VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"Jnana Sangama", Belagavi-590018, Karnataka

A Project Report on

"BRAIN TUMOR DETECTION AND CLASSIFICATION USING PSO AND SVM CLASSIFIER"

Submitted in partial fulfillment of the requirement for the award of degree of

Bachelor of Engineering

In

Computer Science and Engineering

Submitted by

ABHAY P NAIK (4NN16CS001)

ADITHYA REDDY B (4NN16CS003)

ANOOP KUSHALAPPA B D (4NN16CS011)

VISHRUTH M (4NN16CS059)

Under the Guidance of

Mr. Deepak P

Assistant Professor

Dept. of CS & Engineering

ESTD-2008

Department of Computer Science and Engineering

NIE Institute of Technology

Mysuru -570018.

2019-2020

Department of Computer Science and Engineering

NIE Institute of Technology

ESTD-2008

CERTIFICATE

This is to certify that the project work entitled "***BRAIN TUMOR DETECTION AND CLASSIFICATION USING PSO AND SVM CLASSIFIER"*** is carried out by ***ABHAY P NAIK, , ANOOP KUSHALAPPA B D, VISHRUTH M, ADITHYA REDDY B*** bearing **USN: *4NN16CS001, 4NN16CS011, 4NN16CS059, 4NN16CS003*** in partial fulfilment for the Eighth semester of **Bachelor of Engineering in Computer Science & Engineering** of the **Visvesvaraya Technological University, Belagavi** during the academic year **2019-2020**. The project report has been approved as it satisfies the academic requirements with respect to project work prescribed for the Bachelor of Engineering degree.

Signature of the Guide & Coordinator

Mr. Deepak P
Asst. Professor,
Dept. of CS & Engineering
NIEIT, Mysuru

Signature of the HOD

Smt. Vedavathi
Asst. Professor & Head,
Dept. of CS & Engineering
NIEIT, Mysuru

External Viva

Name of the examiners	**Signature with Date**
1……………………....	**1**……………………….
2……………………….	**2**………………………

DECLARATION

We**, Abhay P Naik** of USN **4NN16CS001,Adithya Reddy B** of USN **4NN16CS003,Anoop Kushalappa B D** of USN **4NN16CS011** and **Vishruth M** of USN **4NN16CS059,** hereby undertake that the project work entitled **“BRAIN TUMOR DETECTION AND CLASSIFICATION USING PSO AND SVM CLASSIFIER”** is carried out by us independently under the guidance of **Mr DEEPAK P**, Asst. Professor, Department of Computer Science and Engineering, NIE IT,Mysore-18, in partial fulfillment of the requirement for the award of **Bachelor of Engineering in Computer Science and Engineering** by the **Visvesvaraya Technological University, Belagavi**. The project has been our original work and has not formed the basis for the award of any degree, associate ship, fellowship or any other similar titles.

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without mentioning of the people who made it possible, whose consistent guidance and encouragement crowned our efforts with success.

We consider ourselves proud to be a part of **NIE Institute of Technology Mysuru** family, the institution which stood by our way in all our endeavors.

We wish to express our gratitude to **Prof. Bhansilal,** Principal, **NIE Institute of Technology Mysuru,** for providing a congenial working environment.

We express our sincere thanks to **Smt. Vedavathi**, Assistant Professor & Head, Department of Computer Science and Engineering, for her support and encouragement.

We would like to thank our guides**, Mr. Deepak P** and **Mr. Madhusudhan H S** Assistant Professors, Department of Computer Science and Engineering, for their inspiration, guidance, constant supervision, direction and discussions in successful completion of this project work.

We are thankful to the Department of Computer Science and Engineering staff members and non-staff members for their co-operation extended towards this work.

Finally, heartfelt gratitude to our **family members, brother, relatives** and **friends** for their constant support, motivation and encouragement throughout this project. Above all I would like to thank God, Almighty, for having showered his blessings on me.

ABHAY P NAIK (4NN16CS001)
ANOOP KUSHALAPPA B D (4NN16CS011)
VISHRUTH M (4NN16CS059)
ADITHYA REDDY B (4NN16CS003)

ABSTRACT

The brain is the most important organ in the human body, responsible for controlling and regulating all critical life functions for the body and a tumor is a mass of tissue formed by the accumulation of abnormal cells, which keep on growing. A brain tumor is a tumor which is either formed in the brain or has migrated No primary cause has been identified for the formation of tumors in the brain till date. Though tumors in the brain are not very common (Worldwide brain tumors make up only 1.8% of total reported tumors), the mortality rate of malignant brain tumors is very high due to the fact that the tumor formation is in the most critical organ of the body. Hence, it is of utmost importance to accurately detect brain tumors at early stages to lower the mortality rate. We have thus proposed a computer-assisted radiology system which will assess brain tumors from MRI scans for the management of brain tumor diagnosis. In this study, we have implemented a model that segments images using Otsu algorithm, extracts features using DWT and PCA algorithms, select features using BPSO algorithm and finally classifies the tumors using SVM algorithm with a high degree of accuracy.

CONTENTS

List of Figures:

List of Tables:

Chapter 1

INTRODUCTION

Image Processing has an important role in medical field and medical imaging is a growing and challenging field. Medical imaging is advantageous in diagnosis of the disease. Medical imaging provides proper diagnosis of brain tumor. There are many techniques to detect brain tumor from MRI images.

Brain tumor is an abnormal and undesirable growth of tissue cells in brain which leads to certain neurological disorder for human. Now a days due to environmental and human life style these tumor cases are rapidly increasing. There are two main types of tumors: malignant (cancerous) tumors and benign (non-cancerous) tumors. Cancerous tumors can be divided into primary tumors, which start within the brain, and secondary tumors, which have spread from elsewhere, known as brain metastasis tumors. There are over 120 different types of primary brain tumors. A benign brain tumor consists of very slow-growing cells, usually has distinct borders and rarely spreads. When viewed under a microscope, these cells have an almost normal cellular appearance. A malignant brain tumor is usually rapidly-growing, invasive and life-threatening. Malignant brain tumors are sometimes called brain cancer.

1.1 Aim

The main goal of brain tumor detection and classification is to identify accurate and meaningful information using images with the minimum error possible. MRI is mainly used to get images of the human body and cancerous tissues because of its high resolution and better quality images compared with other imaging technologies. Normally the anatomy of the brain is analysed by MRI scans or CT scans. The proposed system aims to detect if the given MRI scan has a tumor and then it classifies whether the tumor is malignant or benign.

1.2 Purpose

The purpose of this project it to identify the brain tumor at early stage which is very difficult task for doctors to identify. MRI images are more prone to noise and other environmental interference. So, it becomes difficult for doctors to identify tumor and their causes. So here we come up with the system, where system will detect brain tumor from images.

Here we convert image into grayscale image. User has to select the image. System will process the image by applying image processing steps. We applied a unique algorithm to detect tumor from brain image.

1.3 Scope

The scope is to develop an automated system for enhancement, segmentation and classification of brain tumors. The system can be used by neurosurgeons and healthcare specialists. The system incorporates image processing, image segmentation, feature extraction and image classification and is expected to improve the sensitivity, specificity, and efficiency of brain tumor detection.

1.4 Overview

Medical image processing is used in diagnosis of brain tumor using CAD system. There are various steps which are used in paper before getting classification of brain tumor when MRI is given as input. The system is made up of four phases. First phase is image preprocessing. The system uses brain MRI dataset which contains normal, benign, malignant brain MRI images and it will used for training purpose. After an MRI image is taken as input image it undergoes into various steps before passing it for image segmentation. First, we have to convert that image into grey scale image then we have to remove noise. Second phase is segmentation by threshold segmentation. The system uses intensity based segmentation by threshold segmentation. It is based on a threshold value. Initial the value is taken by calculating average value of the total pixel in the input image. After segmentation the system will produce image with dark background and lighting tumor area. Here the system will find the growth stage of tumor by calculating the total number of white pixels. Then in next phase features can be extracted and selected from brain MRI and it can be done PSO (PARTICLE SWARM OPTIMIZATION). Feature selection or reduction is technique in which we select certain feature out of many features which has been extracted based on the property that only these feature leads the effects on the classification process. Final phase is classification with SVM. Image classification is the process by which we categorize any pixel of image to particular class. In present case if computer will predict the tumor then it will also classify between benign and malignant also.

Chapter 2

Literature Survey

A literature survey or a literature review in a project report is that section which shows the various analyses and research made in the field of interest and the results already published, taking into account the various parameters of the project and the extent of the project. It is the most important part of report as it gives a direction in the area of our research. It helps to set a goal for the analysis - thus giving the problem statement.

2.1 Survey Papers

A literature review or narrative review is a type of review article. A literature review is a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic. Literature reviews are secondary sources, and do not report new or original experimental work. We have surveyed the following IEEE papers for our project.

In the paper **[1] Detection of Brain Tumor and Extraction of Features in MRI Images Using K-means Clustering and Morphological Operations** by, Zuliani Zulkoffli, Talha Afzal Shariff in the year 2019. In this paper, they proposed Firstly image is acquired then it Is subjected to pre-processing then k-means clustering algorithm is used to separate the images into 4 clusters of varying intensity levels, then the cluster with the tumor region is found then the tumor is extracted using morphological & region properties operation. The proposed method is divided into 3 sections: image acquisition, pre-processing and post-processing. The k-means clustering algorithm is used on the filtered image with k=4 clusters. This divides the image into four levels of grayscale intensity and separated the tumor and skull from the brain matter, background and cerebrospinal fluid. It was found that the algorithm worked well for 14 out of the 23 images in the database. This indicates a robustness of 60% in extracting only the tumor from the image.

