الجمهورية الجزائرية الديمقراطية الشعبية

Algerian Republic Democratic and Popular

Ministry of Higher Education and Scientific Research

SETIF1 UNIVERSITY - FERHAT ABBAS

FACULTY OF SCIENCES

THESIS

Submitted to Computer Science Department

For the degree of

PHD

Domain : mathematics and computer science

Option: Intelligent System and Machine Learning **Sector: computer science**

BY

YAMINA AZZI

THEME

Deep Learning for Segmentation and Pathology Analysis in Medical Imaging

Held on 29/09/2024 in front of the Jury:

Mohamed Saidi	**MCA**	**Univ. Ferhat Abbas Sétif 1**	**President**
Mohand Tahar Kechadi	**Professor**	**Univ.College Dublin**	**Thesis Director**
Abdelouahab Moussaoui	**Professor**	**Univ. Ferhat Abbas Sétif 1**	**Co-Director**
Kamel Eddine Melkemi	**Professor**	**University of Batna2**	**Examinator**
Mohamed Chaouki Babahenini	**Professor**	**Univ.Mohamed Khider Biskra**	**Examinator**
Layzid Toumi	**MCA**	**Univ. Ferhat Abbas Sétif 1**	**Examinator**

Acknowledgements

I begin by expressing my deepest gratitude to Allah, whose graces and blessings have been my constant source of strength throughout this academic journey.

I extend my appreciation to my esteemed supervisors, Pr.Abdelouahab Moussaoui and Pr.Mohand Tahar Kechadi, for their invaluable guidance, unwavering support, and insightful feedback that have played a pivotal role in shaping this research.

My heartfelt thanks go to all my teachers, from my primary education to university, who have imparted knowledge, instilled a passion for learning, and contributed to my intellectual growth.

I would like to extend my sincer gratitude to the president Dr.Mohamed Saidi and the jury members Pr.Kamel Eddine Melkemi,Pr.Mohamed Chouki Babahenini and Dr.Lyazid Toumi for their insightful feedback. Their invalubale contributions have significantly enhanced the quality of this work.

I am profoundly grateful to my parents for their boundless love, encouragement, and sacrifices. Their unwavering belief in my abilities has been a driving force behind my academic achievements and the bedrock of my success. To my dear siblings and their children, your support and understanding have been a constant source of motivation.

Special appreciation goes to my husband Abdenacer Tezkratt, whose unwavering support, patience, and encouragement have been my pillars of strength. Your belief in my capabilities has been the foundation upon which I built this thesis.

In conclusion, I extend my gratitude to all who have played a role, big or small, in this academic endeavour. Each of you has contributed to completing this PhD thesis, and I am truly thankful for your presence in my life.

Abstract

Since the advent of deep learning, there has been a technological revolution, endowing computers with remarkable abilities in learning, reasoning, thinking, and decision-making. Deep learning has played a pivotal role in facilitating feature extraction, particularly in the realm of computer vision, where it significantly contributes to learning and recognizing image and video content. This impact extends to the field of medical image research, as demonstrated in this thesis. The focus is on the role of deep learning in the precise automatic segmentation of medical images, specifically in segmenting brain glioma tumors and their associated sub-tumoral tissue from multimodal MRI scans. The thesis presents various contributions aimed at achieving accurate segmentation, including a comparative study between machine learning-based methods and deep learning architectures. Additionally, it sheds light on the substantial influence of class imbalance on image segmentation evaluation metrics. Lastly, the thesis proposes a transfer learning-based method to construct a model capable of classifying three types of brain tumors: Meningioma, glioma, and pituitary tumors.

Keywords— Deep learning, Medical Images, Brain Tumor, Image segmentation, Image Classification, Class imbalance

Contents

List of Figures

List of Algorithms

List of Tables

List of Abbreviations

DL . . . . . refers to Deep Learning

ML . . . . . refers to Machine Learning

2D, 3D . . . . two-dimensional or three-dimensional, referring in this thesis to spatial dimensions in an image.

CNN . . . . . Convolutional Neural Network

FCN . . . . . Fully convolutional Neural Network

RF . . . . . Random Forest

SVM . . . . . Support Vector Machines

BRATS . . . . Brain Tumor Segmentation Dataset

WT . . . . . Whole Tumor

TC . . . . . Tumor Core

ET . . . . . Enhancing Tumor

The first step is always the hardest

1 INTRODUCTION

1.1 Introduction

Since the advent of machine learning (ML) techniques, the world of technology has known a great revolution by leaving the explicit programming way where instructions are provided by humans to another way of processing information which is learning and making decisions. Machine Learning is designed to learn from huge quantities of data allowing systems to extract useful knowledge to build predictive models over unseen data. its impact has been particularly significant across diverse sectors, including intelligent robots, self-driving vehicles, enhanced security protocols, and advancements in the healthcare domain[1] one of the most notable advancements within machine learning is the emergence and evolution of Deep learning[2] which is inspired by the human brain to imitate the thinking process. This shift has raised new capabilities in all computer science fields, especially computer vision. The ability of deep learning architectures, notably, Convolutional Neural Networks (CNNs), to automatically learn features from raw data has led to unprecedented achievements in tasks such as image classification, object detection, and semantic segmentation, this has offered Medical image analysis significant advantages through the development of beneficial deep learning basic computer-assisted intervention systems that strengthen diagnostic abilities from pathology structure detection

to disease diagnosis and helped Radiologists to save time and effort. semantic image segmentation is one of the major techniques used to identify and locate pathology structures which is crucial for treatment planning and assessing the progress of the disease, it has recognized an important improvement with the use of deep neural networks where several approaches have been proposed and achieved very competitive results However, they are not accurate enough compared to experts' manual segmentation because of the patient's anatomy variations, the noise, due to artifacts that may occur during the acquisition phase may affect image resolution and the low contrast between tissues, which leads to the absence of well-precise bounds between healthy tissues and abnormalities. However, this process is challenging and still a hot topic.

1.2 Thesis goals

his thesis presents multiple contributions aimed at achieving accurate automatic medical image segmentation, specifically in segmenting brain glioma tumors and their associated sub-tumoral tissue from multimodal MRI scans based on Deep Learning :

- The first contribution is multi-pathway deep learning architecture which has provided promising results against the existing methods
- The second contribution is a deep learning architecture called PU-NET that aims to propose a new way of handling multimodal MRI images, it has provided significant results that outperform the majority of existing methods.
- The third contribution is a comparative study between the Machine learning-based methods here SVM and RF and our deep learning architectures, Furthermore, this thesis also highlights the significant impact of class imbalance on image segmentation evaluation metrics.
- The fourth contribution is a transfer learning-based method from our proposed deep learning architecture used for brain glioma segmentation to build a model

that classifiy three kind of brain tumors :Meningioma,glioma and pituitary tumors.

1.3 Thesis outline

The outline of this dissertation is as follows:

- Chapter 2: Magnetic Resonance Imaging and Brain Pathology: Biomedical Background provides an overview of the medical imaging techniques with a special focus on magnetic resonance imaging techniques and their basic principles and physics behind along with a concise overview of brain anatomy, enhancing our grasp of the brain pathology under examination, which is expounded upon in the concluding sections of this chapter.

- Chapter 3: Image segmentation: state of the art expands on the image analysis techniques used in computer vision, especially the image segmentation technique. It reviews the historical context of image segmentation techniques, particularly focusing on methods that were prevalent before the advent of deep learning and discusses the various traditional approaches, algorithms, and challenges they faced.

- Chapter 4: Deep Learning for image segmentation is devoted to introducing deep learning as a paradigm for image segmentation. Defining key concepts related to deep learning, such as neural networks, convolutional neural networks (CNNs), and the basics of image segmentation using deep learning techniques. Discussing the factors that make deep learning particularly suitable for image segmentation tasks. Also, introduce popular deep learning architectures used for image segmentation, such as U-Net and its variants.

- Chapter 5: Parallel Pathway U-net Architecture for Brain Tumor Segmentation is dedicated to presenting our first contribution by Explaining the rationale behind the architecture, how it differs from traditional U-Net, and the advantages it offers. Providing insights into the training process, hyperparameters,

and modifications made to enhance its performance. Presenting the results obtained from applying this architecture to the BRATS 2018 dataset. and finally, discuss these results and compare them with previous studies.

- Chapter 6: PU-net Architecture for Brain Tumor Segmentation introduces our second proposed solution to address the challenge of brain tumour segmentation. The PU-net architecture represents a significant enhancement and optimization compared to our previous proposal in Chapter 5. We have fine-tuned various aspects, including training parameters, memory utilization, processing speed, and achieved results. This innovative architecture incorporates a novel mechanism for efficiently handling multimodal inputs and has demonstrated remarkable effectiveness when pitted against the existing approaches in the field.

- Chapter 7: Machine Learning for Brain Tumor Segmentation reviews a comparative study between our proposed deep learning architectures in chapter 6 and traditional machine learning models, specifically Support Vector Machine (SVM) and Random Forest. Through this examination, we aim to discern the relative strengths and weaknesses of deep learning vis-à-vis machine learning in the context of brain tumor segmentation. Additionally, we address the pertinent issue of class imbalance, shedding light on how this challenge can significantly impact the outcomes of various evaluation metrics.

- Chapter 8 Transfer Learning from Brain Tumor Segmentation to Brain Tumor Classification: this chapter discusses the transfer learning technique where we propose to use a pre-trained model to segment brain glioma tumors to build a model that classifies brain tumors into meningioma, glioma, and pituitary tumor by just making some changes and taking advantage of the knowledge gained by the pre-trained model.

- Chapter 9 presents a conclusion of the thesis encapsulating the study key findings, and it also offers valuable insights into potential directions for future research expansion.

1.4 Academic Publications

1. Thesis Paper: PU-NET Deep Learning Architecture for Gliomas Brain Tumor Segmentation in Magnetic Resonance Images.(journal class A)

2. International conferences:Class Imbalance and Evaluation Metrics for Medical Image Segmentation with Machine Learning Models(ICSRI-2023)

3. International conferences: Medical Image Analysis and Brain Tumor Segmentation in Magnetic Resonance Images with Deep Learning (ICSRI-2021)

4. International conferences: Semantic Segmentation of Medical Images with Deep Learning: Overview (ICHSMT-2019)+published paper (journal class B)

5. International conferences: Gpu-based pso for bitmap join indexes selection problem in data warehouses (ICAEE-2019)+published paper (master work)

Part I

State of the art

All intelligent thoughts have already been thought;
what is necessary is only to try to think them again.

2

Magnetic Resonance Imaging and Brain Pathology:Biomedical Background

Contents

2.1 Introduction

In the health sector, medical imaging plays a vital role in many disease diagnosing, treatment planning, and patient monitoring, it has transformed the landscape of modern healthcare, enabling clinicians and researchers to peer into the human body with unprecedented clarity and precision. There are several medical imaging modalities, each used for a particular purpose. Magnetic Resonance Imaging (MRI) has emerged as a cornerstone technology for non-invasive visualization of anatomical structures and pathological conditions. With its exceptional soft tissue contrast, versatility, and absence of ionizing radiation, MRI has become an indispensable tool for diagnosing and studying a wide range of medical conditions, particularly those involving brain pathology.

This chapter provides an in-depth exploration of the medical imaging world, with a special focus on the principles of MRI, including the underlying physics and techniques that enable its remarkable imaging capabilities. Furthermore, we introduce the human brain anatomy and vocabulary in addition to the brain gliomas used as the main pathology in this study and other pathologies that are revealed by the MRI modality.

2.2 Medical imaging modalities

since the advent of medical imaging techniques, the health sector has witnessed a great revolution. This revolution is particularly evident in disease diagnosis, tailored treatment applications, and patient monitoring. Medical images offer precise insights into internal bodily structures, thereby aiding in the identification of diverse cancers across various body regions, as well as detecting blood vessel clots, neurological ailments, and fractures in bones. there are several medical imaging techniques spanning from x-rays to magnetic resonance techniques and more [3], some of them are detailed in the next sections.

2.2.1 Radiography

was the first modality in medical imaging discovered by Wilhelm Roentgen after he invented X-rays on November 8, 1895. X-rays are a type of electromagnetic radiation produced by an X-ray tube. A radiography image is obtained by emitted X-rays passed through the patient body and captured by a detector behind him. When the X-rays infiltrate within the body there is a variation in the absorption of X-rays by different tissues for example bones absorb radiation more than soft tissues this is which produces a contrast in the resulting image and allows a clear representation of all structures. it is required to diagnose broken bones, cardiovascular disorders, or some lung pathologies.[3]

2.2.2 Mammography

is breast radiography with an x-ray tube and detector designed specifically for the women's breast, it uses less X-ray energy than general radiography purposes. it aids in revealing lumps that may be benign or malignant like cancers. generally, women after 45 years old are advised to do mammography for early asymptomatic breast cancer screening.[3]

2.2.3 Computed tomography

also called the CT scanner was available in 1970, it is the first modality assisted by a computer. In general CT scanners are built based on X-ray radiation similar to radiography with a different physical design where the patient's body is captured by an x-rays beam applied through various angles by rotating the X-ray tube around the body detectors are opposite to the x-rays tube as usual to collect the results x-rays energies after the absorption phenomena of the body tissues. images are reconstructed to create 2D or 3D images. CT is often used to diagnose: the presence of tumors, cardiac tissue, cardiovascular disease, bone injuries, pulmonary disease, colon health. [3]

2.2.4 Ultrasound

is medical imaging modality completely different from previously mentioned modalities, it is real-time imaging using high sound waves. The ultrasound machine consists of the following components: transducer probe monitor, keyboard, processor (CPU), data storage disk, and printer. The most essential component is the transducer probe. This probe plays a dual role, it produces the sound waves used and receives their echoes back. in an ultrasound exam, a transducer (probe) is placed directly on the skin or inside a body opening. A thin layer of gel is applied to the skin so that the ultrasound waves are transmitted from the transducer through the gel into the body. The ultrasound image is produced based on the reflection of the waves off of the body structures. The strength (amplitude) of the sound signal and the time it takes for the wave to travel through the body provide the information necessary to produce an image. [3]

2.2.5 Nuclear medicine based modalities

Nuclear medicine imaging is the radiology sector in which the specialists use chemical substances of radioactive materials called radiotracers that are given to the patient by injection into the bloodstream, orally or inhaled. Once the tracers reach the appropriate area of the exam and are decayed, they free up energy in the form of gamma rays which are detected by an external radiation detector and reconstructed by a computer. Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) scans are the two most common imaging modalities in nuclear medicine.[3]

2.2.5.1 Single Photon Emission Computed Tomography (SPECT)

is a nuclear imaging equipment with a physical design similar to a CT scan and also assisted by a computer to generate the final images,it is used to capture the distribution of the radioactive tracers in the patient's body using rotated gamma camera detectors that can detect gamma rays emission from the decayed radio-tracers at different angles of the body, The computer translates the collected

information and generates two-dimensional (2D) cross-sections or 3D with stacked 2d cross-sections together.

2.2.5.2 Positron Emission Tomography (PET)

is also similar to SPECT in using radioactive tracers but the main difference is the type of radioactive where in SPECT the decayed radio actives generate direct gamma rays but PET generates particles called positrons and these positrons generate gamma rays photons. Positrons are electrons with positive charge as the electrons have negative charge positrons and electrons annihilate and generate a pair of gamma rays photons that shoot in opposite directions, which will be detected to create 2D or 3D images.

2.3 Magnetic Resonance Imaging

2.3.1 Definition

Magnetic resonance imaging is a non-invasive medical imaging technique that uses non-ionizing radiation to create high detailed 2D or 3D anatomical images to distinguish pathological tissues. It is based on the Nuclear magnetic resonance phenomenon. Generally, MRI is considered as very important test for diagnosing and monitoring diseases in soft tissues. [3, 4, 5]

2.3.2 MRI Components

Magnetic Resonance Imaging (MRI) consists of several components that work together to produce high-quality images of the body's internal structures. These components[4, 5] include:

Magnetic field: A strong, uniform magnetic field is generated by a superconducting magnet or an iron-core magnet, which is the basis for MRI signals.

Radiofrequency (RF) coils: These coils transmit and receive the RF signals that are used to excite and detect the hydrogen protons in the body. There are different types of RF coils, including surface coils, body coils, and head coils, each designed for a specific imaging application.

Gradient coils: These coils generate the gradient magnetic fields that are used to encode the spatial information of the hydrogen protons in the body. The gradient magnetic fields are used to create images with high spatial resolution.

Computer: The computer controls the magnetic field, RF signals, and gradient fields and processes the signals to create images. It also stores and displays the images.

Software: Specialized software is used to control the MRI scanner, acquire images, and perform image analysis, such as image reconstruction, segmentation, and visualization.

Together, these components work to produce high-resolution, non-invasive images of the body's internal structures, which can be used for a variety of medical purposes, including diagnosing diseases, monitoring treatment response, and evaluating anatomy and function.

2.3.3 MRI functioning Principe

Magnetic resonance imaging uses the magnetic resonance property of the protons that are the nucleus of the hydrogen atom which are abundant in the human body as the water (H2O) is the dominant component, especially in soft tissues.[4, 5]

- Proton spin: In chemistry, the hydrogen atom contains a single Proton with a positive charge and one electron revolves around it with a negative charge. The proton turns around its on-axis and forms a tiny magnet like the Earth, which is known as proton spin. In the human body, all protons are randomly aligned in the form that the net magnetization result is equal to 0.

- Precession: when the human body is exposed to a magnetic field B0, all protons are aligned with the B0 axis, some aligned in the same direction to B0, others aligned in the reverse direction but all parallel to the B0 axis and spin around B0. The B0 field affect also the spinning speed which is called precession frequency .

- Excitation and Resonance: radio a frequency is applied equally to the proton's Larmor frequency, and the energy is transferred from the RF to the protons this phenomenon is called resonance once the protons absorb this energy they flip from their longitudinal plane Z axis to its transverse plane XY axis and this is considered as an excitation phase.

- Relaxation: corresponds to the equilibrium state's return of tissue magnetization after stopping the emission of the RF radiofrequency which is escorted by protons liberation RF signal received on a receptive coil to construct the MRI image. but this is not enough to get a high-quality image with contrast because a very potent received signal may lead to a light image and a weak signal leads to a dark image, that is why these two relaxation tissues times were introduced longitudinal relaxation time (T1) and transverse relaxation time (T2) to produce different contrast, and they are different to each tissue composition and depend on the field strength.

- Image construction: The received signal is finally interpreted using Fourier transformation algorithms to generate the MRI image using a computer.

2.3.4 MRI Imaging Sequences and views

MRI sequences depend on some important parameters, that are introduced by the MRI technician in the configuration dashboard[4, 5].

- Repetition Time (TR): is the quantum of time between successive RF pulses applied to the same slice measured in milliseconds.

- Echo Time (TE): is the time between the RF pulse excitation and the receipt of the echo signal. measured in milliseconds.

- Flip angle (θ): is the angle by which the axis of the hydrogen proton shifts from its longitudinal plane (static magnetic field B0) Z axis to its transverse plane XY axis by excitation with RF pulse.

all of these parameters are factors of contrast By tuning them a lot of sequences can be generated for example:

MRI Sequences Parameters	TR (msec)	TE (msec)
T1-Weighted (short TR and TE)	500	14
T2-Weighted (long TR and TE)	4000	90
Flair (very long TR and TE	9000	114

in addition to multiple sequences, the MRI scan produces three anatomical views: axial, sagittal, and coronal

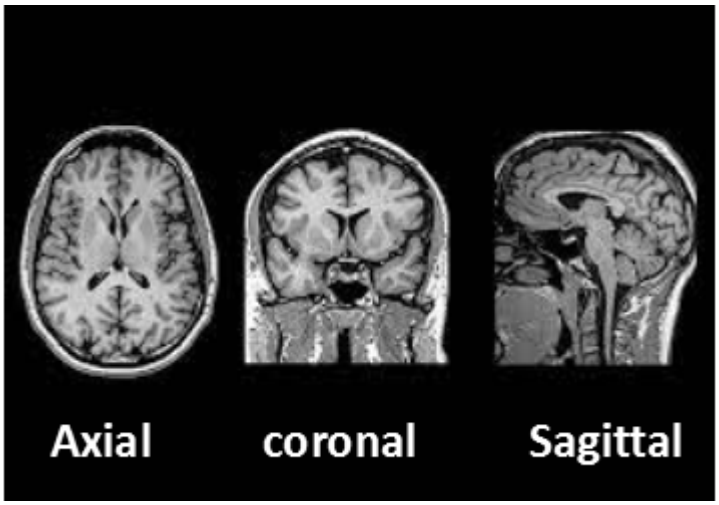


Figure 2.1: The Brain Mri Views

2.4 Brain Anatomy

2.4.1 General Structure

The brain is the central processing and controlling organ in the human body; it is made of 60% fat and the remaining 40% is a combination of water, protein, carbohydrates and salts which are in the form of blood vessels and nerves including neurons and glial cells. It consists of three main parts: cerebrum, cerebellum, and brain stem which is related to the spinal cord to form the central nervous

system that controls all activities of the body. It processes and interprets the information received from the five sensory systems and makes decisions to be executed by the appropriate body part.[6, 7]

- The cerebrum: represents the largest part of the brain, it is split into two halves called hemispheres each controls the opposite half of the body, and also each hemisphere is divided into four regions known as lobes: frontal, parietal temporal, and occipital lobe, that perform an important function.

- The cerebellum: resides in the posterior cranial fossa inferior to the cerebrum, it maintains the coordination of movements and balance.

- Brain Stem: performs the role of a main bridge from the cerebrum and the cerebellum to the spinal cord, it controls many automatic functions like heart rate, breathing, digestion, wake and sleep cycles, and body temperature... It contains also a set of nerve fibers that carry signals controlling muscles and sensations between the cerebrum and the rest of the body [6, 7]

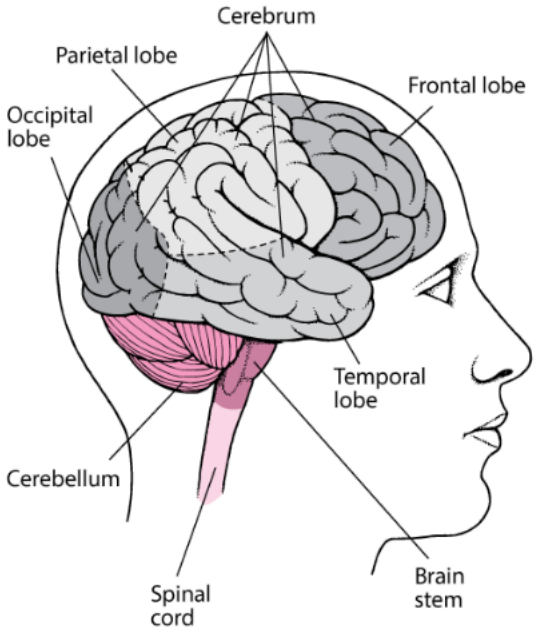


Figure 2.2: The Brain Anatomy

2.4.2 Protective Brain Structures

The human brain is surrounded by a protective system, encased in several layers from outer to inner [6, 7]:

- Skull (Cranium): is the bony protective structure that surrounds the whole brain.
- Meninges: the under skull cover, encompasses three layers of protective membranes from outer to inner: the dura mater, arachnoid mater, and pia mater. The meninges provide additional padding and support for the brain.
- Cerebrospinal Fluid (CSF): is a liquid that surrounds the brain and the spinal cord, it protects them from injuries and shocks and serves as a nutrient and hormone delivery and waste removal system for the brain

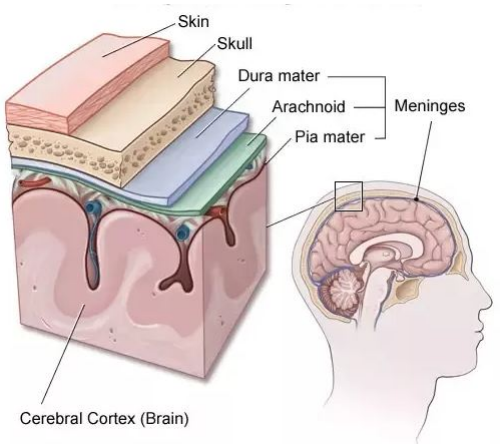


Figure 2.3: The Brain Protective Structure

2.4.3 Cerebral Tissues and Cells

There are two main tissues in the brain white matter and gray matter[6, 7].

- Gray Matter (GM): refers to the darker outer portions, it is a central nervous system tissue part that is made of the neuronal cell bodies also called soma, and the glial cells which provide them nourishment to assist their functioning. It is primarily responsible for the information processing and interpretation

- White Matter (WM): refers to the lighter inner portions, containing the axons, which are the long stems that connect the neurons covered with myelin. It is in charge of ensuring the information transfers between the different parts of the nervous system

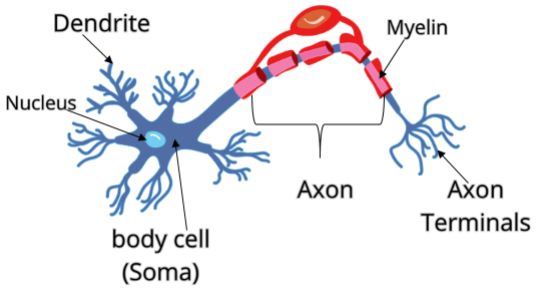


Figure 2.4: The Neuron Anatomy

2.5 Brain Pathologies

As with all human organs, the brain is also expected to several diseases and disorders each with its causes, Symptoms, and treatments like Alzheimer's[8], Parkinson's, Epilepsy, Multiple Sclerosis[9] and brain tumors. This study focused on three types of brain tumors Gliomas, Pituitary tumors, and meningioma tumors.

