IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 3, 2016, accepted October 3, 2016, date of publication October 10, 2016, date of current version January 4, 2017.

Digital Object Identifier 10.1109/ACCESS.2016.2615520

An Efficient and Effective Automatic Recognition System for Online Recognition of Foreign Fibers in Cotton

XUEHUA ZHAO[1, 2], DAOLIANG LI[3], BO YANG[2], SHUANGYIN LIU[4], ZHIFANG PAN[5], AND HUILING CHEN[6]

[1]School of Digital Media, Shenzhen Institute of Information Technology, Shenzhen 518172, China
[2]Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, 130012, China
[3]College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
[4]School of Information Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
[5]Information Technology Center, Wenzhou Medical University, Wenzhou 325035, China
[6]College of Physics and Electronic Information, Wenzhou University, Wenzhou 325035, China

Corresponding author: H. Chen (chenhuiling.jlu@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61303113, Grant 61373053, Grant 61572226, Grant 61402195, Grant 61471133, and Grant 61571444, in part by the Science and Technology Plan Project of Wenzhou of China under Grant G20140048, in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY17F020012 and Grant LY16F030010, in part by the Guangdong Science and Technology Plan Project under Grant 2015A070709015, Grant 2015A020209171, and Grant 2016A040402043 and Guangdong Natural Science Foundation under Grant 2016A030310072, in part by the Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education Foundation of Jilin University under Grant 93K172016K19, and in part by the Jilin Province Natural Science Foundation under Grant 20150101052JC.

ABSTRACT Preventing foreign fibers from being mixed with cotton is essential for producing high-quality cotton textile products. It remains a challenging task to accurately distinguish foreign fibers from cotton. This paper proposes an efficient recognition system to accurately recognize foreign fibers mixed in cotton. The core component of the proposed system is an efficient classifier based on the kernel extreme learning machine (KELM). A two-step grid search strategy, which integrates a coarse search with a fine search, is adopted to train an optimal KELM recognition model. The resultant model is compared with the support vector machine and extreme learning machine on a real data set using tenfold cross-validation analysis. The experimental results show that the proposed recognition system can achieve classification accuracy as high as 93.57%, which is superior to the other two state-of-the-art systems. The results clearly confirm the superiority of the developed model in terms of classification accuracy. Promisingly, the proposed system can serve as a new candidate of powerful foreign fiber recognition systems with excellent performance.

INDEX TERMS Foreign fibers in cotton, online recognition system, kernel extreme learning machine.

I. INTRODUCTION

FOREIGN fibers that can appear in cotton, including hair, binding ropes, plastic film, candy wrappers, and polypropylene twines, have grave impacts on the quality of cotton textile products and lead to great economic loss for cotton textile enterprises [1]. Traditionally, experienced farmers have manually detected these foreign fibers, which is a laborious and time-consuming process. In recent years, machine-vision-based technology for online foreign fiber recognition has attracted researchers' attention [2]. The high recognition accuracy of machine-vision-based online recognition systems is of significant importance. Foreign fiber recognition typically adheres to the following procedure: (1) acquiring high quality images, (2) obtaining foreign fiber objects in the image, (3) extracting the features of these objects, (4) constructing a classification model. The final step (accurate object classification) plays an important role in foreign fiber recognition. A number of classification models have been developed for this purpose. Ji *et al.* [3] used a decision tree support vector machine (SVM) to identify the types of common foreign fibers in cotton. This experiment demonstrated greater than 92% accuracy in identifying different kinds of foreign fibers. Yang *et al.* [4] proposed the use of a one-against-one directed acyclic graph multi-class SVM to

perform classification; a mean accuracy of 92.34% was achieved. Other works have focused on feature selection procedures [5]–[7].

Most of the work on foreign fiber recognition has focused on SVM for constructing the final recognition model. Recently, Huang et al. proposed a new learning algorithm, the extreme learning machine (ELM) [8], for single hidden layer feed-forward neural networks (SLFNs). In this method, the ELM randomly chose input weights and hidden biases, and the output weights were analytically determined using a Moore–Penrose generalized inverse. More recently, in order to tackle the problem of large variation in the classification accuracy found in ELM in different trials, Huang *et al.* [9] proposed a kernel version of ELM (KELM), which required no randomness in assigning connection weights between the input and hidden layers. KELM can achieve comparative or better performance than SVM, is easier to implement, and has a faster training speed in applications like hyperspectral remote-sensing image classification [10]–[12], activity recognition [13], 2-D profile reconstruction [14], biodiesel engine performance optimization [15], [16], disease diagnosis [17], [18], and fault diagnosis [19]. To the best of our knowledge, applications of the KELM method in foreign fiber recognition have not yet been explored.

In this study, we propose an efficient machine-vision-based recognition system of foreign fibers in cotton. The experimental results demonstrate that the proposed approach achieves superior recognition performance over other methods. Our work has three main contributions: (1) An efficient recognition system for cotton foreign fibers is presented, in which KELM is used as the classification model. (2) A two-stage grid search is used to find the optimal parameters of the KELM model. (3) The experimental results show that the proposed approach has superior recognition performance over other methods.

The empirical analysis demonstrates that, the presented system might assist farmers in accurately recognizing foreign fibers in cotton, which will improve cotton textile products and reduce economic losses.

This paper is organized as follows: Section 2 describes the proposed system in detail. Section 3 provides the experimental setting. Section 4 describes and discusses the results. Conclusions and proposals for future work are summarized in Section 5.

II. PROPOSED FOREIGN FIBER RECOGNITION SYSTEM

The aim of our proposed system is to improve the recognition accuracy of identifying foreign fibers in cotton. The system consists of five components: image acquisition, image segmentation, feature extraction, model training, and foreign fiber recognition. The foreign fiber image acquisition system is built to acquire foreign ?ber images. Then, image segmentation is performed to locate the foreign fibers in the images. Next, multiple types of features, including color, shape, and texture are extracted from the images. Then, all of the features are fed into the KELM classifier for training an optimal model

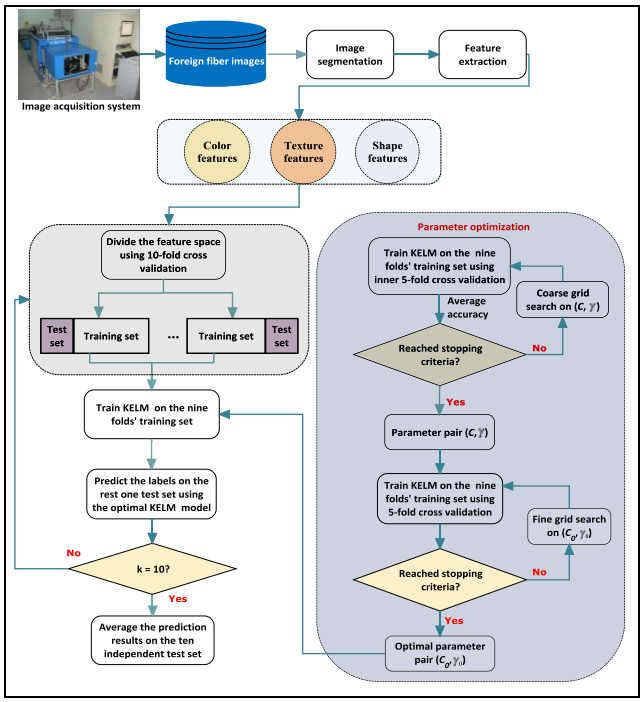

FIGURE 1. Architecture of the proposed system.

using the two-step grid search strategy. Finally, KELM conducts the recognition task using the trained optimal model. The architecture of the proposed foreign fiber recognition system is shown in Figure 1.

