Classification of brain MRI images using support vector machine with various Kernels.

M. Madheswaran and D. Anto Sahaya Dhas*
Centre for Research in Image and Signal Processing, Mahendra Engineering College, Mallasamudram 637503, India
madheswaran.dr@gmail.com
*Department of ECE, Rajas Engineering College, Vadaikkankulam, Tirunelveli 627116, India

Abstract
An enhanced classification system for classification of brain tumor from MR images using association of kernels with support vector machine is developed and presented in this paper. Oriented Rician Noise Reduction Anisotropic Diffusion filter is used for image denoising. A modified fuzzy c-means algorithm termed as Penalized fuzzy c-means algorithm is used for image segmentation. The texture and Tamura features are extracted using GSDM and Tamura method. Genetic algorithm with Joint entropy is adopted for feature selection. The classification is done by support vector machine along with various kernels and the performance is validated. A classification accuracy of 98.83% is obtained using SVM with GRBF kernel.

Keywords: ORNRAD filter, Penalized Fuzzy C-means, genetic algorithm, SVM classifier, MRI images and Brain tumor.

Introduction
In recent years, brain tumor is found as one of the major diseases leading to death of human beings. The MRI is widely used by most of the physicians to identify the brain tumor in the present days. Identification of the tumor regions accurately from the MRI images is considered to be a challenging job for the physicians. The image processing tools are used to accurately classify the tumor. Many researchers reported various techniques for identification of tumor region.

Quantitative analysis on MRI brain image yields significant performance in noise reduction compared to other quality measures. Similarly, an anisotropic diffusion method for ultrasound images was explained using a novel filtering technique that relies on estimation of the standard deviation of the noise proposed by K.Krissian et al.,[1]. From the estimated noise, the metrics of the filter were chosen automatically. This property provided an intuitive filtering by enhancing the convergence rate of the diffusion. The parameters namely planar, volumetric and linear components of the image are combined in this filter.

Anand et al.,[2] proposed a bilateral filtering scheme based on wavelet in order to reduce the noises in MR images. The noisy coefficient utilized was undecimated Wavelet Transform (UWT). The efficiency of the denoising is enhanced by Bilateral filtering of the approximate coefficients. This method of denoising scheme was widely adapted to Rician noise. Moreover, the visualization and the diagnostic quality of the denoised image are enhanced by this filtering scheme. This scheme encompasses greater ability for noise suppression. Among the various denoising methodologies for MR images, tracking algorithm proposed by Jaya et al.,[3] for denoising the MR brain images. According to this algorithm, application of preprocessing technique provides more promising results. The algorithm processes the task in three stages namely, dataset acquisition, preprocessing and noise removal.

Yong Yang[4] proposed a novel fuzzy C-means algorithm named as Penalized fuzzy C-means algorithm. A penalty term is introduced by modifying the objective function of the general FCM. This approach overcomes the noise sensitivity of the FCM. A high speed parallel fuzzy C-means algorithm was introduced by S.Murugavalli et al.,[5]. It combines the advantages of both SFCM and PFCM. This algorithm reduces the execution time for large images.

Earlier methods used Spatial Gray Level Dependence Method (SGLDM) for feature extraction. Though it was successful, the time consumed by this method was higher and also has higher complexity. To reduce the time for computation A.E.Svalos et al.,[6] analyzed the SGLDM

Biomed Res- India 2015 Volume 26 Issue 3
method by a pilot application. A technique was proposed by M.Vasantha et al., [7] for extracting the intensity histogram and Gray Level Co-occurrence Matrix (GLCM) features from MR mammogram image.

Kernel F-force feature selection (KFFS) method was proposed by Kemal Polat et al., [8] for selecting the features. Results showed that the proposed KFFS functions better than the F-score feature selection. Followed by this, Hsieh-Wei Lee et al., [9] proposed a method for extracting the features from the brain images. This was carried out through integrating the Support Vector Machine (SVM) with the feature selection process in the kernel space. A new method was introduced by Bacauskiene.M et al., [10] in order to elect the salient features for classification and further image processing process. Here, paired t-test was used to eliminate the redundant features and a generic search was employed to detect the hyper-parameter, and to elect the salient features.