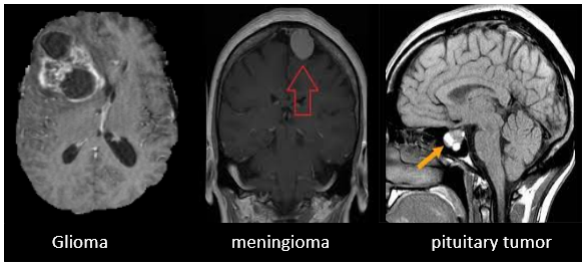


Figure 2.5: Brain Tumors

2.5.1 Gliomas

Gliomas are among the most life-threatening types of cancer that affect the human brain or spinal cord. They are a common type of primary brain tumors, where they

mainly grow in the glial cells opposite to metastasis cancers, which grow anywhere in the body and spread to the brain.[10, 11, 12, 13]

2.5.1.1 Type of Gliomas

Gliomas encompass a diverse range of types, exhibiting variations in growth rates and malignancy. The specific glioma type diagnosed plays a crucial role in determining the severity of the condition and guides the healthcare team in devising an appropriate treatment plan. Gliomas are typically classified into two categories: low-grade gliomas and high-grade gliomas. [14]

1. Low-grade gliomas: are regarded as benign tumors, they are characterised by a slow growth rate and regular morphological shapes in the outward appearance. This category can be treated surgically in the case of early diagnosis and the survival rate may reach up to 10 years, but sometimes, it is possible to be recurrent and grow again, if regular examinations and safety precautions are not applied they may spread and become life threatening. Treatment for Low-Grade Gliomas:

   Surgery: Surgical removal of the tumor is often the first-line treatment for low-grade gliomas. The goal is to remove as much of the tumor as possible while preserving neurological function. Complete removal may not always be feasible, especially if the tumor is located in a critical area of the brain.

   Radiation Therapy: Radiation therapy may be used to target any remaining tumor cells after surgery or as the primary treatment in cases where surgery is not possible.

   Chemotherapy: Some low-grade gliomas may be treated with chemotherapy, either alone or in combination with surgery and/or radiation therapy. Chemotherapy drugs can be administered orally or intravenously.

   Monitoring: Due to the potential for recurrence or transformation into higher-grade tumors, individuals with low-grade gliomas are often closely monitored

through regular imaging studies (such as MRI scans) to detect any changes in the tumor.

2. High-grade gliomas: are more aggressive and still considered untreatable because of their fast-growing, complete restriction is impossible because of their irregular contours in the morphology shape, which means tumor cells quickly affect random nearby healthy cells. Often, a set of treatment options are used after surgeries, such as radiotherapy, chemotherapy, or combined approaches. These tumors have a high recurrent probability and a poor survival rate of 2 to 5 years.

2.5.2 Pituitary tumors

it refers to an abnormal mass that arises in the pituitary gland which is a small, pea-sized gland located at the base of the brain, it is also known as the master gland due to its vital role in regulating various hormonal functions in the body and controlling other glands such as thyroid gland, adrenal glands and reproductive glands (ovaries and testes)
Type of Pituitary tumors: Most pituitary tumors can be classified as benign tumors but they can affect the normal functioning of the gland leading to an overproduction or underproduction of hormones. There are different types of pituitary tumors, and they may be classified based on their size or the specific type of cells they originate from [15]. Some common types include:

1. Prolactinomas: These tumors secrete excessive amounts of prolactin, a hormone that stimulates breast milk production. They are usually benign.

2. Adenomas: The majority of pituitary tumors are adenomas, which are usually benign and arise from the gland's tissue.

3. Growth hormone-secreting tumors: These tumors lead to an excess production of growth hormone, causing acromegaly in adults or gigantism in children.

4. ACTH-secreting tumors: These tumors can cause the overproduction of adrenocorticotropic hormone (ACTH), resulting in Cushing's disease.

5. Non-functioning tumors: These tumors do not produce excess hormones but can cause symptoms due to their size, such as headaches or vision problems.

2.5.3 Meningioma Tumor

it is a kind of tumor that grows in the meninges which are the three protective layers that cover the brain and spinal cord [16].
Type of Meningioma: Meningiomas are mostly benign and characterised by a slow growth rate. However, some can be aggressive and exhibit faster growth. Therefore, they are grouped into three grades based on their histological characteristics.

1. Grade 1: low grade often benign with a slow growth rate

2. Grade 2: Atypical mid-grade tumors with higher recurrence rate after surgical remover.

3. Grade 3: Anaplastic is considered malignant (cancerous) with faster growth speed.

2.5.4 Brain Tumor Diagnostic

Early detection and accurate delineations are crucial for any cancer type, mainly brain tumors, to apply treatments correctly, reduce mortality rate, and ensure good patient quality of life. For this, The healthcare provider takes a detailed medical history, including any symptoms the individual may be experiencing and a neurological examination is conducted to assess motor function, reflexes, sensory perception, and other neurological signs. then, different imaging modalities are used. These include Magnetic Resonance Imaging (MRI) or Computed tomography (CT) to detect these abnormal masses. Currently, MRI is recognized as the best non-invasive diagnostic tool for the majority of brain pathologies[17, 18], due to its potential in soft tissue imaging based on the hydrogen atoms magnetization and its diversity in imaging modalities (T_1, T_2 weighted images, flair or using

contrast enhancement agent gadolinium to obtain T_1 images), which help to exhibit complementary information about tumor segments. It also provides 3D high-resolution detailed anatomical structures in a safe way without radiation fears[19, 20], after that a segmentation process is applied. Segmentation is a crucial step in medical image analysis for identifying and delineating the cancer[21] This process helps in characterizing the tumor, measuring its size, and assisting in treatment planning. There are two main approaches to performing segmentation: manual segmentation and automatic (or computer-aided). In manual segmentation, a trained human expert, often a radiologist or a neuroscientist, visually examines medical images and manually outlines the tumor boundaries. it is advantageous regarding human expertise where Trained professionals can leverage their knowledge and experience to accurately identify and delineate tumor regions. but it can be a tedious and time-intensive process, especially for large datasets, and greatly depends on human experiences and skills where There may be variations in interpretations among different experts. Automatic segmentation relies on computer algorithms and image processing techniques to identify and delineate tumor regions without direct human involvement. This approach offers notable advantages, including high efficiency in analyzing large datasets swiftly, rendering it suitable for high-throughput applications. Moreover, the use of algorithms ensures result consistency, minimizing variability associated with human subjectivity. However, challenges exist, particularly in the context of algorithm performance, as the accuracy of automatic segmentation is contingent on the effectiveness of the specific algorithm employed. Additionally, tumors with irregular shapes or locations may present difficulties for automated methods, highlighting the need for ongoing research and development to enhance the robustness of these segmentation techniques[22, 23]. This which we are going to discuss in the next chapters.

2.6 Conclusion

In conclusion, this chapter has provided a comprehensive exploration of the existing medical imaging modalities focusing on Magnetic Resonance Imaging (MRI) as a

pivotal medical imaging modality, particularly in the context of brain anatomy and pathology By delving into the principles and techniques of MRI, we have gained insight into its unparalleled ability to produce detailed and non-invasive images of the brain's internal structures. the discussion extended to the nuanced world of brain anatomy and pathology, with a specific focus on gliomas. From defining gliomas to elucidating the diagnostic process through MRI, we have unraveled the intricate relationship between this imaging modality and the detection, characterization, and understanding of these tumors.

Don't Wait for Opportunity, Create it

3

Image Segmentation:State of the art

Contents

3.1 Introduction

Image segmentation using deep learning has outstripped the results of traditional methods which rely on manual feature engineering and classical algorithms. This chapter exhibits the most famous image analysis techniques, focusing on image segmentation and its history and state-of-the-art techniques that were prevalent before the advent of neural networks. it also delves into their application on brain tumour segmentation by citing some proposed approaches within this pre-deep

learning epoch to comprehend their role and limitations in tackling the complexities of medical image segmentation.

3.2 image Analysis

in computer vision, image analysis is a branch of knowledge in image processing that is focused on the extraction of useful information from digital images to gain a deeper understanding and context. It is a multi-disciplinary field that combines computer science, mathematics, and statistics to develop algorithms and methods for processing and analyzing images. it is widely used in many application areas, such as medicine, biology, remote sensing, and security. [23] there are three main levels of image analysis:

- Classification: refers to analyzing a set of images and categorizing the entire image into one of several predefined classes, in other words, assigning a label to each image example

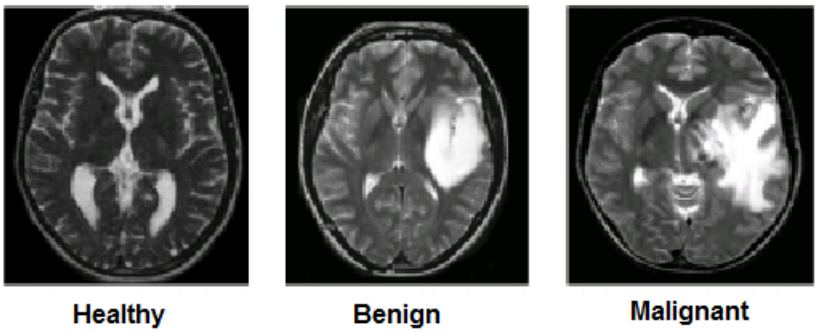


Figure 3.1: Example of image Classification task

- object detection: is to spot the area of interest by drawing a kind of rectangle called a bounding box which means locating and classifying objects into the belonging classes.

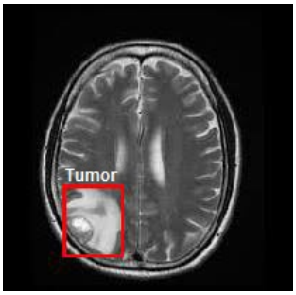


Figure 3.2: Example of Tumor Detection

- Segmentation:is to identify and locate the pixels of the areas of interests as example shows

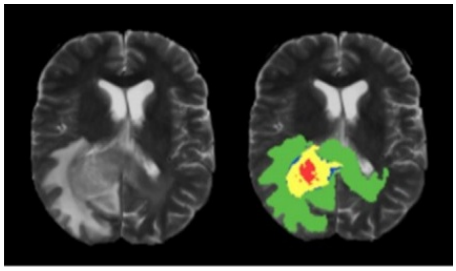


Figure 3.3: Example of Tumor Segementation and localization

3.3 Image Segmentation

3.3.1 Definition

In computer vision, image segmentation is the process of dividing an image into a set of regions to identify and locate the objects that are in the scene to understand the context. , It is used in different domains like content-based image retrieval, medical imaging, robotics, real-time self-driving cars, video surveillance, and automatic traffic control ....

Formal Definition:

The definition of image segmentation provided by Horowitz and Pavlidis [24] is as follows:

Let X be the domain of the image, and f be the function that associates a value f(x, y) with each pixel. A homogeneity predicate P is defined on the set

of subsets of X. The segmentation of X is defined as a partition of X into m subsets $\{S_1, ..., S_m\}$ such that:

- The union of all subsets S_i covers the entire domain X: $X = U_{i=1}^{m} S_i$
- The intersection between any two subsets S_i and S_j is empty for all i $\neq$ j: $S_i \cap S_j = \emptyset \quad \forall i, j : i \neq j$
- The homogeneity predicate P is true for each subset S_i: $P(S_i) = \text{ true} \quad \forall i$
- The homogeneity predicate P is false for the union of any two distinct subsets S_i and S_j: $P(S_i \cup S_j) = \text{ false} \quad \forall i, j : i \neq j$

These conditions ensure that the segmentation forms a valid partition of the image domain, with each subset representing a distinct, homogeneous region. The first condition guarantees coverage of the entire image, the second ensures non-overlapping regions, the third ensures homogeneity within each region, and the fourth prevents the merging of two regions into a homogeneous one. This definition establishes a rigorous framework for the segmentation of images, laying the foundation for meaningful and precise image analysis.

3.3.2 Type of Image Segmentation

There are several image segmentation types:

1. Naïve image segmentation: refers to simple approaches that aim to segment an image into a set of homogeneous regions based on certain low-level features such as intensity, and colors, a human interpretation is necessary. they are characterised by straightforward implementation but they can be with limited abilities in handling complex images of varying intensities and non-uniform object shape
2. Semantic segmentation: opposite to naive approaches, the pixels are grouped into a set of regions according to a specific category, otherwise it focuses on understanding the high-level semantic meaning of the segmented regions by assigning a label

3. Instance segmentation: is a refined version of semantic segmentation the difference lies in Objects of the same category are labeled differently

3.4 Image Segmentation Techniques

Over the years, the image segmentation research area has experienced numerous techniques starting with those based on rigid and classical programming algorithms until those based on machine learning. In the next sections, we exhibit the main techniques for this purpose.[25, 26, 27, 28]

3.4.1 Edge Based Approaches

Aim to locate different objects present in the image scene based on identifying their borders or edges. Edges are a significant local change in the image pixel intensities. generally associated with a discontinuity in the image intensity, there are two main approaches for the edge segmentation derivative method and the Active contour method.[29]

- Derivative-based methods: In these methods, image intensities are represented as a function $f(x, y)$, and edge segmentation is performed based on the first or second derivative of this function. In the first derivative, edges correspond to local maxima in the gradient, while in the second derivative, they correspond to zero-crossing properties (Laplacian).

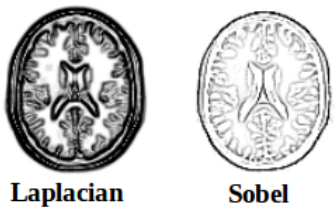


Figure 3.4: The segmentation with edge kernels

- Active contour method: also known as the snake method, this is a very well-known edge segmentation technique. it is a deformable contour model that

can capture the shape of an object by adjusting its shape to fit the object's contour.

3.4.2 Region Based Approaches

Aim to group pixels into a set of homogeneous areas based on pixel intensities and some similarity criterion, there are several techniques like Thresholding, region growing and split and merge Algorithm.[30]

- Thresholding-based segmentation: is among the oldest and the simplest used techniques for image segmentation, it yields a binary image where the pixels are partitioned depending on their intensity value according to a certain value known as a threshold. There are three types of thresholding: global threshold, local threshold, and adaptive threshold .

  The global threshold consists of using a fixed value T and binaries pixel intensities using the equation below the output of this method is a white and dark image.

$$g(x,y) = \begin{cases} 1, & \text{if } f(x,y) > T \\ 0, & \text{if } f(x,y) \leq T \end{cases} \tag{3.1}$$

  The local threshold keeps the same basic idea of the global thresholding, only applied over the partitioned sub-images obtained from the whole image each with its appropriate threshold T. Adaptive threshold or dynamic thresholding is a kind of local thresholding selected in a deeper way based on pixel

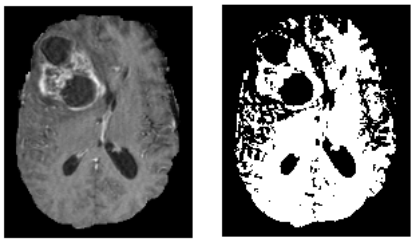


Figure 3.5: The segmentation with Thereshold Algorithm

- region growing: it is a pixel-based segmentation technique that aims to control the growing regions that begin from initial seed pixels and group neighbor pixels according to similar properties. the basic concepts of this algorithm in 1 algorithm.

- split/merge algorithm: this algorithm considers an image as a whole region then a quadtree split is applied where each region satisfies a homogeneity criteria, then merging neighboring regions if their union satisfies the homogeneity criteria. the basic steps are in the algorithm 2

Algorithm 1 Region growing Algorithm

1: Choose a seed pixel as the starting point.
2: Seed Region Formation: Make the seed pixel the initial region.
3: Neighbor Pixel Selection: Select neighboring pixels based on specific criteria.
4: Pixel Connectivity: Define pixel connectivity rules (4 or 8-connectivity).
5: Region Growing: Iteratively add neighboring pixels meeting inclusion criteria to the region.
6: Assign a label to pixels within the region.
7: Termination Condition: Specify when to stop the region growing process.
8: Repeat: If needed, apply region growing for multiple seed pixels.

Algorithm 2 Split/Merge Algorithm

1: Start with the entire image as a single region.
2: Examine the properties of the current region (e.g., variance of intensity values).
3: Split Condition: If the properties indicate heterogeneity above a certain threshold, divide the region into smaller sub-regions.
4: Examine the properties of the sub-regions.
5: If the properties of adjacent sub-regions are sufficiently similar, merge them back into a larger region.
6: Repeat Split and Merge: Continue the split and merge process recursively until no further splits or merges are needed.
7: Define a termination condition, like a minimum region size, to stop the process.
8: Assign labels to the final regions.

3.4.3 Machine Learning Based Approaches

Machine learning (ML) is a branch of Artificial intelligence (AI) that gives electronic systems the ability to learn automatically, improve, and make decisions without being explicitly programmed. The main concept is to learn hidden patterns (features)

from data called training data and build predictive models that can make decisions over unseen data. The predictive models are a kind of classification rules, decision trees, or mathematical formulas...; in the context of image segmentation, the training data is a set of images collected to tackle a specific problem. According to the way of learning from data, ML techniques are categorized into two main types supervised and unsupervised methods[28, 31, 29]

Supervised Learning based techniques In general, the training data is associated with labels, and ML techniques are trained to build models that try to achieve outputs similar to the presented labels with an estimated error and the final objective is to minimize this error during the learning process. Mathematically, we have an input X and an output Y and the ML technique is used to find a function f that verifies the equation: f(X) =Y after this function parameters are tuned it will be used to predict over new data. It exists a panoply of algorithms like k-nearest neighbors, support vector machines, Random Forests, and Artificial Neural Networks. in the context of image segmentation also called semantic segmentation, each image in the train set is associated with a label image or mask also known as ground truth that represents an image with the same size and the same information as the corresponding image in the training set where each pixel intensity is associated with the appropriate category or class label. The model compares the predicted output segmentation with the label image by computing the error and minimizing it until obtaining a precise localization for each object with the exact semantics, the labels can be presented in one of these two forms: class label or one-hot-vector-encoding.

- Class label format: each class is labeled by a definite integer number so each pixel in the ground truth will take the value of a class as the intensity value.

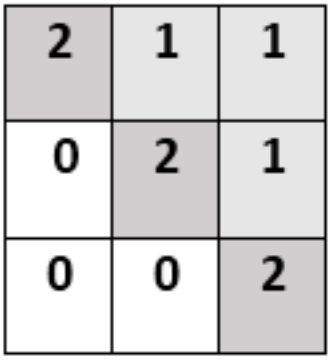


Figure 3.6: Class label format

- One-hot-vector encoding: is a representation of categorical class labels as binary vectors. It requires that the class category be integer. Then each class will be converted to a binary vector of size (height*width of the original image), where all values are zero except the index of the class will take 1. This representation is the most compatible with deep learning frameworks and in general, the class label format is converted implicitly or explicitly to one hot encoding before processing.

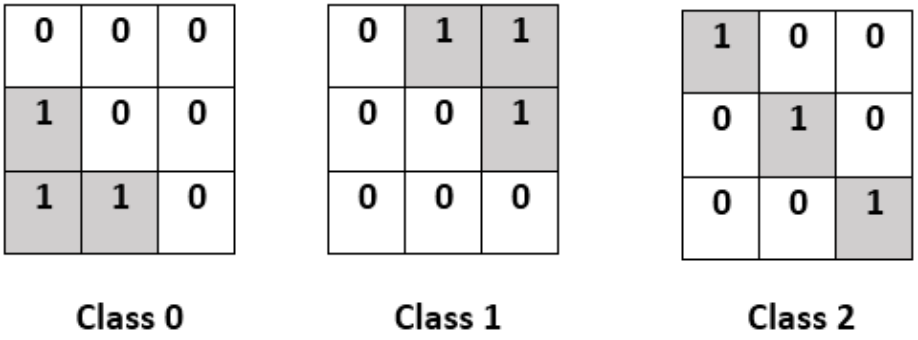


Figure 3.7: One-hot encoding format

Unsupervised learning based techniques Unlike supervised techniques, unsupervised techniques are used to learn from unlabelled datasets in another way the training data is not associated with labels, and ML techniques are used to

discover hidden patterns in the data, this kind of technique requires human expert interpretation.[29] Clustering is very recognized in the image segmentation context, the main idea is to group image pixels or voxels into a set of classes known as clusters using a similarity criterion like the Jaccard coefficient or Euclidian distance. it represents a form of an optimization problem where the objective function is to maximize the similarity of intra-classes and minimize it inter-classes. K-means, Fuzzy C-means(FCM) and Possibilistic c means (PCM) are the most renowned clustering algorithms.

3.5 Application In Brain Tumor Segmentation

In the past, traditional methods for brain tumor segmentation were indeed used, but their efficiency was limited due to the complexity and variability of brain tumor appearances in addition to the multimodal medical images. These conventional approaches often relied on handcrafted features and heuristics, which made them less robust and adaptable to the diverse nature of tumors. They faced challenges in accurately delineating tumor boundaries and distinguishing between different tumor types. With the advent of machine learning techniques, brain tumor segmentation has undergone a transformative evolution, Clustering techniques have been widely used, In [32]the authors suggested using both k-means and fuzzy c-means algorithms to segment brain tumors and compare the final result to estimate the efficiency of each algorithm in terms of segmentation accuracy and execution time. in [33] proposed to segment brain MRI using a fuzzy-possibilistic C-means algorithm in addition to the k-means method to perform the skull stripping to extract the brain region only as a pre-processing step Without forgetting supervised techniques like support vector machine and random forest, k nearest neighbor[34, 35, 36, 37] Random Forest (RF) [38], These methods use feature extraction (edges, texture, colors, etc.), a te- tedious process, as one of their core phases because their accuracy is directly related to the quality of the best discriminative features, this step must be preceded by a time-consuming feature selection. However, both feature selection and

extraction are implicit and fully automatic in DNN, making them more adaptive to learning from various images [39].

Recently, deep learning-based approaches have prevailed in the general medical image segmentation state of the art [40, 41, 42, 43, 44]. And the brain tumor segmentation domain [45], where convolution neural network (CNN) is the fundamental architecture used with the encoder-decoder variants like fully convolutional networks (FCN) and U-net. They return results that outstripped previously mentioned techniques. Deep learning has indeed brought about a revolutionary change in brain tumor segmentation and has demonstrated exceptional capabilities in capturing intricate patterns and features in general medical images which has significantly improved the accuracy of brain tumor segmentation, enabling better diagnosis and treatment planning, this method is better discussed in the next chapter.

3.6 Conclusion

In this chapter, we have discussed the most important techniques for image analysis. Our focus was on the progression of image segmentation methodologies, from traditional algorithms to those based on machine learning. We have thoroughly explored each category of technique by exhibiting its main principles and some renowned algorithms. Additionally, we conducted a literature review that specifically focused on their applications in brain tumor segmentation.

When You Feel Like Quitting, Remember Why You Started

4

Deep Learning for Image Segmentation: Brain Glioma Segmentation

Contents

4.1 Introduction

Automatic image segmentation is a crucial aspect of image analysis that has seen significant advancements and challenges, especially with the emergence of Deep learning. This chapter offers the history of deep neural networks since the invention of artificial neural networks with its detailed steps for feedforward and backward

processes, in addition to direct attention to the most important used techniques and architectures in the context of image segmentation using deep learning.

4.2 Artificial Neural Network (ANN)

Artificial Neural network ANN is a machine learning technique that mimics the brain's cognitive system , it has a long history of development since 1943. like most supervised machine learning techniques ANN can resolve regression or classification problems in any field. as the brain is a massive network of connected neurons, neural networks are first designed as neuron 4.1 with a graph of multiple nodes and connections in layered structure: input layer, hidden layer, and output layer depending on the type of neural network.[46].

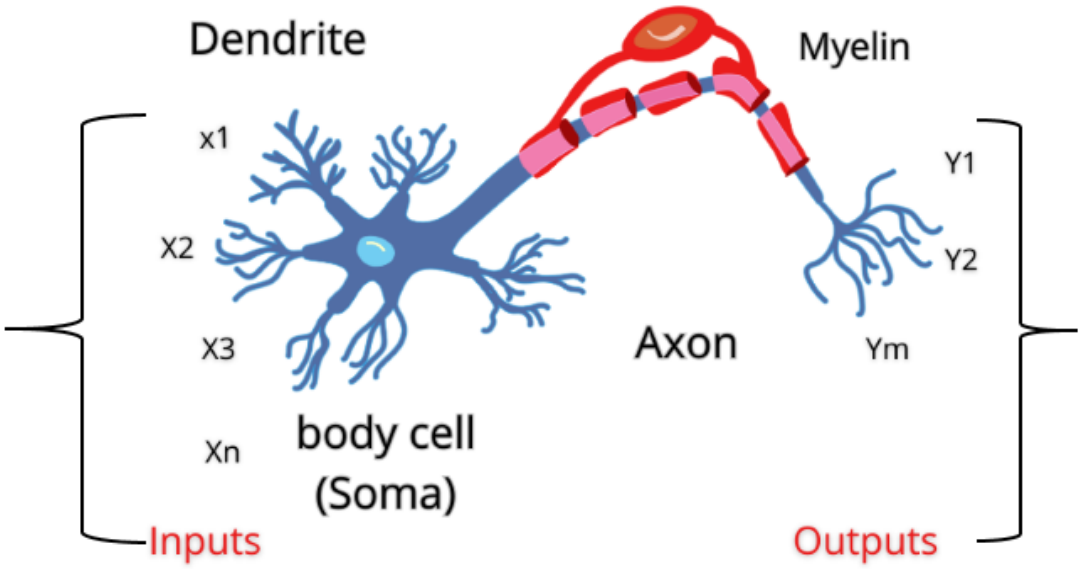


Figure 4.1: From Human Neuron to Neural Network

According to the number of layers, the neural network architectures are classified as follow:

The neural network model is a kind of mathematical function that contains variables, weights, and bias, during the learning process the weights are adjusted

Single Layer Neural Network (Perceptron)	One Input Layer- one Output Layer	
Multilayer Neural Network	Shallow Neural Network	One Input Layer- One Hidden Layer- One Output Layer
	Deep Neural Network	One Input Layer- more Hidden Layers- One Output Layer

according to the variable's input values to reach an output similar to the desired outputs and get the ability to predict over unseen data within a tolerable error this refers to the supervised learning process. The neural network supervised learning is as follow:

1. Random Weights Initialization.
2. The input goes through the network to obtain an output, then calculates the error between the correct and the network output.
3. Adjust the weights to reduce the error using a learning algorithm generally called the backpropagation algorithm(Gradient descent).
4. Repeat steps 2-3 for all training data

it is mainly composed of two main steps feedforward pass(step 1-2) and backward pass(step 3).

