

# A Review on Tissue Segmentation and Feature Extraction of MRI Brain images

D.SELVARAJ

Research scholar, Department of ECE, Sathyabama University,  
Chennai, Tamilnadu, India  
mails2selvaraj@yahoo.com

R.DHANASEKARAN

Dean, Research, Syed Ammal Engineering College,  
Ramanathapuram, Tamilnadu, India  
rdhanashekar@yahoo.com

**Abstract**—Magnetic resonance imaging (MRI) is an important diagnostic imaging technique for the early detection of brain cancer. Brain cancer is one of the most dangerous diseases occurring commonly among human beings. The chances of survival can be increased if the cancer is detected at its early stage. MRI brain image plays a vital role in assisting radiologists to access patients for diagnosis and treatment. Studying of medical image by the Radiologist is not only a tedious and time consuming process but also accuracy depends upon their experience. So, the use of computer aided systems becomes very necessary to overcome these limitations. Even though several automated methods are available, still segmentation of MRI brain image remains as a challenging problem due to its complexity and there is no standard algorithm that can produce satisfactory results. In this review paper, various current methodologies of brain image segmentation using automated algorithms that are accurate and requires little user interaction are reviewed and their advantages, disadvantages are discussed. This review paper guides in combining two or more methods together to produce accurate results.

**Keywords** – Feature extraction, Image segmentation, Clustering, Neural Network, MRI image, Tumour detection

## I. INTRODUCTION

Segmentation of brain into various tissues like gray matter, white matter, cerebrospinal fluid, skull and tumour is very important for detecting tumour, edema, and hematoma. For early detection of abnormalities in brain parts, MRI imaging technique is used. Magnetic resonance imaging is a well established non-invasive diagnostic medical imaging technique based on the nuclear magnetic resonance phenomenon. It provides rich information about the human soft-tissue anatomy. It is mostly used in radiology in order to visualize the structure and function of the human body. It produces detailed images of the body in any direction. Particularly, MRI is useful in neurological (brain), musculoskeletal, and ontological (cancer) imaging because it offers much greater contrast between the diverse soft tissues of the body than the computer tomography (CT). MRI is different from CT, it does not use ionizing radiation, but uses an effective magnetic field to line up the nuclear magnetization of hydrogen atoms in water in the body [39]. Most research in developed countries has exposed that the death rate of people affected by brain tumor has increased over the past three decades [47].

A tumour is a mass of tissue that grows out of control of the normal forces that regulates growth [60]. Tumours can directly destroy all healthy brain cells. It can also indirectly damage healthy cells by crowding other parts of the brain and causing inflammation, brain swelling and pressure within the skull [33]. Brain tumours are of different sizes, locations and positions. They also have overlapping intensities with normal tissues [51]. Brain tumours can be divided according to the place of creation, type of cells or based on their appearance. Tumour can be benign or malignant. Slowly growing tumours are called benign and rapidly growing tumours are called as malignant. Brain tumour can occur in different parts of the brain and may or may not be primary tumours [55]. The most common primary brain tumours are Gliomas, Meningiomas, Glioblastoma, Oligodendroglioma, Pituitary adenomas and Nerve sheath tumours. Identification and segmentation of brain tumour in magnetic resonance images is very crucial in medical diagnosis because its accurate segmentation is very important for detecting tumours, edema and necrotic tissues [8].

Brain tumor is one of the most dangerous diseases occurring commonly among human beings, so study of brain tumor is very crucial. Reference [10] have proposed an image segmentation technique to identify the tumor from the brain magnetic resonance imaging (MRI). Several existing thresholding techniques have produced different result in each image. Thus, to produce a satisfactory result on brain tumor images, they have proposed a technique, where the detection of tumor was done uniquely. As well, reference [7] have proposed a computer-based technique for identifying the tumor region accurately in the brain via MRI images. Here, the classification

has been performed on a brain tumor image for identifying whether the tumor is a benign or malignant one. The steps involved in the proposed algorithm were preprocessing, image segmentation, feature extraction and image classification via neural network techniques. Finally, using the region of interest technique, the tumor area has been located.

Reference [42] have presented a cohesion based self merging (CSM) algorithm for the segmentation of brain MRI in order to find the exact region of brain tumor. CSM has drawn much attention because it gives a satisfactory result when compared to other merging processes. Here, the effect of noise has been reduced greatly and found that the chance of obtaining the exact region of tumor was more and the computation time was very less. Their algorithm was much simpler and computationally less complex.

Reference [12] has proposed a Particle Swarm Optimization (PSO) based clustering algorithm. The proposed algorithm has identified the centroids of number of clusters, where each cluster has grouped together the brain tumor patterns, obtained from MR Images. The results obtained for three performance measures have been compared with those obtained from Support Vector Machine (SVM) and Ada Boost. The performance analysis has shown that the qualitative results of proposed model are analogous with those obtained by SVM. Moreover, the different values of PSO control parameters have been selected in order to acquire better results from the algorithm.

Reference [63] has proposed a robust system for brain tumor diagnosis as well as for brain tumor region extraction. Initially, the proposed system has diagnosed the tumor from the brain MR images by naive bayes classification. After the diagnosis, the K-means clustering and boundary detection techniques have been applied to extract the exact brain tumor region. Here, above 99% accuracy has been achieved for diagnosis. Experimental results have shown that the proposed system has extracted accurate tumor region.

Reference [29] has proposed a technique for segmenting the brain tumors in 3D magnetic resonance images. Their technique was suitable to different kinds of tumors. Initially, the brain has been segmented using the proposed approach. Then, the suspicious areas have been selected with respect to the approximate brain symmetry plane and fuzzy classification for tumor detection. Here, in the segmentation stage, the tumor has been segmented successfully using the combination of a deformable model and spatial relations.

Reference [52] has developed an efficient system, where the Brain Tumor has been diagnosed with higher accuracy using artificial neural network. After the extraction of features from MRI data by means of the wavelet packets, an artificial neural network has been employed to find out the normal and abnormal spectra. Normally, the benefit of wavelet packets is to give richest analysis when compared with the wavelet transforms and thus adding more advantages to the performance of their proposed system.

Reference [91] have proposed a robust medical image segmentation technique, which combines watershed segmentation and Competitive Hopfield clustering network (CHCN) algorithm to minimize undesirable over-segmentation. A region merging method is presented, which is based on employing the region adjacency graph (RAG) to improve the quality of watershed segmentation. The performance of the proposed technique is evaluated through quantitative and qualitative validation experiments on benchmark images.

A new unsupervised MRI segmentation method based on self-organizing feature map was presented by [95]. Their algorithm included extra spatial information about a pixel region by using a Markov Random Field (MRF) model. The MRF term improved the segmentation results without extra data samples in the training set. The cooperation of MRF into SOFM has shown its great potentials as MRF term models the smoothness of the segmented regions. It verified that the neighbouring pixels should have similar segmentation assignment unless they are on the boundary of two distinct regions.

In the early research of medical tumor detection, the algorithms have directly used the classic methods of image processing (Such as edge detection and region growing) based on gray intensities of images. In recent years, the classification of human brain in MRI images is possible via supervised techniques such as k-nearest neighbour, Artificial neural networks and support vector machine(SVM) and unsupervised classification techniques such as self organization map(SOM) and fuzzy C-means algorithm have also been used to classify the normal or pathological  $T_2$  weighted MRI images.

Even though many algorithms are available for brain tumour detection, the detection rate is still not satisfactory. Also, accurate partitioning of an image into meaningful regions is essential key to success or failure of image classification. In this paper, various approaches of MRI brain image feature extraction are discussed in section II and current algorithms used for brain image segmentation are reviewed in section III. Concluding remarks are given in section IV.

## II. FEATURE EXTRACTION

To represent an image, large amount of data is required which occupies large amount of memory and time. In order to reduce the amount of data, memory and time, the features are extracted from an image. The extracted features contain the relevant information of an image. It can be used as an input to the classifier for image classification and segmentation. The type of features extracted from an image is classified as in Fig. 1.

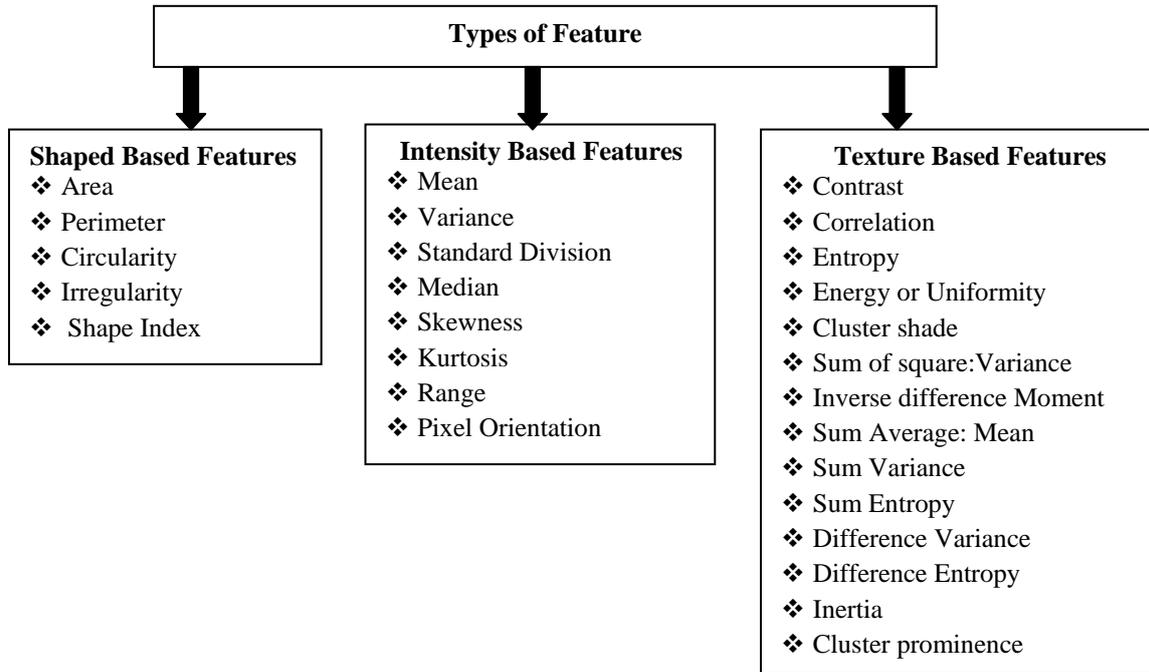


Figure 1. Classification of Feature that can be extracted from an image

### A. The Intensity Based Feature:

Intensity based feature is one of the most widely used feature. The intensity feature such as mean, median, mode, skewness, kurtosis, energy, entropy are considered. Let  $f(x,y)$  be a two dimensional function of an image,  $h(i)$  be the intensity level of an image,  $N_g$  be the total number of gray levels in the entire image and  $p(i)$  be the probability density. Then features that can be extracted are,

