

SVM BASED CLASSIFICATION AND IDENTIFICATION OF **BRAIN TUMOUR FROM MRI USING WAVELET** TRANSFORM

Amrutha. M

Dept. of ECE PSG College of Technology, Coimbatore, Tamilnadu Dr. A. Kannammal Dept. of ECE PSG College of Technology, Coimbatore, Tamilnadu

Abstract-Magnetic Resonance Imaging (MRI) is the most important technique in detecting a brain tumour. The early detection of brain tumours with proper automation algorithm results in assisting doctors to make the decision early for diagnostic purposes. This paper proposes methods to recognize normal and abnormal MR brain images and identification of tumour regions from MRI. MRI is processed through different steps such as image filtering, skull masking, feature extraction, feature reduction and classification. the goal of this paper is to classify MRI brain images into normal and abnormal by SVM. The wavelet transform is used to extract features from images and Principal Component Analysis (PCA) is applied to reduce the dimension of features. These reduced features were given to a Kernel support vector machine. The performance measure of the SVM classifier is done with three different kernels and the tumour region is identified using proper filtering followed by Otsu's binarization method. The results show that the SVM classifier with RBF gives the maximum success rate.

Keywords- Classification, MRI, PCA, SVM, Wavelet Transform DOI Number: 10.48047/nq.2022.20.19.NQ99011 NeuroQuantology 2022; 20(19):104-114

1. INTRODUCTION

Nowadays cancer has become а worldwide public ill-health. There are an estimated 14.1 million cancer cases around the world and this is expected to increase to 24 million by 2035. A brain tumour is the abnormal growth of the cells in the tissues of the brain. Cancer cells will grow without any control and can grow into adjacent tissue. Brain tumours are classified as primary brain tumours and secondary brain tumours. Primary brain tumours are starts in the brain and spinal cord and the latter originates in other parts of the body.

An important view of diagnosis is to provide a differentiated result of these

categories, before further stages of treatments. MRI is the best imaging technique for diagnosing most types of brain tumours and because of the better result in medical analysis of the brain; MRI has opted in many cases. The traditional method in medicine for MR image classification and tumour detection is human inspection. Radiologists observe MRI Images based on visual interpretation to find the presence of tumours. There is a possibility of the wrong diagnosis by the radiologists. Hence the research work here emphasizes the classification and abnormality detection for MR brain images.

Many works are wiped out this area [1], [2], [3], [4], [5], [6]. An automatic



classification and abnormality detection from CT scan brain images by kNN classifier are done in [1]. [2] proposed a new framework called sparse representation segmentation with SVM Classifier. For features, a two level GLCM feature extraction technique is utilized. SVM-kNN hybrid classifier [3], SVM with Wavelet transform [4], feature reduction by PCA [5] are some of the most powerful works in this area. [7] Suggests an efficient method for noise removal from an Image by using the technique Region filling. [8] gives a better idea about wavelet transforms in image processing. [9] proposed a classification technique using feature SVM. Here extraction is performed using GLCM and all the texture features are calculated. They examined linear, Quadratic and Polynomial kernel functions to identify the normal and abnormal brain MRI images.

The proposed method (fig. 1) consist of, Pre-processing such as feature extraction, feature reduction and classification of MR images (SVM) and hence the identification of tumour region. Image pre-processing is used to enhance the quality of images. It includes RGB to greyscale conversion, filtering, morphological operations, feature extraction and feature reduction. It is important to get good quality images for accurate observations since medical images contain different types of noises. The removal of these noises has to be done in such a way that it should not affect the portion of the brain in the image as each portion is the most important part to detect the tumour. Filtering techniques are efficient to avoid noises. The proposed scheme uses Median filtering as it preserves edges efficiently.

Filtering is followed by skull masking which is the process of removal of non-brain tissue (Skull Portion) from MR brain images. Dilation and Erosion are the two morphological operations used for skull masking. This is followed by region filling and image enhancement. Enhanced images provide better contrast to the small print that a picture contains. Feature extraction is one of the important steps in image processing and pattern recognition, which is actually а dimensionality reduction. The process of transforming the input image into a set of features is called feature extraction. Features usually contain information connected to shape, colour, texture or context. In this paper, Wavelet transform is employed for feature extraction [8]. PCA is used for feature reduction. It is a very popular technique for dimensionality reduction [4].

