

# Diagnosis of brain tumor using image processing and determination of its type with RVM neural networks

Elaheh Alipoor<sup>1</sup>, Nasser Lotfivand<sup>2</sup>

<sup>1</sup>Department of Electrical Engineering, Ahar Branch, Islamic Azad University, Ahar, Iran

<sup>2</sup>Department of Electrical Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran

[elahe.alipoor2018@gmail.com](mailto:elahe.alipoor2018@gmail.com), [lotfivand@iaut.ac.ir](mailto:lotfivand@iaut.ac.ir)

## Abstract

*Typically, the diagnosis of a tumor is done through surgical sampling, which is more precise with existing methods. The difference is that this is an aggressive, time consuming and expensive way. In the statistical method, due to the complexity of the brain tissues and the similarity between the cancerous cells and the natural tissues, even a radiologist or an expert physician may also be in error in his diagnosis. Tumor diagnosis is done automatically and various results are achieved. The steps involved in these algorithms can be divided into two sections of the feature discovery and the classification of the samples. The methods generally are that, firstly, the properties of the image are extracted. These characteristics usually include static properties such as entropy, skewness, mean, energy, torque, correlation, etc., or the properties of other algorithms (instant conversion, histogram, etc.). The information obtained at this stage is applied to the sample classification process for decision making. This section is done with an advanced neural network such as RVM. Possible neural networks have the ability to classify more than one class and a kind of radar disease to extract features from MRI images using histogram or satellite conversion techniques, and then selecting appropriate features and ultimately using the system. Fuzzy Neural Network Diagnostics The decision-making system of the fuzzy system is a conclusion that trains with these features and in the output, multiple images are given at different levels. In this research, using image and image processing, we try to find out exactly where the brain is placed. For this purpose, it is initially performed using preventive techniques such as enhancement of contrast, marginalization and morphological functions, and then using the neural network to perform a careful separation of the cancerous parts of the brain health sectors.*

**Keywords:** neuroprotection vector, probability vector, image processing, erotic images, diagnosis of brain tumor

## 1. Introduction

Magnetic resonance imaging (MRI) is a medical imaging technique that provides high quality images of the anatomical structures of the human body, especially the brain, and provides many information for

clinical diagnosis and biology research [1, 5]. MRI diagnostic values are increased by automated and accurate classification of MRI images [6, 8]. Wavelet transform is an effective tool for extracting images from MR brain images, because it allows the analysis

of images at different levels of the resolution according to its multi resolution analysis feature.

However, this technique requires large storage and is expensive to calculate [9]. To reduce the dimensions of the feature vector and increase the discriminatory power, analysis of the main component. (PCA) was used [10]. PCA is attractive because it effectively reduces data and thus reduces computational costs for new data analysis [11]. Then, the problem arises from how the input data is categorized. In recent years, researchers have tried many methods for this purpose, which are divided into two categories.

A categorized monitoring classification, including support vector support (SVM) [12] and unpleasant neighbors (KNN) [13]. The other classifier is unique, [14] including self-organizing charts (SOFM) [12] and physical cmeans [15]. Although all these methods have achieved good results, the classification is better monitored than uncontrolled categories with respect to classification accuracy (speed classification of success).

However, the classification of most existing methods was less than 95%, so the purpose of this article is to use a more precise method. Among the supervised classification methods, SVMs are advanced classification methods based on the theory of machine learning [16, 18]. Compared to other methods such as artificial neural network, decision tree and Bayesian network, SVMs have significant advantages of high precision, elegant mathematical capability, and direct geometric interpretation. Additionally, it does not require a large

number from training samples to prevent overfitting [19].

Main SVMs are linear classifications. In this paper, we introduced the Kernel Squares (KSVMs) that extends the original SVM to nonlinear SVM classifications using the kernel function to replace the dot product form in the original SVM [20]. KSVMs allow us to change the maximum margin in a specified space. This change may be nonlinear and the space is converted to large dimensions; therefore, although the classifier hyperplane in the space of the next features, it may be nonlinear in the main input space [21].

