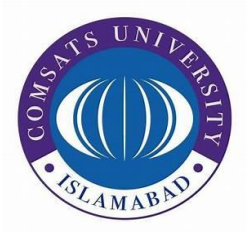


COMSATS University Islamabad
Islamabad, Pakistan

Brain Tumor Classification

By

Fatiha Ilyas CIIT/SP17-BCS-009/ATK

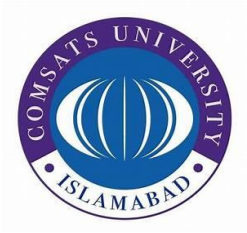
Khizra Mughal CIIT/SP17-BCS-014/ATK

Supervisor

Sir Qasim Khan

Bachelor of Science in Computer Science (2017-2021)

The candidate confirms that the work submitted is their own and appropriate credit has been given where reference has been made to the work of others.



COMSATS University, Islamabad Pakistan

Brain Tumor Classification

A project presented to

COMSATS University, Islamabad

In partial fulfillment

of the requirement for the degree of

Bachelor of Science in Computer Science (2017-2021)

By

Fatiha Ilyas CIIT/SP17-BCS-009/ATK

Khizra Mughal CIIT/SP17-BCS-014/ATK

DECLARATION

We hereby declare that this software, neither whole nor as a part has been copied out from any source. It is further declared that we have developed this software and accompanied report entirely on the basis of our personal efforts. If any part of this project is proved to be copied out from any source or found to be reproduction of some other. We will stand by the consequences. No Portion of the work presented has been submitted of any application for any other degree or qualification of this or any other university or institute of learning.

Student 1

Student 2

CERTIFICATE OF APPROVAL

It is to certify that the final year project of BS (CS) “Brain Tumor Classification” was developed by Fatiha Ilyas (CIIT/SP17-BCS-009) and Khizra Mughal (CIIT/SP17-BCS-014) under the supervision of “Mr. Qasim Khan” and co supervisor “Mr. Umair Khan” and that in (their) opinion; it is fully adequate, in scope and quality for the degree of Bachelors of Science in Computer Sciences.

Supervisor

External Examiner

Head of Department

(Department of Computer Science)

Executive Summary

Computer based systems can improve the diagnostic abilities of radiologists and physicians, optimize the accuracy. Also, it can minimize the time required for accurate results of classification of brain tumor. Brain Tumor is one of the major causes of death in recent years and becoming life-threatening. There are multiple types of brain tumor exist and each tumor type has different structure, placement in head and different nature. Correct treatment of brain tumor type can affect the planning of treatment and increase the survival rate.

In our project, we discussed the automatic classification of brain tumors by using image processing techniques. We performed following steps of deep learning to carry out our project i.e. preprocessing, segmentation, feature extraction and classification. A brain MRI image is input to system, it is first preprocessed by several operations, then it goes for segmentation, then its features are extracted and fed for classification. At the end, the brain tumor label is predicted. Three most common type of brain tumors i.e. Meningioma tumor, Glioma tumor and Pituitary tumor are included in the study and 3064 T1 weighted CE-MRI images are used. We are intended to develop a system that enhance and optimize the tumor diagnostic process. It will help radiologists for decision making and further surgical treatments.

Acknowledgement

All praise is to Almighty Allah who bestowed upon us a minute portion of His boundless knowledge by virtue of which we were able to accomplish this challenging task.

We are greatly indebted to our project supervisor “Mr. Qasim Khan” and our Co-Supervisor “Mr. Umair Khan”. Without their personal supervision, advice and valuable guidance, completion of this project would have been doubtful. We are deeply indebted to them for their encouragement and continual help during this work.

And we are also thankful to our parents and family who have been a constant source of encouragement for us and brought us the values of honesty & hard work.

Student 1 Student 2

--- ---

Table of Contents

List of Figure

List of Table

Chapter 1

INTRODUCTION

1.Introduction

Brain tumor is caused by abnormality in cells of brain. There are many categories of tumor exists, some of them are cancerous and others are non-cancerous. In US around 700,00 people are suffering from brain tumor and more will be diagnosed in future. World Health Organization (WHO) classification systems are working to diagnose the brain tumor types and many researchers proposed the techniques for classification. There are 120 type of brain tumor exists and each tumor type has different size and location. So, it is very difficult task to classify the tumor part. It affects the human health badly, stops the brain function and can be life threatening. It is most treatable and curable if caught in the earliest stages of the disease. Therefore, early treatment of the brain tumor is much needed, and researchers are motivated to solve this problem.

Brain Tumor diagnosis and treatment depends on Radiologists and Doctor's accurate decisions. These decisions are done by seeing and analyzing the tumor location in brain, tumor shape and tumor size. Doctors interpret the MRI of patient and decide that whether the patient has a tumor or not. Diagnosing the correct type of tumor is a challenging task. Human errors can mislead to wrong diagnosing of tumor type and can be misleading for planning the treatment for patients.

The tumor type analysis is still done manually by radiologists, takes several time and effort. Therefore, the development of automatic system is needed both for physicians and patients. This work is driven by the motive to make the effective and efficient system for better classification of the tumor with fast computational speed. It will be very effective for physicians for further diagnosis and decisions. Brain tumor is expected to become the number one cause of death in future. There is large amount of brain tumor images to be process, multiple angles of MRI scan need to be analyze and diagnosis. Sometimes human fatigue and stress can affect the process of

analysis. Physicians sometime fail to detect the early signs. Make software that early detect the tumor type automatically in brain. This will be helpful for radiologist and physicians for deciding the surgical treatment. MRI images provide the clear and sharp view of tumor as shown in figure 1.1. And provide the greater contrast images of brain.

Figure 1-1 MRI of Brain

MRI images have different types depending on the method used to capture them, these types are called Modalities. Some tumor structures are visually clear in one modality more than the others. Also, MRI images do not produce radiations, so they are not harmful. MRI Image captured from different Angles reveals the detail areas inside the brain such as Axial, Sagittal and Coronal. Three Angles of brain are shown in figure 1.2.

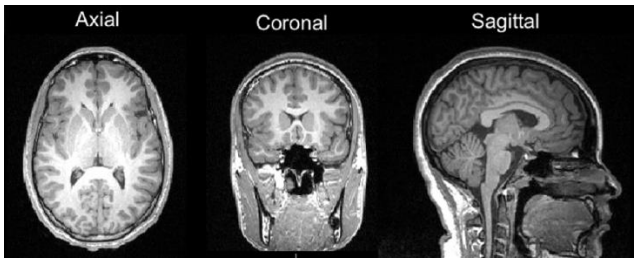


Figure 1-2 Three angles of Brain MRI

1.1 Brief

Medical image processing gains a lot of attention in current research. Brain tumor is caused by unusual growth of cells that originates from the unlimited division of cells. We will use the brain MRI image in our system for detection. It detects that image from all angles. Ultrasound image do not provide enough information, so we use the MRI image. Diagnosis of tumor with the help of MRI image is mandatory but time-taking. This task is difficult for experts when they go for manual method. So, computer supported system are introduced for accurate disease detection from MRI images. It is problematic for radiologists to detect the tumor types manually and make decision due to variation in size, shape and types of abnormal tissues. It can be used as second option or supportive tool. This system proved that it improves accuracy and correctly detect the tumor.

This system reduces the workload and avoids the mistakes that can be done by the radiologist. Therefore, these systems improve the diagnosing ability and have the similar learning mechanism as radiologists. It means that these systems can have the ability to replace the radiologists. They can automatically segment the tumor areas as radiologists do by pointing out the edges. Therefore, these systems are playing vital role in diagnosing and detection of tumor automatically. Nowadays, the use of these systems supporting in the early detection of any type of tumor in recent years. By studying the literature, we have analyzed that system's performance can be better by improving the computational speed and flexible systems for different brain datasets.

To increase the performance, we have focused in building a system that has better computational speed, time complexity and accuracy. We proposed an efficient classification technique able to perform an automated classification of different categories of brain images. We propose a system, based on image segmentation, feature extraction, classification. The proposed system takes brain MRI images and will automatically detect the tumor type accurately. We are intended to make

more efficient and accurate system that can improve the existing tumor detection methods. Convolutional neural network (CNN) technique is used for classifying brain tumor. The classifier is trained on three of tumor types: meningioma, pituitary and glioma

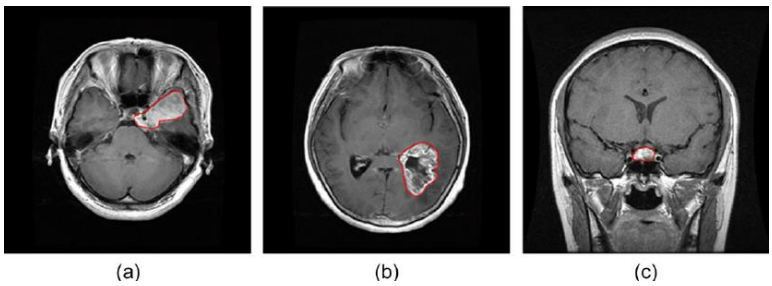


Figure 1-3 Three Tumor Types: Meningioma, Pituitary, Glioma

1.2 Relevance to Course Modules

1.2.1 Artificial Intelligence

In that course all machine learning techniques and algorithms are part of it. The topic that closely relates with my project is neural network. An artificial neuron is developed by machine learning technique to replicate the biological neuron. Neural networks comprise of technologies like machine learning concepts and deep learning and form a part of Artificial Intelligence.

1.2.2 Machine Learning

In this course we have studied about different algorithms which are used to train models to perfume task automatically after getting the training data.

1.2.3 Report Writing Skills

This course is about learning how to write reports and other formal documentation, and, in our project, we need to write our documentation, so this course is helping a lot in this task.

1.2.4 Human Computer Interaction

A system which is interactive is easy and comfortable for the user to use the system and understand it easily and this course is all about designing interacting system following standard rules.

1.3 Project Background

Brain is an important organ of human body that consists of billions of cells. When the growth of cells gets abnormal inside brain then they cause brain tumor. Brain tumor is of two broad types i.e. Malignant and Benign. Malignant tumor is very dangerous and spread to other parts of body and cause immediate death, but benign tumor is not very dangerous and does not cause sudden death. There are three categories of brain tumor as well i.e. meningioma, pituitary and glioma. Meningioma tumor is developed in thin membranes of brain like around spinal cord and it is benign tumor. Pituitary tumor is developed in pituitary gland of brain and it is malignant one. Glioma tumor develops in the substance of brain and is benign one. Magnetic Resonance Imaging (MRI) is a medical technique which is used for detection of brain tumors in clinical approach. An image of brain is taken from three angles and all angles are observed individually. For examining the tumor, the brain MRI is taken to detect the tumor and to diagnose it. Brain MRI image provide much information about features of brain as compared to CT-scan image of brain. MRI image gives a detailed knowledge of brain tissue and its minor features as well.

Radiologists and physicians must examine these MRI images and they consume a lot of their time and energy in coming up with results. If a patient has last stage tumor then it is impossible to save the life because the manual method of detecting the brain tumor takes time. The procedure has all

required steps that must be fulfilled for deducing the result. To get rid of this problem, an automated system is introduced which help radiologist in early detection of brain tumor and to save the precious lives of patients. An automated system is designed purely on basis of artificial intelligence using machine learning techniques and tools.

1.4 Literature Review

With a specific end goal to influence the task, to yield appropriate output, it is important to experience the past investigates and procedures. This heading contains the complete description of the past work and reports that have been completed in such manner. Different approaches have been used in past few decades for diagnosis. We illustrate how these approaches differ from each other, bring to light the merits of each and limitations.

1.4.1 Machine Learning Based Approaches

Machine Learning plays an important role in Computer Diagnosis system and give promising results with improved accuracy .The machine learning also ensure the analysis of the different medical data and machine learning also provide the ability to make classy and good algorithms that help in diagnosing. For examining the bio-medical data, machine learning provides a worthy approach for making efficient algorithms. Following are the approaches that have been proposed by researchers.

1.4.2 Deep Learning

Heba Mohsen.[1] proposed Deep Neural Network classifier (DNN) for classifying 4 types of tumor classes. In model use the walvet transformation and feature extraction tool and over all achieved 96% accuracy for the whole data set. In this technique computation time is high and accuracy is low than other approaches.

1.4.3 Artificial Neural Network

Dipali M. Joshi.[2] proposed neural network approach for classification. In his work, the focus was to classify the tumor as well. He used astrocytoma tumor types for classification, using astrocytoma tumor he performed pre-processing method such as histogram, segmentation, morphological technique. His feature extraction method was Gray Level Co-occurrence Matrix (GLCM). The GLCM based features were then compared already stored features. After that, a Neuro Fuzzy method was developed for classification. Their classification approach required many steps for image preprocessing and Detection which increase the computational time. Also, high computing cost is the draw back that consumes high CPU and use of physical memory.

1.4.4 Support Vector Machine

In the research work proposed by Hari Babu Nandpuru. [3] and Dr. S. S. Salankar, classification algorithm is done using Support Vector Machines (SVM). They proposed brain image classification that use Feature extraction by applying grey scaling and two more features like symmetrical and texture. For diagnosis, they used 50 patient's brain images with two classes such as 'Normal', 'Abnormal'. Using on the SVM classification algorithm they classified the tumor type successfully.

1.4.5 K-Nearest Neighbour (KNN) classifier

[4] used KNN classifier for classification purpose by going through following steps: Determining the k value, calculating distance between the training samples and query instance and, sortation of distance based on the kth minimum distance, majority class assignment, determining the class. Both and color features textural features are being passed in classification algorithms using K-NN give the accuracy of 87.0%.Their method was evaluated on categorized data as Normal Brain Images, Primary Tumor images (or benign), Secondary Tumor Images (or Malignant).

1.4.6 Convolution Neural Network (CNN)

Nyoman Abiwinanda. [5] proposed classification approach which has simple CNN architecture: one each of convolution, max-pooling, and flattening layers, followed by a full connection from one hidden layer. They have used 3064 T1 new publicly available MRI dataset for training the classifier and his classification technique is independent of prior region-based segmentation technique. Without segmentation, he achieved the comparable results. Our project is based on the same approach and dataset they have used to achieve more better accuracy and computational time.

