

# Detection and Classification of Tumor from Mri Brain Images Using Kernel Support Vector Machine

Dr. Shimal Das<sup>1</sup>, Dr. Jhunu DEbbarma<sup>2</sup>

<sup>1,2</sup>Department of Computer Science & Engineering

<sup>1,2</sup>Tripura Institute of Technology, Agartala-799009, India

## Article Info

Page Number: 763 - 775

Publication Issue:

Vol 71 No. 2 (2022)

## Article History

Article Received: 29 November 2021

Revised: 15 December 2021

Accepted: 08 January 2022

Publication: 05 February 2022

## Abstract

Automated and accurate classification of MR brain images is extremely important for medical analysis and interpretation. In this proposed work, we presented a novel method to classify a given MR brain image as normal or abnormal. The proposed method first employed wavelet transform to extract features from images, followed by applying principle component analysis (PCA) to reduce the dimensions of features. The reduced features were submitted to a kernel support vector machine (KSVM). The strategy of K-fold stratified cross validation was used to enhance generalization of KSVM.

**Keywords:** - MRI, Brain Tumor, PCA, Medical Imaging, KSVM, Wavelet Transform, Benign, Malignant..

---

## I. Introduction

In 2008, 12.7 million new cancer cases and 7.6 million cancer deaths occurred worldwide. This accounts for 13% of all deaths for that year, making it the leading cause of death worldwide according to the World Health Organization. It is curious to see a problem so old still an ever present threat even in our current day. In the future, the gold standard of medical care will be personalized medicine. Personalized medicine is the idea that medical decisions and practices will be custom tailored to each individual patient's exact medical situation, resulting in the most direct treatment with the least amount of collateral damage. The key to reaching this goal is, simply stated, the creation, management and analyses of vast amounts of previous information. This is exactly where algorithms, and in a small part this thesis, related into the future. Magnetic resonance imaging (MRI) is an imaging technique that produces high quality images of the anatomical structures of the human body, especially in the brain, and provides rich information for clinical diagnosis and biomedical research [15]. The diagnostic values of MRI are greatly magnified by the automated and accurate classification of the MRI images [68]. Wavelet transform is an effective tool for feature extraction from MRI brain images, because it allows analysis of images at various levels of resolution due to its multi resolution analytic property. However, this technique requires large storage and is computationally expensive [9]. In order to reduce the feature vector dimensions and increase the discriminative power, the principal component analysis (PCA) was used [10]. PCA is appealing since it effectively reduces the dimensionality of the data and therefore reduces the computational cost of analyzing new data [11]. Then, the problem of how to classify on the input data arises. Among supervised classification methods, the SVMs

are state-of-the-art classification methods based on machine learning theory [16-18]. Compared with other methods such as artificial neural network, decision tree, and Bayesian network, SVMs have significant advantages of high accuracy, elegant mathematical tractability, and direct geometric interpretation. Besides, it does not need a large number of training samples to avoid overfitting [19]. Original SVMs are linear classifiers. In this paper, we introduced the kernel SVMs (KSVMs), which extends original linear SVMs to nonlinear SVM classifiers by applying the kernel function to replace the dot product form in the original SVMs [20]. The KSVMs allow us to fit the maximum-margin hyperplane in a transformed feature space. The transformation may be nonlinear and the transformed space high dimensional; thus though the classifier is a hyperplane in the high-dimensional feature space, it may be nonlinear in the original input space [21].

## II. Literature Review

Literature review based on the following criteria:

### A. Relevant Work in Segmentation

Most previous computerized image analysis algorithms for TMAs have involved threshold based schemes [18], [19], [20]. These methods are known to be highly sensitive to even slight changes in color and illumination. Clustering based approaches, including k-means [18], have also been investigated for the analysis of TMAs. However, k-means is a non-deterministic algorithm and is highly sensitive to the initial choice of cluster centers [21]. Active contour schemes [22], while suitable for cell and nuclear segmentation in digital pathology, are not ideally suited to the problem of pixel level classification. Additionally they are typically infeasible for problems where hundreds of objects need to be concurrently segmented on very large images [23]. While supervised learning methods such as Probabilistic Boosting Trees (PBT) [24, 25] have become popular for image classification and segmentation, these methods are constrained by the difficulty [26] in obtaining ground truth segmentations from experts for classifier training of the object of interest. Manual annotation of the data, apart from being time-consuming and laborious, can also be expensive if only a medical practitioner is capable of providing accurate annotations. Additionally, if the target of interest changes, considerable effort might be required to generate new annotations and re-train the classifier.

Normalized Cuts (NCut) [27] is among the final mature descendants from a series of graph cutting techniques ranging from max cut to min cut [28, 29, 30, 31]. It is a popular scheme in spite of its main drawbacks: (1) the large number of calculations needed for determining the affinity matrix and (2) the time consuming Eigen value computation. For large images the computation and overhead of these border on the infeasible [27]. Consequently, a significant amount of research has focused on avoiding their direct calculations [32, 33]. The mean shift algorithm (MS) [34] has been employed and modified in [35] as an unsupervised technique for mode discovery instead of k-means. MS attempts to identify the cluster mean within a predefined bandwidth. By using a steepest gradient approach, a fast convergence to the set of true means of the statistical data can be found [36]. The improved fast Gauss transform (IFGT) implementation of the MS algorithm [37] allowed computation times for large images

to become reasonable. For the rest of this thesis, we will make no distinction between IFGT-MS and MS.

