International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

# Brain Magnetic Resonance Images Classification using Wavelet Transform and SVM

Ms. Swati Jayade
Department of Electronics & Telecommunication,
Govt. Poly. Washim, India

Dr. D. T. Ingole
**Principal** Takshashila Institute of Engineering & Technology
Darapur Amravati

***Abstract*—Accurate and automatic computer system for classification of brain MRI images is very much needed in todays era in the medical field for analysis and prediction of the diseases. Since the invention of MRI scanning system many researchers have developed and presented numerous methods for systematic disease detection. This examination work exhibited another system of order for MRI Brian image in two kinds, for example, ordinary (normal) and irregular (abnormal). In this framework right off the bat discrete wavelet transform is applied on pictures for significant trademark extraction, at that point, the standard segment investigation method is utilized to expel the repetitive information, at last, it gives the measurement diminished element vectors. These highlights at that point given as a contribution to a portion support vector machine for the order of images. In this work, seven normal cerebrum infections pictures are considered as irregular. Altogether 160 MRI pictures have alluded for preparing and testing the framework. This proposed system is tried with four sorts of SVM pieces, and it is discovered that the SVM-GRB portion outflanks the arrangement and its precision is 99.38%.**

**Keywords—MRI, Classification, Brain-disease, KSVM, DWT, Feature Reduction.**

## I. Introduction

Magnetic resonance imaging (MRI) technique is very important tool in medical field for detailed analysis of human body, it produces high resolution images of the anatomical structures, specifically in the brain, and generates rich information for disease diagnosis as well as for further research work [1–3]. The computer oriented automatic and perfect classification of the brain MRI images plays a major role in diagnosis of brain related diseases, specifically brain tumors [5–8].

Because of its multi-goals expository property, wavelet change can be utilized for include extraction from MR cerebrum pictures, as it gives the office to examination of pictures of different goals. Along these lines, it is computationally costly and requires enormous sum stockpiling [9]. The chief part examination (PCA) was utilized to lessen the element vector measurements and to improve the discriminative power [10]. PCA hence diminishes the computational expense of dissecting new information as it viably lessens the dimensionality of the information [11]. At that point, the information arrangement issue emerges .

Many scientists have proposed different ways to deal with accomplish this objective, which can be separated into two classifications. The first is administered characterization, with help vector machine (SVM) [12] and k-Nearest Neighbors (k-NN) [13] and the other is unaided arrangement [14], with self-association highlight map (SOFM) [12] and fuzzy c-means [15]. The unaided classifier has lower execution than regulated one as far as grouping precision . Be that as it may, the grouping exactnesses of most existing techniques were lower than 95%, so the point of this paper is to locate an increasingly precise strategy. The regulated strategies for grouping, the SVMs depend on AI hypothesis [16–18]. Contrasted and different techniques, for example, counterfeit neural system, choice tree, and Bayesian system, SVMs have noteworthy favorable circumstances as it needn't bother with countless preparing tests to stay away from over-fitting [19].

In this paper, the portion SVMs, which stretches out unique straight SVMs to nonlinear SVM classifiers is created. This rendition of SVM classifier is acquired by applying the bit capacity which replaces the dab item structure in the first SVMs [20]. The change might be nonlinear and the changed space high dimensional; however the classifier is a hyperplane in the high-dimensional element space, it might be nonlinear in the first info space. The KSVMs enable us to fit the most extreme edge hyperplane in a changed component space [21].

## II. Preprocessing

This proposed method consists of following steps:

1. Pre-processing of image
   a. feature extraction
   b. feature reduction;
2. Training the kernel Support vector machine classifier
3. Testing the new MRI image on the trained kernel SVM, for classification.