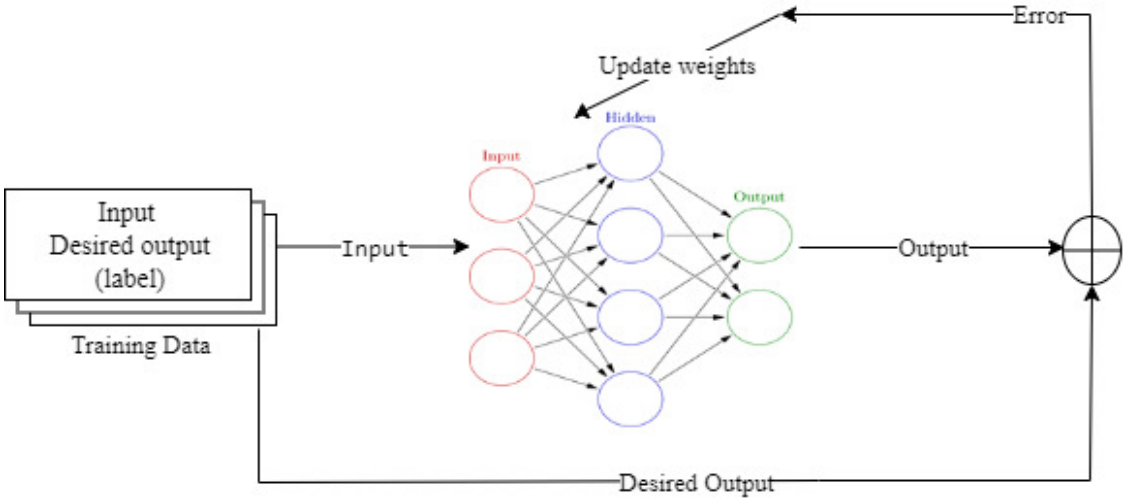


Figure 4.2: The Supervised learning process

4.2.1 The Learning Algorithm

A learning algorithm [1] is a mathematical operations applied over the Network, it is used to update the weights and bias levels of a network to improve the Artificial Neural Network performance in terms of accuracy and precision. this iterative process helps a neural network to learn from the existing conditions and improve. since the development of neural networks, the learning algorithm has known different development stages due to the variety of networks and the learning requirements but the basic idea that was founded on is the same, the gradient descent algorithm or The Delta rule.

4.2.1.1 Gradient Descent

is an optimization algorithm used to find a local minimum/maximum of a differentiable function. in machine learning algorithms, gradient descent is commonly used to train and update the model's parameters to minimize the error function also known as the cost function. Mathematically, the gradient is the rate of change of a function and it is related to the derivative of the function. it represents the amount by which the output function changes for a small change in the input variable. for example in the case of the linear function F(x)=ax, the gradient is a, and this is called the slope. the term gradient is usually used in the case of

multivariate functions which are functions with multiple inputs. the main concept of the gradient descent algorithm is to start from an arbitrary point x and move to the next point by gradient called Δx with a step size called η using the following equation x=x+$\eta\Delta$x and repeat the process until converging towards the stationary point called local minimum if the movement is with a negative gradient or local maximum inversely and this is known as the direction.

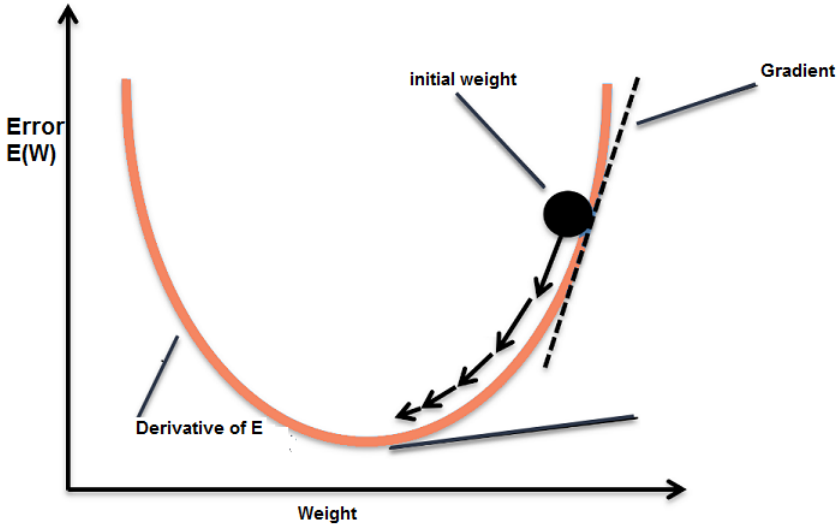


Figure 4.3: Gradient Descent

4.2.2 Single Layer Neural Network: Perceptron

Perceptron is the basic architecture in the simple neural network, it is inspired by biological neuron architecture. in 1943, the neuroscientist Warren McCulloch and the logician Walter Pitts were the first to propose the concept of the neuron functioning in a mathematical computational model as the figures show where the inputs and the outputs formed as binary

In 1957, Frank Rosenblatt introduced the first single-layer perceptron as a supervised linear classifier for binary classes. based on the Muc proposition which supports all types of inputs not only binary ones. it aims to find the appropriate decision linear or hyperplane boundary. the figure below exhibits an example of a single-layer perceptron associated with its parameters

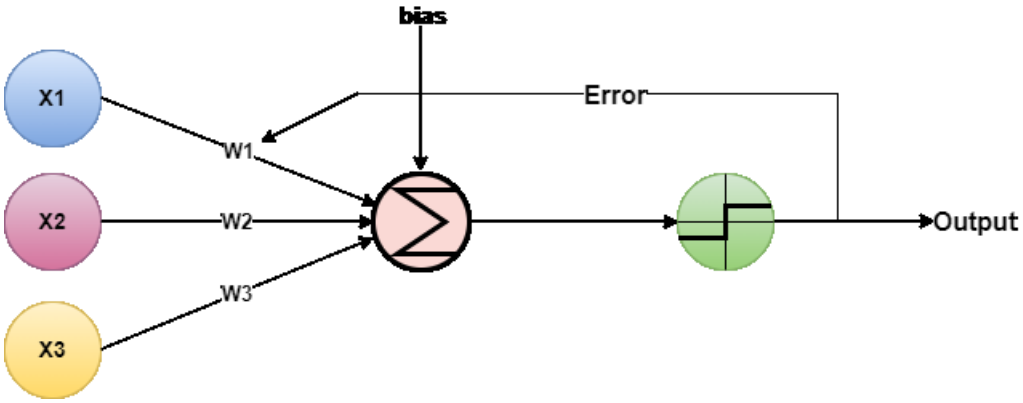


Figure 4.4: Single Perceptron Network

4.2.2.1 Perceptron FeedForward Pass

The figure exposes the perceptron functioning where the nodes (X1..X3) represent the input signals and the(w1..w3) the weights for the corresponding input signals, the bias is used to avoid the zero input when (XW=0) with an output greater than 0.[1] The feedforward process starts as follows: the inputs are first multiplied by the weights before reaching the summing node where they are aggregated in addition to the bias to generate the weighted sum S.

$$S = (w_1 \times x_1) + (w_2 \times x_2) + (w_2 \times x_2) + b \tag{4.1}$$

in general way the weighted sum can be formulated as follows:

$$S = wx + b \tag{4.2}$$

the weighted sum enters into the activation function node to generate the output

$$y = \varphi(v) \tag{4.3}$$

$$y = \varphi(v) = \varphi(wx + b) \tag{4.4}$$

the output of this function is denoted as a predicted output y which must be equal to the desired outcome or the correct label introduced with the input data, the learning process here is supervised which means the perceptron model learns how to reach the correct output within a tolerable error by adjusting the weights through the use of a specific learning algorithm.

4.2.2.2 The single Layer Network Learning Rule: Backward Pass

in the single-layer Network, the learning rule has undergone several changes due to the activation function and error function. for example in the perception, the activation functions are a kind of a step function or threshold and error computed based on a simple equation.

$$e_{ij} = d_{ij} - y_{ij} \tag{4.5}$$

where e is the error , d is the desired output and y the predicted output.

the learning rule applied here is called the perceptron learning rule and is defined based on the delta rule without using derivatives because the step function is not differentiable, and as it is used to update the network weights the learning rule is defined as follows:

$$w_{ij} = w_{ij} + \Delta w_{ij} \tag{4.6}$$

where w_{ij} is the weight between the output node i and input node j and Δw_{ij} considred as the gradient of w_i computed as follow

$$\Delta w_{ij} = \alpha e_i x_j \tag{4.7}$$

where α is the learning rate ($0<\alpha<1$), x_j the output from the input node j,and e_i is the error of the output node i

after that, the Adaline (Adaptive linear perceptron) perceptron was developed by Bernard Widrow and Ted Hoff in 1960 which is similar to the previous but with a small change in the activation function which is the linear activation (identity function), the use of such a continuous function provides the advantage of being differentiable which allows further the use of cost functions like MSE for Mean Squared Error, that can be minimized to update weight using the gradient descent algorithm, the cost function is defined as follow:

$$E_i = \frac{1}{2}\left(d_i - y_i\right)^2 \tag{4.8}$$

example:

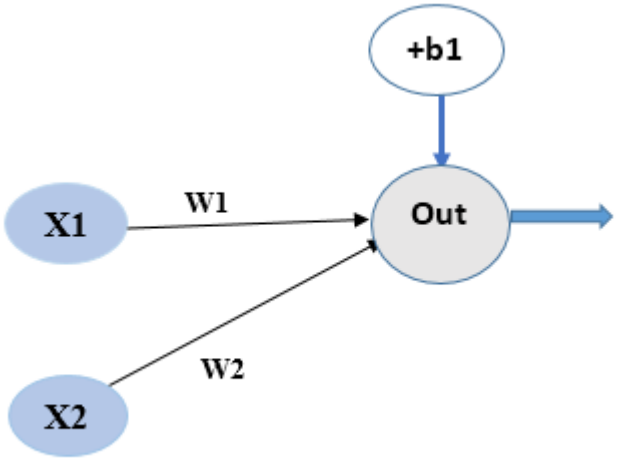


Figure 4.5: Example of Adaline Network with out set to linear activation

the figure 4.5 shows an example of Adaline perceptron where X1 and X2 are considered as inputs,w1 and w2 as weights, and out as output node where a linear activation is applied to generate the prediction Y figure resumes the feedforward process which is similar to the perceptron one.

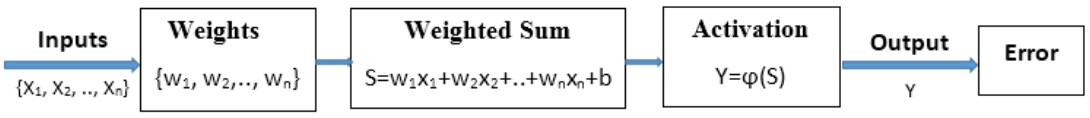


Figure 4.6: The Adaline Feedforward Pass

after that, the error between the desired and the predicted output is calculated using the below equation and the backward pass start to minimize this error by updating the weights using the gradient descent algorithm.

$$E(w) = \frac{1}{2}\sum_{j=1}^{n}\left(d_j - \left(b + \sum_i w_i x_i\right)_j\right)^2 \tag{4.9}$$

in this case, the gradient Δw_{ij} is computed using the partial derivatives because the objective function is multivariate so a vector of gradients is generated as follows:

$$\nabla E(\vec{w}) \equiv \left[\frac{\partial E}{\partial w_0}, \frac{\partial E}{\partial w_1}, \cdots, \frac{\partial E}{\partial w_n}\right] \tag{4.10}$$

where $\frac{\partial E}{\partial w_i}$ is the partial derivative of the function E with respect to w_i computed using the chaine rule as follow:

$$\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial Y} \times \frac{\partial \varphi}{\partial S} \times \frac{\partial S}{\partial w_i} \tag{4.11}$$

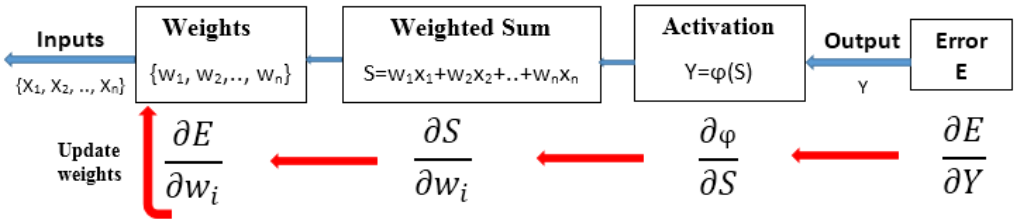


Figure 4.7: The Adaline backward Pass

Figure4.7 resume the backward learning process and after calculating the gradients, each weights is updated using the previous delta rule defined as

$$w_{ij} = w_{ij} + \Delta w_{ij} \tag{4.12}$$

where $\Delta w_i = -\eta \frac{\partial E}{\partial w_i}$ with η as the learning rate.

4.2.3 Multilayer Neural Network

in this case, the network must have at least one hidden layer, and the more hidden layers the more going deeper which is known as a deep neural network. this kind of network was proposed as an extension of a single-layer network to solve more complex problems not only linearly separable problem .[1]

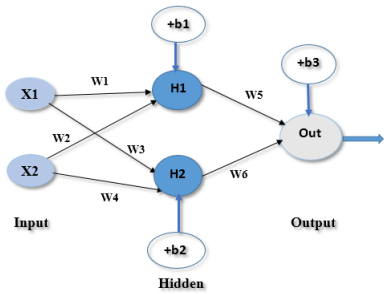


Figure 4.8: Multilayer Neural Network with one Hidden Layer

4.2.3.1 Back-propagation for multilayer Neural Network

it keeps the same key idea of the Adaline learning rule the only difference lies in the multiple hidden layers which make the backward pass more complicated and need more complex chain rule calculus to calculate the gradients, the feedforward and backward passes of the multilayer network in Figure4.9, and from here the backpropagation algorithm has been structured.

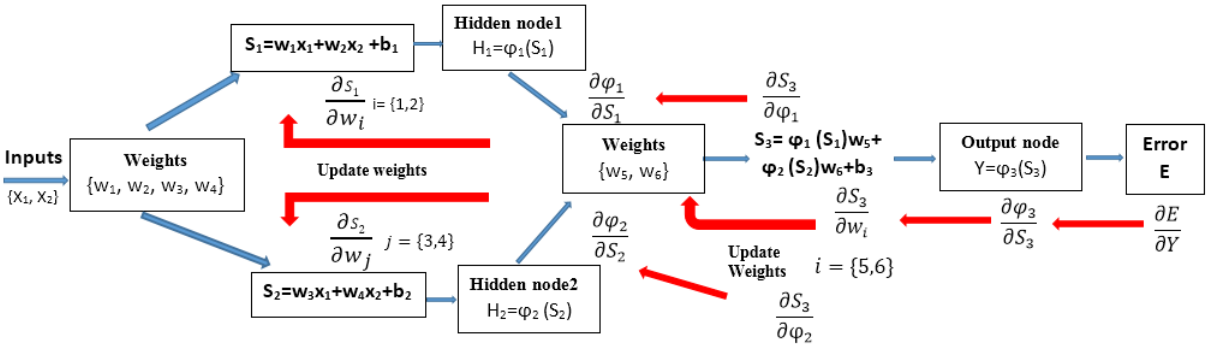


Figure 4.9: Multilayer Backpropagation pass

The Algorithm below recapitulates the main step in the neural network learning process.

Algorithm 3 Backpropagation Algorithm : Feedforwad and backward passes

let X_1 to X_n be inputs
let H_1 to H_k be the hidden nodes.
let w_1 to w_m be the weights between nodes
for each node H_j **do**
    Compute the weighted Sum S_j (Eq.:4.1)
    Compute the activationφ of each S_j (Eq.: 4.3)
    repeat 3-4 until reach the output node and generating the output Y.
end for
Compute the error using E function (Eq.:4.8)
for each node w_i **do**
    backpropagate the error and compute the Gradient using chaine rule.
    Update each weights.
end for

4.2.4 Activation Functions

An activation function is a crucial component of artificial neural networks and deep learning models. It introduces non-linearities into the network, allowing it to learn from and make predictions on complex data.

The primary purpose of an activation function is to determine the output of a neural network node or neuron. Without activation functions (or with only linear activation functions), a neural network, no matter how deep, would behave like a single-layer perceptron, unable to capture complex patterns and relationships in the data.[47]

Here are some common activation functions used in neural networks:

Sigmoid function (Logistic):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{4.13}$$

Output range: (0, 1) Used in the output layer of binary classification models.

Hyperbolic Tangent function (tanh):

$$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \tag{4.14}$$

Output range: (-1, 1) Similar to the sigmoid but with a wider output range.

Rectified Linear Unit (ReLU):

$$\text{ReLU}(x) = \max(0, x) \tag{4.15}$$

Widely used in hidden layers due to its simplicity and effectiveness in training deep neural networks.

Leaky ReLU:

$$LeakyRelu(x) = \begin{cases} \alpha x, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases} \tag{4.16}$$

where

α is a small positive constant, typically 0.01. It addresses the "dying ReLU" problem where neurons can become inactive during training.

4.2.5 Loss Functions

A loss function, also known as a cost function or objective function, measures the difference between the predicted values of a model and the actual values (ground truth) in the training data. The goal during the training of a machine learning model is to minimize this loss function. Different types of problems (classification, regression, etc.) and tasks may require different loss functions. Here are some commonly used loss functions:[48]

Mean Absolute Error :

$$\text{MAE} = \frac{1}{n}{}_{i=1}^{n} |y_i - \hat{y}_i| \tag{4.17}$$

Similar to MSE, but it computes the average absolute difference instead. It is less sensitive to outliers.

Binary Cross-Entropy (Log Loss):

$$\text{Binary Cross-Entropy } = -\frac{1}{n} \quad {}_{i=1}^{n} \left[y_i \log \left(\hat{y}_i \right) + (1 - y_i) \log \left(1 - \hat{y}_i \right) \right] \tag{4.18}$$

Used for binary classification problems. It measures the difference between the true distribution and the predicted distribution for binary outcomes.

Categorical Cross-Entropy:

$$\text{Categorical Cross-Entropy } = -\frac{1}{n} \quad {}_{i=1}^{n} \quad {}_{j=1}^{m} y_{ij} \log \left(\hat{y}_{ij} \right) \tag{4.19}$$

Used for multi-class classification problems. It generalizes binary cross-entropy to more than two classes.

Hinge Loss (SVM Loss):

$$\text{Hinge Loss } = \frac{1}{n} \quad {}_{i=1}^{n} \max \left(0, 1 - y_i \cdot \hat{y}_i \right) \tag{4.20}$$

Commonly used in support vector machines (SVMs) and for binary classification. It penalizes misclassifications.

It's important to choose a loss function that aligns with the objectives of the specific machine learning task.

4.3 Deep Learning for image segmentation

4.3.1 Definition

Deep learning is an artificial intelligence discipline and a machine learning subset and a kind of neural network biologically inspired by the human brain cognitive system There are several (DL) Architectures each used for a specific kind of problem. This technique has gained an important interest in the medical sector and especially for segmentation tasks. [49, 50, 51]

4.3.2 Convolutional Neural Network (CNN)

convolutional neural network is a kind of deep neural network that mimics the brain's visual cortex function in receiving, integrating, processing, and interpreting visual information.[52] CNN offers electronic systems like computers and smartphones a slight human ability to understand the visual context (images and videos). practically, CNN architectures are used for image classification, object recognition, and image segmentation tasks.

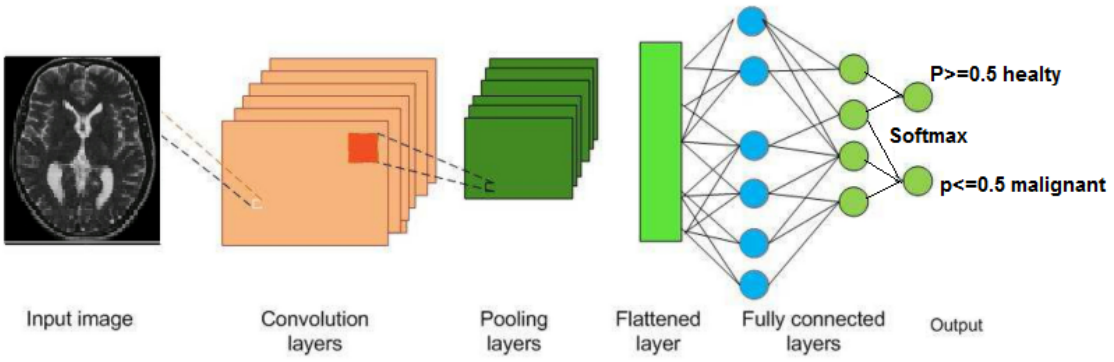


Figure 4.10: The Convolutional Neural Network architecture

regardless of the used method, using the original image directly leads to bad results that's why a processing step is required called feature extraction, for this purpose various techniques have been developed and they are completely independent of machine learning techniques. there are numerous feature extractors each designed to extract a specific type of feature like Robert, Sobel, and Laplacian for edge extraction or Gabor for texture the parameters of those extractors are manually fixed

which requires high cost and time and finally yield low performances, However, CNN turned this manual design into an automated process which is its primary advantage where the feature parameters are initialized randomly and tuned along the training process according to the data. convolutional neural networks are mainly composed of two paths: feature extraction path and classification path. the feature extraction path is composed of stacked convolution and pooling layers to extract numerous features and finished with a fully connected layer to perform the classification. in this section, we describe the most important operations behind CNN

1. Convolution Layer: refers to the convolution operation, it is the mathematical operation that calculates the dot product between two matrices which are the image matrix and the convolution filter. it generates a new abstracted image called a feature map. the filter shape may be 1D,2D, or 3D and the values of the filter matrix play the role of the weights in a simple neural network they are fixed through the training process and that's why the layer is called the convolution layer.

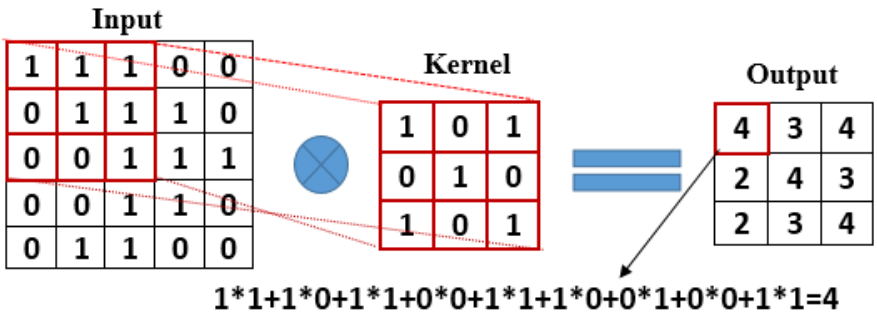


Figure 4.11: The Convolution operation

2. Pooling Layer: this is a mathematical operation that aims to reduce the size of the feature maps and keeps only representative pixel information to minimize computation time and memory consumption. it binds neighboring pixels of a selected square area using a specific pool size in one-pixel value.
   the pixels are bound to either max-pooling operation or Average pooling. the

max-pooling is to select and keep in the feature map the pixel of the maximum value however the average pooling is to compute the average value of the pixels in the selected area example:

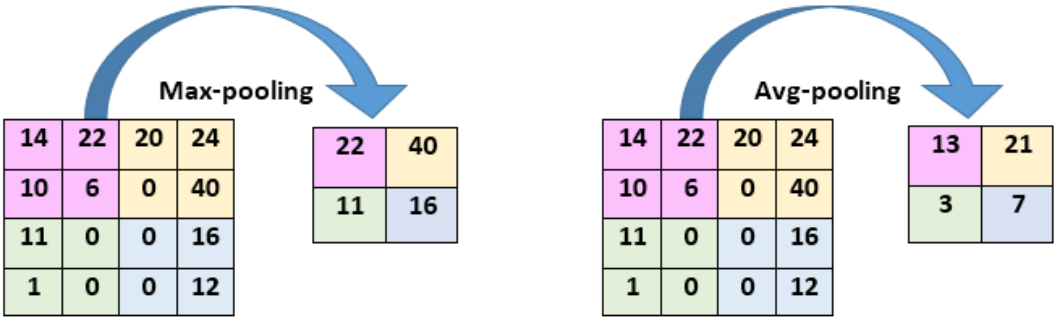


Figure 4.12: The Pooling operation

3. Fully Connected layer: also known as dense layer, it is a type of layer that connects all neurons from the previous layer to every neuron in the current layer. its purpose is to perform the classification task on the extracted features by the previous layers in the network.

4.3.3 Fully convolutional Neural Network FCN

is similar to convolution neural network architecture, but instead of the fully connected layer in the classification path. FCN[53, 2] uses an up-sampling path consisting of layers with opposite operations in the down-sampling path. These operations may be the reverse operations of the convolution(transposed convolution) or the pooling operations (unpooling with max-unpooling, nearest neighbour or bed of nails). In addition to some skip connection which is a merging and concatenating information between layers of the two paths in order to recover the spatial information of the objects "where are these objects located".At the end, an output image with a set of probabilities is generated to identify the classes. as we mentioned above the most important layers to perform in this kind of architecture are the upsampling layers which are used to increase the spatial resolution, their is as follow:

- Up-sampling with Transposed Convolution:(fractionally-strided convolution) is a technique designed to recover or enhance spatial information that might have been lost during downsampling operations, such as max pooling or strided convolution. It involves convolving the input feature map with learnable kernel weights, using a stride greater than 1 and zero-padding to increase the spatial dimensions. During training, learnable weights are updated through backpropagation. By applying this transposed convolution operation, The result is an output feature map with higher resolution, enabling the model to capture finer details and patterns in the data.