Also, a relational vector machine is also used. RVM algorithm is one of the first and most important steps in data modeling based on the linear combination of basic functions that provides a possible structure for judging data. In the RVM algorithm, although judgment is possible about data, it is more logical than SVM. The goal of the RVM algorithm is to find out the probabilities of each of the classes, while the SVM output shows the fairly distributive generalizations and is difficult to use for more than two classes.

The rest of this paper is structured as follows: The next section provides detailed preprocessing instructions, including wavelet transform (DWT) and core component analysis (PCA). Section 3 first introduces the motivation and principles of linear SVM. Section 4 introduces the cross validation of Kfold, the protection of the classifier from overfitting.

Experiments in Section 5 use 160 images as datasets that show the results of extraction

and reduction of the feature. After that, we compare our approach to different kernels to the latest methods in the decade. The final section of Section 6 is devoted to the results and discussions.

## 2. Preprocessing

In summary, the proposed method in this paper is divided into three stages:

Step 1: Preprocessing (including extracting features and reducing features);

Stage 2. SVM and RVM kernel training

Step 3: Submit the new MRI Brain to the SVM and RVM trained kernel, and then export the prediction.

As shown in Fig. 1, this flow diagram is a conventional and standard classification method that has been proven to be the best classification method [22]. We will explain the precise preprocessing methods below.

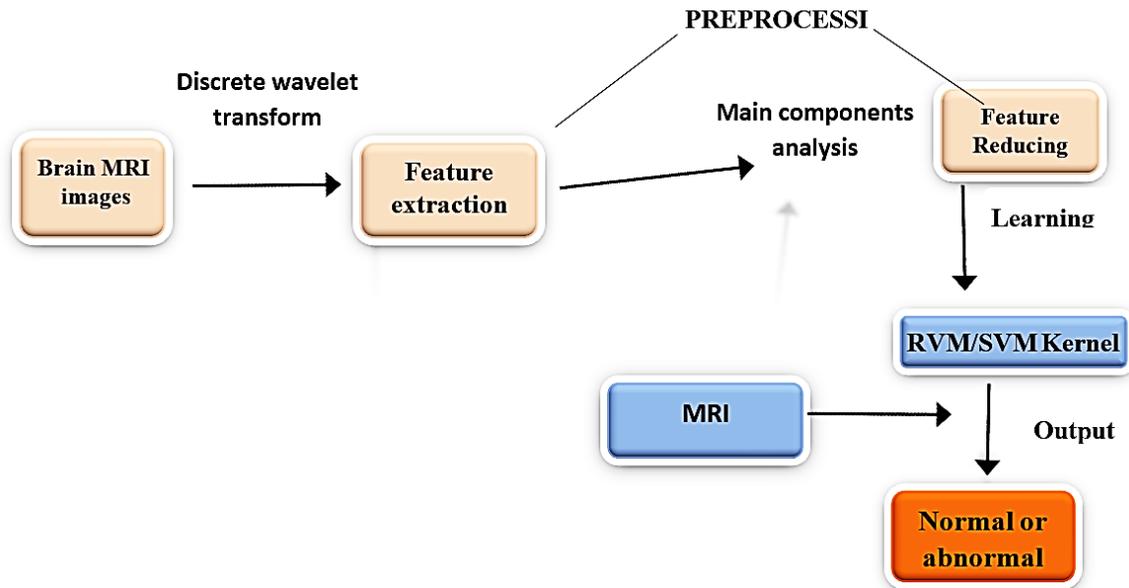


Fig.1. Methodology of our proposed algorithm

## 3. 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. 2. Another advantage of WT is that it adopts "scale" instead of traditional

"frequency", namely, it does not produce a time frequency view but a timescale view of the signal. The timescale 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".

