

Research Article

A Multiple Classifier Fusion Algorithm Using Weighted Decision Templates

Aizhong Mi, Lei Wang, and Junyan Qi

School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China

Correspondence should be addressed to Lei Wang; wang_leiqiy@163.com

Received 2 June 2016; Revised 5 August 2016; Accepted 10 August 2016

Academic Editor: Guo Chen

Copyright © 2016 Aizhong Mi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fusing classifiers' decisions can improve the performance of a pattern recognition system. Many applications areas have adopted the methods of multiple classifier fusion to increase the classification accuracy in the recognition process. From fully considering the classifier performance differences and the training sample information, a multiple classifier fusion algorithm using weighted decision templates is proposed in this paper. The algorithm uses a statistical vector to measure the classifier's performance and makes a weighed transform on each classifier according to the reliability of its output. To make a decision, the information in the training samples around an input sample is used by the k -nearest-neighbor rule if the algorithm evaluates the sample as being highly likely to be misclassified. An experimental comparison was performed on 15 data sets from the KDD'99, UCI, and ELENA databases. The experimental results indicate that the algorithm can achieve better classification performance. Next, the algorithm was applied to cataract grading in the cataract ultrasonic phacoemulsification operation. The application result indicates that the proposed algorithm is effective and can meet the practical requirements of the operation.

1. Introduction

A single classifier was always used in a traditional pattern recognition system. However, in recent years, it has been found that the samples that were wrongly classified by distinct classifiers were usually not the same in many experiments. This finding means that complementary information about the object to be recognized can be potentially offered by different classifiers and effective fusion of the complementary information is expected to considerably improve the performance of a pattern recognition system. When the member classifiers are diverse or complementary, multiple classifier systems can usually obtain higher classification accuracies compared with a single classifier [1]. Through a large number of experiments and applications, it has been proven that fusing classifiers' decisions can achieve better performance than the best single classifier and improve the efficiency and robustness of a pattern recognition system. Currently, many applications areas have adopted the methods of multiple classifier fusion, such as fault diagnosis [2, 3], radar emitter recognition [4], remote sensing image recognition [5–7],

medical diagnosis [8], face recognition [9], and intrusion detection [10]. Some of the fusion methods have good generality and show good classification performance in certain applications. However, similar to the classifier, there is no fusion method that can obtain the optimal classification performance for all of the applications. Therefore, the study of multiple classifier fusion is still an open problem. For a specific application, this approach requires further research on a more suitable algorithm for multiple classifier fusion.

Multiple classifier fusion assumes that all of the classifiers are equally "experienced" over the entire feature space. Thus, all of the outputs of the classifiers are fused in a certain way to achieve the final decision. According to the different outputs of the classifiers, they can be further divided into three categories. When the output is shown in a decision form, Naive Bayes [11], majority voting, and Behavior Knowledge Space (BKS) [12] are the representative fusion algorithms. When the output is shown in a ranked form, Borda Count [13, 14] is a typical fusion algorithm. When the output is shown in a measured form, the fusion algorithms mainly include Max, Min, Sum, Product, Median, Support Vector Machine (SVM)

fusion [15–17], Neural Networks fusion [18], Dempster-Shafer theory [19], and decision templates (DT) [20].

The decision templates (DT) algorithm is a simple and effective fusion algorithm that has measure outputs and that is researched and used widely [21–26]. During the training stage, the DT algorithm calculates the average of the decision profiles that correspond to the training samples that belong to each class according to the outputs of all of the classifiers. The averages are the decision templates (one per class). Then, the input sample's class is determined by evaluating the similarity between its decision profile and various decision templates. The algorithm has the following advantages [27]: there is a simple training process; it requires no strict assumptions compared to probability-based algorithms; it is less sensitive to the size of the training set than other algorithms and overtraining rarely happens; it is also intuitive, having a small number of calculations; and it is not time-consuming. However, the DT algorithm still has two problems: (1) a decision template is only the average of the decision profiles that belong to a class and does not fully reflect the differences in the classifier's performance. (2) It uses only the classification information of the decision template and does not take full advantage of the training samples' information.

Therefore, an improved DT algorithm (called VWDT) is proposed here. This algorithm measures the classifier's performance by a statistical vector and assigns different weights to each classifier according to their outputs. Reliable output is assigned to a large weight in such a way that the output represents a significant share of the decision templates. For a sample that is easily misclassified, the information in the surrounding training samples is used to make a decision in addition to using the similarities between it and the decision templates.

