BRAIN TUMOR MRI IMAGES DETECTION AND CLASSIFICATION ALGORITHM BASED ON DEEP CONVOLUTION NEURAL NETWORK TECHNIQUES

BY

SABAA AHMED YAHYA AL-GALAL

A thesis submitted in fulfillment of the requirement for the degree of Doctor of Philosophy (Computer Science).

Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

MARCH 2023

ABSTRACT

The substantial progress of medical imaging technology in the last decade makes it challenging for medical experts and radiologists to analyse and classify them. Medical images contain massive information that can be used for diagnosis, surgical planning, training, and research. The ability to estimate conclusions without direct human input in healthcare systems using computer algorithms is known as Artificial intelligence (AI) in healthcare. Deep Learning (DL) approaches are already being employed or exploited for healthcare purposes. There is, therefore, a need for a technique that can automatically analyze and classify the images based on their respective contents. DL algorithms open a world of opportunities, and it has been recently used for medical images analysis. Although DL techniques have demonstrated a breakthrough in medical images analysis, research still ongoing to improve the accuracy rate. This research focuses on DL in the context of analysing Magnetic Resonance Imaging (MRI) brain medical images. A comprehensive review of the state-of-the-art processing of brain medical images using DL is conducted in this research. The scope of this research is restricted to three digital databases: (1) the Science Direct database, (2) the IEEEXplore Library of Engineering and Technology Technical Literature, and (3) Scopus database. More than 400 publications were evaluated and discussed in this research. The research focus on both binary classification and multi-class classification. For binary classification, the dataset used is from the brain tumor classification project which contains tumorous and non-tumorous images, and it is available for research and development. For multi-class classification, the dataset contains T1-weighted contrast-enhanced MRI medical images from 233 patients with three types of tumours: meningioma, glioma, and pituitary which is also available for research and development. The proposed neural model is fully automatic brain tumour MRI medical images classification model that uses Convolutional Neural Network (BTMIC-CNN). The model's excellent performance was confirmed using the evaluation metrics and reported a total accuracy of 99%. It outperforms existing methods in terms of classification accuracy and is expected to help radiologists and doctors accurately classify brain tumours' images. This study contributes to goal 3 of the Sustainable Development Goals (SDGs), which involves excellent health and well-being.

ملخص البحث

التقدم الكبير في تكنولوجيا التصوير الطبي في العقد الماضي يجعل من الصعب على الخبراء الطبيين وأخصائي الأشعة تحليلها وتصنيفها. تحتوي الصور الطبية على معلومات ضخمة يمكن استخدامها في التشخيص، والتخطيط الجراحي، والتدريب، والبحث. تُعرف القدرة على تقدير الاستنتاجات دون المدخلات البشرية المباشرة في أنظمة الرعاية الصحية باستخدام خوارزميات الحاسوب باسم الذكاء الاصطناعي (AI) في الرعاية الصحية. يتم بالفعل استخدام مناهج التعلم العميق (DL) لأغراض الرعاية الصحية. لذلك، هناك حاجة إلى تقنية يمكنها تحليل الصور وتصنيفها تلقائيًا بناءً على محتويات كل منها. تفتح خوارزميات DL عالمًا من الفرص، وقد تم استخدامها مؤخرًا لتحليل الصور الطبية. على الرغم من أن تقنيات DL أظهرت تقدمًا كبيرًا في تحليل الصور الطبية، إلا أن الأبحاث لا تزال جارية وهناك العديد من الطرق لتحسين معدل الدقة. يركز هذا البحث على خوارزميات DL في سياق تحليل الصور الطبية للدماغ بالتصوير بالرنين المغناطيسي (MRI) يتم إجراء نظرة عامة شاملة على أحدث معالجة للصور الطبية للدماغ باستخدام الشبكات العصبية العميقة في هذا البحث. يقتصر نطاق هذا البحث على ثلاث قواعد بيانات رقمية: (1) قاعدة بيانات Science Direct، (2) مكتبة IEEEXplore للأدب الفني الهندسي والتكنولوجيا، و (3) قاعدة بيانات Scopus. تم تقييم ومناقشة أكثر من 400 منشور في هذا البحث. يركز البحث على كل من التصنيف الثنائي والتصنيف متعدد الفئات. بالنسبة للتصنيف الثنائي استخدمنا مجموعة البيانات من مشروع تصنيف أورام المخ والذي يحتوي على صور أورام وغير أورام، وهي متاحة للبحث والتطوير. من أجل التصنيف متعدد الفئات، تم استخدام النموذج العصبي المقترح لتصنيف الصور الطبية بالرنين المغناطيسي المعززة بالتباين الموزونة T1 المأخوذة من 233 مريضًا يعانون من ثلاثة أنواع من الأورام: الورم السحائي والورم الدبقية والغدة النخامية. النموذج العصبي المقترح هو نموذج تصنيف صور التصوير الطبي بالرنين المغناطيسي لورم الدماغ التلقائي بالكامل والذي يستخدم الشبكة العصبية التلافيفية (BTMIC-CNN). الدقة الإجمالية قدرها 99٪. حيث تتفوق على الأساليب الحالية من حيث دقة التصنيف ومن المتوقع أن يساعد أطباء الأشعة والأطباء في تصنيف صور أورام الدماغ بدقة. تساهم هذه الدراسة في الهدف 3 من أهداف التنمية المستدامة(SDGs).

APPROVAL PAGE

The thesis of Sabaa Ahmed Yahya Al-Galal has been approved by the following:

Raini Binti Hassan
Supervisor

Imad Fakhri Taha Alshaikhli
Co-supervisor

M. M. Abdulrazzaq
Co-supervisor 2

Marini Binti Othman
Co-supervisor 3

Akram M Z M Khedher
Internal Examiner

Idris El-Feghi
External Examiner

Othman Omran Khalifa
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Sabaa Ahmed Yahya Al-Galal

Signature…… Date…12/4/2023……

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

BRAIN TUMOR MRI IMAGES DETECTION AND CLASSIFICATION ALGORITHM BASED ON DEEP CONVOLUTION NEURAL NETWORK TECHNIQUES

I declare that the copyright holder of this thesis/dissertation are jointly owned by the student and IIUM.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below.

1. Any material contained in or derived from this unpublished research may only be used by others in their writing with due acknowledgement.

2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purpose.

3. The IIUM library will have the right to make, store in a retrieval system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Sabaa Ahmed Yahya Al-Galal

…… 12/4/2023

Signature …….. Date

This thesis is dedicated to my beloved parents for laying the foundation of what I turned out to be in life.

ACKNOWLEDGEMENTS

First and foremost, I want to express my gratitude to Almighty Allah for His divine inspiration and mercy. His gifts, generosity, and kindness are constantly present in my life. He gave me inspiration, enthusiasm, and the ability to work hard. It was through His blessings that I was able to plan, imagine, and put my dreams into action. I would dedicate my accomplishment to Him. To my husband and my family whose untold sacrifices have helped me become where I am today.

I wish to express my appreciation and thanks to my supervisor Dr. Raini Hassan for her continuous support, encouragement, and leadership, and for that, I will be forever grateful. I would like to thank my co-supervisors Dr. Mohd Abdurazzaq and Dr. Marini who have provided their time, effort, and support for this project. To the members of my dissertation committee, thank you for sticking with me. I want to express my gratitude to IIUM and all the employees and members of the Department of Computer Science for their assistance.

A special thanks to Professor Dr. Imad Fakhri Taha Alshaikhli for his guidance throughout my study journey. Prof. Imad's advice, inspirations, and incentives helped me achieve my full potential. He taught me a lot of things. Prof. Imad introduced me to scientific writing, directed me through the skill of critical thinking, taught me to work hard, offered me the flexibility to explore many ideas. I salute his excellence and offer my respects to him. To be honest, I consider myself quite privileged to be working under Professor Imad's guidance. Alhamdulillah.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

AI	Artificial Intelligence
ANN	Artificial Neural Network
BRATS	Brain Tumors Segments Challenges
BTMIC	Brain Tumor MRI Medical Images Classification
CNN	Convolutional Neural Network
CNS	Central Nervous System
CT	Computerized Tomography
DCNN	Deep Convolutional Neural Network
DL	Deep Learning
ISLES	Ischemic Stroke Lesion Segmentation Challenge
ML	Machine Learning
MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
NBTS	National Brain Tumor Society
PET	Positron Emission Tomography
SDGs	Sustainable Development Goals
SPECT	Single-photon emission computed tomography
SVM	Support Vector Machine
TL	Transfer Learning
VGG	Visual Geometry Group
WHO	World Health Organization

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND

A brain tumour is the result of abnormal and uncontrolled development of brain cells. The National Brain Tumour Society reports that ~700,000 people in the United States of America (USA) suffer from brain tumours, and that number will increase by 85,000 in 2021. The world's 10th most common cause of death; ~3460 children under 15 were also diagnosed this year with a brain or central nervous system (CNS) tumour (NBTS, 2021). The early detection and classification of brain tumours is an important research domain in medical imaging, because it aids in the selection of the best treatment choice to save the patients' lives. Medical images are critical towards surgical planning, as it is a crucial information source for many diseases. It can also be used for research and training purposes.

The demand for digital medical images is steadily increasing. For example, CT imaging rates in older adults were 428 per 1000 person-years in 2016, compared to 204 per 1000 in 2000 in US health care systems and 409 per 1000 vs 161 per 1000 in Ontario; MRI rates were 139 per 1000 vs 62 per 1000 in the US and 89 per 1000 vs 13 per 1000 in Ontario; and ultrasound rates were 495 per 1000 vs 324 per 1000 in the US and 580 per 1000 vs 332 per 1000 in Ontario (Smith-Bindman et al., 2019). Therefore, an efficient and precise medical image analysis scheme is needed for operational planning, preparation of medical reports, and medical research and development.

Medical imaging is essential for the early identification and diagnosis of cancer. Cancer is an abnormal and uncontrolled division of cells in any part of the body. If these abnormal cells appear in the brain tissue, it is called a brain tumor. Brain tumors, similar to other types of cancers, can be benign (non-cancerous) or malignant (cancerous). Brain tumors can also be classified into primary and secondary tumors, where the former originates in the brain and is mostly benign, while the latter (known as a metastatic brain tumor) occurs when the abnormal cells spread from other organs to the

brain (Johnson et al., 2017). One common type of primary brain tumor is Glioma, which develops from glial cells (DeAngelis, 2001).

Tumors and strokes have been the world's second and third leading causes of death, respectively, after heart disease. According to the World Health Organization (WHO), 40,000–50,000 people are diagnosed with a brain tumor annually in India, out of which ~20% are infants (Saman & Jamjala Narayanan, 2019). One of the most common forms of cancers is brain cancer, with a ~70% mortality rate. Early detection would help increase the survival rates of brain cancer patients (Brunese, Mercaldo, Reginelli, & Santone, 2020).

There are many types of early detection applicable for medical imaging, such as Single-Photon Emission Computed Tomography (SPECT), Magnetic Resonance Spectroscopy (MRS), Positron Emission Tomography (PET), Computed Tomography (CT), and Magnetic Resonance Imaging (MRI). The Eurostat reports indicated that one person out of ten in Europe is subject to CT imagery annually, one in 13 to MRI, and one out of 200 for PET tomography (Aiello, Cavaliere, D'Albore, & Salvatore, 2019). The majority of prior research involved MRIs (Işın, Direkoğlu, & Şah, 2016; Kong et al., 2019).

The traditional technique of manually analysing medical images is time-consuming, and its interpretation imprecise (prone to human error)(Jyoti Patil & Pradeepini, 2019). Artificial intelligence (AI) is on the verge of transforming medicine in the next years. AI systems will be used frequently to diagnose diseases early, enhance prognoses, and give more effective, individualised treatment regimens, all while saving time and resources. Algorithms that can interpret medical images for doctors in the near future, provide decision assistance for practitioners without specialist expertise, and power AI-driven telemedicine services. Beyond the hospital, AI will be used to continually monitor the health of millions of patients, as well as route them to physician consultations and follow-ups on a massive scale.

Deep Learning is a type of Machine Learning that uses artificial neural networks to learn representations. It is considered a state-of-the-art algorithm for medical image analysis and has been successfully applied in many areas (J. Zhang, Xie, Wu, & Xia, 2019). Deep Learning (DL) has been demonstrated to be significantly superior to

manual image analyses in terms of medical image segmentation, classification, detection, registration, biometric measurement, and quality evaluation (S. Chen, Ding, & Liu, 2019). The current work focus on Deep Convolution Neural Network (DCNN), which has resulted in significant advances in the analysis of medical images (Wachinger, Reuter, & Klein, 2018).

Deep Neural Networks (DNNs) is widely used for brain image analysis. Many studies tended to classify brain diseases, brain tissue segmentation, and anatomy. According to the literature, image analytics are mostly done by DNNs. In Brain Tumors Segments Challenges (BRATS), the longitudinal MSLS 2015, the 2015 Ischemic Stroke Lesion Segmentation Challenge (ISLES), and the MR Brain imaging challenge (MRBrains) in 2013, all of the top teams used CNNs, focusing on the abovementioned techniques of MRI scans of the brain (Litjens et al., 2017). Brain tumor is an abnormal cell growth in the brain. The classification of brain tumours aids in predicting their potential behaviour and improving health-care systems.

Tumors are classified based on their cellular origin and actions, ranging from less active (benign) to aggressive (malignant), malignant tumor are graded from I (least malignant) to IV (most malignant) based on their rate of growth (DeAngelis, 2001). Classification of brain medical images can be binary classification, as the name refers binary classification is for two classes a normal class and an abnormal one. It is used in medical testing to detect whether or not a patient has a certain condition.

In brain MRI medical images for a tumor detection this task involve classifying the images to normal and affected by a tumor. Another task could be classifying the tumor itself to benign or malignant tumor (Kumari & Kr., 2017). Another type of classification is Multiclass classification, this task involves categorising objects into one of multiple classes. In medical imaging this task is widely used to classify different types of diseases. For instance, brain tumor has more than 120 different types. The main goal in this task is to classify one class out of three or more classes (Irmak, 2021; Johnson et al., 2017). Meningiomas are brain tumor that develops in the small walls that generally surround the brain and is mostly non-cancerous, Gliomas are one of the most common types of primary brain tumors. Most pituitary tumors are noncancerous (benign) growths (adenomas). Adenomas remain in the pituitary gland or surrounding tissues and don't spread to other parts of the body (S. Chen et al., 2019). Generally, one

of the life-threatening disorders that can directly damage human lives is brain tumors (Thillaikkarasi & Saravanan, 2019).

One of the most important challenges in computer vision and pattern recognition is image classification. Researchers have previously investigated brain tumor MRI (Magnetic resonance imaging) classification using Machine Learning algorithms. In traditional Machine Learning the classification task need prior step which is features extraction. While in Deep Learning algorithms the process can be completed without the use of handcrafted features, and both the feature extraction and classification steps are combined (Aiello et al., 2019; Brunese et al., 2020). In recent years. In the field of medical image analysis and disease detection, the development of artificial intelligence and Deep Learning-based new technologies has had a significant impact. Deep Convolutional Neural Network (DCNN) is one of the most extensively utilized image processing techniques (Işın et al., 2016; Kong et al., 2019; B. Menze et al., 2014). The use of CNN applications to classify different types of brain tumors has demonstrated a good performance and promising results (Havaei, Davy, Warde-Farley, et al., 2017).

1.2 STATEMENT OF THE PROBLEM

In this technological age, medical specialists are in a position to provide patients with more efficient health care using latest technologies. Brain is the most complicated component of the human body, with millions of cells working together (Guo et al., 2011). Brain tumor is an uncontrolled brain disorder which causes abnormal brain cell groups to develop. MRI employs a powerful magnetic field and waves to generate accurate pictures of the body's organs and cells. The Brain MRI scan is the best means for scientists to detect and follow the progress of the brain tumor (Mamta Mittal et al., 2019; Muhammed Talo, Baloglu, Yıldırım, & Rajendra Acharya, 2019). In high resolution MR images, brain tumor can be detected. But when there are a huge number of images (big data) (Aiello et al., 2019), it is difficult for experts to check and analyse every single image, so here it comes the need of an automatic classification and detection tools. Therefore, the computer-aided methods for analysis and detections of these information must be applied.

Previous studies implemented supervised Machine Learning (ML) algorithms (hand-designed features) for brain tumor classification and segmentation (B. Menze et al., 2014). These techniques use the classical ML algorithm, which first retrieves then provides the features to a classifier whose training does not influence their nature. Alternatively, they develop a task-adapted algorithm, which is a hierarchy of increasingly complicated features to learn directly from within the domain (Havaei, Davy, Warde-Farley, et al., 2017). DL network performed better relative to the classical ML algorithms. Previous studies confirmed that the most efficient way to analyse big datasets is to use DL (Mallick et al., 2019; M. Mittal et al., 2019; M. Talo, Baloglu, Yıldırım, & Rajendra Acharya, 2019). According to (B. H. Menze et al., 2015), there has been an exponential increase in the number of publications dedicated to automated brain tumor classification and segmentation. This study not only highlights the need for automatic analysis of brain tumors, but it also demonstrates that research in this field is ongoing.

The detection and classification of a brain tumor are challenging tasks. It is often not enough to have a good forecast scheme, particularly in medicine, where accountability is essential and can have severe legal implications (Lecun, Bengio, & Hinton, 2015). Another obstacle researchers may face is the lack of a large datasets (Litjens et al., 2017). However, that can be solved with the use of data pre-processing and data augmentation. Classification or detection in medical imaging often is described as a binary: ordinary vs abnormal, object versus context (Rouhi, Jafari, Kasaei, & Keshavarzian, 2015). But it is often a big simplification as both classes can be extremely heterogeneous and so it needs more sophisticated techniques. A closer look to the literature on brain tumor MRI medical images, however, reveals a number of gaps and shortcomings on the accuracy, performance and inconsistencies in data formats and lack of reliable training data, which need to be addressed. Therefore, there is a need to develop more accurate and reliable image detection and classification algorithm.

To sum up, these are the main points:

i. There are a huge number of medical images, it is difficult for experts to check and analyze every single image (Talo, Baloglu, Yıldırım, & Rajendra Acharya, 2019).

ii. The traditional technique of manual medical images analysis has limitation of time consuming and interpretation (human error) (Talo et al., 2019)

iii. Although ML and DL techniques have demonstrated a breakthrough when compared to manual analysis, but research still ongoing to improve the accuracy rate (Menze et al., 2015).

1.3 RESEARCH OBJECTIVES

The main aim of this research is to improve the classification and detection accuracy of brain tumor MRI medical images based on using Deep Learning techniques. The main concerns of the research have been summarized on four objectives as listed below:

1. To identify the current methodologies and algorithms exist to analyze brain MRI medical images.
2. To develop Deep Learning algorithms for detection of a brain tumor MRI medical images through binary classification with a better accuracy.
3. To develop Deep Learning algorithms for classification of a brain tumor MRI medical images through multiclass classification with a better accuracy.
4. To validate and evaluate the proposed algorithms with the current state-of-art algorithms.

1.4 RESEARCH QUESTIONS

The questions of this research are as follow:

1. What are the current methodologies and techniques that are used in analyzing brain MRI medical images?
2. How to develop a Deep Learning algorithm for brain abnormalities MRI medical images detection with a better accuracy?
3. How to develop a Deep Learning algorithm for brain tumor MRI medical images classification with a better accuracy?

4. How effective the proposed algorithms when compared to the other current state-of-art algorithms?

1.5 SIGNIFICANCE OF THE STUDY

This thesis will make a number of significant contributions to the field of medical images analysis. Medical imaging technologies generally have a significant impact on medical diagnosis as well as AI field. Previously, a process was carried out manually, but this technique is deemed costly and time consuming based on the big number of medical images generated daily. Therefore, there is a need to use automatic classification systems for medical image processes, in which these systems can process a great number of images with minimal effort and more accurate result. This research will contribute to develop detection and classification Deep Learning algorithms for brain tumor MRI medical images that may not only imitate experts, but it may outperform them. Another key benefits of the algorithm is that the early detection of the disease that result in saving human life and increasing the survival rate of the patients. This study contributes to goal number 3 of the Sustainable Development Goals (SDGs), which involves excellent health and well-being.

1.6 SCOPE OF THE STUDY

Recent theoretical developments have employed Deep Learning algorithms in the field of medical images analysis. Deep Learning has many architectures, in this study, new architecture using Convolution Neural Network (CNN) is going to be developed. CNN is one of the most prevalent profound Deep Learning algorithms. And among the medical images different types we have decided to apply MRI medical images. So, the analysis will mainly focus on brain tumor MRI medical images. The literature will cover the last five years.

1.7 CHAPTER SUMMARY

On this basis, we conclude that an algorithm of automatic classification of the brain tumor MRI medical images will be developed based on deep neural convolution networks. Our expectation is that results will improve on the current published state-of-art algorithms in term of accuracy and time. Ideally, by having a good structure of CNN, we expect a better accuracy and performance.

1.8 THESIS ORGANIZATION

This thesis is organized into six chapters. The research background, problem statement, principal goal and objectives, research questions, overview of research method, scope of research, research motivation and importance of research are all presented in this chapter. The following is how the rest of the thesis is structured:

Chapter two presents a systematic literature review of the related studies in brain tumor MRI medical image classification. It introduces an overview of techniques used to classify and segment brain tumor MRI medical images. Other related works were also given in order to cover the most recent relevant work that has been undertaken. The chapter presented and concluded solutions based on the analysis of the state of the art.

Chapter three explains the Design Science Research Method (DSRM) that was used to perform this study. It also explains the DSRM's phases, outputs, and evaluation methodologies, as well as how each phase was carried out for this study.

Chapter four covers the methodology and proposed model of both binary and multiclass classification of brain tumor MRI medical images .

Chapter five provides an in-depth analysis of the results obtained from the proposed approaches in terms of classification and retrieval using the correctness rate, recall, precision, and confusion matrix. The proposed approach's results are compared to those of the related works presented. A summary of this chapter is provided at the end.

In Chapter six, the conclusions of the work provided in this thesis are discussed, as well as possible research recommendations for future work in this field.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

Medical imaging techniques such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans are commonly used by oncologists to do early evaluations of brain tumours (Aiello et al., 2019; Işin, Direkoğlu, & Şah, 2016; Kong et al., 2019). These two techniques are commonly utilised to generate highly detailed images of brain structure. Applying artificial intelligence (AI) algorithms on these medical images may assist doctors to improve the accuracy of early cancer diagnosis.

A variety of Machine Learning and Deep Learning algorithms have been employed for the classification, segmentation, and recognition of brain tumours, such as Support Vector Machine (SVM), Artificial Neural Network (ANN) and Convolutional Neural Network (CNN). This research focuses on CNN, which have made substantial advancements in medical image analysis (Wachinger et al., 2018). Generally, there are two main classification directions for the brain tumor medical images. The first one is binary classification for normal and abnormal categories, the second is multi-classification for grading abnormal brain tumor images (Alqudah, Alquraan, Qasmieh, Alqudah, & Al-Sharu, 2019; Kader et al., 2021; Mengash & Hosni Mahmoud, 2021).

There have been numerous studies carried out using Machine Learning algorithms on tumor detection (Wu, Lin, & Chang, 2007; Guo et al., 2011) they used K- means clustering, histogram-clustering, and support vector machines (SVMs) respectively. Recent works shows that DL is the most effective model for analysing large image datasets (W. Liu et al., 2017). According to (Mallick et al., 2019), an image compression method using a Deep Wavelet Autoencoder (DWA) is presented, which combines autoencoder fundamental function reduction properties with wavelet transform image decomposition properties. Another research proposed a way to

automate normal and abnormal brain MR pictures by using profound transfer learning (Muhammed Talo, Baloglu, et al., 2019).

Kong et al. aim to build a single brain network as function display and to make ASD / TC classification using the Deep Neural Network (DNN) classification (Kong et al., 2019). For the segmentation of brain tumour, a Deep Learning method is suggested. The suggested technique involves the notion of Stationary Wavelet Transformation (SWT) and the new Growing Convolution Neural Network (GCNN) (Mamta Mittal et al., 2019). Collectively, most of the prior research has applied Deep Learning for analysing medical images and it shows a great improvement in term of accuracy.

Studies in Table 2.1 focus mainly on brain tumor MRI medical images classification. First and second study in Table 2.1 focus mainly on the segmentation of the brain images and did not include detection and classification problems. Third study on Table 2.1 applied Deep Wavelet Autoencoder (DWA) and did not apply Convolution Neural Network (CNN) which has resulted in a major advance in the analysis of medical images (Wachinger et al., 2018), and it may perform better than DWA. Last study mentioned in the Table 2.1 achieved 100 % accuracy for classification of normal and abnormal MRI brain medical images, but they have used a small dataset merely 613 images and they did not test it in large datasets yet and that may reduce the accuracy rate

Table 2.1 Summary of studies that apply Deep Learning algorithms in analyzing brain tumor images.

Reference	Objectives	Methods	Result	Gaps
(Havaei, Davy, Warde-Farley, et al., 2017)	To develop a fully automated (DNN) method based on brain tumor segmentation. The proposed networks are adapted to the low grade and high-grade glioblastomas illustrated in the MR images.	Algorithms: A new CNN based on Deep Neural Networks (DNNs) which is fully automatic brain tumor segmentation process. Dataset: BRATS 2013	Results from the 2013 BRATS test dataset show that this architecture improves with more than 30 times more rapid results than currently published state-of -the-art.	However, the processing time was improved, the algorithm did not show improvement in term of accuracy rate.
(Mamta Mittal et al., 2019)	To improve conventional system precision.	Algorithms: The new GCNN Stationary Waveslet Transformation (SWT)	This research was conducted in a comparative study with the SVM and CNN. The test findings showed that SWT and CNN were superior to SVM, PSNR, MSE and other performance parameters in the suggested method.	The proposed algorithm need to be tested in other datasets to prove its efficiency.
(Mallick et al., 2019)	In the present paper the Deep Wavelet Autoencoder (DWA) compression technique, which mingles the basic autoencoder reduction feature property together with the transform wavelet's image decomposition property, is proposed. The mixture of both has an enormous impact on the decrease in the size of the feature set by use of DNN for a further classification.	Algorithms: DWA-DNN classifiers. Dataset: RIDER (Reference Image Database to Evaluate Therapy Response)	Consideration was given to the suggested DWA-DNN picture classifier. Compared to other current classifiers such as auto-encoder-DNN or DNN, the performance criterion for the DWA DNN classifier is observed and that the suggested technique outlines the current techniques.	The dataset considered small for DL algorithms.
(Muhammed Talo, Baloglu, et al., 2019)	They proposed a way of classifying normal, abnormal MR brain images automatically by means of deep transfer learning.	Algorithms: (CNN) based ResNet34 model is used as a Deep Learning model. In addition, Current Deep Learning methods have been used for training the model, such as data augmentation, optimal rate determination and fine tuning.	The suggested model accomplished a 5-fold accuracy of 100% classification.	The dataset considered small only 613 images.

The following subsection presents the systematic review protocol for the brain tumor MRI medical images and Deep Learning algorithms.

2.2 SYSTEMATIC LITERATURE REVIEW PROTOCOL

The extent of this review is determined by the use of the keywords ‘Deep Learning’, ‘brain tumor’, and ‘MRI’. The scope was restricted to three digital databases: (1) the Science Direct database providing highly reliable journals with access to science, technology, and articles; (2) the IEEEXplore Library of Engineering and Technology Technical Literature, and (3) Scopus database covering Peer-reviewed journals in essential subject areas: social sciences, physical sciences, and medical sciences.

The process of selecting corresponding studies is relatively tricky, especially when accounting for the many research fields. Filtering papers is crucial and most important when investigating certain subjects. The first step was to remove duplicates, which was achieved using Mendeley. Then, the title and abstract were screened, and irrelevant documents excluded. The next step involves reading the selected papers, sometimes abstract does not give clear reflection of the whole contents, when we went deeper in some papers, we found that they are irrelevant references to the current work. So, we removed those papers, as per Figure 2.1.

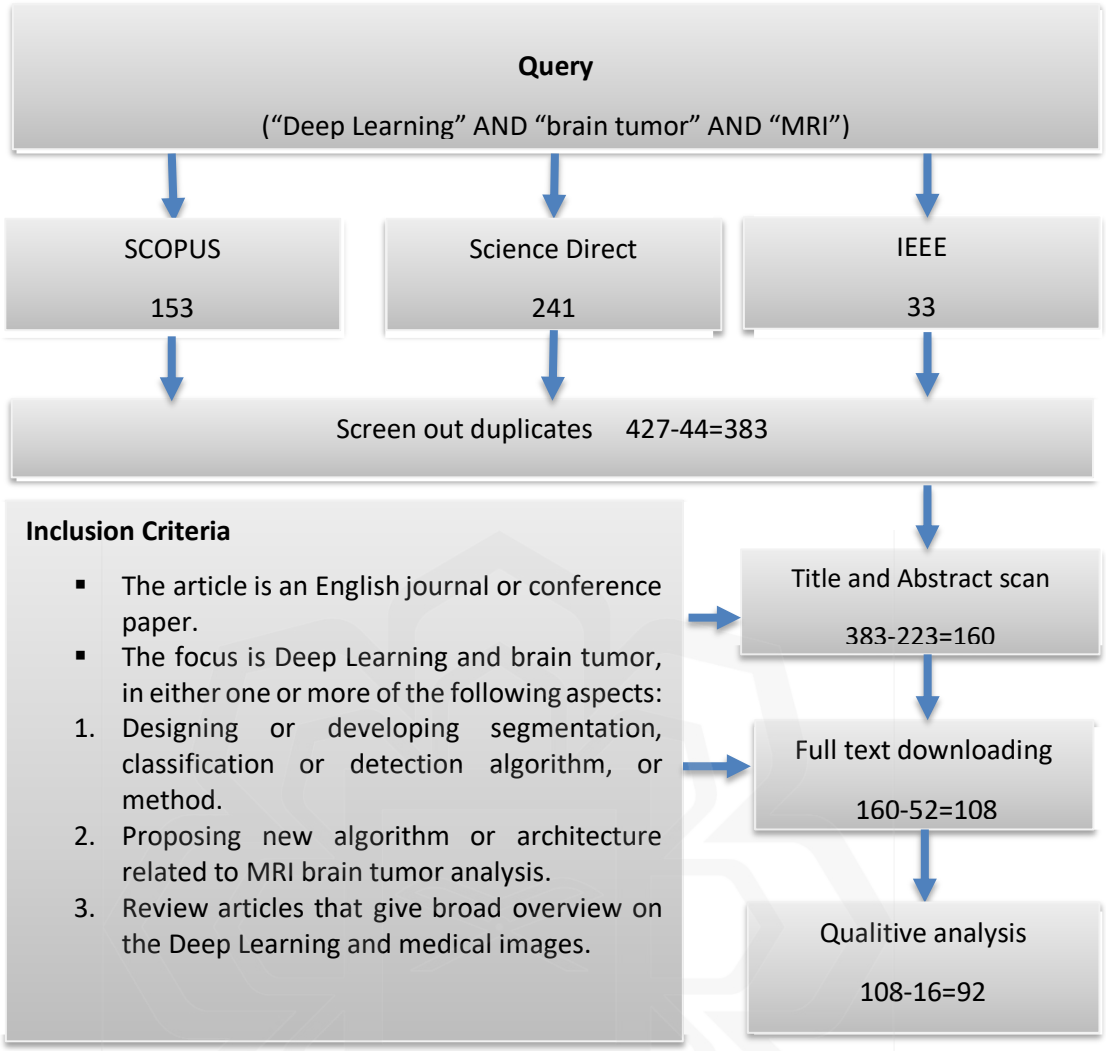


Figure 2.3 The flowchart for research collection includes query words and inclusion criteria.

2.3 TAXONOMY ANALYSIS

The search for the abovementioned query resulted in 427 papers. The field is relatively new, so we focused on recent years (2018-2021). Research alerts were set up for the same query in the three abovementioned databases (IEEE, Science Direct, and Scopus). Forty-four studies were duplicates and were removed, leaving a total of 383 papers. The papers were scanned by reading their titles and abstracts, and we ended up with 160 good papers. The number of papers that were available and downloaded was 108. The

final set included 92 papers that were thoroughly read. The set of articles were divided into two main categories: segmentation and classification. The taxonomic review identified the general categories in Deep Learning for brain tumor medical images analysis. The mapping of the study illustrated in Figure 2.2 represents the area of interest.

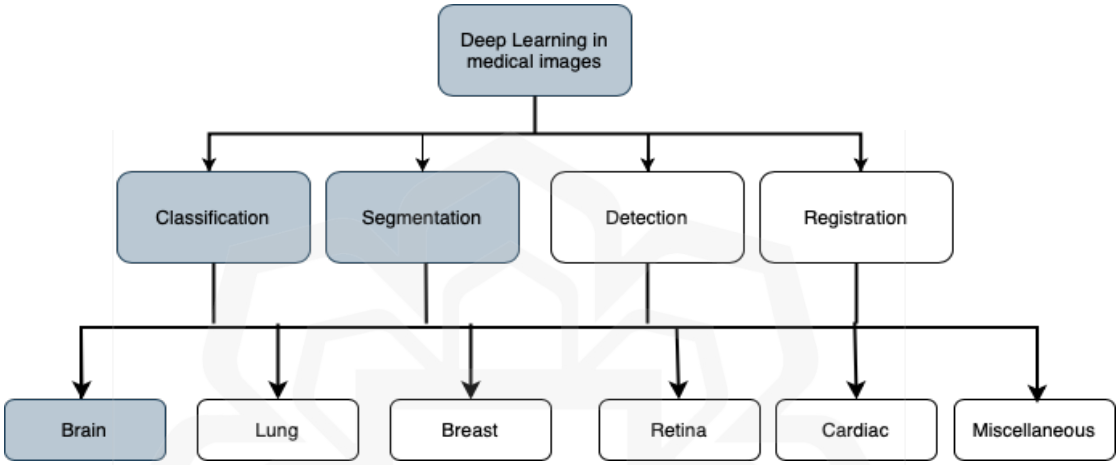


Figure 2.4 Research taxonomy of Deep Learning in medical images showing the pattern recognition tasks and the anatomical regions, the grey highlighted boxes indicate the focus of the research.

2.4 CLASSIFICATION

In Machine Learning there are three main branches: supervised learning, unsupervised learning and reinforcement learning. As illustrated in the Figure 2.3 (Chug, 2018) classification is a main branch under supervised learning. Brain tumor is an abnormal cell growth in the brain. The classification of brain tumours aids in predicting their potential behaviour and improving health-care systems. Tumors are classified based on their cellular origin and actions, ranging from less active (benign) to aggressive (malignant), malignant tumor are graded from I (least malignant) to IV (most malignant) based on their rate of growth (DeAngelis, 2001).

Classification of brain medical images can be binary classification, as the name refers binary classification is for two classes a normal class and an abnormal one. It is used in medical testing to detect whether or not a patient has a certain condition. In brain MRI medical images for a tumor detection this task involves classifying the images to normal and affected by a tumor. Another task could be classifying the tumor itself to benign or malignant tumor (Kumari & Kr., 2017). Another type of classification is Multiclass classification, this task involves categorising objects into one of multiple classes. In medical imaging this task is widely used to classify different types of diseases. For instance, brain tumor has more than 120 different types.

The main goal in this task is to classify one class out of three or more classes (Irmak, 2021; Johnson et al., 2017). Meningiomas are brain tumor that develops in the small walls that generally surround the brain and is mostly non-cancerous, Gliomas are one of the most common types of primary brain tumors. Most pituitary tumors are noncancerous (benign) growths (adenomas). Adenomas remain in the pituitary gland or surrounding tissues and don't spread to other parts of the body (S. Chen et al., 2019). Generally, one of the life-threatening disorders that can directly damage human lives is brain tumors (Thillaikkarasi & Saravanan, 2019).

