate the performance of the segmentation algorithms under this circumstance will be presented in Section 4.4 and 4.5.

4.4 Loss functions of imbalanced segmentation

In medical image analysis including brain tumour segmentation, it is often to encounter the situation where the region of interest (ROI) presents in a small fraction of image pixels compared to the rest of the image as discussed in the previous section. This scenario potentially leads the segmentation algorithm's learning process trapped in the local minimum of the loss optimization, and results in the algorithm that heavily biases its predictions towards the background, and inefficiency detect or entirely miss the foreground (ROI). Various studies have attempted to address this issue through loss functions that prioritize foreground regions during the learning process [96,104,135,136].

Weighted crossed-entropy (WCE) was used to train U-Net for biomedical image segmentation in [96] and can be defined as

$$WCE = -\frac{1}{N}\sum_{n=1}^{N} wr_n \log p_n + (1 - r_n) \log 1 - p_n \tag{4.1}$$

where w is the weight attribute to the foreground class, and r_n, p_n represent the pixel values of the ground truth and the predicted probability map of the n^{th} element, respectively.

Dice Loss (DL) was proposed by [104] for prostate segmentation from 3D volumetric data and has been widely adopted in other medical image segmentation problems. DL was developed based on the Dice Similarity Coefficient and it can be defined as

$$DL = 1 - \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \tag{4.2}$$

where $p_i \in P$ is the predicted binary segmentation volume, $g_i \in G$ is the gound truth, and the sums run over the N voxels.

A variant of DL for binary segmentation was proposed in [135] and can be define as

$$DL_2 = 1 - \frac{\sum_{n=1}^N p_n r_n + \varepsilon}{\sum_{n=1}^N p_n + r_n + \varepsilon} - \frac{\sum_{n=1}^N (1 - p_n)(1 - r_n) + \varepsilon}{\sum_{n=1}^N 2 - p_n - r_n + \varepsilon} \tag{4.3}$$

where r_n, p_n represent the pixel values of the ground truth and the predicted probability map of the n^{th} element, respectively. The ε term is a small positive number used to ensure loss function stability in the cases where ground truth and prediction do not contain the region of interest (ROI).

4.5 Evaluation metrics

Evaluation metrics of semantic segmentation measures the classification accuracy as well as the localization correctness. To improve the algorithm performance, the aims is to score the similarity between the predicted segmentation and the annotated segmentation or the ground truth. Fig. 4.4 illustrates an example of overlapping region of the actual tumour T_1 (blue contour) and predicted tumour region P_1 (red contour). The T_0 is the actual non-tumour region and P_0 is the predicted to be normal brain tissues. These regions will be used to calculate several evaluation metrics.

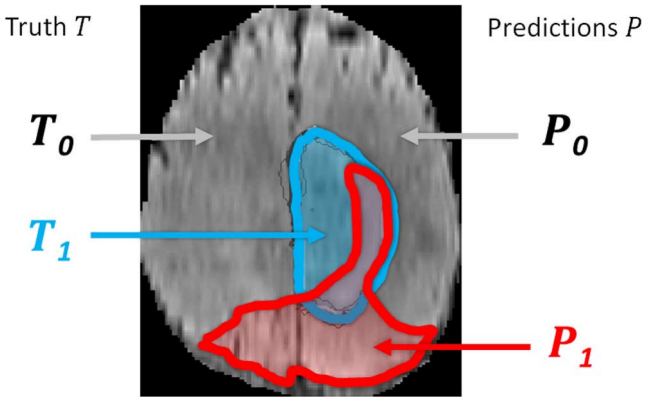
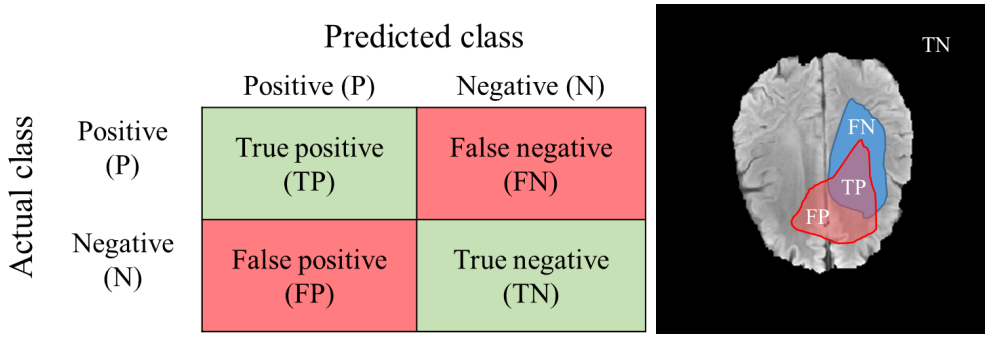


Figure 4.4: An illustration of regions used for calculating evaluation metrics [5].

Fig. 4.5 illustrates another example of the relationship between a confusion matrix and the overlapping areas of the ground truth (blue contour) and predicted segmentation (red contour) of the tumour. The confusion matrix is a performance measurement table that contains four distinct combinations of the actual class or ground truth and the predicted class. In Fig. 4.5(a), the actual class is divided into positive and negative

groups. The positive class represents the ground truth of the annotated ROI, while the negative represent the ground truth of the annotated background. True positive (TP) denotes cases where the algorithm correctly predicted a positive class label, whereas false positive (FP) denotes the cases where the algorithm incorrectly predicted a positive class when the actual class was negative. True negative (TN) represent the cases where the algorithm correctly predicted a negative class label, whiles false negative (FN) represent the cases where the algorithm incorrectly predicted a negative class label when the actual class was positive. Fig. 4.5(b) illustrates the an example of predicted tumour segmentation (red contour) and the corresponding ground truth (blue contour). TP represents the overlapping area between the predicted segmentation and the ground truth, indicating that the algorithm correctly predicted the tumour region. FP refers to the area where the algorithm incorrectly predicted normal brain tissues as tumour region. FN represents the actual tumour region that the algorithm failed to predict, and TN represents the area correctly predicted as non-tumour region and background.



(a) (b)

Figure 4.5: (a) Confusion matrix. (b) The overlapping of the ground truth (blue contour) and predicted segmentation (red contour).

The following metrics can be computed from the confusion matrix:

Accuracy or Rand index is the number of correct predictions including correct positive and negative predictions divided by the total number of predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{4.4}$$

Recall or sensitivity is known as the positive rate. Recall is the number of true positive results divided by the number of all samples that should have been identified as positive. It is a performance metric that assesses the algorithm's capability to accurately detect positive instances among all the actual positive instances in the dataset. In other words, it measures how well the algorithm identifies the true positives.

$$Recall = \frac{TP}{TP + FN} \tag{4.5}$$

Precision measures the correctness of positive predictions and indicates how many of the predicted positive instances are actually true positive. It is calculated as the ratio of the numbers of true positive to the sum of the number of true positive and false positive.

$$Precision = \frac{TP}{TP + FP} \tag{4.6}$$

Specificity is also known as true negative rate. It measures the accuracy of the negative prediction made by the algorithm. It quantifies the algorithm's ability to avoid false positive. It is calculated as the ratio of the number of true negative to the sum of the number of true negative and false positive.

$$Specificity = \frac{TN}{TN + FP} \tag{4.7}$$

Dice Similarity Coefficient (DSC) [137], also known as F1 Score or Sørensen index, is the metrics that calculated from harmonic mean of precision and recall and it

can be define as:

$$DSC = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \tag{4.8}$$

$$= \frac{2}{\frac{1}{\frac{TP}{TP+FP}} + \frac{1}{\frac{TP}{TP+FN}}} \tag{4.9}$$

$$= \frac{2TP}{2TP + FP + FN} \tag{4.10}$$

DSC measure the overlapping between predicted segmentation tumour regions and the ground truth. From the ground truth T_1 and prediction P_1 regions defined in Fig. 4.4, we can also consider equation 4.10 as:

$$DSC = \frac{2 \times |T_1 \cap P_1|}{|T_1| + |P_1|} \tag{4.11}$$

where $\cap$ is the intersection operator and $|\cdot|$ is the size of the set (i.e., number of image pixels belonging to it). The metric range from 0-1 with 0 means that there is no overlapping between the predicted segmentation and the ground truth, while 1 means the predicted segmentation perfectly overlap with the ground truth.

Intersection Over Union (IoU) or Jaccard Index [138] is another metric that measure the overlapping area between the predicted segmentation and the ground truth and can be defined as:

$$IoU = \frac{TP}{TP + FP + FN} \tag{4.12}$$

We can also consider equation 4.12 using ground truth T_1 and prediction P_1 regions defined in Fig. 4.4 as:

$$IoU = \frac{|T_1 \cap P_1|}{|T_1 \cup P_1|} \tag{4.13}$$

where $\cap, \cup$ are the intersection and union operators, and $|\cdot|$ is the size of the set. The difference between IoU and DSC is that IoU penalises under- and over-segmentation more than DSC [132].

In a medical image, where the significant class imbalance between the background and the ROI at the ratio of 9:1, the number of correct predictions for the background class is significantly higher compared to the ROI class. Using metrics that equally weigh true positives and true negatives can produce misleading results. Hence, Accuracy and Specificity that included true negative will produce excessively high score and should therefore be avoided. Metrics that focus solely on true positive prediction without considering true negatives, such as DSC and IoU, are recommended and widely used in the field of medical image segmentation.

4.6 Summary

The real MRI data of glioma patients from the BraTS dataset from 2018-2021 were presented in this chapter. The dataset included multimodal MRI scans (T1, T1Gd, T2, and FLAIR) from both HGG and LGG patients, and the ground truth which are manually segmented by the expert radiologists. The BraTS dataset were published as part of the annually brain tumour segmentation challenge. Examples of the tumour regions and the corresponding labels from the BraTS training dataset were illustrated. The manual segmentation protocol used for the labels annotation of the BraTS dataset was presented. The image data showed that the appearance of tumour structures and the size of the tumour vary in each patient. The data also presented an extreme inequality in labels distribution due to the ROI usually take small percentage of the image compared to the annotated background. The techniques used to mitigate the bias of the segmentation algorithms such as loss functions that prioritize the foreground segmentation during the training process, and the appropriate evaluation metrics for medical image segmentation were discussed.

Further in chapter 5, the development of automatic brain tumour segmentation algorithms using BraTS 2018 dataset will be presented. BraTS 2019 and 2020 dataset

will be used as training and testing data for multi-class segmentation algorithm in chapter 6. Finally, 3D brain tumour region segmentation using BraTS 2021 will be presented in chapter 7. The loss functions and evaluation metrics introduced in this chapter will be used to train and measure the performance of the developed algorithms.

Chapter 5

Automatic brain tumour segmentation using modified U-Net

5.1 Introduction

Previously, treatment options for brain tumour patients were presented in Section 2.2 and the manually segmentation of the ROI in order to deliver the appropriate treatment was presented in Section 3.2. To eliminate the subjectiveness of and make tumour segmentation effective for large-scale studies, there is a need for automated segmentation. Classical approaches for automatic brain tumour segmentation were explored in Section 3.3 but deep learning based models with convolutional neural networks have proved superior in learning essential features and segmenting of the tumour in Section 3.4.

In this chapter, the use of deep neural networks for 2D automatic brain tumour regions segmentation through a modified version of U-Net [96] is proposed. Because the tumour appearance varies from patient to patient, the brain tumour segmentation represent a particular challenging segmentation problem and the underestimation of the tumour can lead to an inefficient treatment planning including the RT. The pro-

posed method introduced the two-pathways local and global feature extractions inspired by [54] to give the model larger context information of region around the image pixel for the segmentation prediction.

The remainder of this chapter is organised as follows. The framework of deep learning based segmentation using 2D MRI slices is described in section 5.2. The details of the proposed method is given in section 5.3 and 5.4. The proposed method performance using evaluation metrics introduced in chapter 4, visualisation of the results and conclusion will be in section 5.5 and 5.6.

5.2 Deep learning based segmentation framework

The process of deep learning based brain tumour segmentation involves model training and model testing as illustrated in Fig. 5.1. The dataset is initially split into training and testing data. The test data is kept aside as an unknown data to evaluate the final model performance. Data normalization is performed for both training and testing data to brings pixel values to a standardized scale and ensure fair comparisons between different image modalities [139]. Subsequently, the normalized training data can be further split into training and validation data. The training data is used as an input for the DL model to extracts essentials high level features from the input and predicts each image pixel into a curtain class of non-tumour and tumour regions. The prediction output is compared with the ground truth segmentation mask or the label. The loss is then calculated using a selected loss function. The process is iterated and the model learns to minimise the loss until the loss is under the desired threshold or the iteration reached the setting number of training epoch. During this training phase, the validation data is used to iteratively evaluate the model and, the weights from the epoch that gives the best performance will be saved to the model. We then observe the model performance from the trained model output using the testing dataset. To visualise the segmentation mask, post-processing is applied to generate the predicted segmentation mask of the brain tumour regions and non-tumour region background.

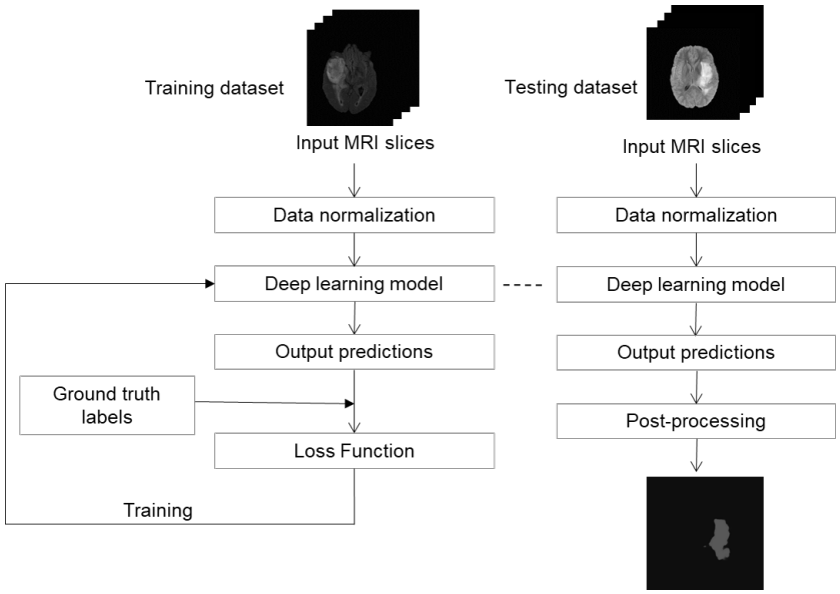


Figure 5.1: Deep learning based segmentation framework from 2D MRI data.

5.3 Modified U-Net

A novel CNN architecture which is based on the U-Net [96] that performs automatic segmentation of the brain tumour regions from 2D MRI slices is proposed. Fig. 5.2 illustrates that the modified U-Net has similar structure to the U-Net comprising down-sampling and up-sampling paths. However, the first block of the conventional U-Net is replaced with two-pathway feature extraction. Two convolutional layers with kernel size of 3x3 and 9x9 are used followed by a ReLU to capture low-level features from 2D axial multimodal MRI input. The multimodal MRI; T1, T2, T1GD, and FLAIR are concatenated to form 4 channels of the input image. We refer to these features obtained from 3x3 and 9x9 kernels as *local* and *global* feature maps. Global and local feature maps are subsequently concatenated. A 2x2 max-pooling with stride of 2 is used to halve the image size. A 3x3 convolutional layer followed by ReLU and 2x2 max-pooling with stride of 2 are repeated until image resolution deceases from 240x240 to 15x15 and feature maps increases from 4 to 512. Two repeated 3x3 convolutional layers followed by a ReLU are used in the fifth block of down-sampling path to double

feature maps to 1024.

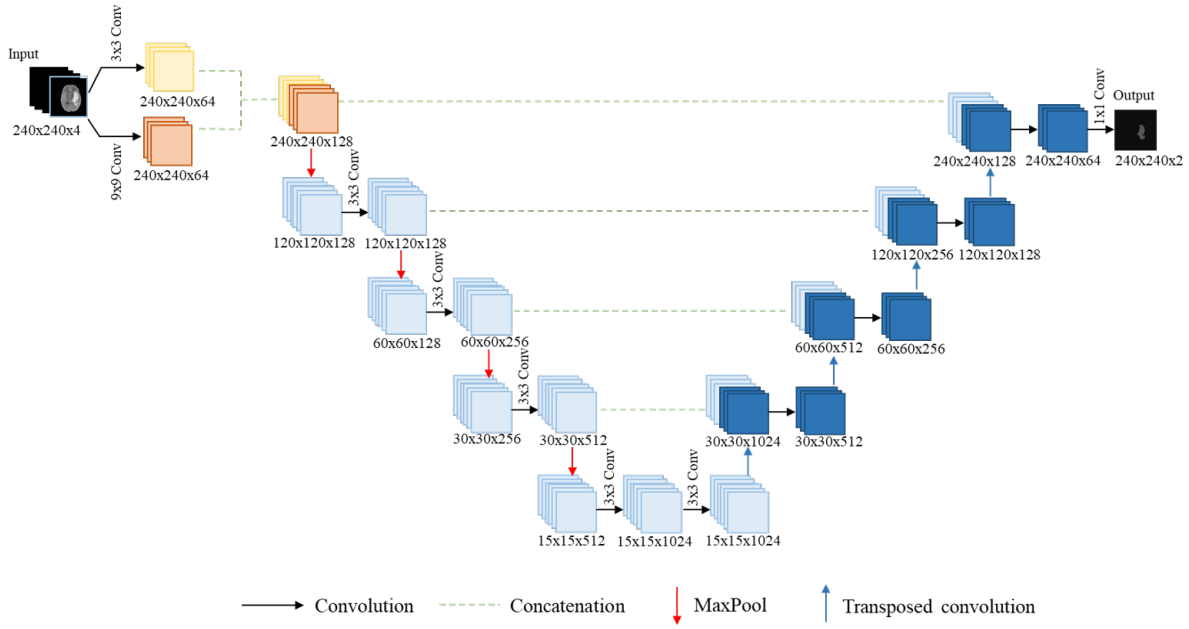


Figure 5.2: Modified U-Net network architecture.

For the up-sampling path, a 3x3 transposed convolution with stride of 2 is used to double image size in both dimensions and to halve the feature maps. The result is then concatenated with the corresponding feature maps from the same level in the downsampling path providing the higher resolution feature map the localization context and allowed more precise segmentation. A 3x3 convolutional layer followed by a ReLU is used to halve the concatenation feature maps. The process is repeated until the image resolution increases from 15x15 to original 240x240 with 64 feature maps. Finally, a 1x1 convolutional [140] layer of length 1x1x64 with Sigmoid function is employed to produce a 240x240x2 output. One for the foreground and one for the background segmentation.