In the paper **[2] An Intelligent System for Early Assessment and Classification of Brain Tumor** by, Keerthana T.K and Shobha Xavier, in the year 2018. This system works on medical image dataset using data mining technique. This system identifies the type of tumor by support vector machine, here we use genetic optimization algorithm for optimizing the features. System classification performance will be improved by feature weighting and SVM parameter optimization. Here we simultaneously optimize the SVM parameter and feature by genetic algorithm. Pre-processing and skull removal process will increase the performance of the system. This System will provide better accuracy with GA-SVM Classifier and will increase the decision making capacity

In the paper **[3] Brain Tumor Classification Using Hybrid Model of PSO And SVM Classifier** by, Arun Kumar, Alaknanda Ashok, M.A Ansari in the year 2018. The process which is followed in classification of brain tumor is very complicated. There are various steps for the medical image processing like image segmentation, image extraction and image classification. Different types of features are extracted from the segmented MRI images like intensity, shapes and texture based features. The feature selection approach is used to select the small subset of features from MRI image which minimize redundancy and maximize relevance to the target. In this paper, online database of MRI images containing brain tumor is taken then a machine learning model is developed by using the Particle Swarm Optimization (PSO) algorithm for feature selection and then Support Vector Machine (SVM) classifier is used to classify the type of tumor in present brain MRI images.

In the paper **[6] An efficient brain tumor detection methodology using k-means algorithm** by, J Vijay, J Subhashani, in the year 2013. In this paper they proposed K-means is the backbone of the paper methodology. k-means algorithm starts clustering by determining k initial points either at random or using some heuristic data, it then groups each image pixel under the central point it is closest to. Next it calculates new central points and repeat the former two steps. Clustering refers to the process of grouping pixels of an image such that pixels which are in the same group (cluster) are similar among them and are dissimilar to the pixels which belong to the other groups (clusters). the number of tumor pixels, K-means clustering gave a better result than the other methods. The clustering algorithm was tested with a database of 100 MRI brain images. K-means clustering achieved about 95% result.

Chapter 3

SYSTEM REQUIREMENT

All computer software needs certain hardware components or other software resources to be present on a computer. These prerequisites are known as (computer) system requirements and are often used as a guideline as opposed to an absolute rule. Most software defines two sets of system requirements: minimum and recommended. With increasing demand for higher processing power and resources in newer versions of software, system requirements tend to increase over time. Industry analysts suggest that this trend plays a bigger part in driving upgrades to existing computer systems than technological advancements. A second meaning of the term of System requirements, is a generalisation of this first definition, giving the requirements to be met in the design of a system or subsystem. Typically, an organisation starts with a set of Business requirements and then derives the System requirements from there.

3.1 Hardware requirements

- PROCESSOR: Intel® core™ i7-8500 CPU
- RAM :4 GB or above
- OS-type: 64-bit
- Hard disk: 256 GB or above
- Processor: 2.00Ghz or above

3.2 Software requirements

- Operating System: WINDOWS 10
- Tools: MATLAB 2018a
- Online dataset from BRATS-2015

3.3 Functional requirements

A **functional requirement** defines a function of a system or its component, where a function is described as a specification of behaviour between outputs and inputs.

- **MATLAB:** MATLAB (matrix laboratory) is a multi-paradigm numerical computing environment and fourth-generation programming language. A proprietary programming language developed by MathWorks, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages.

3.4 Non-functional requirements

A **non-functional requirement** is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviours.

- **Usability**: Usability is the degree to which a software can be used by specified consumers to achieve quantified objectives with effectiveness, efficiency, and satisfaction in a quantified context of use. This project can be used anywhere and by anyone but the only constraint is that the person should be trained.
- **Reliability**: Refers to whether the software can perform a failure-free operation for a specified period of time in a specified environment. The project is reliable, it can be used number of times.
- **Performance**: Ensure software applications will perform well under their expected workload. The response time of the system is good.
- **Supportability**: It refers to the ability of technical support personnel to install, configure, and monitor computer products, identify exceptions or faults, debug or isolate faults to root cause analysis, and provide hardware or software maintenance in pursuit of solving a problem and restoring the product into service. Incorporating serviceability facilitating features typically results in more efficient product maintenance and reduces operational costs and maintains business continuity. The system is cost-effective.

Chapter 4

SYSTEM DESIGN

System Design is the one which is used to give us complete information about project, by considering the physical design, system development, waterfall model, system architecture and class diagrams. It also tells how the project will work and the steps needed for implementation and system requirements needed.

System Design is one of the important roles of software development. System design is the first step to move from problem domain to the solution domain. It is the most critical factor affecting the quality of software and majorly on testing and maintenance. System Design is divided into separate phases such as top-level design and detailed design. It will identify the modules that should be in the system and specifications about the modules and how they interact with each other to produce the desired output.

4.1 System Architecture

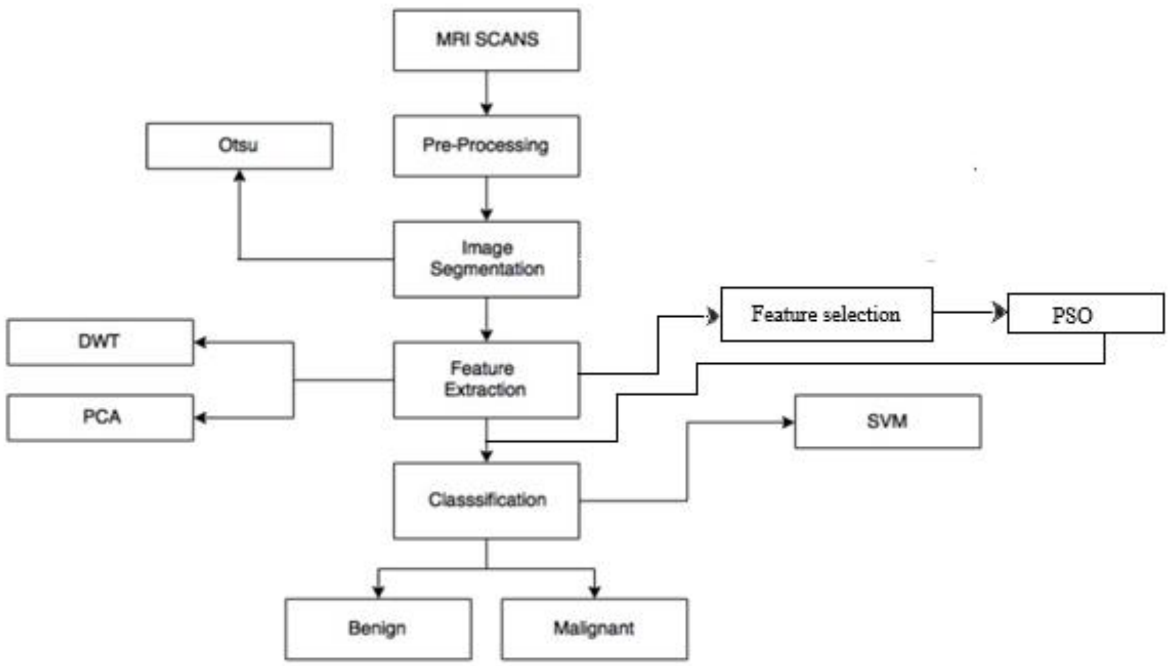

Figure 4.1: System Architecture

Magnetic resonance imaging (MRI) is a imaging technique generally used for brain to provide relevant information, based on which physician or CAD can diagnose whether patient have any tumor or not, if there is tumor detected then they can further differentiate between its types so that proper treatment can be given to that patient. Brain tumor mainly in two categories benign and malignant. Benign tumor are non-cancerous while malignant tumor are more prone to be cancerous which may be developed because of cancer at any part of the body not just only brain.

Medical image processing has provided lots of techniques which deals to automate this task in less time and with more accuracy. Feature extraction and selection, Image segmentation and Image classifications are the very important steps in medical image processing. Feature selection is even more important than feature extraction because optimal feature subset is deemed necessary to improve the image classifier performance and decrease the computation time. In traditional approaches we have Principal Component Analysis (PCA) and Linear discriminant Analysis (LDA) for linear, and Kernel. PCA has been used for nonlinear features. In compare to these type of method Evolutionary computation (EC) techniques gives better results. Out of many EC algorithms Particle swarm Optimization is one of the best techniques which can do Feature selection task efficiently with less computation time.