Justin S. Paul. [6] also proposed study based on the same dataset for classification using CNN. His study was focused on axial images and expanded the data with addition of axial images that has no tumor to increase the number of images provided to the neural network. Experimented with constructing variety of neural networks based on pre-processing of image data. Jun Cheng [7-8] was the initial person who prepared the dataset and proposed classification technique based on data augmentation and partition. He has also proposed a content-based image retrieval in another research using the same dataset.

1.4.7 Comparison of CNN research work based on the Jun Cheng's Brain Tumor dataset

Table 1-1 Comparison of CNN

Author	Method	Target	Dataset	Accuracy
Nyoman Abiwinanda [5]	CNN involves a Rectified Linear Unit (ReLu), a convolution, and a pooling layer	Brain Tumor Classification	Jun Cheng Brain Tumor Dataset	98%
Justin S. Paul. [6]	Deep learning for Brain Tumor Classification	Brain Tumor Classification	Jun Cheng Brain Tumor Dataset	91.43%
Jun Cheng [7]	Brain Tumor Classification by means of tumor Region Augmentation and Partition	Brain Tumor Classification	Jun Cheng Brain Tumor Dataset	Improved 82.31% from 71.39%
Jun Cheng [8]	Retrieval of Brain Tumors by Adaptive Spatial Pooling and Fisher Vector Representation	Content Based image Retrieval	Jun Cheng Brain Tumor Dataset	94.68%

1.5 Analysis from Literature Review

A task is accomplished with specified goal to achieve a desired output, it is necessary to go through past research and techniques for evaluating accuracies. Different approaches have been used in past few decades for diagnosis. These techniques differ from each other, bring to light the merits of each and limitations with respect to our problem. We used Convolution Neural Network approach to carry out our problem. It involves core concepts of machine learning which is a subset of deep learning. Artificial Intelligence is the domain we worked in. In past, machine learning based approaches used for yielding maximum accuracies and appropriate outcomes. The

interpretation of different medical data is ensured by machine learning techniques and it also provide the ability to make classy and good algorithms that help in diagnosing the serious problems. Automatic algorithms and worthy approaches are devised using machine learning.

Deep learning has a classifier named as Deep Neural Network (DNN) classifier that classifies four types of tumor classes. The walvet transformation and feature extraction tool is used in it which yield 96% accuracy overall. The drawback of this classifier is high computation time and low accuracy. Artificial Neural Network (ANN) classifier used astrocytoma tumor types for classification, using astrocytoma tumor. Their classification approach required many steps for image preprocessing and Detection which increase the computational time. The drawback was high computing cost for consuming much memory and resources. Support Vector Machines (SVM) method used feature extraction by applying grey scaling and two more features like symmetrical and texture. K-Nearest Neighbor (KNN) classifier also used for this problem. This algorithm using K-NN give the accuracy of 87.0% and high computing cost.

We have analyzed all these past techniques of machine learning and after reviewing their drawbacks we used best technique that overcomes the high computing cost and yield maximum accuracy. Convolution Neural Network (CNN) has architecture: one each of convolution, max-pooling, and flattening layers, followed by a full connection from one hidden layer. This classification technique is independent of prior region-based segmentation technique. This technique overcame all the defects and gave appropriate results with improved accuracies.

1.6 Methodology and Software Lifecycle for this Project

Project methodology is an important phase of any project because it is a key element and set the overall tone. For this we must first understand the steps that are involved in project methodology. We use iterative approach to develop this application.

1.6.1 Methodology

Software development methodology involves splitting the work into different stages with the goal of better management, planning and understanding Software.

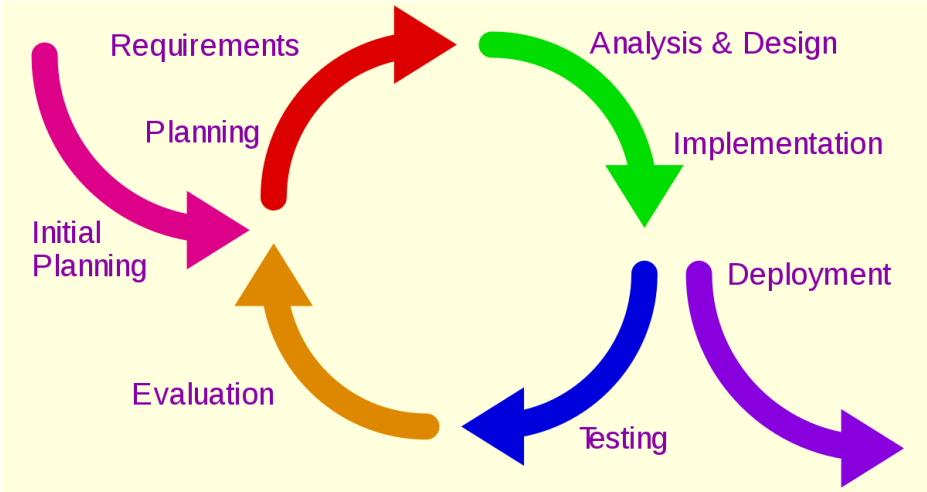


Figure 1-4 Iterative Model

1.6.2 Rationale behind selected methodology

Iterative process is the way to breakdown the whole process to develop a complete software in one go into smaller chunks. The reasons behind this approach is that it is easy measurable and small chunks are easy to test and debug as compare to developing the whole software and then testing it, at this point debugging becomes difficult.

Chapter 2

PROBLEM DEFINITION

2. Problem Definition

Today, the clinics are crowded with tumor patients. The tumor type analysis is still done manually by radiologists, takes several time and effort. The chances of survival may increase if the tumor is correctly detected in its initial stage. Therefore, the development of automatic system is needed both for physicians and patients.

2.1 Problem Statement

Classification of Brain Tumor is needed to overcome the manual analysis, workload, improve computational time, flexibility, and the accuracy. Brain tumor is expected to become the number one cause of death in future. There is large amount of brain tumor images to be process, multiple angles of MRI scan need to be analyzed and diagnosis. Sometimes human fatigue and stress can affect the process of analysis. Physicians sometime fail to detect the early signs. Existing computer aided diagnosing systems are not capable of generating efficient results in less computational time and inflexible to different brain data sets. To accurately classify the brain type, we need accurate systems to accommodate the limitations. Solution to the problem Make software that early detect the tumor type automatically in brain. This will be helpful for radiologist and physicians for deciding the surgical treatment. Automatically classifying the tumor type whether it is meningioma, glioma or pituitary is the required solution.

2.2 Deliverables and Development Requirements

List of deliverables of our project:

- Needed MRI scans of patients.

Development requirements include following software requirements:

- IDE: PyCharm.
- MATLAB.
- MS Word.
- MS PowerPoint.
- Anaconda distribution.
- Dataset based work.
- Programming Language: Python.
- Interface: PyQt Designer.

Chapter 3

REQUIREMENT ANALYSIS

3. Requirement Analysis

Software Requirement Analysis (SRS) is providing the basic understanding of functional as well as non -functional requirements. We can consider it as a starting point of project because it serves as written contract between client and organization about the features and functionalities of the project. With the help of SRS both client and organization make clear to each other about the deliverable project.

3.1 Use Case Diagram

A realistic picture of the connections among the components of a framework is called use case diagram. A use case diagram is a technique utilized as a part of framework investigation to perceive and arrange framework necessities. We design UML (Unified Modelling Language) figures to model a framework in the simple and efficient way. UML describes several types of the diagram to cover all aspects of the framework because it is not enough to define all aspects of the system using single UML diagram.

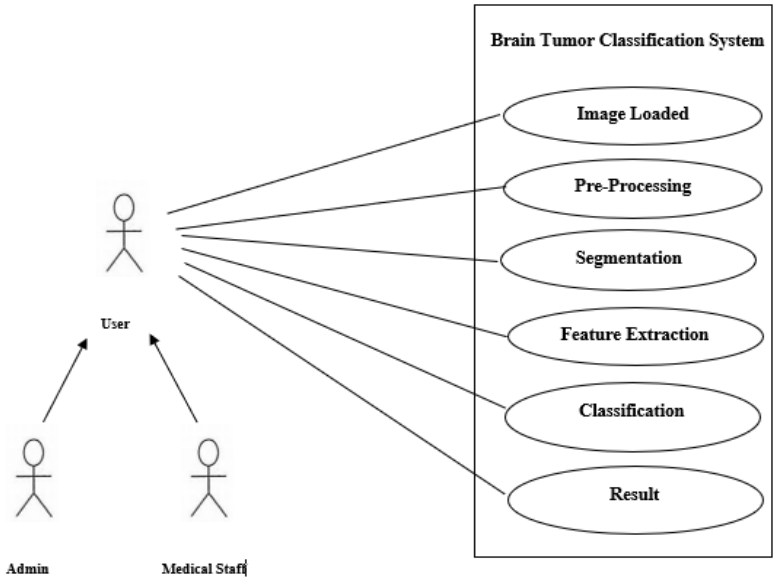


Figure 3-1 Use Case Diagram

3.2 Detailed Use Case Diagram

Use case figure is one of the UML diagrams which models dynamic behavior. It is dynamic in nature, consisting of some inside or outside components for making the connection. It is a set of use cases, actors (primary and secondary actor) and relationships. A use case diagram is used to represent a particular functionality of a system. Many use cases diagrams are used to model the entire system. It is a generic way to define that use case diagram captures all dynamic aspects of the entire system.

1.Use-case name: Input MRI Image.

Action: User.

Description: User will input an image from system directory to application as initial step to use this application.

Pre-condition: User must provide MRI image to a system as an input and it must be in .jpg format.

Post-condition: After giving an input to a system, image is uploaded.

2.**Use-case name**: Pre-Processing.

Action: Computer-aided system.

Description: An input image is going for pre-processing.

Pre-condition: Image must be uploaded.

Post-condition: After giving an input to a system, image is pre-processed, and user can exit, or it go for segmentation process.

3.Use-case name: Segmentation.

Action: Computer-aided system.

Description: A pre-processed image is sent forward for segmentation.

Pre-condition: Input pre-processed images or new image is uploaded.

Post-condition: After segmentation, masks are produced.

3.Use-case name: Feature Extraction.

Action: Computer-aided system.

Description: A preprocessed image is sent forward for feature extraction.

Pre-condition: Input preprocessed images or new image is uploaded.

Post-condition: After feature extraction, features are extracted.

4.Use-case name: Classification

Action: Computer-aided system

Description: An input MRI image is pre-processed, segmented and classifier extracts features and then classify with help of classifier to classify three tumor classes of brain on basis of features.

Pre-condition: All the labels must be assigned to each tumor class and must be placed in separate sub-directories of testing and training directories.

Post-condition: After classification on 80% training and 20% testing, the labels are classified.

5.**Use-case name**: Result

Action: Computer-aided system

Description: After classification of each tumor class, each tumor label is classified with promising accuracies and yield output.

Pre-condition: The labels must classify correctly.

Post-condition: Result is displayed after correct classification of brain tumor classes along with their labels and accuracies.

3.3 Functional Requirement

Functional requirement describes the predictable behavior of the framework. These are main framework requirement and without applying any of these requirements the framework should be fragmented.

- The system will take the input image and perform pre-processing.
- The system will segment the image.
- Based on the segmentation, the system will classify the segments and detect the tumor.

3.3.1 Pre-processing

We convert the images from (.mat) format into .jpg and label them. We make separate directories of each type of tumor class. And then we apply other elements of pre-processing like, image scaling, grey scale, image enhancement, image restoration and noise removal.

3.3.2 Input image

Table3-1 Input Image

Identifier	FR-1
Title	Input image
Requirements	The system take image
Dependencies	None

3.3.3 Segmentation

Term Image segmentation refers to the portioning of an image into groups of pixels which are homogenous with respect to some criteria. The objective of segmentation is to extricate as well as change the portrayal of an image into more significant easy to examine. We will apply segmentation on preprocessed images by following techniques: Threshold method, Watershed Algorithm and Region-Based method.

3.3.4 Feature Extraction

Feature defines the behavior of image. To determine the subset of initial features is known as feature selection. Feature extraction refers to extraction of features based on pixels. To save the computational cost, take the tumor region and extract its features and interpret them. This is done to isolate desired portion from an input image.

3.3.5 Classifier

Table 3-2 Classification

Identifier	FR-2
Title	Classifier
Requirements	The system extracts features, performs the classification, and predict the tumor label.
Dependencies	FR-1

3.3.6 Result

Table 3-3 Result

Identifier	FR-3
Title	Output
Requirements	The classifier predicts the tumor label and displays the accuracy.
Dependencies	FR-2

3.4 Non-Functional Requirement

3.4.1 Performance

Performance of our system is efficient, and it takes very less time to perform the action. Our CAD system saves the time.

3.4.2 Capacity

Our system must meet the agreed capacity.

3.4.3 Availability

The system will be available into any hospital or clinic.

3.4.4 Reliability

Our system is 100% reliable and generate an accurate result.

3.4.5 Maintainability

With the time and needs our system will be maintained properly and any feedback from the user is our priority while performing maintenance.

3.4.6 Efficiency

This system is efficient as it does not require any kind of effort to use it and does not take much time.

3.4.7 Flexibility

The system provides the user to load the image easily.

3.4.8 Usability

This system is very easy to use by radiologist and physicians. The designed interface is very easy to understand.

Chapter 4

DESIGN AND ARCHITECTURE

4. Design and Architecture

After gathering all requirements, the next step is to start planning that how we are going to develop our project, how much resources, costs, time, benefits, and other items are required. After planning we move to the designing and architecture phase that which techniques and methods we can use and how we are going to develop our project. It is the most challenging phase of project development.

4.1 System Architecture

We have used Nyoman Abiwinanda's [5] CNN technique in which he has used image and label from dataset. We start with pre-processing, extracting images, segment them pass to CNN for feature extraction and classification that tells the tumor type in output. Each step is discussed in detail.

