A black background with white text

Description automatically generated with low confidence

MSc Data Science Project

7PAM2002-0509-2023

Department of Physics, Astronomy and Mathematics

**Data Science FINAL PROJECT REPORT**

**Project Title:**

Enhanced MRI Analysis for Brain Tumour Classification using Machine Learning

**Student Name and SRN:**

Umamah Usman

21088746

Supervisor: Alyssa Drake

Date Submitted: Enter the date you are submitting this report.

Word Count: Enter the word count

Declaration Statement

This report is submitted in partial fulfilment of the requirement for the degree of Master of Science in Data Science at the University of Hertfordshire.

I have read the guidance to students on academic integrity, misconduct and plagiarism information at [Assessment Offences and Academic Misconduct](https://www.herts.ac.uk/__data/assets/pdf_file/0007/237625/AS14-Apx3-Academic-Misconduct-v17.0.pdf) and understand the University process of dealing with suspected cases of academic misconduct and the possible penalties, which could include failing the project module or course.

I certify that the work submitted is my own and that any material derived or quoted from published or unpublished work of other persons has been duly acknowledged. (Ref. UPR AS/C/6.1, section 7 and UPR AS/C/5, section 3.6). I have not used chatGPT, or any other generative AI tool, to write the reportor code (other than where declared or referenced).

I did not use human participants or undertake a survey in my MSc Project.

I hereby give permission for the report to be made available on module websites provided the source is acknowledged.

Student Name printed: Umamah Usman

Student Name signature:

A close up of a name

Description automatically generated

Student SRN number: 21088746

UNIVERSITY OF HERTFORDSHIRE

SCHOOL OF PHYSICS, ENGINEERING AND COMPUTER SCIENCE

## Acknowledgement

As I am near the completion of my postgraduate studies, I want to reflect on this incredible learning journey and express my heartfelt thanks to everyone who has supported me along the way.

I am deeply grateful to Almighty God for His constant blessings and for providing me with the strength and confidence to pursue my goals with assurance.

I would like to extend my sincere appreciation to Alyssa Drake, my supervisor, for her invaluable guidance and support throughout this project. Her patience and willingness to entertain my numerous questions have been truly commendable.

I am also thankful to all my professors at the University of Hertfordshire, whose teachings have contributed to my knowledge and understanding of the subjects. Their assistance has been crucial throughout my course.

Lastly, I want to thank my parents, my husband, and my friends for their unwavering encouragement and support. Their belief in me has been a constant source of motivation, and without them, this achievement would not have been possible.

# Abstract

This project focuses on the enhanced classification of MRI brain tumours using a hybrid machine learning technique. By utilizing a vast dataset that includes four categories of brain tumours, this study aims to provide a more detailed and accurate analysis than previous methods. Emphasis will be placed on pre-processing techniques to improve the quality of the input data, ensuring more reliable and precise classification results. The primary objective is to conduct a comparative analysis, evaluating the performance of the proposed hybrid technique against existing approaches. This analysis will highlight the strengths and weaknesses of prior work, demonstrating how our method excels in accuracy and reliability. By addressing the shortcomings of earlier methods, this project aims to contribute significantly to the field of medical imaging, offering improved and early diagnostics for various kinds of brain tumours for on time treatment.

# Contents

## 1 Introduction

The human brain is one of the most essential parts of the body, playing a crucial role in every aspect of daily life. Activities and functions such as sensory perception, emotions, motor skills, responses, and even breathing are all dependent on the brain. However, if a tumour develops in the brain, it can disrupt all these vital functions, severely impacting one's overall health and well-being.(Anantharajan et al., 2024) .

Brain tumours are abnormal masses that form within brain cells, often arising from sudden or irregular growth of brain tissue. The impact of a tumour on a person can vary significantly depending on its size, nature, location within the brain, and the availability of treatment options, (Babu Vimala et al., 2023).

It is suggested to regularly visit the physician so early diagnosis helps to aid the disease much easier, these tumours can be classified into two main categories: benign (non-cancerous) and malignant (cancerous). While benign tumours grow slowly and are less likely to spread, they can still cause significant health issues due to their potential to compress and damage surrounding brain tissue., cancerous brain tumours, on the other hand, are aggressive and can invade nearby tissues or spread to other parts of the brain and spinal cord. (Hemanth et al., 2019) , This makes them particularly challenging to treat and often requires a combination of surgery, radiation therapy, and chemotherapy.