A texture based classification system was proposed by Sidhu et al.,[11] using SVM and wavelet transform. The core concept of this classification system was to identify and analyze the factors that considerably affect the performance of SVM and wavelet transform during the process of texture classification. A batch type learning vector quantization technique for segmentation was proposed by Min-Shen Yang et al., [12]. It provides good accuracy and quality for the accurate measurement of hippocampus volume in MR images. The methodology was compared with the generalized Kohenen’s competitive learning method.

A hybrid technique based on association rule mining with decision tree algorithm was proposed by P.Rajendran et al., [13] for the classification of tumor from CT images. The image was segmented using morphological opening and edge detection techniques. Mubashir Ahmad et al., [14] reported a classification technique based on SVM classifier. The feature extraction was carried out by DAUB-4 wavelet and the features are selected using PCA. Linear kernel and Radial basis kernel functions of SVM were used for the classification. A performance analysis of the SVM classifier on brain tumor diagnosis was reported by P.Shantha Kumar et al., [15]. In this work the GLCM features, LBP features, Gray level features and Wavelet features are extracted and trained using SVM.

Amita Kumari et al., [16] proposed a hybrid method using PSO and SVM. Here the wavelet based texture features are extracted using HAAR wavelet. The features are then selected using PSO and trained with SVM. Anamika Ahirwar [17] provided information about the various segmentation and computation techniques which suits for the classification of brain MRI. Here the results were compared with the Keith’s database to show the predic-

Brain tumor classification using GA and SVM

The general flow diagram for the proposed work is shown in fig.1. The MR brain image under testing undergoes various processes such as preprocessing, feature extraction, feature selection and classification.

![Flow graph of the brain tumor classification system](image)

The preprocessing starts with the denoising of MR images. It is accomplished by Oriented Rician Noise Reduction Anisotropic Diffusion filter. It is a modified form of
anisotropic diffusion filter which removes the rician noise in MRI images without affecting the quality of the interesting regions and preserves the edges. The MRI images are denoted as random vector variables in which the unknown random vector variables are denoted as \( \hat{x} \) and \( x \), while the known random variables are denoted as \( y \). The estimation \( \hat{x} \) of \( x \) is a function of \( y \).

The estimated error vector \( (e) \) is equal to the difference of \( \hat{x} \) and \( x \), while the MSE is equal to the trace of error covariance matrix.

\[
\text{MSE} = \text{tr} \{E[(\hat{x} - x)(\hat{x} - x)^T]\} \tag{1}
\]

where \( T \) is the structural tensor. A Linear MMSE (Minimum Mean Square Error) estimator is applied to minimize the MSE in the MRI images. An extension is applied to the matrix diffusion which possess the statistical properties of the local structure in the image which further improves the overall performance of the filter. Structure tensor can be described by the outer product of the gradient, smoothened by a Gaussian convolution as

\[
T_{\sigma_2} = (\nabla_2 \sigma_2)(\nabla_2 \sigma_2)^T
\tag{2}
\]

Here \( \nabla_2 \sigma_2 \) is a Gaussian kernel of standard deviation \( \sigma_2 \) and \( \nabla_2 \sigma_2 \) is the gradient of the output signal. Let, \( e_1 \geq e_2 \geq e_3 \geq 0 \) are eigenvalues and \( E_1, E_2, \) and \( E_3 \) are corresponding eigenvectors. The eigenvectors \( E_1 \) and \( E_3 \) provide the local orientation of maximal and minimal intensity variation respectively.