The attempt to merge NCuts and mean shift is not new [38]. To overcome the computational issues associated with NCut, a novel approach of combining both the MS and NCut algorithms was presented in [38]. Clustering the image by running the MS algorithm to convergence produced class assignments for the pixels. By taking the average intensity value of the regions obtained via the MS clustering step and using them as the vertices in the NCut algorithm, a significant speed improvement was obtained.

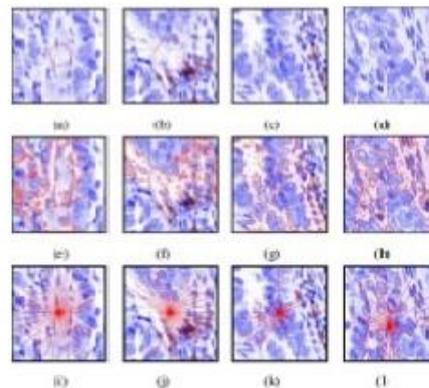
## B. Relevant Work in Localized Scale

The notion of scale in the context of image processing has been routinely employed over the last few decades to facilitate multi-resolution feature analysis; the assumption being that certain pertinent image features are only discernible at certain image scales and hence a spectrum of image resolutions needs to be considered for object recognition. Multi-scale approaches (scale-space [40] and hierarchical pyramids [41]) envisioned image processing operations being applied on a single image at varying levels of resolution; homogeneous regions being operated on at a lower resolution, with more heterogeneous regions being examined at higher resolutions. A limitation of these multi-scale techniques is that an optimal image resolution needs to be selected from within the image pyramid [41]. Additionally, some approaches [42] might require selection of multiple image scales for classification of a single image region.

To overcome these difficulties, the idea of locally adaptive scale emerged [43]. The concept of local scale was introduced to characterize varying levels of image detail so that localized image processing tasks could be performed, yielding an optimal result globally. Pizer et al. [44] suggested that having a locally adaptive definition of scale was necessary even for moderately complex detailed images. By quantifying these image details, an adaptive local scale image could encode implicit information present in the image intensity values. Locally adaptive scale has seen application in a variety of image processing tasks including MRI bias field correction [45], image segmentation [46], image registration [47], and image coding [48].

The common thread among these space scale concepts was that they were defined based on some homogeneity criterion linking the pixels neighboring the spatial location under consideration. Figure 1 reveals that both the b-scale ((a)-(d)) and g-scale ((e)-(h)) representations for a specific spatial location (located in the center of the image) attempts to identify the largest ball and set of pixels, respectively, that is homogeneous with respect to the pixel under consideration. For the stromal regions the associated b-scale ((a), (b)) and g-scale ((e), (f)) regions are large, reflecting the relative image homogeneity in that location of the image. Note however that the corresponding g-scale set is affected by the presence of local heterogeneity (g-scale has multiple cavities). The LMS ((i),(j)) for stromal regions has particles radiating far out in certain directions, but is also locally constrained (to the right) on account of neighboring nuclei. For the tumor regions ((c), (d)), the corresponding b-scale is small, with g-scale resulting in an amorphous shape with multiple cavities. Additionally the g-scale sets for the tumor regions in ((g), (h)) appears dramatically different. The corresponding LMS

((k), (l)) while not constrained by a prior shape model, yields a local structural signature that is consistent across both ((k), (l)) and distinctly different from the corresponding non-tumor LMS signatures ((i), (j)). Note that the initial motivation of both b-scale, and g-scale was from the perspective of noise filtering and bias field correction [45], image processing operations that warranted identification of locally connected homogeneous regions.



**Figure 1: The associated b-scale ((a)-(d)), g-scale ((e)-(h)), and LMS signatures((i)-(l)) shown for a candidate image location on an OCa biopsy image in red.**

### III. Relevant Work in Tumor Identification

#### A. Specialized Staining

One of the common approaches to tackling this challenge revolves around using specialized staining. In some cases, it is possible to stain directly for a specific tumor type of interest allowing for a clear separation of regions. Since this is often not the case, the authors in [51] present an approach which requires specially stained fluorescence images from which they extracted the DAPI (49,6-diamidino-2-phenylindole) channel. They formed cell graphs based on the topological distribution of the tissue cell nuclei and extracted the corresponding graph features. By using topological, morphological and intensity based features they built a supervised classifier using support vector machines which obtains an accuracy of 88% 6.68. We can contrast this with our approach, obtaining on par results with a much smaller variance 86% .000354 while operating on solely industry standard Hematoxylin(H) or Hematoxylin & Eosin(H&E) stained images, allowing broader usage in pre-existing tissue repositories.