This algorithm was compared with the DT algorithm on 15 data sets, from the KDD'99, UCI, and ELENA databases. The experimental results show that the algorithm can achieve better classification performance. Then, the algorithm was applied to cataract grading in the cataract ultrasonic phacoemulsification operation. The application effect indicates that the proposed algorithm is effective and can meet the practical requirements of the operation. Thus, the algorithm can be applied to the computer-aided cataract recognition system of the ultrasonic emulsification apparatus to automatically recognize the hardness grade of the cataract, which thus lowers the operation difficulty level, shortens the study period, and improves the safety of the operation.

The organization of this paper is as follows: Section 2 contains the description of the multiple classifier fusion algorithm proposed in this paper. The experiment test of the algorithm is given in Section 3, and Section 4 provides the validation of the practical application, including the application background and the application description and effect. Finally, Section 5 presents the conclusions and summarizes the contents of this paper.

2. VWDT Algorithm Description

2.1. Relative Expression. The structure of the multiple classifier fusion used in this paper is shown in Figure 1.

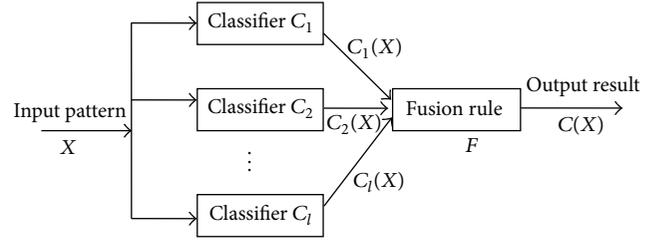


FIGURE 1: The structure of the multiple classifier fusion.

Let R^n be the n -dimensional feature space; let $X = [x_1, x_2, \dots, x_n]^T$ be the n -dimensional feature vector, $X \in R^n$; let $\Omega = \{\omega_1, \omega_2, \dots, \omega_m\}$ be the set of potential class labels; and let $C = \{C_1, C_2, \dots, C_l\}$ be the set of trained classifiers for decision fusion. Given the input pattern X , the output of the i th classifier is denoted as in

$$C_i(X) = [c_{i,1}(X), c_{i,2}(X), \dots, c_{i,m}(X)]^T, \quad (1)$$

where $c_{i,j}(X)$, $i = 1, 2, \dots, l$, $j = 1, 2, \dots, m$, represents the measure value of the possibility that classifier C_i considers that X belongs to class ω_j .

The fused output of l classifiers is constructed as in

$$C(X) = F(C_1(X), C_2(X), \dots, C_l(X)), \quad (2)$$

where F is the fusion rule.

The output of all of the classifiers can be represented as the decision profile, $l \times m$ matrix as shown in

$$DP(X) = \begin{bmatrix} c_{1,1}(X) & \cdots & c_{1,j}(X) & \cdots & c_{1,m}(X) \\ \vdots & & \vdots & & \vdots \\ c_{i,1}(X) & \cdots & c_{i,j}(X) & \cdots & c_{i,m}(X) \\ \vdots & & \vdots & & \vdots \\ c_{l,1}(X) & \cdots & c_{l,j}(X) & \cdots & c_{l,m}(X) \end{bmatrix}. \quad (3)$$

The i th row is the measure layer output of the i th classifier $C_i(X)$, which is given according to (1). The j th column is the possibility measure value that l classifiers consider that the input pattern X belongs to class ω_j . In addition, the fusion result $C(X)$ is an m -dimensional vector that is represented by the measure layer form, which is denoted as in

$$C(X) = [d_1(X), d_2(X), \dots, d_m(X)]^T, \quad (4)$$

where $d_i(X)$, $i = 1, 2, \dots, m$, shows the possibility measure value that X belongs to class ω_i after fusion.

After acquiring the fusion result of (4), a certain rule is used to judge which class the input pattern belongs to. Typically, the rule of using the maximum value is adopted; namely, if $d_K(X) = \max_{i=1}^m d_i(X)$, then it shall be deemed that $X \in \omega_K$.

The simple fusion rules (Max, Min, Sum, Product, and Median) acquire the system output by operating each column of DP(X).

(1) Max rule is

$$d_j(X) = \max_{i=1}^l c_{i,j}(X), \quad j = 1, 2, \dots, m. \quad (5)$$

The Max rule takes the maximum of every DP(X) column as the fused output C(X).