Nowadays several classifiers like k Nearest Neighbour (kNN), Artificial Neural Network (ANN), Support Vector Machine (SVM), Probabilistic Neural Network (PNN), Bayesian Classifiers are used for various applications. In this paper, we used an SVM [2] Classifier with different kernel functions.

2. METHODOLOGY

The proposed method consists of stages,

- Pre-processing: Includes both Feature extraction and Feature reduction
- 2. Classification of MR images (kNN and SVM)
- 3. Identification of tumour region

The block diagram of the proposed method is given in the figure(1) The brain tissue in MRI images can be divided into two main types. Normal tissues including grey matter, white matter and cerebrospinal fluid (CSF) and abnormal tissues usually containing tumours. MR brain images are RGB images so initially,



the images are converted into a greyscale image.







Figure 2: Normal MR Image



Figure 3: Abnormal MR Image

2.1 Pre-processing

Before feature extraction, the image will go through different stages of preprocessing. Image pre-processing is used to get better quality of images.



Figure 4: Pre-processing Stages

2.1.1 Filtering

Filtering is used to remove noise from MR Images. In this research, a median filter is used for noise removal. Median filters are non-linear neighbourhood some operations that can be performed for noise reduction. They preserve edges efficiently than simple smoothing filters. The median filter looks at the neighbourhood of pixels and returns the median values. Thus, without blurring the edges much noise can be removed.

2.1.2 Skull Masking

Skull masking means the removal of nonbrain tissue like scalp, skull, fat, eyes, neck, etc. From MRI brain image. It helps to improve the accuracy and speed of diagnostic and predictive proceedings in medical applications. Dilation and Erosion are the two morphological operations used for skull masking. For erosion and dilation mainly two inputs are needed one is the structuring element and the other is the image to be eroded or dilated. The structuring elements are the tiny binary image that's a little matrix of pixels, with a worth of zero or one for every pixel. There are different types of structuring elements based on their shape and size. Here flat linear structuring element with length 5 is used and is given below,

[1 1 1 1 1] (1) Dilation or Erosion results are influenced by both the size and shape of a structuring



element. Dilation and erosion have opposite effects and they are dual operations.

2.1.3 Image Enhancement

Image enhancement is a very essential image processing task. This process defines us to have a better subjective judgement over the images. Image enhancement simply means, using a transformation T an image f is transforming into an image g. The pixels values in images f and g are denoted by r and s, respectively. They are related by the expression,

s = T(r)

In this work power law transformation is used for image enhancement. Power law transformation function is also known as gamma correction. Different levels of enhancements can be obtained by various values of γ . The expression for transformation is $s = cr^{\gamma}$.

2.1.4 Feature Extraction

Image analysis involves careful examination of image data for a particular application. Feature extraction is one among the important steps in image processing. When the input data is too extensive to be processed and suspected to be redundant then the data is transformed into a reduced feature representation. The process of transforming the input image into a set of features is called feature extraction.

In this work, 3 level decomposition via Harr wavelet was used to extract features.

The discrete wavelet transform (DWT) is an important implementation of the Wavelet

transform (WT) using the dyadic scales and positions. In our algorithm, a 3-level decomposition using Haar wavelet was utilized to extract features. The basics of DWT are introduced as follows. Suppose x(t) is a square integrable function, then the continuous WT of x(t) relative to a given wavelet $\psi(t)$ is defined as,

$$W_{\psi}(a,b) = \int_{-\infty}^{\infty} x(t)\psi_{a,b}(t)dt$$
(2)
Where.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-a}{b}\right)$$
(3)

Here, the wavelet $\psi_{a,b}(t)$ is calculated from the mother wavelet $\psi(t)$ by translation and dilation: a is the dilation factor and b the translation parameter Equation (1) can be discretized by restraining a and b to a discrete lattice $a = 2^b \& n > 0$ to give the DWT, which can be expressed as follows.