#### 4. Discrete Wavelet Transform

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

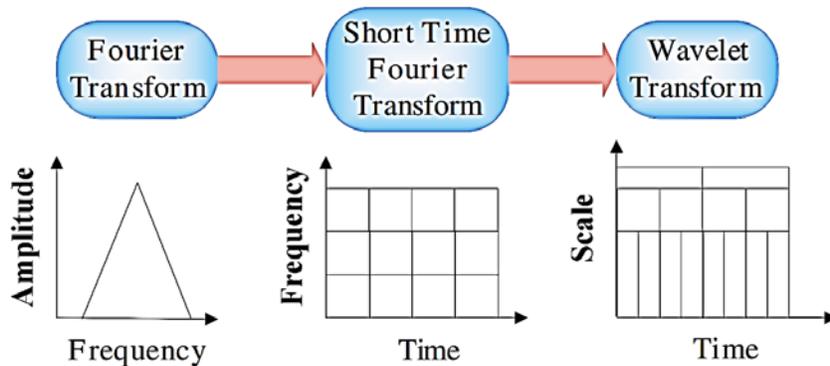


Fig.2.The development of signal analysis.

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:

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

Where

$$\psi_{(a,b)}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-a}{b}\right) \quad (2)$$

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 (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 discretelattice ( $a = 2^j b$  &  $a > 0$ ) to give the DWT, which can be expressed as follows.

$$ca_{j,k}(n) = DS \left[ \sum_n x(n) g_n^*(n - 2^j k) \right] \quad (3)$$

$$cd_{j,k}(n) = DS \left[ \sum_n x(n) h_n^*(n - 2^j k) \right] \quad (4)$$

Here  $ca_{j,k}$ ;  $k$  and  $ca_{j,k}$ ;  $k$  refers 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. Equation (3) is the fundamental of

wavelet decomposes. It decomposes 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 successive approximations being decomposed in turn, so that one signal is broken down into various levels of resolution. The whole process is called wavelet decomposition tree, shown in Fig. 3.

### 5. 2D DWT

In case of 2D images, the DWT is applied to each dimension separately. Fig. 4 illustrates the schematic diagram of 2D DWT. As a result, there are 4 sub-band (LL, LH, HH, and HL) images at each scale. The sub-band LL is used for next 2D DW.

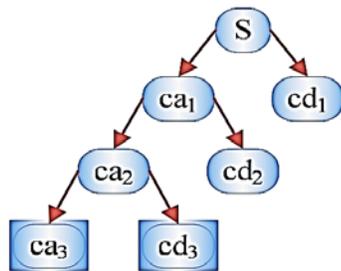


Fig.3.wavelet decomposition tree.

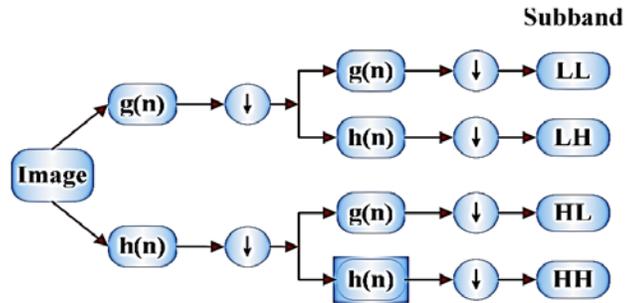


Fig.4.Schematic diagram of 2D DWT.

The LL sub band can be regarded as the approximation component of the image, while the LH, HL, and HH subbands can be regarded as the detailed components of the image. As the level of decomposition increased, compacter but coarser approximation component was obtained. Thus, wavelets provide a simple hierarchical

framework for interpreting the image information. In our algorithm, level-3 decomposition via Harr wavelet was utilized to extract features.

The border distortion is a technique issue related to digital filter which is commonly used in the DWT. As we filter the image, the mask will extend beyond the image at the

edges, so the solution is to pad the pixels outside the images. In our algorithm, symmetric padding method [28] was utilized to calculate the boundary value.

## 6. 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. 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].

## 7. Motivation

Suppose some data points, each of which belong to one of the two classes, are determined, and the purpose of the classification is to determine which new data point will be placed in which class. Here, the data point is viewed as a P-dimensional

vector, here our goal is to create a super-page with (P-1) dimensions. There are a lot of possible super-pages that can successfully classify the data. An appropriate choice as the best super-page is that it represents the maximum distance or margin between the two classes. And we can get the best behavior in response to hidden data during the training. So, we choose the super-page that is the maximum from the point to the nearest point of data on each side [28].