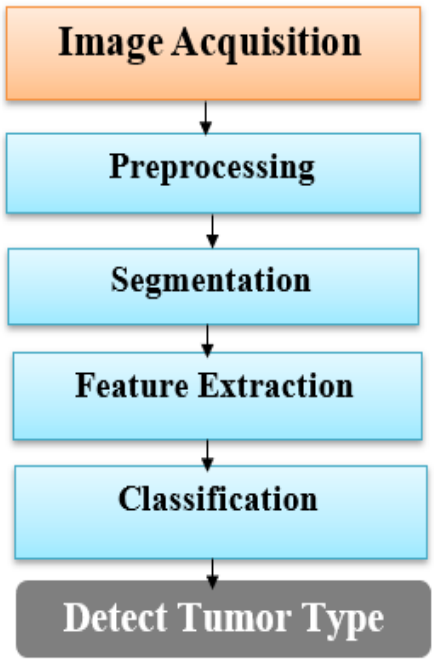


Figure 4-1 System Architecture

4.2 Data Representation [Diagram + Description]

A data flow diagram is graphical interpretation of information move from a data framework is called data flow diagram. A DFD is used for basic step to create an overview of the system without going into great aspect, which can later be elaborated.

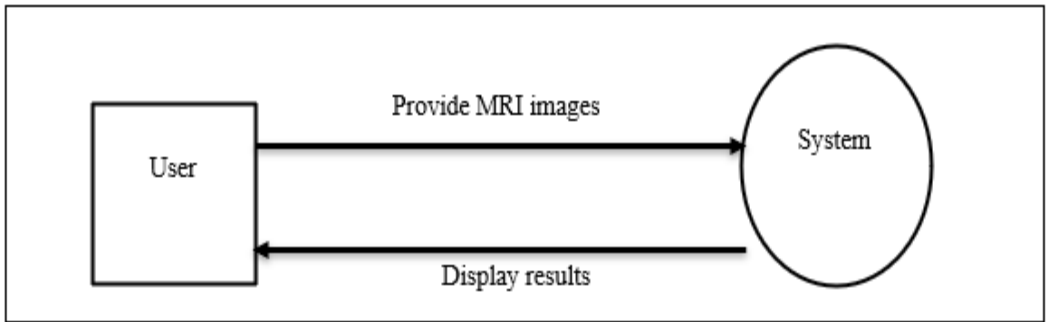


Figure 4-2 DFD Level 0

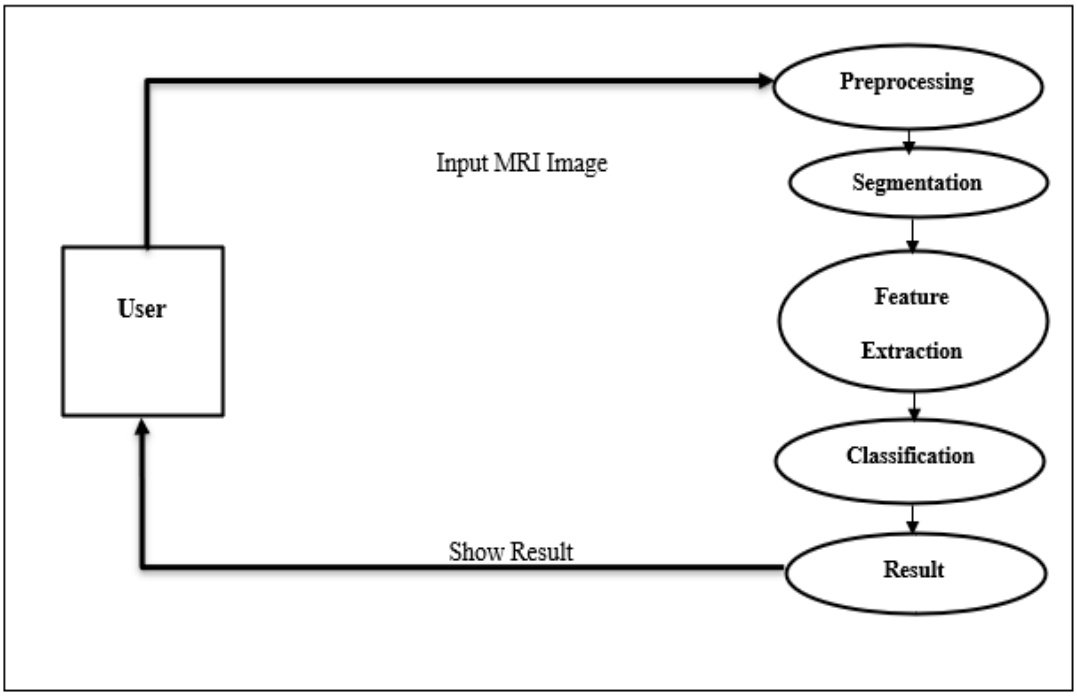


Figure 4-3 DFD Level 1

4.3 Process Flow / Representation

4.3.1 Activity Diagram

An activity diagram is depicting the dynamic component of the framework. An activity diagram is essentially a flowchart to speak from one action to another movement. The movement may be relating to an operation of the framework. This diagram is used to explain the dynamic features of the system. It is more like a flow chart because it shows the flow of data from one activity to other.

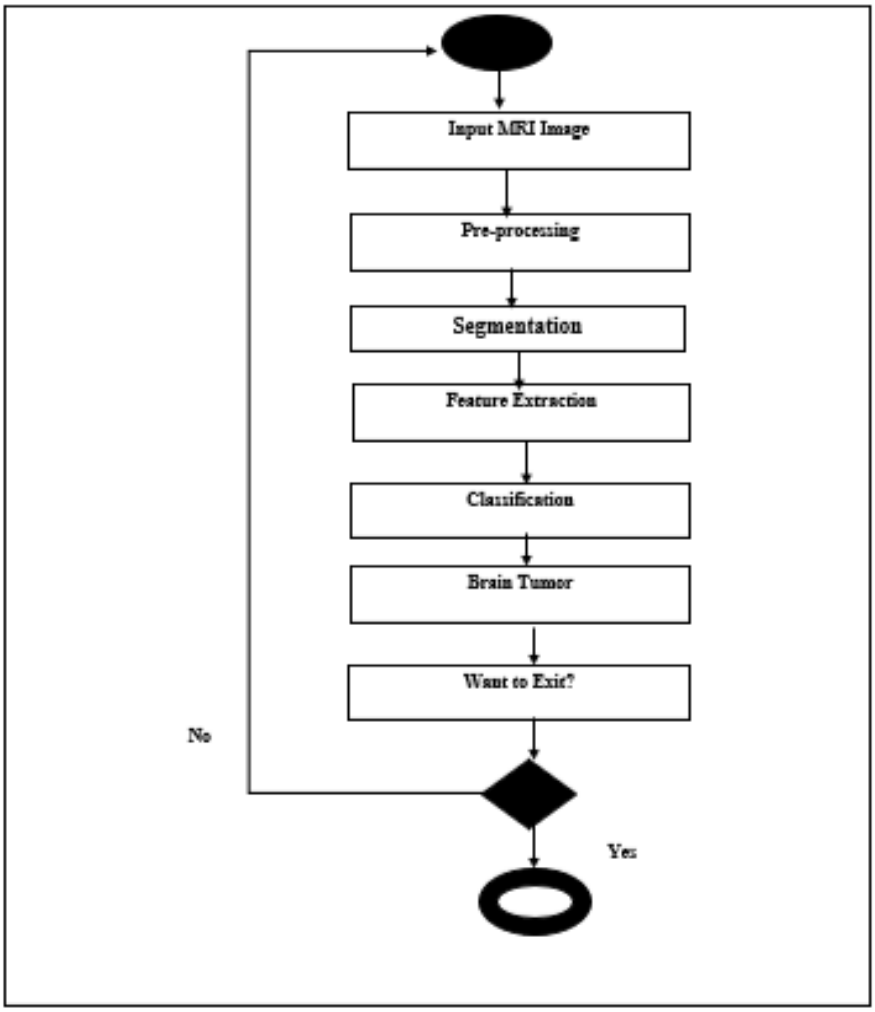


Figure 4-4 Activity Diagram

4.4 Design Models

Sequence Diagram

A cooperation graph that shows how objects work with each other and in what request is called sequence diagram. It is built up of message arrangement outline.

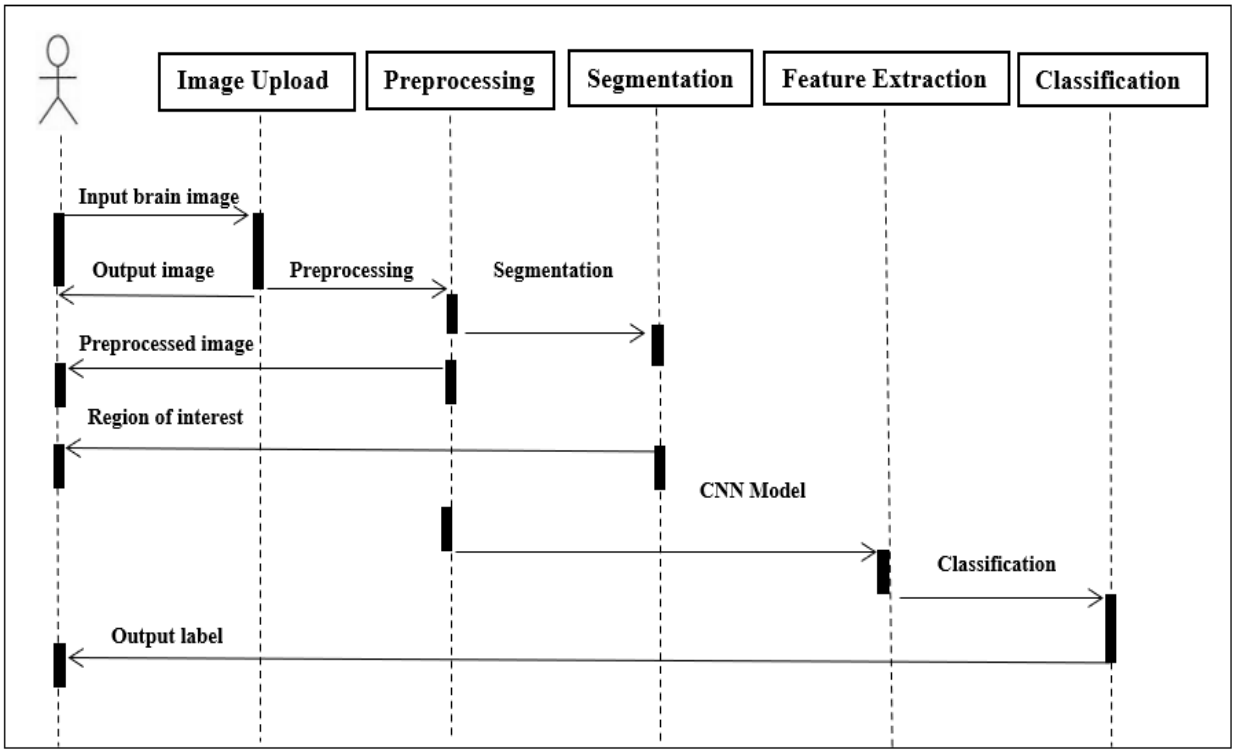


Figure 4-5 Sequence Diagram

CHAPTER 5

IMPLEMENTATION

5. Implementation

In this chapter, we will discuss the algorithms and user interface we use to develop the project. This phase is the most prominent phase of development; as from this step, we transform our idea into meaningful picture. This is significant and challenging step towards learning or developing skills. An application is result of successful implementation of project – the various algorithms, testing approaches and the results.

5.1 Algorithm

We have implemented nine algorithms in this project. A detailed explanation and the various outputs are shown below:

5.1.1 Dataset

The dataset used in this project is in .mat file extension and publicly available by Jun Cheg in 2015. Cheg is the author of the dataset. It consists of 3064 T1 CE-MRI images. Three tumor classes are included in the dataset which are Meningioma as label 1, Glioma as label 2, Pituitary as label 3. It contains Meningioma (708 slices), Glioma (1426 slices), and pituitary (930 slices). The images of 233 patients are collected from different angles of brain i.e. backside (coronal), front side(sagittal) and top (axial). There are 994 axial images, 1045 coronal images and 1025 sagittal images captured. The size of MRI image is 512 x 512 pixels. There are total five attributes in data which are as follows:

- Brain tumor image.
- Labels that specify the tumor type.
- Patient id.

- Tumor border.
- Tumor mask.

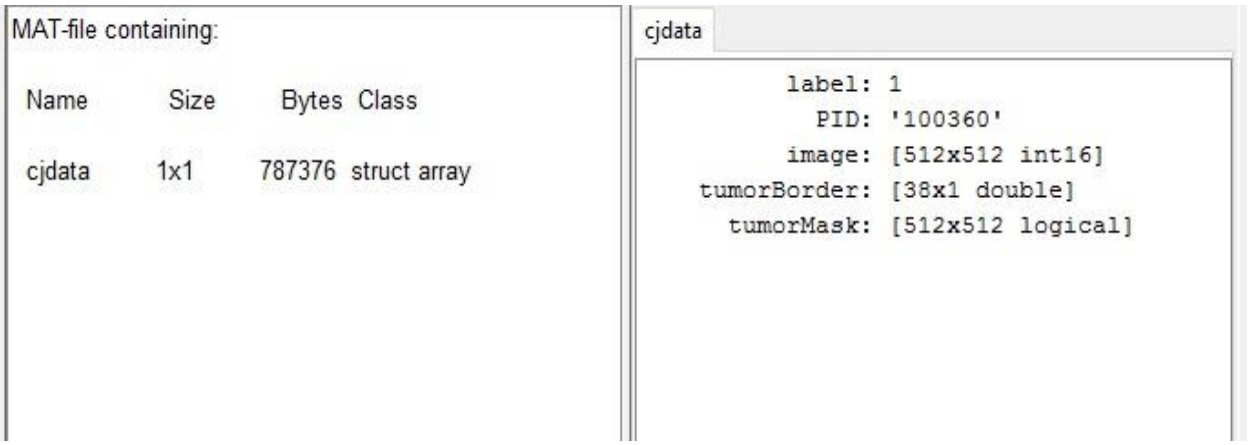


Figure 5-1 (.mat) file

Due to the file size limit of repository, the whole dataset is split into 4 subsets, and archive them in 4 .zip files with each .zip file containing 766 slices.

Figure 5-2 Dataset directories

5.1.2 Preprocessing

Image pre-processing is important task. Before bringing your data into classification, data should be clean, well specified, and consistent. So, we extract MRI image and its label from (.mat) structure class. We converted images into .jpg format and stored them into three directories according to their tumor type. Splitting of dataset into testing and training for classification is performed after the pre-processing. From 3064 images we use 800 images for training from each tumor class and 200 images for testing the data.