FIGURE 2. Foreign fiber image acquisition system.

A. FOREIGN FIBER IMAGE ACQUISITION SYSTEM

The foreign fiber image acquisition system, which is shown in Figure 2, is used to acquire the foreign fiber images in real time. The key components of the system are a lint layer generator, lint layer collector, and detection box. The light

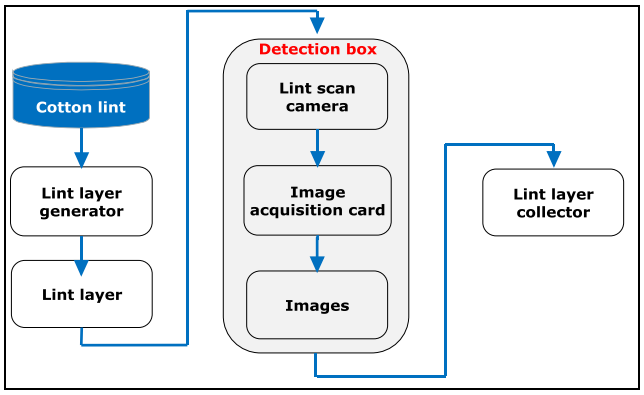

FIGURE 3. Flowchart of foreign fiber images acquisition.

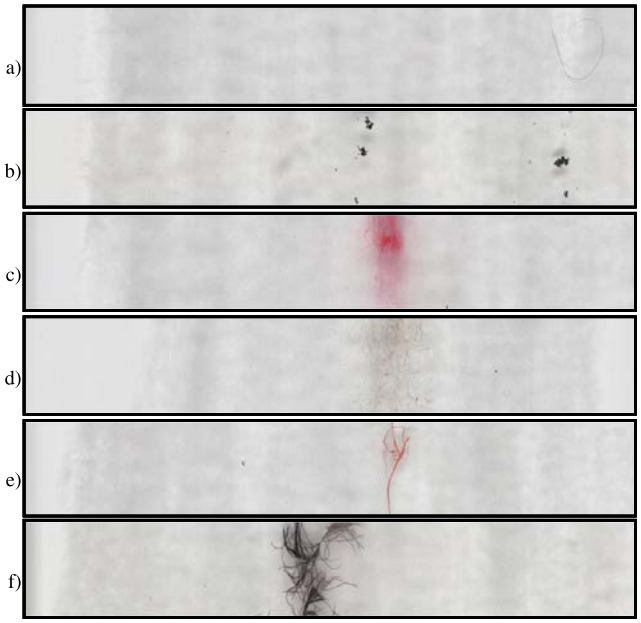

FIGURE 4. Sample of foreign fiber images. (a) hair (b) black plastic film (d) red cloth (d) hemp rope (e) red polypropylene (f) black feather.

source, line scan camera, and image acquisition card are set in the detection box. The process of image acquisition is shown in Figure 3. First, the lint layer generator transfers the lint to a line layer that is about 2mm thick, so that the foreign fibers in the lint can be easily "seen". Then, the line layer goes through the detection box with the LED light source and line scan camera. In the detection box, an image acquisition card produces and assembles the scan lines of the lint layer into an image frame that is 4000 pixels wide and 500 pixels long. The collected images are directly transmitted to a host computer where the images are processed and analyzed. Finally, the detected lint is collected by the lint layer collector.

The resolution of the acquired original images is 4000×500 pixels. The sample of cotton foreign fibers images is shown in Figure 4, where (a)-(f) are, respectively, hair, black plastic film, red cloth, hemp rope, red polypropylene,

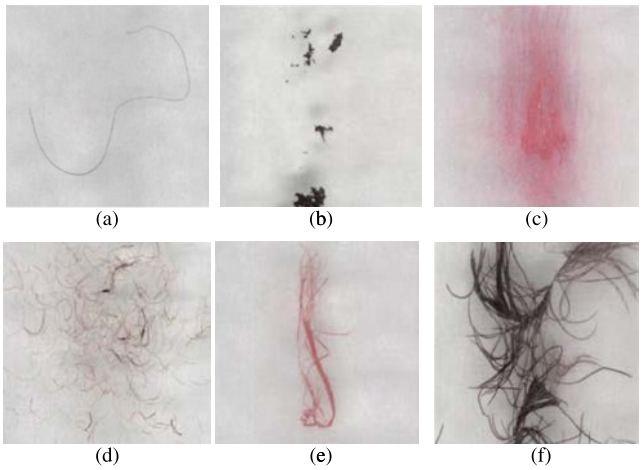

FIGURE 5. Local Cuts of Sample images. (a) hair (b) black plastic film (d) red cloth (d) hemp rope (e) red polypropylene (f) black feather.

and black feathers. For a clearer view of the foreign fibers, Figure 5 shows local cuts of the original images.

B. IMAGE SEGMENTATION

Image segmentation is used to partition an image into a set of meaningful connected components. In this study, the objective of image segmentation is to locate the foreign fibers in the image. The located parts, which only contain foreign fibers, are called the foreign fiber objects. Finding these foreign fiber objects is the precondition of machine vision-based foreign fiber recognition. This can help accurately extract the features of foreign fibers.

A number of different image segmentation methods have been proposed, some of which have been applied to automated visual recognition systems in agriculture [1], [2]. In these methods, the traditional thresholding method is used, which has the advantages of computational simplicity, high speed, and easy implementation. In our study, an improved thresholding method is used for image segmentation, which considers the properties of foreign fiber images in cotton and was proposed in 2011 by our team [1]. In this method, the regularity of the gray distribution of the images is used to enhance the images and the improved Otsu's method is used to partition the images. The closing operation of mathematical morphology is used to fill the gaps generated in the process of image segmentation, and noise produced by small objects is removed by an area calculation. Figure 6 shows the original images before image segmentation. Figure 7 shows the objects of the images corresponding to Figure 6 after image segmentation. As can be seen from the figure, the method is able to accurately locate the foreign fibers in the cotton.

C. FEATURE EXTRACTION

The foreign fibers in cotton vary in terms of color, texture, and shape. To obtain acceptable recognition accuracy, color, shape, and texture features are extracted in this study. For

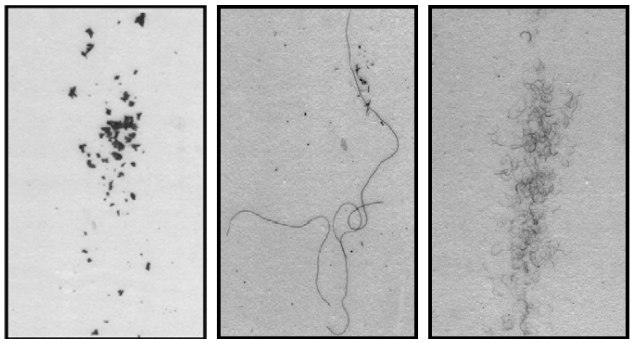

FIGURE 6. Original images. (a) plastic film (b) hair (c) hemp rope.

