Effect Of K-Fold Cross Validation on Mri Brain Images Using Support Vector Machine Algorithm

M.Jaya Lakshmi, S.Nagaraja Rao

Abstract—Recently, exact detection of the cancerous tumor in brain images is a critical task, especially at the early stage of the diseases. Various investigators have used machine-learning methods for the computer-aided diagnosis (CAD) to detect the tumor. In this paper an accurate and an automatic CAD system frame work has been done for verifying, the effect of K-fold cross validation for different values of k. K-means the segmentation technique is in the initial phase of the framework and the image is pre-processed for feature extraction and feature reduction using 2D-DWT and PCA respectively. The reduced features are given to the machine learning algorithm called the kernel support vector machine to classify magnetic resonance images. The K-fold stratified cross validation scheme is utilized to simplify the ability of the suggested strategy. The proposed method uses the different fold cross validation schemes, it is found that the RBF type kernel achieves the maximum classification with k=5 for the given data set. This method of classification of MR brain images, can help radiologists to analyze whether the patient's stage is normal or abnormal.

Keywords: Brain tumor, principal component analysis, feature extraction, classification, segmentation, image de-noising, principal component analysis (PCA), two-dimensional discrete wavelet transform (2D- DWT), kernel support vector machine (KSVM).

I. INTRODUCTION

The growth of unnatural cells which cannot be controlled, within the human body leads to Cancer. A collection or a mass of tissue, with those unnatural cells is known as a brain tumour. While brain tumours are not extremely normal, they are dangerous amongst the most deadly cancers [1], which leads to death.

Depending upon their originating point, brain tumours are classified as "fundamental brain tumours or metastatic brain tumours". In fundamental type of brain tumours, the starting point of the cells is within the brain tissue itself, whereas in the other type of tumours, cells start destructing various body parts and expand into the "brain". "Gliomas" are a kind of tumour cells that start growing from "glial cells", found in the supportive tissue of the brain. The term "glioma" interprets various kinds of "gliomas", extending to next range like "astrocytoma" .These growth of tumour arises from "star-shaped glial cells called astrocytes" and "Oligodendrocyte-gliomas", which grow from cells that make the fatty substance which cover and protect nerves. The foremost widely known essential harmful brain tumour is Glioblastoma multiform (GBM), [2]. Different Medical procedures, "Chemotherapy and Radiotherapy" are utilized for detecting and treating gliomas. [3].

Early stage analysis of gliomas is an important job in enhancing and giving treatment. Restorative Imaging methods, for example, "Positron Emission Tomography (PET), computerised axial Tomography (CT) and Magnetic Resonance Imaging (MRI)" are utilized to give important information like "shape, size, area and digestion of tumours aiding determination". The standard procedure for tumour detection is MRI, as it is more comfortable and does not affect the human body, since there is no emission of harmful radiations. Various "MRI image series" are created by changing excitation and redundancy time while securing the image. These series deliver many kinds of tissue details, like division of tumours aboard their sub-regions. Four standard MRI modalities are utilized to analyse glioma, are shown in Figure. 1.

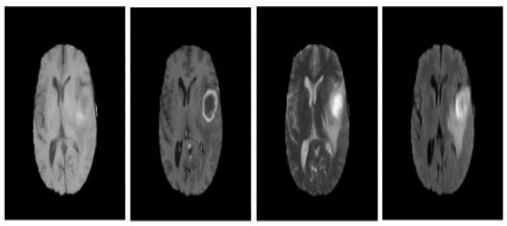

Figure. 1. Four different standard models of MRI used for glioma diagnosis include T1, T2, T1-Gd and FLAIR.

When "MRI images" are acquired from various sources, there is variation from one device to another. Around 150 frames of "2D images" are needed to make "3D brain images". In general, by considering T1, healthy tissues are recognized, even though T2 are used, because they coincide with edema, and which makes the tumour to look bright. In T1-Gd, the border of the tumour is not clearly visible and cannot be recognized by the bright signal, in the dynamic cell area of tumour effected tissue. As "necrotic cells" do not cooperate with the differentiation operator, they are seen by hypo exceptional piece of the tumour core creating it conceivable to effortlessly portion from the dynamic cell space on a similar grouping. In FLAIR, an indication of water atoms are blocked, that helps in recognizing an edema distinct from Cerebrospinal Fluid (CSF) [4].