$$ca_{j,k}(n) = DS\left[\sum_{n} x(n)g_{j}^{*}(n - 2^{j}k)\right]$$
(4)
$$cd_{j,k}(n) = DS\left[\sum_{n} x(n)h_{j}^{*}(n - 2^{j}k)\right]$$
(5)

 $cd_{i,k}$ mention the Here $ca_{i,k}$ and coefficients approximation of the components and therefore the detail components, respectively. g(n) and h(n)refers to the low-pass filter and high-pass filter, respectively. j and k denote the wavelet scale and translation factors, respectively. DS operator means the downsampling. Equations 4 and 5 are the fundamental of wavelet decomposes. It breakdown signal x(n) into two signals, the approximation coefficients ca(n) and the detail components cd(n). This procedure is called one-level decompose.

The above decomposition process can be iterated with consecutive approximations being decomposed in turn, so that one signal is broken down into various levels of resolution. The



whole procedure is called wavelet decomposition tree.



Figure 5: A 3 level wavelet decomposition tree

2D-DWT: In case of the 2D images, the DWT is applied to each dimension separately. The schematic diagram of 2D DWT is shown in figure 6.



Figure 6: Schematic diagram of 2D DWT

The LL sub band can be consider as the approximation component of the image, while the LH, HL, and HH sub bands can be regarded as the detailed components of the image.

Haar wavelet is discontinuous, and bear a resemblance to a step function. It represents the matching wavelet as Daubechies db1. Haar used these functions to grant an example of an orthonormal system for the space of square-integrable function on the unit interval [0, 1]. The Haar wavelet transformation is an easy form of compression which contain averaging and differencing terms, storing detail coefficients, eliminating data, and reconstructing the matrix such that the resulting matrix is similar to the initial matrix. Haar wavelet function $\psi(t)$ is given by,



Figure 7: Haar wavelet function

2.1.5 Feature Reduction

Excessive features will shoot up computation times and storage space. Furthermore, they sometimes create classification more complicated, which is termed the curse of dimensionality. So it is needed to decrease the number of features. In this work, PCA is applied for feature reduction.

PCA is a technique that takes a collection of data and transforms it such that the new data has given statistical properties. The statistical properties are chosen such that the transformation emphasizes the importance of data elements. Thus, the transformed data can be utilized for classification by observing important components of the data. The steps for performing PCA is as follows,

Algorithm 1

- 1. Obtain the feature matrix C_{χ} from the data.
- 2. Compute the covariance matrix \sum_{x} .

elSSN 1303-5150



- 3. Obtain the eigenvalues by solving the characteristic equation $det(\lambda_i I - \sum_x) = 0.$
- 4. Obtain the eigenvectors by solving for w_i in $(\lambda_i I \sum_x) w_i = 0$.
- 5. The transformation W is obtained by considering the eigenvectors as their columns
- 6. Obtain the transform features by computing $c_y = c_x W^T$.
- 7. For classification applications, select the features with large values of λ_i . For compression, reduce the dimensionality of the new feature vectors by setting to zero components with low λ_i values. Features in the original data space can be obtained by $c_x^T = W^T c_y^T$.

3. CLASSIFIER

Image classification is likely the most important part of digital image analysis. Based on the extracted feature vector the classifiers separate the input data into one or several classes. The classifier used in this project is described below.

3.1 SVM classifier

SVM is one of the most popular approaches to data modelling and classification. This algorithm was first developed in 1963 by Vapnik and Lerner. is a binary classifier supported by supervised learning which provides a better result than other classifiers. SVM classifies among two classes by constructing a hyperplane in high dimensional feature space which can be used for classification. SVM may be a classification algorithm, which is predicated on different kernel methods. SVM is classified into two groups.

