

Image Processing for Brain Tumor Segmentation and Classification

Miss. Roshani S. Thombare¹, Mr. Girish D. Bonde²

¹M.Tech Student ²Assistant Professor, Department of Electronics and Telecommunication Engineering, J.T.M. College of Engineering, Faizpur, India

Abstract - In this paper, we have fostered a new transpire for automatic dissection of brain tumors in MR images. The brain tumor segmentation is performed using watershed transform. The propositioned method for brain tumor classification entails of four stages explicitly pre-processing, DWT feature extraction, principal component analysis for feature reduction, feature extraction and classification. In pre-processing adaptive histogram equalization is used for noise reduction and to render the image apposite for extracting the features. In the second stage, DWT features are extracted from the image. In the third stage, Principal Component Analysis (PCA) is depleted to demote the dimensionality of the feature space which results in a more efficient and accurate classification. In the feature extraction stage different texture and statistical features such as contrast, correlation, energy, homogeneity, kurtosis, energy, and entropy are extracted. Finally, in the classification stage, we have evaluated the performance of SVM, KNN and neural network classifiers.

Key Words: Tumor Segmentation, DWT, PCA, SVM Classification, KNN, Neural Networks

1. INTRODUCTION

The Brain is one of the most vital and complex structure in a human body. Brain can be affected from problem which cause change in its normal structure and its normal behavior. This problem is known as brain tumor. Different image techniques like Magnetic Resonance Imaging (MRI) are used by inner structure of the brain and diagnose the tumor. A tumor is an any mass cause abnormal and uncontrolled growth of cell. There are three type of Brain tumors have been classified namely Giloma Tumors, Malignant Tumors, Pituitary Brain Tumor.

Benign tumor is having their boundaries or the edges. They do not spread over the other parts of the body. Malignant tumor is considered to be the most serious one and they develop rapidly. They affect the various necessary organs which even lead to the death.

Brain Tumor is includes Computed Tomography (CT) scan, Magnetic Resonance Imaging (MRI) scan. Classification is the process in which the brain image is classified as Benign or Malignant. In this process, different features are extracted. The classification process is very important. The segmentation process is to perform to the system to detect accurate tumor in image.

Segmentation is used to dividing the image to its parts sharing identical properties like color, texture, contrast and boundaries. Analysis of segmentation of a large set of images and comparisons of these segmentations between relevant

subgroups of images. Adaptive histogram thresholding is implemented for segmentation.

The proposed system suggests using three stages namely Feature Extraction, Feature Reduction and Classification. This approach extracts features using Discrete Wavelet Transform and Feature Reduction is done using Principal Component Analysis (PCA). Different texture and statistical features are extracted. The result is used in Support Vector Machine (SVM) for tumor classification as Benign or Malignant or Pituitary. Wavelet transform is an executive tool for feature extraction from MR brain images. The following step as follows:

Step 1: Input of MRI Image.

Step 2: Generate Preprocessing image.

Step 3: Segmentation is locate the object and pixel image.

Step 4: Feature extraction through DWT

Step 5: Feature reduction through PCA.

Step 6: Texture and statistical features extraction

Step 6: Finally SVM classifier is used to classification from MR Image. It is shown in Figure 1.

This paper is structured as follows: Section 2 relates with related work done in the brain tumor segmentation and classification. Section 3 gives the step by step method followed for classification and segmentation. Section 4 analyses and discusses the results obtained and Section 5 concludes the work done.

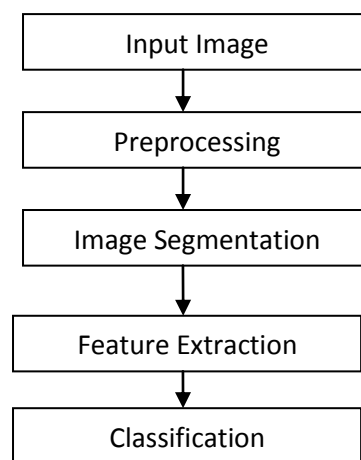


Figure 1 Block Diagram of Proposed System

2. RELATED WORK

Saleck et al [4] introduced a new approach using FCM algorithm, in order to extract the mass from region-of

interested (ROI). The proposed method aims at avoiding problematic of the estimation of the cluster number in FCM by selecting as input data, the set of pixels which are able to provide us the information required to perform the mass segmentation by fixing two clusters only. The Gray Level Occurrence Matrix (GLCM) is used to extract the texture features for getting the optimal threshold, which separate between selected set and the other sets of the pixels that influences on the mass boundary accuracy. The performance of the proposed method is evaluated by specificity, sensitivity and accuracy.