"Image segmentation" for instance, a pixel based appropriation, texture Distribution, and so on., are the

Revised Manuscript Received on April 05, 2019.

M.Jaya lakshmi, Research scholar, JNTUA., Department of ECE, G.Pullareddy Engineering College, Kurnool, India (jayaraju15kumar@gmail.com)

Dr.S.Nagaraja Rao, Professor, Department of ECE, G.Pullareddy Engineering College, Kurnool, India (deanece@gprec.ac.in)

methods which converts an "image" into a limited number of segmented areas, considering a few characteristics which are non-overlapped. Segmentation of restorative images is necessary for some medicinal conclusions like radiation effect treatment, regions of interest (ROI) characterizing the "limit of tumour and intra cerebral brain haemorrhage", etc. [5, 6]. Many methods depend on "fuzzy logic and Neural Networks (NN)" conveyance and so on.... The division of therapeutic images is analysed using restorative analyses methods. Essentially, "image segmentation methods" are been characterized into three major classes: "based on edge, region, texture and pixel "[7].

"K-means clustering" which is a crucial system of "pixel-based techniques". In which pixels are situated in cluster and they are generally low contrasted and unsupervised. The application of this segmentation method when compared to other alternative techniques, is more practicable [8, 9]. This paper suggests automatic strategy to detect, characteristics of Tumour cells utilizing "Morphological procedure" for smoothening the tumour [10-12].

Wavelet Transform is a prominent tool for extracting the features from Brain Images. This is one of the transformation method which results as a multi-resolution analytic property, which allows to study the images at various "levels of resolution" requiring large data storage and it is expensive form of data processing [13]. To overcome this and to increase the efficiency which is due to its dimensionality, in terms of data storage and cost effectiveness, PCA is used [14] which efficiently reduces the spatiality property of the "data" and therefore the "computational cost" for analysing the new data is reduced[17]. Then, the problem is the classification of the input data. At present, to solve this problem researchers have proposed many classification approaches of two different categories where one is supervised, which has "support vector machine (SVM) [16] and k-nearest neighbours (k-NN)" [18]. Self-organization feature map (SOFM) [15] and fuzzy c-means are classified as unsupervised form [19].

Good results are obtained from all these methods, and the supervised performance is better than unsupervised classifier considering the accuracy parameter. However, the good classifier methods have the classification accuracies less than 95%. In this paper among supervised classification methods, more accurate methods are found like "SVMs" classification method based on machine learning theory [20– 22].

Original "SVMs" are linear classifiers and have advantages like it avoids over fitting by considering less number of training samples , have good accuracy, and direct geometric interpretation, when compared to other like "decision tree, ANN, and Bayesian network"[23]. In this paper, we considered a non-linear SVM classifier, kernel SVMs (KSVMs), which is the extension of linear "SVMs" by applying the "kernel function" to replace the "dot product form in the original SVMs" [24]. The "KSVMs" is used to "fit the maximum-margin hyper-plane" in a transformed feature space. Since, the transformation is "nonlinear and the transformed space is of high dimension", the classifier a "hyper-plane" in the high dimensional feature space is created which may be nonlinear within "the original input space".

The remaining part of the paper is arranged in such a way that, Section 2 gives the methodology for this paper. Section 3 introduces the procedure of "K- fold cross validation". Experiments in Section 4 includes a data set consisting of 160 images. Finally, in Section 5 the paper comes to a conclusion.

II. METHODOLOGY

There are several pathological brain tumour detection systems (PBDS) which are utilized by many researchers for the classifying MRI brain image to be normal or abnormal. The PBDS consists of segmentation, pre-processing and classification. Among all the various segmentation algorithms, k-means segmentation algorithm which is generally used on "MRI" by several researchers. The segmented output, from which features are extracted by using "feature extraction method", is done by wavelet transform in this paper. The features which are extracted are then dimensionally reduced using PCA . Then the reduced features are classified using SVM.

Recently, the "brain tumour" is the major cause of death, for different age groups. With the development of PBDS system, it is expected to give more detailed information for the radiologist about the tumour so that accurate decision can be taken for "better healthcare". Using the imaging modality, the tumours can be identified and further processed for accurate classification of tumours for which we operate the image using different image processing tools. Many researchers have suggested various "pre-processing, segmentation and classification algorithms". The "pre-processing" is done with a suitable filter, so that edges of the image are not affected and the noise is removed . The generalized block diagram for the detection of tumour, consisting of pre-processing, segmentation and classification is shown in Figure 2. The images which are to be tested undergo processes like "pre-processing, feature extraction, feature reduction, segmentation and classification of images".