The main novelty of the modified U-Net comes from the local and global feature extractions inspired by TwoPathCNN [54] which aim to capture low-level features from multimodal MRI input and provides the network with contextual information of the essential features. Additional details of modified U-Net and network implementation is presented in the next section.

5.4 Details of modified U-Net

5.4.1 Local and global feature extraction

Fig. 5.3 illustrate the two-pathways local and global feature extraction. It is made of a concatenation of smaller 3x3 kernel and a larger 9x9 kernal. This two-pathways architecture is inspired by TwoPathCNN [54] which aim to provide the segmentation model with two aspects of the input including data visual details of the region around that pixel and its larger context including the localization information.

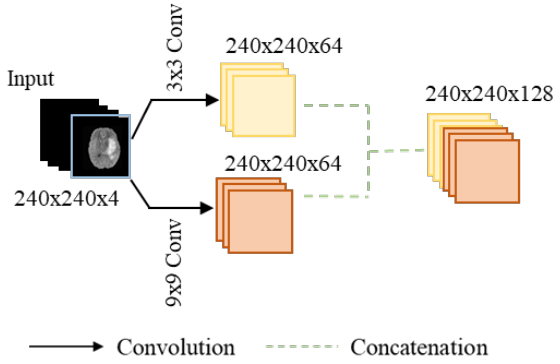
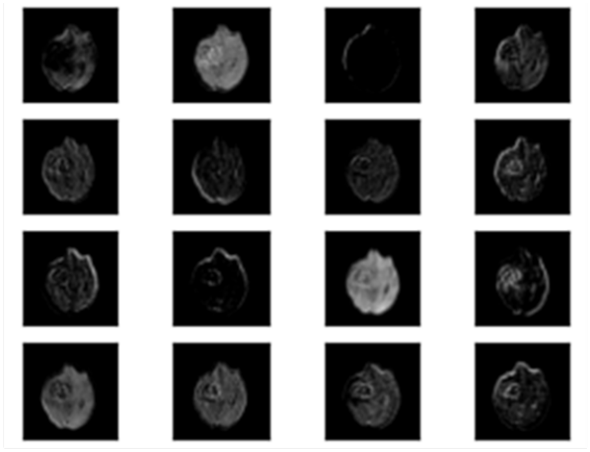
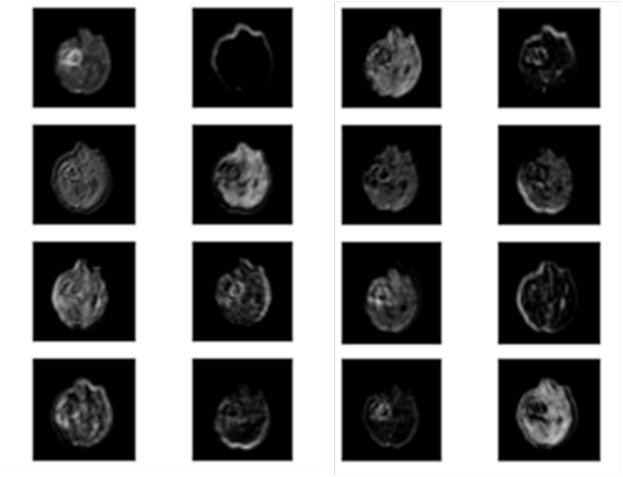


Figure 5.3: Two-pathways feature extraction.

Fig. 5.4 shows an example of feature maps obtained from the two-pathways feature extraction in the first layer of the modified U-Net. Local features from 3x3 kernels and global features from 9x9 kernels are illustrated in Fig. 5.4(a) and Fig. 5.4(b), respectively. These features were extracted from the same MRI slice. While both local and global features represent the brain's structure well, the edges of the brain are more highlighted in the local features, and the fine details of the brain are better preserved in the global features. More example of feature maps obtained from 3x3 and 9x9 kernels can be found in the Appendix A.



(a)



(b)

Figure 5.4: Feature maps obtained from the first layer of modified U-Net using (a) 3x3 kernels, and (b) 9x9 kernels.

5.4.2 Up-sampling: Transposed convolution

Transposed convolution or fractionally strided convolutions [141] is used to restore the original image size of the initial feature. The size of output feature map o from a convolution with kernel k on a given input i with padding p, and stride s is may be written as:

$$o = \frac{i + 2p - k}{s} + 1. \tag{5.1}$$

The convolution defined above has an associate transposed convolution kernel $k' = k$ where $\widetilde{i'}$ is the size of the stretched input obtain by adding $s - 1$ zeros between each input i' with $s' = 1$ and $p' = k - p - 1$. The transposed convolution output size can be written as:

$$o' = s(i' - 1) + k - 2p. \tag{5.2}$$

Fig. 5.5 shows an example of the transpose of 3x3 convolution over a 5x5 input with 1x1 zeros padding boarder using a stride of 2 (i.e., $i = 5, k = 3, s = 2$, and $p = 1$). It is equivalent to convolving a 3x3 kernel over a 3x3 input with 1 zero inserted between inputs, padded with 1 zeros boarder using the stride of 1 (i.e., $i' = 3, \widetilde{i'} = 5, k' = k = 3, s' = 1$, and $p = 1$). In the proposed modified U-Net, we used the 3x3 transposed convolution with zeros padding and a stride of 2 to allowing the model to learn the weights for expanding the resolution of the feature maps to original image size.

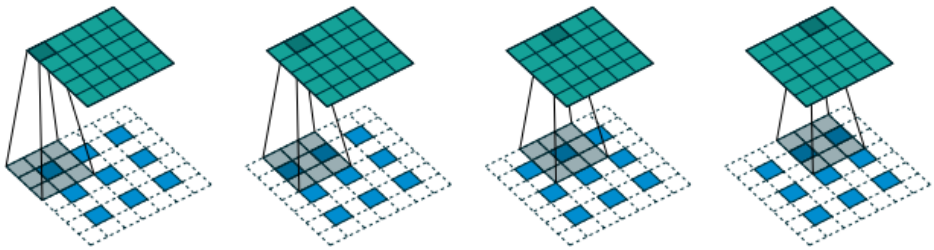


Figure 5.5: The transposed convolution of a 3x3 kernel over a 3x3 input with 1 zero inserted between inputs, padded with 1 zeros padding and using a stride of 1 [141].

5.4.3 Network implementation

An Adam optimizer [142] was used with a learning rate of 0.00001, values of $\beta_1 = 0.9$, $\beta_2 = 0.999$, batch size of 16, and an epoch size of 50. All weights were initialized using He initialization [143] with mean of 0 and standard deviation of $\sqrt{2/n}$, where n is the number of input units in the weight tensor, and all biases were initialized as 0. Dice Loss (DL_2) [104, 135] was used as a loss function and can be defined as

$$DL_2 = 1 - \frac{\sum_{n=1}^{N} p_n r_n + \varepsilon}{\sum_{n=1}^{N} p_n + r_n + \varepsilon} - \frac{\sum_{n=1}^{N} (1 - p_n)(1 - r_n) + \varepsilon}{\sum_{n=1}^{N} 2 - p_n - r_n + \varepsilon} \tag{5.3}$$

where r_n, p_n represent the pixel values of the ground truth and the predicted probability map of the n^{th} element, respectively. The ε term is a small positive number used to ensure loss function stability in the cases where ground truth and prediction do not contain the region of interest (ROI). The experiment was implemented using Tensorflow and Tensorlayer library on PC equipped with NVIDIA GeForce GTX 1070 GPU, an Intel Core i5-8400 CPU 2.80 GHz processor, 16 GB of RAM.

5.5 Experimental results

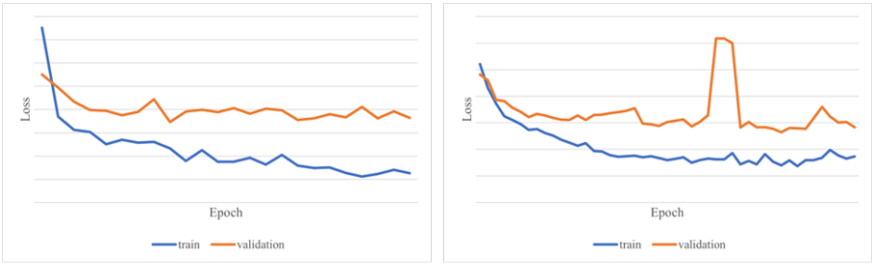
We evaluated the proposed modified U-Net for the segmentation of five tumour regions: whole tumour (WT), tumour core (TC), edema (ED), necrotic and non-enhancing tumour (NCR/NET), and enhancing tumour (ET). WT is the union of all annotated tumour labels (ED+NCR/NET+ET), TC is the union of NCR/NET and ET. The proposed model was trained separately for each tumour region. Multimodal MRI scans (T1, T1Gd, T2, and FLAIR) of 210 HGG and 75 LGG patients from BraTS 2018 was split into the ratio of 0.6:0.2:0.2 for training, validation, and testing data. Data normalization was performed for all dataset. First, the mean and standard deviation of each MRI modality across the dataset are calculated. Subsequently, multimodal MRI of each patient is subtracted with the mean, and divided by the standard deviation of the corresponding modality.

Mean DSC and IOU obtained from the modified U-Net are reported in Table 5.1. The proposed method achieved mean DSC over 0.8 and mean IoU over 0.7 in all tumour regions segmentation from training data. It can be seen that the proposed method achieved the best performance in WT segmentation with 0.81 and 0.83 DSC from validation and testing data, respectively. The segmentation results decreased to over 0.6 DSC for TC, ED, and ET segmentation from validation and testing data, and significantly decreased to over 0.4 DSC for NCR/NET segmentation compared to the results of training data. Example of training curves from WT and ET segmentation are illustrated in Fig. 5.6(a), and Fig. 5.6(b), respectively. It can be observed that the proposed method experienced overfitting [144], especially in ET segmentation where the model learn to fit the training data. The model is not able to generalise to the validation data which leads to the increasing of loss while the training loss is still decreasing as shown in Fig. 5.6(b). To mitigate the overfitting, techniques such as weights regularization, Dropout, and early stopping can be employed [145]. Additionally, the proposed method used multimodal MRI scans from HGG and LGG together and the label distribution is highly imbalanced as discussed in chapter 4. The appearance of NCR/NET and ET regions are small in fraction compared to the annotated background of the image. These reasons combined could affect the performance of the NCR/NET segmentation.

Table 5.1: Segmentation results of the proposed method using BraTS 2018 dataset.

Tumour regions	**DSC**			**IoU**		
	Train	**Val**	**Test**	**Train**	**Val**	**Test**
WT	0.92	0.81	0.83	0.89	0.69	0.69
TC	0.89	0.61	0.62	0.88	0.45	0.47
NCR/NET	0.81	0.40	0.45	0.71	0.29	0.31
ED	0.88	0.66	0.69	0.82	0.51	0.52
ET	0.83	0.68	0.70	0.72	0.54	0.56

Table 5.2 shows the comparative segmentation results obtained from the modified U-Net and the original U-Net using the testing data. The proposed method and the original U-Net achieved over 0.8 DSC for the WT segmentation. However, the proposed method outperformed the U-Net in ET and NCR/NET regions segmentation indicate



(a) (b)

Figure 5.6: Learning curve from (a) WT segmentation, and (b) ET segmentation.

Table 5.2: Brain tumour regions segmentation performance comparison.

Method	**DSC**			
	WT	**ED**	**ET**	**NCR/NET**
U-Net	0.82	0.70	0.62	0.31
Proposed	**0.83**	0.69	**0.70**	**0.45**

that the two-pathways feature extraction introduced in the down-sampling of the proposed method has improved the segmentation performance.

Fig. 5.7 illustrate the visualisation of the WT segmentation (red contours) obtained from the proposed method compared to the corresponding ground truth (blue contours) of patient ID Brats18_TCIA08_469_1. Fig. 5.7(a) shows the predicted segmentation contours and ground truth over the original FLAIR images from slice 40, 70, 90, and 110, respectively. The proposed method successfully segmented the WT in slice 70, 90, and 110. Furthermore, the proposed method was able to identify the slice without ROI (no segmentation predicted) in slice 40, which reduces a chance of an overestimation of the tumour. Fig 5.7(b) and Fig. 5.7(c) illustrate the comparison between the stacked 2D contours of the ground truth (blue) and the predicted WT segmentation (red) from all 155 slices of this patient.

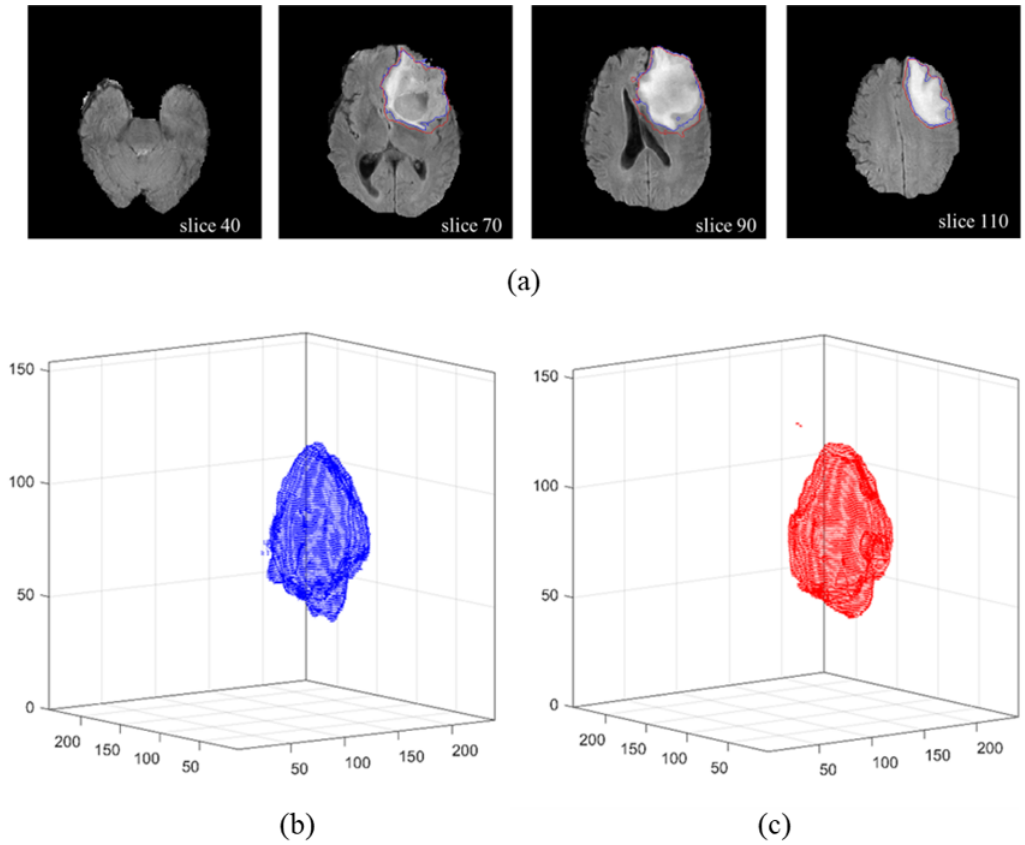


Figure 5.7: Visualisation of the predicted WT segmentation (red) and corresponding ground truth (blue) contours from patient ID Brats18_TCIA08_469_1, (a) tumour contours obtained from the modified U-Net on 2D axial FLAIR images, (b) stacked 2D contours of the ground truth, and (c) stacked 2D contours of the predicted segmentation.

Fig. 5.8 give more examples of different tumour regions segmentation obtained from the proposed method. Each row of the image displays original T1Gd images with corresponding ground truth, and the predicted segmentation masks of different tumour regions from various patients. Fig. 5.8(a)-(c) feature the segmentation of TC region which included all tumour structure except the ED. Fig. 5.8(d)-(l) visualise each intra-structure of the tumour. ET is shown in (d)-(f), ED is shown in (g)-(i), and NCR/NET regions is shown in (j)-(l), respectively. From the first column, TC (a) is not easy to differentiate from ED comparing to ED region that appears darker than normal brain tissue as shown in (g). ET appears as very bright region and it is visible in (d), while NCR/ENT is the dark solid inside the bright ET ring as shown in (i).

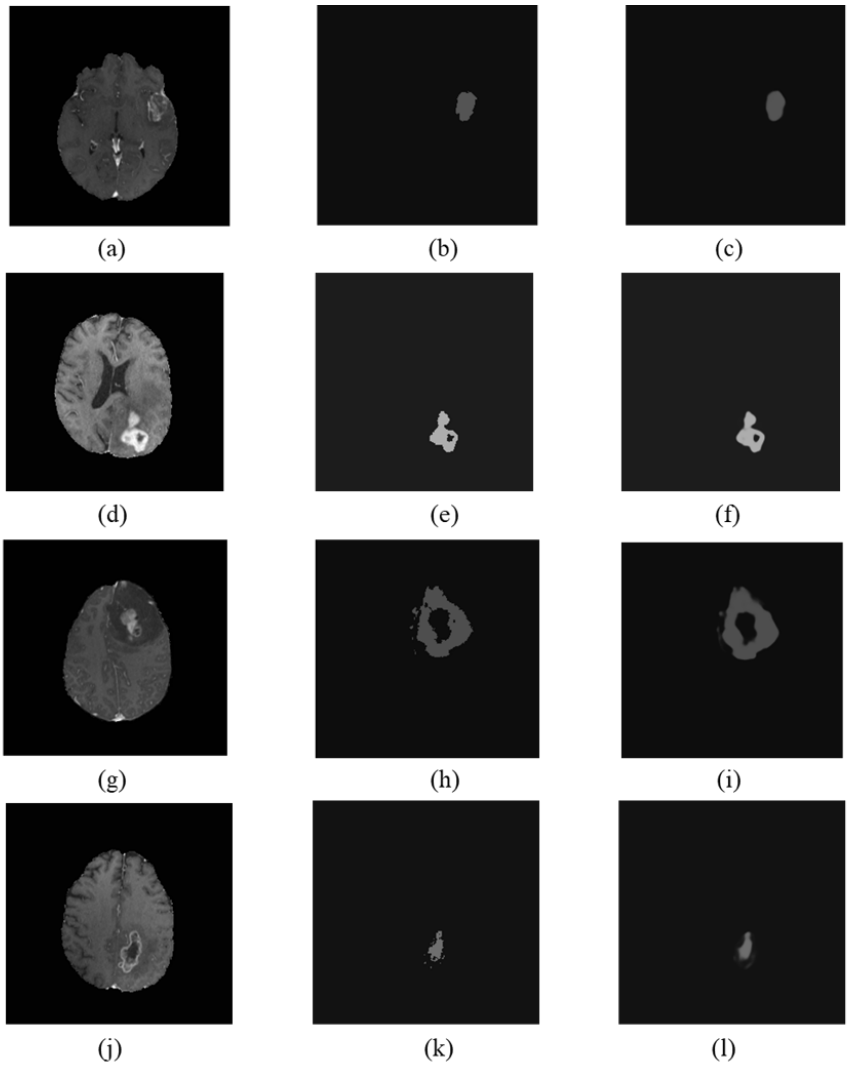


Figure 5.8: Visualisation of original T1Gd image, ground truth, and predicted segmentation of brain tumour regions. The images are from various patient of BraTS 2018 dataset.