A diffusion matrix \( DM \) is designed to share eigenvectors \( E_i, i \in [1, N] \) of \( T_{\sigma_2} \), while the eigenvectors \( E_1, E_2 \) and \( E_3 \) are related to leveling of noise. Then the eigenvectors are written as

\[
E_1 = 1 - K
\tag{3a}
\]

\[
E_2 = 1 - K + \frac{3}{2}(1 - K_p)
\tag{3b}
\]

\[
E_3 = 1 - K + \frac{3}{2}(1 - K_p) + 3(1 - K_3)
\tag{3c}
\]

where \( K \) is the gain coefficient in a local isotropic neighborhood, \( K_3 = K(\langle OP \rangle_{E_3} \langle E_3 \rangle, \text{Var}(OP)_{E_3} \langle E_3 \rangle) \) is the gain in the local planar neighborhood defined by the eigenvectors \( E_2 \) and \( E_3 \), and \( K_3 = K(\langle OP \rangle_{E_3} \langle E_3 \rangle, \text{Var}(OP)_{E_3} \langle E_3 \rangle) \) is the gain in the local linear neighborhood oriented by eigenvector \( E_3 \). Then the local mean value can be calculated as

\[
\langle OP \rangle_{E_2} = \frac{1}{2} \sum_{i=0}^{2} \langle OP \rangle_{E_i} \langle E_i \rangle
\tag{4}
\]

\[
\langle OP \rangle_{E_3} = \frac{1}{7} \sum_{i=0}^{3} \langle OP \rangle_{E_i} \langle E_i \rangle
\tag{5}
\]

Here same sets of neighborhood are used to compute the local variance. The diffusion matrix (DM) can be written upon the basis of \( (E_1, E_2, E_3) \) as

\[
DM = \begin{bmatrix}
E_1 & 0 & 0 \\
0 & E_2 & 0 \\
0 & 0 & E_3
\end{bmatrix}
\tag{6}
\]

Equation (7) shows the corresponding diffusion equation as a sum of three diffusion terms.

\[
\frac{\partial \langle OP \rangle}{\partial t} = \nabla \cdot (D \nabla \langle OP \rangle) = 6 \left[ \frac{1}{5} \nabla \cdot (1 - K_p) \nabla \langle OP \rangle + \frac{3}{2} \nabla \cdot (1 - K_p) \nabla \langle OP \rangle + \frac{1}{5} \nabla \cdot ((1 - K_3) \nabla \langle OP \rangle) \right]
\tag{7}
\]

in an efficient reduction of noise simultaneously preserving the detailed components and the high definition of the interface between various brain tissues.

Segmentation is an important process in MRI image classification in which the testing area is extracted from the whole image. Here it is done by adopting an algorithm termed as Penalized Fuzzy C-Means algorithm (PFCM) which is a modified FCM algorithm. In conventional FCM algorithm, there is no information about spatial context, which cause it to be sensitive to noise and imaging artifacts. In PFCM, a penalty term considers the spatial dependence of the object which is inspired by the Neighborhood Exception maximization algorithm and is modified according to the FCM criterion. The objective function of the PFCM is given by

\[
I_{PFCM} = \sum_{k=1}^{N} \sum_{i=1}^{M} (u_{ik})^q d^2(x_i, v_k) + \gamma \sum_{k=1}^{N} \sum_{j=1}^{M} (u_{ij})^q (1 - u_{ij})^q w_{kj}
\tag{9}
\]

where

\[
w_{kj} = \begin{cases}
1, & \text{if } x_k \text{ and } x_j \text{ are neighbors and } k \neq j \\
0, & \text{otherwise}
\end{cases}
\tag{10}
\]
\(\gamma(\geq 0)\) Controls the effect of the penalty term.

The objective function \(J_{PFM}\) can be minimized with the help of an iterative algorithm which is derived by the valuation of the centroids and membership functions that satisfy a zero gradient condition. When the PFCM algorithm converges, a defuzzification process takes place to convert the fuzzy partition matrix into a crisp partition which is the segmented image.

Visual content in the image are captured by extracting the features present in the image. This information is useful for the indexing and retrieval. Visual features are low-level or primitive features, which can be either domain-specific such as human face, finger prints, etc. or general features like color, shape, and texture. Features can be extracted based on color, texture, or shape. The texture based features are extracted in the present work. The texture can be characterized by structure (spatial relationship) and tone (intensity property). This is implemented through the GSDM and Tamura method. Tamura features of a preprocessed image can be retrieved through constructing a co-occurrence matrix named Gray Level Co-occurrence Matrix (GLCM) also known as GSDM.