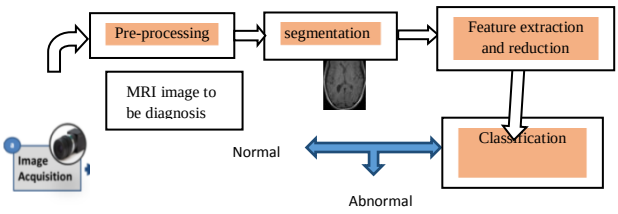

Figure 2. Process for detecting the brain tumour

2.1. preprocessing

This pre-processing method includes feature extraction and feature reduction.

2.1.1. Feature Extraction

The Fourier transform (FT) is the most classical procedure for analysing the signal, and converts a time-domain into sinusoidal components of various

"frequencies", thus, it converts the signal from "time domain to frequency domain". However, it has a disadvantage, it discards some of the time details of the signal. As an example, "Fourier spectrum" does not give the information about the tumour existence for the physician.

To analyse a small portion of the signal at a given selected "time", a method known as Gabor modified FT, called as "short-time Fourier transform" (STFT) or windowing is used. This method applies a mask of particular shape to the signal. This compromises the time and frequency information by providing little information about both. However, the window size restricts the data accuracy. "Wavelet transform" (WT) is also known as a "windowing technique" having "variable size". Thus, it conserves both "frequency and time information of the signal" which is transformed. The signal analysis development of WT is in the Fig. 3.

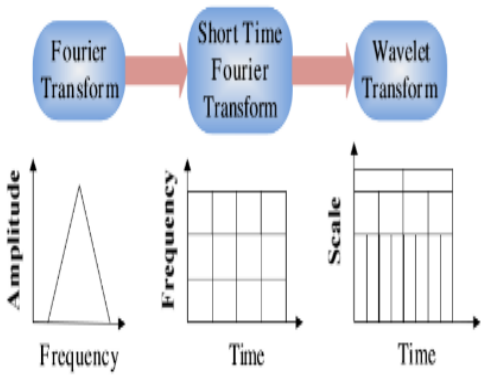

Figure. 3. The development of signal analysis

Additional benefit of "WT" is, rather than the traditional "frequency" parameter. It utilizes "scale", it does not yield "time-frequency" however, it produces a signal of time-scale view, wherever it gives a different view for a data, which is a powerful and more natural way, when we compare to "frequency" & "scale" parameters, it is found that scale parameter is frequently used in the standard of living. Because large/small scale can be conveniently well understood than "in high/low frequency".

2.1.1.1. Discrete Wavelet Transform

The "DWT" is good implementation of "Wavelet Transform" using "the scales and positions". [24]. The fundamentals of "DWT" are discussed. Let square-integrable function *be x* (*t*), then the "continuous WT" of *x* (*t*) is given, "Wψ (*t*)" is mentioned as

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

Where, $\psi_{a,b}(t) = (1/\sqrt{a})\psi\frac{(t-a)}{b}$ ---- (2)

Here, the "wavelet $\psi_{a,b}(t)$" is considered by "translation and dilation of the mother wavelet ψ(t)": "*a"* is the dilation factor and "*b"* the translation parameter, with real positive numbers. There are a number of wavelets which are popular during the wavelet analysis development, out of which the important one is "Harr wavelet", with lot of applications as it is the simplest one [25-27]. The Equation (1) is made discrete by limiting "*a"* and "*b"* to a discrete lattice for "$a = 2^b$ and $a > 0$ to give the DWT", which can be expressed as.

$ca_{j,k}(n) = DS[\sum_n x(n)g * j(n-2jk)i]$

$cd_{j,k}(n) = DS[\sum_n x(n)h * j(n-2jk)i]$----- (3)

With respect to this, approximation-coefficients and detail components are $ca_{j;k}$ and $cd_{j;k}$, respectively. *h* and *g* denote the transfer function of high-pass and low-pass filter, respectively. *k* and *j* determine the wavelet translation and scale factor, *"DS"* is the mathematical operator which implies the down sampling. The Equation (3) is of wavelet decomposition. In one level decomposition process, the signal *x*(*n*) is decomposed into 2 signals, which are the approximation and detail components. The above decomposition method is typically iterative with successive approximation, so that each signal is made into two components with numerous levels of resolution. The complete method is called "wavelet decomposition tree", shown in Fig. 4.