(2) Min rule is

$$d_j(X) = \min_{i=1}^l c_{i,j}(X), \quad j = 1, 2, \dots, m. \quad (6)$$

The Min rule takes the minimum of every DP(X) column as the fused output C(X).

(3) Sum rule is

$$d_j(X) = \sum_{i=1}^l c_{i,j}(X), \quad j = 1, 2, \dots, m. \quad (7)$$

The Sum rule computes the sum of every DP(X) column as the fused output C(X). It is also called the mean rule when it computes the mean; these are simply two forms of the same rule.

(4) Product rule is

$$d_j(X) = \prod_{i=1}^l c_{i,j}(X), \quad j = 1, 2, \dots, m. \quad (8)$$

The Product rule computes the product of every DP(X) column as the fused output C(X).

(5) Median rule is

$$d_j(X) = \text{median}_{i=1}^l c_{i,j}(X), \quad j = 1, 2, \dots, m. \quad (9)$$

The Median rule takes the median of every DP(X) column as the fused output C(X). If l is an even number, then the mean of two medians is taken as the result of a column.

The DT algorithm acquires the system output using the entire DP(X) and calculates the corresponding decision template DT_{*i*}, $i = 1, 2, \dots, m$, on the training set Z, as in

$$DT_i = \frac{1}{N_z} \sum DP(z_j), \quad z_j \in Z, \quad z_j \in \omega_i, \quad (10)$$

where z_j represents the samples that belong to class ω_i in the training set Z and N_z represents the number of z_j .

Then, the input sample's class is determined by evaluating the similarity between its decision profile and various decision templates. Here, the squared Euclidean distance is used for calculating the similarity, but other measures can also be

applied. For a training sample $P \in Z$, the calculation equation is shown in

$$d_E(DP(P), DT_i) = \sum_{j=1}^m \sum_{k=1}^l (c_{k,j}(P) - dt_i(k, j))^2, \quad (11)$$

where $dt_i(k, j)$ represents the element at the intersection of the k th row and j th column in DT_{*i*}.

2.2. Algorithm Idea. Due to the existence of the performance differences of the classifiers, even using the same classifier, the distinguishing capacity for different classes of data in a data set is different. When the classifier output reserves detailed information about a class, the DT algorithm's performance will be quite good. However, when the classifier is very sensitive to certain features of the input space, the result is a major change to certain information in the classes in the output space, and the performance of the algorithm is reduced to an obvious extent. For this problem, the VWDT algorithm measures the performance of various classifiers by a statistical vector and self-adaptively assigns weights for various classifiers according to the specific output status of the classifiers. The output of a classifier that has better performance occupies a larger proportion in the constructing process of the decision template, to acquire a more reliable decision template and improve the classification accuracy.

For certain samples, even being compared with the samples in the same class, they still have obvious differences. These samples are often the outliers of each class of samples, and their decision profiles are quite different from most of the samples. When they are rare, in the DT algorithm, the decision template acquired by the calculation is quite different from their decision profiles. Moreover, the DT algorithm evaluates only the similarity between the decision profile and decision template, which makes some samples from various classes misclassified due to having a small similarity or the fact of being at an overlapping region of multiple classes. With respect to this problem, the VWDT algorithm takes advantage of the information in various training samples that are around the test sample. It searches for several nearest-neighbors of the test sample and combines the training sample's information and the calculation of the decision template similarity together, thus avoiding the deviation that results from total dependence on the decision template to some extent and making the final class judgment more reliable.

2.3. Algorithm Steps. Firstly, a statistical vector is used to measure the performance of various classifiers, and the decision profile takes a weighed transform.

The classification error of the k th classifier C_k can be represented by an m -dimensional statistical vector V^k , as in

$$V^k = [n_1^k, n_2^k, \dots, n_m^k]^T, \quad (12)$$

where the element n_i^k , $i = 1, 2, \dots, m$, represents the number of training samples that are in the i th class that are correctly recognized by classifier C_k . Let the total number of training samples be N . The vector V^k is normalized by dividing

by N . At this time, the meaning of each element changes into its corresponding percentage. The normalized vector is represented as in

$$V^{k/N} = [n_1^{k/N}, n_2^{k/N}, \dots, n_m^{k/N}]^T. \quad (13)$$

Here, the reliability of the output vector of classifier C_k can be weighed by $n_h^{k/N}$, and the value of h should satisfy the following condition, as shown in

$$c_{k,h} = \max_{j=1}^m (c_{k,j}). \quad (14)$$

Thus, the reliable output of classifier C_k can be represented as in

$$C'_k(X) = n_h^{k/N} * C_k(X). \quad (15)$$

From (1), (15) is denoted as in

$$\begin{aligned} & [c'_{k,1}(X), c'_{k,2}(X), \dots, c'_{k,m}(X)] \\ & = n_h^{k/N} * [c_{k,1}(X), c_{k,2}(X), \dots, c_{k,m}(X)]. \end{aligned} \quad (16)$$