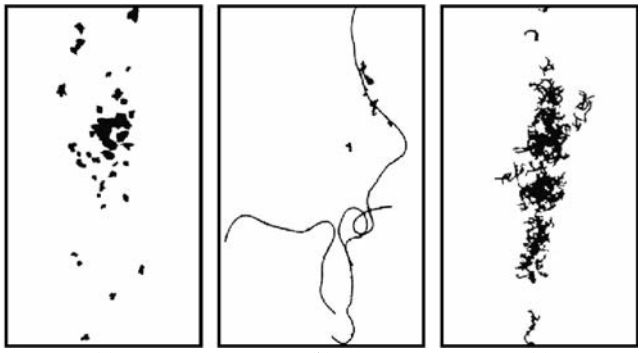

FIGURE 7. Results of image segmentation. (a) plastic film (b) hair (c) hemp rope.

TABLE 1. The extracted color features.

Class	Name
RGB model	RGB mean, R mean, G mean, B mean, RGB variance, R variance, G variance, B variance, RGB third-order moment, R third-order moment, G third-order moments, B third-order moments
HSV model	HSV mean, H mean, S mean, V mean, HSV variance, H variance, S variance, V variance, HSV third-order moment, H third-order moment, S third-order moment, V third-order moments
Gray model	Mean, Variance, Third-order moment

color features, we consider these features in three different models (RGB model, HSV model, and Gray model). There are a total of 27 extracted color features, which are listed in Table 1.

We extract the four classes of texture features: gray-level co-occurrence matrix, gray-gradient co-occurrence matrix, gray-smooth co-occurrence matrix, and gray-level differences. There are 41 total extracted texture features, which are listed in Table 2.

We extract 7 shape features: Area, Euler number, Form factor, Eccentricity, Solidity, Rectangularity, and Sphericity.

The 27 color features, 41 texture features, and 7 shape features, form the 75-dimensional feature vector.

D. FOREIGN FIBER RECOGNITION

The classification model is one of the key components of the recognition system and is used to automatically sort foreign

TABLE 2. The extracted texture features.

Class	Name
Gray-level co-occurrence matrix	(mean, when d=1,θ=0°,45°,90°,135°): Contrast Entropy, Homogeneity, Energy, Cluster shade, Cluster prominent, Gray correlation (mean, when d=2,θ=0°,45°,90°,135°): Contrast, Entropy, Homogeneity, Energy, Cluster shade, Cluster prominent, Gray correlation
Gray-gradient co-occurrence matrix	Small gradient strengths, Large gradient strengths, Gray hole, Uneven gradient distribution, Energy, Average gray level, Average gradient, Gray level variance, Gradient variance, Correlation, Entropy, Inertia, Inverse difference
gray-smooth co-occurrence matrix	Contrast, Entropy, Inverse difference, Energy, Cluster shade, Cluster prominence
Gray-level differences	(d=1 pixel): Contrast, Second moment, Entropy, Average (d=2 pixel): Contrast, Second moment, Entropy, Average

fibers into the correct category. Currently, SVM is extensively used as the classifier model for foreign fiber recognition in cotton because of its high accuracy and acceptable speed. KELM is another advanced machine learning method. Compared to SVM, KELM has fewer optimization constraints, simpler implementation, faster learning, and better generalization performance. Due to these advantages, KELM has attracted attention from many researchers ever since it was developed from single-hidden-layer feed forward neural networks (SLFNs) in 2004 [20]. However, until now, KELM methods have not been used for foreign fiber recognition in cotton. In order to further improve the recognition accuracy of the system, in our work, we study the KELM model and employ a two-stage grid search strategy to tackle the parameter optimization problem found in KELM. To the best of our knowledge, this is the first effort in the literature to apply the KELM method to foreign fiber recognition in cotton. Here, we briefly introduce KELM(for more details refer to [9]).

Given a sample set $\aleph = \{(x_i, t_i)|x_i \in R^n, t_i \in R^m, i = 1, 2, \ldots, N\}$, where $\boldsymbol{x}_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^{\mathrm{T}} \in \mathbf{R}^n$ is the input vector and $\boldsymbol{t}_i = [t_{i1}, t_{i2}, \ldots, t_{im}] \in \mathbf{R}^m$ is the target vector. The standard SLFNs with an activation function $g(x)$ and $\tilde{N}$ hidden neurons can be mathematically modeled as:

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot x_j + b_i) = o_j, \quad j = 1, 2, \ldots, N \tag{1}$$

where $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^{\mathrm{T}}$ denotes the vector of the output weight which connects the i-th hidden neuron and the output neurons, $\boldsymbol{w}_i = [w_{i1}, w_{i2}, \ldots, w_{in}]^{\mathrm{T}}$ denotes the weight vector which connects the i-th neuron and the input neurons, b_i denotes the bias of the i-th hidden neuron, $\boldsymbol{w}_i \cdot \boldsymbol{x}_i$ is the inner product of two vectors, and o_j denotes the output of the j-th sample data.

If SLFNs with $\tilde{N}$ hidden neurons and activation function $g(\mathrm{x})$ can approximate these N samples with zero errors, we will have $\sum_{j=1}^{N} \|o_j - t_j\| = 0$, which means there exists β_i,

$\boldsymbol{w}_i, b_i$ so that

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot x_j + b_i) = t_j, \quad j = 1, 2, \ldots, N \tag{2}$$

Eq. (2) also can be written compactly as follows:

$$\mathbf{H}\beta = \mathbf{T} \tag{3}$$

$$\mathbf{H}(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{\tilde{N}}, b_1, \ldots, b_{\tilde{N}}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) = \begin{bmatrix} g(\boldsymbol{w}_1 \bullet \boldsymbol{x}_1 + b_1) & \cdots & g(\boldsymbol{w}_{\tilde{N}} \bullet \boldsymbol{x}_1 + b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ g(\boldsymbol{w}_1 \bullet \boldsymbol{x}_\mathbf{N} + b_1) & \cdots & g(\boldsymbol{w}_{\tilde{N}} \bullet \boldsymbol{x}_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \tag{4}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \tag{5}$$

$$\mathbf{T} = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{\tilde{N} \times m} \tag{6}$$

where $\mathbf{H}$ is the hidden layer output matrix of the neural network, the i-th column of H is the i-th hidden neuron output with respect to inputs $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$:

Usually, training an SLFN is done to find the specific $\hat{\mathbf{w}}_i, \hat{b}_i, \hat{\beta}_i (i = 1, \ldots, \tilde{N})$ such that

$$\|\mathbf{H}(\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_{\tilde{N}}, \hat{b}_1, \ldots, \hat{b}_{\tilde{N}})\hat{\beta} - \mathbf{T}\| = \min_{\boldsymbol{w}_i, b_i, \beta} \|\mathbf{H}(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{\tilde{N}}, b_1, \ldots, b_{\tilde{N}})\beta - \mathbf{T}\| \tag{7}$$

When $\mathbf{H}$ is unknown, the methods for searching for the minimum of $\|\mathbf{H}\beta - \mathbf{T}\|$ generally are gradient-based learning algorithms. In the searching procedure, $\boldsymbol{w}$, β, b is iteratively adjusted. For feed forward neural networks, the most commonly used learning algorithm is the BP learning algorithm where gradients can be computed efficiently by propagation from the output to the input. However, there are several issues in BP learning algorithms including the presence of local minima, poor generalization performance, and length of time needed. ELM can efficiently avoid these problems.