5.6 Summary

This chapter presented a novel deep learning based automatic brain tumour regions segmentation using 2D multimodal MRI slices from BraTS 2018 dataset. The proposed method is based on U-Net architecture that had proved to work well in biomedical segmentation. We used two-pathway feature extraction with 3x3 and 9x9 kernels to extract local and global feature directly from the input. These features provide the model with two aspects of the visual information from the input data. Two-pathways feature extraction provided the details around the image pixel from local pathway and larger context including localization information from global pathway. The low-level local feature maps obtained from the two-pathways preserved the edges of the brain structures while global feature maintain the fine details of the brain.

The proposed method was trained and evaluated for five tumour regions including WT, TC, NCR/NET, ED, and ET. The training phase required a GPU, and took approximately 12 hours for each tumour region. However, the trained model was able to segment all 155 slices of each patient in under a minute which is significantly less-time consuming compared to the manual segmentation. The segmentation results obtained from the proposed method have achieved over 0.8 DSC for WT segmentation and over 0.6 DSC for TC, ED, and ET segmentation in both validation and testing data. Furthermore, it was able to identify the MRI slices without ROI, reducing the chance of an overestimation. Additionally, it has improved the segmentation performance when compared to the original U-Net, especially the cases of smaller tumour regions including ET and NCR/NET. This improvement indicates that the use of the two-pathway feature extraction benefits the model's learning. However, the proposed method experienced overfitting and still faced the difficulty of NCR/NET region segmentation, which also a challenge in manual segmentation due to inhomogeneity of NCR/NET compared to WT and TC regions.

To tackle the drawbacks experienced in this chapter, an improved two-pathways feature extraction and deep neural networks for multi-class segmentation of the brain tumour will be presented in the next chapter.

Chapter 6

TwoPath U-Net for multi-class 2D segmentation

6.1 Introduction

In the previous chapter, new deep learning based segmentation framework using 2D slices from multimodal MRI data and the modified U-Net for tumour regions segmentation were presented. It was shown that the new method faced difficulty in segmentation of NCR/NET and ET which are smaller tumour regions comparing to WT TC and ED. The segmentation of each tumour region obtained from training the model was carried out separately. In this chapter, we present a new multi-class deep CNN called TwoPath U-Net that provides segmentation of all tumour regions. The proposed method keeps the same U-Net structure as in chapter 5. However, it now comprises a two-pathways local and global feature extraction in all down-sampling blocks.

The remainder of this chapter is organised as follows. The network architecture of the proposed TwoPath U-Net and its details are presented in Section 6.2. The experiments set up including data preparation, data augmentation, parameters setting are discussed in Section 6.3. Further, experimental results including the BraTS 2019 dataset and BraTS 2020 challenge participation are reported and discussed in Section 6.4 and 6.5. Finally, Section 6.6 gives the conclusion of this chapter.

6.2 TwoPath U-Net for multi-class segmentation

A novel CNN architecture which is an extension version of the modified U-Net described in chapter 5 is presented. This is referred to as the TwoPath U-Net, and is illustrated in Fig. 6.1. It comprises three down-sampling blocks, a bottle-neck layer, and three up-sampling blocks. The proposed method uses 2D slices from multimodal MRI scans in the same manner of previous modified U-Net. In every down-sampling block, cascaded two-pathways feature extraction is used to provide local and global feature extractions. The cascaded two-pathway consists of a 3x3 convolution concatenate with 12x12 follows by 9x9 convolution layers with ReLU. These local and global feature extraction paths capture low-level essential features and provide multi-perspective of the essential context from the input to the model. Then, 2x2 max-pooling operator is performed to halve the size of the feature maps dimension and these features maps become the input of the next down-sampling block. The feature extractions are repeated until the feature maps increase from 4 to 512 dense features. Subsequently, two repeated 3x3 convolution layers followed by ReLU are performed. Dropout layer [145] is used after the second convolution in the bottleneck block of the network architecture to reduce overfitting.

For up-sampling blocks, a 3x3 transposed convolution with stride of 2 is used to double images resolution in both dimensions and to halve the feature maps, followed by 3x3 convolution with ReLU. The result of the up-sampling is then concatenated with the corresponding features from the down-sampling side of the same block level as shown in Fig. 6.1. These concatenations or skip connections provide the higher resolution features with local and global localization contextual information to the up-sampling process. The process repeated until the feature resolution increase back to original resolution with 64 feature maps. Finally, 1x1 convolution layer of length 1x1x64 with softmax function is employed to produce the output of segmentation mask with four classes prediction of ED, ET, NCR/NET and background. The details of TwoPath U-Net and its experiments are presented in the next section.

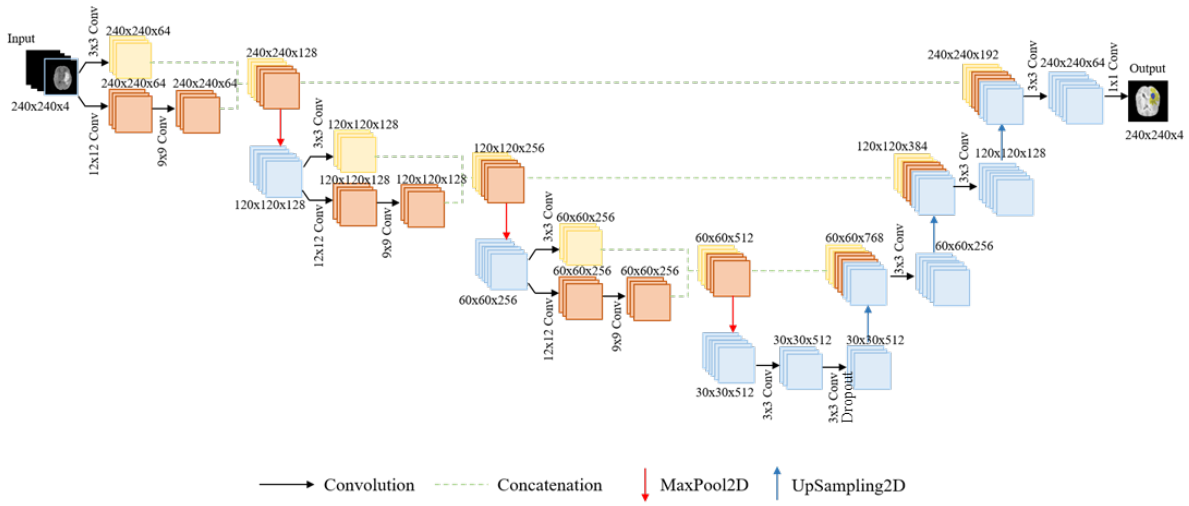


Figure 6.1: TwoPath U-Net network architecture for multiclass segmentation.

6.2.1 Cascaded two-pathway feature extraction

The novelty of the proposed TwoPath U-Net is the cascaded two-pathways of local and global feature extraction. However, instead of using 9x9 convolution as global feature extraction like modified U-Net in chapter 5, we used consecutive 12x12 and 9x9 convolutions as shown in Fig. 6.2 to enlarge the receptive filed and able to extract global feature from input using more neighbour pixels. This two-pathways are used in every down-sampling block. The concatenated feature map obtained from the two-pathways become input of the next down-sampling block, and create cascaded feature extraction architecture.

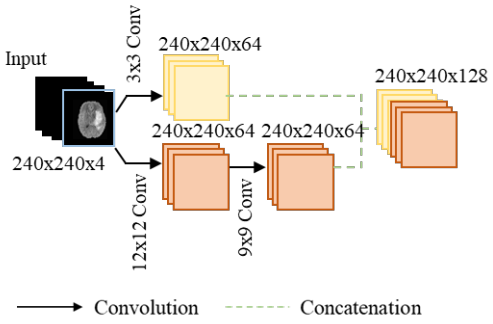


Figure 6.2: Cascaded local and global pathways.

Fig. 6.3 illustrate the feature maps obtained from the cascaded two-pathway feature extraction architecture. Examples of feature maps obtained from the first to the third down-sampling block of the TwoPath U-Net are displayed in Fig. 6.3(a)-(c), respectively. Each column of Fig. 6.3(a)-(c) represents the feature maps obtained from 3x3, 9x9, and 12x12 kernels. It can be observed that the features extracted from the first down-sampling block preserve the structures of the brain as shown in Fig. 6.3(a). Feature maps obtained from the second and third down-sampling blocks present a more abstract representation of the brain and the tumour as shown in Fig. 6.3(b) and Fig. 6.3(c). Additional examples of feature maps obtained from two-pathway feature extraction are featured in Appendix A.

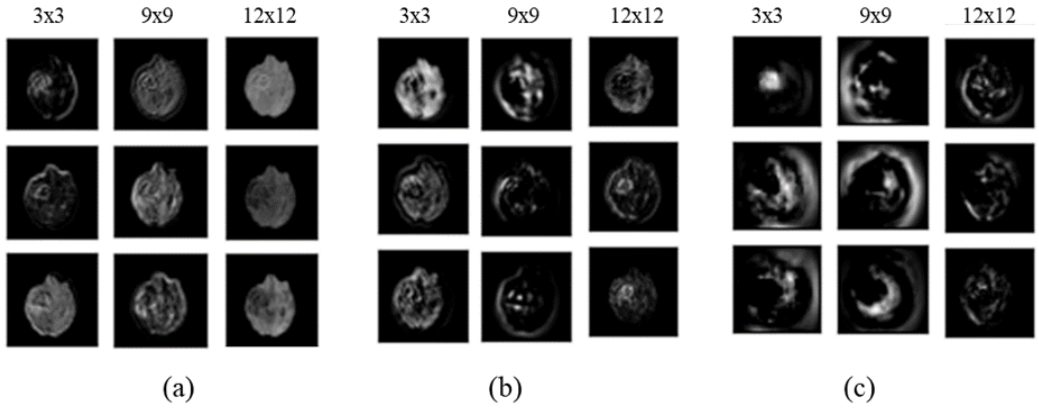


Figure 6.3: Feature maps obtained from cascaded TwoPath feature extraction using 3x3, 9x9, and 12x12 kernels. From (a)-(c) shows feature maps obtained from the first to third down-sampling block of the proposed TwoPath U-Net.

6.3 Experiments

6.3.1 Data preparation

For model training, we implemented the proposed TwoPath U-Net using BraTS 2020 training dataset. Details of the number of HGG and LGG patients available in BraTS 2020 dataset and it corresponding labels were previously presented in chapter 4. The dataset was split into the ratio of 0.8:0.2 for training and validation data. BraTS 2019

validation dataset and BraTS 2020 testing dataset were used as unknown data for model evaluation further in Section 6.4 and 6.5. Note that both BraTS 2019 validation dataset and BraTS 2020 testing dataset do not contain the ground truth segmentation. To evaluate the segmentation results, the predicted segmentations were submitted to CBICA image processing portal [130] to obtain the segmentation results.

Data normalization was performed to all dataset prior model training and testing. First, the mean and standard deviation of each MRI modality across the dataset are calculated. Subsequently, multimodal MRI of each patient is subtracted with the mean, and divided by the standard deviation of the corresponding modality.

Instead of solely utilizing 2D axial view slices, we incorporated multi-view 2D MRI slices to increase the training samples. From left to right, Fig. 6.4 illustrated an example of the original axial view (top to bottom), frontal view (front to back), and sagittal view (left to right) of the image. The corresponding ground truth for each view was also provided. Additionally, Fig. 6.4 shows examples of each class of the ground truth with red, yellow, and blue contours representing NCR/NET, ED, and ET tumour regions, respectively.

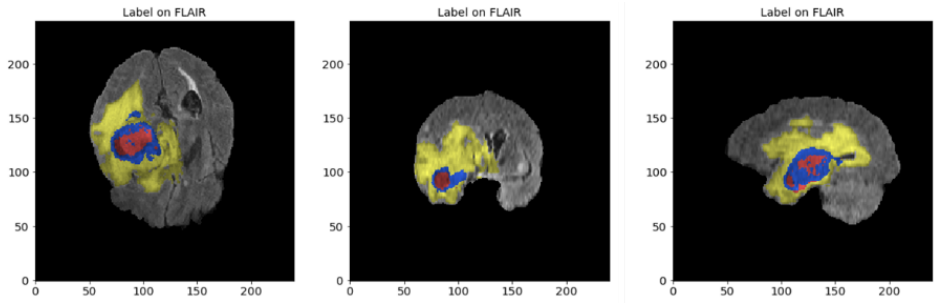


Figure 6.4: Examples of 2D axial view (left), frontal view (middle), and sagittal view (right) from training data with overlaid ground truth labels; NCR/NET (red), ED (yellow), and ET (blue).

We performed experiments using a full-size image and smaller patches for model training. For full-size image training, input image of the size 240x240x4, where 4 is the multimodal MRI (T1, T1Gd, T2, and FLAIR), were used as input of the proposed networks as shown in Fig. 6.1. At the end of the down-sampling path, the resolution of

the input data were halved in both dimension from 240x240 to 30x30 and the feature maps were increase from 4 to 512. For up-sampling path, the dense feature maps resolution from bottleneck of the network were doubled in each block back to original image resolution and final output was the prediction of 4 classes including 3 tumour regions and background .

Since the majority of the image labels are the background and cause the imbalance in label distribution as discussed in chapter 4. We used smaller patches as the input to eliminate the background of the images that do not contain the brain and the tumour extents as shown in Fig. 6.5. The original MRI images were cropped to the size of 160x160x4 using middle crop (Fig. 6.5(a)) and overlapping crop (Fig. 6.5(b)). Therefore, we performed another two experiments implementing the proposed method with middle cropped images and overlapping cropped images as model input. At the end of the down-sampling path, the resolution of the input data were halved in both dimension from 160x160 to 20x20 and the feature maps were increase from 4 to 512. For up-sampling path, the dense feature maps resolution from bottleneck of the network were doubled in each block back to 160x160 resolution. The final output was the prediction of tumour regions and background.

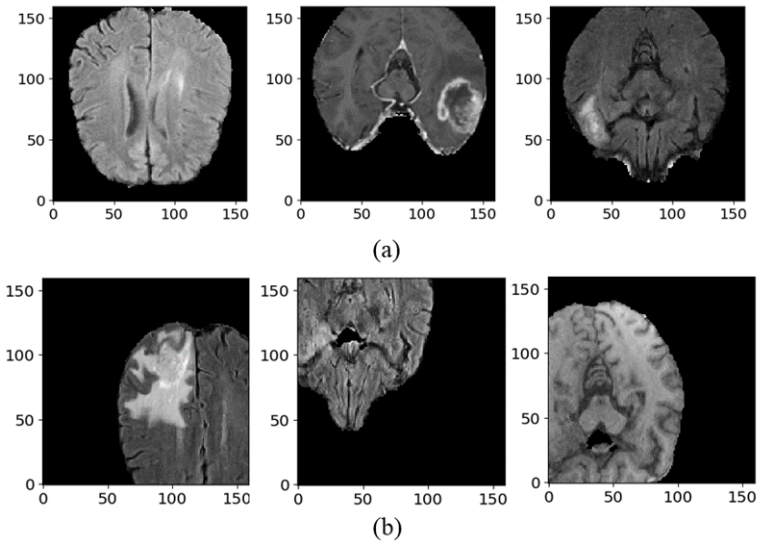


Figure 6.5: Examples of (a) middle cropped images, and (b) overlapping cropped images.

6.3.2 Network implementation

An Adam optimizer was used with a learning rate of 0.00001, values of $\beta_1 = 0.9$, $\beta_2 = 0.999$. All weights were initialized using He initialization [143] with mean of 0 and standard deviation of $\sqrt{2/n}$, where n is the number of input units in the weight tensor, and all biases were initialized as 0. Dropout [145] with probability of 0.5 setting is used to minimise overfitting. We used the training batch of 16 and trained the model for 50 epochs for each segmentation experiments. Categorical cross entropy [146] for multi-class segmentation was used as a loss function and can be defined as

$$L_{CCE}(y, p) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c} \cdot \log\left(p_{i,c}\right) \tag{6.1}$$

where y is ground truth labels, p is a matrix of predicted values for each class, and where indices c and i iterate over all classes and pixels, respectively. The parameters setting and loss function were chosen based on the best performance obtained from series of experiments. Initially, Dice Loss was used to trained the proposed method. However, it did not yielded the segmentation of the tumour regions. Therefore, we used Categorical cross entropy for this multi-class segmentation problem. The proposed method was implemented using Tensorflow and Keras library on PC equipped with NVIDIA GeForce GTX 1070 GPU, an Intel Core i5-8400 CPU 2.80 GHz processor, 16 GB of RAM.

6.4 Experimental results

BraTS 2019 validation dataset was used as unknown testing data to evaluate the trained TwoPath U-Net's performance. The dataset contains multimodal MRI scans (T1, T1Gd, T2, FLAIR) of 125 patients without the ground truth. To obtain the evaluation results, we uploaded the predictions masks to CBICA image processing portal [130] for the segmentation task. We evaluated the proposed method for three tumour regions; WT, TC, and ET. Corresponding labels of each tumour region were previously described in chapter 4 and 5.

Table. 6.1 compares the segmentation result from the classic U-Net, proposed method trained with full-size image (proposed-1), middle cropped images (proposed-2), and overlapping patches (proposed-3). The proposed method with full-size image approach obtained the highest DSC of 0.76, 0.64, and 0.58 for WT, TC, and ET, respectively. However, patches training approach using middle cropped images yields the best performance for identifying the image pixels that contain whole tumour regions with the Sensitivity of 0.85. The specificity values are close to 1 in every tumour regions segmentation because the majority of image pixels are background without ROI.

Table 6.1: Comparative segmentation results using BraTS 2019 validation dataset.

Metric/Tumour region	**DSC**			**Sensitivity**			**Specificity**		
	ET	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**	**WT**	**TC**
Classic U-Net	0.38	0.67	0.46	0.39	0.72	0.43	0.99	0.98	0.99
Proposed-1	**0.58**	**0.76**	**0.64**	**0.69**	0.83	**0.63**	0.99	0.98	0.99
Proposed-2	0.55	0.73	0.62	0.63	**0.85**	0.61	0.99	0.97	0.99
Proposed-3	0.51	0.68	0.56	0.60	0.81	0.56	0.99	0.97	0.99

6.5 Brain tumour segmentation (BraTS) 2020 Challenge

To participate the BraTS 2020 Challenge, we evaluated the proposed method on BraTS 2020 testing dataset. The testing dataset contains multimodal MRI scans of 166 patients without ground truth. The segmentation results validated by the challenge are shown in Table 6.2. We obtained mean dice score of 0.72, 0.66, and 0.64 for the WT, TC, and ET segmentations. The results from BraTS 2020 Challenge has a similar profile to the preliminary results we tested on BraTS 2019 validation dataset. At 50% and 75% quantile of dataset, the proposed method achieved over 0.7 and 0.8mean dice score for all tumour regions segmentation.

