International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

Cumulant Features based Classification of Brain MR Images using ANN and LS-SVM Algorithm

Harikumar Rajaguru, Sannasi Chakravarthy S R

Abstract*: *Automatic classification of magnetic resonance (MR) brain images using machine learning algorithms has a significant role in clinical diagnosis of brain tumour. The higher order spectra cumulant features are powerful and competent tool for automatic classification. The study proposed an effective cumulant-based features to predict the severity of brain tumour. The study at first stage implicates the one-level classification of 2-D discrete wavelet transform (DWT) of taken brain MR image. The cumulants of every sub-bands are then determined to calculate the primary feature vector. Linear discriminant analysis is adopted to extract the discriminative features derived from the primary ones. A three layer feed-forward artificial neural network (ANN) and least square based support vector machine (LS-SVM) algorithms are considered to compute that the brain MR image is either belongs to normal or to one of seven other diseases (eight-class scenario). Furthermore, in one more classification problem, the input MR image is categorized as normal or abnormal (two-class scenario). The correct classification rate (CCR) of LS-SVM is superior than the ANN algorithm thereby the proposed study with LS-SVM attains higher accuracy rate in both classification scenarios of MR images.

Keywords*: *magnetic resonance, brain tumour, cumulant features, wavelet transform, linear discriminant analysis, neural network, support vector machine (SVM).

I. INTRODUCTION

The acquiring procedure of brain magnetic resonance image is a non-invasive with low-risk to provide high quality brain images and produces detailed information for disease diagnosis and research. The MR images are the most common choice of medical imaging technique while soft-tissue delineation is required. This is certainly a right choice for any automatic system that efforts to classify brain diseases [1]. The brain anatomy with a high degree of accuracy and resolution could be perceived positively through MR images. The acquisition is essentially a non-invasive procedure that delivers features like greater soft-tissue differentiation, superior spatial resolution with improved contrast. The MR imaging tool does not provide any radiation impairment to the brain tissues because of not using any ionizing radiation while imaging [2].

The practice of using computer technology in clinical decision support is widespread and pervasive in the medical field such as gastroenterology, cancer research, heart diseases, breast and brain tumors, etc. A complete automatic classification of normal and affected human brain tumours can be designed using the MR images which plays a significant role in both clinical and research studies [3]. In recent times, machine learning algorithms are attaining great attention in the diagnosis of brain tumour. Moreover, the primary thing is to classify the MR image to predict the existence of tumour. If there is an existence of tumour, then it is required to classify its type for better diagnosis [4].

Literature [2-5] implies that a several range of automated clinical tools for diagnosis have been introduced by exploiting necessary signal or image processing methods using transforms along with some computational intelligent algorithms. The classification algorithms used for MR brain images could fall into two groups. The first kind of approach is supervised one such as artificial neural network (ANN), support vector machine (SVM) and k-nearest neighbour algorithms (K-NN). The next kind of approach is unsupervised one such as self-organization map (SOM) and fuzzy c-means (FCM). However, based on the classification accuracy, the supervised type achieves better classification rate than the unsupervised one [5]. It is perceived that the extraction of features from MR images can be done through many popular signal or image examination techniques such as independent component analysis (ICA), wavelet and Fourier transform based techniques, etc. The wavelet transform has turn out to be a wide choice for several medical image analysis and classification problems. The wavelets are superior in extracting frequency space details from a non-stationary signals [6]. The control over the resolution of the analysis is done appropriately by modifying the wavelets in the selected data sequence. The study proposed here utilizes the discrete wavelet transform for the phase of feature extraction since it offers instantaneous details on frequency and time localization together for the image characteristics and is much needed for the computer based analysis. But the methodology will have more computational complexity when it uses higher levels of wavelet coefficients. And if any classification system uses the values of wavelet coefficients, then it is sensitive to rotation of input MR images [7]. The work makes use of one-level 2-D discrete wavelet transform and its robust statistical features that are non-sensitive to rotation of input images. The work involves eight-class classification problem, normal class along with seven different classes (Alzheimer, alzheimer plus visual agnosia, huntington, glioma, meningioma, pick and sarcoma disease) and two-class classification problem where the input MR image is classified as normal class or abnormal class.