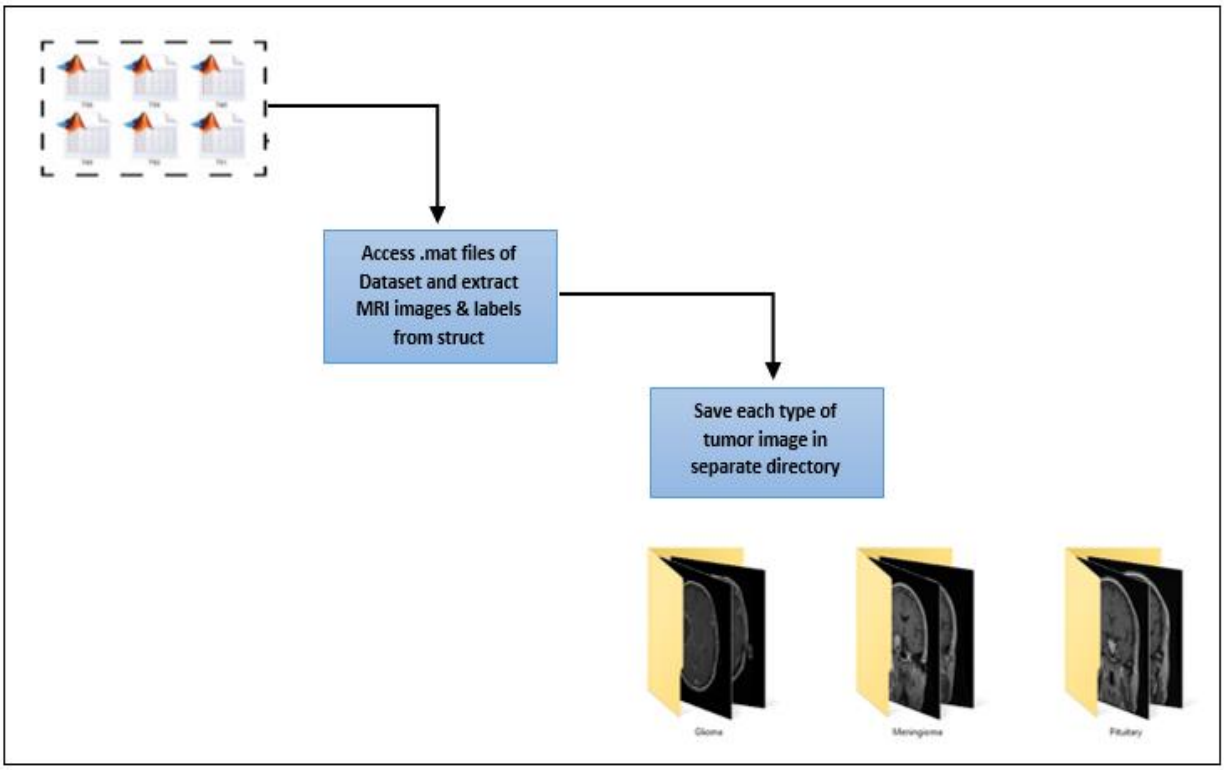


Figure 5-3 Dataset description

Following techniques are applied to clean noise from images:

5.1.2.1 Image Scaling

After taking input image there is need to resize an image. Images provided in dataset are of (512, 512) and we resized them into (64, 64) because machine learning models learn faster on smaller images and saving computational power. All images are rescaled to (64, 64) and model is trained over them. Each input image is rescaled first to that standard dimensions.

5.1.2.2 Grey Scale

Greyscale is applied on an input image as it simplifies the algorithm by reducing computation complexity. It takes less space in memory by minimizing the number of pixels that need to be processed.

5.1.2.3 Image Enhancement

For image enhancement, image normalization is done that changes the range of pixel intensity values to make the image more normal to senses. Its purpose is to increase contrast of image for removing high-frequency and low-frequency noise.

5.1.2.4 Median Filter

Median filter is a non-linear method used to remove (impulse) “salt and pepper” noise quality of while preserving the edges sharp of greyscale image. It works by moving through the greyscale image pixel by pixel, replacing each value with the median value of neighboring pixels. The median is calculated by first sorting all the pixel values into ascending order, and then replacing the pixel with the middle pixel value.

5.1.2.5 Gaussian Filter

Gaussian filter (Gaussian smoothing) is a linear method used to produce blur in the image by gaussian function for removing gaussian noise. It does not preserve the edges sharp like median filter.

5.1.3 Segmentation

Image segmentation is the primary step and the most critical tasks of image analysis. Its purpose is to make segments of an image, highlights the tumor part (region of interest) by discarding other segments of image. The mechanization of medical image segmentation has established wide application in diverse areas such as verdict for patients, treatment management planning, and computer - integrated surgery. Following are the segmentation algorithms that have been implemented:

5.1.3.1 Thresholding

In image processing, binary thresholding algorithm create binary image from a gray image based on the threshold value. The algorithm assumes that the image contains two classes of pixels that are foreground pixels and background pixels, it then calculates the optimum threshold separating the two classes.

5.1.3.2 Edge Detection

Canny edge detection is a technique to extract useful structural information from different vision objects and reduce the amount of data to be processed. The Gaussian filter is applied to smooth the image to remove the noise. Find the intensity gradients of the image. Apply threshold to determine potential edges. Finalize the detection of edges by suppressing all the other edges that are weak and not connected to strong edges. Tumor segment’s edges detected that way.

5.1.3.3 Watershed Algorithm

It is an interactive image segmentation. Label the region which we are sure of being the foreground or object with one color (or intensity), label the region which we are sure of being background or non-object with another color and finally the region which we are not sure of anything, label it with 0. That is the marker. Then our marker will be updated with the labels we give, and boundaries of objects will have a value of -1. So, the tumor region will have different color than other segments of image.

5.1.4 Feature Extraction

Feature extraction is process of extracting quantitative information from an image such as color features, texture, shape, and contrast. Here, Convolutional Neural Network (CNN) will learn the features from image by its own.

5.1.5 Classification

Based on features, classifier compare the input test image with trained data and displays the tumor type in result.

5.1.5.1 CNN Architecture

We implemented simple architecture proposed by Nyoman Abiwinanda [5] . The feature extraction is done by CNN from the pixels of MRI images using two main operations:

5.1.5.1.1 Convolution

Each input image is passed to series of convolution layers with filters. We use 2 convolutions. In first convolution layer we use 32 filters with 3x3 size and in 2nd convolution we use 64 filters with 3 x3 size.

5.1.5.1.2 Pooling

Max pooling is used for down sampling the image. Down sampling operation reduce the dimension and preserve the features of tumor image.

5.1.5.2 Parameter Selection

Table 5-1 Parameter selection

Image Size	64 x 64
Input layer	12544
Hidden Layers	40
Output layer	3
Epochs	10
Activation function	reLU (rectified linear Unit)

5.1.5.3 Labels

Table 5-2 Labels

Labels	**Tumor Class**
1	Meningioma
2	Glioma
3	Pituitary

5.1.5.4 Implementation steps (On dataset)

- **Input Dataset**

  The dataset after pre-processing saved in three sub directories of dataset folder. Each separate folder is loaded using dataloader function of pytorch.

- **CNN Model**

  CNN is a network that use an image, image combine with filters that produce features. Using convolution layers, pooling and ReLu activation function at output layer, the tumor classification is done. Convolutional layers, max pooling and fully connected layers are used.

- **Training**

  In training phase, we used 800 images form each class for training. Using a learning algorithm, we update the filters at the convolution layers. Weights are updated and learning algorithm learns the input using a learning or optimizer algorithm SGD (Stochastics Gradient Descent) to update the filters at the convolution layers. The learning algorithm takes a classification error or loss. The loss act as an input and back propagates the error into the network. Then weights and filters are updated. This process continued. After training, the model we save the model with .pth file extension. This file has all the information of the network and ready to reuse again when we perform test single image for GUI.

- **Testing**

  We used 200 images for testing from each tumor classes. Testing data give a glimpse of CNN was trained and able to classify the tumor type.

- **Output**

  Accuracy of the whole model is returned in output window. Accuracy of each tumor type is also predicted separately. When all steps stated above run successfully, our model return that output as accuracy of whole network, testing accuracy, training accuracy, accuracy of Meningioma, Glioma and Pituitary.

5.1.5.5 Classification Accuracies

Following are the accuracies we have achieved:

Table 5-3 Classification accuracies

Labels	Tumor class	Accuracy	Overall accuracy
1	Meningioma	100 %	100 %
2	Glioma	100 %	
3	Pituitary tumor	100 %	

5.1.5.6 Accuracy Graph

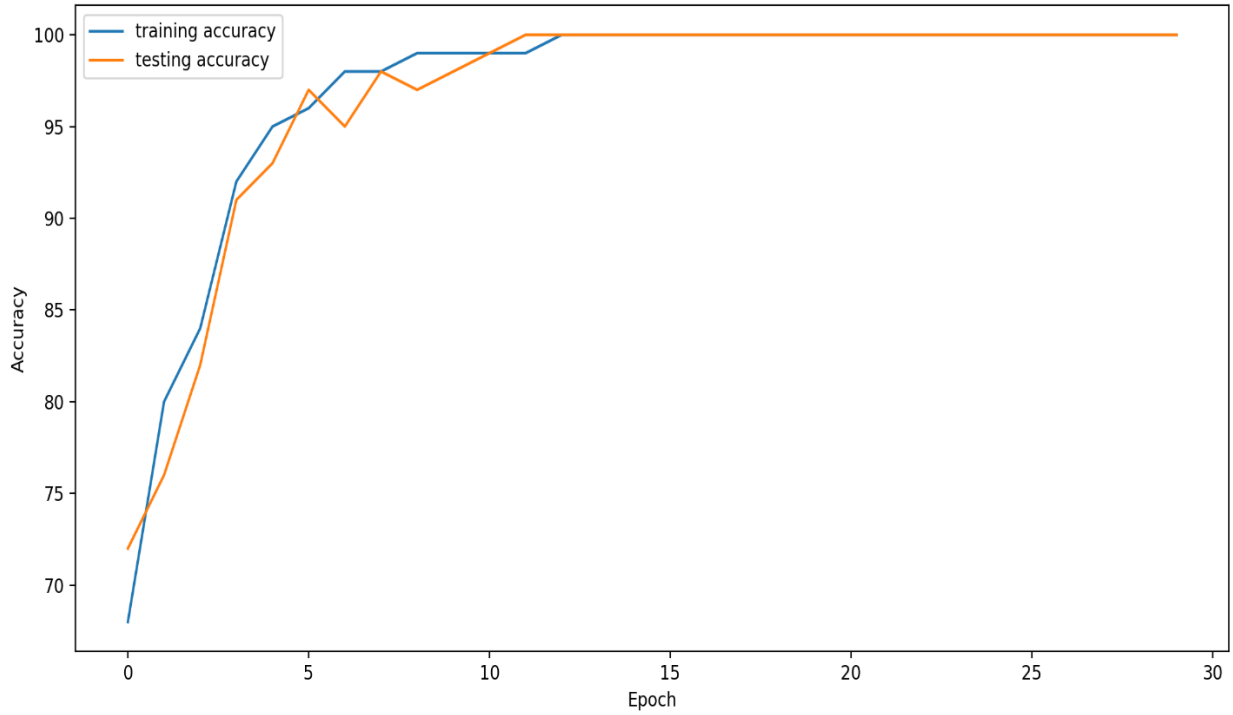


Figure 5-4 Accuracy graph

5.1.5.7 Demo

```
============================================================
Epoch 30 ::
training loss:0.000033 %
training accuracy:100.000000 %
test loss:0.000000 %
test accuracy:100.000000 %
============================================================
Accuracy of the network on the  test images: 100 %
Accuracy of  Meningioma : 100.000000 %
Accuracy of Glioma : 100.000000 %
Accuracy of Pituitary  : 100.000000 %
```

Figure 5-5 Training results

5.2 User Interface

Simple interface will let the user to understand easily. After installation of QtDesigner, interface of application is designed. Buttons allow user for clicking and uploading the image; each button is named as per functionality so that it makes sense. As application will start splash screen will open and loads progress bar. As loading complete main screen will open and allow functionalities of all modules.

Working of each module and interface working has been shown below step by step.