The detail schematic of the work is shown in Fig.1 as a standard classification method [22].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

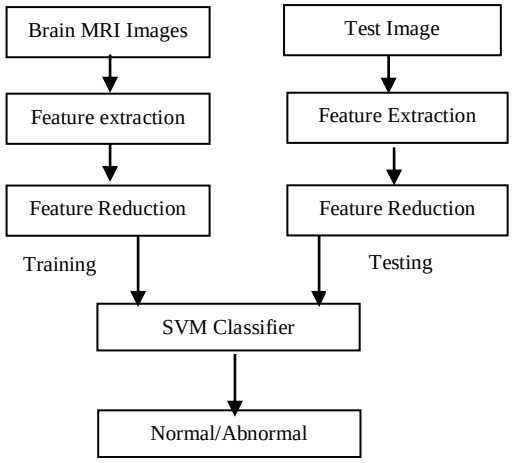


Fig. 1. Methodology of our proposed algorithm.

### A. Feature Extraction

The Fourier transform is utilized for signal investigation, which separates a period space signal into constituent sinusoids of various frequencies. This changes the sign from time area to recurrence space. As FT has a disadvantage of disposing of the time data of the sign. Therefore, the nature of the grouping diminishes as time data is lost. Breaking down just a little segment of the sign at a time is called windowing or brief time Fourier change (STFT), utilized by Gabour [23]. It includes a window of a specific shape to the sign. STFT gives some data about both time and recurrence area however the accuracy of the data is constrained by the size of the window. The following stage is Wavelet change (WT): a windowing system with variable size. It jams both time and recurrence data of the sign. Another advantage of WT is that it produce rather than a time-frequency view of the signal. A time-scale view is a more natural and powerful way, because compared to "frequency", "scale" is commonly used term.

### B. Discrete Wavelet Transform

The discrete wavelet transform (DWT) developed with the help of dyadic scales and positions which makes it powerful for implementation of the WT [24].

The important aspects of DWT are given 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 parent wavelet $\psi(t)$ by translation and dilation: $a$ is the dilation factor and $b$ the translation parameter (both real positive numbers).

The most basic and 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 ($a = 2^b$ & $a > 0$) to give the DWT, which can be expressed as follows.

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

where $ca_{j,k}$ 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. $DS$ stands for down-sampling. Equation (3) is the fundamental of wavelet decomposes. It decomposes signal $x(n)$ into two signals, the detail components $cd$ $(n)$ and the approximation coefficients $ca(n)$. This is called one-level decompose.

The process of wavelet decomposition is depicted with the help of decomposition tree, as shown in Fig.2. The above down sampling process can be iterated with successive approximations being decomposed in all turn, so that one signal is down sampled into many levels of resolution.

### C. 2D DWT

In case of 2D images, the DWT is applied to each dimension separately. Fig.3 shows the schematic diagram of 2D DWT which results in 4 sub-band (LL, LH, HH, and HL) images at each scale. The sub-band LL is used for next 2D DWT.

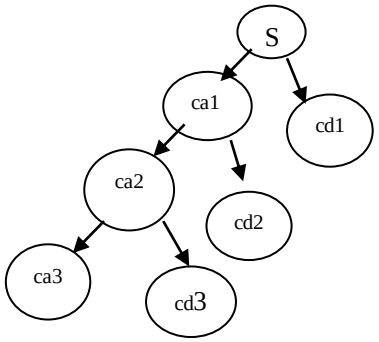


Fig. 2. A 3-level wavelet decomposition tree.

An image is segmented in various sub bands, the LL subband can be approximation component of the image, while the LH, HL, and HH are high frequency components of the image. As the level of decomposition increased, but coarser approximation component was obtained. Wavelet transform gives a simple hierarchical framework for interpreting the image information. In this algorithm, level-3 decomposition is done via Haar wavelet which was utilized to extract features.