3.1.1 Linear SVM

It is the simplest one, in which the training patterns are linearly separable. Given a p-

dimensional N-size training examples of the form,

$$\{(x_n, y_n) \mid x_n \in R^p, y_n \\ \in \{-1, +1\}\} n \\ = 1, 2, \dots, N$$

where y_n is either -1 or 1 corresponds to the class 1 or 2. Each x_n is a p dimensional vector. The maximum margin hyperplane which partitions class 1 from class 2 is the support vector machine we want. Considering that any hyperplane can be written in the form of,

$$w^T \cdot x + b = 0$$

We want to choose the w and b to maximize the margin between the two parallel hyperplanes as large as possible while still separating the data. Optimum hyper plane for linearly separable patterns is given as,

$$w_0^T \cdot x + b_0 = 0$$

where x is the input vector, w_0 is the optimum weight vector and b_0 is the optimum bias. Defining the two parallel hyperplanes by the equations as,

 $w^{T} \cdot x + b = \pm 1$ Maximum
margin
Maximum

Figure 8: Classification of linearly separable data points using SVM

3.1.2 Non Linear SVM

Linear SVM cannot deal with the classification problem of which the different types of data are located at different sides of a hypersurface, that is a nonlinear dataset. It can be utilized for



nonlinear datasets by indirectly mapping the nonlinear inputs into to linear feature space where the maximum margin decision function is approximated. The mapping is done by using a kernel function because for the nonlinear data classification it is difficult to find the separating hyperplane.

In a nonlinear SVM classifier, a nonlinear operator is used to map the input pattern x into a higher dimensional space H. The nonlinear SVM classifier is defined as,

$f(x) = W^T \phi(x) + b$

The output of an SVM classifier in which examples are mapped onto a highdimensional feature space through the use of kernel functions is shown in fig.9. Some commonly using Kernel functions are listed in the Table 1.



Figure 9: Non-Linear SVM Classification

Name	Formula	Parameter	
Polynomial Kernel	$k(x, x_j) = (x^T x_i + c)^d$	d	
Sigmoidal Kernel	$k(x, x_j) = \tanh(ax^T x_i + c)$	C	
Gaussian Radial Basis Function	$k(x, x_j) = \exp(\frac{- x_i - x_j ^2}{\sigma^2})$	σ	

Table 1. Three common Kernels with their formula and naram	otorc

4. TUMOUR IDENTIFICATION

There are several steps of image processing to be considered to detect the tumour in 2D MRI images whereas image de-noising and image segmentation are the two processes employed mainly. Tumour segmentation is done on the abnormal MR images after classification. Thresholding is applied to extract the target from its background by assigning an intensity value T(threshold) for every pixel such each pixel is either classified as an object point or a background point. Thresholding creates a binary image that is binarized. In this project, a Global image threshold using Otsu's method is implemented. Which accept the edge to attenuate the intra-class variance of the black and white pixels.

5. IMPLEMENTATION

The experiments were carried out on a platform of Intel i3 processor and 4GB RAM, running under Windows 10, 64bit operating system. The biostatistical toolbox of Mat lab 7.10.0(R2010a) (The Math works @c) is used for implementing the project. The programs can be run or tested on any computer platform where Mat lab is available. The datasets consist of T1-weighted MR brain images in the axial plane and 256 256 in-plane resolution. The data is collected from, The whole brain atlas provided by Keith A. Johnson, M.D. and J. Alex Becker, Ph.D. (www.med. Harvard.edu/aanl ib/home.html).

We have collected a set of abnormal and normal brain MR images. The samples of each are shown in fig(2) and fig(3). 71 images are selected consisting of 30 normal and 41 Abnormal brain images. The setting of the training images and Testing images is shown in Table II.

Table 2: Dataset

Total no. of Images=71	Normal	Abnormal
Training=49	21	28
Testing=22	9	13

Median filter is used to remove noise from MR images and the filtered image is shown in figure (10).





Figure 10: Filtered Image

The MR image after skull masking is shown in figure (11)and (12). Power law Transformation is used for image enhancement which is fig (13). A 2D DWT can be seen as a 1D wavelet scheme which transform along the rows and then a 1D wavelet transform along the columns. The 2-D DWT operates in a straight forward manner by inserting array transposition between the two 1D DWT. The rows of the array are processed first with only one level of decomposition. This essentially divides the array into two vertical halves, with the first half storing the average coefficients, while the second vertical half stores the detail coefficients. This process is repeated again with the columns, resulting in four sub-bands within the array defined by filter output. The three levels of wavelet decomposition greatly reduce the input image size as shown in Fig. 15. The top left corner of the wavelet coefficients image denotes the approximation coefficients of level-3, whose size is only $38 \times 152 = 5776$.



