Robust Brain Tumor Segmentation for Overall Survival Prediction

Rupal Rashmikant Agravat

A Dissertation Submitted to the Faculty of Ahmedabad University in Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

December 2020

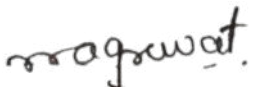
Copyright

Declaration

This is to certify that

1. The dissertation titled, "Robust Brain Tumor Segmentation for Overall Survival Prediction" comprises my original work towards the degree of Doctor of Philosophy at Ahmedabad University and it has not been submitted elsewhere for a degree.

2. Due acknowledgment has been made in the text to all other material used.



Rupal Rashmikant Agravat (135003)

Certificate

This is to certify that the dissertation work titled “Robust Brain Tumor Segmentation for Overall Survival Prediction” in Information and Communication Technology has been completed by Rupal Rashmikant Agravat (135003) for the degree of Doctor of Philosophy at the School of Engineering and Applied Science of Ahmedabad University under our supervision.

Signature of Dissertation Supervisor:

Name of Dissertation Supervisor:

Mehul S. Raval,
Associate Dean - Experiential Learning and Professor,
School of Engineering and Applied Science,
Ahmedabad University.

Signature of Dissertation Co-supervisor:

Name of Dissertation Co-supervisor:

Sanjay R. Chaudhary,
Interim Dean and Professor,
School of Engineering and Applied Science,
Ahmedabad University

Dedication

To my parents

To my son, Nihit

To my friends

Abstract

A Glioma brain tumor is the deadliest brain tumor with high mortality. Treatment planning by human experts depends on a proper diagnosis based on physical symptoms along with a medical image analysis. Highly variability of a brain tumor in terms of size, shape, location, and a high volume of Magnetic Resonance Images makes the analysis time-consuming. Such manual diagnosis is non-reproducible and depends highly on human expertise. Automatic/semi-automatic segmentation methods achieve a reduction in time with excellent reproducible results. Further, proper analysis follows the prediction of the overall survival of patients. The prediction helps in planning post-operative treatments which improve the prognosis. This dissertation develops fully convolutional neural networks for tumor segmentation followed by a random forest regressor to predict the overall survival of the patients.

The first network to address the segmentation problem is the 2D encoder-decoder network that segments the substructures of a tumor separately. A network parameter initialization technique addresses the class imbalance problem and scarcity of the label data. The network is extended to incorporate dense connections between successive layers of the encoder part, which helps the network learn diversified features from the input. In addition, it helps the error gradient to reach the initial layers in case the deep layers are not able to learn from the input. The network is further extended to incorporate a three-dimensional voxel relationship for consistent performance on unseen data.

Tumor segmentation follows classification of patients based on survival days. The classical approaches - random forest classifier and random forest regressor addresses the problem to calculate survival days based on an optimal set of input features, which are extracted from the segmentation network.

Acknowledgements

I would like to take this opportunity to express my gratitude to my guide, my guru, Dr. Mehul S. Raval for his continuous guidance right from day one till the end. I fall short of words to express the way he has guided me throughout; from helping me push my limits to the smallest grammatical correction or boosting my spirits during tough times to monitor my technical progress. He has been like a lamp in the darkness showing me the path on my Doctorate journey. Without his support I would not have acquired the heights I did. It was his constant motivation that has now brought me so far. I sincerely thank him for guiding me professionally and personally with patience, faith, and understanding.

I would like to express my thanks go to my co-supervisor Professor Sanjay Chaudhary for his continuous help and support whilst undertaking this research. I am grateful for his valuable and patient guidance. He has been a great mentor and guide who helped me learn 'how to learn'. I would also like to express my thanks to him for providing me the liberty to take decisions. He provided a healthy and flexible environment during this doctoral research endeavour.

I would like to express my thanks to my Thesis Advisory Committee members Dr. Manjunath Joshi and Dr. N. Padmanabhan. They helped me greatly at each stage of my research by sharing their knowledge, skills, and experience.

I would like to thank my brother Dr. Krutarth Agravat (Medical Officer, Essar Ltd) for clearing my doubts related to medical terminology used for brain tumor. I would also like to thank Mr. Ujjawal Baid for his fruitful discussion in network selection.

I would like to thank NVIDIA Corporation for donating the Quadro K5200 and Quadro P5000 GPU used for this research. My heartfelt thanks to Mr. Rahul Bhathiji and Mr. Adil Shaikh for facilitating the execution of my work on GPUs in the most adverse pandemic situation.

I would like to acknowledge the School of Engineering and Applied Science for providing resources for the experiments. I thank all faculties, staff members, and PhD students of the school for making my research journey more comfortable

and pleasant. In addition, I like to thank Prof. Dharmesh Varade, Chair, PhD programme, and Ms. Isha Ayachit, Administrative Officer for their admistrative support during this journey.

I really appreciate the great support and love which I have got from my son Nihit. He helped me a lot at his level during this journey. I owe special thanks to my husband Amit for his love, care, and understanding. He has been a great companion who gave me all the freedom and comfort to explore myself personally and professionally. I would like to convey my sincerest thanks to my mother-in-law, Mrs. Induben Kapdi, for her total support and encouragement in the pursuit of my academic goals. I would like to thank my sister Jital for pushing me up during the emotional down times in this journey. I would also like to express my deepest thanks to my mother and father. Without their blessings this journey could not have been completed successfully.

Contents

List of Figures **xi**

List of Tables **xv**

1 Introduction **1**

1.1 Brain Tumor 1

1.1.1 Brain Tumor Grades [1] 2

1.2 Medical Image Modalities 4

1.2.1 MRI Physics [2] 4

1.2.2 MRI Modalities 7

1.3 Brain Tumor Treatment 10

1.4 Motivation of the Research 11

1.5 Research Problems and Objectives 12

1.6 Outline of the Dissertation Report 12

2 Preliminaries **14**

2.1 Pre-processing Techniques 14

2.1.1 Image Registration 14

2.1.2 Skull-Stripping 15

2.1.3 Bias Field Correction 15

2.1.4 Intensity Normalization 16

2.2 Dataset 17

2.2.1 Segmentation Task 17

2.2.2 Overall Survival Prediction Task 20

2.3 Post-processing techniques 21

2.3.1 Connected Component Analysis 21

2.3.2 Conditional Random Field 21
2.3.3 Morphological Operations 21
2.4 Evaluation Metrics 21
2.5 Summary 23

3 Literature Review **24**
3.1 Conventional Brain Tumor Segmentation Methods 24
3.1.1 Basic Methods 24
3.1.2 Generative Methods 25
3.1.3 Discriminative Methods 27
3.2 Initial Implementation: K-means Clustering 29
3.3 Limitation of Conventional Methods 30
3.4 Deep Neural Network 31
3.4.1 Evolution of DNN 32
3.4.2 Design Aspects of CNN 33
3.4.3 CNN Architectures 34
3.4.4 Handling Class Imbalance Problem 35
3.4.5 CNN Approaches for Brain Tumor Segmentation 36
3.4.6 Brain Tumor Segmentation using CNN and FCNN 38
3.4.7 Overall Survival Prediction 40
3.5 Summary 42

4 The Development of New Encoder-Decoder Architecture **43**
4.1 Hardware and Software Specifications for Algorithm Development . 43
4.2 Fully Convolutional Neural Network 44
4.3 2D Encoder-Decoder FCNN 45
4.4 2D Encoder-Decoder Dense FCNN 50
4.5 3D Encoder-Decoder Dense FCNN 57
4.5.1 Pre-processing 58
4.5.2 Post-processing 58
4.5.3 Results 59
4.6 Summary 65

5 Overall Survival Prediction **67**
5.1 Radiomics Features 67
5.2 Random Forest Classifier 69
5.3 Random Forest Regressor 70
5.3.1 OS Prediction for BraTS 2019 72
5.3.2 OS Prediction for BraTS 2020 73
5.4 Summary 75

6 Conclusion and Open Research Challenges **76**
6.1 Conclusion 76
6.2 Open Research Challenges 77

Bibliography **79**

Appendices **97**

A List of Abbreviations **97**

B Comparison: Conventional Methods **100**

C Comparison: CNN/FCNN Methods **102**

D Comparison: Classifiers for OS Prediction **109**

List of Figures

1.1 Location based tumor classification [3]. 2

1.2 Healthy brain (A)Medical Image (B) GM (C) WM (D) CSF [4]. . . 3

1.3 Hydrogen spin in the presence of a magnetic field [5]. 5

1.4 Illustration of excitation and relaxation in presence\absence of RF pulse at 90°. 6

1.5 MRI Images (A) T_2 (B) T_1 (C) T_1c (D) FLAIR [6, 7, 8]. 9

2.1 Pre Processing: Image Registration [9] (A) Original Slice (B) Rotated Slice (C) Registered Slice. 15

2.2 Pre-processing: Skull-stripping (A) Slice with skull (B) Slice with skull-stripped [10]. 16

2.3 Pre-processing: Bias Field Correction (A) Original Slice (B) Bias field (C) Corrected Slice [11]. 17

2.4 The appearance of intra-tumoral structures on three imaging modalities with manual annotations. (A) Top: whole tumor (yellow), Bottom: FLAIR, (B) Top: tumor core (red), Bottom: T_2, (C) Top: enhancing core structures (light blue), surrounding the cystic/necrotic core structure (green), bottom: T_1c, (D) Fusion of the three labels [12]. 20

3.1 (A) FLAIR image slice, (B) ground truth slice, (C) K-means segmentation slice, (D) 6-neighborhood post-processed slice, (E) 18-neighborhood post-processed slice, (F) 26-neighborhood post-processed slice. 29

3.2 General pipeline for DNN tumor segmentation. 32

3.3 CNN design approaches. 37

4.1 The proposed 2D encoder-decoder FCNN. 46

4.2 Segmentation output: (A) T2 image slice, (where, Edema is in yellow, Enhancing core is in blue, Necrotic/Non-enhancing core is in green) (B) WT ground truth, (C) WT segmentation, (D) Edema ground truth, (E) Edema segmentation without weight initialization, (F) Edema segmentation with weight initialization, (G) Enhancing tumor ground truth, (H) Enhancing tumor segmentation without weight transfer, (I) Enhancing tumor segmentation with weight transfer, (J) Necrotic ground truth, (K) Necrotic segmentation without weight transfer, (L) Necrotic segmentation with weight transfer. 49

4.3 The proposed 2D encoder-decoder dense FCNN. 50

4.4 Correct segmentation result (A) FLAIR slice, brown-necrosis, gray-enhancing tumor, orange-edema (B) Ground truth slice (C) Segmentation slice. 52

4.5 Incorrect segmentation result (A) FLAIR slice, orange-necrosis, gray-edema (B) Ground truth slice (C) Segmentation slice. 53

4.6 DSC comparison for training and validation set (A) whole tumor (B) tumor core (C) enhancing tumor. 56

4.7 The proposed 3D encoder-decoder dense FCNN. 58

4.8 Image pre-processing (A) FLAIR slice (B) Bias field correction (C) Bias field correction with denoising. 59

4.9 The box plot (A) DSC (B) Hausdorff95 (C) Sensitivity (D) Specificity. 61

4.10 Correct segmentation results of the network, brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) Segmentation without post-processing (D) Segmentation after post-processing. 62

4.11 Incorrect segmentation results of the network, brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) Segmentation without post-processing (D) Segmentation after post-processing. 63
4.12 Brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) 2D FCNN segmentation (D) 3D FCNN segmentation. 63
4.13 Brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) 2D FCNN segmentation (D) 3D FCNN segmentation. 64
4.14 Brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) 2D FCNN segmentation (D) 3D FCNN segmentation. 64
4.15 Orange-necrosis, gray-enhancing tumor (A) FLAIR slice (B) Ground truth (C) 2D FCNN segmentation (D) 3D FCNN segmentation. 65

5.1 Patients' OS days. 68
5.2 Training AUC for various feature sets. 71
5.3 Validation AUC for various feature sets. 71
5.4 Visualization of actual and predicted survival days. 74

List of Tables

2.1 Growth of the BraTS dataset [6, 7, 8]. 19
2.2 The distribution of BraTS dataset [6, 7, 8] features in survival classes. 21

3.1 2D slice DSC for BraTS 2015 dataset. 30
3.2 The proportion of Non Tumor Brain Volume (NTBV) and Non Tumor Background Volume (NTBGV) vs. Tumorous Volume(TV). . . 36

4.1 Test data set evaluation for DSC, sensitivity and positive predictive value (PPV), where A: model without weight initialization, B: model with weight initialization. 48
4.2 BraTS 2019 training set evaluation for DSC and HD95. 52
4.3 BraTS 2019 training set evaluation for sensitivity and specificity. . . 53
4.4 Comparison of DSC and HD95 with state-of-the-art methods. 53
4.5 BraTS 2019 validation set evaluation for DSC and HD95. 54
4.6 BraTS 2019 validation set evaluation for sensitivity and specificity. . 54
4.7 BraTS 2019 test set evaluation for DSC and HD95. 54
4.8 Model variations of 2D encoder-decoder dense FCNN. 55
4.9 BraTS 2020 training set evaluation for DSC and HD95. 60
4.10 BraTS 2020 training set evaluation for sensitivity and specificity. . . 60
4.11 BraTS 2020 validation set evaluation for DSC and HD95. 60
4.12 BraTS 2020 validation set evaluation for sensitivity and specificity. . 60
4.13 BraTS 2020 test set evaluation for DSC and HD95. 61
4.14 DSC comparision with state-of-the-art ensemble methods. 65

5.1 RFC accuracy for various feature sets. 70

5.2 OS accuracy for training, validation and test datasets. 72
5.3 OS comparison with state-of-the-art methods on the BraTS 2019 validation set. 73
5.4 OS accuracy for BraTS 2020 training, validation and test datasets. . 73
5.5 OS accuracy comparison with state-of-the-art methods on the leaderboard of BraTS 2020 for validation set. 74

6.1 OS accuracy on BraTS datasets. 77

B.1 Summarization of conventional brain tumor segmentation methods. 100

C.1 Summarization of segmentation methods using CNN architectures. . 103

D.1 Comparison of OS prediction on the BraTS validation dataset. . . . 109

Chapter 1

Introduction

1.1 Brain Tumor

A tumor of the Central Nervous System (CNS) is a mass formed due to out-of-control brain cell growth. CNS tumors represent the 17th most common cancer with an estimated 297000 cases of CNS cancers worldwide in 2018 [13]. Due to the high case fatality rate, CNS tumors are the 12th most frequent cause of cancer related deaths worldwide [14]. The etiological study of CNS tumors is challenging because of its low incidence rate and high heterogeneity.

Brain cells begin to form tumors due to their genes, environmental carcinogens, other medical conditions like asthma and eczema. Tumor growth pushes the brain's key area, blocking the flow of fluid around the brain. This blockage increases pressure inside the skull and damages tissues in the brain. A tumor in a critical area of the brain can cause severe neurological problems. The spinal fluid plays an important role in the spread of some of the tumors to the distant areas of the brain or the spinal cord. It leads to categorization of the tumor in the following ways:

- A primary brain tumor, which is abnormal brain tissue growth starting in the brain.
- Secondary (Metastatic) tumor where cancerous growth in other body parts spreads to the brain and forms a new tumor.

A primary benign tumor is less dangerous as it does not grow, and proper treatment may be able remove it. On the other hand, a primary cancerous tumor is

more dangerous as it proliferates and may spread to other parts of the brain or the spinal cord. A secondary brain tumor is always cancerous. The tumor results in different symptoms depending on its size and location in the brain. Symptoms may include headaches, dizziness, vomiting, seizures, drowsiness, balance problems, and numbness in the arms, legs, or face.

More than 120 tumors are known, and their nomenclature is based on their location in the brain and the part where they begin to grow [15]. Types of tumors based on the location are as shown in Figure 1.1. Some of the more common brain tumors include Optic Glioma, Astrocytoma, Ependyoma, Meningioma, Oligodendroglioma, Pituitary Tumor, Medulloblastoma, Ependymoma, Craniopharyngioma Brain Stem Glioma.

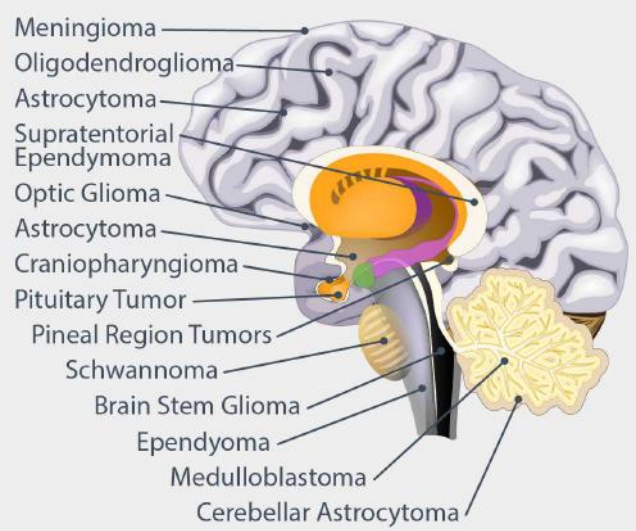


Figure 1.1: Location based tumor classification [3].

1.1.1 Brain Tumor Grades [1]

The grade of the tumor decides its severity and helps doctors in treatment planning. The grade is determined from the appearance and structure of a tumor cell under a microscope. Tumors are divided into the following four grades:

- Grade I. Benign and slow-growing tumor with similar tissue structures as the normal tissues under the microscope. Grade I tumors do not come back after surgical removal and this increase the life span of the patient.

- Grade II. It is a benign tumor similar to grade I, but it can reappear after the surgery.
- Grade III. It is a malignant tumor with different tissue structures than normal ones. It may come back after surgery with more severity and may spread within the brain.
- Grade IV. The most malignant tumor with exponential growth and spread rate which has significantly fewer chances to be cured.

A healthy brain contains three type of tissues namely, 1) Gray Matter (GM) tissues, 2) White Matter (WM) tissues, and Cerebrospinal Fluid (CSF). An image of a healthy brain is shown in the Figure 1.2. The tumor in the brain creates an additional tissue class. The tumor intensity overlaps with normal brain tissues and in some cases the tumor is indistinguishable from the normal brain tissues. Experts analyse medical images of the brain tumor for treatment planning.

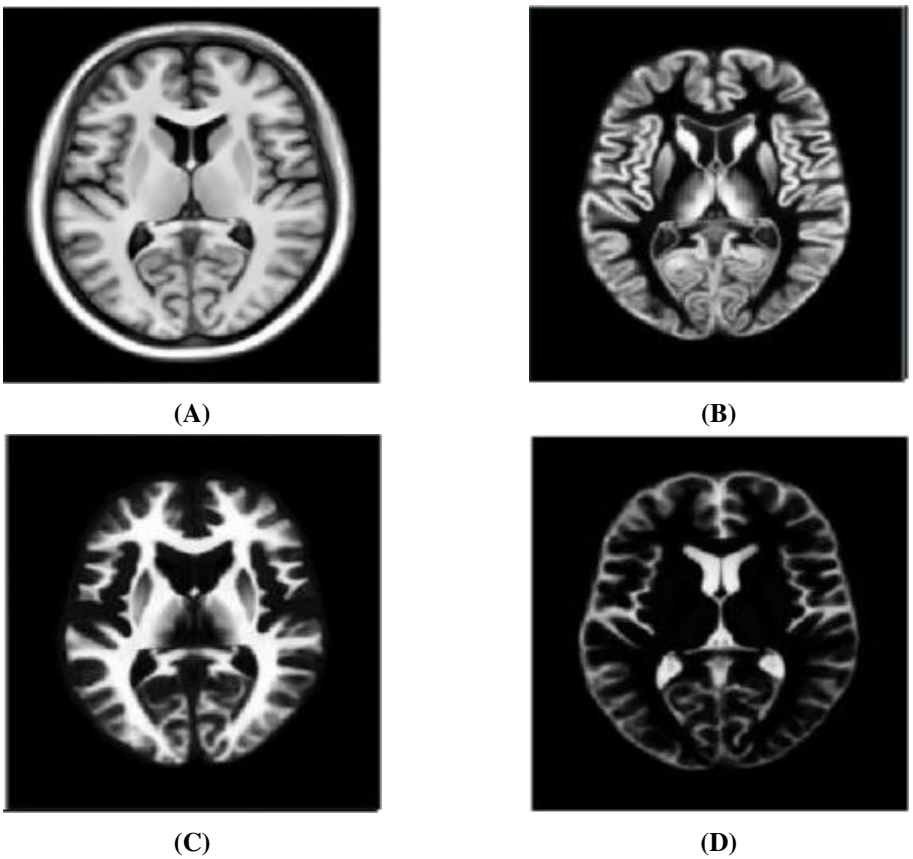


Figure 1.2: Healthy brain (A)Medical Image (B) GM (C) WM (D) CSF [4].

1.2 Medical Image Modalities

Computed Tomography (CT) as well as Magnetic Resonance Imaging (MRI) are two medical imaging techniques that are used most often in the diagnosis as well as the formulation of the treatment plan for brain tumor.

CT imaging that is non-invasive makes use of X-ray beams. These beams rotate around the head of the patient and produce 2D images at specific angles taken from the sagittal, axial and coronal views. A 3D tomographic image is constructed making use of 2D images. Ionizing radiation which is utilized in the process of CT imaging causes harm to the human body. It has a reaction to dyes that are also used during CT imaging.

However, MRI is also extensively utilized for taking images of the brain. As far as Glioma tumors are concerned, MRI imaging is the most preferred technique of observation and is extremely popular. This process makes use of the magnetic property of the hydrogen nuclei, which the human body contains in huge amounts. An MRI image provides detailing of the delicate tissues as well as high contrast between tissues. MRI imaging is more expensive than CT scans, although it provides higher accuracy. 3D images are also supported by MRI imaging in three directions: axial, sagittal and coronal.

1.2.1 MRI Physics [2]

Nuclear Magnetic Resonance (NMR) is the fundamental principle upon which the MRI technology is based. NMR is utilized to examine molecules in spectroscopy.

Hydrogen nuclei, in the form of water exist in the human body. They contain properties that are magnetic. These magnetic properties known as the nuclear spin are what the MRI depends on to function. An electromagnetic field is generated by the hydrogen nuclei when they spin around their axes. As a result of the random spin, the entire magnetic field of the nuclei is null.

Figure 1.3 shows that these hydrogen nuclei spin in one of two ways: either positive or the same direction as the field, or negative or the opposite direction of the field when an external magnetic field B_0 is present. The spins precess at a Larmor

frequency ω_0 around the axis of the magnetic field. This frequency is in proportion to the external field. It is shown as:

$$\omega_0 = \frac{\lambda * B_0}{2 * \pi} \tag{1.1}$$

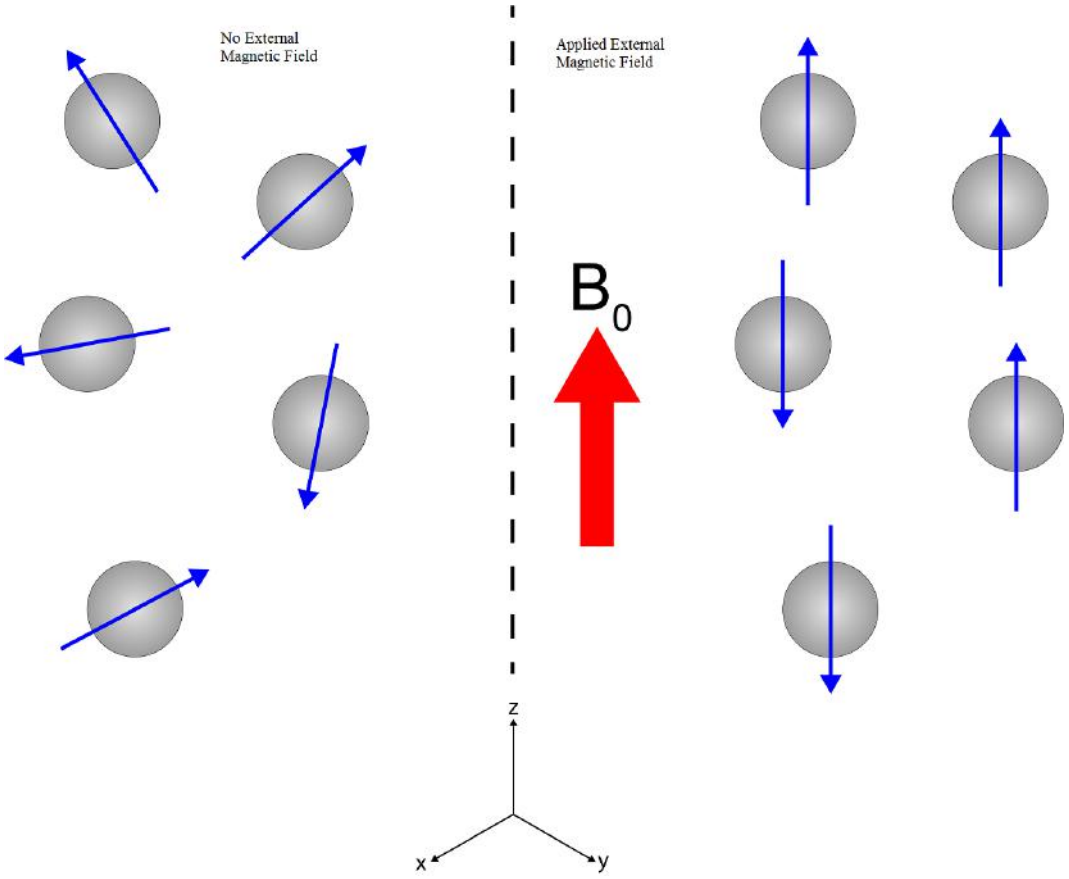


Figure 1.3: Hydrogen spin in the presence of a magnetic field [5].

A longitudinal component M_z, which is parallel to B_0, and a transverse component M_{xy}, which is perpendicular to B_0 express the magnetic vector of the nuclei that are spinning. The nuclei spinning towards the external magnetic field are slightly higher in number than the nuclei spinning in the opposite direction. The result of this is longitudinal magnetization M_z. Null transverse magnetization occurs as the outcome of this out of phase spin precession.

B_1, which is a radiofrequency (RF) pulse and is perpendicular to an external magnetic field B_0, is supplied at the resonance frequency ω_0. As a result of this, the nuclei absorb the energy which in turn causes three things to occur namely: 1)

The nuclei which are spinning in the same direction to the magnetic field change their direction which is opposite to the magnetic field direction, 2) the second event that transpires is that there is in-phase spin precession, and 3) the net magnetization vector tilts away from the longitudinal axis to an axis that is transverse causing the longitudinal magnetic vector to become null at resonance as well as causing the transverse signal to be the strongest possible signal.

The nuclei return to their state of equilibrium because they retransmit the electromagnetic energy that they have absorbed when the RF pulse is turned off and hence, form an NMR signal. Figure 1.4 shows the two ways in which the phenomenon of relaxation can be observed. The two ways are longitudinal and transversal.

