2019-10

Brain Tumor Detection and Classification Using Symmetrical Side Analysis and Thresholding Technique

Tsegaye, Zelalem

BAHIR DAR UNIVERSITY

BAHIR DAR INSTITUTE OF TECHNOLOGY

SCHOOL OF RESEARCH AND POSTGRADUATE STUDIES

FACULTY OF COMPUTING

Brain Tumor Detection and Classification Using Symmetrical Side Analysis and Thresholding Technique

Zelalem Tsegaye Hailu

October 15, 2019

Bahir Dar, Ethiopia

BRAIN TUMOR DETECTION AND CLASSIFICATION USING SYMMETRICAL SIDE ANALYSIS AND THRESHOLDING TECHNIQUE

Zelalem Tsegaye Hailu

A thesis submitted to the school of Research and Graduate Studies of Bahir Dar Institute of Technology, BDU in partial fulfillment

Of

The requirements for the degree of Master of Science in Information Technology in the Faculty of Computing.

Advisor Name: Gebeyehu Belay (Dr.Eng.) Asst Prof

BahirDar, Ethiopia

October 15, 2019

DECLARATION

I, the undersigned, declare that the thesis comprises my work. In compliance with internationally accepted practices, I have acknowledged and refereed all materials used in this work. I understand that non-adherence to the principles of academic honesty and integrity, misrepresentation/ fabrication of any idea/data/fact/source will constitute sufficient ground for disciplinary action by the University and can also evoke penal action from the sources which have not been properly cited or acknowledged.

Name of the student: Zelalem Tsegaye Signature _____________

Date of submission: October 15, 2019

Place: Bahir Dar

This thesis has been submitted for examination with my approval as a university advisor.

Advisor Name: Gebeyehu Belay (Dr,Eng.) Asst Prof

Advisor's Signature: ______________________________

THESIS APPROVAL SHEET

Student:

[illegible] ____ October [illegible], 2019

Name ____ Signature ____ Date

The following graduate faculty members certify that this student has successfully presented the necessary written thesis Document and oral presentation of this Document for partial fulfillment of the thesis-option requirements for the Degree of Master of Science in Information Technology

Approved by Board of Examiners

School Director ____ Signature ____ Date

[illegible] (Dr.) ____ ____ 02/02/2012 E.C

Advisor ____ Signature ____ Date

Fekade Getahun (Dr.) ____ ____ 11/10/2019

External Examiner ____ Signature ____ Date

[illegible] (PhD) ____ ____ [illegible]

Internal Examiner ____ Signature ____ Date

[illegible] ____ ____ 30/01/2012 E.C

Faculty dean ____ Signature ____ Date

To My Family

ACKNOWLEDGEMENT

First and foremost, I **would like to thank God Almighty** for giving me the strength, knowledge, ability and opportunity to undertake this research study and to persevere and complete it. Without his blessings, this achievement **would** not have been possible.

Next, I would like to express my deepest gratitude to my advisor Dr. Gebeyehu for his motivation and constructive guidance right from the moments of problem formulation. Many thanks and appreciations go to him for the discussions with him always made me think that things are possible.

My special thanks go to Mr. Assaminew head of Wuddassie diagnostic laboratory and Dr. Fissha who has helped me in identifying the characteristics and types of brain tumor and has provided me resources that helped me in understanding the anatomy of the human brain and the MRI images.

My special thanks for all staffs of BiT for providing such a conducive environment throughout my study.

At last but not least I would like to thank my family and friends, I wouldn't achieve this without their support and motivation. Thank you so much for everything, for making this possible directly or indirectly.

ABSTRACT

A brain tumor is an abnormal growth of cells in the brain which hinders the proper function of the brain and eventually lead to death if not treated properly. The probability of survival rate can be improved if the tumor is detected early and classified properly. Manually detecting and segmenting brain tumors in today's brain Magnetic Resonance Imaging (MRI), where a large number of MRI scans taken for each patient, is tedious and subjected to inter and intra observer detection and segmentation variability. Therefore, there is a need for computer aided brain tumor detection and segmentation from brain MR images to overcome the tedium and observer variability involved in the manual segmentation. As result a number of methods have been proposed in recent years to fill this gap, but still there is no commonly accepted automated technique by clinicians to be used in clinical floor due to accuracy and robustness issues. In This thesis we proposed a new segmentation method that can overcame the drawback of global thresholding in identifying regions that has lower intensity, to extract brain tumors in MRI images, which is based on symmetrical side analysis combined with automatic thresholding to extract tumor part from the brain. The anticipated method can effectively be useful to distinguish the shape of the tumor and its geometrical measurement. Also in this thesis, we used a histogram of oriented gradient (HOG) to extract the features and Support Vector Machine (SVM) for brain tumor classification. Based on the confusion matrics result observed the proposed method has produced a better result with F-measure of 0.8031 or 80.31% accuracy than the other two segmenting mechanisms which are global thresholding and Sobel edge detection which has f-measure of 0.685 and 0.7021.

KEYWORDS: MRI Imaging, brain tumor, symmetrical side analysis, segmentation, Tumor Detection

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
CAD	Computer-Aided Design
CAT	Computed Axial Tomography
CNS	Central Nervous System
CT	Computed Tomography
DCT	Discrete cosine transform
DWT	Discrete Wavelet Transform
EM	Electromagnet
FCM	Fuzzy C-Mean
FFT	Fast Fourier Transform
GHz	Giga Hertz
GLCM	Grey Level Co-occurrence Matrix
HNN	Hopfield Neural Network
HSOM	Hierarchical Self-Organizing Map
IARC	International Agency for Research on Cancer
IARC	International Agency for Research on Cancer
K-NN	K-Nearest Neighbor
MLP	Multilayer Perception

MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MRMR	Minimum Redundancy Maximum Relevance
NBTF	National Brain Tumor Foundation
PD	Proton Density
PET	Positron Emission Tomography
PNG	Portable Network Graphics
RAM	Random Access Memory
TCIA	the Cancer Image Archive
WHO	World Health Organization

Chapter One: INTRODUCTION

1.1 Brain Tumor Detection System

A brain tumor is an abnormal growth of tissue in the brain or central spine that can disrupt proper brain function (society, 2014). Physicians refer to a tumor based on where the tumor cells arise, and whether they are cancerous *malignant* or not *benign*. All brain tumors can grow to damage areas of normal brain tissue if left untreated, which could lead to disability and possibly be fatal (NBTS, 2014). As the National Brain Tumor Foundation (NBTF) for research in the United States estimates that, in children, brain tumors are the cause of one-quarter of all cancer deaths. Also, NBTF reported most research in developed countries shows that the number of people who develop brain tumors and die from them has increased perhaps as much as 300% over past three decades (El-Dahshan, 2014). According to (Solomon Tessema Memirie, 2018) there are two way of diagnosing brain tumor, which are biopsy and scans and imaging techniques

Biopsy: - is a way of diagnosing a tumor using a surgical procedure by taking a small sample of tissue from the tumor so the cell can be examined under a microscope. There are two biopsy mechanisms that are: **Open Biopsy** where sample tissue is taken by opening the skull. And **Closed Biopsy** (stereotactic or needle biopsy): where a needle is used to access and remove a small selection of tumor tissue from an area that is difficult to reach. **Scans and imaging techniques**: - where different imaging modalities such as Computed Axial Tomography (CAT or CT) scan, Magnetic Resonance Imaging (MRI) scan and Positron Emission Tomography (PET) scan are used. Since this technique does not involve direct contact with the affected tissue, it is preferable by both the patient and the doctor.

Since there is a rapid growth in digital image processing and computer vision it is important to develop an improved image-based detection system that can reduce the risk of physical biopsy as well as manual tumor detection and classification.

Discoveries such as X-rays, radioactivity, magnetic resonance imaging (MRI) or computed tomography and the development of tools that can generate medical images have facilitated the development of some of the most efficient exploration tools in medicine. Such tools are capable of exploring the structure, density, and volume of the diseases that affect the human brain and deals with the cancer-affected region of the brain.

Compared to all other imaging techniques, (John, 2012) MRI is efficient in the application of brain tumor detection and identification, due to the high contrast of soft tissues, high spatial resolution and its safety in being radiation free. Therefore, it is safe to say that MRI imaging is Reliable for fast detection and classification of brain cancer. Detection of a brain tumor in digital image processing using machine learning involves various stages mainly image pre-processing for removal of noise from the image, Segmentation for selecting the region of interest from the whole image, Feature extraction, and classification.

The separation of tumors from magnetic resonance imaging (MRI) is one of the important applications of image segmentation. The manual process of doing the segmentation of brain MR images is a very time consuming and tedious task, and hence it is associated with many challenges such as the depth of knowledge that the radiologist possesses, the mood of the radiologist and the accuracy of the machine. For instance (Bettegowda C1, 2014) stats Diagnostic accuracy of radiologists was reduced significantly after a day of clinical reading, with average areas under the receiver operating characteristic from early to late reading. After a day of image interpretation, error in visual accommodation was greater, and subjective ratings of fatigue were higher. Therefore, we need the assistance of an automated segmentation method for medical imageries. This paper presents an improved segmentation method to extract brain tumors in MRI images.

The main contribution of this is developing tumor segmentation technique based on the symmetrical similarity of the left and right hemispheres of the brain to exactly pinpoint the tumor region. After selecting the region of Interest (ROI)/ tumor detection features are extracted using HOG feature extraction technique to classify the three tumor types and the normal brain. SVM algorithm was used, to classify the MRI image into four main classes, which are meningiomas, gliomas, pituitary, and normal brain.

1.2 Problem statement

A brain tumor is an abnormal growth of brain cells within the brain. Detection of brain tumor is a challenging problem, due to the complex structure of the brain. The incidence rate of all primary malignant and non-malignant brain and other CNS tumors is 23.03 cases per 100,000 for a total count of 392,982 incident tumors (Stats(CBTRUS), 2019). Various researches were carried out in identifying and classification of brain tumor especially in segmenting the area of interest, which in this case is the tumor region however; it is still an open area because of the lack of precision, computation resource and time required for implementing the system. Therefore, it is very important to find a better segmentation technique that can identify and segment the tumor region from brain MR Image.

In addition to the lack of precision in the segmentation technique, most systems classify the segmented image as tumors or non-tumors, which is a vague result because of the verity of brain tumor types that exists in the real world. Therefore, in addition to classifying the MRI image as tumors and non-tumors, it is important to identify the type of tumor it is, at least for those tumor types that do not require additional tests or biopsy for identification.

This provides better and more reliable results for the patients so that more patients can diagnosed and cured. In line with this, early identification of a brain tumor is very useful in encouraging a good chance of patient recovery. There is a need for an automated system in the recognition and classification of brain tumors so that the abuses during diagnosis and treatment reduces. Therefore, in this thesis we initiate a model for human brain tumor detection by employing a new segmentation technique, which can overcome the previous segmentation techniques problem in identifying the ROI by exploring the technology of image processing techniques. In addition to that, the proposed system tried to solve the problem of classifying the MR images as tumors and non-tumors by adding classes of the three most common brain tumor types to the classifier. The ultimate goal is to ease the Physicians role in the recognition of brain tumors and providing better and more reliable results so that more patients can diagnosed. To this end, this study answers the following research question:

- ✓ How to segment a tumor region from the MRI image?

- ✓ How to extract relevant features for building a model for brain tumor detection?
- ✓ How to classify the extracted features for building a model for brain tumor detection?
- ✓ To what extent recognition effectiveness is registered for the brain tumor.

1.3 Objective

1.3.1 General objective

The main purpose of this research is to develop a brain tumor detection and classification system using a new brain MRI image segmentation technique.

1.3.2 Specific objective

To achieve the general objective of this research we have followed these specific objectives

- ✓ Selecting the region of interest (ROI) using symmetrical side analysis and thresholding segmentation method.
- ✓ To determine the features of the segmented image by applying feature extraction Technique.
- ✓ To classify the tumor types by using a classification algorithm.
- ✓ To evaluate the recognition system effectiveness.

1.4 Scope of the study

In this thesis, we explore image-processing techniques on Human brain tumor Detection. The limitation is that since there are various types of brain tumor types, which are classified, based on the region that they occur. Patient history and laboratory report that involves getting sample from spinal fluid, And because performing such autopsy is very dangerous and requires special skill this system only focuses on classifying the image into three tumor types that can be distinguished by only using MRI imaging and normal brain. In addition, since the proposed segmentation method compares the left and right part of the brain to segment tumor region we use only the axial and coronal view of the MRI image.

1.5 Significance of the study

In most third world countries, there is a shortage of expertise in most fields of medicine that require specialized knowledge, radiology being among the top therefore there is a need of automated system that can assist the physicians in reading the images. In addition to that even though manual segmentation and analyzing of MR brain tumor images by radiologists is reliable, with no doubt it is tedious, time-consuming, highly subjective and impractical in today's medical imaging diagnosis where large numbers of images are taken for a single patient. As a result, there is strong demand to automate the tumor detection and segmentation process.

Therefore developing a computer-aided brain tumor detection system that can minimize the need of these experts and reducing radiologist's workload is crucial. We believe even though, this system does not perform at the level of the experts, it is a crucial step in creating one.

1.6 Thesis Organization

This thesis consists of five chapters including this chapter. The contents of the remaining four chapters is described as follows:

Chapter 2: In this chapter, a literature review on topics related to the objective of this research is presented. The areas that are directly related to this research are covered in detail. Brief review of the CAD system, Digital imaging techniques, and previous Brain tumor segmentation is done.

Chapter 3: In this chapter, the methodology that is used to implement the proposed segmentation technique and brain tumor detection system is discussed briefly.

Chapter 4: In this chapter, the implementation of brain tumor detection and classification system is presented and the result and findings of the implementation are discussed.

Chapter 5: Conclusion and recommendation of this thesis work is presented.

Chapter Two: LITERATURE REVIEW

2.1 Overview of brain tumor

Brain tumor, which is one of the most common brain diseases, has affected and devastated many lives. According to the International Agency for Research on Cancer (IARC), it is estimated more than 23,820 people diagnosed for brain tumor per year in the USA only with more than 17,760 mortality (Cancer.net, 2019). In (Solomon Tessema Memirie, 2018) statistics shows a low survival rate of brain tumor patients even though brain tumor disease has been the center of attention of thousands of researchers for many decades, around the world. In recent years, researchers from different disciplines ranging from medical to mathematical and computer sciences have combined their knowledge and efforts to better understand the disease and to find treatments that are more effective.

Cancer has become the second leading cause of death, behind cardiovascular disease, with more than 8.7 million attributable deaths worldwide in 2015 (Solomon Tessema Memirie, 2018). A brain tumor can be counted among the most deadly and intractable diseases. Tumors may be embedded in regions of the brain that are critical to orchestrating the body's vital functions, while they shed cells to invade other parts of the brain, forming more tumors too small to detect using conventional imaging techniques (Yadav, 2012).

Brain cancer's location and ability to spread quickly makes treatment with surgery or radiation is like fighting an enemy hiding out among minefields and caves. In recent years, the occurrence of brain tumors has been on the rise. Unfortunately, many of these tumors are detected too late, after symptoms appear. It is much easier and safer to remove a small tumor than a large one. About 60 percent of glioblastomas, which is one type of brain cancer, start as a lower-grade tumor. However, small tumors become big tumors. Low-grade gliomas become high-grade gliomas. Once symptoms appear, it is generally too late to treat the tumor. Computer-assisted surgical planning and advanced image-guided technology are in use in Neurosurgery to overcome such late detection (Nikita V. Chavan, 2015). Based on the depth of impact we can classify brain tumor into two benign and malignant

2.1.1 Bening tumor.

A benign tumor is the least aggressive type of brain tumor. It originate from cells within or surrounding the brain, do not contain cancer cells, grow slowly, and typically have clear borders that do not spread into other tissue. It may become quite large before causing any symptoms. Removing these tumors entirely will reduce the probability of its return. Still, it can cause significant neurological symptoms depending on their size, and distance to other structures in the brain. However, some benign tumors can progress to become malignant. Therefore, we call those tumors pre-malignant tumors (Weiss, 2013)

2.1.2 Malignant tumor

Malignant brain tumors contain cancer cells and often do not have clear borders. It is considered to be life threatening because it grow rapidly and invade surrounding brain tissue. Although malignant brain tumors very rarely spread to other areas of the body, it can spread throughout the brain or to the spine. One can treat these tumors with surgery, chemotherapy, and radiation, but it may recur after treatment. Based on the origin of the cancerous cell we can classify tumors into two as primary and metastasize (Weiss, 2013).

Whether cancerous or benign, tumors that start in cells of the brain called primary brain tumors. Primary brain tumors may spread to other parts of the brain or the spine, but rarely to other organs. Metastatic or secondary brain tumors begin in another part of the body and then spread to the brain. These tumors are more common than primary brain tumors and are named and treated based on where they originate, such as the lung, breast, colon or skin (Weiss, 2013).