This matrix is a two dimensional matrix \(P_d\), which holds the information about the position of the pixel that has similar brightness value (gray levels). \(P_d\) considers and specifies the relationship among the reference and neighbour pixel at a time. Neighbour pixels, which are presented at the right side of the reference pixel, which is expressed using \(P(i, j)\) relation. It has \(n \times n\) dimension, where \(n\) denotes the number of gray levels in the image. Number of occurrence of the image can be represented in the matrix as \(n_x n\). Likewise, value \((i, j)\) is lying at distance \(d\) in the image. For example, if \(d = (1, 1)\) then they are found from the matrix as shown in Fig. 2.

It contains 16 pairs of pixels that satisfies the separation in spatial. In the example as there are only 3 gray levels the \(P(i, j)\) is of \(3 \times 3\) matrix. Following steps are carried out to construct the co-occurrence matrix.

1. Pairs of pixel are counted, in which the value of \(i\) (i.e. the first pixel) and \(j\) (i.e. \(i\)'s matching pixel) are situated at a distance \(d\).

2. The computed count is inserted into the matrix at \(i^{th}\) row and \(j^{th}\) column Matrix is not symmetric, due to the number of pixel pairs having gray level is not necessarily equal to the number of pairs of pixel having gray level.

\[
\begin{array}{ccc}
2 & 1 & 2 & 0 & 1 \\
0 & 2 & 1 & 1 & 2 \\
0 & 1 & 2 & 2 & 0 \\
1 & 2 & 2 & 0 & 1 \\
2 & 0 & 1 & 0 & 1 \\
\end{array}
\]

\[
P_d=
\begin{bmatrix}
0 & 2 & 2 & 0 \\
2 & 1 & 2 & 1 \\
2 & 3 & 2 & 0 \\
1 & 2 & 2 & 1 \\
\end{bmatrix}
\]

\(\text{Figure 2. Co-occurrence matrix}\)

Different texture and Tamura features can be extracted to predict the risk level of the tumor. Each feature value can be computed from the matrix constructed using their corresponding formulas and they are used to analyze different properties of an image separately. These features explain the spatial ordering of texture constituents.

(i) Contrast

Local variance present in the brain tumor MRI can be measured using contrast. If \(P(i, j)\) in the matrix has more variations then the contrast will be high. The contrast value from the matrix can be obtained from

\[
c_{\text{con}}(x, y) = \sum_i \sum_j (t - f)^2 P(i, j) y
\]
(ii) Homogeneity
The homogeneity of an image can be found from the combination of low and high values of $P(i,j)$ in the co-occurrence matrix. This feature results in spreading the $P(i,j)$ values evenly in the matrix. Mathematically homogeneity can be expressed as

$$C_{	ext{homogeneity}} = \sum_i \sum_j \frac{P(i,j)}{Z_{i,j}}$$

(12)

(iii) Entropy
This is a measure of information content and randomness of intensity distribution. When all the entries of the matrix are of same magnitude then entropy is higher otherwise, it is smaller. The entropy can be measured using the equation (13).

$$C_{\text{Entropy}} = -\sum_i \sum_j P(i,j) \ln P(i,j)$$

(13)

(iv) Correlation
Intensity of an image is measured using Correlation through the equation (14). If an image contains a considerable amount of linear structure then the correlation value will be higher.

$$C_{\text{correlation}} = \sum \sum \frac{IP(i,j) - c_1}{\sigma_1} \frac{IP(i,j) - c_2}{\sigma_2}$$

(14)

Where, $c_1 = \sum IP(i,j)$ and $\sigma_1^2 = \sum i^2 P(i,j) - c_1^2$.