5.2.1 Splash Screen

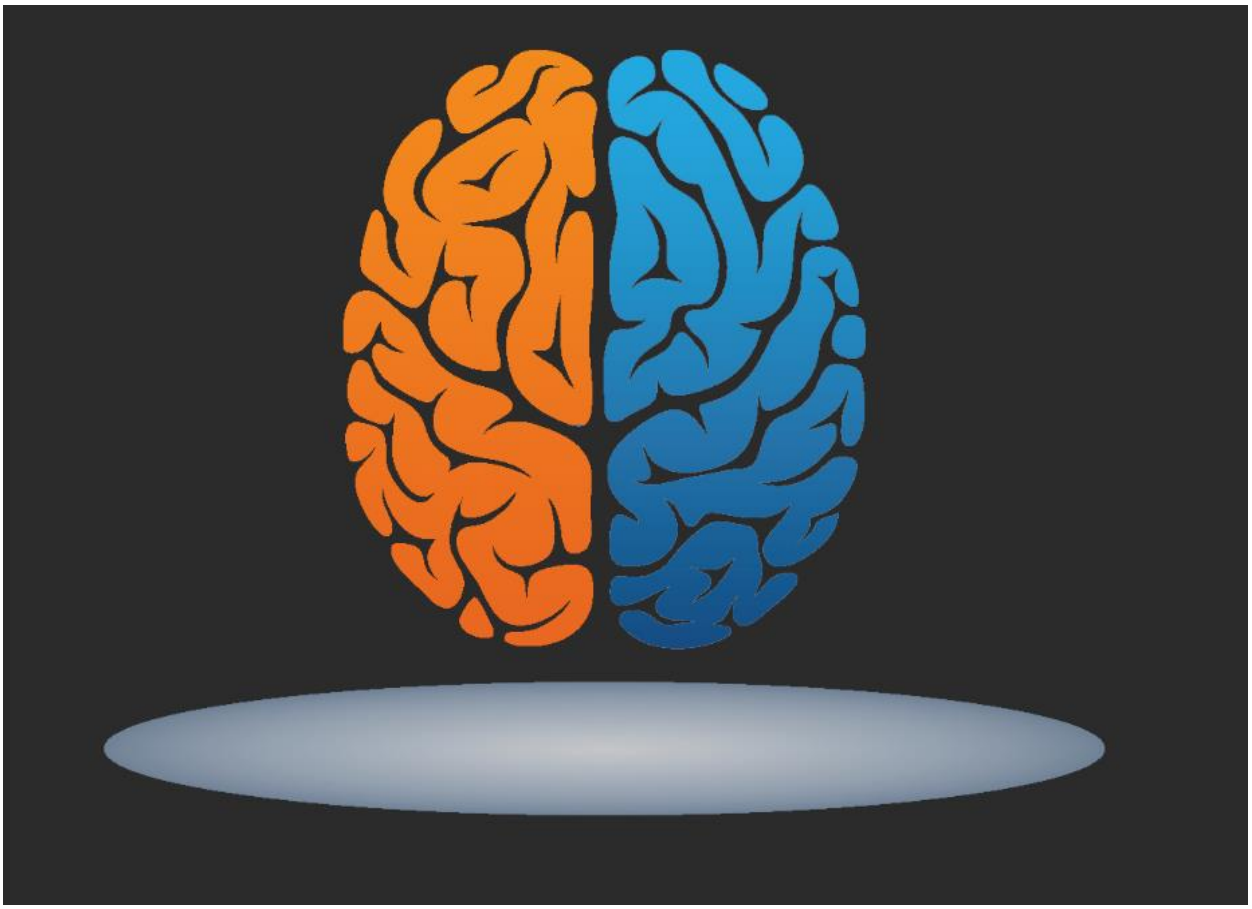


Figure 5-6 Splash Screen

5.2.2 Main Window

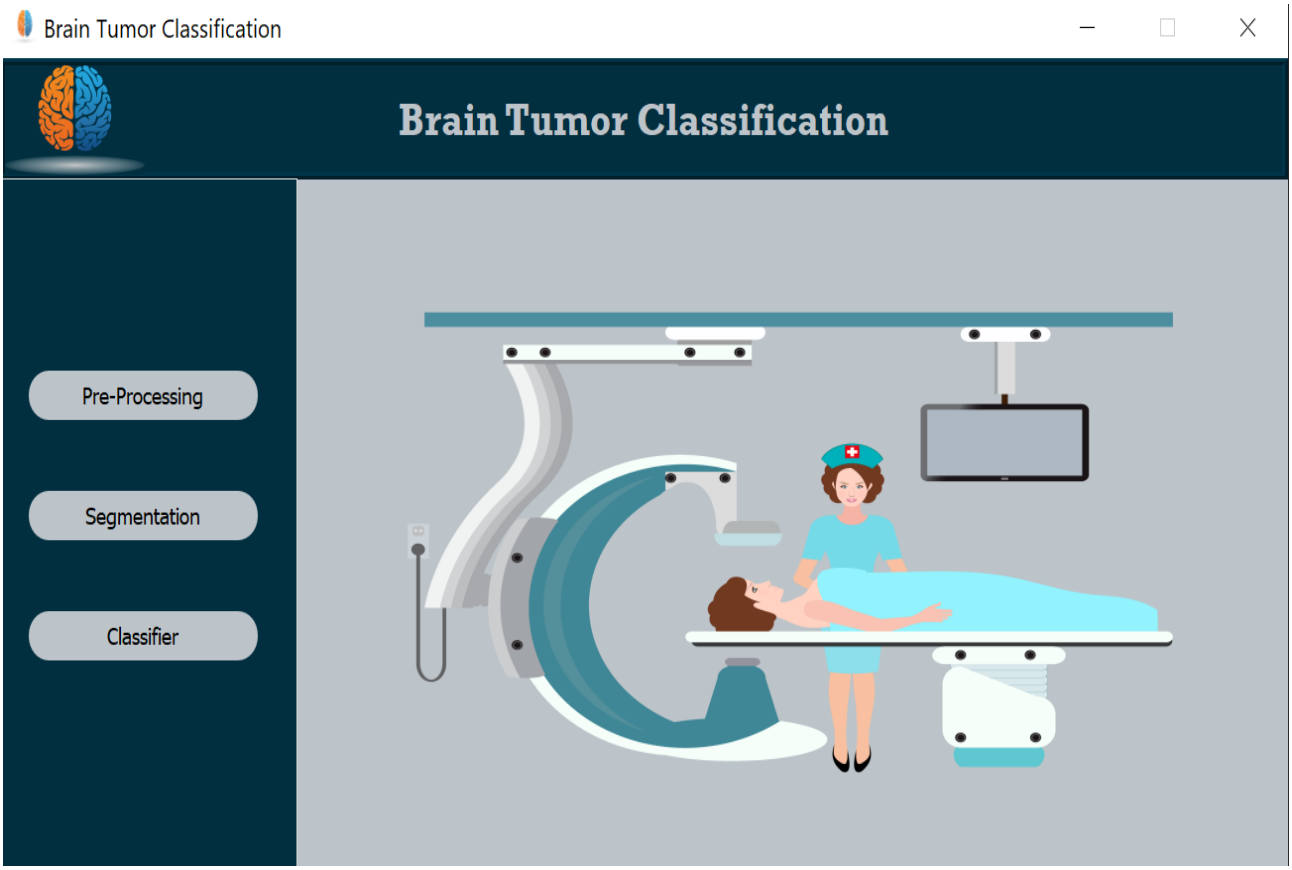


Figure 5-7 Main Window

5.2.3 Preprocessing

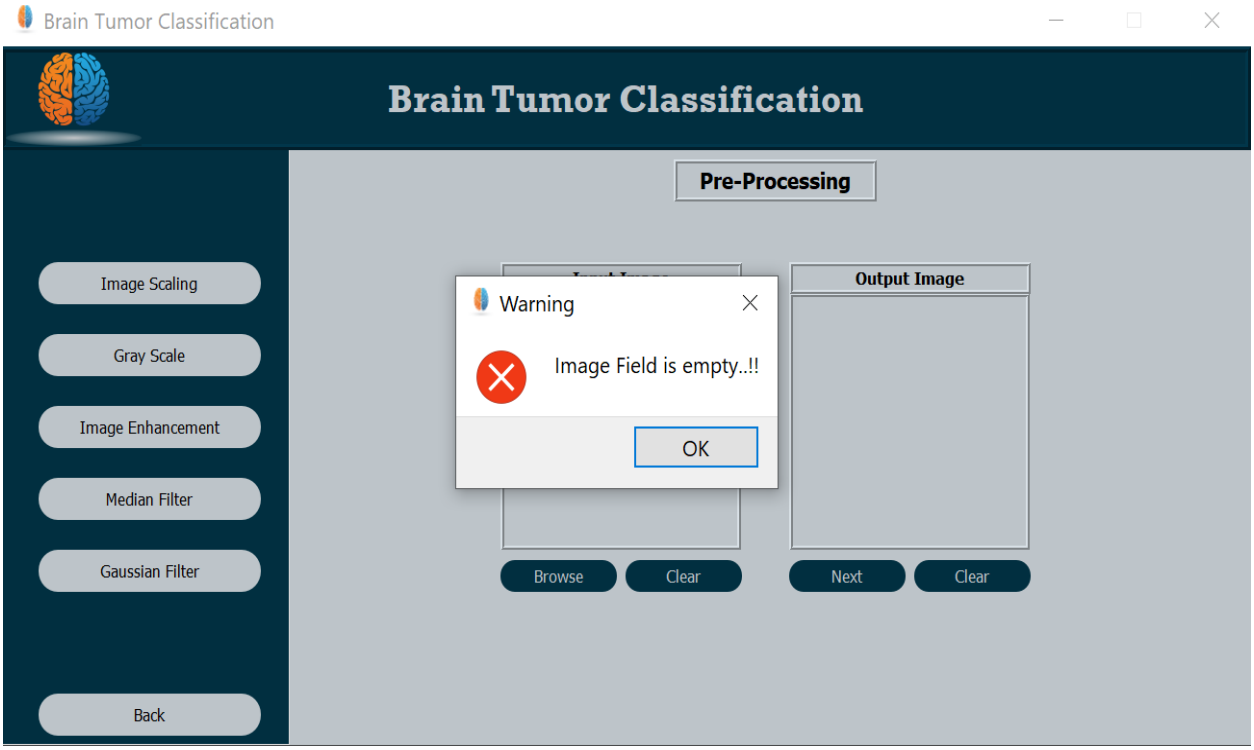


Figure 5-8 Without upload image

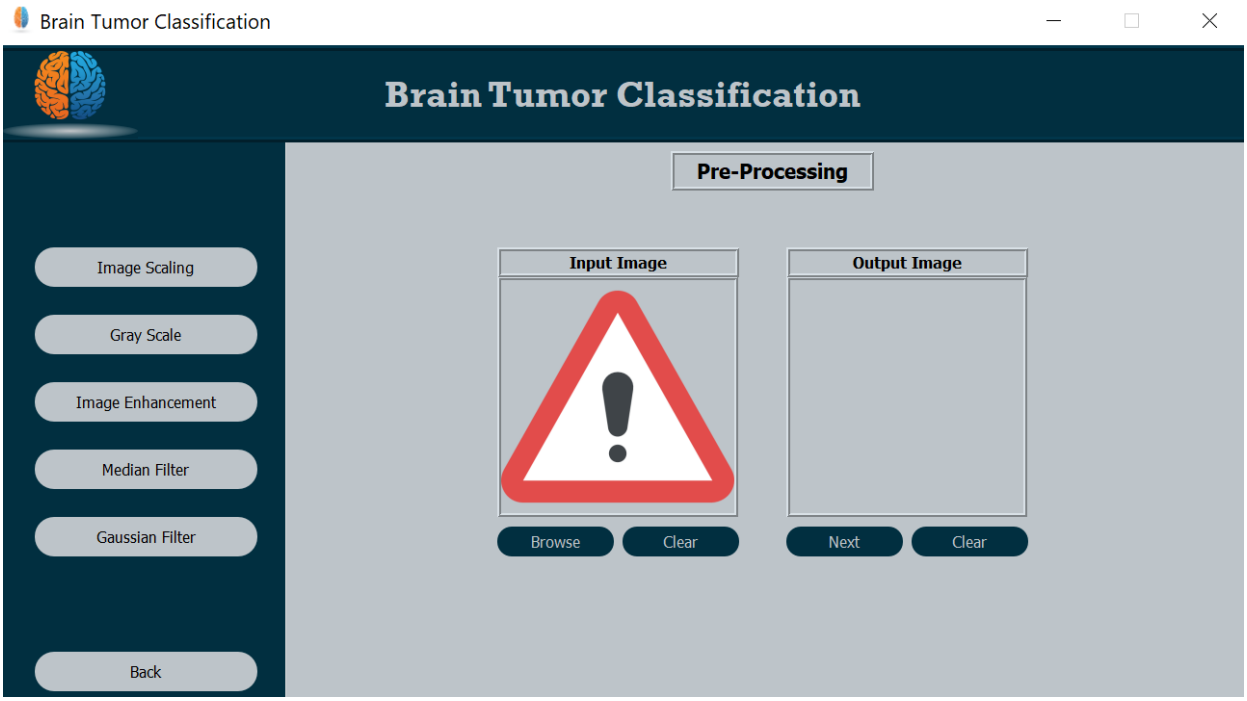


Figure 5-9 Error message

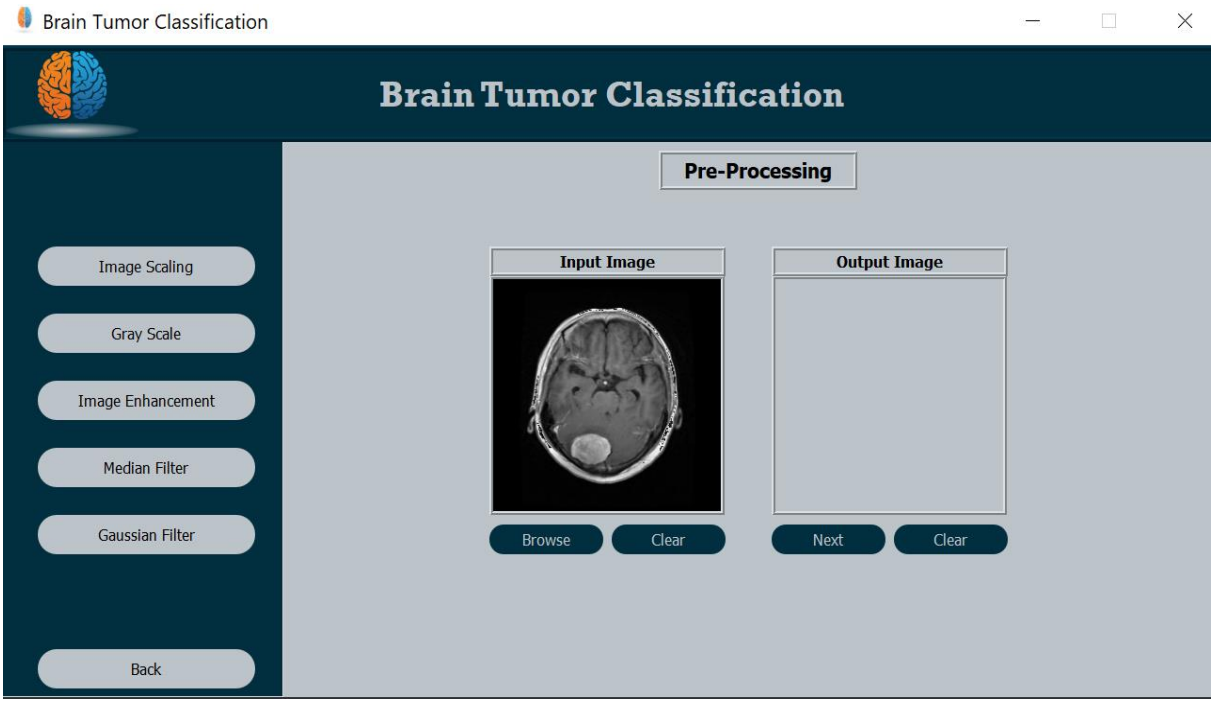


Figure 5-10 Image upload

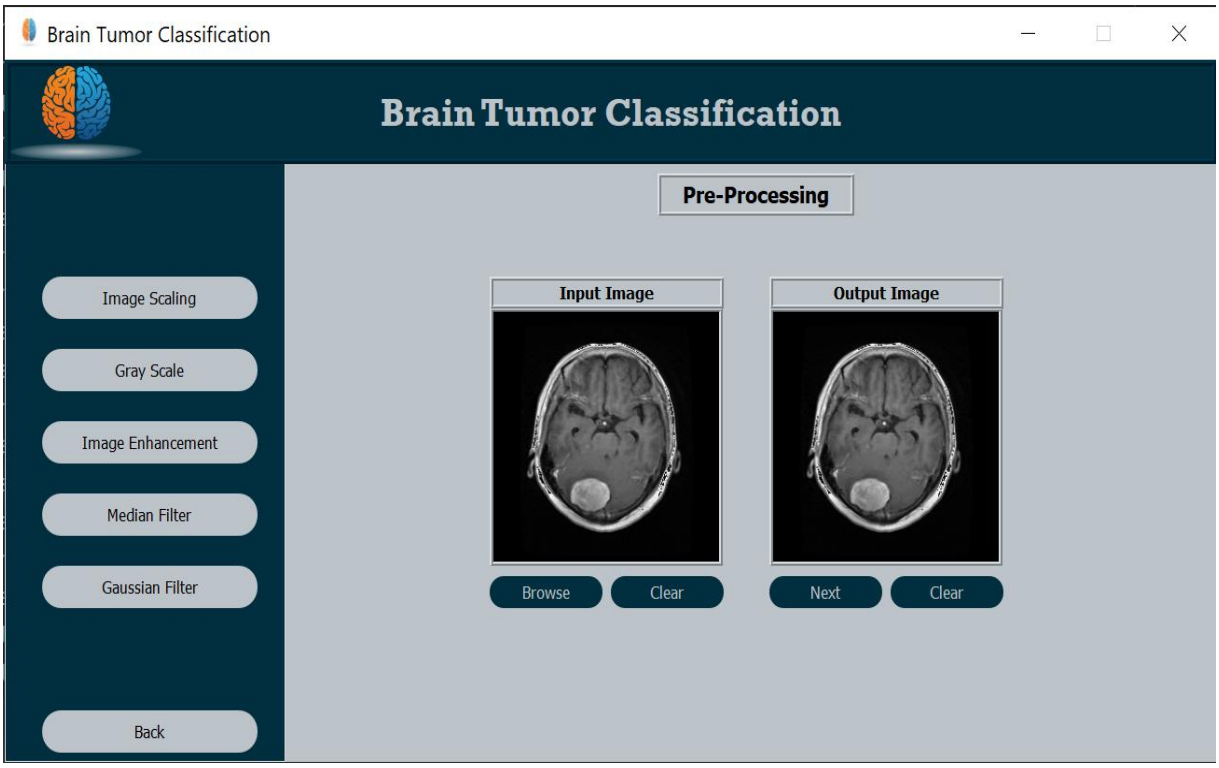


Figure 5-11 Grey Scale

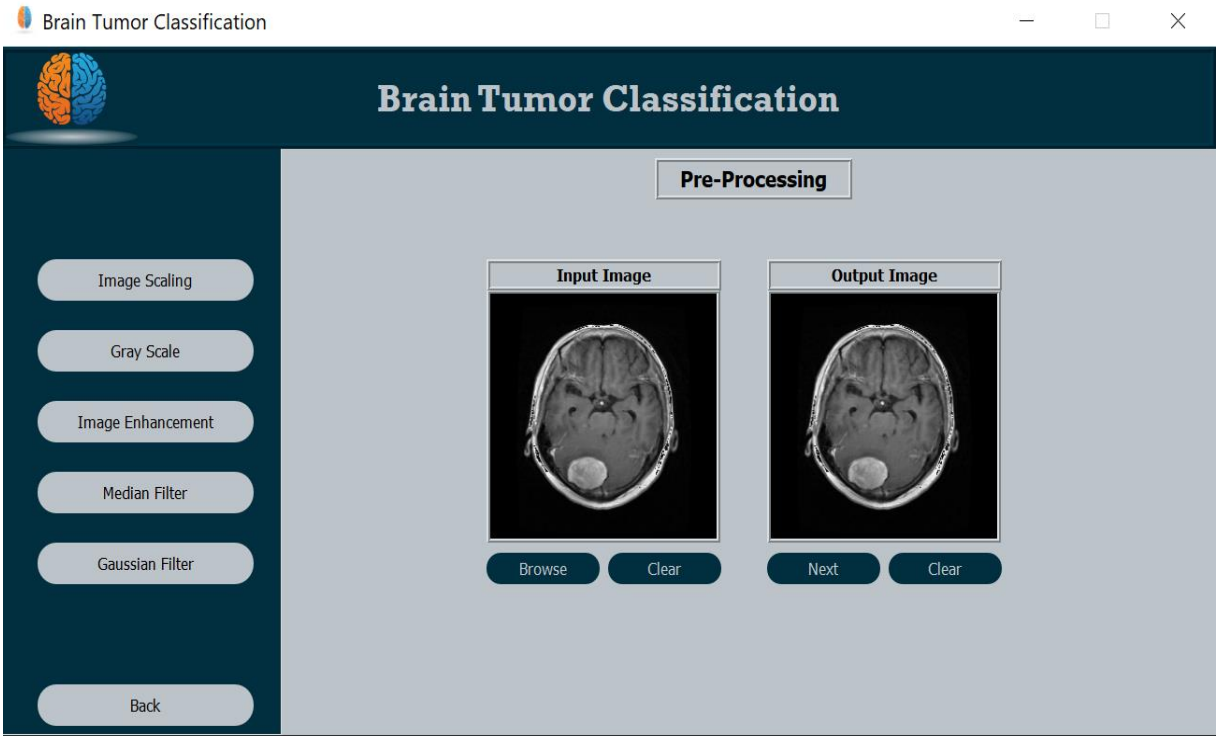


Figure 5-12 Image Enhancement

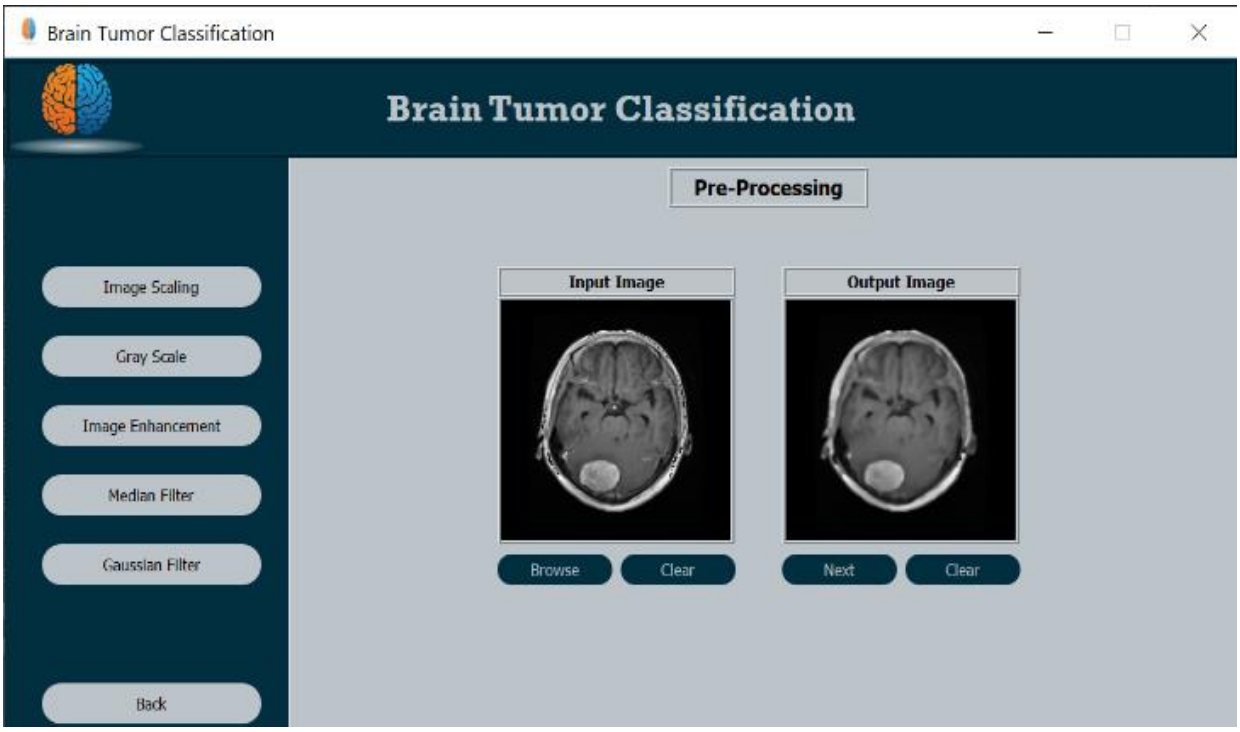


Figure 5-13 Median filter

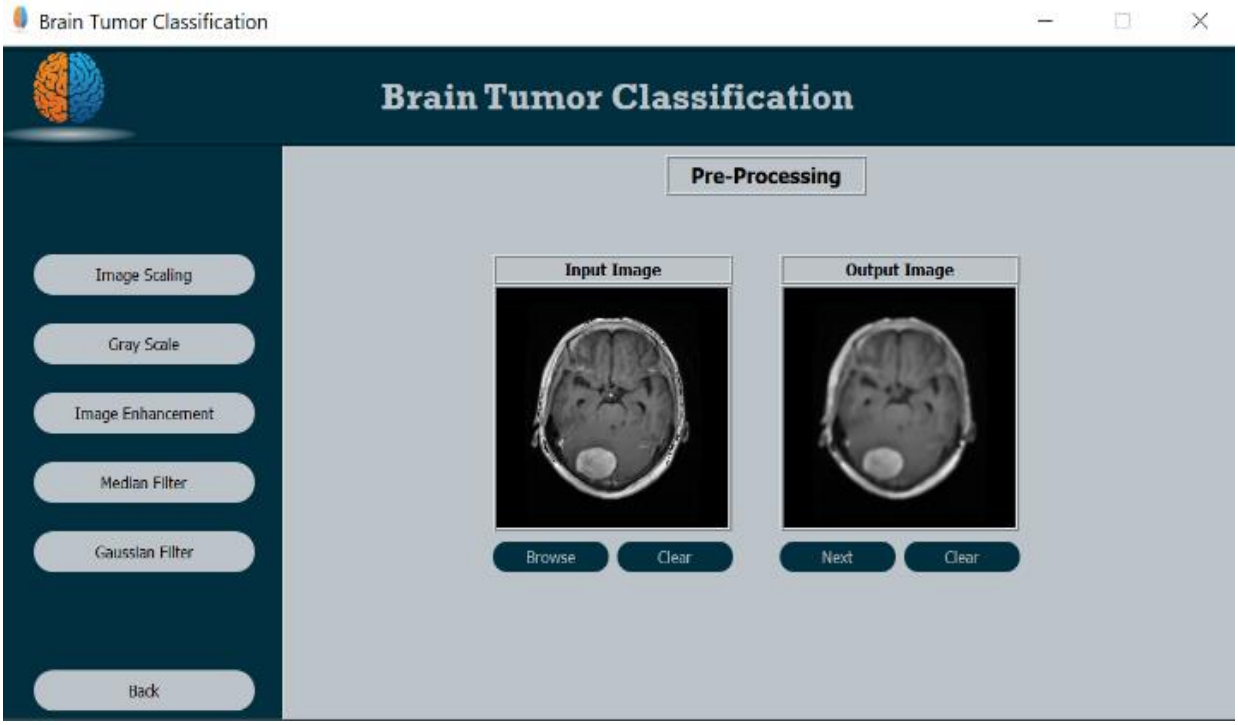


Figure 5-14 Gaussian filter

5.2.4 Segmentation

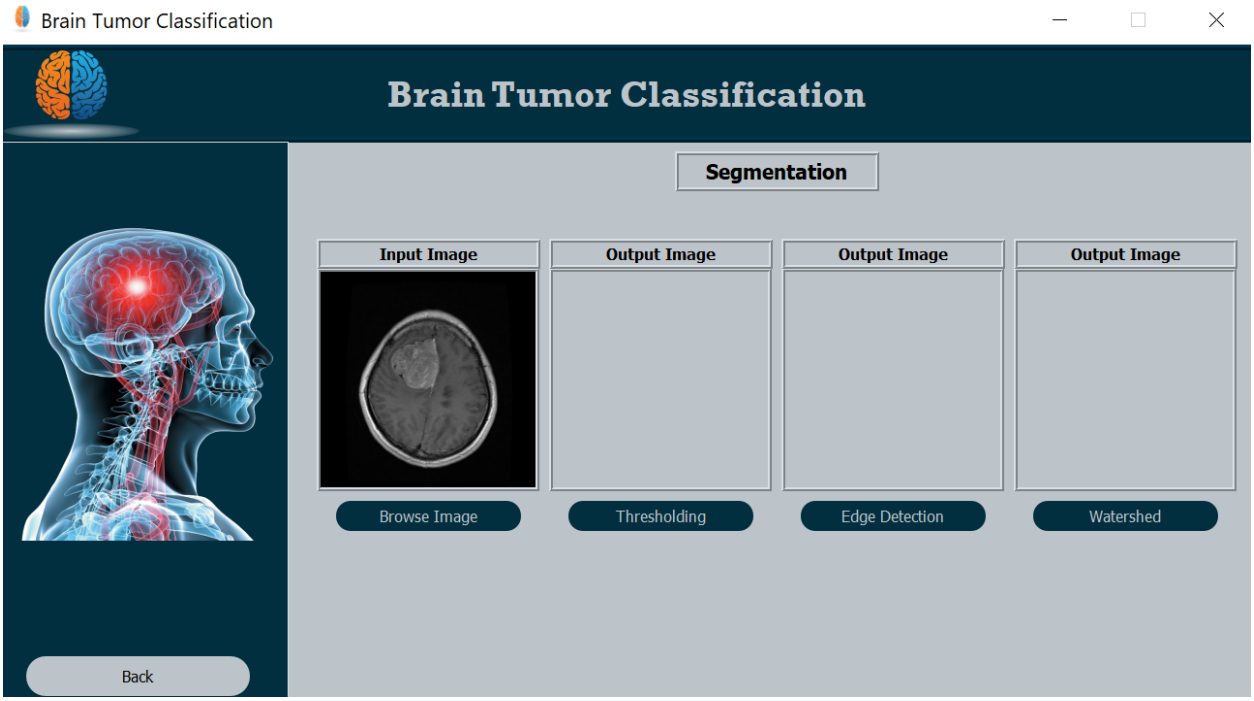


Figure5-15 Segmentation

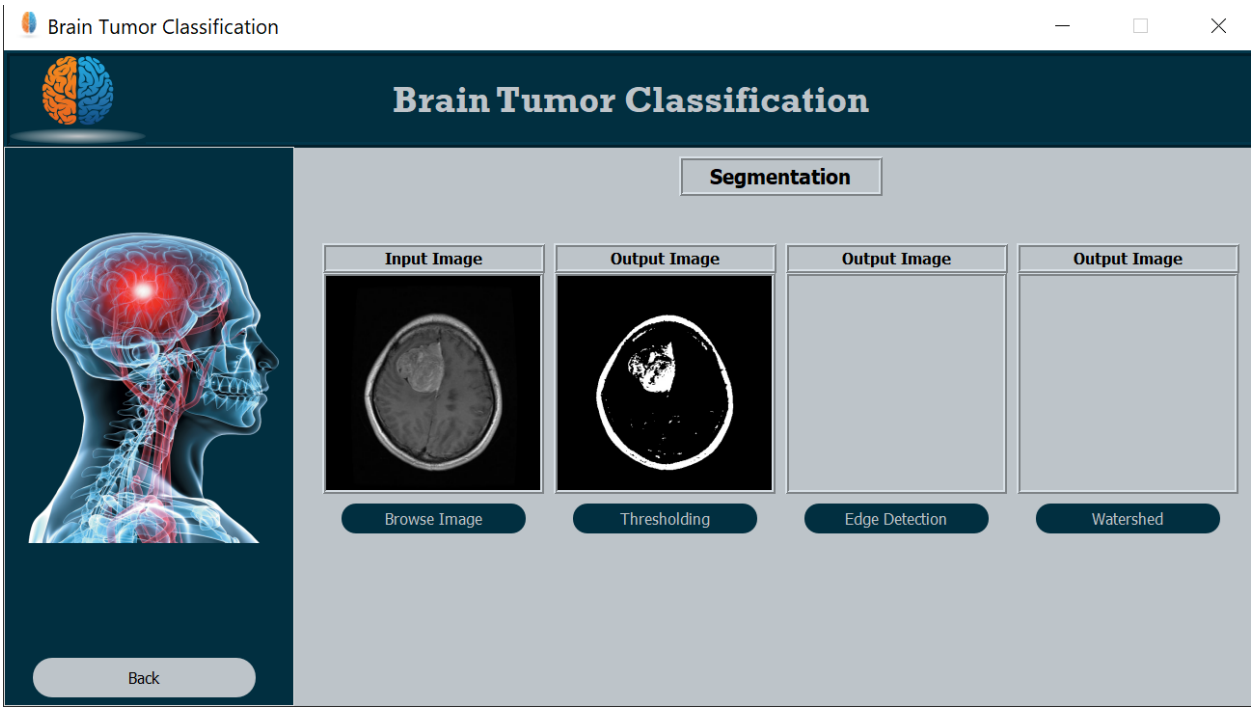


Figure 5-16 Thresholding

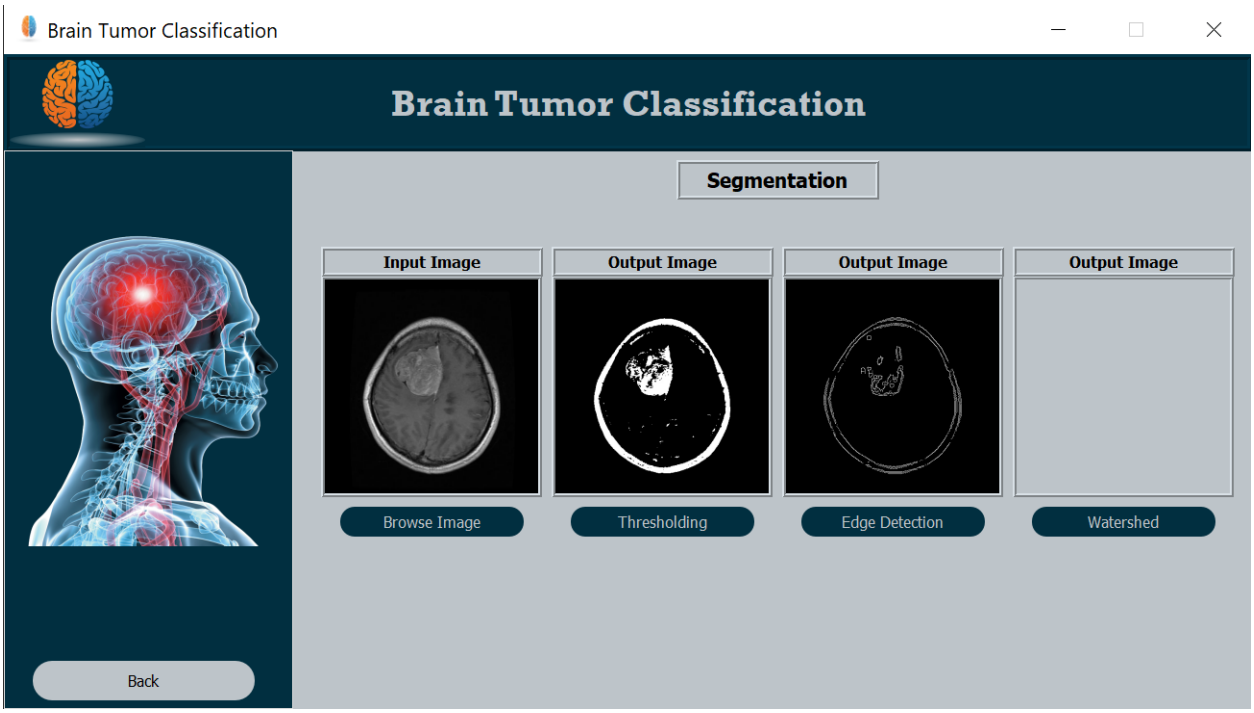


Figure 5-17 Edge detection

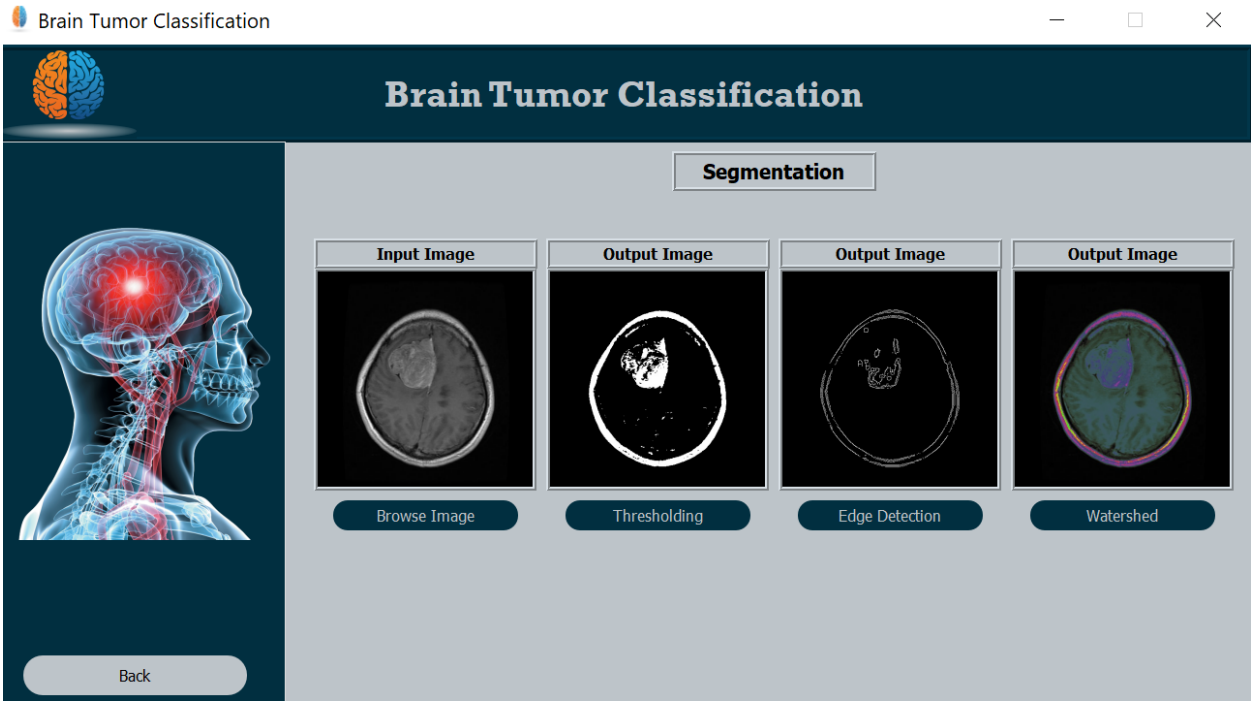


Figure 5-18 Watershed algorithm

5.2.5 Classifier

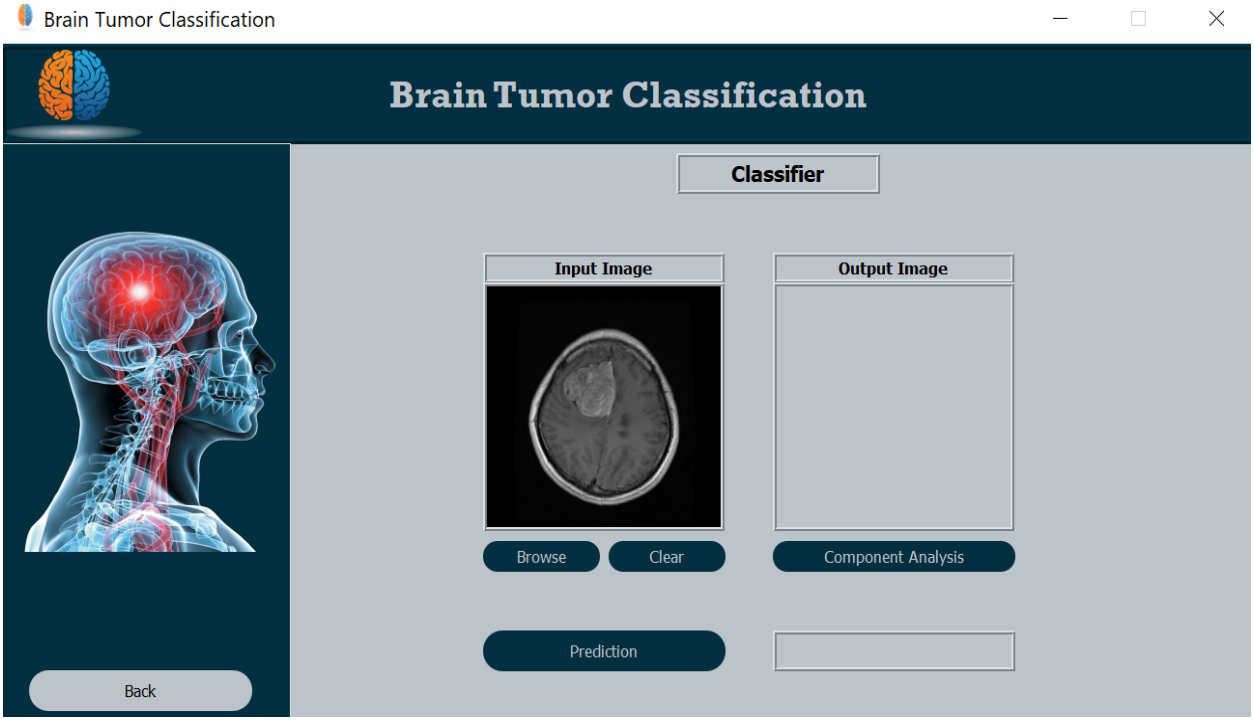


Figure 5-19 Classifier

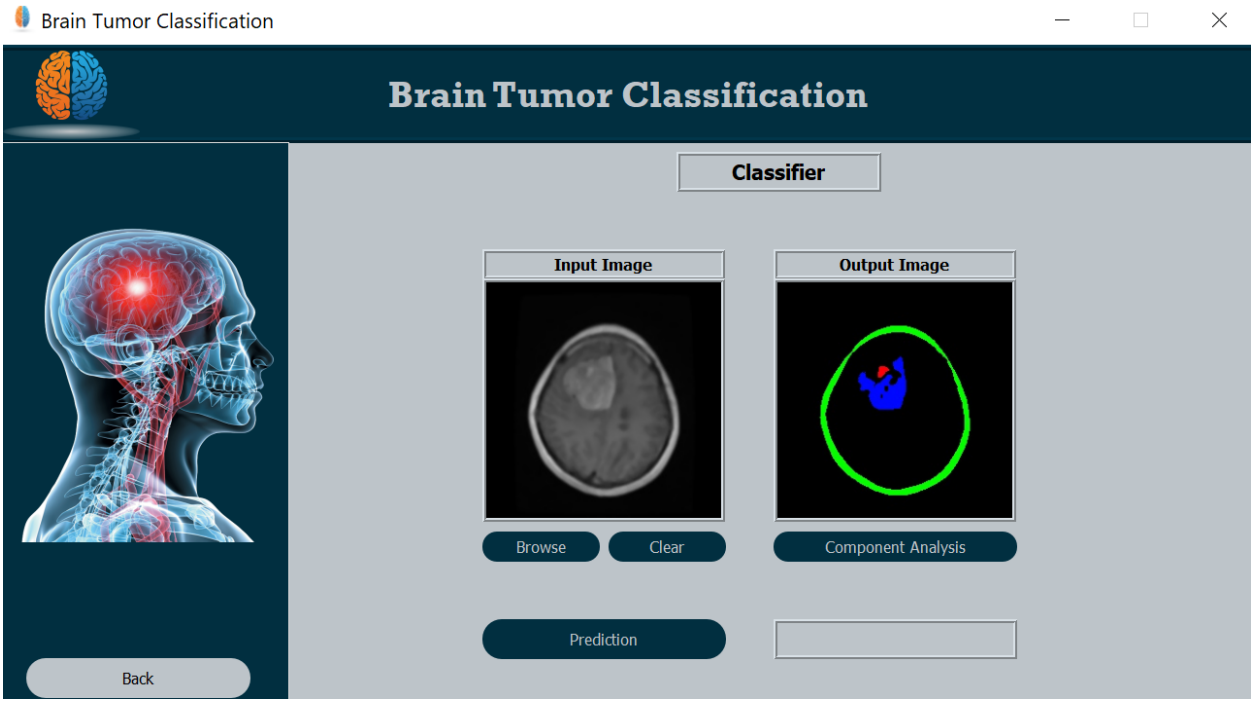


Figure 5-20 Component analysis algorithm

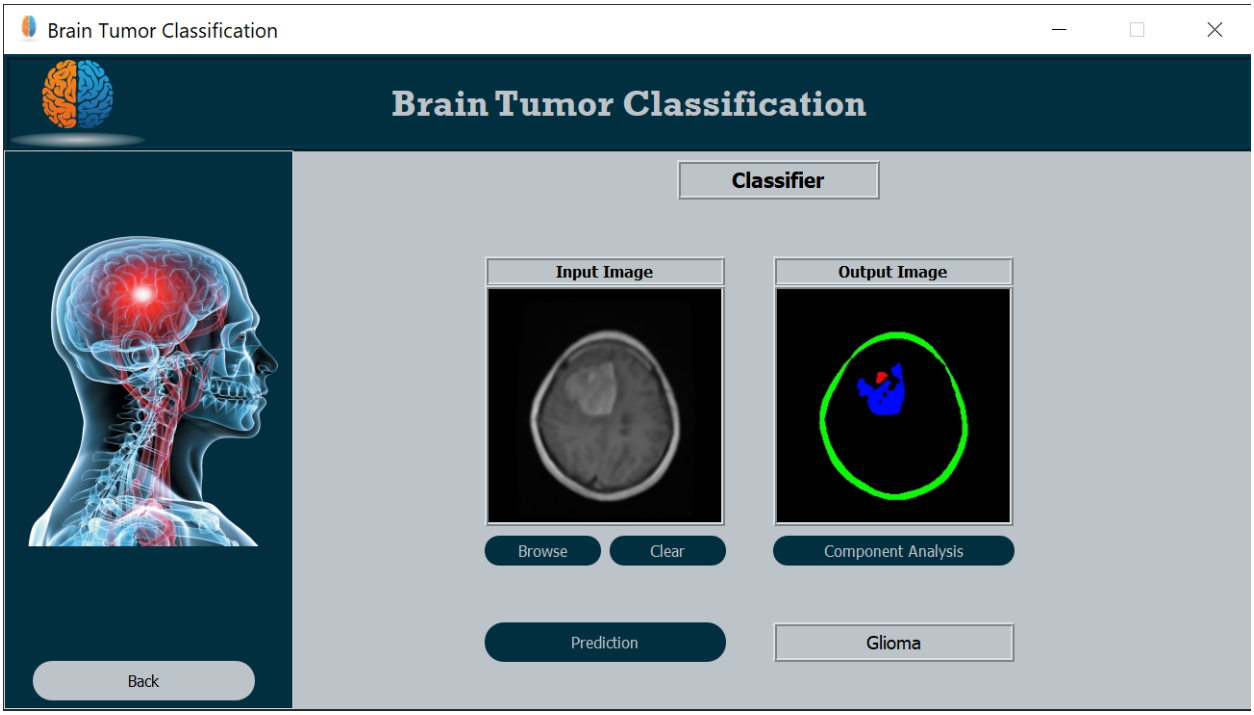


Figure 5-21 Label predicted

CHAPTER 6

TESTING AND EVALUATION

6. Testing and Evaluation

Evaluation is stage where we analyze performance of application by performing different tests to check the effectiveness. Testing is an important phase for any software system. In this phase, we test the functionality of each module of the system in an organized way and test that our system is according to user's requirement or not.

6.1 Manual Testing

In manual testing, there is no use of any automated tool and test cases are executed manually. It helps to find out visible and hidden defects in a system. Firstly, the documentation is observed by tester to know about testing areas of system. She analyses the requirement document to fulfill all needs of physician. Each line of code is examined and then functionality of each module is checked.

6.1.1 System Testing

In system testing, complete system is evaluated on basis of functionality. The Graphical User Interface (GUI) will be tested in system testing. The input will be given to the system through the GUI and the outputs will be checked. It includes both functional and non-functional testing. The MRI brain tumor image is provided to system and classification results are examined.

6.1.2 Unit Testing

In unit testing, we test each unit of our system and check whether the system is according to the user requirements or not. This testing can be performed in each unit. Unit testing is mentioned for whole CNN model.

6.1.2.1 Upload Image

Image uploading done by clicking the button name “Upload” it opens the image dataset (.JPG) format. User can select any image from test folder and that image will be display inside label at interface.

6.1.2.2 Preprocessing

Image and labels are extracted from .mat files. Images have jpg format. No other format is supported. All images are saved with their label name. The five operations of preprocessing i.e. image scaling, grey scale, image enhancement, median filter and gaussian filter are applied when the image is uploaded.

6.1.2.3 Segmentation

When preprocessed images are passed to segmentation algorithms i.e. thresholding, canny edge detection, watershed algorithm then segmented tumor images are examined.