#### 1) Mean:

The mean defines the average level of intensity of the image or texture.

$$\text{Mean, } \mu = \sum_{i=0}^{N_g-1} i.p(i) \tag{1}$$

$$\text{where, } p(i) = \frac{h(i)}{N_x N_y}, i = 0, 1, 2, \dots, N-1 \tag{2}$$

$$h(i) = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} \delta(f(x, y), i), i = 0, 1, 2, \dots, N-1 \tag{3}$$

$$f(x, y) = \begin{bmatrix} f(0, 0) & f(0, 1) & f(0, 2) & \dots & f(0, N_y - 1) \\ f(1, 0) & f(1, 1) & f(1, 2) & \dots & f(1, N_y - 1) \\ f(0, 0) & f(0, 1) & f(0, 2) & \dots & f(0, N_y - 1) \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ f(N_x - 1, 0) & f(N_x - 1, 1) & \dots & \dots & f(N_x - 1, N_y - 1) \end{bmatrix} \tag{4}$$

$$\delta(i, j) = \begin{cases} 1 & ; i=j \\ 0 & ; i \neq j \end{cases} \tag{5}$$

2) *Variance:*

The variance defines the variation of intensity around the mean.

$$\text{Variance, } \sigma^2 = \sum_{i=0}^{N_g-1} (i - \mu)^2 \cdot p(i) \tag{6}$$

3) *Standard Deviation: (SD)*

$$\text{SD or } \sigma = \sqrt{\sum_{i=0}^{N_g-1} (i - \mu)^2 \cdot p(i)} \tag{7}$$

4) *Skewness:*

The skewness defines the symmetry of an image.

$$\mu^3 = \sigma^{-3} \sum_{i=0}^{N_g-1} (i - \mu)^3 \cdot p(i) \tag{8}$$

$$\text{Skewness} = \begin{cases} \mu^3 < 0; & \text{Histogram below the mean} \\ \mu^3 = 0; & \text{Histogram is equal to the mean} \\ \mu^3 > 0; & \text{Histogram above the mean} \end{cases}$$

5) *Kurtosis:*

The kurtosis defines the measure of flatness of the histogram.

$$\text{Kurtosis, } \mu^4 = \sigma^{-4} \sum_{i=0}^{N_g-1} ((i - \mu)^4 \cdot p(i)) - 3 \tag{9}$$

6) *Energy:*

Energy defines the measure of sum of squared elements.

$$\text{Energy, } E = \sum_{i=0}^{N_g-1} [p(i)]^2 \tag{10}$$

7) *Entropy:*

Entropy is defined as a measure of uncertainty in a random variable

$$\text{Entropy, } E_n = - \sum_{i=0}^{N_g-1} p(i) \cdot \log_2[p(i)] \tag{11}$$

8) *Range (R):*

The range has two elements. One is minimum pixel intensity value and other one is maximum pixel intensity value.

Range, R = Minimum value of pixel intensity and maximum value of pixel intensity in a image.

9) *Pixel Orientation (PO):*

$$\text{Pixel orientation, PO} = \text{Tan}^{-1} \{(y - m) / (x - m)\} \tag{12}$$

Where, m – the point which we require to measure the value.

x – point in the x-axis

y – point in the y-axis

**B. Texture Feature:**

Texture is a characteristic of an image that provides higher order description of an image and includes information about the spatial distribution of tonal variations or gray tones. The texture extraction defines the homogeneity or similarity between regions of an image. Texture feature classifies magnetic resonance image of brain in to gray matter, white matter, cerebrospinal fluid and tumour region. The Haralick texture feature [30] are,

1) *Energy (E):*

It is also known as angular second moment or uniformity. This measures the homogeneity.

$$E = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i, j)^2 \quad (13)$$

2) *Contrast (Con):*

It measures the intensity variation between the reference pixel and neighbor pixel.

$$Con = \sum_{n=0}^{N_g-1} n^2 \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i, j)^2 \quad (14)$$

3) *Correlation (Cor):*

It measures, how the reference pixel is related to its neighbor pixel.

$$Cor = \frac{1}{\sigma_x \sigma_y} \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i, j) \cdot p(i, j)^2 - \mu_x \mu_y \quad (15)$$

Where,  $\sigma_x$ ,  $\sigma_y$ ,  $\mu_x$   $\mu_y$  are the standard deviation and mean of  $p_x$  and  $p_y$ .

4) *Sum of squares: Variance ( $\sigma^2$ )*

It measures the gray tone variance.

$$\text{Variance, } \sigma^2 = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - \mu)^2 \cdot p(i, j) \quad (16)$$

5) *Inverse Difference Moment (IDM):*

It measures the local homogeneity of an image.

$$IDM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{1}{1 + (i - j)^2} p(i, j) \quad (17)$$

6) *Sum Average (Mean):*

$$\text{Mean, } \mu = \sum_{i=0}^{2(N_g-1)} i \cdot p_{x+y}(i) \quad (18)$$

7) *Sum Variance (SV):*

$$\text{Sum Variance, } SV = \sum_{i=0}^{2(N_g-1)} (i - SE)^2 p_{x+y}(i) \quad (19)$$

8) *Sum Entropy (SE):*

$$\text{Sum Entropy, } SE = - \sum_{i=0}^{N_g-1} p_{x+y}(i) \cdot \log p_{x+y}(i) \quad (20)$$

9) *Entropy (En):*

Entropy gives the amount of information of an image that are needed for image compression.

$$\text{Entropy, } E = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i, j) \cdot \log p(i, j) \quad (21)$$

10) *Difference Variance (DV):*

$$\text{Difference Variance, } DV = \text{Variance of } p_{x-y} \quad (22)$$

11) *Difference entropy (DE):*

$$\text{Difference Entropy, } DE = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \cdot \log p_{x-y}(i) \quad (23)$$

12) *Inertia (I):*

$$\text{Inertia, } I = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - j)^2 \cdot p(i, j) \quad (24)$$

13) Cluster shade (CS):

$$\text{Cluster Shade, CS} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i + j - \mu_x - \mu_y)^3 \cdot p(i, j) \tag{25}$$

14) Cluster Prominence (CP):

$$\text{CP} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i + j - \mu_x - \mu_y)^4 \cdot p(i, j) \tag{26}$$

C. Shape Features:

Shape provides the geometrical information of an object in an image. The shape features such as centroid, Eccentricity, area, perimeter, circularity, shape index, solidity, orientation, euler number are considered.

III. MRI BRAIN TISSUE SEGMENTATION TECHNIQUES

Various segmentation methods have been cited in the literature for improving the segmentation some of the commonly used methods are summarized as in Fig. 2.