Manuscript published on 30 September 2019.
***Correspondence Author(s)**
Sannasi Chakravarthy S R, Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathyamangalam, India. Email: elektroniqz@gmail.com
Harikumar Rajaguru, Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathyamangalam, India. Email: harikumarrajaguru@gmail.com

The study employs higher-order statistics such as cumulant features which are efficient to classify the different classes of input brain MR images. The figure 1 shows the flow of proposed study. As in the figure 1, the first stage of proposed method involves the computation of one-level 2-D DWT from the input brain MR images and then the cumulants (primary feature vectors) of four sub-bands are calculated. Afterwards, the cumulant features are exposed to data reduction by means of LDA which builds the reduced secondary feature vectors. As a final point, a three layer feed-forward artificial neural network and least square based support vector machine algorithms are adopted to classify the extracted secondary features as normal or suffers from any class of disease.

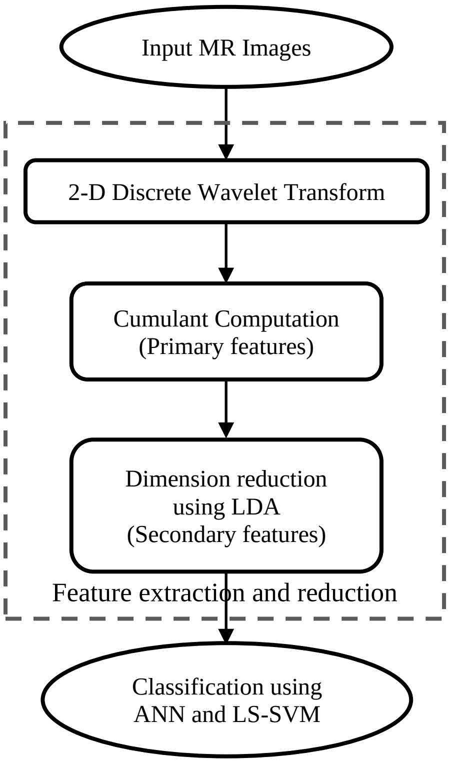

Fig. 1.Proposed Work.

The section II describes the materials and methodology used for the work. The algorithms for classification is discussed in the section III. The section IV shows the results and discussion of the work and the section V describes the concluding remarks.

II. Methods and Materials

A. Dataset

The work employs normal MR images together with seven different brain diseases. Ten MR images are adopted for each class, thus eighty MR brain images are employed for the evaluation. These MR brain images are downloaded from the Harvard medical school website (http://med.harvard.edu/AANLIB/) [8]. All these 80 images are T2-weighted in axial plane with resolution equal to 256 x 256 pixels [8]. Ten normal MR brain images belong to one class and seventy other diseased MR brain images belong to abnormal class in the two-class problem whereas each class of brain disease is taken as a separate class in the eight-class problem.

B. One-level 2-D discrete wavelet transform

The discrete wavelet transform is primarily a linear transform that works on the input data and transforms it as another domain vectors. The transform is operated at each dimension individually if the input is an image; i.e. 2-D DWT. The nature of wavelet transform is that it will also affords localized frequency details of an input image [9]. The work involves the computation of two-dimensional DWT of input MR images using Daubechies-6 basis filters with one level-one decomposition. This results in obtaining distinct sub-bands of first level 2-D DWT for the input brain MR images.

As a result of applying the discrete wavelet transform for the input images, four distinct sub-bands are calculated as LH (low-high), LL (low-low) and HH (high-high) and HL (high-low) sub-bands. The LL sub-band is assumed as the approximation component (low-frequency) of input images whereas the other LH, HH and HL sub-bands indicate the vertical, diagonal and horizontal components of the images respectively. The figure 2 shows the input MR image (Alzheimer brain disease) together with four sub-bands of 2-D DWT at level-one decomposition.