Unlike the general approximation theories which require adjustments to $\boldsymbol{w}$ and b, Huang [8] has shown that input weights $\boldsymbol{w}$ and hidden layer biases b can be randomly assigned if the activation function is infinitely differentiable. It is notable that the input weights $\boldsymbol{w}$ and the hidden layer biases b are not necessarily tuned. The hidden layer output matrix $\mathbf{H}$ can actually remain unchanged once random values have been assigned to these parameters in the beginning of learning. For fixed input weights $\boldsymbol{w}$ and the hidden layer biases b, obtained from Eq. (7), training SLFN is simply equivalent to finding the least square solution $\hat{\beta}$ of the linear system $\mathbf{H}\beta = \mathbf{T}$:

$$\|\mathbf{H}(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{\tilde{N}}, b_1, \ldots, b_{\tilde{N}})\hat{\beta} - \mathbf{T}\| = \min_{\beta} \|\mathbf{H}(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{\tilde{N}}, b_1, \ldots, b_{\tilde{N}})\beta - \mathbf{T}\| \tag{8}$$

If the number $\tilde{N}$ of hidden neurons is equal to the number N of training samples, the matrix $\mathbf{H}$ is square and invertible when the input weight vectors $\boldsymbol{w}_i$ and the hidden biases b_i are randomly chosen, and SLFNs can approximate these training samples with no errors. However, in most cases, the number of hidden neurons is much less than the number of training samples, $\mathbf{H}$ is a nonsquare matrix, and there may not exist $\boldsymbol{w}_i$, b_i; $\beta_i (i = 1, \ldots, \tilde{N})$ satisfying $\mathbf{H}\beta = \boldsymbol{T}$. The minimum norm least squares solution of the above linear system is:

$$\hat{\beta} = \mathbf{H}^\dagger \mathbf{T} \tag{9}$$

where $\mathbf{H}^\dagger$ is the Moor-Penrose generalized inverse of the matrix $\mathbf{H}$. When the feature mapping is unknown [9], we can use a kernel matrix for the ELM according to Eq. (10):

$$\Omega_{ELM} = \mathrm{HH}^T : \Omega_{ELMi,j} = h(x_i) \cdot h(x_j) = K(x_i, x_j) \tag{10}$$

where $h(x)$ is used to map the data from the input space to the hidden-layer feature space. The orthogonal projection method is used to calculate the Moore-Penrose generalized inverse of the matrix, i.e. $\mathbf{H}^\dagger = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}$, and a positive constant C is added to the diagonal of $\mathbf{H}\mathbf{H}^\mathrm{T}$. Accordingly, the output function of ELM becomes the following equation:

$$f(x) = h\beta = h(x)\mathbf{H}^T\left(\frac{I}{C} + \mathbf{H}\mathbf{H}^T\right)^{-1}\mathbf{T} = \begin{bmatrix} K(x, x_1) \\ K(x, x_N) \end{bmatrix}^T \left(\frac{I}{C} + \Omega_{ELM}\right)\mathbf{T} \tag{11}$$

The specific kernel implementation of ELM is named KELM. When the corresponding kernel for the ELM model can be specified, the hidden layer feature mapping does not need to be known to the users. The function of the RBF kernel is often taken as the first choice for the problem at hand, which can be formulated as $K(x, x_i) = \exp(-\gamma \|x - x_i\|^2)$. The two main parameters of the RBF kernel are the penalty parameter C and the kernel bandwidth γ. The penalty parameter C determines the trade-off between model complexity and the fitting error minimization. The parameter γ defines the non-linear mapping from the input space to high-dimensional feature space.

In this study, a two-step grid search strategy is used to determine the values of two parameters, in order to find the optimal penalty parameter C and kernel bandwidth γ. First, a coarse grid search is performed to search for the parameter pair, and then the fine grid search is conducted on the range obtained in the last step. The optimal parameter pair that is acquired last is used to train the model for prediction.

III. EXPERIMENTS

A. EXPERIMENTAL SETUP

All the experiments in this work were conducted on a Dell Intel Core i5-4460 (3.2 GHz) with 4GB of RAM and the Windows 7 operating system. All the methods were implemented using MATLAB. ELM and KELM programs implemented by Huang (http://www3.ntu.edu.sg/home /egbhuang) and LIBSVM, developed by Chang and Lin [21], are used in the experiments.

Before taking the classification, normalization is performed to avoid feature values in greater numerical ranges dominating feature values in smaller numerical ranges. Three main normalization methods are used in data mining: min-max normalization, z-score normalization, and normalization by decimal scaling [22]. After testing three methods in the data of foreign fibers in cotton, we found that there were not obvious differences in the accuracy of classification among the three methods. Finally, we use min-max normalization to preprocess the data in our experiments to scale the data into the interval of $[-1, 1]$.

TABLE 3. Number of foreign fiber objects.

Class	Number of objects
Plastic film objects	34
Cloth objects	36
Hemp rope objects	41
Hair objects	17
Polypropylene objects	44
Feather objects	62
Total	**234**

B. DATA DESCRIPTION

In the experiments, all the samples of foreign fibers in cotton were provided by China Cotton Machinery & Equipment Co., Ltd. These samples fell into six categories: cloth, feather, hair, hemp rope, plastic film and polypropylene. The data was generated according to the following procedure: First, the samples are uniformly mixed into lint. Second, 2000 images were collected by an image acquisition system. Then 234 foreign fiber objects were generated by image segmentation; the number of each class of objects is listed in Table 3. For each object, 75 features were extracted to form a 75-dimension data vector. Finally, the dataset of 234 vectors was obtained from these corresponding objects because each object contributes one sample vector.

In order to ensure valid and unbiased results, the k-fold CV is employed to evaluate the classification performance. In k-fold CV, the whole data set is divided into k parts, with each part containing the same number of samples. The k-1 of the k parts are put together to form a training set and the remaining part is used as the test set. Lastly, the average result across all k trials is computed. The advantage of k-fold CV is that the test sets are independent and the results have high reliability. In most studies, k is set to 10. For KELM and SVM adopting RBF kernel function, the stratified 10-fold CV is employed in our experiments since the penalty parameter C and kernel parameters need to be chosen (that is, model selection must occur) [23]. In the nested stratified 10-fold CV, classification performance evaluation is done in the outer loop where 10-fold CV is adopted. This means that the classifiers are evaluated on the one fold that is independent, and the data from the other nine folds are left to train. Parameter optimization is done in the inner loop where 5-fold CV is taken, which means that KELM and SVM will search for the optimal parameters C and γ on the data of the remaining 9-folds, where the data of the 9-folds is further split into 1-fold data for performance evaluation and 4-fold data for model selection.

C. PARAMETER OPTIMIZATION

For fair comparison, both penalty parameter C and kernel bandwidth γin KELM and SVM are tuned using the two-step grid search strategy and the same grid. The detailed parameter setting for KELM and SVM are as follows. The coarse grid search ranges for the related parameters C and γin both KELM and SVM with RBF kernel vary between $C \in \{2^{-5}, 2^{-3}, \ldots, 2^{15}\}$ and $\gamma \in \{2^{-15}, 2^{-13}, \ldots, 2^{3}\}$. Thus 110 combinations of C and γ are attempted in this step. After an optimal pair (C_0, γ_0) is selected from this coarse grid search, the fine grid search is conducted around (C_0, γ_0), with $C \in \{0.25C_0, 0.5C_0, C_0, 2C_0, 4C_0\}$ and $g \in \{0.25\gamma_0, 0.5\gamma_0, \gamma_0, 2\gamma_0, 4\gamma_0\}$. In this step, 25 combinations of parameter pairs are tried. The final optimal parameter pair is selected from this fine search.

