IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 16, 2016, accepted October 21, 2016, date of current version November 18, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2620996

Multiple Sclerosis Detection Based on Biorthogonal Wavelet Transform, RBF Kernel Principal Component Analysis, and Logistic Regression

SHUI-HUA WANG[1,2,3,10], TIAN-MING ZHAN[1,10], YI CHEN[1,2], YIN ZHANG[4], MING YANG[5], HUI-MIN LU[6], HAI-NAN WANG[2], BIN LIU[7], AND PREETHA PHILLIPS[8,9]

[1]School of Technology, Nanjing Audit University, Nanjing 211815, China
[2]School of Computer Science and Technology, Nanjing Normal University, Nanjing 210023, China
[3]Department of Electrical Engineering, The City College of New York, New York, NY 10031, USA
[4]School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan 430073, China
[5]Department of Radiology, Nanjing Children's Hospital, Nanjing Medical University, Nanjing 210008, China
[6]Department of Mechanical and Control Engineering, Kyushu Institute of Technology, Fukuoka 804-8550, Japan
[7]Department of Radiology, Zhong-Da Hospital of Southeast University, Nanjing 210009, China
[8]School of Natural Sciences and Mathematics, Shepherd University, Shepherdstown, WV 25443, USA
[9]West Virginia School of Osteopathic Medicine, Lewisburg, WV 24901, USA
[10]S. H. Wang and T.M. Zhan contributed equally to this paper

Corresponding author: Y. Chen (cs_chenyi@njnu.edu.cn), Y. Zhang (yin.zhang.cn@ieee.org), M. Yang (yangming19710217@163.com), and H.-M. Lu (dr.huimin.lu@ieee.org).

This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grant BK20150523 and Grant BK20150983, in part by NSFC under Grant 61502206 and Grant 61602250, in part by the Program of Natural Science Research of Jiangsu Higher Education Institutions under Grant 16KJB520025, in part by the Open Fund of Fujian Provincial Key Laboratory of Data Intensive Computing under Grant BD201607, and in part by the Open Fund of Key Laboratory of Statistical Information Technology and Data Mining, State Statistics Bureau under Grant SDL201608.

ABSTRACT To detect multiple sclerosis (MS) diseases early, we proposed a novel method on the hardware of magnetic resonance imaging, and on the software of three successful methods: biorthogonal wavelet transform, kernel principal component analysis, and logistic regression. The materials were 676 MR slices containing plaques from 38 MS patients, and 880 MR slices from 34 healthy controls. The statistical analysis showed our method achieved a sensitivity of $97.12 \pm 0.14\%$, a specificity of $98.25 \pm 0.16\%$, and an accuracy of $97.76 \pm 0.10\%$. Our method is superior to five state-of-the-art approaches in MS detection.

INDEX TERMS Biorthogonal wavelet transform, kernel principal component analysis, logistic regression, multiple sclerosis, computer vision, machine learning.

I. INTRODUCTION

Multiple sclerosis (MS) affects human brain and spinal cord by damaging the insulating covers of neural cells [1]. The cause is unclear, thus the underlying mechanism is either immune system destruction [2] or myelin-producing cell failure [3]. Clinically, MS is associated with depression [4], lower urinary tract symptom [5], fatigue [6], muscle weakness [7], etc.

To detect MS early, the neuroradiologists tend to use magnetic resonance imaging (MRI) technique to scan the patients' brains. Nevertheless, the normal-appearing white matter (NAWM) paradox [8], [9] poses a radical challenge, since the lesions within the white matter may appear the same as healthy white matter.

With the rapid development in computer science, the computer vision (CV) [10], [11] has high probabilities to help neuroradiologists to detect MS. CV studies and mimics the human vision, and thus gaining high-level understanding on digital images and videos. It can implement any tasks that a human visual system can do. Artificial intelligence (AI) [12] can ease the process that CV learns to understand the contents in image and video. Therefore, CV is often combined with AI [13] and its variants, such as machine learning [14], bio-inspired mechanism [15], expert system [16], swarm intelligence [17], etc.

Current CV systems on brain diseases capture more exciting attentions from scholars in both research and industrial domains. For example, Murray et al. [18] extracted features

from MS images. They employed a multiscale amplitude-modulation frequency-modulation (abbreviated as MAMFM) method. Finally, support vector machine (SVM) was used. Phillips et al. [19] suggested a novel feature—wavelet entropy (WE)—for abnormal brain detection. To train the classifier, a **H**ybridization of **B**iogeography-based optimization and **P**article swarm optimization (HBP) was proposed. Siddiqui et al. [20] presented a combined system based on discrete wavelet transform (DWT), principal component analysis (PCA), and least-square support vector machine (LS-SVM). Nayak et al. [21]proposed a novel abnormal MR image detector, based on DWT, probabilistic PCA (PPCA), and random forest (RF). Zhou [22] used stationary wavelet entropy (SWE) to detect MS. They compared three algorithms: decision tree (DT), k-nearest neighbors (kNN), and SVM.

Nevertheless, the accuracy performances of above mentioned references are not satisfying. Besides, their statistical analysis only reported the average result, other than the standard deviation. In this study, we presented a novel MS detection method based on biorthogonal wavelet transform, kernel principal component analysis, and logistic regression. The structure is organized as follows: Section II gives the materials. Section III offers the methodology. Section IV presents the results and discussions. Section V concludes the paper.

II. MATERIALS

In this study, we obtains 676 MR slices containing plaques from 38 MS patients, and 880 MR slices from 34 healthy controls. The detailed description of those data can be found in reference [22]. The MS patients and healthy controls are scanned by different scanners in different position. To ease brain slice comparability, we used the histogram stretching (HS) [23] method to increase the dynamic range of all MS and healthy brain images. The HS was performed as follow:

$$b(i,j) = \frac{a(i,j) - a_{\min}}{a_{\max} - a_{\min}} \tag{1}$$

where (i, j) represents the coordinate of the pixels, a represents original slice, b the HS normalized slice. The $a_{\min}$ and $a_{\max}$ represent the minimum and maximum intensity values, respectively. Figure 1 shows the samples of our used brain slices.

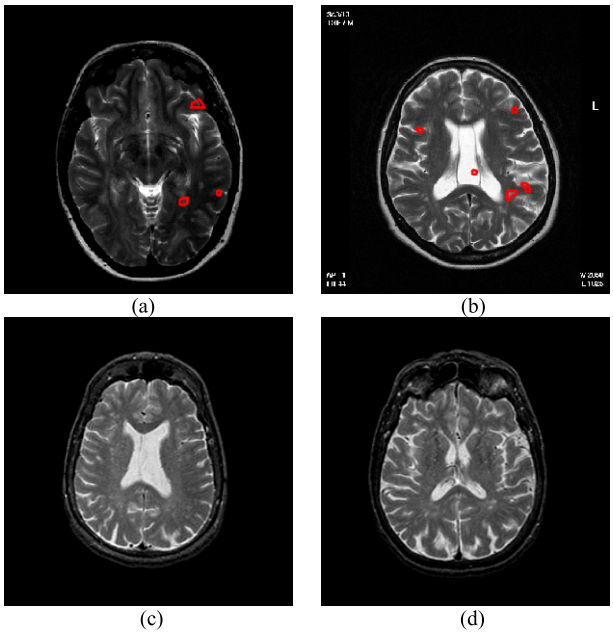


FIGURE 1. Sample of brain slices. (a) A MS slice with 3 plaques. (b) A MS slice with 5 plaques. (c) A healthy brain slice. (d) Another healthy brain slice.