4.2 Methodology

The implementation of Brain Tumor Detection and Classification can be carried out in 5 main stages and can be represented in a data flow diagram as below:

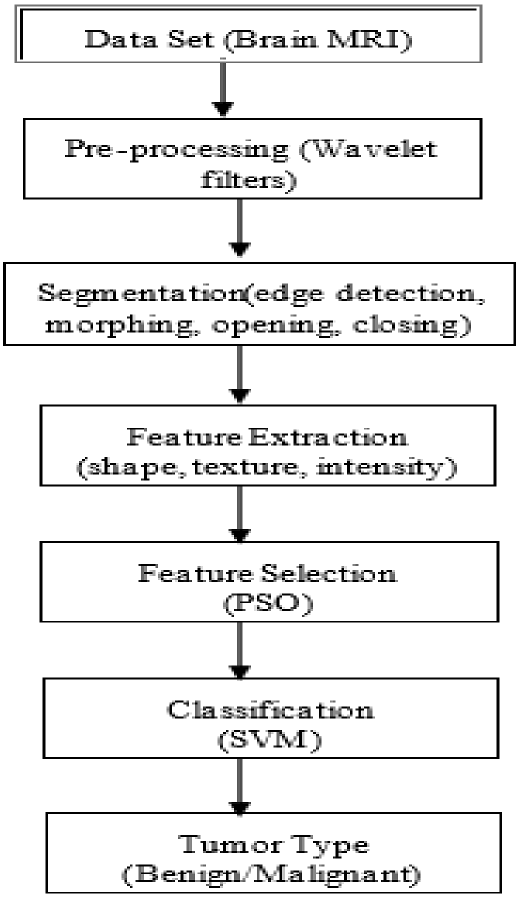

Figure 4.2: Data flow diagram

4.2.1 MRI SCANS

MRI scans of brain are taken as the input for all training and testing purposes. The images are Grayscale and are resized to 200 x200 px for processing. It is assumed that all the images given input to the system are either Benign (Non-cancerous) or Malignant (Cancerous). The images are separated into two labeled folders for training and testing purpose only. These images are further used for preprocessing.

4.2.2 Pre-processing

After an MRI image is taken as input image it undergoes into various steps before passing it for image segmentation. First, we have to convert that image into grey scale image then we have to remove noise or inhomogeneity from that grey scale image.

a) Noise Removal:

First step is conversion of input image in to gray scale image. Gray scale image consists of different levels of gray and it gives 256 possible levels from black to white. Then the unwanted noise can be removed from the gray scale image using median filtering. Median filter is non-linear filtering method and it removes each pixel by median value of its neighbouring pixel. Filter will preserve the edges and provides better image enhancement. The output of this sub-step is noise free brain MRI.

b) Skull Removal:

Skull removal is an important step to detect the brain abnormality. It will increase the processing speed. Noise free brain MRI is taken as the input to skull removal process by morphological operations. It is a non-linear operation related to morphology of features in an image which includes erosion and dilation. It will remove small scale details from input binary image.

4.2.3 SEGMENTATION

Segmentation is the most crucial step in identification of Tumor. This step extracts the tumor from the MRI scans which is then sent for extracting features. Otsu segmentation algorithms that have been implemented.

Otsu Algorithm:

In image processing, Otsu's method, named after Nobuyuki Otsu, is used to automatically perform clustering-based image thresholding. Otsu’s thresholding is a non - linear operation that converts a grayscale image into a binary image where the two levels are assigned to pixel those that are below or above the specified threshold value. Image segmentation is the process of partitioning the whole image into multiple set of pixels and it is easier for further processing. The system uses intensity-based segmentation by threshold segmentation. It is based on a threshold value. Initial the value is taken by calculating average value of the total pixel in the input image. After segmentation the system will produce image with dark background and lighting tumor area. Here the system will find the growth stage of tumor by calculating the total number of white pixels.

The algorithm assumes that the image contains two classes of pixels following bi-modal histogram: Foreground pixels and Background pixels. It then calculates the optimum threshold separating the two classes so that their combined spread (intraclass variance) is minimal, or equivalently (because the sum of pairwise squared distances is constant), so that their inter-class variance is maximal. The algorithm proceeds as follows:

1. Compute histogram and probabilities of each intensity level
2. Set up initial class probabilities and mean.
3. Step through all possible thresholds t = 1,2 ……. maximum intensity
    1. Update class probabilities and mean
    2. Compute inter class variance
4. Desired threshold corresponds to the maximum inter class variance.

4.2.4 FEATURE EXTRACTION

Feature extraction is the method of dimensionality reduction. It transforms the input data in to set of features which are used for classification purpose. Here we can extract the features using GLCM. It works based on the presence of gray levels present in the image. Entries in GLCM are done by finding relationship between two pixels. GLCM is an 8×8 square matrix and the number of rows and columns equal to the number of pixels in the input image.

From this matrix the system extract 13 features as follows: Autocorrelation, Contrast, Correlation, Energy, Dissimilarity, Entropy, Homogeneity, Sum of squares variance, Cluster shade, Cluster prominence, Information measure of correlation1, Information measure of correlation2 and Inverse difference.

Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) with Gray-Level Co-occurrence Matrix Algorithm (GLCM)

The discrete wavelet transform (DWT) is a powerful implementation of the WT using the dyadic scales and positions. The fundamentals of DWT are introduced as follows. Suppose x(t) is a square-integrable function, then the continuous WT of x(t) relative to a given wavelet ψ(t) is defined as

$$W_{\psi}(a,b)=\int_{-\infty}^{\infty}x(t)\psi_{a,b}(t)dt \tag{1}$$

where

$$\psi_{a,b}(t)=\frac{1}{\sqrt{a}}\psi\left(\frac{t-a}{b}\right) \tag{2}$$

Here, the wavelet $\psi_{a,b}(t)$ is calculated from the mother wavelet ψ(t) by translation and dilation: a is the dilation factor and b the translation parameter (both real positive numbers). There are several different kinds of wavelets which have gained popularity throughout the development of wavelet analysis. The most important wavelet is the Harr wavelet, which is the simplest one and often the preferred wavelet in a lot of applications.

GLCM is a widely used method for medical image analysis, classification. This method gives us information about relative position of two pixels with respect to each other. The GLCM is then created by counting the number of occurrences of pixel pairs at a certain distance. To compute the GLCM matrix for an image f (i, j), a distance vector d= (x, y) is defined. The (i,j)th element of the GLCM matrix P is defined as the probability that grey levels i and j occur at distance d and angle θ, then extracting texture features from GLCM matrix P. Four angles (0,45,90,135) and four distances (1,2,3,4) can be used to calculate the co-occurrence matrix.

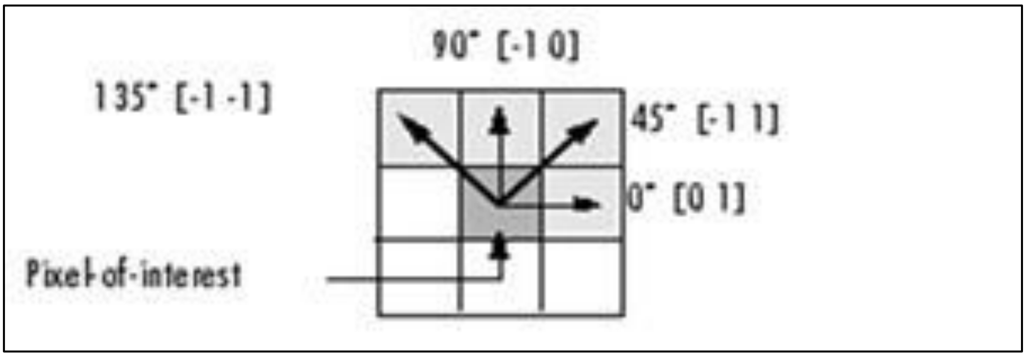

Figure 4.3: Calculation of co-occurrence matrix in GLCM

The extracted features are:

1. Correlation:

It measures the linear dependency of grey levels of neighbouring pixels. It is defined in Eq.1

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2}$$

Equation 1: Correlation

2. Contrast:

Also called the sum of Square Variance. It defers the calculation of the intensity contrast linking pixel and its neighbour over the whole image. It is defined in Eq. 2

$$Contrast = \sum_{i,j=0}^{N-1} P_{ij} (i-j)^2$$

Equation 2: Contrast

3. Energy:

It makes use for the texture that calculates orders in an image. It gives the sum of square elements in GLCM. It is defined in Eq. 3

$$Energy = \sum_{i,j=0}^{N-1} \left(P_{ij}\right)^2$$

Equation 3: Energy

4. Homogeneity:

It passes the value that calculates the tightness of distribution of the elements in the GLCM to the GLCM diagonal. It is defined in Eq. 4

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2}$$

Equation 4: Homogeneity

5. Mean:

Defined as the mean of the pixel values of the input image. It is defined in Eq. 5

$$Mean = \sum_{i,j=0}^{N-1} i(P_{i,j})$$

Equation 5: Mean

6. Standard Deviation:

It is defined as the dispersion of the pixel in consideration from the mean of the pixels of the input image. It is defined in Eq. 6

$$Standard\ Deviation = \sqrt{\boldsymbol{\sigma_i}^2}$$

Equation 6: Standard Deviation

7. Entropy:

It shows the amount of information of the image that is needed for the image compression. Entropy measures the loss of information or message in a transmitted signal and also measures the image information. It is defined in Eq. 7

$$Entropy = \sum_{i,j=0}^{N-1} -\ln(P_{ij})P_{ij}$$

Equation 7: Entropy

8. Root Mean Square (RMS):

RMS is calculated on a set of pixels by taking the square of each pixel, calculating the sum of those squares, and taking the square root. The result is scaled by the number of pixels. RMS gives an accurate measurement of the amount of noise present. It is defined in Eq. 8

$$\text{RMS noise} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \frac{\sum_{i=1}^{n} X_i}{n})^2}{n}}$$

Equation 8: Root Mean Square

9. Variance:

It is the expectation of the squared deviation of a pixel from its mean. It is defined in Eq. 9

$$Variance = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2$$ **Equation 9: Variance**

10. Inverse Difference Movement (IDM):

Inverse Difference Moment (IDM) is the local homogeneity. It is high when local gray level is uniform and inverse GLCM is high. It is defined in Eq. 10

$$\text{IDM} = \frac{\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} P_{i_j}}{1+(i-j)^2}$$ **Equation 10: Inverse Difference Movement**

4.2.5 FEATURE SELECTION

Feature selection or reduction is technique in which we select certain feature out of many features which has been extracted based on the property that only these feature leads the effects on the classification process.