Most medical institutions use the World Health Organization (WHO) classification system to identify brain tumors. The WHO classifies brain tumors by cell origin and how the cells behave, from the least aggressive (benign) to the most aggressive (malignant). Some tumor types are assigned a grade, ranging from Grade I (least malignant) to Grade IV (most malignant), which signifies the rate of growth (Tumor Types: Understanding Brain Tumors, 2019). Most of these tumors requires additional test after the MRI result. However, there are some tumor types that can be detected using MR imaging, most common brain tumor types that do not require further test and used in this work are:

- **Meningioma**: These tumors are technically not brain tumors, as they form in the meninges, which are the membranes that line the skull and vertebral canal. However, their growth may affect the brain by causing various disabilities such as vision and hearing impairment, memory loss, or even seizures. Incidents of meningioma increase with age and the tumors grow slowly, so symptoms could develop gradually over time. Many meningiomas are benign, so doctors may choose to leave asymptomatic cases alone. However, if the tumor starts adversely affecting the quality of life, physicians will either surgically remove it or treat it using radiation therapy. The most common early sign of meningioma is chronic headaches (Penn Medicine Neuroscience blog, 2019)

- **Glioma**: is a term that is used to describe a group of tumors that arise from the glial cells in the brain. These cells support the function of the other main brain cell type — the neuron. Gliomas usually happen in the cerebral hemispheres of the brain. These are the largest, the outermost part of the brain that controls many functions including movement, speech, thinking, and emotions. They can also affect the brain stem, the lower part of the brain that controls functions like breathing, blood pressure, and heartbeat, the optic nerves, and cerebellum, a part of the brain that deals with balance and other non-thinking functions. These tumors can be benign or malignant (Colombia University Department of Neurology, 2019).

- **Pituitary tumors:** are abnormal growths that develop in the pituitary gland. Some pituitary tumors result in too many of the hormones that regulate important functions of the body. Some pituitary tumors can cause the pituitary gland to produce lower levels of hormones. Most pituitary tumors are noncancerous (benign) growths (adenomas). Adenomas remain in the pituitary gland or surrounding tissues and do not spread to other parts of the body (MayoClinin Pituitary tumors, 2019).

2.2 Digital Image processing

An image may be defined as a two-dimensional function f(x, y), where x and y are spatial (plane) coordinates, and the amplitude of (*f)* at any pair of coordinate(x, y) is called the intensity of the gray level of the image at that point. When x, y, and the amplitude value of (*f)* are all finite, discrete quantity, we call the image digital image. These elements are referred to as picture elements, image elements, pels and pixels (Rafael & Paul, 1987).

The digital image processing refers to processing digital image by a means of digital computer. Unlike humans, who are limited to the visual band of the electromagnetic (EM) spectrum, imaging machines covers almost the entire spectrum ranging from gamma to radio waves. They can operate on the image generated by sources that humans are not accustomed to associating with images. These include magnetic resonance images, ultrasound, electron microscopy, and computer-generated images. Thus, digital image processing encompasses a wide and varied field of application.

Since it is difficult to clearly distinguish digital image processing from computer vision, whose main goal is to use computers to emulate human vision including learning and being able to make comparison and take action based on visual input. The authors use a paradigm that considers three types of computerized processes in this continuum which is low, mid and high-level processing where low-level processing is characterized by the fact that both input and output are images. It usually involves primitive operations such as noise removal, contrast enhancement, and image sharpening. Mid-level processing is characterized as the inputs are generally images but the outputs are attributes extracted from those images. It involves segmentation (partitioning an image into regions or objects) description of those objects to reduce them to a form suitable for computer processing. Finally, higher-level processing involves "making sense" of the segmented object. The far end of the continuum, performing the cognitive functions is normally associated with vision (Rafael & Paul, 1987).

Today, there is almost no area of technical endeavor that is not impacted in some way by digital image processing. The application area of image processing is so varied that some form of organization is important in attempting to capture the scope of the field. One of the simplest ways is to classify image processing is based on their principal energy source such as electromagnetic spectrum, ultrasonic and electronic. Images radiate from the electromagnetic spectrum are the most familiar. Especially images in X-ray and visual band of the spectrum.

Figure 2.1 shows Electromagnetic waves, can be conceptualized as the stream of massless particles each traveling in a wave-like pattern and moving at the speed of light. Each massless particle contains a certain amount of energy called a photon. The spectrums ranges from gamma ray (highest energy) at one end to radio waves (lowest energy)

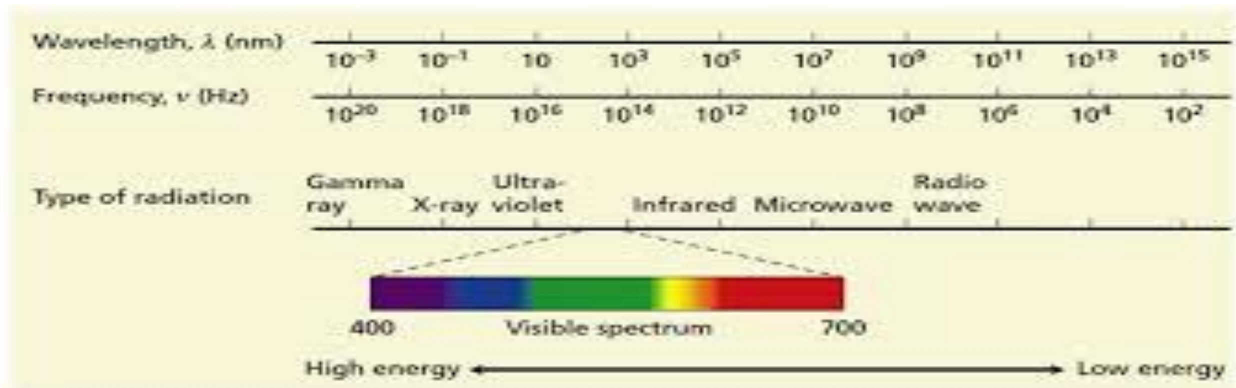

Figure 0.electromagnetic spectrum arranged according to energy per photon

Gamma ray: Used in nuclear medicine and astronomical observation. In nuclear medicine, the patient is injected with a radioactive isotope that emits gamma ray as it decays. Image of this sort is used to locate sites of bone scan obtained by the use of gamma-ray imaging. Major modality of nuclear imaging called positron emission tomography (PET scan).

X-ray: Are among the oldest source of EM radiation used for imaging. The best-known use of X-ray is medical diagnostics. X-rays for medicine and industry imaging are generated using an X-ray tube, which is a vacuum tube with cathode and anode. The cathode is the head causing free electrons to be released. These electrons flow at a high speed to the positively charged anode. When the electron strikes nucleus energy is released in the form of X-ray radiation. The intensity of the X-ray is modified by absorption as they pass through the patient, and the resulting energy failing on the film develops it.

Ultra-violate band: application of ultraviolet "light" includes lithography, industrial inspection, microscopy, lasers, biological imaging, and astronomical evaluation.

Radio band: the main application of imaging in this band are medicine and astronomy. In medicine, radio bands are used in magnetic resonance imaging (MRI). This technique places a patient in a very powerful magnet and passes radio wave through the patient body in short pulse. Each pulse causing a responding pulse of radio waves to be emitted by the patient's tissue. The location from which these signals are originated and strength is determined by a computer. Which produces a two-dimensional image. MRI can produce pictures in any plane

2.3 Medical imaging techniques

The conventional diagnosing method for a brain tumor is biopsy. It is a painful and time-consuming technique. However, by incorporating machine learning and Digital Image Processing, it is possible to do the diagnosis without any physical contact with the patient. This can be implemented using Computer-aided detection (CAD) system. The main idea of CAD is to assist radiologists in interpreting medical images by using dedicated computer systems to provide ‘second opinions'. Medical images used in this system are collected from various sources, such as hospitals, diagnostic centers and online archives that are stored in the database. (Kailash , Pradyumna , & Nagori , 2012). CAD systems and technology show that CAD can help to improve the diagnostic accuracy of radiologists, lighten the burden of increasing workload, reduce cancer missed due to fatigue, overlooked or data overload and improve inter and intra-reader variability (QURAT-UL-AIN, LATIF, SIDRA, M., & ANWAR, 2010)

Medical images nowadays play an essential role in the detection and diagnosis of numerous diseases. Ranging from anatomical information, functional activities, to the molecular and cellular expressions. Medical imaging provides direct visualization means to see through the human bodies and observe the minute anatomical changes and biological processes characterized by different physical and biological parameters (Rafael & Paul, 1987). As described in the digital image processing section there are various types of medical imaging acquisition technique where the most common are X-ray, CT scan, PET scan, MRI

X-ray: An X-ray uses electromagnetic radiation to create images. The image is recorded on a film, called a radiograph. The images appear light or dark, depending on the absorption rates of the various tissues. For example, dense materials, such as bone, show up white on a film, while fat and muscle may appear in varying shades of gray. It is a non-invasive procedure mostly used in chest area and bones for diagnosing disease, monitoring therapy and planning surgical treatment. X-rays may also be used in guiding the placement of medical devices, such as catheters and stents.

CT scan: A CT (Computed Tomography) scan is an X-ray procedure that uses a computer to produce 3-D cross-sectional images of inside the body. Unlike conventional X-rays, CT scans provide exceptionally detailed images of the bones, organs, and tissues. X-rays are taken from many angles and combined to create a cross-sectional image (cancercenter.com, 2019). During a CT scan, a patient rests on a table that slides into a large, tunnel-shaped scanner. Some exams require that a contrast dye be injected into a vein before the procedure. This helps certain areas show up better on the images (cancercenter.com, 2019). It also may be used to pinpoint the location of a tumor, evaluate the extent of cancer in the body and assess whether the disease is responding to treatment. In some cases, CT technology is used to accurately guide cancer treatment during a procedure (cancercenter.com, 2019)

PET scan: A positron emission tomography (PET) scan is an imaging test that allows doctors to check for diseases inside the body. The scan uses a special dye containing radioactive tracers. These tracers are either swallowed, inhaled, or injected into a vein in arm depending on what part of the body is being examined. Certain organs and tissues then absorb the tracer. The tracer will collect in areas of higher chemical activity, which is helpful because certain tissues of the body, and certain diseases, have a higher level of chemical activity. These areas of the disease will show up as bright spots on the PET scan. When detected by a PET scanner, the tracers help the doctor to see how well the organs and tissues are working. In addition, it can measure blood flow, oxygen use, how the body uses sugar, and much more. It is typically an outpatient procedure. This means patients can go about their day after the test is finished (Brian & Mariah, 2019).

MRI: Is an imaging tool designed to create detailed, cross-sectional pictures of the inside of the body. Using radiofrequency waves, powerful magnets and a computer, MRI systems may distinguish between normal and diseased tissue. MRI plays an important role in cancer diagnosis, staging and treatment planning. With MRI, we may distinguish between normal and diseased tissue to precisely pinpoint cancerous cells within the body. It also may be useful for revealing metastases. MRI provides greater contrast within the soft tissues of the body. As a result, it is often used for imaging the brain, spine, muscle, connective tissue and the inside of bones (Cancer treatment center for America, 2019).

Compared to all other imaging modalities, MRI provides superior contrast for different brain tissues and is efficient in the application of brain tumor detection and identification, due to the high contrast of soft tissues, high spatial resolution and its noninvasiveness. Additionally, MR images encapsulate valuable information regarding numerous tissue parameters (proton density (PD), spin-lattice (T1) and spin-spin (T2) relaxation times, flow velocity and chemical shift), which lead to more accurate brain tissue characterization. These unique advantages have characterized MRI as the method of choice in brain tumor studies (El-Dahshan, 2014).

2.4 Image Segmentation techniques

Many techniques have been proposed to automate the brain tumor detection and segmentation in recent years. According to (Megersa, 2012) the Segmentation methods can be broadly classified into two, intelligent based and non-intelligent based. Most notable intelligent based systems are artificial neural network, fuzzy c-means, support vector machine, and hybrid methods. On the other hand, most notable non-intelligent methods include thresholding and region growing. However, there is no clear demarcation between the two, especially intelligent based systems most often use the non-intelligent based ones as a refiner of their output. Some notable methods are reviewed in the following sections.

2.4.1 Thresholding Based Methods

A thresholding procedure attempts to determine an intensity value, called the threshold, which separates the desired classes. The segmentation is then achieved by grouping all pixels with intensity greater than the threshold into one class, and all other pixels into another class. Here are some of the thresholding methods used previously

(Nilesh, Arun, & Har, 2017) Have proposed a new technique in segmentation of the tumor from the brain image. The researchers had used adaptive contrast enhancement base on a modified sigmoid function to improve the signal to noise ratio then the study used a skull stripping technique that is based on a threshold operation to remove skull tissue. After the preprocessing is complete a threshold value of 128 was used and a pixel value above 128 is mapped to white and the rest is mapped to black. To eliminate white pixel an erosion operation of morphology was employed. Then the eroded region and the original image are both divided into two equal regions and the black pixel region extracted from the eroded operation is counted as a brain MR image mask. After the segmentation has completed a Gray Level Co-occurrence Matrix feature extraction method is used to extract features from the segmented image. While extracting the features some of the statistical feature formulas were used. Then finally, SVM is used to classify the MRI images. Even though the researchers have tried to develop a new thresholding mechanism, and produced a better result than the previous works, it still lacks a precision in identifying the ROI exactly. This is because of the thresholding technique that they used converts the whole pixel above 128 to white and the rest to black which might misplace a tumor that has a pixel value under 128. Which remain as a problematic area in brain tumor detection and analysis field using digital imaging processing technique.

(Salem, 2016) has proposed an automatic classification of brain tumor through MRI in which they use a median filter to remove high-frequency noise and high pass filter to remove low frequency noises from the image the propose a new way of segmenting an image using EM algorithm and adaptive thresholding . in which the EM algorithm is used to threshold the whole image using a certain threshold value then after that the adaptive threshold find the local threshold by statistically examining the intensity values of the local neighborhood of each pixel. The statistics, which is most appropriate, depends largely on the input image. After the ROI has been selected Fast Fourier transform (FFT) algorithm was used to extract features by converting the image from spatial domain to frequency domain then the researchers apply Minimal-Redundancy-Maximal-Relevance(MRMR) to select the feature finally the image is classified using SVM. The proposed segmentation method has used the EM algorithm in which the global thresholding value has been used to select the ROI to reduce the computation time of the adaptive thresholding algorithm that will also reduce the accuracy of the adaptive thresholding algorithm which affects the precision of the adaptive thresholding algorithm.

The proposed method by (Balakumar .B, 2014)uses Gaussian filter to improve the quality of the Image by Noise suppression as a preprocessing technique. Then segment the input images using a certain thresholding value that is as indicates most suitable for the present application. Gray Level Co-occurrence Matrix (GLCM) is used for feature extraction. And supervised machine learning algorithms (k-NN) is used to obtain the classification of images under two categories, either normal or a pathological brain. The overall recognition rate or classification accuracy is achieved up to 96.15%, which is more than existing methods. In the proposed system literature survey (Balakumar .B, 2014) has proposed a better way of image preprocessing which is Gaussian filter to improve the quality of the Image by Noise suppression as preprocessing technique. However, it still face the same problem in identifying the ROI when there is a tumor region that has a value less than the threshold it will be considered as a background.

2.4.1.1 Symmetrical analysis base thresholding method

(Pavel , 2013) This article focuses on the detection of a brain tumor location in magnetic resonance images. The aim of this work is not the precise segmentation of the tumor and its parts but only the detection of its approximate location. It is used recommend in future work for more accurate segmentation. For this reason, it also does not deal with detecting of the images containing the tumor. The algorithm expects a2D T2-weighted magnetic resonance image of brain containing a tumor. The detection is based on locating the area that breaks the left-right symmetry of the brain. The created algorithm was tested on 73images containing tumor, tumor with edema or only edema. These pathological structures had various sizes and shapes and were located in various parts of the brain.

(Swathi , Deepa , Vince , & Sankaranarayanan, 2015) In this paper, it is projected to encapsulate and evaluate the method of mechanical recognition of brain tumor through MRI with Histogram Thresholding and Artificial Neural Network. The anticipated method was used to distinguish the shape of the tumor. The appraisal of the adapted ANN classifier concert is deliberate in terms of the guidance performance, classification accuracies and computational time. MRI brain tumor metaphors detection is a difficult mission due to the inconsistency and Convolution of tumors. This paper present two techniques for the exposure purpose; first one is Histogram Thresholding and another one is Artificial Neural Network technique. Histogram is a conspire between number of pixel and pixel intensity. Bar graph was used to plot the histogram. The histogram code operates by first reading the grayscale value at the first entry and upcoming with pixel intensity between 0 and 255. It increases the whole number of pixels and then it will travel to the next row or column entry awaiting it finishes analyzing all the raster data. However, this technique is used to segment only the tumor types that has a higher gray-scale value.

2.4.2 Region Growing Based Methods

Region growing is a technique for extracting a region of the image that is connected based on some predefined criteria. In its simplest form, region growing requires a seed point that is manually selected by an operator and extracts all pixels connected to the initial seed with the same intensity value. To eliminate the dependency on initial seeds and to make the method automatic statistical information and a priori knowledge can be incorporated in the algorithm. Region growing can also be sensitive to noise, causing extracted regions to have holes, or even become disconnected. Conversely, overlapping gray value distribution in MR images can cause separate regions to become connected. Region growing is not often used alone because it is not sufficient to segment brain structures accurately and robustly.

In (Xie, Jie , Z.G. Zhang, & Y.M. Zhu, 2005)region growing is integrated with boundary information by a level set technique. Where region growing is used as a propagation force and boundary information is used as a stopping criterion. The method is applied on a total of axial tumors containing slices obtained from patients and satisfactory results were achieved. However, the method is semi-automatic as it relies on the manual input seed region for region growing.