#### B. Computationally Expensive

There are notable approaches to the domain which given our current technological knowledge and infrastructure are intractable for immediate application in a clinical setting. For example, an N-point correlation function [52] (N-pcfs) for constructing an appropriate feature space for achieving tissue segmentation in histology-stained microscopic images was presented. The Npcfs estimates microstructural constituent packing densities and their spatial distribution in a tissue sample. Afterwards, they represented the multi-phase properties estimated by the N-pcfs in a tensor structure. Using a variant of a higher-order singular value decomposition (HOSVD) algorithm, they realize a classifier that provides a multi-linear

description of the tensor feature space. While the approach was not used directly for tumor versus stroma identification, they showed 90% accuracy of their segmentations in a case-study that focuses on understanding the genetic phenotyping differences in mouse placentae. Unfortunately, the authors note in their discussion section the need to invest additional research in finding more optimal data structures and algorithms to reduce the overall time associated with computations. We show that our feature set can be generated in as little as 0.0058s per sample, motivating the immediate usage in a high-throughput system.

### C. Full-Featured

C-path, as described in [53], first performed an automated, hierarchical scene segmentation that generated thousands of measurements, including both standard morphological descriptors of image objects and higher-level contextual, relational, and global image features. Using the concept of superpixels, they measured the intensity, texture, size, and shape of the superpixel and its neighbors. Afterwards, to produce more biologically meaningful features, they classified superpixels as epithelium or stroma. Using these classified superpixels they created more than 6600 features. This approach found a set that were associated with samples from patients who had a shorter survival period. The key aspect of this analysis was that these features were not predefined by a pathologist as being relevant to cancer; instead, the software itself found the cancer-related features among the very large set of measurements of the image and obtained an accuracy of 89% in the superpixel classification. Their novelty was defined by their successful combination of existing features. Our work presents a significantly lower dimensional novel morphological feature set, which obtains 86% accuracy on the same task, clearly indicating a competitive approach. Additionally, since we have the ability to perform on par using only a single feature, we believe that our approach is far more scalable and thus applicable in a clinical setting.

## IV. BRAIN TUMOR IDENTIFICATION TECHNIQUES

---

Image segmentation is one of the most important and active research area in the medical imaging domain. It can be defined as the delineation of one or several structures of interest within the image. Automated methods are sought in order to avoid the time consuming burden of manually contouring the structures. The problem is particularly difficult in the context of brain tumors. Indeed, most tumors have heterogeneous appearances and their intensity range overlap with the healthy tissues. The presence of a necrotic core is frequent (especially for glioblastomas, but it also occurs for DLGGs) resulting on a strong contrast with the active tumor. Prior information regarding the shape of the tumor cannot be used as they have variable sizes and shapes. DLGGs in particular, have very fuzzy and irregular boundaries due to their infiltrative nature. Edema (swelling of brain tissue around the tumor) and mass effect (tissue displacement induced by the tumor) are quite uncommon due to the slow-growing nature of the DLGGs [Sanai 2011]. In this context, the simplest segmentation methods such as thresholding or region growing are insufficient [Gibbs 1996]. Despite extensive and promising work in the tumor segmentation field, obtaining accurate and reliable segmentations of brain tumors remains a difficult task.

Segmentation methods can be grouped in two categories: surface and region-based approaches. The objective of surface based methods is to find the organ or tumors boundary by propagation a curve/surface with a flow that is determined according to curvature and image constraints (generally the image gradient). Snakes and level sets are typically used in this context. The former defines the objects boundary explicitly as a parametric curve, while the latter defines the contour via an implicit function allowing for more complex geometries and topological changes.

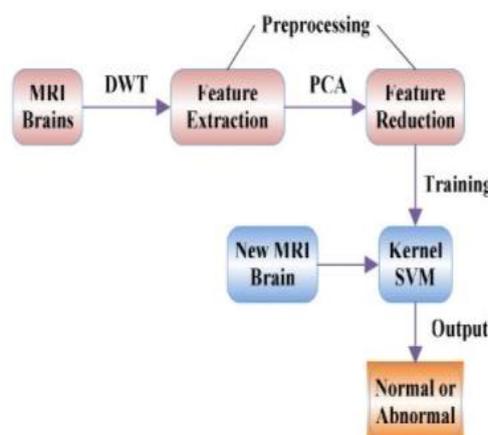
Region based methods consider the segmentation problem from a different angle. Here, the goal is to identify all voxels belonging to the object and separate them from the rest of the image. Early work relied on fuzzy clustering to regroup similar voxels but quickly showed limitations. Supervised statistical pattern classification techniques have been the basis of the majority of recent region based tumor segmentation methods. The voxels are separated by a classification score or probability, the ultimate goal being to label each pixel of the image as tumor or background. The tumor is frequently separated into active tumor.

### A. Magnetic Resonance Imaging

Magnetic Resonance Imaging is one of the most popular medical imaging modalities. There exists a variety of different imaging modalities that enable the study of the brain. This section presents a brief overview of the different imaging methods, then focus on Magnetic Resonance Imaging (MRI) which is the most common modality for brain tumor observation. It has the advantage of being a non invasive and non ionizing technique that produces images with excellent tissue contrast, making it the modality of choice for brain (and more specifically brain tumors) study. MRI is based on the principle of Nuclear Magnetic Resonance (NMR), that is commonly used in spectroscopy to study the physical and chemical properties of molecules.