One of the most important challenges in computer vision and pattern recognition is image classification. Researchers have previously investigated brain tumor MRI (Magnetic resonance imaging) classification using Machine Learning algorithms. In traditional Machine Learning the classification task need prior step which is features extraction. While in Deep Learning algorithms the process can be completed without the use of handcrafted features, and both the feature extraction and classification steps are combined (Guido, 2016; W. Wang et al., 2020). In recent years. In the field of medical image analysis and disease detection, the development of artificial intelligence and Deep Learning-based new technologies has had a significant impact. Deep Convolutional Neural Network (DCNN) is one of the most extensively utilized image processing techniques (Litjens et al., 2017; Muhammed Talo, Yildirim, Baloglu, Aydin, & Acharya, 2019). The use of CNN applications to classify different types of brain tumors has demonstrated a good performance and promising results (Wachinger et al., 2018). In this research we have implemented the Design Science Research Methodology (DSRM) (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007) to

design a new CNN model for the objective of multiclassification of brain tumor MRI medical images from publicly available dataset containing three types of brain tumors namely (meningioma, glioma, and pituitary) (Chug, 2018).

There are different tasks in classification; they are:

1. Binary classification, as the name refers binary classification is for two classes a normal class and an abnormal one. It is used in medical testing to detect whether or not a patient has a certain condition. In brain MRI medical images for a tumor detection this task involves classifying the images to normal and affected by a tumor. Another task could be classifying the tumor itself to benign or malignant tumor.
2. Multiclass classification, this task involves categorising objects into one of multiple classes. In medical imaging this task is widely used to classify different types of diseases.
3. Multilabel classification, in this task each class might have numerous labels assigned to it. Unlike multi-class task, in multilabel classification task There is no limit on how many classes an instance can be assigned to.

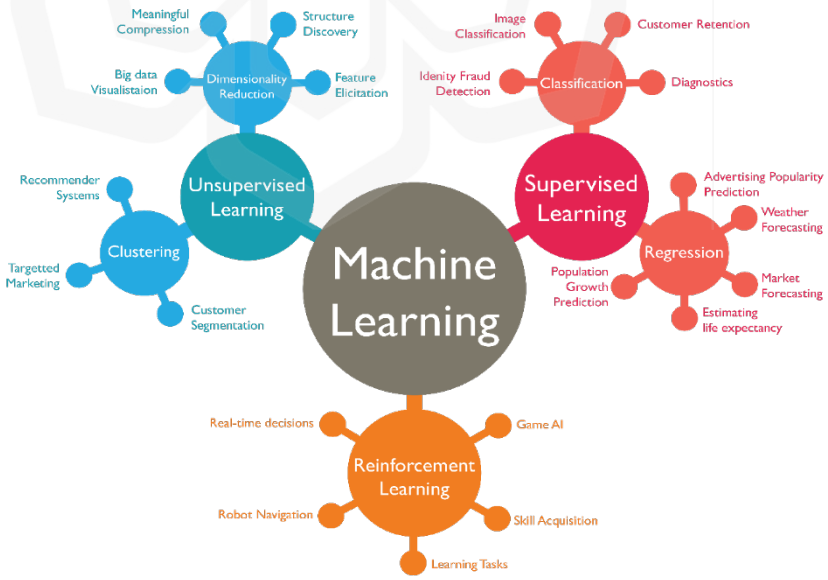


Figure 2.5 A comprehensive view of Machine Learning (Chug, 2018)

Image classification is assigning one or more object labels to a category. It is one of the essential computer vision and pattern recognition tasks. In conventional image classification systems, an image is represented with low or medium-level characteristics, which are used for label assignments by a trainable classifier, while in DL, the process can be carried out without the handcrafted features, and both steps of feature extraction and classification are combined (W. Wang et al., 2020). Many research institutions use the classification systems to classify brain tumors, where ~120 types of tumors are found in the brain and central nervous system (CNS). The main goal in this task is to classify one class out of three or more classes The tumor is classified according to their cellular origin and actions from the less active (benign) to the more aggressive (malignant). Tumors are assigned to grades I (least malignant) to IV (most malignant), representing its growth rate (Johnson et al., 2017).

The use of Machine Learning and Deep Learning algorithms to detect and classify brain tumour images using multiple imaging modalities, particularly MRI, has gained popularity recently. Traditional ML algorithms involve many steps before classification, mainly pre-processing and feature extraction. Feature extraction is a crucial stage in classic ML approaches since classification accuracy is dependent on the extracted features (Kang, Ullah, & Gwak, 2021). In Deep Learning algorithms, extracted features are performed automatically, resulting in significantly improved performance. However, researchers could face obstacles when using Deep Learning models because it requires big data to train the network, which could be overcome using data augmentation, and if it is done appropriately, it can serve as an effective way to enlarge and balance datasets (A. R. Khan et al., 2021).

The process of assigning one or more item labels to a category is known as image classification, which is one of the most critical tasks in computer vision and pattern recognition (Javaria Amin, Sharif, Gul, Yasmin, & Shad, 2020). Most recent studies used the DL methods for the classification of brain tumour medical images. Sultan proposed a network consisting of 16 layers for multiclass classification of brain tumour MR images, and they reported achieving 96.13% and 98.7% accuracy for two datasets, respectively (Sultan, Salem, & Al-Atabany, 2019). Irmak used three datasets, one for binary classification and the other two representing different grades and types of brain tumour MRI medical images; the result showed that brain tumour (binary-

classification) detection is 99.33% accurate, and they reported 92.66% and 98.14 for multi-classification datasets, respectively (Irmak, 2021).

Transfer learning is the approach used to learn different datasets using the knowledge gained via an already trained model (Pan & Yang, 2010). If the dataset size required is insufficient, the network parameters can keep avoiding overfitting. When the objective size is insufficient, transfer learning methods are commonly used. There are popular architectures for this, such as AlexNet (Krizhevsky, Sutskever, & Hinton, 2012), GoogleNet (Lin, Chen, & Yan, 2014), VGG19(Simonyan & Zisserman, 2015), and ResNet (He, Zhang, Ren, & Sun, 2016). Many researchers used transfer learning strategies in DL networks. Swati al. used transfer learning and fine-tuning on CNN (VGG19)(Swati et al., 2019), while Bakr et al. implemented transfer learning on (VGG16)(Badža & Barjaktarović, 2021; Y. D. Zhang et al., 2018). Deepak and Ameer used transfer learning on GoogleNet(Deepak & Ameer, 2020), while Gonella and Binaghi used V-Net to implement transfer learning (Gonella, Binaghi, Nocera, & Mordacchini, 2019). There are many other attempts reported in the literature (Lu, Lu, & Zhang, 2019; Rehman, Naz, Razzak, Akram, & Imran, 2019; Saba, Mohamed, El-Affendi, Amin, & Sharif, 2020; M. Talo et al., 2019).

Multiclass classification of brain tumours MRI images has been previously investigated and reported in the literature. This section elucidates the approaches proposed in related, recent, and significant works found in the literature. Afshar et al. (P. Afshar, Mohammadi, & Plataniotis, 2018) used the capsule network (capsnet) to classify three types of brain tumours, and they reported obtaining 86.56 % for the whole tumour MRI image and 86 % for the segmented tumour images. Pashaei et al. (Pashaei, Sajedi, & Jazayeri, 2018) used three approaches (KELM-CNN-KECNN), and KECNN reported the best accuracy, at 93.68%. Swati et al. (Swati et al., 2019) applied transfer learning of VGG19 and reported an average accuracy of 94.82% under five-fold cross-validation. Similarly in (M. Sajjad et al., 2019) used the same method (transfer learning of VGG19) and reported similar accuracy of 94.58%, sensitivity 88.41%, and specificity 96.12% .

It was proven that using the softmax activation function is more reliable than other functions in multiclass classification, such as the sigmoid function. Zyad et al. (Zyad, Gouskir, & Bouikhalene, 2019) compared softmax and sigmoid activation

functions for the same dataset used in this study, and the method using the softmax output layer outperforms the technique with the sigmoid output layer. However, it should also be pointed out that the sigmoid activation function performs better in binary classification (Kang et al., 2021). Ghassemi et al. (Navid Ghassemi, Shoeibi, & Rouhani, 2020) reported that the random split of the test dataset and training dataset performed better than the introduced split when using the five-fold cross-validation. The accuracy of the introduced split is 93.01%, and the random split reported an accuracy of 95.6% .

The classification of brain tumors helps predict its possible behaviour and improve health systems. One common dataset that many researchers used to train their systems is the 3064 T1-CE MR images, consisting of three types of brain tumors (meningioma, Glioma, and pituitary tumors) obtained from 233 patients(Cheng et al., 2015). Table 2.2 lists the studies involved in applying the abovementioned datasets. They used different DL methods, except one study (Sheela & Suganthi, 2019), which used the snake contour detection algorithm; however, the latter's accuracy was lower than those using the DL algorithms. The best accuracy was reported by (Deepak & Ameer, 2019); however, it suffers from considerable misclassification of samples from the class of meningioma. Another study used the same dataset to compare two CNN archaicities and reported that the method with the Softmax output layer performs better than the method with the sigmoid output layer(Han & Kamdar, 2015).

Table 2.2 Classification studies that used the same dataset.

Ref.	Methods	Result
(N Ghassemi, Shoeibi, & Rouhani, 2020)	CNN	93.01% on the introduced split and 95.6% on a random split.
(M. Sajjad et al., 2019)	VGG19	88.41%, 96.12%, and 94.58% for sensitivity. specificity, and accuracy respectively.
(Sheela & Suganthi, 2019)	Snake Contour detection	DSC was calculated as 0.78, 0.59 and 0.49 For Meningioma, Glioma and Pituitary tumor respectively.

(Zyad et al., 2019)	CNN DENSE SOFTMAX SIGMOID	The method with Softmax output layer performs better than the method with sigmoid output layer.
(Swati et al., 2019)	VGG19	Achieve average accuracy of 94.82% under five-fold cross-validation.
(Deepak & Ameer, 2019)	GoogleNet. CNN SVM.	Classification accuracy of 98%.
(Pashaei et al., 2018)	KELM CNN KE-CNN	The accuracy of KE-CNN method on this dataset is =93.68%
(Parnian Afshar, Mohammadi, & Plataniotis, 2018)	CapsNet	Segmented tumor= 86 %.

The following Table 2.3 summarizes articles involving brain tumor classification using datasets other than the ones introduced earlier. Brunese et al. reported an excellent precision rate, reaching ~99 % for grades (I, II and IV), and 97 % for grade III, however, its method requires handcrafted features (Brunese et al., 2020). Talo et al. and Lu et al. used the same dataset from Harvard Medical School (HMS), with extensive data augmentation for the first mentioned reference. Their results are impressive at an accuracy of 100%, which is due to its limited range of images (Lu et al., 2019; M. Talo et al., 2019). Saxena et al. applied the CNN architecture, consisting of five fully connected layers and the softmax layer on the REMBRANDT dataset, but they obtained a relatively a low accuracy rate relative to that reported in another research (Saxena, Sharma, Joshi, & Rana, 2019). Özyurt et al. used the super-resolution Fuzzy C-Means (SR-FCM) for tumor detection on the TCIA datasets of 500 samples, followed by classification using the CNN extreme learning machine (Özyurt, Sert, & Avcı, 2020). They proved that the accuracy with a super-resolution is 10% better than without super-resolution, but the main drawback of the SR-FCM-CNN strategy is that it relies on the training dataset. Mallick et al. did not use CNN, they implemented the Deep Wavelet Autoencoder (DWA) on RIDER (Reference Image Database to Evaluate Therapy Response) dataset from TCIA and reported an excellent accuracy rate of ~93% (Mallick

et al., 2019). However, researchers recommended using CNN, especially for the classification of medical images.

Table 2.3 Classification studies with different datasets

Ref.	Dataset	Methods	Result
(Brunese et al., 2020)	3 DATASETS: 1. Jefferson University Hospitals (110020) MRI From 130 patients. 2. BRATS 2019. 3. dataset is gathered from the Radiopaedia Repository.	ML algorithms (SVM-KNN-RF-QDA)	A precision of 0.991(Grade I), 0.994 (Grade II), 0.976 (Grade III) and 0.990 (Grade IV).
(Özyurt et al., 2020)	TCIA was preferred. In the relevant database, approximately 500 sample,s were selected per different types of cancer and tissues	SR-FCM squeezeNet ELM classifier	98.33% accuracy rate has been detected in the diagnosis of segmented tumor.
(Lu et al., 2019)	From Harvard medical school contains contain 38 normal and 177 pathological ones	Transfer learning AlexNet	100 % accuracy.
(Saxena et al., 2019)	REMBRANDT which consists of MR scans of 130 patients suffering from glioma tumors of different types and at different stages. From this dataset, a total of 38,952 images.	CNN Softmax Layer	Training accuracy of 63.17%, validation accuracy of 56.67% and test accuracy of 65.24%.

(Isselmou, Xu, Zhang, Saminu, & Javaid, 2019)	20 FLAIR and T1-T2 weight brain images from the Tianjin Medical University Hospital, China.	Fuzzy C-Mean CNN.	Network for 24 epochs achieved Dice similarity 86,785%, and accuracy 98 %.
(Mallick et al., 2019)	RIDER (19 subjects)	DWA-DNN	93% accuracy
(Muhammed Talo, Baloglu, et al., 2019)	613 images only including augmentation from Harvard medical school website.	(CNN) based ResNet34	100% accuracy after 3 stages and 50 epochs
(Sengupta et al., 2018)	9 HGG with pre- and post-surgery MRI data and 9 metastasis patients with pre-surgery MRI data.	SVM	Misclassification error of 8.4%
(C. Ge, Gu, Jakola, & Yang, 2018)	BRATS 2017contains 3D brain volume images from Mayo Clinic in USA	7 layers 2D CNN	Test accuracy of 90.87% for former case, and 89.39 % for the latter case

The TCIA (The Cancer Imaging Archive) is an online resource sharing repository that is publicly available for download from different organizations; therefore, no approval is needed from the Institute Review Board (IRB). It contains many collections of different areas of the human body, with most being either CT/MRI types of medical images. The first published dataset was for Prostate Cancer in June 2011. One example of these collections was a dataset for a brain tumor published in February 2019. This collection contains datasets of 20 newly diagnosed primary glioblastoma patients who underwent surgery and conventional concomitant chemotherapy (CRT) and adjuvant chemotherapy.

For each patient, two IRM examinations were included: within 90 days after completion and advancement of a CRT (clinically determined and based on clinical results or imagery and marked by treatment or intervention change). All images are in the form of DICOM format and included T1w (pre-post contrasting agent), FLAIR, T2w, ADC, normalized cerebral blood flow, normalized relative cerebral blood volume, standardized relative cerebral blood volume and binary tumor masks (created with T1w photographs). The pictures of perfusion were created after a preload of the contra's agent from the dynamic susceptibility comparison (GRE-EPI DSC). The T1+C images are co-recorded for all of the series. The dataset evaluated the efficiency of a detailed learning algorithm for forecasting the development of the tumor (Schmainda et al., 2018). More than 40 studies used this dataset; Table 2.4 includes some detail of these studies.

Table 2.4 The studies that refers to the TCIA Brain-Tumor-Progression datasets.

Ref.	Field	Methodology
(Parthasarathi & Ansari, 2018)	Medical and Biological Engineering	Seed detection algorithm
(Lee Et Al., 2017)	Medicine	Topological data analysis
(Beig Et Al., 2017)	Computer Science	Unsupervised clustering by Euclidean distance.
(Czarnek, Clark, Peters, & Mazurowski, 2017)	Medicine	quantified using algorithmic analysis of MR images.
(Tian Et Al., 2017)	Biomedical Engineering	SVM and a radial basis function
(Kanas Et Al., 2017)	Computer Science	Random forest, Gaussian Naive Bayes, KNN
(Dunn Jr, Hwang, Cooper, Aerts, & Holder, 2016)	Medicine	Velocity AI FSL 3D slicer

(Chaddad, Desrosiers, & Toews, 2016)	Medicine	Rigid registration tool. Gary Level co- occurrence matrix (GLCM)
(Le Reste, Stindel, Morvan, Upadhaya, & Hatt, 2016)	Computer Science	SVM
(Wiest Et Al., 2015)	Medicine	(BraTumIA)
(Nabizadeh & Kubat, 2015)	Computers and Electrical Engineering	SVM-KNN-NSC - k-Means Gaussian-Euclidean city block-Sparse
(Jothi, 2015)	Computer Science	SVM, Neural network
(Reza, Mays, & Iftekharuddin, 2015)	Compute Science	Random Forest (RF)
(Rubin Et Al., 2015)	Medicine	radiologists' toolbox. Consensus matrix, CDF curve, and delta curve for all clusters.
(Han & Kamdar, 2015)	Computer Science	CRNN
(Kirby Et Al., 2014)	Medicine	Measured in terms of overall survival time
(Huang Et Al., 2014)	Medicine	VASARI feature-set
(Kwon, Shinohara, Akbari, & Davatzikos, 2014)	Medicine	Expectation Maximization (EM) algorithm
(Kirby, Jaffe, Et Al., 2014)	Medicine	overall survival (OS) and progression-free survival (PFS) were
(Mazurowski, Zhang, Peters, & Hobbs, 2014)	Medicine	leave-one-out cross validation approach to automatically extract imaging features

2.5 SEGMENTATION

The segmentation method divides the image into areas that are not overlapping by a variety of rules/parameters, such as a collection of identical pixels/characteristics, encompassing contrasts, colours, and textures (L. Zhang & Ji, 2011). The number of publications devoted to automated brain tumor segmentation has grown exponentially over the past decade (B. H. Menze et al., 2015; B. Menze et al., 2014). The use of segmentation in medical images includes relevant information derived from the shape, length, organ location, and symptoms (Anwar et al., 2018). Tables 2.5 and 2.7 list the brain tumor medical images segmentation studies classified based on its used datasets. As per Table 2.5, more than forty studies within the research inclusion criteria used BRATS (more detail on this dataset explained in the upcoming sections), while other studies listed in Table 2.7 used different datasets.

In ML technology, the Support Vector Machine (SVM) and random forest have been used extensively for the automatic segmentation of the brain tumor (Ayachi & Ben Amor, 2009; Liaw & Wiener, 2002). These processes involve the feature engineering process to train the respective ML algorithms as per (Saba et al., 2020), where they achieved a Dice Similarity Coefficient (DSC) of 96 % on BRATS (2015). However, this study depends on handcrafted feature extraction. Pereira et al. applied a hybrid algorithm of random forest algorithm and Restricted Boltzmann Machine (RBM) and achieved a DSC of 84% on the complete, while they obtain ~74% on the core, tested on BRATS (2013) (Pereira et al., 2018). Thaha et al. developed an Enhanced Convolutional Neural Networks (ECNN) with the BAT algorithm by optimizing the loss function for the automatic segmentation process (Thaha et al., 2019). The findings of the experiments confirmed a better efficiency relative to the conventional methods. The parameters compared are precision (CNN: 82%; ECNN: 87%) recall (CNN: 85%; ECNN: 90%) and accuracy (CNN: 89%; ECNN: 92%). However, various selection schemes can be implemented to increase the accuracy.

CNN architecture called U-Net is a breakthrough in the automated segmentation of images (Maier, Syben, Lasser, & Riess, 2019). Ronnebergers's U-Net has a contraction direction and a symmetrical expanding direction with intermediate overhead links. The mirroring method is used for the estimation of the boundary pixels

(Ronneberger, Fischer, & Brox, 2015). Nema and Dudhane achieved a DSC of ~94% on BRATS (2015) and ~94.63% on BRATS (2017). They used the U-Net alongside the skip connection method (Shubhangi Nema, Dudhane, Murala, & Naidu, 2020). Li et al. reported a novel brain tumor segmentation using U-Net and added a skip connection and up skip connection when using two datasets (2015) and BRATS (2017) (H. Li, Li, & Wang, 2019). They achieved a DCS of ~89% and ~86%, respectively. V-net has a similar architecture to U-Net but is associated with 3D images. Gonella et al. applied the V-net algorithm on 3D MRI images and achieved a DCS of ~64 % on a BRATS dataset (Gonella et al., 2019).

Pereira et al. introduced the idea of using small-sized kernels due to its positive impact against overfitting (Pereira, Pinto, Alves, & Silva, 2016). The use of small kernels resulted in a deeper architecture due to the smaller weights in the system. They took part in the on-site BRATS (2015) challenge using their model, placing second with DSC matrices of 0.78, 0.65, and 0.75 for the full, core, and enhancing regions, respectively. Havaei et al. introduced a novel CNN architecture that incorporates a two-way architecture that can absorb local brain information in a broader context (Havaei, Davy, Warde-farley, et al., 2017). The two-phase training method is crucial for resolving unbalanced label distributions. They achieved an accuracy of ~88%. Kamnitsas et al. merged the previously mentioned algorithms (small kernel size and two pathway (normal resolution and low resolution)) and applied the 3D CNN; dense training on image segments and building deeper networks with 3D kernels (Kamnitsas et al., 2017). The results using BRATS (2015) were DSC (~90.1%), precision (~91.9 %), and sensitivity (~89.1%).

Table 2.5 Brain tumor medical images segmentation studies that use BRATS Datasets

Ref.	BRATS dataset	DL methods	Performance measures (DSC)
(Saba et al., 2020)	2015-2016-2017	LGR-DT-KNN-LDA-	DSC= 0.9636 on BRATS 2015

		SVM- VGG19	
(S. Nema, Dudhane, Murala, & Naidu, 2020)	2015-2017	U-Net	DSC= 0.9401 on BRATS 2015 DSC= 0.9463 on BRATS 2017
(H. Li et al., 2019)	2015-2017	U-Net	DSC = 0.89 on BRATS 2015 DSC=0.86 on BRATS 2017
(Sharif, Li, Khan, & Saleem, 2019)	2013-2018	CNN AlexNet	Accuracy = 98.3%
(Deng, Shi, Luo, Yang, & Ning, 2019)	2015	FCNN DMDF	DSC= 90. 89%.
(Kermi, Mahmoudi, & Khadir, 2019)	2018	GDL	DSC= 0.783
(Gonella et al., 2019)	-	SVM V-Net	DSC=64 %, Precision (P)=85% Recall (R)= 53%
(Benson, Pound, French, Jackson, & Pridmore, 2019)	2018	Hourglass U-NET	DSC (complete= 0.82, core=0.72, enh. =0.66)
(Kuzina, Egorov, & Burnaev, 2019)	2018	U-Net-RI U-Net-PR U-Net-DWP	DSC =0.74
(S. Chen et al., 2019)	2015-2017	U-Net DeepMedic MLP	DSC= 85 % on BRATS 2015 DSC= 89.30 % on BRATS 2017

(Mlynarski, Delingette, Criminisi, & Ayache, 2019)	2017	2D-3D Model U-NET	DSC (complete= 0.918, core=0.883, enh. =0.854)
(L. Sun, Zhang, Chen, & Luo, 2019)	2018	CA-CNN U-Net	DSC=61.0%
(Thaha et al., 2019)	2015	CNN ECNN	Precision= (CNN: 82%, ECNN: 87%) Recall= (CNN:85%, ECNN: 90 %) Accuracy= (CNN: 89%, ECNN :92%)
(Havaei, Davy, Warde-Farley, et al., 2017)	2013	CNN	DSC=88%
(Kamnitsas et al., 2017)	2015	3D CNN	DSC= 90.1 % Precision =91.9 % sensitivity =89.1%
(Pereira et al., 2016)	2015	CNN	DSC (complete= 0.78, core=0.65, enh. =0.75)
(Zhai & Li, 2019)	2015-2016	VGG-16	Accuracy= 82.2 %
(J. Sun, Chen, Peng, & Liu, 2019)	2015	3D CNN	DSC (complete= 0.84, core=0.72, enh. =0.62)
(Javeria Amin et al., 2019)	2015-2016-2017 ISLES 2018	Alex/ GoogleNet.	DSC= 0.9891 on BRATS 2015
(Pereira et al., 2018)	2013 SPES	RBM- RF	DSC (complete= 0.84, core=0.74, enh. =0.71)

(X. Kong, Sun, Wu, Liu, & Lin, 2018)	2015-2017	HPU-Net	DSC (complete= 0.92, core=0.80, enh. =0.76) on BRATS 2017 DSC (complete= 0.90, core=0.71, enh. =0.78) on BRATS 2015
(Hoseini, Shahbahrami, & Bayat, 2018)	2016	DCNN	DSC (complete= 0.90, core=0.85, enh. =0.84)
(Y. Hu & Xia, 2018)	2017	Cascade U-Net	DSC (complete= 0.81, core=0.69, enh. =0.55)
(Zhao et al., 2018)	2015 -2016	FCNNs CRF-RNN	DSC (complete= 0.84, core=0.73, enh. =0.62)
(naceur, Saouli, Akil, & Kachouri, 2018)	2017	XCNet-ELOBA_λ	DSC (complete= 0.89, core=0.76, enh. =0.81)
(Ramirez, Martin, & Schiavi, 2018)	2015	MLP/ U-Net	Precision=84 Recall=88 DSC=85
(Charron et al., 2018)	2015	3D-CNN DEEPMEDIC	DSC= 79 %
(D. Liu et al., 2018)	2015	DCR RESNET 50	DSC= 87%
(Iqbal, Ghani, Saba, & Rehman, 2018)	2015	SkipNet SENet IntNet (VGG)	SkipNet DSC= 87% SENet. DSC= 88% IntNet(VGG) DSC= 90%
(Baid et al., 2019)	2013-2015	CNN Linear nexus	DSC (complete= 0.86, core=0.87, enh. =0.90)

(Cui, Mao, Jiang, Liu, & Xiong, 2018)	2015	TLN / ITCN FCN	DSC (complete= 0.89, core=0.77, enh. =0.80)
(Islam & Ren, 2018)	2017	PIXELNET VGG16	DSC=87.6 %
(Pinto, Pereira, Rasteiro, & Silva, 2018)	2013	ERT	DSC=85%
(Qamar, Jin, Zheng, & Ahmad, 2018)	2018	3D CNN	DSC (complete= 0.87, core=0.84, enh. =0.81)
(Mohammadreza Soltaninejad et al., 2018)	2013	RF-DTI	DSC=96%
(M. Soltaninejad et al., 2018)	2017	VGG16.	DSC (complete= 0.86, core=0.78, enh. =0.66)
(G. Wang et al., 2018)	2015	CNN	DSC= 86.13%
(Y. Wang, Li, Zhu, & Zhang, 2019)	2015-2018	WRN RESNET	DSC= 0.91 Sensitivity score= 0.94 PPV = 0.89
(Z. Zhang, Odaibo, Skidmore, & Tanik, 2018)	2017	CNN	DSC=72%.
(Takacs & Manno-Kovacs, 2018)	2015	Saliency map U-Net	Precision=86%

There are several medical images dataset of brain tumors mentioned in the previous sections, but the most common datasets used by researchers are BraTS due to its availability and reliability. More details on the datasets are given in Table 2.6 below.

Table 2.6 Details of BraTS datasets (Bakas et al., 2018)

BraTS	Country	Cases for training	Cases for testing
BraTS 2012	Nice, France	5 LGG, 10 HGG	-
BraTS 2013	Nagoya, Japan	10 LGG, 20 HGG	-
BraTS 2014	Boston, USA	16 LGG, 175 HGG	-
BraTS 2015	Munich, Germany	54 LGG, 220 HGG	110 (HGG, LGG)
BraTS 2016	Athens, Greece	54 LGG, 220 HGG	191(HGG, LGG)
BraTS 2017	Quebec City, Canada	75 LGG, 210 HGG	146(HGG, LGG)
BraTS 2018	Granada, Spain	75LGG, 210 HGG	66(LGG, HGG)
BraTS 2019	Shenzhen, China	-	-
BraTS 2020	Virtual	-	-

In MRI research, BraTS has been the focus of the assessment of state-of-the-art techniques for the segmentation of brain tumors. BraTS use MRT multi-institutional pre-operative scans to target the segmentation of brain tumors, namely gliomas, that are

fundamentally heterogeneous (in their look, shape, and histology). Also, BraTS focuses on the patient's overall survival via integrative radiometric characteristics analysis and ML algorithms to define the clinical importance of this segmentation task. BraTS have eight successful subsequent collections between 2012 - 2020, as per Table 2.6 However, several unprocessed or partially pre-processed images were found in a testing collection for BraTS (2016) (Zhao et al., 2018).

Among all of the reviewed articles of brain tumor segmentation studies, ten did not use the BRATS datasets; those studies are listed in Table 2.7. According to (Thillaikkarasi & Saravanan, 2019), they achieved ~84 % accuracy using a dataset of only 40 MRI images. Similarly, many researchers used different datasets which have a limited number of images for the same purpose (H. Chen, Dou, Yu, Qin, & Heng, 2018; Gottapu & Dagli, 2018; Hongwei Li et al., 2018). However, researchers also reported better accuracies when using DL algorithms as being associated with the usage of a large number of images in the datasets (Litjens et al., 2017).

Table 2.7 Segmentation studies that use different datasets.

Ref.	Dataset	DL methods	Performance measures
(Thillaikkarasi & Saravanan, 2019)	Dataset: 40 MRI mixed images (normal and abnormal).	M-SVM CNN	Accuracy= 84%.
(G. Wang et al., 2019)	FLAIR images in the dataset and only segment the whole tumor. We randomly selected 234 cases for training and used the remaining 40 cases for testing	CNN (P-Net) (R-Net)	DSC= 89.93±6.49%
(J. Li, Yu, Gu, Liu, & Li, 2019)	The MRI data of 5 subjects (2 male and 3 female, varying degrees of atrophy and white matter lesions) were provided	Multi-modality aggregation network	DSC= 86.40%

	as training database. The MRI data of another 15 subjects were collected as the testing database of MRBrainS		
(Mamta Mittal et al., 2019)	2457 images from BRAINIX MRI images	GCNN SWT	Recall =98% Precision =98%.
(Roy & Maji, 2018)	NAMIC Brainweb Turmobase	CNN ROBEX	DSC=95%
(Hongwei Li et al., 2018)	Sixty cases from three centers were released as a public training set	U-Net	DSC= 80%
(H. Chen et al., 2018)	2013 MICCAI MRBrainS , which consist only of 5 training and 15 testing.	RELU BN RESNET	DSC=86%
(Dolz, Desrosiers, & Ben Ayed, 2018)	ISBR dataset ABIDE	3D FCNN	DSC= Thalamus, Caudate, Putamen Pallidum (0.92 ,0.87 0.13 ,0.25)
(Gottapu & Dagli, 2018)	IBSR 2(Internet Brain Segmentation Repository), which consists of 18 manually segmented MR images of the brain	Densenet Growth rate Bottleneck	Accuracy= 92%
(LAKSHMI, FEROZ, & MERLIN, 2018)	256 SAMPLES	SVM- CNN	Accuracy = 67%

2.6 CNN

DL in medical images analysis is triggered by CNN. The second-best image classifying error was halved by a DL algorithm (Convolutional Neural Network) in 2012 in ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) (Krizhevsky et al., 2012). Over the years, scientists developed different CNN architectures leading to superb improvement in the field, The error rate in competition is a good measure of this development such as in ImageNet Large-Scale Visual Recognition Challenge ILSVRC as shown in the Table 2.8.

In this competition the top-5 error rate for images classification fells from over 15% to over 3% in just three years. The top five error rate is the number of test images for which the system's to five predictions did not include the correct answer (Géron ,2017). The Table 2.8 shows that CNN started since 1998 with 60 thousand parameters, on 2012 AlexNet won the ILSVRC challenge with error rate 15.3% and 60 million parameters (Krizhevsky et al., 2012). On 2014 GoogleNet has been introduced and the error rate dropped to only 6.67 % but with only 4 million parameters. And last but not least is ResNet which won the challenge on 2015 and the error rate dropped dramatically to only 3.6% which surpasses human performance 5.1 % and so ResNet considered to be current state-of-art CNN architecture (He, Zhang, Ren, & Sun, 2014).

Table 2.8 ILSVRC top winners for developing CNN.

Year	CNN	By	Error rate	# Parameters
1998	LetNet	Yann Lecun et al.		60 thousand
2012	AlexNet	Alex Krizhevsky et al.	15.3 %	60 million
2013	ZFNet	Matthew Zeiler, Rob Fergus	14.8 %	
2014	GoogleNet	Google	6.67 %	4 million
2014	VGG Net	Simonyan, Zisserman	7.3%	138 million
2015	ResNet	Kaiming He	3.6 %	

It has been previously assumed that object identification in nature is difficult for computers, but CNN has now exceeded even humans in ILSVRC, effectively addressing the classification problem (Trakoolwilaiwan, Behboodi, Lee, Kim, & Choi, 2017). The next list detail some state-of-the-art CNN architectures.

2.6.1 AlexNet

AlexNet is the network that triggered the DL revolution since researchers made outstanding efforts to improve the ILSVRC 2012. It has 60 million parameters and 650,000 neurons, with five convolutional layers, max-pooling layers, activation function (RELU), and 1000-way softmax layer as shown in Figure 2.4 (Krizhevsky et al., 2012).

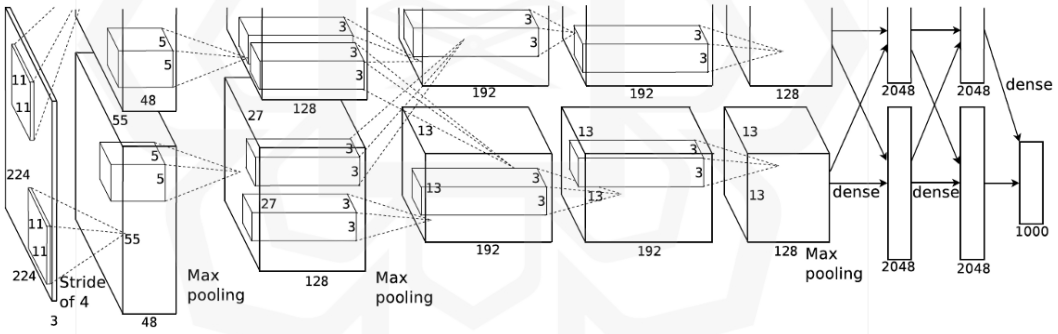


Figure 2.6 AlexNet 2012 Architecture from original article (Krizhevsky et al., 2012)

2.6.2 GoogleNet

The main idea introduced in GoogleNet is the better stacking of CNN layers, similar to the ones in the deep NIN (network-in-network) (Lin et al., 2014). The key concept behind the inception architecture is handling and coping with readily accessible dense components in a convolutional network (Szegedy et al., 2015). GoogleNet is also

famous for reducing the parameters of the network since it does not fully apply the connected layer at the end; instead, it uses the average pooling, which enables it to compete with more complex architectures to achieve a similar accuracy as illustrated in Figure 2.5 (Szegedy et al., 2015).

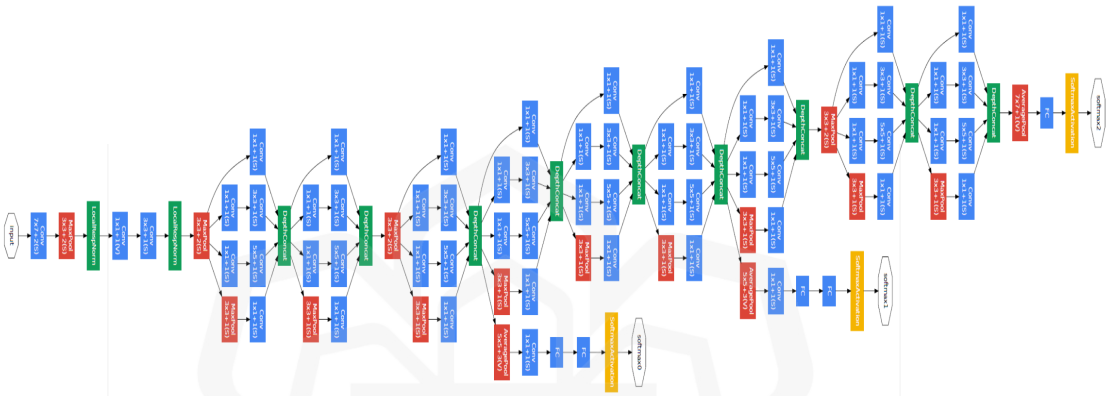


Figure 2.7 GoogleNet Architecture (Szegedy et al., 2015)

2.6.3 VGG19, VGG16

VGG is the culmination of the idea of small kernels and deeper networks, such as VGG 19, which is made up of 19 layers compared to only 7 layers in AlexNet. It comprises mainly of a sequence of convoluted layers using small kernels followed by three full-connected (FC) layers, with the last layer is the soft-max layer as shown in Figure 2.6 below (Simonyan & Zisserman, 2015).