Bhima and Jagan [6] demonstrated the superior accuracy for brain tumor detection in compared to the presented methodologies. Also the major identified bottleneck of the recent research outcomes are limited to detection of brain tumor and the overall analyses of internal structure of the brain is mostly ignored being one of the most important factor for disorder detection.

Vrji and Jayakumari [7] improved brain tumor approximation after a manual segmentation procedure and 2D & 3D visualization for surgical planning and assessing tumor. The tumor identification, the investigations has been made for the potential use of MRI data for improving brain tumor shape approximation. In Preprocessing and Enhancement stage, medical image is converted into standard formatted image. Segmentation subdivides an image into its constituent regions or objects.

Rashid et al [8] investigated the chosen brain MRI image and a method is targeted for more clear view of the location attacked by tumor. An MRI abnormal brain images as input in the introduced method, Anisotropic filtering for noise removal, SVM classifier for segmentation and morphological operations for separating the affected area from normal one are the key stages if the presented method. Attaining clear MRI images of the brain are the base of this method. The classification of the intensities of the pixels on the filtered image identifies the tumor.

Sudharani et al [9] the present paper proposed the classification and identification scores of brain tumor by using k-NN algorithm which is based on training of k. In this work Manhattan metric has applied and calculated the distance of the classifier. The algorithm has been implemented using the Lab View.

Vidyarthi, A., &Mittal [10] proposed a hybrid model which identifies the region of interest using fused results of threshold segmentation and morphological operations. Initially, an abnormal brain MR image is processed with Otsu threshold based segmentation and morphological operations like erosion. Further, both the segmented resultant images are fused with the original MR image to preserve the background and correctly identification of the tumor region.

Li et al [11] proposed framework employs local binary patterns (LBPs) to extract local image features, such as edges, corners, and spots. Two levels of fusion (i.e., feature-level fusion and decision-level fusion) are applied to the

extracted LBP features along with global Gabor features and original spectral features, where feature-level fusion involves concatenation of multiple features before the pattern classification process while decision-level fusion performs on probability outputs of each individual classification pipeline and soft-decision fusion rule is adopted to merge results from the classifier ensemble. Moreover, the efficient extreme learning machine with a very simple structure is employed as the classifier.

Dhanaseely et al [12] presented and investigated two different architectures in this work. The cascade architecture (CASNN) and feed forward neural architecture (FFNN) are investigated. The feature extraction is performed using principal component analysis (PCA) as it reduces the computational burden.

For a given database the features are extracted using PCA. The Olivetti Research Lab (ORL) database is used. The extracted features are divided into training set and testing set. The training data set is used to train both the neural network architectures. Both are tested extensively using testing data.

Liu and Liu [13] proposed an algorithm of HV microscopic image feature extraction and recognition using gray level co-occurrence matrix (GLCM) in order to effectively extract the feature information of human viruses (HV) microscopic images. Firstly, 20 pieces of microscopic images of human virus are obtained by using GLCM, and then the four texture feature parameters, entropy, energy inertia moment and correlation are extracted utilizing the GLCM, and then HV image recognition is carried out.

Parveen and Singh [14] proposed a new hybrid technique based on the support vector machine (SVM) and fuzzy c-means for brain tumor classification.

3. BRAIN TUMOR SEGMENTATION AND CLASSIFICATION

In this study we have used brain tumor dataset containing 3064 T1-weighted contrast-enhanced images from 233 patients with three kinds of brain tumor: meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices) Figure 2 shows brain tumor images from the dataset.