Table. 6.2 also provided the performance comparison between the proposed method and nnU-Net [107] which is the winner of BraTS 2020 challenge. The mean DCS scores obtained from the proposed method are generally lower in all tumour regions compared to the nnU-Net's performance.

Table 6.2: Brain tumour regions segmentation performance comparison.

Metric/Tumour region	**DSC**		
	ET	**WT**	**TC**
Proposed			
Mean	0.64	0.72	0.66
25% quantile	0.60	0.64	0.55
50% quantile	0.74	0.79	0.79
75% quantile	0.81	0.86	0.85
nnUNet [107]			
Mean	0.82	0.89	0.85
25% quantile	0.79	0.88	0.88
50% quantile	0.86	0.93	0.93
75% quantile	0.92	0.95	0.96

Fig. 6.6 gives visualisation of the comparison between ground truth (left) and predicted segmentation (right) with NCR/NET (red contours), ED (yellow contours), and ET (blue contours) overlaid on original FLAIR image. The images are from two different patients. The second row of the figure, show that the proposed method faced difficulty in the ET and NCR/NET regions which considered smaller region when compare to ED.

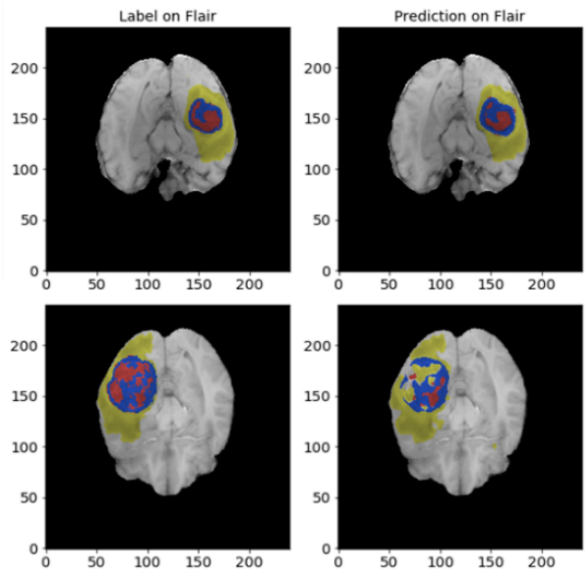


Figure 6.6: Comparison between ground truth (left) and prediction (right) segmentation results on original FLAIR images from training dataset with overlaid labels NCR/NET (red), ED (yellow), and ET (blue) regions.

6.6 Summary

This chapter introduced a novel deep neural networks for multi-class segmentation of the brain tumour regions. The proposed method is called TwoPath U-Net which consists of cascaded local and global feature extraction at every down sampling block of the U-Net like network architecture. The two-pathways feature extraction in this chapter was extended from the Chapter 5 using additional 12x12 kernels to extract global feature from the multimodal MRI input. The proposed method was trained using BraTS 2020 training dataset and evaluate using BraTS 2019 validation dataset.

The proposed method was trained using with different input data strategies including original image resolution of 240x240x4, and cropped image resolution of 160x160x4, where 4 represents four MRI modalities (T1, T1Gd, T2, and FLAIR). The cropped image input aimed to reduce the pixels that belong to the background class. The cropped input also reduced the computation parameters of the model. The training phase of each experiment took approximately 8-12 hours, and the trained model was capable of predicting the segmentation of 155 slices under a minute per patient. The proposed

method obtained the mean DSC of 0.76, 0.64, and 0.58 for WT, TC, and ET segmentation, respectively. Additionally the proposed method reduced the training time compared to the modified U-Net presented in Chapter 5 which trained for each tumour region segmentation separately. However, the proposed method obtained lower DSC for WT and ET segmentation compared to the modified U-Net performance.

The segmentation results obtained from the proposed method were also submitted to participate the BraTS 2020 challenge using BraTS 2020 testing dataset. The results obtained from the proposed method were also compared with the winner of the BraTS 2020. They are overall lower than the winner's performance. The winner of the BraTS 2020 used 3D volumetric input to train the network, while the proposef method used 2D input. The lack of spatial information between slices could be the cause of the overall lower segmentation performance.

Deep neural network for 3D segmentation that combines the advantages of modified U-Net presented in chapter 5 and the proposed method for tumour regions segmentation will be presented in the next chapter.

Chapter 7

3D TwoPath U-Net for brain tumour regions segmentation

7.1 Introduction

In the previous chapters, automatic segmentation algorithms using 2D slices from multimodal MRI data were presented. These networks showed promising results. It is arguably a non-optimal use of the volumetric MRI data. In this chapter, we propose a 3D deep CNN called 3D Two-Path U-Net that give 3D tumour region segmentation. The proposed network keeps the same network architecture of the 2D Two-Path U-Net introduced in the previous chapter but it is extended to a 3D model.

The remainder of this chapter is organised as follows. The framework of deep learning based segmentation using 3D volumetric MRI data and the impact of GPUs in model implementation are described in Section 7.2. The details of the proposed method is given in Section 7.3. Data-preparation and network implementation are presented in Section 7.4. The proposed method performance, visualisation of the results and conclusion are provided in Section 7.5 and 7.6.

7.2 3D brain tumour segmentation

7.2.1 3D brain tumour segmentation framework

3D segmentation of brain tumour considers all sequential MRI data as one volume which makes the process of 3D segmentation framework different from the previously 2D segmentation framework presented in chapter 5 as illustrated in Fig. 7.1. Using 3D deep CNNs based model for segmentation problem significantly increases the training parameters. Hence the original MRI data are usually cropped into a smaller patches and the patch size of 128x128x128 is commonly used among the BraTS challenge participants [101, 107, 108, 111]. Data normalization is performed for both training and testing dataset. Subsequently, the normalised data is used as an input for the 3D DL model to extracts essentials features from the input and predicts each voxel into a curtain class of non-tumour and tumour. The process of model training by minimising the loss calculated from a selected loss function is the same as 2D segmentation framework.

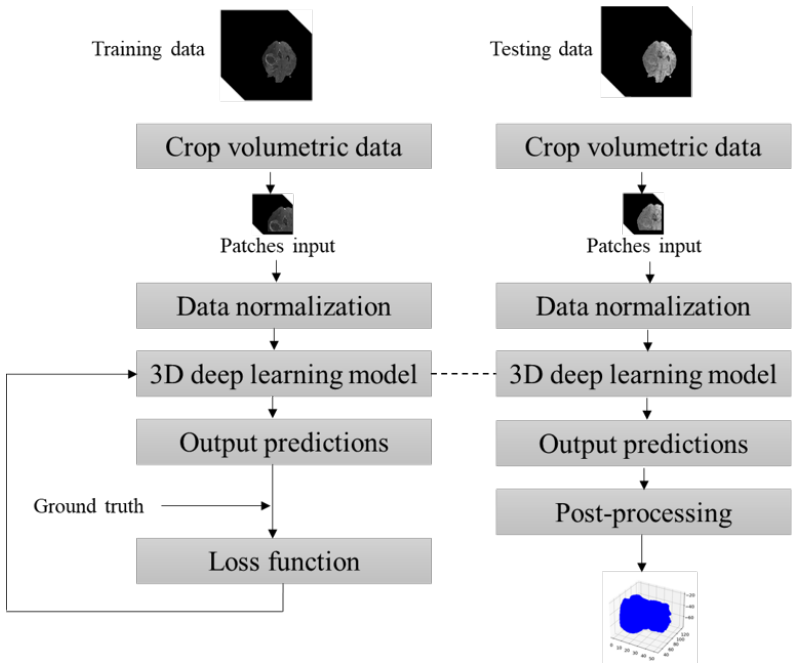


Figure 7.1: 3D TwoPath U-Net network architecture.

We then observe the model performance from the pre-trained model using the testing dataset. To visualise the segmentation output, post-processing is applied to generate the predicted tumour volume. A tumour volume generated from the output of the 3D segmentation model is a preferred representative of the tumour for further quantification analysis. The process of model training for 3D segmentation is usually required an exceedingly longer training time compared to 2D segmentation framework presented in chapter 5. It also requires high-performance computing resource. GPUs acceleration is practically always employed to the implementation of the model. An impact of the GPUs is discussed in the next section.

7.2.2 Impact of GPUs on 3D deep CNNs implementation

Using 3D deep CNNs for segmentation problem comes with increased numbers of training parameters and significant memory and computational requirements [95]. Graphics Processing Unit (GPU) [147] is crucial in 3D segmentation implementation. GPUs were originally developed for 3D graphics but were quickly adopted by the scientific community for powerful parallel computing. Because the computational in CNNs can be represented as tensor or matrix operations and can be effectively parallelised. GPUs enable efficient training of CNNs. Additionally, the most popular deep learning frameworks offer default GPU acceleration support, further enhancing the training process [148]. NVIDIA, a world leading company in AI computing, released a library called cuDNN [149] which is designed to accelerate deep CNNs using GPUs. GPUs also offer high memory capacity which are crucial for working with large 3D volumes [150]. Using GPUs can significantly speed up the model training processes, and enable more efficient and accurate 3D segmentation tasks. In this study, we implemented the proposed method on a GPUs node of ARCHIE-WeST. The specification of ARCHIE-WeST's GPUs node is presented in section 7.4.

7.3 3D TwoPath U-Net

We propose a novel 3D CNNs which is an extension version of 2D TwoPath U-Net introduced in chapter 6. This is referred to as the 3D TwoPath U-Net and is illustrated in Fig. 7.2. It comprises three down-sampling blocks, a bottle-neck layer, and three up-sampling blocks. The proposed method uses multimodal MRI volume (T1Gd, and FLAIR) of 128x128x128 with 2 channels as an input. In every down-sampling block, cascaded 3D two-pathways feature extraction is used to provide local and global feature extractions. The cascaded 3D two-pathway consists of a 3x3x3 convolution concatenate with 12x12x12 followed by 9x9x9 convolution layers with ReLU. Then, 2x2x2 max-pooling operator is performed to halve the size of the feature maps dimension and these features maps become the input of the next down-sampling block. The feature extractions are repeated until the feature maps increase from 2 to 512 dense features. Subsequently, two repeated 3x3x3 convolution layers followed by ReLU are performed. Dropout [145] layer with probability of 0.2 is used after the second 3x3x3 convolution layer of the bottleneck block of the network architecture to reduce overfitting.

For up-sampling blocks, a 3x3x3 transposed convolution with stride of 2 is used to double images resolution in three dimensions and to halve the feature maps, followed by 3x3x3 convolution with ReLU. The result of the up-sampling is then concatenated with the corresponding features from the down-sampling side of the same block level as shown in Fig. 7.2. These concatenations or skip connections provide the higher resolution features with local and global localization contextual information to the up-sampling process. The process repeated until the feature resolution increase back to original resolution with 64 feature maps. Finally, 1x1x1 convolution layer of length 1x1x64 with Sigmoid function is employed to produce the output of tumour volume segmentation and background. Note that in this 3D TwoPath U-Net, we conducted binary segmentation instead of multi-class segmentation presented in Chapter 6. Although, the multi-class segmentation significantly reduced the training time. The binary segmentation was able to made effective use of the dice loss during the training process, and yielded higher DSC, especially for WT and TC segmentation.

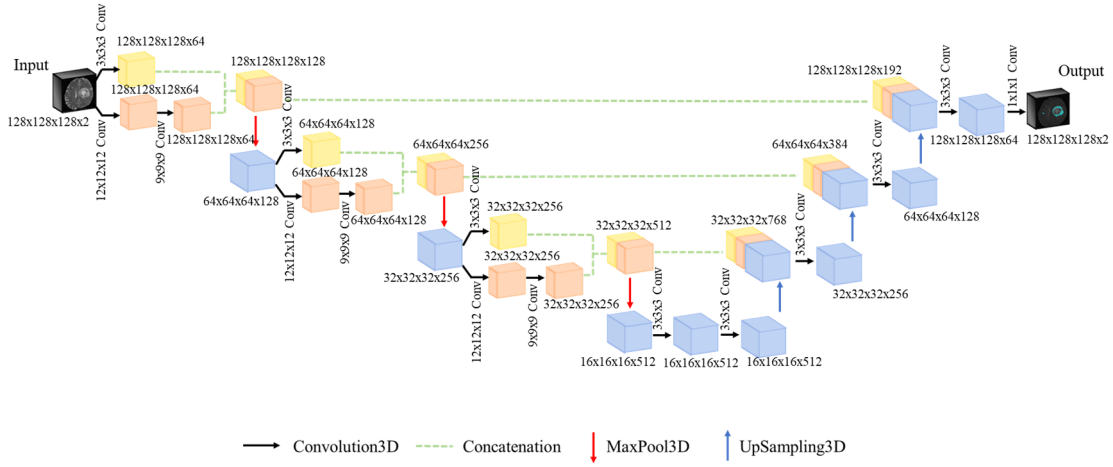


Figure 7.2: 3D TwoPath U-Net network architecture.

7.4 Experiments

7.4.1 3D Data-preparation

We employed the proposed 3D TwoPath U-Net on BraTS 2021 dataset for the segmentation of whole tumour (WT), tumour core (TC), and enhancing tumour (ET) regions instead of learning the labels ED, ET, and NCR/NET to better represent the clinical tumour structure as previously discussed in chapter 4. The dataset were split into the ratio of 0.8:0.2 for cross validation. Volumetric MRI data from BraTS 2021 training dataset were cropped into a volume of the size 128x128x128 using middle cropping to reduce the background. To further reduce the computational parameters, T1Gd and FLAIR scans were selected as multimodal MRI based on the manual segmentation protocol presented in Chapter 4. Data normalization was performed on each MRI volume across the dataset. The proposed model was trained separately for each tumour region. We also used pre-trained weights from WT segmentation as transfer learning for TC and ET segmentation model trainings.

7.4.2 Network implementation

An Adam optimizer was used with a learning rate of 0.00001, values of $\beta_1 = 0.9$, $\beta_2 = 0.999$. All weights were initialized using He initialization [143] with mean of 0 and standard deviation of $\sqrt{2/n}$, where n is the number of input units in the weight tensor, and all biases were initialized as 0. Dropout [145] with probability of 0.2 setting is used to minimise overfitting. We used the training batch of 1 and trained the model for 50 epochs for WT segmentation. The pre-trained model then used to train for another 25 epochs for TC and ET. Dice loss [104] was used as a loss function and can be defined as

$$DL = 1 - \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \tag{7.1}$$

where $p_i \in P$ is the predicted binary segmentation volume, $g_i \in G$ is the gound truth, and the sums run over the N voxels. The proposed method was implemented using Tensorflow and Keras library on GPUs node of ARCHIE-WeST with 4 NVIDIA A100 GPUs, housed in Lenovo SR670 servers, each with 40 cores and 192GB RAM.

7.5 Experimental results

7.5.1 Segmentation results

Table. 7.1 shows the comparative segmentation results using validation data from BraTS 2021 dataset. The proposed method achieved the mean DSC of 0.87, 0.70, and 0.67 for WT, TC, and ET, respectively. By significantly increasing the training data, the proposed method outperformed the 2D TwoPath method introduced in chapter 6 for all tumour regions. Table. 7.1 also shows that segmentation results from Optimized U-Net [111], and NVAUTO [108], the top-rank teams in the BraTS 2021 challenge. Both teams represent NVIDIA, a world leading company in AI computing. Although the proposed method's segmentation accuracy is generally lower than both teams, it was only 5% lower for WT segmentation.

Table 7.1: Segmentation results of the proposed method using validation data from BraTS 2021 dataset.

Method	**DSC**		
	ET	**WT**	**TC**
2D TwoPath (BraTS2020)	0.58	0.76	0.64
Proposed method	0.67	0.87	0.70
Optimized U-Net [111]	0.91	0.91	0.91
NVAUTO [108]	0.89	0.92	0.93

Fig. 7.3 visualises multi-view of the WT segmentations obtained from the proposed method. Each row of the image features the original FLAIR, T1Gd, ground truth, and the predicted segmentation. The first row shows the axial view of the segmentation, the second row shows the sagittal view, and the third row shows the frontal view. The predicted segmentations of axial view and sagittal view show promising results. However, an underestimation of the tumour is presented in the frontal view. The model failed to identify the rest of the smaller tumour region on the left of the larger tumour region.

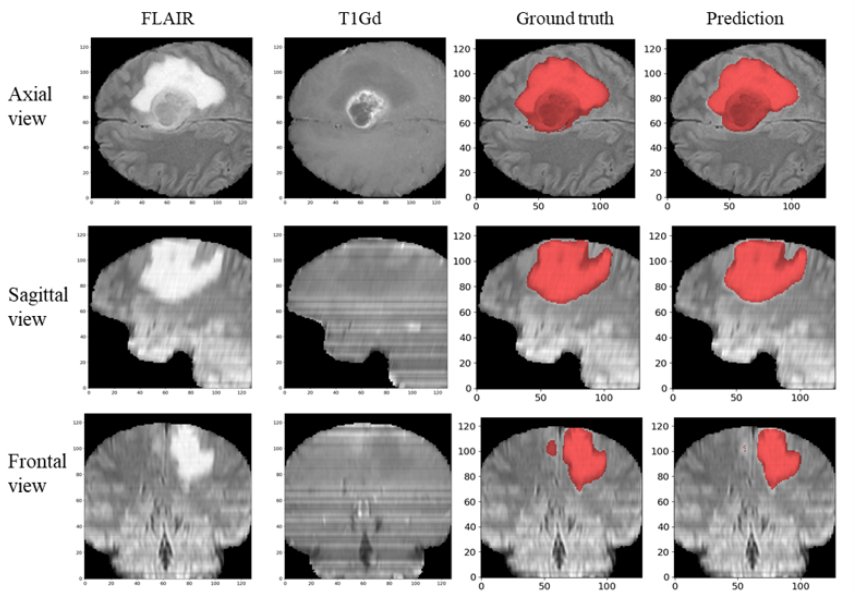


Figure 7.3: Visualisation of multi-view original FLAIR, T1Gd, ground truth and predicted segmentation of WT.

Fig. 7.4 gives the visualisation of predicted segmentation WT volume (red) obtained from the proposed method and the corresponding ground truth (blue). Each row of the image features the segmentation from different patients. An underestimation of the tumour volume can be observed in the top right and the bottom left of the predicted segmentation from the first patient, as well as the top right of the prediction from the second patient. However, the predictions generally show promising results compared to the corresponding ground truth. Furthermore, the 3D tumour output obtained from the proposed method provides a better clinical representation because its ability to incorporate spatial information between the slices compared to the stacked 2D contours where the gaps between slices are presented as shown in Chapter 5.