Figure 11: Eroded Image



Figure 12: Dilated Image





111



The number of extracted features was reduced from 65536 to 5776. It is still too large for calculation. Thus, PCA is used to further reduce the dimensions of features to a higher degree. Here 13 different features are extracted for each image. They are Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, Variance, Smoothness, Kurtosis, Skewness, IDM. Then the extracted feature vectors fed into the classifier and classify the data based on the features. we have used SVM classifier as base classifier and compared its result with results obtained in kNN classifier. The classification process is divided into two parts that is the training and the testing part. Firstly, in the training part known data are given to the classifier for training. Secondly, in the testing part, images are the classifier and given to the classification is performed by using SVM. Here we have used SVM Torch for the implementation of SVM classifier. Three Kernel functions, Polynomial Kernel,



Sigmoidal Kernel and Gaussian Radial Basis Kernel Function are chosen to train SVM. Here the results are expressed in the form of a confusion matrix between the desired classes and target classes.



Figure 14: Input to wavelet decomposition



Figure 15: 3 level Decomposition 6. Results AND DISCUSSION

The author has used around 70 MR images for this project work consisting of 30 normal and 41 abnormal brain images. Firstly, the classification is done with kNN classifier. The confusion matrix obtained after kNN classification at K=6 is given in Table 3,

Table 3: Confusion Matrix for kNN Classifier

	Abnormal	Normal
Abnormal	12	2
Normal	2	7

Aslo, author tested SVM with different kernels (Polynomial kernel and Gaussian Radial Basis Kernel Function). The confusion matrices obtained are shown in Table 4.

Table 4: Confusion matrix of our method	d with different kernels
---	--------------------------

		Normal	Abnormal
Linear	Normal	12	2
	Abnormal	2	7
Polynomial	Normal	12	2
	Abnormal	1	8
Gaussian	Normal	13	1
	Abnormal	0	9

112

6.1 Performance Measures

6.1.1 Error Rate (ERR)

Error Rate is calculated as the number of all incorrect predictions divided by the total number of the dataset. The best error rate is 0.0, whereas the worst is 1.0.

ED	P_ F	P + P	FN	
LKI	$n = \frac{1}{TP + T}$	'N + .	FN + F	⁷ P
FP= False	Positive; FN	= Fals	e Nega	tive
TP= True F	ositive; TN	=True	Negati	ve
FN=False	Negative;	FP=	False	Positive
Results are shown in Table 5				
	Table 5: Fr	rror Ra	tos	

Tuble 51 El	TOT Hates
Methods	Error Rate
kNN	0.17
SVM (Linear)	0.17
SVM (Polinomial)	0.13
SVM (Gaussian)	0.04

6.1.2 Accuracy (ACC)

Accuracy in classification is the rate of correct classification.

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

6.1.3 Sensitivity (SN)

Sensitivity is the ability of a test to correctly classify an individual as diseased. It is also called true positive rate (TPR)

$$SN = \frac{TP}{TP + FN}$$
Specificity (SP)

It is the ability of a test to correctly classify an individual as disease-free. It is also called True Negative Rate(TNR)



6.1.4

$$SP = \frac{TN}{TN + FP}$$

The ACC, SN and SP results in percentage are shown in Table 6

Table 6: Performance Measures				
	Accuracy	Sensitivity	Specificity	
kNN	82.6	85.71	77.77	
SVM (Linear)	82.6	85.71	77.77	
SVM (Polynomial Kernel)	86.95	85.71	88.88	
SVM (Gaussian Kernel)	95.65	92.85	100	

The results show that the SVM Classifier with Gaussian Radial Basis function gives excellent classification. Fig. 17 below shows the affected area of tumour in brain MRI using Otzu's Segmentation method.





Figure 16: Original Image

7. CONCLUSION

The work in this research involves using SVM with different kernel functions to classify the input which is MRI Brain image into normal and abnormal classification. We implemented kNN classifier and SVM classifier with different kernel functions. SVM with Radial Basis Function kernel achieves maximum classification accuracy higher that SVM with polynomial kernel and Sigmoidal kernel. Also, the tumour affected region is extracted by Otsu's binarization method.