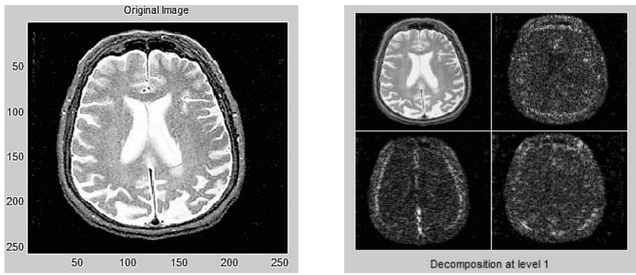

Fig. 2.(a) Input MR image (b) Four sub-bands

C. Cumulant Computation

To make the extracted features as effective, the cumulants of wavelet coefficients for attained four different sub-bands is calculated. The figure 2 (b) implies that utmost all the pixels in LH, HH and HL sub-bands have very lower gray-scale details as compared with the LL sub-band. The pixels in these (LH, HH and HL) sub-bands do not hold any information and so considered to be redundant [10]. Thus before the computation of cumulants, the pixels having absolute gray-scale values inferior than the taken threshold (0.005) are removed. The threshold value of 0.005 has been obtained by experimental analysis [11]. Now computation of cumulants for the remaining pixels of respective sub-band is carried out. Cumulants should give robust statistical with rotation invariant features for the classification of brain MR images. They represent the high order statistics with the shape of distribution [12]. The cumulant features provide the volume of higher order correlation along with distance measurement of a random process for the Gaussian functions. Cumulants are simply a nonlinear mixture of moments and the j^{th} moment of a real-valued zero mean random variable (X) is determined as [13]

Retrieval Number: K24310981119/19©BEIESP
DOI: 10.35940/ijitee.K2431.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

$$M_j = E[X^j] \quad (1)$$

where E[.] is the expectation operator on the random variable. The cumulants of different orders(C2, C3, C4, C5 and C6) are calculated for each sub-band as [13]:

$$C_1 = M_1 \quad (2)$$

$$C_2 = M_2 - M_1^2 \quad (3)$$

$$C_3 = M_3 - 3M_2M_1 + 2M_1^3 \quad (4)$$

$$C_4 = M_4 - 4M_3M_1 - 3M_2^2 + 12M_2M_1^2 - 6M_1^4 \quad (5)$$

$$C_5 = M_5 - 5M_4M_1 - 10M_3M_2 + 20M_3M_1^2 + 30M_2^2M_1 - 60M_2M_1^3 + 24M_1^5 \quad (6)$$

$$C_6 = M_6 - 6M_5M_1 - 15M_4M_2 + 30M_4M_1^2 - 10M_3^2 + 120M_3M_2M_1 - 120M_3M_1^3 + 30M_2^3 - 270M_2^2M_1^2 + 360M_2M_1^4 - 120M_1^6 \quad (7)$$

where the first order cumulant (equation 2)represents the mean that is not considered for classification. In two-class problem, C2, C3, C4 andC5 cumulants are calculated for each sub-band so that the principal feature vector for each input MR brain image has 4 x 4 as 16 primary feature vectors. As in the eight-classproblem, C2, C3, C4, C5 and C6 cumulants are for each sub-band so that the principal feature vector for each input MR brain image has 4 x5as 20 primary feature vectors.

D. Dimension Reduction

The purpose of dimension reduction in the study is to provide a condensed and informative representation of the derived primary features. Afterwards, these primary feature vectors are transformed into a set of new features termed as secondary feature vectors. The study utilizes a linear and supervised form of dimension reduction approach, LDA. It performs the removal of redundancies in the derived primary feature subspace. The key idea of the approach is that LDA always explore the project axes where the data points of distinct classes are distant from other ones and the data points of identical classes are nearer to each ones [14]. The primary cumulants have different dynamic range, so they ought to be normalized to the range between [0, 1] before dimension reduction using LDA.