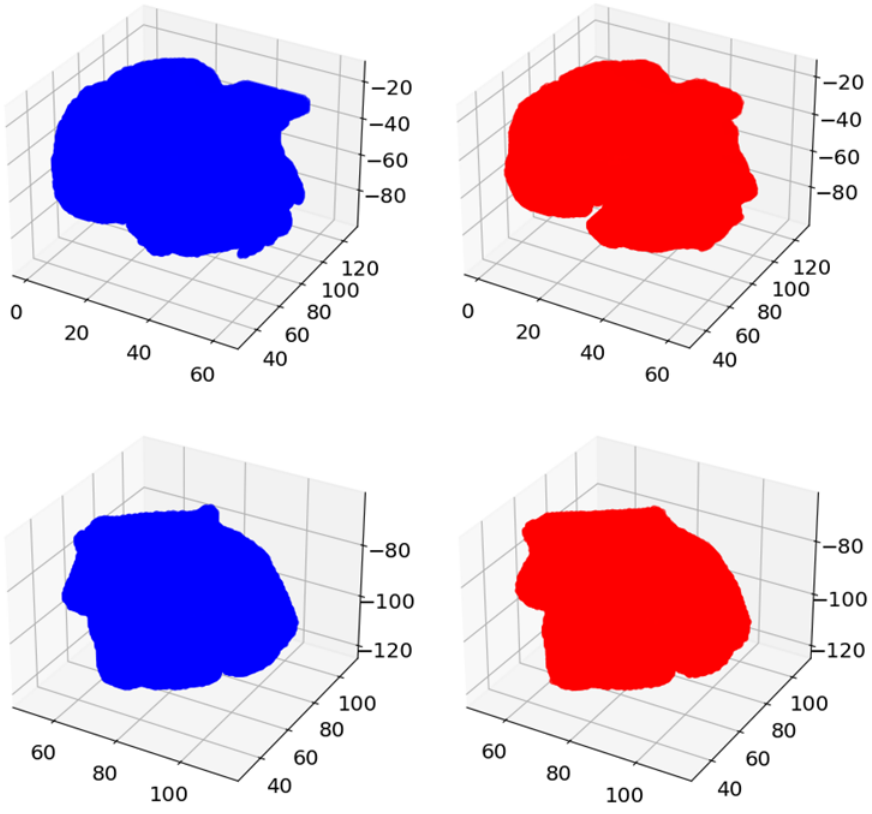


Figure 7.4: Visualisation of ground truth (blue) and predicted segmentation (red) tumour volumes from two patients.

7.5.2 Segmentation time

The proposed method was implemented on the GPUs node of ARCHIE-WeST with specification presented in the section 7.4.2. The model training took approximately 74 hours for WT segmentation, 37 hours for TC segmentation, and 37 hours for ET segmentation. Fig. 7.5 shows example of a terminal for job script submission in the login node of the ARCHIE-WeSt. Command sacct -j followed by job script id was used to display the accurate runtime. The elapsed time of model training for TC segmentation shown in this terminal was 1-13:25:34 indicating 37 hours 25 minutes and 34 seconds.

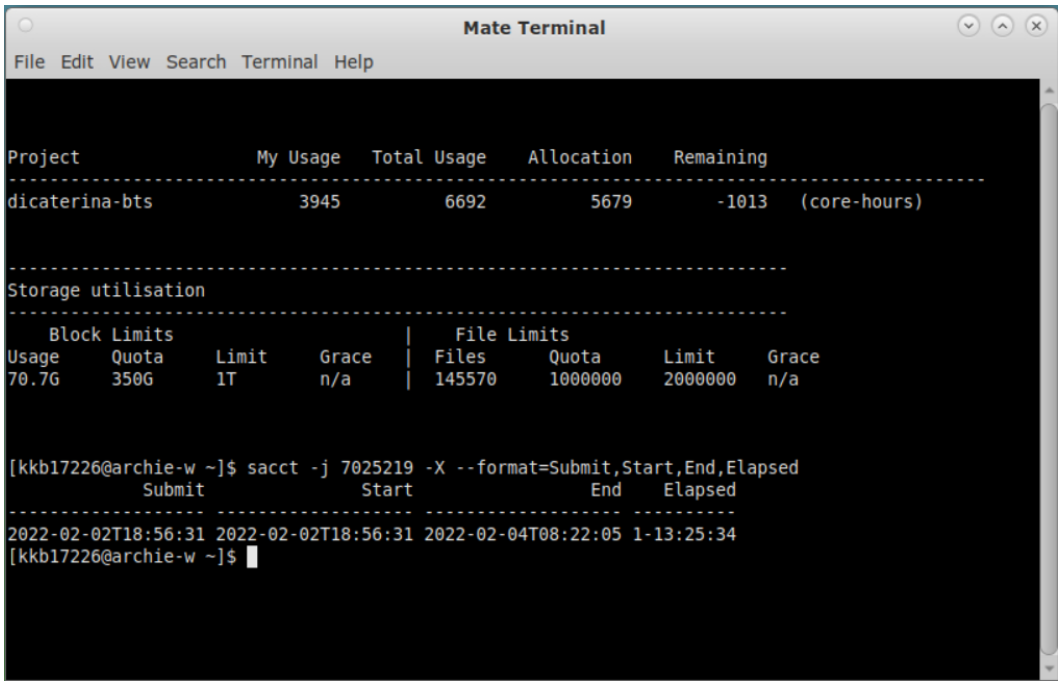


Figure 7.5: Example of training time for TC segmentation task.

Fig. 7.6 gives the predicted WT from the first patient presented in Fig. 7.4 with the prediction time using pre-trained 3D TwoPath U-Net on Google Colab with Tesla V100 GPU. The prediction and visualisation took approximately 45 seconds with the DSC of 0.92. Although the training phase of the proposed method took long hours, the predicted segmentation by the pre-trained model achieved time-efficient comparing to

the manual segmentation presented in chapter 3.

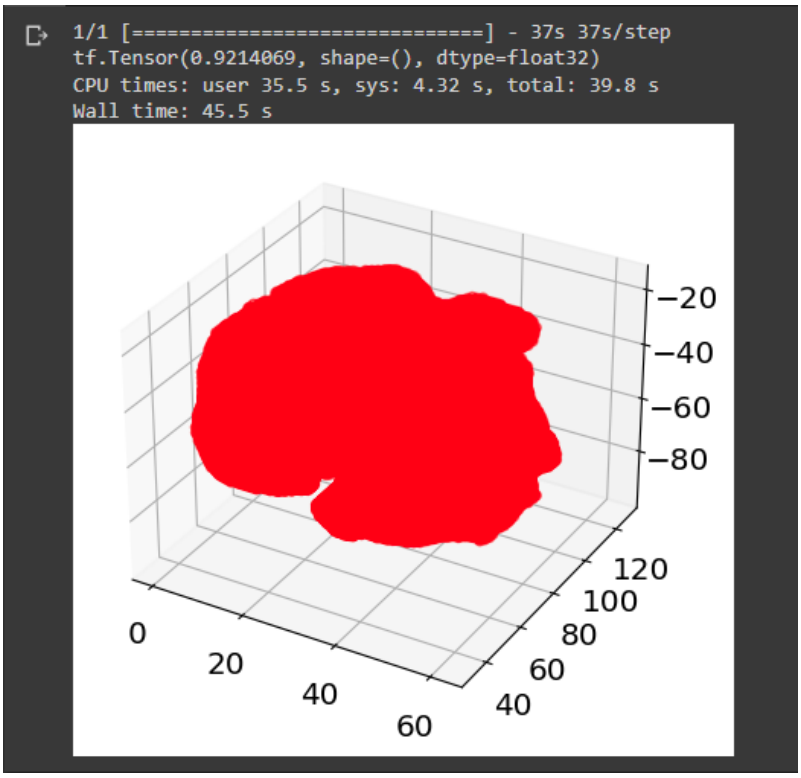


Figure 7.6: Example of WT segmentation with prediction time.

7.6 Conclusion

This chapter presented a novel 3D deep CNN for brain tumour regions segmentation called 3D TwoPath U-Net. The proposed method is an extended version of the 2D TwoPath U-Net introduced in chapter 6. The proposed method keeps the two-pathways feature extraction architecture to provided the model with cascaded local and global features. The proposed method conducted the binary segmentation which trained for WT, TC, and ET segmentation separately. However, Dice Loss and transfer learning techniques were employed during the model training to improve the segmentation performance.

Due to the significantly computational requirement of the 3D deep learning network, cropped voxels from volumetric MRI data were used as model's input. T1Gd and FLAIR scans were selected as multimodal MRI input. This further reduced the computation parameters of the networks. Furthermore, the reduction in MRI modalities required in an automatic segmentation framework could decrease the time patients spend on the MRI imaging process. The proposed method was implemented using a GPUs node of ARCHIE-WeST which is a high performance computing (HPC) facility available at the University of Strathclyde.

The 3D volumetric used as the model input consist of sequential data that leverage the spatial information between slices. The segmentation results obtained from the proposed method achieved the mean DSC of 0.87, 0.70, 0.67 for WT, TC, and ET segmentation, respectively. The results also show the improvement of the segmentation compared to the results obtained from the modified U-Net and 2D TwoPath U-Net presented in the previous chapters. Furthermore, the results were compared with the top rank teams from the BraTS 2021 challenge, showing that the performance of the proposed method are overall lower than the top rank teams in all tumour regions. Additionally, the 3D visualisation of the predicted tumour volumes were obtained using the proposed method. Moreover, the proposed method achieved time-efficiency segmentation, and was capable of giving the predicted segmentation in under a minute, whereas the manual segmentation takes approximately 60 minutes per patient.

Overall, the presented novel 3D CNN approach shows the potential for improved segmentation performance using volumetric MRI data. Future research can focus on refining the model architecture and the quantification of tumour volume using the 3D results obtained from the 3D CNNs.

Chapter 8

Conclusions and future work

8.1 Conclusions

In this thesis, a variety of image processing techniques and deep learning methods have been explored to tackle several challenges of radiotherapy planning. Traditional process of manually GTV delineation requires clinical expertise which is subjective and time-consuming. The main aim of this thesis is to determine the appropriate method for the automatic brain tumour segmentation, and to develop an automatic segmentation of brain tumour that covers whole tumour and its intra-tumour structures, and avoids healthy brain tissues using MRI data.

In Chapter 2, an extensive literature review was presented, focusing on the anatomical structures of the human brain, introducing the brain tumour and its imaging, including the basic concept of the MRI imaging. Various brain tumour segmentation techniques including the conventional methods and the deep learning based methods were presented in Chapter 3. In Chapter 4, the real dataset of the brain tumour patients and its imbalanced in label distribution were discussed. The techniques to mitigate the biases from the imbalanced data includes loss functions and evaluation metrics were also introduced.

The first novel contribution of this thesis was presented in chapter 5. A novel deep neural networks for fully automatic semantic segmentation of brain tumour region from MRI was presented. The proposed method is a modified version of the classical U-Net

architecture. The modification introduced two-pathways feature extraction to extract low-level local and global features directly from the multimodal MRI input. 2D axial slices from four modalities of the MRI including T1, T1Gd, T2, and FLAIR from the BraTS 2018 dataset were used to form four channels of the model input. Transposed convolution was also employed as the up-sampling operator. The novel achievement from this work is the brain tumour regions segmentation results with an improvement of the tumour regions segmentation in higher Dice Similarity Coefficient (DSC) compared to the original U-Net. The proposed method outperforms U-Net especially in the segmentation of ET and NCR/NET which considered difficult regions to segment. Stacked 2D contours obtained from the proposed method provides the visualisation of the gross tumour volume (GTV). However, the proposed method experienced the overfitting during the model training process and the segmentation of NCR/NET could be improved. The segmentation obtained from the validation data using the proposed method were generally poor due to the inhomogeneity of the NCR/NET and ET regions. The proposed method took approximately 12 hours to train for each tumour segmentation, and capable of giving the predicted segmentation under a minute per patient after the model was trained.

The second contribution of this thesis was the development of a multi-class deep neural network called 2D TwoPath U-Net presented in Chapter 6. The proposed method extended the two-pathway feature extraction model to provide cascaded local and global features in every down-sampling step of the network architecture. The cascaded architecture was shown to be able to preserve the global information of the brain structure, image intensity along with its location. To further improve the segmentation, the proposed method returns the output with 4 predicted classes of NCR/NET, ED, ET, and background to generate WT and TC regions. The proposed method was trained using BraTS 2020 dataset and tested using BraTS 2019 validation dataset. Experiments with cropped images input to decrease the background of the images were also conducted. Because the proposed method performed multi-class segmentation, this significantly decreased the model training time. The segmentation results obtained from the 2D TwoPath U-Net are overall higher than original U-Net. However, WT and ET segmen-

tation obtained from the proposed method are lower in DSC compared to the results obtained from the modified U-Net. Moreover, the segmentation results obtained from the proposed method using BraTS 2020 testing dataset were submitted to participate the BraTS 2020 challenge. Although the performance of the proposed method are generally lower than the winner of the challenge, we have made contribution to the automatic brain tumour segmentation research.

The third contribution of this thesis was presented in chapter 7. A novel 3D deep neural networks called 3D TwoPath U-Net was developed as an extension version of the previous 2D TwoPath U-Net. Implementation of the 3D deep neural networks faces the burden of significantly high computational parameters. A GPUs node of ARCHIE-WeST was employed in this work. Cropped voxels of the sequential volumetric MRI data were used as model input. T1Gd and FLAIR scans were selected to generate multimodal MRI. This further decreased the computational parameters of the 3D TwoPath network. Additionally, it could benefits the patient to reduce the time spend during the MRI imaging process. The proposed 3D TwoPath U-Net performed binary segmentation thus enabling the use of dice loss. The novel achievement obtained from this work is the brain tumour regions segmentation with improved DSC compared to the results obtained from the modified U-Net and the 2D TwoPath U-Net. Furthermore, the output of the proposed method gives 3D visualisation of the tumour volume with spatial information without the gaps between MRI slices compared to the stacked 2D contours. Finally, time-efficient segmentation is achieved, tumour regions can be segmented in less than a minute for each patient using the proposed method.

Overall this thesis explored the deep learning based methods for semantic segmentation and developed novel automatic segmentation frameworks for brain tumour segmentation. The novel contributions were achieved through the 2D and 3D segmentation techniques. While the 3D segmentation framework produces the desired segmentation output, it is undeniable that it requires high-performance GPUs and extensive training time (over 37 hours for each tumour region). In situations where high-performance computing facilities are not available, the 2D segmentation framework becomes an efficient segmentation option, as it demands less computational power and shorter training

time. The 2D segmentation only took approximately 8-12 hours using single NVIDIA GeForce GTX 1070 GPU which is more accessible compared to the NVIDIA A100 GPUs node of the ARCHIE-WeSt. Additionally, the 2D segmentation framework can utilize 2D slices from sequential 3D volumetric MRI input data, with further 3D reconstruction for visualising the GTV output.

8.2 Future work

There are various possible works that can be explored for further research. Firstly, train the networks with more data. BraTS 2021 dataset significantly increases the number of MRI data compared to the previous years, and provides an opportunity to enhance the training process. However, computational limitations restrict the use of all available image modalities, leading to the utilization of only T1Gd and FLAIR images in chapter 7. Future work can be conducted to incorporate all four MRI modalities, including T1, T1Gd, T2, and FLAIR. The algorithms can capture more comprehensive understanding of the tumour characteristics and potentially improve segmentation accuracy by using more modalities. To further reduce model complexity, exploring the use of consecutive small kernels instead of large kernels for global feature extraction is worth considering. This adjustment could allow to model to be able to utilize more image modalities without requiring higher computational power. Moreover, further algorithms for quantification of 3D tumour volume obtained from the proposed 3D TwoPath U-Net can be explored.

Secondly, an improvement of the model training technique. Inspired by the winner of BraTS 2019 challenge who used the cascaded U-Net to train the segmentation from coarse to fine. We could stack the proposed 3D TwoPath U-Net in the same manner to possibly improve the segmentation performance especially for NCR/NET and ET regions segmentation.

Thirdly, a crucial aspect of translating the proposed algorithms into practical applications is the development of rapid diagnosis tools with user-friendly interfaces. The algorithms can be transformed into intuitive software applications that integrate into

the existing clinical workflow An example of such interface could be a web-based application that allows clinicians to upload MRI scans, run the segmentation algorithm, and visualise the segmented tumour regions interactively. This could lead to a pilot study in the clinical use with the real patients.

Lastly, we can refine the proposed algorithms to learn from time-dependent data such as MRI scans of the same patient from different time period during the treatment. This could enable clinicians to monitor the effectiveness of the treatment and allow personalised treatment that is appropriate for each patient.

In summary, future research in automatic brain tumour segmentation using deep neural networks can focus on network training with larger and more diverse datasets, improving the network training with cascade model technique, developing user-friendly diagnosis tools, and refining algorithms to learn from time-dependent data. These future works have the potential to revolutionise tumour segmentation and assist clinicians in making more accurate diagnosis and treatment planning.

Appendix A

A.1 Visualisation of feature maps obtained from 2D TwoPath U-Net

The visualisation of feature maps obtained from cascaded two-pathway feature extraction of the 2D TwoPath U-Net were presented in section 6.2.1. More example of feature maps from each down-sampling block are illustrated as follows:

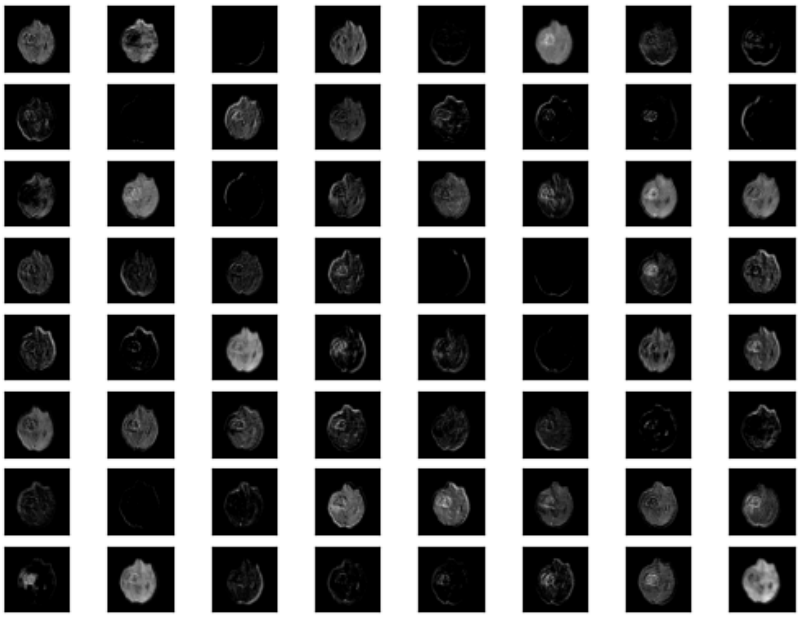


Figure A.1: Example of feature maps obtained from the first down-sampling block of the TwoPath U-Net using 3x3 kernels.