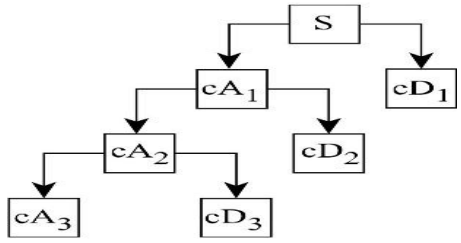

Figure 4. A 3-level decomposition tree of wavelet transforms.

2.1.1.2. 2D DWT

The DWT is applied individually to each dimension. Fig. 5 gives the "schematic diagram of DWT". It consists of 4 sub-bands which are "LL, LH, HH, and HL images" at every level of decomposition. For further processing the LL is used.

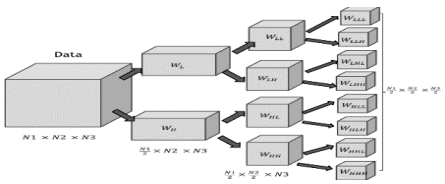

Figure5. The schematic diagram of 2D DWT.

Formation of more sub-signals and detailed information at larger temporal scales are seen sa the level of decomposition is increased. More information would give a better performance of the model, however might reduce the computing efficiency if more input neutrals and the model stability is decreased. However, it is necessary to select a suitable decomposition level for wavelet-neutral modelling. This is the way through which wavelet gives an easy "hierarchical framework" for getting more information. In mentioned algorithm, decomposition at level-3 using "Harr wavelet" was used to get features of image. The step by step process for feature extraction is given in algorithm 1[43].

Algorithm 1. Feature extraction.

Require: N: Total number of brain MR images having size $K \times K$ taken for experiment
Ensure: $FM[1:N, 1:M]$: Feature matrix, M: number of features
‣ Function *waveAppCoeff*() computes the approximation coefficients of level-3 Haar wavelet and MRI_n is the n^{th} brain MR image
1: Initialize $i \leftarrow 1$, $M \leftarrow K/8 \times K/8$ ‣ Total number of features to be extracted from the brain MR image
2: Create an empty matrix $EM[1:K/8, 1:K/8]$ and empty vector $FV[1, 1:M]$
3: **for** $n \leftarrow 1$ to N **do**
4: Get MRI_n
5: $EM_n[1:K/8, 1:K/8] \leftarrow waveAppCoeff(MRI_n)$
6: **while** $i \leq M$ **do**
7: **for** $s \leftarrow 1$ to $(K/8)$ **do**
8: **for** $t \leftarrow 1$ to $(K/8)$ **do**
9: $FV_n[1, i] \leftarrow EM_n[s, t]$
10: $i \leftarrow i+1$
11: **end for**
12: **end for**
13: **end while**
14: $FM[n, 1:M] \leftarrow FV_n[1, 1:M]$
15: **end for**

In the DWT the border distortion technique, digital filter is usually used. Once the digital filter is applied on to the image, at the edges the filter mask could extend beyond the image, the solution for that is to do the padding process to the pixels outside the images, which is done using a symmetric padding method [28].

2.1.2. Feature Reduction

There will be an increase in the computation period and memory storage, if there are more number of features in the image. Further, the classification, complication increases, which are meant as a curse of dimensionality. So there is a need for the researcher to reduce the features.

The most effective tool: PCA is used to reduce the least square reconstruction error and redundancy in the data set which consists of a more number of dependent variables in turn to reduce the dimensionality, while preserving many of the variations . This process can be done by mapping method where the input data set is mapped to a new set of ordered parameters considering their variance. This type of mapping method has three steps: the components of input vectors are normalized so that they do not correlate with each other. The resulting orthogonal components are arranged in such a way, that the elements with large variation are considered first in the data set and others are eliminated". The input vectors are normalized to possess zero mean and unity variance. Steps involved for implementing PCA is shown in figure 6.