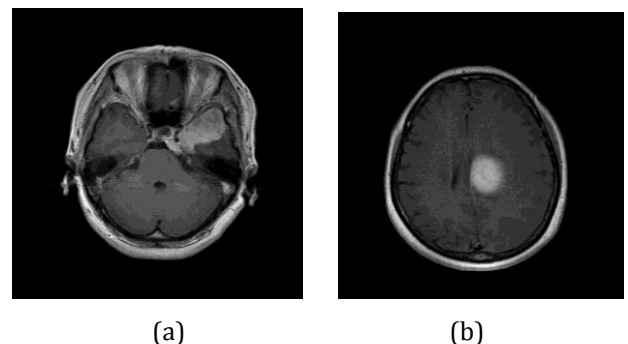


Figure 2 MRI scan image with brain tumor

A. Brain Tumor Segmentation

Segmentation using the watershed technique works well if

the foreground objects and background regions are identified or marked. It is a simple, perceptive method and it is fast. Watershed Division Method extracts seeds, which indicate the nearness of items or foundation at particular areas. There Marker areas are then set to be the local minima inside the topological surface and the watershed calculation is connected. The advantage of watershed division is that it creates an extraordinary answer for a specific picture input. Figure 3 shows the brain tumor segmentation using watershed transform.

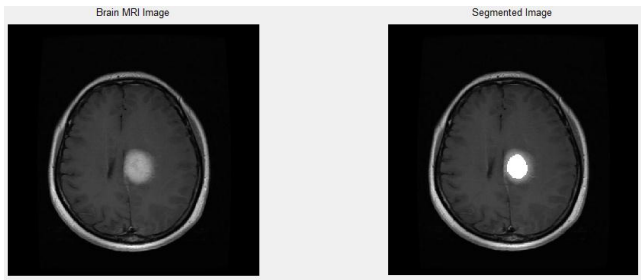


Figure 3 Brain Tumor Segmentation

B. Brain Tumor Classification

There are four steps of Brain Tumor Classification:

- a) Segmentation,
- b) Feature Extraction by Using DWT,
- c) Feature Reduction by Using PCA,
- d) Statistical and Texture Features,
- e) Tumor detection and classification.

a) Threshold segmentation

Image segmentation is typically used to locate objects and boundaries such as lines, curves, etc., in images. Each of the pixels in a region is have some similar characteristic or computed property, such as color, intensity, or texture.

Thresholding technique segments the MR images by a binary partitioning of the image intensity. The segmentation is based on thresholds. Usually MRI images have non-isotropic vowel sizes in order to get isotropic image, compute the threshold of different tissues gray levels and the image histogram as a probability density function of the Image.

b) The Discrete Wavelet Transform (DWT) became a very versatile signal processing tool proposed the multi-resolution representation of signals based on wavelet decomposition. Feature Extraction is used to extract the wavelet coefficient from MR images. DWT is a technique used to extract features of each image from brain MRI, which extracts maximum highlighting pixels present in images to progress results. The main advantage of wavelets is that they provide localized frequency information of classification. In this various statistical features are calculated. There are 1) Mean, 2) Standard Deviation, 3) Entropy.

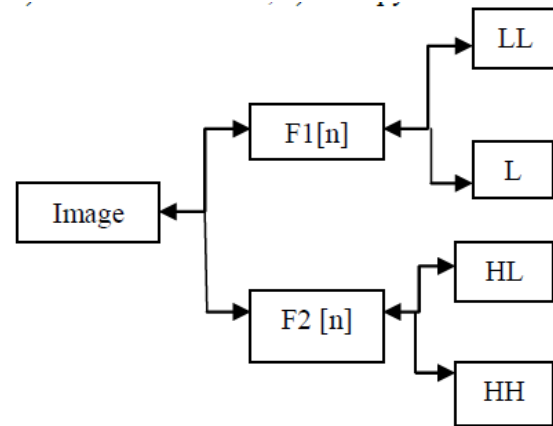


Figure 4 DWT schematically

It is shown in Figure 4. The input image is process by $F1 [n]$ and $F2 [n]$ filters which is the row representation of the original image. A result of this transform there are 4 sub band (LL, LH, HH, HL) images at each scale.

c) Feature reduction by using principal component analysis (pca)

The Principal Component Analysis well- accredited tools for transmuted the presented input features into a new lower dimensional feature space. In proposed method is extract features from brain MR images using PCA. It is commonly used forms of dimensional reduction. The main purpose of using PCA approach is to reduce the dimensionality of the wavelet coefficients.