2.4.3 Neural Networks Based Methods

An artificial neural network is a biologically inspired computational model which consists of processing elements (called neurons) and connections between them with coefficients (weights), and training and recall algorithms attached to the structure. Nowadays, several types of neural networks have been designed and used in medical image segmentation and other fields. The multilayer perceptron and backpropagation learning algorithm (MLP), Hopfield neural networks (HNN) and self-organizing maps (SOM) neural network are some of the algorithms, to mention a few.

In light of (El-Dahshan, 2014) the researchers have proposed a hybrid intelligent machine learning technique for computer-aided detection system for automatic detection of brain tumor through MRI. The proposed technique is based on the feedback pulse-coupled neural network for image segmentation, the discrete wavelet transform for features extraction, the principal component analysis for reducing the dimensionality of the wavelet coefficients, and the feed-forward back-propagation neural network to classify inputs into normal or abnormal. The experiments were carried out on 101 images consisting of 14 normal and 87 abnormal (malignant and benign tumors) from the brain MRI dataset. The classification accuracy on both training and test images was 99%. Even though the researches have a higher accuracy in classifying the images as tumors and non-tumorous. It is unable to identify the kind of tumor that may reduce the accuracy of classification.

One of the earliest applications of MLP to brain tumor segmentation is those of (Mehmed , Benoit, & Robert, 1993). These studies were based on a training procedure that initially uses a single slice from a specific patient. With a training data set generated from this first slice, a MLP model was constructed and was used to segment the adjacent slice of this patient image set. The labels obtained from the adjacent slice were used to generate a second training data set, which was then used to segment the adjacent slice. This process continued until the entire available image data set was segmented. The proposed technique is semi-automatic, needs continuous interaction with the user. Tumor segmentation accuracy was measured using Jaccard's similarity measure between areas delineated as a tumor by a human expert and the proposed automatic methods achieved similarity index that varies from to.

(Logeswari & Karnan, 2010)Used a hierarchical self-organizing map (HSOM), which is the extension of conventional SOM, to detect and visualize a brain tumor. Their system consists of two phases. In the first phase, the MRI brain image is preprocessed to remove artifacts then HSOM is applied for image segmentation. In this paper authors' highlighted the limitation of conventional SOM in which the number of neural units in the competitive layer needs to be approximately equal to the number of regions desired in the segmented image that is not possible to determine a priory the correct number of regions in the segmented image. The HSOM directly addresses the a foresaid shortcomings of the SOM. The performance of the system is measured in terms of weight vector, execution time and tumor pixels detected. However, the accuracy level of segmentation result is not given quantitatively and failed to distinguish the outer layer of the brain, which normally seen in MR brain images

2.4.4 Fuzzy Based Methods

Fuzzy logic is a set of mathematical principles for knowledge representation based on degrees of membership rather than on crisp membership of classical binary logic. In brain tumor segmentation, fuzzy systems permit the development of methods and algorithms that can perform tasks normally associated with intelligent human behavior.

In (Gordillo, Montseny, & Sobrevilla, 2010), expert knowledge and features derived from MR images are used to build fuzzy rules which are used to design fuzzy-based tumor detection and segmentation. The system is fully automatic and exhibits unsupervised learning. To achieve this knowledge extraction was performed by using intensity histogram analysis and a new method for obtaining membership functions to suit the MRI data is introduced. The detection and segmentation result obtained by this method is good, with the lowest score of 71% and the highest score of 93%, though the experiment is performed only on two types of brain tumors; glioblastoma and meningioma multiform. Another, very popular fuzzy-based image segmentation is fuzzy c-means (FCM) clustering algorithm, which was first suggested by Dunn (Dunn, 1973)and later improved by (Bezdek, Robert, & William, 1984). Several researches have been done on implementing FCM and its improved versions for segmenting brain tumor from brain MR images. Although the FCM algorithm is simple, fast and unsupervised, it cannot segment the tumor accurately because of the intensity overlapping of tissues and sensitivity of

FCM to noise and initialization values. Their method is not validated and only tested for one case.

2.5 Feature extracting techniques

The most of the image retrieval systems are based on the shape, color, texture and object layout (Chalechale , Mertins, & Naghdy , 2004). The shape of an object refers to its physical structure. Shape can be represented by the boundary, region, moment, etc. These representations can be used for matching shapes, recognizing objects, or for making measurements of shapes.

Generally, there are two types of shape representation and description techniques: contour based methods and region based methods. Both the methods are further divided into two types: global methods and structural methods, by how they use the contour or region information, whether they use the complete information or partition the contour/region information for representation and description. Global contour based shape descriptor techniques take the whole shape contour as the shape representation. Global region based method consider the whole region for shape representation, so, it effectively use all the pixel information within the region. Grid methods are not rotation invariant for region based shapes because the major axis is sensitive to the noise. Contour based methods are more popular than the region based methods, because human beings are thought to discriminate the shapes mainly by their contour features. In some shape applications the interior content is not important, only the shape contour is important, so, contour based techniques are widely used for that kind of applications. There are so many techniques developed for feature extraction, to make the shape based object recognition easier as well asaccurate. Some of those feature extraction techniques are as follows:

Edge Pixel Neighborhood Information (EPNI): In EPNI method the neighborhood edge pixels are found out that structure of those pixels will be used to make an extended feature vector (Chalechale , Mertins, & Naghdy , 2004). This feature vector is used for matching process for the image retrieval. This method is scale and translation invariant, but not rotation invariant.

Histograms of Edge Directions (HED): In computer vision and image retrieval process edge image matching is widely used for the comparison process. In images with the similar color information and in the absence of the color information this histogram of the edge directions is the significant tool for image retrieval. The edge is extracted using the Canny edge operator for this feature extraction and corresponding edge directions are, subsequently, quantized into 72 bins of 50 each (Mitisha & Purvi, 2016). HED is also useful for shape representation.

Angular radial partitioning (ARP): In ARP method, the images in the stored database are converted to grayscale and edge detection is performed. To achieve the scale invariance property, the edge image will be partitioned by the surrounding circles and the intersection points of the edge and surrounding circle are found and angles will be measured for feature extraction process which will be used for the comparison process in image retrieval process. The algorithm uses the surrounding circle of edge of an object and also make the n number of radial partition for that edge of the object image i.e. after making the surrounding circle equidistant circles will be made to extract the features for achieving scale invariance.

Histogram of Oriented Gradients (HOG): is a feature descriptor that is often used to extract features from image data. It is widely used in computer vision tasks for object detection. Histogram of the gradient is a feature descriptor, which is most commonly used in computer vision for object recognition. It counts the gradient orientation occurrence in confined area of an image. Histogram gradient technique is a well-known method used for the extraction of patterns. Patches of multiple scales of images is analyzed by the gradient histogram algorithm at many image locations. The patches will be cropped and further resized. The horizontal and vertical gradients are calculated by the algorithm. The absolute value of x-gradient, absolute value of y-gradient and finally the magnitude of gradient is calculated in order to extract the features (Seemanthini & Dr.Manjunath, 2018).

2.6 classification techniques

In machine learning and statistics, classification is a supervised learning approach in which the computer program learns from the data input given to it and then uses this learning to classify new observation. This data set may simply be bi-class or it may be multi-class too. Some examples of classification problems are: object identification, speech recognition, handwriting recognition, bio metric identification, document classification etc. some of the most common classification algorithms in Machine Learning are:

Naive Bayes Classifier (Generative Learning Model): It is a classification technique based on Bayes' Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability. Naive Bayes model is easy to build and particularly useful for very large data sets (Types of classification algorithms in Machine Learning, 2019).

Nearest Neighbor: The k-nearest-neighbors algorithm is a classification algorithm, and it is supervised: it takes a bunch of labelled points and uses them to learn how to label other points. To label a new point, it looks at the labelled points closest to that new point (those are its nearest neighbors), and has those neighbors vote, so whichever label the most of the neighbors have is the label for the new point (Types of classification algorithms in Machine Learning, 2019).

Decision Trees: Decision tree builds classification or regression models in the form of a tree structure. It breaks down a data set into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes. A decision node has two or more branches and a leaf node represents a classification or decision. The topmost decision node in a tree which corresponds to the best predictor called root node. Decision trees can handle both categorical and numerical data. (Types of classification algorithms in Machine Learning, 2019)

Neural Network: A neural network consists of units (neurons), arranged in layers, which convert an input vector into some output. Each unit takes an input, applies a (often nonlinear) function to it and then passes the output on to the next layer. Generally the networks are defined to be feed-forward: a unit feeds its output to all the units on the next layer, but there is no feedback to the previous layer. Weightings are applied to the signals passing from one unit to another, and it is these weightings which are tuned in the training phase to adapt a neural network to the particular problem at hand (Tumor Types: Understanding Brain Tumors, 2019).

Support Vector Machines (SVM): Vapnik proposed statistical learning theory based machine learning method which is known as Support vector machine (SVM) (Vapnik , 2004). SVM has considered as one of the highest prominent and convenient technique for solving problems related to classification of data and learning and prediction Support vectors are the data points that lie closest to the decision surface (Xiao , Peng , Wang , & Li , 2007)It executes the classification of data vectors by a hyper plane in immense dimensional space. SVM helps to determine the most simple classification problem of linear separable training data with binary classification. The maximal margin classifier used to find the hyper plane with maximal margin in real world complications. The main advantage of SVM is its capability to deal with wide variety of classification problems includes high dimensional and not linearly separable problems.

2.7 Related work

(Nilesh, Arun, & Har, 2017)Have proposed a new technique in segmenting the tumor from the brain image. The researchers had used adaptive contrast enhancement base on a modified sigmoid function to improve the signal to noise ratio then the study used a skull stripping technique that is based on a threshold operation to remove skull tissue. After the preprocessing is complete a threshold value of 128 was used and a pixel value above 128 is mapped to white and the rest is mapped to black. To eliminate white pixel an erosion operation of morphology was employed. Then the eroded region and the original image are both divided into two equal regions and the black pixel region extracted from the eroded operation is counted as a brain MR image mask. After the segmentation has completed a Gray Level Co-occurrence Matrix feature extraction method is used to extract features from the segmented image. While extracting the features some of the statistical feature formulas were used. Then finally, SVM was used to classify the MRI images. Even though the researchers have tried to develop a new thresholding mechanism for achieving a better result in tumor segmentation and produced a better result than the previous works it still lacks a precision in identifying the ROI. Because of the thresholding technique that they used converts the whole pixel above 128 to white and the rest to black which might misplace a tumor that has a pixel value under 128 which still remain as a problematic area in brain tumor detection and analysis field using digital imaging processing technique.

(B.Devkota, Abeer, P.W.C.Prasada, A.K.Sing, & EA.Elchouemi, 2018) Has introduced a new segmentation of the region of interest by applying a mathematical morphological operation. The researchers first preprocess the image using median filter then strip the skull. Then to segment normalized global threshold was found and brain MRI image is converted from grey-scale to binary. Then the resultant image is opened by reconstruction. This process involves identifying an appropriate structuring element; determine the size of the structuring element to obtain an eroded image. Then the eroded image is morphologically closed. Finally, the reconstructed and morphologically closed image is multiplied to obtain a segmented image. After the segmentation part is completed they used a principal component analysis to extract features from the image and used SVM to classify it. Even though the proposed system has introduced a new segmentation technique they still face the same problem while thresholding the image and converting it to a binary image in which the tumor couldn’t be identified if it has a low-intensity value.

(Miss. Priyanka Katti, 2015)Used region growth method to segment brain MRI images. This method includes several steps. At first, some initial points (grains) are selected automatically. These points pertain to regions that need to be separated from the background. Then, starting from these points and considering neighboring points, other points are checked. If points belong to the first region, based on selected similarity criterion, then they are added to that region. Genetic algorithm was exploited in which after selecting the initial population and defining appropriate fitness function, optimal initial points are searched. Standard deviation criterion was used to check similarity. Finally, the proposed method was applied to the MRI image. Comparing the results of the proposed method and the result of region growth method with manual selection has improved brain MRI image segmentation. The accuracy of the segmentation technique proposed in this research depends on the similarity criterion which might vary and affect the precision of the segmentation in addition to that comparison and checking each point require high computational resource and run time.

(Valentina , et al., 2012) In this work a fully automatic algorithm to detect brain tumors by using symmetry analysis is proposed. The researchers starts their work on the assumption that, an healthy brain has a strong sagittal symmetry, that is weakened by the presence of tumor. The comparison between the healthy and ill hemisphere, considering that tumors are generally not symmetrically placed in both hemispheres, was used to detect the anomaly. A clustering method based on energy minimization through Graph-Cut is applied on the volume computed as a difference between the left hemisphere and the right hemisphere mirrored across the symmetry plane. And, through a histogram analysis the ill hemisphere is recognized. even though, in this paper a symmetrical analysis is performed the researchers used a histogram analysis to detect the ill hemisphere of the brain which can only show whether there is tumor or not, and lacks the ability to exactly pin point the tumor part so that the size, shape, and location of the tumor be determine in which we use those information’s to determine the class of the tumor. In addition to the lack of the ability to determine the shape and size of the tumor the research was carried for only glial tumor.

2.8 Summary

This study emphasizes on detection and classification of some brain tumor types from a static image using image processing techniques, which involves preprocessing, symmetrical side analysis and tumor detection of specific tumor type classification. As it is tried to review different papers thoroughly, in the area of Brain tumor detection many works are done by different researchers and since the technique and dataset used in each study are different comparing them is a challenge, and also still identification of tumor type is a challenging task due to the difference exist in the size, location and shape of the tumor.

Based on the reviewed literature it is able to understand the limitation and strength of detection and classification of brain tumor. To successfully detect tumor, most of the work done in this area relies on the shape of the tumor. However, it is also important to determine the location and size of the tumor to clearly detect the tumor and then classify its type. So to solve this problem, shape, size and location is important. Combining size, location and shape features it is possible to reduce the number of regions that could correspond to tumor. Therefore, we used this three property of the image to properly classify the tumor type.

For the classification phase to have a better performance as recommended by different researchers Support Vector Machines (SVM), which is trained using Histogram of Oriented Gradient (HOG) features is used. Because SVM classification is, fast, highly accurate, and which has a very low computational time for real-time application compared to other classification methods including neural network (Mirmehdi, 2012), (Ayoub ELLAHYANI, 2016). Also, the feature extraction technique HOG is very simple, fast and so suitable for classification of brain tumor because tumors are composed of strong geometric shapes and high-contrast edges, (Praveena & Rama , 2018)

Chapter Three: METHODOLOGY AND DESIGN OF BRAIN TUMOR DETECTION

3.1 Overview of Brain tumor Detection

The task of classification occurs in a wide range of human activities. The problem of classification is concerned with the construction of a procedure that are applied to differentiate items, in which each new item must be assigned to one of a set of predefined classes based on the observed attributes or features. Accordingly, image analysis or computer vision is used in the recognition of brain tumors to predefined classes. The predefined classes are the feature or attributes computed from brain images. These observed features of brain image were used to decide the class or the type of Brain tumor.

Figure 0.1 logical view of tumor recognition

Then, the final output process gives the type of brain tumor if there is, of the given MR images. Hence, in this research, the main interest is to differentiate the type of brain tumor by using image analysis techniques to maximize the curability of the disease. As (El-Dahshan, 2014) Stated that many different techniques are used for developing a CAD scheme and various techniques have been summarized in several review papers. Generally, CAD systems are executable on all imaging modalities and all kinds of examinations. To create a CAD system, the integration of various image-processing operations such as image preprocessing, image segmentation, feature extraction and selection, and classification are essential.

Our system workflow is divided into three different phases, which are preprocessing, detection, and classification and *figure 3.3* shows each phase of learning and testing the system. In the learning phase, various Brain MRI images are used to train the classifier. The detection phase uses region and shape-based methods to detect potential class, after detecting ROI features are extracted from these images to create a master feature vector using HOG feature extraction, and this master feature vector is then fed to the SVM classifier as an input to construct the classification model. In addition, the detail process of each phase is discussed in more detail in subsequent sections.