TABLE 4. Results obtained by KELM with optimal parameters.

Fold	Coarse search (C, γ)	Classification accuracy	Fine search (C_0, γ_0)	Classification accuracy
#1	(32, 8)	0.8261	(8, 32)	1.0000
#2	(1/8, 8)	0.8750	(1/2, 4)	0.9583
#3	(8, 8)	0.9167	(8, 32)	0.9583
#4	(2, 8)	1.0000	(1/2, 4)	0.9583
#5	(1/8, 2)	0.9583	(1/32, 1)	0.9167
#6	(1/8, 8)	0.9565	(1/2, 2)	0.8696
#7	(8, 8)	0.9565	(4, 8)	1.0000
#8	(32, 8)	0.8696	(8, 4)	0.9565
#9	(32, 8)	0.9565	(8, 8)	0.8696
#10	(8, 8)	0.8696	(16, 32)	0.8696
Avg.		**0.9185**		**0.9357**
Dev.		**0.0556**		**0.0513**

IV. RESULTS AND DISCUSSION

A. RESULTS OF KELM BASED RECOGNITION SYSTEM

The detailed results of KELM optimized by the two-step grid search strategy via 10-fold CV are detailed in Table 4. As shown in Table 4, KELM with the coarse search achieved a mean classification accuracy of 91.58% and a standard deviation of 0.0556. After a fine search, the performance of KELM is further improved. KELM has achieved good performance with a mean accuracy of 93.57% and a standard deviation of 0.0513 across over 10 independent folds. Contrasted to coarse search methods, the fine search has increased the mean accuracy by 2%. This indicates that the designed two-step grid search strategy can effectively determine the optimal parameter values for C and γ. Detailed values of C and γ are recorded in Table 4.

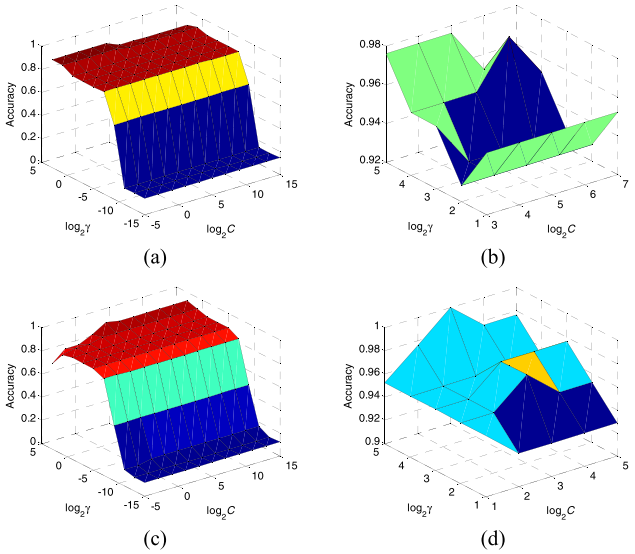

FIGURE 8. Training accuracy surface of KELM by two-step grid search. (a) coarse search in fold 1, (b) fine search in fold 1, (c) coarse search in fold 3, (d) fine search in fold 3.

Figure 8 shows the training classification accuracy surface of fold 1 and fold 3 in the two-step grid search procedure, where the x-axis, y-axis, and z-axis are log2C, log2γ, and accuracy, respectively. Each mesh point in the (x, y) plane of the training accuracy stands for a parameter combination and the z-axis denotes the obtained classification accuracy value for each parameter combination. As can be seen in Figure 8, in the coarse search stage, the change of accuracy of KELM is smooth in the large space [0, 1], but in the fine search stage the change of accuracy of KELM is sharp in the small space [0.9, 1].

TABLE 5. Frequency of optimal parameters for each fold.

C	Times	γ	Times
32	3	32	8
16	1	16	5
8	5	8	5
4	6	4	8
2	9	2	2
1	3	0.5	1
0.25	1	0.015625	1
0.125	1		
0.0078125	1		

To further analyze the relationship between parameters and accuracy, we randomly run the algorithm 10 times via 10-fold CV, and select and analyze the 30 parameter combinations of C and γ which correspond to these folds, with more than 91% testing accuracy. The statistical results of the parameter values are listed in Table 5. We can see that (1) the value of the parameter C varies largely when there is high accuracy. This indicates that the optimal C value is not unique. (2) 9 times the C value corresponds to 2, which is the largest value listed in Table 5. (3) The parameter γ has similar characteristics to the parameter C.

TABLE 6. KELM classification results.

Foreign fiber class	Classification results (1)	(2)	(3)	(4)	(5)	(6)	Number of samples	Accuracy (%)
(1) plastic film	34	0	0	0	0	0	34	100
(2) cloth	0	35	0	0	0	1	36	97.22
(3) hemp rope	0	1	36	0	1	3	41	87.80
(4) hair	0	0	1	14	1	1	17	82.35
(5) polypropylene	0	0	2	0	42	0	44	95.45
(6) feather	1	3	0	0	0	58	62	93.55
Mean								**93.57**

Table 6 illustrates the classification results of KELM in terms of a confusion matrix. The recognition rates of the plastic film, cloth, polypropylene, feather, hemp rope, and hair are 100%, 97.22%, 95.45%, 93.55%, 87.80%, and 82.35%, respectively. Of the six classes of samples the performance of KELM is the best for the plastic film samples and the worst for the hair samples. The plastic film can be correctly classified by KELM. For the other kinds of foreign fibers, one sample of cloth, five samples of hemp rope, three samples of hair, two samples of polypropylene, and four samples of feather were incorrectly recognized as other types. Overall, the total recognition rate was 93.57%, which is very promising for automatic foreign fiber recognition.

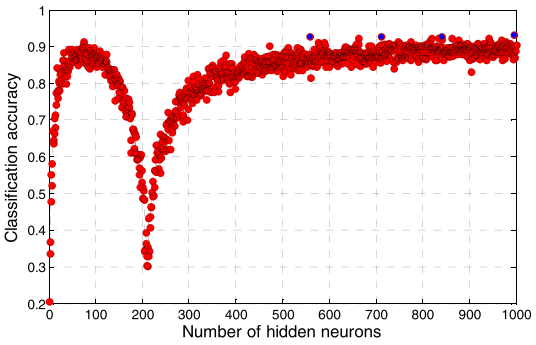

FIGURE 9. Relationship between the accuracy and number of hidden neurons for ELM.