Algorithm

Step1: Input Image $I [x, y]$.

Step2: Arrange in two column vector.

Step3: Calculate Empirical means values are subtracted.

Step4: Find the covariance Matrix.

Step5: Compute Eigen vector and Eigen values of result vector.

d) Statistical and Texture Features

After the segmentation is performed on tumor region, the segmented modules are used for feature extraction. Feature extraction is one of the most important steps in this system. A feature is a significant piece of information extracted from an image which provides more detailed understanding of the image. A feature is defined as a function of one or more measurements, the values of some quantifiable property of an object, computed so that it quantifies some significant characteristics of the object.

i. GLCM Features

A gray level co-occurrence matrix is a second order statistical measure. GLCM is the gray-level co-occurrence matrix (GLCM), also known as the gray level spatial dependence matrix. The Gray-Level Co-occurrence Matrix

(GLCM) is based on the extraction of a gray-scale image. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. Statistical parameters calculated from GLCM values are as follows:

The matlab function `glcm = graycomatrix(I)` creates a gray-level co-occurrence matrix (GLCM) from image I. `graycomatrix` creates the GLCM by calculating how often a pixel with gray-level (grayscale intensity) value i occurs horizontally adjacent to a pixel with the value j. (You can specify other pixel spatial relationships using the 'Offsets' parameter -- see Parameters.) Each element (i, j) in `glcm` specifies the number of times that the pixel with value i occurred horizontally adjacent to a pixel with value j.

The matlab function `graycomatrix` calculates the GLCM from a scaled version of the image. By default, if I is a binary image, `graycomatrix` scales the image to two gray-levels. If I is an intensity image, `graycomatrix` scales the image to eight gray-levels.

`stats = graycoprops(glcm, properties)` calculates the statistics specified in `properties` from the gray-level co-occurrence matrix `glcm`. `glcm` is an m-by-n-by-p array of valid gray-level co-occurrence matrices. If `glcm` is an array of GLCMs, `stats` is an array of statistics for each `glcm`. `graycoprops` normalizes the gray-level co-occurrence matrix (GLCM) so that the sum of its elements is equal to 1.

Each element (r,c) in the normalized GLCM is the joint probability occurrence of pixel pairs with a defined spatial relationship having gray level values r and c in the image. `graycoprops` uses the normalized GLCM to calculate properties.

Table 1 GLCM properties

Property	Description	Formula
'Contrast'	Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image. Range = [0 (size(GLCM,1)-1)^2] Contrast is 0 for a constant image.	$\sum_{i,j} i - j ^2 p(i, j)$

Property	Description	Formula
'Correlation'	Returns a measure of how correlated a pixel is to its neighbor over the whole image. Range = [-1 1] Correlation is 1 or -1 for a perfectly positively or negatively correlated image. Correlation is NaN for a constant image.	$\sum_{i,j} \frac{(i - \mu_i)(j - \mu_j) p(i, j)}{\sigma_i \sigma_j}$
'Energy'	Returns the sum of squared elements in the GLCM. Range = [0 1] Energy is 1 for a constant image.	$\sum_{i,j} p(i, j)^2$
'Homogeneity'	Returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Range = [0 1] Homogeneity is 1 for a diagonal GLCM.	$\sum_{i,j} \frac{p(i, j)}{1 + i - j }$

Along with GLCM following features are extracted

ii. Mean

`M = mean(A)` returns the mean values of the elements along different dimensions of an array. If A is a vector, `mean(A)` returns the mean value of A. If A is a matrix, `mean(A)` treats the columns of A as vectors, returning a row vector of mean values.

iii. Standard Deviation

`s = std(X)`, where X is a vector, returns the standard deviation using (1) above. The result s is the square root of an unbiased estimator of the variance of the population from which X is drawn, as long as X consists of independent,

identically distributed samples. If X is a matrix, $\text{std}(X)$ returns a row vector containing the standard deviation of the elements of each column of X . If X is a multidimensional array, $\text{std}(X)$ is the standard deviation of the elements along the first non-singleton dimension of X .

iv. Entropy

$E = \text{entropy}(I)$ returns E , a scalar value representing the entropy of grayscale image I . Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Entropy is defined as

$$-\sum (p \cdot \log_2(p))$$

Where p contains the histogram counts returned from `imhist`. By default, entropy uses two bins for logical arrays and 256 bins for `uint8`, `uint16`, or `double` arrays.