In this system symmetric padding method [28] was utilized to calculate the boundary value. When we filter the image, the mask will extend beyond the image at the edges,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

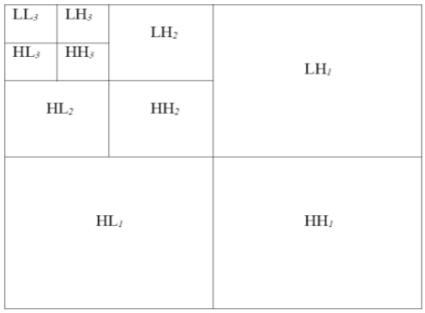


Fig. 3. 2D DWT decomposion

so the solution is to pad the pixels outside the images. As the border distortion is a technique issue related to digital filter which is commonly used in the DWT.

### *D. Feature Reduction.*

Exorbitant highlights make characterization progressively muddled just as increment calculation times and capacity memory. It is required to diminish the quantity of highlights. PCA is best to diminish the element of an informational collection. It is accomplished by changing the informational collection to another arrangement of requested factors as indicated by their differences or importance. PCA comprising of countless interrelated factors while holding the vast majority of the varieties. This works by three way first, it orthogonalizes the segments of the info vectors so uncorrelated with one another, second requesting of the subsequent symmetrical parts so that those with the biggest variety start things out, lastly takes out those segments with least variety in the informational index.

## III. KERNEL-SVM

The appearance of the support vector machine (SVM) is a milestone inside the subject of the framework contemplating. The gifts of SVMs incorporate inordinate exactness, rich numerical tractability, and direct geometric elucidation [29]. As of late, two or three ventured forward SVMs have developed quickly, among which the part SVMs are the most well known and successful. Piece SVMs have the accompanying advantages [30]: (1) well working by and by and have been surprisingly fruitful in such various fields as common language arrangement, bioinformatics and pc inventive and judicious; (2) have barely any tunable parameters; and (3) preparing routinely involves raised quadratic advancement [31]. Subsequently, arrangements are worldwide and for the most part explicit, as needs be keeping off the assembly to close by minima displayed by methods for various factual acing structures, comprehensive of neural systems. Information is seen as a p-dimensional vector, and our undertaking is to make a (p 1)- dimensional hyperplane. One sensible decision as the best hyperplane is the one that speaks to the biggest contrast between the two classes, for better conduct because of concealed information during preparing, i.e., better characterization.

Hence , in this work the hyperplane method is used so that the distance from the computed classification line to the closest sample data point on each side is maximized [32]. Fig. 4 shows the schematic of linear SVMs, where H, H1, H2 are three hyperplanes lines, it can classify the two classes successfully, however, but H1 and H2 does not have the largest margin, so they will not give good results to new test data. The H hyperplane has the maximum difference value to the support vectors, so it is selected as the best vector for classification [33].

### *A. Principles of Linear SVMs*

Given a *p*-dimensional *N* -size training dataset of the form

$$\{(x_n, y_n)|x_n \in R^p, y_n \in \{-1, +1\}\}, \quad n = 1, \ldots, N \tag{4}$$

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

$$\mathbf{w} \cdot \mathbf{x} - b = 0 \tag{5}$$

where it denotes the dot product and **w** the normal vector to the hyperplane. It is needed to choose the **w** and *b* to maximize the margin between the two parallel (as shown in Fig. 5) hyperplanes as large as possible while still separating the data. So we define the two parallel hyperplanes by the equations as

$$\mathbf{w} \cdot \mathbf{x} - b = \pm 1 \tag{6}$$

In this manner, the errand can be changed to an improvement issue, i.e., we need to amplify the separation between the two parallel hyperplanes, subject to avoid information falling into the edge.

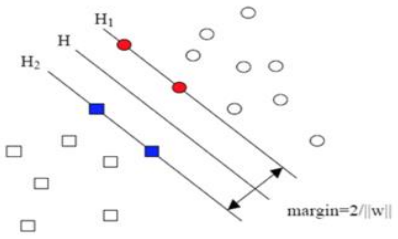


Fig. 4. Schematic of linear SVMs (H denotes for the hyperplane, S denotes for the support vector) [43].