Binary Particle Swarm Optimization

Particle Swarm Optimization (PSO) was originally applied to solve problems in continuous-numbers search space. However, feature selection, as in many other optimization problems, occur in discrete search spaces. Kennedy and Eberhart developed a binary version of the particle swarm algorithm (BPSO) to tackle optimization problems in discrete domains. In BPSO, the velocity is still updated in the same fashion as in the standard PSO. However, variables x_{id}, p_{id}, and p_{gd} can only have the values 0 or 1. In doing so, velocity would indicate the probability of a particle in the position vector to take the value 1.

In BPSO, the position of the current particle is updated as in Eq. (3) based on the probability value T (Vt) obtained from Eq. (4).

$$x(t+1) = \begin{cases} 1 \ if \ rand < S(v(t+1)) \\ 0 \ Otherwise \end{cases} \qquad (3)$$

where S(v(t)) is the Sigmoid function as depicted

$$S\,(v(t)) = \frac{1}{(1 + e - v(t))} \qquad (4)$$

where rand is a random number ∈ [0, 1].

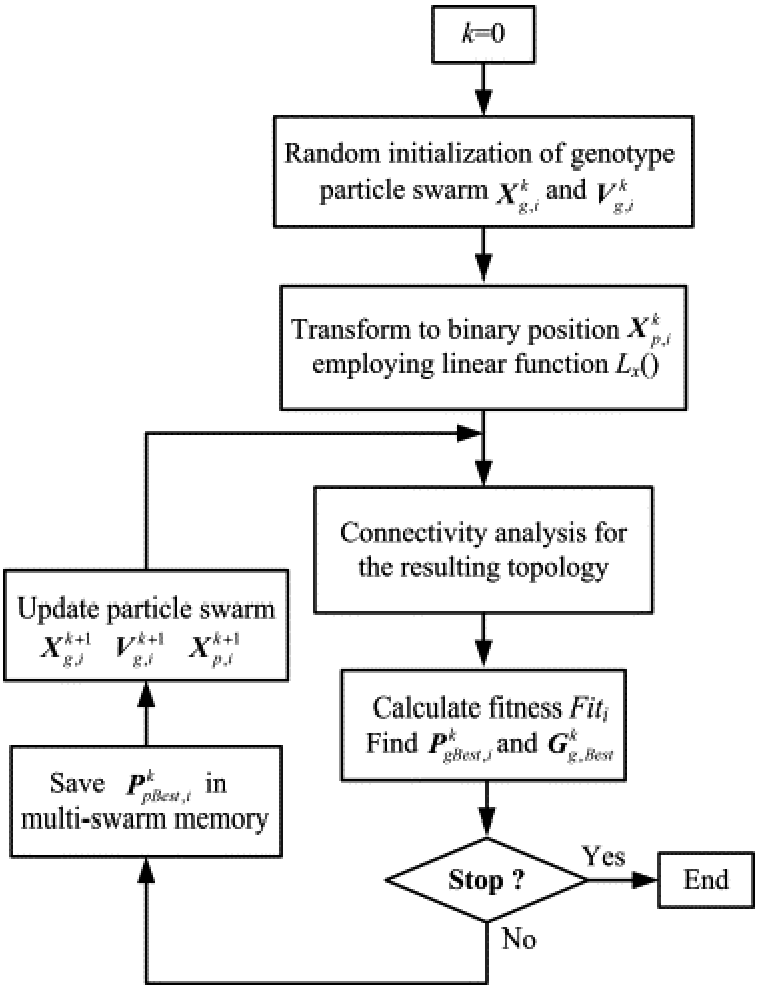

Figure 4.4: BPSO Algorithm

4.2.6 CLASSIFICATION

Support Vector Machine (SVM)

Support vector machines (SVMs) are a type of supervised learning models along with associated learning algorithms that analyse data and recognize various patterns, used for classification analysis. The basic SVM takes a set of input data and predicts, for each given input, which of two possible classes, malignant and benign forms the output, making it a non-probabilistic binary linear classifier. Now that there are set of training examples at hand, each marked as belonging to one of two categories, an SVM training algorithm constructs a model that assigns new examples into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. Newer examples are then plotted into it and then predicted to belong to a category based on which side of the gap they fall on.

More formally, a support vector machine constructs a hyper plane or set of hyper planes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyper plane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

SVM classifier is widely used for two class and multi class classification. Multiclass SVM classification is possible with multi SVM. It is a non-linear classification and it is done by doing series of binary classification based on one-versus-rest approach. Multi SVM can be used her for classification and selection of hyper plane is based on RBF (Radial Basis Function) kernel function. Here the system consists of three classes and it make three binary classification for building the multi SVM classifier.

An SVM takes a set of feature vectors as input, generates a training model after scaling, selecting and validating, and generates a training model as the output. The following figure represents the training process of an SVM:

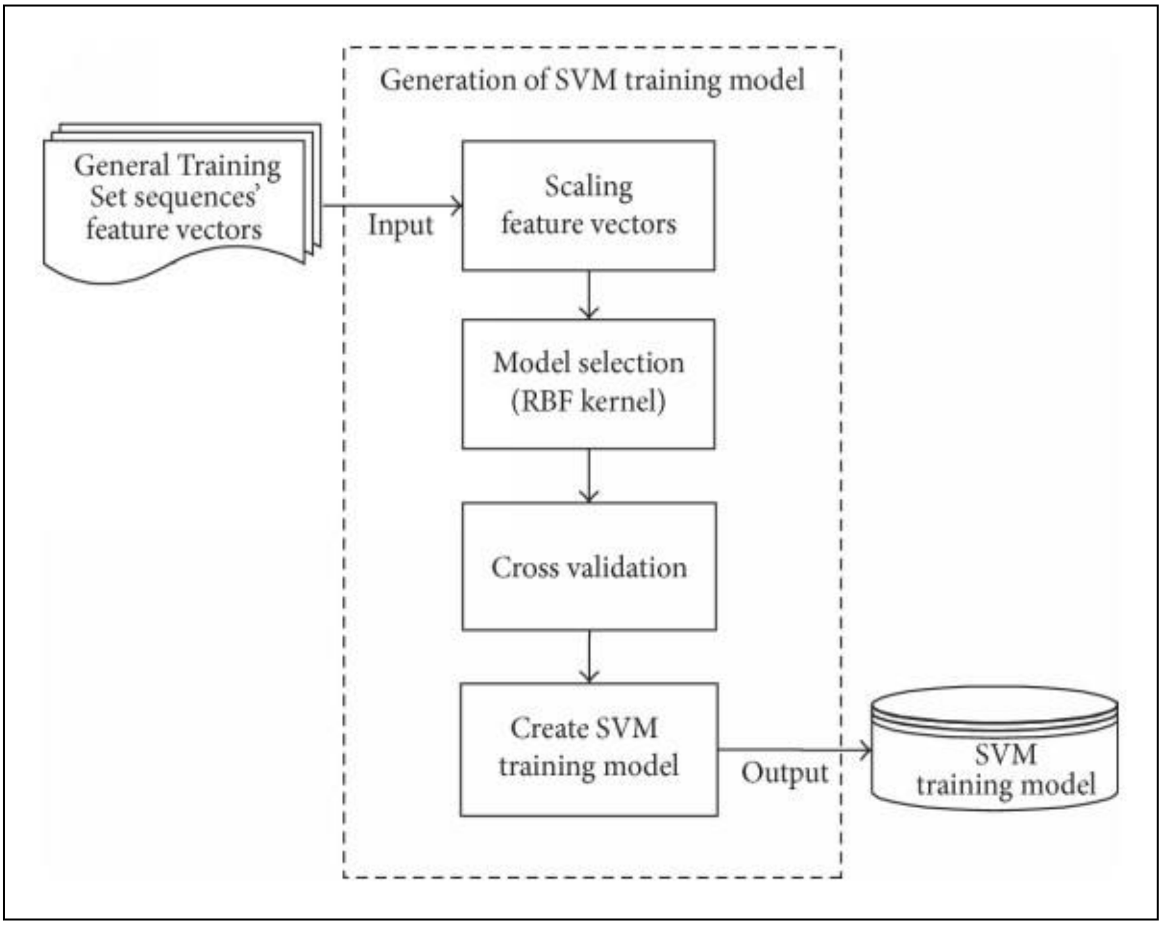

Figure 4.5: Training an SVM Model

4.2.7 TUMOR TYPE

We have two different type of tumor class benign and malignant in the given dataset. After training SVM classifier with feature selection done by PSO we compare the accuracy of classifier with the case where all the features have been used as Specification and specificity are shown. Total 6 Features out of 14 which are selected by PSO shows the both types of classified tumor benign and malignant tumor.

4.3 Sequence Diagram

A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios.

A sequence diagram shows, as parallel vertical lines (*lifelines*), different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. This allows the specification of simple runtime scenarios in a graphical manner.