III. METHODOLOGY

A. DISCRETE WAVELET TRANSFORM

In numerical analysis, the discrete wavelet transform (DWT) is an effective way to extract global features from images or videos. It is also used in JPEG 2000—an image compression standard and coding system [24] and the fingerprint identification systems [25]. In academic fields, DWT is applied in various fields, e.g., classification of MR image [26], hearing loss detection [27], pathological brain detection [28], video watermarking [29], abnormal brain detection [30], infant cry detection [31], dendrite spine detection [32], biometric template generation [33], tea classification [34], etc.

Mathematically, the DWT of a brain image x is obtained by passing it through a series of filters. The discrete samples of $x$are passed through a low-pass filter g and a high-pass filter h, resulting the approximation coefficients (AC) and detail coefficients (DC), respectively. The filters output are usually down-sampled by a factor of 2.

$$AC(n) = \sum_{m=-\infty}^{+\infty} x(m)g(2n-m) \tag{2}$$

$$DC(n) = \sum_{m=-\infty}^{+\infty} x(m)h(2n-m) \tag{3}$$

These two filters are known as the quadrature mirror filter. Figure 2 shows the diagram of passing through filters.

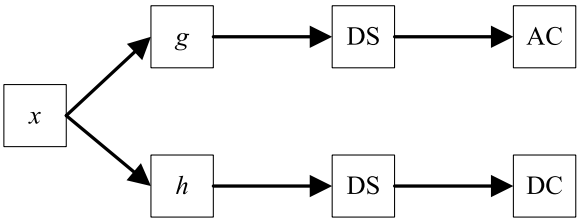


FIGURE 2. Diagram of passing through filters (DS = down-sampling, AC = approximation coefficient, DC = detail coefficient).

B. BIORTHOGONAL WAVELET TRANSFORM

There are many wavelet families, such as Haar [35], db [36], and others. In this study, we chose the biorthogonal wavelet. The advantage of orthogonal wavelet is the associate wavelet transform is orthogonal, thus, the inverse wavelet transform

is the adjoint of the wavelet transform. The advantage of biorthogonal wavelet transform (BWT) is it allows more degrees of freedom compared to orthogonal wavelet [37].

In this study, we chose the biorthogonal 4.4 wavelet. Its filters and functions for decomposition are shown in Figure 3 and Figure 4, respectively. The corresponding filters and functions for reconstruction are not presented, since our task only uses decomposition.

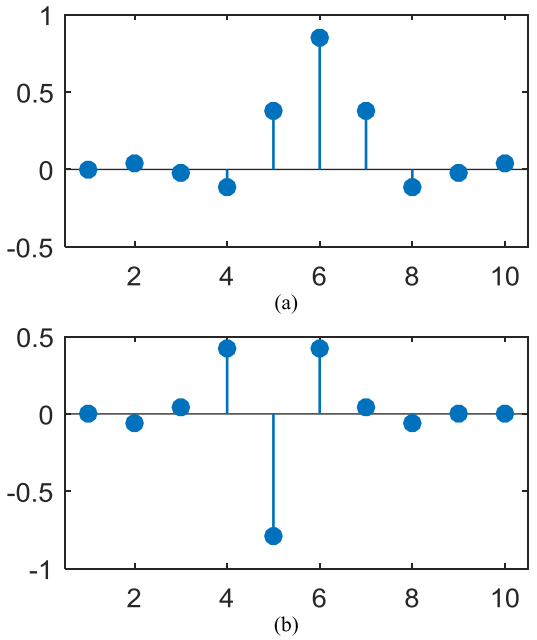


FIGURE 3. Filters of biorthogonal 4.4: (a) low-pass filter *g*; (b) high-pass filter *h*.

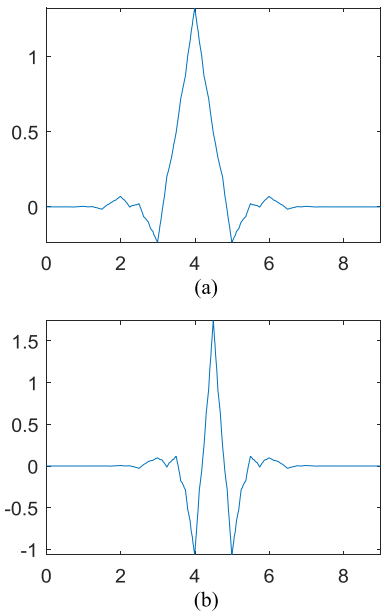


FIGURE 4. Functions of biorthogonal 4.4: (a) scaling function; (b) wavelet function.

Besides BWT, there are many other excellent wavelet transform variants, such as wavelet packet transform [38], relative wavelet energy [39], wavelet energy [40], scale-discretized wavelet transform [41], stationary wavelet transform [42], spherical wavelet transform [43], exponential wavelet transform [44], dual-tree complex wavelet transform [45], etc. Those advanced wavelet transforms are also expected to give better performance than standard DWT. In the future, we shall test their performances.

C. PRINCIPAL COMPONENT ANALYSIS

As an effective dimensionality reduction tool, principal component analysis (PCA) can reduce the size of wavelet coefficients from MR brain images [46]. Assume there is a dataset C with size of N and dimension of d, first we calculate the sample mean m_j of j-th feature as

$$m_j = \frac{1}{N}\sum_{i=1}^{N} C(i,j) \tag{4}$$

Next, we calculate the zero-mean dataset B as

$$B = C - em^T \tag{5}$$

Here e represents an $N \times 1$ vector of all ones [47].

Third, the $d \times d$ covariance matrix Z is generated

$$Z = \frac{B^*B}{N-1} \tag{6}$$

Fourth, the covariance matrix Z has an eigen decomposition expression as

$$Z = XYX^{-1} \tag{7}$$

here X represents the eigenvector matrix, and Y represents the eigenvalue matrix, which is also a diagonal matrix[48].

$$Y = \begin{bmatrix} Y(1,1) & & & \\ & Y(2,2) & & \\ & & \ddots & \\ & & & Y(d,d) \end{bmatrix} \tag{8}$$

Fifth, we rearrange X and Y, so that the eigenvalue is in a decreasing way.

$$Y(1,1) \geq Y(2,2) \geq \cdots \geq Y(d,d) \tag{9}$$

Sixth, we calculate cumulative variance for each eigenvector by

$$G(k) = \sum_{i=1}^{k} Y(i,i) \tag{10}$$

Thus, we can form a vector as

$$G = \begin{bmatrix} G(1)\ G(2) \cdots G(d) \end{bmatrix} \tag{11}$$

Seventh, assume the threshold is T, and thus we select L^* that satisfies

$$L^* = \arg\min \left\{ L \left| \frac{G(L)}{G(d)} \geq T \right. \right\} \tag{12}$$

Finally, we output L^* most important principal components.