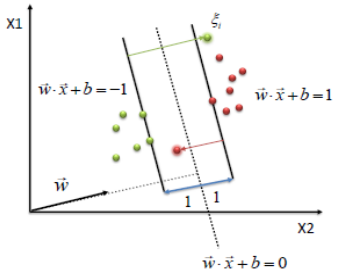


Fig. 5. : The concept of parallel hyperplanes (*w* denotes the weight, and *b* denotes the bias).

Using simple mathematical knowledge, the problem can be formulated as:

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

$$\min_{\mathbf{w},b} \|\mathbf{w}\|$$
$$s.t.\ y_n\ (\mathbf{w} \cdot x_n - b) \geq 1,\ \ n = 1, \ldots, N \qquad (7)$$

In practical situations the $\|\mathbf{w}\|$ is usually be replace by

$$\min_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2$$
$$s.t.\ y_n\ (\mathbf{w} \cdot x_n - b) \geq 1,\ \ n = 1, \ldots, N \qquad (8)$$

The reason is that the ||w|| is involved in a square root calculation. After it is lowered with formula (8), the solution will will remain as it is, but the problem is altered into a quadratic programming optimization that is easy to solve by using Lagrange multipliers [34] and standard quadratic programming techniques and programs [35, 36].

Regular SMVs gives a hyperplane to classify data, so it cannot deal with classification problem of which the different types of data located at different sides of a hypersurface, the kernel strategy is applied to SVMs [37]. This algorithm is similar, except that every dot product is replaced by a nonlinear kernel function. The kernel is related to the transform ϕ(xi) by the equation k(xi, xj) = ϕ(xi)ϕ(xj). The value w is also in the transformed space, with

$$w = \sum_i \alpha_i \gamma_i \phi(x_i).$$

Dot product with *w* for classification can be computed by

$$w \cdot \phi(x) = \sum_i \alpha_i \gamma_i k(x_i, x).$$

By the other side KSVMs allow to fit the maximum-margin hyperplane in a transformed feature space. The transformed space is higher dimensional and the transformation may be nonlinear and; thus though the classifier is a hyperplane in the higher-dimensional feature space, it may be nonlinear in the original input space. For each kernel, there should be at least one adjusting parameter so as to make the kernel flexible and tailor itself to practical data. Three common kernels [38] are listed in Table 1.

TABLE I. THREE COMMON KERNELS (HPOL, IPOL, AND GRB) WITH THEIR FORMULA AND PARAMETERS.

| Name | Formula |
|---|---|
| Homogeneous Polynomial | $k(xi, xj) = (xi \cdot xj)^d$ |
| Inhomogeneous Polynomial | $k(x_i, x_j) = (x_i \cdot x_j + 1)^d$ |
| Gaussian Radial Basis | $k(x_i,x_j)=\exp -\gamma\|x_i-x_j\|$ |

## IV. CROSS VALIDAION

As for training of classifier some data set is used, so there is possibility to achieve high accuracy for classification only for this training dataset not yet other independent datasets. For this overfitting avoidance, we need to incorporate cross validation into our method. Cross validation will not improve the final classification accuracy, but it will make the classifier reliable and can be generalized to other independent datasets.

Cross approval incorporates three sorts of strategies: Random subsampling, K-fold cross approval and forget about one approval. The K-overlay cross approval is executed because of its straightforwardness, effortlessness, and utilizing all information for preparing and approval. The philosophy is to make a K-fold segment of the entire dataset, K times rehashed to utilize K- 1 folds for preparing and a left overlay for approval, and finally, the normal mistake is determined for K tests.

The K folds might be absolutely haphazardly partitioned, as, a few folds may have been dispersed differently when contrasted with different folds. Subsequently, stratified K - crease cross approval was actualized, in which each overlay has appropriated about in comparative classes [39]. The following test is to locate the number of folds. In the event that K is set excessively enormous, the calculation will be tedious on the grounds that the inclination of the genuine mistake rate estimator will be little, and the change of the estimator will be huge. On the other hand, if K is set excessively little, the fluctuation of the estimator naturally turns out to be little, yet the inclination of the estimator will be enormous, henceforth the calculation time will diminish, [40]. In this study, initially K is determined as 5 for working on trial-and-error basis, which means that with increasing step as 1and if parameter K varying from 3 to 10, we can train the SVM by each value. At the end we will select the optimal value of K belonging to the highest classification accuracy.