### V. Methodology

The proposed method in figure 2 consists of three stages: Step 1. Preprocessing (including feature extraction and feature reduction); Step 2. Training the kernel SVM; Step 3. Submit new MRI brains to the trained kernel SVM, and output the prediction.



**Figure 2: Proposed Methodology**

#### A. Preprocessing

## 1. Feature Extraction

The most conventional tool of signal analysis is Fourier transform (FT), which breaks down a time domain signal into constituent sinusoids of different frequencies, thus, transforming the signal from time domain to frequency domain. However, FT has a serious drawback as discarding the time information of the signal. For example, analyst cannot tell when a particular event took place from a Fourier spectrum. Thus, the quality of the classification decreases as time information is lost.

Gabor adapted the FT to analyze only a small section of the signal at a time. The technique is called windowing or short time Fourier transform (STFT) [23]. It adds a window of particular shape to the signal. STFT can be regarded as a compromise between the time information and frequency information. It provides some information about both time and frequency domain. However, the precision of the information is limited by the size of the window.

Wavelet transform (WT) represents the next logical step: a windowing technique with variable size. Thus, it preserves both time and frequency information of the signal. The development of signal analysis is shown in Fig. .

Another advantage of WT is that it adopts scale instead of traditional frequency, namely, it does not produce a time-frequency view but a time-scale view of the signal. The time-scale view is a different way to view data, but it is a more natural and powerful way, because compared to frequency, scale is commonly used in daily life. Meanwhile, in large/small scale is easily understood than in high/low frequency.

## 2. Discrete Wavelet Transform

The discrete wavelet transform (DWT) is a powerful implementation of the WT using the dyadic scales and positions [24]. The fundamentals 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

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

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

Here, the wavelet  $\psi_{a,b}(t)$  is calculated from the other wavelet  $\psi(t)$  by translation and dilation:  $a$  is the dilation factor and  $b$  the translation parameter (both real positive numbers). There are several different kinds of wavelets which have gained popularity throughout the development of wavelet analysis. The most important wavelet is the Harr wavelet, which is the simplest one and often the preferred wavelet in a lot of applications [25,27].

Equation (1) can be discretized by restraining  $a$  and  $b$  to a discrete lattice ( $a = 2^j$  &  $a > 0$ ) to give the DWT, which can be expressed as follows.

$$ca_{j,k}(n) = DS[\sum_n x(n) g^*_{j(n-2^j k)}]. \quad \text{Eq.2}$$

$$cd_{j,k}(n) = DS[\sum_n x(n) h^*_{j(n-2^j k)}], \quad \text{Eq.3}$$

Here  $ca_{j,k}$  and  $cd_{j,k}$  refer to the coefficients of the approximation components and the detail components, respectively.  $g(n)$  and  $h(n)$  denote for the low-pass filter and high-pass filter,

respectively.  $j$  and  $k$  represent the wavelet scale and translation factors, respectively. DS operator means the down sampling. This procedure is called one-level decompose. The above decomposition process can be iterated with successive approximations being decomposed in turn, so that one signal is broken down into various levels of resolution.

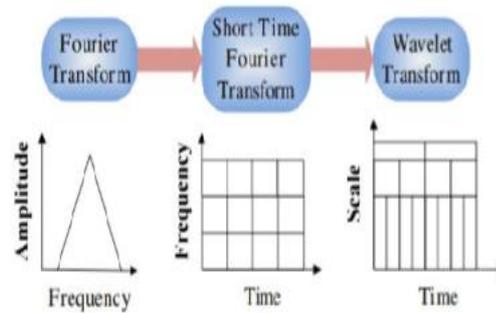


Figure 3: The development of signal analysis.

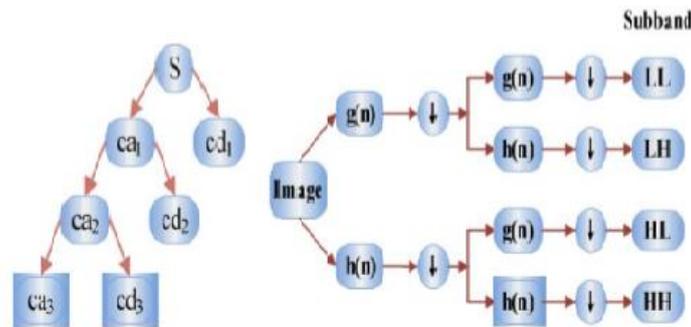


Figure 4: The development of signal analysis in 2DWT

### 3. Feature Reduction

Excessive features increase computation times and storage memory. Furthermore, they sometimes make classification more complicated, which is called the curse of dimensionality. It is required to reduce the number of features.

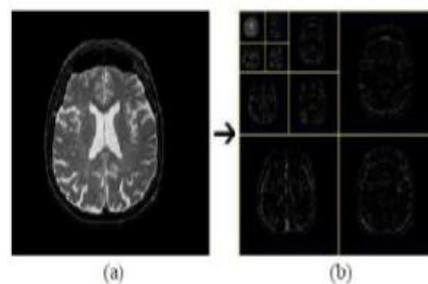
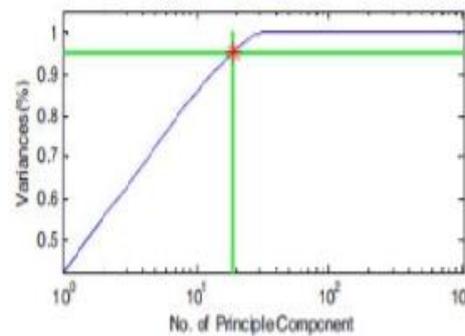


Figure 5 The procedures of 3-level 2D DWT: (a) normal brain MRI; (b) level-3 wavelet coefficients.