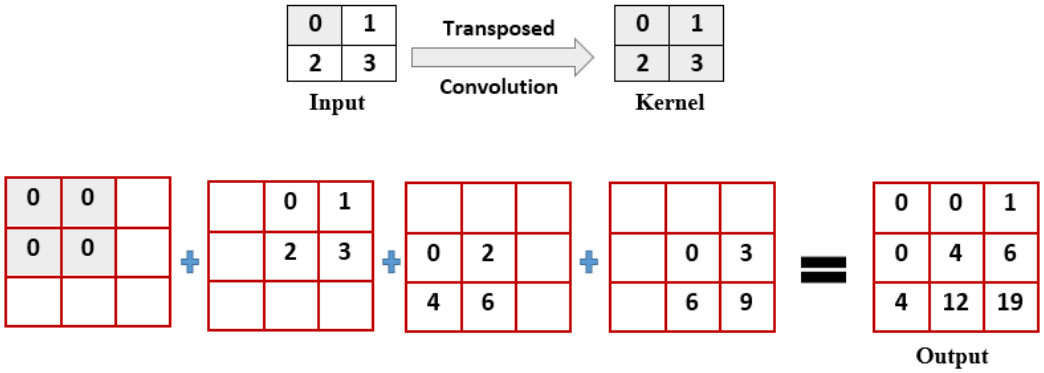


Figure 4.13: Example of Transposed Convolution

- Up-sampling with max-unpooling: during the max-pooling operation in the feature extraction step the indexes of the max values are saved to be used later in the upsamling where the input pixels from the feature map are mapped to the saved index and filling zeros the remain indexes. example

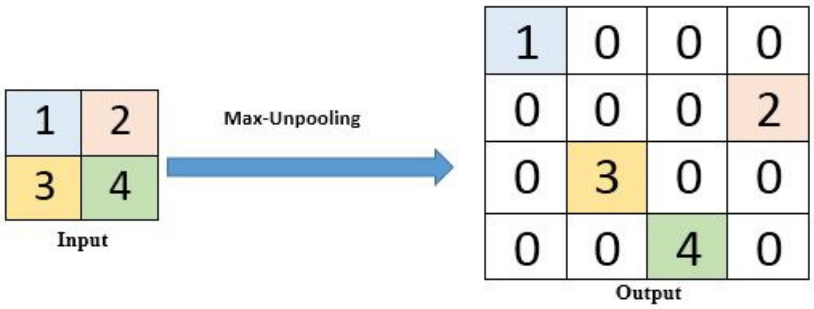


Figure 4.14: Up-sampling with max-unpooling

- Up-sampling with nearest neighbor: is a simple upsampling approach where the input feature map pixel are copied to the corresponding output and the k-nearest neighbors are filled with the same pixel value example

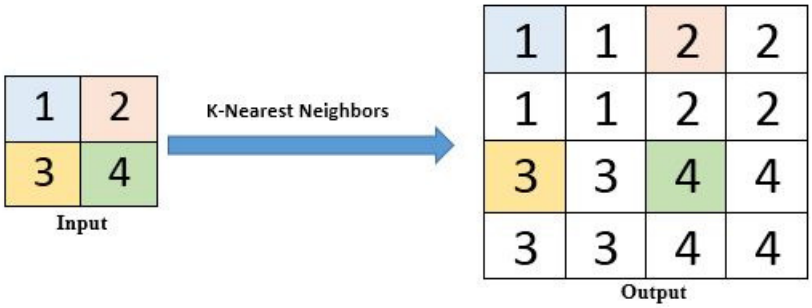


Figure 4.15: Up-sampling with nearest neighbor

- Up-sampling with bed of nails: same as the nearest neighbor approach but the the k-nearest neighbors are filled with zeros example:

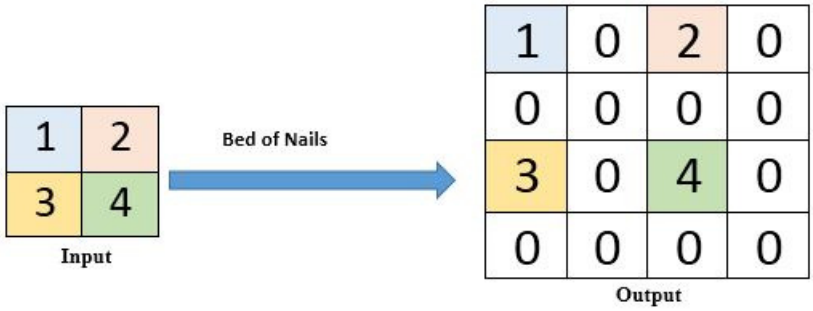


Figure 4.16: Up-sampling with bed of nails

4.3.4 Popular Architecture for Image Segmentation

4.3.4.1 SegNet Architecture

SegNet is a deep learning architecture specifically designed for semantic segmentation tasks in computer vision first introduced by Vigy et al in for real world segmentation images 2017[54] and later used for medical image segmentation [55]. it is a symmetrical encoder-decoder architecture where the encoder path represents the vgg16 convolution parts with its 13 layers. the decoder path consists of Upsampling layers that uses the max-unpooling which is one of the distinctive features of SegNet, it use max-pooling indices from the encoder stage ,Rather than storing the actual pixel values during max-pooling, this information is later used in the decoder to perform precise upsampling.

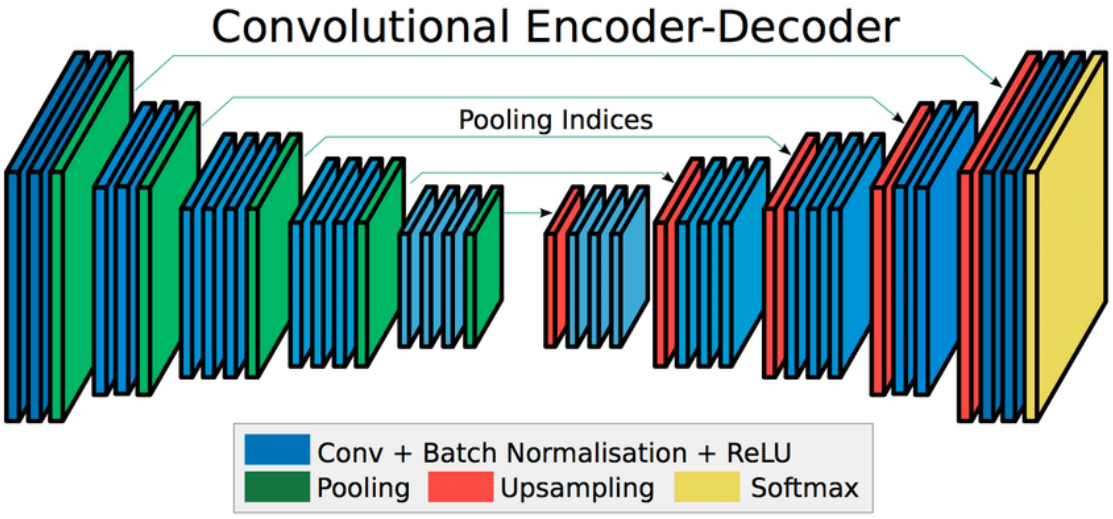


Figure 4.17: The SegNet Architecture

4.3.4.2 DeepLab Architecture

DeepLab is a series of popular semantic segmentation models designed by Google in open-sourced and since the appearance of the first version DeepLab V1 multiple improvement have been made DeepLab V2,V3 and V3+.

The first version named DeepLab v1[56] was built upon a mixture of convolution neural network CNN and conditional Random Field (CRF). the CNN part is a kind of traditional feature-extracting path composed of a series of convolution and

pooling operations or pre-trained model but with a specific change in last layers where the Atrous convolution layers are integrated.

Atrous convolution (dilated convolution), the name atrous comes from the French word “à trous “ which means with holes. These holes are gaps introduced between the elements of the convolution kernel (filter) to increase the receptive field or the field of view without augmenting the number of trainable parameters in the kernel. This gap is defined by a dilation rate greater or equal two to when it is set to one, it is referred to as the ordinary convolution to allow the capturing of contextual information from a broader region around each pixel and gathering multi-scale features using different dilation rates..

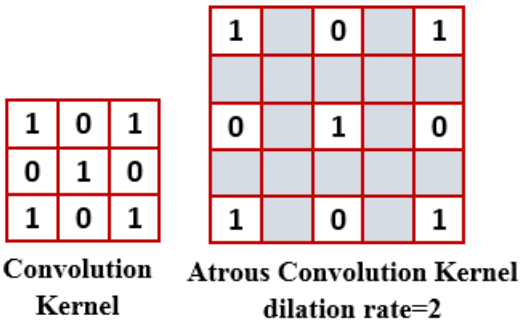


Figure 4.18: The Atrous Convolution Kernel

in DeepLabv1, this kind of convolution is used to get a higher sampling density to calculate the feature map and then use bilinear interpolation to restore the feature to the original image size and fully connected Conditional Random Field CRFs used to enhance the model’s ability to capture details, ensuring the accurate split of the boundary positions. One of the significant enhancements in DeepLab v2 [56] is the introduction of Atrous Spatial Pyramid Pooling (ASPP) module which involves parallel atrous (dilated) convolutions with different dilation rates to capture multi-scale contextual information and helps the model to have a more comprehensive perception of the input image at different scales.

The third version DeepLab v3[57] similar to the second version but with CRF removed and replaced by(1*1) convolution in addition to some updates in the ASPP

like adding batchnormalization layers and global average pooling with pre-trained VGG19 or Resnet as backnone architecture.

The last version DeepLabV3+[58] was changed to FCN architecture with encoder(DeepLabV3) and decoder netowrk with the usage of separble atrous convolutions .

4.3.4.3 U-Net Architecture

U-net was the first architecture introduced to deal with semantic segmentation of biomedical images and achieved great success in medical image segmentation, first proposed [59] in 2015 .It is based on a fully convolutional neural network(FCN). the architecture was designed as a symmetric U-shape composed of two paths a contracting path and an expansive path based on the encoder-decoder aspect with a bottleneck between them as an additional touch.

1. Contracting Path: often referred to as the encoder path, consists of a series of convolutional and pooling layers. These layers are responsible for extracting and capturing features at various scales and reducing the spatial resolution of the input image.

2. Expansive path: known adecoder path,consists of a series of transposed convolutions (also known as "deconvolutions" or "up-sampling") and skip connections which are a critical component of UNet and help improve the segmentation accuracy. They connect feature maps from the contracting path to the corresponding feature maps in the expansive path. This allows the network to use both low-level and high-level features in the segmentation process, which is particularly useful for capturing fine details. This path is responsible for gradually increasing the spatial resolution of the feature map while combining information from the contracting path through skip connections.

3. Bottleneck: The architecture includes a bottleneck layer that connects the encoder and decoder parts. It acts as a bridge for information to pass from the encoding to the decoding part of the network.

4. Final Convolutional Layer: The expansive path eventually leads to a final convolutional layer that produces the segmentation mask. This layer typically has a single channel output for each class that the network is designed to segment.

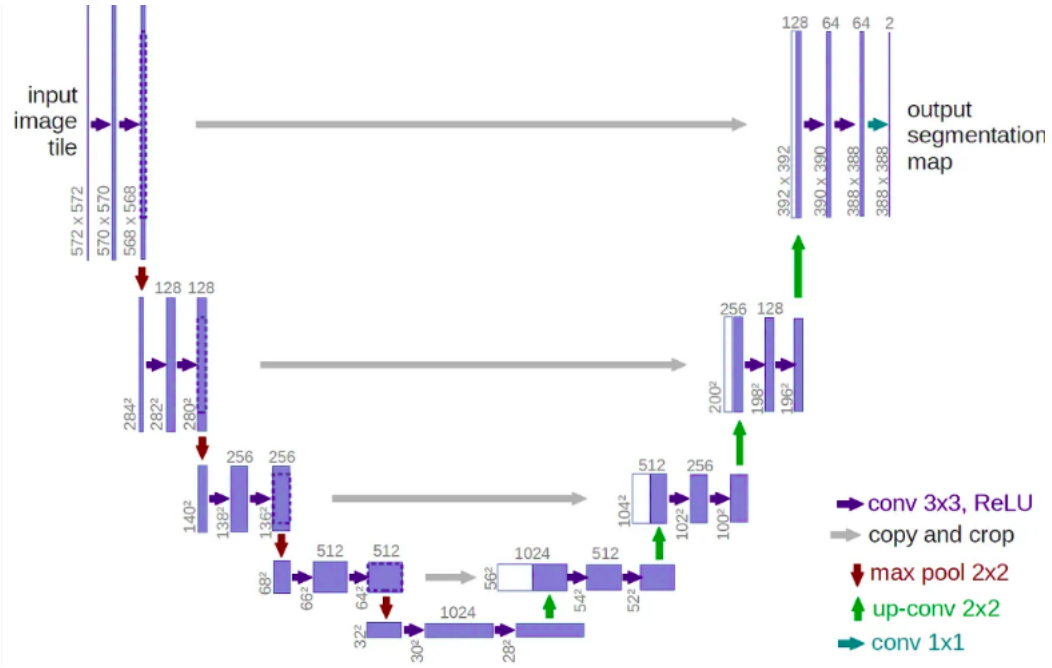


Figure 4.19: The U-Net Architecture

4.4 Conclusion

This chapter summarized the historical trajectory of deep neural networks with special attention to the convolutional neural network and the fully convolutional neural network and their use in the image segmentation task. It provided details about the architecture design with layers operation, learning process, and the most well-known architecture used to perform image segmentation.

Part II

Contribution

Great Things Never Come From Comfort Zones

5

Parallel Pathway U-NET Architecture for Brain Tumor Segmentation

Contents

5.1 introduction

Automatic medical image segmentation is one of the main tasks for many organs and pathology structures delineation.It is also a crucial technique in the posterior clinical examination of brain tumors, like applying radiotherapy or tumor restrictions. Various image segmentation techniques have been proposed and applied to different

image types. Recently, it has been shown that the deep learning approach accurately segments images, and its implementation is usually straightforward. This chapter introduces an innovative deep-learning methodology tailored to segment brain gliomas from multimodal MRI data. Our approach introduces a novel strategy for effectively handling multiple input modalities, presenting advancements in training strategies, evaluation metrics, and experimental results. Through a comprehensive comparative analysis, we discuss the findings, highlighting both the advantages and limitations of our proposed architectures. Furthermore, we delve into the implications of our research, suggesting potential future directions for refining and expanding the application of our novel approach in the dynamic field of medical image analysis.

5.2 Related works

In the last few years, brain tumor segmentation has received great interest [60], and various challenges have been documented under the BRATS title acronym for Brain Tumor Segmentation. The majority of contributions involved the utilization of deep learning techniques and achieved the top ranks. For example, in BRATS 2017 release,Kamnitsas et al[61] proposed an approach that explores multiple deep network architectures trained separately and Ensemble prediction maps to construct the final segmentation results; this achieved the first place on the challenge. Cascaded neural networks with anisotropic and dilated convolutions in the second place introduced by Wang et al [62], where the cascaded design is used to segment the whole tumor bounding box first, which was used to feed the next model to segment the tumor core, and finally, Enhancing the tumor is segmented with the same process.

In BRATS 2018,Myronenko[63] was the first winner with its proposed 3D encoder-decoder architecture, accompanied by a variational auto-encoder part starting from the encoder endpoint. It aims to reconstruct the original input image to regularize the shared encoder. The second winner was Isensee et al [64], who proposed using 3D baseline U-Net with minor modification and focused on region-based prediction and a co-training process that uses additional data. McKinleyet al[65] was the

third winner with its preposition, which is a shallow network similar to the famous U-Net architecture that uses a densely connected block of dilated convolutions and introduces a new loss function based on binary cross-entropy loss function to calculate the label's uncertainty. Wang et al [66] propounded a work to segment brain glioma structures. It aims to use test time augmentation by augmenting images with 3D rotation, flipping, scaling, and adding noise to training and test images using different underpinning 3D network structures to perform multi-class segmentation and cascaded networks. A.Albiol et al [67]preferred to get inspired by the well-known two-dimensional (2D) deep learning models like VGG, inception2, inception3, and dense-like models and extend them to 3D versions. All these models were used to segment brain gliomas separately in addition to a final ensemble result. RuiHua et al[68] developed a cascaded 3D V-Nets framework to handle the problem. The segmentation was performed in a cascaded way where the whole tumor was segmented first by three ensembled V-Nets. The detected tumor is segmented to the other tumor parts necrosis, edema, and enhancing tumor using two ensembled V-Nets. Kermi et al[69] proposed four channels of 2D U-Net architecture to segment gliomas and used residual blocks instead of plain blocks in the original U-Net. To address the class imbalance problem, Weighted Cross-Entropy (WCE) and generalized Dice loss were used. Marcinkiewicz et al [70] suggested segmenting brain gliomas in 2D two cascaded stages based on a convolutional neural network inspired by U-Net, the first stage applied to detect regions of interest and the second stage used to perform multi-class classification. In the next BRATS challenges 2019, the top-ranked approaches [71, 72, 73] were also based on deep learning architectures. Many additional proposed methods exist in [74, 75]of different proposed 2D and 3D-based networks. Outside the BRATS challenge, a multitude of research studies has explored deep learning-based approaches [76, 77, 78, 79], which have exhibited favourable achievements too. These works often employ a 2D and 3D U-Net architecture as a foundation and incorporate customized enhancements.

The significant limitations of these existing works are the increased computational complexity, high memory requirements, and long-running time due to using the

3D format and complex operations in deep learning architectures like dilated convolutions and residual blocks, which entail processing many parameters.

5.3 Methodology

In this work, we propose to segment the brain gliomas tumour from multi-modal MRI images (T1CE, T1, T2, Flair) using a deep learning architecture called PP-UNET. Our architecture is inspired from the well-known deep learning architecture used for biomedical image segmentation U-NET[59], where we proposed a customized U-NET used in multi-pathway mode to deal with MRI modalities. Each image modality goes through a processing U-net pathway and before generating the final prediction map the final pathways results are merged or concatenated. The customized U-NET is designed symmetrically as follows: the contraction path: also called the encoder path, consists of three levels each level is made of two consecutive (3*3)convolution layers of 16 learnable filters doubled for each level, each followed by an Activation Layer with LeakyRELU function parameters, then the result (R) of these layers go through a (2*2) pooling layer followed by batch normalization layer and Dropout layer to prevent overfitting. The expanding path: also called the decoder path, consists of three levels too each level is made of a transposed convolution layer, then a concatenation layer known also as skip connection between the two symmetric path where the result (R) is concatenated with the transposed convolution result to get more information, then a dropout layer and two consecutive (3*3)convolution layers each followed by an Activation Layer with LeakyRELU function parameters are applied. The paths are related between them by two consecutive (3*3)convolution layers each followed by an Activation Layer with LeakyRELU function parameters called a bridge. The final outputs of the expanding paths from each modality are concatenated and then the resulting map goes through a (1*1) convolution layer with a sigmoid activation function to generate the final prediction map.

Figure 5.1: The parallel pathways architecture PP-UNET

5.4 Experimentations

5.4.1 Data :

In our work we propose to use the brain tumor segmentation (BRATS) dataset 2018 version [80, 81, 12], it s composed of two datasets training data and validation data, the training data contains a set of MRI images of 285 patients diagnosed with high-grade gliomas (Glioblastoma) and low-grade gliomas. Each patient file encompasses four co-registered MRI modalities native pre-contrast (T1), post-contrast T1-weighted (T1c), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR). In addition to the manual segmentation file where all pixels were segmented into four classes with four rates (0, 1, 2, 4) summarized in Table 5.1. The data was acquired with various scanners at 19 institutions and with different clinical protocols, All Brats multimodal scans are available as Nifty files (.nii.gz) and they are pre-processed: co-registered to the same anatomical template, interpolated to the same resolution ($1mm^3$) and skull-stripped. Tables 5.1 and 5.2 below summarize the training dataset information:

Table 5.1: The Brats 2018 training data information

MICCAI_BraTS_2018_Data_Training		
Gliomas grade	High Grade Gliomas (HGG)	Low Grade Gliomas (LGG)
# of patients	210	75
Modalities	T1, T1CE, T2, FLAIR	
Dimension	(240,240,155)	
Format	Nifty (Nii.gz)	
Ground truth	Available in seg.nii.gz file	
Labels	0 =background and healthy tissue	
	1= Necrotic /non-enhancing (NCR/NET).	
	2= Peritumoral edema (ED)	
	4 =Enhancing tumour (ET).	

Table 5.2: The Brats 2018 Validation data information

MICCAI_BraTS_2018_Data_Validation	
Gliomas grade	Unknown
# of patients	66
Modalities	T1, T1CE, T2, FLAIR
Dimension	(240,240,155)
Format	Nifty (Nii.gz)
Ground truth	Not available

5.4.2 Data Pre-processing:

Normalizing the data is an essential step in all machine learning tasks, to avoid the domination of some features to others and gain the right information from all relevant features. In our case we propose to use z-score normalisation [82, 83] to get a normal distribution of all voxels in each image modality by computing the mean and standard deviation only for the brain region (non zero part). [84] and update each voxel value by a new value computed by equation 5.1.

$$Z = \frac{v - \mu}{\sigma} \tag{5.1}$$

Where v is the voxel value and μ, σ are the mean and the standard deviation voxels respectively.

Figure 5.2: Example of normalized slice with the Z-score formula on slice 102 extracted from T1contrast enhanced file of patient named "Brats18_TCIA09_141_1"

5.4.3 Implementation details:

Our deep learning proposed approach was implemented using python3 over the Tensorflow Keras library and executed on UB2-HPC (University of BATNA 2)

GPU node which contains 4 GPUs configured with CUDA 10.0 and CUDNN 5.6.7. In this case, the official Brats training data was splited randomly into Train set 90% and Valid set 10% with 42 as a random state to build the model, which is later tested over the official validation set, provided by the organisation of 66 patients which is used in our case as test set.

Table 5.3: The PP-UNET Model used datasets

Datasets	# of patients	Data source
Train set	256	MICCAI_BraTS_2018_Data_Training (90%)
Valid set	29	MICCAI_BraTS_2018_Data_Training (10%)
Test set	66	MICCAI_BraTS_2018_Validation_Data

To obtain results that are more significant in very efficient way we proposed resolving the problem as binary segmentation problem rather than multi-class segmentation, where we turned the original label classes to the proposed classes in the challenge, we trained the model for each class separately, and the results were fused. The new classes are exhibited in Table 5.4 :

Table 5.4: The segmentation Classes

New Class	Labels
Whole Tumor (WT)	All the classes labels (NCR/NET 1 +edema 2 +Enhancing Tumor 4)
Tumor Core (TC)	Is (NCR/NET 1 +Enhancing Tumor 4)
Enhancing Tumor (ET)	Is the enhancing Tumor 4 label

for each class each model is trained over 2D MRI images in the axial view each cropped from(240,240,155) to (160,160,128) with the following indices [40:200,40,200,12:140] the slices are organized as follows:

- Each slice in Each modality must be from the same patient and the same index to respect the parallel pathways inputs
- the slices in the same modality are charged as follows: each index slice is charged from all patient's files to guarantee the perfect data shuffle.

Table 5.5: The PP-UNET architecture Layers information

PP_UNET	Layers	Config	Input Size	Output Size
Input	Input Layer	-	160*160*1	-
Encoder	Conv+ BN+LR+	Conv(3*3)	160*160*1	160*160*16
	Conv+ BN+LR+	Padding=same	160*160*16	160*160*16
	Maxpooling +Dropout	Pool(2*2)	160*160*16	80*80*16
	Conv+ BN+LR+	Conv(3*3)	80*80*16	80*80*32
	Conv+ BN+LR+	Padding=same	80*80*32	80*80*32
	Maxpool +Dropout	Pool(2*2)	80*80*32	40*40*32
	Conv+ BN+LR+	Conv(3*3)	40*40*32	40*40*64
	Conv+ BN+LR+	Padding=same	40*40*64	40*40*64
	Maxpooling +Dropout	Pool(2*2)	40*40*64	20*20*64
Bridge	Conv+ BN+LR+	Conv(3*3)	20*20*64	20*20*128
	Conv+ BN+LR+	Padding=same	20*20*128	20*20*128
Decoder	Conv2dtranspose+concat+	Kernel(3*3)	20*20*128	40*40*64
	Dropout+ Conv+ BN+LR+	Stride=2	40*40*64	40*40*64
	Conv+ BN+LR	Padding=same		
	Conv2dtranspose+concat+	Kernel(3*3)	40*40*64	80*80*32
	Dropout+ Conv+ BN+LR+	Stride=2	80*80*32	80*80*32
	Conv+ BN+LR	Padding=same		
	Conv2dtranspose+concat+	Kernel(3*3)	80*80*32	160*160*16
	Dropout+ Conv+ BN+LR+	Stride=2	160*160*16	160*160*16
	Conv+ BN+LR	Padding=same		
Concatenation	Concat (path1,path2,path3,path4)	Output of each path	160*160*16*4	160*160*64
Output	Convolution	Conv(1*1) sigmoid	160*160*64	160*160*1

Table 5.6: The network parameters

Parameters	values
Views	Axial
# Training Samples	32768
# Vaidation Samples	3712
Initial filters N°	16
Batch size	16
Dropout rate	0.4
Optimizer	Adam (default)
Epochs	100
Model checkpoint	Active
Early stopping	Active (patience epochs=10)
Trainable params	2147649

The final Network details and parameters are in tables 5.5 and 5.6 Initial filters N° represents the number of convolution kernels and doubles for each network plain block. Batch size is the number of training samples used in one iteration. Epochs are the number of times that the algorithm train over the whole samples. The dropout rate is the proportion of randomly selected nodes to set to zero with early stopping used to stop the learning process if the validation dice score didn't improve after 10 patience epochs and Model checkpoint active to save the best model with the best validation dice score,all used to prevent over-fitting. Finally, the Adam optimizer with default parameters was used to fine-tune the network weights and Trainable params as the nubmer of trainable parameters in the network.