The graph 3-10 illustrates the geometric insertion of a linear SVM. Here H1, H2, and H3 are cloud-based, which can be two classes to be categorized successfully, although H2 and H3 do not have the greatest distance and therefore cannot perform well the new data tests. And the H1 has the most margin and can support vectors (S11, S12, S13, S21, S22, S23) and is therefore chosen as the best classifier super-page. The same routine is also repeated in the relevance vector machine neural networks. However, due to the nature of the approximate and the percentage of the output of the network, and the output is displayed as a percentage relationship with the target in different classes, the output of another measurement criterion is added to the output of the other used algorithms.

Meaning that the output of the RVM network is formed with the attention to the characteristics extracted, including the percentage of brightness and the boundary variance between the class probabilities, for example, if the output of RVM is 0.6 compared to benign and 0.5% malignancy should be considered as a white color and turbidity of the tumor. Be the best

## 8. Results

In this article, we have written for the accurate examination of the proposed methodology of a program in MATLAB software to diagnose and accurately determine the location of a tumor in images. Images in the standard databases used in most articles have been used.

To check the accuracy of the proposed method, we provided the selected photos as inputs to our program in the MATLAB environment. First, the processes described in the previous sections, which relate to determining the position of the tumor in the images, were taken on selected images. Then, in order to calculate the accuracy of the system's performance, the second part of our program was executed on the images of the processing results, which are described in the next section of the procedure. Let's explain experiments on a 3.1 GHz Intel processor, 8 GB RAM core i7, running Windows 10 operating system. The algorithm has been expanded through the Toolbox Wavelet, the Biostatistical Matlab Toolkit 2016b (MathWorks ©).

We used the Toll Box core in our program to provide SVT to the SVM core and the RVM core, a program that was not available on MATLAB itself, using a mathematical math function and a series of ready-made kernels on the MathWorks site to complete this type of network. And we have expanded on this issue to classify the brain's MR images. Applications can be run or tested on any platform that is available.

## 9. Database

The database consists of images of T2 weighted MR brains on the page axis and a 256 x 256 resolution on the page, which is located on the website of the Faculty of Medicine and the following databases (<http://www.med.harvard.edu/aanlib/>)Harvard (<http://adni.loni.usc.edu/>)ADNI (<http://www.oasis-brains.org>) OASIS

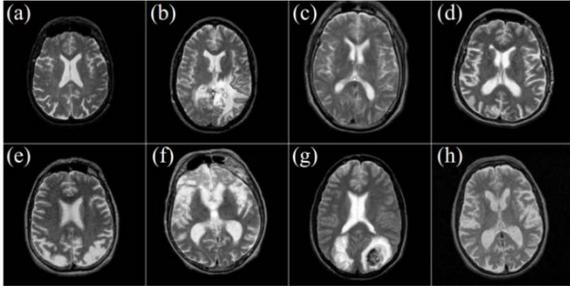
We choose T2 model since T2 images are of higher-contrast and clearer vision compared to T1 and PET modalities.

Abnormal MRI images of the brain include glioma, meningitis, Alzheimer's disease, Alzheimer's disease, visual perception, sarcoma, and pixus disease. An example of any disease is shown in Figure 5 We randomly selected 20 images of each kind of brain. Since there is a type of normal brain and several types of abnormal brain in the database, 160 images are selected including 20 normal brain and 140 abnormal brain images (7 types of illness \* illness / 20 images). A set of education and validation illustrations is shown in Table 2 categories of validation have been used.

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. 5.

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 × 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.



**Fig.5.**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.

## 10. Morphological Operations

A morphological operation refers to an operation that is applied to binary values and aims to modify or correct components within a binary image. This operation is usually performed one step before the final processing operation. The purpose of the final processing operation is to operate in That information is extracted from the image. For example, the environment or area of the image components is computed. Among these operations are four of the most important ones. Expansions, dilation, erosion, open, close Dilation is an operation in which objects grow in a binary image or become thicker in the term. The behavior and size of this thickening process is controlled by the structural member. To perform this operation in the software of the article the command imdilate is used. The Erosion operator is used to shorten or thin objects in a binary image. As in the case of

Dilation, Erosion behavior and mode of operation are determined by a structural member. This operation is performed in the application software using the *imerode* command.