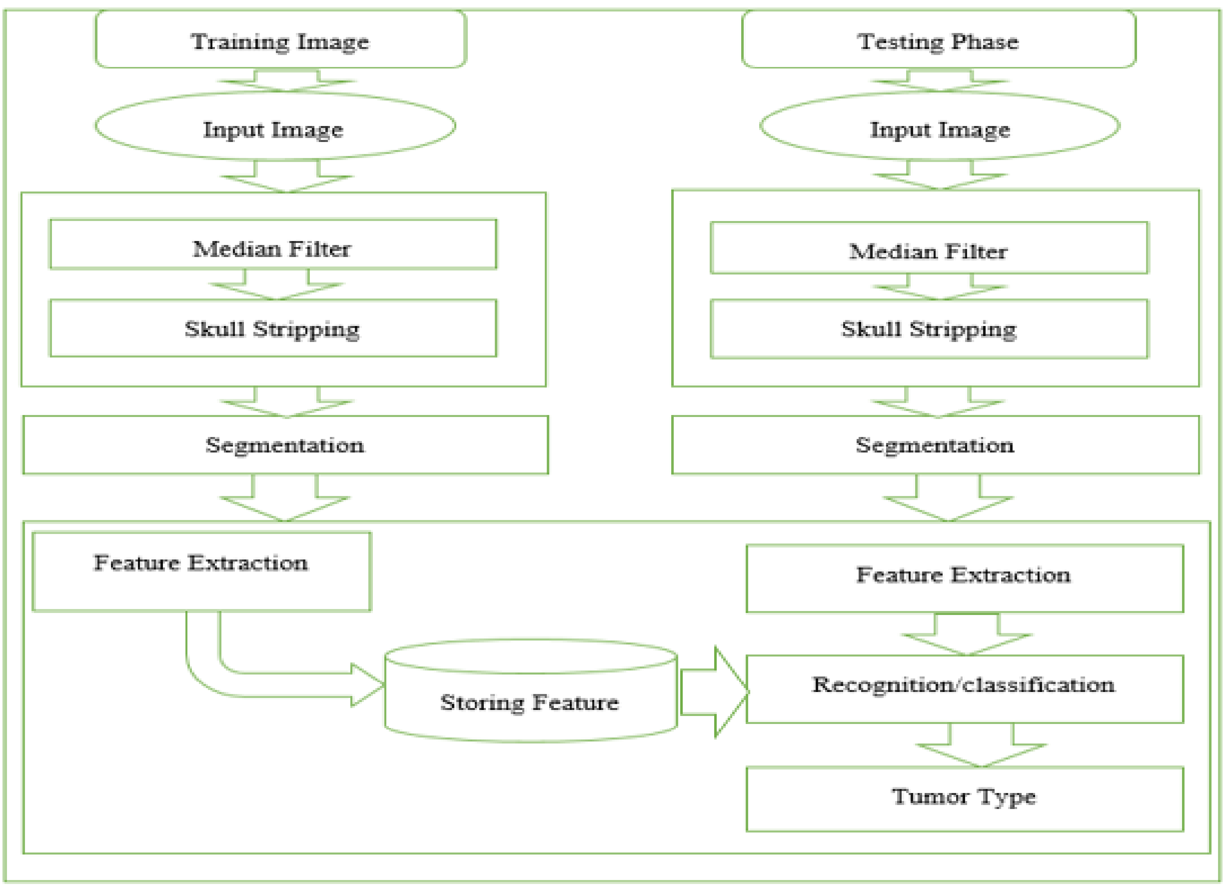

Figure 0.2 System Architecture

As showed in *figure 3.3,* the general steps in creating a CAD system for human brain tumor detection involves five major steps where in the first step an MRI image is going to be collected.

Then the collected image is preprocessed to remove the skull part of the brain and the noises that might affect the classification process. In the third step, the tumor part of the image will be segmented from the collected and preprocessed image. after the tumor region is segmented different features of the image are extracted in ordered to proceed to the next step which is classification this step involves accepting the extracted feature, learn the main object property and then classify an input in to a certain group.

3.2 Image acquisition/collection:

Image acquisition is the starting point of image processing. Medical images are pictures of distributions of physical attributes stored by an image acquisition system. To obtain a real brain MR images and use in a research is a very complex because of a privacy issue in patient information disclosure. However as (Kalpana & Prof. Kapse, 2016) stated the state-of-the-art of now-a day's technologies of digital medical image acquisition are improving tremendously, which gives images of high quality and resolution while the noise on the images is still an issue.

We mainly obtained the three types of tumorous images from Fig share (http://dx.doi.org/10.6084/m9.figshare.1512427). Which is brain tumor archive where there are images of meningiomas, gliomas and pituitary tumor in (.mat) file format. In addition, the normal brain MRI images ware collected in (.JPEG) file format from kaggle.com (https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection) these images are first converted to (.png) format for performing the image processing operations. After the image has been converted to (.png) format, different image processing operations carried out on MATLAB 2018a on Windows 8 operating system. The prepared image dataset contains four different classes of Brain MRI images, 3 of them are tumorous, and one is a normal brain MRI. Namely, those tumorous MRI images are meningiomas tumor type, gliomas tumor type, pituitary tumor type. Based on the collected and converted data we have prepared the data set. We have categorized the data into two main groups the first one is training data and the second one is used to test the trained system therefor we have used 218 images for training classes of the system. And on the second group, we have used 63 images for testing the system.

Table 0.1 Dataset used in the system

	Brain MRI image	Training data	Testing Data	Total
1	Meningiomas	60	20	80
2	gliomas,	60	21	81
3	Pituitary	47	12	59
4	Normal brain mri	51	10	61
Total				281

3.3 Used Tools

For this study, MATLAB version 2018a is used as the main development tool, with Image processing, Computer Vision, Statistics and Machine Learning Toolboxes. In addition, specifically, we choose MATLAB because it is more powerful and reach of libraries for image processing and the following more qualities listed below:

- A very large (and growing) database of built-in algorithms for image processing and computer vision applications
- Clearly written functions with many examples, as well as online resources such as web seminars
- allows you to work interactively with your data, helps you to keep track of files and variables, and simplifies common programming/debugging tasks
- A large user community with lots of knowledge sharing

3.4 Image preprocessing:

Image Preprocessing is a process that is used to boost the precision and interpretability of an image. Image preprocessing is a noteworthy and challenging task in the CAD system. In medical image processing, preprocessing of an image is very important so that the extracted image does not have any impurities, and it is accomplished to be better for the forthcoming process such as segmentation, feature extraction, etc. Only the correct segmentation of the tumor will yield an accurate result. Accurate detection leads to precise feature extraction and in turn, gives perfect classification. The accurate tumor segmentation is promising, only if the image is preprocessed clearly. To clean this noise we have to preprocess it. In this stage of medical image analysis, image is enhanced in the way that finer details are improved and noise is removed from the image. Therefore, first we have to remove the skull part of the brain from the image to select the brain part alone for further segmentation and other image processing phases. After that even though there are several image de-noising approaches that can be used, we select median filter to clear the noise. Median filter is the most common de-noising method used to reduce paper and salt noise and improve the quality in an image while preserves the edges of the images (Kalpana & Prof. Kapse, 2016).

3.4.1 Noise removal

Medical image analysis requires the pre-processing of the image, because, the noise may be added to the MR images due to imaging devices malfunction, signal interference, and patient maladjustment. To clean this noise we have to preprocess it. In this stage of medical image analysis image is enhanced in the way that finer details are improved and noise is removed from the image. Even though there are several image de-noising approaches that can be used, we select median filter to clear the noise because, median filter is most common de-noising method used to reduce paper and salt noise and improve the quality in an image while preserves the edges of the images (Kalpana & Prof. Kapse, 2016).

Once, the skull stripping is done it is necessary to use a filtering algorithm to remove noise from the image, as a result, it enhances the classification process. In addition, in this study, median filter is used to smooth the image and to fill up the unexpected area. Filtering is a neighborhood operation, in which the value of any given pixel in the output image is determined by applying some algorithm to the values of the pixels in the neighborhood of the corresponding input pixel, and median filter is a nonlinear method, which is used to remove noise like salt and pepper noises from images.

Salt and pepper noises are dark pixels in bright regions and bright pixels in dark regions caused due to analog-to-digital converter errors, bit errors in transmission, etc. and median filter has the power of removing noise very effectively while preserving edges and is highly effective in removing salt-and pepper noise. This is one of its most powerful features and this is why median filter is frequently chosen in image segmentation. (Jyoti, Kumar , & Ahmad , 2015)

The working methodology of this filter is: it moves through the image pixel by pixel replacing each value by median value of the neighboring pixels. The pattern of the neighbors is called the window, which slides, pixel by pixel, over the entire image. The median is then calculated by first sorting all the pixel values from the window into numerical order. At last, replace the considering pixel with the median pixel value (Fanuel, 2018).

The median filter works by moving through the image pixel by pixel, replacing each value with the median value of neighboring pixels. And the pattern of neighbors is called the "window", which slides, pixel by pixel over the entire image, and the median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the middle (median) pixel value.

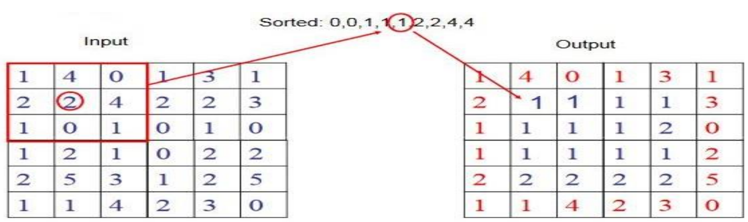

Figure 0.3 Median filter

3.4.2 Skull Stripping of MR Brain Images

The Extraction of brain tissue from non-brain tissues in MR images, which is referred to as skull stripping, is an important step in many neuroimaging studies. According to (Kalavathi & Surya Prasath, 2016) Skull stripping algorithms can generally be categorized into five types: mathematical morphology-based methods, intensity-based methods, deformable surface-based methods, atlas based methods, and hybrid methods.

As indicated in (Megersa, 2012)skull stripping that uses intensity thresholding followed by morphological operations to remove narrow connections is the most common.

In this thesis work, we used an automatic threshold value selection using Otsu's algorithm to automatically select threshold value. Then, mathematical morphology operations on a binarised image are applied stage by stage to achieve acceptable skull stripped brain images. The skull stripping method comprises four steps as shown in Figure 3-4. Initially, image binarisation is performed using a threshold value obtained from Otsu's threshold selection algorithm then, largest connected component from the binarised image is selected by considering the fact that the brain is the largest connected structure inside the head. Thirdly, mathematical morphology operations such as filling holes and dilation are carried out on the selected largest binarised image. Finally, we obtained the skull stripped brain image using mathematical logical operations on binarised and original images. The above steps are described in the following figure and sections after that in detail.

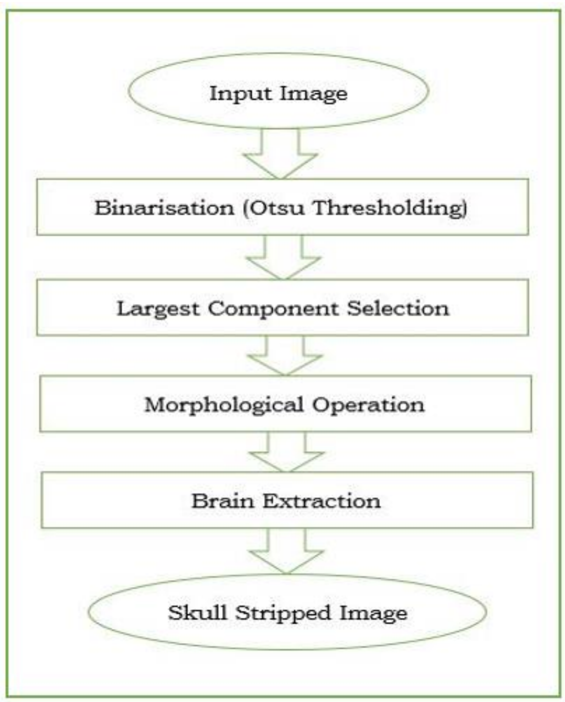

Figure 0.4 skull striping framework

3.4.2.1 Binarization

Image binarization is the simplest method of image segmentation that transforms an image gray level into only two values. By selecting optimum threshold value it is possible to separate brain MR image into two classes, one class is background: formed by the very low-intensity pixels corresponding to air, part of cerebrospinal fluid; and the other class is composed of foreground tissues including the gray matter (GM) and White matter (WM) of the brain.

The main issue in image binarization is selecting an optimum threshold value, in an image with a maximum gray level value. In this paper, we used Otsu's algorithm (Otsu, 1979)to automatically select the threshold value.

3.4.2.2 Largest Connected Component Selection

Binarization on brain MR images classifies the image into background and foreground, leaving the foreground into several connected components. By assuming that the brain is the largest connected structure inside the head, we select the component whose area is the biggest.

3.4.2.3 Morphological Operations

Mathematical morphology is a theory of image transformations, which is based on set-theoretical, geometrical, and topological concepts. Morphological operations implemented in this work include dilation and filling. The two most common morphological operations, which are used as a base for all other morphological operations are dilation and erosion (Rafael & Paul, 1987).

Dilation operation makes the areas of foreground pixels grow in size while reducing the holes within those regions. The final morphological operation performed in this work is the fill operation that is used to fill remaining holes in the binary image.

3.4.2.4 Brain Extracting

The earlier steps enable us to obtain the binary mask for head MRI scan then the brain is extracted by performing bitwise AND operation between the original head MRI scans and the masked one.

3.5 Segmentation:

Several image-processing methods are required before the brain images can be explored. Image processing covers various techniques that are applicable to a wide range of applications, among which segmentation is an essential and important process in medical image processing and analysis (Prasath, 2015).

1. In Image segmentation process, an image is partitioned into multiple segments i.e. sets of pixels also known as super pixels. Segmentation subdivides an image into its constituent regions or objects. The level to which the subdivision takes place depends on the problem being solved i.e. segmentation should not stop until the objects of interest in an image have been isolated. Segmentation aim is to simplify or change the representation of an image so that it can become more meaningful and easier to analyze. In other words, segmentation can be described as a process of separation of suspicious regions from the background image. Image segmentation algorithms are generally based on one of two basic properties of intensity values: Discontinuity and Similarity

Image segmentation is required to delineate the boundaries of the region of interest (ROIs) ensuring, in our case, that tumors are outlined and labeled consistently across subjects. Segmentation can be performed manually, automatically, or semi-automatically. The manual method is time-consuming and its accuracy highly depends on the domain knowledge of the operator. Specifically, various approaches have been proposed to deal with the task of segmenting brain tumors in MR images.

3.5.1 Proposed thresholding Technique

At present, threshold-based methods are classified into global and local thresholding (Kazi, Chowhan, & Kulkarni, 2017). If an image contains objects with homogeneous intensity or the contrast between the objects and the background is high, global thresholding is the best option to segment the objects and the backgrounds. When the contrast of an image is low, threshold selection will become difficult. Therefore, to overcome the difficulty Local thresholding can be determined by estimating a threshold value for the different regions from the intensity histogram (Jin , Min , Jianxin , Fangxiang , & Tianm, 2014). As shown in figure 3.6 the proposed segmentation technique used automatic threshold value selection using symmetrical side gray-level intensity value to automatically choose threshold value. To achieve this first the input image is split into two equal parts (left and right). Then the second image is flipped to match the size of the first image. After that, we compare the two images and threshold those by using a local thresholding mechanism based on the intensity value difference of the images. After segmentation is complete, we then flip back the second image to its original shape and merge it with its other half.

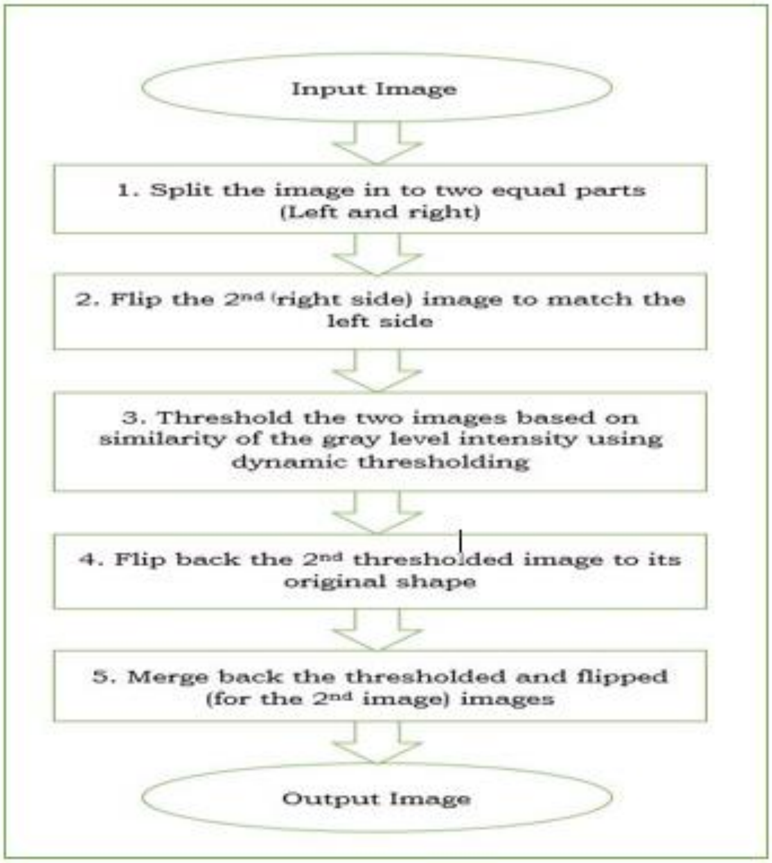

Figure 0.5 work flow of the proposed segmentation method

3.5.2 Global thresholding technique

Global Thresholding (Static thresholding) is simpler and straightforward approach. A pre-determined selected threshold value is used for segmentation of an image. It is effective when the background conditions in which image is captured are well known and they do not change. It is the process of converting images into binary images using specific constant thresholding value, and in this study, global thresholding is used for segmenting images with a value of 128. We use this thresholding technique for comparing the efficiency of the result with the proposed technique

3.5.3 Sobel edge detection

The Sobel method of edge detection for image segmentation finds edges using the Sobel approximation to the derivative. It precedes the edges at those points where the gradient is highest. The Sobel technique performs a 2-D spatial gradient quantity on an image and so highlights regions of high spatial frequency that correspond to edges. In general, it is used to find the estimated absolute gradient magnitude at each point in n input grayscale image. In conjecture at least the operator consists of a pair of 3x3 complication kernels as given away in under Figure.