Sorting of eigenvalues of normalized feature vectors in descending order is done to determine the effective feature vectors for classification. The normalized cumulative total ($\lambda_{CS}(i)$) of sorted eigenvalues relative to the ith one of sorted feature is determined as [15],

$$\lambda_{CS}(i) = \frac{\sum_{j=1}^{i} \lambda(j)}{\sum_{j=1}^{N} \lambda(j)} \quad (8)$$

where the value of i ranges from 1 to N, N represents the total amount of primary feature vectors. The amount of efficient features is found once λ_{CS} reaches to unity, then the features related to those normalized eigenvalues will create the secondary features.

III. Classification Algorithms

A. Artificial Neural Network

Feed forward artificial neural network is widely used for the applications of patternclassification [16]. The firstinput layer consists of input neurons, the middle layer with hidden neurons and the final layer with output neurons. Thefirst input layer has 12 nodes related to each applied input, the intermediate layer contains 10 nodes andthe final output layerhas eight nodes and two nodes for eight-class and two-class classification respectively. The classification using ANN is initiated by assigning the corresponding random weights.Thenthe reduced feature sets are given to the neural network to determineits response in the networkoutput layer. Depending upon the output class, the error is calculated and back-propagated to update the neural network weights. The above process is repeated once the value of mean squarederror (MSE) falls below a threshold [17]. After the completion of training the neural network, the testing data is given to the already trained neural networkso that the output is analysed and classifier performanceis assessed.

For training the classifiers, three MR input images are randomly selected from the normal category and other seven categories. Then testing of classifiers is done with the remaining set of MR images. Thus 24 training set and 56 testing set of MR images is considered for both classification scenario. There are various ways for choosing three images from ten images. In this paper, we consider 50 different sets.

B. Least Square – SVM

Support Vector Machine is a highly non-linear network used widely for binary and multi-class classification. SVM has superior generalization capability to classify any unknown input data accurately [18]. It works simply by reducing the structural risk and increasing the distance among the input features to the class that separating the hyper-plane respectively. Generally if the feature vectors are not linearly separable, then the vectors are transformed to a higher-order dimensional space so that they will become linearly separable and termed as least square based SVM [19]. The LS-SVM are least squares versions of SVM. The aim of LS-SVM is to put on the minimization of sum of squared errors (SSE) to the considered objective function and it is a kernel-based learning technique [20]. The work employs LS-SVM with radial basis function (RBF) kernel function for classification.

IV. Results and Discussion

The evaluation ofthe study is attained for training and testing set using correct classification rate (CCR). The CCR is calculated as the addition of number on the diagonal of confusion matrix divided by the numberof samples in test data [21].The minimum CCR, maximum CCR, mean CCR and standard deviation of CCR of all the input setsare computed.The figure 3 shows the normalized cumulative total (λ_{CS}) of sorted eigenvalues for both the classification scenarios, since they provide $\lambda_{CS} =$ 1 as discussed in the section 2.4.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