I can be a multidimensional image. If I have more than two dimensions, the entropy function treats it as a multidimensional grayscale image and not as an RGB image.

v. RMS

The RMS block computes the true root mean square (RMS) value of the input signal. The true RMS value of the input signal is calculated over a running average window of one cycle of the specified fundamental frequency

vi. Variance

$V = \text{var}(X)$ returns the variance of X for vectors. For matrices, $\text{var}(X)$ is a row vector containing the variance of each column of X . For N -dimensional arrays, var operates along the first nonsingleton dimension of X . The result V is an unbiased estimator of the variance of the population from which X is drawn, as long as X consists of independent, identically distributed samples. var normalizes V by $N - 1$ if $N > 1$, where N is the sample size. This is an unbiased estimator of the variance of the population from which X is drawn, as long as X consists of independent, identically distributed samples. For $N = 1$, V is normalized by 1.

vii. Kurtosis

$k = \text{kurtosis}(X)$ returns the sample kurtosis of X . For vectors, $\text{kurtosis}(x)$ is the kurtosis of the elements in the vector x . For matrices $\text{kurtosis}(X)$ returns the sample kurtosis for each column of X . For N -dimensional arrays, kurtosis operates along the first nonsingleton dimension of X .

viii. Skewness

$y = \text{skewness}(X)$ returns the sample skewness of X . For vectors, $\text{skewness}(x)$ is the skewness of the elements of x . For matrices, $\text{skewness}(X)$ is a row vector containing the sample skewness of each column. For N -dimensional arrays, skewness operates along the first nonsingleton dimension of X .

e) Classification using support vector machine (SVM)

SVM is a supervised machine learning algorithm. It can be used for classification and regression. It is mostly used in classification problems. It is state of the art pattern recognition technique grown up from statistical learning theory.

The basic idea of applying SVM for solving classification problems: a) through a non-linear mapping function to transform the input space to higher dimension feature space. b) The separating hyper plane with maximum distance from the closest points of the training set will be constructed. Using a SVM classifier to accurately classify features of every image in dataset.

The last step for diagnosis and which is specifically used for classification of tumors. There are two important commands `svm train` and `svm classify` used for this process. In Kernel space the training data to map by using Kernel function `svm train`. The kernel function can be following the function: linear, quadratic, polynomial. SVM `Strut` is used to classify each row of the data in sample, a matrix of data and using the information, created using the `svm train` function.

4. RESULT ANALYSIS

A. Brain Tumor Segmentation

The experiments are conducted on the brain tumor detection system where the inputs are MRI images of brain. MRI image is successfully processed at each step in brain tumor detection system and the desired result is obtained. MRI image of brain is given to watershed transform. For the purpose of image segmentation marker- controlled watershed segmentation is used. Resultant output from the segmentation method are generated and evaluated. Obtained results are shown in Figure 5.

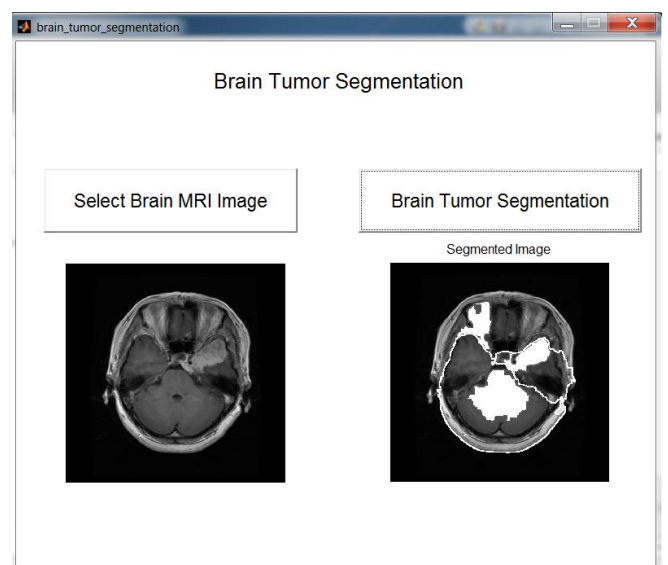


Figure 5 Brain Tumor Segmentation

B. Brain tumor Classification

The MATLAB based GUI designed for brain tumor classification is depicted in figure 6.