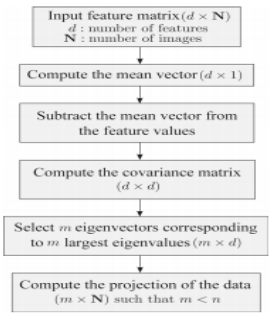

Figure.6. steps for PCA implementation.

2.2. K-means based segmentation

This method is broadly used for segmentation in biomedical applications among the different "clustering methods that are based on minimizing a formal objective function". Modifications are done to the "K-means clustering method" which is unsupervised, statistical, non-deterministic and iterative, which makes the PBDS system faster, and more efficient. A main technique in pixel-based methods is this clustering technique. Compared with all different region or edge, pixel-based methods with clustering methods are simple and with low complexity for computation, the application is more practicable. The flow chart of "K-means based segmentation method" is in figure 7.

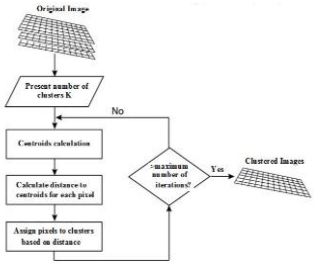

Figure.7-Working flowchart of K-Means clustering method

The main objective is to decrease the intensive distance computation, that takes place at iteration of this segmentation algorithm between all cluster centres and each data point. An easy mechanism is used to reduce the intensive distance between these two is, at each iteration, the space between them is computed and recorded in a data structure. Thus, in the successive iterations, "the distance between each data point and its previous nearest cluster is calculated again in an iterative manner".

2.3. KERNEL SVM

An image of brain has 4 regions, "white matter (WM), background, grey matter (GM), and cerebra spinal fluid (CSF)". To avoid the wrong classification risk, an input image is to be separated into four categories. In the field of a "machine learning", the SVM is a landmark. The advantages of SVM contain high accuracy to solve the problem of classification, control to reduce over-fitting, good performance and direct geometric interpretation to classify accurately [29]. There is a rapid growth in advanced SVMs in the area of machine learning, among them KSVMs are popular and effective.

Advantages of KSVMs [30]: they are symmetric, in practice, they work very well and so they have been a success in bioinformatics, language categorization and computer vision which is remarkable; it has very few tuneable parameters, and training which often include optimization which are quadratic [31-36]. Later, solutions are global and distinctive, thus local minima displayed using alternative statistical learning methods which avoid convergence, like NNs. To classify data, SVMs constructs a hyper-plane, so that it does not have any classification

problem to locate the data on different sides of the hyper-plane. The kernel strategy is applied to this method [37, 38], the algorithm is same, except that a nonlinear kernel function is placed at inner product. The kernel is related to the transform function $\psi(x_i)$ is given by $k(x_i; x_j) = \psi(x_i)\psi(x_j)$.

The value *"w"* is available in the "transformed space", with w value $\sum_i \alpha i \gamma i k(xi, x)$. Classification can be calculated with Dot products "w" by

$w.\psi(x) = \sum_i \alpha i \gamma i k(xi, x)$.

In another perspective, this method permits to" fit to the greatest-margin on the hyper-plane of a feature space which is transformed", which is of large dimension and the transformation may be nonlinear and however, the classifier consists of hyper plane within the large-dimensional feature space, in the first input, the space is nonlinear. For every kernel, there would be at a minimum of one change in the parameter to do the kernel more flexible and change the data which is practical.

III. K- FOLD STRATIFIED CROSS VALIDATION

The given dataset is trained to the classifier, so the classification accuracy is high for this trained dataset when compared to other datasets. To escape from overfitting, we utilize cross validation process in our model. Because of which there will not be any increase the final classification accuracy. However, classification reliability can be improved and can be added to other independent datasets.

There are three validation methods: "*K*-fold cross validation, Random subsampling validation, and leave-one-out validation". The properties of the first validation method are simple and easy to apply, and complete data is used for "training and validation" by researchers. To make a *K*-fold partition of the complete dataset, *K* times it is to be repeated to use "*K*-1 folds for training and a rest for validation, and finally average the error rates of *K* different experiments". In this method, as *K* folds are often randomly partitioned, but, some of the folds could have different distributions compared to other folds. Where each fold has almost the same sort of distribution [39].figure 8 give the basics structure of "k-fold cross validation method".

Figure 8. Basic k-folds for cross validation.