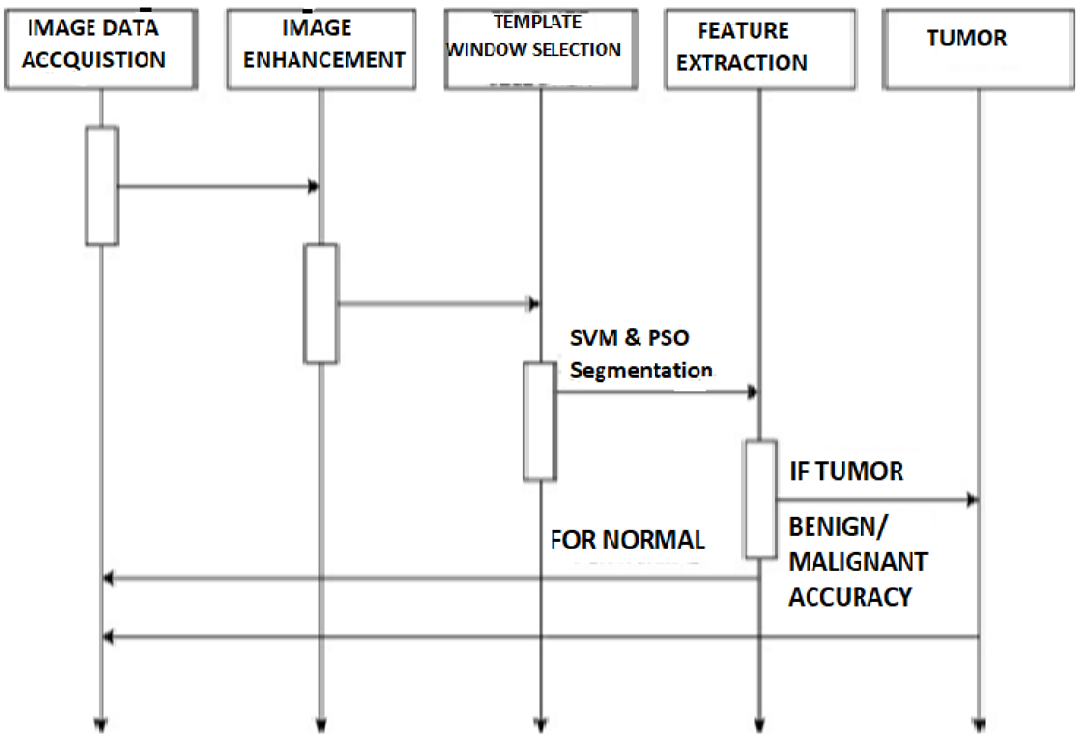

Figure 4.6: Sequence Diagram

The above sequence diagram shows the interaction between various modules in the project. The first stage involves collecting MRI image is taken as input image it undergoes various steps. Image Enhancement is the process of dividing the image into different homogenous part. In present case brain tumor portion will be required to be segmented from complete MRI of brain which will become our region of interest

In further step the process of taking initial data and then deriving relevant information in terms of key factors known as feature vectors so that these features vector can be used instead of whole data for further processing. Feature selection or reduction is technique in which we select certain feature out of many features which has been extracted based on the property that only these feature leads the effects on the classification process. Image classification is the process by which we categorize any pixel of image to particular class e.g. in any remote sensing satellite image of earth we can classify any pixel that whether it belong to water body area or land area depending upon its characteristics. Similarly, computer can also classification if it is trained about different characteristics of classes which may occurs. This training can be supervised or unsupervised. The PSO and SVM segmentation is used for the classification of tumor so that this combination of algorithm gives the result that the gives MRI Images are either cancerous or non-cancerous.

Chapter 5

IMPLEMENTATION

5.1 Data collection and preparation

The data collection will contain the predicted data and the actual perceived data, as a percept sequence. The database consists of MRI brain scans taken as input data, that are loaded to the system for training and classification, which are updated itself every time the patient database increases. Only the physician and neurosurgeons will have access rights to modify this database so that sensitive information does not leak. To avoid any latency in operation or lag in performance of the application system, it is necessary to ensure processor is not slow, or that an outdated version of MATLAB is under use currently.

```
cd images;
df =[];
for i=1: 30

    str1=int2str(i);
    str2=strcat(str1,'.jpg');
    nor=imread(str2);
    I = imresize(nor,[200,200]);
    [feat,img]=feature_extr(I);
    df=[df;feat];
end
handles.train= df;
grouptrain={'BENIGN'; 'BENIGN' ;'BENIGN'; 'BENIGN'; 'BENIGN' ;'BENIGN'
;'BENIGN' ;'BENIGN' ;'BENIGN' ;'BENIGN' ;'BENIGN' ;'BENIGN' ;'BENIGN'
;'BENIGN' ;'BENIGN' ;'MALIGNANT'; 'MALIGNANT'; 'MALIGNANT'; 'MALIGNANT';
'MALIGNANT';'MALIGNANT';'MALIGNANT';'MALIGNANT';'MALIGNANT';'MALIGNANT';
'MALIGNANT';'MALIGNANT';'MALIGNANT';'MALIGNANT';'MALIGNANT';};
xdata =df;
label=grouptrain;
N=30; T=50; c1=2; c2=2; Vmax=10; Wmax=0.9; Wmin=0.4;
[sFeat,Sf,Nf,curve]=jBPSO(xdata,label,N,T,c1,c2,Wmax,Wmin,Vmax);
df=sFeat;
handles.train=df
|
handles.index=Sf
handles.group=grouptrain;
cd ..;
f=msgbox('Data Prepared');
```

Figure 5.1: Data Preparation Code

5.2 Getting Image Data

Image should be loaded for testing. After loading image, the pre-processing is done. Pre-processing are used to improve the detection of the suspicious region from Magnetic Resonance Image (MRI). This section presents the gradient - based image enhancement method for brain MRI images which is based on the first derivative and local statistics. Feature extraction is process of extracting quantitative information from an image such as colour features, texture, shape and contrast. Here, we have used discrete wavelet transform (DWT) for extracting wavelet coefficients and gray-level co-occurrence matrix (GLCM) for statistical feature extraction.

```
[FileName,PathName] = uigetfile('*.jpg;*.png;*.bmp','Pick an MRI Image');
if isequal(FileName,0)||isequal(PathName,0)
   warndlg('User Press Cancel');
else
  P = imread([PathName,FileName]);
  P = imresize(P,[200,200]);
  axes(handles.input);
  imshow(P);
  [feat,img]=feature_extr(P);

  handles.image=img;

  Sf=handles.index
  featl=[];
  for i=1:length(Sf)
       featl(i)=feat(Sf(i));
  end
  handles.imgdata= featl;
 guidata(hObject, handles);
```

Figure 5.2: Load Image

```
function [feat,img]=feature_extr(I)
  J = rgb2gray(I);
  level = graythresh(I);
  img = im2bw(J,0.7);
  img = bwareaopen(img,80);
  signal1 = img(:,:);
  [cA1,cH1,cV1,cD1] = dwt2(signal1,'db4');
  [cA2,cH2,cV2,cD2] = dwt2(cA1,'db4');
  [cA3,cH3,cV3,cD3] = dwt2(cA2,'db4');
  DWT_feat = [cA3,cH3,cV3,cD3];
  G = pca(DWT_feat);
  g = graycomatrix(G);
  stats = graycoprops(g,'Contrast Correlation Energy Homogeneity');
  Contrast = stats.Contrast;
  Correlation = stats.Correlation;
  Energy = stats.Energy;
  Homogeneity = stats.Homogeneity;
  Mean = mean2(G);
  Standard_Deviation = std2(G);
  Entropy = entropy(G);
  RMS = mean2(rms(G));
  Variance = mean2(var(double(G)));
  a = sum(double(G(:)));
  Smoothness = 1-(1/(1+a));
  Kurtosis = kurtosis(double(G(:)));
  Skewness = skewness(double(G(:)));
  m = size(G,1);
  n = size(G,2);
  in_diff = 0;
  for i = 1:m
    for j = 1:n
        temp = G(i,j)./(1+(i-j).^2);
        in_diff = in_diff+temp;
    end
  end
  IDM = double(in_diff);
  feat = [Contrast,Correlation,Energy,Homogeneity, Mean,
Standard_Deviation, Entropy, RMS, Variance, Smoothness, Kurtosis, Skewness,
IDM];
```

Figure 5.3: Feature Extraction

5.3 Image Segmentation and Training Model

Image segmentation is the primary step and the most critical tasks of image analysis. Its purpose is that of extracting from an image by means of image segmentation. The mechanization of medical image segmentation has established wide application in diverse areas such as verdict for patients, treatment management planning, and computer - integrated surgery. SVM classification technique applied on different fields such as face recognition, text categorization, cancer diagnosis, glaucoma diagnosis, microarray gene expression data analysis. SVM utilizes binary classification of brain MR image as tumor affected. SVM divides the given data into decision surface, (i.e. a hyperplane) which divides the data into two classes. The prime objective of SVM is to maximize the margins between two classes of the hyper-plane. Dimensionality reduction and precise feature set given as input to the SVM on the duration of training part as well as during the testing part.

```
feat=handles.imgdata;


xdata = handles.train;
group =handles.group;
svmStruct1 = fitcsvm(xdata,group,'KernelFunction', 'RBF');
species = predict(svmStruct1,feat);



 if strcmpi(species,'MALIGNANT')
     set(handles.type,'string','MALIGNANT');
     disp(' Malignant Tumor ');
 else
     set(handles.type,'string','BENIGN');
     disp(' Benign Tumor ');
 end

  axes(handles.segmented);
  imshow(handles.image);
  guidata(hObject, handles);
```

Figure 5.4: Training Model

Chapter 6

TESTING

System testing is any activity aimed at evaluating an attribute or capability of a program or system and determining that it meets its required results. Although crucial to software quality and widely deployed by programmers and testers, software testing still remains an art, due to limited understanding of the principles of software. The difficulty in software testing stems from the complexity of software: we cannot completely test a program with moderate complexity.