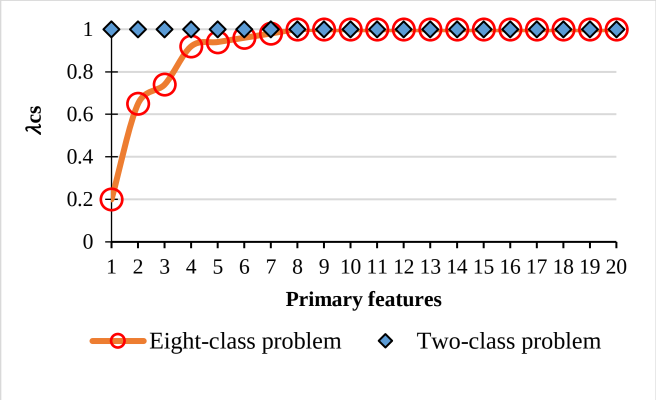

Fig. 3. Normalized cumulative total (λ_{CS}) of sorted eigenvalues

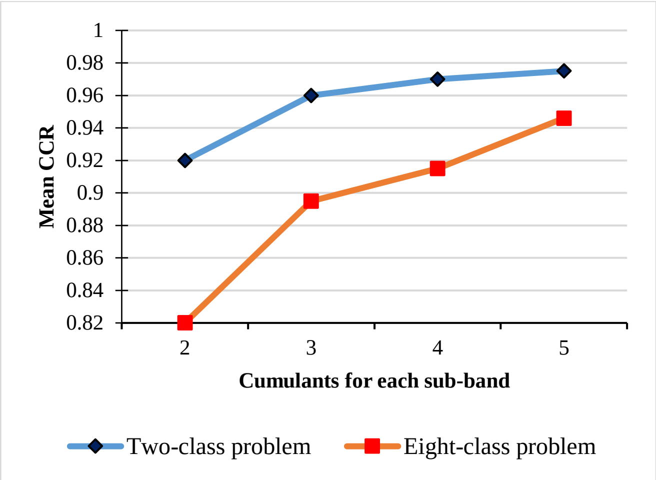

Fig. 4. Performance of ANN classifier for different number of cumulants

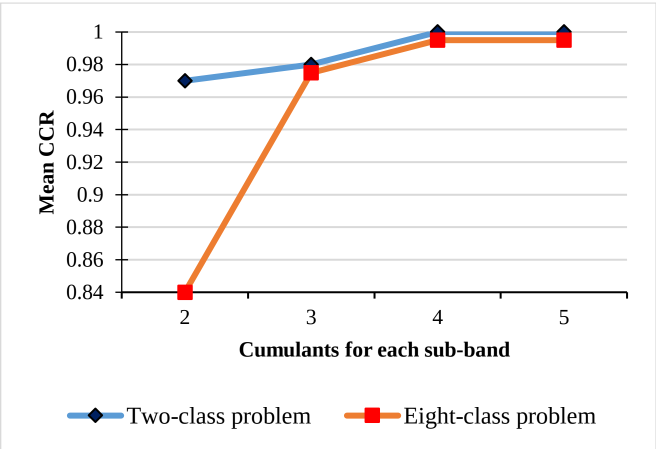

Fig. 5. Performance of LS-SVM classifier for different number of cumulants

The figures 4 and 5 show the calculated mean value of obtained CCR of ANN and LS-SVM classifiers for distinct number of cumulants for each wavelet sub-band. In two-class problem, four cumulants for each wavelet sub-band attain the superior CCR value so that the primary vector should have 16 features and in the eight-class problem, five cumulants for each wavelet sub-band provide the superior CCR value so that the primary vector should have 20 features.

The table I gives the statistics of values of CCR for different sets using ANN and LS-SVM classifiers. As in the table I, it is noted that in two-class problem, LS-SVM attains 100% CCR for the input sets and in eight-class problem, LS-SVM algorithm attains higher accuracy and lower standard deviation than ANN.

Table- I: Statistic Values of CCR of Classifiers

Classifiers	Problem	CCR			
		Minimum	***Maximum***	***Mean***	***Standard Deviation***
ANN	Two-class (4 cumulants for each sub-band)	0.943	0.9614	0.8593	0.2851
	Eight-class (5 cumulants for each sub-band)	0.8911	0.9517	0.8291	0.8226
LS-SVM	Two-class (4 cumulants for each sub-band)	1	1	1	0
	Eight-class (5 cumulants for each sub-band)	0.9751	1	0.9929	0.0064

As from the above table I, LS-SVM attains superior performance over ANN algorithm for both the classification scenarios. Also, the confusion matrix for the LS-SVM for binary classification is simply a 2 x 2 identity matrix since it provides 100% CCR as in the table I. The confusion matrix of ANN and LS-SVM for eight-class problem is given in table II and III where the confusion matrix of all other input sets are averaged.