**Figure 6: Variances against No. of principle components (x axis is log scale).**

PCA is an efficient tool to reduce the dimension of a data set consisting of a large number of interrelated variables while retaining most of the variations. It is achieved by transforming the data set to a new set of ordered variables according to their variances or importance. This technique has three effects: it orthogonalizes the components of the input vectors so that uncorrelated with each other, it orders the resulting orthogonal components so that those with the largest variation come first, and eliminates those components contributing the least to the variation in the data set.

It should be noted that the input vectors be normalized to have zero mean and unity variance before performing PCA. The normalization is a standard procedure. Details about PCA could be seen in Ref. [10].

## B. KERNEL SVM

The introduction of support vector machine (SVM) is a landmark in the field of machine learning. The advantages of SVMs include high accuracy, elegant mathematical tractability, and direct geometric interpretation [29]. Recently, multiple improved SVMs have grown rapidly, among which the kernel SVMs are the most popular and effective. Kernel SVMs have the following advantages [30]: (1) work very well in practice and have been remarkably successful in such diverse fields as natural language categorization, bioinformatics and computer vision; (2) have few tunable parameters; and (3) training often involves convex quadratic optimization [31]. Hence, solutions are global and usually unique, thus avoiding the convergence to local minima exhibited by other statistical learning systems, such as neural networks.

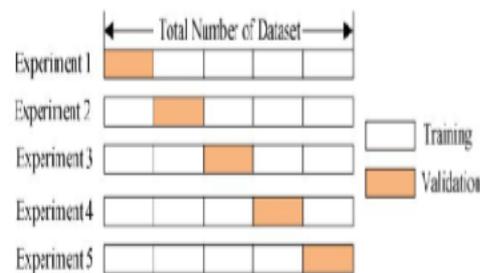
### 1. Motivation

Suppose some prescribed data points each belong to one of two classes, and the goal is to classify which class a new data point will be located in. Here a data point is viewed as a  $p$ -dimensional vector, and our task is to create a  $(p-1)$ -dimensional hyperplane. There are many possible hyperplanes that might classify the data successfully. One reasonable choice as the best hyperplane is the one that represents the largest separation, or margin, between the two classes, since we could expect better behavior in response to unseen data during training, i.e., better generalization performance. Therefore, we choose the hyperplane so that the distance from it to the nearest data point on each side is maximized [32]. Fig. 4 shows the geometric

interpolation of linear SVMs, here H1, H2, H3 are three hyperplanes which can classify the two classes successfully, however, H2 and H3 does not have the largest margin, so they will not perform well to new test data. The H1 has the maximum margin to the support vectors (S11, S12, S13, S21, S22, and S23), so it is chosen as the best classification hyperplane [33].

## 2. K-FOLD STRATIFIED CROSS VALIDATION

Since the classifier is trained by a given dataset, so it may achieve high classification accuracy only for this training dataset not yet other independent datasets. To avoid this overfitting, we need to integrate cross validation into our method. Cross validation will not increase the final classification accuracy, but it will make the classifier reliable and can be generalized to other independent datasets.



**Figure 7: A 5 fold cross validation.**

## VI. Experiment Results and Discussion

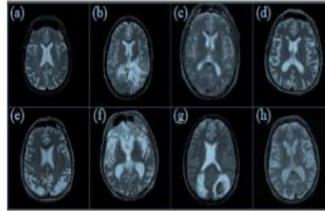
The experiments were carried out on the platform of i3 Intel with 2.4 GHz processor and 4GB RAM, running under Windows 8.1 operating system. The algorithm was in-house developed via the wavelet toolbox, the biostatistical toolbox of Matlab 2013a. The open SVM toolbox is downloaded and extended to Kernel SVM, and applied it to the MR brain images classification.

### A. Database

The datasets consists of T2-weighted MR brain images in axial plane and 256 X 256 in plane resolution, which were downloaded from the website of Harvard Medical School (URL: med.harvard.edu/AANLIB/), OASIS dataset (URL: www.oasis-brains.org/), and ADNI dataset (URL: http://adni.loni.ucla.edu/). We choose T2 model since T2 images are of higher contrast and clearer vision compared to T1 and PET modalities.

The abnormal brain MR images of the dataset consist of the following diseases: glioma, meningioma, Alzheimer's disease, Alzheimer's disease plus visual agnosia, Pick's disease, sarcoma, and Huntington's disease. The samples of each disease are illustrated in Fig. 6

We randomly selected 20 images for each type of brain. Since there are one type of normal brain and seven types of abnormal brain in the dataset, 160 images are selected consisting of 20 normal and 140 (= 7 types of diseases X 20 images/diseases) abnormal brain images. The setting of the training images and validation images is shown in Table 2 since 5-fold cross validation was used.