6.1.2.4 Dataset

- **Training dataset**

  Our training data has three further directories. All images have (64,64) dimensions. We made sure that each MRI image has its correct label and saved in correct directory. Train folder is divided in three separate folders as Meningioma, Glioma and Pituitary. Total 800 images are taken from each folder for training.

- **Testing dataset**

  Testing images are placed in test folder. All images have (64,64) dimensions Test folder is divided in three separate folders as Meningioma, Glioma and Pituitary. There are 200 images

from each directory of tumor class. Then our train folder is ready to load in our CNN model. System can only process images in 64x64 size.

6.1.2.5 Classification

For classification, first we made a CNN model that has been trained on our dataset. Our pretrained model is saved in (.pth) extension which we gave to our system and test the given image, the input image is compared with pre trained CNN load model. Then it detects the tumor type if the tumor image is matched then our system displays the result.

6.1.3 Functional Testing

In functional testing, we test the functionality of the system that the functions are working according to the requirements or not. For functional testing, we generate the test cases. They are performed to check usefulness of application from user point of view.

Test Case 1

Table 6-1 Upload image

No.	Test case	Attribute and value	Expected result	Result
1	Upload image	The user is required to select image from folder to know identified tumor.	Selected image is ready to be displayed.	Selected image successfully displayed on screen.

Test Case 2

Table 6-2 Without upload image

No.	Test case	Attribute and value	Expected result	Result
2	Without uploading image	If user click button of any operation.	Error message to be shown.	Error message shown instead of picture.

Test Case 3

Perform the extraction of images from (.mat) structures to (.jpg) format. JPG images should be extracted along with labels and five operations are performed.

Table 6-3 Preprocessing

No.	Test case	Attribute and value	Expected result	Result
1	Image Scaling	Test image should be available in directory.	Image should be scaled and shown in output.	Image scaled and shown successfully.

2	Greyscale	Test image should be available in directory	Image should be shown in output.	Image greyscale and shown successfully.
3	Image Enhancement	Test image should be available in directory	Image should be shown in output.	Image enhanced and shown successfully.
4	Median Filter	Test image should be available in directory	Image should be shown in output.	Image noise removed and shown successfully.
5	Gaussian Filter	Test image should be available in directory	Image should be shown in output.	Image noise removed and shown successfully.

Test Case 4

Table 6-4 Segmentation

No.	Test case	Attribute and value	Expected result	Result
1	Edge Detection	Test image should be available in directory	Tumor mask to be generated in output.	Tumor region is highlighted successfully in output.
2	Thresholding	Test image should be available in directory	Tumor mask to be generated in output.	Tumor region is highlighted successfully in output
3	Watershed	Test image should be available in directory	Tumor mask to be generated in output.	Tumor region is highlighted successfully in output

Test Case 6

Table 6-5 Classify Tumor