(v) Energy
The Texture energy is measured by

$$C_{\text{Energy}} = \sum_i \sum_j (C(i,j))^2$$

(15)

(vi) Maximum Probability
This feature corresponds to the strongest response. This can be expressed mathematically as

$$C_{\text{PM}} = \max C_{\text{PM}}$$

(16)

(vii) Local Homogeneity, Inverse Difference Moment (IDM)
This is influenced by the feature homogeneity of an image. IDM values are low for the inhomogeneous images and high for homogeneous images. It can be measured as

$$C_{\text{IDM}} = \sum_{i=1}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1+(i-j)^2} P(i,j)$$

(17)

(viii) Sum of square, variance
The element whose values differ greatly from the $P(i,j)$’s average value than for such elements the feature values are relatively high value. It can be computed as

$$C_{\text{var}} = \sum_{i=1}^{G-1} \sum_{j=0}^{G-1} (i - \bar{i})^2 P(i,j)$$

(18)

(ix) Auto correlation
Coarseness and regularity of an images texture can be analyzed through Kaizer. Spatial relationship among the primitives can be measured using

$$C_{\text{Kar}}(p,q) = \frac{MN}{(M-p)(N-q)} \sum_{x=p}^{M-p} \sum_{y=q}^{N-q} (IP(x,y) - \bar{IP})^2$$

(19)

(x) Directionality
Total degree of directionality is calculated for the neighbours that are non-overlapping using

$$C_{\text{Dir}} = \sum_{p} \sum_{q} (\theta - \phi)^2 H_{\theta}(0)$$

(20)

(xi) Coarseness
This feature is calculated for each pixel $(x, y)$ in the image using the equation (21). This represents the direct relationship to the repetition rate and scale.

$$C_{\text{coarseness}}(x,y) = \sum_{i=x-2}^{x+2} \sum_{j=y-2}^{y+2} f(i,j)/2u$$

(21)

Other features
1. Cluster Shade:
$$C_{\text{CSh}} = \sum_{i=x-2}^{x+2} (i - \bar{i})^3 H_{\bar{i}}$$

(22)

Where, $\bar{i} = \frac{1}{2} \sum_{i=x-2}^{x+2} i H_{\bar{i}}$

2. Cluster Prominence:
$$C_{\text{CPrem}} = \sum_{i=x-2}^{x+2} \sum_{j=y-2}^{y+2} (i + j - \bar{i} - \bar{j})^6 \times P(i,j)$$

(23)

3. Inertia :
$$I_{i,j} = \sum_{i=x-2}^{x+2} \sum_{j=y-2}^{y+2} (i - j)^2 \times P(i,j)$$

(24)

4. Cluster tendency:
$$C_{\text{Tend}} = \sum_{i=x-2}^{x+2} \sum_{j=y-2}^{y+2} (i + j - 2\mu)^2 \times P(i,j)$$

(25)

The subset of features can be selected for dimensionality reduction from the extracted features. This is normally carried out in order to eliminate redundant and irrelevant features that are extracted. The selection process is carried out using joint entropy and genetic algorithm. The entropy is estimated for the features and for the target of the selected image that required to be predicted. This value is determined from a grayscale image, which measures the randomness present in the image to characterize the input image’s texture. Having entropy values determined the mutual information among the features and targets are determined. This information is used to estimate and measure how a random variable is able to describe and impact on other variable.

With this statistical information, genetic algorithm is implemented to estimate the fitness values for all the extracted features from the population initialization. The determined fitness values are analyzed to determine whether the minimum relevance and maximum redundancy exits between the features that are extracted. If it fails to choose a feature subset then they are processed again by the genetic algorithm. Otherwise, the chosen features are grouped to form a subset based on which prediction process is carried out.

Classification is the process of classifying the given input by training with a suitable classifier. Support Vector Machine (SVM) classifier is one of the best classifier suggested by many researchers which can be opted for the brain tumor classification from MR images. It is independent of dimensionality and feature space. SVM transforms the input space to a higher dimension feature space through a non-linear mapping function and construct the separating hyperplane with maximum distance from the closest points of the training set. SVM classifier along
with linear and non-linear kernel functions produces best results in classification. It is found that the seven kernels mentioned here can be used for MR image classification.