-1	-2	-1
0	0	0
+1	+2	+1

G_x

-1	0	-1
-2	0	+2
-1	0	+1

G_y

Figure 0.6 Sobel convolution kernels

3.6 Feature extraction

Features are said to be properties that describe the whole image. The purpose of feature extraction is to reduce the original dataset by measuring certain features. In medical images, feature extraction has to focus on specific regions and capture not only shape but also structural and internal volume properties that can be useful for building a classification model.

Feature extraction is a method by which unique features of a specific image are extracted and represented in a compact feature vector, and this process is essential step of image to obtain fast speed of detection because it reduces the original data set by measuring certain properties or features, which distinguish one input pattern from another. In image processing technique there are many different ways of extracting a feature from an image, and for this study, HOG feature extraction is used. HOG feature extraction is the most common and popular feature extraction technique in CAD systems, as many different literatures tells us a top performance can be reached by using a HOG features, and here are some additional reasons for choosing HOG feature extraction:

- HOG is very simple and fast
- Work well for small resolutions
- Invariant to small shifts In this stage

once ROI is detected and confirmed, features such as size, location, and shape features are extracted from the image using the HOG feature extraction technique, and the extracted features from the training data set are fed to the classifier algorithm as an input, to be used as classification model during testing.

3.7 Classification

Image classification is a fundamental part in pattern recognition. (Aycheh, 2008)Pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest, make sound and reasonable decisions about the categories of the patterns. Patterns are any entity or object.

In classification, the objective is to categorize objects in the scene from a set of measurements of the objects. The measured values are the features of the pattern. A set of similar objects or patterns possessing more or less identical features are said to belong to a certain category called classes.

The classification step involves two major phases the first one is training and the second one is the testing phase. Once the region of interest (ROI) is determined and a tumor is detected using the different techniques discussed above, it should be recognized using a predetermined database of all the tumor types in the system. The above-mentioned feature extraction process outcome is a numeric feature vector. And this vector is fed to the classification algorithm as an input. So by taking those input features, this recognition phase is responsible to classify four different types of brain MRI images to their corresponding classes.

As it is discussed in the previous sections this system is limited only to 4 different classes. So for the classification purpose machine learning technique SVM (Support Vector Machines) is selected and used with the HOG feature extraction technique. SVM is a useful machine learning technique for classification, and it is considered as easier and faster with less computational time than Neural Networks. In addition, additionally, it has also many qualities, which makes it more preferable than the others, just to mention SVM is highly accurate, and fast which has a very low computational time for real-time application and in addition to this it is less prone to overfitting compared to other many classification methods.

SVM is extremely fast-supervised machine learning algorithm for solving multiclass classification problems. It constructs a hyperplane to separate data into classes. It learns from training data, where each data instance has n data points followed by a class of the instance. In SVM, classes are separated by applying an optimum hyper plane by decreasing the distance between classes. These hyper planes are known as support vector, where each side of the vector contains the instances of different classes (site). Therefore, the recognition is done using this technique for the input image and the result is displayed in the consule.

3.8 Performance measurement parameters

The comparison of detection classes is done based on detection performance parameters. The performance measurement parameters are computed from the confusion matrix. This study model performance is measured using the following performance measurement parameters. Which are correctly classified instances, incorrectly classified instances, precision, recall, and F-measure.

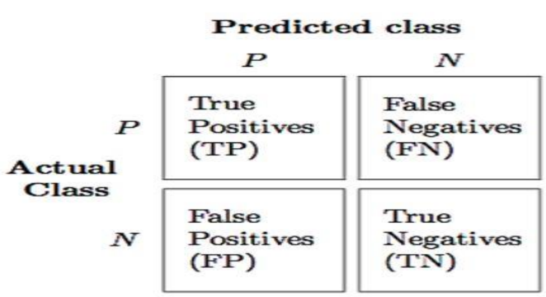

Figure 0.7confusion matrix

False Positive Rate (FPR): Tumor is not detected correctly

False Negative Rate (FNR): Tumor is detected as a non-Tumor region.

True Positive Rate (TPR): Tumor is correctly detected.

True Negative Rate (TNR): A non-tumor region is correctly recognized as a non-Tumor region.

According to (Zelalem, 2017)each performance measure parameters are computed as follows:

Precision (P) (positive predictive value)-proportion of predicted positives which are actual positive.

$$\text{Precision (P)} = TP / (TP+FP)$$

True positive rate (TPR) is the proportion of datasets that were classified as a class, among all datasets which truly have class, i.e. how much part of the class was achieved which is equal to Recall. In the confusion matrix, this is the diagonal element divided by the sum over the relevant row. Recall (true positive rate) is the proportion of actual positives, which are predicted positive so-called sensitivity.

$$\text{Recall (R)} = TP / (TP + FN)$$

F-Measure is the harmonic mean between precision and recall.

$$\text{F Measure (FM)} = 2 * P * R / (P + R)$$

Generally, the total detection accuracy is measured in terms of a true positive, true negative, false positive and false negative.

Chapter Four: EXPERMENT, RESULT AND DISCUSSION

4.1 Introduction

This chapter describes the implementation of the brain tumor detection system, which was specified in detail in the previous chapter. As described in section 3.1, the detection of brain tumor has six components. They were image acquisition, image preprocessing, image segmentation, feature extraction, classification, and testing.

4.2 Description of dataset

As described in the previous chapters to implement the brain tumor detection system there must be a dataset in which the proposed system steps are going to implement. And as described in section 3.1 the medical images are collected from figshare and kaggle image archive.

Based on the collected and converted data we have prepared the data set. We have categorized the data into two main groups the first one is training data and the second one is used to test the trained system therefor we have used an average of 54 images for training each classes of the system. In addition, on the second group, we have used an average of 15 images for each class to test the system.

Table 0.1 Number of images used in the system

	Brain MRI image	Training data	Testing Data	Total
1	Meningiomas	60	20	80
2	gliomas,	60	21	81
3	Pituitary	47	12	59
4	Normal brain MRI	51	10	61
Total				281

Data used in this thesis are T1-weighted and T2-weighted head MRI scans and experiments have been done on the axial and coronal views of the MRI scans.

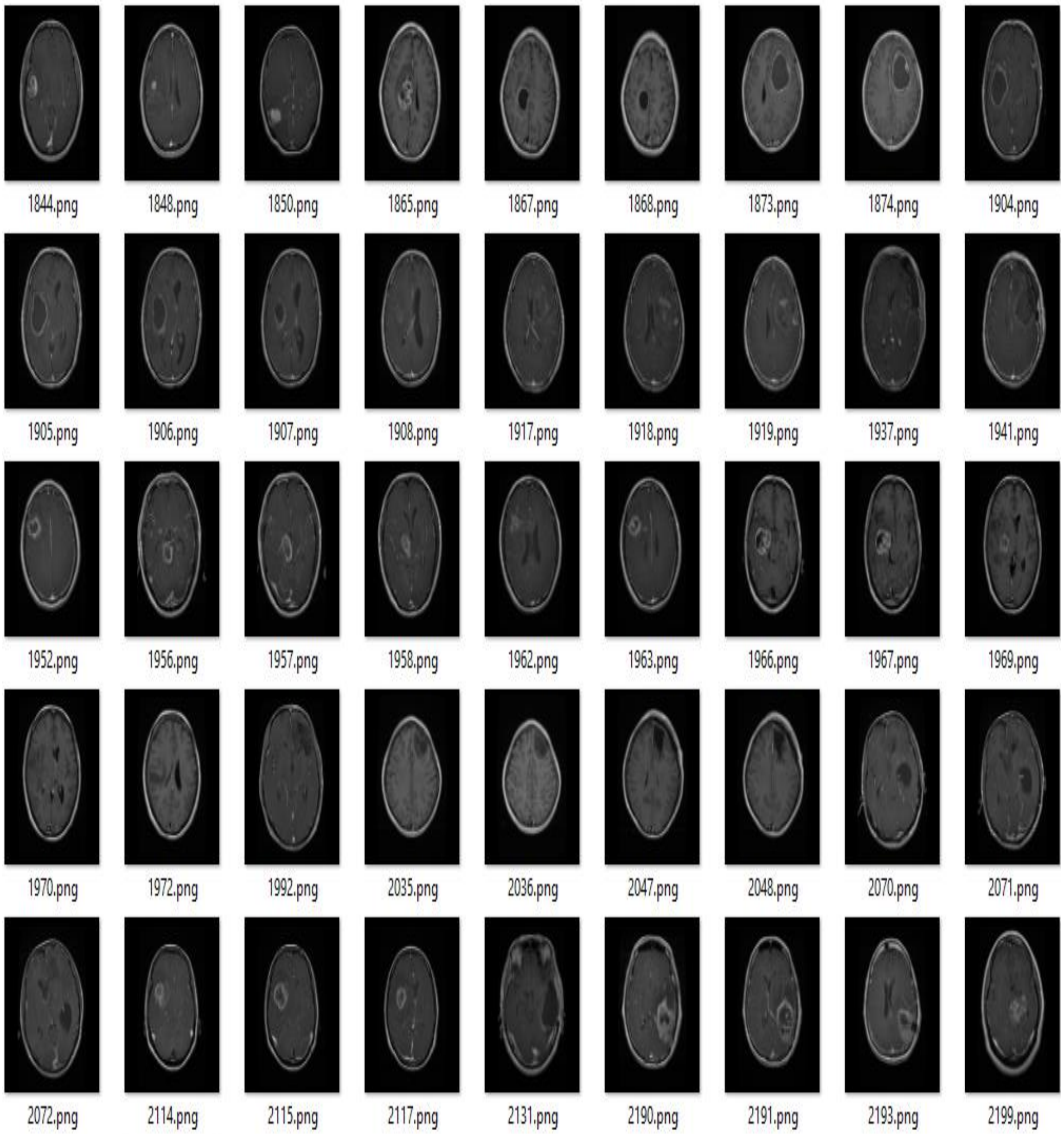

Figure 0.1 brain tumor images

And as it is shown in the above table the total dataset is composed of 281 images, and from the total collected images in each class, 75% of it is used for training and the remaining 25% for testing purpose.

4.3 Experimental steps

In this study different necessary image processing techniques are integrated and used to develop the system. In addition, to implement these Image processing techniques a software with the capability of Computer Vision, Statistics and Machine Learning Toolboxes is required. Therefore, we have used MATLAB R2018a to implement the system.

To run MATLAB a computer with a resource like high memory and powerful processing speed is required. Therefore we have used HP laptop with a specification of 6 GB RAM, 4^{th} gen 1.6 GHZ core I3 Intel processor with 1 TB hard disk.

4.4 Implementation phases

In this part, we have tried to implement the system that we have proposed in the previous section so that we can evaluate the result and make a suggestion for future research endeavors.

To implement the whole system we have tried to break down the system into smaller unite (phases) that have been described in the previous chapters. The reason for dividing the system into smaller steps is that we can see each step performance in detail and to control unwanted change and alteration of the system.

4.4.1 Image preprocessing

4.4.1.1 Median Filtering

As described in section 3.4.2 MRI image is subjected to various noises. These noises can cause inaccuracy in classification. Therefore, to avoid that, images are subjected to various image-processing techniques. One of the key element in image processing is the filtering of an image. Preprocessing is done to removes the noise, and improve the edge in the image. For smoothing image from noise, median filtering is used. Median filtering is a common step in image processing. Median filtering is used for minimizing the influence of small structures caused due to signal interference.

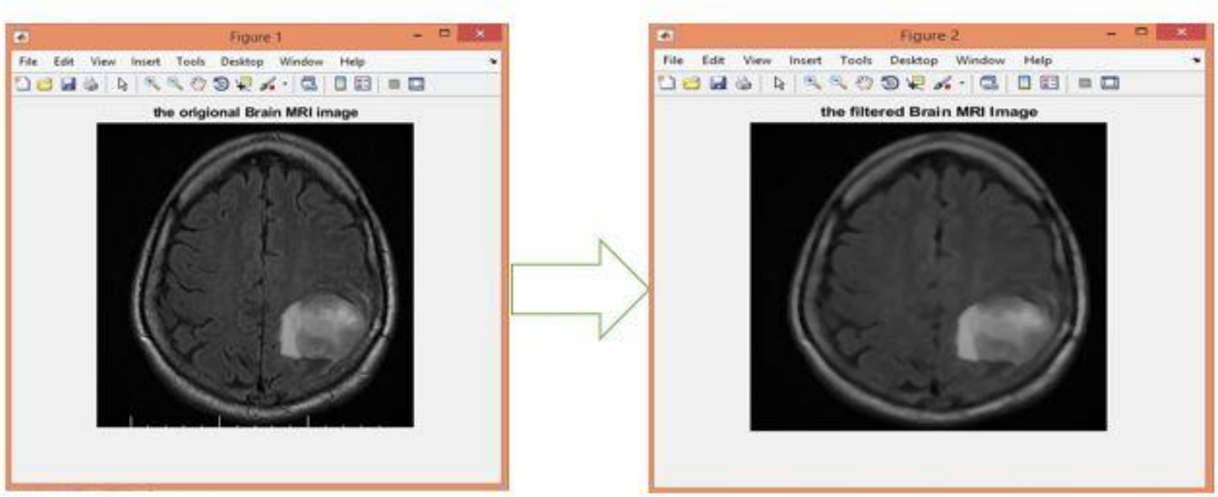

Figure 0.2 the comparison of original and filtered brain MRI image

4.4.1.2 Skull striping

The outer body (skull part) must be removed from the image to avoid any ambiguity for the classifier and to make the segmentation finer and more accurate. Therefore, we have tried to reduce the impact of the skull in the classification algorithm. As described in sec 3.5.2 the skull-stripping phase has three steps the first step is binerization of the image in which we use OTSU thresholding. Then we have selected the largest component of the image then a morphological operation was performed to fill holes and identify the full shape of the brain and then we use bitAND function to merge the masked and original image for extracting the brain.

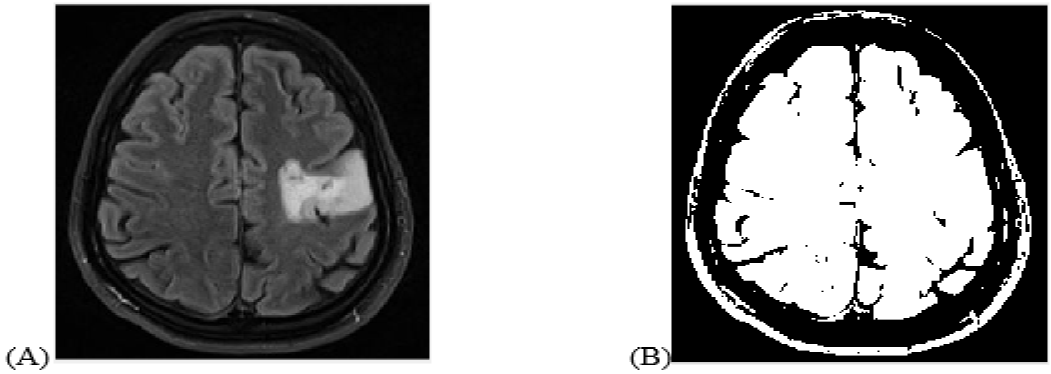

Figure 0.2 thresholding image. A) An axial view of brain MRI image. B) Threshold image using OTSU thresholding

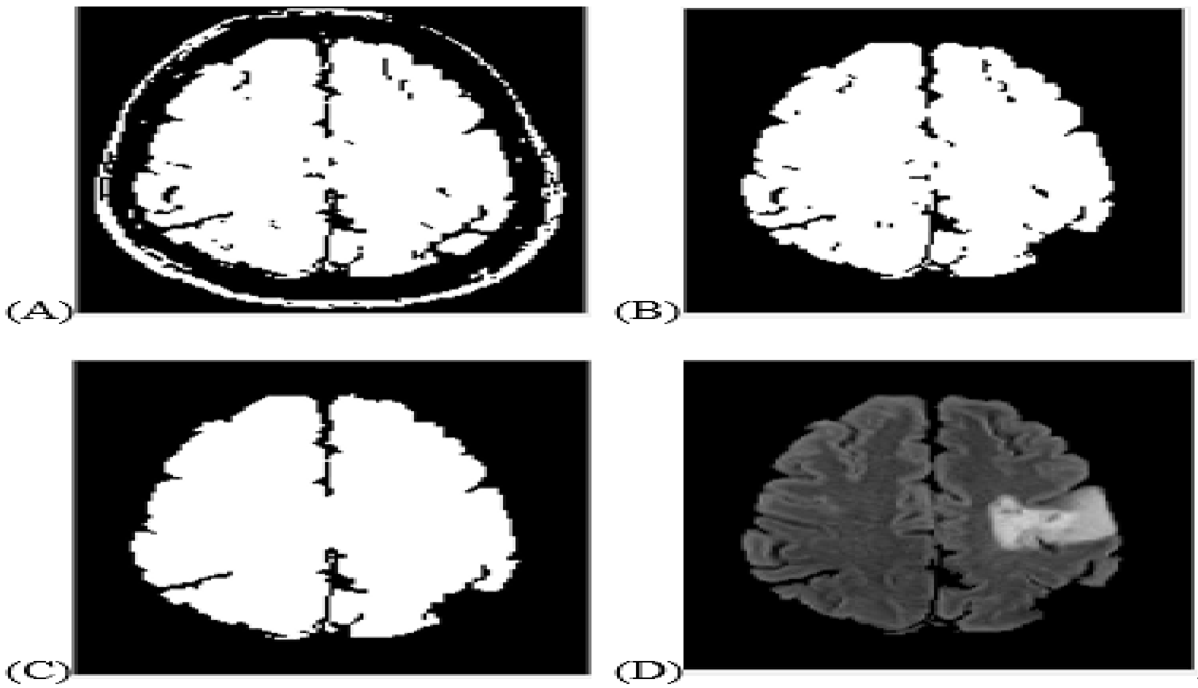

Figure 0.3 skull stripping process A) Threshold image using OTSU threshold B) The largest component extracted from the threshold image C) Masked image after morphological operation D) the final skull stripped image

4.4.2 Image segmentation

4.4.2.1 Segmentation using symmetric side analysis and thresholding technique

To segment the region of the tumor, we have used a dynamic thresholding mechanism in which we have first split the image into two separate parts. Then we compare the grayscale intensity of the two images and create another image that has a grayscale value of either zero or 1. This way will enable us to identify the tumor part from the entire image. Zero for the point on the new image where the difference between the two images on that specific point is equal or less than 35. And one for the point on the new image where the difference between the two images is greater than 35. This way we can clearly segment the tumor region from the rest of the image. We choose 35 as a separating value because of the better result that we obtained compared to other values. The following steps show how the proposed system works.