Table- II: Confusion Matrix for ANN Classifier for eight-class problem

Actual Class	Predicted Class							
	1	***2***	***3***	***4***	***5***	***6***	***7***	***8***
1	0.8543	0	0	0	0.1457	0	0	0
2	0	0.8561	0	0	0	0	0.1439	0
3	0	0	0.8915	0	0.1085	0	0	0
4	0	0	0.0768	0.8366	0	0.0866	0	0
5	0	0	0	0	1	0	0	0
6	0	0	0.1533	0	0	0.8467	0	0
7	0	0	0	0.2072	0	0	0.7928	0
8	0	0	0	0	0	0.2029	0	0.7971

Table- III: Confusion Matrix for LS-SVM Classifier for eight-class problem

Actual Class	Predicted Class							
	1	***2***	***3***	***4***	***5***	***6***	***7***	***8***
1	1	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0
4	0	0	0	1	0	0	0	0
5	0	0	0	0	1	0	0	0
6	0	0	0.0154	0	0	0.9846	0	0
7	0	0	0	0	0.0187	0	0.9813	0
8	0	0	0	0	0	0.0226	0	0.9774

V. CONCLUSION AND FUTURE

The work shows that the cumulants as higher-order spectra features can afford efficient representation of input brain MRimages to decide the appropriateclass. The subbands of 2-D wavelet are calculated; fourth and fifth order cumulants of the pixels areobtained after removing the redundant pixels corresponding to two-class scenario and eight-class scenario. Then primary features with 20 and 16 vectorsareapplied to LDA so that the secondary features with one and seven vectorsarecalculatedforthe bothclassification scenarios. The obtained results shows that the LS-SVM algorithmprovidesunity classification accuracy for two-class scenario and gives superior classification accuracy (≈ 1) for eight-class problem.The results of performancecomparison between ANN and LS-SVM indicates that the higher-order cumulants with LS-SVM algorithm provides higher accuracy than ANN in both classification scenarios.