Testing is more than just debugging. Software testing is an investigation conducted to provide stakeholders with information about the quality of the product or service under test. Test techniques include, but are not limited to the process of executing a program or application with the intent of finding software bugs (errors or other defects). The purpose of testing can be quality assurance, verification and validation, or reliability estimation.

Software testing can be stated as the process of validating and verifying that a computer program/application/product:

- Meets the requirements that guided its design and development
- Works as expected
- Can be implemented with the same characteristics
- Satisfies the needs of stakeholders.

Testing can be used as a generic metric as well. Software testing is usually a trade off between budget, time and quality. Software testing, depending on the testing method employed, can be implemented at any time in the software development process. Traditionally most of the test effort occurs after the requirements have been defined and the coding process has been completed, but in the agile approaches most of the test effort is on- going. As such, the methodology of the test is governed by the chosen software development methodology.

6.1 Stages in Implementation of Testing

6.1.1 Unit Testing

Unit testing verification efforts on the smallest unit of software design, module. This is known as Module Testing. The modules are tested separately. This testing is carried out during programming stage itself. In these testing steps, each module is found to be working satisfactorily as regard to the expected output from the module.

6.1.2 Integration Testing

Integration testing is a systematic technique for constructing tests to uncover error associated within the interface. In the project, all the modules are combined and then the entire programmer is tested as a whole. In the integration-testing step, all the error uncovered is corrected for the next testing steps.

6.1.3 System Testing

Once individual module testing completed, modules are assembled to perform as a system. Then the top down testing, which begins from upper level to lower level module testing, has to be done to check whether the entire system is performing satisfactorily.

After unit and integration testing are over then the system as whole is tested. There are two general strategies for system testing.

They are:

- Code Testing
- Specification Testing

6.1.4 Acceptance Testing

When the system has no measure problem with its accuracy, the system passes through a final acceptance test. This test confirms that the system needs the original goal, Objective and requirements established during analysis. If the system fulfils all the requirements, it is finally acceptable and ready for operation.

6.1.5 Test Plan

When the system has no measure problem with its accuracy, the system passes through a final acceptance test. This test confirms that the system needs the original goal, Objective and requirements established during analysis. If the system fulfils all the requirements, it is finally acceptable and ready for operation.

6.2 Testing Objectives

These are several guidelines that can save as testing purpose, they are:

- Testing is a procedure of running program with intent of finding errors.
- A better test case is one that has a high possibility of finding an undiscovered error.

Testing procedures for the project is done in the following sequence

- System testing is done for checking the server name of the machines being connected between the consumer and decision-making.
- The product information on condition that by the company to the executive is tested against the validation with the centralized data store.
- System testing is also done for verifying the executive availability to connected to the server.
- The server name authentication is checked and availability to the customer.
- Following are the some of the testing methods applied to this effective project.

6.3 Unit Testing:

Each module will be tested to check whether it gives the desired output. The segmentation algorithms will be executed and the output images will be checked for the segmented tumor. In case of feature extraction, it will be ensured that the features are stored in an excel sheet and don’t result in NaN values. The classification algorithms will be trained on the datasets and then tested on the training datasets to obtain the accuracy of the classifier.

Test Case:	Average Accuracy
SVM Accuracy	**87.25**

Table 6.1: Unit Test Cases

6.4 Integration Testing:

All the possible combinations of the individual modules will be integrated and tested to ensure that it produces correct output. Each segmentation module will be combined with feature extraction module and then combined with classification module for testing.

Test Case:	Average Accuracy
SVM+PSO Accuracy	**92.15**

Table 6.2: Integration test cases

6.5 System Testing:

The GUI will be tested in system testing. The input will be given to the system through the GUI and the outputs will be checked.

Test Id	Test Case	Expected Output	Actual Output	Pass/Fail
1	User Screen	After running the code, a user interface screen is to be displayed.	A user interface screen is displayed.	Pass
2	Prepare Data.	After clicking on “Prepare Data” button, SVM model for images using PSO algorithm is trained and pop up box of data prepared is to be displayed	After clicking on “Prepare Data” button, SVM model with feature extraction and segmentation is trained for images and pop up box of data prepared is displayed.	Pass

3	Load Image	After clicking on "Load Image" button, a folder is opened to select an image from tab and click ok.	A folder is opened to select an image and image loaded in user screen.	Pass
4	Detect Tumor	After clicking on "Detect Tumor" button, a segmented imaged and type of tumor to be displayed.	After clicking detect tumor button, a segmented image and tumor type is displayed.	Pass
5	Accuracy	After clicking on "Accuracy" button, a popup window showing evaluating 100 iteration is displayed and accuracy is shown on text box.	Accuracy percentage is displayed in text box.	Pass

Table 6.3: System Testing

Chapter 7

CONCLUSION AND FUTURE ENHANCEMENTS

7.1 Conclusion

Abnormal growth of tissue in the brain which affect the normal functioning of the brain is considered a brain tumor. The main goal of medical image processing is to identify accurate and meaningful information using algorithms with minimum error possible. Brain tumor detection and classification through MRI images can be categorized into four different sections: pre-processing, image segmentation, feature extraction and image classification. Various segmentation methodologies are explored in the project. It can be concluded that the algorithms and the parameters used in the proposed system are all meant to increase the efficiency of the system by achieving better results.

The boundary approach and the edge-based approach for segmentation are very common but the region growing approach gives better results. It is found that the particle swarm optimization algorithm gives the most accurately segmented tumor. Features extracted by using GLCM method help to increase efficiency as minute details of tumor by using various features can be extracted. Of the various classification methods studied, it was experimentally found that the SVM give the best classification accuracy. Accuracy and reliability are of utmost importance in tumor diagnosis, as a patient’s life depends on the results predicted by the system. Thus, the proposed methodology helps in increasing the accuracy and obtaining the desired result.

7.2 Future Enhancements

Encouraged by these results, future work will involve the improvement of classification result and overall accuracy. The number of output classes can also be increased if more data is available. With a more extensive and diverse dataset, the overall classification accuracy can be dramatically increased.

Another approach to improve the result would be to increase the number of hidden layers of the neural network. By increasing the number of hidden layers, the weights will be better adjusted and thus increase the classification. One can also do fine tuning and transfer learning approaches to better tune the model on the basis of already trained models.