5.4.4 Evaluation Metrics

Evaluation metrics are quantitive measures used to assess the performance and the effectiveness of machine learning models. They are frequently computed based on the confusion matrix.

All evaluation metrics used are those proposed by the challenge organization [12]. The evaluation metrics used are Dice Score, Sensitivity, Specificity, and Hausdorff 95 distance, those metrics are highly recommended for medical image segmentation[85]. Most of the time the healthy tissue pixels are larger than the tumours ones which means that the data suffer from the class imbalance distribution problem which prevent the right learning and lead the model to learn only about the frequent class label where the images segmentation problem focuses on the infrequent class[86]. In such cases, all these metrics are well adapted to address the class imbalance problem. They only consider the segmentation class and not the background class[87].

Dice Score: it measures the overlap of the predicted mask and the ground truth. it does not account for the background class which make it more suitable in the class imbalance problem. The dice score outputs is a score in the range [0,1] where 1 is a perfect overlap. Thus, (1-Dice) can be used as a loss function.

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \tag{5.2}$$

Where A and B are the prediction mask and the Ground truth respectively.

Sensitivity: known also as true positive rate refers to the proportion of positives that are correctly predicted comparing to the ground truth in our case measures the proportion of true predicted and segmented tumor pixels which is very important for patients diagnosis

$$Sensitivity = \frac{TP}{TP + FN} \tag{5.3}$$

Where TP is the number of true positives and FN the number of false negatives.

Specificity: or true negative rate is the proportion of negatives that were correctly predicted.

$$Specificity = \frac{TN}{TN + FP} \tag{5.4}$$

Where TN is the number of true negatives and FP is the number of false positives.

Hausdorff distance: it measures the greatest distance between the predicted tumor area A and the ground truth tumor area B[88] Formally, it is defined as follows:

$$H(A, B) = \max(h(A, B), h(B, A)) \tag{5.5}$$

Where A and B are the prediction map and the Ground truth respectively.

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \tag{5.6}$$

Where a and b are points from the sets A and B respectively, and $\|.\|$ is any metric between these points ; for example the Euclidian distance between a and b, $h(A, B)$ called directed Hausdorff distance from A to B which compute the maximum distance of a point a to the nearest point b and oppositely in $h(B, A)$. The Hausdorff95 indicates the distance of the tow point sets is with a percentile value of 95.

$$h(A, B) = \max_{a \in A}(95\%) \min \|a - b\| \tag{5.7}$$

5.5 Results

The table 5.7 exhibits the results of the model PP-UNET over the training and the validation set.

Table 5.7: The PP-UNET training and validation results

Metrics	Dice			Hausdorff95			Sensitivity			Specificity		
Classes	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET
Train Set	0.912	0.888	0.863	7.112	7.008	5.973	0.924	0.908	0.892	0.968	0.986	0.994
Valid Set	0.912	0.874	0.864	6.566	6.007	4.914	0.931	0.906	0.854	0.963	0.985	0.994

The tables 5.8, 5.9 shows the PP-Unet detailed results over the test set, all results were collected from the online submission system CBICA server `https://ipp.cbica.upenn.edu/` which is an Image Processing Portal available for authorized users to access the Center for Biomedical Image Computing and Analytics computing cluster and imaging analytics.

Table 5.8: The PP-UNET Dice and Hausdorff95 Test results

Metrics	Dice			HD95(mm)		
classes	WT	TC	ET	WT	TC	ET
Mean	0.890	0.787	0.745	6.109	10.620	8.046
Median	0.908	0.873	0.857	3.741	5.656	2.236
25quantile	0.881	0.724	0.780	2.828	3.0	1.414
75quantile	0.935	0.925	0.886	6.123	15.367	4.414

Table 5.9: The PP-UNET Sensitivity and Specificity Test results

Metrics	Sensitivity			Specificity		
classes	WT	TC	ET	WT	TC	ET
Mean	0.912	0.821	0.791	0.992	0.996	0.998
Median	0.945	0.927	0.863	0.993	0.997	0.998
25quantile	0.907	0.768	0.769	0.990	0.995	0.997
75quantile	0.968	0.958	0.916	0.996	0.998	0.999

The following figures 5.3, 5.4 and 5.5 are examples from the test set Table 5.1 segmented by our PP-UNET

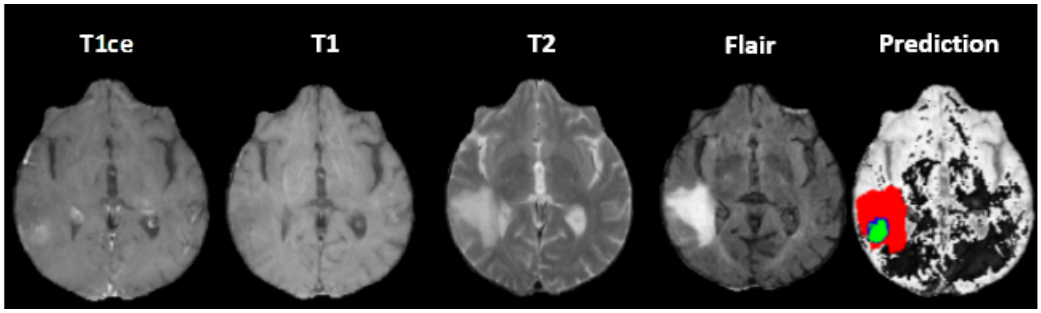


Figure 5.3: Segmentation result of the brain tumor structures by the Proposed PP-UNET on patient named "Brats18_TCIA02_400_ 1"/slice 70 on the axial view

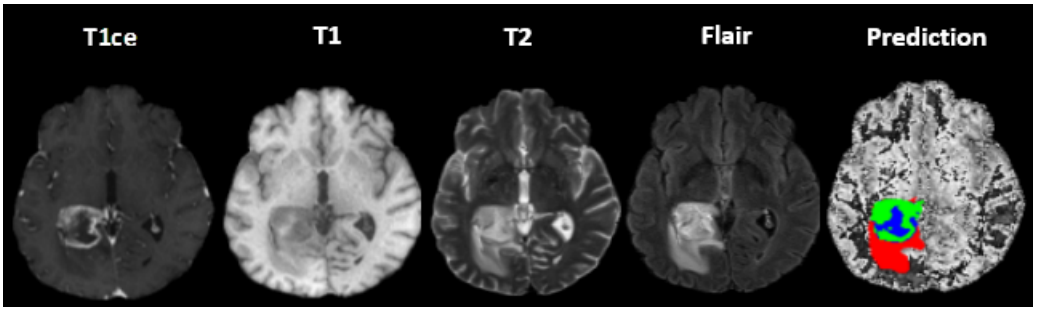


Figure 5.4: Segmentation result of the brain tumor structures by the Proposed PP-UNET on patient named "Brats18_CBICA_BHF_1 "/slice 69 on the axial view

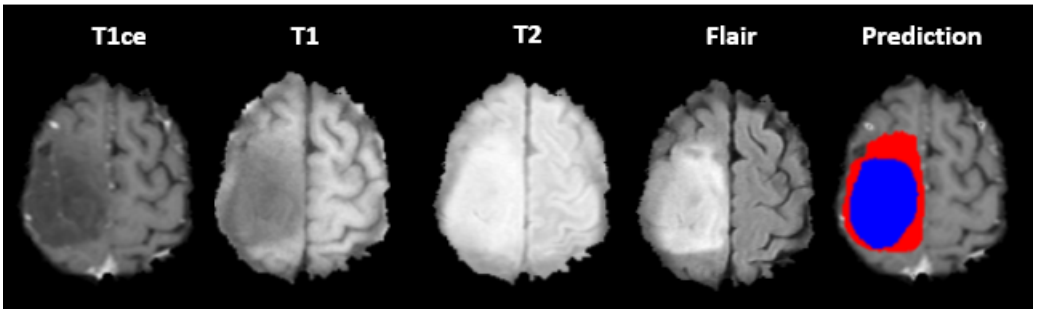


Figure 5.5: Segmentation result of the brain tumor structures by the Proposed PP-UNET on patient named "Brats18_TCIA09_248_ 1"/slice 125 on the axial view

5.5.1 PP-UNET vs RGBA

in this section, we propose to compare the results of The PP-UNET architecture against the RGBA version to evaluate the effectiveness of the models. The RGBA model is one branch version, it involves stacking each corresponding slice from each modality (T1ce, T1, T2, and Flair) on the axial view as RGBA color images

and provide them directly to the U-NET part of our proposed model as input the whole architecture is provided in the following figure.

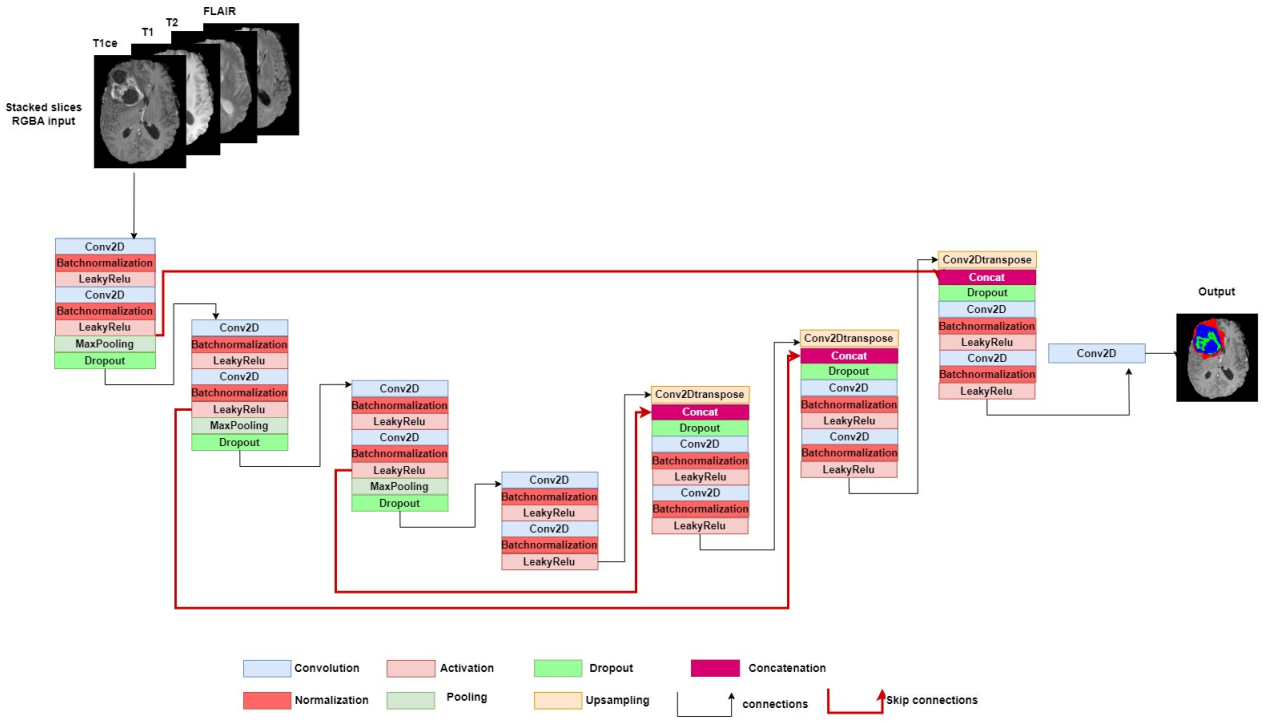


Figure 5.6: The RGBA architecture

The Table 5.10 provide the results of the RGBA model against the PP-UNET results over the test.

Table 5.10: The PP-UNET Vs RGBA results

Metrics	Dice			HD95(mm)			Sensitivity			Specificity		
Classes	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET
PP_UNET	0.890	0.787	0.745	6.109	10.62	8.046	0.912	0.821	0.791	0.992	0.996	0.998
RGBA	0.891	0.786	0.761	8.210	14.113	5.584	0.909	0.800	0.798	0.992	0.996	0.998

5.5.2 PP-UNET against existing Methods

in this section, 5.11 summarizes the results of the state-of-the-art techniques that refer to the currently existing approach or method that has achieved high performance in the field to compare them against our PP-UNET results to analyze the strengths and weaknesses of both methods and determine if our approach outperforms state-of-the-art or provides any notable improvements.

Table 5.11: PP-UNET and state-of-the-art techniques comparison results

Method	type	DICE			HD95(mm)		
		WT	TC	ET	WT	TC	ET
[63]	3D	0.910	0.866	0.823	4.51	6.85	3.92
[64]	3D	0.912	0.863	0.808	4.27	6.52	2.41
[65]	3D	0.903	0.847	0.796	3.55	4.17	4.93
	3D U-Net+ TTA	0.873	0.783	0.754	5.90	8.03	4.53
[66]	3D WNet+TTA	0.895	0.730	0.770	4.92	11.13	4.44
	3D Cas+TTA	0.902	0.858	0.797	6.18	6.37	3.13
	3D VGG	0.872	0.760	0.751	_	_	_
	3D inception2	0.877	0.773	0.753	_	_	_
[67]	3D inception3	0.873	0.776	0.781	_	_	_
	3D Densely	0.874	0.755	0.729	_	_	_
	Ens	0.881	0.777	0.773	_	_	_
[68]	3D	0.904	0.836	0.776	5.17	6.27	3.51
[76]	3D	0.895	0.797	0.777	9.13	8.67	3.90
[79]	3D	0.900	0.820	0.800	6.0	7.5	4.4
[69]	2D	0.868	0.805	0.783	8.12	9.84	3.72
[70]	2D	0.898	0.811	0.751	_	_	_
[77]	2D	0.895	0.823	0.813	4.05	6.34	2.93
[78]	2D	0.872	0.808	0.772	5.62	8.36	3.57
PP-UNET	2D	**0.890**	**0.787**	**0.745**	**6.10**	**10.62**	**8.04**

5.6 Discussion

From the provided results in tables 5.7, 5.8and 5.9, the PP-UNET model is performing and proves its capability of generalization over the unseen data which is the Test set here and the results are close to each other which proves the absence of overfitting and the efficiency of the used dropout layers, From the provided results in Table 5.10, the RGBA model results are slightly raiser than the PP-UNET ones but without any significance, both results are comparably similar however, in terms of memory consumption, training time, and trainable parameters the RGBA version is more optimized than the PP-UNET from the idea of one branch in RGBA against four branches in the PP-UNET which is its major limit. as an advantage of this experiment, we have got an important insight into our proposal that our customized

changes to the U-net network were a good choice and beneficial.

from the table 5.11 which exposes the results of the existing works, our 2D PP-UNET performs well compared to the 2D proposed works, the margin between the results is not very significant for both whole tumor and tumor core classes but in the case of the ET classes the majority of works reached significant rates of dice scores. when compared to the 3D approaches, our approach still performs well for both whole tumor and tumor core classes with the same note of the ET classes the majority of works reached significant rates of dice scores especially the top-ranked approaches highlighted in red color. the advantage of this approach compared to the existing work is the simplicity and straightforward implementation with the same limited mentioned when compared to the RGBA version.

5.7 Conclusion

In this work, we investigated an approach inspired by the reputed U-net architecture to tackle the brain tumour segmentation problem from multimodal MRI images called PP-UNET, we suggested using multiple pathways of an updated U-net each dealing with a specific 2D slice extracted from a specific modality. we suggested exploring this task in RGBA form where we used one pathway fed with the four slices stacked as an RGBA color image . The results were also verified against the proposed works where our method achieved interesting results and using additional improvement seemed as the next verified step with our method for more results enhancement.

Action is the Foundational Key to Success

6

PU-NET Architecture for Brain Tumor Segmentation

Contents

6.1 Introduction

In this chapter we introduce a new perspective to our deep-learning architecture which provides significant improvement in the segmentation of the brain glioma tumor. we proposed a novel approach, called PU-NET, for automatic brain tumor segmentation in multi-modal magnetic resonance images(MRI). We introduced an input processing block to a customized fully convolutional network derived from

the U-Net network to handle the multi-modal inputs. We performed experiments over the Brain Tumor Segmentation(BRATS) dataset collected in 2018 too and achieved Dice scores of 90.5%,82.7%, and 80.3% for the whole tumor, tumor core, and enhancing tumor classes, respectively. This study provides promising results compared to the deep learning methods used in this context[89].

6.2 Methodology

In this work, we propose a deep-learning architecture for brain tumor segmentation based on the encoder-decoder aspect and inspired by the reputed network used to segment biomedical images called U-Net [59]with significant changes. This architecture is called PU-NET, which refers to its two main parts: an input processing part and a customized U-Net network part. They are based on a central block known as a plain block. It encompasses two consecutive convolution layers of (3×3) kernel size with stride(1,1), each followed by a Batch normalization layer added to ensure the network data normalization and improve the training convergence and speed, followed by a LeakyRelu layer as an activation function with a slope α set to 0.01 Eq.6.1.

$$LeakyRelu(x) = \begin{cases} \alpha x, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases} \tag{6.1}$$

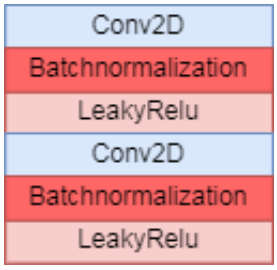


Figure 6.1: The architecture of the plain block

In contrast with the original proposition, all used convolutions are padded to preserve and mitigate the loss of information at the image borders. This loss is significant as the pooling operation applied later will induce information loss.

We also proposed using the LeakyRelu activation instead of the Relu function in the original proposition to prevent the dying Relu [90, 91] problem that may happen where most of the Relu Eq. 6.2 neurons only output zero because of negative inputs during the learning process, which may cause the inactivity of a significant part of the network neurons and negatively affect the results.

$$Relu(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases} \tag{6.2}$$

The input processing part handles the multi-modal MRI scans quickly and efficiently. It consists of four input layers, each for a specific modality 2D slice (T1ce, T1, T2, Flair), each followed by a plain block to extract relevant features from each modality separately as a first step. Then, a concatenation layer is used to merge the four outputs and transfer the output to the U-Net part as the primary input.

The U-Net network part keeps the same divisions as the original proposal [59]. It comprises three parts: An encoder path(contracting path), a decoder path(an expansive path), and a simple bridge between them. The encoder path is similar to the convolution neural network (CNN) feed-forward pass. It contains three levels against four in the original U-Net to alleviate the network's training parameters and reduce memory consumption, complexity, and running time. Each level includes a plain convolution block, except for the first one, where the concatenation output replaces it, followed by a max-pooling layer with stride two and a dropout layer that regularizes the network to avoid over-fitting. This path is used for extracting relevant features and capturing the context of the input images to enable the segmentation task.

The expansive path is symmetric to the contracting path in the number of levels, each comprising a stack of layers. It starts with a padded convolution transpose of (3×3) kernel size with stride(2,2) used to recover some loosed information in previous convolution layers from the feature map. It is also followed by a concatenation layer that merges the output of the convolution plain block in the contracting side with the corresponding output of the convolution transpose to get more information

about the spatial resolution after a convolution plain block is applied, followed by a dropout layer. The two paths are related by a bridge, which is a plain block.

Finally, a (1×1) convolution with sigmoid activation is applied to generate the output probabilities for the segmentation map. Figure 6.2 exposes the detailed PU-NET architecture layers.

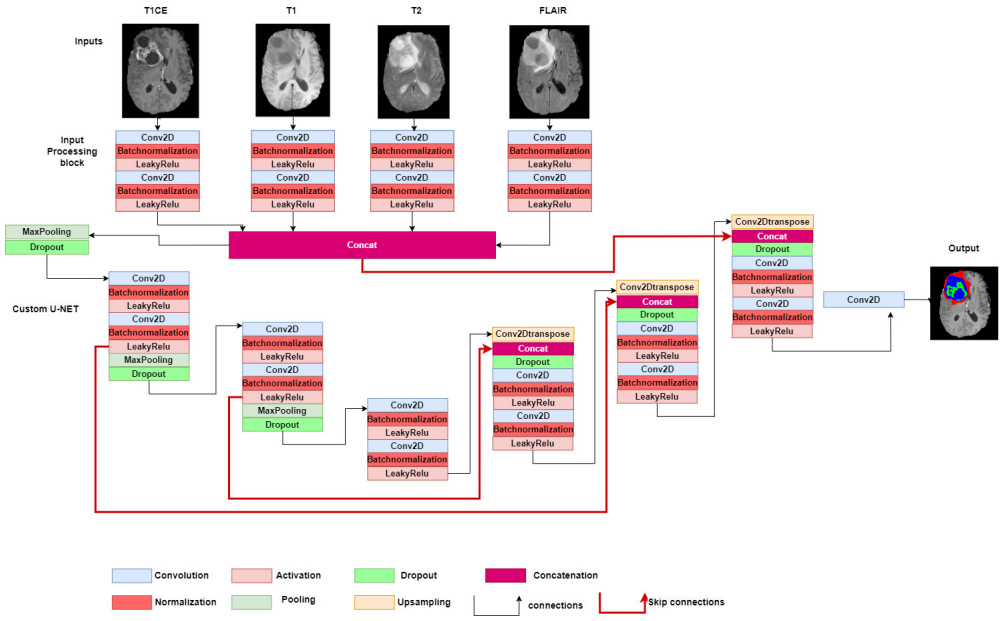


Figure 6.2: The PU-NET model architecture used to segment brain tumor structures

6.3 Experimentations

In this proposition we suggest using the same Dataset Brats 2018 in section 5.4.1 with the same pre-processing operation in section 5.4.2 .

6.3.1 Implementation details:

Our deep learning proposed approach was implemented using python3 over the Tensorflow Keras library and executed on UB2-HPC (University of BATNA 2) GPU node which contains 4 GPUs configured with CUDA 10.0 and CUDNN 5.6.7. In our case, the official Brats training data was splited randomly into Train set 80% and Valid set 20% with 42 as a random state to build the model, which is later

tested over the official validation set, provided by the organisation of 66 patients without Ground truth which is used in our case as test set.

Table 6.1: The PU-NET Model used datasets

Datasets	# of patients	Data source
Train set	228	MICCAI_BraTS_2018_Data_Training (80%)
Valid set	57	MICCAI_BraTS_2018_Data_Training (20%)
Test set	66	MICCAI_BraTS_2018_Validation_Data

similar to the previous study, we proposed resolving the problem as a binary segmentation problem rather than multi-class segmentation, where we turned the original label classes to the proposed classes in the challenge, we trained the model for each class separately, and the results were fused. The classes are exhibited in Table 5.4.

We propose also to use this training process over the three MRI views (Axial, Coronal, and Sagittal) separately and ensemble the results over models to ensure the collection of the 3D contextual information. All image modalities and their corresponding ground-truth were cropped from (240,240,155) size to (128,128,128) with the following indices [56:184, 56:184, 12:140] to reduce memory consumption. Table 6.2 shows our network parameters and table 6.3 shows our network details:

Table 6.2: PU-NET network parameters

Parameters	Values
Views	Axial/Coronal/Sagittal
Training Samples	29184
Validation Samples	7296
Initial filters N°	16
Batch size	32
Epochs	200
Dropout rate	0.4
Early stopping	Active (patience epochs=10)
Model checkpoint	Active
Optimizer	Adam (default)
Trainable params	565281

Table 6.3: The PU-NET architecture Layers information

PU-NET	Layers	Config	Input Size	Output Size
Input1,input2,input3,input4	Input Layer	-	128*128*1	-
Input processing	c1=Plain Block(input1) c2=Plain Block(input2) c3=Plain Block(input3) c4=Plain Block(input4)	Conv(3*3) Padding=same	128*128*1	128*128*16
Concatenation	concat (c1,c2,c3,c4)		128*128*16	128*128*64
Encoder	Maxpooling +Dropout	Conv(3*3) Padding=same Pool(2*2)	128*128*64	64*64*64
	Plain Block Maxpool +Dropout	Conv(3*3) Padding=same Pool(2*2)	64*64*64 64*64*32	64*64*32 32*32*32
	Plain Block Maxpooling +Dropout	Conv(3*3) Padding=same Pool(2*2)	32*32*32 32*32*64	32*32*64 16*16*64
Bridge	Plain Block	Conv(3*3) Padding=same	16*16*64 16*16*128	16*16*128 16*16*128
Decoder	Conv2dtranspose+ concat+Dropout+ Plain Block	Kernel(3*3) Stride=2 Padding=same	16*16*128 32*32*64	32*32*64 32*32*64
	Conv2dtranspose+ concat+Dropout+ Plain Block	Kernel(3*3) Stride=2 Padding=same	32*32*64 64*64*32	64*64*32 64*64*32
	Conv2dtranspose+ concat+Dropout+ Plain Block	Conv(3*3) Stride=2 Padding=same	64*64*32 128*128*16	128*128*16 128*128*80
Output	Convolution	Conv(1*1) sigmoid	128*128*80	128*128*1

6.3.2 Evaluation Metrics

The same previous evaluation metrics are used, detailed in section 5.4.4

6.4 Results

6.4.1 Training Results

Tables 6.4 and 6.5 exhibit the results of our PU-NET approach on the training and validation data respectively(Table 6.1) .the results are reported also over the MRI views separately in addition to the validation ensemble results realized by multi-view label fusion using majority voting which prove the efficiency of the multi-view exploration and fusion in the result's improvement without overffiting .