REFERENCES

1. McDonald, Robert J., Jennifer S. McDonald, David F. Kallmes, Mark E. Jentoft, Michael A. Paolini, David L. Murray, Eric E. Williamson, and Laurence J. Eckel. "Gadolinium deposition in human brain tissues after contrast-enhanced MR imaging in adult patients without intracranial abnormalities." Radiology 285, no. 2 (2017): 546-554.
2. Wiesinger, Florian, Laura I. Sacolick, Anne Menini, Sandeep S. Kaushik, SangtaeAhn, Patrick Veit-Haibach, Gaspar Delso, and Dattesh D. Shanbhag. "Zero TE MR bone imaging in the head." Magnetic resonance in medicine 75, no. 1 (2016): 107-114.
3. Yang, Gelan, Yudong Zhang, Jiquan Yang, Genlin Ji, Zhengchao Dong, Shuihua Wang, Chunmei Feng, and Qiong Wang. "Automated classification of brain images using wavelet-energy and biogeography-based optimization." Multimedia Tools and Applications 75, no. 23 (2016): 15601-15617.
4. Zhang, Yudong, Shuihua Wang, Ping Sun, and Preetha Phillips. "Pathological brain detection based on wavelet entropy and Hu moment invariants." Bio-medical materials and engineering 26, no. s1 (2015): S1283-S1290.
5. Moeskops, Pim, Manon JNL Benders, Sabina M. Chiţă, Karina J. Kersbergen, Floris Groenendaal, Linda S. de Vries, Max A. Viergever, and Ivana Išgum. "Automatic segmenta-tion of MR brain images of preterm infants using supervised classification." NeuroImage 118 (2015): 628-641.
6. Abirami, C., R. Harikumar, and SR Sannasi Chakravarthy. "Performance analysis and de-tection of micro calcification in digital mammograms using wavelet features." In 2016 Inter-national Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 2327-2331. IEEE, 2016.
7. Sannasi Chakravarthy, SR, Rajaguru, H. Detection and classification of microcalcification from digital mammograms with firefly algorithm, extreme learning machine and non-linear regression models: A comparison. Int J Imaging Syst Technol. 2019; 1– 21. https://doi.org/10.1002/ima.22364
8. Harvard Medical School, data available at www.med.harvard.edu/AANLIB/.
9. Gupta, Dipalee, and Siddhartha Choubey. "Discrete wavelet transform for image pro-cessing." International Journal of Emerging Technology and Advanced Engineering 4, no. 3 (2015): 598-602.
10. Lidong, Huang, Zhao Wei, Wang Jun, and Sun Zebin. "Combination of contrast limited adaptive histogram equalisation and discrete wavelet transform for image enhancement." IET Image Processing 9, no. 10 (2015): 908-915.
11. Kalbkhani, H., Shayesteh, M.G. and Rashidi, A., 2019, April. Brain Tumor Classification from MRI Images Based on Cumulant Features. In 2019 27th Iranian Conference on Electri-cal Engineering (ICEE) (pp. 1746-1750). IEEE.
12. Elhaj, Fatin A., Naomie Salim, Arief R. Harris, Tan Tian Swee, and Taqwa Ahmed. "Ar-rhythmia recognition and classification using combined linear and nonlinear features of ECG signals." Computer methods and programs in biomedicine 127 (2016): 52-63.
13. Orosco, Eugenio, Pablo Diez, Eric Laciar, Vicente Mut, Carlos Soria, and Fernando Di Sciascio. "On the use of high-order cumulant and bispectrum for muscular-activity detection." Biomedical Signal Processing and Control 18 (2015): 325-333.
14. Tharwat, Alaa, Tarek Gaber, Abdelhameed Ibrahim, and Aboul Ella Hassanien. "Linear discriminant analysis: A detailed tutorial." AI communications 30, no. 2 (2017): 169-190.
15. Sharma, Alok, and Kuldip K. Paliwal. "Linear discriminant analysis for the small sample size problem: an overview." International Journal of Machine Learning and Cybernetics 6, no. 3 (2015): 443-454.
16. Carleo, Giuseppe, and Matthias Troyer. "Solving the quantum many-body problem with artificial neural networks." Science 355, no. 6325 (2017): 602-606.
17. Van Gerven, Marcel, and Sander Bohte. "Artificial neural networks as models of neural information processing." Frontiers in Computational Neuroscience 11 (2017): 114.
18. Tehrany, MahyatShafapour, Biswajeet Pradhan, ShattriMansor, and Noordin Ahmad. "Flood susceptibility assessment using GIS-based support vector machine model with differ-ent kernel types." Catena 125 (2015): 91-101.
19. Kisi, Ozgur, and Kulwinder Singh Parmar. "Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution." Journal of Hydrology 534 (2016): 104-112.
20. Deo, Ravinesh C., Ozgur Kisi, and Vijay P. Singh. "Drought forecasting in eastern Aus-tralia using multivariate adaptive regression spline, least square support vector machine and M5Tree model." Atmospheric Research 184 (2017): 149-175.
21. Sasaki, JefferEidi, Amanda Hickey, John Staudenmayer, Dinesh John, Jane A. Kent, and Patty S. Freedson. "Performance of activity classification algorithms in free-living older adults." Medicine and science in sports and exercise 48, no. 5 (2016): 941.

AUTHORS PROFILE

Sannasi Chakravarthy S Ris pursuing his research inthe area of medical image processing, machine learningand artificial intelligence.

Harikumar Rajaguruhas completed his doctoral degree in the area of bio-signal processing and hisresearch area includes medical image processing,machine learning, soft computing and artificialintelligence.