SNAPSHOTS

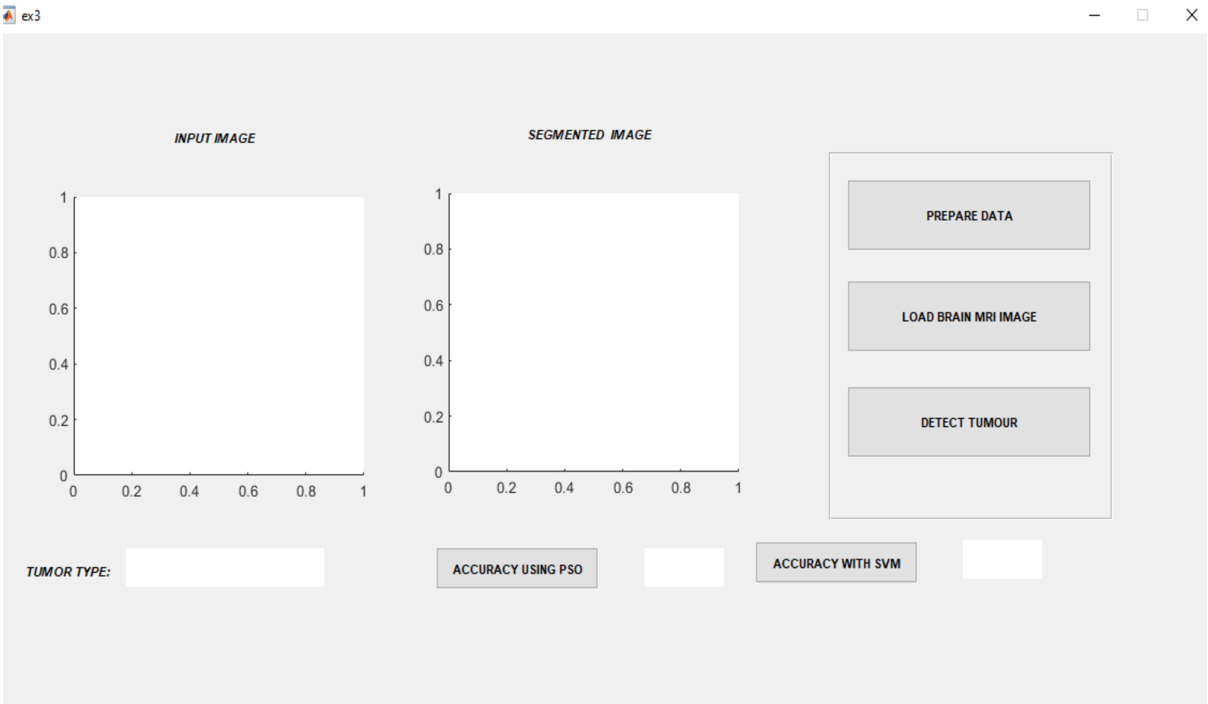

Figure 7.1: Graphical User Interface

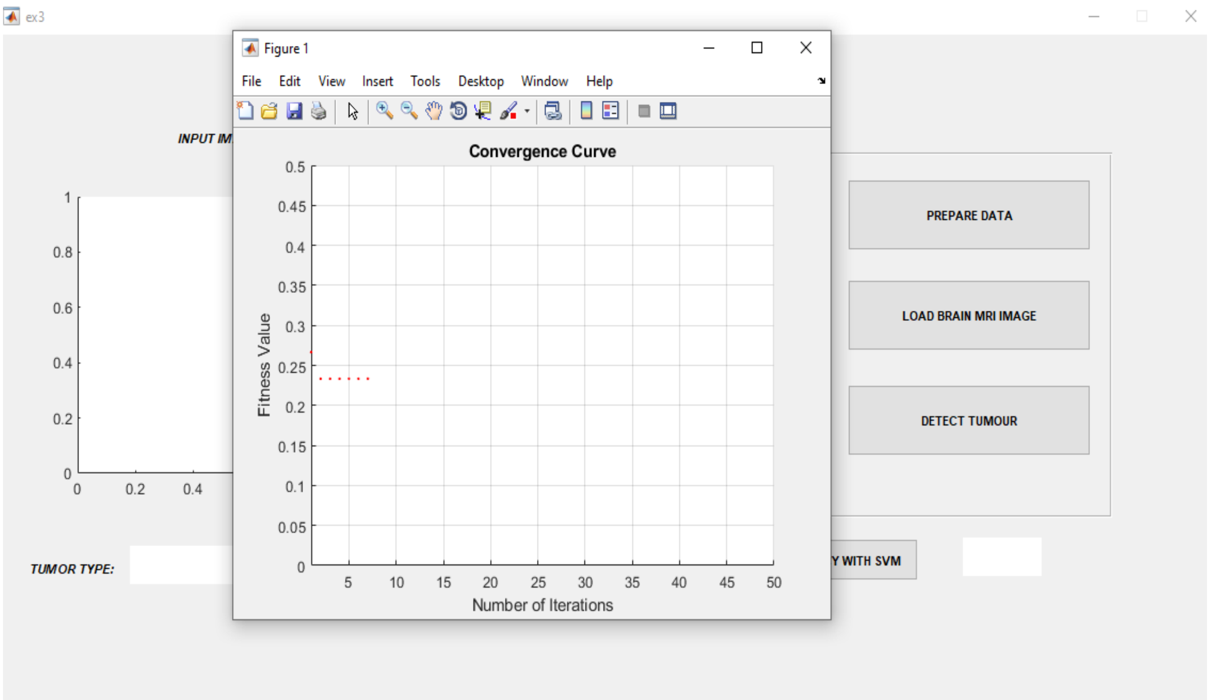

Figure 7.2: Training PSO model

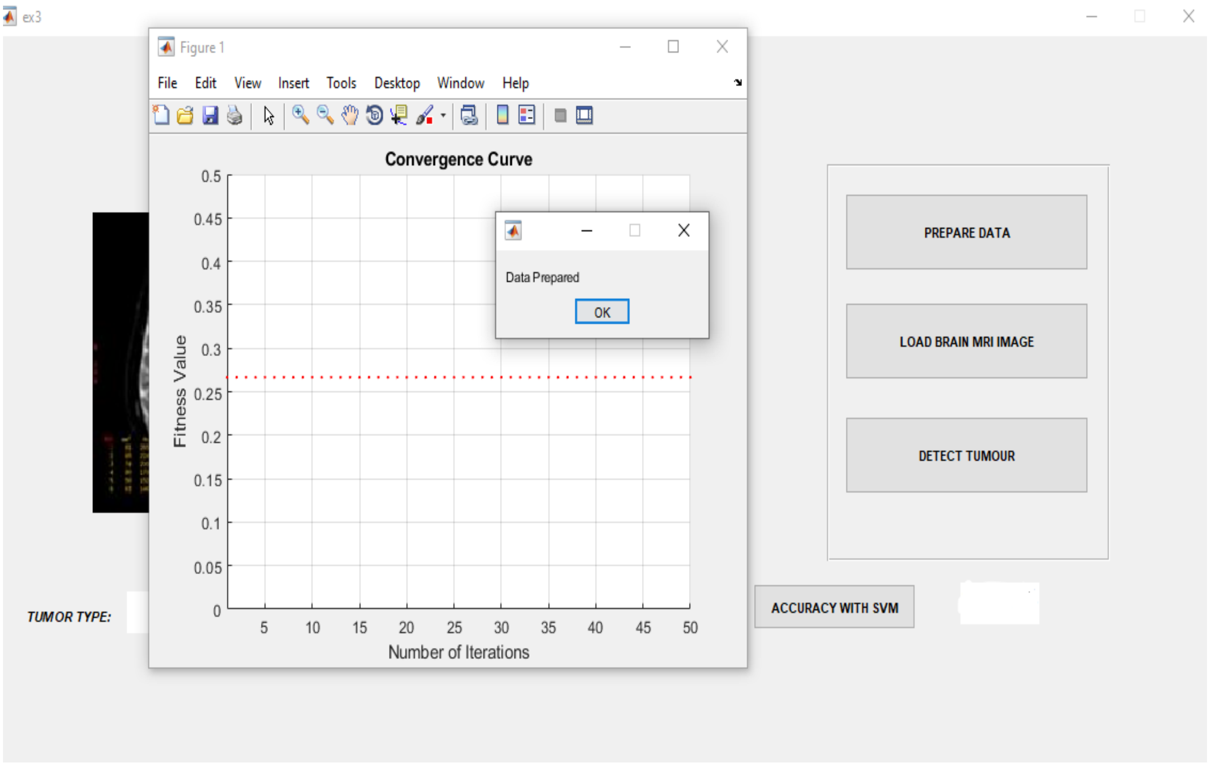

Figure 7.3: Features selected

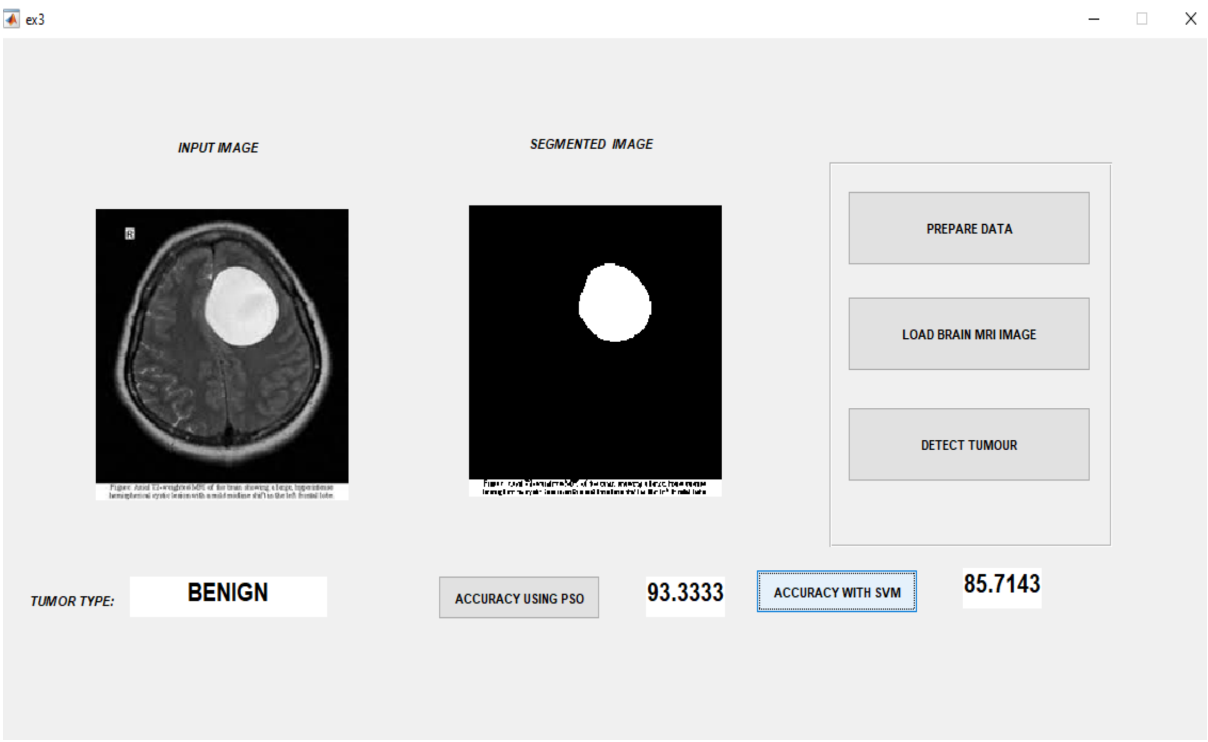

Figure 7.4: Classified output (Benign)