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure 2.8 VGG architecture from original article (Simonyan & Zisserman, 2015)

2.6.4 ResNet

This CNN architecture was developed by Kaiming He et al., who won first place at the ILSVRC 2015 (He et al., 2016). One of ResNet's key advantages is that it enables the creation of an incredibly deep neural network by implementing the skip connection principle to skip one/more layers to solve the degrading precision and vanishing/exploding gradients issue (Abdelaziz Ismael, Mohammed, & Hefny, 2020). ResNet was used to develop extensive computer vision applications, but little work has been performed involving the segmentation of medical images (J. Sun et al., 2019). ResNext is the extension version of ResNet introduced by (Xie, Girshick, Dollár, Tu, & He, 2017) which combines the inception model from GoogleNet and skip connection from ResNet as shown in Figure 2.7 (Hara, Kataoka, & Satoh, 2018).

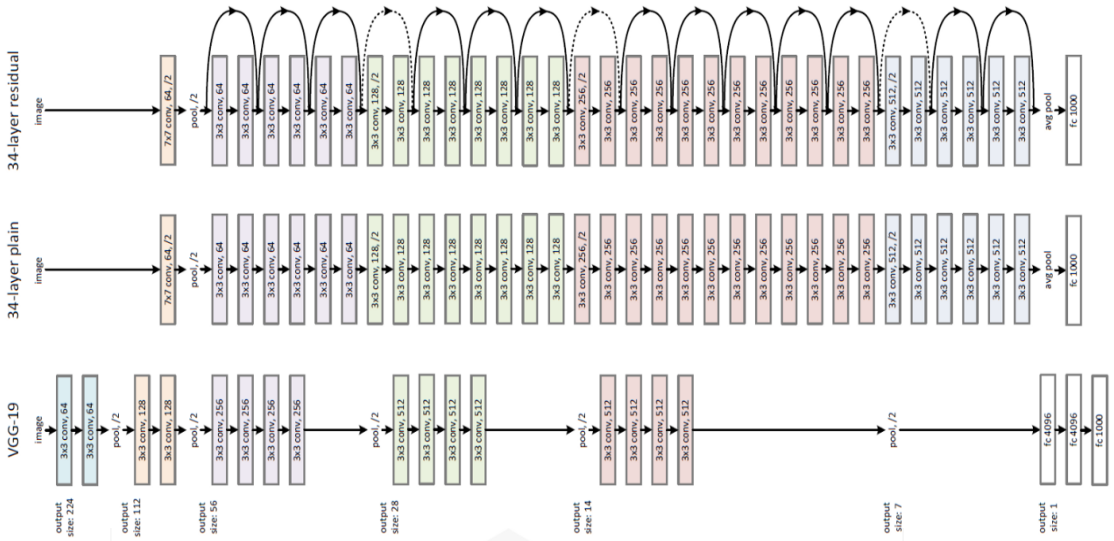


Figure 2.9 Graph from original article shows the different between Resnet and VGG19 and plain network (He et al., 2016)

2.6.5 U-Net

It is a popular and efficient 2D image segmentation network. The image is first processed using a conventional CNN architecture, and once an image is obtained, a symmetric network is used for the up-sampling path to revert the image to its original size. This network also applies the skip connection technique, similar to the way it is used in ResNet as in Figure 2.8 (Ronneberger et al., 2015). The 3D version of U-Net is called V-Net (Milletari, Navab, & Ahmadi, 2016).

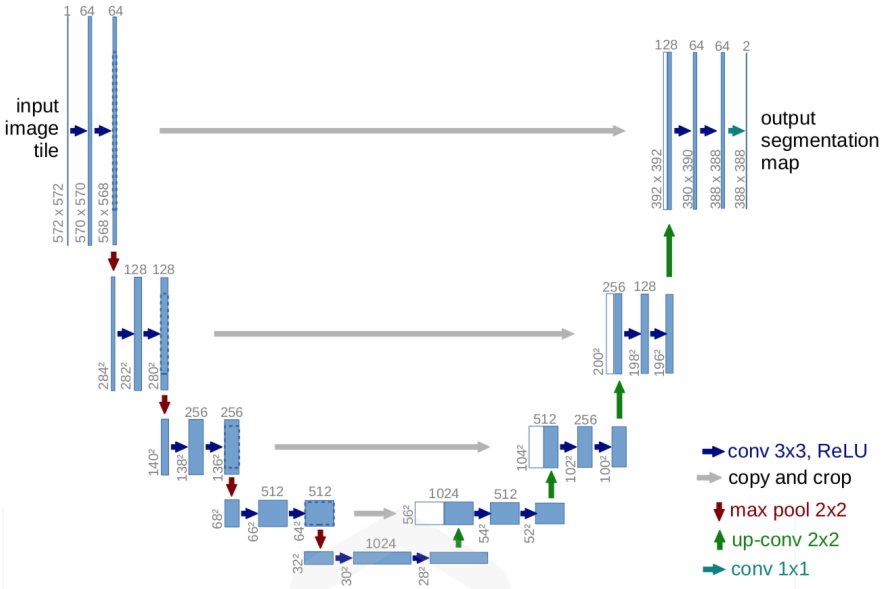


Figure 2.10 U-net Architecture from original article (Ronneberger et al., 2015)

2.6.6 SE-Net

The Squeeze and Execution (SE) network won the ILSVRC 2017 challenge. It focuses on the ResNext but added easy-to-train parameters that can be used by the network to measure each feature map, while previous networks only implemented it. They SE-Net layers enable the network to individually model the channel and space data to increase the model's capability. The Squeeze-Net layers can be applied to any CNN model, but doing so will increase its competitional costs (J. Hu, Shen, & Sun, 2018). Although ML and DL techniques are better than manual analyses, research is ongoing on improving both approaches in terms of its accuracy (B. H. Menze et al., 2015). The challenges of DL technology due to insufficient data and restricted labels are addressed via strategies such as data augmentation and transfer learning. Since DNN rely on convoluted complex structural models of the training data for predictions, it makes it difficult to interpret these predictions. The concern with the DNN is "black box," which means that it can make exact predictions, but the extracted features are ambiguous (Litjens et al., 2017; Selvikvåg Lundervold & Lundervold, 2018).

2.7 DATA AUGMENTATION

A broad dataset is essential for the success of the DL model. By increasing the existing data, the model's efficiency can be increased. Data augmentation allows researchers to increase the variety of data for learning models without actively acquiring new data. Some conventional techniques for data augmentation include rotation, cropping, padding, horizontal flipping, Gaussian blur, sharpen, edge detection (Muhammad Sajjad et al., 2019a), cross-modality image generation (Chenjie Ge, Gu, Store Jakola, & Yang, 2019), and synthetic data (Frid-Adar, Klang, Amitai, Goldberger, & Greenspan, 2018; Shin et al., 2018). Other researchers also suggested image translation (Krizhevsky et al., 2012), but it could lead to a wrong class to the patch for segmentation (Pereira et al., 2016). Data augmentation is an excellent method for increasing the datasets in DL applications as long as it is carefully used.

2.8 TRANSFER LEARNING

Transfer learning is a collection of ML techniques that can be used to save knowledge from a problem/dataset for application to a similar problem. If the intended dataset size is not enough, the network parameters can be preserved to prevent overfitting (Pan & Yang, 2010). Transfer learning methods are standard practice when the goal size is insufficient, and many researchers used this technique in DL architectures. Two transfer learning techniques can be utilized when the processed data is broad, as is the case in MRI medical images: first is the pre-trained network for extracting functions, and second is the fine-tuning a pre-trained network with the data (Litjens et al., 2017). Kuzina and Egorov's experiments have been conducted using a basic U-Net (transfer learning of the U-Net) model and using a subset of BraTS 2018(Kuzina et al., 2019). Transfer learning has also been applied to AlexNet and GoogleNet by Amin et al. (Javeria Amin et al., 2019), Talo and Baloglu applied transfer learning using the pre-trained CNN ResNet34, while Swati and Zhao applied transfer learning and fine-tuning on CNN (VGG19) (Abd-Ellah, Awad, Khalaf, & Hamed, 2018; Ahammed Muneer, Rajendran, & Paul Joseph, 2019; Saba et al., 2020; Swati et al., 2019). Deepak and Ameer applied transfer learning on GoogleNet (Deepak & Ameer, 2019), while Gonella

and Binaghi applied transfer learning on V-Net (Gonella et al., 2019). Many others adopted similar approaches (Banerjee, Crawley, Bhethanabotla, Daldrup-Link, & Rubin, 2018; Banzato, Bernardini, Cherubini, & Zotti, 2018; Lu et al., 2019; Rehman et al., 2019; Saba et al., 2020; M. Talo et al., 2019).

2.9 CHAPTER SUMMARY

The vast number of medical images make it difficult for accurate expert analyses. Manual analyses of medical images analysis are limited in their effectiveness (prone to human error) and are time-consuming (Kermi et al., 2019; Kokkalla, Kakarla, Venkateswarlu, & Singh, 2021; M. Talo et al., 2019). Despite these challenges, DL networks for the analyses of medical images continue to persist. Medical attention can also be leveraged in the area of Computer-Aided Diagnosis (CAD) systems for the improvement of the general attitude of computer science researchers. The significant advantage of the techniques of DL is its computability and consistency with many conventional techniques (Chlap et al., 2021). This merger has excellent advantages in the context of innovations. Typically, CNN architectures are known to be computationally very powerful (Hashemzehi, Mahdavi, Kheirabadi, & Kamel, 2020).

A systematic study of DL approaches for MRI brain medical images processing has been conducted. DL methods focusing on Convolutional Neural Networks (CNN) are more applicable to all sub-fields of medical image processing, such as classification, identification, and segmentation. The issues associated with DL approaches due to minimal data and labels are addressed using strategies such as data augmentation and transfer learning. Nonetheless, it is also assumed that other issues can be addressed with the collection of resources provided in this research. Based on the literature DL is likely to remain an essential field of research in MRI brain tumor medical image analyses.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 INTRODUCTION

Image classification is a complex task in computer vision and pattern recognition. Machine Learning techniques were previously used to examine brain tumour MRIs (Magnetic Resonance Imaging). The classification problem in traditional Machine Learning requires a preliminary step, which is feature extraction. However, Deep Learning algorithms can be accomplished without handcrafted features, as the feature extraction and classification processes are integrated (Muhammed Talo, Yildirim, et al., 2019; W. Wang et al., 2020).

Recently, Artificial Intelligence (AI) and Deep earning (DL) have substantially impacted medical image analyses and disease identification. One of the most widely used image processing algorithms is CNN and its applications. They are widely used to classify different brain tumours. They have demonstrated excellent performance and reported promising results (Litjens et al., 2017; Wachinger et al., 2018). Design Science Research Methodology (DSRM) (Peffers et al., 2007) is implemented to create a novel CNN model to multi-class classify MRI medical images of brain tumours from an available dataset for research and development purposes. It contains three types of brain tumours, namely (meningioma, glioma, and pituitary) (Chugh, 2018).

The "Design Science Research Methodology (DSRM)" is one of the most widely used research methods for problem-solving in information technology and data science. It is formed by two steps: design and evaluation (Peffers et al., 2007). Figure 3.1. shows that both steps are repeated multiple times until it satisfies the evaluated model. The primary goal of this study is to create an efficient model for classifying brain MRI medical images. The DSRM was employed as the research method for this purpose. This chapter discusses the research approaches used to complete the project. The following section details the DSRM phases, results, and the primary assessment methodologies and work principles for each research phase.

3.2 DESIGN SCIENCE RESEARCH METHODOLOGY (DSRM)

DSRM is a formal measure and metric for software reuse rates that has been designed and evaluated. These artefacts provide a valid and useful measure for evaluating and assessing the effectiveness and performance of software reuse activities at the organizational and project levels in development practice (Peffers et al., 2007). As shown in the Figure 3.1 this model first step is problem identification and research motivation followed by defining objectives and questions, third step is designing and development of the model then demonstration, evaluation and finally contribution and publication.

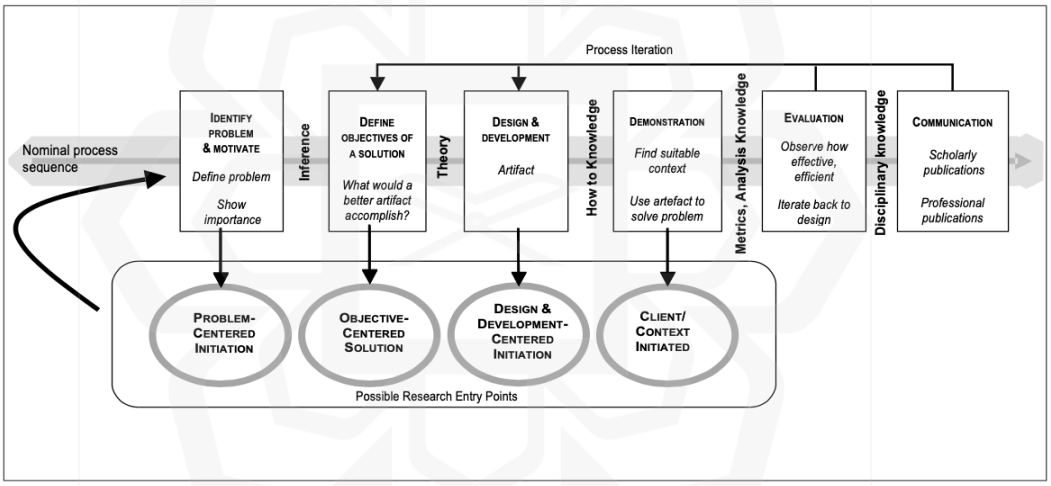


Figure 3.1 Design Science Research Methodology (DSRM) Process Model retrieved from (Peffers et al., 2007)

The design, development, implementation, and usage of an information system must support the solution of a relevant problem to be a viable research activity. The research methods are the primary working plan that allows researchers to conduct their research, and must be carefully recognised, set, and described for any research type.

In information systems, numerous research approaches can be used. The methodology, processes, and basic concepts of these methods differ. DSRM is a formal measure and metric for software reuse rates that have been previously designed and evaluated. These artefacts provide a valid and valuable measure for evaluating and assessing the effectiveness and performance of software reuse activities at the organisational and project levels in development practices (Peffers et al., 2007).

The first step in this model, per Figure 3.1, is identifying the problem and research motivation, then defining objectives and questions, then designing and developing the model, demonstrating, and evaluating the model, and finally contribution and publication. The modification of the DSRM model for adaptation into this work is illustrated in Figure 3.2. Each step is explained in the following subsections.

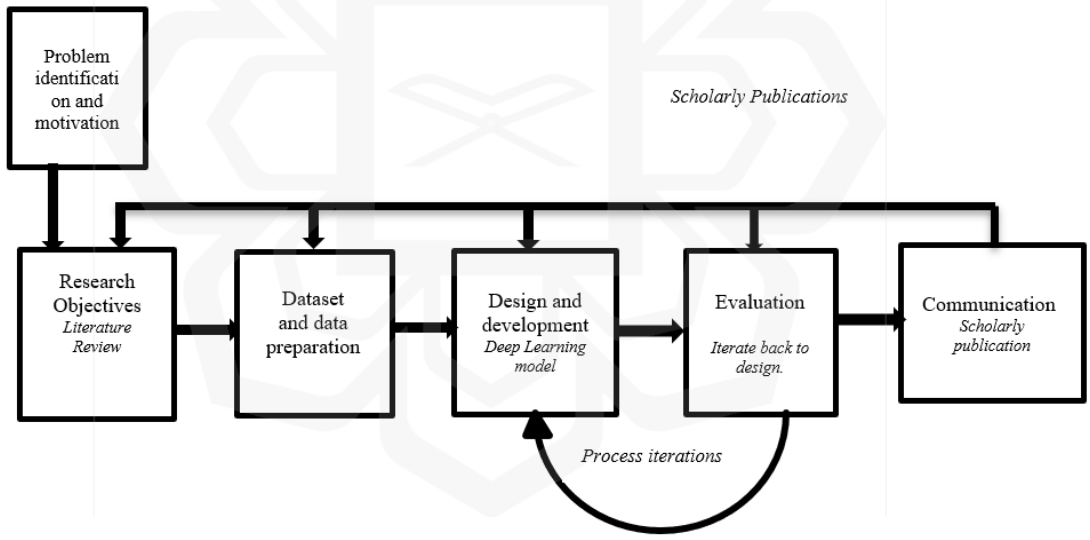


Figure 3.2 Adapted DSRM for the proposed model.

3.2.1 Problem Identification and Motivation

A brain tumour is an uncontrolled brain disease that results in aberrant brain cell groupings. MRI uses a powerful magnetic field and waves to produce accurate brain

images. The best way for scientists to discover and track the progression of a brain tumour is to use a brain MRI scan. A brain tumour can be identified in high-resolution MR imaging. However, specialists cannot verify and evaluate every single image when there are many of them, necessitating automatic categorisation and detection techniques.

Brain tumour's identification and classification are a complex undertaking in medicine, and having a reliable prediction method is frequently insufficient, as accountability, in this case, is critical and making a mistake could result in serious legal ramifications (Lecun et al., 2015). Another problem for researchers is the shortage of significant datasets (Litjens et al., 2017), which can be rectified via data preparation and augmentation. In addition, classification in medical imaging is frequently binary: normal versus abnormal. Therefore, more advanced approaches are typically required because both classes can be very similar. However, a deeper examination of the literature on brain tumour MRI medical images indicates several gaps and inadequacies in accuracy, performance, and data format discrepancies, all of which must be addressed.

AI in medical imaging has a significant influence on medical diagnosis. Previously, the procedure was done manually, but due to the large number of medical images produced daily, this approach is costly and time-consuming. Therefore, automatic classification systems for medical image processing are needed, as these systems can analyse large number of images accurately. The proposed model aims to improve classification accuracy and efficiency.

3.2.2 Research Objectives

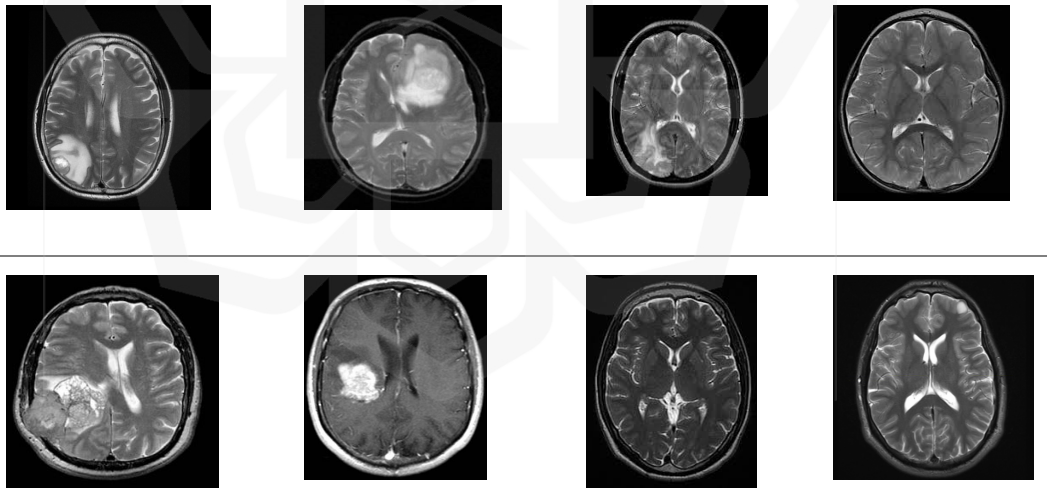
The main objective of this project is to develop a DL model capable of detecting and classifying brain tumour MRI images and elucidating existing models and algorithms for analysing brain MRI medical images and improving the classification accuracy. The developed model is expected to be helpful to pathologists.

3.2.3 Datasets and Data Preparation

The two datasets used for the purpose of brain tumor MRI medical images classification is presented in this chapter. First dataset is for binary classification (tumorous vs non-tumorous) and the second dataset is used for multi-class classification for three different types of brain tumors.

3.2.3.1 The binary dataset

The dataset has been downloaded from the Kaggle website (Chakrabarty, 2019), which contains 253 tumorous and non-tumorous MRI brain images. The tumorous folder contains 155 images in a folder called “yes”, while the non-tumorous images contain 98 images in a folder called “no”. Some of these images are shown in Figure 3.3.

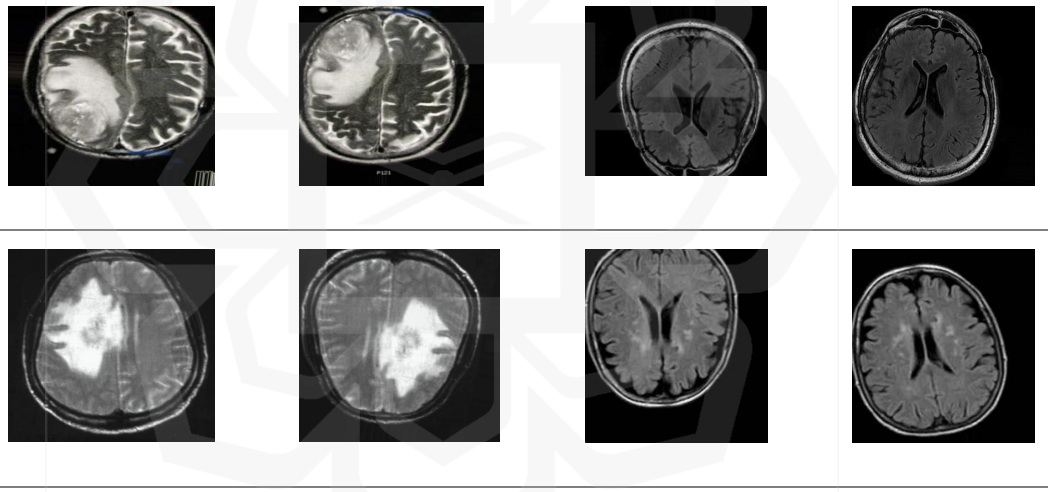


(a) Tumorous MRI-images **(b) Non-tumorous MRI-images**

Figure 3.3 Brain MRI images from the dataset

3.2.3.1.1 Data Augmentation

A large dataset is required for the DL model to succeed. The model's efficiency can be improved by increasing the amount of data available. Researchers can use data augmentation to improve the range of data available to the learning models without actively acquiring more data. The tumorous images have been augmented seven times, resulting in 1085 images, while the non-tumorous images have been augmented ten times, resulting in 980 images. The reason that non-tumorous images have been augmented more is to balance the dataset. The total number of the augmented dataset is 2065 images. Some of the used operations include rotation, shifting, horizontal, and vertical flips. Figure 3.4 show examples of the augmented dataset.



(a) Augmented tumorous MRI brain images **(b)Augmented non-tumorous MRI brain images**

Figure 3.4 Brain MRI images from the augmented dataset

3.2.3.2 The multiclass dataset

This dataset contains 3064 T1-weighted contrast-enhanced (CE) MRI medical images from 233 patients with three different forms of brain tumours. The 3064 slices in this collection include meningiomas (708 slices), gliomas (1426 pieces), and pituitary

tumours (930 slices) (sagittal, coronal, and axial). Figure 3.5 and Table 3.1 shows details of tumour images. The dataset is open to researchers and developers (Cheng et al., 2015). It was collected from Tianjing Medical University between 2005-2012, Nanfang Hospital, General Hospital, and Guangzhou in China.

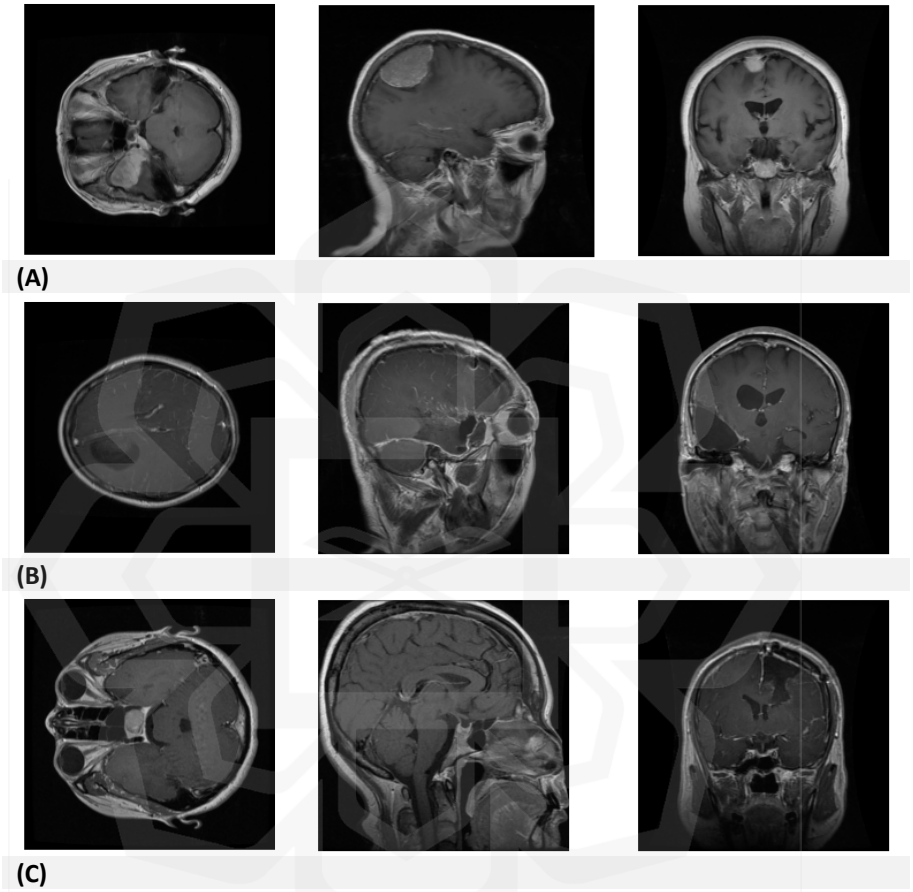


Figure 3.5 Sample of the dataset's images(A) Meningioma for (Axial, Sagittal, and Coronal) views, (B) Glioma for (Axial, Sagittal, and Coronal) views, and (C) Pituitary for (Axial, Sagittal, and Coronal) views

Table 3.1 Details of the dataset

Tumor type	Number of subjects	Number of MRI	MRI view	Number of each view
Glioma	89	1426	Axial Coronal	494 437

			Sagittal	495
Meningioma	82	708	Axial Coronal Sagittal	209 268 231
Pituitary	62	930	Axial Coronal Sagittal	291 319 320
Total	233	3064		3064

3.2.3.2.1 Data pre-processing and preparation

The dataset is organised into four folders, each with 766 files. The data is stored in the MATLAB format (.mat file). Each picture file is labelled as follows:

1 indicates meningioma, 2 indicates glioma, and 3 shows a pituitary tumour in cjdata.label.

cjdata.PID is the patient's unique identifier.

image data shape (512, 512) pixels in cjdata.image.

Inputs to the model are MRI images which are down sampled to 256 × 256 from

512 × 512, to reduce the number of parameters in the model and decrease the training time.

Border: a vector that stores the coordinates of discrete spots along the tumour's edge.

cjdata.tumor is a binary image with ones denoting the tumour location is used as a mask.

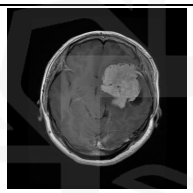
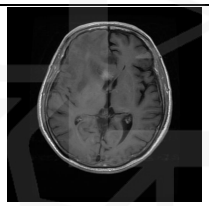
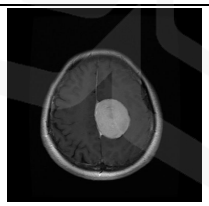
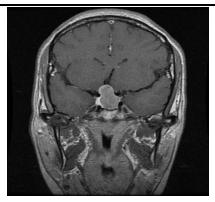
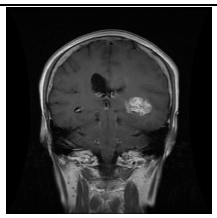
Depending on the label, the image data has been extracted to new folders as (.jpg) files. As a result, we had three directories (meningioma, glioma, and pituitary).

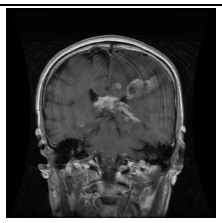
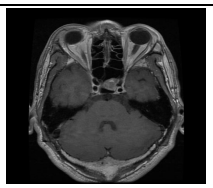
This dataset considers big enough to be tested by Deep Learning models and no need for data augmentation.

3.2.3.2.2 One-Hot encoding

One-hot encoding is a technique for converting categorical information into a format that can be fed into Machine Learning algorithms to improve prediction accuracy. In Machine Learning algorithms when supplying prediction values for different groups example is categorical labelling (meningioma, glioma, and pituitary), we convert them into integer values. This type of encoding converts the data into a form that is acceptable by Machine Learning algorithms as illustrated in the following Table 3.2.

Table 3.2 Integer values assigned to the types of brain tumor.

Samples	Image	Type of tumor	Integer value assigned
1		Meningioma	1
2		Glioma	2
3		Meningioma	1
4		Pituitary	3
5		Glioma	2

6		Glioma	2
7		Pituitary	3

To convert this integer data to one-hot encoding, Boolean column is created for each category and only one of these columns can take the value of (1) for each sample hence the term (one-hot encoding) as reflected in Table 3.3

Table 3.3 One-hot encoding example for different types of brain tumor

Sample	Meningioma	Glioma	Pituitary
1	1	0	0
2	0	1	0
3	1	0	0
4	0	0	1
5	0	1	0
6	0	1	0
7	0	0	1

3.2.3.2.3 Data Augmentation

The dataset has been augmented to its triple size. Glioma, pituitary and meningioma datasets has been augmented 3 times resulting on (4278) for glioma tumor, (2790)

images for pituitary tumor and (2124) images for meningioma tumor; the new dataset total number of images is (9192) images as shown in Figure 3.6. The operation was randomly selected from the following code list:

```
datagen = ImageDataGenerator(rotation_range=10,
                             width_shift_range=0.1,
                             height_shift_range=0.1,
                             shear_range=0.1,
                             brightness_range=(0.3, 1.0),
                             horizontal_flip=True,
                             vertical_flip=True,
                             fill_mode='nearest')
```

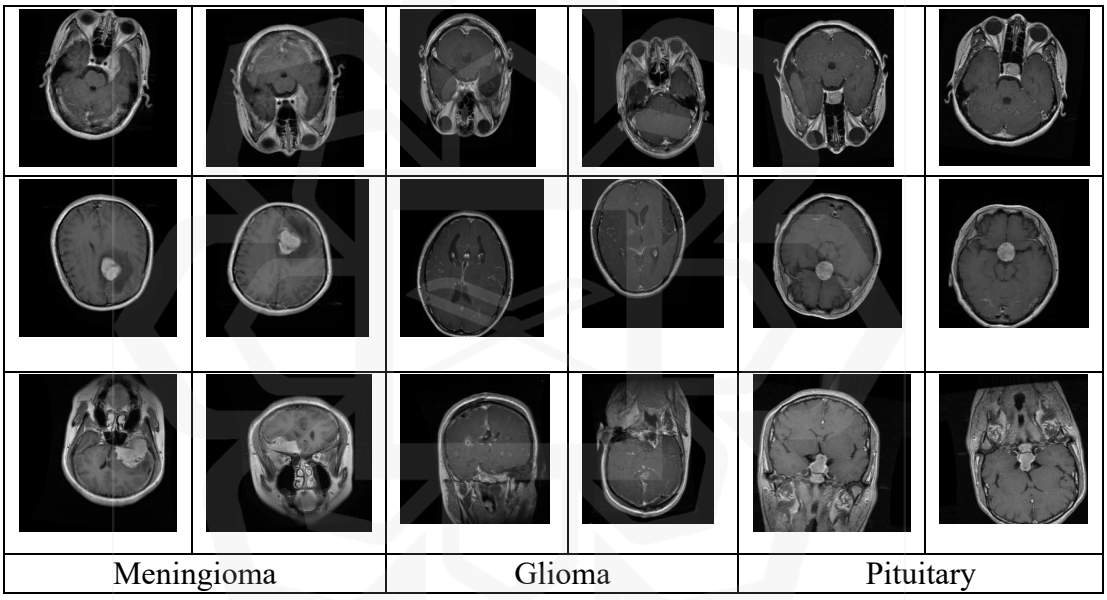


Figure 3.6 Examples from the augmented dataset

3.2.4 Design and Development

The computability and consistency of DL methods with many traditional techniques render its combination significantly advantageous in terms of innovation. The application of DL approaches for analysing MRI brain medical images has been substantially researched and reported in the literature. At the same time, Convolutional Neural Networks (CNNs)-based Deep Learning techniques are more applicable to all

fields of computer vision research, including segmentation, identification, and classification. The CNN's design for a specific task considers the work's under-representation, criteria, and the optimum utilisation of the model's computational and resource use strategy. CNN's early designs relied on a simple layering system. More advanced CNN designs are increasingly sophisticated, and each evolution updates its developments by relying on prior technological concepts and experiences. The proposed CNN model is described in chapter 4. The architecture of CNNs shows distinct preferences, and the layers and their respective preferences are described next:

3.2.4.1 The convolutional layer

Convolutional layers are the fundamental components of CNN; they use tiny input data to learn features by preserving pixel relationships (Trakoolwilaiwan et al., 2017). It is a mathematical process with two inputs (the kernel applied on an image matrix). It is the first layer (the kernel) of image processing, which captures the functionality of an input vector. Convolution learns visual features by retaining the link between pixels using tiny squares of input data (Trakoolwilaiwan et al., 2017). It is a two-input mathematical function with an image matrix and a filter/kernel as illustrated in Figure 3.7. A convolution layer is a method for extracting high-level features from an input image, and its use should not be limited to a single layer. Low-level features such as colour, borders, gradients, and orientation, are collected in the first Conv-Layer. As more layers are added, the design adapts to the needs of the high-level features.

ConvNet's reduce the images to a more usable format without compromising the essential pixels required to make accurate predictions, which is crucial for designing a network for learning and scalable to large datasets (S. H. Wang et al., 2018) .

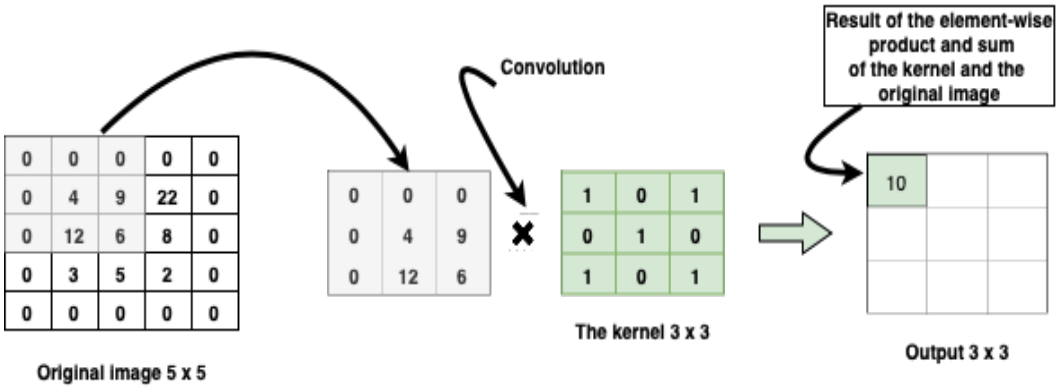


Figure 3.7 Conv-layer example of input (3 x 3) and kernel (3 x 3)

3.2.4.2 Activation function

A non-linear activation function is used to process neurons during the activation process. The typical shape of function activation is a sigmoid, as shown below:

$$f_{sigmoid} = \frac{1}{1+exp^{(-x)}} \qquad (1)$$

Saturation of the sigmoid function generated unsatisfactory results when training the network. The Rectified Linear Unit (ReLU) was populated to address this problem since the stochastic gradient curve convergence accelerated relative to the sigmoid function (Y. D. Zhang et al., 2018). We can also use an activation map with a zero threshold. The following is the definition of (ReLU):

$$f_{relu} = max(0, x) \qquad (2)$$

As shown in Figure 3.8, there are various activation functions; however, we will focus on Rectified Linear Units (ReLU) in this research. The ReLU function is the most extensively utilised activation function in recent neural networks (S. Wang, Jiang, Hou, Cheng, & Du, 2017).