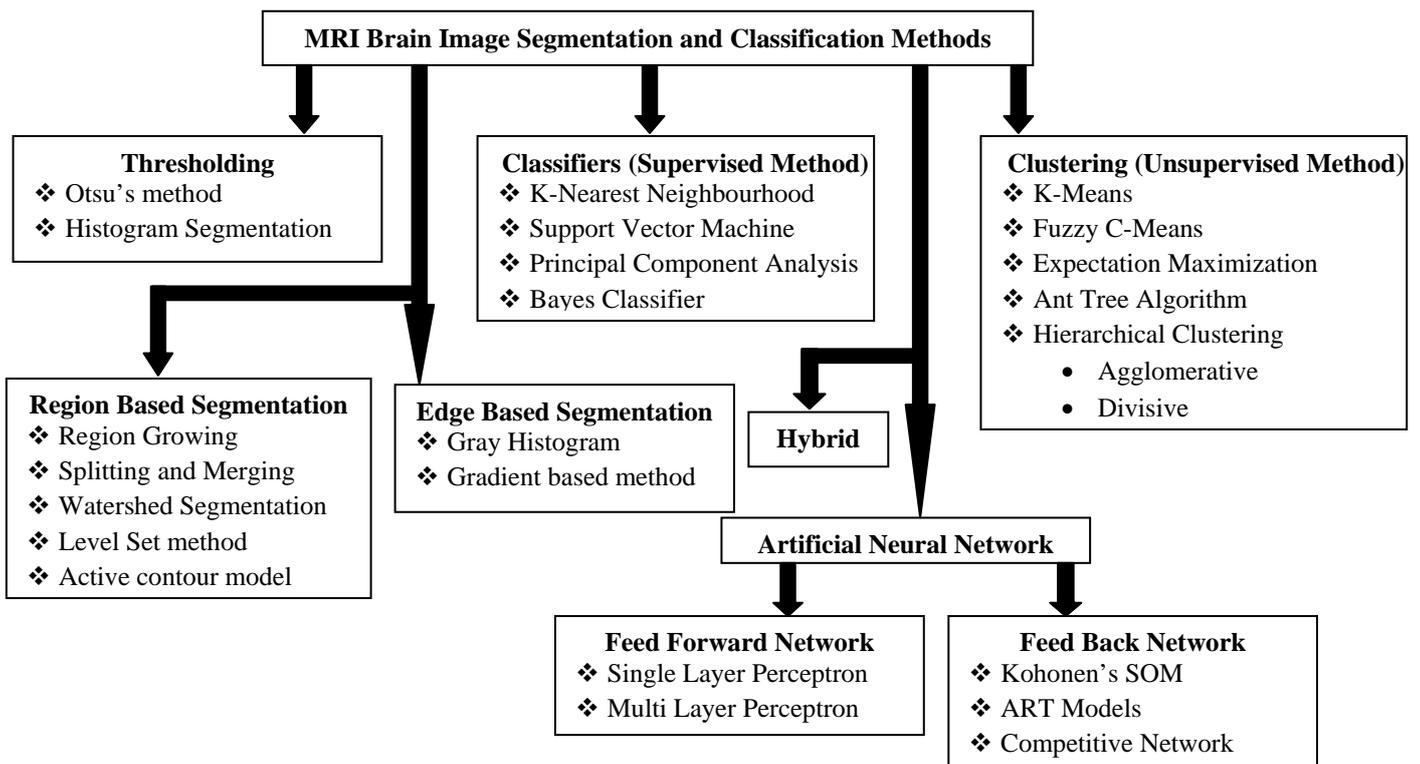


Figure 2. Classification of MRI Brain image segmentation methods

A. Thresholding

Thresholding is one of the frequently used method for image segmentation. This method is effective for images with different intensities. Using this method, the image is partitioned directly into different regions based on the intensity values.

Thresholding is defined mathematically [76] as Eq. (27)

Let  $f(x,y)$  be the input image and 'T' be the threshold value then the segmented image  $g(x,y)$  is given by,

$$g(x, y) = \begin{cases} 1, & \text{if } f(x, y) > T \\ 0, & \text{if } f(x, y) \leq T \end{cases} \tag{27}$$

Using the above Eq. (27), the image can be segmented into 2 groups. If we want to segment the given image into multigroups then we should have multi threshold point.

If we have 2 threshold values, then the above equation becomes as Eq. (28) and this equation segments the image into 3 groups

$$g(x, y) = \begin{cases} a, & \text{if } f(x, y) > T_2 \\ b, & \text{if } T_1 < f(x, y) \leq T_2 \\ c, & \text{if } f(x, y) \leq T_1 \end{cases} \quad (28)$$

Algorithm:

The algorithm for the thresholding is given by [27],

Step 1: Select an initial estimate for the global threshold,  $T$ .

Step 2: Segment the image using  $T$  in equation (27) to get 2 groups of pixels:  $G_1$  consisting of all pixels with intensity values  $>T$  and  $G_2$  consisting of pixels with values  $\leq T$ .

Step 3: Compute the average mean intensity values  $m_1$  and  $m_2$  for the pixels in  $G_1$  and  $G_2$  respectively.

Step 4: Compute a new threshold value

Step 5: Repeat steps 2 through 4 until the difference between values of  $T$  in successive iterations is smaller than a predefined parameter  $\Delta T$ .

A survey on available thresholding techniques are provided in the literature [5, 16, 67, 74, 93].

#### 1) Global thresholding

Global thresholding method selects only one threshold value for the entire image. Global thresholding is used for bimodal images. It is simple and faster in computation time only if the image has homogeneous intensity and high contrast between foreground and background.

#### 2) Otsu's thresholding

Reference [87] has developed improved threshold image segmentation algorithm based on the otsu method. This method ensures that both the variance of the object and the variance of the background keep away from the variance of the whole image. The improved method produces satisfactory results for both the image with histogram of bimodal and unimodal distributions.

Otsu method is the commonly used thresholding technique. Otsu method is simple and effective to implement. Otsu's thresholding technique is based on a discriminant analysis which partitions the image into two classes  $C_1$  and  $C_2$  at gray levels 'k' such that  $C_1 = \{0, 1, 2, 3, \dots, k\}$  and  $C_2 = \{k+1, k+2, \dots, L-1\}$  where, 'L' is the total number of gray levels of the image. Let 'n' be the total number of pixels in the given image and ' $n_i$ ' be the number of pixels at the  $i^{\text{th}}$  gray level.

The probability of occurrence of gray level is defined as,

$$P_i = \frac{n_i}{n} \quad (29)$$

' $C_1$ ' and ' $C_2$ ' are two classes representing the region of interest and the background.

The probabilities of classes  $C_1$  and  $C_2$  are,

$$P_1(k) = \sum_{i=0}^k P_i \quad (30)$$

$$P_2(k) = \sum_{i=k+1}^{L-1} P_i = 1 - P_1(k) \quad (31)$$

The mean intensity values of these two classes are

$$m_1(k) = \sum_{i=0}^k i.P(i/C_1) \quad (32)$$

$$m_1(k) = \sum_{i=0}^k i.P(i/C_1).P(i)/P(C_1)$$

$$m_1(k) = \frac{1}{P_1(k)} \sum_{i=0}^k i.P_i \quad (33)$$

$$m_2(k) = \sum_{i=k+1}^{L-1} i.P(i/C_2)$$

$$m_2(k) = \frac{1}{P_2(k)} \sum_{i=k+1}^{L-1} i.P_i \quad (34)$$

where,  $m_1(k)$  and  $m_2(k)$  are object's center gray and background's center gray.

Let  $\sigma_B^2$  and  $\sigma_T^2$  be the between-class variance and total variance. An optimal threshold  $k^*$  can be obtained by maximizing the between-class variance.

$$K^* = \text{Arg} \left\{ \max_{0 < k \leq L-1} \left( \frac{\sigma_B^2}{\sigma_T^2} \right) \right\} \quad (35)$$

The between-class variance  $\sigma_B^2$  and total variance  $\sigma_T^2$  is defined as,

$$\sigma_B^2 = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2 \quad (36)$$

$$\sigma_T^2 = \sum_{i=0}^{L-1} (i - m_G)^2 \cdot P_i \quad (37)$$

The cumulative (average intensity) up to level 'k' is given by

$$m(k) = \sum_{i=0}^k i \cdot P_i \quad (38)$$

The total mean of the whole image is defined as,

$$m_G = \sum_{i=0}^{L-1} i \cdot P_i \quad (39)$$

An equivalent and simpler formula for obtaining threshold  $k^*$  is given as,

$$k^* = \max_{0 \leq k \leq L} \left\{ w_0(\mu_0 - \mu_T)^2 + w_1(\mu_1 - \mu_T)^2 \right\} \quad (40)$$

Algorithm:

Step1: Compute probabilities of each intensity level

Step2: Compute various thresholds  $T = 1, 2, \dots$ , max intensity and i) upgrade  $w_i$  and  $\mu_i$  ii) Compute

$$\sigma_B^2(k)$$

Step 3: Select threshold value having max  $\sigma_B^2(k)$

Disadvantages:

1. Otsu method can segment only larger objects from background
2. Otsu method fails if the histogram is unimodal or close to unimodal

### 3) Local thresholding

Threshold values are selected locally by dividing an image into sub-images and calculate a threshold value for each part. A local threshold takes more computation time than global thresholding. Its result is satisfactory in background variations in an image. It can extract only small regions [56].

#### a) Histogram thresholding

Histogram thresholding segmentation is based up on thresholding of histogram features and gray level thresholding. Threshold is defined mathematically as Eq. (27)

Algorithm [69, 76].

Step1: The MRI brain image is divided into two equal halves around its central axis and the histogram of each part drawn.

Step2: Threshold point of the histogram is calculated based on the comparison technique made between 2 histograms.

Step3: Segmentation is done using the threshold point for both the halves.

Step4: The detected image is cropped along its contour to find out the physical dimension of the tumour.

Step5: Create an image of the original size, check the segmented image pixel value. If its value is greater than threshold value then assign 255 else 0.

Step6: Segment the tumour area.

Step7: Calculate the area of the tumour

### B. Edge based segmentation

Edge based segmentation methods partition an image based on rapid changes in intensity near edges [27, 58]. The result is a binary image. Based on theory there are two main edge based segmentation methods- gray histogram and gradient based method [37].

#### 1) Gray Histogram Technique

The result of edge detection technique mainly depends upon selection of threshold T [37].

#### 2) Gradient Based Method

In gradient based method, the difference between two neighbouring pixel values is considered. So, when there is an abrupt change in intensity near edge and there is little image noise then gradient based method works well [37]. This method involves convolving gradient operators with the image. Common edge detection operators used in gradient based method are sobel operator, canny operator, Laplace operator, Laplacian of Gaussian (LOG) operator & so on, canny is most promising one [27], but takes more time as compared to sobel operator.

Edge detection methods require a balance between detecting accuracy and noise immunity. If the level of detecting accuracy is too high, noise may bring in fake edges making the outline of images unreasonable and if the degree of noise immunity is too excessive [37], some parts of the image outline may get undetected and the position of objects may be mistaken. Thus, edge detection algorithms are suitable for images that are simple and noise-free as well often produce missing edges or extra edges on complex and noisy images [86].

To detect boundaries between 2 distinct regions, edge based segmentation is used.

Algorithm:

Step1: Apply the derivative operator to detect edges of the image

Step2: Find the magnitude at the edges

Step3: Retain all edge having magnitude greater than threshold value T

Step4: Find the position of crack edge, the crack edge is either retained or rejected.

Step5: Repeat step 3 and step 4 with different values of threshold so as to find out the closed boundaries, segmentation of an image is obtained.