D. KERNEL PCA

The shortcoming of PCA is it cannot extract non-linear structure information [49]. To solve this problem, scholars have proposed a powerful variant of PCA—kernel PCA (KPCA). The KPCA implements the same as PCA except transforming the dataset C into a higher-dimensional space [50].

Two different KPCAs were studied. One is the polynomial kernel PCA (PKPCA) defined as

$$k(x, y|\text{PKPCA}) = [a(x \times y) + b]^c \tag{13}$$

where a, b, and c are kernel parameters

The other is the RBF kernel PCA (RKPCA) [51] defined as

$$k(x, y|\text{RKPCA}) = \exp\left(-\frac{\|x - y\|^2}{d^2}\right) \tag{14}$$

where d represents the scaling factor.

The optimal estimation of hyper parameters a, b, c, and d can be obtained by grid search (GS) algorithm. GS is also named as parameter sweep. It is an exhaustive searching method within a manually specified subset of the hyper parameter space.

E. LOGISTIC REGRESSION

Traditional regression analysis help the users understand the relationship between a dependent variable and on or more independent variables. Logistic regression is an improved regression model that can handle the situation where dependent variable is categorical [52]. In this study, we predict a bran MR image as either MS or healthy. This prediction output belongs to a binary categorical variable.

For a binary logistic regression, the output is usually encoded as either 0 or 1 [53]. Following common convention, we encode the particular noteworthy output as 1, here the MS patient. We also encode the contrary output as 0, here as the healthy. Table 1 shows the encoding strategy for the output.

TABLE 1. Output encoding.

Value	Meaning
1	MS patient
0	Healthy Subject

Suppose we have L principal components as $[x_1, x_2, \ldots, x_L]$ as the independent variable, and we have one dependent variable y either 0 or 1 indicating healthy or MS patient. Then, we can create the binary logistic regression model as to find the optimal vector $\beta = [\beta_0, \beta_1, \beta_2, \ldots, \beta_L]$ that best fits

$$y = \begin{cases} 1 & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_L x_L + \varepsilon > 0 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

here ε represents the unobservable error.

To achieve above model, a challenge arise as to smash the input (with values from negative to positive infinity) to the output (with values between 0 and 1). The logistic function

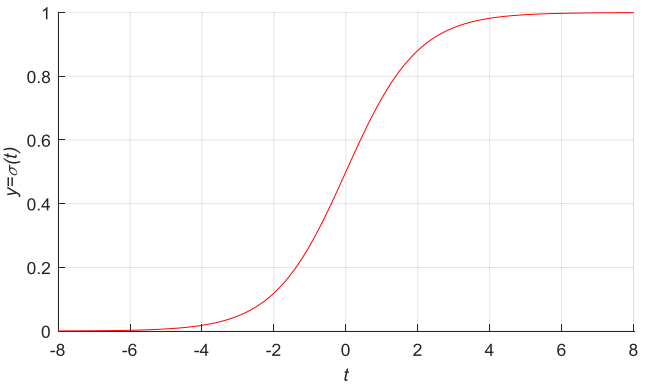


FIGURE 5. Logistic function.

$\sigma(t)$ can solve this problem [54]. $\sigma(t)$ is defined below with a curve plot shown in Figure 5.

$$\sigma(t) = \frac{1}{1 + \exp(-t)} \tag{16}$$

In this study, t can be regarded as a latent variable, which is a linear weighted combination of independent variable x as

$$t \leftarrow \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_L x_L \tag{17}$$

Thus, the binary logistic model is:

$$F(x) = \frac{1}{1 + \exp\left[-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_L x_L)\right]} \tag{18}$$

where $F(x)$ represents the probability of dependent variable $y = 1$, i.e., corresponding to a MS patient. β_0 is the intercept. $[\beta_1, \beta_2, \ldots, \beta_L]$ represents the regression coefficient for $[x_1, x_2, \ldots, x_L]$.

There are other advanced classifiers besides LR, such as feed forward neural network [55], association rule learning [56], decision tree [57], dynamic Bayesian network [58], nonparallel support vector machine [59], reinforcement learning [60], twin support vector machine [61], extreme learning machine [62], etc. Those classifiers have a radically different mechanism with the LR, but they may also give satisfying performances. In the future, we shall apply them to MS detection.

F. STATISTICAL ANALYSIS

Before we step into the experiment, we need to point out the importance of statistical analysis and its relationship to the hyper parameters of $[a, b, c, d]$.

We used a ten-fold cross validation (TFCV) as shown in Fig. 6. The whole dataset was segmented into 10 folds (A to J). In every trial, eight folds out of 10 folds were used for training, one fold for validation, and the final fold for test. The purposes of the three sets are listed in Table 2. Note that the classifier needs to be retrained for each trial. To further reduce the randomness, we ran the TFCV ten times, and report the average and the standard deviation in terms of sensitivity, accuracy, and specificity.

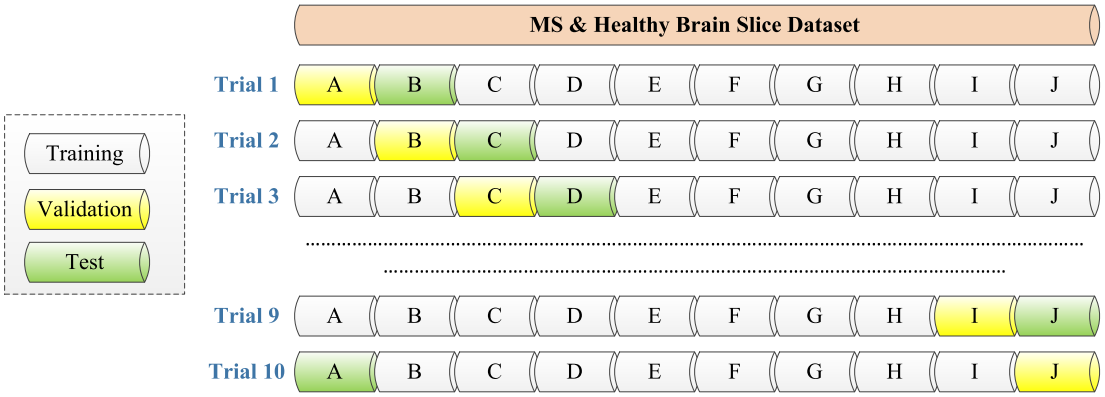


FIGURE 6. Illustration of TFCV.

TABLE 2. Purpose of training, validation and test sets.