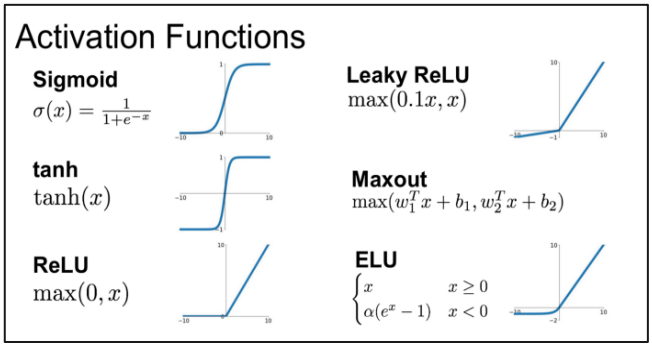


Figure 3.8 Activation functions illustration

Softmax activation function is preferred for multi-classification problem and there is evidence that proof its better performance in this type of problems (Paul, Plassard, Landman, & Fabbri, 2017; M Sajjad et al., 2019) while in binary classification the sigmoid activation function is preferred (Zyad et al., 2019).

3.2.4.3 Pooling Layer

In this layer, each image's dimensionality is reduced while critical information is preserved (Y. D. Zhang et al., 2018). For example, Max Pooling returns the maximum value of the part of the image projected by the kernel. Another example of the pooling layer is average pooling, which returns the average value.

The pooling layers decrease the number of pixels of the input image. Spatial pooling, also known as sub-sampling/sampling, reduces the dimension of the feature while preserving crucial data. The pooling layer, like the convolution layer, is essential for the dimension of the convolved features. Decreasing dimension reduces the number of computational resources needed to handle input (S. Wang et al., 2017). Max Pooling (Hang & Aono, 2017) outputs a maximum value of a portion of the image. On the other hand, Average Pooling is the cumulative total of values in the Kernel part of the image, as seen in Figure 3.9.

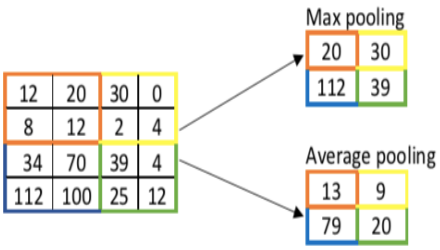


Figure 3.9 Example of Pooling layer

3.2.4.4 Batch Normalisation

The main goal of this layer to normalise the output of the previous layers. Batch normalisation enhances learning efficiency and can also be used to prevent model overfitting as a regularisation strategy.

3.2.4.5 Flatten layer

In a flatten layer, data are converted from a 2-dimensional (2-D) array to flatten the one-dimension (1-D) array connected to the final classification layers (fully connected layers) as shown in Figure 3.10.

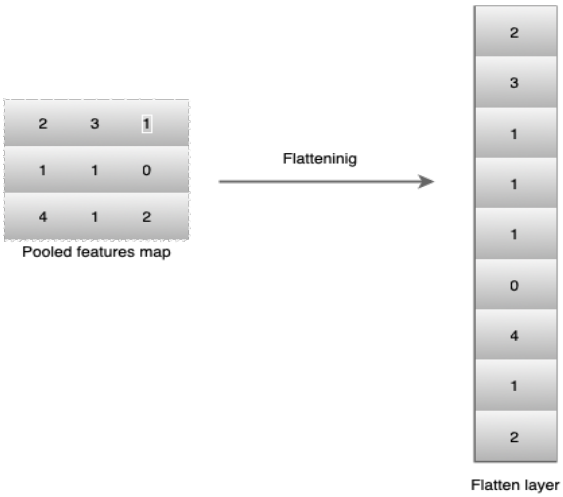


Figure 3.10 Example of flatten layer

3.2.4.6 Optimisation

In the BTMIC-CNN model we utilise the optimiser (Adam). It is ideal for a wide range of data and/or parameters because its implementation is straightforward and effective in processing and memory requirements. To decrease the loss in deep neural networks, there are several optimization strategies that include altering parameters such as weights and learning rates. Diederik Kingma (Kingma & Ba, 2015) introduced the Adaptive Moment Estimation (Adam) optimizer. RMSprop and Stochastic Gradient Descent with momentum are combined in the Adam optimizer. As described in the following equations for updated weights and bias in Adam:

$$W = W - \eta \times (V_{dw}/\sqrt{S_{dw}+\epsilon}) \quad (3)$$

$$b = b - \eta \times (V_{db}/\sqrt{S_{db}+\epsilon}) \quad (4)$$

W and b are the derivative of weight and bias respectively as per the following equations:

$$W = W - \eta \times dW \quad (5)$$

$$b = b - \eta \times db \quad (6)$$

Epsilon (ϵ) is a number to prevents zero division (ϵ = 10-8).

(η) is a learning rate with a different range of values.

V is the moving mean of our gradients as illustrated bellow:

$$V_{dw} = \beta \times V_{dw} + (1 - \beta) \times dW \quad (7)$$

$$V_{db} = \beta \times V_{db} + (1 - \beta) \times db \quad (8)$$

In RMSProp (Wong & Shen, 1995), S is the exponential moving mean square of gradients is calculated as per RMSProp equations:

$$S_{dw} = \beta \times S_{dw} + (1 - \beta) \times dW^2 \quad (9)$$

$$S_{db} = \beta \times S_{db} + (1 - \beta) \times db^2 \quad (10)$$

The hyperparameter β is used to adjust exponentially weighted means (H. A. Khan, Jue, Mushtaq, & Mushtaq, 2020).

3.2.4.7 Dropout layer

Dropout is a layer that drops out some neurons (setting to zero) to the following layer while leaving all others unchanged. A Dropout layer can sometimes be applied to the training dataset, dropping out some of its properties; however, it can also be applied to a dense layer, dropping out certain hidden neurons. Dropout layers are critical in CNN learning since they minimize overfitting that can occur in the training data The dropout rate is the percentage of features that are wiped out; it is often set between 0.2 and 0.5. (Chollet, 2017). On the other hand, the extensive use of dropout layers in the network can leads to losing out some important feature. Figure 3.11illustrates the dynamic of this layer.

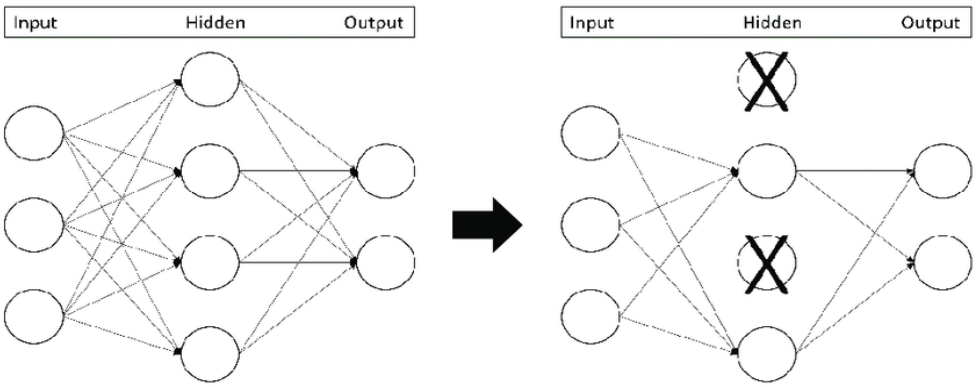


Figure 3.11 Example of dropout layer

3.2.4.8 Connected Layer (FC Layer)

Fully connected or densely connected are used interchangeably. Every input of the input vector affects every output of the output vector, implying that all possible connections layer-to-layer are present. Neurons in these layers have complete links to all activations in the previous layer. The activations of these neurons can be calculated by multiplying the matrix with an offset bias.

3.2.5 Evaluation

Multiclass classification findings are generally more difficult to comprehend than binary classification results. The confusion matrix and the classification report, including (F1-score, precision, and recall), are two typical techniques for performance indicators, in addition to accuracy (Pereira et al., 2016).

$$F1_{score} = 2\ x\ \frac{(Precision\ x\ Recall)}{(Precision+Recall)} \qquad (11)$$

Where:

$$Precision = \frac{(TP)}{(TP+FP)} \qquad (12)$$

$$Recall = \frac{(TP)}{(TP+TN)} \quad (13)$$

The details of the abbreviation in the equations are as follows: true positive (TP) represents the number of instances classified correctly as defected, false positive (FP) represents the number of cases classified incorrectly as defected, and true negative (TN) represents the number of instances classified correctly as non-defected (Anwar et al., 2018). Loss curves are essential to show the model performance throughout the training iterations. The curves' shapes vary for different models. Listed below are some terms describing various curves for better understanding (Chollet, 2017).

i. Good fit happens when training and validation are converging.
ii. Underfitting happens when validation and training error high model cannot learn from the training data.
iii. Overfitting happens when the validation error is higher than the training error model and is not generalised to work for a real-life dataset.
iv. Uncertainty. The test set's uncertainty can be relatively high, to the point that various test sets can produce significantly different results.
v. Dropout curve (testing accuracy is more than training accuracy). Although, to some extent is not logical and may refer to an unbalanced dataset. Sometimes, it is considered acceptable because the model is more robust during testing.
vi. Test sets can help unbiased estimate of generalisation performance.
vii. Generalisation Models are ready for another dataset if the accuracy is good.

3.2.6 Contribution

The conclusion and communication activities are the last part of any design scientific research project (Peffers et al., 2007). Contribution and communication activities must conclude any design scientific research. The primary goal of the contributing activity is to exhibit the research findings as will be explained in the next chapters. Whereas the communication goal is to illustrate the relevance of the solved problem and its influence on the research community via the publication of numerous academic articles.

3.3 CHAPTER SUMMARY

The conducted research methodologies were described and analysed in detail in this chapter. The DSRM was employed in this study because it provides a widely accepted method for doing effective design science research in the domain of information systems. The technique offered is made up of process models, which are divided into five parts. This chapter looked at how those phases were adapted and used in this project. In addition, this chapter shows how to use Deep Learning algorithms to create a framework for automatic medical image classification. The fundamental goal is to provide a model for image classification that is both efficient and accurate.

CHAPTER FOUR

BRAIN TUMOR MRI MEDICAL IMAGES CLASSIFICATION ALGORITHM IMPLEMENTATION BASED ON CNN (BTMIC-CNN) ANALYSIS

4.1 INTRODUCTION

This chapter explains the proposed model for brain tumor MRI images classification including binary and multiclass classification. The proposed model is based on DCNN architectures where it will be able to identify brain tumor based on MRI medical images.

4.2 THE PROPOSED MODEL (BTMIC-CNN)

The main objective is to propose the design of fully automatic CNN models for brain tumour MRI images classification (BTMIC-CNN). Basically, any CNN architectures consist of convolutions, pooling, dense layers. Other layers such as normalization, dropout can be added to optimize the network and get the best classification performance as will be detailed in the following sections. Design and development followed by evaluation in the DSRM model is the main key to optimize BTMIC-CNN model.

4.2.1 Experimental set-up

The potential experimental set - up for classifications is presented in this section. Due to the availability of the most common Machine Learning and neural network modules, the hybrid model was created in Python 3.6.9. Tensorflow and Keras libraries were utilized in this research, and it can be executed on both the CPU and the GPU (GPU). Google Colaboratory (also known as Colab) is a cloud service based on Jupyter Notebooks for teaching and researching Machine Learning (Carneiro et al., 2018). It comes with 12.7 GB of RAM, 33 GB of hard disc, a Deep Learning runtime, and free

access to a powerful GPU. Nvidia K80, T4s, P4s, and P100 GPUs are frequently accessible at Colab. Table 4.1 illustrate the details of the GPU used and the memory allocated for the experiments.

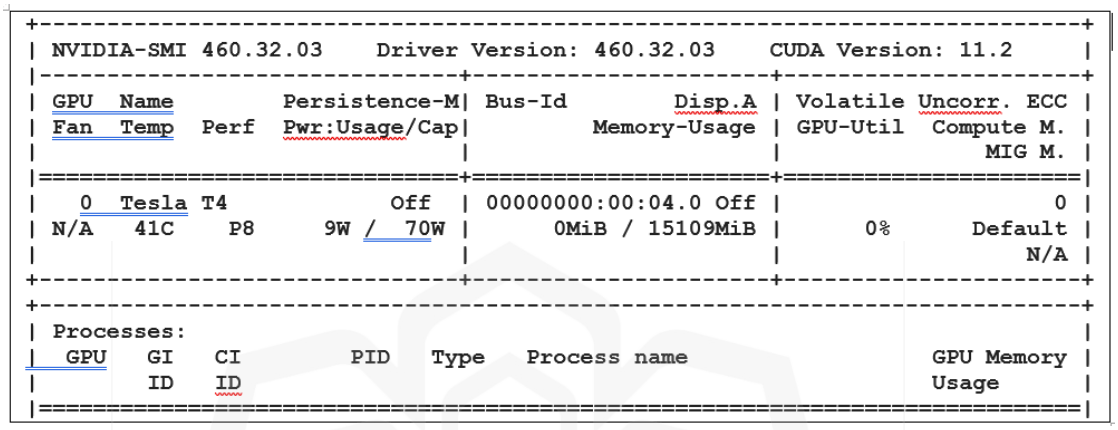


Figure 4.1 Details of the GPU used.

The BTMIC-CNN model is designed for two main classification tasks (Multiclass BTMIC-CNN and Binary BTMIC-CNN) as will be explained in the following sections:

4.2.2 BTMIC-CNN Network Architecture

The network architecture of the proposed model (BTMIC-CNN) for binary task is shown in Figure 4.1 and for multiclass task is shown in Figure 4.2. The model consists of ten main blocks (Input layer, six Convolutional layers, two dense layers and output layers). The first layer is input layer which contains the resized image 256 x 256. All convolutional layers are followed by activation function (RELU). Followed by max-pooling layers (2 x 2) which helped to reduce the input image size by half for each block. Some convolutional blocks are followed by normalization layer as the following order: (1 CONV layer- Normalization layer, 2 CONV layers -Normalization layer, 2

CONV layers -Normalization layer, 1 CONV layer- Normalization layer). We did not include the dropout layers in the convolutional layers as we want to keep all the important features to be entered to the fully connected layers. The flatten layer is the transformation layer between convolutional blocks and the fully connected layers. It transforms the 2D array to 1D flatten array that is ready for the dense layers.

The model consists of two dense layers in which the first one is followed by a dropout layer to help reducing the overfitting and optimize the network. The second dense layer is connected to the output layer. Sigmoid was chosen to be the activation function for the binary task, while SoftMax activation function was implemented for the multiclass classification task. Ultimately, the design of the architecture of any deep neural network will go of trial and error in which we call (design and evaluation) in DSRM model. Here are some specific points that have been applied for the proposed BTMIC-CNN model.

i. Pre-processed and clean data gives meaningful result.
ii. Convolutions layers Are followed by normalization layers to optimize the network.
iii. Dense layers are followed by dropout layers to reduce the overfitting.
iv. Small kernel size to ease computation process.

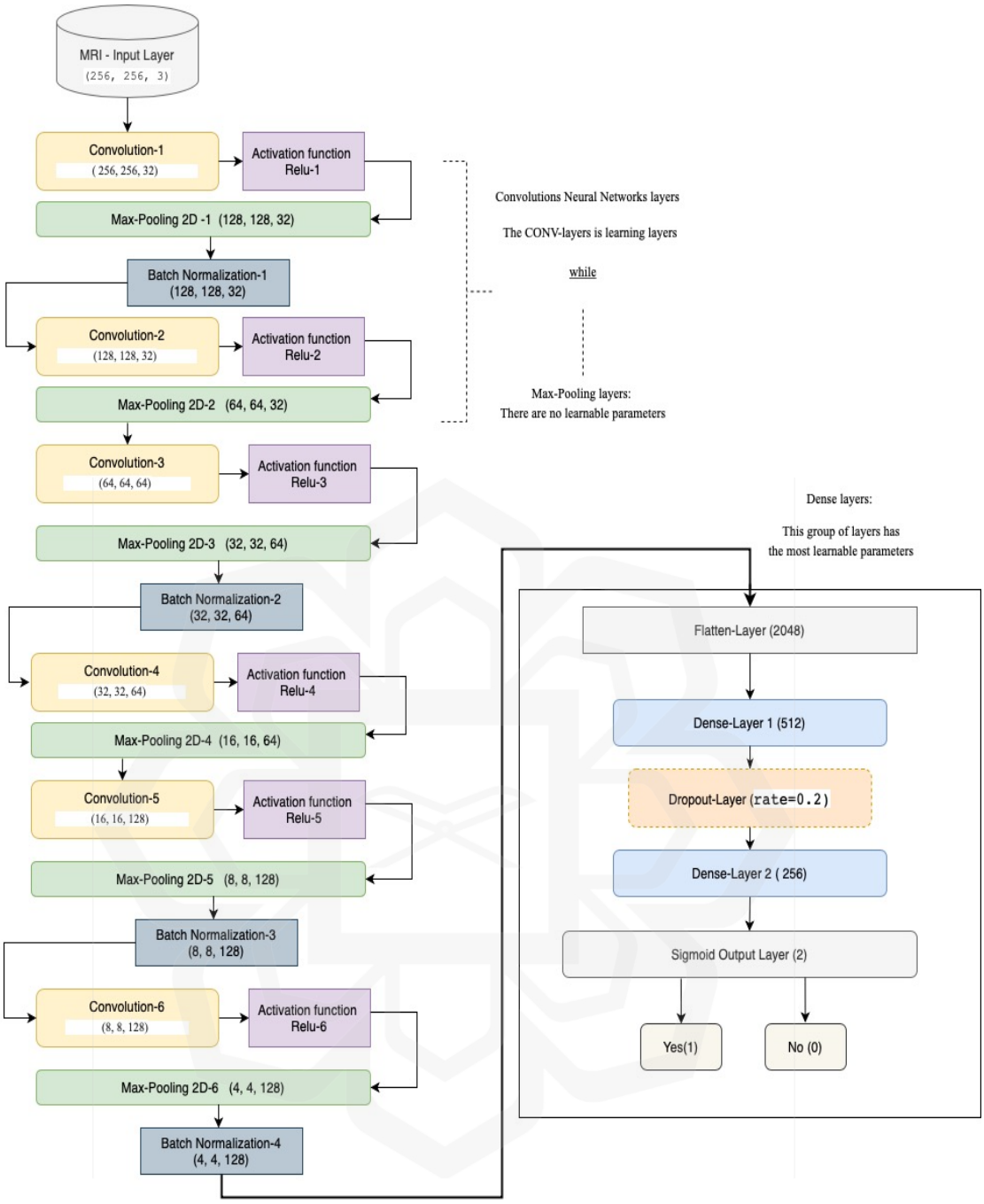


Figure 4.2 The proposed model (BTMIC-CNN) architectures for binary dataset

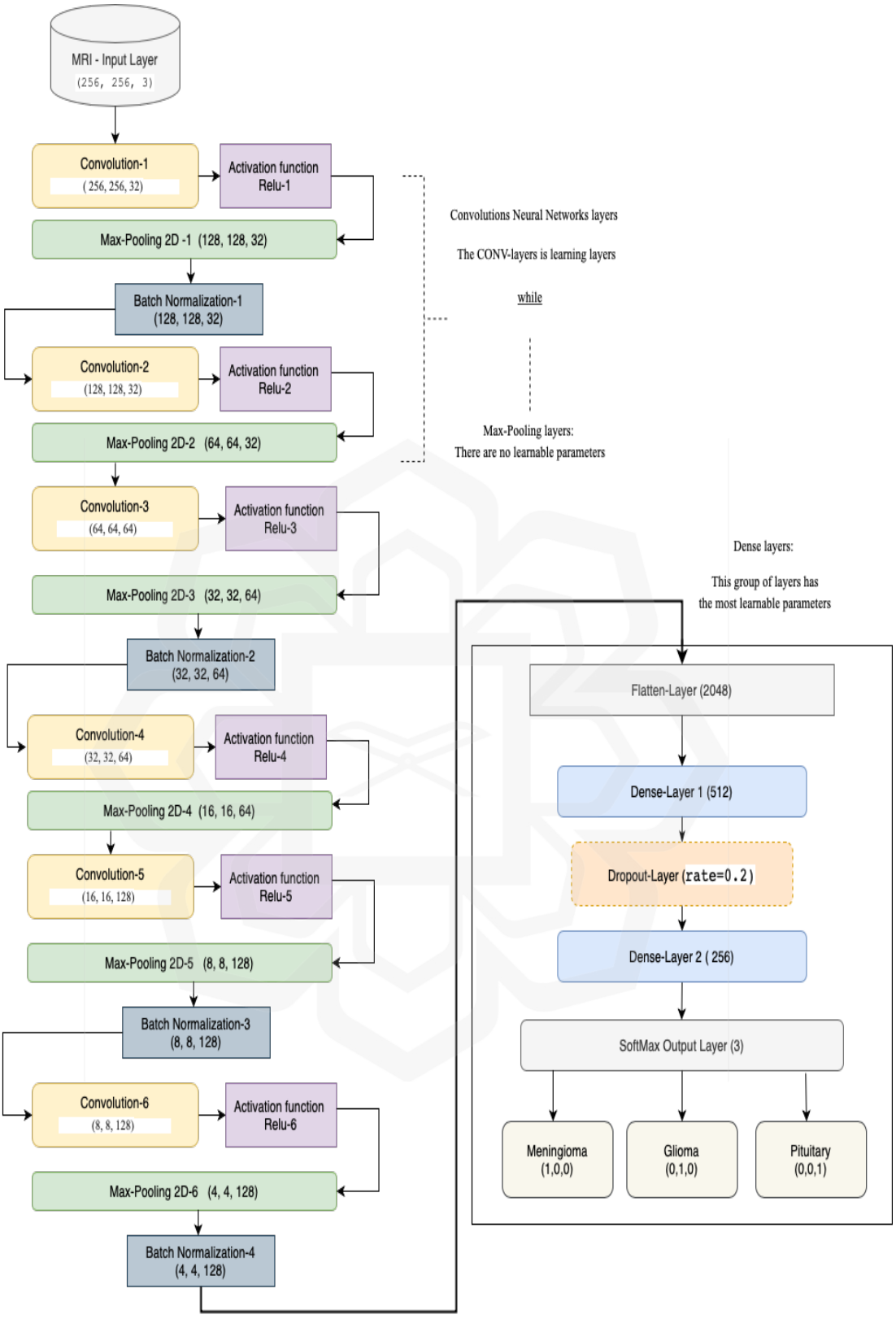


Figure 4.3 The proposed model (BTMIC-CNN) architectures for multiclass dataset

4.2.3 Design of the experiments

Figure 4.3 and Figure 4.4 show that the BTMIC-CNN model have been applied for the datasets with and without data augmentation. The dataset has been augmented for two reasons; first, when the dataset comprises an equal number of samples from each class, classification accuracy is an effective measure for characterising performance. However, the dataset used in this paper is uneven, which means that it needs more thorough assessment of the proposed system using additional performance indicators. Therefore, we used data augmentation techniques in order to balance the dataset. The second reason is that DL neural network performs better in a larger dataset as explained in chapter 3.

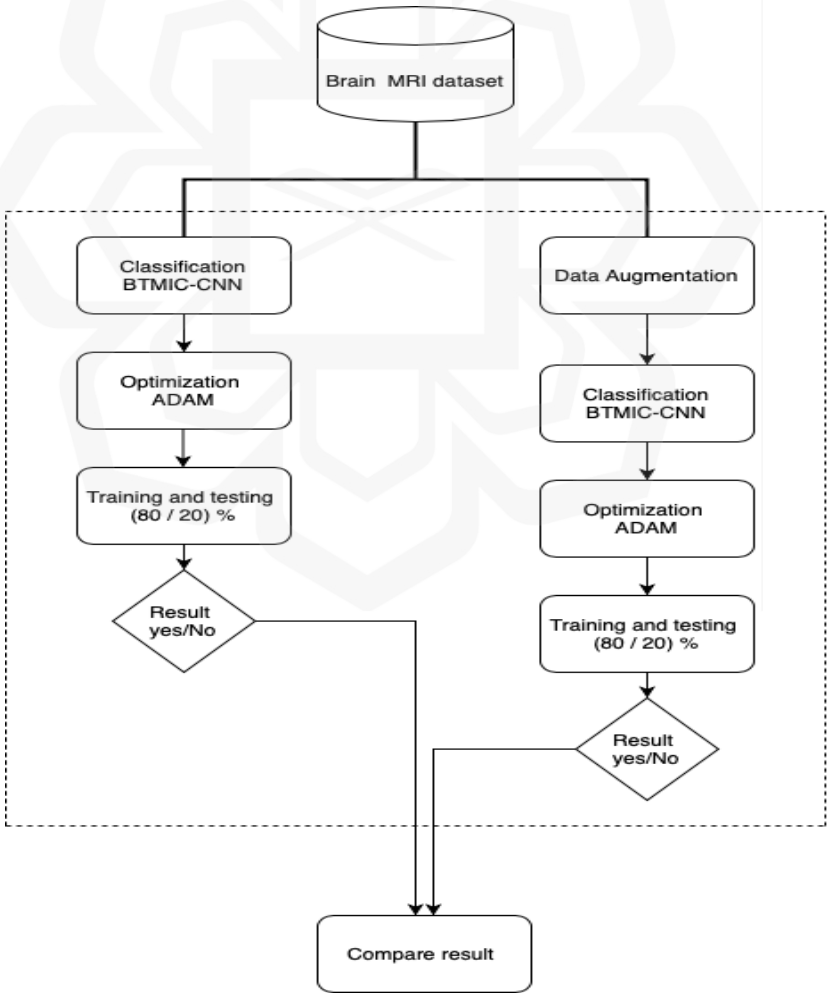


Figure 4.4 Flowchart of the BTMIC-CNN on binary task

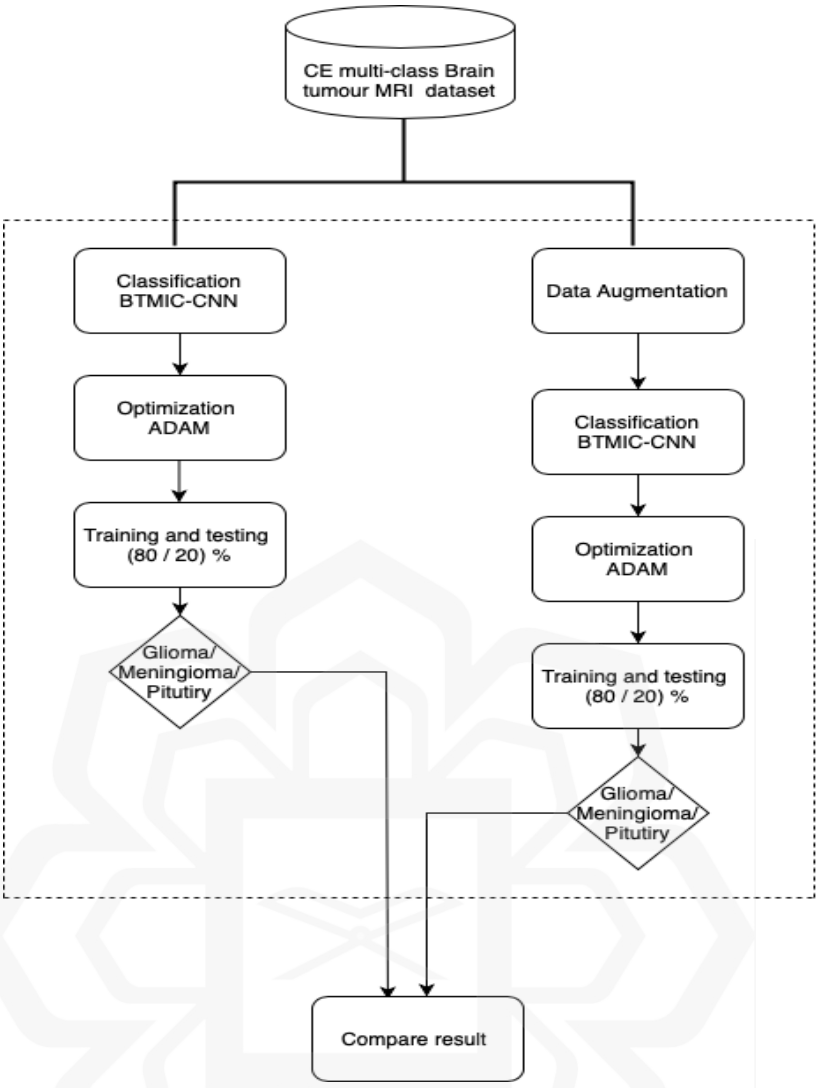


Figure 4.5 Flowchart of the BTMIC-CNN on multiclass task

4.2.4 Learning parameters calculation

The BTMIC-CNN model has been set to classify different types of tumours. The CONV layer is learning layers. Therefore, weight matrices will undoubtedly be present. Simply multiply the form of length L, width W, the preceding layer's filter F, and accounting for all these filter's T in the current layer to obtain the learnable parameters. For every filter, bias factor must be calculated. So, the expression to calculate number of the parameters is **((L * W * F) +1) * T).**

Input layer: The input layer has nothing to learn; all it does is supply the form of the input image. As a result, there are no parameters that can be learned. So, the number of parameters equals zero.

Pooling layers: There are no learnable parameters in this layer since it only calculates a single number.

Dense layers: This group of layers has the most learnable parameters when compared to the other layers. Because each neuron is linked to all other neurons.

The output layer of the second CNN model consists of three neurons because it is developed to classify images into one of three categories. Finally, the softmax classifier receives the fully connected layer, a 3D vector, and generates the tumour type prediction. Table 4.2 and Table 4.3 illustrates the architecture and design of the BTMIC-CNN model for both tasks. The last column shows the number of parameters in each layer.

Table 4.1 Details of BTMIC-CNN proposed model for binary dataset

	CNN-Layer	Shape	Activation Size	Total parameters
1	Input-Layer	256 x 256 x 3	196608	0
2	Conv2d-1	256 x 256 x 32	2097152	896
3	MaxPooling 2D-1	128 x 128 x 32	524288	0
4	BatchNormalization-1	128 x 128 x 32	524288	128
5	Conv2d-2	128 x 128 x 32	524288	9248
6	MaxPooling 2D-2	64 x 64 x 32	131072	0
7	Conv2d-3	64 x 64 x 64	262144	18496
8	MaxPooling 2D-3	32 x 32 x 64	65536	0
9	BatchNormalization-2	32 x 32 x 64	65536	256
10	Conv2d-4	32 x 32 x 64	65536	36928
11	MaxPooling 2D-4	16 x 16 x 64	16384	0
12	Conv2d-5	16 x 16 x 128	32768	73856

13	MaxPooling 2D-5	8 x 8 x 128	8192	0
14	BatchNormalization-3	8 x 8 x 128	8192	512
15	Conv2d-6	8 x 8x 128	8192	147584
16	MaxPooling 2D-6	4x 4 x 128	2048	0
17	BatchNormalization-4	4 x 4 x 128	2048	512
18	Flatten	2048	2048	0
19	Dense-layer 1	512	512	1049088
20	Dropout	512	512	0
21	Dense-layer 1	256	256	131328
22	Output-layer (Sigmoid)	2	2	514
Total params: 1,469,346 **Trainable params: 1,468,642** **Non-trainable params: 704**				

Table 4.2 Details of BTMIC-CNN proposed model for multiclass dataset

	CNN-Layer	Shape	Activation Size	Total parameters
1	Input-Layer	256 x 256 x 3	196608	0
2	Conv2d-1	256 x 256 x 32	2097152	896
3	MaxPooling 2D-1	128 x 128 x 32	524288	0
4	BatchNormalization-1	128 x 128 x 32	524288	128
5	Conv2d-2	128 x 128 x 32	524288	9248
6	MaxPooling 2D-2	64 x 64 x 32	131072	0
7	Conv2d-3	64 x 64 x 64	262144	18496
8	MaxPooling 2D-3	32 x 32 x 64	65536	0
9	BatchNormalization-2	32 x 32 x 64	65536	256
10	Conv2d-4	32 x 32 x 64	65536	36928
11	MaxPooling 2D-4	16 x 16 x 64	16384	0
12	Conv2d-5	16 x 16 x 128	32768	73856

13	MaxPooling 2D-5	8 x 8 x 128	8192	0
14	BatchNormalization-3	8 x 8 x 128	8192	512
15	Conv2d-6	8 x 8x 128	8192	147584
16	MaxPooling 2D-6	4x 4 x 128	2048	0
17	BatchNormalization-4	4 x 4 x 128	2048	512
18	Flatten	2048	2048	0
19	Dense-layer 1	512	512	1049088
20	Dropout	512	512	0
21	Dense-layer 1	256	256	131328
22	Output-layer (SoftMax)	3	3	771
Total params: 1,469,987 **Trainable params: 1,469,091** **Non-trainable params: 896**				

4.2.5 Proposed model (BTMICT-CNN) Performance Testing

The datasets for each task were divided into three sections for training (80%), validation (10%), and testing (10%). The training data is used to learn the model, while the validation data is used to evaluate the model and tune the parameters. The test data will be used to evaluate the proposed model in the end. The flowchart of the BTMIC-CNN dataset distribution in general is shown in Figure 4.5. The result and performance of the proposed model will be illustrated in the next chapter.