### C. Region based segmentation

Compared to edge detection method, segmentation algorithms based on region are relatively simple and more immune to noise [37, 98]. Edge based methods partition an image based on rapid changes in intensity near edges whereas region based methods, partition an image into regions that are similar according to a set of predefined criteria [27, 36]. Segmentation algorithms based on region mainly include following methods:

#### 1) Region growing

Region growing method is one of the popular segmentation methods. This method starts with a seed pixel and grows the region by adding the neighbouring pixels based on threshold value. When the growth of a region stops, another seed pixel which does not belong to any other region is chosen and the process is repeated [50]. The region growing is stopped when all pixels belongs to some region.

Region growing segmentation is particularly used for delineation of small, simple structures such as tumours and lesions [24, 70, 72]. The various limitations in using this method are, 1. Sometimes, manual interaction is required to select the seed point. 2. Sensitive to noise so it produces holes or over segmentation in the extracted regions. The discontinuity in the extracted image can be removed by using homotopic region growing algorithm [11].

#### 2) Region splitting and merging

In this method, the image is splitted into various regions depending on some criterion and then it is merged. The whole image is initially taken as a single region and then internal similarity is computed using standard deviation. If too much variety occurs then the image is split into regions using thresholding. This is repeated until no more splits are further possible. Quadtree is a common data structure used for splitting. Then comes the merging phase, where two regions are merged if they are adjacent and similar. Similarity can be measured by comparing the mean gray level or using statistical tests. Two regions  $R_1$  and  $R_2$  are merged into  $R_3$  if,

$$H(R_1 \cup R_2) = \text{True and} \quad (41)$$

$$|m_1 - m_2| < T \quad (42)$$

Where,  $m_1$  and  $m_2$  are the mean gray level values in the regions  $R_1$  and  $R_2$ ; T is some appropriate threshold [27]. Merging is repeated until no more further merging is possible. The major advantage of this technique is guaranteed connected regions. The drawbacks of the split and merge technique are, 1. The results depend on the position and orientation of the image, 2. Regular division leads to over segmentation by splitting. This drawback can be overcome by reducing number of regions by using Normalized cuts method.

Improved Quadtree method for split and merge is introduced in [21]. In this improved method they have used three steps first splitting the image, second initializing neighbours list and the third step is merging splitted regions. The survey says split and merge algorithm is a fast computation method. Its drawback is lack of sensitivity to image semantics [24, 92]

### 3) *Watershed segmentation*

Watershed segmentation algorithm can be used if the foreground and the background of the image can be identified. Watershed algorithm is also used to capture weak edges. Selection of seed point is the main drawback of this approach. Random selection of seed point may lead to inappropriate results and increases convergence rate. In water segmentation, an image looks like a surface where bright pixels are considered as mountain tops and dark pixels are considered as valleys. Some valleys have punctures which are slowly merged into water that will be poured and then it will start to fill the valleys. But if water comes from different punctures, it is not allowed to be mixed. So, the dam is built at contact points which make dams work as boundaries of water and image objects [22, 64].

The steps followed in the watershed segmentation [76] are,

Step1: Compute segmentation function

Step2: Compute gradient magnitude using derivative operator

Step3: Compute foreground markers

Step4: Compute background markers

Step5: Modify the segmentation function to have minimum values at the foreground and background marker locations.

### 4) *Snakes*

Active contours or snakes are computer generated curves [34] that move within the image to find object boundaries under the influence of internal and external forces.

The steps followed in active contour:

Step1: Snake is placed near the contour of Region of Interest (ROI).

Step2: The Snake is attracted towards the target by an iterative process (by various internal and external forces within the image) [38].

Step3: An energy function is constructed which consisting of internal and external forces is constructed to measure the appropriateness of the Contour of ROI

Step4: Minimize the energy function

The internal forces are responsible for smoothness while the external forces guide the contours towards the contour of ROI. It requires user interaction, which consists of determining the curve around the detected object [38]. Snake should be placed usually near the boundary of ROI [100]

### 5) *Level Set Model*

Many of the PDEs [34] used in image processing are based on moving curves and surfaces with curvature based velocities. The curves or surfaces are represented as the zero level set of a higher dimensional hyper surface. It solves the problem of corner point producing and curve breaking. Since the edge-stopping function depends on the image gradient, only objects with edges defined by gradients can be segmented.

## D. *Classifiers*

### 1) *KNN*

K-Nearest Neighbour (k-NN) classification technique is the simplest technique that provides good classification accuracy [89]. The k-NN algorithm is based on a distance function and a voting function in k-Nearest Neighbours, the metric employed is the Euclidean distance [25]. The k-NN has higher accuracy and stability for MRI data than other common statistical classifiers, but has a slow running time [17]. The segmentation steps of k-NN are,

Step1: Determine k value

Step2: Calculate distance between the query instance and the training samples

Step3: Sort the distance based on the kth minimum distance

Step4: Assign the majority class

Step5: Determine the class

Step6: Segment the brain abnormalities

## 2) SVM

Support Vector Machine is a supervised classifier with associated learning algorithm derived from the statistical theory [82, 83]. It was first developed as an extension of the Generalized Portrait algorithm [84, 85]. It works under the principle of structure risk reduction from statistical learning theory. Instead of minimizing an objective function based on the training samples (such as mean square error), the SVM attempts to minimize the bound on the generalization error (i.e., the error made by the learning machine on the test data not used during training). As a result, an SVM tends to perform well when applied to data outside the training set. SVM achieves this advantage by focusing on the training examples that are most difficult to classify. These “borderline” training examples are called support vectors. Vapnik’s SVM formulation is modified [73, 78, 79] by adding a least squares term in the cost function. This variant circumvents the need to solve a more difficult quadratic programming problem and only requires the solution of a set of linear equations. This approach significantly reduces the complexity and computation in solving the problem. The SVM classifier is based on the hyperplane that maximizes the separating margin between the two classes.

The basic SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the output, making it a non-probabilistic binary linear classifier. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. To maximize the margin between the classes [99], its kernel is used to control the empirical risk and classification. There are many kernel functions such as linear, polynomial of degree and Radial basis function (RBF). Among these kernel functions, a radial basis function proves to be better for MRI brain images [99].

SVM has two stages; training and testing stage. SVM trains itself by features given as an input to its learning algorithm. During training SVM selects the suitable margins between two classes. Features are labeled according to class associative with particular class. Artificial neural network has a few issues having local minima and number of neurons selection for each problem. In order to resolve this problem SVM occupies no local minima and overhead to neurons selection by initiating the idea of hyper planes.

SVM is an attractive and systematic method for two class problems. In this research work, many authors [1, 13, 14, 40, 57] are classifying MRI brain images into two separate classes such as normal and abnormal using Support Vector Machine. Support Vector machine based classification of various levels of MR glioma images are performed [96]. This method claimed to be the better than rule based systems but the accuracy is low.

## 3) PCA

Principal Component Analysis (PCA) is used to extract principal features. It uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. PCA is mainly used to reduce the large dimensionality of the data by removing the redundant features [23, 77]. PCA is also called as discrete karhunen-Loeve transform (KLT), the Hotelling transform or proper orthogonal decomposition (POD).

The steps involved in extracting the Principal Components using PCA are,

Step 1: Convert the 2D image into 1D image for both test images and database images.

Step 2: Find the mean value for each one dimensional

Step 3: Find the difference matrix for each image.

Step 4: Find the covariance matrix by Covariance (L) = A\*A

Step 5: Find the eigen vector for 1D image by [V, D] = eig(L)

Step 6: Find the eigen face of 1D image, Eigen Face = Eigen vector\*A

The results of a PCA are usually discussed in terms of component scores called as factor scores [23].

## E. Clustering

A Clustering is one of the most useful techniques in MRI Segmentation, where it classifies pixels into classes, without knowing previous information or training. It classifies pixels with highest probability into the same class. Clustering technique training is done by using pixel features with properties of each class [59, 61, 68, 88].

### 1) K-means

K-means clustering algorithm is the simplest unsupervised learning algorithm that can solve clustering problem. The procedure followed to classify a given set of data through a certain number of clusters is very simple. In K-means ‘K’ centres are defined, one for each cluster. These clusters must be placed far away from each other. The next step is to take a point belonging to a given data set and associate it to the nearest centre. When no point is pending, the first step is completed and early grouping is done. The second step is to recalculate ‘k’ new centroids as barycentre of the clusters resulting from the previous step. After having ‘K’ new centroids a new binding has to be done between the same data set points and the nearest new centre. A loop has

been generated. As a result of this loop, the k centres change their location step by step until centres do not move any more. Finally this algorithm aims at minimizing an objective function known as squared error function given by,

$$J(V) = \sum_{i=1}^c \sum_{j=1}^{c_i} \|x_i - v_j\|^2 \tag{43}$$

Where,  $\|x_i - v_j\|$  is the Euclidean distance between  $x_i$  and  $v_j$

' $C_i$ ' is the number of data points in  $i^{th}$  cluster.

' $C$ ' is the number of cluster centres.

Algorithmic steps for K-means clustering:

Let  $X = \{x_1, x_2, x_3, \dots, x_n\}$  be the set of data points and  $V = \{v_1, v_2, v_3, \dots, v_c\}$  be the set of centres.

Step1: Randomly select ' $c$ ' cluster centres

Step2: Calculate the distance between each data point and cluster centres.

Step3: Assign the data point to the cluster centre whose distance from the cluster centre is minimum of all the cluster centres.

Step4: Recalculate the new cluster centre using

$$v_i = \left( \frac{1}{C_i} \right) \sum_{j=1}^{C_i} x_j \tag{44}$$

Where ' $C_i$ ' represents the number of data points in  $i^{th}$  cluster.

Step5: Recalculate the distance between each data point and newly obtained cluster centres.

Step6: If no data point was reassigned then stop, otherwise repeat from step 3.

K-means algorithm is fast, robust and easier to understand. It also gives better result when data set are well separated from each other. But, if there are 2 highly overlapping data then k-means will not be able to resolve that there are 2 clusters.