Table 6.4: The PU-NET Dice and Hausdorff95 training and validation results

Metrics		**Dice**			**Hausdorff95**		
	Classes	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**
Train set	**Axial**	0.927	0.934	0.877	3.524	2.388	1.913
	Coronal	0.924	0.918	0.864	3.579	2.448	1.954
	Sagittal	0.937	0.910	0.880	3.636	2.564	1.958
Valid set	**Axial**	0.917	0.806	0.859	3.388	2.323	1.756
	Coronal	0.918	0.835	0.861	3.423	2.339	1.792
	Sagittal	0.916	0.812	0.857	3.496	2.433	1.800
	Ensemble	0.925	0.839	0.868	3.546	2.360	1.771

Table 6.5: The PU-NET Sensetivity and Speceficity training and validation results

Metrics		**Sensetivity**			**Speceficity**		
	Classes	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**
Train set	**Axial**	0.922	0.935	0.885	0.953	0.979	0.991
	Coronal	0.933	0.921	0.895	0.951	0.979	0.990
	Sagittal	0.933	0.923	0.893	0.952	0.979	0.991
Valid set	**Axial**	0.908	0.792	0.836	0.950	0.981	0.993
	Coronal	0.921	0.830	0.861	0.949	0.981	0.992
	Sagittal	0.910	0.805	0.837	0.950	0.981	0.992
	Ensemble	0.920	0.820	0.849	0.950	0.981	0.992

6.4.2 Test Results

Tables 6.6 and 6.7 show the results obtained by our approach tested on the Test set in Table 6.1 in each MRI view, in addition to the ensemble result, which was realized by multi-view label fusion using majority voting too in addition to the result of ensembled 5-fold validation,this results proves the ability of the model to generalize over the test set without overfitting. all results were collected from the online submission system CBICA server `https://ipp.cbica.upenn.edu/` which is an Image Processing Portal available for authorized users to access the Center for Biomedical Image Computing and Analytics computing cluster and imaging analytics.

Table 6.6: The PU-NET Dice and Hausdorff95 test results

Metrics	**Dice**			**Hausdorff95**		
Classes	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**
Axial	0.890	0.791	0.781	15.912	12.288	6.231
Coronal	0.889	0.787	0.782	12.134	18.774	11.002
Sagittal	0.894	0.792	0.776	11.174	13.732	3.486
ensemble	0.904	0.815	0.803	5.054	7.732	3.019
5 fold	0.905	0.827	0.803	4.638	8.679	3.573

Table 6.7: The PU-NET Sensetivity et Speceficity test results

Metrics	**Sensetivity**			**Speceficity**		
Classes	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**
Axial	0.907	0.807	0.810	0.993	0.997	0.997
Coronal	0.923	0.811	0.821	0.991	0.996	0.997
Sagittal	0.917	0.829	0.803	0.992	0.996	0.997
ensemble	0.924	0.817	0.820	0.993	0.997	0.997
5 fold	0.921	0.817	0.815	0.993	0.997	0.998

The following figures 6.3, 6.4 and 6.5 are examples from the test set Table 6.1 segmented by our PU-NET.

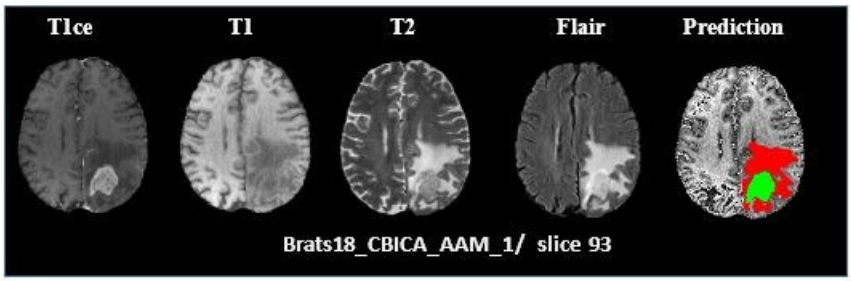


Figure 6.3: Segmentation result of the brain tumor structures by the Proposed PU-NET on patient named "Brats18_CBICA_AAM_ 1"/slice 93 on the axial view.

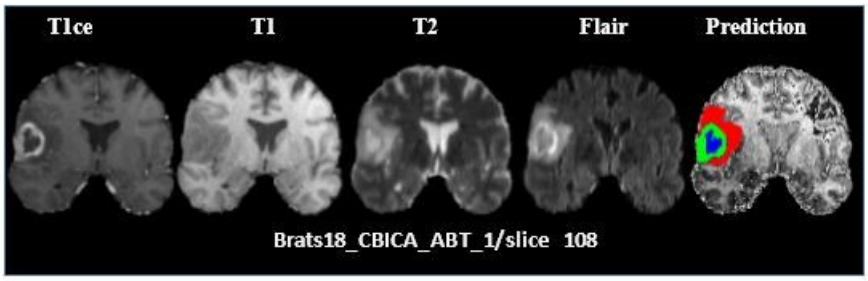


Figure 6.4: Segmentation result of the brain tumor structures by the Proposed PU-NET on patient named "Brats18_CBICA_ABT_1"/slice 108 on the coronal view.

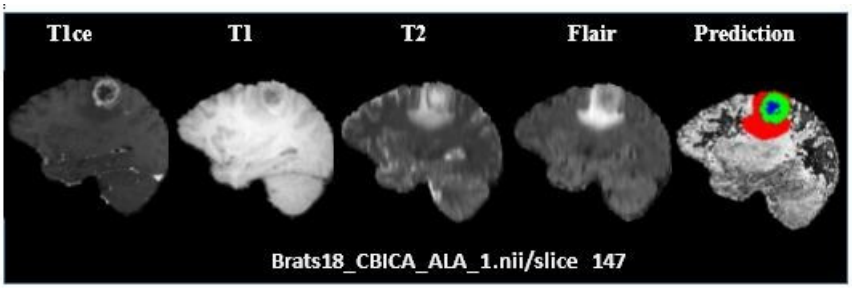


Figure 6.5: Segmentation result of the brain tumor structures by the Proposed PU-NET on patient named "Brats18_CBICA_ALA_1"/slice 108 on the sagittal view

6.4.3 PU-NET vs RGBA vs 3D

In this section, our primary goal is to evaluate the effectiveness of our 2D PU-NET model by specifically examining the impact of the input processing block.We compare baseline techniques that propose different approaches for processing multi-modal MRI data.The first baseline technique involves stacking each corresponding slice from each modality (T1ce, T1,T2, and Flair) as RGBA,The second baseline

technique involves providing the multi-modal MRI data in the 3D format to the PU-NET architecture. This means the data is preserved in its volumetric structure instead of processed as 2D slices. Additionally, the architecture operations are modified to accommodate the 3D format. Tables 6.8 and 6.9 expose the results of RGBA model, the results were reported from separate views.

Table 6.8: The RGBA Dice and Hausdroff95 Results on the test set

Metrics	**Dice**			**Hausdroff95**		
Classes	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**
Axial	0.891	0.786	0.761	8.210	14.113	5.584
Coronal	0.892	0.793	0.761	13.670	18.720	12.491
Sagittal	0.890	0.789	0.724	11.597	18.722	10.457

Table 6.9: The RGBA Sensetivity and Specificity Results on the test set

Metrics	**Sensetivity**			**Specificity**		
Classes	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**
Axial	0.909	0.800	0.798	0.992	0.996	0.998
Coronal	0.918	0.795	0.816	0.992	0.997	0.997
Sagittal	0.918	0.800	0.815	0.992	0.996	0.997

Tables 6.10 and 6.11 reports the results the ensemble version of the RGBA above models and compare them against the 3D PU-NET version and to our 2D PU-NET approach.

Table 6.10: PU-NET and Baseline techniques comparaison Dice and Hausdroff95 results

Metrics	**Dice**			**Hausdroff95**		
Classes	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**
RGBA	0.903	0.801	0.772	5.442	6.996	3.967
3D PU-NET	0.884	0.710	0.734	7.898	28.930	29.695
PU-NET	**0.905**	**0.827**	**0.803**	**4.638**	**8.679**	**3.573**

Table 6.11: PU-NET and Baseline techniques comparaison Sensetivity and Specificity results

Metrics	Sensetivity			Specificity		
Classes	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**
RGBA	0.921	0.766	0.781	0.993	0.998	0.998
3D PU-NET	0.911	0.780	0.750	0.980	0.980	0.982
PU-NET	**0.921**	**0.817**	**0.815**	**0.993**	**0.997**	**0.998**

these results provide compelling evidence that our PU-NET model performs exceptionally well when compared to stacking slices as RGBA color images. These results further validate the efficiency and effectiveness of our input processing block.However, it is worth noting that the 3D version of our model exhibits some drawbacks. Specifically, it requires a significantly longer training time, exceeding 20 hours under GPU for each class label. In contrast,our 2D PU-NET model takes a maximum of 1 hour for execution. Furthermore, the results obtained from the3D version diverge significantly from those of our 2DPU-NET model, particularly in the TC and ET classes.

6.4.4 PU-NET Versus the Proposed Methods

In this section, table 7.4 summarizes the results of the state-of-the-art techniques that refer to the currently existing approach or method that has achieved high performance in the field to comparethem against our PU-NET results to analyze the strengths and weaknesses of both methods and determine if our approach outperforms state-of-the-art or provides any notable improvements

Table 6.12: PU-NET and state-of-the-art techniques comparison results

Method	type	DICE			HD95(mm)		
		WT	TC	ET	WT	TC	ET
[63]	3D	0.910	0.866	0.823	4.51	6.85	3.92
[64]	3D	0.912	0.863	0.808	4.27	6.52	2.41
[65]	3D	0.903	0.847	0.796	3.55	4.17	4.93
	3D U-Net+ TTA	0.873	0.783	0.754	5.90	8.03	4.53
[66]	3D WNet+TTA	0.895	0.730	0.770	4.92	11.13	4.44
	3D Cas+TTA	0.902	0.858	0.797	6.18	6.37	3.13
	3D VGG	0.872	0.760	0.751	_	_	_
	3D inception2	0.877	0.773	0.753	_	_	_
[67]	3D inception3	0.873	0.776	0.781	_	_	_
	3D Densely	0.874	0.755	0.729	_	_	_
	Ens	0.881	0.777	0.773	_	_	_
[68]	3D	0.904	0.836	0.776	5.17	6.27	3.51
[76]	3D	0.895	0.797	0.777	9.13	8.67	3.90
[79]	3D	0.900	0.820	0.800	6.0	7.5	4.4
[69]	2D	0.868	0.805	0.783	8.12	9.84	3.72
[70]	2D	0.898	0.811	0.751	_	_	_
[77]	2D	0.895	0.823	0.813	4.05	6.34	2.93
[78]	2D	0.872	0.808	0.772	5.62	8.36	3.57
PU-NET	2D	**0.905**	**0.827**	**0.803**	**4.63**	**8.67**	**3.57**

6.4.5 Discussion

From Table 7.4, our proposed method PU-NET outperforms the 2D approaches in terms of both Dice and Hausdorff95 metrics. While it is not away from the top-ranked approaches in the BRATS 2018 challenge and this is due to the fine parameters tuning; it is important to consider, too, that Myronenko[63], Isenseeet al[64], and McKinley et al [65] utilized data augmentation techniques and additional data in the co-training process. In contrast, our work did not incorporate any additional or augmented data. Furthermore, when computing the statistical p-values between the dice scores, we found values of 0.61,0.72 and 0.66, respectively, suggesting that the results are not highly significant compared to our findings.

Regarding Wang et al [66], our 2D approach outperformed all of their proposed

3D architectures with data augmentation, except for the cascaded architecture, which exhibited slightly better performance than our PU-NET in TC class segmentation, similar to the findings of RuiHua et al[68]. However, it is important to note that the p-values of 0.87 and 0.90 suggest these differences are not statistically significant. Moreover, our PU-NET approach demonstrated greater efficiency than the extended architectures proposed by A.Albiol et al [67], which are among the popular deep learning architectures used in the context of image segmentation, both [76, 79] works which are out of the BRATS 2018 challenge proposed 3D-based architectures with complex U-Net designs. Still, our results surpass them and support our the idea that sometimes complex architectures and using 3D format can increase the computational complexity without necessarily being efficient, so it is justifiable to shift the research focus towards 2D simple architectures.

In our perspective, the strength of our PU-NET model lies in incorporating the input processing block as an initial separate feature extractor for each MRI modality. It serves as a robust mechanism for integrating the multi-modal complementary information and improving the tumor segmentation accuracy. In addition, the simplicity of our U-Net architecture part and the use of 2D slices address major limitations in deep learning-based approaches, specifically regarding time and memory requirements, which are advantageous when working with large-size medical datasets. As with any approach, there are certain limitations associated with our proposed method; the input processing block in our approach is specifically designed to handle datasets with multi-modal images. While this benefits such datasets, it may not be suitable or optimal for datasets with single-modality images. Similar to other deep learning architectures, the interpretability of our approach is limited. Deep learning models, including our PU-NET, often function as complex black boxes, challenging understanding of the underlying decision-making process. Interpretability is an ongoing area of research in deep learning, and further efforts are needed to enhance the transparency and explainability of models like ours. Indeed, while our approach may have limitations, it is important to recognize

that its performance remains competitive. The results obtained with our 2D PU-NET architecture demonstrate promise and suggest potential avenues for further improvement in future research.

6.5 Conclusion

In this work, we investigated a novel approach inspired by the reputed U-net architecture to tackle the brain tumour segmentation problem from multimodal MRI images called PU-NET, we introduced an input processing block to an updated U-net that deal with multi-input 2D images collected from the different modalities. we suggested to explore this task in the three MRI views coronal, sagittal and axial and aggregate the final predictions to generate the final segmentation and benefit from the 3D contextual information .The results were compared later against the RGBA and the 3D version, also verified against the old machine learning techniques support vector machines and random forest and the proposed works where our method achieved interesting results and more using data augmentation and additional training data seemed as the next verified steps with our method for more results enhancement.

In the End We Only Regret the Chances We Didn't Take

7 Machine Learning for Brain Tumor Segmentation

Contents

7.1 Introduction

Before the emergence of deep learning, machine learning techniques like k-nearest neighbours, clustering, support vector machines, random forests, and others constituted a research-focused domain. These traditional machine learning algorithms were extensively studied and applied across various fields, providing valuable solutions for classification, clustering, regression, and pattern recognition tasks. Researchers

dedicated considerable efforts to optimizing and fine-tuning these algorithms to achieve better performance and generalization on diverse datasets.

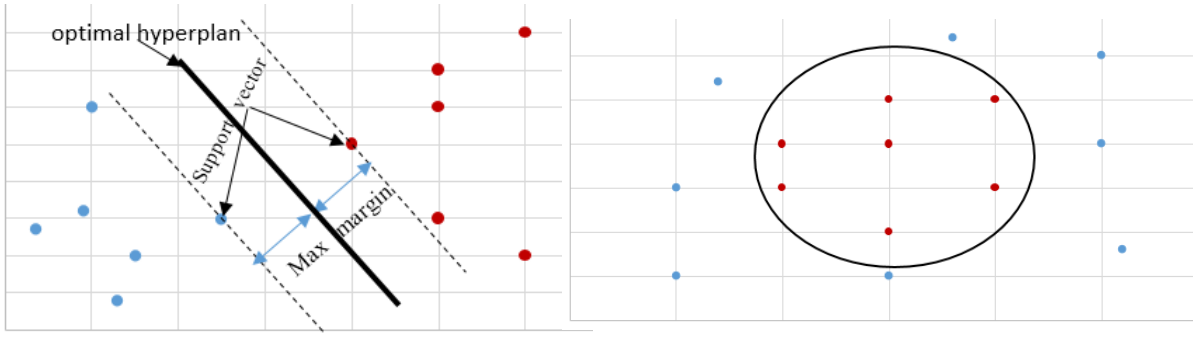
However, with the evolution of deep learning, there has been a notable shift in focus within the machine-learning community. Deep learning, particularly fueled by advancements in neural network architectures and the availability of large-scale labelled datasets, has demonstrated unprecedented success. this chapter presents a comparative study between our proposed support vector machine and random forest models and our previous PU-NET deep learning architecture in the brain tumour segmentation task to discuss the advantages and limits of each methodology in addition to an introduction of the class imbalance problem and its impact on image segmentation.

7.2 Support Vector Machines

Support vector machine[92] is one of the early supervised Machine learning techniques that was originally developed by Vladimir Vapnik and his colleagues in the 1990s. it is used for solving classification as well as Regression problems. the main goal is to find a hyperplane that best separates data points from different classes in a high-dimensional feature space. The hyperplane is chosen in such a way that it maximally separates the classes with the largest margin. The data points that are closest to the hyperplane and affect its position are known as support vectors, which gives the algorithm its name It is frequently used for classification where the goal is to create the best line called the decision boundary that separates n-dimensional data points in the space into classes. SVM can be in two types:

1. Linear SVM: used for linearly separable data which means the data points/vectors are separated

2. Non-Linear SVM: used in the case of non-linearly separable data using the kernel trick.Using a kernel function, SVM can still find a separating hyperplane by mapping the data to higher dimensional space. common kernels include: linear, polynomial, radial basis function and sigmoid kernels



(a) Linear SVM **(b)** non-Linear SVM

Figure 7.1: Support Vector Machines

7.3 Random Forest

Random Forest [93]is an ensemble learning technique that combines the predictions of multiple individual decision trees where each tree is trained on a subset of the training data to make a more accurate and robust model in terms of predictions,it uses a technique called bagging or bootstrapping to create diverse subsets of training data for each tree and a final voting is performed between the predictions of individual trees to determine the final prediction . The decision tree is a supervised machine-learning technique frequently used for both classification and regression tasks. it is a kind of graphical hierarchical representation of decisions and their outcomes.The tree is composed of

- Root Node: the top node of the tree that represents the entire dataset.
- Internal Node: nodes inside the tree that refer to a discriminating attribute or feature to split the data into subsets using a split criterion like Gini index and entropy for classification tasks and mean squared error MSE in regression.
- Branches: the edge connection between nodes is a test over attributes where each branch is a possible decision.
- Leaf Node: or terminal nodes is the final class label reached.

The process of building a decision tree involves recursively splitting the dataset at internal nodes based on the attribute that provides the best information gain.

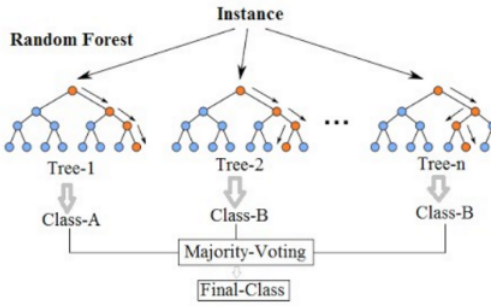


Figure 7.2: Random Forest Architecture

7.4 Methodology

For both SVM and RF, we suggest using predefined scikit-learn python models with customized feature selection and extraction steps. Feature extraction is a primordial step in this context, from the fact that each pixel has characteristics used for the discrimination between the tumor and the healthy pixels. For both SVM and RF a feature selection step was performed naively, the first main feature was the image pixel intensities in addition to a texture feature computed using the Gabor filter and some edge features computed by Sobel, Roberts, Canny, Scharr, prewit Gaussian and median kernels. in this case, the image pixels are represented as multidimensional data points stored in vectors where each case is a pixel feature.

7.5 Experimentation

In this section, we propose to evaluate the efficiency of our deep learning (PU-NET) architecture compared to ordinary machine learning techniques like Support Vector Machines (SVM) and Random Forest (RF).

7.5.1 Implementation Details

in this case, we suggested using the Train set in Table 6.1 of 228 patients as training data and the Valid set in Table 6.1 of 57 patients as a Test set and due to limited available resources and high memory consumption from SVM and RF, we customized both of datasets as follow:

- Each modality was pre-processed as explained in Section 5.4.2

- Each modality and ground truth were cropped from (240,240,155) to (128,128,30) with the following indices [56:184,56:184,sl-15:sl+15] where sl is the index of the slice in the ground truth which contains the maximum non zero pixels(or the large tumor region)
- The models were trained over the axial view only

and the remain information are reported in the table below:

Table 7.1: The Data Information

Modalities	Dataset	# of samples
T1ce, T1, T2, Flair	Train set	6840
	Test set	1710

Configuration Parameters:
RF and SVM were implemented in Python using numpy, opencv and sklearn libraries The configuration parameters for each model are:

- SVM: the Stochastic Gradient Descent (SGD) classifier was used with Radial Basis Function (RBF) kernel approximation and maximum number of iterations set to 10000,n_job set to -1 for parallel CPU cores execution and class_weight set to Balanced to address the class imbalance problem.
- RF: The Random Forest Classifier was used with number of estimators set to 10,n_job=-1 -1,class_weight set to Balanced and max_samples set to 0.2 and gini impurity criteria.
- PU-NET: keep the same number of epochs, dropout rate, early stopping, model checkpoint and optimizer parameters in Table 6.2 only the number of initial convolution filters and batch size was set to eight.

7.5.2 Results

Table 7.2 exhibits the training results of the three models evaluated with Dice Score , Sensetivity and speceficity previously defined in section 5.4.4

Table 7.2: PU-NET and Machine Learning techniques training comparison results

Models	SVM			RF			PU-NET		
Classes	WT	TC	ET	WT	TC	ET	WT	TC	ET
DICE	0.655	0.472	0.372	0.756	0.557	0.372	0.933	0.846	0.852
Sen	0.716	0.783	0.792	0.759	0.782	0.810	0.937	0.854	0.848
Spec	0.828	0.847	0.909	0.853	0.880	0.907	0.991	0.990	0.995

We can see that the performance of our deep learning approach is the best and from the experiments, we have noticed some limits of SVM and RF compared to PU-NET:

- The requirement for customizing the training data samples to 6840 is due to the SVM and RF training phase which requires high memory availability which affect also the small number of estimators 10 in RF model .
- The feature extraction step before launching the training which consume the memory.
- A tedious Feature Extraction step due to the manual naive choice.

whereas the PU-NET in 6.2 support 29184 samples from each modality as training data in the same execution environment because of the advantage of feature extraction during training time and the training over batch. without forgetting the flexibility in automatic extracting features using convolution kernels with trainable parameters.

The aim of machine learning techniques is to build models that have the ability to predict and generalize over unseen data that is why we propose to evaluate the previous models and test them on the Test set selected samples.

Table 7.3: PU-NET and Machine Learning techniques testing comparison results

Models	SVM			RF			PU-NET		
Classes	WT	TC	ET	WT	TC	ET	WT	TC	ET
DICE	0.668	0.463	0.354	0.747	0.505	0.343	0.925	0.788	0.799
Sen	0.707	0.787	0.827	0.736	0.727	0.818	0.931	0.778	0.900
Spec	0.828	0.845	0.910	0.851	0.879	0.908	0.989	0.984	0.998

From exhibited results in Table 7.3 the PU-NET method still performing better than SVM and RF and has a significant ability to generalize over the unseen data, especially in segmenting the small tumour regions TC and ET, this is maybe due to the quality of naively selected features in both SVM and RF.However, in the PU-net approach, this step is full automatic,trainable and adaptable to the training data and this which justify the abundance of deep learning based approaches in medical image segmentation and in general image analysis tasks.

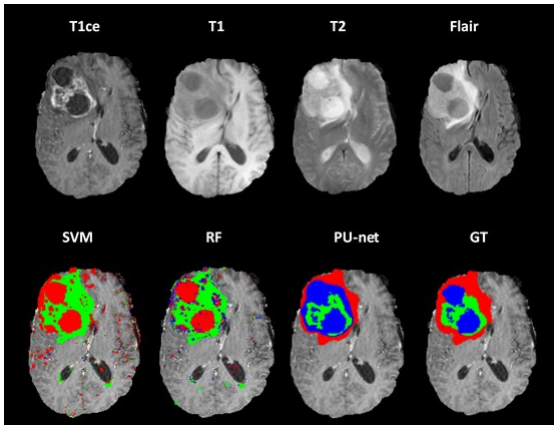


Figure 7.3: Segmentation results of the brain tumor structures of two slices from an example patients (Brats18_TCIA08_469_1/slice 78) in the test set by the proposed method PU-NET against SVM , RF comparing to the Ground Truth (GT)

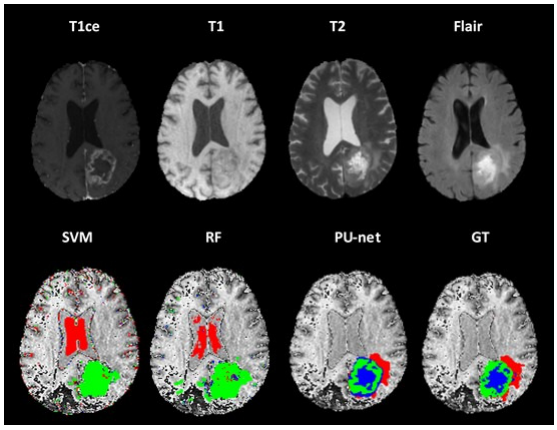


Figure 7.4: Segmentation results of the brain tumor structures of two slices from an example patient (Brats18_CBICA_ATB_1/slice 92) in the test set by the proposed method PU-NET against SVM , RF comparing to the Ground Truth (GT)

7.6 Class Imbalance and Evaluation Metrics

The class imbalance[94] is regarded as a reputed challenge that faces image segmentation with machine learning techniques[95], It refers to a situation in which the distribution of the label classes is significantly imbalanced, this means that one class often the majority class is represented by a remarkably higher number of instances and predominate other classes which can impact the model learning process, leading to a highly biased model toward the majority class and a neglect of the minority class due to the difficulty to effectively capturing its feature. The class imbalance in medical images refers to the unequal distribution of pixels or voxels belonging to different classes or structures within the images where the minority class typically is the region of interest such as tumors and lesions which is the aim region of the segmentation process and the remaining areas such as normal healthy tissues and background as the majority class. This imbalance can markedly affect the learning process, potentially resulting in a model that predominantly segments normal and background pixels only Figure 7.5 Due to this issue, selecting metrics for evaluating machine learning model performance becomes challenging, because it highly influences metrics that heavily weigh the majority class or consider true negatives in evaluations, potentially yielding misleading results.