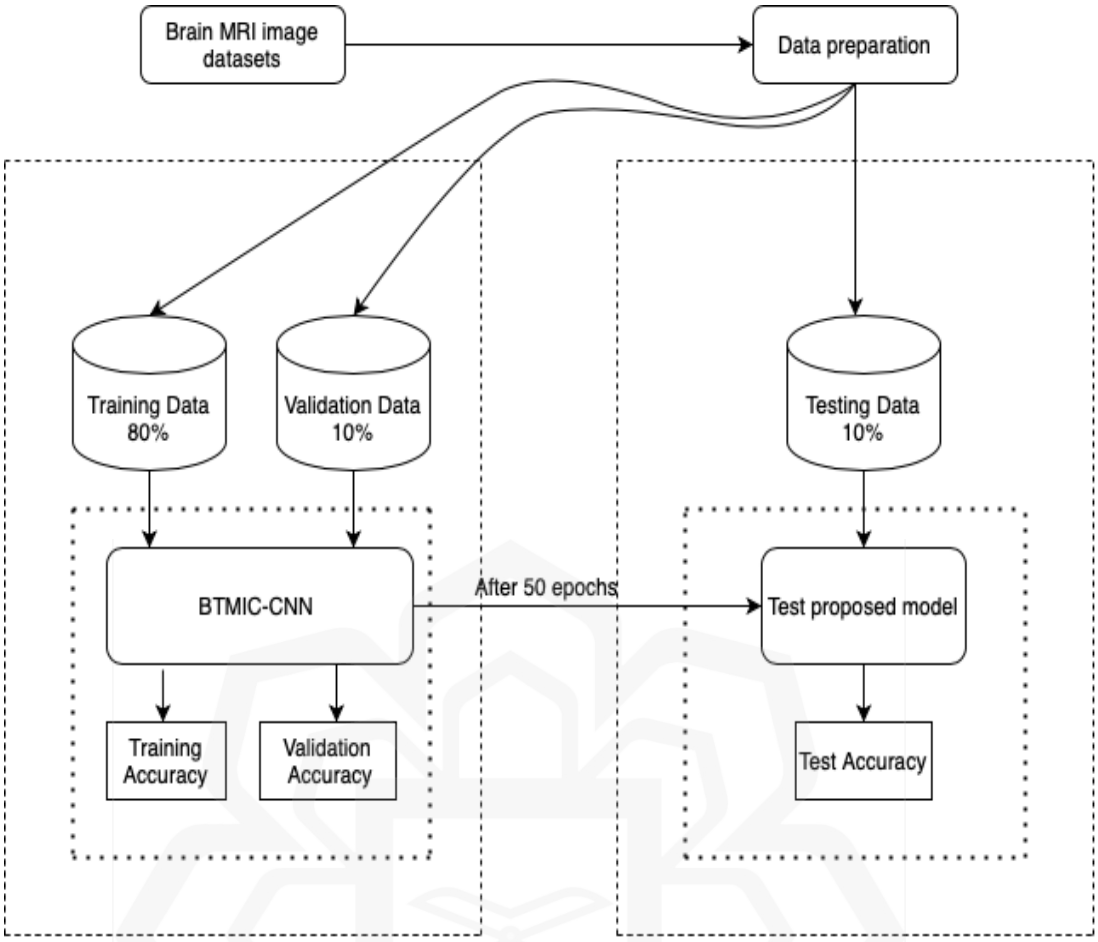


Figure 4.6 Flowchart illustrates the datasets distribution for training, validation, and testing sets for BTMIC-CNN proposed model

4.3 TRANSFER LEARNING OF VGG16

This research will investigate VGG16 on the binary and multiclass datasets. According to the literature VGG16 is one from best CNN algorithms for brain tumor MRI -medical image classification (Ahammed Muneer et al., 2019; A. R. Khan et al., 2021; H. A. Khan et al., 2020; Manikanta, Yashwanth, Deepthimurthy, & Santosh, 2021; Muhammad Sajjad et al., 2019b). Transfer learning builds on past knowledge instead of starting from scratch. We utilised a pre-trained VGG-16 Convolutional Neural Network model that was fine-tuned by freezing parts of the layers to avoid overfitting.

Transfer learning was optimized to enhance the CNN model. The goal of this approach was to reduce learning time and increase accuracy. With different node SoftMax layers and classification layers, the transfer learning approach used full connection (FC) as an image representation. The original model's other parameters were kept and utilised as initializations. Figure 4.6 and Figure 4.7 depicts the architecture of the deep transfer learning VGG16 for the binary and multi-classification task respectively.

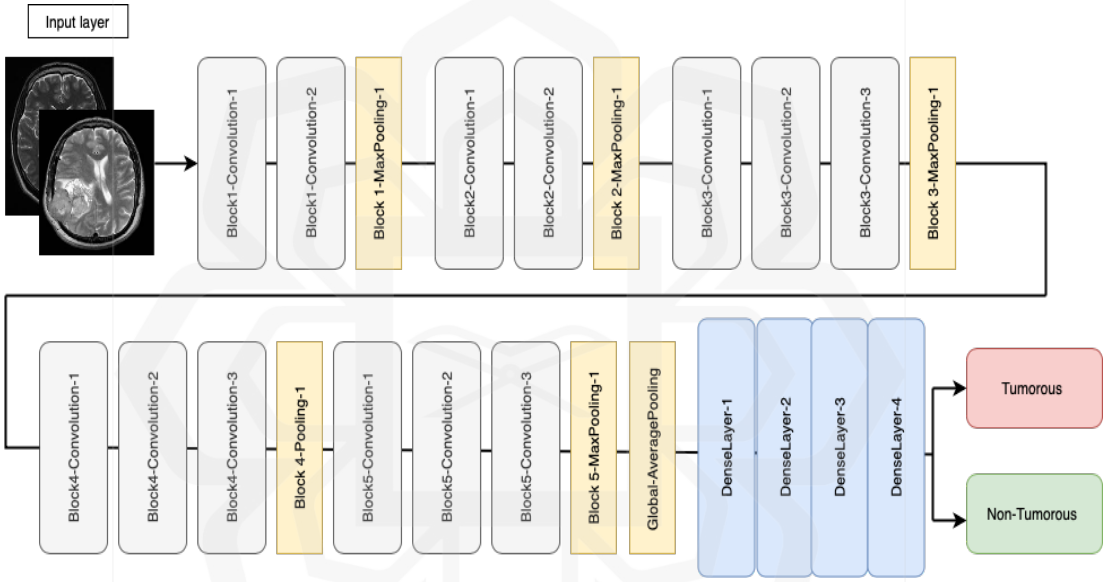


Figure 4.7 Transfer learning (VGG16) architectures for binary classification task

Table 4.4 and Table 4.5 contains the layers information as well as number of learnable parameters for each classification task. Table 4.4 is for the transfer learning of VGG16 for the binary classification task and it is reflected in Figure 4.6. Convolutional layers have the most number learning parameters in the convo blocks, while in pooling layers there is no learning parameters.

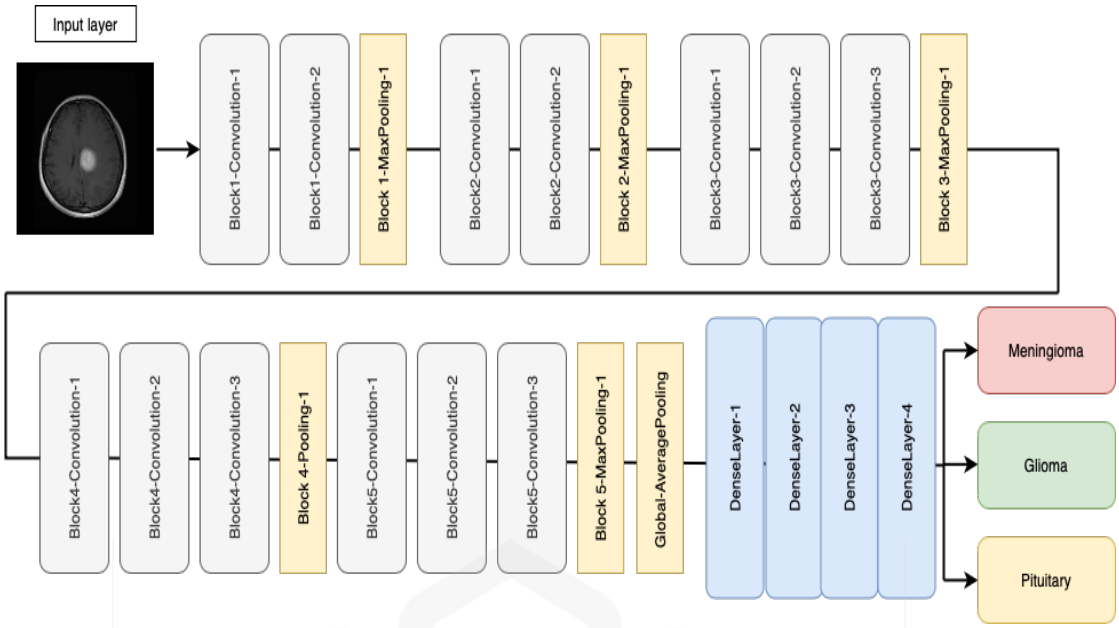


Figure 4.8 Transfer learning of VGG16 for multiclass classification task

Global average pooling is the transfer point between convolutional blocks and dense layers. The Tables 4.4 show that the greatest number of learning parameters are generated in the fully connected layers (dense layers). Total learning parameters for brain tumor MRI medical images binary classification using transfer learning of VGG16 is ***(16,815,426)*** all are trainable parameters.

Similarly in Table 4.5 which present the layer prosperities for the brain tumor MRI medical images multiclass classification task using transfer learning of VGG16. The layers of both models are same except a modification in the output layer. In the multiclass classification task softmax activation function was implemented while in the binary classification task sigmoid activation function was implemented. The loss function is also different. For the multiclass classification task categorical cross-entropy was implemented while for binary classification task binary cross-entropy was implemented. Total learning parameters for brain tumor MRI medical images multiclass classification using transfer learning of VGG16 is (16,815,939) all are trainable parameters.

Table 4.3 Layer and number of parameters in VGG16 for binary classification task

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(224 x 224 x 3)]	0
block1_conv1 (Conv2D)	(224 x 224 x 64)	1792
block1_conv2 (Conv2D)	(224 x 224 x 64)	36928
block1_pool (MaxPooling2D)	(112 x 112 x 64)	0
block2_conv1 (Conv2D)	(112 x 112 x 128)	73856
block2_conv2 (Conv2D)	(112 x 112 x 128)	147584
block2_pool (MaxPooling2D)	(56 x 56 x 128)	0
block3_conv1 (Conv2D)	(56 x 56 x 256)	295168
block3_conv2 (Conv2D)	(56 x 56 x 256)	590080
block3_conv3 (Conv2D)	(56 x 56 x 256)	590080
block3_pool (MaxPooling2D)	(28 x 28 x 256)	0
block4_conv1 (Conv2D)	(28 x 28 x 512)	1180160
block4_conv2 (Conv2D)	(28 x 28 x 512)	2359808
block4_conv3 (Conv2D)	(28 x 28 x 512)	2359808
block4_pool (MaxPooling2D)	(14 x 14 x 512)	0
block5_conv1 (Conv2D)	(14 x 14 x 512)	2359808
block5_conv2 (Conv2D)	(14 x 14 x 512)	2359808
block5_conv3 (Conv2D)	(14 x 14 x 512)	2359808
block5_pool (MaxPooling2D)	(7 x 7 x 512)	0
global_average_pooling2d(Gl)	(512)	0
dense (Dense)	(1024)	525312
dense_1 (Dense)	(1024)	1049600
dense_2 (Dense)	(512)	524800
dense_3 (Dense)	(2)	1026

Table 4.4 Transfer learning of VGG16 for the multi-class classification task

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(256 x 256 x 3)]	0
block1_conv1 (Conv2D)	(256 x 256 x 64)	1792
block1_conv2 (Conv2D)	(256 x 256 x 64)	36928
block1_pool (MaxPooling2D)	(128 x 128 x 64)	0
block2_conv1 (Conv2D)	(128 x 128 x 128)	73856
block2_conv2 (Conv2D)	(128 x 128 x 128)	147584
block2_pool (MaxPooling2D)	(64 x 64 x 128)	0
block3_conv1 (Conv2D)	(64 x 64 x 256)	295168
block3_conv2 (Conv2D)	(64 x 64 x 256)	590080
block3_conv3 (Conv2D)	(64 x 64 x 256)	590080
block3_pool (MaxPooling2D)	(32 x 32 x 256)	0
block4_conv1 (Conv2D)	(32 x 32 x 512)	1180160
block4_conv2 (Conv2D)	(32 x 32 x 512)	2359808
block4_conv3 (Conv2D)	(32 x 32 x 512)	2359808
block4_pool (MaxPooling2D)	(16 x 16 x 512)	0
block5_conv1 (Conv2D)	(16 x 16 x 512)	2359808
block5_conv2 (Conv2D)	(16 x 16 x 512)	2359808
block5_conv3 (Conv2D)	(16 x 16 x 512)	2359808
block5_pool (MaxPooling2D)	(8 x 8 x 512)	0
global_average_pooling2d (GlobalAveragePooling2D)	(512)	0
dense (Dense)	(1024)	525312
dense_1 (Dense)	(1024)	1049600
dense_2 (Dense)	(512)	524800
dense_3 (Dense)	(3)	1539

The dataset was divided into three parts: training (80%), validation (10%), and testing (10%). Validation data is used to test the model and fine-tune the parameters, while the training data is used to learn the model. Finally, the suggested model will be evaluated using the test data. Figures 4.8 and 4.9 depict the flowchart of VGG16's TL for both tasks. The Figures illustrate that the TL of VGG16 is implemented on the augmented and the original datasets without augmentation. The result will be compared as will be discussed in the next chapter.

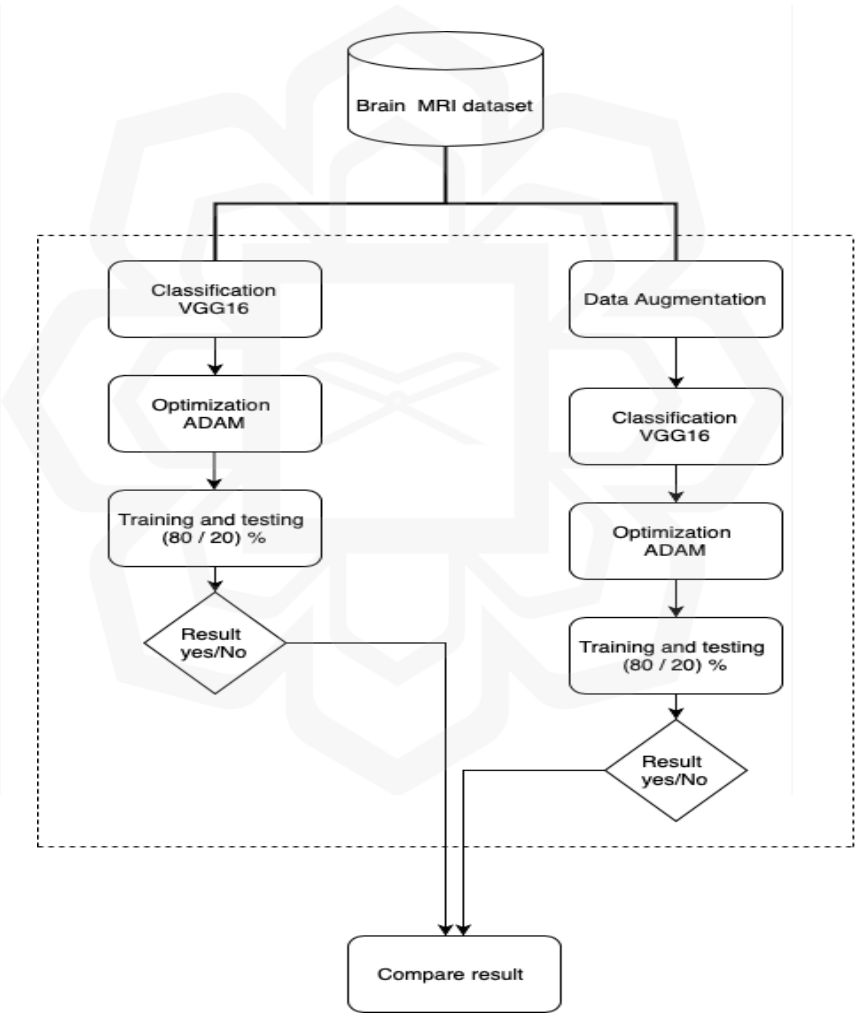


Figure 4.9 Flowchart of the TL-VGG16 on binary classification of brain MRI images

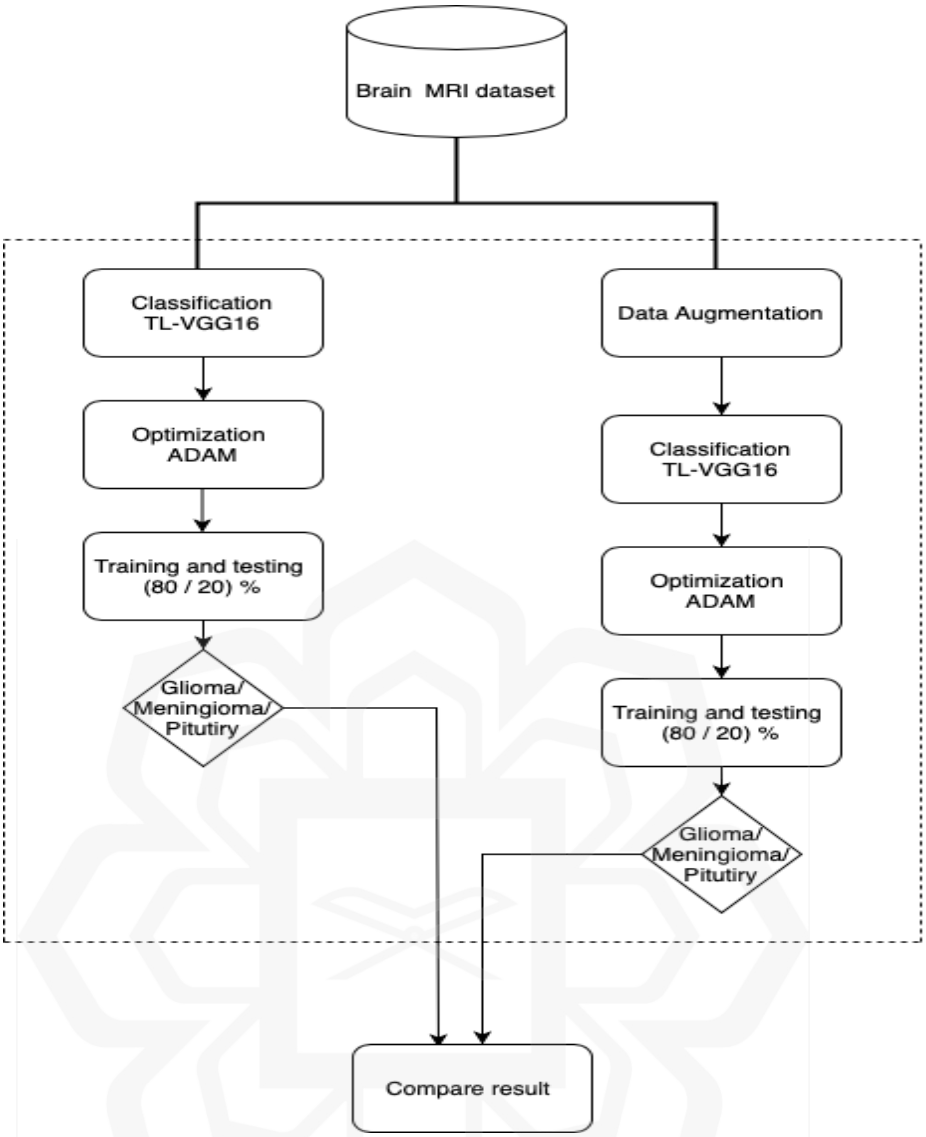


Figure 4.10 Flowchart of the TL-VGG16 on binary classification of brain MRI images

4.4 CHAPTER SUMMARY

This chapter explains the design and development of the proposed model Brian Tumor MRI medical Images Classification based on Convolution Neural Network (BTMIC-CNN). The model is designed for binary image classification(detection) containing tumorous and non-tumorous MRI medical images as well as multiclass classification for three types of brain tumor MRI medical images. We have presented also transfer learning of VGG16 for both classification tasks. Result and evaluation of the

performance proposed model (BTMICT-CNN) and transfer learning of VGG16 will be illustrated in the following chapter.

CHAPTER FIVE

RESULT AND DISCUSSION

5.1 INTRODUCTION

Throughout this study, several experiments were carried out to fulfill the objectives of the research. Both detection and multiclass classification results and evaluation based on the proposed BTMIC-CNN model and transfer learning of VGG16 are discussed in this chapter.

5.2 BINARY CLASSIFICATION

This task includes detection of the brain tumor (binary classification) from the first dataset which includes (tumorous and non-tumorous) brain MRI images. Four experiments have been performed using BTMIC-CNN and transfer learning of VGG16 for augmented and non-augmented dataset as will be explained in the upcoming sections.

5.2.1 Experiment-1 (The proposed BTMIC-CNN Model without data augmentation)

Task 1 involves the application of the proposed model (BTMIC-CNN) on the binary dataset which contains two main classes (tumorous and non-tumorous) brain tumor MRI medical images as detailed in chapter 3. The first experiments in this section applied BTMIC-CNN on the original dataset which contains only 253 MRI images without data augmentation. The dataset was split: 80 % training dataset, 10% validation dataset, and 10% test dataset. This dataset considered a small dataset especially when applying deep neural network which requires big data. The accuracy and loss curves are illustrated in Figure 5.1 for this experiment. However, the training accuracy is reaching

100%, The test accuracy is 88.57% which considered low in comparison with the upcoming experiments. The second experiment will explain how the accuracy improved after expanding the dataset with data augmentation.

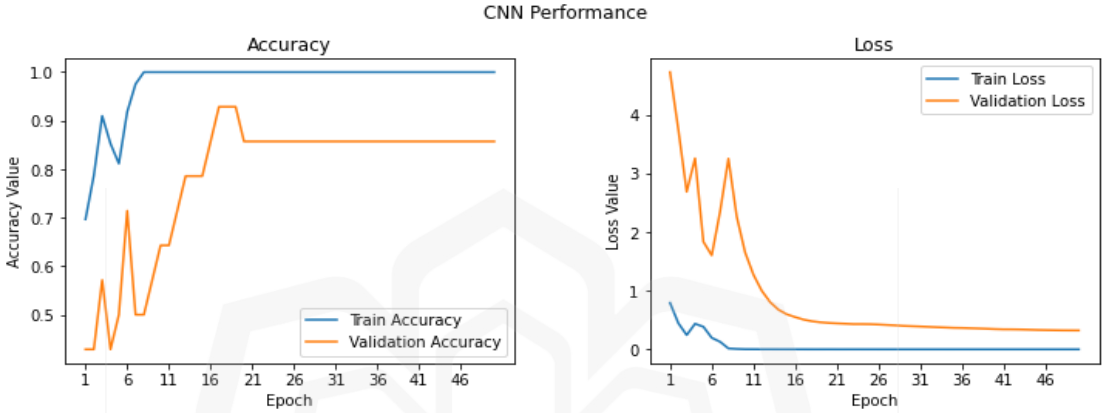


Figure 5.1 Accuracy and loss curves of the proposed BTMIC-CNN model for binary classification without data augmentation

Confusion matrix has been calculated for this experiment as shown in Figure 5.2. the Figure shows that 149 out of 155 MRI images have been detected correctly as tumorous MRI images. Similarly, 93 out of 98 non-tumorous MRI images have been classified to the correct class.

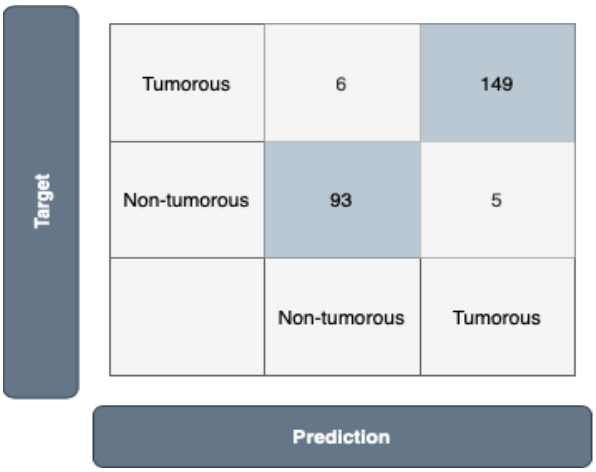


Figure 5.2 Confusion matrix showing number of images and the accuracy for each class of the proposed model (BTMIC-CNN) without data augmentation.

Further evaluation for this experiment has been performed. Precision, recall and F1-Score have been calculated for each type of images as illustrated in Table 5.1. The Table shows that the overall accuracy is ~ 96%. The prediction result and actual result for some random images has been attached in the appendix A.

Table 5.1 Classification report of the proposed (BTMICT-CNN) for task 1

Metrics/Tumor	***Precision***	***Recall***	***F1-score***	Total images
Yes	96%	95%	96%	155
No	95%	97%	96%	98
Average	96%	96%	96%	253
Accuracy	96%	96%	96%	253

5.2.2 Experiment-2 (The proposed BTMIC-CNN Model with data augmentation)

The second experiment applied BTMIC-CNN model on the augmented binary dataset which contains 2015 MRI brain medical images. Obviously, the test accuracy has been improved from the first experiment reaching 97.02% as illustrated in Figure 5.3 below. This proof that data augmentation is a good method to solve the overfitting issues caused by small datasets and improve classification accuracy.

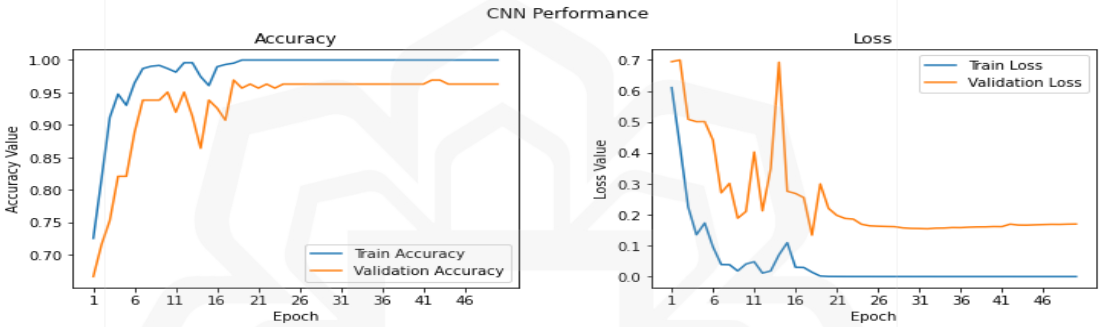


Figure 5.3 Accuracy and loss curves of the proposed BTMIC-CNN model for binary classification with data augmentation

The confusion matrix shows how many images exactly have been classified correctly for each class and vice versa. Figure 5.4 illustrates that 924 MRI images out of 930 was classified correctly in the no (non-tumorous images) class which is nearly ~ 99.35%. On the other hand, 1073 MRI images out of 1085 was classified correctly in the yes(tumorous) class reaching 98.89%.

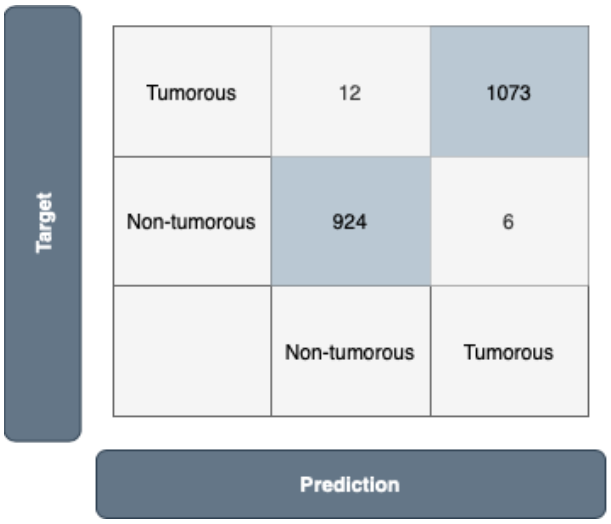


Figure 5.4 Confusion matrix showing number of images and the accuracy for each class of the proposed model (BTMIC-CNN).

The classification report for experiment 2 has been summarized in Table 5.2 below. It shows three important metrics (precision, recall and F1-score) as well as the accuracy for both classes of brain tumors MRI images. The overall accuracy is 99 %. In comparison with previous work with the same dataset this considered among the highest accuracies for the given dataset. Appendix B contains random MRI images for actual and predicted result for this experiment.

Table 5.2 Classification report of the proposed (BTMICT-CNN)

Metrics/Tumor	***Precision***	***Recall***	***F1-score***	Total images
Yes	99%	99%	99%	1085
No	99%	99%	99%	930
Average	99%	99%	99%	2015
Accuracy	99%	99%	99%	2015

5.2.3 Experiment-3 (TL of VGG16 without data augmentation)

Transfer learning of VGG16 has been employed on the binary dataset with and without augmentation to study and compare the accuracy changes for both experiments. This experiment showed that train and validation accuracy kept increasing throughout the training iterations; however, the training accuracy for this experiment is ~100 %, the testing accuracy is 94.44%, as per Figure 5.5.

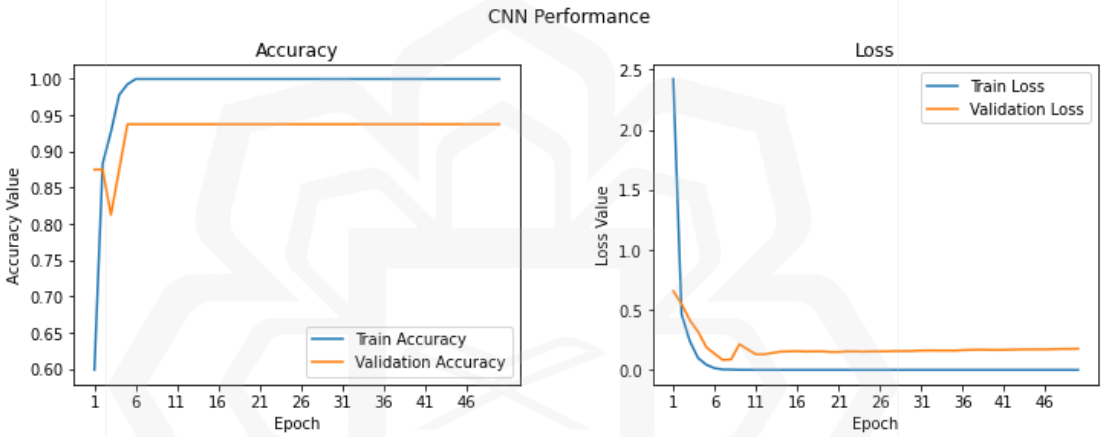


Figure 5.5 Training progress: The Accuracy and the loss history represent each set validation and training history of the dataset without augmentation.

Confusion matrix for this experiment is shown in Figure 5.6. The number of tumorous images that have been classified correctly is 150 images out of 155. While the non-tumorous images have all been classified to the correct class.

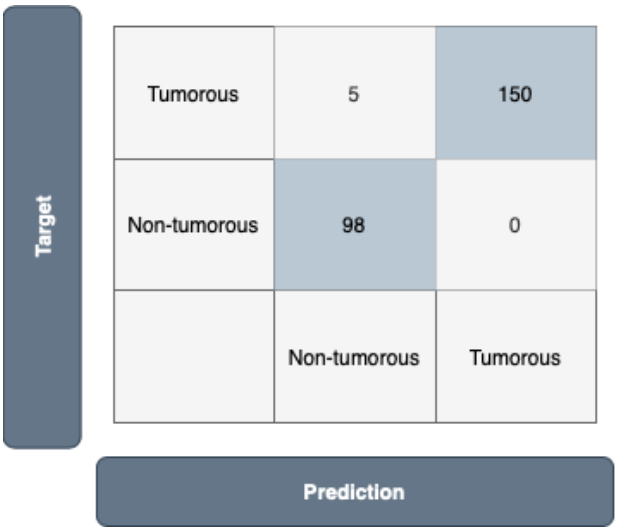


Figure 5.6 Confusion matrix showing number of images and the accuracy for each class using TL of VGG16

The Table below shows the evaluation for this experiment. It shows that Precision is 100% for tumorous class and 98% for non-tumorous class. F1-score is 99% for both classes. While recall is 98% for tumorous class and 100% for non-tumorous class. The overall accuracy is ~99% as shown in the Table 5.3. The prediction value vs actual value for random selected images is shown in appendix C.

Table 5.3 Classification report of binary classification using TL of VGG16 without data augmentation.

Metrics/Tumor	***Precision***	***Recall***	***F1-score***	Total images
Yes	100%	98%	99%	155
No	98%	100%	99%	98
Average	99%	99%	99%	253
Accuracy	99%	99%	99%	253

5.2.4 Experiment-4 (TL of VGG16 with data augmentation)

The fourth experiment involve the dataset with data augmentation. The results showed that both training and validation accuracy increased throughout the training time to reach ~100% accuracy, as per Figure 5.7. Obviously, the accuracy has improved and proved that data augmentation in case of leak data is a good way to improve the performance of Deep Learning models and it can also solve overfitting issues.

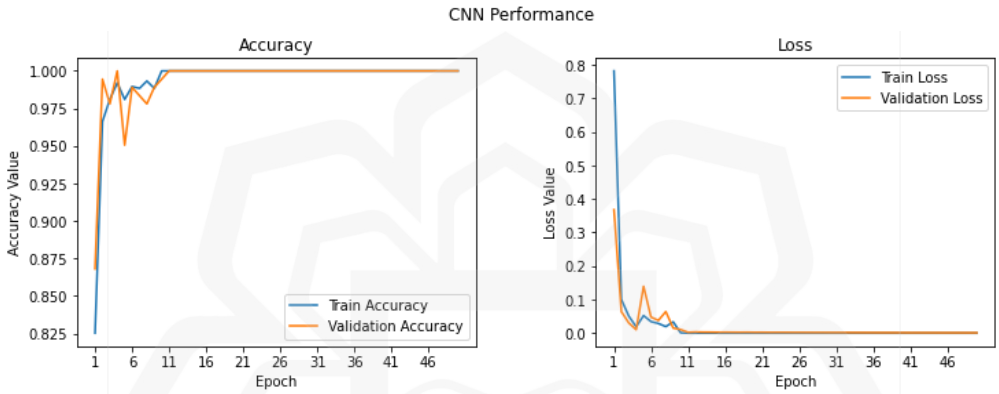


Figure 5.7 Training progress: The Accuracy and the loss history represent each set validation and training history of the augmented dataset.

The confusion matrix is a good indicator of the model performance as presented in Figure 5.8. The matrix shows the number of images classified correctly and their percentage of both tumorous and non-tumorous MRI images. The prediction and target of tumorous augmented MRI images is 100% correctly classified (1085 MRI images). Similarly for the non-tumorous augmented MRI images is 100% correctly classified (930 MRI images). This result is also reflected in the appendix D which shows that probability of the prediction class is exactly same to the actual one.

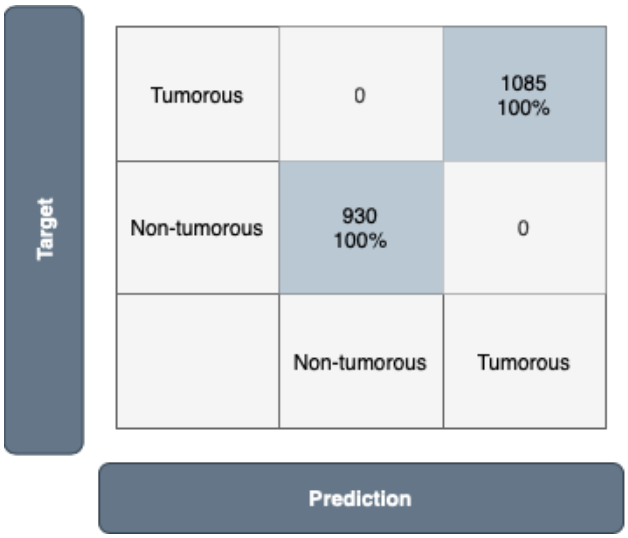


Figure 5.8 Confusion matrix showing number of images classified correctly for the binary dataset.

Further analysis was carried out. The classification report is summarized in Table 5.4 which shows the percentages of precision, recall and F1-score reach ~ 100% for both tumorous and non-tumorous images.

Table 5.4 Classification report of the proposed transfer learning of VGG16

Metrics / Tumor	Precision	Recall	F1-score	Total images
Tumorous	100%	100%	100%	1085
Non-tumorous	100%	100%	100%	930
Average	100%	100%	100%	2015
Accuracy	100%	100%	100%	2015

5.2.5 Comparison with others work

This set of experiments on this kind of dataset demonstrate the efficiency and effectiveness of the proposed BTMIC-CNN algorithm in term of brain MRI medical images binary classification and its result can compete to one of the most powerful known classification algorithms (VGG16) which has been performed through transfer learning in the experiments 3 and 4. It can also be concluded that data augmentation is an excellent method to improve the model's accuracy and balance the dataset. Table 5.5 illustrates a compression between the proposed model and previous works that used the same dataset.

Table 5.5 Performance of different models on the same dataset

Ref.	Algorithm	Accuracy
(Bakr Siddiaue et al., 2020)	DCNN (VGG16)	96 %
(Sane, 2021)	KNN-RF	96 %
Experiment 1	BTMIC-CNN without data augmentation	96%
Experiment 2	BTMIC-CNN with data augmentation	99%
Experiment 3	TL-VGG16 without data augmentation	99%
Experiment 4	TL-VGG16 with data augmentation	100 %

5.3 MULTICLASS CLASSIFICATION

This section illustrates the result and discussion of the multiclass dataset for both the proposed BTMIC-CNN model and transfer learning of VGG16.