According to the reliable output vectors of all of the classifiers calculated by (16), the new decision profile $DP'(X)$ can be acquired as in

$$DP'(X) = \begin{bmatrix} c'_{1,1}(X) & \cdots & c'_{1,j}(X) & \cdots & c'_{1,m}(X) \\ \vdots & & \vdots & & \vdots \\ c'_{i,1}(X) & \cdots & c'_{i,j}(X) & \cdots & c'_{i,m}(X) \\ \vdots & & \vdots & & \vdots \\ c'_{l,1}(X) & \cdots & c'_{l,j}(X) & \cdots & c'_{l,m}(X) \end{bmatrix}. \quad (17)$$

Using $DP'(X)$ to replace $DP(X)$ and adopting the Euclidean distance as the similarity measure, the VWDT algorithm is described as follows.

2.3.1. Training Process

- (1) Use each classifier to classify the training samples, acquire each classifier's statistical vector V^k , $k = 1, 2, \dots, l$, and construct the set D that corresponds to the decision profile $DP'(X)$; from (10), the corresponding decision template DT'_i , $i = 1, 2, \dots, m$, can be acquired, as shown in

$$DT'_i = \frac{1}{N_z} \sum DP'(z_j), \quad z_j \in Z, \quad z_j \in \omega_i. \quad (18)$$

- (2) For a training sample $P \in Z$ and from (11), the square value of the Euclidean distance between the decision profile $DP'(P)$ and DT'_i can be acquired, as shown in

$$d_E(DP'(P), DT'_i) = \sum_{j=1}^m \sum_{k=1}^l (c'_{k,j}(P) - dt'_i(k, j))^2. \quad (19)$$

- (3) Calculate the possibility measure values of training sample P belonging to various classes, as shown in

$$d_i(P) = 1 - \frac{1}{l \times m} d_E(DP'(P), DT'_i). \quad (20)$$

If $d_K(P) = \max_{i=1}^m d_i(P)$, then it shall be deemed that $P \in \omega_K$.

- (4) Repeat steps (2)-(3) until all of the training samples are classified. According to the classification results, acquire a weighted decision profile set D_e ($D_e \subseteq D$) that corresponds to the training samples that are misclassified.
- (5) Set $D_e \neq \Phi$, calculate the average values of all of the elements in set D_w as an additional decision template, which shall be denoted as DT'_{m+1} , and use the set D_e to train a k -NN classifier C_{knn} .
- (6) Return all of the decision templates DT'_i , $i = 1, 2, \dots, m+1$, the statistical vector V^k , $k = 1, 2, \dots, l$, and the classifier C_{knn} .

2.3.2. Classification Process

- (1) For an input pattern X , use the statistical vectors that are acquired in the above training process to construct its decision profile $DP'(X)$.
- (2) Calculate the squared value of the Euclidean distance between $DP'(X)$ and DT'_i by (19), and, then, calculate the possibility measure value $d_i(X)$, $i = 1, 2, \dots, m+1$, of the various classes that the sample X belongs to by (20).
- (3) Judge the class of X :
 - (a) Let $d_K(X) = \max_{i=1}^{m+1} (d_i(X))$; if $K \leq m$, then it shall be deemed that $X \in \omega_K$.
 - (b) Otherwise, use C_{knn} to classify X , and let its measure layer output be $[c_1(X), c_2(X), \dots, c_m(X)]^T$ if $d_K(X) = \max_{i=1}^m (d_i(X) + c_i(X))$, $X \in \omega_K$.