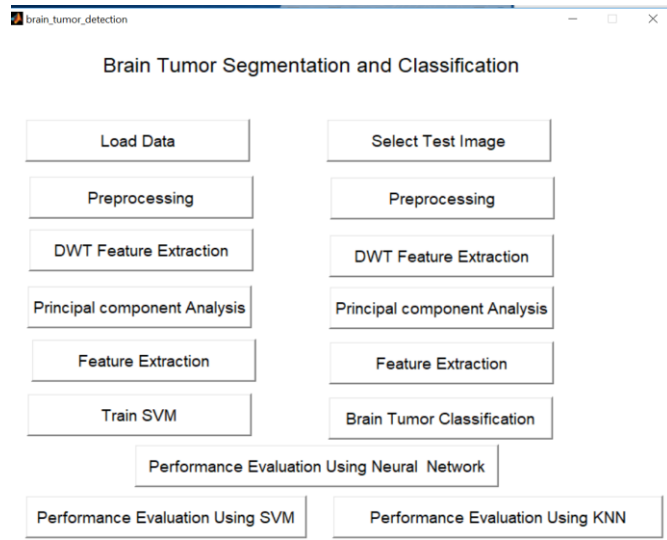


Figure 6 Brain Tumor Classification

In this study we have used brain tumor dataset containing 3064 T1-weighted contrast-enhanced images from 233 patients with three kinds of brain tumor: meningioma (708 slices), glioma (1426 slices) and pituitary tumor (930 slices)

The performance is evaluated using SVM, KNN and Neural networks classifiers.

C. Accuracy arithmetic

MATLAB is a tool for the analysis and check the efficiency of the algorithm and proposed model. The algorithm's performance can be evaluated in terms of accuracy, sensitivity, and specificity. The confusion matrix defining the terms TP, TN, FP, and FN from the predicted class and actual class result for the calculation of accuracy as shown in Table 2

Table 2 Confusion matrix defining the term TN, FP, FN and TN

Evaluation Parameter	Number of Test Images(919)		
	NNs	SVM	K-NN
True negative	538	363	479
False positive	128	137	152
True positive	164	187	124
False negative	89	232	164
Accuracy (%)	76.39	60.85	65.61

Where,

TP is the number of true positives, which is used to indicate the total number of abnormal cases correctly classified,

TN is the number of true negatives, which is used to indicate normal cases correctly classified;

FP is the number of false positive, and it is used to indicate wrongly detected or classified abnormal cases; when they are actually normal cases and

FN is the number of false negatives; it is used to indicate wrongly classified or detected normal cases; when they are actually abnormal cases.

All of these outcome parameters are calculated using the total number of samples examined for the detection of the tumor. The quality rate parameter accuracy is defined as the ratio of all the correct predictions divided by total number of cases examined. Formula to calculate accuracy is given below.

$$\text{Accuracy (Quality parameter)} = \frac{\text{All the correct prediction}}{\text{Total no cases examined}}$$

$$= \frac{TP+TN}{TP+TN+FN+FP}$$

The test performance of the classifier determined by the computation of the statistical parameters such as accuracy in comparison with different classifier techniques is shown in Table 3. Furthermore, higher values of accuracy indicate better performance.

Table 3 illustrates the comparison of various classifier techniques with accuracy the proposed system gives 76.39% accuracy for Neural Network, 65.61% for k-Nearest Neighbor and 60.85% for Support Vector Machine. The detailed analysis of performance measures is shown in Table 3 and through the performance measure, it is depicted that the performance of the proposed methodology has significantly improved the tumor identification compared with the NNs, SVM and K-NN based classification techniques.