Set	Purpose
Training	Reduce training error
Validation	Optimize hyperparameters
Test	Give unbiased performance estimation

TABLE 3. Comparison of different dimensionality reduction methods (Threshold = 95% of total variance).

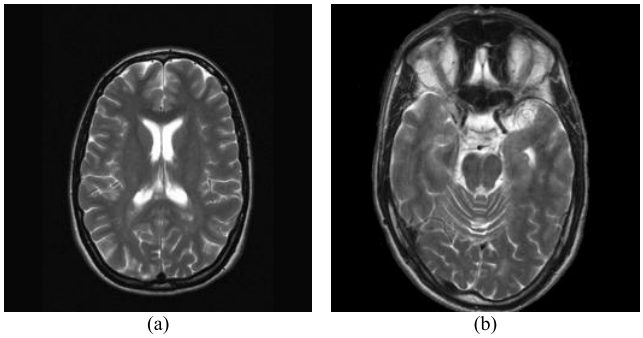
Method	No. of PCs	Ratio
PCA	424	0.58%
PKPCA	405	0.55%
RKPCA	396	0.54%

TABLE 4. Sensitivities over 10 runs.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Sum
R1	95.52	97.06	100.00	92.54	100.00	98.53	95.59	97.01	97.06	97.06	97.04
R2	100.00	94.12	95.59	94.03	100.00	98.51	100.00	92.65	98.53	95.59	97.04
R3	100.00	98.53	94.03	97.01	100.00	98.51	95.59	97.06	98.53	94.12	97.34
R4	98.53	94.12	100.00	100.00	98.51	100.00	97.06	92.54	97.06	94.03	97.19
R5	95.59	91.04	98.53	97.06	98.53	97.06	98.53	100.00	100.00	97.01	97.34
R6	97.06	97.06	100.00	89.55	100.00	100.00	94.12	97.06	97.01	98.51	97.04
R7	100.00	97.01	94.12	98.53	98.53	91.04	94.12	100.00	100.00	97.01	97.04
R8	94.03	100.00	98.51	97.01	97.06	97.06	100.00	94.12	98.53	94.12	97.04
R9	94.03	100.00	95.59	95.59	100.00	97.06	95.59	97.01	97.06	97.01	96.89
R10	95.59	100.00	100.00	94.03	94.12	100.00	100.00	94.12	98.51	94.12	97.19
Ave.											97.12 ±0.14

TABLE 5. Specificities over 10 runs.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Sum
R1	100.00	100.00	100.00	97.73	95.45	95.45	96.59	97.73	100.00	98.86	98.41
R2	98.86	96.59	95.45	98.86	96.59	100.00	100.00	100.00	97.73	97.73	98.30
R3	100.00	97.73	100.00	97.73	100.00	98.86	98.86	97.73	93.18	97.73	98.18
R4	96.59	98.86	100.00	96.59	100.00	100.00	97.73	97.73	97.73	98.86	98.52
R5	95.45	100.00	96.59	100.00	95.45	100.00	96.59	98.86	97.73	98.86	98.07
R6	97.73	100.00	100.00	97.73	95.45	97.73	96.59	100.00	96.59	100.00	98.18
R7	96.59	96.59	100.00	100.00	96.59	100.00	98.86	96.59	96.59	100.00	98.41
R8	98.86	100.00	100.00	94.32	96.59	100.00	100.00	100.00	97.73	95.45	98.52
R9	100.00	98.86	96.59	98.86	100.00	96.59	100.00	97.73	95.45	100.00	98.52
R10	98.86	98.86	98.86	96.59	96.59	100.00	98.86	96.59	97.73	100.00	98.30
Ave.											98.25 ±0.16



(a) (b)

FIGURE 7. Two samples for biorthogonal decomposition. (a) Sample I. (b) Sample II.

IV. RESULTS AND DISCUSSIONS

Our experiment was performed on Dell laptop with 3.20 GHz i5-3470 CPU and 4GB RAM. Programs were developed in-house and ran on Windows 10 Operating System.

A. BIORTHOGONAL DECOMPOSITION

Figure 7 presents two samples. We performed three-level bior 4.4 decomposition over these two sample images. The decompositions results are offered in

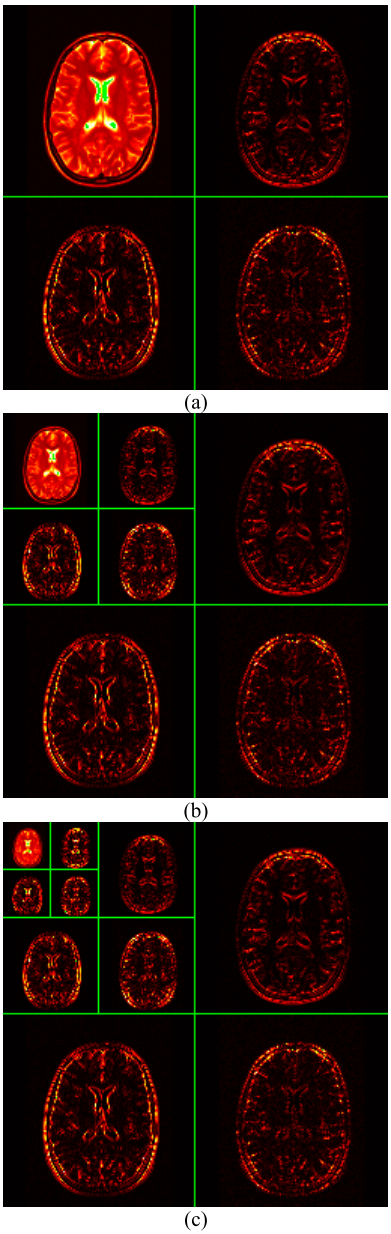


FIGURE 8. BWT Decomposition of Sample I (Hot color map was added). (a) 1-level. (b) 2-level. (c) 3-level.

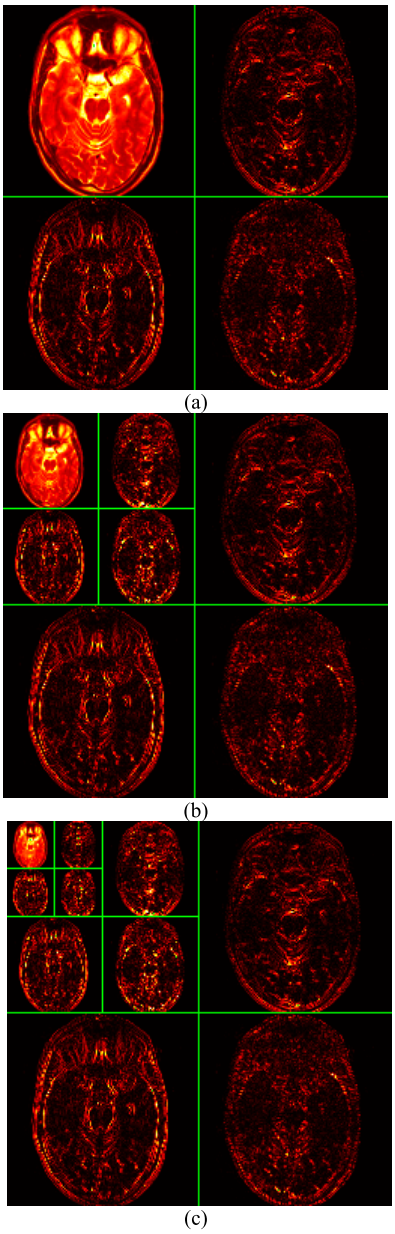


FIGURE 9. BWT Decomposition of Sample II (Hot color map was added). (a) 1-level. (b) 2-level. (c) 3-level.