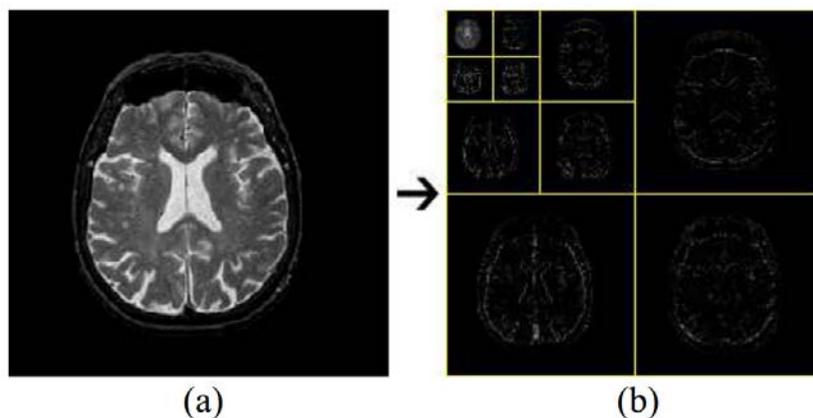
## 11. Feature Extraction

The three levels of wavelet decomposition greatly reduce the input image size as shown in Fig. 6. The top left corner of the wavelet coe-cients image denotes the approximation coe-cients of level-3, whose size is only  $32 \times 32 = 1024$ .

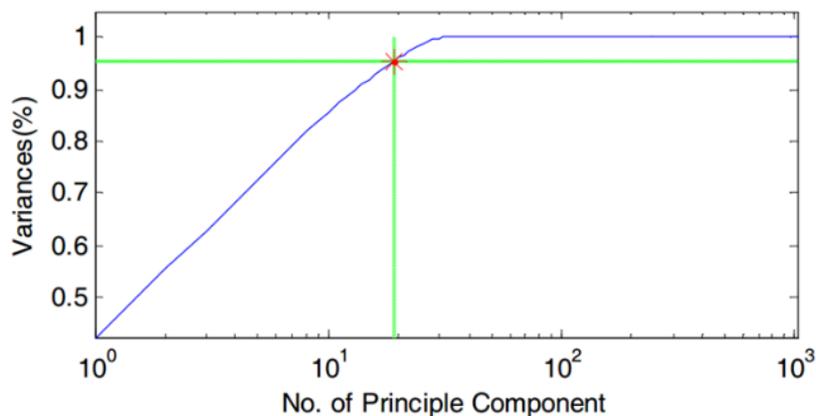
## 12. Feature Reduction

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. 7. The variances versus the number of principle components from 1 to 20 are listed in Table 3. It shows that only 19 principle components (bold font in table), which are only 1.86% of the original features, could preserve 95.4% of total variance.

We tested four SVMs with different kernels (GRB, IPOL, HPOL, LIN) and used the RVM network in the Gaussian kernel. Regarding the use of the linear kernel, KSVM decomposes into a major linear SVM. We compute one hundred simulations to estimate the optimal parameters of core functions such as the D-order in the HPOL core, IPOL, and the Y-factor factor in the GRB kernel.



**Fig.6.**The procedures of 3-level 2D DWT:  
(a) normal brain MRI;(b) level-3 wavelet coefficients



**Fig.7.**Variations against No. of principle components (x axis is log scale).

**Table 2.** Setting of training and validation images (5-fold stratified cross validation)

Total No. of images	Training (128)		Validation (32)	
	Normal	Abnormal	Normal	Abnormal
160	16	112	4	28

**Table 3.** Detailed data of PCA

No. of Prin. Comp.	1	2	3	4	5	6	7	8	9	10
Variance (%)	42.3	55.6	62.4	68.1	72.3	76.2	79.3	82.1	84.0	85.6
No. of Prin. Comp.	11	12	13	14	15	16	17	18	19	20
Variance (%)	87.3	88.6	89.8	91.0	92.0	93.0	93.9	94.6	<b>95.4</b>	96.1