Table 3 Comparison of accuracies in different classifier

Actual Class \ Prediction Class	Positive	Negative	Row Total
	Positive	TP	FP
Negative	FN	TN	FN+TN
Column Total	TP+FN	FP+TN	TP+FP+FN+TN

5. CONCLUSIONS

In this paper, the concepts related to image segmentation and Classification of brain tumor detection is discussed.

The watershed based thresholding approach is used for image segmentation which helps to recognize the portion of tumor in MRI image. Then using the approach of Discrete Wavelet Transform for Feature Extraction and Reduction of Principle Component Analysis, The Features Extracted are Mean, Standard deviation, Kurtosis, Skewness, Entropy, Contrast, Variance, Smoothness, Correlation and Energy. The accuracy of the brain tumor segmentation is also measured. Using this algorithm the brain tumors are accurately segmented and classification from an MR brain image. Then, Support Vector Machine is used to classify the tumors into benign, pituitary and malignant.

REFERENCES

- [1] Borole, V. Y., Nimbhore, S. S., & Kawthekar, D. S. S. (2015). Image Processing Techniques for Brain Tumor Detection: A Review. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, 4(5), 2.
- [2] Bahadure, N. B., Ray, A. K., & Thethi, H. P. (2017). Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. *International journal of biomedical imaging*, 2017
- [3] Coatrieux, G., Huang, H., Shu, H., Luo, L., & Roux, C. (2013). A watermarking-based medical image integrity control system and an image moment signature for tampering characterization. *IEEE journal of biomedical and health informatics*, 17(6), 1057-1067.
- [4] Saleck, M. M., ElMoutaouakkil, A., & Mouçouf, M. (2017, May). Tumor Detection in Mammography Images Using Fuzzy C-means and GLCM Texture Features. In *Computer Graphics, Imaging and Visualization, 2017 14th International Conference on* (pp. 122-125). IEEE.
- [5] Zulpe, N., & Pawar, V. (2012). GLCM textural features for brain tumor classification. *International Journal of Computer Science Issues (IJCSI)*, 9(3), 354.
- [6] Bhima, K., & Jagan, A. (2016, April). Analysis of MRI based brain tumor identification using segmentation technique. In *Communication and Signal Processing (ICCSP), 2016 International Conference on* (pp. 2109-2113). IEEE.
- [7] Vrji, K. A., & Jayakumari, J. (2011, July). Automatic detection of brain tumor based on magnetic resonance image using CAD System with watershed segmentation. In *Signal Processing, Communication, Computing and Networking Technologies (ICSCCN), 2011 International Conference on* (pp. 145-150). IEEE.
- [8] Rashid, M. H. O., Mamun, M. A., Hossain, M. A., & Uddin, M. P. (2018, February). Brain Tumor Detection Using Anisotropic Filtering, SVM Classifier and Morphological Operation from MR Images. In *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)* (pp. 1-4). IEEE.
- [9] Sudharani, K., Sarma, T. C., & Rasad, K. S. (2015, December). Intelligent Brain Tumor lesion classification and identification from MRI images using k-NN technique. In *Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2015 International Conference on* (pp. 777-780). IEEE.
- [10] Vidyarthi, A., & Mittal, N. (2013, December). A hybrid model for extraction of brain tumor in MR images. In *Signal Processing and Communication (ICSC), 2013 International Conference on* (pp. 202-206). IEEE.
- [11] Li, W., Chen, C., Su, H., & Du, Q. (2015). Local Binary Patterns and Extreme Learning Machine for Hyperspectral Imagery Classification. *IEEE Trans. Geoscience and Remote Sensing*, 53(7), 3681-3693.
- [12] Dhanaseely, A. J., Himavathi, S., & Srinivasan, E. (2012). Performance comparison of cascade and feed forward neural network for face recognition system.
- [13] Liu, Q., & Liu, X. (2013, December). Feature Extraction of Human Viruses Microscopic Images Using Gray Level Co-occurrence Matrix. In *Computer Sciences and Applications (CSA), 2013 International Conference on* (pp. 619-622). IEEE.
- [14] Singh, A. (2015, February). Detection of brain tumor in MRI images, using combination of fuzzy c-means and SVM. In *Signal Processing and Integrated Networks (SPIN), 2015 2nd International Conference on* (pp. 98-102). IEEE.