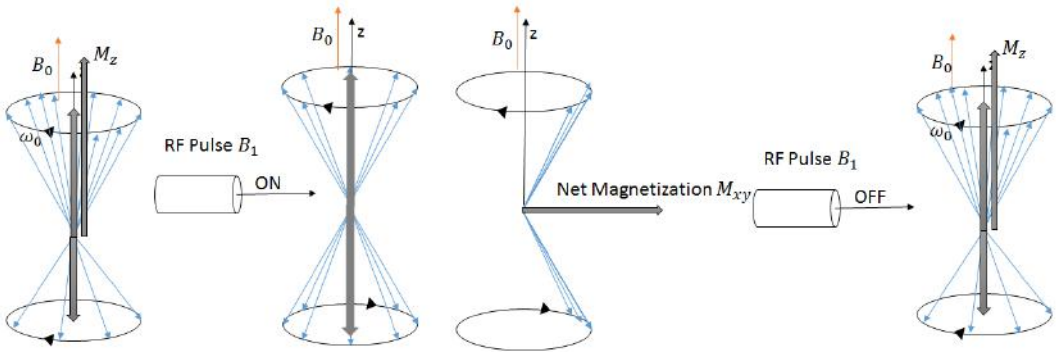


Figure 1.4: Illustration of excitation and relaxation in presence\absence of RF pulse at 90°.

Longitudinal relaxation: Also known as spin-lattice interaction is an occurrence where the nuclei spin gives away the lattice energy and thereby returning to their state of equilibrium. Also, they come into alignment with the magnetic field B_0. Recovery of magnetism that is longitudinal adheres to a a tissue-specific exponential curve with T_1 as the time constant. Beyond this, there is recovery of the longitudinal magnetism up to 63% of the final value. This is shown in Equation 1.2 as:

$$M_z = M_0(1 - \exp(-\frac{t}{T_1})) \tag{1.2}$$

M_0 is the net magnetization at equilibrium. It relies on the density of the proton and the strength of the external magnetic field.

Transverse relaxation: Also known as spin-spin interaction relates to the precession of the nuclei that is out of phase. The recovery of the transverse magnetism follows an exponential curve that is tissue-specific with T_2 as the time constant as shown in Equation 1.3 below:

$$M_{xy} = M_{xy0} \exp(-\frac{t}{T_2}) \tag{1.3}$$

M_{xy0} is the amplitude of the transverse signal following the RF pulse. M_{xy0} is equal to M_0 whenever the RF pulse is perpendicular to B_0. In practice, T_2 is always smaller than T_1. It is these two constants that are very significant to the definition and contrast of the images of the various MRI sequences.

There are two main parameters that are basic to an MRI sequence. Echo Time (TE) is the primary one. This is the time between the measurement of the signal and the RF pulse. Time that is in between the two RF pulses is the second parameter of the MRI sequence. It is known as Repetition Time (TR). During relaxation, subsequent to every RF pulse, the NMR signal is received and how the transverse component evolves is displayed by it. The signal oscillates at resonance frequency and its envelope decays exponentially. There is a correlation between the magnitude of the longitudinal component and the initial amplitude which impacts the status of its recovery. In the event that the longitudinal component recovers well, the signal will be substantial.

1.2.2 MRI Modalities

The response time of various brain tissues along with the response of the tissues to chemicals injected in the patient generate various modalities.

- T_2-weighted image: T_2 refers to the time taken by hydrogen nuclei to possess their original precession phase which is disturbed by the presence of radiofrequency signal. T_2-weighted images which are also known as T_2 images for short, have more sensitivity to water content. Hence, the pathology and the CSF appear to be more hyper-intense in these images.

- T_1-weighted image: During the MRI imaging process, when the hydrogen nuclei that are present in a tissue have to go back to the initial magnetization state that is given by the static magnetic field, it is referred to as T_1. Anatomical details are better provided by simple T_1 weighted images (also known as T_1 images for short) rather than T_2-weighted images. However, T_1 images do not generate adequate information to support brain tumor investigation. Nevertheless, they are made use of along with an injection of a contrast agent fluid into the vascular system of the patient. The flow of the blood is highlighted in the T1-weighted images because of the contrast agent. The area of the tumor that is active along with the blood vessels seem to be hyper-intense and can be distinguished with ease from the tissues that surround the area. The presence of an active tumor is often an indication of malignancy to the experts during a tumor investigation. MRI images like this are known as contrast-enhanced T_1-weighted images (T_1c).
- FLAIR image: Fluid-Attenuated Inversion Recovery (FLAIR) is a sequence that causes suppression of fluids and hence, in brain imaging it is utilized for suppression of CSF. The construction of this sequence is undertaken by flipping the longitudinal magnetic field M_z to begin with. It is flipped in the opposite direction making use of an extra electromagnetic pulse that is oriented at 180 degrees from the original field B_0. A longitudinal relaxation comes next after the pulse in order to go back to equilibrium, passing through the null value. The addition of a time constraint inversion time decides when the perpendicular RF pulse should be added. The signal of specific tissues can by suppressed by opting the inversion time that corresponds to when the magnitude of the signal is null. When using the FLAIR sequences, the setting of inversion time is undertaken for the suppression of the high CSF signal that occurs during T_2 imaging. Currently, in the visualization of brain tumors, the modality that shows the best contrast is the FLAIR sequence. In T_2 images, lesions remain hyper-intense and CSF becomes hypo-intense. In order to distinguish the lesions from the CSF, the FLAIR sequence is utilized. It is for

this reason that this sequence is used extensively in brain tumor imaging.

Figure 1.5 compares T_2, T_1, T_1c and FLAIR images with the existence of a tumor. One can observe the hyper-intense active tumor shown in the T_1c image, hyper-intense tumor and edema that are shown in the T_2 and FLAIR images, and the hypo-intense CSF in the FLAIR image.

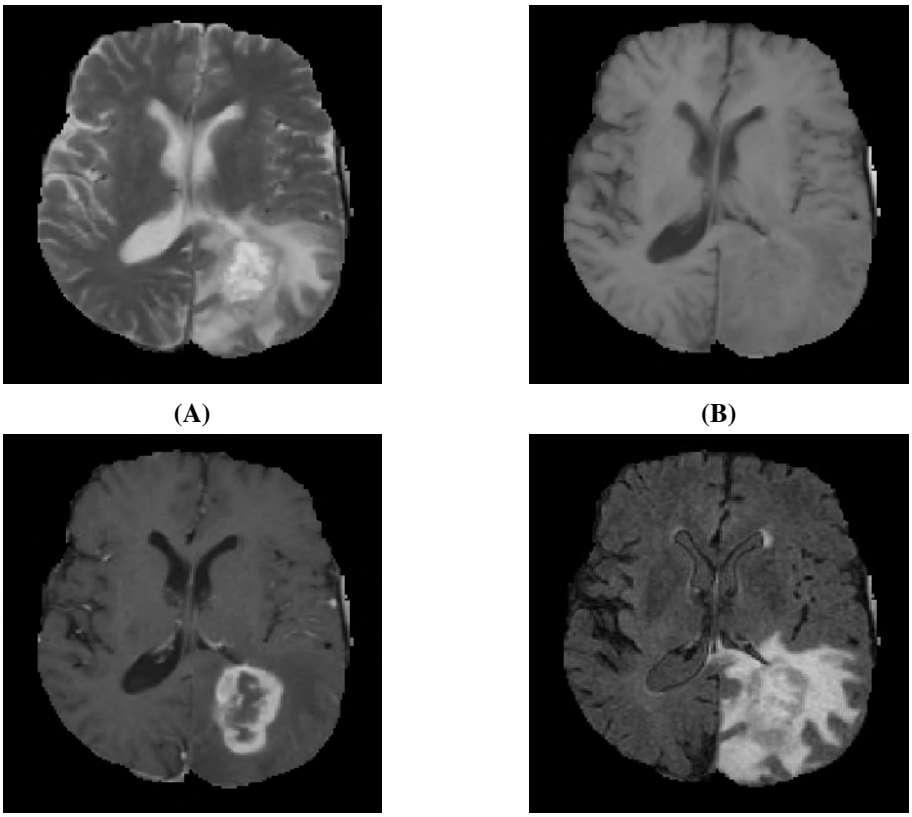


Figure 1.5: MRI Images (A) T_2 (B) T_1 (C) T_1c (D) FLAIR [6, 7, 8].

Lately newer techniques of MRI imaging such as Diffusion Tensor Imaging (DTI) have been developed. It is this modality that enables reconstruction of the white matter tracts connecting the different parts of the neural networks of the brain. This is done by taking a measurement of the anisotropic diffusion of water that is inside the tissues. The existence of a tumor can have a direct impact on the structure of the fiber by creating disruption, displacement or infiltration. Additional information used in diagnosing and studying tumors in the brain is provided by modalities like the MR spectroscopy which measures significant metabolites in the

tissue of the tumor and perfusion MRI which measures the relative cerebral blood volume making use of a contrast agent.

1.3 Brain Tumor Treatment

The tumor treatment recommendation depends on the its size and type, growth rate, and the patient's general health. Treatment options include:

- Surgery(surgical resection): This involves the removal of a part of the tumor or the total tumor from the skull which in turn eases the pressure on the brain and alleviates the symptoms of the tumor.
- Radiation therapy: Traditional therapy involving radiation makes use of external beams of gamma rays, X-rays or protons that are directed towards the tumor to eliminate cancerous cells and reduce the size of brain tumors. Such radiation therapy can be undertaken in a single treatment or several treatments that are spread out over several weeks.
- Targeted biological therapy(immunotherapy): This treatment involves making use of the body's own defense mechanism, namely, the immune system to battle cancer. This type of treatment relies on markers on the surface of tumor cells being identified.

Any of the above options can be used in tandem with the other options.

Medical image analysis is expert-dependent therefore it is subject to intra observer variability. Algorithms specifically developed for brain tumor investigation can identify the features of the tumor and delineate it in complete three-dimensional (3D) image. The aims of any computer-based automated analysis are as follows:

1. The results that are generated by the algorithms must be reproducible and consistently accurate.
2. The volume of the data should not overwhelm the algorithm or the computer.
3. Using a Graphical Processing Unit (GPU) along with algorithms helps accelerate the process of analysis.
4. Automated segmentation of the tumor reinforces human experts. This leads to appropriate treatment-planning as well as follow-up.

5. Data of this type can be captured at any remote location, but the processing can be done at a centralized location.

1.4 Motivation of the Research

Human experts (neurologists, radiologists) analyze and interpret volumetric brain MRI images to segment various brain tissues and locate the tumor. This analysis is time-consuming. Besides, this type of segmentation is non-reproducible. Accuracy of brain tumor segmentation with utmost precision is desirable to plan proper treatment like medication or surgery, and depends on human experts to a very great extent. Computer-aided analysis helps a human expert locate the tumor in less time and regenerates the analysis results. The intended analysis by computerized methods requires appropriate input with correct working methods. However, input to the method may face the following challenges:

- Low Signal to Noise Ratio (SNR) and artifacts in raw MRI data are mainly due to electronic interferences in the receiver circuits, radiofrequency emissions due to the thermal motion of the ions in the patient's body and coils, and electronic circuits in MRI scanners. This random fluctuation reduces the image contrasts due to signal-dependent data bias [16].
- Non-uniformity is an irrelevant additional intensity variation throughout the MRI signal. Possible causes of non-uniformity are radiofrequency coils, acquisition pulse sequence and the geometry and nature of the sample.
- Unwanted information acquired by the MR machines is information on the skull, fat, and the skin.
- The intensity profile of MR images may vary due to a variety of MRI machine configurations.
- Publicly available brain tumor images for computer-aided analysis are significantly less. The collection of MR images from various hospitals has privacy or confidentiality issues.
- Class imbalance problem is another major issue in medical image analysis. Abnormal class images might be challenging to find because abnormal classes are rare compared to normal classes.

1.5 Research Problems and Objectives

Research Problems

- Design an end-to-end fully convolutional neural network for robust brain tumor segmentation.

  This thesis aims to design a fully convolutional neural network for the segmentation of Glioma brain tumors from 3D MRI images without human interaction and with less computation complexity.
- Predict the overall survival of High Grade Glioma (HGG) patients.

  The objective is to extract the discriminative features of tumor after segmentation to predict the life expectancy for better prognosis and further treatment planning.

Research Objectives

1. Designing of 2D encoder-decoder architecture for brain tumor segmentation.
2. Inductive transfer learning-based parameter initialization for subregion network training.
3. Designing of 3D encoder-decoder architecture for brain tumor segmentation.
4. Network training with a focal loss to weigh the hard samples for better network training in class imbalance.
5. Tumor features selection and patients' overall survival prediction.

1.6 Outline of the Dissertation Report

The thesis is divided into six chapters, as follows:

- Chapter 1 introduces the domain.
- Chapter 2 shows the pre-processing techniques, dataset, post-processing techniques, and evaluation metrics.
- Chapter 3 covers the literature review on non-deep learning and deep learning methods for tumor segmentation.
- Chapter 4 presents three proposed approaches for brain tumor segmentation.

- Chapter 5 demonstrates the overall survival prediction approach with the focus on feature selection.
- Chapter 6 includes the conclusion and future direction.

Chapter 2

Preliminaries

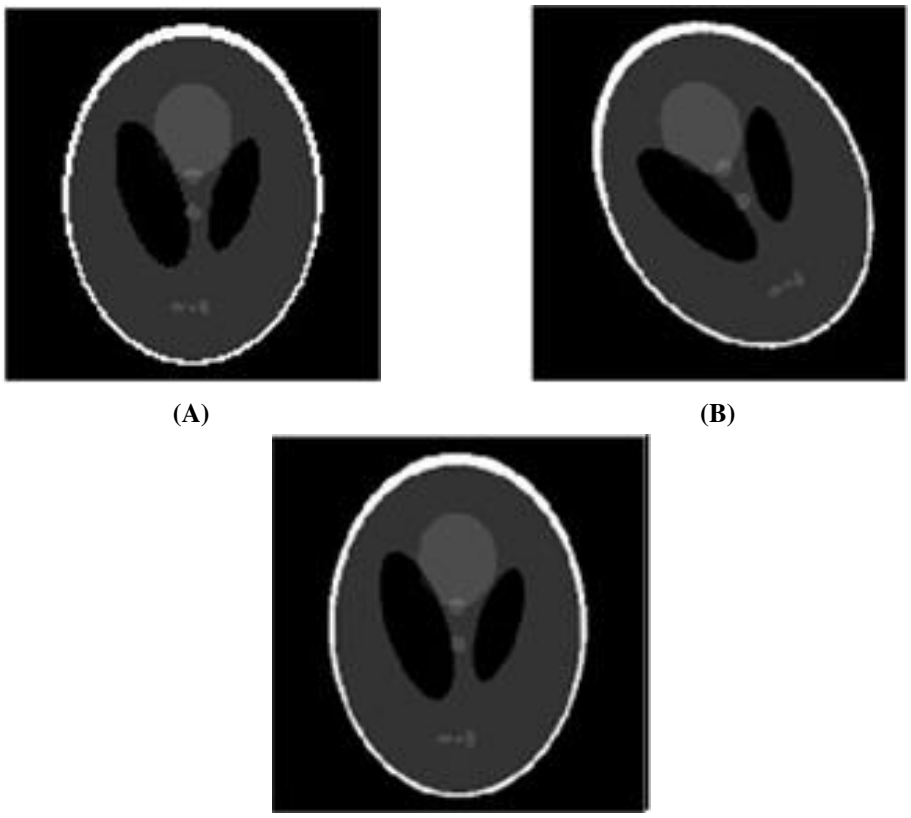
This chapter focuses on MRI image pre-processing and post-processing techniques, availability of benchmark datasets, and various evaluation measures.

2.1 Pre-processing Techniques

MRI image pre-processing techniques prepare data at the lowest level of abstraction, improving the image for segmentation. Pre-processing removes unwanted information from the input, cleans noisy images, resamples and orients the images with a reference image for feature extraction of the same location from multiple images. Pre-processing typically consists of: 1) image registration or image resampling, 2) bias field correction, 3) brain extraction or skull stripping, and finally 4) intensity normalization. In the case of publicly available datasets that are typically partially pre-processed, further pre-processing is carefully chosen based on the need.

2.1.1 Image Registration

The head movement of the patient at the time of acquisition leads to the need for image registration. Image registration resamples the image and orients it to the same image space. It helps to capture the features from the exact location of different images. Image registration is required in two cases: 1) orientation variation of the input images, 2) orientation variation of multiple modalities of the same input image. In the first case, all the images of the input are resampled to the reference anatomical template. Figure 2.1 illustrates the co-registered image with the reference image.



(A) **(B)**

(C)

Figure 2.1: Pre Processing: Image Registration [9] (A) Original Slice (B) Rotated Slice (C) Registered Slice.

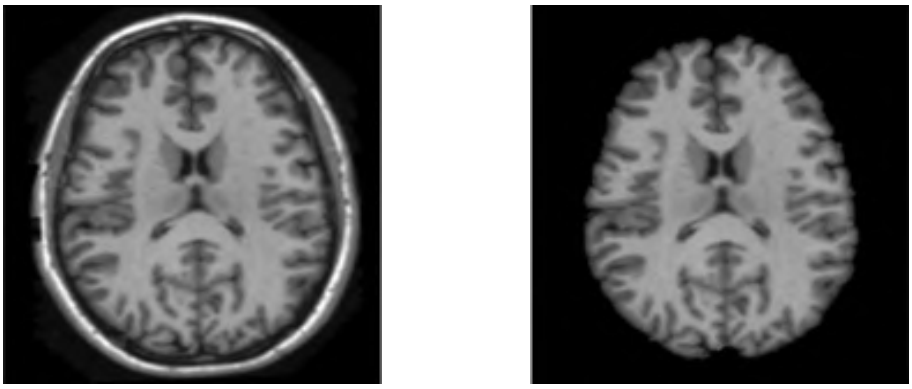
In the second case, the image co-registration is applied with affine registration, where initially all modality images are registered among themselves, followed by registration to the reference anatomical template.

2.1.2 Skull-Stripping

The brain image contains non-brain and unnecessary tissues of the head like skin, fat, skull. Skull-stripping eliminates such non-brain tissues from the MRI images and speeds up the prognostic and diagnostic procedures. Figure 2.2 shows the brain image with skull and with skull-stripping.

2.1.3 Bias Field Correction

The bias field is a low-frequency smooth intensity variation across the image, which changes the same tissue's appearance at different locations. Bias field can be introduced by 1) imperfect image acquisition process of scanner 2) anatomy especially



(A) **(B)**

Figure 2.2: Pre-processing: Skull-stripping (A) Slice with skull (B) Slice with skull-stripped [10].

due to shape, position, orientation, and magnetic permeability of the patient [11]. In general, the bias field is assumed to be additive or multiplicative. The additive form stems from the superposition of the magnetic field in the MRI. The multiplicative form comes from the sensitivity of the reception coils in the MRI [17]. The N3 bias field correction algorithm is an iterative scheme on the model based on B-spline least-squares fitting. The improvement to N3 bias field correction is N4 bias field correction [18], which iteratively calculates the residual bias fields instead of the total bias fields. N4 bias field correction employs multiresolution approximation with faster execution time. Figure 2.3 illustrates the N4 bias field correction of a FLAIR image, together with their extracted additive bias field.

2.1.4 Intensity Normalization

Individual voxel intensity in MR images does not reflect any physical quantities, and interpretation is possible within specific intensity distribution only. The intensity distribution of MR images varies depending on the scanning protocol and various other scanner specific parameters, which leads to the different intensity distribution of the same object. If segmentation or classification algorithms are based on voxel intensities and intensity distributions, it is necessary to standardize the intensity range. Intensity normalization pre-processing normalizes, scales, or aligns the voxel intensity distribution to the customized range. Typically used intensity normalization techniques are z-score normalization, intensity scaling, and

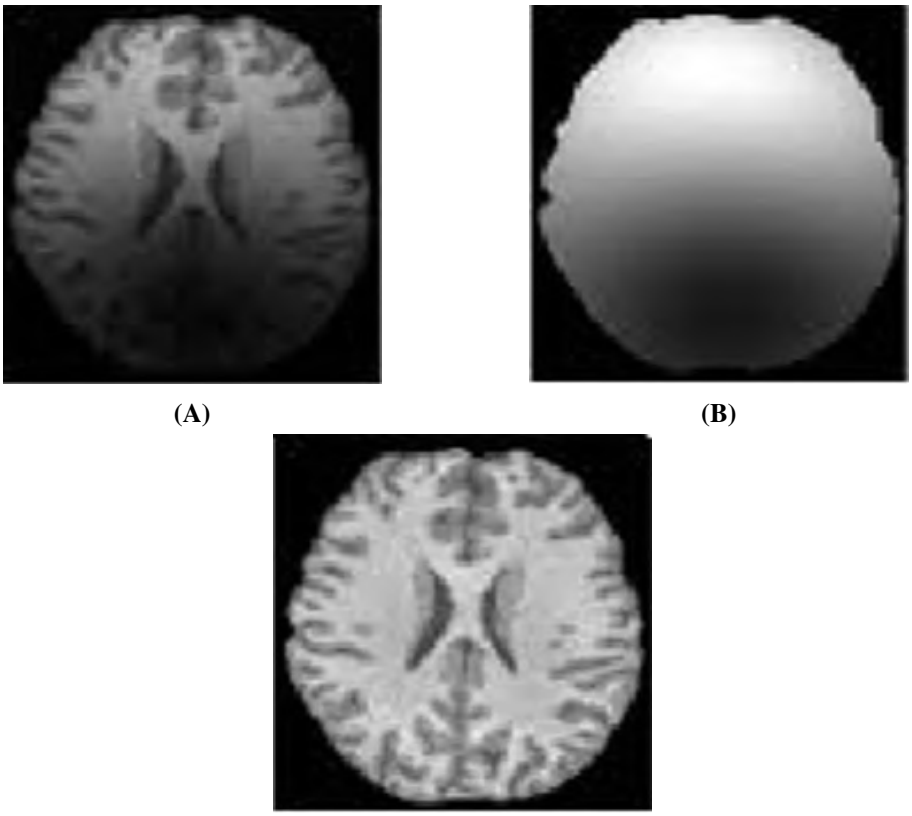


Figure 2.3: Pre-processing: Bias Field Correction (A) Original Slice (B) Bias field (C) Corrected Slice [11].

histogram-matching.

2.2 Dataset

2.2.1 Segmentation Task

Literature proposes various methods for brain tumor segmentation. All the methods claim their superiority and usefulness in some way. Initially, all such techniques work on the images taken from some hospitals or radiology laboratories, which are private. Disclosure of those images to the other researchers was not allowed, deterring a comparison of different methods. Publicly available datasets and evaluation frameworks compare and evaluate methods on the same measure. The Multimodal Brain Tumor Segmentation [6, 7, 8] challenge dataset, widely known as BraTS is one such publicly available dataset for brain tumor related challenges. This dataset widely used for segmentation task and overall survival prediction task. It contains

Low Grade Glioma (LGG) and High Grade Glioma images. The dataset has the following characteristics:

- It contains multi-parametric MRI scans in T_1, T_1c, T_2, and T_2-FLAIR volumes. Initially the scans were pre and post-operative. Later on the post-operative scans were omitted since 2014.
- The dataset contains images with 2D or 3D clinical protocols and 1.5T or 3T scanners from multiple institutions.
- The dataset set includes images with pre-processing for harmonization and standardization without affecting the apparent image information.
- The individual scan is co-registered to the T_1 reference anatomical template and interpolated to a uniform isotropic resolution ($1mm^3$) with skull-stripping.

Initially, the dataset's clinical images were significantly few. It was challenging to compare methods based on the results of such a small number of images. Comparison is possible with an increase in the number of sample images and accurate generation of the ground truth images. The ground truth is generated based on the evaluation by more than one expert to avoid inter-observer variability. Each scan has the dimension of 240x240x155 (Height (H) x Breadth (B) x Depth (D)) . The growth of the dataset from its inception is as shown in Table 2.1.

Four different intra-tumoral structures are useful for ground truth generation: edema, enhancing core, non-enhancing core, and necrotic core, as shown in Figure 2.4. The expert raters had annotated each case manually. Those segmentation results of each case were fused to obtain a single unanimous segmentation for the ground truth.

The segmentation task focuses on the proper segmentation of the tumor substructures. The validation of the segmentation methods is based on: 1) Whole Tumor (WT): all intra-tumoral substructures, 2) Tumor Core (TC): enhancing, necrotic, and non-enhancing core substructures, and 3) Enhancing Tumor (ET): enhancing core substructure.

Table 2.1: Growth of the BraTS dataset [6, 7, 8].

Year	Total Images	Training Images	Validation Images	Test Images	Tasks	Type of Data
2012	Clinical:45 Synthetic:65	Clinical data: 30(20HGG + 10LGG), Synthetic data: 50(25HGG + 25LGG)	N/A	Clinical data:15 Synthetic data:15	Segmentation	Pre and post operative scans
2013	65	51 HGG, 14 LGG	N/A	Clinical:25 Challenge: 10	Segmentation	Pre and post operative scans
2014	238	200	N/A	38	Segmentation, Disease Progression	Pre-operative, Longitudinal
2015	384	220 HGG, 54 LGG	N/A	110	Segmentation, Disease Progression	Pre-operative, Longitudinal
2016	465	220 HGG, 54 LGG	N/A	191	Segmentation, Disease Progression	Pre-operative, Longitudinal
2017	477	210 HGG, 75 LGG	46	146	Segmentation, Survival Prediction	Pre-operative, Longitudinal
2018	542	210 HGG, 75 LGG	66	191	Segmentation, Survival Prediction	Pre-operative, Longitudinal
2019	626	259 HGG, 76 LGG	125	166	Segmentation, Survival Prediction, Uncertainty Prediction	Pre-operative, Longitudinal
2020	660	293 HGG, 76 LGG	125	166	Segmentation, Survival Prediction, Uncertainty Prediction	Pre-operative, Longitudinal

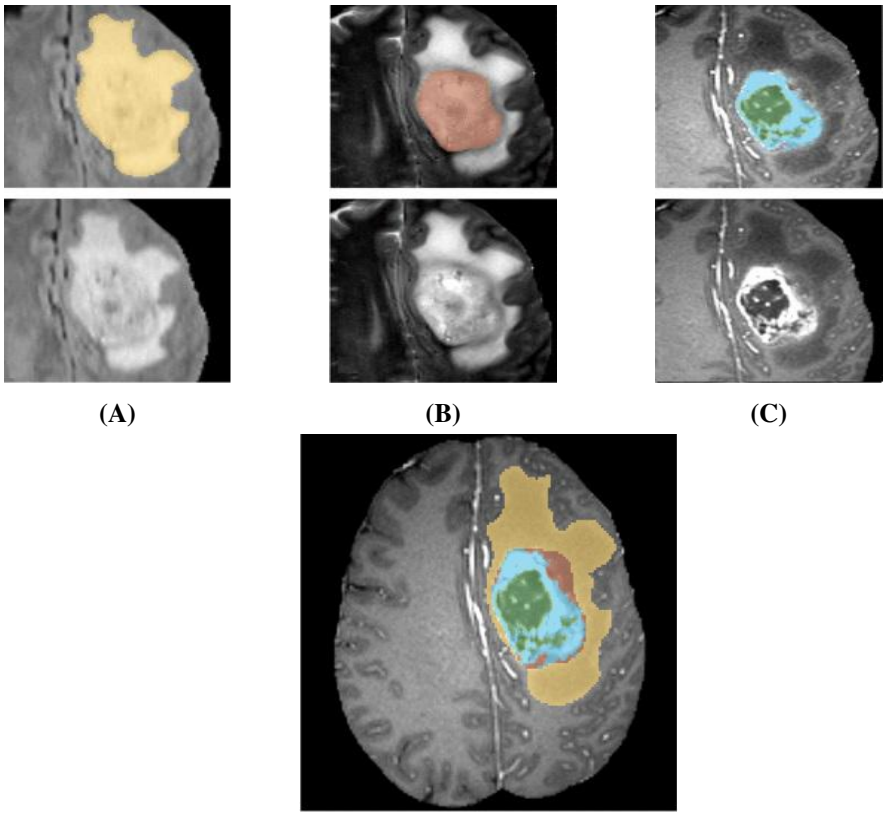


Figure 2.4: The appearance of intra-tumoral structures on three imaging modalities with manual annotations. (A) Top: whole tumor (yellow), Bottom: FLAIR, (B) Top: tumor core (red), Bottom: T_2, (C) Top: enhancing core structures (light blue), surrounding the cystic/necrotic core structure (green), bottom: T_1c, (D) Fusion of the three labels [12].