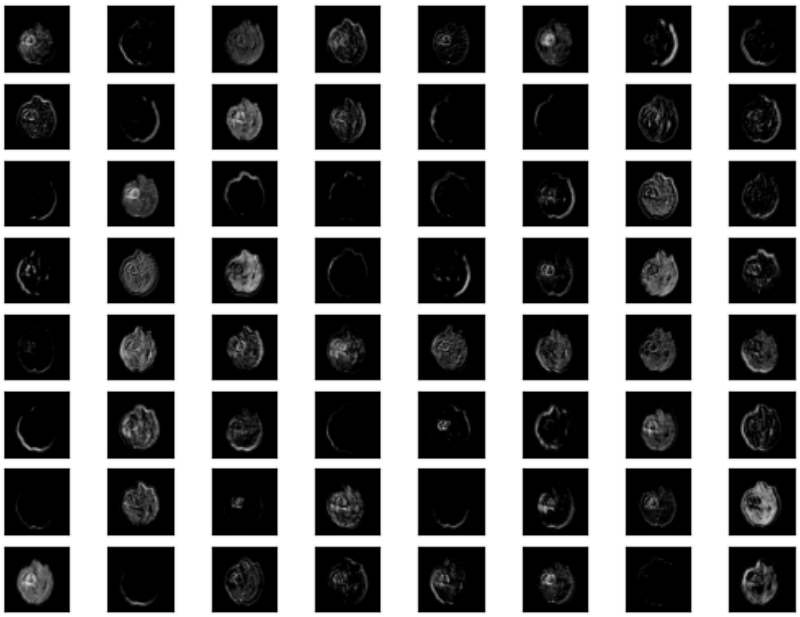


Figure A.2: Example of feature maps obtained from the first down-sampling block of the 2D TwoPath U-Net using 9x9 kernels.

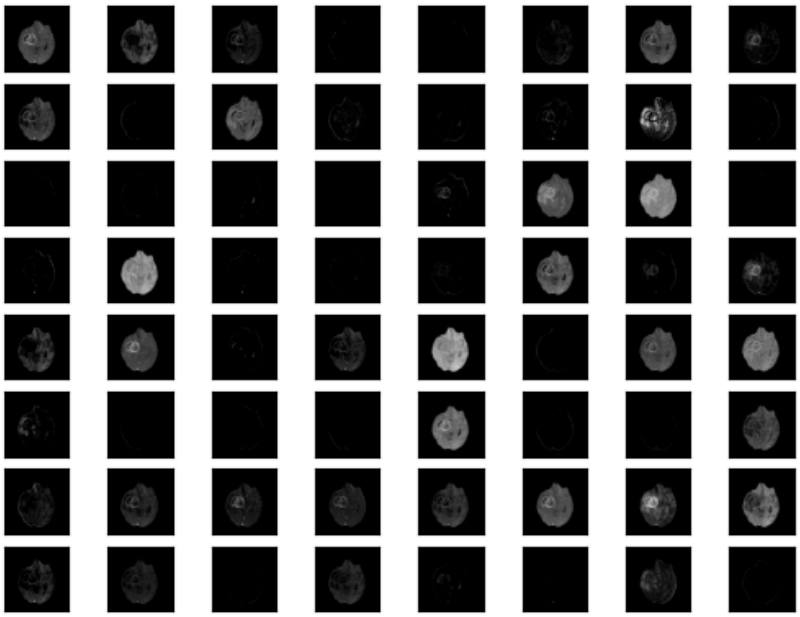


Figure A.3: Example of feature maps obtained from the first down-sampling block of the 2D TwoPath U-Net using 12x12 kernels.

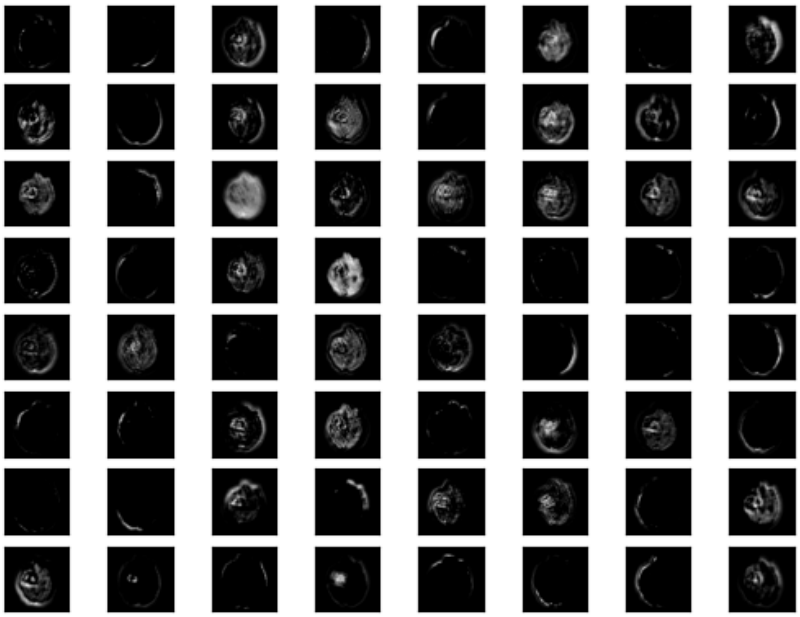


Figure A.4: Example of feature maps obtained from the second down-sampling block of the 2D TwoPath U-Net using 3x3 kernels.

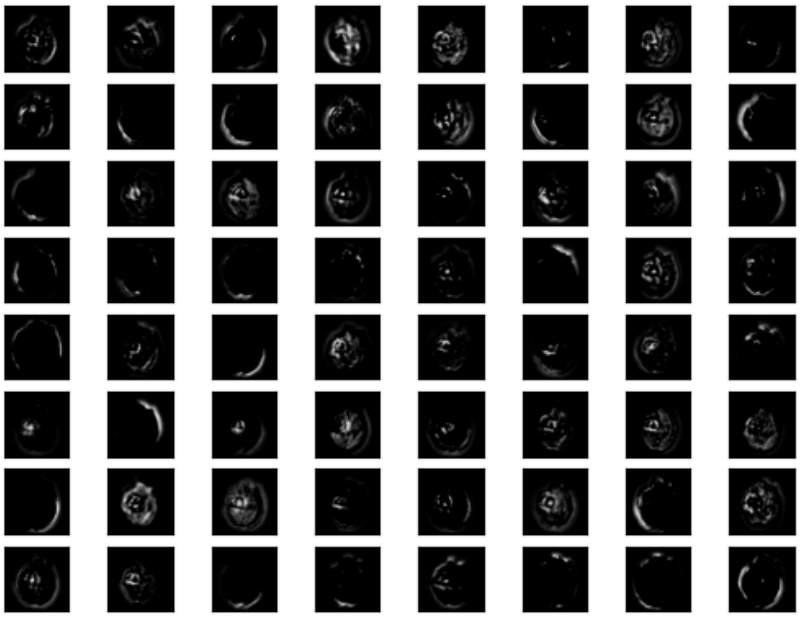


Figure A.5: Example of feature maps obtained from the second down-sampling block of the 2D TwoPath U-Net using 9x9 kernels.

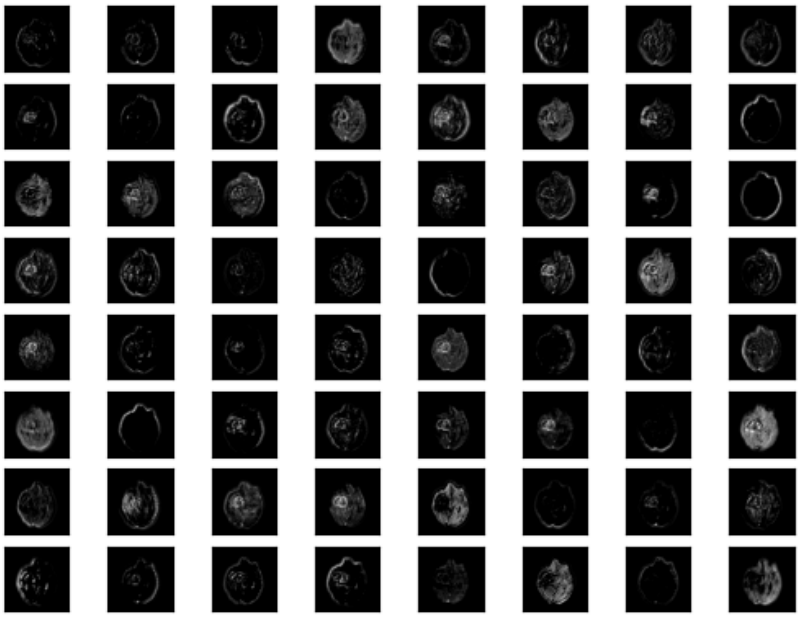


Figure A.6: Example of feature maps obtained from the second down-sampling block of the 2D TwoPath U-Net using 12x12 kernels.

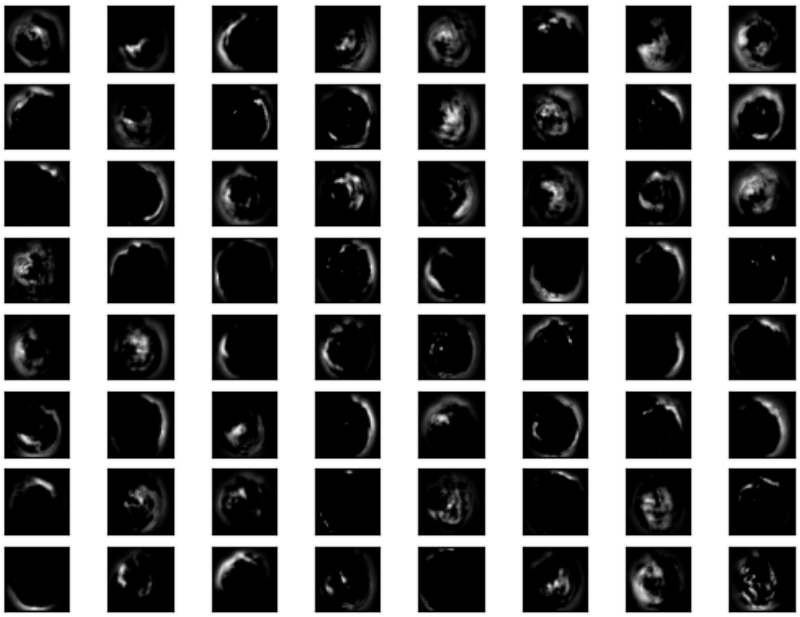


Figure A.7: Example of feature maps obtained from the third down-sampling block of the 2D TwoPath U-Net using 3x3 kernels.

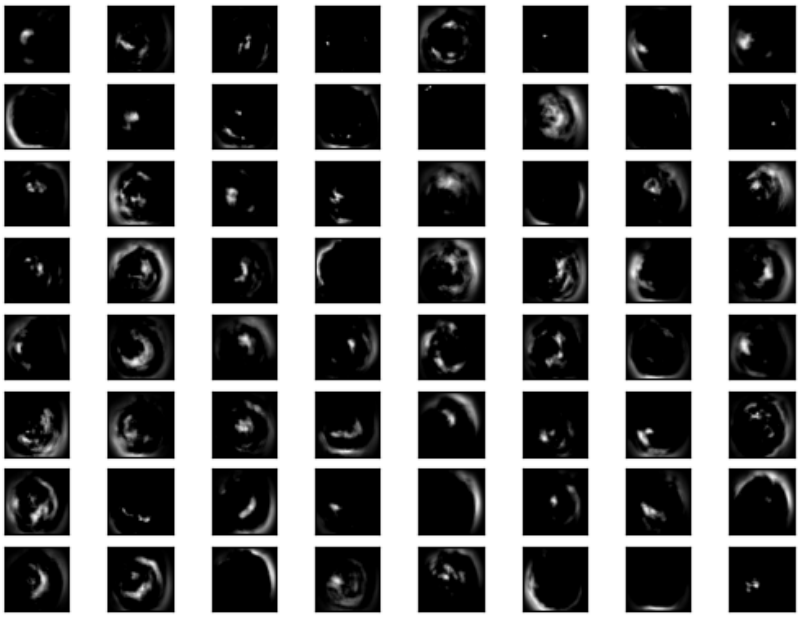


Figure A.8: Example of feature maps obtained from the third down-sampling block of the 2D TwoPath U-Net using 9x9 kernels.

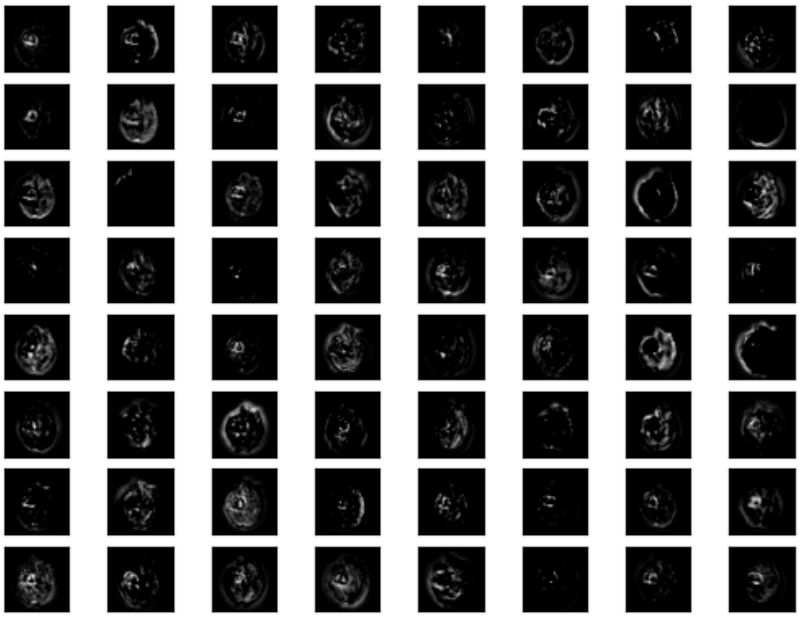


Figure A.9: Example of feature maps obtained from the third down-sampling block of the 2D TwoPath U-Net using 12x12 kernels.

Appendix B

ARCHIE-WeST

B.1 ARCHIE-WeST access

To access GPU node on ARCHIE-WeSt HPC system, we used ThinLinc remote desktop. ThinLinc allows remote access to Linux server for Windows, MAC and Linux desktops. Fig. B.1 shows an example of the ThinnLinc login interface. We used thinlinc.hpc.strath.ac.uk server and the user's Strathclyde DS username and password to login. After successfully login, user can submit job scripts via the server login node terminal as shown in Fig. B.2.

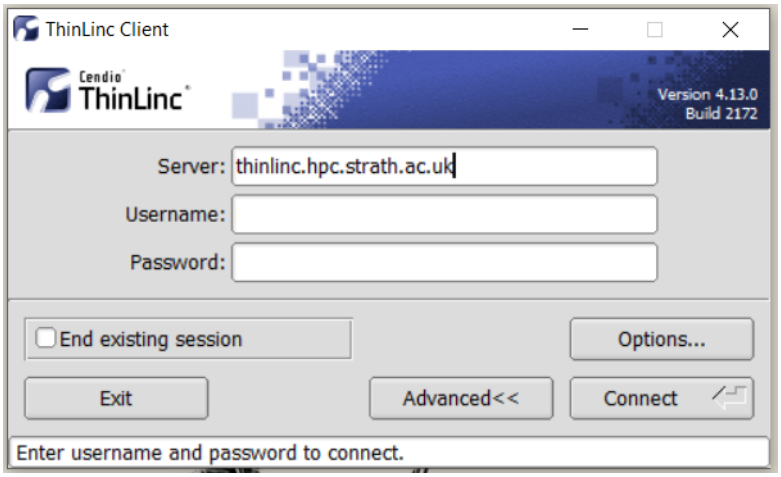


Figure B.1: ThinLinc remote desktop access.

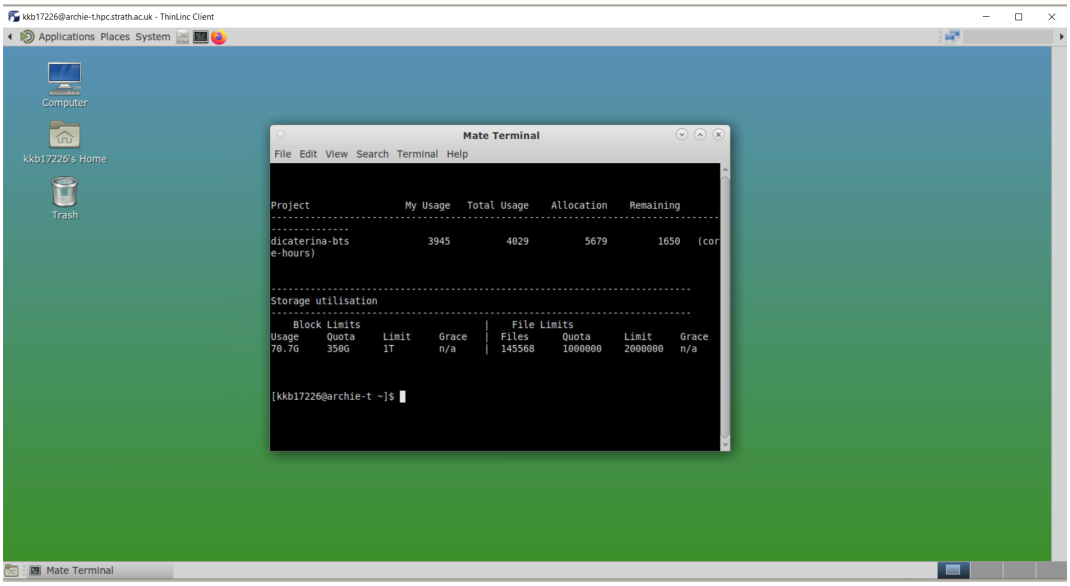


Figure B.2: The login node and the example of job script submission terminal.