**Figure 8: Sample of brain MRIs: (a) normal brain; (b) glioma; (c) meningioma; (d) Alzheimer's disease; (e) Alzheimer's disease with visual agnosia; (f) Pick's disease; (g) sarcoma; (h) Huntington's disease.**

### B. Feature Extraction

The three levels of wavelet decomposition greatly reduce the input image size as shown in Fig. 5.2. The top left corner of the wavelet coefficients image denotes the approximation coefficients of level-3, whose size is only  $32 \times 32 = 1024$ .

### C. Feature Reduction

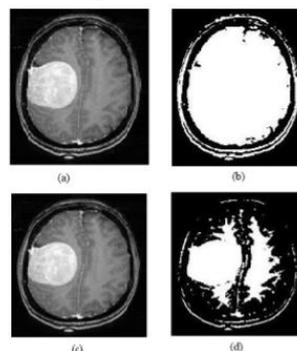
As stated above, the number of extracted features was reduced from 65536 to 1024. However, it is still too large for calculation. Thus, PCA is used to further reduce the dimensions of features to a higher degree. The curve of cumulative sum of variance versus the number of principle components is shown in Fig. 5.3.

### D. Classification Accuracy

We tested four SVMs with different kernels (LIN, HPOL, IPOL, and GRB). In the case of using linear kernel, the KSVM degrades to original linear SVM.

We computed hundreds of simulations in order to estimate the optimal parameters of the kernel functions, such as the order  $d$  in HPOL and IPOL kernel, and the scaling factor in GRB kernel.

The results showed that the proposed DWT+PCA+KSVM method obtains quite excellent results on both training and validation images.



**Figure 9: (a)Input Brain MRI Image; (b) Thesholding technique to extract foreground; (c) Result after Clustering; (d)Extracted Feature**

**Few Results of Classification Accuracy and Calculated Features**

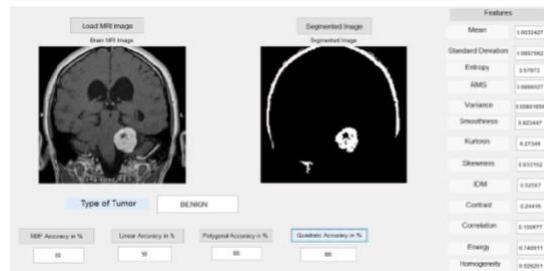


Figure 10

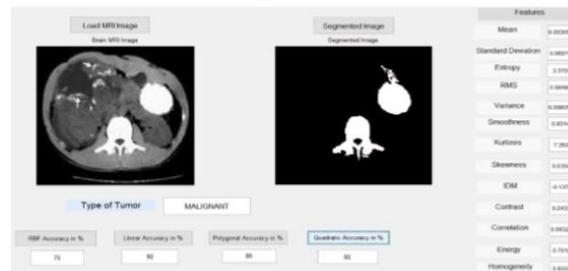


Figure 11

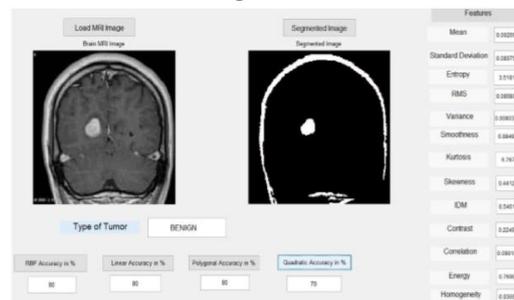


Figure 12

## Vii. Conclusion And Discussions

In this study, we have developed a novel DWT+PCA+KSVM method to distinguish between normal and abnormal MRIs of the brain. We picked up four different kernels as LIN, HPOL, IPOL and GRB. The experiments demonstrate that the GRB kernel SVM obtained 90.00% classification accuracy on the 160MR images, higher than HPOL, IPOL and GRB kernels, and other popular methods in recent literatures. The proposed DWT+PCA+KSVM with GRB kernel method shows superiority to the LIN, HPOL, and IPOL kernels SVMs. The reason is the GRB kernel takes the form of exponential function, which can enlarge the distance between samples to the extent that HPOL can't reach. Therefore, we will apply the GRB kernel to other industrial fields.

This technique of brain MRI classification based on PCA and KSVM is a potentially valuable tool to be used in computer assisted clinical diagnosis.

Future work should focus on the following four aspects: First, the proposed SVM based method could be employed for MR images with other contrast mechanisms such as T1-weighted, Proton Density weighted, and diffusion weighted images. Second, the computation time could be accelerated by using advanced wavelet transforms such as the lift-up wavelet. Third, Multi classification, which focuses on specific disorders studied using brain MRI, can also be explored. Forth, novel kernels will be tested to increase the classification accuracy.

There are some other excellent feature transformation methods such as ICA, manifold learning. In the future, focus will be on investigating the performance of these algorithms.