## V. EXPERIMENTAL RESULTS

### A. Databaset

The datasets used for experimentation consists of brain MRI images having resolution 256 by 256. The abnormal brain MRI images of the dataset consist of the seven types of diseases viz: glioma, meningioma, Alzheimer's disease, Alzheimer's disease plus visual agnosia, Pick's disease, sarcoma, and Huntington's disease. The samples images of each disease are given in Fig. 6. The confusion matrix for training and validation purpose of images is given in Table 2.

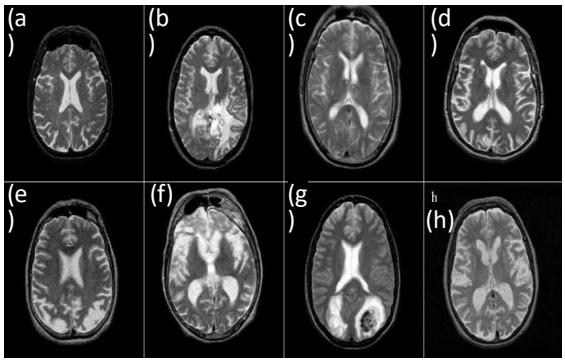


**Fig** 6 :Sample of MRI images: (a) normal brain; (b) glioma;(c) meningioma; (d) Alzheimer's disease; (e) Alzheimer's disease with visual agnosia; (f) Pick's disease; (g) sarcoma; (h) Huntington's disease[43].

### B. Feature Extraction

To extract features from MRI images .wavelet transform is applied to the third level, it gives different chaacteristics,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

first level transform is shown in Fig. 7. The approximate coefficient matrix is used for input to get the prominent features only.

### C. Feature Reduction

After getting the approximate wavelet coefficient image, PCA is applied to reduce the dimensions of features.

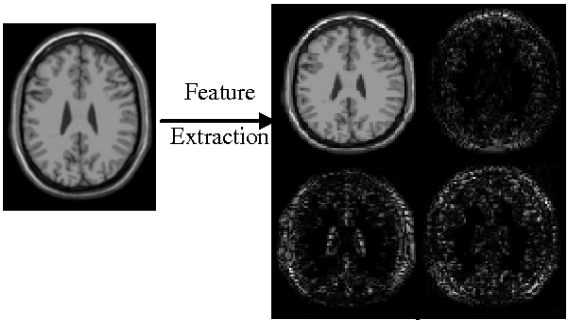


Fig 7: The procedures of 1-level 2D DWT: Brain MRI; level-1 wavelet transformed coefficients.

TABLE II. CONFUSION MATRIX OF OUR DWT+PCA+KSVM METHOD (KERNEL CHOSE LIN, HPOL, IPOL, AND GRB [13].

| 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 |

### D. Classification Accuracy

SVM is tested with four different kernels (LIN, HPOL, IPOL, and GRB). Result indicates that the case of using linear kernel, the KSVM degrades to original linear SVM.

Hundreds of simulations are computed 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 $\gamma$ in GRB kernel. The confusion matrices of our methods are listed in Table 2. The element of $i^{th}$ row and $j^{th}$ column represents the classification accuracy belonging to class $i$ are assigned to class $j$ after the supervised classification. The empirical study shows that the proposed combined DWT+PCA+KSVM method obtains good results on training and validation ofimage data. For LIN kernel, the classification accuracy is 94.58%; for HPOL kernel, it is 96.88%;for IPOL kernel, it is 98.12%; and for the GRB kernel, it is 99.38%. It indicates that, the GRB kernel based SVM is giving better results than other three kernels. These methods are compared with other techniques given in literature like DWT+SOM [12], DWT+SVM with linear kernel [12], DWT+PCA+ANN [41], DWT+PCA+$k$ NN [41]. The comparative results are shown in Fig. 8. It indicates that the method DWT+PCA+KSVM with GRB kernel performed best than other ten methods.