B. COMPARISON WITH ELM AND SVM

In order to validate the KELM model, the core component of the proposed system, we compare ELM and SVM under the same situation. For ELM, the number of hidden neurons, which is one of the key parameters, has a critical impact on the performance of ELM. In order to determine the optimal number of hidden neurons, we run ELM with the number of hidden neurons in the foreign fiber dataset varying from 1 to 1000, and record the classification accuracy versus the number of hidden neurons. As shown in Figure 9, it is notable that with the increase of hidden neurons, ELM first reaches local peak validation accuracy, and then increases gradually

and achieves global optimal validation accuracy. When the numbers of hidden neurons are 558, 711, 840, and 996, high validation accuracy can be achieved. Considering the tradeoff between accuracy and complexity of the model, 558 hidden neurons are adopted to train the ELM classifier in the experiment. For the activation function, the sigmoid activation function is used to compute the hidden layer output matrix.

TABLE 7. Classification results of ELM.

Foreign fiber class	Classification results						Number of samples	Accuracy (%)
	(1)	(2)	(3)	(4)	(5)	(6)		
(1) plastic film	34	0	0	0	0	0	34	100
(2) cloth	0	31	1	0	0	4	36	86.11
(3) hemp rope	0	1	34	1	2	3	41	82.93
(4) hair	0	0	0	15	0	2	17	88.24
(5) polypropylene	0	0	1	0	41	2	44	93.18
(6) feather	1	2	3	0	0	56	62	90.32
Mean								**90.17**

TABLE 8. Classification results of the SVM.

Foreign fiber class	Classification results						Number of samples	Accuracy (%)
	(1)	(2)	(3)	(4)	(5)	(6)		
(1) plastic film	33	0	0	0	0	1	34	97.06
(2) cloth	0	33	0	0	0	3	36	91.67
(3) hemp rope	0	1	34	0	2	4	41	82.93
(4) hair	0	0	0	16	0	1	17	94.12
(5) polypropylene	0	0	1	2	41	0	44	93.18
(6) feather	1	4	0	0	1	56	62	90.32
Mean								**91.07**

Detailed classification results for ELM and SVM are recorded in terms of a confusion matrix as shown in Tables 7 and 8. From the two tables, we can see that ELM and SVM have average accuracies of 90.17% and 91.07%, respectively. Compared with ELM and SVM, KELM has achieved, respectively, 3.39% and 2.5% higher recognition rates. The detailed comparison results are illustrated in Figure 10. As can be seen, KELM outperforms the other two systems in most of the folds. Notably, the standard deviation obtained by KELM is smaller than SVM and ELM in the experiment. This indicates that the stability and robustness of KELM is better than ELM and SVM, and that KELM could become a promising recognition tool for foreign fiber recognition in cotton.

To validate KELM, we compare it to BP, PNN, and kNN [24], [25]. For BP, the number of hidden neurons is set to 4 according to the experience. For PNN, the spread is set to 1. For kNN, we set k to 1, 2, and 3. The results via 10-fold CV are shown in Figure 11. KELM had the best average accuracy among the eight algorithms. The kNN-2 had the worst performance in the data of foreign fibers in cotton.

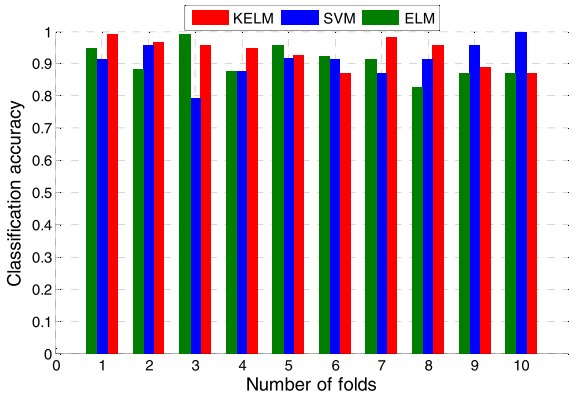

FIGURE 10. Classification accuracy of KELM, ELM and SVM.

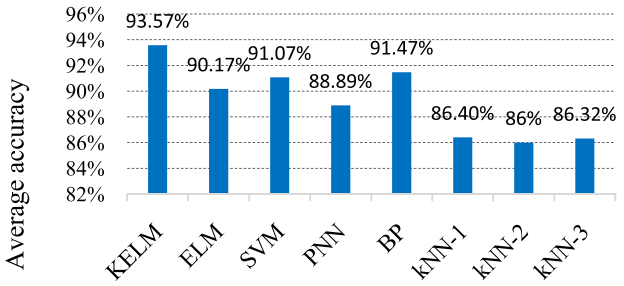

FIGURE 11. Comparison of average accuracy of eight algorithms.

C. KELM WITH FEATURE SELECTION

To determine the optimal feature set and improve the performance of system, in the next experiments, we adopt a feature selection approach based on the Fisher Score to select the optimal feature set for identifying foreign fibers in cotton. Then, we build the KELM based on the optimal classification subset.

The Fisher Score is an effective supervised feature selection algorithm and has been widely applied. Given class labels $y = \{y1, \ldots, yn\}$, Fisher Score [26] selects features that assign similar values to the samples from the same class and different values to samples from different classes. The Fisher Score can be defined as:

$$FS(f_i) = \frac{\sum_{j=1}^{c} n_j(\mu_{i,j} - \mu_i)^2}{\sum_{j=1}^{c} n_j \sigma_{i,j}^2} \tag{12}$$

where μ_i is the mean of the feature f_i, n_j is the number of samples in the jth class, and $\mu_{i;j}$ and $\sigma_{i;j}$ are the mean and the variance of f_i in class j.

In the experiments, the Fisher Score is evaluated independently for each feature and the features with the top-k values are selected as the relevant feature set. k is set as a value from 1 to 75 in turn, resulting in the generation of 75 feature sets. We then run the KELM in each feature set via 10-fold CV.

Figure 12 shows the mean classification accuracy of the subsets corresponding to different k values. It can be seen that the accuracy mean of the subsets increase in stability when the value of k is greater than 19. When the value of k is 38,

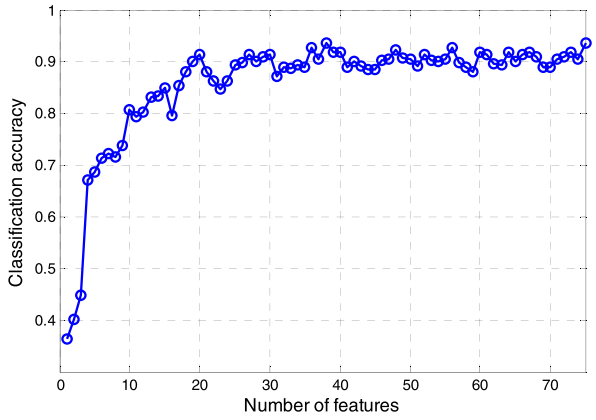

FIGURE 12. Curves of average accuracy of top-k feature set.

TABLE 9. The feature subset with 38 features.

Class	Number	Name
Color	18	RGB mean, R mean, G mean, B mean, R variance, RGB third-order moment, R third-order moment, G third-order moments, B third-order moments, HSV mean, S mean, V mean, H variance, S variance, V variance, HSV third-order moment, H third-order moment, Third-order moment
Texture	19	*Gray-level co-occurrence matrix*: (mean, when d=1,θ=0°,45°,90°,135°):Homogeneity, Gray correlation, (mean, when d=2,θ=0°,45°,90°,135°):Contrast, Entropy, Cluster prominent, Gray correlation *Gray-gradient co-occurrence matrix*: Uneven gradient distribution, Average gray level, Gray level variance, Gradient variance, Correlation, Entropy, Inertia, Inverse difference *gray-smooth co-occurrence matrix*: Inverse difference, Energy, Cluster shade, Cluster prominence *Gray-level differences*: Contrast (d=1 pixel),
Shape	1	Sphericity

the accuracy mean reaches 93.57%. This indicates that (1) Feature selection can help to find the optimal subsets that are small in size and have acceptable accuracy. In contrast to the original set with 75 features, the subset with 38 features can efficiently reduce the recognition speed of the system without losing accuracy. (2) It is difficult to obtain good performance for the recognition system using only a few features. (3) We can use the feature subset with 19 features as the feature set of the recognition system when focusing on considering the recognition speed and acceptable accuracy.