1. Linear kernel:
The linear kernel is the simplest kernel. It is defined by
\[ K(x_i, x_j) = x_i^T x_j \] (26)

2. Polynomial kernel:
The polynomial kernel is suited for problems with normalized training data. It is given by
\[ K(x_i, x_j) = (\alpha x_i^T x_j + c)^d \] (27)
where \( \alpha \) is the slope, \( c \) is a constant and \( d \) is the polynomial degree.

3. Gaussian Radial Based Function kernel:
The GRBF kernel is defined as
\[ K(x_i, x_j) = \exp \left( -\frac{||x_i - x_j||^2}{2\sigma^2} \right) \] (28)
where \( \sigma \) is an adjustable parameter. It is a non-linear kernel and is very sensitive to noise.

4. Exponential Radial Based Function kernel:
The ERBF kernel is close to GRBF with only the square of the norm removed. It is defined as
\[ K(x_i, x_j) = \exp \left( -\frac{||x_i - x_j||^2}{2\sigma^2} \right) \] (29)

5. ANOVA kernel:
The ANOVA kernel is also a radial based function kernel. It is defined as
\[ K(x_i, x_j) = \sum_{k=1}^{n} \exp(-\sigma(x_i^k - x_j^k)^2)^d \] (30)

6. Multilayer Perceptron kernel:
Multilayer Perceptron kernel is also called as Hyperbolic Tangent kernel or Sigmoid kernel. It is defined as
\[ K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r) \] (31)
where \( \gamma \) is the slope and \( r \) is the intercept constant.

7. Fisher kernel: Fisher kernel is defined as
\[ K(x_i, x_j) = U_i^T I^{-1} U_j \] (32)

### Results and Discussion

The classification of MRI brain images using PFCM, GA and SVM with kernels is carried out using image processing tools. The images are taken from the databases namely MR-TIP, NCIGT, BraTS, BITE and TCIA. The images are preprocessed for noise removal, segmented for separation of interesting area and the features are extracted for classification. The performance of the ORNARD filter is estimated and compared with various filters as shown in Table 1 and is further used for validation.

### Table 1. Performance comparison of various filters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>MSE</th>
<th>PSNR</th>
<th>CNR</th>
<th>IQI</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMF</td>
<td>47.6</td>
<td>31.42</td>
<td>53.23</td>
<td>0.935</td>
<td>2.97</td>
</tr>
<tr>
<td>LMMSE</td>
<td>84.8</td>
<td>28.75</td>
<td>45.52</td>
<td>0.15</td>
<td>4.2</td>
</tr>
<tr>
<td>HOHM</td>
<td>85.6</td>
<td>29.02</td>
<td>45.35</td>
<td>0.88</td>
<td>3.62</td>
</tr>
<tr>
<td>NLM</td>
<td>49.2</td>
<td>31.25</td>
<td>58.2</td>
<td>0.89</td>
<td>3.18</td>
</tr>
<tr>
<td>ORNARD</td>
<td>30</td>
<td>33.35</td>
<td>52.56</td>
<td>0.92</td>
<td>1.25</td>
</tr>
</tbody>
</table>

![Figure 3. Noisy input image and denoised output image of ORNARD filter](image-url)

---

510

**Biomed Res- India 2015 Volume 26 Issue 3**
The denoised image is then segmented for extracting the area of interest from the image. The skull region and other unwanted regions were removed using Penal- ized Fuzzy C-means algorithm. The segmentation accuracy of PFCM algorithm for 10 images is compared with other Fuzzy based segmentation algorithms and the result is given in Table 2. The PFCM algorithm performs well by consuming a minimal time compared to other fuzzy algorithms. Then the segmented image is subjected to feature extraction. The extracted texture and Tamura features are then subjected to feature selection which is done by Genetic algorithm along with Joint entropy. The extracted feature values for a sample image are given in Table 3.