The present challenging situation is to take a decision about the number of folds required. If *K* value is taken too large, then bias will be small of the "true error rate estimator", but large will be the variance of the estimator and its time-consumption is more. Alternatively, if small the *K*, there will be a decrease in time computation, variance of the estimator is small, but more will be the estimator bias [40]. During this, we change the "*K* values from 3 to 10" by varying with a factor 1, and there after the SVM is trained for every value. As a result, we hereby we take the best value of *K*, where we are supposed to get the "largest classification accuracy".

IV. RESULTS AND DISCUSSIONS

The process is done with T2, as these images possess high contrast and vision is clear when compared to that of T1 and PET modals.20 images are randomly selected from each type of brain. Since, we have different kinds of MRI images within the dataset collected, one hundred sixty images are considered, out of which 20 are normal and rest of them are abnormal brain images.

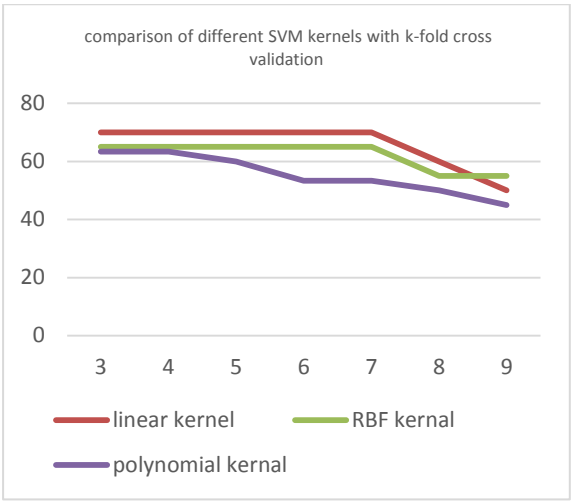

Figure 9. A Comparison of different SVM kernels accuracy on y-axis and different k-fold cross validation on x-axis.

The statistical measures for analyzing are:

Sensitivity (true positive fraction), measures the proportion of actual positives that are correctly identified, which is given by the formula,

$$\frac{TP}{TP + FN}$$

Specificity (true negative rate fraction), measures the proportion of actual negatives that are correctly identified, which is given by the formula,

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

Accuracy: The degree to which the result of a measurement, to the correct value or a standard.

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

By observing the plot in figure 9, which is the effect of the cross validation on different SVM kernels. We can say that the 5 cross validation of all the SVM kernel behaviour is horizontal and it is utilized as it gives high classification accuracy for the given data set as shown in the figure 10.

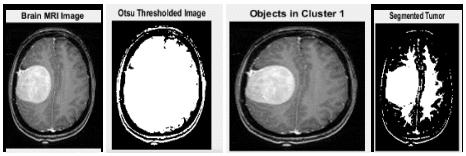

Figure 10.The MRI images with the DWT+PCA+KSVM with 5-cross validation.

V. CONCLUSIONS

In the article, "DWT with PCA and KSVM technique" for checking the effect of k-fold cross validation for various MRIs of the brain have been used. For which three completely different kernels have been chosen which are LIN, IPOL and GRB. Wherever GRB-KSVM is obtained, classification accuracy of 99.38% on 160 MR images, is more when compared to IPOL and GRB kernels.

The future of this work can be further extended considering these four analyzations: The SVM is the primary method which may be utilized for "MR images" of other kinds. The other one is by using different advanced transformation methods for the execution and consumption time to be reduced, Multi-classification that centers on specific diseases to avoid any mistreatment. Finally by the usage advanced kernels, the classification accuracy can be increased.

The information extracted from the images using DWT is more efficient with less loss. Spatial resolution is the advantage of DWT over DFT, after that, feature reduction is to be done if not, we would like a large search space which can cause significant computation risk that leads to decrease in accuracy of classification [41,42]. There are many good feature extracting methods such as Curvelet, Ridgelet, extreme learning etc. in Future one can specialise the work and the performance of those algorithms considering parameter metrics.

In this proposed DWT+PCA+KSVM Model used to extract brain tumour is with GRB kernel technique is superior to other kernels SVMs. As GRB kernel is an exponential function, because of which it can increase the distance between samples for which HPOL cannot. Therefore, GRB kernel is the better one to apply for other mechanical fields.