2.2.2 Overall Survival Prediction Task

The additional task of Overall Survival (OS) prediction was introduced in the BraTS challenge in 2017. This task focuses on the OS prediction of HGG patients. The dataset includes age and survival days along with resection status: Gross Total Resection (GTR) or Sub Total Resection (STR) information for HGG patients in addition to the images. The task is to classify the patients into long-term survivors (OS days $>$15 months), mid-term survivors (OS days between 10 to 15 months), and short-term survivors (OS days $<$10 months) [19]. A detailed description of the OS information available since BraTS 2017 challenge is given in Table 2.2.

Table 2.2: The distribution of BraTS dataset [6, 7, 8] features in survival classes.

Year	# Records	Features	Short-term Survivors (<10 months)			Mid-term Survivors (between 10 to 15 months)			Long-term Survivors (> 15 months)		
			Count	Age ($\mu \pm \sigma$)	OS Days($\mu \pm \sigma$)	Count	Age($\mu \pm \sigma$)	OS Days($\mu \pm \sigma$)	Count	Age ($\mu \pm \sigma$)	OS Days($\mu \pm \sigma$)
2017	163	Age	65	65.44 ± 10.68	147.44 ± 83.08	50	58.70 ± 11.26	394 ± 49.32	48	55.11 ± 12.19	826.23 ± 370.91
2018	163	Age	65	65.44 ± 10.68	147.44 ± 83.08	50	58.70 ± 11.26	394 ± 49.32	48	55.11 ± 12.19	826.23 ± 370.91
2019	212	Age, Resection status	82	66.66 ± 11.42	150.21 ± 84.72	54	59.14 ± 10.98	377.43 ± 40.44	76	57.16 ± 11.84	796.38 ± 354.32
2020	236	Age, Resection status	89	66.37 ± 11.15	149.64 ± 82.03	60	59.27 ± 10.67	375.36 ± 43.42	87	57.29 ± 11.58	796.40 ± 343.39

2.3 Post-processing techniques

The segmentation output generated by computer-assisted methods may contain false segmentation in the image due to improper or incorrect feature selection. The segmentation improves by applying post-processing techniques like:

2.3.1 Connected Component Analysis

Connected Component Analysis (CCA): It groups the voxels based on the connectivity depending on similar voxel intensity values. Connected components that are very small are excluded from the result. Such components are considered to be false positives due to spurious segmentation results.

2.3.2 Conditional Random Field

Conditional Random Field (CRF): Typically classifier predicts the voxel class based on the features related to that voxel, which does not depend on the neighbouring relationship of that voxel with other nearby voxels. CRF considers this relationship and builds a graphical model to implement the dependencies between the predictions.

2.3.3 Morphological Operations

Such operations are applied to adjust the voxel value based on its neighbourhood according to the size and shape of the tumor.

2.4 Evaluation Metrics

The standard evaluation framework for tumor segmentation and OS prediction includes the following metrics.

1. Dice Similarity Coefficient (DSC) (or F1 measure): It is the overlap of two objects divided by the total size of both the objects. True Positive (TP) is the outcome where the model correctly predicts the positive class. In contrast, False Positive (FP) is the outcome where the model incorrectly predicts the negative class to be positive. False Negative (FN) is when the model incorrectly predicts the positive class to be negative. This is represented by Equation 2.1 below:

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{2.1}$$

2. Jaccard Similarity Coefficient: It is known as the intersection over the union of two different sets as shown in Equation 2.2 below:

$$Jaccard = \frac{TP}{TP + FP + FN} \tag{2.2}$$

3. Sensitivity: It is a measure that correctly identifies tumorous voxels. This is shown in Equation 2.3 below:

$$Sensitivity = \frac{TP}{TP + FN} \tag{2.3}$$

4. Hausdorff Distance(HD): It measures how far two subsets of a metric space are from each other. If x and y be two non-empty subsets of a metric space (M,d), then their Hausdorff distance $d_H(x,y)$ can be defined by:

$$d_H(x,y) = max\{\sup_{x\in X}\inf_{y\in Y} d(x,y), \sup_{y\in Y}\inf_{x\in X} d(x,y)\} \tag{2.4}$$

   Where *sup* represents the supremum and *inf* the infimum.

5. Accuracy: It is defined to be the quality of being precise. True Negative (TN) is the outcome where the model correctly predicts the negative class as shown in Equation 2.5 below:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{2.5}$$

2.5 Summary

The chapter includes the basic pre-processing techniques: image registration, skull-stripping, bias field correction, and intensity normalization. The BraTS dataset contains the images which are co-registered and skull-stripped. The anatomy and scanner specific image intensity variation is addressed by the bias field, which is removed widely by N4 bias field correction [18]. The intensity normalization techniques scale the image intensities to the specified range. The z-score is a widely used normalization technique for MRI images. Post-processing improves the segmentation results by fine-tuning the segmentation map and correcting the false segmentation. The chapter introduces the BraTS incremental dataset [6, 7, 8], a publicly available dataset containing HGG and LGG images, with 45 images in the 2012 dataset to 660 images in the 2020 dataset. Initially, the dataset was used to address only the segmentation task, but gradually, the growth of dataset has incorporated the survival prediction task since 2017. The dataset also provides features like age, resection status, and survival days for HGG images with GTR and STR as resection status to address the OS prediction task. The most widely used evaluation measures for segmentation (DSC, specificity, Hausdorff distance) and OS prediction (accuracy) is covered at the end of the chapter.

Chapter 3

Literature Review

Automated or semi-automated brain tumor segmentation methods facilitate image interpretation, which is less time-consuming and reproducible. From the late 90s, brain tumor segmentation has gained the focus of researchers. Various algorithms have been developed to address the task of brain tumor segmentation which has been divided into conventional methods and Fully Convolutional Neural Network (FCNN) methods.

3.1 Conventional Brain Tumor Segmentation Methods

Conventional methods can be categorized into three groups: 1) the first group is made up of basic methods, 2) the second one is made up of generative methods, and 3) the third one is made up of discriminative methods [20]. Basic and generative methods can be interactive and non-interactive. Interactive methods require input from the user for processing, like tumor seed point, tumor diameter, or boundary selection. Non-interactive methods do not require user input.

3.1.1 Basic Methods

Threshold-based methods [21] use either the gray level of a pixel or the histogram of an image. Many thresholds are utilized for each distinct region in the event that the image contains multiple regions, each having a broad range of pixel values.

The basis of Edge methods [22, 23] is detecting the boundaries that exist between each of the region of the images. Edge based features play an important role for region characterization based on grayscale. Such features represent a discontinuity. The basis of the discontinuity is one of the following three: gray level, colour

or texture. Multiple operators like Sobel, Laplacian, Laplacian of Gaussian (LoG) are available to detect edges.

Region methods [24, 25] utilize values of pixels for the definition of homogeneity criteria for these particular methods. Texture methods [26, 27] make use of structural, spectral, or statistical features. Pixel intensity properties form the basis of the tone of a texture, and the structure relies on the spatial pixel relationship. Statistical features are procured from the values of pixels. These features includes variance, mean, median, kurtosis, and other higher-order statistics features.

3.1.2 Generative Methods

One of the most powerful techniques of segmenting brain tumors is the atlas based segmentation [28]. It forms an atlas or a look up table by using data on the size, shape and features of various organs. It is very much like techniques based on correlation and it conducts classification and segmentation in parallel.

In [29], the authors initially identified abnormal brain tissues by registering the tumorous brain image with healthy brain atlas image. This step was followed by identifying the presence of an edema using T2 images, and finally, geometric and spatial constraints were applied to detect the tumor and edema regions.

In order to match the patient globally, the authors in [30] applied an affine transformation to the atlas image. The Adaptive Template Moderated Spatially Varying statistical Classification (ATMSVC) algorithm was brought into use for the segmentation of the lesion. An expert seeded this atlas manually making use of just one voxel that was placed on the estimated origin of the patient's lesion. The next step was to use a non-linear registration algorithm with the model of the growth of the lesion deforming the seeded atlas in order to match the patient. The implementation of the model was conducted making use of four volumes of the contrast enhanced agent with meningioma.

The paper [31] applied a semi-automatic method, which required user input to give the seed point for the tumor, a radius for each tumor, and a seed point for each regular tissue class. The random walk generated tumor priors using initial tumor seeds. The patient-specific atlas was modified to accommodate tumor classes,

using a tumor growth model. The empirical Bayes model used the Expectation Maximization (EM) framework to update the posterior of the tumor, growth model parameters, and a patient-specific atlas. This work was then extended in [32]. Instead of a single seed point for various labels, multiple seed points were considered to find the intensity mean and variance of a specific label. This work focused on preoperative MRI scans. It was further extended in [33] to include post-operative scans and additional features to the Gaussion Mixture Model (GMM). Here the need for manual selection of the seed point was also omitted.

Model-based techniques [28] depend on the fact that the structure of the brain has a repetitive geometry. Hence, this makes it possible for the method to be modelled probabilistically in order to capture the variations in geometry and shape. The training data also needs to be registered. Internal forces and external forces are utilized by deformable models for the segmentation of the tumor on the basis of physics, geometry as well as data based on approximation. However, there are some concerns with these methods and they are: 1. Decision about the initial model and the selection of its parameters needs manual intervention; 2. low convergence properties.

The techniques based on models use random fields. For example, in order to label a voxel, use is made of the Markov random field taking into account its reliance on the voxels in the neighbourhood. An undirected graph can be used to represent this type of dependency. These types of models are based on the assumption the neighbouring voxels rely on each other and that the modelling takes on a form that is simply factorized. In [34], a Random Walk (RW) based interactive and an iterative method was applied to fine-tune the tumor boundary. RW was applied as an edge-weighted graph in the discrete feature space based on the variation of the distribution density of the voxels in the feature space. The user made an initial tumor seed selection for tumors as well as edema. Afterwards, RW was applied to the feature space as well as on the image. If the user did not approve the results, then the segmentation process was reinitiated.

In [35], a Hidden Markov Random Field (HMRF) based model was used with

modified Pott's Model to panelize the neighbouring pixels belonging to different classes. In [36], various modality intensity images and their neighbourhood voxel intensities were fed into the map-reduced Hidden Markov Model (HMM), and the model was corrected iteratively based on class labels.

In [37], a Gabor filter bank based Bayesian classification was followed by an MRF classification. Initially, each voxel was divided into its constituent class by applying the Gabor filter bank to the input vector (made up of intensities of four modalities at a voxel) to classify the voxel into five different classes (GM, WM, CSF, tumor, and edema). Next, an MRF based classifier was applied to the tumor as well as edema classes. It used voxel intensity and spatial intensity differences over the neighbouring voxels. [38] used a Non-negative matrix to find voxel clusters, which showed the tumor and level set methods to fine-tune the region boundary.

3.1.3 Discriminative Methods

Soft computing forms the basis for discriminative segmentation techniques. They can be categorized into two types: unsupervised and supervised. Most of the unsupervised forms are based on the idea of clustering. It learns the voxel reltionship as well as gain the knowledge about decision boundaries from unlabelled data. It depends on natural feature grouping in multidimensional space without utilizing the domain knowledge. K-means clustering is the widely used unsupervised approach which makes crisp sets by partitioning brain tissues, whereas fuzzy C means clustering considers partial volume effect of the voxels for clustering. The shape and size of the clusters may differ. The limitation of the fuzzy C-means exists because the sensitivity of the partition matrix gets stuck in the local minima.

The supervised form requires the presence of training data for model learning. Artificial Neural Network (ANN) [39] is made up of numerous neurons that are interconnected. Adaptation of the free parameters of ANN during the process of training is how the ANN learns. Later on in the process, this knowledge is used to classify test patterns without any concern for the input-output functional relationship. In [40], tensor features were extracted along with mean, entropy, and standard

deviation features. Authors in [41] extracted features for super voxels from multi-scale images and created sparse feature vectors to segment the whole tumor.

Random Decision Forest (RDF) has achieved remarkable success for tumor segmentation. Authors in [42] used context-aware spatial features and tissue appearance probability generated by GMM to train the decision forest. [43] worked on Random forest classification with CRF regularization to predict tissue probability in multiple classes, i.e., GM, WM, CSF, edema, necrotic core, and enhancing tumor. A 28-dimensional feature vector included the intensity of each modality along with first-order statistics like mean, variance, skewness, kurtosis, energy, and entropy computed from local patches around each voxel in each modality.

In [44], Extremely Randomized Trees (ExtraTrees) was used, which introduces more randomness at training time. The classifier was trained on 208 features extracted from all the four modalities, including intensity values, local histograms, first-order statistics, second-order statistics, and basic histogram-based segmentation. In the paper, the ExtraTrees trained with the best threshold rather than the individual features threshold. In [45], pixel classification was done with ten random forests with ten trees each. They were trained to reduce the training time and finally merge into a single forest with Gini impurity. One thousand samples for the tumorous class and 1000 samples for the non-tumorous class train the RF. Classification forest in [46] used 237 features, which included appearance specific features (image intensities, first-order texture features, and gradient features) and context-sensitive features (ray feature, symmetry intensity difference features).

In [40], tensor features were extracted along with mean, entropy, and standard deviation features. Authors in [41] extracted features for super voxels from multi-scale images and created sparse feature vectors to segment the whole tumor. Substructures of the tumor were then separated using CRF.

Appendix B, Table B.1 includes the comparison of the conventional methods for the pre-processing, dataset, number of images and DSC.

3.2 Initial Implementation: K-means Clustering

K-means clustering is made use of for partitioning data into K-clusters. Based on similar features, the K-means clustering groups data points. It creates distinct clusters for dissimilar data points. Distance measures like Euclidean distance or Manhattan distance are used to group feature vectors.

Post-processing is applied in the removal of voxels that are non-tumorous. The size of a given neighbourhood is considered in the removal of these voxels. For the process of segmentation, three types of neighbourhoods are taken into account namely, 6, 18 and 26 [47]. The assigning of a label to a voxel of interest is done by majority voting amongst its neighbours. In the process of assigning a label, the majority voting for 6, 18, and 26 neighbourhoods make use of 5, 17 and 25 pixels in a neighbourhood. Figure 3.1 depicts the output of the K-means clustering.

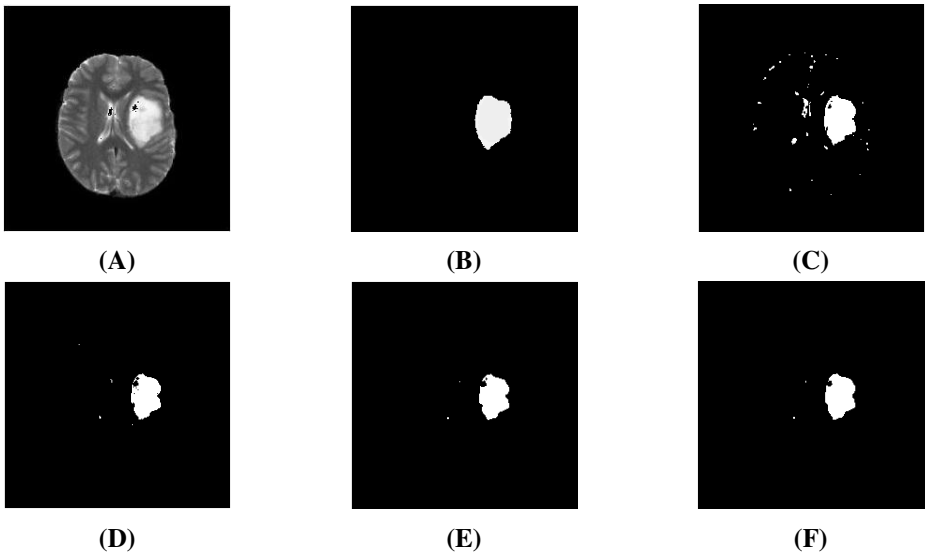


Figure 3.1: (A) FLAIR image slice, (B) ground truth slice, (C) K-means segmentation slice, (D) 6-neighborhood post-processed slice, (E) 18-neighborhood post-processed slice, (F) 26-neighborhood post-processed slice.

As can be seen, segmentation that is done with no post-processing is extremely noisy. The segmentation results in this case are generated with numerous false positives. Noise can be removed with the application of K-means. The second observation that is made is that as the size of the neighbourhood increases, the process of cleaning greatly improves. It is the fact that voxels in the neighbourhood

Table 3.1: 2D slice DSC for BraTS 2015 dataset.

	Six Neighbourhood T_2	Six Neighbourhood FLAIR	Eighteen Neighbourhood T_2	Eighteen Neighbourhood FLAIR	Twenty Six Neighbourhood T_2	Twenty Six Neighbourhood FLAIR
Patient 1	0.447	0.172	0.448	0.171	0.450	0.168
Patient 2	0.420	0.442	0.421	0.442	0.425	0.445
Patient 3	0.337	0.471	0.338	0.473	0.344	0.477
Patient 4	0.000	0.066	0.000	0.066	0.000	0.066
Patient 5	0.447	0.327	0.446	0.329	0.447	0.326

have a tendency to be alike that leads to improvisation because of K-means. The application of K-means to BraTS 2015 images can be seen in a different experiment. Table 3.1 shows the DSC outcome of the experiment on five images with varying neighbourhoods.

An increase in the size of the neighbourhood brings about an improvement in the results of the segmentation. However, the segmentation shows no improvement overall since this technique localizes to the neighbourhood. This confirms that K-means on image pixels does not lend itself well for segmentation of brain tumors.

3.3 Limitation of Conventional Methods

The limitations associated with the methods working on handcrafted features are as follows:

1. Identifying tissue probability classes: Tumor tissue intensities overlap with that of the healthy tissues; in such a case identifying the probable class for tumorous tissue is quite challenging.
2. Atlas matching (healthy or tumorous atlas): Usually, the brain atlas contains the normal brain tissue distribution map. Due to the deformation of the healthy tissue by the tumor, the atlas matching of a tumorous brain may result in the wrong map.

3. Manual seed point identification for the tumor or its subparts: Almost all semi-automated methods require some initial selection for the tumorous voxel, its diameter, or its rough outer boundary. The selection depends on the expert. Its repetition over all the slices of the brain is a time-consuming task.
4. Feature extraction from the images: RF training depends on the features extracted from the brain images. All the MRI modalities contain different biological information. This variation in the information complicates the tasks of feature extraction as well as selection, to training RF.
5. Discontinuity: The results generated by such methods are spurious, which increases the chances of false segmentation. Proper post-processing techniques are required to fine-tune the generated results.

The limitations of the conventional methods raise the need for methods which are non-interactive, fast, and can extract features automatically from the given input. The next section addresses the solution to this.

3.4 Deep Neural Network

Deep Neural Network (DNN) is an artificial intelligence function that does data processing and creates patterns for decision making by extracting and learning the features from the input. The success of DNN can mainly be attributed to the following four reasons:

1. The DNN models solve problems in an end-to-end manner from simple to complex feature learning from the input. Automatic feature learning has eliminated the need for domain expertise.
2. Computational capabilities of the hardware in terms of GPU and efficient implementation of the GPU model with various open-source libraries have made the training of the DNN faster than CPU.
3. Efficient optimization techniques for robust learning contribute to the success of DNN for optimal network performance.

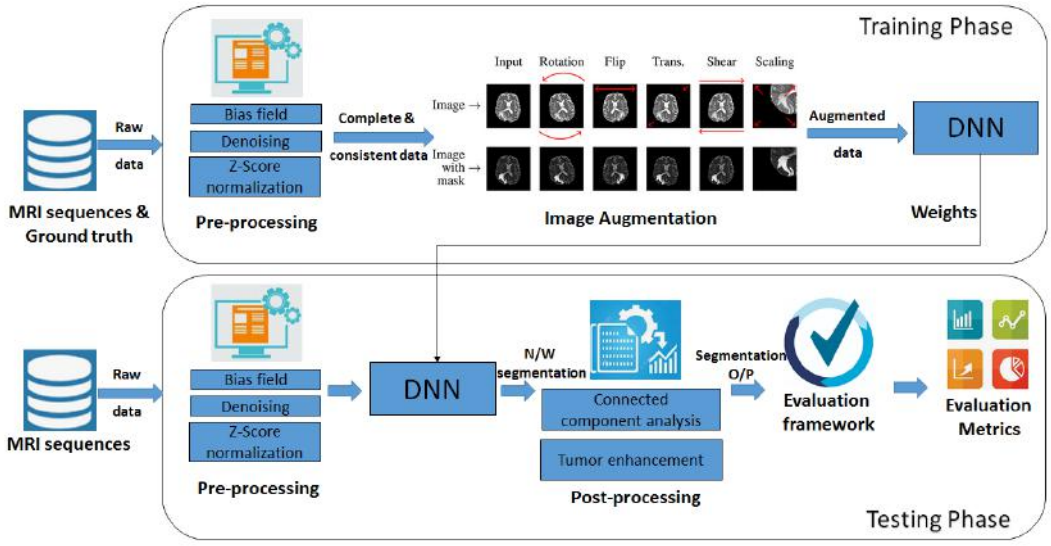


Figure 3.2: General pipeline for DNN tumor segmentation.

4. The availability of benchmark datasets allows training and testing of various deep learning models to be implemented successfully.

The general block diagram of any DNN is, as shown in Figure 3.2. The crucial task is to get a labelled data set. After the availability of the dataset, it is divided into training and validation sets, followed by appropriate pre-processing techniques as per the task on hand. Actual DNN applies to the training data, which makes the network learn the network parameters. The output of DNN is spurious in some brain areas, and post-processing fine-tunes the segmentation result. Finally, the evaluation framework measures the performance of the network.

3.4.1 Evolution of DNN

In medical image analysis, a semantic segmentation task is common, e.g., organ segmentation, lesion. Convolutional Neural Network (CNN), a specific type of DNN architecture, gained its popularity since 1990 with the architecture of LeNet [48]. The authors suggested a two layer CNN architecture. After the availability of fast GPUs and other computing facilities, over fifteen years later, AlexNet was proposed by authors in [49] with five convolutional layers.

3.4.2 Design Aspects of CNN

The CNNs are designed with: 1) various layers like convolution layers, non-linearity layers, pooling layers, fully connected layers 2) regularization techniques, 3) optimization techniques, 4)loss functions, and 5) network parameter initializations and normalization. Authors in [50, 51] have explained the architectural elements of CNN nicely, which are as follows:

- **Convolution layer:** extracts representative features from the input. It achieves: 1)weight-sharing mechanism, 2)exploits local connectivity of input, and 3) provides shift invariance to some extent.
- **Non-linearity layer:** provides a sparse representation of the input space, which achieves data variability invariance and computationally efficient representation. Types of non-linearity layers are Rectified Linear Unit (ReLU), Leaky ReLU (LReLU), Parametric ReLU (PReLU), S-shaped ReLU (SReLU), Maxout and its variants, Exponential Linear Unit (ELU) and its variants.
- **Pooling/subsampling layer:** extracts prominent features from the non-overlapping neighbourhood. It is used to: 1) reduce the number of parameters, 2) reduce over-fitting, and 3) achieve translation invariance. There are two pooling methods that are used very commonly and they are average pooling and max pooling.
- **Fully connected layer:** converts 2D features to a 1D feature vector. It helps to predict the input image class label.
- **Loss functions:** improve the learning process by improving within-class similarity and between-class separability.
- **Regularizations:** deal with over-fitting issues. Commonly used regularization techniques are L_1 and L_2 regularizations, dropout, early stopping, batch normalization.
- **Optimization:** used for proper updates of network parameters during back-propagation. Various techniques of optimization include Nesterov accelerated gradient descent, adaptive gradient algorithm (Adagrad) and Root Means

Square Propagation (RMSProp).

- **Weight initialization and normalization:** boost the learning process by helping the weight update with proper initial values.

The convolution layers extract features from the input by applying kernels to it. The output feature map depends on the type of kernel and its size. At the initial layers, simple features are extracted from the input like edge or lines. The gradual increase of the network depth requires a higher number of feature maps to extract complex shapes [20]. The activation function is applied to the feature maps to learn the non-linear relationship within the data and it allows the errors to back-propagate to the initial layers for accurate parameter updates. An increase in the network depth exponentially increases the network parameters, which is computationally very expensive. Pooling layers are introduced to down-sample the input feature maps and reduce its spatial size by considering only prominent features. Fully connected layers at the end of the network flatten the results of the input layers before actual classification. The loss function at the classification layer calculates the error in the prediction. Based on this error, the network parameters are updated using gradient descent methods by back-propagation. Commonly used loss functions are:

- Cross Entropy loss function:

$$J = -\frac{1}{N}\left(\sum_{voxels} y_{true} \cdot log\, \hat{y}_{pred}\right) \tag{3.1}$$

- Dice Loss function:

$$J = 1 - \frac{2\sum_{voxels} y_{true} y_{pred} + \varepsilon}{\sum_{voxels} y_{true}^2 + \sum_{voxels} y_{pred}^2 + \varepsilon} \tag{3.2}$$

Here, $N = Number\ of\ voxels$, $y_{true} = ground\ truth\ label$, and $y_{pred} = network\ predicted\ label$.

3.4.3 CNN Architectures

The CNNs, with convolution layers followed by fully-connected layers, classify the entire image in a single category. GoogleNet(Inception) [52] and InceptionV3 [53] networks have introduced the inception module, which implements kernels of

different sizes to reduce network parameters. ResNet [54] has introduced a residual connection between the convolution layers such that it learns the identity function, which allows practical training of deeper networks. In DenseNet [55], layers are very narrow and add significantly fewer feature maps to the network, allowing them to design deeper architectures and training. Each layer has direct access to the gradient of the loss function.

For semantic segmentation, CNN can classify each voxel of the image individually by presenting it with several patches extracted around the particular voxel. Each voxel of the image is classified with the same process, resulting in segmentation of the entire image. This 'sliding-window' approach repeats the convolution operations for adjacent patches of neighbouring voxels. The improvement of this approach is replacing fully-connected layers with convolution layers, which generates the probability map of the entire input image rather than generating output for a single voxel. Such networks are known as FCNN. FCN [56] is a type of FCNN where skip connections are introduced to reconstruct a high-resolution image.

U-net [57] is a very well known, highly adapted network architecture for tumor segmentation. It has taken the encoder-decoder approach where every encoding layer is connected with its peer decoding layer with skip connection to reconstruct the dimension and get detailed spatial information from the encoding layer. SegNet [58] and DeepLab [59] are the other types of FCNN architectures adopted to solve brain tumor segmentation.