## 2) Fuzzy C-means (FCM)

FCM clustering is an unsupervised method for the data analysis. This algorithm assigns membership to each data point corresponding to each cluster centre on the basis of distance between the cluster centre and the data point. The data point near to the cluster centre has more membership towards the particular centre. Generally, the summation of membership of each data point should be equal to one. After each iteration, the membership and cluster centres are updated according to the formula

$$V_j = \frac{\left( \sum_{i=1}^n (\mu_{ij})^m x_i \right)}{\left( \sum_{i=1}^n (\mu_{ij})^m \right)}, \forall j = 1, 2, \dots, c \tag{45}$$

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{d_{ij}}{d_{ik}} \right)^{\left( \frac{2}{m-1} \right)}} \tag{46}$$

Where,

' $n$ ' is the number of data points

' $V_j$ ' represents the  $j^{th}$  cluster centre

' $m$ ' is the fuzziness index  $m \in [1, \infty]$

' $c$ ' represents the number of cluster centre

' $\mu_{ij}$ ' represents the membership of  $i^{th}$  data to  $j^{th}$  cluster centre.

' $d_{ij}$ ' represents the Euclidean distance between  $i^{th}$  data and  $j^{th}$  cluster centre.

' $x_i$ ' is the  $i^{th}$  of  $d$ -dimensional measured data

' $c_j$ ' is the  $d$ -dimension centre of the cluster

$\| * \|$  is any norm expressing the similarity between any measured data and the centre.

$$d_{ij} = \|x_i - c_j\|, d_{ik} = \|x_i - c_k\|$$

The main objective of fuzzy c-means algorithm is to minimize

$$J(U, V) = \sum_{i=1}^n \sum_{j=1}^c (\mu_{ij})^m \|x_i - v_j\|^2, 1 \leq m < \infty \tag{47}$$

Where,

$\|x_i - v_j\|$  is the Euclidean distance between  $i^{\text{th}}$  data and  $j^{\text{th}}$  cluster centre.

Algorithmic steps for fuzzy C-means clustering:

Let  $X = \{x_1, x_2, x_3, \dots, x_n\}$  be the set of data points and  $V = \{v_1, v_2, v_3, \dots, v_c\}$  be the set of cluster centres.

Step1: Randomly select ‘c’ cluster centres

Step2: Calculate the fuzzy membership ‘ $\mu_{ij}$ ’ using the equation

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{d_{ij}}{d_{ik}} \right)^{\left( \frac{2}{m-1} \right)}} \tag{48}$$

Step3: Compute the fuzzy centres ‘ $v_j$ ’ using

$$V_j = \frac{\left( \sum_{i=1}^n (\mu_{ij})^m x_i \right)}{\left( \sum_{i=1}^n (\mu_{ij})^m \right)}, \forall j = 1, 2, \dots, c \tag{49}$$

Step4: Repeat step2 and step3 until the minimum ‘J’ value is achieved or  $\|U^{(k+1)} - U^{(k)}\| < \beta$

Where,

‘k’ is the iteration step

‘ $\beta$ ’ is the termination criterion between [0,1]

‘ $U = (\mu_{ij})_{n \times c}$ ’ is the fuzzy membership matrix

‘J’ is the objective function

The first loop of the algorithm calculates membership values for the data points in clusters and the second loop recalculates the cluster centres using these membership values. When the cluster centre stabilizes the algorithm ends.

The FCM algorithm gives best result for overlapped data set and also gives better result than k-means algorithm. Here, the data point can belong to more than one cluster centre. The main drawback of FCM is 1) the sum of membership value of a data point  $x_i$  in all the clusters must be one but the outlier points has more membership value. So, the algorithm has difficulty in handling outlier points. 2) Due to the influence of all the data members, the cluster centres tend to move towards the centre of all the data points [19]. It only considers image intensity thereby producing unsatisfactory results in noisy images [28]. A bunch of algorithms are proposed to make FCM robust against noise and in homogeneity but it’s still not perfect [28, 48, 2, 81, 20, 97].

a) *Modified fuzzy c-means*

Many approaches have been made to modify the existing standard FCM algorithm to improve its performance. Each of the modified FCM algorithms proposes a new membership function for calculating the membership of data points in clusters.

**FCM WITH MODIFIED DISTANCE FUNCTION**

A new distance function based on dot product instead of the conventional Euclidean distance [26]. The introduced new membership function is given in Eq.(50).

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{d^2(v_j, x_k)}{d^2(v_i, x_k)} \right)^{\left( \frac{1}{m-1} \right)}} \tag{50}$$

### MODIFIED C-MEANS FOR MRI SEGMENTATION

A new approach for segmentation of MRI images that have been corrupted by intensity inhomogeneities and noise is also proposed [45]. The algorithm is formulated by modifying the objective function of the standard fuzzy C-means method to compensate for intensity inhomogeneities [45]. Here the membership function is given as Eq.(51)

$$\mu_{jk} = \frac{1}{\sum_{l=1}^c \left( \frac{\delta_{jk} + \gamma_k}{\delta_{jl} + \gamma_l} \right)^{\frac{1}{m-1}}} \quad (51)$$

Here, 'δ' is the distance and 'γ' denotes the influence on a pixel by the neighbouring membership values.

Other improved versions of FCM by the modification of the objective function were introduced by [3] and [97]. In the proposed IFCM algorithm of [75] during clustering, each pixel attempts to attract its neighbouring pixels towards its own cluster. Here the attraction depends up on the intensity of the pixel. In MRI brain images, the presence of noise will alter the pixel intensity. Due to this the segmentation may be affected so, instead of modifying the objective function, the measurement of similarity was extended by considering neighbourhood attraction [71].

### ADAPTIVE FUZZY CLUSTERING

The adaptive fuzzy clustering algorithm [19] is a modified version of standard FCM. The membership values in this method are calculated using Eq. (52)

$$\mu_{ij} = \frac{n * \left( \frac{1}{d_{ji}} \right)^{\frac{1}{m-1}}}{\sum_{k=1}^p \sum_{z=1}^n \left( \frac{1}{d_{kz}} \right)^{\frac{1}{m-1}}} \quad (52)$$

This algorithm is efficient in handling data with outlier points. In comparison with FCM algorithm it gives very low membership for outlier points [19].

#### 3) Expectation maximization

A model based tumour segmentation technique was implemented by [53]. This approach uses a modified Expectation Maximization (EM) algorithm to differentiate the healthy and the tumorous tissues. A set of tumour characteristics are presented in this paper which is highly essential for accurate segmentation. But the drawback of this work is the lack of quantitative analysis on the extracted tumour region.

EM steps are demonstrated in the following steps [8]:

Step1: Initialize mean and Covariance matrix using K-means.

Step2: Calculate membership probability of each training data.

Step3: Compute mean and variance of each gaussian component using membership function obtained in step

2. The step 2 and 3 are repeated until convergence.

Gauss mixture vector of each class is obtained by EM training data for that class. The applications of the EM algorithm to brain MR image segmentation were reported by [90] and [44]. A common disadvantage of EM algorithm is that the intensity distribution of brain images is modeled as a normal distribution.

#### 4) Hierarchical clustering

Hierarchical clustering method works by grouping data objects into a tree of clusters. Hierarchical clustering doesn't require specifying the number of clusters. Hierarchical clustering is deterministic. There are 2 types of hierarchical clustering. 1. Agglomerative clustering, 2. Divisive Clustering.

The basic process of hierarchical clustering for a given set of N items to be clustered and NxN distance matrix is given as, [43]

Step 1: Start by assigning each item to a cluster i.e., N items will have N clusters. Let the distance between the clusters be the same as the distance between the items they contain.

Step 2: Find the similar pair of clusters and merge them into a single cluster to reduce the number of clusters.

Step 3: Compute the distance between the new cluster and each of the old cluster.

Step 4: Repeat steps 2 and 3 until all items are clustered into a single cluster of size N.

a) *Agglomerative Hierarchical clustering*

In Agglomerative clustering, each element is treated as a singleton cluster and then merged (agglomerated) until all merge in a single cluster, which results in dendograms formation [43]. Dendograms are horizontal lines which when cut at a point you get a specific part or element and explains how clustering helps forming an image.

Algorithm:

Step 1: Compute NxN similarity matrix C

Step 2: Execute N-1 steps of merging the currently most similar clusters.

Step 3: In each iteration, the most similar clusters are merged and the rows and columns of the merged cluster i in 'C' are updated.

Step 4: The clustering is stored as a list of merges in A.

Step 5: I indicates which clusters are still available to be merged.

Step 6: The function SIM(i, m, j) computes the similarity of cluster j with the merge of clusters i and m.

Step 7: For some Hierarchical algorithm, SIM(i, m, j) is simply a function of C[j][i] and C[j][m] for example, the maximum of these 2 values for single link.

Disadvantages:

They do not scale well: time complexity of at least  $O(n^2)$ , where n is the number of total objects. They can never undo what was done previously.

b) *Divisive Hierarchical clustering*

Divisive hierarchical is more efficient than Agglomerative hierarchical clustering [43]. Divisive Hierarchical Clustering can be stopped when the goal is achieved. The divisive Hierarchical Clustering algorithm is given as [43].

Algorithm:

Step 1: Consider the whole image as one cluster.

Step 2: Find the most dissimilar point in the image and divide the image into 2 clusters.

Step 3: Repeat step 2& 3 for each cluster.

Step 4: Form a tree like structure.

Step 5: Continue until the tumour cluster is obtained.

5) *Ant tree algorithm*

The Ant tree algorithm [6] was introduced for MRI brain image segmentation. The general Ant-Tree algorithm cannot be used in MRI Brain image segmentation directly [94, 46]. The general Ant-Tree algorithm produces a hierarchical structure in an incremental manner in which the ants join together. In this algorithm, each ant represents as a single datum from the data set and it moves in the structure according to the similarity sim (i, j) with the other ants already connected in the tree under construction. In order to make a partition of the whole data set, a more appropriate decision is made when determining the place in which each ant will be connected, either to the main support (Generates a new cluster) or to another ant (refines an existing cluster).