## References

- [1] Zhang, Y., L. Wu, and S. Wang, Magnetic resonance brain image classification by an improved artificial bee colony algorithm, *Progress In Electromagnetics Research*, Vol. 116, 6579, 2011.
- [2] Mohsin, S.A., N.M. Sheikh, and U. Saeed, MRI induced heating of deep brain stimulation leads: Effect of the air-tissue interface, *Progress In Electromagnetics Research*, Vol. 83, 8191, 2008.
- [3] Golestanirad, L., A. P. Izquierdo, S. J. Graham, J. R. Mosig, and C. Pollo, Effect of realistic modeling of deep brain stimulation on the prediction of volume of activated tissue, *Progress In Electromagnetics Research*, Vol. 126, 116, 2012.
- [4] Mohsin, S.A., Concentration of the specific absorption rate around deep brain stimulation electrodes during MRI, *Progress In Electromagnetics Research*, Vol. 121, 469484, 2011.
- [5] Oikonomou, A., I. S. Karanasiou, and N. K. Uzunoglu, Phased array near field radiometry for brain intracranial applications, *Progress In Electromagnetics Research*, Vol. 109, 345360, 2010.
- [6] Scapaticci, R., L. Di Donato, I. Catapano, and L. Crocco, A feasibility study on microwave imaging for brain stroke monitoring, *Progress In Electromagnetics Research B*, Vol. 40, 305324, 2012.
- [7] Asimakis, N. P., I. S. Karanasiou, P. K. Gkonis, and N. K. Uzunoglu, Theoretical analysis of a passive acoustic brain monitoring system, *Progress In Electromagnetics Research B*, Vol. 23, 165180, 2010.
- [8] Chaturvedi, C.M., V.P. Singh, P. Singh, P. Basu, M. Singaravel, R.K. Shukla, A. Dhawan, A. K. Pati, R. K. Gangwar, and S. P. Singh. 2.45 GHz (CW) microwave irradiation alters circadian organization, spatial memory, DNA structure in the brain cells and blood cell counts of male mice, *mus musculus*, *Progress In Electromagnetics Research B*, Vol. 29, 2342, 2011.
- [9] Emin Tagluk, M., M. Akin, and N. Sezgin, Classification of sleep apnea by using wavelet transform and artificial neural networks, *Expert Systems with Applications*, Vol. 37, No. 2, 16001607, 2010.
- [10] Zhang, Y., L. Wu, and G. Wei, A new classifier for polarimetric SAR images, *Progress in Electromagnetics Research*, Vol. 94, 83 104, 2009.
- [11] Camacho, J., J. Pico, and A. Ferrer, Corrigendum to The best approaches in the on-line monitoring of batch processes based on PCA: Doesthemodellingstructurematter?[*Anal. Chim. Acta* Volume 642 (2009) 59-68], *Analytica Chimica Acta*, Vol. 658, No. 1, 106106, 2010.
- [12] Chaplot, S., L. M. Patnaik, and N. R. Jagannathan, 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, 8692, 2006.
- [13] Cocosco, C.A., A.P. Zijdenbos, and A.C. Evans, A fully automatic and robust brain MRI tissue classification method, *Medical Image Analysis*, Vol. 7, No. 4, 513527, 2003.

- [14] Zhang, Y. and L. Wu, Weights optimization of neural network via improved BCO approach, *Progress In Electromagnetics Research*, Vol. 83, 185198, 2008.
- [15] Yeh, J.-Y. and J. C. Fu, A hierarchical genetic algorithm for segmentation of multi-spectral human-brain MRI, *Expert Systems with Applications*, Vol. 34, No. 2, 12851295, 2008.
- [16] 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, 10301037, 2005.
- [17] Wang, F.-F. and Y.-R. Zhang, The support vector machine for dielectric target detection through a wall, *Progress In Electromagnetics Research Letters*, Vol. 23, 119128, 2011.
- [18] Xu, Y., Y. Guo, L. Xia, and Y. Wu, An support vector regression based nonlinear modeling method for Sic mesfet, *Progress In Electromagnetics Research Letters*, Vol. 2, 103114, 2008.
- [19] Li, D., W. Yang, and S. Wang, 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, 274279, 2010.
- [20] Gomes, T. A. F., et al., Combining meta-learning and search techniques to select parameters for support vector machines, *Neurocomputing*, Vol. 75, No. 1, 313, 2012.
- [21] Hable, R., Asymptotic normality of support vector machine variants and other regularized kernel methods, *Journal of Multivariate Analysis*, Vol. 106, 92117, 2012.
- [22] Ghosh, A., B. Uma Shankar, and S. K. Meher, A novel approach to neuro-fuzzy classification, *Neural Networks*, Vol. 22, No. 1, 100109, 2009.
- [23] Gabor, D., Theory of communication. Part 1: The analysis of information, *Journal of the Institution of Electrical Engineers Part III: Radio and Communication Engineering*, Vol. 93, No. 26, 429441, 1946.
- [24] Zhang, Y. and L. Wu, Crop classification by forward neural network with adaptive chaotic particle swarm optimization, *Sensors*, Vol. 11, No. 5, 47214743, 2011.
- [25] Zhang, Y., S. Wang, and L. Wu, A novel method for magnetic resonance brain image classification based on adaptive chaotic PSO, *Progress In Electromagnetics Research*, Vol. 109, 325343, 2010.
- [26] Ala, G., E. Francomano, and F. Viola, A wavelet operator on the interval in solving Maxwells equations, *Progress In Electromagnetics Research Letters*, Vol. 27, 133140, 2011.
- [27] Iqbal, A. and V. Jeoti, A novel wavelet-Galerkin method for modeling radio wave propagation in tropospheric ducts, *Progress In Electromagnetics Research B*, Vol. 36, 3552, 2012.
- [28] Messina, A., Refinements of damage detection methods based on wavelet analysis of dynamical shapes, *International Journal of Solids and Structures*, Vol. 45, Nos. 1415, 40684097, 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. 12, 3238, 2009.