Table 2. Performance comparison of various Fuzzy based segmentations

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Number of Misclassified pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FCM</td>
</tr>
<tr>
<td>Image 1</td>
<td>163</td>
</tr>
<tr>
<td>Image 2</td>
<td>87</td>
</tr>
<tr>
<td>Image 3</td>
<td>123</td>
</tr>
<tr>
<td>Image 4</td>
<td>56</td>
</tr>
<tr>
<td>Image 5</td>
<td>73</td>
</tr>
<tr>
<td>Image 6</td>
<td>158</td>
</tr>
<tr>
<td>Image 7</td>
<td>192</td>
</tr>
<tr>
<td>Image 8</td>
<td>133</td>
</tr>
<tr>
<td>Image 9</td>
<td>92</td>
</tr>
<tr>
<td>Image 10</td>
<td>125</td>
</tr>
<tr>
<td>Consuming time (approx..)</td>
<td>5 sec</td>
</tr>
</tbody>
</table>

Table 3. Extracted values for the features

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Features</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Contrast</td>
<td>2.3752</td>
</tr>
<tr>
<td>2</td>
<td>Homogenity</td>
<td>0.7623</td>
</tr>
<tr>
<td>3</td>
<td>Entropy</td>
<td>2.1436</td>
</tr>
<tr>
<td>4</td>
<td>Correlation</td>
<td>0.2136</td>
</tr>
<tr>
<td>5</td>
<td>Energy</td>
<td>0.1561</td>
</tr>
<tr>
<td>6</td>
<td>Maximum probability</td>
<td>0.7539</td>
</tr>
<tr>
<td>7</td>
<td>Inverse difference moment</td>
<td>1.1752</td>
</tr>
<tr>
<td>8</td>
<td>Variance</td>
<td>0.2257</td>
</tr>
<tr>
<td>9</td>
<td>Auto Correlation</td>
<td>47.130</td>
</tr>
<tr>
<td>10</td>
<td>Directionality</td>
<td>0.8349</td>
</tr>
<tr>
<td>11</td>
<td>Coarseness</td>
<td>2.376</td>
</tr>
<tr>
<td>12</td>
<td>Cluster shade</td>
<td>56.23</td>
</tr>
<tr>
<td>13</td>
<td>Cluster prominenec</td>
<td>3.2732</td>
</tr>
<tr>
<td>14</td>
<td>Inertia</td>
<td>0.125</td>
</tr>
<tr>
<td>15</td>
<td>Cluster tendency</td>
<td>43.239</td>
</tr>
</tbody>
</table>

Sensitivity is the probability of positive for a diagnostic test. It is also termed as true positive fraction. The percentage of sensitivity is given by

\[
Sensitivity = \frac{T_p}{T_p + F_n} \times 100\%
\]

where \( T_p \) is the True positive and \( F_n \) is the False negative.

Specificity is the probability of negative for a diagnostic test. It is also termed as true negative fraction. The percentage of specificity is given by

\[
Specificity = \frac{T_n}{T_n + F_p} \times 100\%
\]

where \( T_n \) is the True negative and \( F_p \) is the False positive.

Accuracy is the probability that a diagnostic test is correctly performed. It is calculated by

\[
Acc = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \times 100\%
\]

The performance of the SVM classifier with various kernels is given in Table 4. From the table it is found that the SVM classifier with GRBF kernel performs better than the other kernels. It is also estimated 100% sensitivity, 90% specificity and accuracy of 98.83% for the SVM classifier with GRBF kernel. This may be due to the inherent property of gradient in GRBF kernel.

The classifications of the MR image using various kernels are given in Table 5. It is found that the GRBF kernel classifies the tumor more accurately than the other kernels.
Table 4. Performance of various kernels with SVM