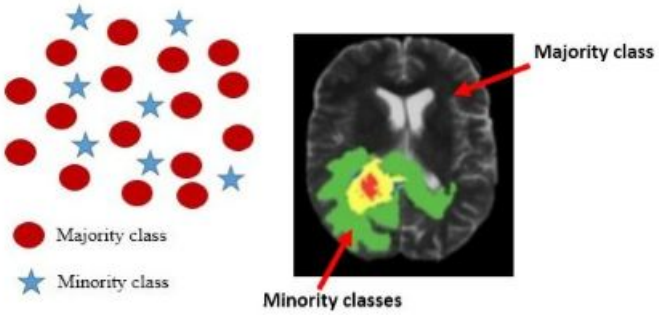


Figure 7.5: Class Imbalance Example

To handle this problem several solutions have been proposed such as data augmentation, oversampling, under-sampling, patch-based segmentation and the appropriate choice of evaluation metrics to detect the correct performance of the

model; numerous works have focused on this last point to expose the most suitable evaluation metrics for image segmentation tasks[96, 97, 98]. This section aims to exhibit the impact of the class imbalance problem on machine learning evaluation metrics in medical image segmentation in which we propose to use the accuracy metric as an additional metric behind the dice score, sensitivity and specificity to evaluate the performance of the previous models SVM, RF and PU-NET.

7.6.1 Evaluation Metric

All The used evaluation metrics except accuracy, are widely used in the context of image segmentation and highly recommended to address the class imbalance problem because of their focus on the infrequent classes or True positives; however, traditional metrics like accuracy are still used and can be misleading in imbalanced datasets. In this case a model might achieve high accuracy by simply predicting the majority class or the healthy tissue and background pixel and this, which is proved through this experiment. The Accuracy metric is defined as follow:

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{7.1}$$

7.6.2 Results

Tables 7.4 and 7.5 shows the previous models results in addition to the Accuracy metrics

Table 7.4: PU-NET and Machine Learning techniques training comparison results with accuracy metric

Models	SVM			RF			PU-NET		
Classes	WT	TC	ET	WT	TC	ET	WT	TC	ET
DICE	0.655	0.472	0.372	0.756	0.557	0.372	0.933	0.846	0.852
ACC	0.980	0.970	0.990	0.980	0.980	0.970	0.990	0.990	0.990
Sen	0.716	0.783	0.792	0.759	0.782	0.810	0.937	0.854	0.848
Spec	0.828	0.847	0.909	0.853	0.880	0.907	0.991	0.990	0.995

Table 7.5: PU-NET and Machine Learning techniques testing comparison results with accuracy metric

Models	SVM			RF			PU-NET		
Classes	WT	TC	ET	WT	TC	ET	WT	TC	ET
DICE	0.668	0.463	0.354	0.747	0.505	0.343	0.925	0.788	0.799
ACC	0.980	0.970	0.990	0.980	0.980	0.970	0.990	0.990	0.990
Sen	0.707	0.787	0.827	0.736	0.727	0.818	0.931	0.778	0.900
Spec	0.828	0.845	0.910	0.851	0.879	0.908	0.989	0.984	0.998

7.7 Discussion

this comparative study encompasses two main experiments. The first was an SVM and RF-based proposition used to evaluate the performance of deep learning against old machine learning techniques in the medical image segmentation context (glioma segmentation) which conducted to a fact results that deep learning-based approaches are more powerful and from the experiment phase both SVM and RF are considered as tedious and time-consuming regarding the conception step and the training step in addition to the high memory consumption. Furthermore, the low quality of the segmentation is due to the naive feature selection and extraction step where this step is fully automatic in deep learning. The second experiment was the introduction of the accuracy metric to discuss the impact of class imbalance on the evaluation metrics where in general medical images are regarded as highly imblanced. Table7.4 and Table 7.5 exhibit the summary results of the models on the training and test sets respectively. A large panoply of metrics is used for the performance evaluation of the Machine Learning models. Still, there are some constraints because they are not all suited for all tasks where each task may require specific metrics; for example, in this experiment, which tackles the glioma segmentation problem, we proposed using Dice score, Accuracy, sensitivity, and specificity. From the exhibited results in Tables 7.4 and 7.5 , we can see the accuracy is always 99%; however, the Dice varies between 35% to 90%. The Dice score is considered as a similarity measure which measures the overlap between the predicted map and the ground truth, it depends on the true positive pixels only which are the objective of the segmentation or the wanted area of interest, which represents the tumour area in our case but the accuracy is the

percentage of pixels in the image that are classified correctly, it depends on the true positive and true negative pixels, in this study the true negative pixels represent the healthy tissue and the background, which is the large class in terms of pixels that dominate the image that is why accuracy is near 100% and this is due to the class imbalance, the gap is highly noted in the SVM and Random forest model, where there is a significant difference between the accuracy and the Dice coefficient and this raises concerns about its reliability and precision which justify that is not adaptable for such problem because high accuracy does not imply high model segmentation ability. As we mentioned previously dice score focuses on true positives only which is the infrequent class and penalize false positives which makes it more sensitive and suitable metrics for this kind of task and a good solution for the class imbalance problem. [8]. For the sensitivity and the specificity, as we previously mentioned, sensitivity refers to the proportion of true positives that are correctly segmented compared to the ground truth, which is, in our task, the rate of indeed predicted tumorous pixels. Inversely, specificity refers to the proportion of true negatives that mean the proportion of pixels that are predicted healthy and background exactly like the ground truth, both metrics work with one class separately so they are not affected by the class imbalance. They are adaptable for this task. They evaluate the model's capability in detecting tumorous pixels and healthy ones, respectively. The limitation of this experiment is attributed to the use of a limited number of evaluation metrics. As future directions for this work, it is recommended to consider alternative solutions, including the exploration of a broader range of evaluation metrics, as well as the implementation of techniques such as under-sampling and patch-based segmentation to effectively address this challenge.

7.8 Conclusion

This chapter provided a comparative study between traditional machine learning techniques using support vector machines(SVM) and Random Forest (RF) in brain glioma segmentation against our PU-NET deep learning approach which proved the efficiency of deep learning and justified the abundant use of this technique in

nowadays. In addition, an introduction to the class imbalance problem in image segmentation and its impact on evaluation metrics especially accuracy through an experiment that shows the evaluation of an image segmentation model using this metric can be misleading .

It always seems impossible until it's done.

8

Transfer Learning from Brain Tumors Segmentation to Brain Tumors Classification

Contents

8.1 Introduction

In the area of medical imaging and diagnosis, the task of accurately classifying brain tumours holds immense significance. A pivotal approach in this domain involves

leveraging advanced machine learning techniques, specifically transfer learning, to enhance the efficiency and effectiveness of the classification process. In this section, we propose harnessing the capabilities of our brain glioma segmentation model by applying transfer learning for brain tumor classification. We aim to categorize tumors into meningioma, glioma, and pituitary tumors using a novel dataset. This transfer learning methodology has demonstrated remarkable results, overcoming challenges in a remarkably short training period with a straightforward implementation.

8.2 Transfer Learning

Transfer Learning[99] is a technique used in machine learning and widely used in deep learning approaches, it aims to exploit the knowledge gained from a previous task to improve generalization about another. The fundamental concept involves leveraging the expertise gained by a model through extensive training on a task with abundant labeled data. Rather than starting the learning process from scratch, the approach entails applying the acquired patterns or features to different tasks that lack substantial data. By initiating the model with the knowledge gleaned from addressing a related task, it aims to enhance performance in scenarios where training data is limited. it can be applied in various domains, including computer vision, natural language processing, and speech recognition. Transfer learning typically involves two main steps:

1. Pre-trained model on Source Task:

   A model trained on a large dataset for a source task that is typically related to the target task. This pre-trained model learns general features and representations from the source task.

2. Fine-tuning on a Target Task:

   The pre-trained model is then adapted or fine-tuned on a smaller dataset for the target task. The model's parameters are adjusted to make it more specialized for the target task while retaining the knowledge gained during pre-training.

8.3 Related works

Similar to the significant interest in brain tumour segmentation, there has been a notable focus on the exploration of brain tumour classification that helps in pathology diagnosis and further explorations and decisions regarding proposed treatments. Since the advent of deep learning and convolutional neural networks (CNN), several medical image datasets have been collected and numerous approaches have been suggested and employed. In 2015 (Cheng et al) proposed a collection of MRI brain tumours of three types meningioma, gliomas, and pituitary tumours last updated in 2017. He proposed enhancing the classification performance of these brain tumours by incorporating tumor region augmentation and partitioning through the extraction of diverse features, including a bag-of-words (BOW), an intensity histogram, and a grey-level co-occurrence matrix (GLCM). The classification of these features was carried out using a support vector machine (SVM), resulting in an impressive accuracy of 91.28%[100]. Tahir et al [101] explored different pre-processing techniques to enhance classification outcomes. Their study involved three pre-processing methods: noise reduction, contrast enhancement, and edge detection. These techniques were systematically tested in various combinations on diverse test sets. The authors argue that utilizing a diverse range of preprocessing schemes yields more benefits than relying on a single preprocessing approach. Utilizing the Figshare dataset, they applied the SVM classifier, achieving an accuracy of 86%. However, deep learning-based models offer better flexibility and comfort in terms of handcrafted feature extraction by making this process fully automatic and achieving high classification accuracy which is considered its robust advantage. Abiwinanda et al. crafted a model for brain tumour classification employing convolutional neural networks (CNN)[102]. They devised seven unique CNN variations, the second variant of their model achieved the highest testing accuracy, reaching 84.19%, on the Figshare brain MRI dataset. Finally, Z.N.K.Swati et al[103] suggested using a well-known pre-trained model Vgg19 to classify tumours using transfer learning techniques and fine-tuning strategy block-wise in supposing that transfer

learning is more adaptable for small datasets, they used five-fold cross-validation and reached an average accuracy of 94.82.

8.4 Methodology

In this instance, we advocate for the implementation of multiclass brain tumor classification by building a Convolutional Neural Network(CNN)model through the utilization of transfer learning from our previously introduced RGBA U-NET in section 5.5.1, specifically designed for glioma segmentation. This strategic approach leverages the pre-trained encoder and the features extracted during the segmentation process. The construction of our classification model follows a systematic design:

1. Retaining the RGBA Encoder Model: The first step involves preserving the RGBA encoder model, which has been initially tailored for glioma segmentation.

2. Integrating a Global Averaging Pooling Layer: Subsequently, we augment the model architecture by incorporating a global averaging pooling layer. This addition serves to condense the spatial dimensions of the extracted features, facilitating a more compact representation.

3. Introducing a Dense Layer for Multiclass Classification: The next layer comprises a dense layer equipped with an appropriate number of nodes(classes), tailored to the multiclass nature of the brain tumor classification task (in this case, three classes: meningioma, glioma, and pituitary tumors). The activation function utilized is softmax, allowing for the generation of probabilities across multiple classes.

By adhering to this architectural design, we capitalize on the knowledge acquired during glioma segmentation, enabling our model to effectively classify brain tumors into distinct categories with enhanced accuracy and efficiency. Figure 8.1 exhibit the final architectural design of the CNN model.

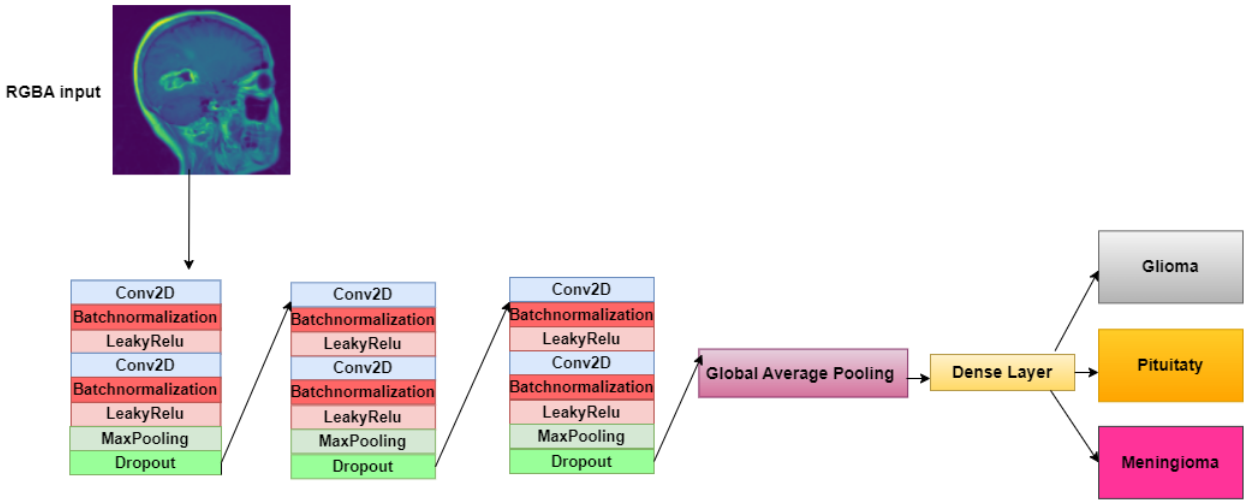


Figure 8.1: The CNN Transfer Learning Architecture

8.5 Experimentations

8.5.1 Data

In this research, our transfer Convolutional Neural Network (CNN) is trained using a dataset comprising 3064 T-1 weighted contrast-enhanced MRI (CE-MRI) images of brain tumors. The dataset, originally provided by Jun Cheng and utilized in his previous paper [100] , encompasses images of gliomas, meningiomas, and pituitary tumors. This data is organized in matlab data format. Each file stores a struct containing the following fields for an image:

- cjdata.label: 1 for meningioma, 2 for glioma, 3 for pituitary tumor cjdata.PID: patient ID
- cjdata.image: image data
- cjdata.tumorBorder: a vector storing the coordinates of discrete points on tumor border. For example, [x1, y1, x2, y2,...] in which x1, y1 are planar coordinates on tumor border. It was generated by manually delineating the tumor border. So we can use it to generate binary image of tumor mask.

- cjdata.tumorMask: a binary image with 1s indicating tumor region in our work we focus only on the image data and label to elaborate the classification task the following table exhibits the remain data information:

Table 8.1: The Data Information

Brain Tumor Dataset			
Tumor Types	Meningioma	Gliomas	Pituitary
# of Slices	708	1426	930
Modalities	T1CE		
Dimension	(512,512)		
Format	Matlab (.mat)		
Labels	0 =Meningioma 1=Gliomas 2=Pituitary		

8.5.2 Data Pre_Processing

As a normalization technique, we suggest using the min-max normalization[104]also known as feature scaling or min-max scaling, which is a technique used in data pre-processing to scale the values of a numerical feature to a specific range[0,1]. same as other normalisation techniques its purpose is to bring all features to a similar scale, preventing certain features from dominating due to their larger magnitudes. The min-max normalization process involves the following steps for each feature:

1. the minimum (min) and maximum (max) values of pixels in each image.

2. update each pixel using the following formula :

$$\mathrm{X} = \frac{x - \min}{\max - \min} \tag{8.1}$$

where X is the new pixel value and x is the original pixel value

8.5.3 Implementation Details

Our deep learning proposed approach was implemented using python3 over the Tensorflow Keras library and executed on Colab GPU node Tesla T4 of 15GB RAM configured with CUDA 12.0. the data was split randomly into Train set

80% , valid set 10% and Test set 10% , we also propose to resize images from (512,512) to (128,128) and for the reason of transfer learning, the architecture requires that the images must be in the RGBA format of (128,128,4) shape, more details are exposed in the table below:

Table 8.2: The Architecture Parameters

Parameters	values
Training Samples	2451
Validation Samples	306
Test Samples	307
Initial filters N°	16
Batch size	8
Dropout rate	0.4
Optimizer	Adam (default)
Epochs	250
Model checkpoint	Active
Early stopping	Active (patience epochs=25)

8.5.4 Evaluation Metrics

we suggest evaluate the performance of our image classification model using the performance metrics outlined in prior studies like [100, 102],they are the most frequently used metrics in such tasks. The evaluation encompassed key metrics such as sensitivity/recall(equation5.3), precision, F1-score, and accuracy(equation7.1. all of these metrics are computed based on the confusion Matrix.

- Precision : measures how often a model can correctly predicts the positive class

$$Precision = \frac{TP}{TP + FP} \tag{8.2}$$

- F1-Score: The F1 score is the harmonic mean of the precision and recall,The highest possible value maximum 1.0, indicat perfect precision and recall to ensure that a classifier makes good predictions of the relevant class (good precision) in sufficiently large numbers (good recall) on a target dataset

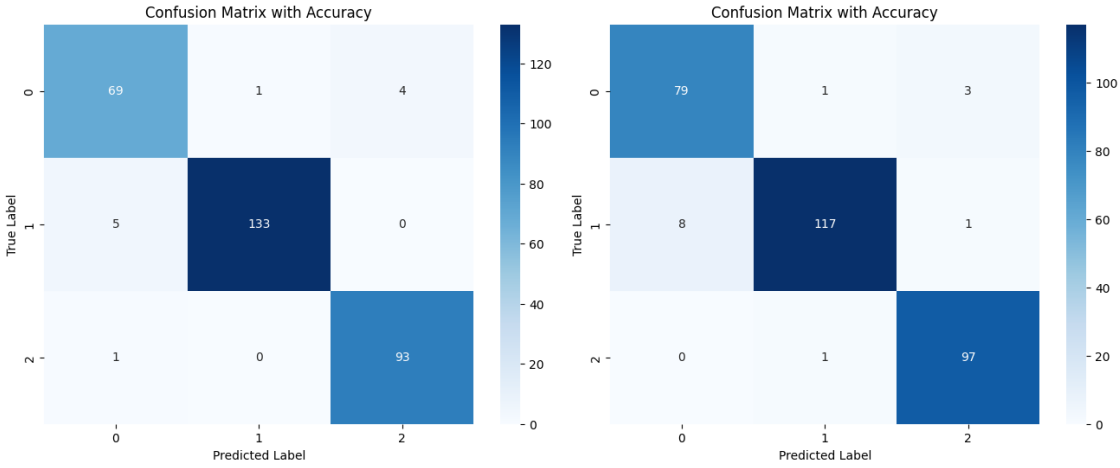
$$F1 - Score = \frac{2 * Precision * Recall}{Precision * Recall} \tag{8.3}$$

8.6 Results

The table 8.3 exhibits the results of our transfer learning model on the validation and the test set mentioned in the table using the proposed evaluation Metrics in addition to the confusion Matrices used for computing the evaluation metrics.

Table 8.3: Transfer Learning classification Validation and Test Results

Metrics	Validation			TEST		
	Tumor Classes			Tumor Classes		
	Meningioma	Glioma	Pituitary	Meningioma	Glioma	Pituitary
Accuracy	0,9324	0,9638	0,9894	0,9518	0,9286	0,9898
Precision	0,9200	0.9925	0.9588	0,9080	0.9832	0.9604
Recall	0,9324	0,9638	0.9894	0,9518	0,9286	0.9898
F1-Score	0,9262	0.9779	0,9738	0,9294	0.9551	0,9749



(a) Validation **(b)** Test

Figure 8.2: The Confusion Matrix for Classification Details

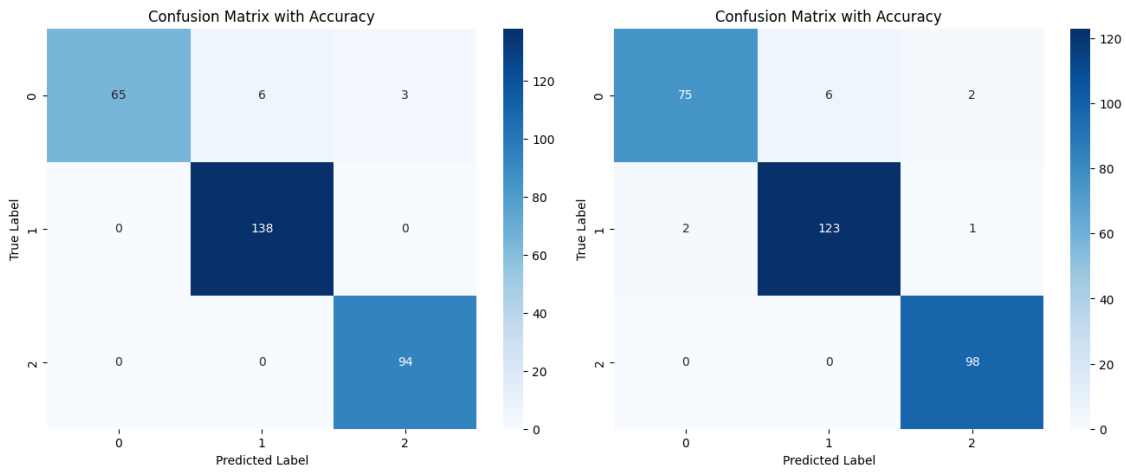
8.6.1 Data Augmentation

In this section, we propose the implementation of a data augmentation technique to enhance the diversity of the dataset. The Python script utilizes the OpenCV and NumPy libraries to apply a range of random transformations to RGBA images. These transformations include random horizontal and vertical flips, rotations within the -45 to 45-degree range, random adjustments to brightness and contrast with a factor between 0.5 and 1.5 and potential Gaussian blur with randomly selected

kernel size in the range [3,5,7]. By specifying the number of augmentations per image, this process systematically augments the dataset, generating a more varied set of images. The augmented images, along with their corresponding labels, are then stored and concatenated with the original dataset. This data augmentation process is instrumental in improving the dataset's diversity, ultimately contributing to the enhanced performance and generalization of machine learning models trained on the augmented dataset. we propose to keep the same data split in the section8.5.3 and augment the training data only where number of training samples has become after augmentation 4902 instead of 2451, results are reported in table 8.4

Table 8.4: Transfer Learning classification Validation and Test Results with data augmentation

Metrics	Validation			TEST		
	Tumor Classes			Tumor Classes		
	Meningioma	Glioma	Pituitary	Meningioma	Glioma	Pituitary
Accuracy	0.8784	1.000	1.000	0.9036	0.9762	1.000
Precision	1.000	0.9583	0.9691	0.9740	0.9535	0.9703
Recall	0.8784	1.000	1.000	0.9036	0.9762	1.000
F1-Score	0,9353	0.9787	0,9843	0.9375	0.9647	0.9849



(a) Validation **(b)** Test

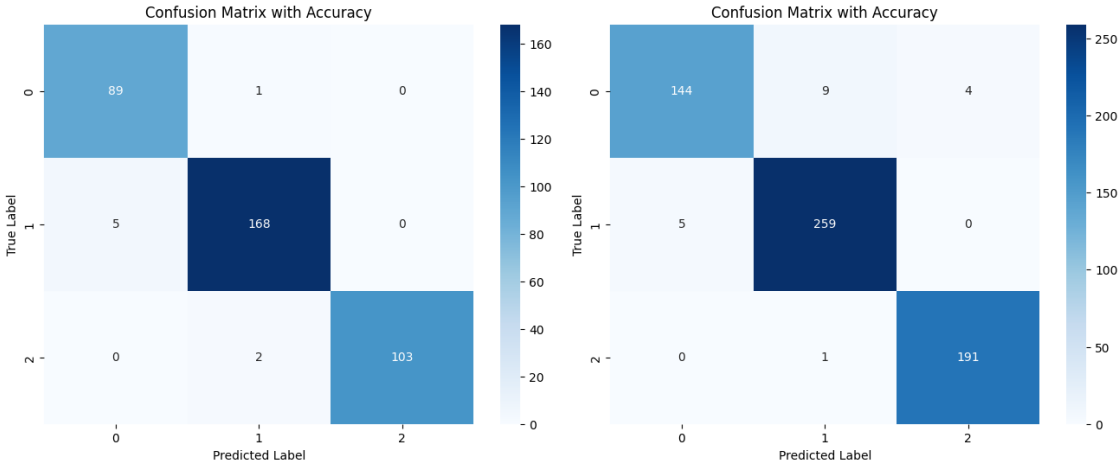
Figure 8.3: The Confusion Matrix for Classification Details with augmented data

Table 8.5 expose the results of another experiment, in which we propose to extend the test set and segmenting the original dataset into 20% as a test set and

15% validation set rather than(10%,10% previously), and apply the same data augmentation process to the remaining training data.

Table 8.5: Transfer Learning classification Validation and Test Results with data augmentation

Metrics	Validation			TEST		
	Tumor Classes			Tumor Classes		
	Meningioma	Glioma	Pituitary	Meningioma	Glioma	Pituitary
Accuracy	0.9889	0,9711	0,9810	0,9172	0,9811	0,9948
Precision	0,9468	0.9825	1.0000	0,9664	0.9628	0.9795
Recall	0.9889	0,9711	0,9810	0,9172	0,9811	0,9948
F1-Score	0,9674	0.9767	0,9904	0,9412	0.9719	0,9871



(a) Validation **(b)** Test

Figure 8.4: The Confusion Matrix for Classification Details with extended test set and augmented data

8.6.2 Comparaison Against Existing Works

Table8.6 exposes the results achived by the state of the art and our achived results in the three previous experiments.