Each year, between 7 and 11 out of every 100,000 people across different age groups are diagnosed with brain tumours. Tragically, this devastating illness results in approximately 227,000 deaths annually. Additionally, around 7.7 million survivors are living with the long-term impacts and disabilities caused by brain tumours.(Fernandes et al., 2020).

The location of the tumour is also a critical factor in determining its impact. Tumours in certain areas of the brain can affect essential functions such as speech, movement, vision, and cognitive abilities. For instance, a tumour in the frontal lobe can impact decision-making and personality, while one in the occipital lobe can affect vision. Treatment availability and effectiveness depend on various factors, including the type of tumour, its stage at diagnosis, and the overall health of the patient. Advanced imaging techniques like MRI and CT scans have significantly improved the ability to detect and diagnose brain tumours early, which is crucial for effective treatment. On the other hand, ongoing research in the field of neuro-oncology is continuously improving treatment modalities, aiming to increase survival rates and quality of life for patients with brain tumours.(Mayo clinic, n.d.).

To begin with, a radiologist must capture an image of the affected area for a manual diagnosis. Following this, an experienced physician analyses the image and devises a treatment plan, but research into the accuracy of manually diagnosing brain tumours has revealed inconsistencies among the experts reviewing the data. Reportedly, the agreement rate among specialists for manually diagnosing brain tumours is between 90% and 95%. This consensus drops significantly when diagnosing mixed types of tumours, such as mixed glioma and medulloblastoma, where the agreement rates fall to 77% and 58%, respectively.(Rehman et al., 2019).

The current research on brain tumour classification and segmentation highlights several significant challenges, such as managing class imbalance in datasets, limited data availability, and the limitations of existing architectures for accurate tumour diagnosis. These issues hinder the achievement of high accuracy and efficiency in diagnosis. Moreover, dealing with class imbalance datasets remains problematic. Therefore, there is a pressing need for innovative architectures and techniques to improve the performance of brain tumour classification and segmentation models. The hybrid model is a combination of Machine learning and its subset deep learning EDN-SVM. Due to the complexity of the brain's anatomy and the variability of tumours, classifying and segmenting brain tumours has become a challenging task in medical imaging. Conventional approaches often fall short in terms of accuracy and effectiveness.(Shah et al., 2023) .

This project aims to implement the EDN-SVM model, a hybrid approach that combines neural networks with support vector machines (SVMs) to overcome the limitations of traditional SVMs. This model enhances reliability and accuracy in tumour classification. Previously, the EDN-SVM model was used with a limited dataset and only two variables: tumours versus non-tumours. In this project, a larger dataset will be utilized, including four different categories of brain tumours: pituitary, meningiomas, gliomas, and a healthy (no tumour) class. A comparative analysis will be conducted to evaluate the model's accuracy across these diverse categories. Additionally, various pre-processing techniques will be applied, and the best method will be selected to optimize the model's input.

"Machine learning allows us to build software solutions that exceed human understanding and shows us how AI can innervate every industry." - SpaceX and Tesla board member, Steve Jurvetson

## 2 Literature Review

Literature Review will be sub divided into two subsections: Brain Tumour Segmentation and Brain Tumour Classification:

# 2.1 Brain Tumour Segmentation:

Many researchers have implemented various techniques for segmentation of brain tumours, as segmentation can be helpful in detecting the tumour early and can lead to faster diagnosis and treatment. The aim of the segmentation is to locate and segment the area of the brain where tumour lies with the help of binary or multi class segmentation map, that accurately detects the region and presence of the tumour, segmentation gives a colourful image of the brain which reflects the affected area clearly. (*Tumor Segmentation - an Overview | ScienceDirect Topics*, n.d.). There are several research done regarding brain tumour segmentation, some of which are worth comparing and mentioning in the table 01. The table below compares the different segmentation techniques used.

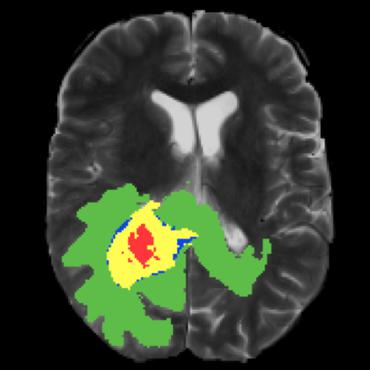


Figure A Brain Tumour Segmentation

|  |  |  |  |
| --- | --- | --- | --- |
| Papers | Dataset Used | Algorithm or/and Classifier Used | Results |
| Brain Tumour segmentation with Deep Learning.(Rao et al., 2015) | BRATS 2015 | convolutional neural network (CNN) and extracted features are fed into Random Forrest Model to train the classifier | Accuracy= 67%  Training data loss= 2.9% with Random Forrest Classifier |
| A Systematic Approach for MRI Brain Tumor Localization and Segmentation Using Deep Learning and Active Contouring.(Gunasekara et al., 2021) | Figshare MRI dataset | R-CNN Region based Convolutional Neural Network and Chan-Vese segmentation algorithm was used | The Accuracy of this segmentation technique reached up to 94% |
| Big data analysis for brain tumour detection: deep convolutional neural networks. (Amin et al., 2018) | ISLES and BRATS datasets | lightweight dense neural network for pixel-wise prediction | method achieved a 99.8% dice index on Flair data, 98% on T2-weighted data, and 97.4% on T1 data. |
| Table 1Research papers and their comparison (Brain Tumour Segmentation) | | | |

From the papers in Table 1, it would be relevant to discuss the papers and their achieved accuracy to create a clear picture of what techniques have previously been used in this field for segmenting brain tumours. The results are purely based on the nature of dataset used and other parameters, as shown in table 1 a convolutional neural network (CNN) was proposed for segmenting brain neoplasms using the BRATS 2015 dataset. This network utilizes multiple convolutional, pooling, and normalization operations to extract effective features. By incorporating non-linearity, the CNN establishes a mapping between the input and output. The features extracted from each convolutional layer are then combined and fed into a random forest model for training the classifier. This approach achieved an accuracy of 67% and a training data loss of 2.9% with the random forest classifier.(Rao et al., 2015)

Another architecture mentioned in the above table is R-CNN that that combines a convolutional neural network (CNN) for classification with a region-based CNN (R-CNN) for segmentation is described in [27]. This approach utilizes the Chan-Vese segmentation algorithm to accurately delineate the boundaries of the neoplasm. By integrating these techniques, the segmentation method achieves an impressive accuracy of 94%.(Gunasekara et al., 2021) .

Finally, the last paper mentions that the researchers utilized a lightweight dense neural network for pixel-wise prediction in the segmentation of brain neoplasms, leveraging the ISLES and BRATS datasets. Their approach demonstrated remarkable performance, achieving a 99.8% dice index for Flair images, 98% for T2-weighted images, and 97.4% for T1-weighted images, all within a processing time of 5.502 seconds and an overall accuracy of 98.65%. Furthermore, their technique achieved a dice index of 95.4% on T1-weighted contrast-enhanced images. This high level of accuracy underscores the effectiveness of their lightweight dense neural network model for detailed and precise brain tumour segmentation across various MRI modalities.

The various approaches discussed above illustrate different methods used to perform brain tumour segmentation, aiming to achieve highly accurate results. The following sections will provide a detailed discussion on the techniques and their efficacy in brain tumour classification and their need.

# 2.2 Brain tumour Classification:

Brain tumour segmentation focuses on region-based segmentation, which isolates the tumorous part of the brain by providing a colourful image. However, the challenge lies in the classification of tumours. As previously discussed in this report, tumours can be either benign or malignant, making it crucial to identify the type of tumour present in the brain. This necessity has driven researchers to develop methods not only to segment and locate the tumour but also to classify its type for appropriate treatment. Various techniques have been employed for brain tumour classification, some of which are discussed in the table below for a more precise and valid discussion.

|  |  |  |  |
| --- | --- | --- | --- |
| Papers | Dataset Used | Algorithms and/or Classifier used | Results |
| Modified local ternary patterns technique for brain tumour segmentation and volume estimation from MRI multi-sequence scans with GPU CUDA machine.(Sriramakrishnan, 2019) | BraTS 2013 dataset and  BraTS 2015 dataset | SVM (Support Vector Machine) used Fuzzy C-mean technique | Accuracy= 94.3 % |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Table 2 Research papers and their Comparison (Brain tumour Classification) | | | |

The first paper mentions the Fuzzy C-Means (FCM) technique was employed to extract neoplasm regions, followed by the implementation of the Probabilistic Local Ternary Patterns algorithm to identify the neoplasm substructure. For classifying the neoplasm on the extracted slices, a Support Vector Machine (SVM) was utilized. This classification approach achieves a dice index of 76% on the BraTS 2013 dataset and 81% on the BraTS 2015 dataset for the entire neoplasm. A Convolutional Neural Network (CNN) was used additionally to classify neoplasms into pituitary, glioma, and meningioma categories, attaining an accuracy of 94.39%.(Sriramakrishnan, 2019)