**Table 4.** Confusion matrix of our DWT+PCA+KSVM method (Kernel chose LIN, HPOL, IPOL, and

LIN	Normal (o)	Abnormal (o)
Normal (T)	17	3
Abnormal (T)	5	135
HPOL	Normal (o)	Abnormal (o)
Normal (T)	19	1
Abnormal (T)	4	136
IPOL	Normal (o)	Abnormal (o)
Normal (T)	18	2
Abnormal (T)	1	139
GRB	Normal (o)	Abnormal (o)
Normal (T)	20	0
Abnormal (T)	1	139
(o denotes for output; T denotes for Target)		

The confusion metric of our methods is shown in Table 5. The elements of the i-th line and the j column represent the classification accuracy assigned to the class i defined for the j class after the classification

by supervision. The results show that the proposed DWT + PCA + KSVM method has very fast results in educational and validation images.

**Table 5.** Classification accuracy comparison of 10 different algorithms for the same MRI dataset and same number of images.

Approach from literatures	Classification Accuracy (%)
DWT+SOM [12]	94
DWT+SVM with linear kernel [12]	96
DWT+SVM with RBF based kernel [12]	98
DWT+PCA+ANN [41]	97
DWT+PCA+kNN [41]	98
DWT+PCA+ACPSO+FNN [25]	98.75
Approach from this paper	Classification Accuracy (%)
DWT + PCA + KSVM (LIN)	95%
DWT+PCA+KSVM (HPOL)	96.88%
DWT+PCA+KSVM (IPOL)	98.12%
DWT+PCA+KSVM (GRB)	99.38%

In the case of the RVM network, the results were faster than SVM, but not very large and almost the same, only in relation to the interrelationship network, these percentages were a percentage of the malignancy

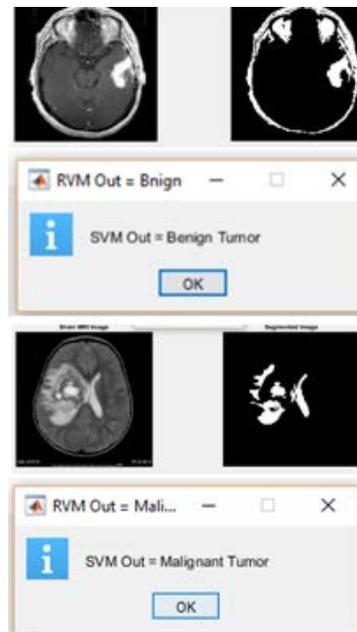
probability and a few percent of the benign probability that we analyzed by re-analysis of this output with a statistical contrast ratio in the output The final result is for the LIN kernel, the overall classification accuracy

is% 95 = 160 / (17 + 135) for the Hpol kernel% 88.96 = 160 / (19 + 136) and for the kernel IPOL% 98.12 = 160 / (18 + 139) and for the GRB kernel% 99.38 = 160 / (20 + 139). It is seen that GRB SVM kernel and RVM neural network are better than the other three SVM kernels. In addition, we compared our method with six methods [DWT + SOM], DWT + SVM with linear kernel, DWT + SVM based on kernel [12], DWT + PCA + ANN [41], DWT + PCA + KNN [41 ], DWT + PCA + ACP SO + FNN [25] described in previous studies using the same MRI database and the same number of images. The results show that the proposed method of DWT + PCA + KSVM with the GRB kernel is best from the middle 10 methods have been taken and accuracy of 98.75% has been achieved. The second method, DWT + PCA + ACP SO + FNN, has a precision of 99.38% and the third is the DWT + PCA + KSVM with the IPOL kernel with a precision of 98.12%.