5.3.1 Experiment-5 (The proposed BTMIC-CNN Model without data augmentation)

First, we will discuss the result of the original dataset without data augmentation. The dataset was split: 80 % training dataset, 10% validation dataset, and 10% test dataset. The validation dataset (10%) was executed during training the model, while the testing dataset (10%) was the final validation of an unseen dataset to confirm that the DL system was appropriately trained, as shown in Figure 5.10. The training accuracy curves show that training accuracy reaches ~100% while the testing accuracy is 96.41%. The validation accuracy converges with the training curve, reaching 98.37 % at the end of the training iterations. The loss curve of the proposed model shows a good fit model since training and validation loss is decreasing and converging during the training iterations, as seen in the Figure 5.9 below.

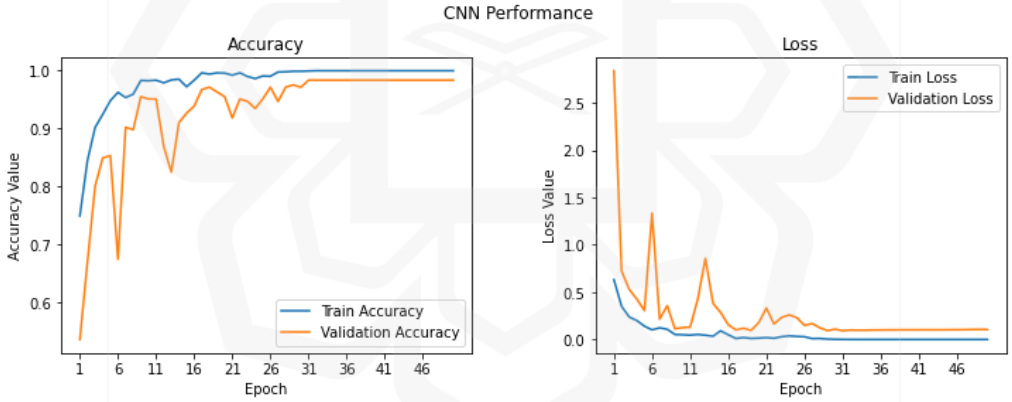


Figure 5.9 Accuracy and loss curves of the proposed BTMIC-CNN model

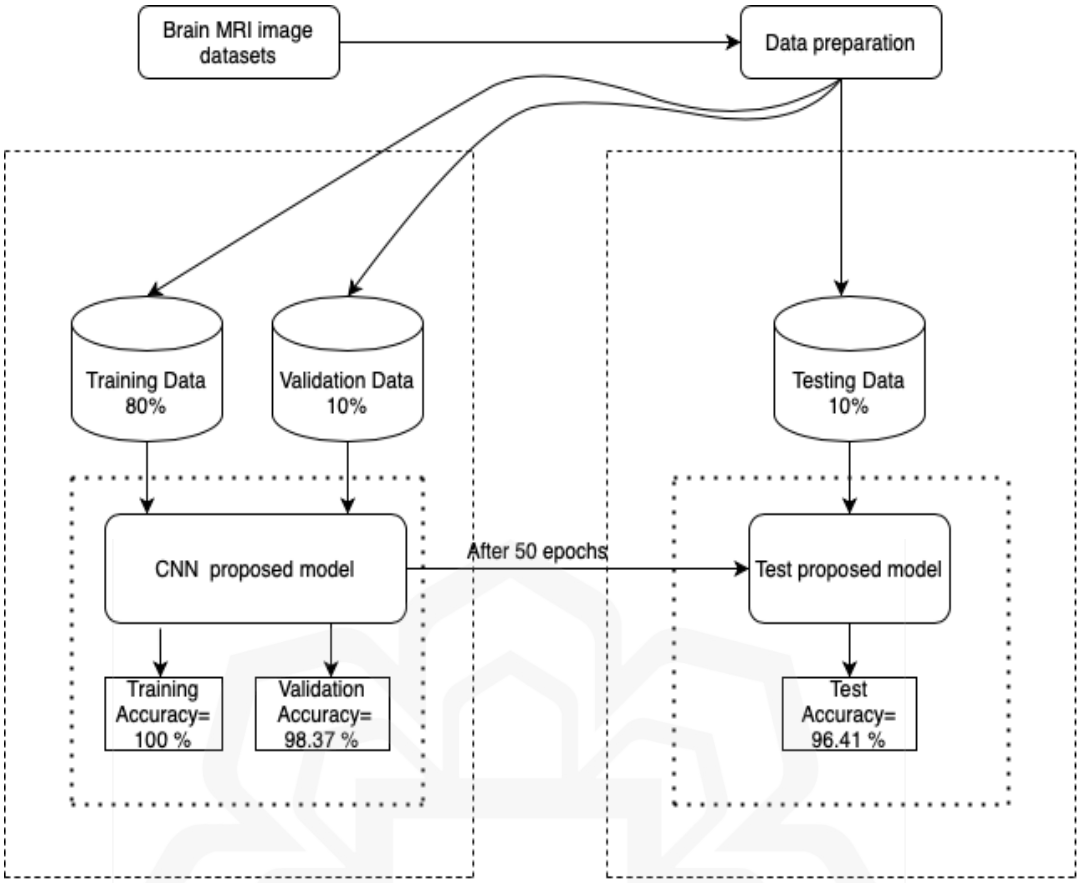


Figure 5.10 Flowchart of data distribution and the accuracy result

The confusion matrix and classification evaluation metrics have been calculated for each type of brain tumour image individually. Figure 5.11 shows the BTMIC-CNN model's confusion matrix for Glioma class 1416 images out of 1426 correctly identified, resulting in of 99.29 % accuracy. In the Meningioma class 695 images out of 708 total images have been classified correctly which reflect 98.16 % accuracy. Third class is pituitary tumor, in which 927 images out of 930 total images is classified correctly which reflect 99.68 % accuracy for this class.

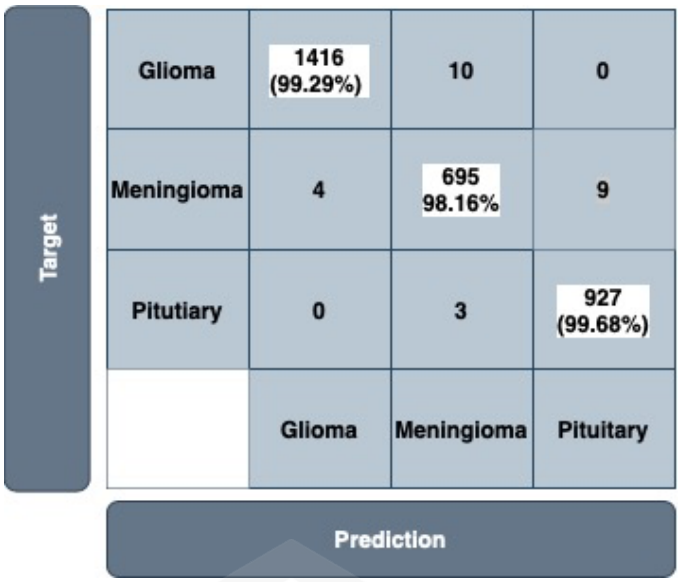


Figure 5.11 Confusion matrix showing number of images and the accuracy for each tumor type of the proposed model (BTMIC-CNN)

As demonstrated in Table 5.6, the suggested BTMIC-CNN model performed well in the evaluation. The precision, recall and F1-score are reported in Table 5.6. Glioma class show an excellent report in which it reach ~100% for precision and F1-score, and 99% for recall. The Meningioma tumour report is ~98% for precision , recall and F1-score. The third class is the pituitary tumour which report 99% for precision and F1-score and 100% for recall. Per Table 5.6, the average is ~99 for the precision, recall, and F1-score. Finally, the accuracy has been calculated for all the images together, The overall accuracy of the predicted test is ~99%. Appendix E shows the actual result vs predicted result for this experiment.

Table 5.6 Classification report of the proposed (BTMICT-CNN) without data augmentation

Metrics/Tumor	***Precision***	***Recall***	***F1-score***	Total images
Glioma	100%	99%	100%	1426
Meningioma	98%	98%	98%	708
Pituitary	99%	100%	99%	930
Average	99%	99%	99%	3064
Accuracy	99%	99%	99%	3064

5.3.2 Experiment-6 (The proposed BTMIC-CNN Model with data augmentation)

This experiment involved the application of the proposed algorithm (BTMIC-CNN) on the augmented dataset. The dataset contains 9192 MRI images. It has been distributed: 80 % training dataset, 10% validation dataset, and 10% test dataset as proposed earlier. The accuracy and loss curves are shown in Figure 5.12. The training accuracy reach to ~100% after 50 epochs while the test accuracy converges to be 96.79% which is slightly higher than the testing accuracy of the original dataset which was 96.41 % as mentioned in the previous experiment. The proposed model's loss curve indicates a good model since training and validation losses are falling and converging over training iterations as shown in the Figure 5.12.

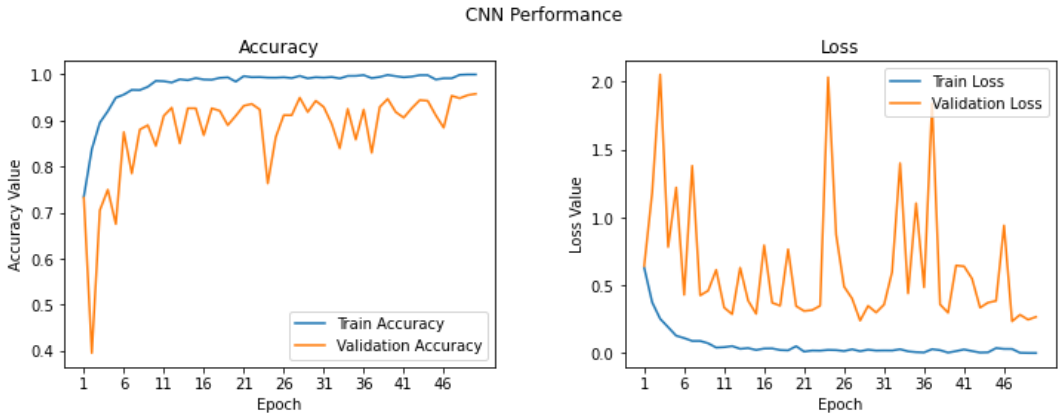


Figure 5.12 Accuracy and loss curves of the proposed BTMIC-CNN model for the augmented dataset

The performance of BTMIC-CNN model demonstrated again it is efficiency in the augmented dataset as illustrated in Table 5.7. The confusion matrix and classification evaluation metrics have been calculated for each type of brain tumour image separately as illustrated in Figure 5.13. the evaluation of both experiments showed the capability of the proposed algorithm (BTMIC-CNN) in the multiclass classification problems for brain tumors MRI medical images for small dataset as well as augmented dataset.

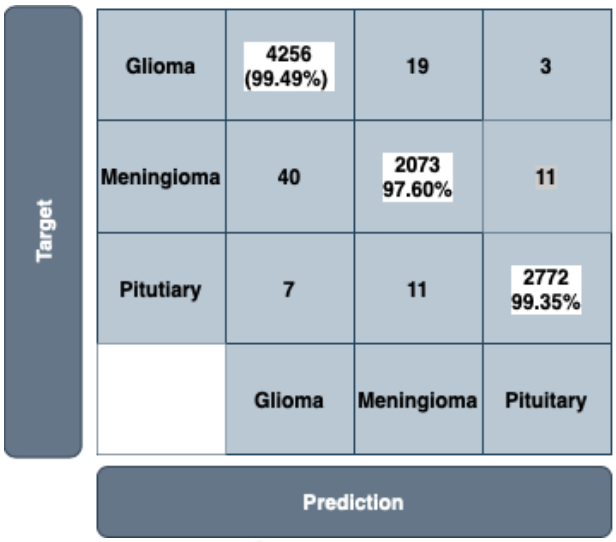


Figure 5.13 Confusion matrix showing number of images and the accuracy for each tumor type of the proposed model (BTMIC-CNN) of the augmented dataset.

The classification report in Table 5.7 summarizes the evaluation metrics for this task. It shows that both glioma and pituitary classes are achieving 99% for precision, recall and F1-score. While in meningioma class it shows that 98% was achieved for precision, recall and F1-score. The overall accuracy in this task as shown in Table 5.7 is 99% which demonstrates the efficiency of the proposed model BTMIC-CNN for the augmented dataset. Appendix F contains prediction and actual label for randomly selected images from this dataset.

Table 5.7 Classification report of the proposed (BTMICT-CNN) of the augmented dataset

Metrics/Tumor	Precision	Recall	F1-score	Total images
Glioma	99%	99%	99%	4278
Meningioma	98%	98%	98%	2124
Pituitary	99%	99%	99%	2790

Average	99%	99%	99%	9192
Accuracy	99%	99%	99%	9192

In this section, one more experiment was added to check the accuracy rate of proposed model on the unseen augmented multiclass dataset with 70% training, 20% testing and 10% validation. The performance of the model was good after 50 epochs as depicted in Figure 5.14. The training and validation accuracy fluctuated initially but stabilized around 40 epochs, achieving around 99% and 90% respectively. The test accuracy for this experiment reported is 94.31%. The loss curves exhibited a consistent trend with the accuracy curve throughout the training epochs.

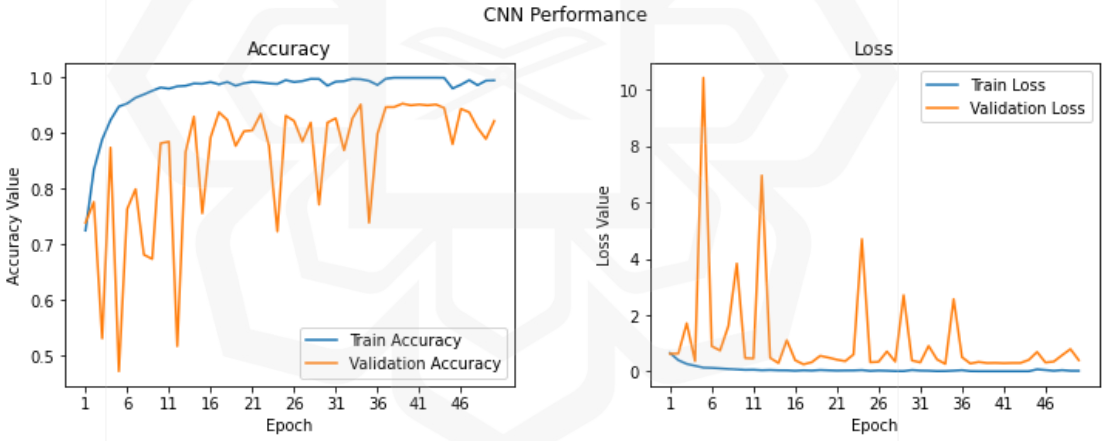


Figure 5.14 Accuracy and loss curves of the proposed BTMIC-CNN model for the augmented dataset 70% training 20% testing and 10%validation

The performance of BTMIC-CNN model demonstrated again it is efficiency in the augmented dataset as illustrated in Table 5.8. The confusion matrix and classification evaluation metrics have been calculated for each type of brain tumour image separately as illustrated in Figure 5.15. the evaluation of both experiments

showed the capability of the proposed algorithm (BTMIC-CNN) in the multiclass classification problems for brain tumors MRI medical images for small dataset as well as augmented dataset.

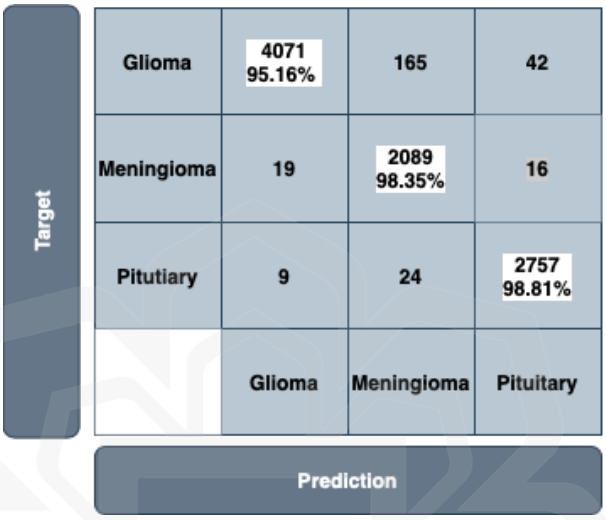


Figure 5.15 Confusion matrix showing number of images and the accuracy for each tumor type of the proposed model (BTMIC-CNN) of the augmented dataset 70% training 20% testing and 10% validation.

Table 5.8 presents an overview of the evaluation metrics for this particular task. The results demonstrate overall accuracy of 97%, which highlights the effectiveness of the BTMIC-CNN model on the augmented dataset of 70% training 20% testing and 10%validation. which indicates the efficiency of BTMIC-CNN model for different dataset, different sizes and different distribution of the training and testing datasets.

Table 5.8 Classification report of the proposed (BTMICT-CNN) of the augmented dataset 70% training 20% testing and 10%validation

Metrics/Tumor	Precision	Recall	F1-score	Total images
Glioma	99%	95%	97%	4278
Meningioma	92%	98%	95%	2124
Pituitary	98%	99%	98%	2790
Average	96%	97%	97%	9192
Accuracy	97%	97%	97%	9192

5.3.3 Experiment-7 (TL of VGG16 without data augmentation)

The dataset was split: 80 % training dataset, 10% validation dataset, and 10% test dataset. This experiment involves Transfer learning of VGG16 on the multiclass dataset. The result shows a good performance after 50 epochs as shown in Figure 5.16. The training and validation accuracy was fluctuated during the training until it stabilized in the 40 epochs to reach ~ 99 % for the training accuracy and ~ 92.50% for the test accuracy. The loss curves show a consistent result with the accuracy curve during the training epochs as illustrated in figure 5.16.

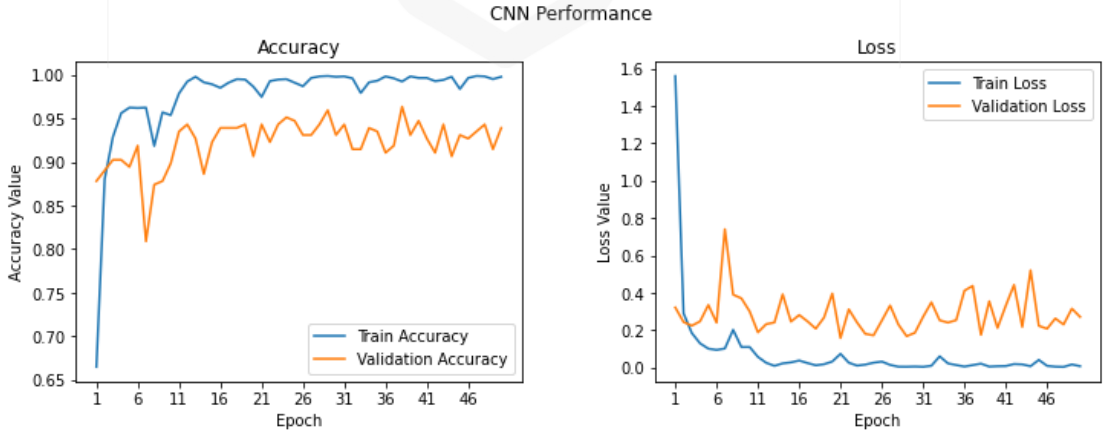


Figure 5.16 Accuracy and loss curves for TL-VGG16 on multiclass CE-dataset

Figure 5.17 shows the confusion matrix for the three types of the brain tumors (meningioma, glioma and pituitary). The Figure illustrates the number of images classified correctly as well as their percentage. For glioma 1399 images out of 1426 images were classified correctly which represent 98.10 % accuracy. While meningioma, which has the least number of images among other types in this dataset, 683 images out of 708 images was classified correctly which represent 96.46% accuracy. The third type is pituitary as illustrated in the confusion matrix 917 images out of 930 images was classified correctly reaching 98.60 % accuracy. This experiment will be followed by another experiment using same model (TL-VGG16) but with augmented dataset to check further improvement of the accuracy for meningioma class since it has the lowest number of images and lowest accuracy among other types in this dataset.

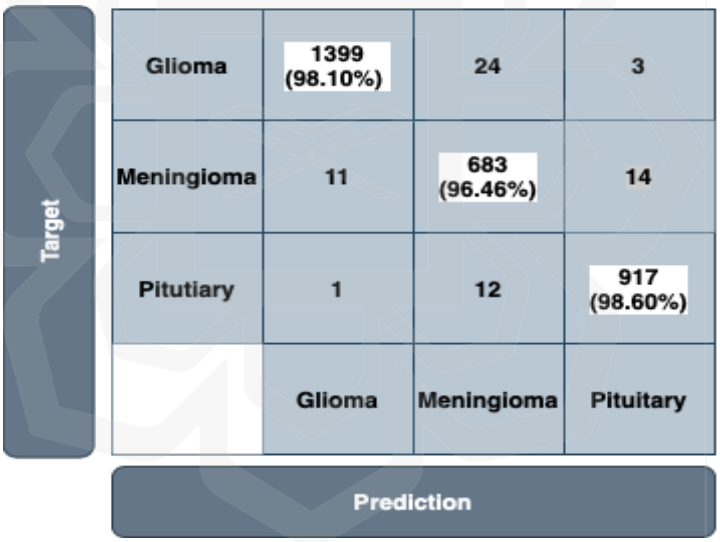


Figure 5.17 Confusion matrix of Glioma, Meningioma and Pituitary brain tumor using TL-VGG16

The classification report shows the result of the evaluation matrix as reported in Table 5.9. The precision in Glioma class is 99%, in pituitary 98 %, while 95% in Meningioma class. Similarly, Recall and F1-score represent higher percentages for Glioma and Pituitary and lowest was for Meningioma class. The overall accuracy for

the 3064 images is ~ 98% for this model. Appendix G contains randomly selected images for this experiment.

Table 5.9 Classification report of TL-VGG16 without data augmentation

Metrics / Tumor	Precision	Recall	F1-score	Total images
Glioma	99%	98%	99%	1426
Meningioma	95%	96%	96%	708
Pituitary	98%	99%	98%	930
Average	97%	98%	98%	3064
Accuracy	98%	98%	98%	3064

5.3.4 Experiment-8 (TL of VGG16 with data augmentation)

This experiment involves the augmented dataset which contains 9192 images. The experiment has been performed in same procedures as previous experiment in which the data has been divided to 80% training, 10% validation and 10 % testing. The accuracy result showed improvement from the previous experiment especially for the meningioma class. The test accuracy is 95.87%.

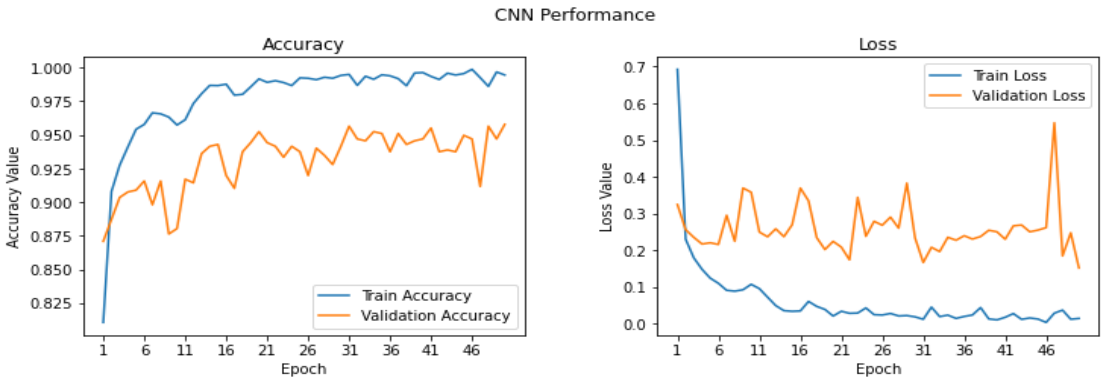


Figure 5.18 Accuracy and loss curves for TL-VGG16 on multiclass CE-dataset with data augmentation

The following Figure 5.19 shows the confusion matrix for the number of images that have been classified correctly in this experminet-2. Meningioma class classification accuracy has been improved as well as pituitary class. While glioma class has slightly decreased in term of accuracy of total number of images that have been classified correctly.

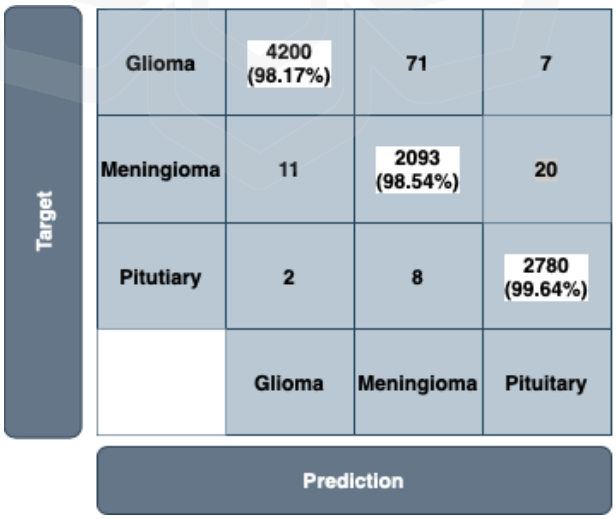


Figure 5.19 Confusion matrix of Glioma, Meningioma and Pituitary brain tumor using TL-VGG16 with data augmentation.

The classification report in Table 5.10 showed an improvement for precision, recall and F1-score with data augmentation. The precision for meningioma class has improved from 95% without augmentation to 97% with data augmentation. Similarly for recall and F1-score they have improved from 96% to 98% and 99% respectively. Precision and F1-score have also improved from 98% without data augmentation to 99% with data augmentation for pituitary while recall improved to ~100%. Glioma class shows improvement in precision from 99% to 100%. The overall accuracy has also showed improvement from 98% without data augmentation to 99% with data augmentation as reported in Table 5.10.

Table 5.10 Classification report of TL-VGG16 with data augmentation

Metrics / Tumor	Precision	Recall	F1-score	Total images
Glioma	100%	98%	99%	4278
Meningioma	96%	99%	97%	2124
Pituitary	99%	100%	99%	2790
Average	98%	99%	99%	9192
Accuracy	99%	99%	99%	9192

5.3.5 Comparison with others work

Table 5.11 summarises the comparisons between the proposed model and state-of-the-art methods that employed the same dataset. The best result was achieved using the proposed BTMIC-CNN model which reach 99%. The experiments performed demonstrated the efficiency of the proposed model BTMIC-CNN for both types (binary and multiclass) of classification problems in brain tumor MRI medical images.

Table 5.11 Comparison with related studies that used the T1-CE MRI dataset

Author	**Method**	**Accuracy**
(Parnian Afshar et al., 2018)	CapsNet	86.56 %
(Cheng et al., 2015)	SVM	91.28 %
(Fuad et al., 2021)	GoogleNet	92 %
(Irmak, 2021)	CNN	92.66 %
(Pashaei et al., 2018)	KE-CNN	93.68 %
(M. Sajjad et al., 2019)	VGG19	94.58 %
(Fuad et al., 2021)	AlexNet	94.6 %
(Swati et al., 2019)	VGG19	94.82 %
(Díaz-Pernas et al., 2021)	CNN	97.3 %
(Deepak & Ameer, 2019)	Deep CNN-SVM	98 %
Proposed model	**BTMIC-CNN**	**99 %**

5.4 CHAPTER SUMMARY

This chapter described the result of the fully automated brain tumour classification of MRI medical images model based on a CNN (BTMIC-CNN) for binary and multiclass classification tasks. The datasets used is available for research and development. The evaluation of the proposed model proved it is efficient and is highly accurate, at ~99%, making it helpful to radiologists in binary and multiclass classifying brain tumour images based on their different types. This research provides a brain tumour MRI medical images classification method that is accurate and automatic, with minimal pre-processing. Deep transfer learning of VGG16 was used to compare the classification result with the proposed system. In addition, the dataset was augmented and balanced to improve the performance. Relative to other works, the system reported the highest classification accuracy. Also, train and validation accuracy measures were employed to assess the system's performance and robustness, and as a result, the proposed models could be implemented for the early detection of brain tumor from MRI medical images.

CHAPTER FIVE

CONCLUSION AND FUTURE DIRECTIONS

6.1 INTRODUCTION

Understanding how to mix Deep Learning predictions with human judgments is a critical condition for using Deep Learning technology in healthcare procedures. Simply sharing the Deep Learning system's prognosis with the doctor is one simple kind of collaboration. There are critical considerations in this model as well, such as the form the prediction should take and how the forecast should be conveyed to the doctor. The massive growth of medical images in the last decade makes it difficult for medical expertise and radiologist to analyse and classify them.

Medical images have huge information that can be used for medical diseases detection, surgical planning, training and research. Thus, there is a need for a technique that can analyse and classify these images based on their contents automatically. Deep Learning techniques have been recently used in the medical image analysis domain. The main aim of this research is to develop algorithm to detect and classify brain tumor images with high performance and better accuracy based on using DL techniques. Moreover, a set of experiments is performed for evaluation using MRI medical image datasets.

6.2 REVIEW OF OBJECTIVES

1. **Objective 1.** To identify the current methodologies and algorithms exist to analyze brain tumor MRI medical images.

   A systematic review was used to achieve this goal as detailed in chapter two. More than 426 papers were reviewed, and more than 160 papers were studied in detail. Studies was grouped and summarized based on the

objectives, algorithms and datasets used. The conclusion drawn from related previous work helped to shape an idea of the proposed model.

2. **Objective 2.** To improve a Deep Learning algorithm for detection of a brain tumor MRI medical images.

   This objective was met in identifying tumorous images and non-tumours images throughout chapter 3, 4 and 5. The dataset used to meet this objective contains two type of MRI medical images which are tumorous and non-tumours. Data augmentation was used to increase and balance the dataset. The proposed BTMIC-CNN model proved its efficiency in brain tumor detection through binary classification resulting in accuracy of ~99%.

3. **Objective 3.** To improve a Deep Learning algorithm for classification of a brain tumor MRI medical images.

   This objective was met by proposed BTMIC-CNN model that is capable of classifying different types of brain tumors. This model is explained throughout chapter 3, 4 and 5. The dataset used contains three type of brain tumors and it is available for research and development. Dataset understanding, preparation and pre-processing was an initial step in order to feed it to the proposed model. Many trails were carried out until we reach a high-performance level and high accuracy classification rate which reach to ~ 99% and this result outperformed previous studies that employed the same dataset.

4. **Objective 4.** To validate and evaluate the proposed algorithms with the current state-of-art algorithm.

   Validation and evaluation of the proposed algorithms were performed in chapter 5. The classification reports including accuracy, confusion matrix, precision, recall and F1-Score for all experiments performed was reported and summarized in different Figures and Tables in chapter 5. The overall performance was compared to previous research.

6.3 CONTRIBUTION

Overall, the research shows improvement in Deep Learning for classification of brain tumor MRI medical images employing a combination of achievements in:

1. Improvements in research that analyze the performance of Deep Learning algorithms in a systematic way. The systematic review identified recent algorithms and techniques that are used to classify brain MRI medical images. The responses to research question 1 are reflected in this contribution.
2. Developments in Deep Learning algorithms in the context of detection of tumorous brain medical images. Dataset contains tumorous and non-tumorous MRI medical images. The dataset has been improved through data augmentation. Binary classification algorithm showed an excellent classification accuracy proposed BTMIC-CNN model. The responses to research question 2 are reflected in this contribution.
3. Developments in Deep Learning algorithms in the context of identifying the type of brain tumor through MRI medical images dataset that contains three types of tumors. BTMICT-CNN model has been introduced and the result shows a great classification accuracy compared to recent works that used the same dataset. The responses to research question 3 are reflected in this contribution.
4. Evaluation of the algorithms in the defined datasets and comparison with other state-of-art algorithms. The responses to research question 4 are reflected in this contribution.

6.4 FUTURE WORK AND RECOMMANDATIONS

The seamless integration of Deep Learning technology into clinical operations is the future of AI in medicine. Despite the fact that the study done here reveals that the objectives have been met, there is still much space for improvement. Considering the following topics for further research, the proposed model might be used in the future to evaluate more complicated medical images databases, such as those with more

categories. There are several Deep Learning classification approaches, as discussed in Chapter 2. Developing hybrid classifiers might be beneficial. This hybrid system may result in a good performance.

Deep learning algorithms can become better at analysing medical images, they may be able to help identify patients who are at risk for certain conditions and provide personalized treatment recommendations. For example, deep learning algorithms could analyse medical images to determine which patients are more likely to benefit from a particular treatment or which patients are more likely to experience certain side effects.

It is also possible to test the same classifier with other datasets of different tumor's parts of the body. Generally, there's more research to be done in the medical area to build a fully functional application that can be utilised in hospitals for its well-organized framework and open standards.

REFERENCES

Abd-Ellah, M. K. K., Awad, A. I. I., Khalaf, A. A. M. A. M., & Hamed, H. F. A. F. A. (2018). Two-phase multi-model automatic brain tumour diagnosis system from magnetic resonance images using Convolutional Neural Networks. *Eurasip Journal on Image and Video Processing*, *2018*(1). https://doi.org/10.1186/s13640-018-0332-4

Abdelaziz Ismael, S. A., Mohammed, A., & Hefny, H. (2020). An enhanced Deep Learning approach for brain cancer MRI images classification using residual networks. *Artificial Intelligence in Medicine*, *102*(December). https://doi.org/10.1016/j.artmed.2019.101779

Afshar, P., Mohammadi, A., & Plataniotis, K. N. (2018). Brain Tumor Type Classification via Capsule Networks. *Proceedings - International Conference on Image Processing, ICIP*, 3129–3133. https://doi.org/10.1109/ICIP.2018.8451379

Ahammed Muneer, K. V., Rajendran, V. R., & Paul Joseph, K. (2019). Glioma Tumor Grade Identification Using Artificial Intelligent Techniques. *Journal of Medical Systems*, *43*(5). https://doi.org/10.1007/s10916-019-1228-2

Aiello, M., Cavaliere, C., D'Albore, A., & Salvatore, M. (2019). The Challenges of Diagnostic Imaging in the Era of Big Data. *Journal of Clinical Medicine*, *8*(3), 316. https://doi.org/10.3390/jcm8030316

Alqudah, A. M., Alquraan, H., Qasmieh, I. A., Alqudah, A., & Al-Sharu, W. (2019). Brain tumor classification using Deep Learning technique - A comparison between cropped, uncropped, and segmented lesion images with different sizes. *International Journal of Advanced Trends in Computer Science and Engineering*, *8*(6), 3684–3691. https://doi.org/10.30534/ijatcse/2019/155862019

Amin, Javaria, Sharif, M., Gul, N., Yasmin, M., & Shad, S. A. (2020). Brain tumor classification based on DWT fusion of MRI sequences using Convolutional Neural Network. *Pattern Recognition Letters*, *129*, 115–122. https://doi.org/10.1016/j.patrec.2019.11.016

Amin, Javeria, Sharif, M., Yasmin, M., Saba, T., Anjum, M. A., & Fernandes, S. L. (2019). A New Approach for Brain Tumor Segmentation and Classification Based on Score Level Fusion Using Transfer Learning. *Journal of Medical Systems*, *43*(11), 326. https://doi.org/10.1007/s10916-019-1453-8

Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical Image Analysis using Convolutional Neural Networks: A Review. *Journal of Medical Systems*, *42*(11), 1–13. https://doi.org/10.1007/s10916-018-1088-1

Ayachi, R., & Ben Amor, N. (2009). Brain tumor segmentation using support vector machines. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5590 LNAI*, 736–747. https://doi.org/10.1007/978-3-642-02906-6_63

Badža, M. M., & Barjaktarović, M. Č. (2021). Segmentation of Brain Tumors from MRI Images Using Convolutional Autoencoder. *Applied Sciences*, *11*(9), 4317. https://doi.org/10.3390/app11094317

Baid, U., Talbar, S., Rane, S., Gupta, S., Thakur, M. H., Moiyadi, A., … Mahajan, A. (2019). Deep Learning radiomics algorithm for gliomas (DRAG) model: A novel approach using 3D UNET based deep Convolutional Neural Network for predicting survival in gliomas. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-030-11726-9_33

Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., … Menze, B. (2018). *Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge*. (November). Retrieved from http://arxiv.org/abs/1811.02629

Bakr Siddiaue, M. A., Sakib, S., Rahman Khan, M. M., Tanzeem, A. K., Chowdhury, M., & Yasmin, N. (2020). Deep Convolutional Neural Networks model-based brain tumor detection in brain MRI images. *Proceedings of the 4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020*, 909–914. https://doi.org/10.1109/I-SMAC49090.2020.9243461

Banerjee, I., Crawley, A., Bhethanabotla, M., Daldrup-Link, H. E., & Rubin, D. L. (2018). Transfer learning on fused multiparametric MR images for classifying histopathological subtypes of rhabdomyosarcoma. *Computerized Medical Imaging and Graphics*, *65*, 167–175. https://doi.org/https://doi.org/10.1016/j.compmedimag.2017.05.002

Banzato, T., Bernardini, M., Cherubini, G. B., & Zotti, A. (2018). A methodological approach for Deep Learning to distinguish between meningiomas and gliomas on canine MR-images. *BMC Veterinary Research*, *14*(1). https://doi.org/10.1186/s12917-018-1638-2

Beig, N., Patel, J., Prasanna, P., Partovi, S., Varadan, V., Madabhushi, A., & Tiwari, P. (2017). *Radiogenomic analysis of hypoxia pathway reveals computerized MRI descriptors predictive of overall survival in glioblastoma*. *10134*, 101341U. https://doi.org/10.1117/12.2255694

Benson, E., Pound, M. P., French, A. P., Jackson, A. S., & Pridmore, T. P. (2019). Deep hourglass for brain tumor segmentation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-030-11726-9_37

Brunese, L., Mercaldo, F., Reginelli, A., & Santone, A. (2020). An ensemble learning approach for brain cancer detection exploiting radiomic features. *Computer Methods and Programs in Biomedicine*, *185*, 105134. https://doi.org/https://doi.org/10.1016/j.cmpb.2019.105134

Carneiro, T., Da Nobrega, R. V. M., Nepomuceno, T., Bian, G. Bin, De Albuquerque, V. H. C., & Filho, P. P. R. (2018). Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. *IEEE Access*, *6*, 61677–61685. https://doi.org/10.1109/ACCESS.2018.2874767

Chaddad, A., Desrosiers, C., & Toews, M. (2016). Radiomic analysis of multi-contrast brain MRI for the prediction of survival in patients with glioblastoma multiforme. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, *2016-Octob*, 4035–4038. https://doi.org/10.1109/EMBC.2016.7591612

Chakrabarty, N.(2019). Brain MRI images dataset for brain tumor detection. Available from: https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection.