Step One: As described in figure 3.6 the first step in the proposed thresholding method is to split the image in to two equal parts.

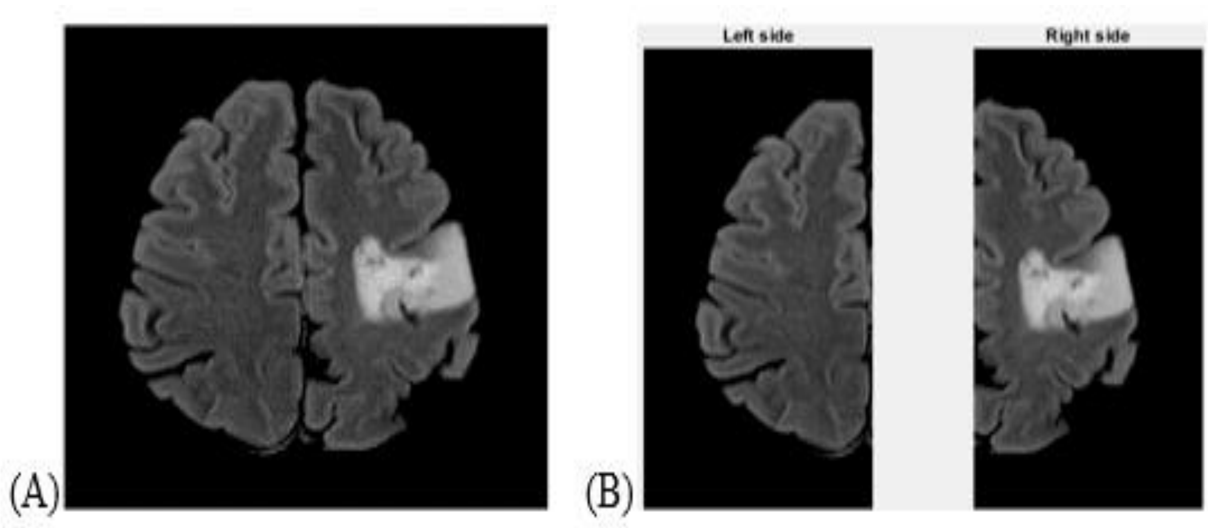

Figure 0.4 Splitting the image A) the original skull stripped image B) the split left side and right side of the image

Step Two: After we split the image we flipped the right side of the image so that we can compare the two sides

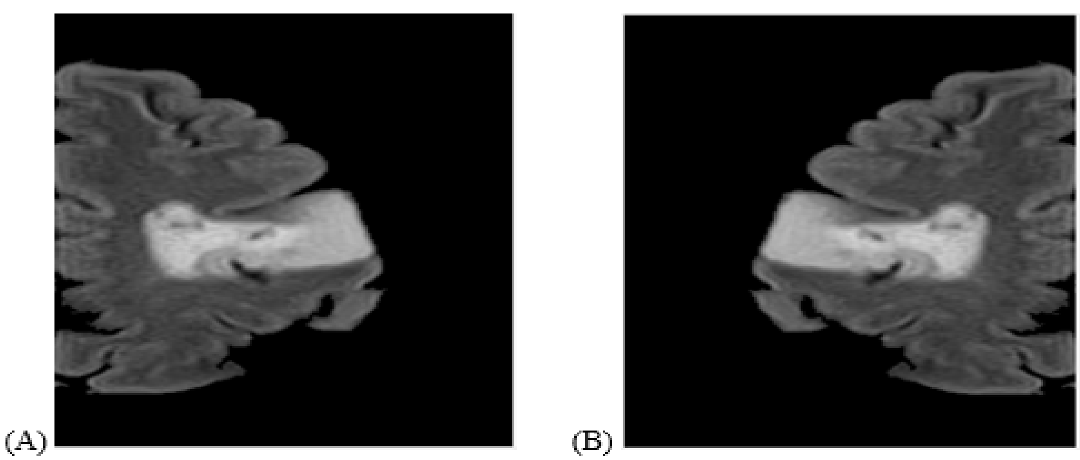

Figure 0.5 the right side of the brain A) the original right side of the image B) the flipped right side of the image

Step Three: After we flipped the right side of the image as shown in figure 4.6 we subtract the flipped right side of the image from the left side of the image and set the value to zero if the difference is less than 35, and set the value to 1 if the difference is greater than 35. In this way, we have the new threshold left side of the image. We have used the value 35 to be al local thresholding point because of the result that we have observed is higher when we use 35 than other values from 20 to 70. And we used this range because if the value is greater than greater than 70 it will be difficult to find tumor that has very low gray-scale difference from the actual brain and if the value is less than 20 the probability of counting noises as a tumor would be high.

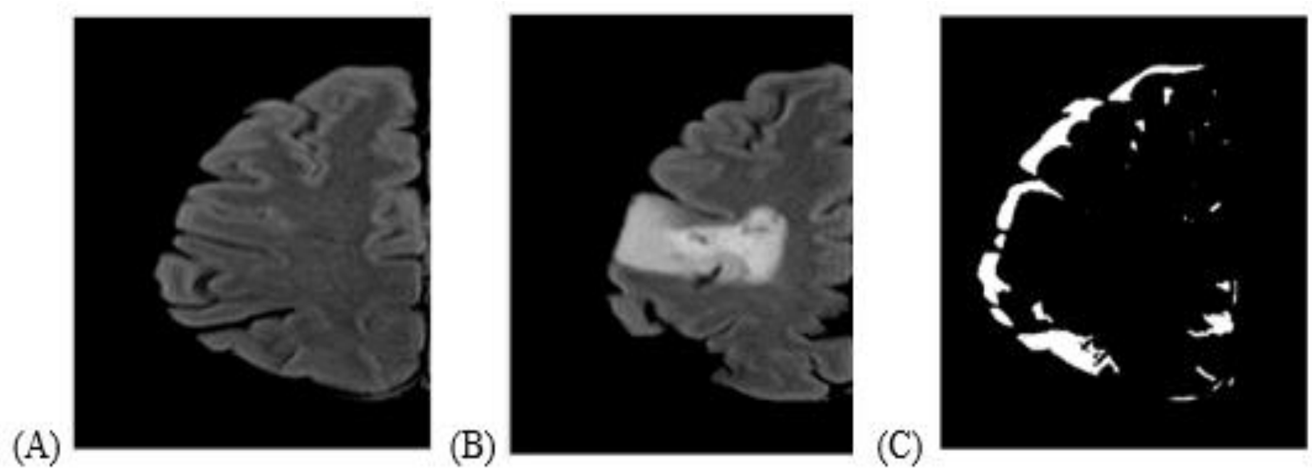

Figure 0.6 segmenting left side of the image A) the original left side of the image B) the flipped right side of the original image C) the new segmented left side of the image

We have used the following code for merging the two images

```
[rows,columns]=size(IMAGE1);

for p=1:rows
for q=1:columns
    SUBVAL=IMAGE1(p,q)-IMAGE2FLIP1(p,q);
if SUBVAL >= 35
BinaryImage1(p,q)=1;
else
BinaryImage1(p,q)=0;
end
end
end
```

Figure 0.7 MATLAB code used for segmenting the left side of the image

Step Four: just like we did in step three we subtract the left side of the image from the flipped right side so that we can get the segmented right side. After we found the left side of the image we then subtract the left side from the right side. By putting 1 where the difference is greater than 1 and, 0 where the difference between the two coordinate is less than or equal to 35. The following figure shows this step.

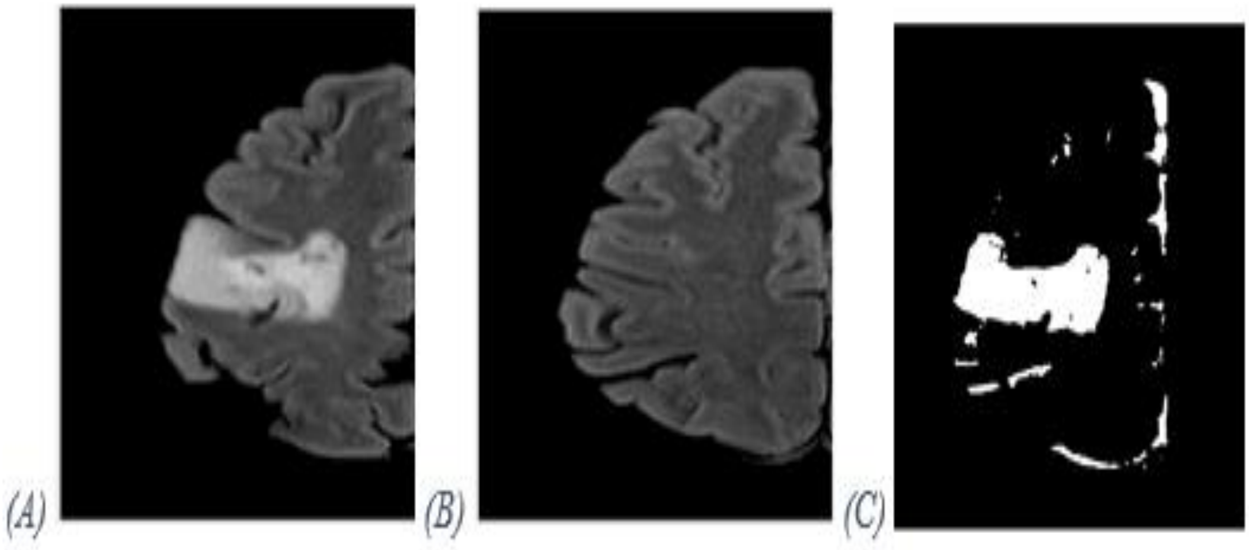

Figure 0.8 segmenting right side of the image A) the flipped left side of the image B) the original left side of the image C) the segmented right side of the image

```
[row1,column1]=size(IMAGE2FLIP1)
for x=1:row1
for y=1:column1
    SUBVAL1=IMAGE2FLIP1(x,y)-IMAGE1(x,y);
if SUBVAL1 >= 35
BinaryImage2(x,y)=1;
else
BinaryImage2(x,y)=0;
end
end
end
```

Figure 0.9 MATLAB code used for segmenting the right side of the image

Step Five: After we found the new right side of the image we then flip it back to its original position and then we merge the two new left side and right side. This way we can find the segmented image.

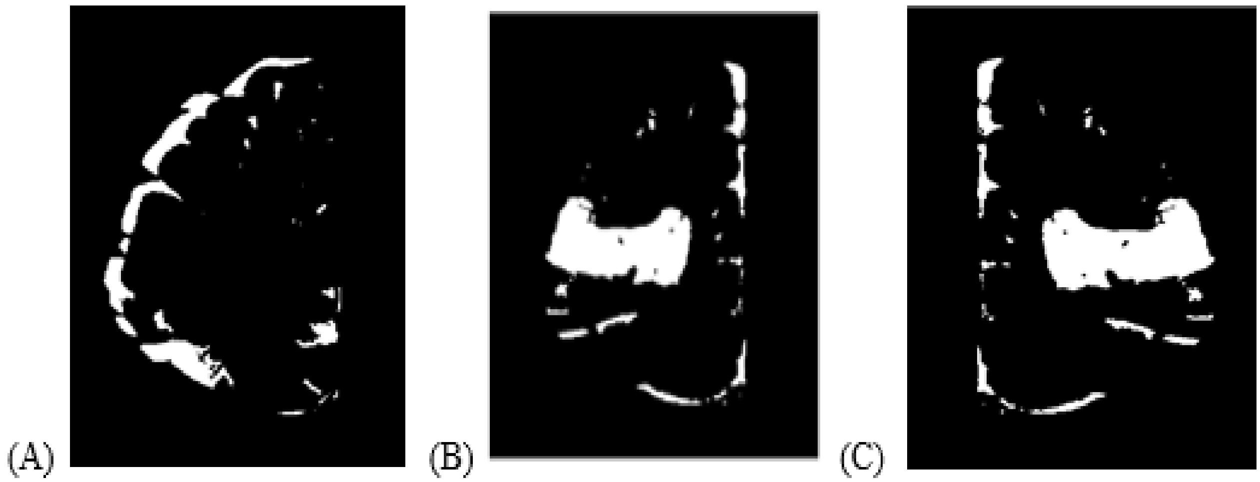

Figure 0.10 segmented left and right side of the image. A) the segmented left side of the image B) the segmented right side of the brain C) the segmented and flipped right side of the image

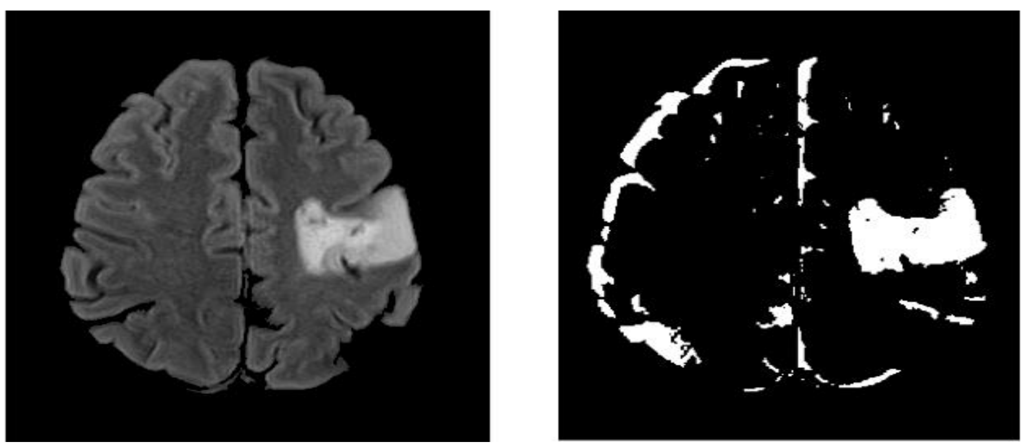

Figure 0.11 original skull stripped image and segmented d image

4.4.2.2 Segmentation using Global threshold value

Global thresholding is the process of converting an image into binary by taking a single constant value. In this thesis we use a global threshold value of 128 which is the around the midpoint of the gray scale. We have used this method so that we can compare the result of the proposed thresholding system with the preexisting technique.

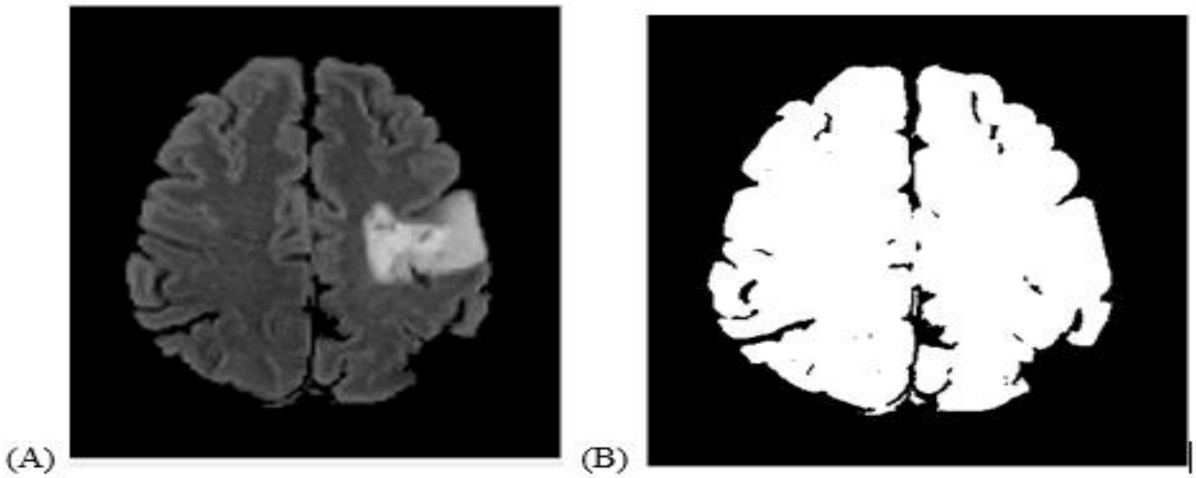

Figure 0.12 Global thresholding A) original skull stripped image B) segmented image using global thresholding

4.4.2.3 Segmentation using Sobel edge detection

The second technique we used to check the proposed segmentation technique with the preexisting segmentation methods is edge detection. We used Sobel edge detection to identify and segment the connected regions. The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of high spatial frequency that correspond to edges. Typically it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image.