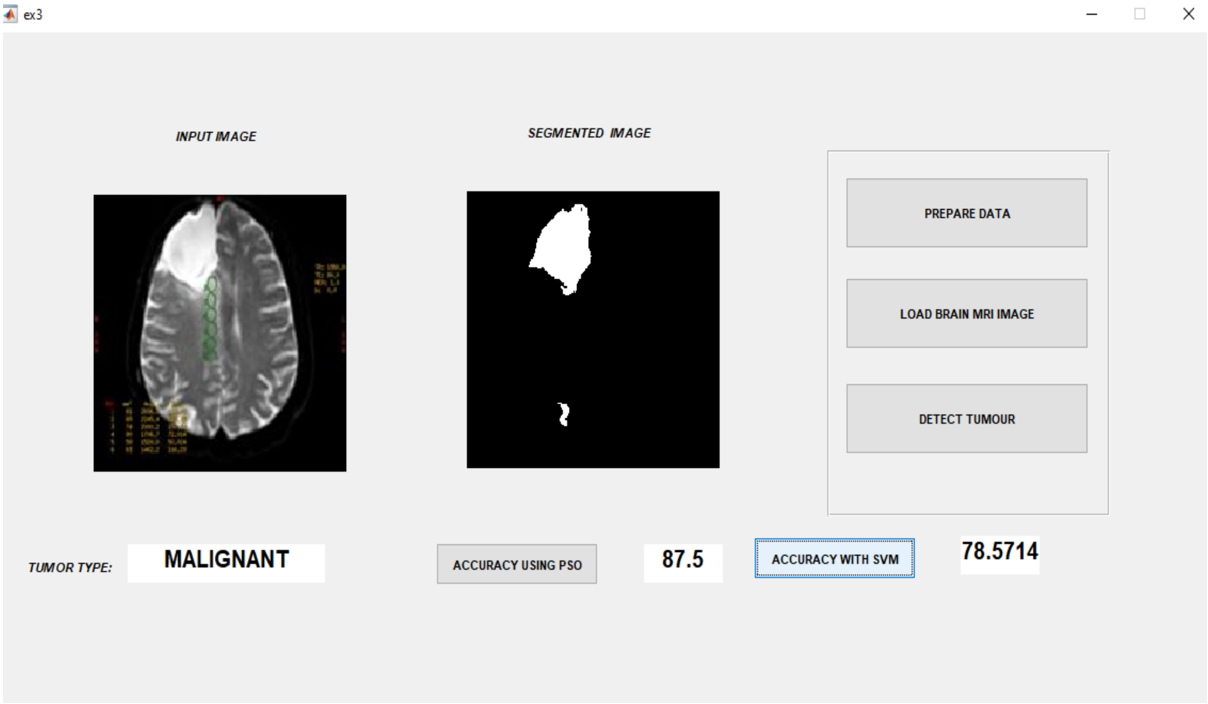

Figure 7.5: Classified output (Benign)

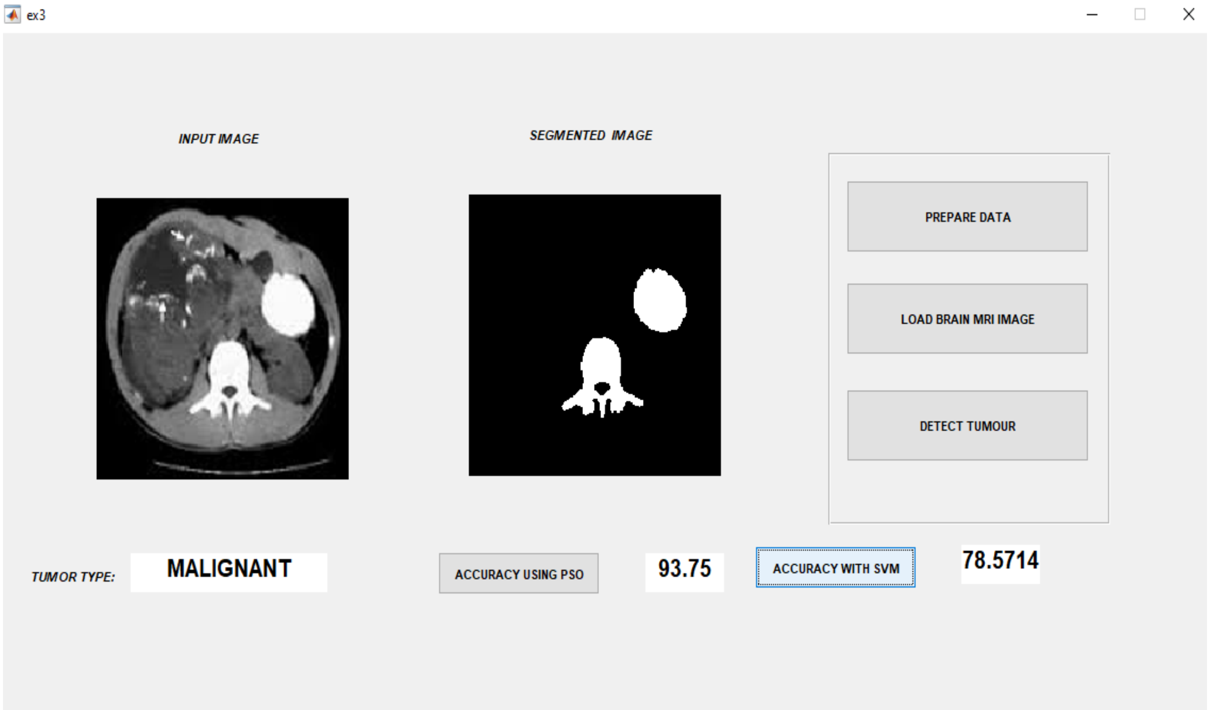

Figure 7.6: Classified output (Malignant)

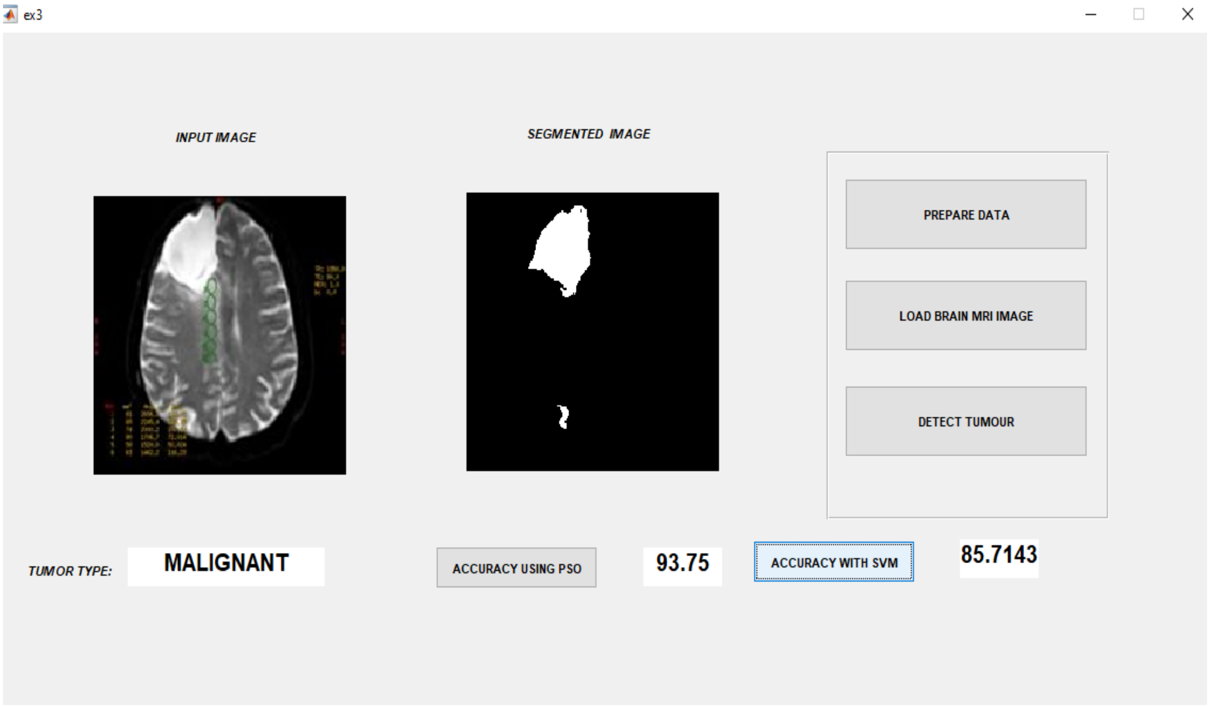

Figure 7.7: Classified output (Malignant)

REFERENCES

[1] Zuliani Zulkoffli and Talha Afzal Shariff "Detection of Brain Tumor and Extraction of Features in MRI Images Using K-means Clustering and Morphological Operations" IEEE 2019

[2] Keerthana T K and Sobha xavier "An Intelligent System for Early Assessment and Classification of Brain Tumor" IEEE 2018

[3] Arun Kumar, Alaknanda Ashok and M. A. Ansari "Brain Tumor Classification Using Hybrid Model of PSO And SVM Classifier" IEEE 2018

[4] Majid, M. A., et al. "A comparative study on the application of binary particle swarm optimization and binary gravitational search algorithm in feature selection for automatic classification of brain tumor MRI." Journal of Fundamental and Applied Sciences 10.2S (2018)

[5] Taie, Shereen A., and Wafaa Ghonaim. "CSO-based algorithm with support vector machine for brain tumor's disease diagnosis." Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on. IEEE, 2017.

[6] J Vijay and J Subhashani "An efficient brain tumor detection methodology using k-means algorithm" IEEE 2013