- [30] Bermejo, S., B. Monegal, and J. Cabestany, Fish age categorization from otolith images using multi-class support vector machines, *Fisheries Research*, Vol. 84, No. 2, 247253, 2007.
- [31] Muniz, A. M. S., et al., Comparison among probabilistic neural network, support vector machine and logistic regression for evaluating the effect of subthalamic stimulation in Parkinson disease on ground reaction force during gait, *Journal of Biomechanics*, Vol. 43, No. 4, 720726, 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*, Springer-Verlag New York, Inc., 1995.
- [34] Jeyakumar, V., J. H. Wang, and G. Li, Lagrange multiplier characterizations of robust best approximations under constraint data uncertainty, *Journal of Mathematical Analysis and Applications*, Vol. 393, No. 1, 285297, 2012.
- [35] Cucker, F. and S. Smale, On the mathematical foundations of learning, *Bulletin of the American Mathematical Society*, Vol. 39, 149, 2002.
- [36] Poggio, T. and S. Smale, The mathematics of learning: Dealing with data, *Notices of the American Mathematical Society (AMS)*, Vol. 50, No. 5, 537544, 2003.
- [37] Acevedo-Rodriguez, J., et al., Computational load reduction in decision functions using support vector machines, *Signal Processing*, Vol. 89, No. 10, 20662071, 2009.
- [38] Deris, A. M., A. M. Zain, and R. Sallehuddin, Overview of support vector machine in modeling machining performances, *Procedia Engineering*, Vol. 24, 308312, 2011.
- [39] May, R. J., H. R. Maier, and G. C. Dandy, Data splitting for artificial neural networks using SOM-based stratified sampling, *Neural Networks*, Vol. 23, No. 2, 283294, 2010.
- [40] Armand, S., et al., Linking clinical measurements and kinematic gait patterns of toewalking using fuzzy decision trees, *Gait & Posture*, Vol. 25, No. 3, 475484, 2007.
- [41] El-Dahshan, E.-S. A., T. Hosny, and A.-B. M. Salem, Hybrid intelligent techniques for MRI brain images classification, *Digital Signal Processing*, Vol. 20, No. 2, 433441, 2010.
- [42] Evans, A. C., et al., Brain templates and atlases, *NeuroImage*, Vol. 62, No. 2, 911922, 2012.
- [43] H. Fatakdawala, J. Xu, A. Basavanahally, et al. Expectation maximization driven geodesic active contour with overlap resolution (EMaGACOR): Application to lymphocyte segmentation on breast cancer histopathology. *IEEE Transactions on Biomedical Engineering*, 57(7):16761689, July 2010.
- [44] L. Cohen and R. Kimmel. Global minimum for active contour models: A minimal path approach. *IEEE International Journal on Computer Vision*, 24(1):5778, August 1997.
- [45] Z. Tu. Probabilistic boosting-tree: learning discriminative models for classification, recognition, and clustering. *IEEE International Conference on Computer Vision*, 2:15891596, October 2005.
- [46] P. Tiwari, M. Rosen, G. Reed, and et al. Spectral embedding based probabilistic boosting tree (ScEPTre): Classifying high dimensional heterogeneous biomedical data. *Medical Image Computing and Computer Assisted Intervention*, 1:844851, 2009.

- [47] B. Carneiro, G. and Georgescu, S. Good, and D. Comaniciu. Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree. *IEEE Transactions on Medical Imaging*, 27(9):13421355, September 2008
- [48] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888905, August 2000.
- [49] V. Vazirani. *Approximation Algorithms*. Springer, March 2004.
- [50] M. R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.
- [51] Z. Wu and R. Leahy. An optimal graph theoretical approach to data clustering: theory and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):11011113, November 1993.
- [52] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:359374, September 2001.
- [53] S. Dhillon, Y. Guan, and B. Kulis. Weighted graph cuts without eigen vectors: A multilevel approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):1944 1957, November 2007.
- [54] S. Chandran, S. Hebbar, V. Mamanian, and A. Sawa. Improved cut-based foreground identification. In *The Indian Conference on Computer Vision, Graphics and Image Processing*, pages 447454, 2004.
- [55] K. Fukunaga and L. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, 21(1):32 40, January 1975.
- [56] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603619, May 2002.
- [57] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):790799, August 1995
- [58] C. Yang, R. Duraiswami, N. Gumerov, and L. Davis. Improved Fast Gauss transform and efficient kernel density estimation. *IEEE International Conference on Computer Vision*, 1:664671, October 2003.
- [59] W. Tao, H. Jin, and Y. Zhang. Color image segmentation based on mean shift and normalized cuts. *IEEE Transactions on Systems, Man, and Cybernetics*, 37(5):13821389, 2007.