Figure 8 and Figure 9, respectively. For better view of coefficients, we use hot pseudo color map.

Some literature combined entropy with discrete wavelet transform [63], [64]. In this condition, the entropy operation can be regarded as a means to reduce features. Nevertheless, we already used KPCA in this study; thus it is unnecessary for us to perform entropy operations.

B. PCA VERSUS KPCA

In this section, we compared PCA with KPCA. The wavelet coefficients of each brain image were realigned as a row vector with length of 73056. This value is a bit more than 256^2= 65536 due to the border and down sampling. Afterwards, the 1556 images will form a matrix with size of 1556×73056.

Using three dimensionality reduction methods (PCA, PKPCA, RKPCA) and setting the threshold as 95%, we plot the cumulated explained variances versus selected PCs in Figure 10. Here we know that PCA selects 424 PCs, PKPCA selects 405 PCs, and RKPCA selects 396 PCs. Dividing them by the total coefficients, we know that PCA selects 0.58% of total coefficients, PKPCA selects 0.55%, and RKPCA selects 0.54%, which are listed in Table 3. Therefore, we find that RKPCA selects the least number of PCs while attaining the same threshold.

C. STATISTICAL ANALYSIS

The sensitivities, specificities, and accuracies over all 10 runs are listed below in Table 4, Table 5, and Table 6, respectively.

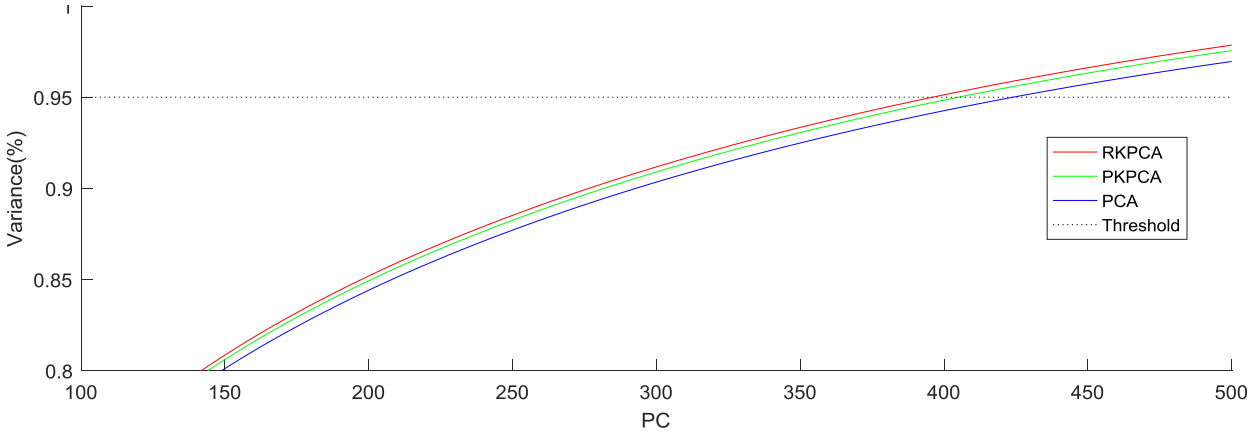


FIGURE 10. Selected PCs by different algorithms.

TABLE 6. Accuracies over 10 runs.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Sum
R1	98.06	98.72	100.00	95.48	97.42	96.79	96.15	97.42	98.72	98.08	97.62
R2	99.36	95.51	95.51	96.77	98.06	99.35	100.00	96.79	98.08	96.79	97.69
R3	100.00	98.08	97.42	97.42	100.00	98.71	97.44	97.44	95.51	96.15	97.75
R4	97.44	96.79	100.00	98.06	99.35	100.00	97.44	95.48	97.44	96.77	97.94
R5	95.51	96.13	97.44	98.72	96.79	98.72	97.44	99.35	98.71	98.06	97.75
R6	97.44	98.72	100.00	94.19	97.42	98.72	95.51	98.72	96.77	99.35	97.62
R7	98.06	96.77	97.44	99.36	97.44	96.13	96.79	98.08	98.08	98.71	97.81
R8	96.77	100.00	99.35	95.48	96.79	98.72	100.00	97.44	98.08	94.87	97.88
R9	97.42	99.36	96.15	97.44	100.00	96.79	98.08	97.42	96.15	98.71	97.75
R10	97.44	99.35	99.36	95.48	95.51	100.00	99.35	95.51	98.06	97.44	97.75
Ave.											97.76 ±0.10

Here F means fold, R means run. We can observe that our proposed method yields a sensitivity of 97.12±0.14, a specificity of 98.25±0.16, and an accuracy of 97.76±0.10.

D. COMPARISON TO STATE-OF-THE-ART METHODS

We submitted a 1556x396 matrix to the classifier of LR. Here 1556 is the number of total image, 396 is the number of PCs. Two Matlab commands "mnrfit" and "mnrval" were used to accelerate the program developing.

We compared our BWT + RKPCA + LR method with five state-of-the-art approaches: MAMFM + SVM [18], WE + HBP [19], DWT + PCA + LS-SVM [20], DWT + PPCA + RF [21], and SWE + DT [22]. Table 7 presents the comparison results. Note that our method runs 10 times, so we also report the standard deviation. The unit of data in Table 7 is percentage.

TABLE 7. Algorithm comparison.

Method	Sensitivity	Specificity	Accuracy
MAMFM + SVM [18]	94.08	93.64	93.83
WE + HBP [19]	96.15	97.16	96.72
DWT + PCA + LS-SVM [20]	95.86	96.48	96.21
DWT + PPCA + RF [21]	96.01	96.70	96.40
SWE + DT [22]	96.75	**98.30**	97.62
BWT + RKPCA + LR(Our)	**97.12±0.14**	98.25±0.16	**97.76±0.10**

(Bold means the best)

From the data in Table 7, we see that our BWT + RKPCA + LR method achieves the highest sensitivity and accuracy of all six algorithms. For the specificity, our method achieves an average value of 98.25%, slightly lower than the SWE + DT [22] method of 98.30%. It is worthy to note that sensitivity is more important than specificity, since detecting MS can provide early treatment. We can conclude that our method is superior to other five state-of-the-art approaches.

V. CONCLUSION

In this study, our team presents a novel MS detection method on the basis of BWT, RKPCA, and LR. The experiments results showed that this BWT + RKPCA + LR method was superior to five state-of-the-art methods.