Table 8.6: Comparaison between our model and existing methods

Authors	Cheng et al	Tahir et al	Abiwinanda et al	Z.N.K.Swati et al	Ours
Avg Accuracy	91.28	86	84.19	94.82	95 % 96 % 97 %

8.7 Discussion

Transfer learning is an optimization strategy that seeks to capitalize on the knowledge acquired by models trained on extensive datasets. Its goal is to efficiently train models on smaller datasets, achieving high accuracy with reduced effort and time. This approach is particularly advantageous when there's a resemblance between tasks. In this practical application, the primary model undergoes training for brain glioma segmentation. Subsequently, this acquired knowledge is exploited to enhance the performance of a related task specifically, the classification of brain tumors. The three experiments demonstrate highly significant results , confirming the efficacy of transfer learning in terms of reduced training time, simplified implementation, and enhanced generalization. In terms of classification quality, each of the three experiments yielded remarkably high results, boasting accuracy averages of 95%, 96%, and 97%, respectively. Notably, data augmentation played a pivotal role, particularly evident in the third experiment, where the test dataset size expanded to 613 from 307 in the preceding experiments. This expansion highlighted the model's robustness, emphasizing its superior ability to generalize, in term of precision and recall all experiments reached high scores which prove the effectiveness of the model in correctly identifiying instances of diseases which is a key indecator of its power and reliability. when comparing our model results against some existing methods, our model outperforms them without any complex model and time-consuming training process.

8.8 Conclusion

In conclusion, the presented chapter is a critical study in the domain of brain tumor classification using transfer learning and deep neural network architecture. We applied transfer learning using our RGBA model constructed for brain glioma segmentation as a source task and classifying the brain tumor type from glioma, meningioma, and pituitary using the Figshare dataset as a target task. we suggested making different experiments with and without data augmentation and the results were very significant with less time complexity and straightforward implementation.

However difficult life may seem, there is always something you can do and succeed at.

9 Conclusions

9.1 Results

This thesis introduces a contribution to medical image analysis through the use of deep learning. One of the most well-known image analysis techniques is image segmentation which is the process of dividing an image into multiple regions to identify and locate its objects in medical images, it mainly aims to locate pathology structures to facilitate diagnosing and patient assisting. unfortunately, all proposed works are still not enough to reach the best results due to the challenges of such sensitive images as image low quality and noise, etc...

This thesis focuses on the segmentation of one of the most life-threatening types of primary brain cancers called glioma from multimodal MRI images using deep learning.

Our first motivation was to propose a straightforward strategy to handle the multimodal MRI input images by using multiple pathways, one path for each modality, and each pathway is a kind of deep learning architecture called U-NET with result concatenation before the generation of the prediction probability map. This model achieves competitive results but is a little bit far from the top-ranked approaches which motivates us toward a second approach.

Our second approach was to propose an optimized and efficient architecture by

suggesting a new way of handling multimodal MRI images where we incorporated an input processing block into an updated U-NET. it has provided significant results that outperform the majority of existing methods.
The third motivation was to establish a comparative study between old machine-learning techniques and our proposed architecture which proved the real efficiency of deep learning.
finally, we had the motivation to perform a brain tumor classification task based on knowledge transfer from our built segmentation model using the transfer learning technique which succeeded in reaching high accuracy in a very simple implementation and record training time.

9.2 Future Works

Deep learning requires High running resources availability and large data sets to guarantee better results which we hope to get one day, after discovering the power of deep learning in the future we could investigate more tasks like :

- Using deep learning for other pathology structure segmentation like breast cancer lung cancer Etc
- Keep improving our models which prove that better results could be reached in a very simple way.
- Explore More domains using deep learning rather than the medical one.

Résumé

Depuis l'avènement de l'apprentissage profond, une révolution technologique s'est opérée, dotant les ordinateurs de capacités remarquables en matière d'apprentissage, de raisonnement, de réflexion et de prise de décision. L'apprentissage profond a joué un rôle central dans la facilitation de l'extraction de caractéristiques, en particulier dans le domaine de la vision par ordinateur, où il contribue significativement à l'apprentissage et à la reconnaissance du contenu des images et des clips vidéo. Cet impact s'étend au domaine de la recherche en imagerie médicale, comme le démontre cette thèse. L'accent est mis sur le rôle de l'apprentissage profond dans la segmentation automatique précise d'images médicales, notamment dans la segmentation des tumeurs cérébrales de gliome et de leurs tissus sous-tumoraux associés à partir d'images par résonance magnétique multimodales. La thèse présente diverses contributions visant à obtenir une segmentation précise, comprenant une étude comparative entre les méthodes basées sur l'apprentissage machine et les architectures d'apprentissage profond. De plus, elle met en lumière l'influence substantielle du déséquilibre de classes sur les métriques d'évaluation de la segmentation d'images. Enfin, la thèse propose une méthode basée sur le transfert d'apprentissage pour construire un modèle capable de classifier trois types de tumeurs cérébrales : méningiome, gliome et tumeurs hypophysaires.

Mot-clés : Apprentissage profond, Images médicales, Tumeur cérébrale, Segmentation d'images, Image Classification, Déséquilibre de classe

ملخص

منذ ظهور التعلم العميق، حدثت ثورة في ميدان التكنولوجيا، حيث اكتسب الحاسوب قدرات استثنائية في التعلم والاستدلال والتفكير واتخاذ القرارات. لعب التعلم العميق دورًا حاسمًا في تسهيل استخراج الخصائص، خاصة في مجال الرؤية الحاسوبية، حيث ساهم بشكل كبير في تعلم والتعرف على محتوى الصور ومقاطع الفيديو. يمتد هذا التأثير إلى ميدان البحث في الصور الطبية، كما يظهر ذلك في هذه الرسالة. التركيز هو على دور التعلم العميق في تقسيم الصور الطبية تلقائيًا بدقة، وتحديداً في تقسيم أورام الدماغ من نوع الأورام الدبقية وأنسجتها الفرعية المرتبطة بها من صور الرنين المغناطيسي المتعدد الوسائط. تقدم الرسالة مساهمات متنوعة تهدف إلى تحقيق تقسيم دقيق، بما في ذلك دراسة مقارنة بين طرق التعلم الآلي وهندسات التعلم العميق. بالإضافة إلى ذلك، تسلط الضوء على تأثير التفاوت الكبير في الفئات على مقاييس تقييم تقسيم الصور. وأخيرًا، تقترح الرسالة طريقة مبنية على التعلم الانتقالي لإنشاء نموذج قادر على تصنيف ثلاثة أنواع من أورام الدماغ: الورم السحائي، والأورام الدبقية، وورم الغدة النخامية.

الكلمات المفتاحية : التعلم العميق، الصور الطبية، ورم الدماغ، تجزئة الصورة،تصنيف الصورة، عدم التوازن الطبقي

Bibliography

[1] Ahmed Fawzy Gad. *Practical Computer Vision Applications Using Deep Learning with CNNs: With Detailed Examples in Python Using TensorFlow and Kivy*. Apress, 2018.

[2] Holger R. Roth, Chen Shen, Hirohisa Oda, Masahiro Oda, Yuichiro Hayashi, Kazunari Misawa, and Kensaku Mori. Deep learning and its application to medical image segmentation. *ArXiv*, abs/1803.08691, 2018.

[3] Harris A. Ahmad, Hui Jing Yu, and Colin G. Miller. *Medical Imaging Modalities*, page 3–26. Springer London, December 2013.

[4] Donald W. McRobbie, Elizabeth A. Moore, Martin J. Graves, and Martin R. Prince. *MRI from Picture to Proton*. Cambridge University Press, sep 2006.

[5] Magnetic resonance imaging (mri) of the brain and spine: Basics. `https://case.edu/med/neurology/NR/MRI%20Basics.htm`, Accessed september,2023.

[6] Brain basics: Know your brain. `https://www.ninds.nih.gov/health-information/public-education/brain-basics/brain-basics-know-your-brain`, Accessed september,2023.

[7] Brain anatomy and how the brain works. `https://www.hopkinsmedicine.org/health/conditions-and-diseases/anatomy-of-the-brain`, Accessed september,2023.

[8] Armand S. Schachter and Kenneth L. Davis. Alzheimer's disease. *Dialogues in Clinical Neuroscience*, 2(2):91–100, 2000. PMID: 22034442.

[9] Kelley M. Swanberg, Abhinav V. Kurada, Hetty Prinsen, and Christoph Juchem. Multiple sclerosis diagnosis and phenotype identification by multivariate classification of in vivo frontal cortex metabolite profiles. *Scientific Reports*, 12(1), August 2022.

[10] Kristin Huntoon and J. Bradley Elder. *High-Grade Gliomas*. Oxford University Press, October 2018.

[11] Nancy Ann Oberheim Bush and Susan Chang. Treatment strategies for low-grade glioma in adults. *Journal of Oncology Practice*, 12(12):1235–1241, 2016.

[12] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, and Et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE T Med Imaging*, 34(10):1993–2024, oct 2015.

[13] Jahangir Moini and Pirouz Piran. Chapter 1 - histophysiology. In *Functional and Clinical Neuroanatomy*, pages 1–49. Academic Press, 2020.

[14] Lawrence D. Recht and Mark Bernstein. Low-Grade Gliomas. *Neurologic Clinics*, 13(4):847–859, nov 1995.

[15] Kamal Thapar, Kalman Kovacs, Bernd Scheithauer, and Ricardo V. Lloyd. *Diagnosis and Management of Pituitary Tumors*. Humana Press, November 2000.

[16] Robin A Buerki, Craig M Horbinski, Timothy Kruser, Peleg M Horowitz, Charles David James, and Rimas V Lukas. An overview of meningiomas. *Future Oncology*, 14(21):2161–2177, September 2018.

[17] Junfeng Zhang, Heng Liu, Haipeng Tong, Sumei Wang, Yizeng Yang, Gang Liu, and Weiguo Zhang. Clinical applications of contrast-enhanced perfusion mri techniques in gliomas: Recent advances and current challenges. *Contrast Media and amp; Molecular Imaging*, 2017:1–27, 2017.

[18] Kedar R. Mahajan and Daniel Ontaneda. The role of advanced magnetic resonance imaging techniques in multiple sclerosis clinical trials. *Neurotherapeutics*, 14(4):905–923, October 2017.

[19] Pope WB Brandal G. Conventional and advanced magnetic resonance imaging in patients with high-grade glioma. *Q J Nucl Med Mol Imaging*, 62(3):239, 2018.

[20] Neeraj Sharma, AmitK Ray, KK Shukla, Shiru Sharma, Satyajit Pradhan, Arvind Srivastva, and LalitM Aggarwal. Automated medical image segmentation techniques. *J. Med. Phys*, 35(1):3, 2010.

[21] T. Manochandar and P. Kumaraguru Diderot. Deep learning-based magnetic resonance image segmentation and classification for alzheimer's disease diagnosis. *International Journal of Image and Graphics*, August 2023.

[22] D.J. Withey and Z.J. Koles. Medical Image Segmentation: Methods and Software. In *Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging*, volume 112, pages 140–143. IEEE, oct 2007.

[23] Ali Işin, Cem Direkoğlu, and Melike Şah. Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods. *Procedia Comput. Sci.*, 102(August):317–324, 2016.

[24] Steven L. Horowitz and Theodosios Pavlidis. Picture segmentation by a tree traversal algorithm. *J. ACM*, 23(2):368–388, apr 1976.

[25] Song Yuheng and Yan Hao. Image segmentation algorithms overview. *ArXiv*, abs/1707.02051, 2017.

[26] Jiss Kuruvilla, Dhanya Sukumaran, Anjali Sankar, and Siji P Joy. A review on image processing and image segmentation. In *2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)*, pages 198–203, 2016.

[27] Klaus D. Toennies. *Guide to Medical Image Analysis: Methods and Algorithms*. Springer London, 2012.

[28] Dzung L Pham, Chenyang Xu, and Jerry L Prince. Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2(1):315–337, 2000.

[29] Dinesh D Patil and Sonal G Deore. Medical image segmentation: a review. *International Journal of Computer Science and Mobile Computing*, 2(1):22–27, 2013.

[30] M. Mary Synthuja Jain Preetha, L. Padma Suresh, and M. John Bosco. Image segmentation using seeded region growing. In *2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET)*, pages 576–583, 2012.

[31] Yamina Azzi, Abdelouahab Moussaoui, and Mohand-Tahar Kechadi. Semantic segmentation of medical images with deep learning: Overview. *Medical Technologies Journal*, 4(3):568–575, Jul. 2023.

[32] K M Nimeesha and Rajaram M Gowda. Brain Tumour Segmentation Using K-Means And Fuzzy C-Means Clustering Algorithm. (March 2013):1–6, 2015.

[33] Abhishek Bal, Minakshi Banerjee, Punit Sharma, and Mausumi Maitra. Brain Tumor Segmentation on MR Image Using K-Means and Fuzzy-Possibilistic Clustering. In *2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, number May, pages 1–8. IEEE, may 2018.

[34] Raouia Ayachi and Nahla Ben Amor. Brain tumor segmentation using support vector machines. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 5590 LNAI, pages 736–747, 2009.

[35] Tian Chi Zhang, Jing Yang, and Jian Pei Zhang. SVM Methods in Image Segmentation. (c):62–65, 2016.

[36] J Kowski Rajan, D Ferlin Deva Shahil, M Tamil Elakkiya, and Rajesh Prasad. Image Segmentation Using SVM Pixel Classification. 3(3):45–51, 2017.

[37] Vijay Wasule and Sonar Poonam. Classification of Brain MRI Using SVM and KNN Classifier. pages 218–223, 2017.

[38] Zoltán Kapás, László Lefkovits, and László Szilágyi. Automatic Detection and Segmentation of Brain Tumor Using Random Forest Approach. In Vicenç Torra, Yasuo Narukawa, Guillermo Navarro-Arribas, and Cristina Yañez, editors, *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, volume 9880 of *Lecture Notes in Computer Science*, pages 301–312. Springer International Publishing, Cham, 2016.

[39] Ashnil Kumar, Lei Bi, Jinman Kim, and David Dagan Feng. Machine learning in medical imaging. In *Biomedical Information Technology*, pages 167–196. Elsevier Inc., 2020.

[40] Brahim Ait Skourt, Abdelhamid El Hassani, and Aicha Majda. Lung CT image segmentation using deep neural networks. *Procedia Computer Science*, 127:109–113, 2018.

[41] Shuyi Li, Min Dong, Guangming Du, and Xiaomin Mu. Attention Dense-U-Net for Automatic Breast Mass Segmentation in Digital Mammogram. *IEEE Access*, 7:59037–59047, 2019.

[42] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis, 2017.

[43] Intisar Rizwan I Haque and Jeremiah Neubert. Deep learning approaches to biomedical image segmentation. *Informatics in Medicine Unlocked*, 18:100297, 2020.

[44] Li Wen, Fucang Jia, and Qingmao Hu. Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks. *Journal of Computer and Communications*, 03(11):146–151, 2015.

[45] Tirivangani Magadza and Serestina Viriri. Deep Learning for Brain Tumor Segmentation : A Survey of. 2021.

[46] Phil Kim. *Matlab deep learning*, volume 130. Springer, 2017.

[47] Shiv Ram Dubey, Satish Kumar Singh, and Bidyut Baran Chaudhuri. Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, 2022.

[48] Qi Wang, Yue Ma, Kun Zhao, and Yingjie Tian. A comprehensive survey of loss functions in machine learning. *Annals of Data Science*, pages 1–26, 2020.

[49] June-Goo Lee, Sanghoon Jun, Younghoon Cho, Hyunna Lee, Guk Bae Kim, Joon Beom Seo, and Namkug Kim. Deep learning in medical imaging: General overview. *Korean Journal of Radiology*, 18:570 – 584, 2017.

[50] Geert J. S. Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen 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.

[51] S. Hussain, Syed Muhammad Anwar, and Muhammad Majid. Segmentation of glioma tumors in brain using deep convolutional neural network. *ArXiv*, abs/1708.00377, 2017.

[52] Baris Kayalibay, Grady Jensen, and Patrick van der Smagt. Cnn-based segmentation of medical imaging data. *ArXiv*, abs/1701.03056, 2017.

[53] Xiaomei Zhao, Yihong Wu, Guidong Song, Zhenye Li, Yazhuo Zhang, and Yong Fan. A deep learning model integrating fcnns and crfs for brain tumor segmentation. *Medical Image Analysis*, 43:98–111, 2017.

[54] 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, December 2017.

[55] Nagaraj Yamanakkanavar, Jae Young Choi, and Bumshik Lee. Sm-segnet: A lightweight squeeze m-segnet for tissue segmentation in brain mri scans. *Sensors*, 22(14):5148, July 2022.

[56] 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, April 2018.

[57] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation, 2017.

[58] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. *Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation*, page 833–851. Springer International Publishing, 2018.

[59] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *IEEE Access*, 9:16591–16603, may 2015.

[60] Mina Ghaffari, Arcot Sowmya, and Ruth Oliver. Automated brain tumor segmentation using multimodal brain scans: A survey based on models submitted to the brats 2012–2018 challenges. *IEEE Rev. Biomed. Eng.*, 13:156–168, 2020.

[61] Konstantinos Kamnitsas, Wenjia Bai, Enzo Ferrante, Steven McDonagh, Matthew Sinclair, Nick Pawlowski, Martin Rajchl, Matthew Lee, Bernhard Kainz, Daniel Rueckert, and Ben Glocker. Ensembles of Multiple Models and Architectures for Robust Brain Tumour Segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2017,Lect Notes Comput Sc*. Springer International Publishing, nov 2018.

[62] Guotai Wang, Wenqi Li, Sebastien Ourselin, and Tom Vercauteren. Automatic Brain Tumor Segmentation using Cascaded Anisotropic Convolutional Neural Networks. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2017. Lect Notes Comput Sc*, pages 1–12, sep 2017.

[63] Andriy Myronenko. 3D MRI Brain Tumor Segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018.Lect Notes Comput Sc*, pages 311–320. Springer, Cham, 2019.

[64] Fabian Isensee, Philipp Kickingereder, Wolfgang Wick, Martin Bendszus, and Klaus H. Maier-Hein. No new-net. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc*, volume 11384, pages 234–244, 2019.

[65] Richard McKinley, Raphael Meier, and Roland Wiest. Ensembles of Densely-Connected CNNs with Label-Uncertainty for Brain Tumor Segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018,Lect Notes Comput Sc*, volume 2, pages 456–465. Springer, Cham, 2019.

[66] Guotai Wang, Wenqi Li, Sébastien Ourselin, and Tom Vercauteren. Automatic brain tumor segmentation using convolutional neural networks with test-time augmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc*, volume 11384, pages 61–72, 2019.

[67] Alberto A.Albiol, Albiol Antonio, and Albiol Francisco. Extending 2D Deep Learning Architectures to 3D Image Segmentation Problems. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries.BrainLes 2018. Lect Notes Comput Sc*, volume 1, pages 73–82. Springer International Publishing, 2019.

[68] Rui Hua, Quan Huo, Yaozong Gao, Yu Sun, and Feng Shi. Multimodal Brain Tumor Segmentation Using Cascaded V-Nets. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018.Lect Notes Comput Sc*, volume 11384, pages 49–60. Springer, Cham, 2019.

[69] Adel Kermi, Issam Mahmoudi, and Mohamed Tarek Khadir. Deep Convolutional Neural Networks Using U-Net for Automatic Brain Tumor Segmentation in Multimodal MRI Volumes. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc*, volume 2, pages 37–48. Springer International Publishing, 2019.

[70] Michal Marcinkiewicz, Jakub Nalepa, Pablo Ribalta Lorenzo, Wojciech Dudzik, and Grzegorz Mrukwa. Segmenting Brain Tumors from MRI Using Cascaded Multi-modal U-Nets. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc*, volume 2, pages 13–24. Springer, Cham, 2019.

[71] 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. BrainLes 2019. Lect Notes Comput Sc)*, volume 11992, pages 231–241. Springer International Publishing, 2020.

[72] Richard McKinley, Michael Rebsamen, Raphael Meier, and Roland Wiest. Triplanar ensemble of 3d-to-2d cnns with label-uncertainty for brain tumor segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and*

Traumatic Brain Injuries. BrainLes 2019, Lect Notes Comput Sc), volume 11992, pages 379–387. Springer International Publishing, 2020.

[73] Yuan-Xing Zhao, Yan-Ming Zhang, and Cheng-Lin Liu. Bag of Tricks for 3D MRI Brain Tumor Segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019, Lect Notes Comput Sc*, volume 11992, pages 210–220. Springer International Publishing, 2020.

[74] Alessandro Crimi, Spyridon Bakas, Hugo Kuijf, Bjoern Menze, and Mauricio Reyes, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, volume 10670 of *Lect Notes Comput Sc*. Springer International Publishing, Cham, 2018.

[75] Alessandro Crimi, Spyridon Bakas, Hugo Kuijf, Farahani Keyvan, Mauricio Reyes, and Theo van Walsum, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, volume 11384 of *Lect Notes Comput Sc*. Springer International Publishing, Cham, 2019.

[76] Agus Subhan Akbar, Chastine Fatichah, and Nanik Suciati. Single level unet3d with multipath residual attention block for brain tumor segmentation. *Journal of King Saud University-Computer and Information Sciences*, 34(6):3247–3258, 2022.

[77] Mehrdad Noori, Ali Bahri, and Karim Mohammadi. Attention-guided version of 2d unet for automatic brain tumor segmentation. In *2019 9th international conference on computer and knowledge engineering (ICCKE)*, pages 269–275. IEEE, 2019.

[78] Jianxin Zhang, Zongkang Jiang, Jing Dong, Yaqing Hou, and Bin Liu. Attention gate resu-net for automatic mri brain tumor segmentation. *IEEE Access*, 8:58533–58545, 2020.

[79] Chandan Ganesh Bangalore Yogananda, Bhavya R Shah, Maryam Vejdani-Jahromi, Sahil S Nalawade, Gowtham K Murugesan, Frank F Yu, Marco C

Pinho, Benjamin C Wagner, Kyrre E Emblem, Bjørnerud, Madhuranthakam AJ Fei B, and Maldjian JA. A fully automated deep learning network for brain tumor segmentation. *Tomography*, 6(2):186–193, 2020.

[80] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S. Kirby, and Davatzikos C Freymann JB, Farahani K. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Scientific Data*, 4(1):170117, dec 2017.

[81] 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*, nov 2018.

[82] Kuntal Kumar Pal and K. S. Sudeep. Preprocessing for image classification by convolutional neural networks. In *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, pages 1778–1781. IEEE, may 2016.

[83] Jacob C. Reinhold, Blake E. Dewey, Aaron Carass, and Jerry L Prince. Evaluating the impact of intensity normalization on mr image synthesis. *Proc SPIE Int Soc Opt Eng*, 10949, 2019.

[84] Leon Weninger, Oliver Rippel, Simon Koppers, and Dorit Merhof. Segmentation of Brain Tumors and Patient Survival Prediction: Methods for the BraTS 2018 Challenge. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc*, volume 11384, pages 3–12, Cham, 2019. Springer International Publishing.

[85] Abdel Aziz Taha and Allan Hanbury. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *Bmc Med Imaging*, 15(1):29, dec 2015.

[86] Harriet Small and Jonathan Ventura. Handling Unbalanced Data in Deep Image Segmentation. 2017.

[87] Shruti Jadon. A survey of loss functions for semantic segmentation. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2020.

[88] D.P. Huttenlocher, G.A. Klanderman, and W.J. Rucklidge. Comparing images using the Hausdorff distance. *IEEE T Pattern Anal*, 15(9):850–863, 1993.

[89] Yamina Azzi, Abdelouahab Moussaoui, and Mohand-Tahar Kechadi. Pu-net deep learning architecture for gliomas brain tumor segmentation in magnetic resonance images. *Image Analysis and Stereology*, 42(3):197–206, Jan. 2024.

[90] Lu Lu, Yeonjong Shin, Yanhui Su, and George Em Karniadakis. Dying relu and initialization: Theory and numerical examples. *ArXiv*, abs/1903.06733, 2019.

[91] Stamatis Mastromichalakis. Alrelu: A different approach on leaky relu activation function to improve neural networks performance. *ArXiv*, abs/2012.07564, 2020.

[92] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20:273–297, 1995.

[93] Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

[94] Y Azzi, A Moussaoui, and MT Kechadi. Class imbalance and evaluation metrics for medical image segmentation with machine learning models.

[95] Justin M. Johnson and Taghi M. Khoshgoftaar. Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1), March 2019.

[96] Hattie Thompson Small and Brown. Handling unbalanced data in deep image segmentation. 2017.

[97] Abdel Aziz Taha and Allan Hanbury. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*, 15, 2015.

[98] Dominik Müller, Iñaki Soto-Rey, and Frank Kramer. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes*, 15, 2022.

[99] Padmavathi Kora, Chui Ping Ooi, Oliver Faust, U Raghavendra, Anjan Gudigar, Wai Yee Chan, K Meenakshi, K Swaraja, Pawel Plawiak, and U Rajendra Acharya. Transfer learning techniques for medical image analysis: A review. *Biocybernetics and Biomedical Engineering*, 42(1):79–107, 2022.

[100] Jun Cheng, Wei Huang, Shuangliang Cao, Ru Yang, Wei Yang, Zhaoqiang Yun, Zhijian Wang, and Qianjin Feng. Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PLoS ONE*, 10, 2015.

[101] Bilal Tahir, Sajid Iqbal, M Usman Ghani Khan, Tanzila Saba, Zahid Mehmood, Adeel Anjum, and Toqeer Mahmood. Feature enhancement framework for brain tumor segmentation and classification. *Microscopy research and technique*, 82(6):803–811, 2019.

[102] Nyoman Abiwinanda, Muhammad Hanif, S. Tafwida Hesaputra, Astri Handayani, and Tati Rajab Mengko. Brain tumor classification using convolutional neural network. *IFMBE Proceedings*, 2018.

[103] Zar Nawab Khan Swati, Qinghua Zhao, Muhammad Kabir, Farman Ali, Zakir Ali, Saeed Ahmed, and Jianfeng Lu. Brain tumor classification for mr images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, 75:34–46, 2019.

[104] SGOPAL Patro and Kishore Kumar Sahu. Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*, 2015.