<table>
<thead>
<tr>
<th>Kernel used</th>
<th>Total no. of images</th>
<th>TP</th>
<th>FN</th>
<th>TN</th>
<th>FP</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>172</td>
<td>147</td>
<td>3</td>
<td>17</td>
<td>5</td>
<td>98.02</td>
<td>95.34</td>
<td>96.51</td>
</tr>
<tr>
<td>Polynomial</td>
<td>172</td>
<td>149</td>
<td>3</td>
<td>17</td>
<td>3</td>
<td>98.02</td>
<td>95.34</td>
<td>96.51</td>
</tr>
<tr>
<td>GRBF</td>
<td>172</td>
<td>152</td>
<td>0</td>
<td>18</td>
<td>2</td>
<td>98.05</td>
<td>95.34</td>
<td>96.51</td>
</tr>
<tr>
<td>ERBF</td>
<td>172</td>
<td>151</td>
<td>3</td>
<td>14</td>
<td>4</td>
<td>98.05</td>
<td>95.34</td>
<td>96.51</td>
</tr>
<tr>
<td>ANOVA</td>
<td>172</td>
<td>151</td>
<td>2</td>
<td>16</td>
<td>3</td>
<td>98.69</td>
<td>95.34</td>
<td>96.51</td>
</tr>
<tr>
<td>MLP</td>
<td>172</td>
<td>140</td>
<td>8</td>
<td>15</td>
<td>9</td>
<td>94.59</td>
<td>95.34</td>
<td>96.51</td>
</tr>
<tr>
<td>Fisher</td>
<td>172</td>
<td>144</td>
<td>6</td>
<td>16</td>
<td>6</td>
<td>96.00</td>
<td>95.34</td>
<td>96.51</td>
</tr>
</tbody>
</table>

Table 5. Classification of MR image based on the kernels

<table>
<thead>
<tr>
<th>Kernel used</th>
<th>No. of images</th>
<th>Normal</th>
<th>Benign</th>
<th>Malignant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>172</td>
<td>17</td>
<td>103</td>
<td>44</td>
</tr>
<tr>
<td>Polynomial</td>
<td>172</td>
<td>17</td>
<td>104</td>
<td>45</td>
</tr>
<tr>
<td>GRBF</td>
<td>172</td>
<td>18</td>
<td>107</td>
<td>45</td>
</tr>
<tr>
<td>ERBF</td>
<td>172</td>
<td>14</td>
<td>107</td>
<td>44</td>
</tr>
<tr>
<td>ANOVA</td>
<td>172</td>
<td>16</td>
<td>107</td>
<td>44</td>
</tr>
<tr>
<td>MLP</td>
<td>172</td>
<td>15</td>
<td>103</td>
<td>45</td>
</tr>
<tr>
<td>Fisher</td>
<td>172</td>
<td>16</td>
<td>102</td>
<td>44</td>
</tr>
</tbody>
</table>

Table 6. Comparison of Classification Accuracy Obtained by Various Methods

<table>
<thead>
<tr>
<th>Author</th>
<th>Technique used</th>
<th>Classification</th>
<th>Classification Accuracy in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jafari et al [18]</td>
<td>Denoising: Gabor filter</td>
<td>Feature extraction: HAAR wavelet</td>
<td>Feature selection: GA</td>
</tr>
<tr>
<td>El-Dishan et al [19]</td>
<td>-</td>
<td>-</td>
<td>HAAR wavelet</td>
</tr>
<tr>
<td>Mahmood et al [20]</td>
<td>-</td>
<td>-</td>
<td>HAAR wavelet</td>
</tr>
<tr>
<td>Present Work</td>
<td>ORNRAD filter</td>
<td>PFCM</td>
<td>GSDM and Tamura method</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The performance of the developed classification system is compared with the existing techniques is given in the Table 6.

It is seen that the SVM classifiers outperform other classifiers. In particular, the SVM classifier with GRBF kernel produces a higher classification accuracy of 98.83%. This is due to the combination of the ORNRAD filter smoothing characteristics and GRBF kernel for classification.

Conclusion

An efficient brain MR images classification technique using support vector machine with kernels is proposed in this paper. The features are extracted using Penalized fuzzy C-means algorithm and the optimized features are selected using genetic algorithm along with joint entropy. The performance of SVM classifiers with various kernels was estimated and found that the GRBF kernel performs well compared to other kernels. The classification accuracy was also found to be high for SVM classifier with GRBF kernel.

References


Correspondence to:
M. Madheswaran
Centre for Research in Image and Signal Processing
Mahendra Engineering College
Mallasamudram 637503
India