### 13- Analysis Time

Another important factor in determining the classification is the "calculation time". The time for SVM training was not considered. Since SVM and RVM parameters remain unchanged after training, we use all 160 images for classification. The recording of the calculation time comes from the average amount of time consumed depicted in the various stages described in the previous chapter in the manninology and the chart of this method. For each  $256 \times 256$  image, the average calculation time for feature extraction, the reduction of the SVM classification feature of the RVM

classification is 0.00305 s, 0.0031s, 0.0187s, 0.023s, respectively. The extraction step with the 0.023s has the highest usage time and the cost reduction feature is 0.0187s. The cost of *svm* was reduced at a minimum of 0.0031s. The total computing time for each image is  $256 \times 256$  in about 0.0448 seconds, and it has enough speed to detect real time. The following screenshots will be shown in the final picture of the program, on the left side of the image, and on the right side of the final image and the extracted tumor. In the first 11 photos, the tumors are benign and the program is in the guide box, both the SVM network output and the output of the RVM network is shown below. In the tables below, the output accuracy values of the RBF or Gaussian kernel, the linear kernel, the polygonal polygon, and the quadratic kernel are visible.



**Fig. 9.** The second image is malignant

In the following screenshots, it is clear that the diagnosis of the tumor is moderate to the

95% precision, which is a very high figure, due to the application of hybrid algorithms. The first image shows a benign tumor. In this program, we used the Matlab's dialog box to display the output of each network, which displays the SVM, RVM output in the top and middle of the box.

### **15-Analysis of Results:**

Experiments show that the accuracy of GRB, SVM kernel classification accuracy was 99.38% in 160 MR images, with a higher accuracy than kernels, IPOL and HPOL GRB, and the RVM neural network has the same but higher value, but the strength of this type of neural network is more than classical It is of two types, and also the output percentage that enables the physician to estimate the different probabilities and the type of it. The DWT can efficiently extract information from original MR images with less significant reductions. The benefits of DWT are greater than the Fourier resolution of spatial resolution; DWT acquires a pair of frequencies and local information.

The importance of PCA is shown in the discussion section. If we remove PCA, we will encounter a large search space shown in Table 3 The PCA will reduce the next 1024 search space to the next 19-dimensional search space, which will make the computation and accuracy of the classification better. The proposed method of DWT + PCA + KSVM and PCA + RVM + DWT is superior to the GRB kernel method than the LIN method, HPOL with the IPOL, SVM kernel. The reason is that the GRB kernel is a function representation that can

increase the distance between specimens to its final limit, which HPOL cannot reach. Therefore, we apply the GRB kernel to other industrial areas. This thesis describes the wavelet transform, the PCA, the SVM kernel, and the RVM kernel. The purpose of this thesis is to propose a method that combines them as a powerful tool for detecting the brain's natural brain MRI from abnormal MRI

Meanwhile, we tested four kernels, one of the most successful kernels of the GRB kernel. This method of classification of brain MRI is based on PCA and RVM, KSVM, which is a valuable tool. The accuracy of this method has reached 99% and the accuracy of the output is about 98%, and the sensitivity of this method is 96.8% Which in general has a better performance than previous compilation methods

### **Conclusion**

In this study, we developed a new method of DWT + PCA + KSVM and PCA + RVM + DWT, combining several techniques from processing techniques and image optimizers to find the difference between normal and abnormal brain MRIs. We chose four different kernels of LIN, HPOL, IPOL and GRB in the SVM neural network and the RBF kernel in the RVM neural network.

### **References**

- [1] A P Zijdenbos, R Forghani, and A C Evans. s.l: pipeline Analysis of 3-D MRI Data for Clinical Trials: Application to Multiple Sclerosis". IEEE Transactions on Medical Imaging, 2002. 7.
- [2] Khadem, Mohammad Shajib. Göteborg, Sweden: MRI Brain image segmentation using graph cuts. Master of Science Thesis, 2010.