In improved Ant tree algorithm [46], the tree structure of ant tree algorithm ensures that the new ants don't need to search the entire data set when they are connected to the structure. Usually each sub-tree is directly connected to the support as a cluster. So, the hierarchical structure of the tree contains a large number of redundant information. To overcome this problem, a new method of cluster centre was introduced in [46] to improve the tree model of the ant tree algorithm. In [46] it is proved improved ant tree clustering algorithm can produce better results than k-means and FCM. In addition, the processing speed of the improved ant tree algorithm is much faster than the other two clustering algorithms. But, the algorithm is complex.

Ant Tree Algorithm [46]:

Step1: Find the peaks from the histogram of the image as the cluster centre.

Step2: For each grayscale k, Find the nearest gray scale l:

$$\begin{aligned} \text{Diff}(k,l) &= |\text{Gray}(k) - \text{Gray}(l)| \\ &= \text{Min}[|\text{Gray}(k) - \text{Gray}(l)|] \end{aligned}$$

Step3: If  $\text{Diff}(k, l) < T$  then merge the two cluster centre and use greedy algorithm to update the centre.

Step4: Repeat step 2 to 3 until no cluster centre can be merged.

The ant tree algorithm successfully segments the MR brain into white matter, gray matter and cerebrospinal fluid in a better way than FCM and k-means algorithm [46]. In addition, the processing speed of the ant tree

algorithm is better than the other two algorithms (FCM, k-means) which are suitable for segmenting large scale MR brain images instantly.

#### F. Artificial Neural Network

Neural Network based segmentation is totally different from conventional segmentation algorithms. In this, an image is firstly mapped into a Neural Network. Where every Neuron stands for a pixel [37, 58], thus image segmentation problem is converted into energy minimization problem. The neural network was trained with training sample set in order to determine the connection and weights between nodes. Neural network segmentation includes two important steps feature extraction and image segmentation based on neural network. Feature extraction is very crucial as it determines input data of neural network [100], firstly some features are extracted from the images, such that they become suitable for segmentation and then they were the input of the neural network. All of the selected features compose of highly non-linear feature space of cluster boundary.

Neural network based segmentation have three basic characteristics:-

- (i). High parallel ability and fast computing [37].
- (ii). Improve the segmentation results when the data deviates from the normal situation [100].
- (iii). Reduced requirement of expert intervention during the image segmentation process.

However there are some drawbacks of neural networks based segmentation either, such as:-

- (a). Some kind of segmentation information should be known beforehand.
- (b). Neural network should be trained using learning process beforehand [37]
- (d). Period of training may be very long, and overtraining should be avoided at the same time.

Based on the architecture, Artificial Neural Network can be grouped into 2 categories.

- i. Feed-Forward Network
- ii. Feed-Backward Network or Recurrent Network

In feed-forward network, the neurons are arranged in layers that have unidirectional connections between them. Feed forward networks produce only one set of output values. It is called as static network because, the output values does not depend on previous output values. The output values are produced only based on current input. Feed forward network is also called as memory less network. In feedback network, the neurons are arranged in layers that have bidirectional connections between them [68, 101]. Feedback networks produce a set of values which is updated based on the output values fed. Feedback network is also known as dynamic network because the output values depend up on the previous state values.

##### 1) Back Propagation Algorithm:

Back Propagation algorithm [32, 65] is used in layered feed-forward ANN. This algorithm works for feed-forward networks with continuous output. In this network, the neurons are organized in layers and send their signals in the forward direction. The errors generated are propagated in the backward direction. The network receives the input by neurons in the input layer and the output of the network is given by the neurons on an output layer. The network consists of one or more intermediate hidden layers. Supervised learning is used in Back Propagation algorithm i.e., Examples of both the input and the output to be computed is given to the algorithm. The error between the input and the computed output is calculated.

The network is trained with random weights and later the weights are adjusted to get the minimal error. The network will be perfect if the error is minimal. In back propagation, the weights and thresholds are changed each time an example is presented, such that the error gradually reduces. This is repeated until there is no change in the error.

In ANN, the activation function of the artificial neuron is a weighted sum.

$$A_j(\bar{x}, \bar{w}) = \sum_{i=0}^n x_i W_{ji} \quad (53)$$

The most commonly used output function, sigmoid function is given as

$$O_j(\bar{x}, \bar{w}) = \frac{1}{1 + e^{-A_j(\bar{x}, \bar{w})}} \quad (54)$$

The sigmoid function is '1' for +ve number 0.5 for zero and '0' for -ve number. The Error function for the output of each neuron is given as

$$E_j(\bar{x}, \bar{w}, \bar{d}) = (O_j(\bar{x}, \bar{w}) - d_j)^2 \quad (55)$$

The Error function of the network is the sum of errors of all the neurons in the output layer.

$$E_j(\bar{x}, \bar{w}, \bar{d}) = \sum_j (O_j(\bar{x}, \bar{w}) - d_j)^2 \quad (56)$$

After finding the error depending up on the output, input and weights, the weights are adjusted based on gradient descent.

$$\Delta W_{ji} = -\eta \frac{\partial E}{\partial W_{ji}}$$

$$\Delta W_{ji} = -2\eta(O_j - d_j)O_j(1 - O_j)x_i \quad (57)$$

A maximum likelihood method was compared with a neural network technique for tissue classification [18]. In this study, neural network provided better boundary definition than maximum likelihood. Reference [28] compared MR-image segmentation techniques based on supervised multilayered neural networks [65] and the unsupervised fuzzy C-means algorithm [9]. These segmentation techniques were tested on MR images from healthy volunteers and selected patients with brain tumours surrounded by edema. The supervised and unsupervised segmentation techniques used in this study produced broadly similar results.

MR image segmentation based on feed forward neural networks relies heavily on the training set used for their supervised training. The training set is constructed by selecting feature vectors from a single MR image or an ensemble of MR images and reflects the judgment of the human experts who assign labels to the feature vectors according to the tissues they represent [31]. Many researchers proved that the multi-layer propagation with 3 layers can perform arbitrarily complex classification.

## 2) SOM

Self-organizing maps (SOM) is an unsupervised clustering network that maps inputs which can be high dimensional to one or two dimensional discrete lattice of neuron units [80]. The input data is organized into several patterns according to a similarity factor like Euclidean distance and each pattern assigns to a neuron. Each neuron has a weight that depends on the pattern assigned to that neuron [80]. Input data is classified according to their grouping in input space and neighbouring neuron. SOP consists of two layers: Input layer and competitive layer. The number of neurons in first layer is equal to dimension of input and the number of neurons in the second layer is equal to number of classes or clusters.

Each connection from input layer to a neuron in competitive layer is assigned with a weight factor. The SOM functions in two steps. Firstly, finding the winning neuron i.e. the most similar neuron to input by a similarity factor like Euclidean distance, and secondly, updating the weight of winning neuron and its neighbour pixels based on input.

The famous unsupervised approach developed by [41] uses a linear update rule for the weights. This network is used for classification of tissues in MR brain images. Reference [15] previously introduced kohonen's SOM for segmentation of dual-echo MR images. When [15] used the SOFM of kohonen, only normal tissues (white matter, gray matter and cerebrospinal fluid) were classified in the MR brain images. Their technique was not able to classify the abnormal tissue in the brain images. To overcome this drawback, [4] classified the unknown tissue as a tumour.

## G. Hybrid Techniques

Several hybrid neuro-fuzzy approaches for MRI brain image analysis are reported in the literature. A combinational approach of SOM, SVM and fuzzy theory implemented by [35] performed superiorly when compared with other segmentation techniques. The SOM is combined with FCM for brain image segmentation [54] but this technique is not suitable for tumours of varying size and convergence rate is also very low.

A hybrid approach such as combination of wavelets and support vector machine (SVM) for classifying the abnormal and the normal images is used by [13]. This report revealed that the hybrid SVM is better than the kohonen neural networks interms of performance measures. But, the major drawback of this system is the small size of the dataset used for implementation and the classification accuracy results may reduce when the size of the dataset is increased. A modification of conventional SVM such as Least square SVM for brain tumour recognition is proposed by [49].

A Hybrid approach for pattern classification is reported by [47]. The combination of SVM and fuzzy rules is experimented in this work. The results revealed that the proposed hybrid approach is accurate, fast and robust.

#### IV. CONCLUSION

Many image segmentation methods have been developed in the past several decades for segmenting MRI brain images, but still it remains a challenging task. A segmentation method may perform well for one MRI brain image but not for the other images of same type. Thus it is very hard to achieve a generic segmentation method that can be commonly used for all MRI brain images. In this work, the merits and demerits of various automated techniques for brain tumour identification is analyzed in detail. Several novel hybrid approaches may be developed through the ideas conveyed in this paper. The survey shows that BPN classifier gives fast and accurate classification that can be effectively used for segmenting MRI brain images with high level of accuracy.