Charron, O., Lallement, A., Jarnet, D., Noblet, V., Clavier, J.-B., & Meyer, P. (2018). Automatic detection and segmentation of brain metastases on multimodal MR images with a deep Convolutional Neural Network. *Computers in Biology and Medicine*, *95*, 43–54. https://doi.org/https://doi.org/10.1016/j.compbiomed.2018.02.004

Chen, H., Dou, Q., Yu, L., Qin, J., & Heng, P.-A. (2018). VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images. *NeuroImage*, *170*, 446–455. https://doi.org/https://doi.org/10.1016/j.neuroimage.2017.04.041

Chen, S., Ding, C., & Liu, M. (2019). Dual-force Convolutional Neural Networks for accurate brain tumor segmentation. *Pattern Recognition*, *88*, 90–100. https://doi.org/https://doi.org/10.1016/j.patcog.2018.11.009

Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z., … Feng, Q. (2015). Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PLoS ONE*, *10*(10), 1–13. https://doi.org/10.1371/journal.pone.0140381

Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L., & Haworth, A. (2021). A review of medical image data augmentation techniques for Deep Learning applications. *Journal of Medical Imaging and Radiation Oncology*, *65*(5), 545–563. https://doi.org/10.1111/1754-9485.13261

Chollet, F. (2017). *Deep Learning with python -Keras -book builds understanding through intuitive explanations and practical examples*. https://doi.org/10.1017/CBO9781107415324.004

Cui, S., Mao, L., Jiang, J., Liu, C., & Xiong, S. (2018). Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network. *Journal of Healthcare Engineering*, *2018*. https://doi.org/10.1155/2018/4940593

Chugh, J. (2018). Types of Machine Learning and Top 10 Algorithms Everyone Should Know. *URL: https://blogs. oracle. com/ai/types-ofmachine-learning-and-top-10-algorithms-everyone-should-know*.

Czarnek, N., Clark, K., Peters, K. B., & Mazurowski, M. A. (2017). Algorithmic three-dimensional analysis of tumor shape in MRI improves prognosis of survival in glioblastoma: a multi-institutional study. *Journal of Neuro-Oncology*, *132*(1), 55–62. https://doi.org/10.1007/s11060-016-2359-7

DeAngelis. (2001). *BRAIN TUMORS*. *344*(2), 114–123.

Deepak, S., & Ameer, P. M. (2019). Brain tumor classification using deep CNN features via transfer learning. *Computers in Biology and Medicine*, *111*. https://doi.org/10.1016/j.compbiomed.2019.103345

Deepak, S., & Ameer, P. M. (2020). Retrieval of brain MRI with tumor using contrastive loss based similarity on GoogLeNet encodings. *Computers in Biology and Medicine*, *125*(September). https://doi.org/10.1016/j.compbiomed.2020.103993

Deng, W., Shi, Q., Luo, K., Yang, Y., & Ning, N. (2019). Brain Tumor Segmentation Based on Improved Convolutional Neural Network in Combination with Non-quantifiable Local Texture Feature. *Journal of Medical Systems*, *43*(6). https://doi.org/10.1007/s10916-019-1289-2

Díaz-Pernas, F. J., Martínez-Zarzuela, M., González-Ortega, D., & Antón-Rodríguez, M. (2021). A Deep Learning approach for brain tumor classification and segmentation using a multiscale Convolutional Neural Network. *Healthcare (Switzerland)*, *9*(2). https://doi.org/10.3390/healthcare9020153

Dolz, J., Desrosiers, C., & Ben Ayed, I. (2018). 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. *NeuroImage*, *170*, 456–470. https://doi.org/https://doi.org/10.1016/j.neuroimage.2017.04.039

Dunn Jr, W. D., Hwang, S. N., Cooper, L. A., Aerts, H. J. W. L., & Holder, C. A. (2016). Assessing the Effects of Software Platforms on Volumetric Segmentation of Glioblastoma. *Journal of Neuroimaging in Psychiatry and Neurology*, 64–72. https://doi.org/10.17756/jnpn.2016-008

Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). Synthetic data augmentation using GAN for improved liver lesion classification. *Proceedings - International Symposium on Biomedical Imaging*, *2018-April*, 289–293. https://doi.org/10.1109/ISBI.2018.8363576

Fuad, M. S., Anam, C., Adi, K., & Dougherty, G. (2021). Comparison of two Convolutional Neural Network models for automated classification of brain cancer types. *AIP Conference Proceedings*, *2346*(March). https://doi.org/10.1063/5.0047750

Ge, C., Gu, I. Y.-H., Jakola, A. S. S., & Yang, J. (2018). Deep Learning and Multi-Sensor Fusion for Glioma Classification Using Multistream 2D Convolutional Networks. *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, *2018*, 5894–5897. https://doi.org/10.110 Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media. (2019).9/EMBC.2018.8513556

Ge, Chenjie, Gu, I. Y.-H., Store Jakola, A., & Yang, J. (2019). Cross-Modality Augmentation of Brain Mr Images Using a Novel Pairwise Generative Adversarial Network for Enhanced Glioma Classification. *2019 IEEE International Conference on Image Processing (ICIP)*, 559–563. https://doi.org/10.1109/icip.2019.8803808

Ghassemi, N, Shoeibi, A., & Rouhani, M. (2020). Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images. *Biomedical Signal Processing and Control*, *57*. https://doi.org/10.1016/j.bspc.2019.101678

Gonella, G., Binaghi, E., Nocera, P., & Mordacchini, C. (2019). Investigating the behaviour of Machine Learning techniques to segment brain metastases in radiation therapy planning. *Applied Sciences (Switzerland)*, *9*(16). https://doi.org/10.3390/app9163335

Gottapu, R. D., & Dagli, C. H. (2018). DenseNet for Anatomical Brain Segmentation. *Procedia Computer Science*, *140*, 179–185.

https://doi.org/https://doi.org/10.1016/j.procs.2018.10.327

Guido, S. (2016). *Introduction to Machine Learning with Python*.

Guo, L., Zhao, L., Wu, Y., Li, Y., Xu, G., & Yan, Q. (2011). Tumor detection in MR images using one-class immune feature weighted SVMs. *IEEE Transactions on Magnetics*, *47*(10), 3849–3852. https://doi.org/10.1109/TMAG.2011.2158520

Han, L., & Kamdar, M. R. (2015). MRI to MGMT: predicting methylation status in glioblastoma patients using convolutional recurrent neural networks. *Anal Chem.*, *25*(4), 368–379. https://doi.org/10.1142/9789813235533_0031

Hang, S. T., & Aono, M. (2017). Bi-linearly weighted fractional max pooling: An extension to conventional max pooling for deep Convolutional Neural Network. *Multimedia Tools and Applications*, *76*(21), 22095–22117. https://doi.org/10.1007/s11042-017-4840-5

Hara, K., Kataoka, H., & Satoh, Y. (2018). Can Spatiotemporal 3D CNNs Retrace the History of 2D CNNs and ImageNet? *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 6546–6555. https://doi.org/10.1109/CVPR.2018.00685

Hashemzehi, R., Mahdavi, S. J. S., Kheirabadi, M., & Kamel, S. R. (2020). Detection of brain tumors from MRI images base on Deep Learning using hybrid model CNN and NADE. *Biocybernetics and Biomedical Engineering*, *40*(3), 1225–1232. https://doi.org/10.1016/j.bbe.2020.06.001

Havaei, M., Davy, A., Warde-farley, D., Biard, A., Courville, A., Bengio, Y., … Larochelle, H. (2017). Brain tumor segmentation with Deep Neural Networks. *Medical Image Analysis*, *35*, 18–31. https://doi.org/10.1016/j.media.2016.05.004

He, K., Zhang, X., Ren, S., & Sun, J. (2014). Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classificatio. *Microsoft Research*, 1026–1034. https://doi.org/10.1109/ICCV.2015.123

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem*, 770–778.

https://doi.org/10.1109/CVPR.2016.90

Hoseini, F., Shahbahrami, A., & Bayat, P. (2018). An Efficient Implementation of Deep Convolutional Neural Networks for MRI Segmentation. *Journal of Digital Imaging*, *31*(5), 738–747. https://doi.org/10.1007/s10278-018-0062-2

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-Excitation Networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 7132–7141. https://doi.org/10.1109/CVPR.2018.00745

Hu, Y., & Xia, Y. (2018). 3D deep neural network-based brain tumor segmentation using multimodality magnetic resonance sequences. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-75238-9_36

Huang, E., Gutman, D. A., Jilwan-Nicolas, M., Hwang, S. N., Jain, R., Rubin, D., … Wintermark, M. (2014). Imaging genomic mapping of an invasive MRI phenotype predicts patient outcome and metabolic dysfunction: a TCGA glioma phenotype research group project. *BMC Medical Genomics*, *7*(1), 1–9. https://doi.org/10.1186/1755-8794-7-30

Iqbal, S., Ghani, M. U., Saba, T., & Rehman, A. (2018). Brain tumor segmentation in multi-spectral MRI using Convolutional Neural Networks (CNN). *Microscopy Research and Technique*, *81*(4), 419–427. https://doi.org/10.1002/jemt.22994

Irmak, E. (2021). Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework. *Iranian Journal of Science and Technology - Transactions of Electrical Engineering*, *45*(3), 1015–1036. https://doi.org/10.1007/s40998-021-00426-9

Işin, A., Direkoğlu, C., & Şah, M. (2016). Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods. *Procedia Computer Science*, *102*(August), 317–324. https://doi.org/10.1016/j.procs.2016.09.407

Islam, M., & Ren, H. (2018). Multi-modal PixelNet for brain tumor segmentation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-75238-9_26

Isselmou, A. E. K., Xu, G., Zhang, S., Saminu, S., & Javaid, I. (2019). Deep Learning algorithm for brain tumor detection and analysis using MR brain images. *ACM International Conference Proceeding Series*, 28–32. https://doi.org/10.1145/3348416.3348421

Johnson, D. R., Guerin, J. B., Giannini, C., Morris, J. M., Eckel, L. J., & Kaufmann, T. J. (2017). 2016 updates to the WHO brain tumor classification system: What the radiologist needs to know. *Radiographics*, *37*(7), 2164–2180. https://doi.org/10.1148/rg.2017170037

Jothi, N. V. S. N. and J. A. A. (2015). Automatic Classification of Brain MRI Images Using SVM and Neural Network Classifiers. *Advances in Intelligent Systems and Computing*, *320*, 621–631. https://doi.org/10.1007/978-3-319-11218-3

Jyoti Patil, M. S., & Pradeepini, G. (2019). Development of Deep Learning algorithm for brain tumor segmentation. *International Journal of Engineering and Advanced Technology*, *9*(1), 2800–2803. https://doi.org/10.35940/ijeat.A9784.109119

Kader, I. A. El, Xu, G., Shuai, Z., Saminu, S., Javaid, I., & Ahmad, I. S. (2021). Differential deep Convolutional Neural Network model for brain tumor classification. *Brain Sciences*, *11*(3). https://doi.org/10.3390/brainsci11030352

Kamnitsas, K., Ledig, C., Newcombe, V. F. J. J., Simpson, J. P., Kane, A. D., Menon, D. K., … Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, *36*, 61–78. https://doi.org/10.1016/j.media.2016.10.004

Kanas, V. G., Zacharaki, E. I., Thomas, G. A., Zinn, P. O., Megalooikonomou, V., & Colen, R. R. (2017). Learning MRI-based classification models for MGMT methylation status prediction in glioblastoma. *Computer Methods and Programs in Biomedicine*, *140*, 249–257. https://doi.org/10.1016/j.cmpb.2016.12.018

Kang, J., Ullah, Z., & Gwak, J. (2021). Mri-based brain tumor classification using ensemble of deep features and Machine Learning classifiers. *Sensors*, *21*(6), 1–21. https://doi.org/10.3390/s21062222

Kermi, A., Mahmoudi, I., & Khadir, M. T. T. (2019). Deep Convolutional Neural Networks using U-Net for automatic brain tumor segmentation in multimodal MRI

volumes. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-030-11726-9_4

Khan, A. R., Khan, S., Harouni, M., Abbasi, R., Iqbal, S., & Mehmood, Z. (2021). Brain tumor segmentation using K-means clustering and Deep Learning with synthetic data augmentation for classification. *Microscopy Research and Technique*, (November 2020), 1–11. https://doi.org/10.1002/jemt.23694

Khan, H. A., Jue, W., Mushtaq, M., & Mushtaq, M. U. (2020). Brain tumor classification in MRI image using Convolutional Neural Network. *Mathematical Biosciences and Engineering*, *17* (5), 6203–6216. https://doi.org/10.3934/MBE.2020328

Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–13.

Kirby, J., Colen, R., Rubin, D. L., Hu, Y., Buetow, K., Mikkelsen, T., … Meerzaman, D. (2014). Addition of MR imaging features and genetic biomarkers strengthens glioblastoma survival prediction in TCGA patients. *Journal of Neuroradiology*, *42*(4), 212–221. https://doi.org/10.1016/j.neurad.2014.02.006

Kirby, J., Jaffe, C. C., Poisson, L. M., Mikkelsen, T., Flanders, A., Rao, A., … Freymann, J. (2014). Outcome Prediction in Patients with Glioblastoma by Using Imaging, Clinical, and Genomic Biomarkers: Focus on the Nonenhancing Component of the Tumor. *Radiology*, *272* (2), 484 – 493. https://doi.org/10.1148/radiol.14131691

Kokkalla, S., Kakarla, J., Venkateswarlu, I. B., & Singh, M. (2021). Three-class brain tumor classification using deep dense inception residual network. *Soft Computing*, (2019). https://doi.org/10.1007/s00500-021-05748-8

Kong, X., Sun, G., Wu, Q., Liu, J., & Lin, F. (2018). Hybrid pyramid u-net model for brain tumor segmentation. In *IFIP Advances in Information and Communication Technology* (Vol. 538). https://doi.org/10.1007/978-3-030-00828-4_35

Kong, Y., Gao, J., Xu, Y., Pan, Y., Wang, J., & Liu, J. (2019). Classification of autism

spectrum disorder by combining brain connectivity and deep neural network classifier. *Neurocomputing*, *324*, 63–68. https://doi.org/10.1016/j.neucom.2018.04.080

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, *2*, 1097–1105.

Kumari, R., & Kr., S. (2017). Machine Learning: A Review on Binary Classification. *International Journal of Computer Applications*, *160*(7), 11–15. https://doi.org/10.5120/ijca2017913083

Kuzina, A., Egorov, E., & Burnaev, E. (2019). Bayesian generative models for knowledge transfer in MRI semantic segmentation problems. *Frontiers in Neuroscience*, *13*(JUL). https://doi.org/10.3389/fnins.2019.00844

Kwon, D., Shinohara, R. T., Akbari, H., & Davatzikos, C. (2014). Combining generative models for multifocal glioma segmentation and registration. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8673 LNCS*(PART 1), 763–770. https://doi.org/10.1007/978-3-319-10404-1_95

LAKSHMI, V. K., FEROZ, C. A., & MERLIN, J. A. J. (2018). Automated Detection and Segmentation of Brain Tumor Using Genetic Algorithm. *2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 583–589. https://doi.org/10.1109/ICSSIT.2018.8748487

Le Reste, P.-J., Stindel, E., Morvan, Y., Upadhaya, T., & Hatt, M. (2016). Prognosis classification in glioblastoma multiforme using multimodal MRI derived heterogeneity textural features: impact of pre-processing choices. *Medical Imaging 2016: Computer-Aided Diagnosis*, *9785*, 97850W. https://doi.org/10.1117/12.2217151

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lee, J. K., Wang, J., Sa, J. K., Ladewig, E., Lee, H. O., Lee, I. H., … Nam, D. H. (2017). Spatiotemporal genomic architecture informs precision oncology in glioblastoma.

Nature Genetics, *49*(4), 594–599. https://doi.org/10.1038/ng.3806

Li, H., Li, A., & Wang, M. (2019). A novel end-to-end brain tumor segmentation method using improved fully convolutional networks. *Computers in Biology and Medicine*, *108*, 150–160. https://doi.org/10.1016/j.compbiomed.2019.03.014

Li, Hongwei, Jiang, G., Zhang, J., Wang, R., Wang, Z., Zheng, W.-S., & Menze, B. (2018). Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images. *NeuroImage*, *183*, 650–665. https://doi.org/https://doi.org/10.1016/j.neuroimage.2018.07.005

Li, J., Yu, Z. L., Gu, Z., Liu, H., & Li, Y. (2019). MMAN: Multi-modality aggregation network for brain segmentation from MR images. *Neurocomputing*, *358*, 10–19. https://doi.org/https://doi.org/10.1016/j.neucom.2019.05.025

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18–22.

Lin, M., Chen, Q., & Yan, S. (2014). Network in network. *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings*, 1–10.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., … Sánchez, C. I. (2017). A survey on Deep Learning in medical image analysis. *Medical Image Analysis*, *42*, 60 –88. https://doi.org/https://doi.org/10.1016/j.media.2017.07.005

Liu, D., Zhang, H., Zhao, M., Yu, X., Yao, S., & Zhou, W. (2018). Brain Tumor Segmention Based on Dilated Convolution Refine Networks. *Proceedings - 2018 IEEE/ACIS 16th International Conference on Software Engineering Research, Management and Application, SERA 2018*, 113–120. https://doi.org/10.1109/SERA.2018.8477213

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*(October 2016), 11–26. https://doi.org/10.1016/j.neucom.2016.12.038

Lu, S., Lu, Z., & Zhang, Y.-D. (2019). Pathological brain detection based on AlexNet and transfer learning. *Journal of Computational Science*, *30*, 41–47.

https://doi.org/https://doi.org/10.1016/j.jocs.2018.11.008

Maier, A., Syben, C., Lasser, T., & Riess, C. (2019). A gentle introduction to Deep Learning in medical image processing. *Zeitschrift Fur Medizinische Physik*, *29*(2), 86–101. https://doi.org/10.1016/j.zemedi.2018.12.003

Mallick, P. K., Ryu, S. H., Satapathy, S. K., Mishra, S., Nguyen, N. G., & Tiwari, P. (2019). Brain MRI ImageClassification for Cancer Detection using Deep Wavelet Autoencoder based Deep Neural Network. *IEEE Access*, *PP*(c), 1–1. https://doi.org/10.1109/access.2019.2902252

Manikanta, Y., Yashwanth, B. S., Deepthimurthy, T. S., & Santosh, H. G. (2021). *Detection Of Brain Tumour Using Convolutional Neuralnetwork*. *12*(12), 2330–2338.

Mazurowski, M. A., Zhang, J., Peters, K. B., & Hobbs, H. (2014). Computer-extracted MR imaging features are associated with survival in glioblastoma patients. *Journal of Neuro-Oncology*, *120*(3), 483–488. https://doi.org/10.1007/s11060-014-1580-5

Mengash, H. A., & Hosni Mahmoud, H. A. (2021). Brain cancer tumor classification from motion-corrected MRI images using Convolutional Neural Network. *Computers, Materials and Continua*, *68*(2), 1551–1563. https://doi.org/10.32604/cmc.2021.016907

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., … Van Leemput, K. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, *34*(10), 1993–2024. https://doi.org/10.1109/TMI.2014.2377694

Menze, B., Jakab, A., Bauer, S., Kalpathy-cramer, J., Farahani, K., Kirby, J., … Leemput, K. Van. (2014). *Benchmark (BRATS) To cite this version : HAL Id : hal-00935640 The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)*.

Milletari, F., Navab, N., & Ahmadi, S. A. (2016). V-Net: Fully Convolutional Neural Networks for volumetric medical image segmentation. *Proceedings - 2016 4th International Conference on 3D Vision, 3DV 2016*, 565–571. https://doi.org/10.1109/3DV.2016.79

Mittal, M., Goyal, L. M., Kaur, S., Kaur, I., Verma, A., & Jude Hemanth, D. (2019). Deep Learning based enhanced tumor segmentation approach for MR brain images. *Applied Soft Computing Journal*, *78*, 346–354. https://doi.org/10.1016/j.asoc.2019.02.036

Mittal, Mamta, Goyal, L. M., Kaur, S., Kaur, I., Verma, A., & Jude Hemanth, D. (2019). Deep Learning based enhanced tumor segmentation approach for MR brain images. *Applied Soft Computing Journal*, *78*, 346–354. https://doi.org/10.1016/j.asoc.2019.02.036

Mlynarski, P., Delingette, H., Criminisi, A., & Ayache, N. (2019). 3D Convolutional Neural Networks for tumor segmentation using long-range 2D context. *Computerized Medical Imaging and Graphics*, *73*, 60–72. https://doi.org/10.1016/j.compmedimag.2019.02.001

Nabizadeh, N., & Kubat, M. (2015). Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features. *Computers and Electrical Engineering*, *45*, 286–301. https://doi.org/10.1016/j.compeleceng.2015.02.007

naceur, M. Ben, Saouli, R., Akil, M., & Kachouri, R. (2018). Fully Automatic Brain Tumor Segmentation using End-To-End Incremental Deep Neural Networks in MRI images. *Computer Methods and Programs in Biomedicine*, *166*, 39–49. https://doi.org/10.1016/j.cmpb.2018.09.007

Nema, S., Dudhane, A., Murala, S., & Naidu, S. (2020). RescueNet: An unpaired GAN for brain tumor segmentation. *Biomedical Signal Processing and Control*, *55*. https://doi.org/10.1016/j.bspc.2019.101641

Nema, Shubhangi, Dudhane, A., Murala, S., & Naidu, S. (2020). RescueNet: An unpaired GAN for brain tumor segmentation. *Biomedical Signal Processing and Control*, *55*, 101641. https://doi.org/10.1016/j.bspc.2019.101641

Özyurt, F., Sert, E., & Avcı, D. (2020). An expert system for brain tumor detection: Fuzzy C-means with super resolution and Convolutional Neural Network with extreme learning machine. *Medical Hypotheses*, *134*. https://doi.org/10.1016/j.mehy.2019.109433

Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on*

Knowledge and Data Engineering, *22* (10), 1345 – 1359. https://doi.org/10.1109/TKDE.2009.191

ParthaSarathi, M., & Ansari, M. A. (2018). Multimodal Retrieval Framework for Brain Volumes in 3D MR Volumes. *Journal of Medical and Biological Engineering*, *38*(2), 261–272. https://doi.org/10.1007/s40846-017-0287-4

Pashaei, A., Sajedi, H., & Jazayeri, N. (2018). Brain tumor classification via Convolutional Neural Network and extreme learning machines. *2018 8th International Conference on Computer and Knowledge Engineering, ICCKE 2018*, 314–319. https://doi.org/10.1109/ICCKE.2018.8566571

Paul, J. S., Plassard, A. J., Landman, B. A., & Fabbri, D. (2017). Deep Learning for brain tumor classification. *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, *10137*(2), 1013710. https://doi.org/10.1117/12.2254195

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24* (3), 45 – 77. https://doi.org/10.2753/MIS0742-1222240302

Pereira, S., Meier, R., McKinley, R., Wiest, R., Alves, V., Silva, C. A., & Reyes, M. (2018). Enhancing interpretability of automatically extracted Machine Learning features: application to a RBM-Random Forest system on brain lesion segmentation. *Medical Image Analysis*, *44*, 228–244. https://doi.org/https://doi.org/10.1016/j.media.2017.12.009

Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). *Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images*. *35*(5), 1240–1251.

Pinto, A., Pereira, S., Rasteiro, D., & Silva, C. A. (2018). Hierarchical brain tumour segmentation using extremely randomized trees. *Pattern Recognition*, *82*, 105–117. https://doi.org/https://doi.org/10.1016/j.patcog.2018.05.006

Qamar, S., Jin, H., Zheng, R., & Ahmad, P. (2018). 3D Hyper-Dense Connected Convolutional Neural Network for Brain Tumor Segmentation. *Proceedings - 2018 14th International Conference on Semantics, Knowledge and Grids, SKG*

2018, 123–130. https://doi.org/10.1109/SKG.2018.00024

Ramirez, I., Martin, A., & Schiavi, E. (2018). Optimization of a variational model using Deep Learning: An application to brain tumor segmentation. *Proceedings - International Symposium on Biomedical Imaging*, *2018-April*(Isbi), 631–634. https://doi.org/10.1109/ISBI.2018.8363654

Rehman, A., Naz, S., Razzak, M. I., Akram, F., & Imran, M. (2019). A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning. *Circuits, Systems, and Signal Processing*. https://doi.org/10.1007/s00034-019-01246-3

Reza, S. M. S., Mays, R., & Iftekharuddin, K. M. (2015). Multi-fractal detrended texture feature for brain tumor classification. *Medical Imaging 2015: Computer-Aided Diagnosis*, *9414*, 941410. https://doi.org/10.1117/12.2083596

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9351*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Systems with Applications*, *42*(3), 990–1002. https://doi.org/10.1016/j.eswa.2014.09.020

Roy, S., & Maji, P. (2018). An accurate and robust skull stripping method for 3-D magnetic resonance brain images. *Magnetic Resonance Imaging*, *54*, 46–57. https://doi.org/https://doi.org/10.1016/j.mri.2018.07.014

Rubin, D. L., Westbroek, E. M., Gevaert, O., Achrol, A. S., Rodriguez, S., Loya, J. J., … Feroze, A. H. (2015). Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities. *Science Translational Medicine*, *7*(303), 303ra138-303ra138. https://doi.org/10.1126/scitranslmed.aaa7582

Saba, T., Mohamed, A. S., El-Affendi, M., Amin, J., & Sharif, M. (2020). Brain tumor detection using fusion of hand crafted and Deep Learning features. *Cognitive*

Systems Research, *59*, 221–230. https://doi.org/https://doi.org/10.1016/j.cogsys.2019.09.007

Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A., & Baik, S. W. (2019). Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *Journal of Computational Science*, *30*, 174–182. https://doi.org/10.1016/j.jocs.2018.12.003

Saman, S., & Jamjala Narayanan, S. (2019). Survey on brain tumor segmentation and feature extraction of MR images. *International Journal of Multimedia Information Retrieval*, *8*(2), 79–99. https://doi.org/10.1007/s13735-018-0162-2

Sane, U. (2021). *AND ENGINEERING TRENDS ADVANCED BOOST BRAIN TUMOR CLASSIFICATION WITH RANDOM TREE & KNN SEGMENTATION*. *6*(2), 11–17.

Saxena, N., Sharma, R., Joshi, K., & Rana, H. S. (2019). Identification of glioma from MR images using Convolutional Neural Network. In *Advances in Intelligent Systems and Computing* (Vol. 880). https://doi.org/10.1007/978-3-030-02686-8_44

Schmainda, K. M., Prah, M. A., Rand, S. D., Liu, Y., Logan, B., Muzi, M., … Quarles, C. C. (2018). Multisite concordance of DSC-MRI analysis for brain tumors: Results of a National Cancer Institute Quantitative Imaging Network Collaborative Project. *American Journal of Neuroradiology*, *39*(6), 1008–1016. https://doi.org/10.3174/ajnr.A5675

Selvikvåg Lundervold, A., & Lundervold, A. (2018). An overview of Deep Learning in medical imaging focusing on MRI. *Zeitschrift Fur Medizinische Physik*, (Mmiv). https://doi.org/10.1016/j.zemedi.2018.11.002

Sharif, M. I., Li, J. P., Khan, M. A., & Saleem, M. A. (2019). Active Deep neural Network Features Selection for Segmentation and Recognition of Brain Tumors using MRI Images. *Pattern Recognition Letters*. https://doi.org/https://doi.org/10.1016/j.patrec.2019.11.019

Sheela, C. J. J., & Suganthi, G. (2019). Automatic Brain Tumor Segmentation from MRI using Greedy Snake Model and Fuzzy C-Means Optimization. *Journal of*

King Saud University - Computer and Information Sciences. https://doi.org/https://doi.org/10.1016/j.jksuci.2019.04.006

Shin, H. C., Tenenholtz, N. A., Rogers, J. K., Schwarz, C. G., Senjem, M. L., Gunter, J. L., … Michalski, M. (2018). Medical image synthesis for data augmentation and anonymization using generative adversarial networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11037 LNCS*, 1–11. https://doi.org/10.1007/978-3-030-00536-8_1

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.

Smith-Bindman, R., Kwan, M. L., Marlow, E. C., Theis, M. K., Bolch, W., Cheng, S. Y., … Miglioretti, D. L. (2019). Trends in Use of Medical Imaging in US Health Care Systems and in Ontario, Canada, 2000-2016. *JAMA - Journal of the American Medical Association*, *322*(9), 843–856. https://doi.org/10.1001/jama.2019.11456

Soltaninejad, M., Zhang, L., Lambrou, T., Yang, G., Allinson, N., & Ye, X. (2018). MRI brain tumor segmentation and patient survival prediction using random forests and fully convolutional networks. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-75238-9_18

Soltaninejad, Mohammadreza, Yang, G., Lambrou, T., Allinson, N., Jones, T. L., Barrick, T. R., … Ye, X. (2018). Supervised learning based multimodal MRI brain tumour segmentation using texture features from supervoxels. *Computer Methods and Programs in Biomedicine*, *157*, 69–84. https://doi.org/10.1016/j.cmpb.2018.01.003

Sultan, H. H., Salem, N. M., & Al-Atabany, W. (2019). Multi-Classification of Brain Tumor Images Using Deep Neural Network. *IEEE Access*, *7*, 69215–69225. https://doi.org/10.1109/ACCESS.2019.2919122

Sun, J., Chen, W., Peng, S., & Liu, B. (2019). DRRNet: Dense Residual Refine Networks for Automatic Brain Tumor Segmentation. *Journal of Medical Systems*,

43(7). https://doi.org/10.1007/s10916-019-1358-6

Sun, L., Zhang, S., Chen, H., & Luo, L. (2019). Brain tumor segmentation and survival prediction using multimodal MRI scans with Deep Learning. *Frontiers in Neuroscience*, *13*(JUL). https://doi.org/10.3389/fnins.2019.00810

Swati, Z. N. K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., & Lu, J. (2019). Brain tumor classification for MR images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, *75*, 34–46. https://doi.org/10.1016/j.compmedimag.2019.05.001

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., … Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *07-12-June*, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

Takacs, P., & Manno-Kovacs, A. (2018). MRI brain tumor segmentation combining saliency and convolutional network features. *Proceedings - International Workshop on Content-Based Multimedia Indexing*, *2018-Septe*, 1–6. https://doi.org/10.1109/CBMI.2018.8516544

Talo, Muhammed, Baloglu, U. B., Yıldırım, Ö., & Rajendra Acharya, U. (2019). Application of deep transfer learning for automated brain abnormality classification using MR images. *Cognitive Systems Research*, *54*, 176–188. https://doi.org/10.1016/j.cogsys.2018.12.007

Talo, Muhammed, Yildirim, O., Baloglu, U. B., Aydin, G., & Acharya, U. R. (2019). Convolutional Neural Networks for multi-class brain disease detection using MRI images. *Computerized Medical Imaging and Graphics*, *78*, 101673. https://doi.org/https://doi.org/10.1016/j.compmedimag.2019.101673

Thaha, M. M., Kumar, K. P. M., Murugan, B. S., Dhanasekeran, S., Vijayakarthick, P., & Selvi, A. S. (2019). Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *Journal of Medical Systems*, *43*(9), 1240–1251. https://doi.org/10.1007/s10916-019-1416-0

Thillaikkarasi, R., & Saravanan, S. (2019). An Enhancement of Deep Learning Algorithm for Brain Tumor Segmentation Using Kernel Based CNN with M-

SVM. *Journal of Medical Systems*, *43*(4), 1–7. https://doi.org/10.1007/s10916-019-1223-7

Tian, Q., Wang, L., Liu, Y., Li, B., Liang, Z., Gao, P., … Liu, Y. (2017). Relationship between Glioblastoma Heterogeneity and Survival Time: An MR Imaging Texture Analysis. *J Magn Reson Imaging*, *38*(9), 1695–1701. https://doi.org/10.1002/jmri.2596010.3174/ajnr.A5279

Trakoolwilaiwan, T., Behboodi, B., Lee, J., Kim, K., & Choi, J.-W. (2017). Convolutional Neural Network for high-accuracy functional near- infrared spectroscopy in a brain– computer interface. *Neurophoton*, *5*(1). https://doi.org/10.1117/1.NPh.5.1

Wachinger, C., Reuter, M., & Klein, T. (2018). DeepNAT: Deep Convolutional Neural Network for segmenting neuroanatomy. *NeuroImage*, *170*, 434–445. https://doi.org/https://doi.org/10.1016/j.neuroimage.2017.02.035

Wang, G., Li, W., Zuluaga, M. A., Pratt, R., Patel, P. A., Aertsen, M., … Vercauteren, T. (2018). Interactive Medical Image Segmentation Using Deep Learning with Image-Specific Fine Tuning. *IEEE Transactions on Medical Imaging*, *37*(7), 1562–1573. https://doi.org/10.1109/TMI.2018.2791721

Wang, G., Zuluaga, M. A. A., Li, W., Pratt, R., Patel, P. A. A., Aertsen, M., … Vercauteren, T. (2019). DeepIGeoS: A Deep Interactive Geodesic Framework for Medical Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *41*(7), 1559–1572. https://doi.org/10.1109/TPAMI.2018.2840695

Wang, S. H., Phillips, P., Sui, Y., Liu, B., Yang, M., & Cheng, H. (2018). Classification of Alzheimer's Disease Based on Eight-Layer Convolutional Neural Network with Leaky Rectified Linear Unit and Max Pooling. *Journal of Medical Systems*, *42*(5), 85. https://doi.org/10.1007/s10916-018-0932-7

Wang, S., Jiang, Y., Hou, X., Cheng, H., & Du, S. (2017). Cerebral Micro-Bleed Detection Based on the Convolution Neural Network with Rank Based Average Pooling. *IEEE Access*, *5*, 16576–16583. https://doi.org/10.1109/ACCESS.2017.2736558

Wang, W., Liang, D., Chen, Q., Iwamoto, Y., Han, X.-H., Zhang, Q., … Chen, Y.-W. (2020). *Medical Image Classification Using Deep Learning BT - Deep Learning in Healthcare: Paradigms and Applications* (Y.-W. Chen & L. C. Jain, eds.). https://doi.org/10.1007/978-3-030-32606-7_3

Wang, Y., Li, C., Zhu, T., & Zhang, J. (2019). Multimodal brain tumor image segmentation using WRN-PPNet. *Computerized Medical Imaging and Graphics*, *75*, 56–65. https://doi.org/10.1016/j.compmedimag.2019.04.001

Wiest, R., Aerts, H. J. W. L., Rios Velazquez, E., Meier, R., Reyes, M., Alexander, B., … Bauer, S. (2015). Fully automatic GBM segmentation in the TCGA-GBM dataset: Prognosis and correlation with VASARI features. *Scientific Reports*, *5*(1), 1–10. https://doi.org/10.1038/srep16822

Wong, W. H., & Shen, X. (1995). A Stochastic Approximation Method. *Statistics*.

Wu, M. N., Lin, C. C., & Chang, C. C. (2007). Brain tumor detection using color-based K-means clustering segmentation. *Proceedings - 3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2007.*, *2*, 245–248. https://doi.org/10.1109/IIHMSP.2007.4457697

Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017). Aggregated residual transformations for deep neural networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-Janua*, 5987–5995. https://doi.org/10.1109/CVPR.2017.634

Zhai, J., & Li, H. (2019). An Improved Full Convolutional Network Combined with Conditional Random Fields for Brain MR Image Segmentation Algorithm and its 3D Visualization Analysis. *Journal of Medical Systems*, *43*(9). https://doi.org/10.1007/s10916-019-1424-0

Zhang, J., Xie, Y., Wu, Q., & Xia, Y. (2019). Medical image classification using synergic Deep Learning. *Medical Image Analysis*, *54*, 10–19. https://doi.org/10.1016/j.media.2019.02.010

Zhang, L., & Ji, Q. (2011). A Bayesian Network Model for Automatic and Interactive Image Segmentation. *IEEE Transactions on Image Processing*, *20*(9), 2582–2593. https://doi.org/10.1109/tip.2011.2121080

Zhang, Y. D., Hou, X. X., Chen, Y., Chen, H., Yang, M., Yang, J., & Wang, S. H. (2018). Voxelwise detection of cerebral microbleed in CADASIL patients by leaky rectified linear unit and early stopping. *Multimedia Tools and Applications*, *77*(17), 21825–21845. https://doi.org/10.1007/s11042-017-4383-9

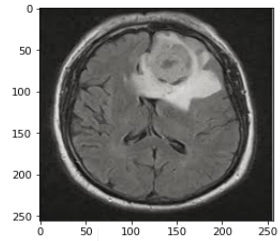
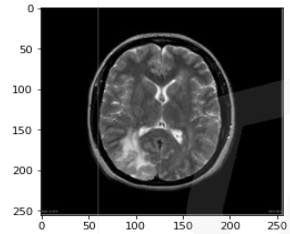
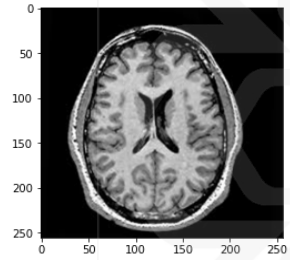
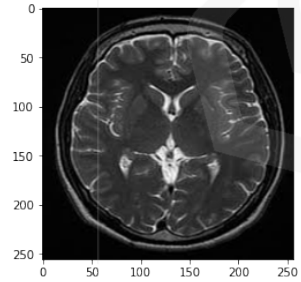
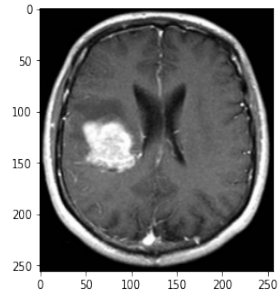
Zhang, Z., Odaibo, D., Skidmore, F. M., & Tanik, M. M. (2018). A big data analytics approach in medical imaging segmentation using deep Convolutional Neural Networks. In *Big Data and Visual Analytics*. https://doi.org/10.1007/978-3-319-63917-8_10

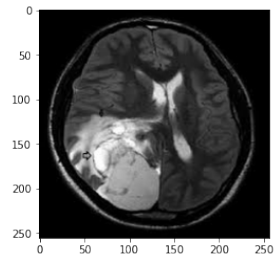
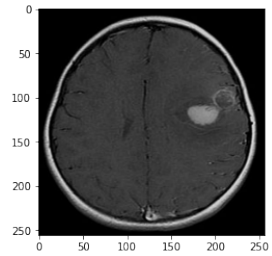
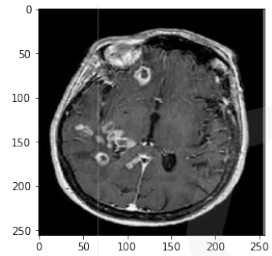
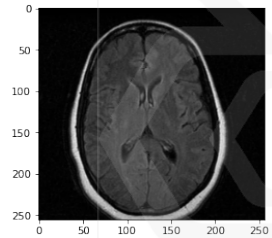
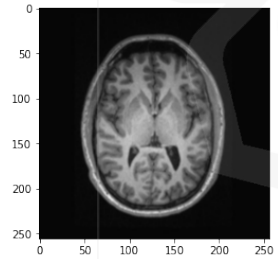
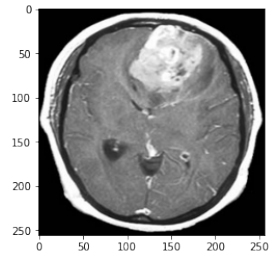
Zhao, X., Wu, Y., Song, G., Li, Z., Zhang, Y., & Fan, Y. (2018). A Deep Learning model integrating FCNNs and CRFs for brain tumor segmentation. *Medical Image Analysis*, *43*, 98–111. https://doi.org/https://doi.org/10.1016/j.media.2017.10.002

Zyad, M. A., Gouskir, M., & Bouikhalene, B. (2019). Classification of brain tumor from magnetic resonance imaging using Convolutional Neural Networks. *International Journal of Advanced Science and Technology*, *126*, 31–38. https://doi.org/10.33832/ijast.2019.126.04

APPENDIX A

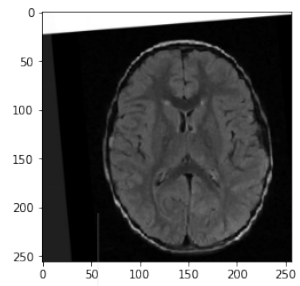
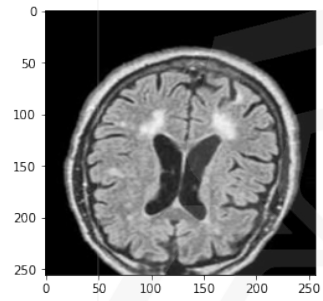
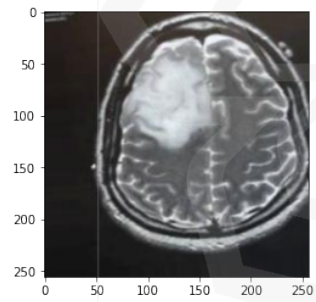
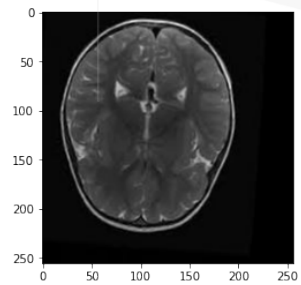
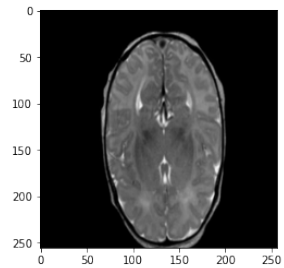
Experiment 1- Prediction of some random Binary MRI image dataset without data augmentation using (BTMIC-CNN)

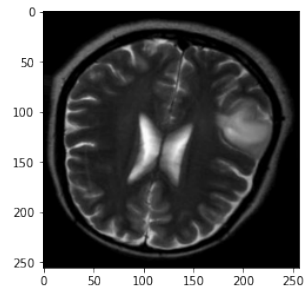
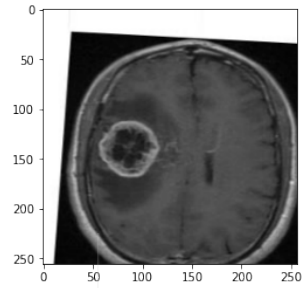
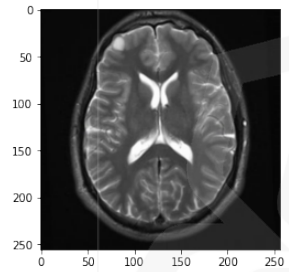
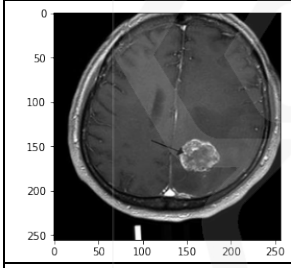
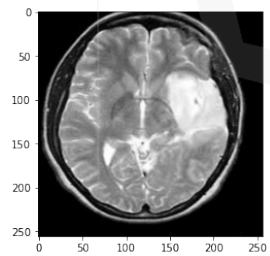
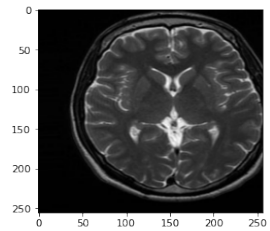
Image	Address	Actual	Prediction
	10	[0. 1.]	[[0.11873625 0.84457403]]
	20	[1. 0.]	[[1.0000000e+00 1.6856283e-08]]
	30	[1. 0.]	[[1.00000e+00 5.20205e-09]]
	33	[1. 0.]	[[9.9999797e-01 2.5061677e-06]]
	25	[0. 1.]	[[8.907078e-06 9.999901e-01]]

	15	[0. 1.]	[[0.7542042 0.23336393]]
	1	[0. 1.]	[[7.7638506e-05 9.9974924e-01]]
	27	[0. 1.]	[[0.01117427 0.98929554]]
	22	[1. 0.]	[[0.95701545 0.01986801]]
	17	[1. 0.]	[[0.9976101 0.00764243]]
	9	[0. 1.]	[[2.9045436e-08 1.0000000e+00]]

APPENDIX B

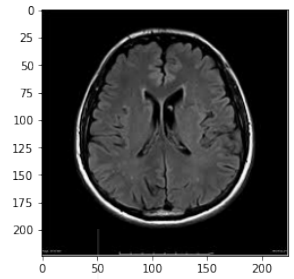
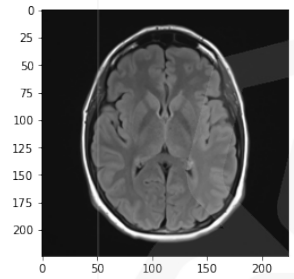
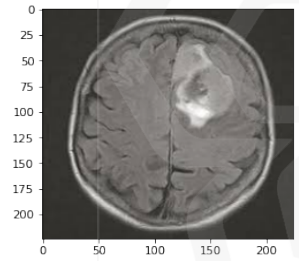
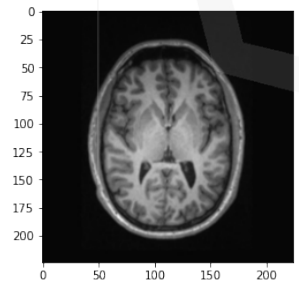
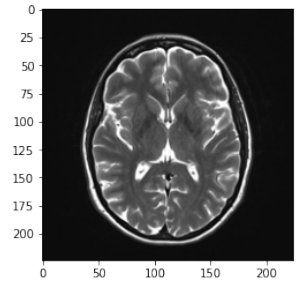
Experiment 2- Prediction of some random Binary MRI image dataset with data augmentation using (BTMIC-CNN)

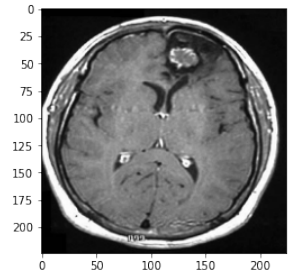
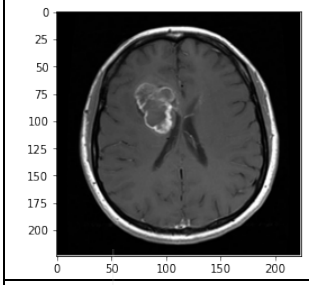
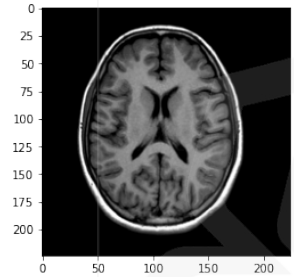
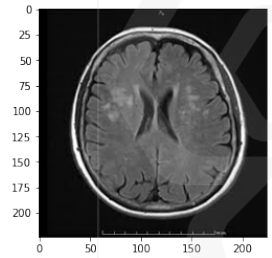
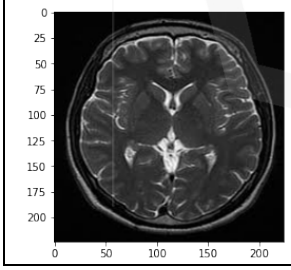
Image	Address	Actual	Prediction
	10	[1. 0.]	[[9.9999464e-01 4.8165662e-06]]
	100	[1. 0.]	[[9.9996257e-01 6.6268680e-05]]
	200	[0. 1.]	[[7.4002806e-08 9.9999988e-01]]
	166	[1. 0.]	[[1.0000000e+00 2.8693805e-07]]
	130	[1. 0.]	[[1.0000000e+00 4.1902276e-14]]

	300	[1. 0.]	[[9.9991047e-01 1.0159253e-04]]
	400	[0. 1.]	[[3.7777347e-11 1.0000000e+00]]
	350	[1. 0.]	[[1.000000e+00 3.090302e-08]]
	401	[0. 1.]	[[6.889129e-08 9.999999e-01]]
	220	[0. 1.]	[[1.559243e-05 9.999937e-01]]
	89	[1. 0.]	[[9.999999e-01 4.199458e-07]]

APPENDIX C

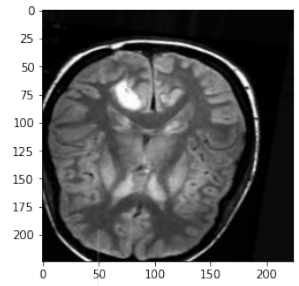
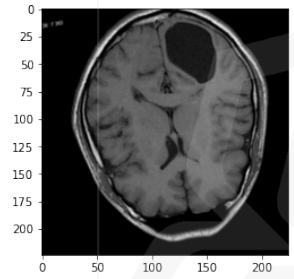
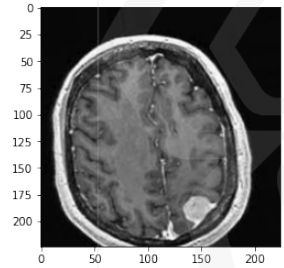
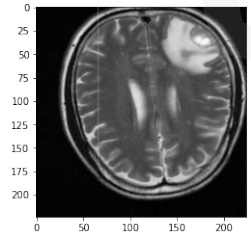
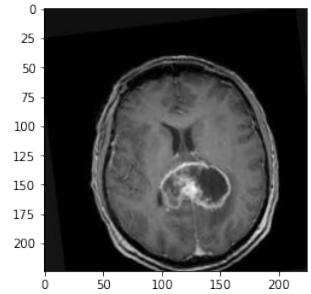
Experiment 3- Prediction of some random Binary MRI image dataset without data augmentation using (TL of VGG16)

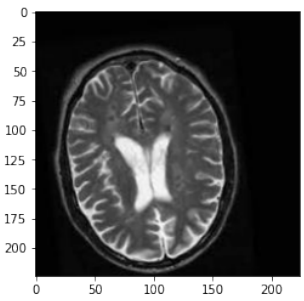
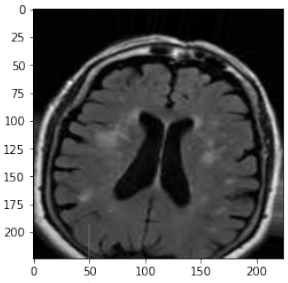
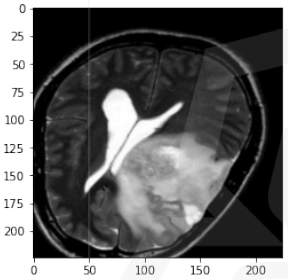
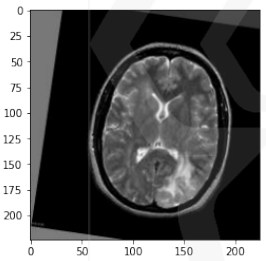
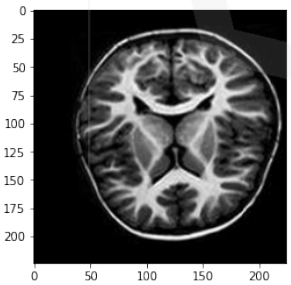
Image	Address	Actual	Prediction
	10	[0. 1.]	[[2.301599e-09 1.000000e+00]]
	9	[0. 1.]	[[1.3738847e-07 9.9999988e-01]]
	3	[1. 0.]	[[9.9997199e-01 2.9563942e-05]]
	8	[0. 1.]	[[9.957298e-06 9.999937e-01]]
	5	[0. 1.]	[[2.9456271e-05 9.9999046e-01]]

	16	[1. 0.]	[[9.9999893e-01 1.3422114e-06]]
	17	[1. 0.]	[[0.69432116 0.41049862]]
	13	[0. 1.]	[[2.8200074e-11 1.0000000e+00]]
	4	[0. 1.]	[[6.960827e-05 9.999578e-01]]
	7	[0. 1.]	[[2.7578457e-08 1.0000000e+00]]

APPENDIX D

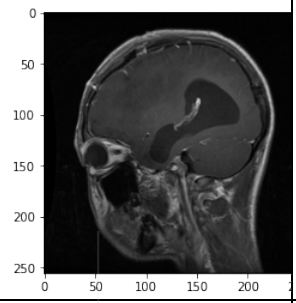
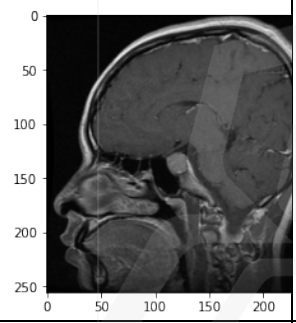
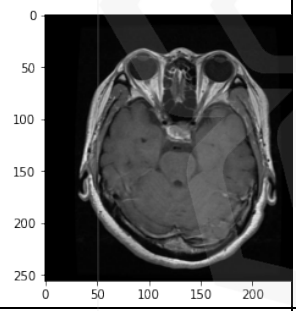
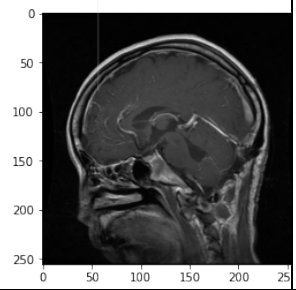
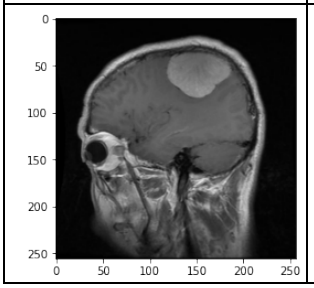
Experiment 4- Prediction of some random Binary MRI image dataset with data augmentation using (TL of VGG16)

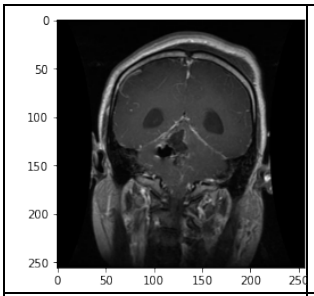
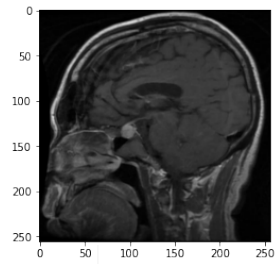
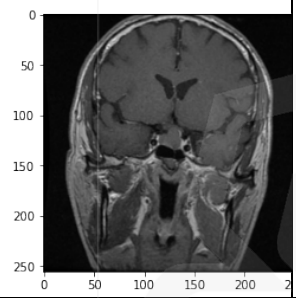
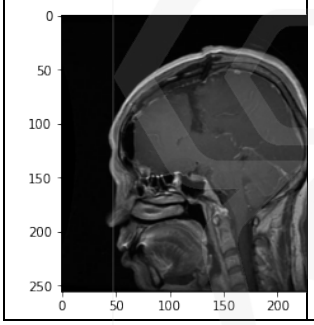
Image	Address	Actual	Prediction
	100	[1. 0.]	[[1.0000000e+00 4.7192855e-10]]
	64	[0. 1.]	[[1.9121908e-13 1.0000000e+00]]
	108	[1. 0.]	[[1.000000e+00 4.781294e-10]]
	27	[1. 0.]	[[1.0000000e+00 2.4672342e-10]]
	130	[1. 0.]	[[1.0000000e+00 2.5611335e-09]]

	125	[0. 1.]	[[1.247185e-17 1.000000e+00]]
	193	[0. 1.]	[[5.5115756e-14 1.0000000e+00]]
	179	[1. 0.]	[[1.0000000e+00 1.7440843e-13]]
	181	[0. 1.]	[[1.1711432e-18 1.0000000e+00]]
	172	[0. 1.]	[[1.1241518e-27 1.0000000e+00]]

APPENDIX E

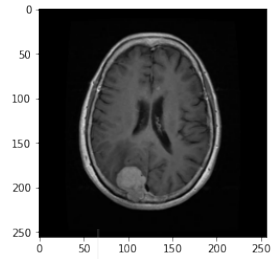
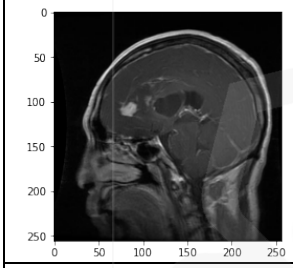
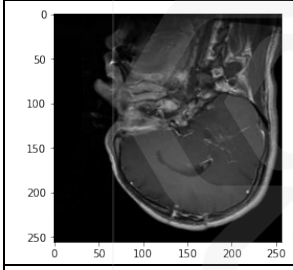
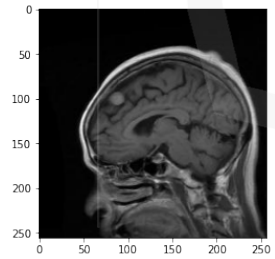
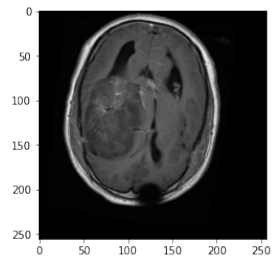
Experiment 5- Prediction of some random Binary MRI image dataset without data augmentation using (BTMIC-CNN)

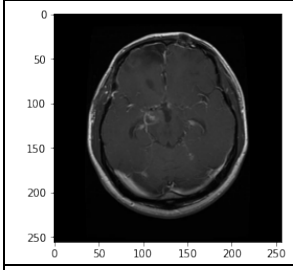
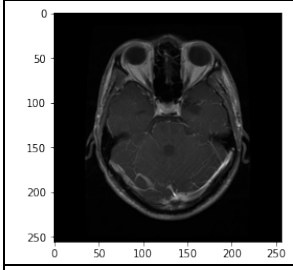
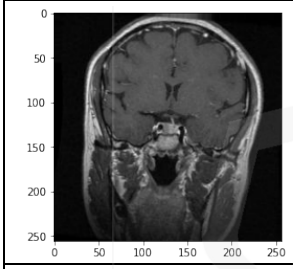
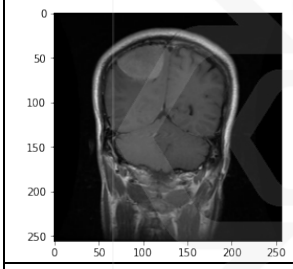
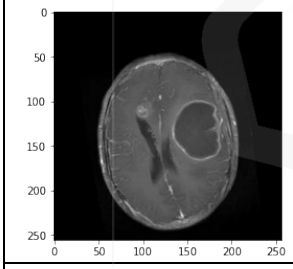
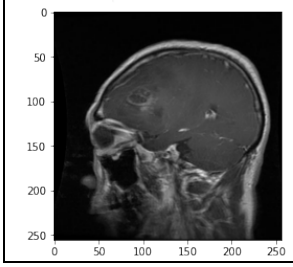
Image	Address	Actual	Prediction
	222	[1. 0. 0.]	[[1.00000000e+00 4.73646816e-12 1.08174564e-13]]
	333	[0. 0. 1.]	[[1.8402617e-10 2.3304723e-08 1.0000000e+00]] :
	540	[0. 0. 1.]	[[5.9212204e-09 5.6509872e-08 9.9999988e-01]]
	450	[1. 0. 0.]	[[1.0000000e+00 4.8893908e-13 5.7800766e-16]]
	612	[0. 1. 0.]	[[9.2343092e-01 7.6486655e-02 8.2466926e-05]]

	600	[1. 0. 0.]	[[1.0000000e+00 4.8416436e-17 1.9562445e-18]]
	590	[0. 0. 1.]	[[5.034775e-07 8.729625e-08 9.999994e-01]]
	390	[0. 0. 1.]	[[3.9615334e-18 5.4337746e-16 1.0000000e+00]] The actual label for this image is:
	239	[1. 0. 0.]	[[1.0000000e+00 3.2832503e-10 2.7844792e-12]]

APPENDIX F

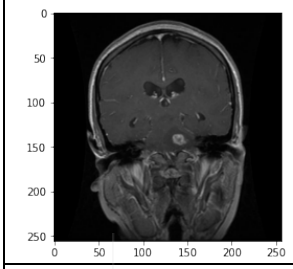
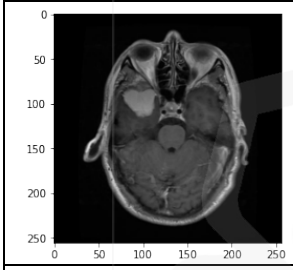
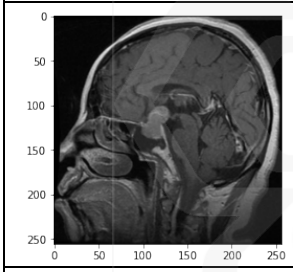
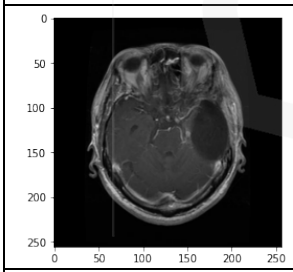
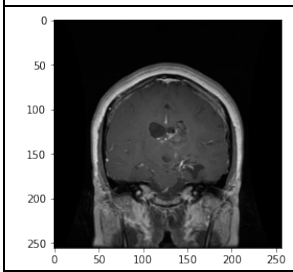
Experiment 6- Prediction of some random Binary MRI image dataset with data augmentation using (BTMIC-CNN)

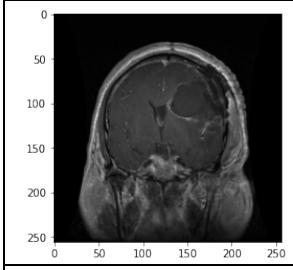
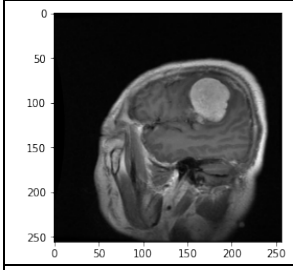
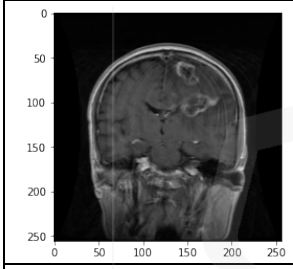
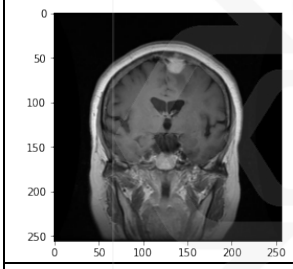
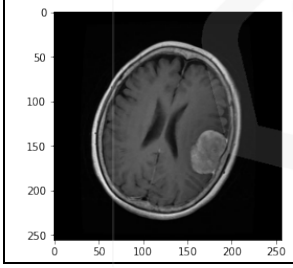
Image	Address	Actual	Prediction
	980	[0. 1. 0.]	[[1.2708116e-09 1.0000000e+00 5.0799018e-11]]
	1788	[1. 0. 0.]	[[9.9999845e-01 1.4127004e-06 1.2982900e-07]]
	1529	[1. 0. 0.]	[[9.6820962e-01 3.1662140e-02 1.2820223e-04]]
	1678	[0. 1. 0.]	[[0.02320074 0.9705729 0.00622631]]
	521	[1. 0. 0.]	[[0.3756106 0.19635317 0.42803624]]

	742	[1. 0. 0.]	[[9.9999988e-01 1.1578497e-07 2.5640204e-10]]
	1042	[1. 0. 0.]	[[9.9999988e-01 6.0138618e-08 7.2555884e-10]]
	1800	[0. 0. 1.]	[[2.3205418e-15 1.8030814e-14 1.0000000e+00]]
	1750	[0. 1. 0.]	[[7.0736322e-05 9.9992907e-01 2.4311134e-07]]
	721	[1. 0. 0.]	[[9.9999940e-01 5.6609446e-07 4.5863158e-10]]
	672	[1. 0. 0.]	[[1.0000000e+00 7.2926428e-15 2.8205462e-14]]

APPENDIX G

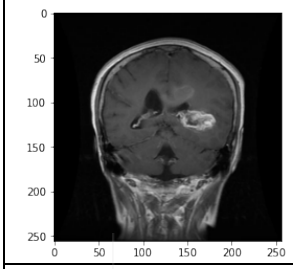
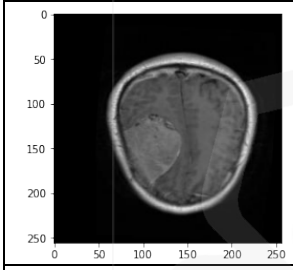
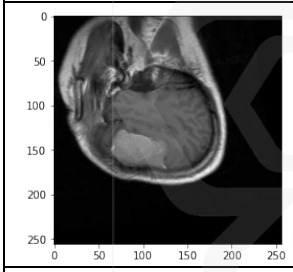
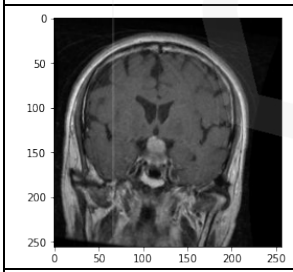
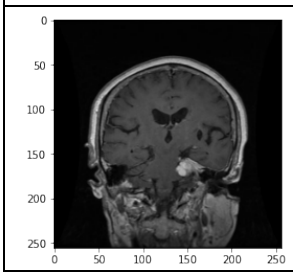
Experiment 7- Prediction of some random Binary MRI image dataset without data augmentation using (TL of VGG16)

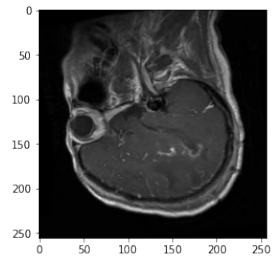
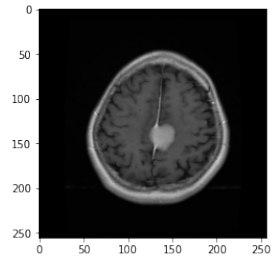
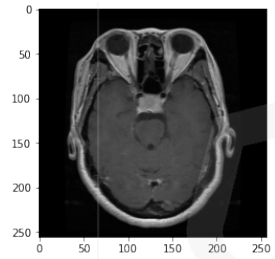
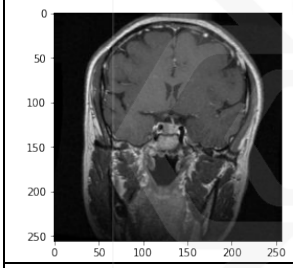
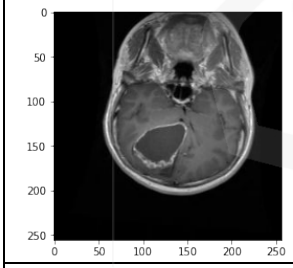
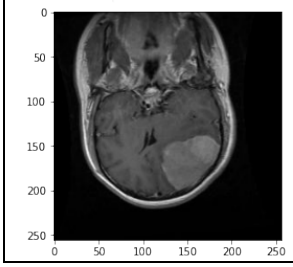
Image	Address	Actual	Prediction
	374	[1. 0. 0.]	[[9.9963343e-01 3.2440299e-04 4.2142623e-05]]
	71	[0. 1. 0.]	[[4.915585e-08 1.000000e+00 7.431027e-09]]
	159	[0. 0. 1.]	[[2.0065771e-09 9.2310574e-07 9.9999905e-01]]
	476	[1. 0. 0.]	[[1.0000000e+00 8.3575126e-16 8.0782492e-17]]
	378	[1. 0. 0.]	[[1.0000000e+00 1.3922032e-10 6.1089970e-12]]

	155	[1. 0. 0.]	[[1.0000000e+00 6.6016975e-12 1.8064112e-13]]
	564	[0. 1. 0.]	[[7.2310976e-04 9.9927622e-01 7.1340429e-07]]
	406	[1. 0. 0.]	[[9.9998724e-01 1.2388574e-05 4.0876114e-07]]
	242	[0. 1. 0.]	[[7.0847626e-08 9.9999988e-01 1.9557918e-08]]
	42	[0. 1. 0.]	[[8.1350054e-06 9.9999189e-01 3.5753267e-08]]

APPENDIX H

Experiment 8- Prediction of some random Binary MRI image dataset with data augmentation using (TL of VGG16)

Image	Address	Actual	Prediction
	1442	[1. 0. 0.]	[[1.8186584e-01 8.1813413e-01 4.9721741e-08]]
	266	[0. 1. 0.]	[[1.3810693e-07 9.9999988e-01 3.3426353e-09]] :
	1200	[0. 1. 0.]	[[3.3216392e-07 9.9999964e-01 2.1305974e-08]]
	660	[0. 0. 1.]	[[2.1699154e-16 2.9274386e-10 1.0000000e+00]]
	540	[0. 1. 0.]	[[5.1382125e-07 9.9999952e-01 3.3826426e-09]]

	78	[1. 0. 0.]	[[1.0000000e+00 4.4737942e-09 3.9011976e-09]]
	965	[0. 1. 0.]	[[1.7903016e-04 9.9982089e-01 1.7486765e-07]]
	875	[0. 0. 1.]	[[2.160393e-12 5.656467e-06 9.999944e-01]]
	1800	[0. 0. 1.]	[[2.3881614e-18 2.3176846e-11 1.0000000e+00]]
	730	[1. 0. 0.]	[[9.9999964e-01 4.1545209e-07 9.4279098e-12]]
	1756	[0. 1. 0.]	[[2.8923427e-04 4.0047742e-02 9.5966297e-01]]