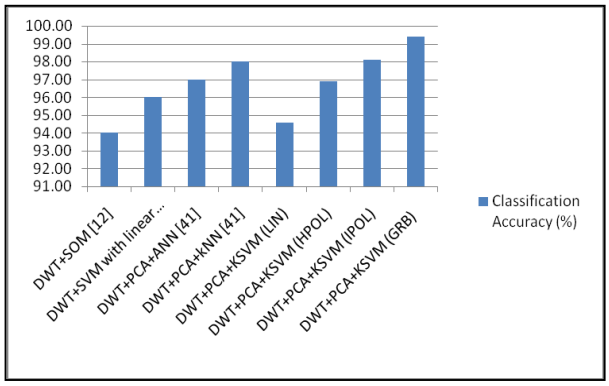


Fig 8: The comparative results of the existing systems(first four) with implemented methods (las four).

## VI. CONCLUSION AND DISCUSSION

There are many techniques developed based on wavelet transforms, PCA, and kernel SVMs. The wavelet transform extracts the features from brain MRI images. The major contribution of this research paper is to present a novel method which uses DWT+PCA+KSVM for classification of normal MRI or abnormal MRI images. Four different SVM kernels as LIN, HPOL, IPOL and GRB are implemented. The experimental results indicates that the SVM-GRB kernel gives 99.38% classification accuracy on the dataset. It is better than the other existing techniques.

## 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, 65–79, 2017.

[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, 81–91, 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, 1–16, 2018.

[4] Mohsin, S. A., "Concentration of the specific absorption rate around deep brain stimulation electrodes during MRI," *Progress In Electromagnetics Research*, Vol. 121, 469–484, 2015.

[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, 345–360, 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, 305–324, 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, 165–180, 2017.

[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, 23–42, 2011.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

[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, 1600–1607, 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. Pic´o, and A. Ferrer, "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.

[12] Chaplot, S., L. M. Patnaik, and N. R. Jagannathan, "Classifica- tion 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.

[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, 513–527, 2003.

[14] Zhang, Y. and L. Wu, "Weights optimization of neural network via improved BCO approach," *Progress In Electromagnetics Research*, Vol. 83, 185–198, 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, 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 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.

[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, 1946.

[24] Zhang, Y. and L. Wu, "Crop classification by forward neuralnetwork with adaptive chaotic particle swarm optimization,"*Sensors*, Vol. 11, No. 5, 4721–4743, 2015.

[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, 325–343, 2010.

[26] Ala, G., E. Francomano, and F. Viola, "A wavelet operator on the interval in solving Maxwell's equations," *Progress In Electromagnetics Research Letters*, Vol. 27, 133–140, 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, 35–52, 2017.

[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, 2018.

[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 categorization from otolith images using multi-class support vector 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 effect 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*, 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, 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 with data," *Notices of the American Mathematical Society (AMS)*, Vol. 50, No. 5, 537–544, 2003.

[37] Acevedo-Rodr´ıguez, J., et al., "Computational load reduction in decision functions using support vector machines," *Signal Processing*, Vol. 89, No. 10, 2066–2071, 2009.

[38] 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.

[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, 283–294, 2017

[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, 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, 2018.

[42] Evans, A. C., et al. , "Brain templates and atlases, " *NeuroImage*, Vol. 62, No. 2, 911–922, 2018.

[43] Y. Zhang and L. Wu, "An MR Brain Images Classifier Via Principal Component Analysis And Kernel Support Vector Machine", *Progress In Electromagnetics Research*, Vol. 130, 369-388, 2012