3.4.4 Handling Class Imbalance Problem

In medical image analysis, the class imbalance issue exists at two levels. At the first level it exists because the number of abnormal images is very small compared to the normal images, which results in class imbalance. At the second level, within the abnormal images found, even on a single brain slice, the volume of the tumorous tissue compared to the non-tumorous tissue is very small, again resulting in a class imbalance issue. Also, finding the number of abnormal images compared to the normal images is difficult as an abnormality like the tumor is rare. All images in this thesis are of the tumor. The proportion of the Brain Volume (BV) and Background

Volume (BGV) with respect to the tumor volume for the BraTS 2020 dataset is as shown in Table 3.2. It can be observed that TV proportion is extremely small.

Table 3.2: The proportion of Non Tumor Brain Volume (NTBV) and Non Tumor Background Volume (NTBGV) vs. Tumorous Volume(TV).

	% NTBV	% Necrosis	% Edema	% ET
BV	93.10	1.53	4.00	1.37
BGV	98.88	0.25	0.65	0.22

The following approaches address data imbalance problem.

- Patch sampling: The patch sampling-based method can mitigate the imbalanced data problem. The sampling process includes equiprobable patches from all the tumorous regions as well as the non-tumorous region.
- Improvement in loss functions: Some of the loss functions, when used in their raw form, may not suit the tumor segmentation task, as they consider balanced datasets. These functions adopt an imbalanced dataset with some modifications. Some of those functions are weighted cross-entropy, generalized dice loss, focal loss functions.
- Augmentation techniques: Mostly, a lot of the labels for training are unavailable for multiple reasons. In order to label the dataset, an expert in this field is required, which turns out to be costly as well as time-consuming. One method for reducing over-fitting and increasing the volume of the training data is data augmentation. By transforming images in the training dataset, it creates new images. Such images are translated, flipped, rotated, scaled, distorted and also some noise such as Gaussian noise is added.

3.4.5 CNN Approaches for Brain Tumor Segmentation

The classification of CNNs for brain tumor segmentation depends on the input and output type, CNN design with or without ensemble approach. The possible combination for network design is as shown in Figure 3.3.

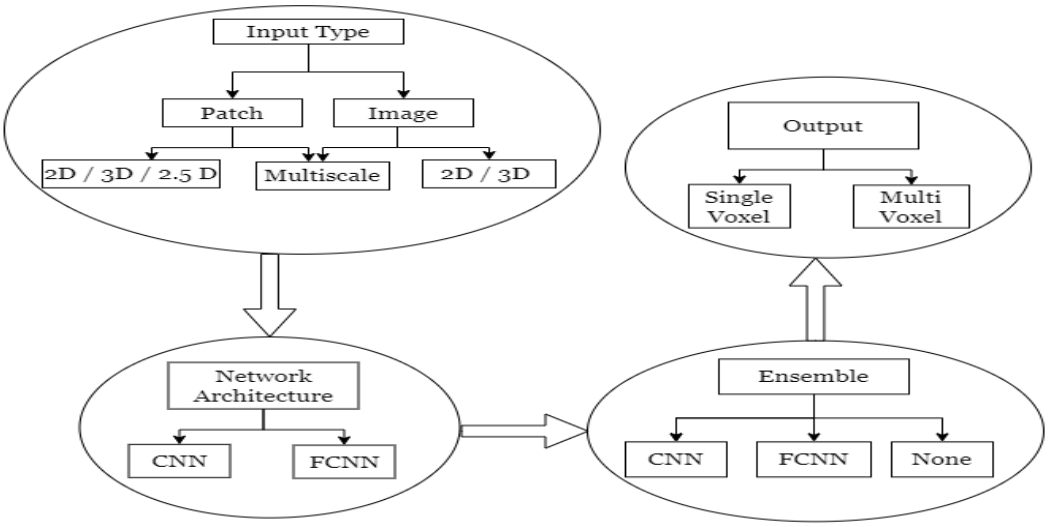


Figure 3.3: CNN design approaches.

Input type: The network may take 2D/3D input in the form of patches or images. With fully-connected layers, CNN classifies the centre voxel of the patch, whereas FCNN predicts multiple or all voxels of the patch/image. The network may take multi-scale patches to extract coarse and fine details of the input.

Output Type: The output of the network depends on the problem to solve. It predicts a single output for the classification problem and multiple voxel outputs for the semantic segmentation problem.

Type of network: The CNN approach indicates the convolution network with fully-connected layers at the end, whereas FCNN indicates the network with all convolution layers.

Ensemble Approach: The Ensemble approach can be classified into cascaded and parallel approaches. In the cascaded approach, multiple networks combine in series to fine-tune the end output. The input of one network depends on the output of the other. In the parallel approach, multiple networks work in parallel and take the same/different input to gather the comprehensive details of the input. The final output of the network is decided based on the majority voting or averaging all the network outputs.

3.4.6 Brain Tumor Segmentation using CNN and FCNN

Initially, shallow CNN performs voxel-based image segmentation. The authors in [60] proposed a voxel-wise classification using CNN multipath way architecture. One pathway used 2D patches of size 32x32, and the other used fully-connected input of 5x5 patch size having the center pixel the same as a 32x32 patch. Patch selection was made such that the labels are equiprobable. L_1 and L_2 regularizations overcome overfitting. In [61], voxel-wise class probability prediction used separate 3D CNNs for HGG and LGG images. The final probability classified the voxel into six classes. In [62], a five layer deep 2D CNN architecture performed voxel-wise classification.

Gradually the depth of CNN had increased to accommodate more layers in the network. In [63], a 2D deep CNN with fully connected output layers separated HGG and LGG. This approach was further extended by [64] with a two-phase process and a weighted loss function-initially. The network trained using equiprobable patches that followed actual patch training without the class imbalance problem. In [65], the authors designed a 2D Input Cascaded CNN, which took the output of a Two Path CNN to train other 2D CNNs with input images. After successfully implementing an FCN the authors in [56, 66] proposed a two pathway architecture, where both pathways included residual connections and trained on different input patch sizes. As the network was fully convolutional, output labels of multiple voxels could be generated at a time.

In [67], the authors used the 2D FCNN approach along with CRF. The FCNN trained on patches and CRF on slices. In [68], cascaded encoder-decoder like FCNNs along with residual connections were proposed. The first FCNN segmented the whole tumor, followed by internal tumor region segmentation by the second FCNN. Authors in [69] proposed an encoder-decoder FCNN based architecture to segment various tumor subregions. In [70], the authors proposed three different FCNN architectures and showed that the architectures with multiresolution features performed better than single-resolution architectures. Authors in [71] implemented Dilated

Residual Network for patch-based training where equiprobable patches were supplied to the training network.

Authors in [72] adopted a U-net architecture for brain tumor segmentation. Authors in [73] modified the U-net, which took 3D input, and the depth of the network reduced to three. Authors in [74] optimized the training of the network proposed in their previous work [75]. In [76], the authors proposed a novel encoder-decoder architecture that worked well on multiple biomedical image segmentation problems.

An ensemble of CNNs performs better compared to a single CNN, as in [77], where authors implemented an ensemble of seven networks using DeepMedic, FCN, and U-net along with variations of those three networks. They also tried three different approaches for pre-processing on all these networks. The output of individual networks with all the pre-processing generated the final label. Authors in [78] extended their work proposed in [69], where the dense module and dilated module were introduced in the encoder-decoder cascaded architecture of two networks. The pooling layers were replaced with dilated convolution layers.

The authors in [79] implemented an ensemble of ten encoder-decoder based architectures, which included the auto-encoder stream to reconstruct the original image for additional information and regularization purposes. Authors in [80] extended their approach proposed in [81]. They used a combination of U-net and Densenet with U-net like architecture containing dense blocks of dilated convolution. The network of [67] was extended in [82] to create an ensemble of three networks that were trained on three image views. Two network cascaded paths were used in [83], where Coarse Segmentation Network segmented WT, and Fine Segmentation Network segmented the sub-regions of the network. Both the networks used 3D U-net with four levels of deep architecture.

Authors in [84] used three networks (WNet, TNet, and ENet) to prepare the cascaded path. The multi-scale prediction of the voxels was averaged for the private network. These networks trained on three-views of images (axial, coronal, sagittal) and the result from the averages generated the final segmentation output. In [85],

a cascaded network was proposed, which initially segmented the whole tumor followed by the tumor core, and enhancing tumor segmentation refinement. In [86], the authors proposed a cascaded U-net with three networks. The network processed downsampled input, and the generated output was passed to the next network in the cascaded sequence. Authors in [87] proposed multi-scale mask 3D U-nets with atrous spatial pyramid pooling layers. WT segmentation generated by the first network was passed to the second for TC generation, which in sequence passed to the final network to generate ET output.

Authors in [88] used an ensemble of 12 3D networks with patches as input. In [89], the authors used various 3D networks with a variation in input patches. Authors of [90] extended their work with instance normalization instead of batch normalization.

Appendix C, Table C.1 shows the comparison of the FCNN based segmentation methods. The comparison mainly includes the pre-processing techniques, MRI modality and augmentation techniques, the type of the input to the network, input view, network architecture, number of networks used, ensemble approach in case of more than one network, loss function, post-processing technique, and DSC.

3.4.7 Overall Survival Prediction

In [91], an ensemble of RF and CNN segmented the tumor and random forest regressor (RFR) was used to predict the OS days using 240 features out of 1366 different features (Kaplan-Meier estimator was used to find relevant and useful features). Authors in [81] modified U-net with a Full Resolution Residual Network (FRRN) and Residual Unit (RU) units along with weight scaling dropout. The survival prediction ANN worked with a linear activation function on four selected features.

The variant of U-net was used in [75], which took a 3D input and included a context module and a localization module in each architecture level. The segmentation result was generated based on element-wise summation of the output from the decoder layers. Survival prediction was the average of RFR and Multi Layer Perceptron (MLP). RFR trained on 517 features extracted from three tumor subregions using the radiomics package [92]. The output of RFR and MLP averaged

the OS prediction. 15 MLPs were designed with three hidden layers, each with 64 neurons.

Authors in [93] implemented 3D U-net with three stages of encoder-decoder architecture. Regression model based radiomics features selection trained the MLP for OS prediction. In [94], two 3D U-nets used four-stage encoder-decoder architecture; the first network segmented the whole tumor, and the second one segmented the tumor sub-region. In addition to the four conventional modalities, they used an additional image as an input, which the T_1c-T_1 subtracted. This image provided additional information for the tumor core region. They used only the age feature to predict the OS using linear regressor. The approach presented in [95] used an FCNN named FCRN-101, which was derived from pre-trained SegNet and U-net architecture. A three path network combined the result of three views, i.e., axial, coronal, and sagittal. The OS prediction used SPNet, the fully-connected CNN, which took four modalities and the network segmentation result as input to predict the probability of OS prediction.

Authors in [96] used an ensemble of six 3D U-net like networks with variation in the input size, the number of encoding/decoding blocks, and feature maps at every layer. The OS prediction used linear regression with ground truth segmentation, OS information, and surface area. Features were input to the network after z-score normalization. In [97], the authors implemented FCN and generated results for three axes and used majority voting to generate final segmentation results. For OS prediction, ten features (focusing on necrosis and active tumor) from the segmentation results were generated, and mean PCA and standard deviation PCA were used to train the RF on the GTR images.

In [98], an ensemble of three networks (U-net [57], DFKZNet [75], and CACNN [84]) was used, and majority voting applied for final segmentation. The OS prediction used RF with 14 radiomics features selected from various modality images, Laplacian of Gaussian Images, and wavelet decomposed images. Authors in [99] implemented a 2D U-net architecture with three stages for tumor segmentation and age, volumetric, and the shape features of the whole tumor were used to predict

OS. Authors in [100] implemented a 2D U-net of three stages with dense blocks at every encoder level and the feature set of [99] of a necrosis tumor sub-region for OS prediction.

3.5 Summary

This chapter gives an overview of various methods to address the segmentation and OS prediction tasks. The segmentation was initially handled by conventional methods, which include basic, generative, and discriminative methods. The limitations of such methods are identifying tissue probability class, manual tissue class selection, atlas matching, and feature extraction. The DNN methods, with a particular focus on the CNN/FCNN, are covered to address the limitations as mentioned earlier. These methods are categorized as per the approaches shown in Figure 3.3. The end-to-end processes use CNN/FCNN for tumor segmentation and conventional methods like SVM, RF, and linear regression for OS prediction. Out of all traditional prediction of OS methods, RF gives a promising prediction outcome. The deep learning methods are not used for OS prediction as the available dataset for it contains limited samples (236), which are not sufficient to train any deep learning methods.

Chapter 4

The Development of New Encoder-Decoder Architecture

This chapter depicts the development of a new encoder-decoder FCNN architecture that helps to segment a brain tumor efficiently with less computation complexity.

4.1 Hardware and Software Specifications for Algorithm Development

The proposed networks are designed using the following open source python packages:

- SimpleITK and Nibabel: These packages help read and write the MRI images in .nii format. SimpleITK is also used for bias field correction.
- Scikit-image: The package helps apply the affine transformation on the image for pre-processing and connected component analysis for post-processing.
- TensorFlow: This package helps design and train the network as this package supports library management, has better debugging and scalable functionalities and is designed for pipelining to support various backends like CPU and GPU.

The networks are trained and tested on the following workstations:

- HP Z240: Intel Xeon Processor E3-1245 v5 3.50GHz, 4 Cores, Quadro K5200 8GB GDDR5 GPU memory, 256-bit memory interface, 193 GB/s memory bandwidth.

- HP Z840: Intel Xeon Processor E5-2643 v4 3.40GHz, 6 Cores, Quadro P5000 16GB GDDR5X GPU memory, 256-bit memory interface, up to 288 GP/s memory bandwidth.

4.2 Fully Convolutional Neural Network

In healthcare, a medical image is the most reliable treatment plan for acute diseases like cancer, fracture, heart problems, and chest infection. Segmentation helps identify the diseased area from the images. Though medical images are the largest source in the field, it is difficult to analyse these images without experts like radiologists. Artificial intelligence-based deep learning algorithms are introduced in this field for fast, automated, and independent image analysis. Moreover, a Glioma brain tumor affects a patient's mortality, which requires robust tumor segmentation for better OS prediction.

As mentioned in Chapter 3, researchers have attempted the task of brain tumor segmentation with conventional and deep learning approaches. The CNNs have proven robust in comparison to the conventional methods.

The important part of CNN architecture is the convolution layer [101], which helps in sparse interactions, network parameter sharing, and translation invariant feature representations. This layer generates the set of feature maps resulting from applying a convolution filter to the given input. The feature maps represent the features learned during the training phase. The initial layers learn basic features from the input. Deeper layers successively learn complex features. The design of CNN requires special attention for the following points:

- The number of convolution filters in a layer and its size: The convolution filter works as a feature detector and generates the explanatory factors in feature maps by looking at the input. The number of filters in a layer decides how many different features must be learned by the layer. Moreover, the kernel's size in a filter decides the neighborhood voxels relationship, which is considered for convolution operation to generate the feature map. The number of convolution filters and filter size may vary depending on the application,

problem complexity, and image resolution. Empirically, the more number of filters means the more the network learns. However, saturation and convergence of the model matter the most. These hyperparameters are specified before training.

- The initialization of convolution filter parameters: For the imbalanced data set, the convolution filter initialization plays an essential role in iterative training convergence. Improper parameter initialization may lead the training converge in platue or saddle point and does not converge to a global optimum.
- The memory requirement for CNN training depends on three factors:

    1. The intermediate volume sizes: These numbers represent the amount of memory used for activations at each layer and its gradient, which are needed for backpropagation.
    2. The parameter sizes: These are the numbers that hold the network parameters and their gradients for backpropagation. Additional parameters are required for advanced parameter optimizations like momentum, Adagrad, or RMSProp.
    3. The miscellaneous memory: It is required to store the image data batches and their augmented versions, if any.

The CNN and FCNN methods of Chapter 3 are computationally expensive and memory inefficient. The proposed architectures are made simple, computationally efficient, and require less memory. Besides, the focus is to address the accurate OS prediction requires the features extracted from the network segmented results.

4.3 2D Encoder-Decoder FCNN

Figure 4.1 shows the proposed FCNN with eleven convolution layers. The arrangement of layers makes the two layers encoder-decoder architecture where the combination of two successive convolution layers creates two encoder, one bottleneck, and two decoder modules, respectively. The network design considers the following points:

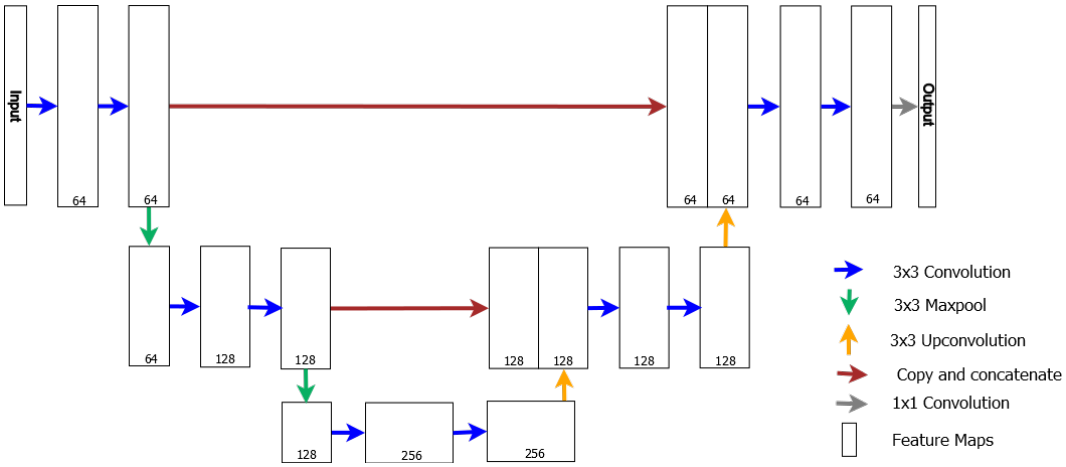


Figure 4.1: The proposed 2D encoder-decoder FCNN.

- Number of convolution filters: The heterogeneous tumor appearance demands learning of various features from the input image. This requirement is addressed by the number of convolution filters at each layer. The first encoder module uses 64 filters at each convolution layer. This filter count is successively doubled until the bottleneck layer, which results in 64, 128, and 256 filters in the first encoder module, second encoder module, and bottleneck layer. At the decoder module, the filters are halved at each layer, which results in 128 and 64 filters at the first and second decoder modules.
- Kernel size: The kernel size at each convolution layer is 3x3, which is the empirically chosen smallest size to reduce computational complexity. If the kernel size is 5x5, then a single convolution operation results in 25 multiplication operations, whereas 3x3 results in 9 such operations reduce computations by approximately 2.8 times. Further, two consecutive 3x3 convolution operation results in 18 operations give a wider receptive field the same as 5x5 and reduce the computation by 1.4 times. Also, it reduces the noise, if any, and preserves the sharpness of the median filter. The last convolution layer uses a 1x1 kernel size. This layer applies channel-wise pooling for output map generation.
- Max pooling: It follows each encoder module. It keeps the most prominent feature from the previous feature map, summarizing the features learned by

convolution layers. It also reduces the number of other passed parameters to the successive layers, reducing the network computation cost.

- Upconvolution: It upscales the summarized information received from previous layers. The upscaling is performed to match the input image size required to establish a voxel-wise relationship between network input and output.
- Long skip connections [102]: These connections are between the peer modules of encoder-decoder path, which work as a highway to pass low-level features from the encoder module to the decoder module. Such connections have advantages like 1) it helps model convergence by resolving the vanishing gradient problem, 2) they help improve information reconstruction by providing the low-level information learned at the encoder modules.

The ReLU activation follows all the convolution layers. The network uses the dice loss function for network parameter learning. Every two filter maps have the dimension of 240x240, 120x120, 60x60, 120x120, and 240x240, respectively, followed by the output probability map. The segmentation map is learned using the one-vs-all approach, which leads to separate training for individual subregions. The network input is made up of 2D T_1, T_1c, T_2, and FLAIR modality slices. The scarcity of the tumor images for networking learning is handled with the data augmentation techniques. It generates images from the available data set images to help geometric transformation and filtering techniques [103]. The geometric transformation techniques preserve the labels post-transformation. These techniques include rotation, translation, flipping. The other transformation includes the elastic transformation, which is the filter-based transformation. The disadvantage of this augmentation is that it does not preserve the label post-transformation.

The small amount of labelled data including augmentated one for substructures specifically for enhancing and necrosis substructures causes network training failure. The solution gets stuck in the local minima, which necessitates network training to be re-initiated. Inductive transfer learning [104] for parameter initialization resolves the above issue. Here the substructure network initializes with whole tumor weights instead of random initialization. It serves two purposes: 1) It handles

Table 4.1: Test data set evaluation for DSC, sensitivity and positive predictive value (PPV), where A: model without weight initialization, B: model with weight initialization.

	DSC %		**Sensitivity %**		**PPV %**	
Whole Tumor	78		76		91	
	A	B	A	B	A	B
Necrosis	58	**65**	56	**67**	69	**70**
Enhancing	58	**60**	**56**	54	69	**74**
Edema	63	**71**	56	**66**	79	**83**

the dearth of the labelled data, and 2) It also provides substructure area localization. The functioning of network training has been greatly improved because of weight transfer. The network training uses random parameter initialization from normal distribution for whole tumor segmentation. The procured weights that are obtained, are in turn, conveyed to the substructure networks for the initialization of the parameters. The network undergoes training for 50 epochs during each iteration. The DSC comparison is as shown in Table 4.1.

Figure 4.2 depicts sample 2D slice results for the segmentation of the entire tumor along with three substructures (including inductive transfer learning and also with no inductive transfer learning). A network which is trained without parameter initialization results in more false outputs as compared to a network which is trained with whole tumor parameter initialization. Many such false results are marked with arrows in the Figure 4.2.

The advantages of the proposed 2D FCNN network are:

1. The translation invariance capabilities of the model can locate the tumor anywhere within the brain.
2. fewer network parameters speed up the segmentation and it takes less than 50s to generate the result.
3. subcomponent network initialization learns the parameters in a better way.

Limitations of the proposed 2D FCNN network are:

1. It does not learn the diversified features.

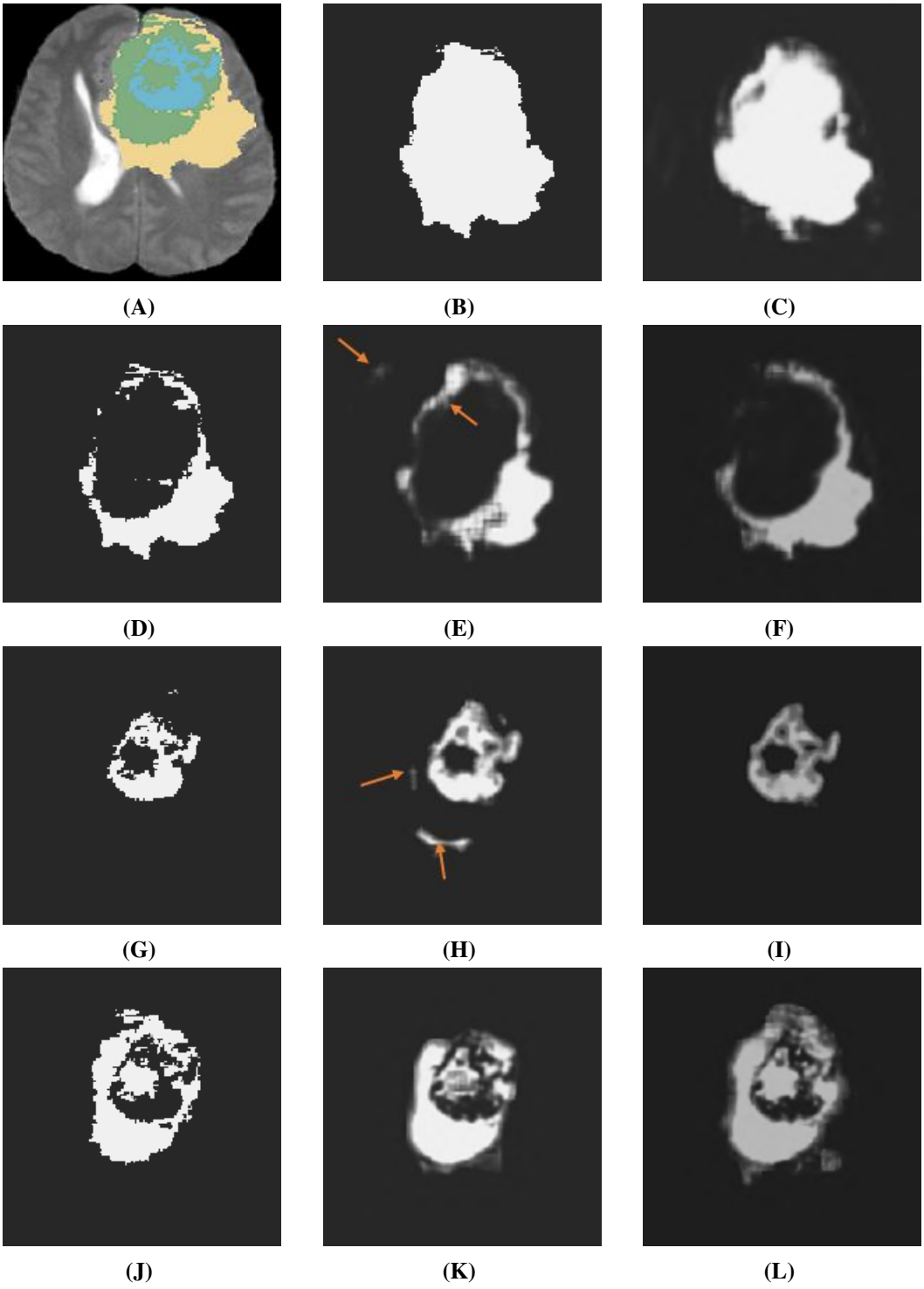


Figure 4.2: Segmentation output: (A) T2 image slice, (where, Edema is in yellow, Enhancing core is in blue, Necrotic/Non-enhancing core is in green) (B) WT ground truth, (C) WT segmentation, (D) Edema ground truth, (E) Edema segmentation without weight initialization, (F) Edema segmentation with weight initialization, (G) Enhancing tumor ground truth, (H) Enhancing tumor segmentation without weight transfer, (I) Enhancing tumor segmentation with weight transfer, (J) Necrotic ground truth, (K) Necrotic segmentation without weight transfer, (L) Necrotic segmentation with weight transfer.

2. A smaller receptive field does not learn global features accurately.
3. In extreme cases, the network may get stuck into local minima or platue and the gradient does not flow back to the initial layers resulting in no parameter learning.
4. The 2D training does not allow for learning the 3D relationship between the voxels across slices.
5. Four networks are trained for the whole tumor and the tumor substructures, which takes around eight days for the training.

A dense FCNN overcomes some of these limitations of the proposed 2D FCNN. The following section covers a dense FCNN.

4.4 2D Encoder-Decoder Dense FCNN

The proposed FCNN makes use of a two-layer architecture of 4.3, where the encoder module is replaced with the dense module. The network architecture is as shown in Figure 4.3.

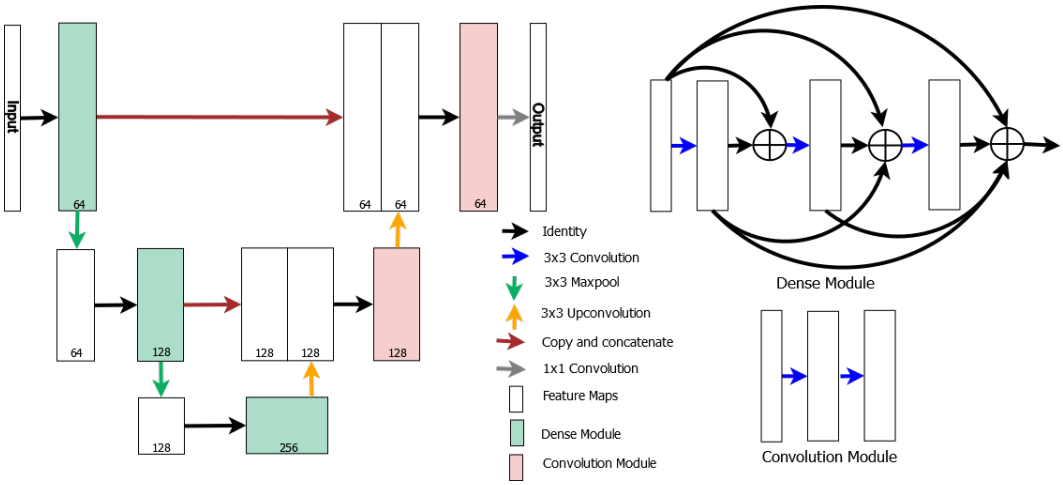


Figure 4.3: The proposed 2D encoder-decoder dense FCNN.

The dense module is made up of successive convolution layers with dense connections in between. The dense connections create the short skip connections between the layers of a module. These short skip connections between convolution layers have multiple advantages:

- They allow extra inputs to construct collective knowledge of the earlier layers and pass on their feature-maps to all the successive layers.
- They allow gradient flow to the preceding layers for in-depth supervision by the classification layer.
- They provide heterogeneous features to the subsequent layers, which results in a better ability for identifying patterns.

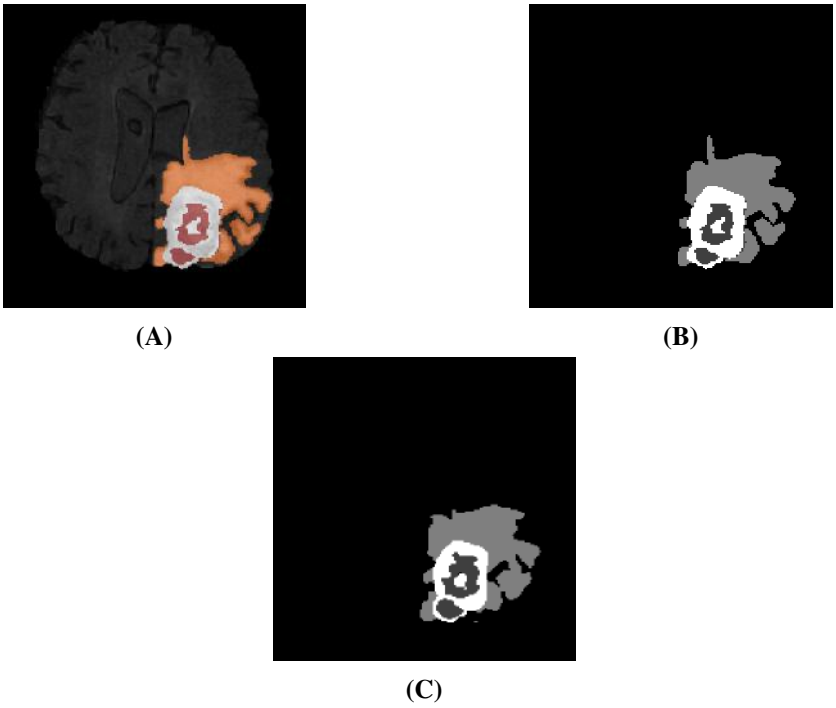
ReLU activation follows each convolution layer. The parameter initialization and data augmentation techniques of section 4.3 to initialize the substructure network training are used for this network. The first phase completes the whole tumor training, followed by substructure training with whole tumor training weights. The data set is divided into 85% and 15% for training and validation. Input to the network is 2D slices of T_1, T_2, T_1c, and FLAIR modalities of 240x240 size each. The blank slices of the images are removed to speed up the process. Further, the focal loss function for network training handles the misclassification error as shown in Equation 4.1 below:

$$\begin{aligned} Loss_{FL}(p_t) &= -\alpha_t(1-p_t)^{\gamma} log(p_t) \\ p_t &= \begin{cases} p_t & if\, y_i = 1 \\ 1-p_t & if\, y_i = -1 \end{cases} \end{aligned} \tag{4.1}$$

where $y \in [1,-1]$ is the ground-truth label, and $p_t \in [0,1]$ is the predicted probability of the class with label $y = 1$. The weighting parameter α deals with an imbalanced data set. The focusing parameter γ makes an adjustment in the rate at which easy examples are down-weighted very easily. Setting $\gamma > 0$ can result in a reduction in a relative loss for examples that are well classified. It moves the focus on examples that are challenging and misclassified. The focal loss is the original cross-entropy loss when $\gamma = 0$.

The successful segmentation of a network for a sample slice is shown in Figure 4.4. In some cases, the network is unable to distinguish between the substructures which is shown in Figure 4.5. In some tumor cases, the enhancing tumor surrounds the necrosis, whereas in other cases it penetrates the necrosis. These variations are

not learned by the network which degrades the performance on ET. Table 4.2 shows the achieved DSC and Hausdorff95, and Table 4.3 shows sensitivity and specificity of the network on the training set. The values in the tables show that the network performs accurately on the given data. Various current techniques have been compared and depicted in Table 4.4. The suggested network performance compares well for ET and WT and outperforms for TC for both DSC and Hausdorff95 measures.

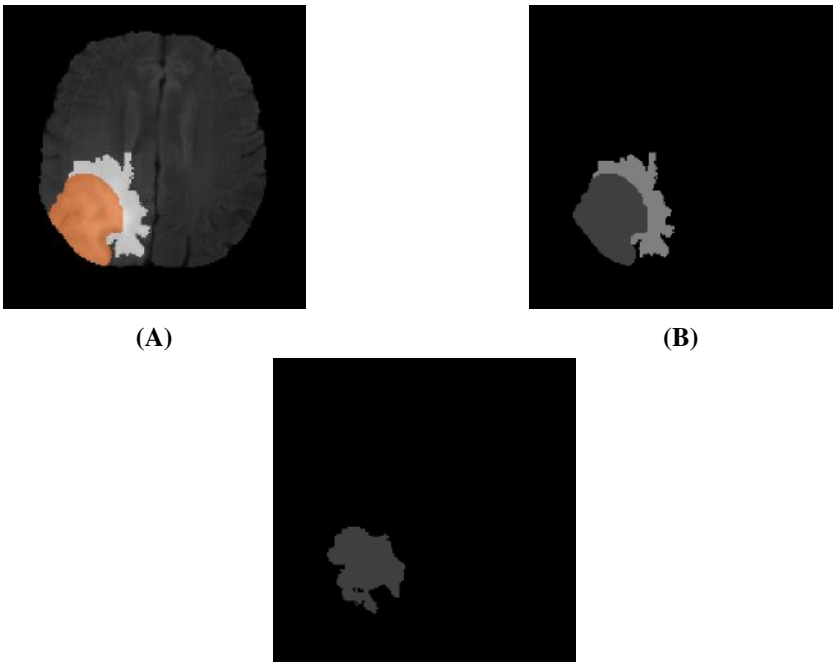


(A) **(B)** **(C)**

Figure 4.4: Correct segmentation result (A) FLAIR slice, brown-necrosis, gray-enhancing tumor, orange-edema (B) Ground truth slice (C) Segmentation slice.

Table 4.2: BraTS 2019 training set evaluation for DSC and HD95.

	DSC%			**HD95**		
	ET	WT	TC	ET	WT	TC
Mean	79.32	91.58	89.86	04.07	04.23	03.75
StdDev	24.58	08.57	12.32	11.67	06.39	07.79
Median	87.36	94.53	93.31	01.41	02.24	02.00
25quantile	81.90	91.07	89.28	01.00	01.41	01.41
75quantile	91.83	95.87	95.55	01.73	04.24	03.00



(A) **(B)**

(C)

Figure 4.5: Incorrect segmentation result (A) FLAIR slice, orange-necrosis, gray-edema (B) Ground truth slice (C) Segmentation slice.

Table 4.3: BraTS 2019 training set evaluation for sensitivity and specificity.

	Sensitivity%			**Specificity%**		
	ET	WT	TC	ET	WT	TC
Mean	79.10	89.69	88.04	99.90	99.65	99.81
StdDev	20.59	12.10	13.82	00.22	00.33	00.36
Median	84.74	93.92	92.32	99.95	99.74	99.91
25quantile	77.24	89.45	87.76	99.87	99.55	99.83
75quantile	90.73	95.98	95.01	99.98	99.86	99.97

Table 4.4: Comparison of DSC and HD95 with state-of-the-art methods.

References	**DSC%**			**HD95**		
	ET	WT	TC	ET	WT	TC
[89]	79.70	90.00	87.30	03.64	04.43	04.16
[90]	77.12	90.53	84.47	03.71	07.09	05.05
[105]	75.61	92.18	88.88	04.01	03.70	03.92
[106]	78.62	90.31	82.90	03.37	05.04	05.56
[107]	79.17	90.94	83.22	04.02	03.80	05.65
[108]	82.62	91.65	88.80	03.07	04.01	03.67
Proposed	79.30	91.60	**89.60**	03.90	03.76	**03.56**

On the training set, the network functions well, but on the validation and test sets the outcome does not compare well. The observation for validation set is shown in Table 4.5 and Table 4.6 whereas for test set is shown in 4.7.

Table 4.5: BraTS 2019 validation set evaluation for DSC and HD95.

	DSC%			**HD95**		
	ET	WT	TC	ET	WT	TC
Mean	59.17	72.52	65.09	09.62	12.80	15.37
StdDev	34.14	23.64	30.36	15.83	16.86	19.91
Median	76.28	83.63	78.43	03.61	07.48	07.81
25quantile	29.44	64.92	51.45	01.93	04.58	04.00
75quantile	85.44	88.60	88.18	07.98	12.81	16.16

Table 4.6: BraTS 2019 validation set evaluation for sensitivity and specificity.

	Sensitivity%			**Specificity%**		
	ET	WT	TC	ET	WT	TC
Mean	58.83	67.38	63.67	99.80	98.51	98.79
StdDev	33.09	24.97	31.05	00.78	05.90	05.31
Median	71.35	75.16	76.48	99.96	99.70	99.90
25quantile	33.46	53.62	41.13	99.84	99.31	99.61
75quantile	85.64	87.82	88.12	99.99	99.94	99.98

Table 4.7: BraTS 2019 test set evaluation for DSC and HD95.

	DSC%			**HD95**		
	ET	WT	TC	ET	WT	TC
Mean	63.75	71.83	66.44	055.11	41.30	057.12
StdDev	32.82	29.24	35.98	125.68	99.06	122.21
Median	78.39	84.41	85.71	002.24	06.40	005.51
25quantile	55.06	67.08	50.95	001.41	03.81	002.45
75quantile	86.39	90.50	92.47	010.48	13.24	018.45

Minor modifications are applied to the architecture, input as well as loss functions to observe the performance on the training and validation sets. These changes along with DSC are as shown in the Table 4.8. The variations in the network include:

Table 4.8: Model variations of 2D encoder-decoder dense FCNN.

Model	Architectural change	Input modalities	# Feature maps	Input size	Loss function	Training set DSC(%)	Validation set DSC(%)
1	Original	T1, T1c, T2, FLAIR	32, 64, 128, 64, 32	240 x 240	dice	WT:90 TC:86 ET:75	WT:72 TC:63 ET:55
2	Original	T1, T1c, T2, FLAIR	32, 64, 128, 64, 32	156 x 200	dice	WT:90 TC:86 ET:75	WT:72 TC:63 ET:55
3	Dense module + Convolution layer	T1, T1c, T2, FLAIR	32, 64, 128, 64, 32	156 x 200	dice	WT:89 TC:84 ET:71	WT:71 TC:61 ET:56
4	Dense module + Convolution layer	T1, T1c, T2, FLAIR	32, 64, 128, 64, 32	156 x 200	focal	WT:93 TC:90 ET:83	WT:75 TC:65 ET:60
5	Dense module + Convolution layer	T1c, T2, FLAIR	32, 64, 128, 64, 32	156 x 200	focal	WT:87 TC:81 ET:67	WT:71 TC:58 ET:51
6	Dense module + Convolution layer	T1c, T2, FLAIR	32, 64, 128, 64, 32	156 x 200	focal	WT:92 TC:89 ET:79	WT:76 TC:66 ET:60

- The architectural changes: the number of feature maps at each layer and inclusion of convolution layer before the dense module at the encoder side.
- The input changes: number of input modalities and the size of the 2D slice extracted from the input images as the bounding box. The bounding box is identified to remove the blank area in the image.
- Loss functions: The network is trained using either dice loss or focal loss.

The DSC comparison of both, the training as well as validation sets for these variations is shown in Figure 4.6. The segmentation results show that the method overfits the training data and does not generate good results for the unknown validation set. The performance degradation of the proposed architecture is due to the following model limitations:

- The limited network depth restricts the abstract level feature learning, which results in improper feature reconstruction at the decoder modules.
- The augmentation is applied to an entire image, which does not solve the imbalance between tumorous and non-tumorous slices.

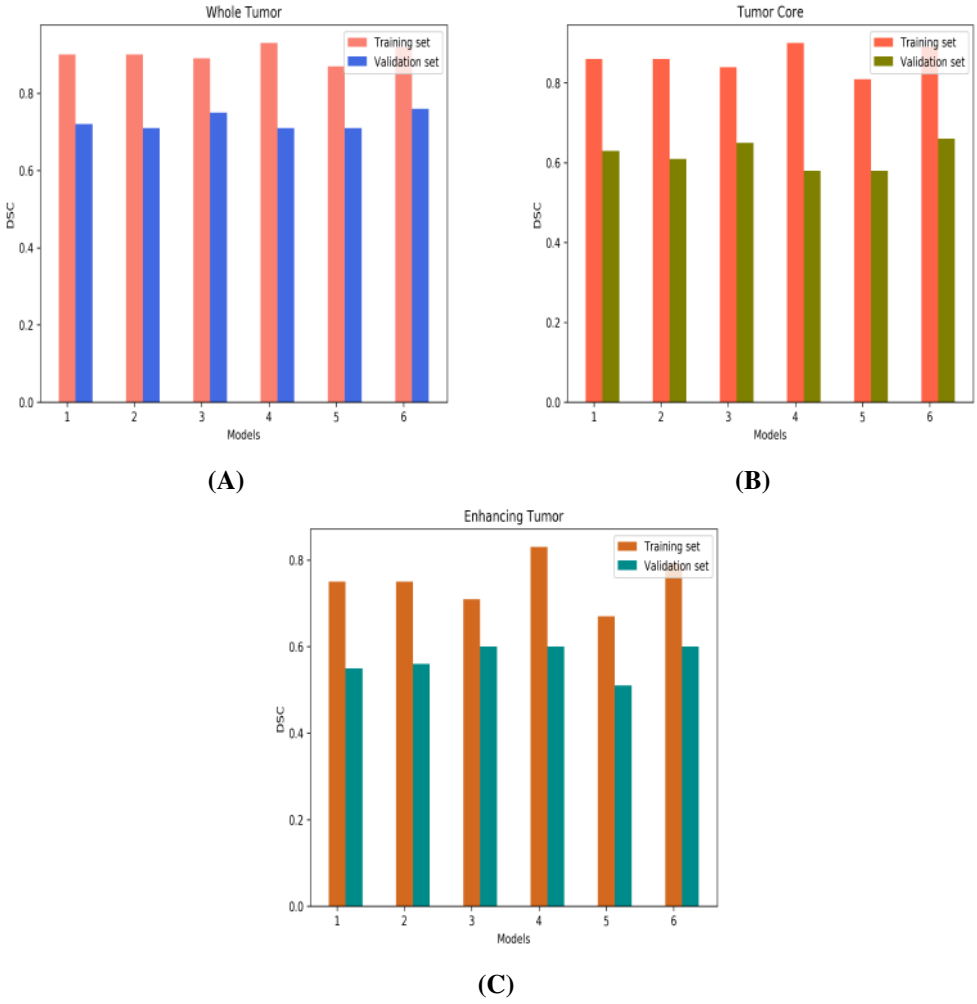


Figure 4.6: DSC comparison for training and validation set (A) whole tumor (B) tumor core (C) enhancing tumor.

- The dense encoding module learns various features, which are not correctly reconstructed at the decoding module as these modules are comparatively more straightforward than the encoding module. This leads to the network overfitting issue.
- It cannot learn the voxel relationship in three-dimension, which results in improper segmentation for unknown images.
- Four networks are trained for a whole tumor and the tumor substructures, which takes around eight days for the training.

The proposed 3D FCNN model in Section 4.5 overcomes these limitations.

4.5 3D Encoder-Decoder Dense FCNN

The 3D FCNN gathers a spatial relationship between the voxels. It is based on the network proposed in Section 4.4 with a 3D variation, as shown in Figure 4.7. The architectural changes to the network are as follows:

- The network depth is increased by one layer, which helps the network learn abstract features.
- The decoding modules are replaced with the dense module. This helps the network learn proper feature reconstruction at the decoding modules.
- The batch normalization is applied after the convolution to reduce the internal covariance shift [109], which introduces strong regularization in the network. Parametric ReLU activation follows batch normalization, which accelerates network training.
- Input to the network is equiprobable 3D patches from a background (nontumorous region) and foreground (tumorous region). This addresses the imbalance class issue.

The encoder dense modules generate 2^6, 2^7, 2^8 and 2^9 feature maps, respectively. The decoder dense modules generate 2^8, 2^7 and 2^6 feature maps followed by 1x1x1 convolution as the last operation to generate a one probability map which includes labels for multi-class classification.

An equal number of patches from tumor and non-tumor brain areas are extracted from the images to guarantee that the network does not overlearn the background voxels. Further, the network trains with a combination of the dice function and the focal loss function. The network takes training patches from all the four modalities, and all the training set images are taken as input. Four hundred training patches are extracted from each training image. The network trains for 610 epochs with batch size 1. A sliding window approach generates the output for each subject. The stride size is reduced to half of the training window size to overcome the unstable prediction issue of the boundary voxels. The output of the original patch and flipped patch is predicted and averaged to generate the final output.

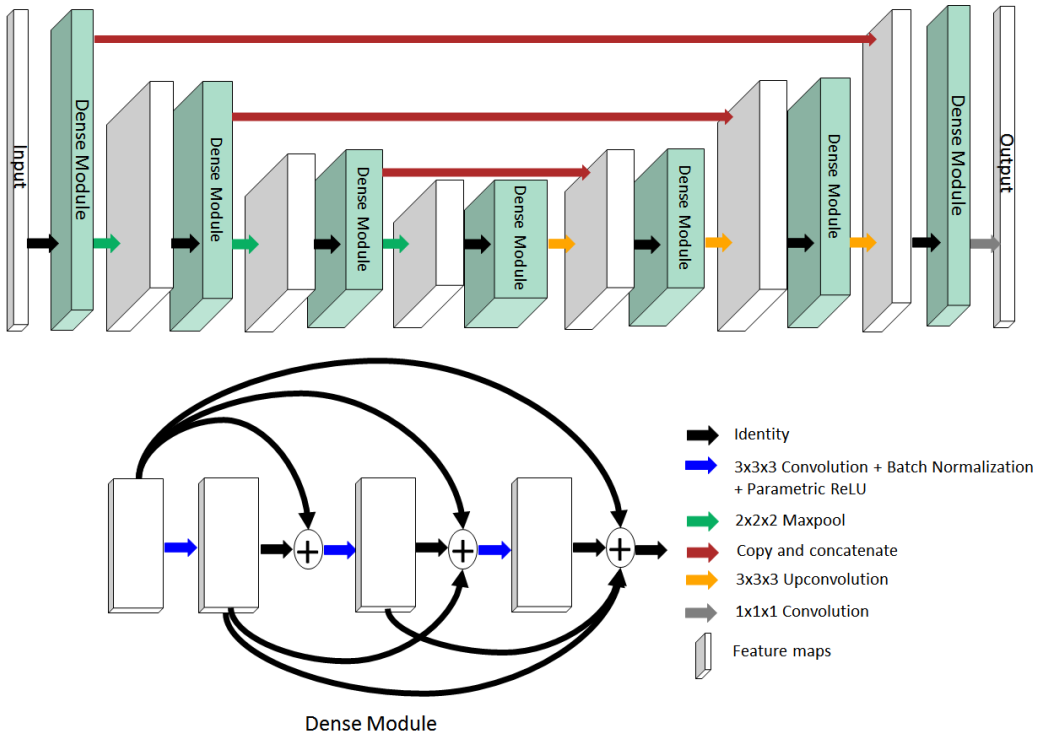


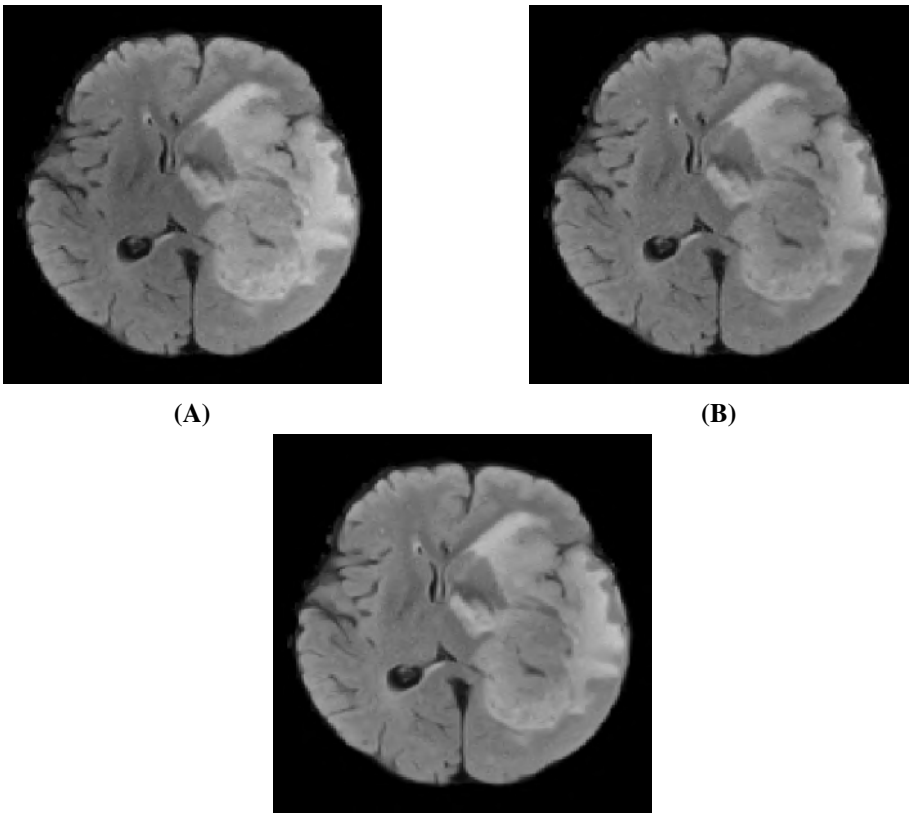
Figure 4.7: The proposed 3D encoder-decoder dense FCNN.

4.5.1 Pre-processing

The pre-processing enhances network training and performance. All four modality images are bias field corrected followed by denoising and z-score normalization. The pre-processing is applied on individual MR sequence. Data augmentation happens by flipping the patches around the vertical axis. Figure 4.8 shows the bias field corrected and denoised images. The z-score normalization is applied at the time of training.

4.5.2 Post-processing

Post-processing includes two methods: 1) The connected component analysis removes the tumor with volume less than thousand voxels, 2) an enhancing tumor is formed surrounding the necrosis and its size cannot be very small in HGG. Such small sized enhancing tumor is converted to necrosis. The empirically chosen threshold for the conversion is 300.



(A) **(B)** **(C)**

Figure 4.8: Image pre-processing (A) FLAIR slice (B) Bias field correction (C) Bias field correction with denoising.

4.5.3 Results

Various evaluation metrics; DSC, Hausdorff95, Sensitivity and Specificity for the training set are in Table 4.9 and Table 4.10. The results for validation set are given in Table 4.11 and Table 4.12. The comparable performance of network is also observed for test set, which is as shown in Table 4.13. Figure 4.9 depicts the box plot of the evaluation metrics, where the red marked crosses show the below average segmentation. The failure of the network is observed for: 1) small size of a tumor, 2) small size of the necrosis, and 3) absence or small size of an enhancing tumor.

Figure 4.10 shows successful segmentation of the tumor. False positive segmentation voxels are removed in post-processing. The network fails to segment the tumor for some HGG images and many LGG images. One such segmentation failure is shown in Figure 4.11.

The segmentation improvement of 3D FCNN over 2D FCNN is as shown in

Table 4.9: BraTS 2020 training set evaluation for DSC and HD95.

	DSC%			**HD95**		
	ET	WT	TC	ET	WT	TC
Mean	78.22	88.15	83.22	29.27	06.23	06.99
StdDev	24.68	11.66	19.10	94.96	08.13	20.51
Median	87.18	91.97	90.97	01.41	03.46	03.32
25quantile	77.89	86.23	80.50	01.00	02.24	02.00
75quantile	91.19	94.28	94.38	03.16	06.94	07.87

Table 4.10: BraTS 2020 training set evaluation for sensitivity and specificity.

	Sensitivity%			**Specificity%**		
	ET	WT	TC	ET	WT	TC
Mean	78.29	84.44	80.19	99.98	99.94	99.97
StdDev	25.20	15.80	20.91	00.03	00.09	00.06
Median	86.36	89.55	88.25	99.97	99.97	99.99
25quantile	76.20	79.53	76.33	99.97	99.93	99.98
75quantile	92.45	94.50	93.07	99.99	99.98	99.99

Table 4.11: BraTS 2020 validation set evaluation for DSC and HD95.

	DSC%			**HD95**		
	ET	WT	TC	ET	WT	TC
Mean	76.29	87.28	75.28	27.70	07.04	10.87
StdDev	25.92	09.83	26.28	90.92	09.35	33.82
Median	85.21	90.77	87.79	02.24	03.74	04.69
25quantile	75.08	85.64	71.06	01.41	02.45	02.24
75quantile	89.89	93.51	92.60	04.24	06.48	11.05

Table 4.12: BraTS 2020 validation set evaluation for sensitivity and specificity.

	Sensitivity%			**Specificity%**		
	ET	WT	TC	ET	WT	TC
Mean	75.90	84.69	71.34	99.97	99.91	99.97
StdDev	27.27	14.93	28.83	00.03	00.13	00.07
Median	85.25	89.66	84.13	99.99	99.96	99.97
25quantile	71.60	80.93	60.66	99.97	99.88	99.97
75quantile	93.25	95.45	92.17	99.99	99.98	99.99

Table 4.13: BraTS 2020 test set evaluation for DSC and HD95.

	DSC%			HD95		
	ET	WT	TC	ET	WT	TC
Mean	77.92	87.49	81.49	27.08	08.30	21.61
StdDev	23.19	11.24	25.02	92.55	30.01	74.65
Median	84.67	91.01	91.26	01.73	03.46	03.16
25quantile	76.01	85.49	83.36	01.10	02.00	01.80
75quantile	90.83	93.50	94.82	02.73	06.16	07.86

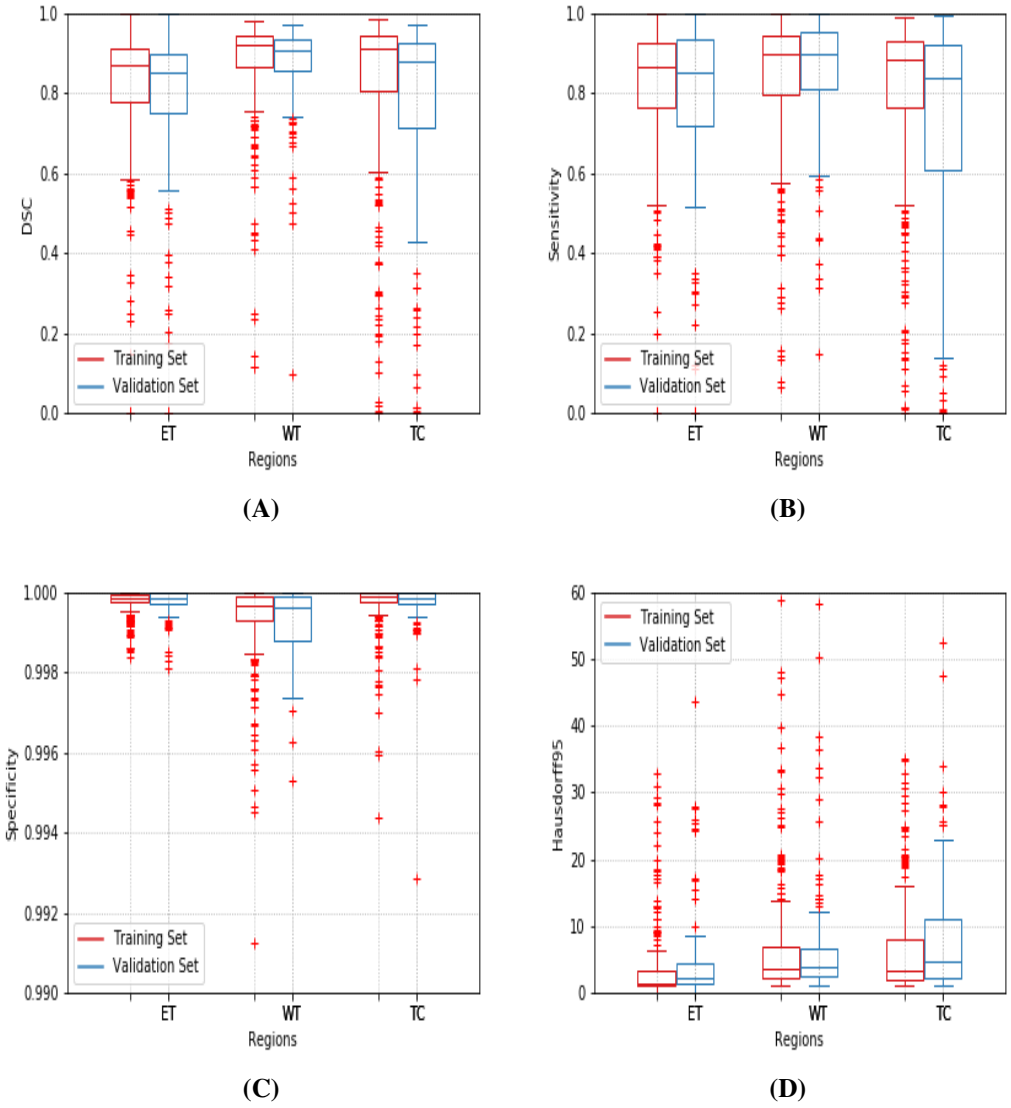


Figure 4.9: The box plot (A) DSC (B) Hausdorff95 (C) Sensitivity (D) Specificity.

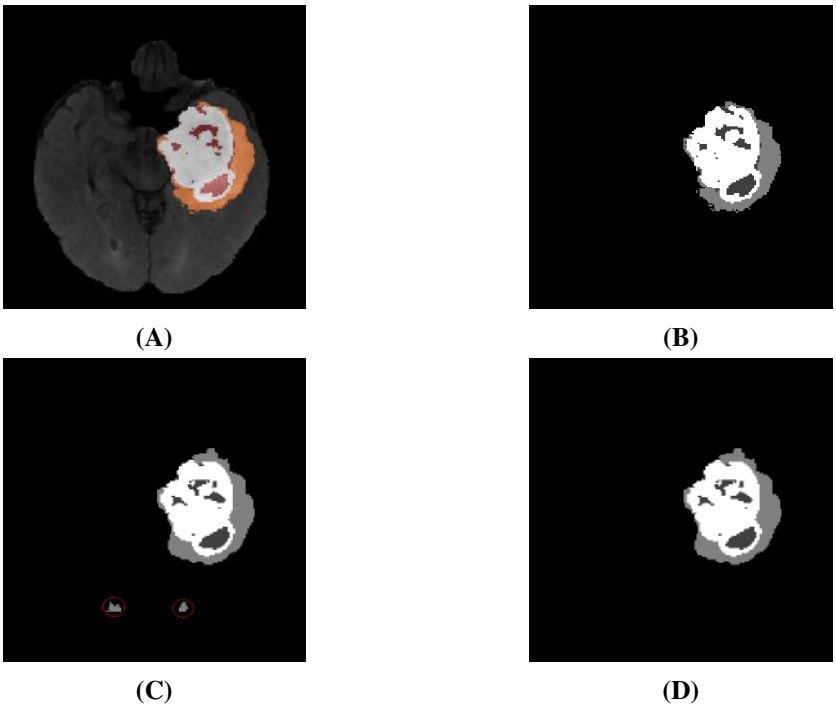
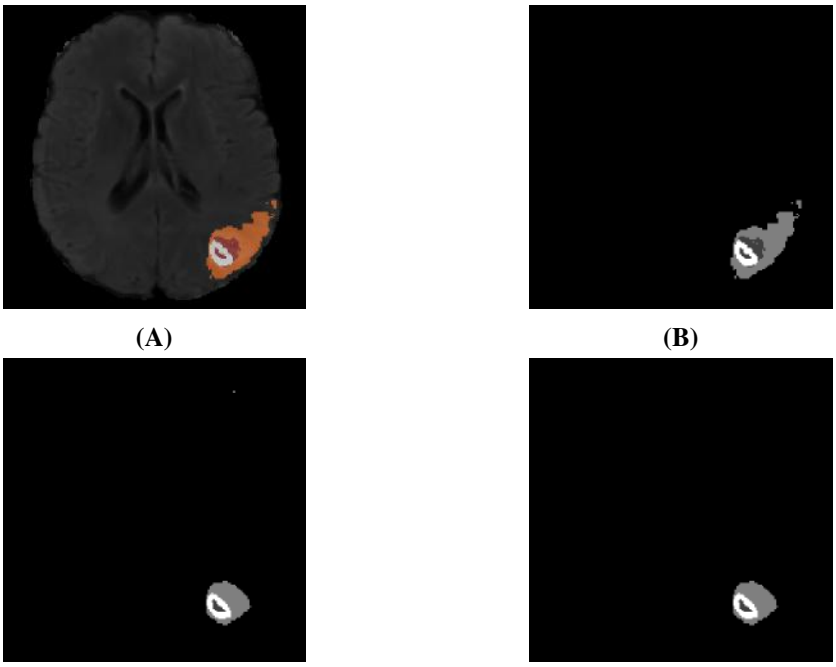


Figure 4.10: Correct segmentation results of the network, brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) Segmentation without post-processing (D) Segmentation after post-processing.

Figures 4.12 to 4.15. Both the networks segment the tumor irrespective of location, but the voxel relationship is learned better by 3D FCNN than 2D FCNN. The 3D FCNN improves the result in the absence of enhancing tumor, which is generally the case in LGG images, as shown in Figure 4.15.

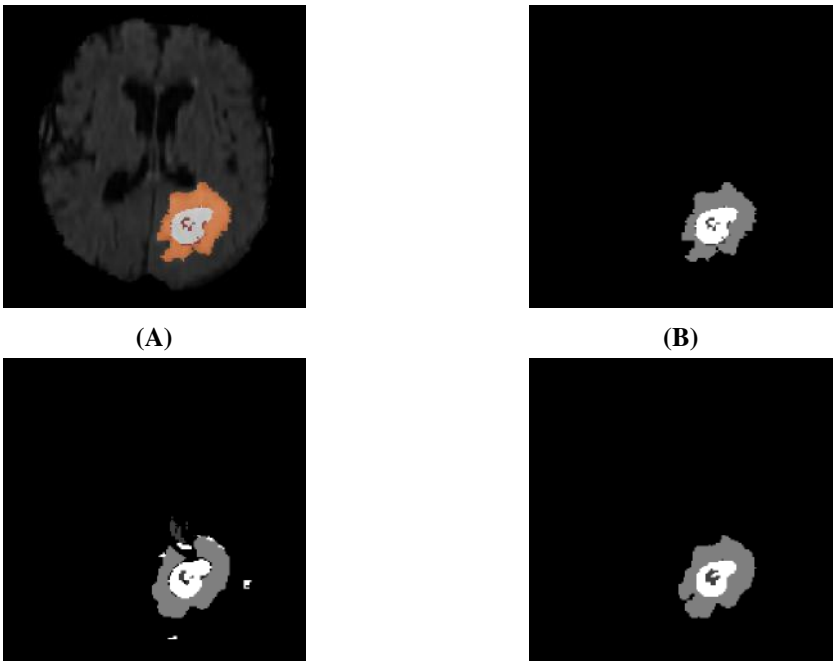
The proposed methods are compared with other advanced ensemble techniques and the comparison is listed in Table 4.14. All the methods use an ensemble of FCNN which works in parallel, cascade, parallel of cascades or cascades of parallel. Although the ensemble approaches provide robust segmentation, they have certain limitations: (1) ensemble methods are computationally expensive. Hence, learning time and prediction time are added to the problem along with memory constraints, (2) using ensemble methods reduces model interpretability due to increased complexity and makes it very difficult to understand. The proposed network aims performance similar to ensemble approaches and gives consistent output with less computational complexity.



(A) **(B)**

(C) **(D)**

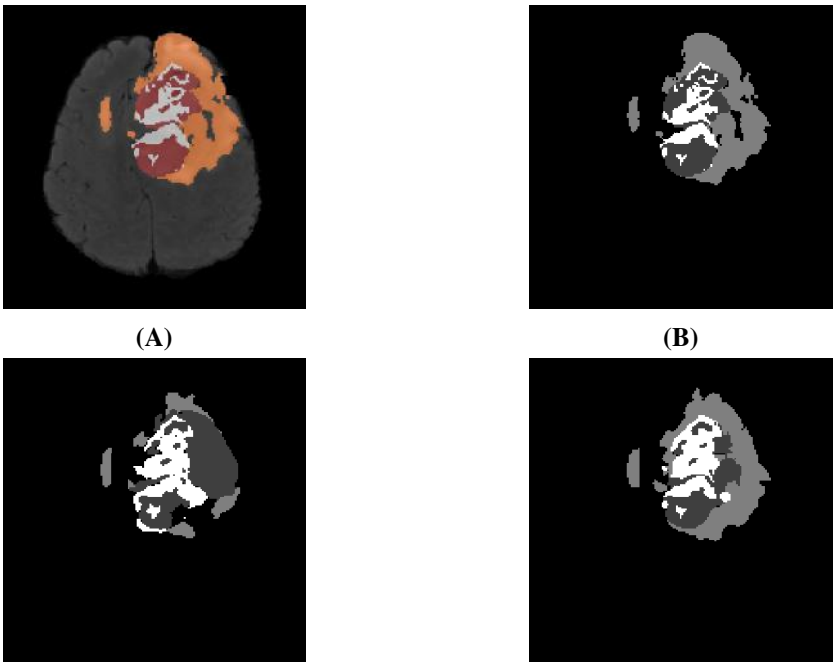
Figure 4.11: Incorrect segmentation results of the network, brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) Segmentation without post-processing (D) Segmentation after post-processing.



(A) **(B)**

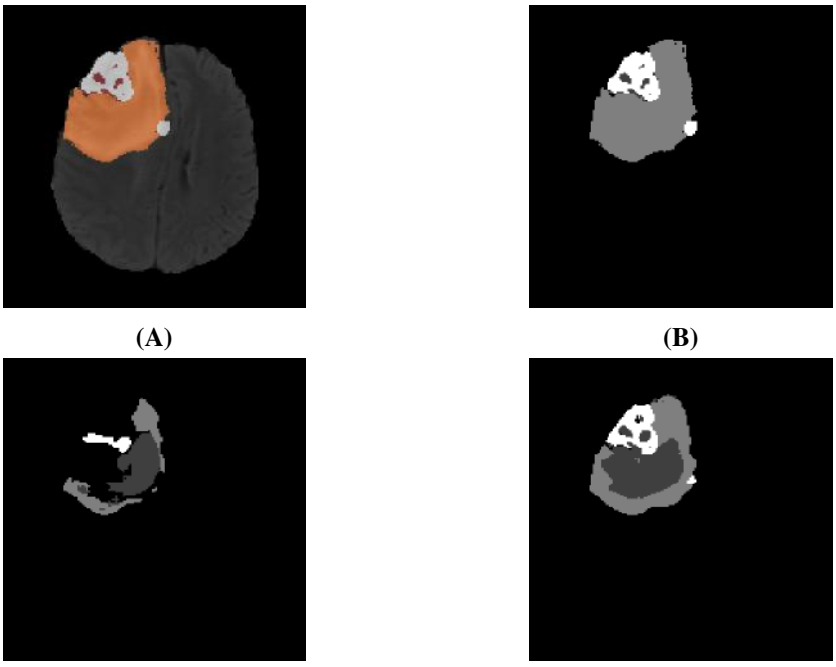
(C) **(D)**

Figure 4.12: Brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) 2D FCNN segmentation (D) 3D FCNN segmentation.



(A) **(B)** **(C)** **(D)**

Figure 4.13: Brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) 2D FCNN segmentation (D) 3D FCNN segmentation.



(A) **(B)** **(C)** **(D)**

Figure 4.14: Brown-necrosis, gray-enhancing tumor, orange-edema (A) FLAIR slice (B) Ground truth (C) 2D FCNN segmentation (D) 3D FCNN segmentation.

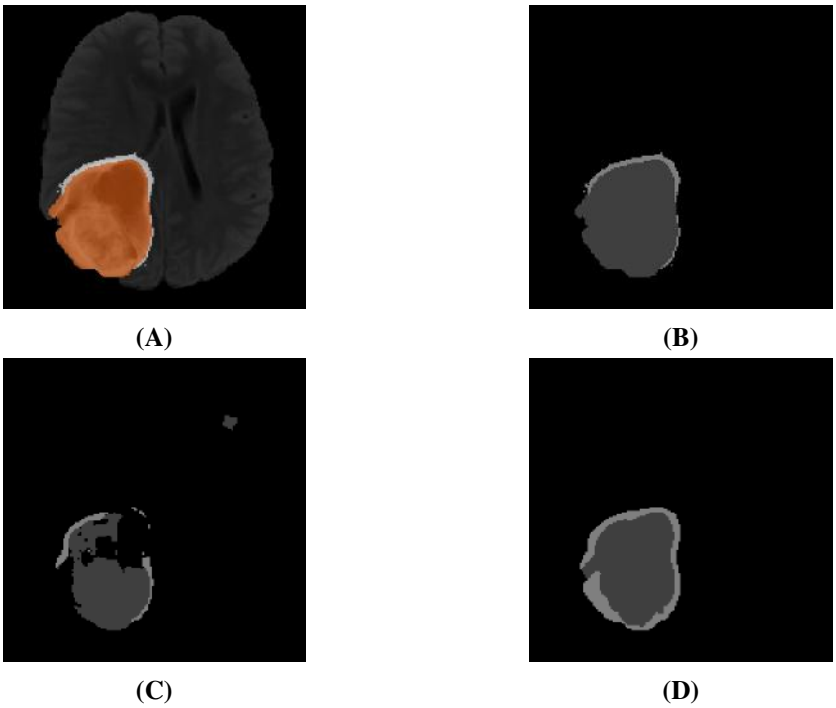


Figure 4.15: Orange-necrosis, gray-enhancing tumor (A) FLAIR slice (B) Ground truth (C) 2D FCNN segmentation (D) 3D FCNN segmentation.

Table 4.14: DSC comparision with state-of-the-art ensemble methods.

Reference	Input	# FCNN	# Conv layers	DSC%		
				ET	WT	TC
[110]	3D image	3	32	78	78	79
[111]	2D image and 3D patches	4	72	67	87	76
[112]	2D image	15	22	71	85	71
[113]	3D patches	2	18	59	80	69
[114]	3D patches	4	14	67	87	73
[115]	3D patches	2	22	62	85	76
[116]	2D images	4	72	68	84	73
Proposed [100]	3D patches	**1**	21	76	87	75

4.6 Summary

The chapter shows three FCNN architectures to address tumor segmentation. The 2D encoder-decoder FCNN is a two-layer architecture that trains on 2D slices of

BraTS 2018 training dataset. The parameter initialization technique improves the segmentation of the tumor substructures. The 2D encoder-decoder dense FCNN is also a two-layer architecture where encoder modules use dense connections between the convolution layers. The network trains on 2D slices of the BraTS 2019 training dataset. The network further improves the substructure segmentation with a focal loss function. The 3D encoder-decoder dense FCNN is a three-layer architecture that trains on equiprobable 3D patches of the BraTS 2020 training dataset. Further, the network introduces a combination of dice and focal loss function. The 2D FCNN and 2D dense FCNN use z-score normalization pre-processing, whereas 3D FCNN uses bias field correction, denoising, and z-score normalization pre-processing techniques. The 3D FCNN adds the connected component analysis post-processing to remove spurious segmentation and conversion of enhancing substructure to necrosis if it is below three hundred.

Chapter 5

Overall Survival Prediction

Overall Survival prediction gives the number of days the patient will survive after the treatment is given in the existing state of the tumor. The automated OS prediction includes tumorous feature processing. This feature set may include age, tumor status, gene expression, and other available clinical data. The availability of all such information is a bit difficult in the healthcare domain. The BraTS training dataset provides age, resection status and survival days to help OS prediction. As per Section 2.2.2, the patient's survival is divided in three classes namely, short-term survivors, mid-term survivors and long-term survivors. Figure 5.1 shows the spread of the HGG patients' survival days with age for the BraTS 2020 dataset.

Although the average age of the short-term survivors, mid-term survivors and long-term survivors is 66, 59 and 57 respectively, there is an overlap of age between each class. The need for proper feature selection from the MRI images plays an important role in training the model for accurate prediction. The chapter depicts features extracted from the MRI images and focuses on the feature selection for OS prediction task.

5.1 Radiomics Features

The tumor-related features from various MRI modality images include radiomics features [92]. The radiomics method extracts many features from radiographic images; these features uncover disease characteristics that are not visible to the eye. These features are calculated from the radiographic image for the label

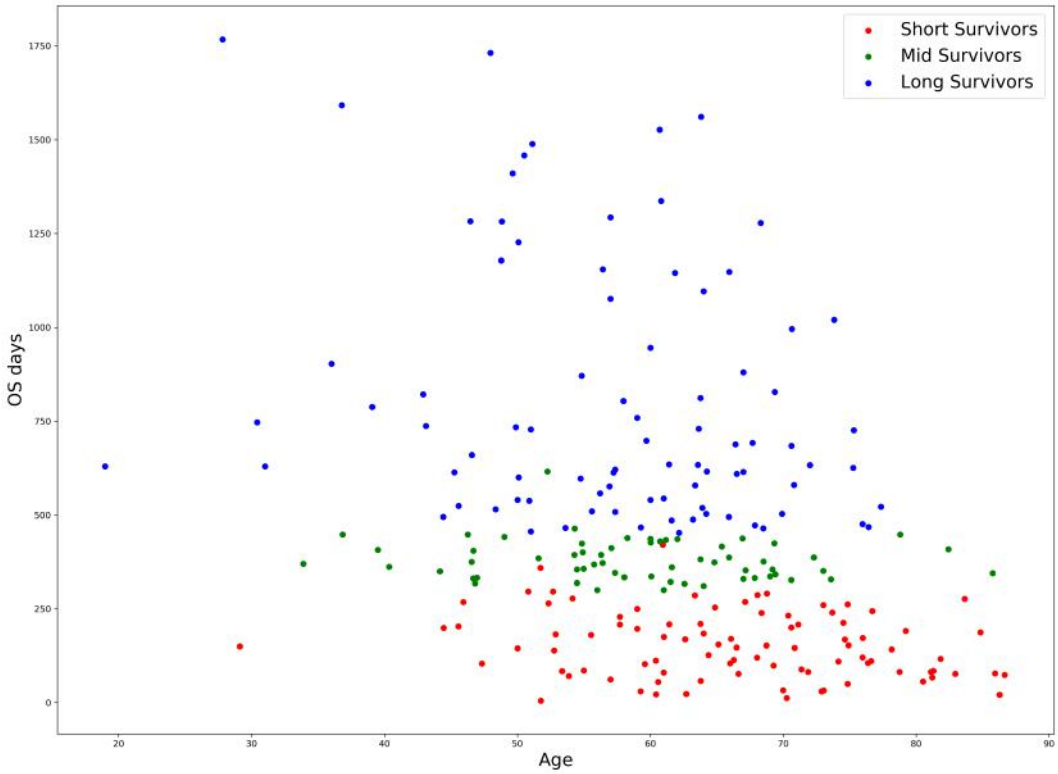


Figure 5.1: Patients' OS days.

mask(segmentation mask) applied to it. The most widely used radiomics features are:

- First order features (19): describe the distribution of voxel intensities within the image region.
- Shape features (16): describe the three-dimensional shape and size of the ROI.
- Gray Level Co-occurrence Matrix (GLCM) (26).
- Gray Level Size Zone Matrix (GLSZM) (16).
- Gray Level Run Length Matrix (GLRLM) (16).
- Neighbouring Gray Tone Difference Matrix (NGTDM) (5).
- Gray Level Dependence Matrix (GLDM) (15).

All radiomics features can be calculated either from the original image or the derived image (obtained after applying filters). The shape features are not dependent on gray value, and are extracted from the labelled mask. These features are common

across all the modality images. Other statistics related to the tumor are: number of tumorous voxels, number of necrosis voxels, number of edema voxels, number of enhancing tumor voxels, the spread of a tumor, amount of tumor. After extracting such features from the image, it is highly desirable to select an appropriate set of features for accurate OS prediction.

5.2 Random Forest Classifier

Authors in [117] tested 179 classifiers on 121 different data sets and identified that random forest is the best classifier amongst the other classifiers. The same finding is supported by OS prediction accuracy as shown in Appendix D, Table D.1, where various classifiers are compared based on number of features on the BraTS data sets. Section 5.1 describes various features which can be extracted either from the tumor label images or from the MRI images. According to [118], the radiomics features are multicollinear which introduces redundancy. This requires proper selection of discriminative features to train RFC for accurate OS prediction. Widely used features for OS prediction are:

- Age (Available with dataset)
- Volumetric (statistical features extracted from the tumor)
- Shape Features (gray level independent features): According to [119], the shape features help in the accurate study of Glioma tumors with the morphological features to study shape irregularities.
- Other Radiomics features (extracted based on the tumor mask)

A RFC of 20 decision trees with information gain as the splitting criteria is used with different combination of above mentioned features to classify the OS prediction in three different classes, i.e., short-term survivors(0), mid-term survivors(1), and long-term survivors(2). The training and test sets contain 130 and 33 patients, respectively, from the BraTS 2018 dataset. RFC is trained on the whole tumor features extracted from the network's segmentation results explained in section 4.3.

Out of various combinations of the features which have been experimented, the feature combination of age, shape, and volumetric features is found to be the

Table 5.1: RFC accuracy for various feature sets.

Set No.	Features	Training set Accuracy (%)	Test set Accuracy(%)
1	Volumetric	89.2	31.3
2	Shape	86.9	37.5
3	Radiomics	90.0	25.0
4	Age and Volumetric	89.2	34.4
5	Age and Shape	85.4	43.8
6	Age and Radiomics	87.7	31.2
7	Volumetric, Shape and Radiomics	86.9	40.6
8	Age, Shape, Volumetric and Radiomics	86.9	37.5
9	**Age, Volumetric and Shape**	87.7	59.4

best possible discriminative feature set which achieves an accuracy of 59.4% on the test set using RFC. The accuracy of the RFC for different feature sets is as shown in Table 5.1.

In addition to the accuracy, the Area Under the receiver operating characteristic Curve (AUC) is another metric to measure the performance of the classifier. The AUC measures the prediction's rank and not its absolute value. It also measures the quality of the prediction irrespective of the classification threshold. The AUC for the training set is approximately equal for all sets, as shown in Figure 5.2. This shows that all feature sets perform equally well. Figure 5.3 shows the AUC for the test set which clearly indicates the better performance of the RFC on the Age, Volumetric and Shape features.

5.3 Random Forest Regressor

The BraTS challenge requires the submission of exact number of OS days, which are further classified in the short-term, mid-term and long-term survivors by the online tool [1]. Hence, Random Forest Regressor (RFR) is used instead of RFC to

[1] https://ipp.cbica.upenn.edu/

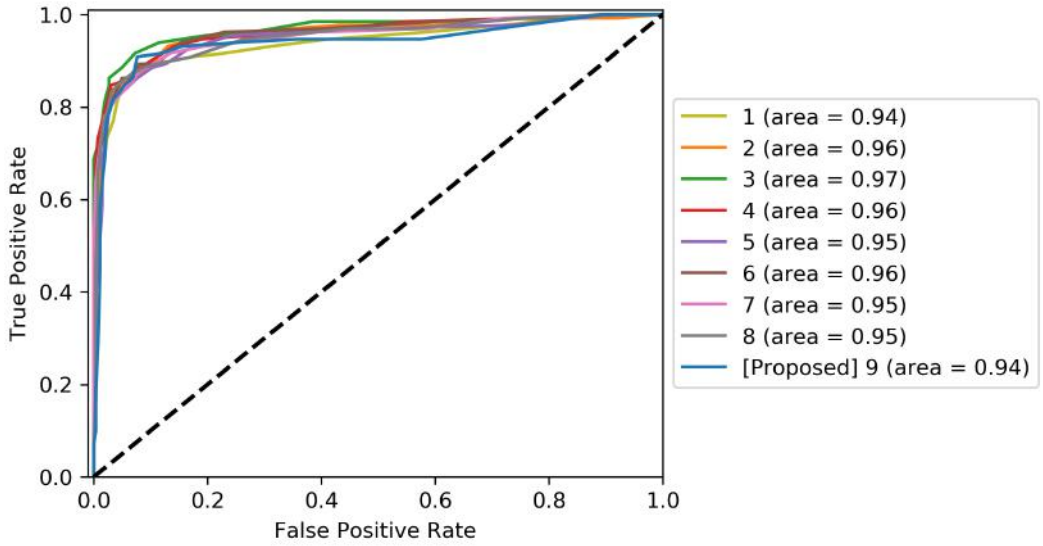


Figure 5.2: Training AUC for various feature sets.

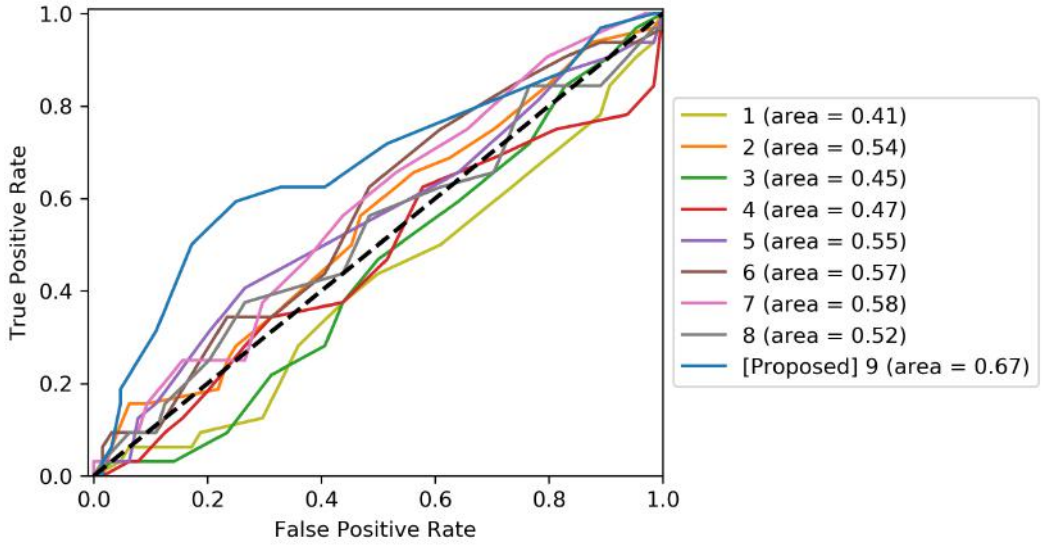


Figure 5.3: Validation AUC for various feature sets.

Table 5.2: OS accuracy for training, validation and test datasets.

Dataset	Accuracy(%)	MeanSE	MedianSE	StdSE	SpearmanR
Training	56.4	073144.54	022891.69	0136542.53	0.60
Validation	58.6	105061.87	016460.89	0188752.44	0.40
Test	57.9	374998.78	046483.36	1160428.92	0.43

generate prediction of OS days. Moreover, the feature are extracted from necrosis tumor substructure instead of whole tumor as:

- the necrotic tumor substructure can be easily resected as this is the collection of dead brain tissues. However, the enhancing and edema substructures do not necessarily show the operable tumor tissues.
- the tumor diagnosis and prognosis depend on the necrotic substructure [120].
- The RFR training includes images with both GTR and STR, which again depends on the necrosis.

The features which are mentioned in 5.2 are extracted from necrotic substructure of ground truth segmentation to train RFR. The OS prediction for the validation set considers the trained RFR.

5.3.1 OS Prediction for BraTS 2019

RFR training uses features extracted from 213 ground truth segmentation images. This trained RFR uses the features of the network segmented images for OS prediction. The feature extractor marks all the features as zero except age in the case of absence of necrotic core within the image. Table 5.2 depicts the OS accuracy for the training set as well as the validation set of those images whose resection status is GTR. The accuracy for all the datasets is consistent but the mean squared error (MSE) increases. This is due to the error in the predicted days and actual days.

Table 5.3 shows a comparison between the proposed method with the other existing state-of-the-art techniques on the BraTS 2019 validation set. This method outperforms other methods for accuracy and median squared error (medianSE).

Table 5.3: OS comparison with state-of-the-art methods on the BraTS 2019 validation set.

Model	Accuracy (%)	MeanSE	medianSE	stdSE	SpearmanR
[106]	48.3	088872.50	034369.32	121078.39	0.33
[107]	31.0	107639.33	077906.27	109586.73	0.20
[108]	44.8	100303.72	049284.00	134504.19	0.25
Proposed	**58.6**	105061.87	**016460.89**	188752.44	**0.40**

Table 5.4: OS accuracy for BraTS 2020 training, validation and test datasets.

Dataset	Accuracy (%)	MeanSE	MedianSE	StdSE	SpearmanR
Training	56.8	083165.96	021481.53	0181697.87	0.60
Validation	51.7	116083.48	043974.09	0168176.16	0.22
Test	47.7	382492.36	046612.81	1081670.06	0.33

5.3.2 OS Prediction for BraTS 2020

RFR training uses features extracted from 237 ground truth segmentation images. This trained RFR uses the features of the network segmented images for OS prediction. The feature extractor marks all the features as zero except age in the case of absence of necrotic core within the image. Table 5.4 shows the OS accuracy for the training as well as validation datasets of those images whose resection status is GTR. RFR performs better for all the measures on the training set but gives a poor result on the validation set. This is due to the high error in the prediction of the OS days for the long survivors as shown in Figure 5.4.

Table 5.5 lists a comparison between the proposed classifier and other state-of-the-art techniques on the leaderboard[2] of BraTS 2020. The table includes the results from the BraTS 2020 leaderboard. The methodology for the OS prediction of the mentioned methods is unknown as the leaderboard does not disclose it. The method performance for accuracy, MSE, MedianSE and SpearmanR is better compared to the first ranked method 'redneucon'.

[2]https://www.cbica.upenn.edu/BraTS20/lboardValidationSurvival.html

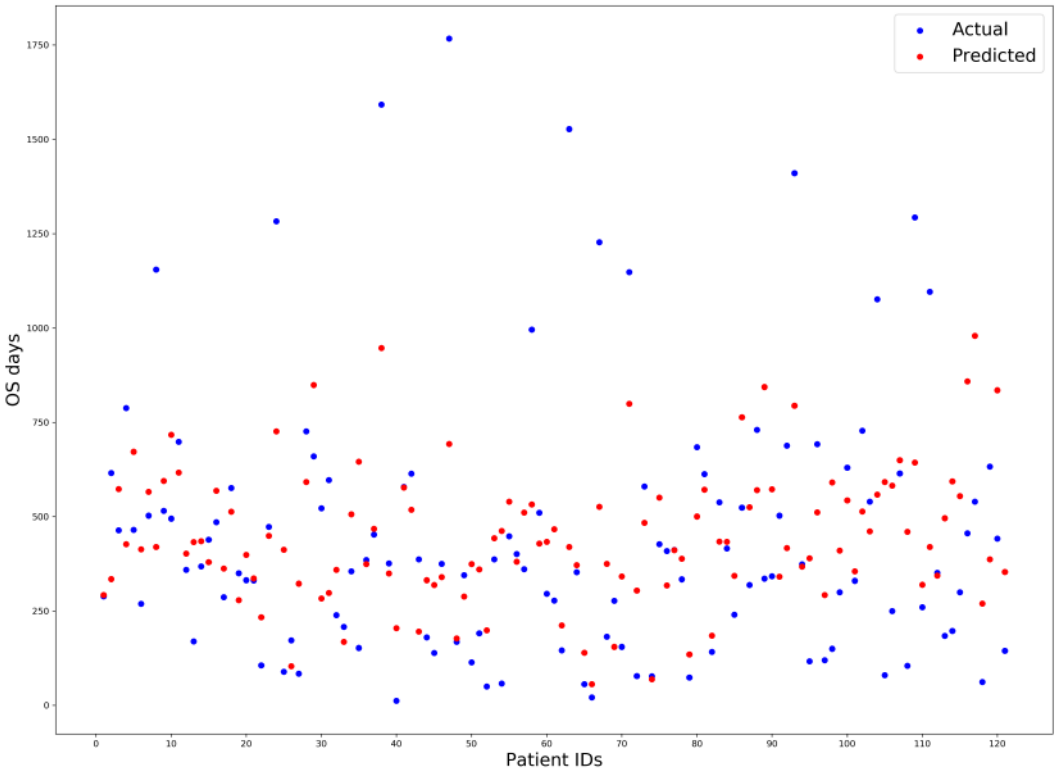


Figure 5.4: Visualization of actual and predicted survival days.

Table 5.5: OS accuracy comparison with state-of-the-art methods on the leaderboard of BraTS 2020 for validation set.

Team Name	Accuracy (%)	MeanSE	medianSE	stdSE	SpearmanR
SCAN	41.4	098704.66	036100.00	152175.57	0.25
redneucon	51.7	122515.76	070305.26	157674.00	0.13
VLB	37.9	093859.54	067348.26	102092.41	0.28
COMSATS-MIDL	48.3	105079.42	037004.93	146375.99	0.13
Proposed	51.7	116083.48	403974.09	168176.16	0.22

5.4 Summary

The chapter focuses on discriminative feature selection from age, statistical, shape, and other radiomic features. The feature selection is made using RFC on BraTS 2018 dataset. RFC is trained on the components extracted for the entire tumor using the network segmentation mask. The features which give the highest accuracy with better ROC are age, statistical, and shape features. The same feature set is used to train RFR for BraTS 2019 and 2020 datasets. As the necrosis is the core substructure of the tumor, the prognosis and other treatment planning depend on the characteristics of this substructure; the RFR is trained on the features selected from the necrosis label of ground truth. In the absence of necrosis, the RFR predicts the OS days with age and marks remaining features to be zero.

Chapter 6

Conclusion and Open Research Challenges

The overall focus of the research was to develop a robust brain tumor segmentation model to address accurate OS prediction.

6.1 Conclusion

The contribution of the work is divided into three parts:

1. Design a 2D encoder-decoder FCNN with better parameter initialization to improve the learning of substructures with proper feature selection for OS prediction.

2. 2D encoder-decoder FCNN with dense connection, improve training for false segmentation results, and handle the class imbalance problem along with OS prediction.

3. 3D encoder-decoder FCNN with dense connection and a proper combination of loss functions to handle the class imbalance issue. This part also includes OS prediction.

The encoder-decoder FCNN proves to perform better where long skip connections between peer layers help in proper feature reconstruction on the decoder side. The first 2D encoder-decoder FCNN trains on the whole tumor; later, the subcomponent network training is initialized with the whole tumor network parameters. This parameter initialization has significantly improved the subcomponent network learning for BraTS 2018 dataset with 285 training images. The second 2D encoder-decoder dense FCNN has incorporated the dense connections to learn the

diversified features from all the preceding layers. Also, the focal loss function of the network training has proved to improve the segmentation results for the BraTS 2019 dataset with 335 training images. Both these networks work on 2D input, and as a result, these networks give quite a poor result for validation sets, which shows that the networks cannot learn the 3D relationship between the voxels. This issue is addressed by a 3D encoder-decoder dense FCNN with an extension to the 2D encoder-decoder dense FCNN. The network has densely connected modules at all the layers with a combined loss function to penalise wrongly classified voxels even within a tiny tumor region in the image. The 3D encoder-decoder dense FCNN segments the images of the BraTS 2020 dataset with 369 training images and 125 validation images with approximately the same evaluations for both the sets.

A variety of tumor radiomics features were extracted from the MRI images for OS prediction. The combination of age, volumetric, and shape features has proven to be the most appropriate feature set to train the RFC and RFR for better OS prediction. The features are extracted for the whole tumor to train the RFC for BraTS 2018 dataset. Training and validation sets are considered from the training set. The features are extracted for the necrosis of ground truth to train the RFR for BraTS 2019 and 2020 datasets. The accuracy of the three datasets is as shown in Table 6.1.

Table 6.1: OS accuracy on BraTS datasets.

	# Images		Accuracy (%)	
Dataset	Training set	Validation set	Training set	Validation set
BraTS 2018	130	30	87.7	59.4
BraTS 2019	213	29	56.4	58.6
BraTS 2020	237	29	56.8	51.7

6.2 Open Research Challenges

The direction of the future research can be two ways: dataset based and model-based.

Dataset: The association of demographic and geographic information with the dataset improves the segmentation as well as OS prediction as tumor occurrence mainly depends on it; e.g., a malignant glioma tumor is more common in whites and less common in African Americans [13]. Besides, it is quite common in men than in women. The inclusion of such information in the dataset can improve the network result and help in a better prognosis. The availability of post-treatment information can improve OS accuracy.

Model: Although the deep learning techniques effectively aid human experts in diagnosing medical images, its black-box nature restricts its clinical use as such networks lack transparency, understandability, and explainability. These limitations do not gain the trust of human experts or even patients. Moreover, explainability study of the network is highly required for domain specific feature selection, which influences the prognosis and treatment planning decision. The explainability of the network can be addressed by:

- Saliency methods: These methods can visualize the insides of the network. They help to identify the weight activation based on the input. E.g., it gives input voxels for the weight activation of intermediate layers of the network. Such methods also help to identify a specific set of convolution layers to find a region of interest.
- The attention mechanism: The neurological disorder relates to the abnormal functionality of a specific brain region. The networks which can identify such biomarkers will improve the explainability of the network. Such networks may work as pre-processing to focus on the brain region to locate the tumor area; this result can be finetuned further to delineate the entire tumor.

The other limitation of deep networks is their significantly high computational cost, which limits their deployability. This cost is introduced mainly due to convolution layers of the network, which takes up the resources for inference. The design optimization of convolution operations and layers can reduce cost of such processes and make them suitable for real time inference and deployment.

Publications by Author

Journal

1. Agravat, Rupal R., Raval, Mehul S.: Brain Tumor Segmentation and Survival Prediction. In: Crimi A., Bakas S. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke, and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science, vol 11992. pp. 338-348. Springer, Cham (2020).
2. Agravat, Rupal R., Raval, Mehul S.: 3D Semantic Segmentation of Brain Tumor for Overall Survival Prediction. In: Crimi A., Bakas S. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke, and Traumatic Brain Injuries. BrainLes 2020. Lecture Notes in Computer Science, Springer, Cham (2021).
3. Agravat, Rupal R., Raval, Mehul S.: A survey on Automated Glioma Brain Tumor Segmentation and Overall Patient Survival Prediction. In: Archives of Computational Methods in Engineering, Springer [Under review].

Conference

1. Agravat, Rupal R., Raval, Mehul S.: Prediction of overall survival of brain tumor patients. In: 2019 IEEE Region 10 Conference(TENCON). pp. 31-35. IEEE (2019).

Book Chapter

1. Agravat, Rupal R., Raval, Mehul S.: Deep learning for automated brain tumor segmentation in MRI images. In: Soft Computing Based Medical Image Analysis. pp. 183-201. Elsevier (2018).

Magazine

1. Agravat, Rupal R., Raval, Mehul S.: Brain Tumor Segmentation. In: CSI Communication, pp. 31-35. (2016)

Bibliography

[1] Department of Neurology Columbia University. *Brain Tumors: Grades*, 2020 (accessed December 3, 2020).

[2] Stuart Currie, Nigel Hoggard, Ian J Craven, Marios Hadjivassiliou, and Iain D Wilkinson. Understanding mri: basic mr physics for physicians. *Postgraduate medical journal*, 89(1050):209–223, 2013.

[3] MIAMI NEUROSCIENCE CENTER. *Brain Tumor Types*, 2020 (accessed October 12, 2020).

[4] Yiming Xiao. *NeuroImaging and Surgical Technologies Lab*, 2020 (accessed October 12, 2020).

[5] Angelika Sebald and Aneurin Kennerley. *Detailed Information on MRI*, 2020 (accessed October 23, 2020).

[6] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S Kirby, John B Freymann, Keyvan Farahani, and Christos Davatzikos. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data*, 4:170117, 2017.

[7] 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-gbm collection. the cancer imaging archive, 2017.

[8] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin Kirby, John Freymann, Keyvan Farahani, and Christos Davatzikos. Segmentation labels and radiomic features for the pre-operative scans of the tcga-lgg collection. *The Cancer Imaging Archive*, 286, 2017.

[9] MathWorks. *Automatic 2d Rigid Body Image Registration*, 2020 (accessed October 12, 2020).

[10] Jiahui Wang, Clement Vachet, Ashley Rumple, Sylvain Gouttard, Clémentine Ouziel, Emilie Perrot, Guangwei Du, Xuemei Huang, Guido Gerig, and Martin Andreas Styner. Multi-atlas segmentation of subcortical brain structures via the autoseg software pipeline. *Frontiers in neuroinformatics*, 8:7, 2014.

[11] Shuang Song, Yuanjie Zheng, and Yunlong He. A review of methods for bias correction in medical images. *Biomedical Engineering Review*, 1(1), 2017.

[12] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993–2024, 2014.

[13] CP Wild, E Weiderpass, and BW Stewart. World cancer report: cancer research for cancer prevention. *Lyon: International Agency for Research on Cancer*, 2020.

[14] Christopher P Wild, Bernard W Stewart, and C Wild. *World cancer report 2014*. World Health Organization Geneva, Switzerland, 2014.

[15] David N Louis, Arie Perry, Guido Reifenberger, Andreas Von Deimling, Dominique Figarella-Branger, Webster K Cavenee, Hiroko Ohgaki, Otmar D Wiestler, Paul Kleihues, and David W Ellison. The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta neuropathologica*, 131(6):803–820, 2016.

[16] Bhawna Goyal, Sunil Agrawal, and BS Sohi. Noise issues prevailing in various types of medical images. *Biomedical & Pharmacology Journal*, 11(3):1227, 2018.

[17] László Lefkovits, Szidónia Lefkovits, and Mircea-Florin Vaida. An atlas based performance evaluation of inhomogeneity correcting effects. *MACRo 2015*, 1(1):79–90, 2015.

[18] Nicholas J Tustison, Brian B Avants, Philip A Cook, Yuanjie Zheng, Alexander Egan, Paul A Yushkevich, and James C Gee. N4itk: improved n3 bias correction. *IEEE transactions on medical imaging*, 29(6):1310–1320, 2010.

[19] Spyridon Bakas, Mauricio Reyes, Andras Jakab, Stefan Bauer, Markus Rempfler, Alessandro Crimi, Russell Takeshi Shinohara, Christoph Berger, Sung Min Ha, Martin Rozycki, et al. 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.

[20] Rupal R Agravat and Mehul S Raval. Deep learning for automated brain tumor segmentation in mri images. In *Soft Computing Based Medical Image Analysis*, pages 183–201. Elsevier, 2018.

[21] N Ramesh, J-H Yoo, and IK Sethi. Thresholding based on histogram approximation. *IEE Proceedings-Vision, Image and Signal Processing*, 142(5):271–279, 1995.

[22] Yuh-Tay Liow. A contour tracing algorithm that preserves common boundaries between regions. *CVGIP: Image Understanding*, 53(3):313–321, 1991.

[23] Hermann Ney. A comparative study of two search strategies for connected word recognition: Dynamic programming and heuristic search. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (5):586–595, 1992.

[24] HP Ng, SH Ong, KW Foong, and WL Nowinski. An improved watershed algorithm for medical image segmentation. In *Proceedings 12th International Conference on Biomedical Engineering*, 2005.

[25] Antonio M López, Felipe Lumbreras, Joan Serrat, and Juan J. Villanueva. Evaluation of methods for ridge and valley detection. *IEEE Transactions on pattern analysis and machine intelligence*, 21(4):327–335, 1999.

[26] Milan Sonka, Vaclav Hlavac, and Roger Boyle. *Image processing, analysis, and machine vision*. Cengage Learning, 2014.

[27] Ludvík Tesař, Akinobu Shimizu, Daniel Smutek, Hidefumi Kobatake, and Shigeru Nawano. Medical image analysis of 3d ct images based on extension of haralick texture features. *Computerized Medical Imaging and Graphics*, 32(6):513–520, 2008.

[28] DJ Withey and ZJ Koles. Three generations of medical image segmentation: Methods and available software. *International Journal of Bioelectromagnetism*, 9(2):67–68, 2007.

[29] Marcel Prastawa, Elizabeth Bullitt, Sean Ho, and Guido Gerig. A brain tumor segmentation framework based on outlier detection. *Medical image analysis*, 8(3):275–283, 2004.

[30] Meritxell Bach Cuadra, Claudio Pollo, Anton Bardera, Olivier Cuisenaire, J-G Villemure, and J-P Thiran. Atlas-based segmentation of pathological mr brain images using a model of lesion growth. *IEEE transactions on medical imaging*, 23(10):1301–1314, 2004.

[31] D Kwon, H Akbari, X Da, B Gaonkar, and C Davatzikos. Multimodal brain tumor image segmentation using glistr. *MICCAI brain tumor segmentation (BraTS) challenge manuscripts*, pages 18–19, 2014.

[32] Spyridon Bakas, Ke Zeng, Aristeidis Sotiras, Saima Rathore, Hamed Akbari, Bilwaj Gaonkar, Martin Rozycki, Sarthak Pati, and Christos Davatzikos.

Glistrboost: combining multimodal mri segmentation, registration, and biophysical tumor growth modeling with gradient boosting machines for glioma segmentation. In *BrainLes 2015*, pages 144–155. Springer, 2015.

[33] Ke Zeng, Spyridon Bakas, Aristeidis Sotiras, Hamed Akbari, Martin Rozycki, Saima Rathore, Sarthak Pati, and Christos Davatzikos. Segmentation of gliomas in pre-operative and post-operative multimodal magnetic resonance imaging volumes based on a hybrid generative-discriminative framework. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 184–194. Springer, 2016.

[34] Yang Xiao and Jie Hu. Hierarchical random walker for multimodal brain tumor segmentation. *MICCAI Challenge on Multimodal Brain Tumor Segmentation*, 2012.

[35] 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*, pages 18–22, 2013.

[36] Thomas Taylor, Nigel John, Patricia Buendia, and Michael Ryan. Map-reduce enabled hidden markov models for high throughput multimodal brain tumor segmentation. *Multimodal Brain Tumor Segmentation*, 43, 2013.

[37] N Subbanna and T Arbel. Probabilistic gabor and markov random fields segmentation of brain tumours in mri volumes. *Proc MICCAI Brain Tumor Segmentation Challenge (BRATS)*, pages 28–31, 2012.

[38] Dimah Dera, Fabio Raman, Nidhal Bouaynaya, and Hassan M Fathallah-Shaykh. Interactive semi-automated method using non-negative matrix factorization and level set segmentation for the brats challenge. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 195–205. Springer, 2016.

[39] C Vijayakumar, Gharpure Damayanti, R Pant, and CM Sreedhar. Segmentation and grading of brain tumors on apparent diffusion coefficient images

using self-organizing maps. *Computerized Medical Imaging and Graphics*, 31(7):473–484, 2007.

[40] Halandur Nagaraja Bharath, S Colleman, Diana Maria Sima, and Sabine Van Huffel. Tumor segmentation from multimodal mri using random forest with superpixel and tensor based feature extraction. In *International MICCAI Brainlesion Workshop*, pages 463–473. Springer, 2017.

[41] JP Serrano-Rubio and R Everson. Brain tumour segmentation method based on supervoxels and sparse dictionaries. brainles 2018. *Springer LNCS*, 11384:210–221, 2019.

[42] Darko Zikic, Ben Glocker, Ender Konukoglu, Antonio Criminisi, Cagatay Demiralp, Jamie Shotton, Owen M Thomas, Tilak Das, Raj Jena, and Stephen J Price. Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel mr. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 369–376. Springer, 2012.

[43] Stefan Bauer, Thomas Fejes, Johannes Slotboom, Roland Wiest, Lutz-P Nolte, and Mauricio Reyes. Segmentation of brain tumor images based on integrated hierarchical classification and regularization. In *MICCAI BraTS Workshop. Nice: Miccai Society*, page 11, 2012.

[44] Michael Goetz, Christian Weber, Josiah Bloecher, Bram Stieltjes, Hans-Peter Meinzer, and Klaus Maier-Hein. Extremely randomized trees based brain tumor segmentation. *Proceeding of BRATS challenge-MICCAI*, pages 006–011, 2014.

[45] Jens Kleesiek, Armin Biller, Gregor Urban, U Kothe, Martin Bendszus, and F Hamprecht. Ilastik for multi-modal brain tumor segmentation. *Proceedings MICCAI BraTS (brain tumor segmentation challenge)*, pages 12–17, 2014.

[46] Raphael Meier, Stefan Bauer, Johannes Slotboom, Roland Wiest, and Mauricio Reyes. Appearance-and context-sensitive features for brain tumor segmentation. *Proceedings of MICCAI BRATS Challenge*, pages 020–026, 2014.

[47] Junichiro Toriwaki and Hiroyuki Yoshida. *Fundamentals of three-dimensional digital image processing*. Springer Science & Business Media, 2009.

[48] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[49] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[50] Jose Bernal, Kaisar Kushibar, Daniel S Asfaw, Sergi Valverde, Arnau Oliver, Robert Martí, and Xavier Lladó. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. *Artificial intelligence in medicine*, 95:64–81, 2019.

[51] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88, 2017.

[52] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[53] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In

Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2818–2826, 2016.

[54] Sasha Targ, Diogo Almeida, and Kevin Lyman. Resnet in resnet: Generalizing residual architectures. *arXiv preprint arXiv:1603.08029*, 2016.

[55] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[56] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[57] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[58] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

[59] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2017.

[60] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. Brain tumor segmentation with deep neural networks. *Medical image analysis*, 35:18–31, 2017.

[61] Gregor Urban, M Bendszus, F Hamprecht, and J Kleesiek. Multi-modal brain tumor segmentation using deep convolutional neural networks. *MICCAI BraTS (brain tumor segmentation) challenge. Proceedings, winning contribution*, pages 31–35, 2014.

[62] Darko Zikic, Yani Ioannou, Matthew Brown, and Antonio Criminisi. Segmentation of brain tumor tissues with convolutional neural networks. *Proceedings MICCAI-BRATS*, pages 36–39, 2014.

[63] Sérgio Pereira, Adriano Pinto, Victor Alves, and Carlos A Silva. Deep convolutional neural networks for the segmentation of gliomas in multi-sequence mri. In *BrainLes 2015*, pages 131–143. Springer, 2015.

[64] Ramandeep S Randhawa, Ankit Modi, Parag Jain, and Prashant Warier. Improving boundary classification for brain tumor segmentation and longitudinal disease progression. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 65–74. Springer, 2016.

[65] Mohammad Havaei, Francis Dutil, Chris Pal, Hugo Larochelle, and Pierre-Marc Jodoin. A convolutional neural network approach to brain tumor segmentation. In *BrainLes 2015*, pages 195–208. Springer, 2015.

[66] Konstantinos Kamnitsas, Enzo Ferrante, Sarah Parisot, Christian Ledig, Aditya V Nori, Antonio Criminisi, Daniel Rueckert, and Ben Glocker. Deepmedic for brain tumor segmentation. In *International workshop on Brainlesion: Glioma, multiple sclerosis, stroke and traumatic brain injuries*, pages 138–149. Springer, 2016.

[67] Xiaomei Zhao, Yihong Wu, Guidong Song, Zhenye Li, Yong Fan, and Yazhuo Zhang. Brain tumor segmentation using a fully convolutional neural network with conditional random fields. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 75–87. Springer, 2016.

[68] Peter D Chang. Fully convolutional deep residual neural networks for brain tumor segmentation. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 108–118. Springer, 2016.

[69] Richard McKinley, Rik Wepfer, Tom Gundersen, Franca Wagner, Andrew Chan, Roland Wiest, and Mauricio Reyes. Nabla-net: A deep dag-like convolutional architecture for biomedical image segmentation. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 119–128. Springer, 2016.

[70] Adria Casamitjana, Santi Puch, Asier Aduriz, and Verónica Vilaplana. 3d convolutional neural networks for brain tumor segmentation: a comparison of multi-resolution architectures. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 150–161. Springer, 2016.

[71] Marc Moreno Lopez and Jonathan Ventura. Dilated convolutions for brain tumor segmentation in mri scans. In *International MICCAI Brainlesion Workshop*, pages 253–262. Springer, 2017.

[72] Hao Dong, Guang Yang, Fangde Liu, Yuanhan Mo, and Yike Guo. Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. In *annual conference on medical image understanding and analysis*, pages 506–517. Springer, 2017.

[73] Xue Feng and Craig Meyer. Patch-based 3d u-net for brain tumor segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2017.

[74] Fabian Isensee, Philipp Kickingereder, Wolfgang Wick, Martin Bendszus, and Klaus H Maier-Hein. No new-net. In *International MICCAI Brainlesion Workshop*, pages 234–244. Springer, 2018.

[75] Fabian Isensee, Philipp Kickingereder, Wolfgang Wick, Martin Bendszus, and Klaus H Maier-Hein. Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge. In *International MICCAI Brainlesion Workshop*, pages 287–297. Springer, 2017.

[76] Liang Chen, Paul Bentley, Kensaku Mori, Kazunari Misawa, Michitaka Fujiwara, and Daniel Rueckert. Drinet for medical image segmentation. *IEEE transactions on medical imaging*, 37(11):2453–2462, 2018.

[77] Konstantinos Kamnitsas, Wenjia Bai, Enzo Ferrante, Steven McDonagh, Matthew Sinclair, Nick Pawlowski, Martin Rajchl, Matthew Lee, Bernhard Kainz, Daniel Rueckert, et al. Ensembles of multiple models and architectures for robust brain tumour segmentation. In *International MICCAI Brainlesion Workshop*, pages 450–462. Springer, 2017.

[78] Richard McKinley, Alain Jungo, Roland Wiest, and Mauricio Reyes. Pooling-free fully convolutional networks with dense skip connections for semantic segmentation, with application to brain tumor segmentation. In *International MICCAI Brainlesion Workshop*, pages 169–177. Springer, 2017.

[79] Andriy Myronenko. 3d mri brain tumor segmentation using autoencoder regularization. In *International MICCAI Brainlesion Workshop*, pages 311–320. Springer, 2018.

[80] Richard McKinley, Raphael Meier, and Roland Wiest. Ensembles of densely-connected cnns with label-uncertainty for brain tumor segmentation. In *International MICCAI Brainlesion Workshop*, pages 456–465. Springer, 2018.

[81] Alain Jungo, Richard McKinley, Raphael Meier, Urspeter Knecht, Luis Vera, Julián Pérez-Beteta, David Molina-García, Víctor M Pérez-García, Roland Wiest, and Mauricio Reyes. Towards uncertainty-assisted brain tumor segmentation and survival prediction. In *International MICCAI Brainlesion Workshop*, pages 474–485. Springer, 2017.

[82] Xiaomei Zhao, Yihong Wu, Guidong Song, Zhenye Li, Yazhuo Zhang, and Yong Fan. 3d brain tumor segmentation through integrating multiple 2d fcnns. In *International MICCAI Brainlesion Workshop*, pages 191–203. Springer, 2017.

[83] RG Rodríguez Colmeiro, CA Verrastro, and Thomas Grosges. Multimodal brain tumor segmentation using 3d convolutional networks. In *International MICCAI Brainlesion Workshop*, pages 226–240. Springer, 2017.

[84] Guotai Wang, Wenqi Li, Sébastien Ourselin, and Tom Vercauteren. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In *International MICCAI brainlesion workshop*, pages 178–190. Springer, 2017.

[85] Chenhong Zhou, Shengcong Chen, Changxing Ding, and Dacheng Tao. Learning contextual and attentive information for brain tumor segmentation. In *International MICCAI Brainlesion Workshop*, pages 497–507. Springer, 2018.

[86] Dmitry Lachinov, Evgeny Vasiliev, and Vadim Turlapov. Glioma segmentation with cascaded unet. In *International MICCAI Brainlesion Workshop*, pages 189–198. Springer, 2018.

[87] Yanwu Xu, Mingming Gong, Huan Fu, Dacheng Tao, Kun Zhang, and Kayhan Batmanghelich. Multi-scale masked 3-d u-net for brain tumor segmentation. In *International MICCAI Brainlesion Workshop*, pages 222–233. Springer, 2018.

[88] Zeyu Jiang, Changxing Ding, Minfeng Liu, and Dacheng Tao. Two-stage cascaded u-net: 1st place solution to brats challenge 2019 segmentation task. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Revised Selected Papers, Part I 5*, pages 231–241. Springer International Publishing, 2020.

[89] Yuan-Xing Zhao, Yan-Ming Zhang, and Cheng-Lin Liu. Bag of tricks for 3d mri brain tumor segmentation. In *International MICCAI Brainlesion Workshop*, pages 210–220. Springer, 2019.

[90] Richard McKinley, Michael Rebsamen, Raphael Meier, and Roland Wiest. Triplanar ensemble of 3d-to-2d cnns with label-uncertainty for brain tumor segmentation. In *International MICCAI Brainlesion Workshop*, pages 379–387. Springer, 2019.

[91] Zeina A Shboul, Lasitha Vidyaratne, Mahbubul Alam, and Khan M Iftekharuddin. Glioblastoma and survival prediction. In *International MICCAI Brainlesion Workshop*, pages 358–368. Springer, 2017.

[92] Joost JM Van Griethuysen, Andriy Fedorov, Chintan Parmar, Ahmed Hosny, Nicole Aucoin, Vivek Narayan, Regina GH Beets-Tan, Jean-Christophe Fillion-Robin, Steve Pieper, and Hugo JWL Aerts. Computational radiomics system to decode the radiographic phenotype. *Cancer research*, 77(21):e104–e107, 2017.

[93] Ujjwal Baid, Sanjay Talbar, Swapnil Rane, Sudeep Gupta, Meenakshi H Thakur, Aliasgar Moiyadi, Siddhesh Thakur, and Abhishek Mahajan. Deep learning radiomics algorithm for gliomas (drag) model: a novel approach using 3d unet based deep convolutional neural network for predicting survival in gliomas. In *International MICCAI Brainlesion Workshop*, pages 369–379. Springer, 2018.

[94] Leon Weninger, Oliver Rippel, Simon Koppers, and Dorit Merhof. Segmentation of brain tumors and patient survival prediction: Methods for the brats 2018 challenge. In *International MICCAI Brainlesion Workshop*, pages 3–12. Springer, 2018.

[95] Yuexiang Li and Linlin Shen. Deep learning based multimodal brain tumor diagnosis. In *International MICCAI Brainlesion Workshop*, pages 149–158. Springer, 2017.

[96] Xue Feng, Nicholas Tustison, and Craig Meyer. Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic features. In *International MICCAI Brainlesion Workshop*, pages 279–288. Springer, 2018.

[97] Elodie Puybareau, Guillaume Tochon, Joseph Chazalon, and Jonathan Fabrizio. Segmentation of gliomas and prediction of patient overall survival: a simple and fast procedure. In *International MICCAI Brainlesion Workshop*, pages 199–209. Springer, 2018.

[98] Li Sun, Songtao Zhang, and Lin Luo. Tumor segmentation and survival prediction in glioma with deep learning. In *International MICCAI Brainlesion Workshop*, pages 83–93. Springer, 2018.

[99] Rupal R Agravat and Mehul S Raval. Prediction of overall survival of brain tumor patients. In *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, pages 31–35. IEEE, 2019.

[100] Rupal R Agravat and Mehul S Raval. Brain tumor segmentation and survival prediction. In *International MICCAI Brainlesion Workshop*, pages 338–348. Springer, 2019.

[101] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series, the handbook of brain theory and neural networks, 1998.

[102] Nikolas Adaloglou. Intuitive explanation of skip connections in deep learning. *https://theaisummer.com/*, 2020.

[103] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1):60, 2019.

[104] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2009.

[105] Zeyu Jiang, Changxing Ding, Minfeng Liu, and Dacheng Tao. Two-stage cascaded u-net: 1st place solution to brats challenge 2019 segmentation task. In *International MICCAI Brainlesion Workshop*, pages 231–241. Springer, 2019.

[106] Shuo Wang, Chengliang Dai, Yuanhan Mo, Elsa Angelini, Yike Guo, and Wenjia Bai. Automatic brain tumour segmentation and biophysics-guided survival prediction. *arXiv preprint arXiv:1911.08483*, 2019.

[107] Xue Feng, Quan Dou, Nicholas Tustison, and Craig Meyer. Brain tumor segmentation with uncertainty estimation and overall survival prediction. In *International MICCAI Brainlesion Workshop*, pages 304–314. Springer, 2019.

[108] Feifan Wang, Runzhou Jiang, Liqin Zheng, Chun Meng, and Bharat Biswal. 3d u-net based brain tumor segmentation and survival days prediction. In *International MICCAI Brainlesion Workshop*, pages 131–141. Springer, 2019.

[109] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

[110] Minglin Chen, Yaozu Wu, and Jianhuang Wu. Aggregating multi-scale prediction based on 3d u-net in brain tumor segmentation. In *International MICCAI Brainlesion Workshop*, pages 142–152. Springer, 2019.

[111] Soopil Kim, Miguel Luna, Philip Chikontwe, and Sang Hyun Park. Two-step u-nets for brain tumor segmentation and random forest with radiomics for survival time prediction. In *International MICCAI Brainlesion Workshop*, pages 200–209. Springer, 2019.

[112] Sebastian Starke, Carlchristian Eckert, Alex Zwanenburg, Stefanie Speidel, Steffen Löck, and Stefan Leger. An integrative analysis of image segmentation and survival of brain tumour patients. In *International MICCAI Brainlesion Workshop*, pages 368–378. Springer, 2019.

[113] Woo-Sup Han and Il Song Han. Multimodal brain image segmentation and analysis with neuromorphic attention-based learning. In *International MICCAI Brainlesion Workshop*, pages 14–26. Springer, 2019.

[114] Xutao Guo, Chushu Yang, Ting Ma, Pengzheng Zhou, Shangfeng Lu, Nan Ji, Deling Li, Tong Wang, and Haiyan Lv. Brain tumor segmentation based on attention mechanism and multi-model fusion. In *International MICCAI Brainlesion Workshop*, pages 50–60. Springer, 2019.

[115] Thibault Buatois, Élodie Puybareau, Guillaume Tochon, and Joseph Chazalon. Two stages cnn-based segmentation of gliomas, uncertainty quantification and prediction of overall patient survival. In *International MICCAI Brainlesion Workshop*, pages 167–178. Springer, 2019.

[116] Krzysztof Kotowski, Jakub Nalepa, and Wojciech Dudzik. Detection and segmentation of brain tumors from mri using u-nets. In *International MICCAI Brainlesion Workshop*, pages 179–190. Springer, 2019.

[117] Manuel Fernández-Delgado, Eva Cernadas, Senén Barro, and Dinani Amorim. Do we need hundreds of classifiers to solve real world classification problems? *The journal of machine learning research*, 15(1):3133–3181, 2014.

[118] Stefania Rizzo, Francesca Botta, Sara Raimondi, Daniela Origgi, Cristiana Fanciullo, Alessio Giuseppe Morganti, and Massimo Bellomi. Radiomics: the facts and the challenges of image analysis. *European radiology experimental*, 2(1):1–8, 2018.

[119] Ahmad Chaddad, Christian Desrosiers, Lama Hassan, and Camel Tanougast. A quantitative study of shape descriptors from glioblastoma multiforme phenotypes for predicting survival outcome. *The British journal of radiology*, 89(1068):20160575, 2016.

[120] Moo-Kon Song, Joo-Seop Chung, Ho-Young Yhim, Sung-Nam Lim, Seong-Jang Kim, Yeon-Hee Han, Hye-Kyung Shim, Sung-Hoon Jung, Je-Jung Lee,

and Deok-Hwan Yang. Tumor necrosis and complete resection has significant impacts on survival in patients with limited-stage upper aerodigestive tract nk/t cell lymphoma. *Oncotarget*, 8(45):79337, 2017.

[121] László G Nyúl and Jayaram K Udupa. On standardizing the mr image intensity scale. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 42(6):1072–1081, 1999.

[122] Po-Yu Kao, Thuyen Ngo, Angela Zhang, Jefferson W Chen, and BS Manjunath. Brain tumor segmentation and tractographic feature extraction from structural mr images for overall survival prediction. In *International MICCAI Brainlesion Workshop*, pages 128–141. Springer, 2018.

[123] Mohammadreza Soltaninejad, Lei Zhang, Tryphon Lambrou, Guang Yang, Nigel Allinson, and Xujiong Ye. Mri brain tumor segmentation and patient survival prediction using random forests and fully convolutional networks. In *International MICCAI Brainlesion Workshop*, pages 204–215. Springer, 2017.

[124] Rui Hua, Quan Huo, Yaozong Gao, Yu Sun, and Feng Shi. Multimodal brain tumor segmentation using cascaded v-nets. In *International MICCAI Brainlesion Workshop*, pages 49–60. Springer, 2018.

[125] Subhashis Banerjee, Sushmita Mitra, and B Uma Shankar. Multi-planar spatial-convnet for segmentation and survival prediction in brain cancer. In *International MICCAI Brainlesion Workshop*, pages 94–104. Springer, 2018.

[126] Lutao Dai, Tengfei Li, Hai Shu, Liming Zhong, Haipeng Shen, and Hongtu Zhu. Automatic brain tumor segmentation with domain adaptation. In *International MICCAI Brainlesion Workshop*, pages 380–392. Springer, 2018.

Appendix A

List of Abbreviations

- CNS: Central Nervous System
- GM: Gray Matter
- WM: White Matter
- CSF: Cerebrospinal Fluid
- CT: Computed Tomography
- MRI: Medical Resonance Imaging
- NMR: Nuclear Magnetic Resonance
- RF: Radiofrequency
- TE: Echo Time
- TR: Repetition Time
- FLAIR: Fluid-Attenuated Inversion Recovery
- DTI: Diffusion Tensor Imaging
- GPU: Graphical Processing Unit
- SNR: Signal to Noise Ratio
- HGG: High Grade Glioma

- LGG: Low Grade Glioma
- WT: Whole Tumor
- TC: Tumor Core
- ET: Enhancing Tumor
- OS: Overall Survival
- GTR: Gross Total Resection
- STR: Sub Total Resection
- CCA: Connected Component Analysis
- CRF: Conditional Random Field
- DSC: Dice Similarity Coefficient
- FN: False Negative
- FP: False Positive
- TP: True Positive
- TN: True Negative
- HD: Hausdorff Distance
- FCNN: Fully Convolutional Neural Network
- LoG: Laplacian of Gaussian
- ATMSVC: Adaptive Template Moderated Spatially Varying statistical Classification
- EM: Expectation Maximization
- GMM: Gaussion Mixture Model
- RW: Random Walk

- HMRF: Hidden Markov Random Field
- HMM: Hidden Markov Model
- ANN: Artificial Neural Network
- RDF: Random Decision Forest
- ExtraTrees: Extremely Randomized Trees
- DNN: Deep Neural Network
- CNN: Convolutional Neural Network
- ReLU: Rectified Linear Unit
- LReLU: Leaky ReLU
- PReLU: Parametric ReLU
- SReLU: S-shaped ReLU
- ELU: Exponential Linear Unit
- RMSProp: Root Means Square Propogation
- BV: Brain Volume
- BGV: Background Volume
- NTBV: Non Tumor Brain Volume
- NTBGV: Non Tumor Background Volume
- TV: Tumorous Volume
- RFR: Random Forest Regressor
- FRRN: Full Resolution Residual Network
- RU: Residual Unit
- MLP: Multi Layer Perceptron

Appendix B

Comparison: Conventional Methods

Table B.1 covers the comparison of conventional methods for brain tumor segmentation. The comparison is based on the pre-processing techniques, dataset, number of images and DSC. The pre-processing techniques are:

- BFC - Bias field correction
- IN - Image normalization
- HM - Histogram matching
- NR - Noise Removal

Table B.1: Summarization of conventional brain tumor segmentation methods.

Ref.	**Pre-processing**	**Dataset**	**# images**	**DSC (%)**
[31]	BFC, IN	BraTS 2014	200	Validation WT:86,TC:79,ET:59 Test WT:88,TC:83, ET:72
[32]	BFC, IN	BraTS 2015	186	Training WT:88,TC:77,ET:68
[33]	NR, HM	BraTS 2016	200	WT:89,TC:77,ET:67

Table B.1 – continued ...

Ref.	Pre-processing	Dataset	# images	DSC (%)
[34]	NR	BraTS 2012	30	TC:53,ED:25
[35]	-	BraTS 2013	30	HGG WT:84,TC:54,ET:67 LGG WT:81,TC:54, ET:11
[36]	BFC	BraTS 2013	30	HG TC:62,ED:59
[37]	-	BraTS 2012	28	TC:66,ED:56
[40]	IN, HM	BraTS 2017	285	Validation WT:79,TC:67,ET:61 Test WT:77,TC:61, ET:50
[41]	-	BraTS 2018	285	Validation WT:80,TC:63,ET:57 Test WT:73,TC:58,ET:50
[42]	-	Custom	40	TC:85,NC:75,ET:80
[43]	BFC, IN, HM	BraTS 2012	30	HGG TC:62,ED:61 LGG TC:49,ED:35
[44]	BFC, HM	BraTS 2013	208	WT:83,TC:71,ET:68
[45]	HM, IN	BraTS 2014	200	Training WT:84,TC:68,ET:72 Valid/Test WT:87,TC:76, ET:64
[46]	NR, IN, BFC	BraTS 2013	30	Training WT:83,TC:66, ET:58 Challenge WT:84,TC:73, ET:68

Appendix C

Comparison: CNN/FCNN Methods

The comparison of the methods summarized in Table C.1 is based on the pre-processing techniques, FCNN architectures, activation functions, loss functions, post-processing, and DSC achieved. The pre-precessing techniques considered for the comparison are:

1. Intensity clipping: 1 % of highest and lowest frequencies are clipped.
2. Bias field correction.
3. Z-score normalization: $Z = (x - \mu)/\sigma$.
4. Histogram matching: Histogram of all the images is match with the reference histogram.
5. Image normalization: Min-max normalization.
6. Intensity standardization with Nyul approach [121].
7. Image denoising: applies noise filtering for e.g. Gaussian noise filtering.
8. Intensity rescaling : rescaling the intensity range between some specific limits.

The post-processing techniques used for segmentation result improvement are:

1. Connected component analysis: Analyse the connected components and removes the component with the volume below some threshold.
2. Conditional random field.
3. Morphological pperators to remove false positives and fill the holes.
4. Relabelling the output label: Enhancing tumor labels below some threshold are relabelled as necrosis.

Table C.1: Summarization of segmentation methods using CNN architectures.

Ref.	Pre-processing	Input Modality + Augmentation	Patch/ Image	Input View	Network Architecture	# Networks	Ensemble Type	Loss Function	Post-processing	Dataset	DSC(%)
[60]	1, 2 on (T1, T1c), 3	T1, T2, T1c, FLAIR	2D multiscale patches	axial	2D CNN	3	cascaded, parallel	softmax	-	BraTS 2013	Training-HGG: WT:79,TC:68,ET:57 Training-LGG: WT:81,TC:75,ET:54
[61]	5	T1, T2, T1c, FLAIR	3D patches	axial	3D CNN	1	-	softmax	1	BraTS 2013	Test set: WT:87,TC:77,ET:73
[62]	2, downsampling	T1, T2, T1c, FLAIR	2D patches	axial	2D CNN	1	-	softmax	-	BraTS 2013	Training: WT:84,TC:74,ET:69

Table C.1 – continued ...

Ref.	Pre-processing	Input Modality + Augmentation	Patch/ Image	Input view	Network Architecture	# Networks	Ensemble Type	Loss Function	Post-processing	Dataset	DSC(%)
[63]	2, 6	T1, T2, T1c, FLAIR, Rotation	2D patches	axial	2D CNN	2	-	cross entropy	3	BraTS 2015	Training: WT:87,TC:73,ET:68
[64]	1,2,3	T1, T2, T1c, FLAIR, flipping	2D patches	axial	2D CNN	2	-	weighted cross entropy	1	BraTS 2016	Test: WT:87,TC:75,ET:71
[65]	1,2,3	T1, T2, T1c, FLAIR	multiscale patches	axial	2D CNN	4	cascaded, parallel	softmax	1	BraTS 2013	Test: WT:88,TC:79,ET:73
[66]	3	T1, T2, T1c, FLAIR, flipping	3D multiscale patches	axial	3D FCNN	2	parallel	-	2	BraTS 2015	Training: WT:90,TC:76,ET:72
[67]	1, 5	T1c, T2, FLAIR	2D multiscale patches	axial	2D FCNN	2	cascaded	-	1	BraTS 2013	Validation: WT:86,TC:73,ET:62 Test: WT:87,TC:83,ET:76
[68]	5	T1, T2, T1c, FLAIR, Rotation, additional images	2D patches	axial	2D FCNN	2	cascaded	softmax	-	BraTS 2016	Test: WT:89,TC:83,ET:78

Table C.1 – continued ...

Ref.	Pre-processing	Input Modality + Augmentation	Patch/ Image	Input view	Network Architecture	# Networks	Ensemble Type	Loss Function	Post-processing	Dataset	DSC(%)
[69]	-	T1, T2, T1c, FLAIR	2D patches	axial	2D FCNN	1	-	sigmoid	-	BraTS 2012	Training: WT:87,TC:69,ET:56
[70]	-	T1, T2, T1c, FLAIR	3D multiscale patches	axial	3D FCNN	2	parallel	softmax	-	BraTS 2015	Training: 3Dnet1 WT:90,TC:77,ET:63 3DNet2 WT:92,TC:70,ET:74 3DNet3 WT:92,TC:84,ET:77
[71]	5	T1, T2, T1c, FLAIR	2D patches	axial	2D FCNN	1	-	dice	-	BraTS 2017	Validation: WT:78,TC:69,ET:57 Test: WT:69,TC:61,ET:51
[72]	3	T1, T2, T1c, FLAIR	2D images	axial	2D FCNN	1	-	dice	-	BraTS 2015	Training: WT:86,TC:86,ET:65
[73]	1,2,3	T1, T2, T1c, FLAIR	3D patches	axial	3D FCNN	1	-	-	-	BraTS 2017	Validation: WT:84,TC:75,ET:66

Table C.1 – continued ...

Ref.	Pre-processing	Input Modality + Augmentation	Patch/ Image	Input view	Network Architecture	# Networks	Ensemble Type	Loss Function	Post-processing	Dataset	DSC(%)
[74]	3	T1, T2, T1c, FLAIR	3D patches	axial	3D FCNN	1	-	weighted dice + cross entropy	4	BraTS 2018	Validation: WT:91,TC:86,ET:81 Test: WT:88,TC:81,ET:78
[76]	-	T1, T2, T1c, FLAIR	2D images	axial	2D FCNN	1	-	cross entropy	-	BraTS 2017	Training: WT:83,TC:73,ET:65
[77]	3,2,5 and its different combination	T1, T2, T1c, FLAIR	3D multiscale patches	axial	3D FCNN	7	parallel	-	1	BraTS 2017	Validation: WT:90,TC:80,ET:74 Test: WT:87,TC:79,ET:73
[78]	3	T1, T2, T1c, FLAIR	2D images	axial	2D FCNN	2	cascaded	-	1	BraTS 2017	Validation: WT:88,TC:76,ET:71 Test: WT:86,TC:71,ET:71
[79]	3, intensity scale, intensity shift, flip	T1, T2, T1c, FLAIR	3D images	axial	3D FCNN	1	-	weighted average of L2, Dice,KL	-	BraTS 2018	Validation: WT:91,TC:87,ET:82 Test: WT:88,TC:82,ET:77

Table C.1 – continued ...

Ref.	Pre-processing	Input Modality + Augmentation	Patch/ Image	Input view	Network Architecture	# Networks	Ensemble Type	Loss Function	Post-processing	Dataset	DSC(%)
[80]	3 on the individual volume	T1, T2, T1c, FLAIR	2D images	axial, coronal, sagittal	2D FCNN	3	parallel	-	-	BraTS 2018	Validation: WT:90,TC:85,ET:80 Test: WT:89,TC:80,ET:73
[82]	1, 5	T1c, T2, FLAIR	2D images	axial, coronal, sagittal	2D FCNN	3	parallel	-	1	BraTS 2017	Validation: WT:89,TC:79,ET:75 Test: WT:88,TC:75,ET:76
[83]	5,3	T1, T2, T1c, FLAIR	3D patches	axial	3D FCNN	2	cascaded	dice	-	BraTS 2017	Validation: WT:86,TC:69,ET:66 Test: WT:82,TC:67,ET:60
[84]	3	T1, T2, T1c, FLAIR	3D patches	axial, coronal, sagittal	2D FCNN	3	parallel of cascades	dice	-	BraTS 2017	Validation: WT:90,TC:83,ET:78 Test: WT:87,TC:77,ET:78
[85]	-	-	3D multi scale patches	axial	-	7	parallel	-	1, 4	BraTS 2018	Validation: WT:91,TC:87,ET:81 Test: WT:88,TC:80,ET:78

Table C.1 – continued ...

Ref.	Pre-processing	Input Modality + Augmentation	Patch/ Image	Input view	Network Architecture	# Networks	Ensemble Type	Loss Function	Post-processing	Dataset	DSC(%)
[86]	3,1	T1, T2, T1c, FLAIR	2D images	axial	2D FCNN	3	cascaded	mean dice	-	BraTS 2018	Validation: WT:91,TC:84,ET:77
[87]	2,3	T1, T2, T1c, FLAIR	3D patches	axial	3D FCNN	3	cascaded	cross entropy	-	BraTS 2018	Validation: WT:90,TC:83,ET:80
[88]	5	T1, T2, T1c, FLAIR	3D patches	axial	3D FCNN	-	parallel of cascaded	dice	CCA	BraTS 2019	Validation: WT:91,TC:87,ET:80 Test: WT:89,TC:84,ET:83
[89]	5	T1, T2, T1c, FLAIR	3D patches	axial	3D FCNN	-	parallel of parallel	dice	-	BraTS 2019	Validation: WT:91,TC:84,ET:75 Test: WT:88,TC:86,ET:81
[90]	5	T1, T2, T1c, FLAIR	3D patches	axial	3D FCNN	-	-	edge weighted + dice + binary cross	-	BraTS 2019	Validation: WT:90,TC:82,ET:0.78 Test: WT:89,TC:83,ET:81

Appendix D

Comparison: Classifiers for OS Prediction

Table D.1: Comparison of OS prediction on the BraTS validation dataset.

Ref.	Method	# features/ type of network	Dataset	Accuracy (%)
[75]	RFR and MLP	66	BraTS 2017	52.6
[81]	ANN	4	BraTS 2017	42.4
[95]	2D CNN	-	BraTS 2017	55
[96]	Linear Regression Model	9	BraTS 2018	32.1
[98]	RF	14	BraTS 2018	46.4
[93]	MLP	468	BraTS 2018	57.1
[94]	Linear Regressor	1	BraTS 2018	50
[99]	RF	13	BraTS 2018	59
[100]	RF	13	BraTS 2019	58.6
[122]	SVM with linear kernel	-	BraTS 2018	35.7
[123]	RF	4	BraTS 2017	48.5
[124]	ensemble of Xboost, SVM, MLP, DT, RF, LDA	900	BraTS 2018	51.9
[125]	MLP	83	BraTS 2018	54
[126]	Xboost	195	BraTS 2017	50