#### REFERENCES

- [1] N. Abdullah, U. K. Ngah, S. A. Aziz, "Image Classification of brain MRI using Support Vector Machine," *Imaging systems and Techniques (IST)*, 2011, IEEE International conference on 17-18 May 2011.
- [2] S.T Acton, D.P Mukherjee, "Scale space classification using area morphology," *IEEE Trans Image Process* 9(4), 2000, pp.623-635.
- [3] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, T. Moriarty, "A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data," *IEEE transaction medical imaging*, Vol.21, No.3, 2002, pp.193-199.
- [4] Javad Alirezaie, M. E Jernigan, C. Nahmias, "Automatic segmentation of cerebral MR images using Artificial Neural Network," *IEEE transactions on nuclear science*, 1998, vol 45, no.4.
- [5] Anam., Mustaqeem., Ali Javed.,Tehseen Fatim, "An Efficient brain tumour detection algorithm using watershed and thresholding based segmentation," *International journal of image, Graphics and Signal processing(IJIGSP)*, Vol.4, No.10, 2012, pp. 34-39.
- [6] H. Azzag, N. Monmarche, M. Slimane, G. Venturini, "Ant Tree: A New model for clustering with Artificial Ants," *IEEE*, 2003, pp.2642-2647.
- [7] E. F. Badran, E. G. Mahmoud, N. Hamdy, "An algorithm for detecting brain tumors in MRI images", *Proceedings of the International Conference on Computer Engineering and Systems (ICCES)*, Cairo, 2010, pp: 368 - 373.
- [8] M. A. Balafar, A. R. Ramli, M. I. Saripan, S. Mashohor, "Review of brain MRI segmentation methods," *Artificial Intelligence review*, vol.33, 2010, pp.261-274.
- [9] J. C. Bezdek, "Pattern Recognition with Fuzzy objective function algorithms" New York, 1981.
- [10] D. Bhattacharyya, Kim Tai-hoon, "Brain Tumor Detection Using MRI Image Analysis", *Communications in Computer and Information Science*, Vol: 151, 2011, pp: 307-314.
- [11] B. H. Brinkmann, A. Manduca, R. A. Robb, "Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction," *IEEE T. Med. Imag.*, 17, 1998, pp.161-171.
- [12] S. Chandra, R. Bhat, H. Singh, "A PSO based method for detection of brain tumors from MRI", *Proceedings of the World Congress on Nature & Biologically Inspired Computing*, Coimbatore, 2009, pp. 666 - 671.
- [13] S. Chaplot, L. M. Patnaik, "Classification of magnetic resonance brain images using wavelets as input to support vector machines and neural networks," *Biomedical Signal Processing and Control*, 2006, pp. 86-92.
- [14] S. Chaplot, L. M. Patnaik, "Brain Tumor Diagnosis with wavelets and Support Vector Machine," *proceeding of 3rd international Conference on Intelligent Systems and Knowledge Engineering*, 2008.
- [15] T. Chou, C. Chen, W. Lin, "Segmentation of dual-echo MR images using neural networks", *Proceeding SPIE, medical imaging*, 1993, pp.220-227.
- [16] M. H. Chowdhury, W. D. Little., "Image thresholding techniques" *IEEE pacific Rim conference on communications, computers and signal processing*, proceedings 17-19 may 1995, 1995, pp.585-589.
- [17] L. P. Clarke, R. P. Velthuizen, S. Phuphanich, J. D. Schellenberg, J. A. Arrington, M. Silbiger, "MRI: Stability of Three Supervised Segmentation Techniques", *Magnetic Resonance Imaging*, 11: pp. 95-106, 1993.
- [18] L. P. Clarke, "MR image segmentation using MLM and artificial Neural Network," *Medical physics*, Vol.18, No.3, 1991, pp.673.
- [19] E. Cox, " Fuzzy Modeling and Genetic Algorithms for Data mining and Exploration," Elsevier, 2005.
- [20] R. N. Dave, "Characterization and detection of noise in clustering," *Pattern Recogn Lett* 12(11):657-664, 1991.
- [21] Deeplai Kelkar and Surendra Gupta "Improved Quadtree method for split merge image segmentation", *Emerging Trends in Engineering and Technology*, 2008, ICETET.
- [22] S. Deorah, L. Charles, A. Zita, R. Timothy, "Trends in brain cancer incidence and survival in the United States: Surveillance", *Epidemiology and End results program. Neurosurgery Focus* 2006; 2006, vol.20, pp.1-7.
- [23] K. I. Diamantaras, S. Y. Kung, "Principal Component Neural Networks: Theory and Applications", Wiley, 1996.
- [24] L. Pharm Dzung , Xu Chenyang, Jerry L. Prince, "A Survey of Current Methods in Medical Image Segmentation," *Technical Report JHU /ECE 99-01*, Department of Electrical and Computer Engineering, 1998.
- [25] A. El-Sayed, El-Dahshan, M. S. Abdel-Badeeh, H. Y. Tamer, "A Hybrid Technique for Automatic MRI Brain Images Classification", *digital Signal Processing*, Vol. 20, Issue 2, March 2010, 2010, pp. 433-441.
- [26] Frank Klawonn and Annette keller, "Fuzzy Clustering based on Modified Distance measures", Available: [http://citeseer.istpu.edu/fuzzy\\_clustering\\_62](http://citeseer.istpu.edu/fuzzy_clustering_62)
- [27] Rafael C. Gonzalez, E. Woods. Richard, "Digital Image Processing", Pearson Education., 3<sup>rd</sup> Edition.2007
- [28] L. O. Hall, A. M. Bensaid, L. P. Clarke, R. P. Velthuizen, M. S. Silbiger, J. Bezdek. "A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain," *IEEE Trans Neural Network*, Vol.3, 1992, pp.672-682
- [29] Hassan Khotanlou, Olivier Colliot, Jamal Atif, Isabelle Bloch,(2009): "3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models", *Fuzzy Sets and Systems*, Vol: 160, No: 10, pp: 1457-1473.
- [30] R. M. Haralick, I. Dinstein and K. Shanmugam, "Textural features for image Classification," *IEEE Trans on systems man and cybernetics*. Vol. 3, 1973, pp. 610 - 621.
- [31] N. Jabbar, M. Mehrotra, "Application of fuzzy neural network for image tumor description," *Proc. of World Academy of Science, Engineering and Technology*; 34, 2008, pp.575-77.
- [32] A. K. Jain, J. Mao, "Artificial Neural Networks: A Tutorial", *IEEE Press*, 1996, pp. 31-44
- [33] J. Jaya, K. Thanushkodi, "Exploration on selection of medical images employing New Transformation Techniques," *IJCSI International journal of computer science issues*, Vol.7, Issue 3, No.4, 2010.
- [34] X. Jiang, R. Zhang, S. Nie, "Image Segmentation Based on PDEs Model: a Survey", *IEEE conference*, pp. 1-4, 2009.
- [35] C. Juang, S. Chiu, S. Chang, "A self organizing TS-type fuzzy network with support vector learning and its application to classification problems," *IEEE Trans. on Fuzzy Systems* 2007; 15, 2007, pp.998-1008.
- [36] H. G. Kaganami, Z. Beij, "Region Based Detection versus Edge Detection", *IEEE Transactions on Intelligent information hiding and multimedia signal processing*, 2009, pp. 1217-1221.