B.2 Example of single GPU job script

The following script is an example of job script for submitting our 3D brain tumour segmentation python codes to GPU node of the ARCHIE-WeST HPC system. In this script, we set the GPU node to NVIDIA A100 with 48 hours runtime limit.

```
#!/bin/bash
# Propogate environment variables to the compute node
#SBATCH --export=ALL
# Run in the standard partition (queue)
#SBATCH --partition=gpu --gres=gpu:A100
# Specify project account
#SBATCH --account=dicaterina-bts
# Specify (hard) runtime (HH:MM:SS)
#SBATCH --time=48:00:00
#SBATCH --job-name=brats2020
# Output file
#SBATCH --output=slurm-%j.out
```

```
module purge
module load anaconda/python-3.8.8/2021.05
module load nvidia/sdk/21.3

#============================================================
# Prologue script to record job details
#============================================================
/opt/software/scripts/job_prologue.sh
#------------------------------------------------------------

source /opt/software/anaconda/python-3.8.8/2021.05/etc/profile.d/conda.sh
. ${HOME}/setup-conda
#conda init bash
conda activate keerati
cd /users/kkb17226/Brats20
python3 /users/kkb17226/Brats20/train_brats2021_1000_T1ceFlair_ET.py

#============================================================
# Epilogue script to record job endtime and runtime
# Do not change the line below
#============================================================
/opt/software/scripts/job_epilogue.sh
#------------------------------------------------------------
```

Appendix C

C.1 Automatic segmentation results from 3D TwoPath U-Net

The visualisation of multi-view segmentation obtained from 3D TwoPath was presented in section 7.5. More example of TC and ET segmentation results are shown in Fig. C.1 and Fig. C.2, respectively.

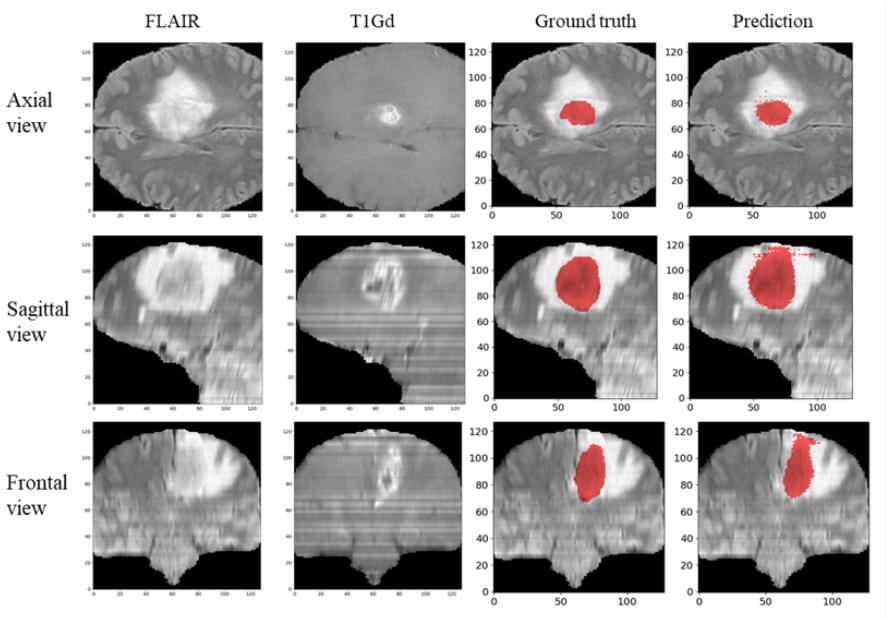


Figure C.1: Visualisation of multi-view original FLAIR, T1Gd, ground truth and predicted segmentation of TC.

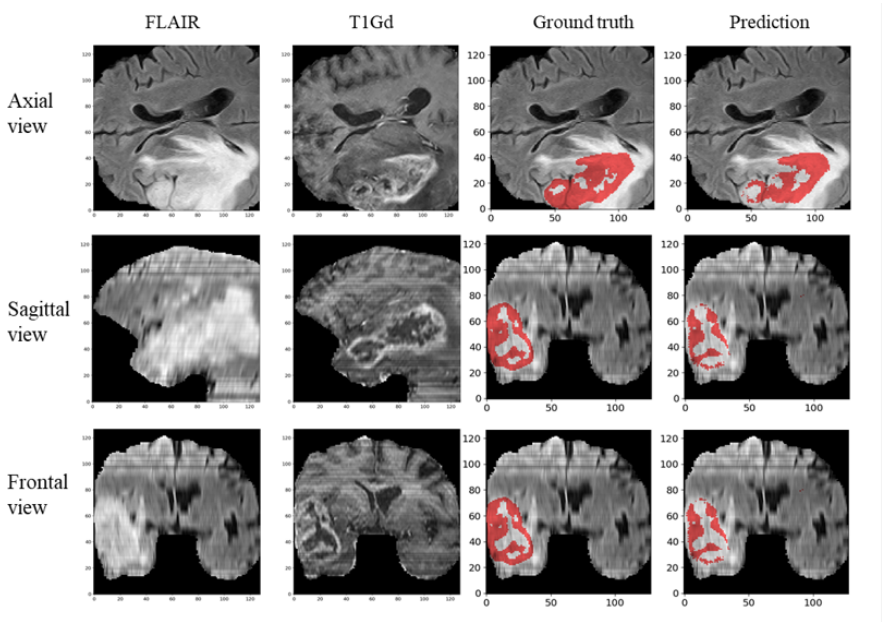


Figure C.2: Visualisation of multi-view original FLAIR, T1Gd, ground truth and predicted segmentation of ET.

Bibliography

[1] Cancer Research UK, "Brain, other CNS and intracranial tumours statistics," 2023. [Online]. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/brain-other-cns-and-intracranial-tumours{#}heading-Zero

[2] ——, "Brain, other CNS and intracranial tumours statistics," 2022. [Online]. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/brain-other-cns-and-intracranial-tumours

[3] D. N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary," *Acta Neuropathologica*, vol. 131, no. 6, pp. 803–820, 2016. [Online]. Available: https://doi.org/10.1007/s00401-016-1545-1

[4] I. Goetz and A.-L. Grosu, "Advanced Imaging Modalities and Treatment of Gliomas: Radiation Therapy BT - Brain Tumor Imaging," E. Hattingen and U. Pilatus, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 135–142. [Online]. Available: https://doi.org/10.1007/174{_}2014{_}1022

[5] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, E. Gerstner, M. Weber, T. Arbel, B. B. Avants, N. Ayache, P. Buendia, D. L. Collins, N. Cordier, J. J. Corso, A. Criminisi, T. Das, H. Delingette, Ç. Demiralp, C. R. Durst,

M. Dojat, S. Doyle, J. Festa, F. Forbes, E. Geremia, B. Glocker, P. Golland, X. Guo, A. Hamamci, K. M. Iftekharuddin, R. Jena, N. M. John, E. Konukoglu, D. Lashkari, J. A. Mariz, R. Meier, S. Pereira, D. Precup, S. J. Price, T. R. Raviv, S. M. S. Reza, M. Ryan, D. Sarikaya, L. Schwartz, H. Shin, J. Shotton, C. A. Silva, N. Sousa, N. K. Subbanna, G. Szekely, T. J. Taylor, O. M. Thomas, N. J. Tustison, G. Unal, F. Vasseur, M. Wintermark, D. H. Ye, L. Zhao, B. Zhao, D. Zikic, M. Prastawa, M. Reyes, and K. V. Leemput, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.

[6] D. L. Pham, C. Xu, and J. L. Prince, "Current Methods in Medical Image Segmentation," *Annual Review of Biomedical Engineering*, vol. 2, no. 1, pp. 315–337, aug 2000. [Online]. Available: https://doi.org/10.1146/annurev.bioeng.2.1.315

[7] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, and Y. Pan, "A survey of MRI-based brain tumor segmentation methods," *Tsinghua Science and Technology*, vol. 19, no. 6, pp. 578–595, 2014.

[8] N. Gordillo, E. Montseny, and P. Sobrevilla, "State of the art survey on MRI brain tumor segmentation," *Magnetic Resonance Imaging*, vol. 31, no. 8, pp. 1426–1438, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0730725X13001872

[9] H. Noh, S. Hong, and B. Han, "Learning Deconvolution Network for Semantic Segmentation," *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1520–1528, 2015.

[10] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240–1251, may 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7426413/

[11] A. In, C. Direkolu, and M. ah, "Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods," *Procedia Computer Science*, vol. 102, pp. 317–324, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S187705091632587X

[12] M. Niyazi, M. Brada, A. J. Chalmers, S. E. Combs, S. C. Erridge, A. Fiorentino, A. L. Grosu, F. J. Lagerwaard, G. Minniti, R.-O. Mirimanoff, U. Ricardi, S. C. Short, D. C. Weber, and C. Belka, "ESTRO-ACROP guideline target delineation of glioblastomas," *Radiotherapy and Oncology*, vol. 118, no. 1, pp. 35–42, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167814015006611

[13] M. P. A. Starmans, S. R. van der Voort, J. M. Castillo Tovar, J. F. Veenland, S. Klein, and W. J. Niessen, "Chapter 18 - Radiomics: Data mining using quantitative medical image features," in *The Elsevier and MICCAI Society Book Series*, S. K. Zhou, D. Rueckert, G. B. T. H. o. M. I. C. Fichtinger, and C. A. Intervention, Eds. Academic Press, 2020, pp. 429–456. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128161760000235

[14] M. Crispino and E. Crispino, *Brain*. Cham: Springer International Publishing, 2015, pp. 1–27. [Online]. Available: https://doi.org/10.1007/978-3-319-10750-9{_}1

[15] Mayfield Clinic, "Anatomy of the Brain." [Online]. Available: http://www.mayfieldclinic.com/pe-anatbrain.htm

[16] C. Henley, *Foundations of Neuroscience*. Michigan State University Libraries, 2021.

[17] "Medical gallery of Blausen Medical 2014," *WikiJournal of Medicine*, vol. 1, no. 2, 2014. [Online]. Available: https://en.wikiversity.org/wiki/WikiJournal{_}of{_}Medicine/Medical{_}gallery{_}of{_}Blausen{_}Medical{_}2014

[18] T. brain tumour charity, "Brain cells." [Online]. Available: https://www.thebraintumourcharity.org/understanding-brain-tumours/symptoms-and-information/brain-cells/

[19] OpenStax, "Textbook OpenStax Anatomy and Physiology," 2016. [Online]. Available: https://commons.wikimedia.org/wiki/File:1209{_}Glial{_}Cells{_}of{_}the{_}CNS-02.jpg

[20] Mayfield Clinic, "Glioma brain tumors (astrocytoma, oligodendroglioma, glioblastoma)." [Online]. Available: http://www.mayfieldclinic.com/pe-glioma.htm

[21] Cancer Reseach UK, "Glioma," 2019. [Online]. Available: https://www.cancerresearchuk.org/about-cancer/brain-tumours/types/glioma-adults

[22] D. A. Forst, B. V. Nahed, J. S. Loeffler, and T. T. Batchelor, "Low-grade gliomas." *The oncologist*, vol. 19, no. 4, pp. 403–413, apr 2014.

[23] R. Stupp, M. Brada, M. J. Van Den Bent, J.-C. Tonn, and G. Pentheroudakis, "High-grade glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up," *Annals of oncology*, vol. 25, pp. iii93–iii101, 2014.

[24] N. G. Burnet, S. J. Thomas, K. E. Burton, and S. J. Jefferies, "Defining the tumour and target volumes for radiotherapy." *Cancer imaging : the official publication of the International Cancer Imaging Society*, vol. 4, no. 2, pp. 153–161, oct 2004.

[25] M. C. Mabray, R. F. Barajas Jr, and S. Cha, "Modern brain tumor imaging," *Brain tumor research and treatment*, vol. 3, no. 1, pp. 8–23, apr 2015. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/25977902https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4426283/

[26] Johns Hopkins Medicine, "Computed Tomography (CT or CAT) Scan of the Brain." [Online]. Avail-

able: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/computed-tomography-ct-or-cat-scan-of-the-brain

[27] Mayo Clinic, "CT scan," 2022. [Online]. Available: https://www.mayoclinic.org/tests-procedures/ct-scan/about/pac-20393675{#}:{~}:text=Aspecialdyecalledcontrastmaterialisneededforsome,Bymouth.

[28] O. Keunen, T. Taxt, R. Grüner, M. Lund-Johansen, J.-C. Tonn, T. Pavlin, R. Bjerkvig, S. P. Niclou, and F. Thorsen, "Multimodal imaging of gliomas in the context of evolving cellular and molecular therapies," *Advanced Drug Delivery Reviews*, vol. 76, pp. 98–115, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169409X14001513

[29] J. Wölfer and W. Stummer, "Advanced Imaging Modalities and Treatment of Gliomas: Neurosurgery BT - Brain Tumor Imaging," E. Hattingen and U. Pilatus, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 143–154. [Online]. Available: https://doi.org/10.1007/174{_}2014{_}1023

[30] "Spin and the Nuclear Magnetic Resonance Phenomenon BT - How Does MRI Work? An Introduction to the Physics and Function of Magnetic Resonance Imaging," D. Weishaupt, V. D. Köchli, and B. Marincek, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 1–5. [Online]. Available: https://doi.org/10.1007/978-3-540-37845-7{_}1

[31] "The Basics BT - MRI in Clinical Practice," G. Liney, Ed. London: Springer London, 2006, pp. 3–25. [Online]. Available: https://doi.org/10.1007/1-84628-162-8{_}1

[32] "Relaxation BT - How Does MRI Work? An Introduction to the Physics and Function of Magnetic Resonance Imaging," D. Weishaupt, V. D. Köchli, and B. Marincek, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 7–10. [Online]. Available: https://doi.org/10.1007/978-3-540-37845-7{_}2

[33] "Image Contrast BT - How Does MRI Work? An Introduction to the Physics and Function of Magnetic Resonance Imaging," D. Weishaupt, V. D. Köchli,

and B. Marincek, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 11–20. [Online]. Available: https://doi.org/10.1007/978-3-540-37845-7{_}3

[34] P. Sprawls, "Magnetic Resonance Imaging: Principles, Methods, and Techniques," *Radiology*, vol. 221, no. 2, pp. 462–462, nov 2001. [Online]. Available: http://pubs.rsna.org/doi/10.1148/radiol.2212012537B

[35] K. Krupa and M. Bekiesińska-Figatowska, "Artifacts in magnetic resonance imaging." *Polish journal of radiology*, vol. 80, pp. 93–106, 2015.

[36] A. Murphy and U. Bashir, "Gibbs and truncation artifacts," in *Radiopaedia.org*. Radiopaedia.org, jan 2012. [Online]. Available: http://radiopaedia.org/articles/16567

[37] L. Erasmus, D. Hurter, M. Naude, H. Kritzinger, and S. Acho, "A short overview of MRI artefacts," *South African Journal of Radiology*, vol. 8, no. 2, p. 13, jun 2004. [Online]. Available: https://sajr.org.za/index.php/sajr/article/view/127

[38] J. Juntu, J. Sijbers, D. Van Dyck, and J. Gielen, "Bias Field Correction for MRI Images BT - Computer Recognition Systems," M. Kurzyński, E. Puchała, M. Woźniak, and A. żołnierek, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 543–551.

[39] R. A. Peters, "A new algorithm for image noise reduction using mathematical morphology," *IEEE Transactions on Image Processing*, vol. 4, no. 5, pp. 554–568, 1995.

[40] S. A. Sarra, "Algorithm 899: The Matlab Postprocessing Toolkit," *ACM Trans. Math. Softw.*, vol. 37, no. 1, jan 2010. [Online]. Available: https://doi.org/10.1145/1644001.1644011

[41] U. Vovk, F. Pernus, and B. Likar, "A Review of Methods for Correction of Intensity Inhomogeneity in MRI," *IEEE Transactions on Medical Imaging*, vol. 26, no. 3, pp. 405–421, 2007.

[42] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4ITK: improved N3 bias correction," *IEEE transactions on medical imaging*, vol. 29, no. 6, pp. 1310–1320, jun 2010. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/20378467https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3071855/

[43] D. Leung, X. Han, T. Mikkelsen, and L. B. Nabors, "Role of MRI in primary brain tumor evaluation." *Journal of the National Comprehensive Cancer Network : JNCCN*, vol. 12, no. 11, pp. 1561–1568, nov 2014.

[44] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Physics in Medicine & Biology*, vol. 58, no. 13, p. R97, 2013. [Online]. Available: https://dx.doi.org/10.1088/0031-9155/58/13/R97

[45] 3D Slicer, "3D Slicer image computing platform." [Online]. Available: https://www.slicer.org/

[46] Z. Liu, L. Tong, L. Chen, Z. Jiang, F. Zhou, Q. Zhang, X. Zhang, Y. Jin, and H. Zhou, "Deep learning based brain tumor segmentation: a survey," *Complex & Intelligent Systems*, vol. 9, no. 1, pp. 1001–1026, 2023. [Online]. Available: https://doi.org/10.1007/s40747-022-00815-5

[47] X. Guo, L. Schwartz, and B. Zhao, "Semi-automatic segmentation of multimodal brain tumor using active contours," *Multimodal Brain Tumor Segmentation*, vol. 27, 2013.

[48] A. Hamamci, N. Kucuk, K. Karaman, K. Engin, and G. Unal, "Tumor-Cut: Segmentation of Brain Tumors on Contrast Enhanced MR Images for Radiosurgery Applications," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, pp. 790–804, mar 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6112681/

[49] B. H. Menze, K. Van Leemput, D. Lashkari, M.-A. Weber, N. Ayache, and P. Golland, "A generative model for brain tumor segmentation in multi-modal images."

Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention, vol. 13, no. Pt 2, pp. 151–159, 2010.

[50] ——, "Segmenting glioma in multi-modal images using a generative model for brain lesion segmentation," *Proc MICCAIBRATS (Multimodal Brain Tumor Segmentation Challenge)*, vol. 8, 2012.

[51] S. Doyle, F. Vasseur, M. Dojat, and F. Forbes, "Fully automatic brain tumor segmentation from multiple MR sequences using hidden Markov fields and variational EM," *Procs. NCI-MICCAI BraTS*, pp. 18–22, 2013.

[52] E. Geremia, B. H. Menze, and N. Ayache, "Spatial decision forests for glioma segmentation in multi-channel MR images," *MICCAI Challenge on Multimodal Brain Tumor Segmentation*, vol. 34, pp. 14–18, 2012.

[53] S. Pereira, J. Festa, J. A. Mariz, N. Sousa, and C. A. Silva, "Automatic brain tissue segmentation of multi-sequence MR images using random decision forests," *Proceedings of the MICCAI grand challenge on MR brain image segmentation (MRBrainS'13)*, 2013.

[54] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle, "Brain tumor segmentation with Deep Neural Networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841516300330

[55] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the royal statistical society: series B (methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[56] G. Celeux, F. Forbes, and N. Peyrard, "EM procedures using mean field-like approximations for Markov model-based image segmentation," *Pattern recognition*, vol. 36, no. 1, pp. 131–144, 2003.