In the future, we shall apply our method to brain CT [65], retinal image [66], low-dose X-ray [67], PET, and SPECT. The structure extraction [68] method will also be tested.

CONFLICT OF INTEREST

We have no conflict of interest to disclose, with regard to the subject matter of this paper.

REFERENCES

[1] D. Alali, K. Ballard, and H. Bogaardt, "Treatment effects for dysphagia in adults with multiple sclerosis: A systematic review," *Dysphagia*, vol. 31, pp. 610–618, Oct. 2016.

[2] L. S. Tejedor, T. Skripuletz, M. Stangel, and V. Gudi, "Mesenchymal stem cells require the peripheral immune system for immunomodulating effects in animal models of multiple sclerosis," *Neural Regenerat. Res.*, vol. 11, pp. 90–91, Jan. 2016.

[3] T. D. Faizy *et al.*, "Heterogeneity of multiple sclerosis lesions in multislice myelin water imaging," *PLoS ONE*, vol. 11, Mar. 2016, Art. no. e0151496.
[4] A. P. Kallaur *et al.*, "Immune-inflammatory and oxidative and nitrosative stress biomarkers of depression symptoms in subjects with multiple sclerosis: Increased peripheral inflammation but less acute neuroinflammation," *Molecular Neurobiol.*, vol. 53, no. 8, pp. 5191–5202, Oct. 2016.
[5] A. Lúcio *et al.*, "Pelvic floor muscle training with and without electrical stimulation in the treatment of lower urinary tract symptoms in women with multiple sclerosis," *J. Wound Ostomy Continence Nurs.*, vol. 43, pp. 414–419, Jul./Aug. 2016.
[6] N. Aghaei, S. Karbandi, M. A. H. Gorji, M. B. Golkhatmi, and B. Alizadeh, "Social support in relation to fatigue symptoms among patients with multiple sclerosis," *Indian J. Palliative Care*, vol. 22, pp. 163–167, Apr./Jun. 2016.
[7] A. Manca *et al.*, "Effect of contralateral strength training on muscle weakness in people with multiple sclerosis: Proof-of-concept case series," *Phys. Therapy*, vol. 96, pp. 828–838, Jun. 2016.
[8] A. Beer *et al.*, "Tissue damage within normal appearing white matter in early multiple sclerosis: Assessment by the ratio of T1-and T2-weighted MR image intensity," *J. Neurol.*, vol. 263, pp. 1495–1502, Aug. 2016.
[9] I. de Kouchkovsky, E. Fieremans, L. Fleysher, J. Herbert, R. I. Grossman, and M. Inglese, "Quantification of normal-appearing white matter tract integrity in multiple sclerosis: A diffusion kurtosis imaging study," *J. Neurol.*, vol. 263, pp. 1146–1155, Jun. 2016.
[10] Y. Zhang and L. Wu, "Classification of fruits using computer vision and a multiclass support vector machine," *Sensors*, vol. 12, no. 9, pp. 12489–12505, 2012.
[11] C. Nguyen, W. C. Feng, and F. Liu, "Hotspot: Making computer vision more effective for human video surveillance," *Inf. Visualizat.*, vol. 15, pp. 273–285, Oct. 2016.
[12] S. Balochian, "Artificial intelligence and its applications," *Mathematical Problems in Engineering*, vol. 2014, Apr. 2014, Art. no. 840491, doi: 10.1155/2014/840491.
[13] J. G. Liu and H. Y. Wang, "Research on the computer vision imaging techniques based on artificial intelligence," in *Proc. 3rd Int. Conf. Adv. Soc. Sci., Humanities, Manage.*, Guangzhou, China, 2015, pp. 1180–1188.
[14] Y. Zhang and S. Wang, "Detection of Alzheimer's disease by displacement field and machine learning," *PeerJ*, vol. 3, Sep. 2015, Art. no. e1251.
[15] N. V. K. Medathati, H. Neumann, G. S. Masson, and P. Kornprobst, "Bio-inspired computer vision: Towards a synergistic approach of artificial and biological vision," *Comput. Vis. Image Understand.*, vol. 150, pp. 1–30, Sep. 2016.
[16] R. J. Friedman *et al.*, "The diagnostic performance of expert dermoscopists vs a computer-vision system on small-diameter melanomas," *Arch. Dermatol.*, vol. 144, pp. 476–482, Apr. 2008.
[17] Y. Zhang, P. Agarwal, V. Bhatnagar, S. Balochian, and X. Zhang, "Swarm intelligence and its applications 2014," *Sci. World J.*, vol. 2014, Jun. 2014, Art. no. 204294.
[18] V. Murray, P. Rodriguez, and M. S. Pattichis, "Multiscale AM-FM demodulation and image reconstruction methods with improved accuracy," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1138–1152, May 2010.
[19] Y. Zhang, S. Wang, Z. Dong, P. Phillip, G. Ji, and J. Yang, "Pathological brain detection in magnetic resonance imaging scanning by wavelet entropy and hybridization of biogeography-based optimization and particle swarm optimization," *Prog. Electromagn. Res.*, vol. 152, pp. 41–58, Mar. 2015.
[20] M. F. Siddiqui, A. W. Reza, and J. Kanesan, "An automated and intelligent medical decision support system for brain MRI scans classification," *PLoS ONE*, vol. 10, Aug. 2015, Art. no. e0135875.
[21] D. R. Nayak, R. Dash, and B. Majhi, "Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests," *Neurocomputing*, vol. 177, pp. 188–197, Feb. 2016.
[22] Y. Zhang, "Comparison of machine learning methods for stationary wavelet entropy-based multiple sclerosis detection: Decision tree, *k*-nearest neighbors, and support vector machine," *Simulation*, vol. 92, pp. 861–871, Sep. 2016.
[23] S. S. Negi and Y. S. Bhandari, "A hybrid approach to image enhancement using contrast stretching on image sharpening and the analysis of various cases arising using histogram," in *Proc. Recent Adv. Innov. Eng. (ICRAIE)*, Jaipur, India, 2014, pp. 1–6.
[24] D. Barina, P. Musil, M. Musil, and P. Zemcik, "Single-loop approach to 2-D wavelet lifting with JPEG 2000 compatibility," in *Proc. Int. Symp. Comput. Archit. High Perform. Comput. Workshop (SBAC-PADW)*, Florianópolis, Brazil, 2015, pp. 31–36.
[25] T. Çevik, A. M. A. Alshaykha, and N. Çevik, "Performance analysis of GLCM-based classification on wavelet transform-compressed fingerprint images," in *Proc. 6th Int. Conf. Digit. Inf. Commun. Technol. Appl.*, Konya, Turkey, 2016, pp. 131–135.
[26] D. R. Nayak, R. Dash, and B. Majhi, "Classification of brain MR images using discrete wavelet transform and random forests," in *Proc. 5th Nat. Conf. Comput. Vis., Pattern Recognit., Image Process. Graph.*, Patna, India, 2015, pp. 1–4.
[27] S. Wang *et al.*, "Wavelet entropy and directed acyclic graph support vector machine for detection of patients with unilateral hearing loss in MRI scanning," *Frontiers Comput. Neurosci.*, vol. 10, Oct. 2016, Art. no. 160.
[28] Y.-D. Zhang *et al.*, "Pathological brain detection in MRI scanning by wavelet packet Tsallis entropy and fuzzy support vector machine," *SpringerPlus*, vol. 4, Nov. 2015, Art. no. 716.
[29] S. Kadu, C. Naveen, V. R. Satpute, and A. G. Keskar, "Discrete wavelet transform based video watermarking technique," in *Proc. Int. Conf. Microelectron., Comput. Commun. (MicroCom)*, Durgapur, India, 2016, pp. 6–11.
[30] Y. Zhang, "Preliminary research on abnormal brain detection by wavelet-energy and quantum- behaved PSO," *Technol. Health Care*, vol. 24, pp. S641–S649, Apr. 2016.
[31] A. Chaiwachiragompol and N. Suwannata, "The features extraction of infants cries by using discrete wavelet transform techniques," in *Proc. Int. Elect. Eng. Congr.*, Amsterdam, The Netherlands, 2016, pp. 285–288.
[32] S. Wang *et al.*, "Detection of dendritic spines using wavelet-based conditional symmetric analysis and regularized morphological shared-weight neural networks," *Comput. Math. Methods Med.*, vol. 2015, Sep. 2015, Art. no. 454076, doi: 10.1155/2015/454076.
[33] R. Thanki and K. Borisagar, "Multibiometric template generation using CS theory and discrete wavelet transform based fusion technique," in *Proc. 5th Nirma Univ. Int. Conf. Eng.*, Ahmedabad, India, 2015, pp. 143–149.
[34] S. Wang, X. Yang, Y. Zhang, P. Phillips, J. Yang, and T.-F. Yuan, "Identification of green, Oolong and black teas in China via wavelet packet entropy and fuzzy support vector machine," *Entropy*, vol. 17, no. 10, pp. 6663–6682, 2015.
[35] M. Schiurunack, S. Nguyen, and P. Mercorelli, "Anatomy of Haar Wavelet Filter and Its Implementation for Signal Processing," *IFAC Papers Online*, vol. 49, no. 6, pp. 99–104, 2016.
[36] A. K. Abbas, R. Bassam, and R. M. Kasim, "Mitral regurgitation PCG-signal classification based on adaptive Db-wavelet," in *Proc. 4th Kuala Lumpur Int. Conf. Biomed. Eng.*, Kuala Lumpur, Malaysia, 2008, pp. 212–216.
[37] P. M. K. Prasad, D. Y. V. Prasad, and G. S. Rao, "Performance analysis of orthogonal and biorthogonal wavelets for edge detection of X-ray images," in *Proc. 4th Int. Conf. Recent Trends Comput. Sci. Eng.*, Amsterdam, The Netherlands, 2016, pp. 116–121.
[38] Y. Zhang, Z. Dong, S. Wang, G. Ji, and J. Yang, "Preclinical diagnosis of magnetic resonance (MR) brain images via discrete wavelet packet transform with Tsallis entropy and generalized eigenvalue proximal support vector machine (GEPSVM)," *Entropy*, vol. 17, no. 4, pp. 1795–1813, 2015.
[39] M. García, J. Ródenas, R. Alcaraz, and J. J. Rieta, "Application of the relative wavelet energy to heart rate independent detection of atrial fibrillation," *Comput. Methods Programs Biomed.*, vol. 131, pp. 157–168, Jul. 2016.
[40] G. Yang *et al.*, "Automated classification of brain images using wavelet-energy and biogeography-based optimization," *Multimedia Tools Appl.*, pp. 1–17, May 2015, doi: 10.1007/s11042-015-2649-7.
[41] J. D. McEwen, P. Vandergheynst, and Y. Wiaux, "On the computation of directional scale-discretized wavelet transforms on the sphere," *Proc. SPIE*, vol. 8858, Sep. 2013, Art. no. 88580I.
[42] Y. Zhang, "Magnetic resonance brain image classification via stationary wavelet transform and generalized eigenvalue proximal support vector machine," *J. Med. Imag. Health Informat.*, vol. 5, pp. 1395–1403, Nov. 2015.
[43] M. Zimbres, R. A. Batista, and E. Kemp, "Using spherical wavelets to search for magnetically-induced alignment in the arrival directions of ultra-high energy cosmic rays," *Astroparticle Phys.*, vol. 54, pp. 54–60, Feb. 2014.
[44] Z. Dong, P. Phillips, G. Ji, and J. Yang, "Exponential wavelet iterative shrinkage thresholding algorithm for compressed sensing magnetic resonance imaging," *Inf. Sci.*, vol. 322, pp. 115–132, Nov. 2015.
[45] S. Wang, S. Lu, Z. Dong, J. Yang, M. Yang, and Y. Zhang, "Dual-tree complex wavelet transform and twin support vector machine for pathological brain detection," *Appl. Sciences*, vol. 6, no. 6, 2016, Art. no. 169.