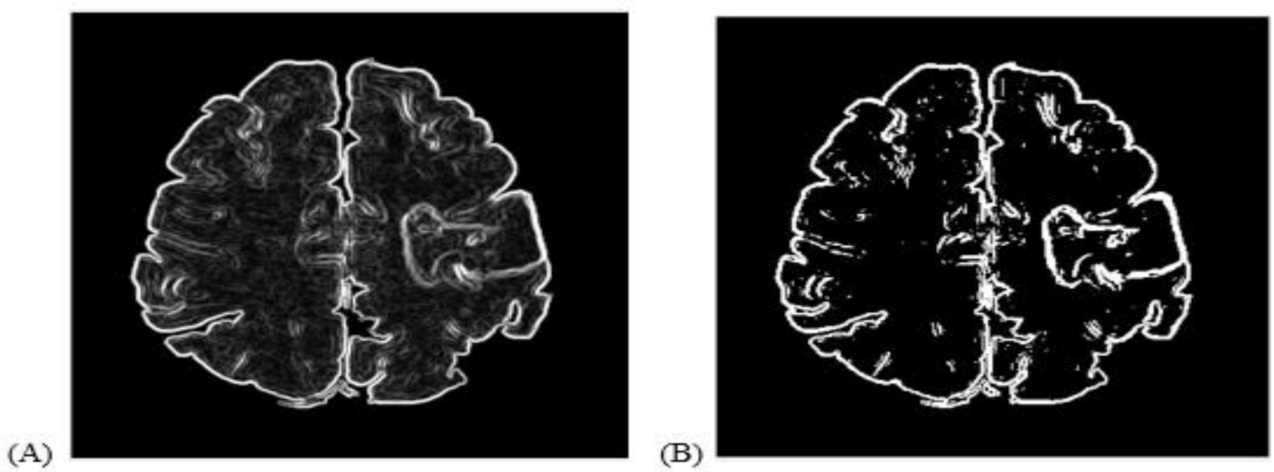

Figure 0.13Sobel Edge Detection A) Sobel Gradient of the original image B) Soble edge detection with global threshold

4.4.3 HoG Feature Extraction

Feature extraction is a technique by which unique features of images are extracted. An image feature is a distinguishing basic characteristic or attribute of an image. Due to the huge size of digital images, it can be time-consuming if an image is to be analyzed in its original form. So to make it simple and less time consuming, quantitative information is extracted from the objects to be analyzed in the image (Fanuel, 2018).

Varying the HOG cell size parameter differs the amount of shape information encoded in the feature vector. For example, the cell size of [8 8] does not encode much shape information, while a cell size of [2 2] encodes a lot of shape information but increases the dimensionality of the HOG feature vector significantly. So a good compromise is 4-by-4 cell size. So by using 4x4 cell size for the detected tumor a HOG feature vector and this feature vector is used to classify the tumor into their respective classes

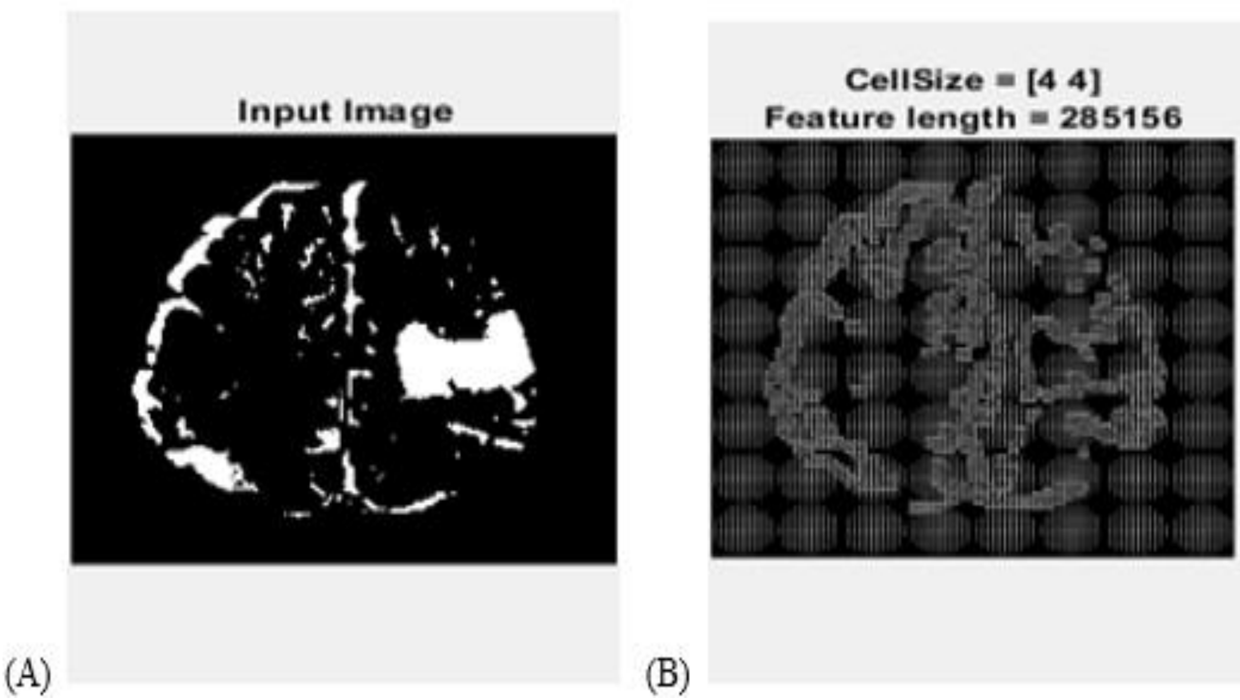

Figure 0.14 HOG feature extraction A) Segmented image using proposed segmentation technique B) Feature extraction using HOG

4.4.4 SVM Classification

In this system, once a region of interest (ROI) which contains the tumor part is determined, classification of the tumor type is done using SVM algorithm. The outcome of the above-discussed feature extraction process (HOG) is a numeric feature vector. And this vector is fed to the classification algorithm as an input. So by taking those features, this classification phase is responsible to classify three different types of brain tumor and normal brain into their corresponding classes.

Brain tumor detection is a multiclass classification problem, where you have to classify an image into one out of the many possible brain tumor classes. In this study, the “fitcecoc” function from the Statistics and Machine Learning Toolbox™ is used to create a multiclass classifier using binary SVMs. SVM is well-known machine learning technique for classification, and it is a highly accurate and fast which has a very low computational time for real-time applications.

The procedures of system testing are almost similar to the training process except in the classification module. To test an image using the trained brain tumor detection system all the pre-processing and detection steps are applied to the image and finally, it will use the features of the trained images for the classification process. And by doing so the performance of the system is measured for all test images which are about 25% of the total dataset and confusion matrix which shows classification accuracy is generated.

4.4.5 Result and discussion

In this study, the implemented brain tumor detection system is done on three different tumor types and normal brain. Each class contains images ranging from 48 to 60 for training and 10 to 21 for testing individually and 281 images totally in four classes. In addition in this work, the system is trained and tested with three different types of thresholding techniques including the proposed technique, using HOG feature extraction and SVM classification technique, so that the result is different for each method and it is discussed below for each. The used and implemented combinations are:

I. Segmentation using the proposed symmetrical side analysis thresholding with HOG feature extraction and SVM classifier.

II. Segmentation using global thresholding with HOG feature extraction and SVM classifier.

III. Segmentation using edge detection with HOG feature extraction and SVM classifier.

4.4.5.1 Segmentation using the proposed symmetrical side analysis thresholding method

This is an experiment in which detection and Classification is done on the proposed segmentation technique to find the thresholding values, and for testing the system 63 images are used, which is 25% of the total dataset. The following table shows the confusion matrix that indicates the correct and wrong classifications of all the test images that are divided into four categories. The columns of the matrix represent the predicted labels, while the rows represent the known labels.

Table 0.2 Summery result of the classification of the proposed method

Predicted Class / Actual Class	Meningiomas	Gliomas	Pituitary	Normal Brain	Total
Meningiomas	16	0	0	4	20
Gliomas	0	21	0	0	21
Pituitary	2	0	7	3	12
Normal brain	1	0	1	8	10

Segmentation using the proposed symmetrical side analysis thresholding for the brain tumor detection and classification system, as it is shown in the above table from the total 63 test images 52 images or 82.53% are classified correctly, and 11 images or 17.47% are misclassified, and for each class the accuracy of classification result shows :

80% - for Meningiomas tumor

100% - for Gliomas

58% - for Pituitary and

80% - for normal Brain MRI

But since the dataset used is not balanced the above accuracy value couldn't tell us the models performance correctly, instead of measuring the system using F-cross metric is more preferable with imbalanced data, so the table below shows classification performance of the model in terms of recall, precision, and F-measure:

Table 0.3 precision and recall result of the proposed classification

Detection Class	**Precision**	**Recall**
Meningiomas	0.842	0.8
Gliomas	1	1
Pituitary	0.875	0.583
Normal brain	0.533	0.8

Having precision and recall of each class F-score of the model is calculated as: F measure = 2*(Average Precision)*(Average Recall) / (Average Precision + Average Recall) and the result is **0.8031**

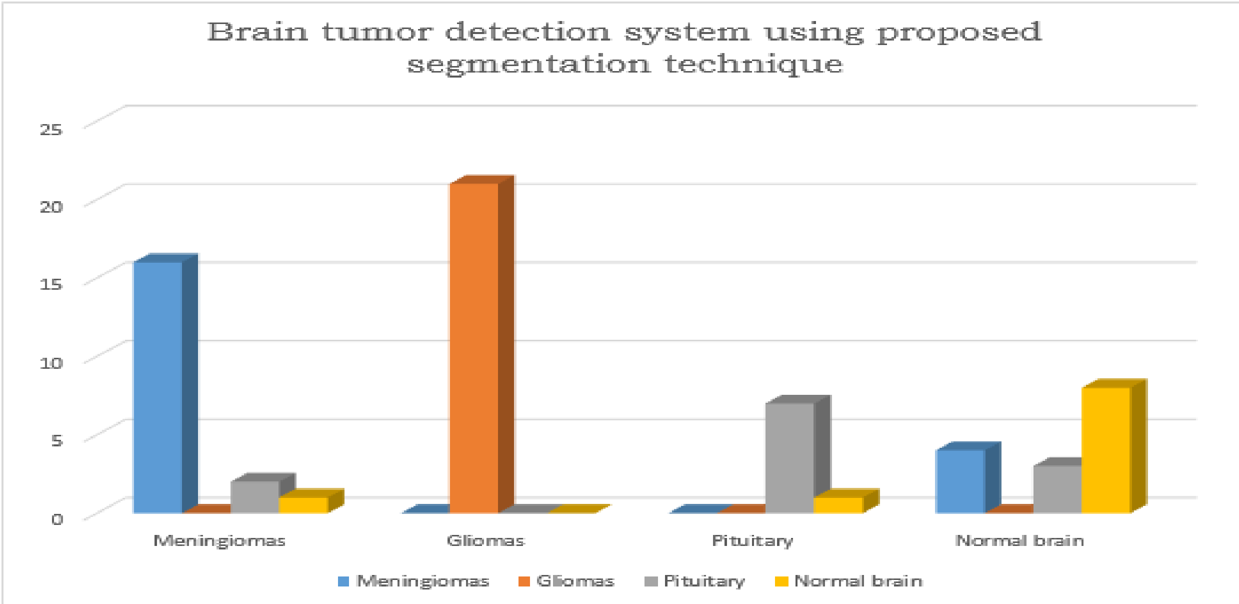

Figure 0.15 Classification Result using proposed segmentation technique

4.4.5.2 *Segmentation using global thresholding*

This is an experiment in which detection and Classification is done on with global thresholding values, and for testing the system 63 images are used, which is 25% of the total dataset. The following table shows the confusion matrix that indicates the correct and wrong classifications of all the test images, which are divided into four categories. The columns of the matrix represent the predicted labels, while the rows represent the known labels.

Table 0.4 Summary result of classification using global threshold segmentation

Predicted Class / Actual Class	Meningiomas	Gliomas	Pituitary	Normal Brain	Total
Meningiomas	14	0	0	6	20
Gliomas	3	18	0	0	21
Pituitary	2	0	4	6	12
Normal brain	0	0	2	8	10

Using global thresholding values of 0.5 for the brain tumor detection and classification system, as it is shown in the above table from the total 63 test images 44 or 69.84% are classified correctly, and 19 images or 30.16% are misclassified, and for each class the accuracy of classification result shows:

70% - for Meningiomas tumor

86% - for Gliomas tumor

33% - for Pituitary tumor and

80% - for normal brain MRI

But since the dataset used is not balanced the above accuracy value couldn't tell us the models performance correctly, instead of measuring the system using F-cross metric is more preferable with imbalanced data, so the table below shows classification performance of the model in terms of recall, precision, and F-measure

Table 0.5 Precision and recall result of classification using global thresholding for segmentation

Detection Class	**Precision**	**Recall**
Meningiomas	0.736	0.7
Gliomas	1	0.857
Pituitary	0.66	0.33
Normal brain	0.4	0.8

Having precision and recall of each class F-score of the model is calculated as: F measure = 2*(Average Precision)*(Average Recall) / (Average Precision + Average Recall) and the result is **0.685**

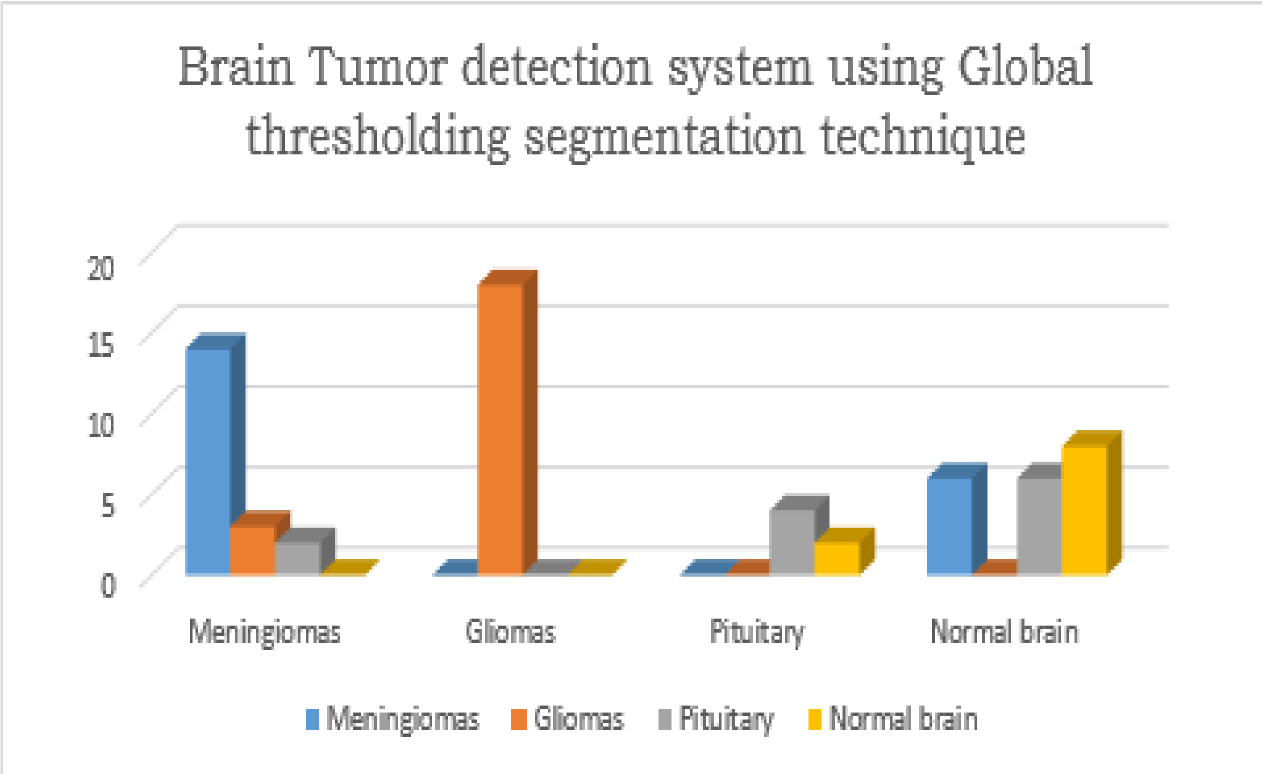

Figure 0.16 Classification Result using Global thresholding segmentation technique

4.4.5.3 *Segmentation using edge detection*

This is an experiment in which detection and classification is done on with Sobel edge detection and for testing the system 63 images are used, which is 25% of the total dataset. The following table shows the confusion matrix that indicates the correct and wrong classifications of all the test images, which are divided into four categories. The columns of the matrix represent the predicted labels, while the rows represent the known labels.