No.	Test case	Attribute and value	Expected result	Result
6	Prediction	Compare input image with pretrained load model.	Predict the label of image.	Label of image predicted successfully.

Test Case 7

Table 6-6 Another organ MRI

No.	Test case	Attribute and value	Expected result	Result
7	Another organ MRI	Any object appears.	The image should not be classified.	Classification unsuccessful.

6.1.4 Integration Testing

When all modules are integrated and when frontend linked to backend, then the result shown is correct.

Table 6-7 Integration testing

No.	Test case	Attribute and value	Expected result	Result
1	Upload image	The user is required to select image from folder to know identified tumor.	Selected image is ready to be shown.	Selected image successfully shown on screen.
2	Without uploading image	If user click button of any operation.	Error message to be shown.	Error alert shown instead of picture.
3	Pre-processing	Output image available in directory.	JPG images should be extracted along with labels and five operations are performed.	Image preprocessed successfully.
4	Segmentation	Image should be preprocessed.	Tumor mask to be generated.	Tumor region is highlighted successfully.

5	Classify tumor	Compare input image with pretrained load model.	Predict the label of image.	Label of image predicted successfully.
6	Another organ MRI	Any object appears.	The image should not be classified.	Classification unsuccessful.

6.2 Automated Testing

We only performed manual testing.

6.3 Tools

Table 6-8 Tools

Tool Name	**Tool Description**	**Applied on [list of related tests cases / FR / NFR]**	**Results**
Matlab	MATLAB 17b	Preprocessing	Images and labels extracted from (.mat) files.
PyCharm	JetBrains PyCharm Community Edition 2019.2.1	All test cases.	Code is written over it.

Anaconda distribution	Anaconda3-2019.10-Windows-x86_64	All test cases	It supports libraries
PyQt Designer	Comes from library of python, PyQt5.	Used for designing interface.	It displays frontend of system.
Pytorch	Python library module	All test cases.	It supports libraries.
Microsoft Word	WORD 2016	Used for documentation.	The documentation is made.
Microsoft PowerPoint	POWERPOINT 2016	Used for presentation slides.	The presentation slides are made.

CHAPTER 7

CONCLUSION AND FUTURE WORK

7. Conclusion and Future Work

7.1 Conclusion

Abnormal growth of tissue in the brain which affect the normal functioning of the brain is considered a brain tumor. The main goal of medical image processing is to identify accurate and meaningful information using algorithms with minimum error possible. Brain tumor classification through MRI images can be categorized into four different sections: pre-processing, image segmentation, feature extraction and image classification. It can be concluded that the algorithms and the parameters used in the proposed system are all meant to increase the efficiency of the system by achieving better results. Accuracy and reliability are of utmost importance in tumor diagnosis, as a patient's life depends on the results predicted by the system.