[46] Y. Zhang and L. Wu, "An MR brain images classifier via principal component analysis and kernel support vector machine," *Prog. Electromagn. Res.*, vol. 130, pp. 369–388, Sep. 2012.

[47] Y. Ikemiya, K. Itoyama, and K. Yoshii, "Singing voice separation and vocal F0 estimation based on mutual combination of robust principal component analysis and subharmonic summation," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 24, no. 11, pp. 2084–2095, Nov. 2016.

[48] M. R. Mowla, S. C. Ng, M. S. A. Zilany, and R. Paramesran, "Single-trial evoked potential estimation using iterative principal component analysis," *IEEE Sensors J.*, vol. 16, no. 18, pp. 6955–6960, Sep. 2016.

[49] D. Feng, M. Xiao, Y. Liu, H. Song, Z. Yang, and L. Zhang, "A kernel principal component analysis–based degradation model and remaining useful life estimation for the turbofan engine," *Adv. Mech. Eng.*, vol. 8, no. 5, May 2016, Art. no. 1687814016650169.

[50] O. Taouali, I. Jaffel, H. Lahdhiri, M. F. Harkat, and H. Messaoud, "New fault detection method based on reduced kernel principal component analysis (RKPCA)," *Int. J. Adv. Manuf. Technol.*, vol. 85, pp. 1547–1552, Jul. 2016.

[51] A. A. Joseph, T. Tokumoto, and S. Ozawa, "Online feature extraction based on accelerated kernel principal component analysis for data stream," *Evol. Syst.*, vol. 7, pp. 15–27, Mar. 2016.

[52] C. Thrane, "Norwegian students' package trip propensity in 2007 and 2014—A logistic regression analysis," *Tourism Econ.*, vol. 22, pp. 1141–1150, Oct. 2016.