REFERENCES

1 De Angelis L M. Brain Tumors. N. Engl. J. Med. 2001; 344:114-23.

2 Deimling A. Gliomas. Recent Results in Cancer Research vol 171. Berlin: Springer; 2009.

3 Stupp R. Malignant glioma: ESMO clinical recommendations for diagnosis, treatment and follow-up. Ann Oncol 2007; 18(Suppl 2):69-70.

4 Drevelegas A and Papanikolou N. Imaging modalities in brain tumors Imaging of Brain Tumors with Histological Correlations. Berlin: Springer; 2011; chapter 2:13-34.

5 M. Masroor Ahmed and Dzulkifli Bin Mohamad "Segmentation of Brain MR Images for Tumor Extraction by Combining Kmeans Clustering and Perona Malik Anisotropic Diffusion Model" International Journal of Image Processing, Volume(2)Issue(l),2010.

6 Arash Azim Zadeh Irani and Bahari Belaton "A K-means Based Generic Segmentation System" 2009 Sixth International Conference on Computer Graphics, Imaging and Visualization.

7 Ming-Ni Wu, Chia-Chen Lin and Chin-Chen Chang "Brain Tumor Detection Using Color-Based K-Means Clustering Segmentation"InternationalConference onComputer Graphics,2009.

8 S.Sathish Kumar, M.Moorthi, M.Madhu3and Dr.R.Amutha "An Improved method of Segmentation using Fuzzy-Neuro Logic"Second International ConferenceonComputerResearchanddevelopment,2009.

9 S.Murugavalliand,V.Rajamani"An improved Implementation of Brain tumor Detection Using Segmentation Based on Neuro Fuzzy Technique", Journal of Computer Science,3,2007october.

10 Kharrat, A. Benamrane, N. Ben Messaoud, M. Abid, Computer. & Embedded System Lab.(CES),Nat.Eng.Sch.of Sfax,Tunisia,Nov.2009.

11 M. Schmidt, I. Levner, R. Greiner, A. Murtha, and A. Bistritz. "segmenting Brain Tumors using Alignment-Based Features". International Conference on Machine Learning and Applications, Los Angeles,December 2005.

12 S.R. Kannan ,Jan. 2005 "Segmentation of MRl using new Unsupervised Fuzzy Cmeansalgorithm,"ICGST-GVIPJournal,Volume5,Issue2.

13 Emin Tagluk., et al., "Classification of sleep apnea by using wavelet transform and artificial neural networks, " Expert Systems with Applications, Vol. 37, No. 2, 1600–1607, 2010.

14 Zhang, Y., et al., "A new classifier for polarimetric SAR images, " Progress in Electromagnetics Research, Vol. 94, 83– 104, 2009.

15 Camacho, J., et al., "Corrigendum to "The best approaches in the on-line monitoring of batch processes based on PCA: Does the modelling structure matter?" [Anal. Chim. Acta Volume 642 (2009) 59-68], " Analytica Chimica Acta, Vol. 658,No. 1, 106–106, 2010.

16 Chaplot,S., et al., "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network, " Biomedical Signal Processing and Control, Vol. 1, No. 1, 86–92, 2006.

17 Cocosco, C. A., et al., "A fully automatic and robust brain MRI tissue classification method, " Medical Image Analysis, Vol. 7, No. 4, 513–527, 2003.

18 Zhang,Y. ., et al., "Weights optimization of neural network via improved BCO approach, " Progress In Electromagnetics Research, Vol. 83, 185–198, 2008.

19 Yeh, ., et al., "A hierarchical genetic algorithm for segmentation of multi-spectral human-brain MRI, " Expert Systems with Applications, Vol. 34, No. 2, 1285–1295, 2008.

20 Patil, N. S., et al., "Regression models using pattern search assisted least square support vector machines, " Chemical Engineering Research and Design, Vol.83, No.8, 1030–1037, 2005.

21 Wang, F., et al., "The support vector machine for dielectric target detection through a wall, " Progress In Electromagnetics Research Letters, Vol. 23, 119–128, 2011.

22 Xu, Y., et al., "An support vector regression based nonlinear modeling method for Sic mesfet, " Progress In Electromagnetics Research Letters, Vol. 2, 103–114, 2008.

23 Li, D., et al., "Classification of foreign fibers in cotton lint using machine vision and multi-class support vector machine, " Computers and Electronics in Agriculture, Vol.4, No.2, 274–279, 2010.