In this work, DWT is used for feature extraction. DWT can effectively extract the information from original MR images with slight loss. DWT captures both frequency and location information. Haar wavelet is used in this work. The most important contribution of this paper is to propose a method that combines them as a powerful tool for identifying normal MR brain images from abnormal MR brain images. We tested three kernels and found that the GRB kernel is the most successful one.

REFERENCES

- [1]. Dr. R J. Ramteke and Khachane Monali Y, "Automatic Medical Image Classification and Abnormality Detection Using K Nearest Neighbor", International Journal of Advanced Computer Research, Volume-2 Number-4 Issue-6 December-2012.
- [2].S.L Jany Shabu, Dr.C. Jayakumar "Brain Tumour Classification with MRI brain Images using 2-level GLCM features and Sparse representation based segmentation", 3rd International Conference on Intelligent Sustainable Systems (ICISS),IEEE Explore,2020
- [3]. Ketan Machhale, Hari Babu Nandpuru, Vivek Kapur, Laxmi Kosta, "MRI Brain Cancer Classification Using Hybrid Classifier (SVM-KNN)". 2015 International Conference on Industrial Instrumentation and (ICIC) Control College of Engineering Pune, India. May 28-30,2015
- [4].Y. Zhang , L. Wu, "An MR Brain Images Classifier via Principal Component Analysis and Kernel

Support Vector Machine "Progress In Electromagnetics Research, Vol. 130, 369-388, 2012

- [5]. Lee Wee Chun, Noramalina Abdullah and Umi Kalthum Ngah."Improvement of MRI Brain Classification Using Principal Component Analysis". International conference on Control system and computing engineering 2011 IEEE, 2011.
- [6]. Dipali M. Joshi, Dr. N.K Rana, and V.M Mishra. "Classification of Brain Cancer Using Artificial Neural Network". International conference on electronic computer technology (ICECT), 2010 international Conference IEEE, 2010.
- [7]. Daizy Deb ,Bahnishikha Dutta, Sudipta Roy "A Noble Approach for Noise Removal from Brain Image using Region Filling"IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)2014
- [8]. Dipalee Gupta, Siddhartha Choubey "Discrete Wavelet Transform for Image Processing" International Journal of Emerging Technology and Advanced Engineering, Volume 4, Issue 3, March 2015
- [9]. Hari Babu Nandpuru, Dr. S. S. Salankar and Prof. V. R Bora, "MRI brain Classification using Support Vector Machine", Students Conference on Electrical, Electronics and Computer Science 2014. IEEE, 2014.
- [10]. SK.Hasane Ahammad, V. Rajesh, MD.ZIA UR Rahman, "Fast and Accurate Feature Extraction based Segmenttaion framework for Spinal Cord Injury Severity

Classification", IEEE Transactions on Medical Imaging, 2020

- [11]. Nilesh Bhaskarrao Bahadure, Arun Kumar, Har Pal Thethi, "Image Analysis for MRI based Brain Tumor Detection and Feature Extraction Using Biologically Inspired BWT and SVM", International Journal of Biomedical Imaging, Hindawi,2017
- [12]. Meiyan Huang, Wei Yang, Yao Wu, Jun Jiang, Wufan Chen, "Brain Tumor Sementation Based on Local Independent Projection based Classification", IEEE Transactions on Biomedical Engineering, Vol.61, No. 10,2014
- Parveen, Amritpal Singh,
 "Detection of Brain Tumor in MRI Images using Combination of Fuzzy C Means & SVM, International Conference on Signal Processing and Integrated Networks, 2015
- Bashayer Fouad
 Marghalani, Muhammad Arif,
 "Automatic Classification of Brain
 Tumor and Alzheimer's Disease in
 MRI, International Learning &
 Technology Conference-2019
- [15]. Zar Nawab Khan Swati, Qinghua Zhao, Muhammad Kabir, Farhan Ali, Zakir Ali, Saeed Ahmed, "Brain Tumor Classification for MR images using transfer learning and finetuning, Computerized Medical Imaging and Graphics, Vol.75, page 34-46-2019

Conflict of Interest Statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper