

Article

Identification of Green, Oolong and Black Teas in China via Wavelet Packet Entropy and Fuzzy Support Vector Machine

Shuihua Wang^{1,4,†}, Xiaojun Yang^{2,†}, Yudong Zhang^{1,4,†,*}, Preetha Phillips³, Jianfei Yang⁴ and Ti-Fei Yuan^{5,*}

¹ School of Computer Science and Technology, Nanjing Normal University, 210023 Nanjing, China; E-Mail: wangshuihua@njnu.edu.cn

² Department of Mathematics and Mechanics, China University of Mining and Technology, 221008 Xuzhou, China; E-Mail: dyangxiaojun@163.com

³ School of Natural Sciences and Mathematics, Shepherd University, Shepherdstown, 25443 West Virginia, WV, USA; E-Mail: preetha.phillips@gmail.com

⁴ Jiangsu Key Laboratory of 3D Printing Equipment and Manufacturing, 210042 Nanjing, China; E-Mail: jianfei_yang_3d@163.com

⁵ School of Psychology, Nanjing Normal University, 210008 Nanjing, China

† Those authors contributed equally to this paper.

* Author to whom correspondence should be addressed;

E-Mails: zhangyudong@njnu.edu.cn (Y.Z.D.); ytf0707@126.com (T.-F.Y.).

Academic Editor: Carlo Cattani

Received: 5 August 2015 / Accepted: 21 September 2015 / Published: 25 September 2015

Abstract: To develop an automatic tea-category identification system with a high recall rate, we proposed a computer-vision and machine-learning based system, which did not require expensive signal acquiring devices and time-consuming procedures. We captured 300 tea images using a 3-CCD digital camera, and then extracted 64 color histogram features and 16 wavelet packet entropy (WPE) features to obtain color information and texture information, respectively. Principal component analysis was used to reduce features, which were fed into a fuzzy support vector machine (FSVM). Winner-take-all (WTA) was introduced to help the classifier deal with this 3-class problem. The 10 × 10-fold stratified cross-validation results show that the proposed FSVM + WTA method yields an overall recall rate of 97.77%, higher than 5 existing methods. In addition, the number of reduced features is only five, less than or equal to existing methods. The proposed method is effective for tea identification.

Keyword: tea identification; wavelet packet entropy; Shannon entropy; wavelet analysis; support vector machine (SVM); fuzzy SVM; information theory

1. Introduction

Tea is a pleasant beverage commonly brewed by pouring (near) boiling water over curled leaves of the *Camellia sinensis*, which is an evergreen bush indigenous to Asia. Tea is an extensively consumed beverage worldwide with an expanding market [1].

Tea originated in China having been recognized for its healing properties, and in recent times, there has been more and more evidence to show that specific substances in tea can help resist diseases. For example, researchers found antioxidants in tea may be capable of protecting against diseases: Alzheimer's disease [2], Parkinson's disease [3], neurodegenerative disease [4], blood pressure and cardiovascular disease [5], colon cancer [6], breast cancer [7], lung cancer [8], *etc.*

There are many different species of tea in the world. At least six different types have been produced and are listed in Table 1. This study focuses on tea in China, where green, Oolong, and black tea are the most popular classes [9].

Table 1. Tea Categories.

Category	Characteristics
White Tea	wilted and unoxidized
Yellow tea	unwilted and unoxidized, with sweltering
Green tea	unwilted and unoxidized
Oolong tea	wilted, bruised, and partially oxidized
Black tea	wilted and fully oxidized
Post-fermented tea	fermented green tea

The classification of tea categories is very important for controlling fermentation time and obtaining specific types of tea, which strongly influences the commercial market. In the past, human experts measured the color variation during fermentation by visual checks; however, this method has various shortcomings such as irreproducibility, high labor cost, and inconsistency. The colorimeter was then introduced for objective quality evaluations. In addition to examining color features, sieves of different pore sizes were used. However, those traditional methods are rather rough; hence, they cannot accurately assess, predict, and control tea quality.

Two categories of automatic tea-classification methods were proposed: (1) to develop new measurement devices and (2) to develop new algorithms based on computer vision. For the former category, Herrador and Gonzalez [10] selected eight metals (Al, Zn, Ca, Mn, Ba, Mg, Cu, and K) as chemical descriptors to differentiate three tea classes. They proved the back propagation-artificial neural network (BP-ANN) achieved an almost 95% recall rate. Zhao *et al.* [11] utilized near-infrared (NIR) spectroscopy for fast identification of green, oolong, and black tea. They extracted five principal components (PCs), which were sent to SVM classifiers. For all categories, their identification accuracies were no less than 90%. Chen *et al.* [12] proposed using NIR reflectance spectroscopy to identify three types of tea. They used RBF-SVM as the classifier, and found the best identification for green, black,

and oolong teas were 90%, 100%, and 95%, respectively. Wu *et al.* [13] put forward a new nondestructive method based on multispectral digital image texture feature. The sample images were obtained via a red waveband, NIR waveband and green waveband multispectral digital imager. They combined DCT and LS-SVM as the classifier. Chen *et al.* [14] developed a portable electronic nose, by an odor imaging sensor array, with the aim of tea classification of three different fermentation degrees. Liu *et al.* [15] used electronic tongue technique to analyze 43 samples of green and black tea. A class of metallic oxide-modified nickel foam electrodes (SnO_2 , ZnO , TiO_2 , Bi_2O_3) was compared. The signals obtained by cyclic voltammetry underwent multivariate data analysis that consisted of principal component analysis (PCA) and SVM.

The above methods can obtain good identification results; however, they need expensive signal acquiring devices with time-consuming procedures. On the other hand, the latter category (*i.e.*, computer-vision based systems) is attaining support in the food industry due to its rapid speed, low cost, consistency, reliability, and high accuracy. For instance, Borah *et al.* [16] described a novel texture-feature estimation technique, with the goal of discriminating images of eight dissimilar grades of CTC tea. Chen *et al.* [17] used 12 color feature variables and 12 texture feature variables. PCA and linear discriminant analysis (LDA) were employed to build the identification model for five varieties of Chinese green tea. Features were reduced to 11. Jian *et al.* [18] used computer-vision techniques in order to classify and grade a particular tea sample, based on the parameters of color and shape. Two kinds of features were obtained: (1) The color features were extracted by transforming the color tea image to HSI model. (2) The shape features were extracted after the color images were degraded to binary images. Genetic neural-network (GNN) was used as the classification method. It achieved acceptable identification performance with eight features of both shape and color. Gill *et al.* [19] presented a survey of versatile computer vision techniques, which are related to color and texture analysis, with a tendency towards tea grading and monitoring. They pointed out that computer vision and image analyses were harmless for sorting tea. They predicted computer vision based techniques will become more and more popular in tea classification. Laddi *et al.* [20] acquired the images of tea granules using 3-CCD color camera under dual ring light. In all, 10 graded tea samples were obtained and analyzed. Their acquired features, *i.e.*, (energy, entropy, contrast, correlation, and homogeneity) were reduced by principal component analysis (PCA). Their experimental results showed that the best discrimination was dark-field illumination (variance = 96%). In contrast, the bright-field illumination showed weak discrimination (variance = 83%). Zhang *et al.* [21] used the combination of different features (color, shape, and texture) and a feedforward neural network (FNN) with the aim of classifying fruits. The method can be directly used for tea identification.

Those methods used cheap digital camera as the main system; however, their classification accuracy does not meet the standard of practical usage. In information theory, the Shannon entropy [22] is commonly based on compression in a quantization process, and this can be investigated by using the wavelet compression. Hence, wavelet entropy (WE) is proposed and used in many applications. From another point of view, WE is the minimization of the feature space for data analysis. To augment the performance, we proposed replacing wavelet decomposition with wavelet packet transform (WPT), and thus the WE is replaced with wavelet packet entropy (WPE).

Based on WPE and machine learning techniques, we proposed a novel tea-identification method, with the aim of developing an automatic identification system with better identification accuracy of green,

black, and oolong teas. To implement it, we employed the latest developments in signal processing. The proposed methods consist of three stages following common convention: (i) Feature extraction: we combined color features (obtained by color histogram) with texture features (obtained by WPE); (ii) Feature reduction: we employed PCA to reduce the feature dimensions; (iii) Classification: we introduced a fuzzy support vector machine (FSVM).

The remainder of the paper is organized in the following way: Section 2 describes the sample preparation, image-acquiring procedure, the proposed methodology, and the statistical setting. Experiments in Section 3 compare the proposed methods with state-of-the-art methods. Finally, Section 4 concludes the paper and outlines future research directions. For ease of reading, we explain the nomenclatures in Abbreviation (please refer to the end of this work).

2. Proposed Method

2.1. Tea Preparation

Three hundred samples consisting of three categories were prepared with origins from various provinces in China. Each tea category contained different brands to increase the generalization ability of the identification system. All tea samples were bought on stock within the period of four months. Table 2 shows the characteristics of tea samples used.

Table 2. Characteristics of tea samples.

Category	Number	Origins
Green tea	100	Henan; Guizhou; Jiangxi; Anhui; Zhejiang; Jiangsu
Black tea	100	Yunan; Hunan; Hubei; Fujian
Oolong tea	100	Fujian; Guangdong

2.2. Image Acquiring

A flowchart of the used computer vision system is illustrated in Figure 1, which consists of 5 basic parts: a digital camera, an illumination platform, a capture board (digitizer or frame grabber), computer software, and computer hardware [23]. The image acquiring procedures are as follows:

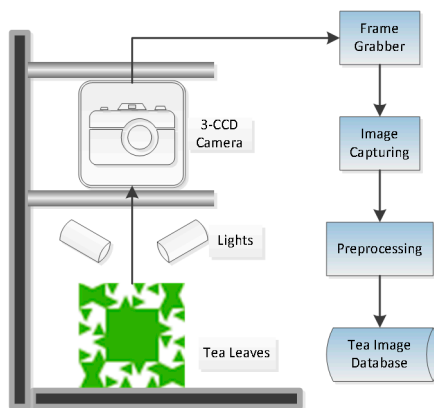


Figure 1. Computer vision based system to obtain the tea image database.

Tea images were grabbed after spreading tea leaves uniformly by a 3-CCD digital camera, which uses three separate charge-coupled devices (CCD). For each one, it takes a segregated mensuration of the primary colors: red (R), green (G), and blue (B) light, respectively [24]. The optical system then splits the light, which emits through the lens, by a prism assembly. The appropriate wavelength ranges of light are directed to the corresponding CCDs. In general, a 3-CCD camera provides better image quality through enhanced resolution and lower noise [20] than a 1-CCD camera.

The intensity and nature of illumination affects the performance of the computer vision system. The lighting placement is arranged as either back or front lighting [25]. The back lighting is used for producing a silhouette image to augment tea edges; meanwhile, front lighting is used to enhance tea surface features [19]. In this study, front lighting was used.

2.3. Feature Processing

Here, we presented a hybrid feature set that consisted of color and wavelet-based features. For a tea image with size of 256×256 , its total number of features is 65,536, because each pixel can be regarded as a feature. We obtained 64 color-based features and 16 wavelet packet entropy features. Then, principal component analysis (PCA) was utilized in order to decrease the number of total features so that the remained PCs should explain more than 99.9% variances of original 65,536-feature set. Figure 2 shows a diagram of the proposed feature processing method.

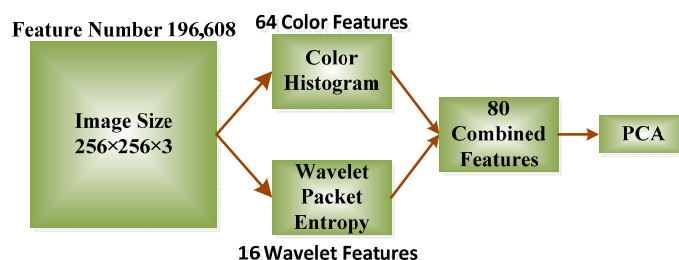


Figure 2. Flowchart of feature processing.

2.3.1. Color Histogram

The color histogram (CH) was harnessed over wide application areas, to count the distribution of colorations from a given tea image [26]. The CH counts the number of particular pixels, which have similar colors within a fixed range that cover the predefined color space. Usually, the CH is generated by two steps: (i) color discretization into $4 \times 4 \times 4 = 64$ bins (that means four bins for each channel), and (ii) counting the number of pixels in each bin [21].

The color histogram (CH) offers a concise summarization of the color distribution. Obviously, CH is relatively invariant with both rotation and translation about the inspecting axis. By extracting and comparing CH features of two images, the CH is found to be especially suitable for problems of object detection of unidentified position and unidentified rotation angles inside a scene. CH is used as an extremely important feature in many applications.

2.3.2. Discrete Wavelet Packet Transform

In the domain of signal processing, the discrete wavelet transform (DWT) is an outstanding tool that has various successful applications [27]. Further, the Discrete Wavelet packet transform (DWPT) is a powerful extension of DWT. The difference is that all nodes in the tree structure are allowed to split further at any decomposition level for a DWPT, which is forbidden in a conventional DWT [28].

In a detailed way, conventional DWT passes only the previous approximation subband to the next decomposition procedure [29,30]. Nevertheless, DWPT passes both the detail and approximation subbands to the next decomposition; hence, it can generate a full binary tree (See Figure 3).

From another point of view, DWPT features are provided on the basis of both detail and approximation subbands at various levels, so it yields more information than the conventional DWT does.

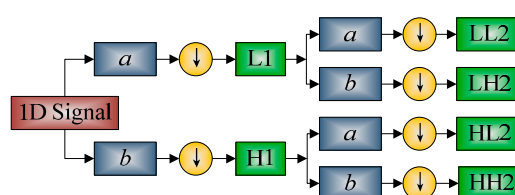


Figure 3. Diagram of two-level one-dimensional wavelet packet transform. Here, *a* and *b* represent the low-pass and high-pass filters, respectively. L and H represent the low-frequency and high-frequency subbands, respectively.

It is necessary to reconstruct DWPT coefficients in image domain using zero-padding for each decomposition subband, since the following entropy-based feature extraction technique is for the image domain not the wavelet coefficient domain. 2D-DWPT is implemented by applying 1D-DWPT along the *x*- and *y*-axis, respectively. 3D-DWPT is carried out in a similar way [31].

2.3.3. Shannon Entropy

In the past, entropy was used to measure the randomness of systems by statisticians. It is then generalized to measure uncertainty of the information content of a given system with the professional name of Shannon entropy (SE) [32]

$$S = -\sum_{n=1}^G h_n \log_2(h_n) \tag{1}$$

where *S* represents the Shannon entropy, *n* the greylevel of any subband, *h_n* the probability of *n*-th greylevel, and *G* the total number of greylevels [33].

In this study, entropies of both approximation and detail components at 2-level DWPT were computed, and termed as wavelet packet entropy (WPE). The pseudocode of calculating WPE is listed in Table 3. In the initial phase, each pixel was regarded as a feature, since pixels do provide some information. Hence, an image of 256 × 256 is considered containing 65,536 features. After feature extraction, we reduce the 65,536 features to only 16 WPE features (See Table 3).

Table 3. Pseudocode of calculating wavelet packet entropy (WPE) (Suppose 2 decomposition level).

Pseudocode of WPE	
Step A	Input Image. Read the 2D Image.
Step B	1D-DWPT. Pass the image through low-pass and high-pass filters and perform downsampling along <i>x</i> -axis and <i>y</i> -axis in sequence. Obtain four subbands.
Step C	2D-DWPT. For the four subbands obtained in 1D-DWPT, we continually implement 1D-DWPT to each subband, and finally obtained 16 subbands.
Step D	WPE. Extract Shannon entropy from the 16 subbands obtained by 2D-DWPT, and output the final feature vector of 16 elements.

Literatures showed that WPE is able to capture the wavelet-based features (including shape and texture) in an efficient way and has many successful applications in various fields [34–36].

2.3.4. Principal Component Analysis

In total, there are 80 features (64 color and 16 WPE features) extracted from a prescribed tea image. Those 80 features will increase computation resources and cost a mass of storage memory, which may aggravate the performance of the classifier. A valid strategy is to decrease the feature number by feature-reduction techniques [37].

Principal component analysis (PCA) is an effective utensil that is commonly used not only to reduce the interrelated variables, but also to retain the most substantial principal components (PC). The implementation of PCA is achieved by transforming the sample vectors to a new set of new variables sorted by the variance degree in a decreasing way.

Principal component analysis (PCA) has three benefits [38]: (i) It can orthogonalize the variables of the input vectors, in order to decorrelate each other. (ii) It sorts the resulting orthogonal variables, to guarantee the principal components with the largest variation come first and those with the smallest variation come last. (iii) It completely removes the variables from the dataset that impart the least variation.

Note that the input dataset is suggested to take a normalization as zero-mean and unity-variance before implementing a principal component analysis. Users can type the “PCA” command in Matlab platform to implement a canonical PCA procedure.

2.4. Classification

Support vector machine (SVM) is the most popular learning model that analyzes data and recognizes patterns for supervised classification [39]. However, SVM cannot deal with outliers and noises, *i.e.*, its performance will decrease sharply when the data set either contains outliers or was contaminated by noises. Fuzzy SVM (FSVM) was an effective variant with the advantages of reducing the effect from outliers and noises. There are some other advanced variants of SVM, such as generalized eigenvalue proximal SVM [40], twin SVM, least-square SVM

2.4.1. Support Vector Machine

Let us suppose there is an N -size training samples of z -dimensional vector, and suppose the goal is to create a hyperplane of $(z - 1)$ -dimension. Assume the dataset takes the form of [41]

$$\{(p_n, y_n) \mid p_n \in \mathbb{R}^z\}, n = 1, \dots, N \tag{2}$$

where p_n denotes a training point that is a z -dimensional vector, y_n is the realistic class of p_n taking the value of either +1 or -1, which corresponds to the class 2 or 1, respectively [42]. The hyperplane with maximum-margin that separates the two classes is the desired SVM. Considering that any hyperplane is in the form of

$$\mathbf{w}p - b = 0 \tag{3}$$

where \mathbf{w} represents the weights and b the bias. We need to select the optimal values of \mathbf{w} and b to maximize the distance between the two parallel hyperplanes to the full degree, while it can yet separate the data of the two classes.

$$\begin{aligned} \min_{\mathbf{w}, b} & \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{s.t. } & y_n (\mathbf{w}p_n - b) \geq 1, n = 1, \dots, N \end{aligned} \tag{4}$$

Positive slack vector $\xi = (\xi_1, \dots, \xi_n, \dots, \xi_N)$ are added to measure the misclassification degree of sample p_n . Hence, the mathematical formula of the optimal SVM can be deduced by solving:

$$\begin{aligned} \min_{\mathbf{w}, \xi, b} & \frac{1}{2} \|\mathbf{w}\|^2 + Le^T \xi \\ \text{s.t. } & \begin{cases} y_n (\mathbf{w}^T p_n - b) \geq 1 - \xi_n, n = 1, \dots, N \\ \xi_n \geq 0 \end{cases} \end{aligned} \tag{5}$$

where L represents the error penalty and e a vector of ones of N -dimension. Therefore, the optimization turns to a trade-off between a small error penalty and a large margin. The constraint optimization problem is solved using ‘‘Lagrange multiplier’’:

$$\min_{\mathbf{w}, \xi, b} \max_{\alpha, \beta} \left\{ \frac{1}{2} \|\mathbf{w}\|^2 + ce^T \xi - \sum_{n=1}^N \alpha_n [y_n (\mathbf{w}^T p_n - b) + \xi_n - 1] - \sum_{n=1}^N \xi_n \beta_n \right\} \tag{6}$$

The min-max problem is not easy to solve, so *dual form* technique is commonly proposed to solve it as

$$\begin{aligned} \max_{\alpha} & \sum_{n=1}^N \alpha_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_m \alpha_n y_m y_n p_m^T p_n \\ \text{s.t. } & \begin{cases} L \geq \alpha_n \geq 0 \\ \sum_{n=1}^N y_n \alpha_n = 0, n = 1, \dots, N \end{cases} \end{aligned} \tag{7}$$

The main merit of the dual form is that the slack variables ξ_n disappear, with only the constant L be an additional constraint on the Lagrange multipliers.

2.4.2. Fuzzy SVM

Fuzzy support vector machine (FSVM) is effective than simple SVM models especially in predicting or classifying real-world data, because several training samples are more substantial than others. It makes sense to require the meaningful training samples must be recognized perfectly meanwhile to neglect some meaningless points like noises or outliers [43].

Fuzzy support vector machine (FSVM) applies a fuzzy membership function (FMF) s to every training point [44], such that the training samples are transferred to fuzzy training samples, which can be expressed as

$$\{(p_n, y_n, s_n) \mid p_n \in \mathbb{R}^z, 0 < s_n \leq 1\}, n = 1, \dots, N \tag{8}$$

where s_n denotes the altitude of the corresponding training point towards one class and $(1 - s_n)$ is the attitude of meaning less. The optimal hyperplane problem of FSVM is defined as:

$$\begin{aligned} \min_{\mathbf{w}, \xi, \mathbf{b}} & \frac{1}{2} \|\mathbf{w}\|^2 + \mathbf{c}\mathbf{s}^T \xi \\ \text{s.t.} & \begin{cases} y_n (\mathbf{w}^T p_n - \mathbf{b}) \geq 1 - \xi_n, & n = 1, \dots, N \\ \xi_n \geq 0 \end{cases} \end{aligned} \tag{9}$$

where $\mathbf{s} = (s_1, s_2, \dots, s_N)$ represents the membership vector of FMF. A smaller s_n decreases the influence of the parameter ξ_n , such that the corresponding sample p_n is regarded less substantial. In a similar way, the Lagrangian is constructed as:

$$\min_{\mathbf{w}, \xi, \mathbf{b}} \max_{\alpha, \beta} \left\{ \frac{1}{2} \|\mathbf{w}\|^2 + \mathbf{c}\mathbf{s}^T \xi - \sum_{n=1}^N \alpha_n [y_n (\mathbf{w}^T x_n - \mathbf{b}) + \xi_n - 1] - \sum_{n=1}^N \xi_n \beta_n \right\} \tag{10}$$

Again, “dual form” is used to transform Problem (10) to

$$\begin{aligned} \max_{\alpha} & \sum_{n=1}^N \alpha_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_m \alpha_n y_m y_n p_m^T p_n \\ \text{s.t.} & \begin{cases} 0 \leq \alpha_n \leq s_n C \\ \sum_{n=1}^N \alpha_n y_n = 0, & n = 1, \dots, N \end{cases} \end{aligned} \tag{11}$$

Therefore, it is clear that the task becomes merely a function of the support vectors, which are the subset of the training data lying on the margins.

2.4.3. Fuzzy Membership Function

We set the fuzzy membership function (FMF) to the distance between the point and its class center. Suppose the mean of class +1 and class -1 as p_+ and p_- , respectively. Then, we can get the radius of class +1 and class -1 as

$$r_+ = \max_{\{p_n: y=1\}} |p_+ - p_n| \tag{12}$$

$$r_- = \max_{\{p_n: y=-1\}} |p_- - p_n| \tag{13}$$

where r_+ and r_- represents the radius of class +1 and -1, respectively. The fuzzy membership s_n is defined

as a function of both the radius and the mean of both classes [43]

$$s_n = \begin{cases} 1 - |p_+ - p_n| / (r_+ + \delta) & y_n = +1 \\ 1 - |p_- - p_n| / (r_- + \delta) & y_n = -1 \end{cases} \quad (14)$$

where $\delta > 0$ is used to guarantee $s_n > 0$.

2.4.4. Multiclass Technique

Support vector machines (SVMs) and its variants were originally developed for a two-class problem. However, we needed to predict three classes: green, oolong, and black tea. Several methods have already been proposed for multi-class problems via SVMs, among which the most popular approach was to break down the multiclass task into multiple two-class tasks. We chose the Winner-Takes-All (WTA) method.

Suppose $V (>2)$ classes exist in the task. WTA strategy classifies new instances based on the idea of one-versus-all. At first, we train V different individual binary classifiers (SVM or its variants). The n -th individual classifier is aimed to distinguish the data in class n from the data of all the remaining classes $(1, 2, \dots, n-1, n+1, \dots, V)$. A new test sample will be sent to all the V individual classifiers, and the individual classifier that outputs the largest value is chosen.

2.5. Statistical Setting

It will yield an optimistically biased assessment if the whole dataset is used as a validation set, which is dubbed “in-sample estimate”. Therefore, the whole dataset is divided into two sets: training set and test (or validation) set, and the evaluation performance of the test set is reported, which is dubbed “out-of-sample estimate”. However, for small-size dataset problems, the “out-of-sample estimate” will increase the variance of estimation of classification performance [45].

In this study, we used a more advanced technique to calculate the out-of-sample performance of the proposed identification system: the K -fold stratified cross validation (SCV) technique. The original samples were randomly segmented into K mutually exclusive subsets with closely equal length. Then, $K-1$ subsets were used for training and the rest for validation.

The abovementioned procedure repeated K runs, such that each subset was used once and only once for validation. The K validation results from the K runs were then merged together, in order to generate an out-of-sample estimation over the whole dataset. We assigned K with a value of 10 following common convention.

The 10-fold SCV was repeated 10 times, to further reduce the variance of estimation. Stratification technique was used so that each subset contained roughly the same proportions of different tea classes. As per Figure 4, the purpose of the proposed system is bi-fold: offline learning for training the classifier, and online prediction for predicting the category of query tea images.

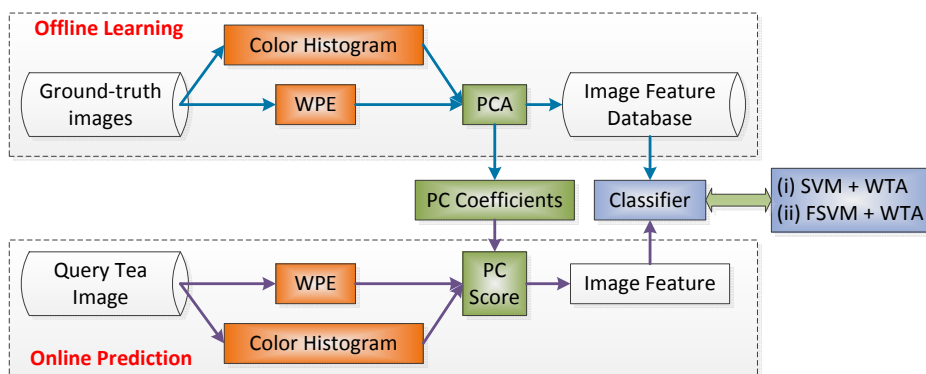


Figure 4. Diagram of the proposed automatic tea classification system.

3. Results and Discussions

The experiments were carried out on the IBM machine with 3 GHz core i3 processor and 8 GB random access memory (RAM), running on the Windows 7 operating system. The algorithms were developed in-house via Matlab 2015a (The Mathworks ©, Natick, Massachusetts, USA).

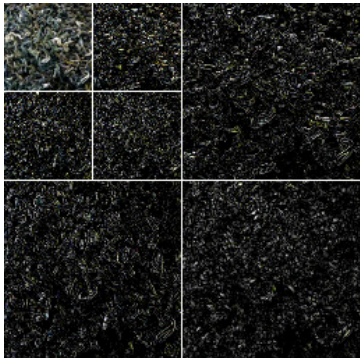
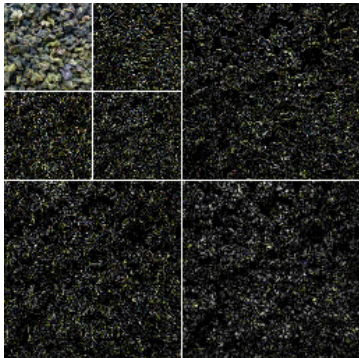
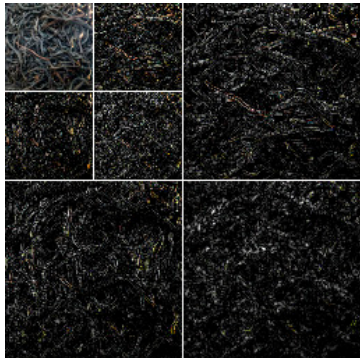
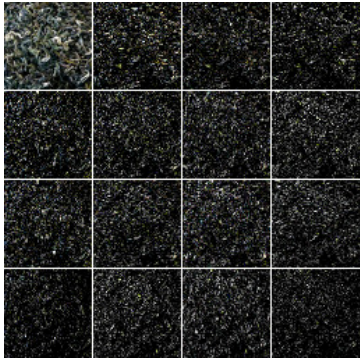
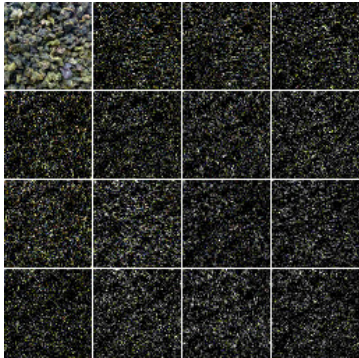
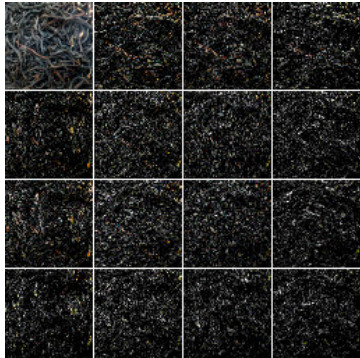
3.1. Feature Extraction

The second row of Table 4 shows the sample of each category of teas. Indeed, their colors and textures were distinct from each other using human vision. Next, we tried to extract their color information by CH. The third row of Table 4 shows the corresponding histogram, which clearly indicates the distribution of CHs of three categories of teas were different; hence, the CH is an effective feature.

Table 4. Feature extraction of three category of tea.

	Green	Oolong	Black
Sample Image			
Color Histogram			

Table 4. Cont.

	Green	Oolong	Black
Discrete Wavelet Transform (DWT)			
Discrete Wavelet Packet Transform (DWPT)			

The fourth and fifth rows of Table 4 compare the results of DWT decomposition with DWPT decomposition. There were three channels (RGB) of each image, so we performed decomposition for each channel and combined them to output the final decomposition result. It was clearly observed the DWPT decomposed the detail component that DWT did not decompose. Therefore, DWPT can give more multi-resolution information than DWT. Entropy was extracted on the 16 subbands of DWPT.

3.2. PCA Result

The 80 features extracted from each tea image were aligned to be a row vector, and features of all images were aligned row-wise to generate a 2D matrix, on which PCA was performed. Figure 5 shows the curve of variance explained against number of PCs, and the detailed results are listed in Table 5, from which it is clear one PC preserves 93.91% of total variances, two PCs preserve 99.08% of total variances, three PCs preserve 99.49% of total variances, four PCs preserve 99.78% of total variances. Finally, five PCs obtains more than 99.90% of total variance, which meet our criterion.

Table 5. Detailed data of accumulated explained variance.

# of PC	1	2	3	4	5
Variance Explained	93.91%	99.08%	99.49%	99.78%	99.90%

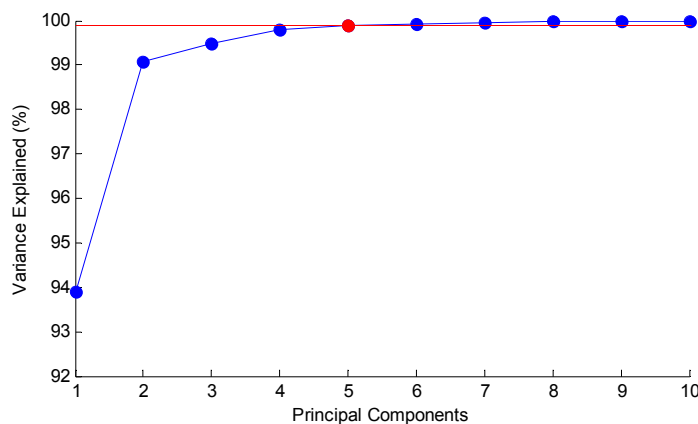


Figure 5. Curve of accumulated variance explained against number of principal components (PCs).

3.3. Classification Performance Comparison

We compared the proposed two classifiers (SVM + WTA; and FSVM + WTA), with state-of-the-art methods (BP-ANN [10], SVM [12], LDA [17], GNN [18], and FSCABC-FNN [21]), by averaging the results of 10×10 -fold SCV. The classification-performance comparison results are listed in Table 6. Please refer Abbreviation to see the full names of abbreviations.

Table 6. Average sensitivity/recall rate over a 10×10 -fold stratified cross validation (SCV).

Existing Approaches							
# of Original Features	# of Reduced Features	Classifier	Green Tea	Oolong Tea	Black Tea	Overall	Rank
8 metal features	8	BP-ANN [10]	N/A	N/A	N/A	95%	6
3735 spectrum features	5	SVM [12]	90%	100%	95%	95%	6
12 color + 12 texture features	11	LDA [17]	96.7%	92.3%	98.5%	95.8%	5
2 color + 6 shape features	8	GNN [18]	95.8%	94.4%	97.9%	96.0%	4
64 CH + 7 texture + 8 shape features	14	FSCABC-FNN [21]	98.1%	97.7%	96.4%	97.4%	2
Proposed Approaches							
64 CH + 16 WPE features	5	SVM + WTA	95.7%	98.1%	97.9%	97.23%	3
64 CH + 16 WPE features	5	FSVM + WTA	96.2%	98.8%	98.3%	97.77%	1

Studies in literatures [10] and [12] differentiated the same categories of green tea, oolong tea, and black tea as in this study. Therefore, we directly listed their results on test set, which obtained the out-of-sample evaluation, the same as the K -fold SCV used in this study.

Table 6 shows that the proposed two methods (“SVM + WTA” and “FSVM + WTA”) achieve good recall rates. The “SVM + WTA” method obtained 95.7%, 98.1%, and 97.9% on green, oolong, and black tea, respectively. The “FSVM + WTA” method obtained 96.2%, 98.8%, and 98.3% on the same categories. The overall recall rate of former method was 97.23%, while the latter one obtained 97.77% overall recall rate. From the results, we can deduce the FSVM was more effective than SVM. It can increase the overall recall rate slightly (about 0.54%). This finding aligns with past publications that stated FSVM is better than SVM [46,47].

In summary, the proposed “FSVM + WTA” yields 97.77% overall recall rate over an average of 10 runs, which exceeds not only the proposed “SVM + WTA” method with 97.23%, but also the state-of-the-art methods including BP-ANN [10] with 95%, SVM [12] with 95%, LDA [17] with 95.8%, GNN [18] with 96.0%, and FSCABC-FNN [21] with 97.4%. The reasons were two-fold: (i) Compared to traditional shape features and texture features, the WPE has better a ability to analyze transient features of non-stationary signals, which are abundant in tea leaves. (ii) The FSVM is a rather novel variant of SVM. FSVM can reduce the effect of outliers and noises, hence, increasing the overall recall rate.

From another point of view, few features expend less time. The number of the final reduced features of the proposed method was only five, less than or equal to BP-ANN [10] with eight, SVM [12] with five, LDA [17] with 11, GNN [18] with eight, and FSCABC-FNN [21] with 14. The excellent recall rate proved the effectiveness of those five features used in this study.

3.4. Optimal Wavelet

To find which wavelet performs best for this problem, we tested six different wavelets including db1, db2, db3, bior2.2, bior3.3, and bior4.4. The classification method was “FSVM + WTA”. Decomposition level was set to 2. A 10×10 -fold SCV was employed. The average overall recall-rates by different wavelets are listed in Table 7.

Table 7. Average Overall Recall-Rate of different wavelets.

Wavelet	Recall Rate
db1	96.37%
db2	96.60%
db3	96.57%
bior2.2	96.53%
bior3.3	97.10%
bior4.4	97.77%

Table 7 shows the bior4.4 wavelet achieves the highest average overall recall-rate among all wavelets. This explains why we chose the bior4.4 wavelet in our research work. In addition, db1 wavelet performs the worst for tea-category identification. To find the reason, we show the scaling function, wavelet function, low-pass filter (LPF) and high-pass filter (HPF) of both decomposition and reconstruction functions of bior 4.4 in Figure 6 and Figure 7, respectively.

We find that the scaling functions and wavelet functions of bior4.4 are more similar to gray-level changes along the textures and edges in tea images. Hence, this may be the reason why bior4.4 is better than other wavelets. This result aligns with past publications: Manthalkar *et al.* [48] stated “*Bior4.4 gives the best classification performance*” and Prabhakar and Reddy [49] also found bior4.4 performed the best among all wavelets.

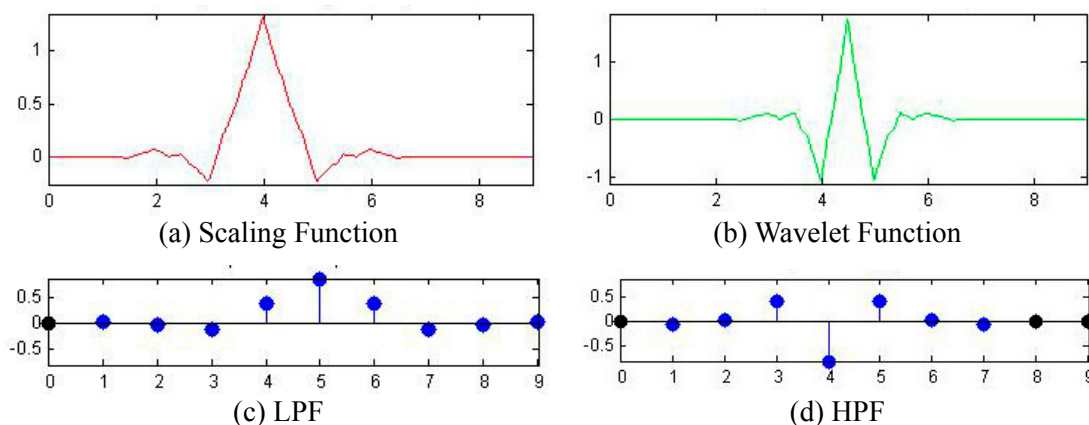


Figure 6. Decomposition functions for bior4.4, (a) Scaling Function, (b) Wavelet Function, (c) low-pass filter (LPF), (d) high-pass filter (HPF).

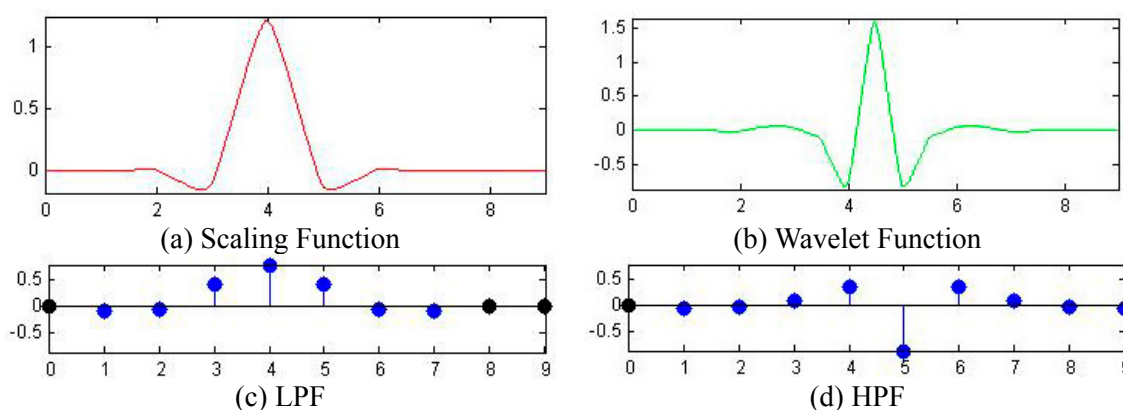


Figure 7. Reconstruction functions for bior4.4, (a) Scaling Function, (b) Wavelet Function, (c) low-pass filter (LPF), (d) high-pass filter (HPF).

3.5. Further Discussion of Methods

Here, we revisit the proposed methods. Firstly, there were three reasons why we used CH: (i) Color information is an effective tool for recognizing objects, which is the case of this study; (ii) The CH related to the physical properties of the teas; (iii) CH not only can reflect illumination conditions and tea color, but also ascertains image geometry and surface roughness.

Secondly, we have two reasons for choosing WPE: (i) Wavelet packet transform yields more information related to high-frequency subbands than wavelet transform does, which was essential in feature extraction; (ii) Entropy is effective in describing the uncertainty and complexity of 1D signals and 2D images. Although DWPT took more time than DWT, it is worthy considering its excellent performance.

Thirdly, FSVM is employed since the influences of training samples vary. Usually, one part of training samples is more important than the rest. Hence, it is natural to expect the important training points to be recognized correctly, and that other training samples (like noises) are not properly considered by the classifier. Generally, using FSVM, each sample no longer only belongs to one class. This is the fundamental concept of FSVM.

4. Conclusion

The goal of this work is to develop an automatic tea-category classification system with high sensitivity/recall rate. The contributions of this study consisted of the following aspects: (i) We proposed using WPE as a novel feature; (ii) We introduced FSVM that can reduce the effect from either outliers or noises, compared to conventional SVM; (iii) We used a WTA technique to cope with the three-class problem; (iv) The proposed method achieved a higher identification rate than five state-of-the-art methods while using the least features.

In the future, we would like to extract fractional features [50–52] from tea images. We will also try to apply the FSVM method to pathological brain detection [53–55], fruit classification [56], neuroscience [57,58], and CT Image Processing [59–61]. In addition, its stability and convergence [62,63] will be analyzed.

Acknowledgment

This paper was supported by NSFC (61273243, 51407095), Natural Science Foundation of Jiangsu Province (BK20150982, BK20150983), Jiangsu Key Laboratory of 3D Printing Equipment and Manufacturing (BM2013006), Key Supporting Science and Technology Program (Industry) of Jiangsu Province (BE2012201, BE2013012-2, BE2014009-3), Program of Natural Science Research of Jiangsu Higher Education Institutions (13KJB460011, 14KJB520021), Special Funds for Scientific and Technological Achievement Transformation Project in Jiangsu Province (BA2013058), Nanjing Normal University Research Foundation for Talented Scholars (2013119XGQ0061, 2014119XGQ0080).

Author Contributions

S. Wang & X. Yang conceived study, Y. Zhang & P. Phillips designed the model, Y. Zhang acquired the data, J. Yang & T.F. Yuan processed the data, S. Wang, Y. Zhang, J. Yang interpreted the results, Y. Zhang & P. Phillips developed the programs, Y. Zhang wrote the draft, T.F. Yuan gave critical revisions. All authors have read and approved the final manuscript.

Abbreviation

(A)(BP)(F)NN	(Artificial) (Back-propagation) (Feed-forward) neural network
(D)W(P)T	(Discrete) wavelet (packet) transform
(FSC)ABC	(Fitness-scaled Chaotic) Artificial bee colony
(F)SVM	(Fuzzy) support vector machine
(L)(H)PF	(Low-) (High-) Pass Filter
CCD	charge-coupled device
CH	Color Histogram
FMF	Fuzzy membership function
GNN	Genetic neural-network
LDA	Linear discriminant analysis
NIR	Near-infrared
SCV	Stratified cross validation

SE	Shannon entropy
WPE	Wavelet packet entropy
WTA	Winner-Takes-All

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Yang, C.S.; Landau, J.M. Effects of tea consumption on nutrition and health. *J. Nutr.* **2000**, *130*, 2409–2412.
2. Lim, H.J.; Shim, S.B.; Jee, S.W.; Lee, S.H.; Lim, C.J.; Hong, J.T.; Sheen, Y.Y.; Hwang, D.Y. Green tea catechin leads to global improvement among Alzheimer's disease-related phenotypes in NSE/hAPP-C105 Tg mice. *J. Nutr. Biochem.* **2013**, *24*, 1302–1313.
3. Qi, H.; Li, S.X. Dose-response meta-analysis on coffee, tea and caffeine consumption with risk of Parkinson's disease. *Geriatr. Gerontol. Int.* **2014**, *14*, 430–439.
4. Sironi, E.; Colombo, L.; Lompo, A.; Messa, M.; Bonanomi, M.; Regonesi, M.E.; Salmona, M.; Airolidi, C. Natural compounds against neurodegenerative diseases: Molecular characterization of the interaction of catechins from Green Tea with A β 1-42, PrP106–126, and Ataxin-3 Oligomers. *Chem. Eur. J.* **2014**, *20*, 13793–13800.
5. Bøhn, S.K.; Croft, K.D.; Burrows, S.; Puddey, I.B.; Mulder, T.P.J.; Fuchs, D.; Woodman, R.J.; Hodgson, J.M. Effects of black tea on body composition and metabolic outcomes related to cardiovascular disease risk: A randomized controlled trial. *Food Funct.* **2014**, *5*, 1613–1620.
6. Hajiaghaalipour, F.; Kanthimathi, M.S.; Sanusi, J.; Rajarajeswaran, J. White tea (*Camellia sinensis*) inhibits proliferation of the colon cancer cell line, HT-29, activates caspases and protects DNA of normal cells against oxidative damage. *Food Chem.* **2015**, *169*, 401–410.
7. Ch Yiannakopoulou, E. Green Tea catechins: Proposed mechanisms of action in breast cancer focusing on the interplay between survival and apoptosis. *Anti-Cancer Agents Medicinal. Chem.* **2014**, *14*, 290–295.
8. Wang, L.F.; Zhang, X.W.; Liu, J.; Shen, L.; Li, Z.Q. Tea consumption and lung cancer risk: A meta-analysis of case-control and cohort studies. *Nutrition* **2014**, *30*, 1122–1127.
9. Horanni, R.; Engelhardt, U.H. Determination of amino acids in white, green, black, oolong, pu-erh teas and tea products. *J. Food Compos. Anal.* **2013**, *31*, 94–100.
10. Herrador, M.A.; González, A.G. Pattern recognition procedures for differentiation of green, black and oolong teas according to their metal content from inductively coupled plasma atomic emission spectrometry. *Talanta* **2001**, *53*, 1249–1257.
11. Zhao, J.W.; Chen, Q.S.; Huang, X.Y.; Fang, C.H. Qualitative identification of tea categories by near infrared spectroscopy and support vector machine. *J. Pharm. Biomed Anal.* **2006**, *41*, 1198–1204.
12. Chen, Q.S.; Zhao, J.W.; Fang, C.H.; Wang, D.M. Feasibility study on identification of green, black and oolong teas using near-infrared reflectance spectroscopy based on support vector machine (SVM). *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2007**, *66*, 568–574.

13. Wu, D.; Chen, X.J.; He, Y. Application of multispectral image texture to discriminating tea categories based on DCT and LS-SVM. *Spectrosc. Spectr. Anal.* **2009**, *29*, 1382–1385.
14. Chen, Q.S.; Liu, A.P.; Zhao, J.W.; Ouyang, Q. Classification of tea category using a portable electronic nose based on an odor imaging sensor array. *J. Pharm. Biomed. Anal.* **2013**, *84*, 77–83.
15. Liu, N.; Liang, Y.Z.; Bin, J.; Zhang, Z.M.; Huang, J.H.; Shu, R.X.; Yang, K. Classification of green and black teas by PCA and SVM analysis of cyclic voltammetric signals from metallic oxide-modified electrode. *Food Anal. Method* **2014**, *7*, 472–480.
16. Borah, S.; Hines, E.L.; Bhuyan, M. Wavelet transform based image texture analysis for size estimation applied to the sorting of tea granules. *J. Food Eng.* **2007**, *79*, 629–639.
17. Chen, Q.; Zhao, J.; Cai, J. Identification of tea varieties using computer vision. *Trans. ASABE* **2008**, *51*, 623–628.
18. Jian, W.; Xian, Y.Z.; Shi, P.D. Identification and grading of tea using computer vision. *Appl. Eng. Agric.* **2010**, *26*, 639–645.
19. Gill, G.S.; Kumar, A.; Agarwal, R. Monitoring and grading of tea by computer vision —A review. *J. Food Eng.* **2011**, *106*, 13–19.
20. Laddi, A.; Sharma, S.; Kumar, A.; Kapur, P. Classification of tea grains based upon image texture feature analysis under different illumination conditions. *J. Food Eng.* **2013**, *115*, 226–231.
21. Zhang, Y.D.; Wang, S.H.; Ji, G.L.; Phillips, P. Fruit classification using computer vision and feedforward neural network. *J. Food Eng.* **2014**, *143*, 167–177.
22. Cattani, C. Fractional calculus and Shannon wavelet. *Math. Probl. Eng.* **2012**, *2012*, doi:10.1155/2012/502812.
23. Zhang, Y.D.; Wang, S.H.; Ji, G.L.; Dong, Z.C. Exponential wavelet iterative shrinkage thresholding algorithm with random shift for compressed sensing magnetic resonance imaging. *IEEJ Trans. Electr. Electron. Eng.* **2015**, *10*, 116–117.
24. Atangana, A.; Goufo, E.F.D. Computational analysis of the model describing HIV infection of CD4(+)T cells. *Biomed Res. Int.* **2014**, *2014*, doi:10.1155/2014/618404.
25. Tai, Y.-H.; Chou, L.-S.; Chiu, H.-L. Gap-type a-Si TFTs for front light sensing application. *J. Disp. Technol.* **2011**, *7*, 679–683.
26. De Almeida, V.E.; da Costa, G.B.; Fernandes, D.D.S.; Diniz, P.H.G.D.; Brandão, D.; de Medeiros, A.C.D.; Vêras, G. Using color histograms and SPA-LDA to classify bacteria. *Anal. Bioanal. Chem.* **2014**, *406*, 5989–5995.
27. Cattani, C. Harmonic wavelet approximation of random, fractal and high frequency signals. *Telecommun. Syst.* **2010**, *43*, 207–217.
28. Zhang, Y.D.; Wang, S.H.; Phillips, P.; Dong, Z.C.; Ji, G.L.; Yang, J.Q. Detection of Alzheimer's disease and mild cognitive impairment based on structural volumetric MR images using 3D-DWT and WTA-KSVM trained by PSOTVAC. *Biomed. Signal Process. Control* **2015**, *21*, 58–73.
29. Fang, L.T.; Wu, L.N.; Zhang, Y.D. A novel demodulation system based on continuous wavelet transform. *Math. Probl. Eng.* **2015**, *2015*, doi:10.1155/2015/513849.
30. Zhang, S.; Wang, S.; Wu, L. A novel method for magnetic resonance brain image classification based on adaptive chaotic PSO. *Prog. Electromagn. Res.* **2010**, *109*, 325–343.
31. Abdon, A. Numerical analysis of time fractional three dimensional diffusion equation. *Therm. Sci.* **2015**, *19*, 7–12.

32. Cattani, C.; Pierro, G.; Altieri, G. Entropy and multifractality for the myeloma multiple TET 2 gene. *Math. Probl. Eng.* **2012**, *2012*, doi:10.1155/2012/193761.
33. Zhang, Y.D.; Wang, S.H.; Sun, P.; Phillips, P. Pathological brain detection based on wavelet entropy and Hu moment invariants. *Bio-Med. Mater. Eng.* **2015**, *26*, 1283–1290.
34. Li, Y.-B.; Ge, J.; Lin, Y.; Ye, F. Radar emitter signal recognition based on multi-scale wavelet entropy and feature weighting. *J. Cent. South Univ.* **2014**, *21*, 4254–4260.
35. Moshrefi, R.; Mahjani, M.G.; Jafarian, M. Application of wavelet entropy in analysis of electrochemical noise for corrosion type identification. *Electrochem. Commun.* **2014**, *48*, 49–51.
36. Yang, Y.H.; Li, X.L.; Liu, X.Z.; Chen, X.B. Wavelet kernel entropy component analysis with application to industrial process monitoring. *Neurocomputing* **2015**, *147*, 395–402.
37. Zhang, Y.D.; Wang, S.H.; Phillips, P.; Ji, G.L. Binary PSO with mutation operator for feature selection using decision tree applied to spam detection. *Knowl.-Based Syst.* **2014**, *64*, 22–31.
38. Wang, S.H.; Zhang, Y.D.; Dong, Z.C.; Du, S.D.; Ji, G.L.; Yan, J.; Yang, J.Q.; Wang, Q.; Feng, C.M.; Phillips, P. Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. *Int. J. Imaging Syst. Techn.* **2015**, *25*, 153–164.
39. Zhang, Y.D.; Dong, Z.C.; Wang, S.H.; Ji, G.L.; Yang, J.Q. 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* **2015**, *17*, 1795–1813.
40. Zhang, Y.D.; Dong, Z.C.; Liu, A.J.; Wang, S.H.; Ji, G.L.; Zhang, Z.; Yang, J.Q. Magnetic resonance brain image classification via stationary wavelet transform and generalized eigenvalue proximal support vector machine. *J. Med. Imaging Health Inform.* **2015**, *5*, 1395–1403.
41. Zhang, Y.D.; Wang, S.H.; Dong, Z.C. Classification of alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree. *Prog. Electromagn. Res.* **2014**, *144*, 171–184.
42. Zhang, Y.D.; Wang, S.H.; Ji, G.L.; Dong, Z.C. An MR brain images classifier system via particle swarm optimization and kernel support vector machine. *Sci. World J.* **2013**, *2013*, doi:10.1155/2013/130134.
43. Lin, C.-F.; Wang, S.-D. Fuzzy support vector machines. *Neural Netw. IEEE Trans.* **2002**, *13*, 464–471.
44. Xian, G.-M. An identification method of malignant and benign liver tumors from ultrasonography based on GLCM texture features and fuzzy SVM. *Expert Syst. Appl.* **2010**, *37*, 6737–6741.
45. Zhang, Y.D.; Dong, Z.C.; Phillips, P.; Wang, S.H.; Ji, G.L.; Yang, J.Q. Exponential wavelet iterative shrinkage thresholding algorithm for compressed sensing magnetic resonance imaging. *Inform. Sci.* **2015**, *322*, 115–132.
46. Kumar, S.; Ghosh, S.; Tetarway, S.; Sinha, R.K. Support vector machine and fuzzy C-mean clustering-based comparative evaluation of changes in motor cortex electroencephalogram under chronic alcoholism. *Med. Biol. Eng. Comput.* **2015**, *53*, 609–622.
47. Abe, S. Fuzzy support vector machines for multilabel classification. *Pattern Recognit.* **2015**, *48*, 2110–2117.
48. Manthalkar, R.; Biswas, P.K.; Chatterji, B.N. Rotation and scale invariant texture features using discrete wavelet packet transform. *Pattern Recognit. Lett.* **2003**, *24*, 2455–2462.

49. Prabhakar, B.; Reddy, M.R. HVS scheme for DICOM image compression: Design and comparative performance evaluation. *Eur. J. Radiol.* **2007**, *63*, 128–135.
50. Atangana, A.; Alkahtani, B.S.T. Analysis of the Keller–Segel model with a fractional derivative without singular kernel. *Entropy* **2015**, *17*, 4439–4453.
51. Yang, X.-J.; Srivastava, H.M.; He, J.-H.; Baleanu, D. Cantor-type cylindrical-coordinate method for differential equations with local fractional derivatives. *Phys. Lett. A* **2013**, *377*, 1696–1700.
52. Yang, X.J.; Baleanu, D.; Khan, Y.; Mohyud-Din, S.T. Local fractional variational iteration method for diffusion and wave equations on cantor sets. *Romanian. J. Phys.* **2014**, *59*, 36–48.
53. Zhang, Y.D.; Dong, Z.C.; Phillips, P.; Wang, S.H.; Ji, G.L.; Yang, J.Q.; Yuan, T.-F. Detection of subjects and brain regions related to Alzheimer’s disease using 3D MRI scans based on eigenbrain and machine learning. *Front. Comput. Neurosci.* **2015**, *9*, doi:10.3389/fncom.2015.00066.
54. Zhang, Y.D.; Dong, Z.C.; Ji, G.L.; Wang, S.H. Effect of spider-web-plot in MR brain image classification. *Pattern Recognit. Lett.* **2015**, *62*, 14–16.
55. Zhang, Y.D.; Wang, S.H.; Dong, Z.C.; Phillips, P.; Ji, G.L.; Yang, J.Q. Pathological brain detection in magnetic resonance imaging scanning by wavelet entropy and hybridization of biogeography-based optimization and particle swarm optimization. *Prog. Electromagn. Res.* **2015**, *152*, 41–58.
56. Wang, S.H.; Zhang, Y.D.; Ji, G.L.; Yang, J.Q.; Wu, J.G.; Wei, L. Fruit classification by wavelet-entropy and feedforward neural network trained by fitness-scaled chaotic ABC and biogeography-based optimization. *Entropy* **2015**, *17*, 5711–5728.
57. Yuan, T.-F.; Hou, G.L. The effects of stress on glutamatergic transmission in the brain. *Mol. Neurobiol.* **2015**, *51*, 1139–1143.
58. Zhang, Y.D.; Wang, S.H. Detection of Alzheimer’s disease by displacement field and machine learning. *PeerJ.* **2015**, *3*, doi:10.7717/peerj.1251.
59. Chen, Y.; Ma, J.; Feng, Q.; Luo, L.; Chen, W.; Shi, P. Nonlocal Prior Bayesian Tomographic Reconstruction. *J. Math. Imag. Vis.* **2008**, *30*, 133–146.
60. Chen, Y.; Huang, S.; Pickwell-MacPherson, E. Frequency-wavelet domain deconvolution for terahertz reflection imaging and spectroscopy. *Optic. Express* **2010**, *18*, 1177–1190.
61. Chen, Y.; Chen, W.; Yin, X.; Ye, X.; Bao, X.; Luo, L.; Feng, Q.; Li, Y. Improving Low-dose Abdominal CT Images by Weighted Intensity Averaging over Large-scale Neighborhoods. *Eur. J. Radiol.* **2011**, *80*, e42–e49.
62. Atangana, A. On the stability and convergence of the time-fractional variable order telegraph equation. *J. Comput. Phys.* **2015**, *293*, 104–114.
63. Atangana, A. Convergence and stability analysis of a novel iteration method for fractional biological population equation. *Neural Comput. Appl.* **2014**, *25*, 1021–1030.