24 Gomes, T. A. F., et al., "Combining meta-learning and search techniques to select parameters for support vector machines, " Neuro computing, Vol. 75, No. 1, 3–13, 2012.

25 Hable, R., "Asymptotic normality of support vector machine variants and other regularized kernel methods, " Journal of Multivariate Analysis, Vol. 106, 92–117, 2012.

26 Howard Lee ., et al., "Image based computer aided diagnosis system for cancer detection", Expert Systems with Applications 42 (2015) 5356–5365.

27 Zhang Y ., et al., "Feature extraction of brain MRI by stationary wavelet transform," in Proc. Int. Conf. Biomed. Eng. Comput. Sci., Wuhan, China, 2010, pp. 1–4.

28 Messina, A., "Refinements of damage detection methods based on wavelet analysis of dynamical shapes," International Journal of Solids and Structures, Vol. 45, Nos. 14–15, 4068–4097, 2008.

29 Martiskainen, P., et al., "Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines," Applied Animal Behaviour Science, Vol. 119, Nos. 1–2, 32–38, 2009.

30 Bermejo, S., B. Monegal, and J. Cabestany, "Fish age categorizationfromotolithimagesusingmulti-classsupportvector machines," Fisheries Research, Vol. 84, No. 2, 247–253, 2007.
31 Muniz, A. M. S., et al., "Comparison among probabilistic neural network, support vector machine and logistic regression for evaluating the eff ect of subthalamic stimulation in Parkinson disease on ground reaction force during gait," Journal of Biomechanics, Vol. 43, No. 4, 720–726, 2010.
32 Bishop, C. M.,"Pattern Recognition and Machine Learning" (Information Science and Statistics), Springer-Verlag New York, Inc., 2006.
33 Vapnik, V., "The Nature of Statistical Learning Theory," SpringerVerlag New York, Inc., 1995.
34 Jeyakumar, V., et al., "Lagrange multiplier characterizations of robust best approximations under constraint data uncertainty," Journal of Mathematical Analysis and Applications, Vol. 393, No. 1, 285–297, 2012.
35 Cucker, F. and S. Smale, "On the mathematical foundations of learning," Bulletin of the American Mathematical Society, Vol. 39, 1–49, 2002.
36 Poggio, T. and S. Smale, "The mathematics of learning: Dealing withdata," Notices of the American Mathematical Society (AMS), 388 Zhang and WuVol. 50, No. 5, 537–544, 2003.
37 Acevedo-Rodr´iguez, J., etal., "Computational load reduction in decision functions using support vector machines," Signal Processing, Vol. 89, No. 10, 2066–2071, 2009.
38 Deris, A. M., ., et al., "Overview of support vector machine in modeling machining performances," Procedia Engineering, Vol. 24, 308–312, 2011.
39 May, R. J., et al.,, "Data splitting for artificial neural networks using SOM-based stratified sampling," Neural Networks, Vol. 23, No. 2, 283–294, 2010.
40 Armand, S., et al., "Linking clinical measurements and kinematic gait patterns of toe-walking using fuzzy decision trees," Gait & Posture, Vol. 25, No. 3, 475–484, 2007.
41 El-Dahshan, ., et al., "Hybrid intelligent techniques for MRI brain images classification," Digital Signal Processing, Vol. 20, No. 2, 433–441, 2010.
42 Evans, A. C., et al., "Brain templates and atlases," NeuroImage, Vol. 62, No. 2, 911–922, 2012.
43 D. R. Nayak, R. Dash, B. Majhi," Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests", Neurocomputing,Vol. 177, 188–197, 2016.

BIBLIOGRAPHY

Dr.S.Nagaraja Rao,Professor & Dean of ECE dept., G.Pulla Reddy Engg College, kurnool. Received his Ph.D. degree from JNTUA, Anantapuramu, Andhra Pradesh, in 2011. His area of interest is communication systems, Image & Signal processing,. He is a professional member of ISTE and ISOI.He authorised for more than 50 papers in journals and conferences at international level.

M.Jayalakshmi, Research scholar received her M.Tech degree in the year 2010 from SK University, Anatapuram, India. She is currently pursuing Ph.D. degree from JNTUA, Anantapuramu, and Andhra Pradesh. Her current research includes biomedical image processing.