A lot of machine learning work has been done for development of this application. There is lot of work, which can incorporate in future. This is just beginning not the ending point. it was experimentally found that Convolution Neural Network generate the promising results and they are considered as state of art in computer vision and medical imaging. In this project, we implemented the CNN based computer-aided system which classify the brain tumor type which does not required region segmentation. Without prior segmentation, the CNN returns the good results. We achieved the 100% accuracy of the network. Three types of tumors: Meningioma, Glioma and Pituitary are used for prediction. Meningioma tumors have 100% classification accuracy. They are hard to predict due to their complex position in brain. We achieved 100% accuracy for pituitary and glioma tumor images. All implementation is done using machine learning frameworks mentioned previously. This application will take a minute or two to identify tumor type rather than must wait long.

7.2 Future Work

Encouraged by these results, future work will involve the improvement of classification result and overall accuracy. There are multiple tumor types exist and need predictions. The number of tumor classes can also be increased if more data is available. With a more extensive and diverse dataset, the overall classification accuracy can be dramatically increased. Also, this automatic system can be implemented in clinics for doctors' supportive tool to make decisions and treatments. Moreover, complex architectures, complex brain images, region-based CNN and combinations of multiple algorithms can enhance the performance of CNN.

CHAPTER 8

REFERENCES

8. References

1. https://www.sciencedirect.com/science/article/pii/S2314728817300636#fig1
2. https://ieeexplore.ieee.org/abstract/document/5479975
3. https://ieeexplore.ieee.org/document/6804439
4. https://pdfs.semanticscholar.org/3889/dc9221de455f782a3b2b55a925454783969f.pdf.
5. Abiwinanda N., Hanif M., Hesaputra S.T., Handayani A., Mengko T.R. (2019) Brain Tumor Classification Using Convolutional Neural Network. In: Lhotska L., Sukupova L., Lacković I., Ibbott G. (eds) World Congress on Medical Physics and Biomedical Engineering 2018. IFMBE Proceedings, vol 68/1. Springer, Singapore.
6. Justin S. Paul, Andrew J. Plassard, Bennett A. Landman, and Daniel Fabbri "Deep learning for brain tumor classification", Proc. SPIE 10137, Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging, 1013710 (13 March 2017); doi: 10.1117/12.2254195;https://doi.org/10.1117/12.2254195.
7. Cheng J, Huang W, Cao S, Yang R, Yang W, et al. (2015) Correction: Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. PLOS ONE 10(12): e0144479.
8. Cheng J, Yang W, Huang M, Huang W, Jiang J, Zhou Y, et al. (2016) Retrieval of Brain Tumors by Adaptive Spatial Pooling and Fisher Vector Representation. PLoS ONE 11(6): e0157112. https://doi.org/10.1371/journal.pone.0157112.
9. https://ieeexplore.ieee.org/document/6804439
10. https://pytorch.org/tutorials/beginner/data_loading_tutorial.html
11. https://figshare.com/articles/brain_tumor_dataset/1512427
12. https://ieeexplore.ieee.org/document/8934603.

13. https://arxiv.org/abs/2001.08844.
14. https://www.sciencedirect.com/science/article/pii/S2314728817300636.
15. https://www.sciencedirect.com/science/article/abs/pii/S1877750318307385.
16. https://www.ijser.org/researchpaper/Efficient-Brain-Tumor-Detection-Using-Image-Processing-Techniques.pdf.
17. https://www.researchgate.net/publication/308757676_A_Survey_on_Techniques_for_Brain_Tumor_Segmentation_from_Mri.
18. https://ieeexplore.ieee.org/document/7939554.
19. https://www.ijettcs.org/Volume4Issue5(2)/IJETTCS-2015-10-01-7.pdf.