[57] U. Ilhan and A. Ilhan, "Brain tumor segmentation based on a new threshold approach," *Procedia Computer Science*, vol. 120, pp. 580–587, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050917324948

[58] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, jan 1979. [Online]. Available: http://ieeexplore.ieee.org/document/4310076/

[59] M. Sujan, N. Alam, S. A. A. Noman, and M. J. Islam, "A Segmentation based Automated System for Brain Tumor Detection," *International Journal of Computer Applications*, vol. 153, pp. 41–49, 2016.

[60] I. Maiti and M. Chakraborty, "A new method for brain tumor segmentation based on watershed and edge detection algorithms in HSV colour model," in *2012 NATIONAL CONFERENCE ON COMPUTING AND COMMUNICATION SYSTEMS*, 2012, pp. 1–5.

[61] W. Deng, W. Xiao, H. Deng, and J. Liu, "MRI brain tumor segmentation with region growing method based on the gradients and variances along and inside of the boundary curve," in *2010 3rd International Conference on Biomedical Engineering and Informatics*, vol. 1, 2010, pp. 393–396.

[62] I. S. Bajwa, M. N. Asghar, and M. A. Naeem, "Learning-based improved seeded region growing algorithm for brain tumor identification," *Proceedings of the Pakistan Academy of Sciences: Part A*, vol. 54, no. 2, pp. 127–133, 2017. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85028315139{&}partnerID=40{&}md5=f9b781096e05bbcb965e131cfa4132a8

[63] C. J. J. Sheela and G. Suganthi, "Morphological edge detection and brain tumor segmentation in Magnetic Resonance (MR) images based on region growing and performance evaluation of modified Fuzzy C-Means (FCM) algorithm," *Multimedia Tools and Applications*, vol. 79, no. 25, pp. 17 483–17 496, 2020. [Online]. Available: https://doi.org/10.1007/s11042-020-08636-9

[64] J. Kim, S. Lee, G. Lee, Y. Park, and Y. Hong, "Using a Method Based on a Modified K-Means Clustering and Mean Shift Segmentation to Reduce File Sizes and Detect Brain Tumors from Magnetic Resonance (MRI) Images," *Wireless Personal Communications*, vol. 89, no. 3, pp. 993–1008, 2016. [Online]. Available: https://doi.org/10.1007/s11277-016-3420-8

[65] B. Singh and P. Aggarwal, "Detection of brain tumor using modified mean-shift based fuzzy c-mean segmentation from MRI Images," in *2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2017, pp. 536–545.

[66] A. Wadhwa, A. Bhardwaj, and V. Singh Verma, "A review on brain tumor segmentation of MRI images," *Magnetic Resonance Imaging*, vol. 61, pp. 247–259, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0730725X19300347

[67] R. Ranjbarzadeh, A. Caputo, E. B. Tirkolaee, S. Jafarzadeh Ghoushchi, and M. Bendechache, "Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools," *Computers in Biology and Medicine*, vol. 152, p. 106405, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482522011131

[68] A. Aslam, E. Khan, and M. M. S. Beg, "Improved Edge Detection Algorithm for Brain Tumor Segmentation," *Procedia Computer Science*, vol. 58, pp. 430–437, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050915021687

[69] J. Vijay and J. Subhashini, "An efficient brain tumor detection methodology using K-means clustering algoriftnn," in *2013 International Conference on Communication and Signal Processing*, 2013, pp. 653–657.

[70] D. M. Kumar, D. Satyanarayana, and M. N. G. Prasad, "An improved Gabor wavelet transform and rough K-means clustering algorithm for MRI brain tumor

image segmentation," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 6939–6957, 2021. [Online]. Available: https://doi.org/10.1007/s11042-020-09635-6

[71] R. Khilkhal and M. Ismael, "Brain Tumor Segmentation Utilizing Thresholding and K-Means Clustering," in *2022 Muthanna International Conference on Engineering Science and Technology (MICEST)*, 2022, pp. 43–48.

[72] S. N. Shivhare, N. Kumar, and N. Singh, "A hybrid of active contour model and convex hull for automated brain tumor segmentation in multimodal MRI," *Multimedia Tools and Applications*, vol. 78, no. 24, pp. 34 207–34 229, 2019. [Online]. Available: https://doi.org/10.1007/s11042-019-08048-4

[73] C. J. J. Sheela and G. Suganthi, "Accurate MRI brain tumor segmentation based on rotating triangular section with fuzzy C- means optimization," *Sdhan*, vol. 46, no. 4, p. 226, 2021. [Online]. Available: https://doi.org/10.1007/s12046-021-01744-8

[74] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. USA: Prentice-Hall, Inc., 2006.

[75] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.

[76] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

[77] J. C. Bezdek, L. O. Hall, and L. Clarke, "Review of MR image segmentation techniques using pattern recognition." *Medical physics*, vol. 20, no. 4, pp. 1033–1048, 1993.

[78] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[79] N. E. A. Khalid, "MRI Brain Abnormalities Segmentation using K-Nearest Neighbors (k-NN)," 2011.

[80] M. D. Steenwijk, P. J. W. Pouwels, M. Daams, J. W. van Dalen, M. W. A. Caan, E. Richard, F. Barkhof, and H. Vrenken, "Accurate white matter lesion segmentation by k nearest neighbor classification with tissue type priors (kNN-TTPs)," *NeuroImage: Clinical*, vol. 3, pp. 462–469, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213158213001332

[81] R. Ayachi and N. Ben Amor, "Brain Tumor Segmentation Using Support Vector Machines BT - Symbolic and Quantitative Approaches to Reasoning with Uncertainty," C. Sossai and G. Chemello, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 736–747.

[82] T. S. Kumar, K. Rashmi, S. Ramadoss, L. K. Sandhya, and T. J. Sangeetha, "Brain tumor detection using SVM classifier," in *2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS)*, 2017, pp. 318–323.

[83] T. Zhang, Y. Xia, and D. D. Feng, "Hidden Markov random field model based brain MR image segmentation using clonal selection algorithm and Markov chain Monte Carlo method," *Biomedical Signal Processing and Control*, vol. 12, pp. 10–18, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809413001110

[84] C. S. Rao and K. Karunakara, "A comprehensive review on brain tumor segmentation and classification of MRI images," *Multimedia Tools and Applications*, vol. 80, no. 12, pp. 17 611–17 643, 2021. [Online]. Available: https://doi.org/10.1007/s11042-020-10443-1

[85] K. Held, E. R. Kops, B. J. Krause, W. M. Wells, R. Kikinis, and H. W. Muller-Gartner, "Markov random field segmentation of brain MR images," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 878–886, 1997.

[86] I. Liljeqvist, "The Essence of Artificial Neural Networks." [Online]. Available: https://medium.com/@ivanliljeqvist/the-essence-of-artificial-neural-networks-5de300c995d6

[87] "CS231n Convolutional Neural Networks for Visual Recognition," 2023. [Online]. Available: https://cs231n.github.io/

[88] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[89] Y. LeCun, C. Cortes, and C. Burges, "The MNIST Database of Handwritten Digits. New York, USA." 1998. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[90] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. Burges, L. Bottou, and K. Q. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[91] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, dec 2015. [Online]. Available: http://link.springer.com/10.1007/s11263-015-0816-y

[92] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv e-prints*, p. arXiv:1409.1556, sep 2014.

[93] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, and D. Shen, "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *NeuroImage*, vol. 108, pp. 214–224, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811914010660

[94] J. Kleesiek, G. Urban, A. Hubert, D. Schwarz, K. Maier-Hein, M. Bendszus, and A. Biller, "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," *NeuroImage*, vol. 129, pp. 460–469, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811916000306

[95] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1361841516301839

[96] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation BT - Medical Image Computing and Computer-Assisted Intervention MICCAI 2015," N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.

[97] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," p. arXiv:1411.4038, nov 2014. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2014arXiv1411.4038L

[98] Tensorflow, "Data augmentation," 2023. [Online]. Available: https://www.tensorflow.org/tutorials/images/data{_}augmentation

[99] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," p. arXiv:1606.06650, jun 2016. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2016arXiv160606650C

[100] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks," in *Medical Image Understanding and Analysis*, M. Valdés Hernández and V. González-Castro, Eds. Cham: Springer International Publishing, 2017, pp. 506–517.

[101] Z. Jiang, C. Ding, M. Liu, and D. Tao, "Two-Stage Cascaded U-Net: 1st Place Solution to BraTS Challenge 2019 Segmentation Task BT - Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries," A. Crimi and S. Bakas, Eds. Cham: Springer International Publishing, 2020, pp. 231–241.

[102] Y. Chen, K. Wang, X. Liao, Y. Qian, Q. Wang, Z. Yuan, and P.-A. Heng, "Channel-Unet: A Spatial Channel-Wise Convolutional Neural Network for Liver and Tumors Segmentation," *Frontiers in Genetics*, vol. 10, 2019. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fgene.2019.01110

[103] C. Guo, M. Szemenyei, Y. Yi, W. Wang, B. Chen, and C. Fan, "SA-UNet: Spatial Attention U-Net for Retinal Vessel Segmentation," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 1236–1242.

[104] F. Milletari, N. Navab, and S. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 565–571.

[105] Y. Chen, B. Shi, Z. Wang, P. Zhang, C. D. Smith, and J. Liu, "Hippocampus segmentation through multi-view ensemble ConvNets," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017, pp. 192–196.

[106] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, "CE-Net: Context Encoder Network for 2D Medical Image Segmentation," *IEEE Transactions on Medical Imaging*, vol. 38, no. 10, pp. 2281–2292, 2019.

[107] F. Isensee, P. F. Jäger, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, "nnU-Net for Brain Tumor Segmentation BT - Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries," A. Crimi and S. Bakas, Eds. Cham: Springer International Publishing, 2021, pp. 118–132.

[108] M. M. R. Siddiquee and A. Myronenko, "Redundancy Reduction in Semantic Segmentation of 3D Brain Tumor MRIs," in *BrainLes@MICCAI*, 2021.

[109] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. Roth, and D. Xu, "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images," p. arXiv:2201.01266, jan 2022. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2022arXiv220101266H

[110] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021. [Online]. Available: https://doi.org/10.1038/s41592-020-01008-z

[111] M. Futrega, A. Milesi, M. Marcinkiewicz, and P. Ribalta, "Optimized U-Net for Brain Tumor Segmentation," p. arXiv:2110.03352, oct 2021. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2021arXiv211003352F

[112] L. M. Molera, "Deep Learning for Computer Vision using Python and MATLAB," 2022. [Online]. Available: https://blogs.mathworks.com/deep-learning/2022/01/03/deep-learning-for-computer-vision-using-python-and-matlab/

[113] S. Raschka, Y. H. Liu, V. Mirjalili, and D. Dzhulgakov, *Machine Learning with PyTorch and Scikit-Learn: Develop machine learning and deep learning models with Python.* Packt Publishing Ltd, 2022.

[114] NumPy, "NumPy." [Online]. Available: https://numpy.org/

[115] SciPy, "SciPy." [Online]. Available: https://scipy.org/

[116] Scikit-learn, "Scikit-learn. Machine learning in python." [Online]. Available: https://scikit-learn.org/stable/

[117] Matplotlib, "Matplotlib: visualization with Python." [Online]. Available: https://matplotlib.org/

[118] Keras, "Keras." [Online]. Available: https://keras.io/about/

[119] TensorFlow, "Introduction to TensorFlow." [Online]. Available: https://www.tensorflow.org/learn

[120] TensorLayer, "TensorLayer," 2019. [Online]. Available: https://tensorlayer.readthedocs.io/en/latest/https://github.com/tensorlayer/tensorlayer/blob/master/docs/index.rst

[121] PyTorch, "PyTorch." [Online]. Available: https://pytorch.org/

[122] M. Paluszek and S. Thomas, "Software for Machine Learning," in *MATLAB Machine Learning*. Berkeley, CA: Apress, 2017, pp. 25–31. [Online]. Available: http://link.springer.com/10.1007/978-1-4842-2250-8{_}3

[123] MathWorks, "Getting Started with Image Segmenter App," 2022. [Online]. Available: https://uk.mathworks.com/help/images/image-segmentation-using-the-image-segmenter-app.html

[124] ARCHIE-WeSt, "ARCHIE-WeSt User Guide: Introduction." [Online]. Available: https://docs.hpc.strath.ac.uk/user-guide/

[125] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features," *Scientific Data*, vol. 4, no. 1, p. 170117, 2017. [Online]. Available: https://doi.org/10.1038/sdata.2017.117

[126] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R. T. Shinohara, C. Berger, S. M. Ha, and M. Rozycki, "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge," *arXiv preprint arXiv:1811.02629*, 2018.

[127] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, J. Freymann, K. Farahani, and C. Davatzikos, "Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection," *The cancer imaging archive*, vol. 286, 2017.

[128] S. Bakas, H. Akbari, and A. Sotiras, "Segmentation labels for the pre-operative scans of the TCGA-GBM collection. The Cancer Imaging Archive," 2017.

[129] U. Baid, S. Ghodasara, S. Mohan, M. Bilello, E. Calabrese, E. Colak, K. Farahani, J. Kalpathy-Cramer, F. C. Kitamura, S. Pati, L. M. Prevedello, J. D. Rudie, C. Sako, R. T. Shinohara, T. Bergquist, R. Chai, J. Eddy, J. Elliott, W. Reade, T. Schaffter, T. Yu, J. Zheng, A. W. Moawad, L. Otavio Coelho, O. McDonnell, E. Miller, F. E. Moron, M. C. Oswood, R. Y. Shih, L. Siakallis, Y. Bronstein, J. R. Mason, A. F. Miller, G. Choudhary, A. Agarwal, C. H. Besada, J. J. Derakhshan, M. C. Diogo, D. D. Do-Dai, L. Farage, J. L. Go, M. Hadi, V. B. Hill, I. Michael, D. Joyner, C. Lincoln, E. Lotan, A. Miyakoshi, M. Sanchez-Montano, J. Nath, X. V. Nguyen, M. Nicolas-Jilwan, J. Ortiz Jimenez, K. Ozturk, B. D. Petrovic, C. Shah, L. M. Shah, M. Sharma, O. Simsek, A. K. Singh, S. Soman, V. Statsevych, B. D. Weinberg, R. J. Young, I. Ikuta, A. K. Agarwal, S. C. Cambron, R. Silbergleit, A. Dusoi, A. A. Postma, L. Letourneau-Guillon, G. J. Guzman Perez-Carrillo, A. Saha, N. Soni, G. Zaharchuk, V. M. Zohrabian, Y. Chen, M. M. Cekic, A. Rahman, J. E. Small, V. Sethi, C. Davatzikos, J. Mongan, C. Hess, S. Cha, J. Villanueva-Meyer, J. B. Freymann, J. S. Kirby, B. Wiestler, P. Crivellaro, R. R. Colen, A. Kotrotsou, D. Marcus, M. Milchenko, A. Nazeri, H. Fathallah-Shaykh, R. Wiest, A. Jakab, M.-A. Weber, A. Mahajan, B. Menze, A. E. Flanders, and S. Bakas, "The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification," p. arXiv:2107.02314, jul 2021. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2021arXiv210702314B

[130] "CBICA Image Processing Portal." [Online]. Available: https://ipp.cbica.upenn.edu/

[131] "BraTS 2021 challenge," 2021. [Online]. Available: https://www.synapse.org/{#}!Synapse:syn25829067/wiki/610863

[132] D. Müller, I. Soto-Rey, and F. Kramer, "Towards a guideline for evaluation metrics in medical image segmentation," *BMC Research Notes*, vol. 15, no. 1,

p. 210, dec 2022. [Online]. Available: https://bmcresnotes.biomedcentral.com/articles/10.1186/s13104-022-06096-y

[133] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Medical Imaging*, vol. 15, no. 1, p. 29, 2015. [Online]. Available: https://doi.org/10.1186/s12880-015-0068-x

[134] Y.-H. Nai, B. W. Teo, N. L. Tan, S. O'Doherty, M. C. Stephenson, Y. L. Thian, E. Chiong, and A. Reilhac, "Comparison of metrics for the evaluation of medical segmentations using prostate MRI dataset," *Computers in Biology and Medicine*, vol. 134, p. 104497, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482521002912

[135] C. H. Sudre, W. Li, T. K. M. Vercauteren, S. Ourselin, and M. J. Cardoso, "Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations," *Deep learning in medical image analysis and multimodal learning for clinical decision support : Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, held in conjunction with MICCAI 2017 Quebec City, QC,...*, vol. 2017, pp. 240–248, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:21957663

[136] W. Crum, O. Camara, and D. Hill, "Generalized Overlap Measures for Evaluation and Validation in Medical Image Analysis," *IEEE Transactions on Medical Imaging*, vol. 25, no. 11, pp. 1451–1461, nov 2006. [Online]. Available: http://ieeexplore.ieee.org/document/1717643/

[137] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297–302, jun 1945. [Online]. Available: http://www.jstor.org/stable/1932409

[138] P. Jaccard, "The Distribution of the Flora in the Alpine Zone," *The New Phytologist*, vol. 11, no. 2, pp. 37–50, jun 1912. [Online]. Available: http://www.jstor.org/stable/2427226

[139] A. Depeursinge, J. Fageot, and O. S. Al-Kadi, "Fundamentals of Texture Processing for Biomedical Image Analysis," in *Biomedical Texture Analysis*. Elsevier, 2017, pp. 1–27. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/B9780128121337000016

[140] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv:1312.4400*, 2013.

[141] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," *arXiv preprint arXiv:1603.07285*, 2016.

[142] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[143] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," feb 2015. [Online]. Available: http://arxiv.org/abs/1502.01852

[144] A. Amidi and S. Amidi, "Deep Learning Tips and Tricks cheat-sheet." [Online]. Available: https://stanford.edu/{~}shervine/teaching/cs-230/cheatsheet-deep-learning-tips-and-tricks

[145] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.

[146] M. Yeung, E. Sala, C.-B. Schönlieb, and L. Rundo, "Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation," *Computerized Medical Imaging and Graphics*, vol. 95, p. 102026, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0895611121001750

[147] M. Pandey, M. Fernandez, F. Gentile, O. Isayev, A. Tropsha, A. C. Stern, and A. Cherkasov, "The transformational role of GPU computing and deep learning

in drug discovery," *Nature Machine Intelligence*, vol. 4, no. 3, pp. 211–221, 2022. [Online]. Available: https://doi.org/10.1038/s42256-022-00463-x

[148] H. Kim, H. Nam, W. Jung, and J. Lee, "Performance analysis of CNN frameworks for GPUs," in *2017 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, 2017, pp. 55–64.

[149] N. developer, "NVIDIA cuDNN." [Online]. Available: https://developer.nvidia.com/cudnn

[150] C. Li, Y. Yang, M. Feng, S. Chakradhar, and H. Zhou, "Optimizing Memory Efficiency for Deep Convolutional Neural Networks on GPUs," in *SC '16: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2016, pp. 633–644.