- [3] R Khayati, M Vafadust, F Towhidkhalah, and S M Nabavi. s.l: Fully Automatic Segmentation of Multiple Sclerosis Lesions in Brain MR FLAIR Images using Adaptive Mixtures method and Markov Random Field Model. *Computers in Biology and Medicine*, 2008
- [4] Kamber, M. s.l: Model-based 3D segmentation of multiple sclerosis lesions in magnetic resonance brain images. *IEEE Trans. Med. Imag.*, 1995
- [5] Udupa, J.K. s.l: Fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation. *Graphical Models Image Process*, 1996
- [6] J.K. Udupa, L. Wei, S. Samarasekera, Y. Miki, M.A. van Buchem, R.I. Grossman. s.l: Multiple sclerosis lesion quantification using fuzzy connectedness principles. *IEEE Trans. Med. Imag.*, 1997, Vol. 598.
- [7] P. Anbeek, K.L. Vincken, M.J.P. van Osch, R.H.C. Bisschops, J. van der Grond. s.l: Probabilistic segmentation of white matter lesions in MR imaging. *NeuroImage* 21, 2004, Vols. 1037-1044
- [8] C. Pachai, Y.M. Zhu, J. Grimaud, M. Hermier, A. Dromigny-Badin, A. Boudraa, G. Gimenez, C. Confavreux, J.C. Froment. s.l: A pyramidal approach for automatic segmentation of multiple sclerosis lesions in brain MRI. *Comput. Med. Imag. Graphics*, 1998.
- [9] Acevedo-Rodriguez, J., et al., \Computational load reduction in decision functions using support vector machines," *Signal Processing*, Vol. 89, No. 10, 2066-2071, 2009.
- [10] Deris, A. M., A. M. Zain, and R. Sallehuddin, \Overview of support vector machine in modeling machining performances," *Procedia Engineering*, Vol. 24, 308-312, 2011.
- [11] 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, 283-294, 2010.
- [12] 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.
- [13] 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, 433-441, 2010.
- [14] Evans, A. C., et al., \Brain templates and atlases," *NeuroImage*, Vol. 62, No. 2, 911-922, 2012.
- [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, 1285-1295, 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, 1030-1037, 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, 119-128, 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, 103-114, 2008.
- [19] Li, D., W. Yang, and S. Wang, Classification of foreign fliers in cotton lint using machine vision and multi-class support vector machine," *Computers and Electronics in Agriculture*, Vol. 4, No. 2, 274-279, 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, 3-13, 2012.
- [21] Hable, R., \Asymptotic normality of support vector machine variants and other regularized kernel methods," *Journal of Multivariate Analysis*, Vol. 106, 92-117, 2012.
- [22] Ghosh, A., B. Uma Shankar, and S. K. Meher, \A novel approach to neuro-fuzzy classification," *Neural Networks*, Vol. 22, No. 1, 100-109, 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, 429-441, 194

- [24] Ahmad Bijar, Rasoul Khayati, Antonio Pen, alver Benavent. s.l: Increasing the Contrast of the Brain MR FLAIR Images Using Fuzzy Membership Functions and Structural Similarity Indices in Order to Segment MS Lesions. University of Maryland, College Park, 2012.
- [25] R.C. Gonzalez, R.E. Woods and S.L.Eddins. Digital image processing using MATLAB. USA: Second edition, Gatesmark publishing, 2009
- [26] Yamamoto, Daisuke, et al. "Computer-aided detection of multiple sclerosis lesions in brain magnetic resonance images: False positive reduction scheme consisted of rule-based, level set method, and support vector machine."Computerized Medical Imaging and Graphics 34.5 (2010): 404-413.
- [27] Daisuke Yamamoto, Hidetaka Arimura, Shingo Kakeda, Taiki Magome, Yasuo. s.l: ComputerizedMedical Imaging and Graphics, 2010.02.001.
- [28] Kamber, M. s.l: Model-based 3D segmentation of multiple sclerosis lesions in magnetic resonance brain images. IEEE Trans. Med. Imag, 1995.
- [29] Udupa, J.K. s.l: Fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation. Graphical Models Image Process, 1996.
- [30] P. Anbeek, K.L. Vincken, M.J.P. van Osch, R.H.C. Bisschops, J. van der Grond. s.l.: Probabilistic segmentation of white matter lesions in MR imaging. NeuroImage 21, 2004, Vols. 1037-1044.
- [31] C. Pachai, Y.M. Zhu, J. Grimaud, M. Hermier, A. Dromigny-Badin,A. Boudraa, G. Gimenez, C. Confavreux, J.C. Froment. s.l.: A pyramidal approach for automatic segmentation of multiple sclerosis lesions in brain MRI. Comput. Med. Imag. Graphics, 1998.