Table 0.6 Summary result of classification using Sobel edge detection for segmentation

Predicted Class / Actual Class	Meningiomas	Gliomas	Pituitary	Normal Brain	Total
Meningiomas	15	0	0	5	20
Gliomas	6	15	0	0	21
Pituitary	2	0	4	6	12

Normal brain	2	0	0	8	10

Using Sobel edge detection for the brain tumor detection and classification system, as it is shown in the above table from the total 63 test images 42 or 66.66% are classified correctly, and 21 images or 33.44% are misclassified, and for each class the accuracy of classification result showed that:

75% - for Meningiomas tumor

71% - for Gliomas tumor

33% - for Pituitary tumor and

80% - for normal brain MRI

But since the dataset used is not balanced the above accuracy value couldn't tell us the models performance correctly, instead of measuring the system using F-cross metric is more preferable with imbalanced data, so the table below shows classification performance of the model in terms of recall, precision, and F-measure:

Table 0.7 Precision and recall result of classification using Sobel edge detection for segmentation

Detection Class	**Precision**	**Recall**
Meningiomas	0.6	0.75
Gliomas	1	0.714
Pituitary	1	0.333
Normal brain	0.421	0.8

Having precision and recall of each class F-score of the model is calculated as: F measure = 2*(Average Precision)*(Average Recall) / (Average Precision + Average Recall) and the result is **0.7021**

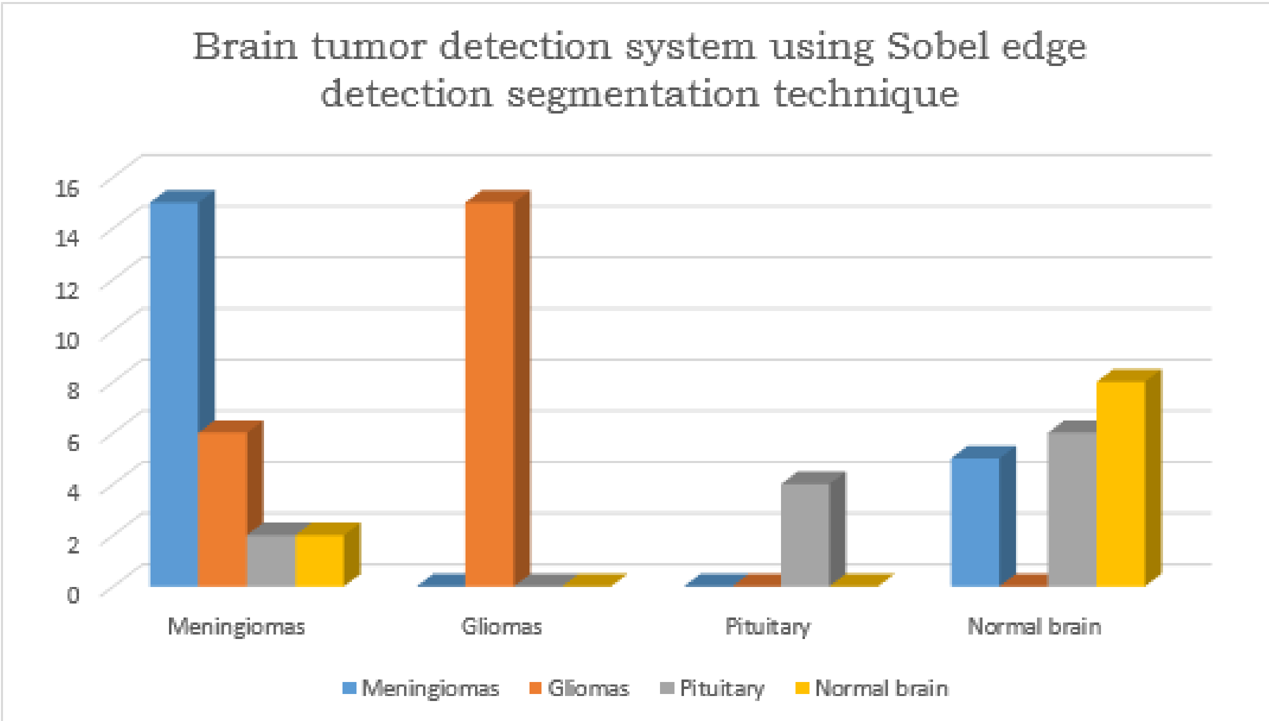

Figure 0.17 Classification Result using Sobel edge detection for segmentation technique

4.5 Discussion

As we have presented in detail in the previous section, the experiments were conducted under three scenarios by using proposed segmentation technique, global thresholding and Sobel edge detection. Then, features were extracted using HOG feature extraction technique and we have used SVM classifier to classify the tumor type. The total number of data sets is 281. Out of these, 75% were used for training and the remaining 25% were used for testing. In general, the overall result showed that the proposed segmentation technique has a better result than the other segmentation techniques. And also the system has a higher classifying efficiency for Gliomas tumor type. And low classifying efficiency for classifying Pituitary tumor type. And this is because of the location of the Pituitary tumor is in the middle of the brain which makes it difficult to be identified while we split the images into two parts.

4.6 Factors that affect brain tumor detection and classification system

Such systems are dependent on different factors, some of the factors that hindered us from achieving the maximum result are:

- **Size and shape of the image**: Since the proposed segmentation technique compares the left and right hemisphere of the brain, the image alignment must be perfect. Therefore having an image that tilt in any side or an image that is not symmetrical could result in misrepresentation.

- **Size and shape of the tumor**: the proposed system uses the comparison of the left and right side of the brain and while we are comparing these sides if the tumor is small the probability of misclassifying it will be high.

- **Location of the tumor**: the proposed system uses the comparison of the left and right side of the brain if the tumor is located near to the center of the brain the probability of it to be classified as pituitary tumor or normal brain MRI is high.

- **Quality of the MRI image**: even though we have used a filtering technique to reduce the noise there might still be some noises that might confuse the system, therefore, having a clear image could result in a better performance of the system.

- **Used Dataset**: Accuracy of the system is also directly related to the number of images used as a dataset. A good accuracy will be achieved if a large amount of dataset is used.

Chapter Five: CONCLUSION AND RECOMMENDATION

5.1 Conclusion

A brain tumor is a collection, or mass, of abnormal cells in the brain. Any growth inside small and rigged space can cause problems. MRI is most effective way of acquiring an image of the brain for tumor detection. However, detection and extraction of Brain tumor via MRI is a challenging task for physicians in a medical field. Therefore, a CAD system that can detect and classify a brain tumor using the MR image is very important. Even though there have been many CAD system that have been develop to clearly segment and classify a brain tumor there is a still a lake of precision. Therefore In this paper a brain tumor detection and classification system is successfully implemented using a novel thresholding technique.

In this study an implementation of brain tumor detection and classification system is done. This system has 3 parts which are: preprocessing, detection and classification, and is implemented for 3 tumor types and normal brain image. As a dataset, a collection of 281 images is used and from that 75% is applied for training and the rest 25% is used as testing data. The detection module of the system is done using three different segmentation approaches, which are: segmentation using the proposed segmentation using symmetrical side analysis and dynamic threshold value, segmentation using global thresholding value and segmentation using sobel edge detection technique. In all those three experiments HOG feature extraction with SVM classification technique are used for classification phase and their F-measure is 80.31%, 68.5%, and 70.21% respectively. Therefore, classification using the proposed segmentation symmetrical side analysis and dynamic threshold value with HOG feature extraction and SVM classifier shown better result than the other two.

5.2 Recommendation

In this study as it is described in the previous section, the tumor detection system is done using a new proposed and two previously used segmentation technique with same feature extraction and classification technique for three tumor types. To make the system more efficient it needs some additional works, some of them are:

- This thesis work is done only for 3 tumor types with a small dataset as a prototype, so for the future for increasing its accuracy it should have a large data set including other types of tumor.

- In this thesis work, only one feature extraction and classification method is used, so for the future, it is better to do it with the latest alternative algorithms to increase its performance.

- After implementing the proposed thresholding technique we have only compared its performance with non-intelligent based segmentation techniques, so for the future it is better to test the system performance compared to the intelligent based segmentation techniques

References

Asra Aslam, E. K. (2015). Improved Edge Detection Algorithm for Brain Tumor Segmentation. *Second International Symposium on Computer Vision and the Internet (VisionNet'15)*, 430 – 437.

Aycheh, H. M. (2008). IMAGE ANALYSIS FOR ETHIOPIAN COFFEE CLASSIFICATION. *bahir dar university thesis*, 1-108.

B.Devkota, Abeer, A., P.W.C.Prasada, A.K.Sing, & EA.Elchouemi. (2018). Image Segmentation for Early Stage Brain Tumor Detection using Mathematical Morphological Reconstruction. *Procedia Computer Science Volume 125*, 115-123.

Balakumar .B, M. S. (2014). An Automatic Brain Tumor Detection and Segmentation Scheme for Clinical Brain Images. *International journal of imerging technologies in computetional and applide science*, 37-42.

Bettegowda C1, S. M. (2014). detection of circulating tumore DNA in early and late srage human malignancies. *science translation medicine*, 1-25.

Bezdek, J. C., Robert, E., & William, F. (1984). FCM: The Fuzzy C-Means Clustering Algorithm. *Computer and Geosciences, Vol. 10, No. 2-3*, 193-203.

Brian, K., & Mariah, A. (2019, 03 25). *healthline.com/health/pet-scan*. Retrieved from healthline web site: https://www.healthline.com/health/pet-scan

Cancer treatment center for America. (2019, 05 13). Retrieved from Cancer treatment center for america web site: https://www.cancercenter.com/diagnosing-cancer/diagnostic-imaging

Cancer.net. (2019, 04 21). Retrieved from cancer.net/cancer-types/brain-tumor/statistics: https://www.cancer.net/cancer-types/brain-tumor/statistics

cancercenter.com. (2019, 03 05). Retrieved from Cancer treatment center for America: https://www.cancercenter.com/diagnosing-cancer/diagnostic-imaging/ct-scans

Chalechale , A., Mertins, A., & Naghdy , G. (2004). Edge image description using angular radial partitioning. *EE Proc.-Vis. Image Signal Process., Vol. 151, No. 2.*

Chih-Wei , H. (2016). A Practical Guide to Support Vector Classification.

Colombia University Department of Neurology. (2019, 05 11). Retrieved from Colombia University Department of Neurology: http://www.columbianeurology.org/neurology/staywell/document.php?id=42006

Dunn, J. C. (1973). A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics, Vol. 3, No.3*, 32-57.

El-Dahshan, E.-S. &.-B. (2014). Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm. *Expert Systems with Applications*, 5526–5545.

F. B. (2018). *Voice-Based Automatic Traffic Sign Detection and Recognition System.* Bahir Dar.

Gordillo, N., Montseny, E., & Sobrevilla, P. (2010). A New Fuzzy Approach to Brain Tumor Segmentation. *IEEE International Conference on Fuzzy Systems (FUZZ)*, 1-8.

Jin , L., Min , L., Jianxin , W., Fangxiang , W., & Tianm. (2014). A Survey of MRI-Based Brain Tumor Segmentation Methods. *TSINGHUA SCIENCE AND TECHNOLOGY Volume 19, Number 6*, 578-595.

John, P. (2012). Brain Tumor Classification Using Wavelet and Texture Based Neural Network. *International Journal of Scientific & Engineering Research Volume 3*, 1-7.

Jyoti, D. B., Kumar , A. G., & Ahmad , F. K. (2015). omparing the Performance of L*A*B* and HSV Color Spaceswith Respect to Color Image Segmentation. *International Journal of Emerging Technology and Advanced Engineering Volume 5, Issue 2*, 192-203.

Kailash , K. D., Pradyumna , K. P., & Nagori , M. B. (2012). Brain Tumor Classification Using Neural Network Based Methods. *International Journal of Computer Science and Informatics ISSN (PRINT): 2231 –5292, Vol-1*, 85-90.

Kalavathi, P., & Surya Prasath, V. B. (2016). Methods on Skull Stripping of MRI Head Scan Images—a Review. *Society for Imaging Informatics in Medicine* , 365-379.

Kalpana , R. U., & Prof. Kapse, Y. D. (2016). MATLAB Based Brain Tumor Extraction Using Artificial Neural Network. . *International Journal on Recent and Innovation Trends in Computing and Communication. Volume: 4 Issue: 3*, 1-9.

Kazi, I. M., Chowhan, S. S., & Kulkarni, U. V. (2017). MRI Brain Image segmentation using Adaptive Thresholding and K-means. *International Journal of Computer Applications Volume 167 – No.8*, 11-15.

Logeswari, T., & Karnan, M. (2010). An improved implementation of brain tumor detection. *Journal of Cancer Research and*, 6-14.

MayoClinin Pituitary tumors. (2019, 05 11). Retrieved from mayoclinic.org: https://www.mayoclinic.org/diseases-conditions/pituitary-tumors/symptoms-causes/syc-20350548

Megersa, Y. (2012, 11). Brain Tumor Detection and Segmentation Using Hybrid Intelligent Algorithms: Design and Implementation. pp. 1-86.

Mehmed , O., Benoit, D. M., & Robert, M. J. (1993). Neural Network Based Segmentation of Multi-Modal Medical Images: A comparative and Prospective Study. *IEEE Transaction on Medical Imaging, Vol. 12, No. 13*, 534-544.

Miss. Priyanka Katti, M. V. (2015). Implementation of Classification System for Brain Tumor using Probabilistic Neural Network. *International Journal of Advanced Research in Computer and Communication Engineering Vol. 4*.

Mitisha, N. P., & Purvi, T. (2016). A Survey on Feature Extraction Techniques for Shape based Object Recognition. *nternational Journal of Computer Applications (0975 –8887)Volume 137–No.6*, 16-20.

Nikita V. Chavan, B. P. (2015). Detection andClassification of Brain Tumors. *International Journal of Computer Applications (0975 –8887) Volume 112 –No. 8*, 48-53.

Nilesh, B. B., Arun, K. R., & Har, P. T. (2017). Image Analysis for MRI Based Brain Tumor Detection and Feature Extraction Using Biologically Inspired BWT and SVM. *International Journal of Biomedical Imaging*, 1-12.

Otsu, N. (1979). A threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No.1*, 62-66.

Pavel , D. (2013). Detection of Brain Tumors Based on Automatic Symmetry Analysis. *18thComputer Vision Winter WorkshopWalter G. Kropatsch, Fuensanta Torres, Geetha Ramachandran (eds.)Hernstein*.

Penn Medicine Neuroscience blog. (2019, 05 11). Retrieved from Penn Medicien: https://www.pennmedicine.org/updates/blogs/neuroscience-blog/2018/november/what-are-the-most-common-types-of-brain-tumors

Prasath, P. K. (2015). Methods on Skull Stripping of MRI Head Scan Images—a Review. *J Digit Imaging.*, 365-379.

QURAT-UL-AIN, LATIF, G., SIDRA, K. B., M., J. A., & ANWAR, M. M. (2010). Classification and Segmentation of Brain Tumor using Texture Analysis. *RECENT ADVANCES in ARTIFICIAL INTELLIGENCE, KNOWLEDGE ENGINEERING and DATA BASES*, 147-155.

Rafael, G. C., & Paul, W. (1987). *Digital Image Processing.* New Jersey: Prentice Hall.

Salem, M. A.-b. (2016). An authomatic classification of brain tumor through MRI using SVM. *Egyptian computer science Journal*, 11-22.

Seemanthini , K., & Dr.Manjunath, S. S. (2018). Human Detection and Tracking using HOG for Action Recognition. *International Conference on Computational Intelligence and Data Science (ICCIDS 2018)*, 1317-1326.

society, n. b. (2014). *understanding brain tumor*. National brain tumor society.

Solomon Tessema Memirie, M. K. (2018). Estimates of Cancer Incidence in Ethiopia in 2015 Using Population-Based Registry Data. *Journal of global onchology an american socity of clinical oncology journal*.

Stats(CBTRUS), C. B. (2019, 04 15). *Central Brain Tumor Registry of United Stats*. Retrieved from http://www.cbtrus.org: http://www.cbtrus.org/factsheet/factsheet.html

Swathi , P. S., Deepa , D., Vince , P., & Sankaranarayanan, P. N. (2015). Brain Tumor Detection and Classification Using Histogram Thresholding and ANN . *(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (1)*, 173-176.

Tumor Types: Understanding Brain Tumors. (2019, 05 11). Retrieved from National brain tumor socity: https://braintumor.org/brain-tumor-information/understanding-brain-tumors/tumor-types/

Types of classification algorithms in Machine Learning. (2019, 04 12). Retrieved from https://medium.com/: https://medium.com/@Mandysidana/machine-learning-types-of-classification-9497bd4f2e14

Valentina , P., Elisabetta , B., Sergio , B., Alessandro, D., Emanuele, M., & Renzo , M. (2012). Glial brain tumor detection by using symmetry analysis. *The International Society for Optical Engineering*, 141-150.

Vapnik , V. (2004). The Nature of Statistical Learning Theory.

Weiss, E. (2013). *frankly speaking about cancer brain tumors*. USA: National Brain Tumor Socity.

Xiao , H., Peng , F., Wang , L., & Li , H. (2007). Ad hoc-based feature selection and support vector machine classifier for intrusion detection. *IEEE International Conference on Grey Systems and Intelligent Services*, 1117-1121.

Xie, k., Jie , Y., Z.G. Zhang, & Y.M. Zhu. (2005). Semi-automated brain tumor and edema. *European Journal of Radiology, Vol. 56 , Issue 1*, 12 -19.

Yadav, M. K. (2012). Brain Tumor Detction and Segmentation Using Histogram Thresholding. *International Journal of Engineering and Advanced Technology (IJEAT)*, 16-20.

Zelalem, H. (2017). *APPLYING MEDICAL IMAGE PROCCESSING TECHNIQUES FOR THE DETECTION OF MOST FREQUENTLY OCCURRING PULMONARY DISEASES.* Bahir Dar.