- [37] W. X. Kang, Q. Q. Yang, R. R. Liang, "The Comparative Research on Image Segmentation Algorithms", IEEE Conference on ETCS, 2009, pp. 703-707.
- [38] P. Karch, I. Zolotova, "An Experimental Comparison of Modern Methods of Segmentation", IEEE 8th International Symposium on SAMI, 2010, pp. 247-252.
- [39] H. B. Kekre, K. Sarode Tanuja., M. Gharge Saylee, "Detection and Demarcation of Tumor using Vector Quantization in MRI images", IJEST, Vol: 1, No: 2, 2009, pp: 59-66.
- [40] A. Kharrat, K. Gasmi, M. B. Messaoud, N. Benamrani, M. Abid, "A hybrid Approach for Automatic classification of Brain MRI using Genetic Algorithm and Support vector Machine," Leonardo journal of sciences, ISSN: 1582-0233, 2009, pp.71-82, issue-17.
- [41] T. Kohonen. "Self organizing maps," springer-verlag berlin, 5.
- [42] S. Koley, A. Majumder, "Brain MRI segmentation for tumor detection using cohesion based self merging algorithm", Proceedings of the IEEE 3rd International Conference on Communication Software and Networks (ICCSN), Xi'an, 2011, pp: 781 - 785.
- [43] Kshitij bhagwat, Dhanshri more, sayali shinde, Akshay Daga, "Comparative study of brain tumour detection using k means, Fuzzy c means and Hierarchical clustering algorithms," International journal of scientific and engineering research, vol. 4, issue 6, 2013, pp. 626-632.
- [44] K. V. Leemput, F. Maes, D. Vandermeulen, P. Suetens, "Automated model-based tissue classification of MR images of brain," IEEE Transaction medical imaging, Vol.18, 1999, pp.897-908.
- [45] Lei Jiang and Wenhui Yang, "A modified Fuzzy C-means Algorithm for segmentation of Magnetic Resonance Images" Proc. VII digital Image Computing: Techniques and Applications, PP. 225-231, 10-12 Dec. 2003. Sydney.
- [46] Li chenling, Zeng wenhua, Zhuang Jiahe, "An Improved Ant Tree Algorithm for MRI Brain Segmentation," Proceedings of 2008 IEEE International symposium on IT in medicine and Education, pp.679-683.
- [47] C. Lin, C. Yeh, S. Liang, J. Chung, N. Kumar, "Support-vector based fuzzy neural network for pattern classification," IEEE Trans. on Fuzzy Systems; vol. 14, 2006, pp.31-41.
- [48] P. L. Lions, J. M. Morel, T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion". SIAM J Numer Anal 29(1), 1992, pp.182-193
- [49] J. Luts, A. Heerschap, A. Johan, S. Huffel, "A combined MRI and MRSI based multiclass system for brain tumour recognition using LS-SVMs with class probabilities and feature selection," Artificial Intelligence in Medicine; vol. 40, 2007, pp.87-102.
- [50] M. Mancas, B. Gosselin, B. Macq, "Segmentation using a Region Growing Thresholding," Proceedings of the Electronic imaging Conference of the International society for optical imaging (SPIE/EI 2005), San Jose, 2005.
- [51] G. Mirajkar, B. Barbadekar, "Automatic Segmentation of Brain tumour from MR images using undecimated wavelet transform and Gabor wavelets," IEEE international conference on Electronics, circuits and systems, pp.702-705, 2010.
- [52] R. Mishra, "MRI based brain tumor detection using wavelet packet feature and artificial neural networks", Proceedings of the International Conference and Workshop on Emerging Trends in Technology, 2010.
- [53] N. Moon, E. Bullitt, K. Leemput, G. Gerig, G "Model based brain and tumor segmentation," Int. Conf. on Pattern Recognition 2002, 2002, pp.528-531,
- [54] S. Murugavalli, V. Rajamani, "An improved implementation of brain tumor detection using segmentation based on neuro fuzzy technique," 2007, Journal of Computer Science; 3:841-46.
- [55] V. J. Nagalkar, S. S. Asole, "Brain tumour detection using digital image processing based on soft computing," Journal of signal and image processing, vol.3, Issue 3, 2012, pp.102-105.
- [56] O. Wirjadi (2007): "Survey of 3D image Segmentation Methods," Technical report.
- [57] M. F. B. Othman, N. B. Abdullah, N. F. B. Kamal, " MRI brain classification using Support vector machine," IEEE, Centre for Artificial Intelligence & Robotics (CAIRO), Universiti Teknologi Malaysia, 2011.
- [58] N. R. Pal, S. K. Pal, "A Review on Image Segmentation Techniques", Pattern Recognition, Vol. 26, No. 9, 1993, pp. 1277-1294.
- [59] Pan Lin ,Yong Yang., Chong-Xun Zheng., Jian-Wen Gu "An Efficient Automatic Framework for Segmentation of MRI Brain Image", Proceedings of the Fourth International Conference on Computer and Information Technology(CIT'04), IEEE, 2004.
- [60] Panos Kotsas, "Non-rigid Registration of medical image using an Automated method," World Academy of Science, Engineering and Technology, 2005.
- [61] D. Pham, C. Xu, L. Prince, "A survey of Current methods in medical image segmentation", Annual review of Biomedical Engineering, vol 2, 1998, pp. 315-338.
- [62] D. Pham, C. Xu, L. Prince, "Current methods in medical image segmentation", Annual review of Biomedical Engineering, 2000, pp. 315-337.
- [63] Qurat-ul Ain., Irfan Mehmood., M. Naqi Syed., M. Arfan Jaffar, "Bayesian Classification Using DCT Features for Brain Tumor Detection", Lecture Notes in Computer Science, Vol:6276, 2010, pp: 340-349.
- [64] J. Rogowska, "Overview and Fundamentals of medical image segmentation," Handbook of medical imaging: Processing and Analysis Management, I.Bankman, Ed.pp.69-85, 2009.
- [65] D. Rumelhart, J. McClelland, "Parallel Distributed Processing", MIT Press, Cambridge, Mass.1986
- [66] D. E. Rumelhart, G. E. Hinton, R. J. Williams, "Learning internal representations by error propagation," Parallel Distributed processing, M.A: MIT. 1986, Vol. 1, 1986, pp.318-362.
- [67] P. K. Sahoo, S. Soltani, A. K. C. Wong, Y. C. Chen, "A Survey of thresholding techniques," Computer vision Graphics image process (CVGIP), Vol.41, 1988, pp.233-260.
- [68] D. Selvaraj, R. Dhanasekaran, "MRI brain tumour detection using feed-forward neural network," Proceedings of EXCITE 2012, 2012, pp.171-174.
- [69] D. Selvaraj, R. Dhanasekaran, "Automatic detection of brain tumour from MRI brain image by histogram thresholding and segmenting it using region growing," Karpagam journal of computer science, vol.7, issue 2, 2013, pp.1-7.
- [70] D. Selvaraj, R. Dhanasekaran, "MRI brain tumour detection by histogram and segmentation by modified GVF model," IJECET, Vol.4, issue 1, 2013, pp.55-68.
- [71] D. Selvaraj, R. Dhanasekaran, "MRI brain tumour segmentation using IFCM and comparison with FCM and k-means," Proceedings of NC-Event 2013. 2013, pp. 47-52.
- [72] D. Selvaraj, R. Dhanasekaran, "Segmentation of cerebrospinal fluid and internal brain nuclei in brain magnetic resonance images," IRECOs, Vol.8, issue 5, 2013, pp.1063-1071.
- [73] H. Selvaraj, S. Thamaraiselvi, D. Selvathi, L. Gewali, "Brain MRI slices classification using Least Squares Support Vector Machine," IC-MED, Vol-1., No.1, Issue 1, 20013, pp. 21 of 33.
- [74] M. Sezgin, B. Sankar, "Survey over image thresholding techniques and Quantitative performance evaluation" J. Electron imaging 13(1), 2004, pp.146-165.00
- [75] Shah shen., William sandham., Malcolm granat., Annette sterr "MRI fuzzy segmentation of brain tissue using neighbourhood attraction with neural network optimization" IEEE transactions on information technology in biomedicine, Vol. 9, No.3, 2005.

- [76] P. K. Srimani, Mahesh Shanthi, "A Comparative study of different segmentation techniques for brain tumour detection", IJETCAS, 2013, pp. 192-197.
- [77] N. Sumitra, Rakesh kumar saxena, "Brain tumour classification using Back Propagation Neural Network", International journal of Image, Graphics and Signal Processing, 2013, 2, PP.45-50.
- [78] J. A. K. Suykens, T. J. Van Gestel, De. Brabanter, B. Moor, J. Vandewalle, "Least Squares Support Vector Machines", World Scientific Publishing Co., Singapore, 2002.
- [79] J. A. K. Suykens, J. Vandewalle, J., "Least Squares Support Vector Machine Classifiers", Neur. Proc. Lett. 9(3), 1999, pp.293- 300.
- [80] D. Tian, L. Fan, "A brain MR images segmentation method based on SOM neural network," 1st international conference on bioinformatics and biomedical engineering, 2007, pp 686–689.
- [81] Y. A. Tolias, S. M. Panas, "On applying spatial constraints in fuzzy image clustering using a fuzzy rule-based system," IEEE Signal Process Letter 5(10), 1998, pp.245–247
- [82] V. Vapnik, "Statistical Learning Theory," Wiley, New York, 1998.
- [83] V. Vapnik, "The Nature of Statistical Learning Theory", Springer, New York, 1995.
- [84] V. Vapnik, A. Chervonenkis, "A note on class of perceptron", Automation and Remote Control, 24, 1964.
- [85] V. Vapnik, A. Lerner, "Pattern recognition using generalized portrait method", Automation and Remote Control, 24, 1964.
- [86] S. Varshney, N. Rajpal, R. Purwar, "Comparative Study of Image Segmentation Techniques and Object Matching using Segmentation", Proceeding of International Conference on Methods and Models in Computer Science, 2009, pp. 1-6.
- [87] Wang Hongzhi., Dong Ying, "An Improved image Segmentation Algorithm based on otsu Method," International symposium on photoelectronic detection and Imaging 2007, Proceedings of SPIE, Vol. 6625, 2007.
- [88] J. Wang, J. Kong, L. Yinghua, Q. Miao B. Zhang, "A modified FCM algorithm for MRI brain image segmentation using both local and non-local spatial constraints", Computerized medical imaging and Graphics, Vol.32, 2008, pp.685-698.
- [89] S. Warfield, Duda, "K-Nearest Neighbour Classification", 2001.
- [90] W. M. Wells, W. E. L. Grimson, R. Kikinis, F. A. Jolesz, "Adaptive segmentation of MRI data," IEEE transaction medical imaging, Vol.15, No.4, 1996, pp.429-442.
- [91] Wen-Feng Kuo., Chi-Yuan Lin., Yung-Nien Sun., "Brain MR images segmentation using statistical ratio: Mapping between watershed and competitive Hopfield clustering network algorithms", Computer Methods and Programs in Biomedicine, Vol: 91, No: 3, 2008, pp: 191-198.
- [92] D. Withey, Z. Koles, "Medical image segmentation: Methods and software, Noninvasive functional source imaging of the brain and heart and the international conference on functional biomedical imaging," 6<sup>th</sup> international symposium, 2007, pp.140-143.
- [93] S. Xavierrockiaraj, K. Nithya, R. Maruni Devi, "Brain tumour Detection using Modified Histogram thresholding – Quadrant approach," Journal of computer application s, Vol.5, No.1, 2012, pp.21-25.
- [94] Xiaochun Yang, Weidong zhao, Li pan, "An Improved clustering algorithm based on Ant-Tree," IEEE, 2008, pp.1855-1858.
- [95] Li, Yan, Chi Zheru, "MR Brain Image Segmentation Based on Self-Organizing Map Network", International Journal of Information Technology, vol. 11, No. 8, 2005.
- [96] J. Yang, G. Li, Ye. C. Geng, "Degree prediction of malignancy in brain glioma using support vector machines", Computers in Biology and Medicine 2006; 36, 2006, pp.313-25.
- [97] D. Q. Zhang, S. C. Chen, "A Novel kernelized fuzzy c-means algorithm with application in medical image segmentation," Artificial intelligence, Vol.32, 2004, pp.37-52.
- [98] H. Zhang, J. E. Fritts, S. A. Goldman, "Image Segmentation Evaluation: A Survey of unsupervised methods", computer vision and image understanding, 2008, pp. 260-280.
- [99] K. Zhang, H. X. CAO, H. Yan, "Application of support vector machines on network abnormal intrusion detection", Application Research of Computers, Vol.5, 2006, pp.98-100.
- [100] C. Zhu, C. Ni, Y. Li, G. Gu, "General Tendencies in Segmentation of Medical Ultrasound Images", International Conference on ICICSE, 2009, pp. 113-117.
- [101] D. Selvaraj, R. Dhanasekaran, "Combining Tissue Segmentation and Neural Network for Brain tumour Detection", IAJIT, Vol.12. No.1, In press.