[53] F. M. Ali *et al.*, "Ordinal logistic regression and Monte Carlo simulation in the mapping of DLQI scores to EQ-5D utility values," *J. Invest. Dermatol.*, vol. 136, pp. S163–S163, Sep. 2016.

[54] N. Kyurkchiev and S. Markov, "On the Hausdorff distance between the Heaviside step function and Verhulst logistic function," *J. Math. Chem.*, vol. 54, pp. 109–119, Jan. 2016.

[55] Y. Zhang, P. Phillips, S. Wang, G. Ji, J. Yang, and J. Wu, "Fruit classification by biogeography-based optimization and feedforward neural network," *Expert Syst.*, vol. 33, no. 3, pp. 239–253, 2016.

[56] F. Erlandsson, P. Bródka, A. Borg, and H. Johnson, "Finding influential users in social media using association rule learning," *Entropy*, vol. 18, May 2016, Art. no. 164.

[57] Y. Zhang, S. Wang, P. Phillips, and G. Ji, "Binary PSO with mutation operator for feature selection using decision tree applied to spam detection," *Knowl.-Based Syst.*, vol. 64, pp. 22–31, Jul. 2014.

[58] J. J. Dabrowski, C. Beyers, and J. P. de Villiers, "Systemic banking crisis early warning systems using dynamic Bayesian networks," *Expert Syst. Appl.*, vol. 62, pp. 225–242, Nov. 2016.

[59] Y.-D. Zhang, S. Chen, S.-H. Wang, J.-F. Yang, and P. Phillips, "Magnetic resonance brain image classification based on weighted-type fractional Fourier transform and nonparallel support vector machine," *Int. J. Imag. Syst. Technol.*, vol. 25, pp. 317–327, Nov. 2015.

[60] B. Tegelund, H. Son, and D. Lee, "A task-oriented service personalization scheme for smart environments using reinforcement learning," in *Proc. Int. Conf. Pervas. Comput. Commun. Workshops*, Sydney, NSW, Australia, 2016, pp. 6–11.

[61] S. Wang, "Morphological analysis of dendrites and spines by hybridization of ridge detection with twin support vector machine," *PeerJ*, vol. 4, Jul. 2016, Art. no. e2207.

[62] S. Lu and Z. Lu, "A pathological brain detection system based on kernel based ELM," *Multimedia Tools Appl.*, pp. 1–14, May 2016, doi: 10.1007/s11042-016-3559-z.

[63] D. R. Nayak, R. Dash, B. Majhi, and J. Mohammed, "Non-linear cellular automata based edge detector for optical character images," *Simulation*, vol. 92, pp. 849–859, Aug. 2016.

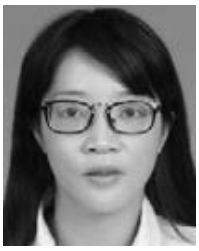
[64] P. Sun, "Pathological brain detection based on wavelet entropy and Hu moment invariants," *Bio-Medical Mater. Eng.*, vol. 26, pp. 1283–1290, 2015.

[65] Y. Chen *et al.*, "CT metal artifact reduction method based on improved image segmentation and sinogram in-painting," *Math. Problems Eng.*, vol. 2012, Jun. 2012, Art. no. 786281.

[66] Y. Zhang *et al.*, "Image processing methods to elucidate spatial characteristics of retinal microglia after optic nerve transection," *Sci. Rep.*, vol. 6, Feb. 2016, Art. no. 21816.

[67] Y. Chen *et al.*, "Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior," *Comput. Med. Imag. Graph.*, vol. 33, pp. 495–500, Oct. 2009.

[68] Y. Chen *et al.*, "Curve-like structure extraction using minimal path propagation with backtracking," *IEEE Trans. Image Process.*, vol. 25, no. 2, pp. 988–1003, Feb. 2016.



SHUI-HUA WANG received the B.S. degree from Southeast University in 2008, the M.S. degree from The City University of New York in 2012, and the Ph.D. degree from Nanjing University in 2016. She is currently an Assistant Professor with Nanjing Normal University.

TIAN-MING ZHAN received the Ph.D. degree in pattern recognition and intelligence system from the Nanjing University of Science and Technology, Nanjing, China, in 2013. He is currently an Associate Professor with the school of technology, Nanjing Audit University. His research interests include in medical imaging processing and analysis.

YI CHEN received the Ph.D. degree from the School of Computer Science and Technology, Nanjing University of Science and Technology, in 2012. His current research interests include pattern recognition, compute vision, and image processing.

YIN ZHANG was a Post-Doctoral Fellow with the School of Computer Science and Technology, Huazhong University of Science and Technology, China. He is currently an Assistant Professor with the School of Information and Safety Engineering, Zhongnan University of Economics and Law, China. He has authored over 30 prestigious conference and journal papers. His research interests include data analysis, data mining, healthcare big data, and social network. He serves as a Guest Editor of the IEEE Sensors Journal and the *N*ew Review of Hypermedia and Multimedia. He also served as the TPC Co-Chair of the CloudComp 2015 and the Local Chair of the TRIDENTCOM 2014. He is the Vice-Chair of the IEEE Computer Society Big Data STC.

MING YANG received the M.D. degree from Southeast University in 2011. She is currently a Radiologist with Nanjing Children's Hospital. Her research interest is fMRI.

HUIMIN LU received the B.S. degree in electronics information science and technology from Yangzhou University in 2008, the M.S. degree in electrical engineering from the Kyushu Institute of Technology, the M.S. degree in electrical engineering from Yangzhou University in 2011, and the Ph.D. degree in electrical engineering from the Kyushu Institute of Technology in 2014. He is currently an Associate Professor with the Kyushu Institute of Technology and also serves as an Excellent Young Researcher of MEXT-Japan. His current research interests include computer vision, robotics, artificial intelligence, and ocean observing.

HAINAN WANG received the B.S. degree from Yancheng Normal University in 2014. She is currently pursuing the master's degree with the School of Computer Science and Technology, Nanjing Normal University. She is also with the Key Laboratory of Symbolic Computation and Knowledge Engineering of the Ministry of Education, Jilin University. Her research interest is image processing and machine learning.

BIN LIU received the M.D. degree and the master's degree of Medical Imaging from Southeast University, Nanjing, China, in 1990 and 2004, respectively. His research interests focus on Functional MRI.

PREETHA PHILLIPS received the B.Sc. degree from Shepherd University in 2016. She is currently pursuing the Doctor of Osteopathic Medicine degree with the West Virginia School of Osteopathic Medicine. She has conducted research in Shepherd University in investigating the effects of Roundup on the sex hormones in the aquatic snail, Lymnaea palustris. She received an honorable mention for a poster presentation for this research in the West Virginia Academy of Science that took place in Marshall University in 2016. She has also conducted research on using magnetic resonance spectroscopy in diagnosing panic disorder Columbia Presbyterian Medical Center. She has also helped conduct research analyzing pain mechanisms in rats in Stony Brook University.