Tables 9 and 10 list the specific features included in the top-38 features and the top-19 features. Both of the subsets consist of three classes of features: color, texture, and shape. This indicates that ideal feature sets with high accuracy are combinations of the three classes of features and each class of feature has a different contribution for classification.

D. COMPARISONS WITH OTHER DESCRIPTORS

Here, we make comparisons with feature sets obtained by other descriptors (http://hexaco.org/), such as local binary pattern (LBP), uniform local binary pattern (ULBP), local binary pattern histogram fourier features (LBPHF), rotation invariant description with local binary pattern histogram fourier features (RILBPHF), local configuration pattern (LCP), local phase quantization (LPQ), rotation invariant local phase quantization (RILPQ) and binarized statistical image features (BSIF).

TABLE 10. The feature subset with 19 features.

Class	Number	Name
Color	10	RGB mean, R mean, G mean, B mean, RGB third-order moment, R third-order moment, G third-order moments, B third-order moments, HSV mean, H variance
Texture	8	*Gray-level co-occurrence matrix*: Gray correlation(mean, when d=1,θ=0°,45°,90°,135°) *Gray-gradient co-occurrence matrix*: Gray level variance, Correlation, Entropy, Inertia *gray-smooth co-occurrence matrix*: Energy, Cluster shade, Cluster prominence
Shape	1	Sphericity

TABLE 11. Comparisons of discriminative ability of nine feature sets.

Feature sets	Mean of accuracy	Standard deviation of accuracy
Original feature set	**0.9357**	**0.0084**
Feature set of LBP	0.5993	0.0184
Feature set of ULBP	0.6303	0.0193
Feature set of LBPHF	0.5622	0.0145
Feature set of RILBPHF	0.5599	0.0198
Feature set of LCP	0.7609	0.0158
Feature set of LPQ	0.5744	0.0197
Feature set of RILPQ	0.5647	0.0200
Feature set of BSIF	0.5507	0.0124

Firstly, we respectively build the datasets based on the eight descriptors and run KELM on the eight datasets. The results are listed in Table 11. As we see, the original feature set, that is combination of color, texture and shape features, has the best discriminative ability among nine feature sets. This indicates that any one of feature sets obtained by these descriptors cannot provide the excellent discriminative ability for recognition of cotton foreign fibers.

Secondly, to validate the performance of texture descriptors, we respectively replace the texture features in original feature sets with the features obtained by each descriptor and build the new eight feature sets. We respectively run KELM on the eight datasets based on new feature sets. The results are listed in Table 12. We can see that the original feature set still has the best discriminative ability among nine feature sets. This further indicates that the good recognition ability is difficult to achieve only using texture features, and the feature sets combining many classes of features are still current the best way to obtain the good recognition performance of cotton foreign fiber.

TABLE 12. Comparisons of discriminative ability of nine feature sets with three class of features.

Feature sets	Mean of accuracy	Standard deviation of accuracy
Original feature sets	**0.9357**	**0.0084**
Feature set including LBP	0.8958	0.0133
Feature set including ULBP	0.9172	0.0182
Feature set including LBPHF	0.8958	0.0130
Feature set including RILBPHF	0.8862	0.0189
Feature set including LCP	0.8892	0.0172
Feature set including LPQ	0.8363	0.0174
Feature set including RILPQ	0.8360	0.0129
Feature set including BSIF	0.8144	0.0252

V. CONCLUSION AND FUTURE WORK

In this work, we have developed an efficient recognition system for distinguishing foreign fibers that can appear in cotton. The core component of the proposed system is the KELM classifier, whose key parameters C and g are determined by a two-step grid search strategy via cross validation analysis. In order to verify the effectiveness of the KELM in identifying foreign fibers in cotton, other advanced classifiers such as SVM and ELM are used for comparison. The simulation results demonstrate that the KELM is more robust and stable than the other methods, and has a higher recognition rate. In addition, to finding optimal feature sets, feature selection technology is tested in our experiments. The results shows that feature selection can efficiently reduce the number of features needed to recognize foreign fibers in cotton without compromising accuracy. We can conclude that, the developed recognition system might assist farmers in accurately detecting foreign fibers in cotton, improving the cotton textile products and reducing economic losses.

In future works, we plan to study the parameter optimization of the KELM model using swarm intelligence algorithms with excellent global search ability like the genetic algorithm (GA), ant colony optimization (ACO), particle swarm optimization (PSO), and the Wolf Algorithm (WA). We will also attempt to put feature selection and model optimization into the same optimization process, in order to develop a recognition system that balances accuracy and complexity can be built for the detection of foreign fibers in cotton.

REFERENCES

[1] W. Yang, D. Li, L. Zhu, Y. Kang, and F. Li, "A new approach for image processing in foreign fiber detection," *Comput. Electron. Agricult.*, vol. 68, pp. 68–77, Aug. 2009, doi: 10.1016/j.compag.2009.04.005.

[2] D. Li, W. Yang, and S. Wang, "Classification of foreign fibers in cotton lint using machine vision and multi-class support vector machine," *Comput. Electron. Agricult.*, vol. 74, pp. 274–279, Nov. 2010, doi: 10.1016/j.compag.2010.09.002.

[3] R. Ji, D. Li, L. Chen, and W. Yang, "Classification and identification of foreign fibers in cotton on the basis of a support vector machine," *Math. Comput. Model.*, vol. 51, pp. 1433–1437, Jun. 2010, doi: 10.1016/j.mcm.2009.10.007.

[4] W. Yang, S. Lu, S. Wang, and D. Li, "Fast recognition of foreign fibers in cotton lint using machine vision," *Math. Comput. Model.*, vol. 54, pp. 877–882, Aug. 2011, doi: 10.1016/j.mcm.2010.11.010.

[5] X. Zhao, D. Li, B. Yang, C. Ma, Y. Zhu, and H. Chen, "Feature selection based on improved ant colony optimization for online detection of foreign fiber in cotton," *Appl. Soft Comput.*, vol. 24, pp. 585–596, Nov. 2014, doi: 10.1016/j.asoc.2014.07.024.

[6] X. Zhao, D. Li, W. Yang, and G. Chen, "Feature selection based on ant colony optimization for cotton foreign fiber," *Sensor Lett.*, vol. 9, no. 3, pp. 1242–1248, 2011.

[7] W. Yang, D. Li, and L. Zhu, "An improved genetic algorithm for optimal feature subset selection from multi-character feature set," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 2733–2740, 2011, doi: 10.1016/j.eswa.2010.08.063.

[8] G.-B. Huang, "What are extreme learning machines? Filling the gap between Frank Rosenblatt's dream and John von Neumann's puzzle," *Cognit. Comput.*, vol. 7, no. 3, pp. 263–278, 2015, doi: 10.1007/s12559-015-9333-0.

[9] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012, doi: 10.1109/TSMCB.2011.2168604.

[10] M. Pal, A. E. Maxwell, and T. A. Warner, "Kernel-based extreme learning machine for remote-sensing image classification," *Remote Sens. Lett.*, vol. 4, no. 9, pp. 853–862, 2013, doi: 10.1080/2150704X.2013.805279.

[11] C. Chen, W. Li, H. Su, and K. Liu, "Spectral-spatial classification of hyperspectral image based on kernel extreme learning machine," *Remote Sens.*, vol. 6, no. 6, pp. 5795–5814, 2014, doi: 10.3390/rs6065795.

[12] P. Du, A. Samat, P. Gamba, and X. Xie, "Polarimetric SAR image classification by boosted multiple-kernel extreme learning machines with polarimetric and spatial features," *Int. J. Remote Sens.*, vol. 35, no. 23, pp. 7978–7990, 2014, doi: 10.1080/2150704X.2014.978952.

[13] W.-Y. Deng, Q.-H. Zheng, and Z.-M. Wang, "Cross-person activity recognition using reduced kernel extreme learning machine," *Neural Netw.*, vol. 53, pp. 1–7, May 2014, doi: 10.1016/j.neunet.2014.01.008.

[14] B. Liu, L. Tang, J. Wang, A. Li, and Y. Hao, "2-D defect profile reconstruction from ultrasonic guided wave signals based on QGA-kernelized ELM," *Neurocomputing*, vol. 128, pp. 217–223, Mar. 2014, doi: 10.1016/j.neucom.2012. 11.053.

[15] P. K. Wong, K. I. Wong, C. M. Vong, and C. S. Cheung, "Modeling and optimization of biodiesel engine performance using kernel-based extreme learning machine and cuckoo search," *Renew. Energy*, vol. 74, pp. 640–647, Feb. 2015, doi: 10.1016/j.renene.2014.08.075.

[16] K. I. Wong, P. K. Wong, C. S. Cheung, and C. M. Vong, "Modelling of diesel engine performance using advanced machine learning methods under scarce and exponential data set," *Appl. Soft Comput.*, vol. 13, no. 11, pp. 4428–4441, 2013, doi: 10.1016/j.asoc.2013.06.006.

[17] H.-L. Chen, G. Wang, C. Ma, Z.-N. Cai, W.-B. Liu, and S.-J. Wang, "An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson's disease," *Neurocomputing*, vol. 184, pp. 131–144, Apr. 2015, doi: 10.1016/j.neucom.2015.07.138.

[18] C. Ma, J. Ouyang, H.-L. Chen, and X.-H. Zhao, "An efficient diagnosis system for Parkinson's disease using kernel-based extreme learning machine with subtractive clustering features weighting approach," *Comput. Math. Method Med.*, vol. 2014, Nov. 2014, Art. no. 985789, doi: 10.1155/2014/985789.

[19] Y. Jiang, J. Wu, and C. Zong, "An effective diagnosis method for single and multiple defects detection in gearbox based on nonlinear feature selection and kernel-based extreme learning machine," *J. Vibroeng.*, vol. 16, no. 1, pp. 499–512, 2014.

[20] G.-B. Huang, X. Ding, and H. Zhou, "Optimization method based extreme learning machine for classification," *Neurocomputing*, vol. 74, pp. 155–163, Dec. 2010, doi: 10.1016/j.neucom.2010.02.019.

[21] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.* vol. 2, no. 3, 2011, Art. no. 27.

[22] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Amsterdam, The Netherlands: Elsevier. 2011.

[23] H.-L. Chen *et al.*, "A novel bankruptcy prediction model based on an adaptive fuzzy *k*-nearest neighbor method," *Knowl.-Based Syst.*, vol. 24, pp. 1348–1359, Dec. 2011, doi: 10.1016/j.knosys.2011.06.008.

[24] Y. C. Lin, D.-D. Chen, M.-S. Chen, X.-M. Chen, and J. Li, "A precise BP neural network-based online model predictive control strategy for die forging hydraulic press machine," *Neural Comput. Appl.*, Sep. 2016.

[25] N. Settouti, M. E. A. Bechar, and M. A. Chikh, "Statistical comparisons of the top 10 algorithms in data mining for classification task," *Int. J. Interact. Multimedia Artif. Intell., Special Issue Artif. Intell. Underpinning*, vol. 4, pp. 46–51, Sep. 2016, doi: 10.9781/ijimai.2016.419.
[26] Z. Zhao, F. Morstatter, S. Sharma, S. Alelyani, A. Anand, and H. Liu, "Advancing feature selection research," Dept. Comput. Sci. Eng., Arizona State Univ., ASU Feature Selection Repository, Tech. Rep., 2010, pp. 1–28.

XUEHUA ZHAO received the Ph.D. degree from the College of Computer Science and Technology, Jilin University, in 2014. He is currently a Lecturer with the School of Digital Medium, Shenzhen Institute of Information Technology. His main research interests are related to machine learning and data mining.

DAOLIANG LI received the Ph.D. degree from the College of Information and Electrical Engineering, China Agricultural University, in 1999. He is currently a Professor with the College of Information and Electrical Engineering, China Agricultural University. His current research interests are in the areas of data mining and knowledge engineering, with applications to agricultural engineering.

BO YANG received the B.Sc., M.S., and Ph.D. degrees in computer science from Jilin University in 1997, 2000, and 2003, respectively. He is currently a Professor with the College of Computer Science and Technology, Jilin University, China. He is also the Head of the Key Laboratory of Symbolic Computation and Knowledge Engineering, Ministry of Education, China. He has authored over 60 international journal articles including the IEEE/ACM Transactions and conference papers including IJCAI and AAAI. His current research interests are in the areas of social network analysis theory, data mining, multi-agent systems with applications to knowledge discovery and engineering, and intelligent software design.

SHUANGYIN LIU received the Ph.D. degree from the College of Information and Electrical Engineering, China Agricultural University, in 2014. He is currently a Professor with the School of Information Science and Technology, Zhongkai University of Agriculture and Engineering. His current research interests are in the areas of intelligent information system of agriculture, artificial intelligence, software engineering, and computational intelligence.

ZHIFANG PAN was born in Wenzhou, China, in 1973. He received the bachelor's degree in physics from Wenzhou University, China, in 1996, the master's degree in medicine from Wenzhou Medical University, China, in 1999, and the Ph.D. degree from the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China. In 1999, he joined Wenzhou Medical University, where he is currently a Deputy Director of Information Technology Center and an Associate Professor with the Department of Computer and Information Management. His research interests include analysis of medical image and biomedical data.

HUILING CHEN received the Ph.D. degree from the School of Computer Science and Technology, Jilin University, China. He is currently a Lecturer with the Department of Computer Science and Technology, Wenzhou University, China. His present research interest centers on machine learning, pattern recognition, and data mining and their applications, including medical diagnosis, bankruptcy prediction, gene selection, and face recognition. He is currently a Reviewer for several international journals including the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, Part B, *Computer Methods and Programs in Biomedicine* and *Future Generation Computer Systems*. He has authored over 50 papers in international journals and conference proceedings, including *Expert Systems with Applications*, *Knowledge-Based Systems*, and *Soft Computing*.