The algorithm measures the classifier's performance by a statistical vector, which is acquired from the statistical data on the training sample set. Different weights are self-adaptively assigned to the classifiers' output vectors according to their judgments, which makes the decisions from different classifiers and the decisions of different classes from a classifier have different proportions in the decision templates. The weights can be acquired independently according to the prior information of the training samples and the classifiers' decisions on the current input sample. According to the statistical vector, if a classifier has better classification performance for a certain class of samples, then its output can be considered to be more reliable when it judges that the current input sample belongs to this class, and it is assigned a larger weight. If so, more reliable output has a higher proportion in the decision templates, which makes them more reliable. The training samples that are wrongly classified by the multiple

classifier system are used to construct an additional decision template. When an input sample is nearest to the center of this template, it can be reasonably considered that the sample is easily misclassified by the decision templates. At this time, the nearest-neighbor rule is adopted to make the judgment, which combines the information in the training samples near the input sample and the distance data between the input sample and various templates, which thus improves the accuracy.

3. Experimental Analyses

Currently, there are many practical classification algorithms, such as the nearest-neighbor algorithm, linear classifiers, the minimum distance algorithm, support vector machines, and artificial neural networks. The nearest-neighbor algorithm is very intuitive, wherein the basic idea is to find a sample from the training set that has the minimum distance to the object to be recognized and take the sample's class as the recognition result. The k -nearest-neighbor algorithm, as a typical representative of nonparametric algorithms, is used as the base classifier in this paper because it is simple and reliable. The experiments in this paper were realized with Matlab. During the programming process of the specific codes, the pattern recognition toolbox "PRTools4" [28] was used. In the experiments, the code of the k -nearest (k -NN) classifier is from the toolbox. To simplify the experiments, the parameter k is selected automatically and optimally by the leave-one-out method. The measure of the classifier's output is the posterior probability. To sufficiently test the performance and generality of the proposed algorithm, 15 data sets from the KDD'99, UCI, and ELENA databases were selected for the experiments. These databases are comparatively authoritative test data from related research fields, and the selected data sets show relatively good representation.

In the experiment results, if the sample data has separate training set and test set and the test result is unchanged, the statistical equation of the error rate is shown in

$$E = \frac{N_e}{N_t}, \quad (21)$$

where E represents the error rate, N_e represents the number of misclassified test samples, and N_t represents the number of test samples.

If the sample data has only one set and M -fold cross-validation is used, the test result is the average of the error rates of M validations, and the statistical equation of the average error rate is shown in

$$\bar{E} = \frac{1}{M} \sum_{i=1}^M E_i, \quad (22)$$

where \bar{E} represents the average error rate.

3.1. Using the KDD'99 Data. The samples are from the KDD'99 intrusion detection database (<http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>). The available database includes a large number of network connection patterns

TABLE 1: Description of the KDD data sets.

Data set name	Number of classes	Number of samples	Number of features
ftp5c	5	A 800	29
ftp2c	2	B 5212	
http4c	4	A 410	28
http2c	2	B 3380	

TABLE 2: Performance on the KDD data sets (in %).

Data set name	Error rate	
	DT	VWDT
ftp5c	5.794	5.411
ftp2c	1.055	1.036
http4c	3.047	0.237
http2c	0.237	0.148

that were extracted from TCP/IP packets. Each connection pattern is a 41-dimensional feature vector and belongs to one out of five classes (i.e., normal, R2L attacks, DoS attacks, U2R attacks, and probing attacks). The feature can be divided into three categories: intrinsic feature, traffic feature, and content feature.

The connection patterns related to the ftp service were selected from two data sets with corrected labels. Each pattern was processed by dimension reduction (these features whose values were always constants were abandoned) and quantification of the symbolic features and normalization (a linear transformation was applied to make each feature within the range $[0, 1]$). Training set A and test set B were obtained correspondingly. This group of data sets with five-class data was denoted by ftp5c. Combining the four attack classes as an abnormal class, a group of data sets with two-class data was obtained and denoted by ftp2c.

Through a similar process, the connection patterns that were related to the http service were also selected. Two groups of data sets were acquired, which were denoted by http4c and http2c. The former was four-class data, and the latter was two-class data. Because the number of http connection patterns was large, a restricted condition in the selection process was added: the duration was above zero. Table 1 is the description of the KDD data sets. Three k -NN classifiers were trained using a distinct category of features on each training set A and constituted the classifier set. The original test results of the DT and VWDT algorithms are shown in Table 2 and Figure 2 (ER represents the error rate).

It can be observed that on the four data sets, the performance of the VWDT algorithm is better than that of the DT algorithm. Especially on the multiple-class data sets of ftp5c and http4c, the error rate of the algorithm is reduced more obviously. The reason is that the more the classes that exist, the more obvious the difference that is in the classifier's performance to various classes, which is reflected in the decision template, and the reliability of the decision template is higher. At the same time, the difference between the misclassified samples and other samples is more

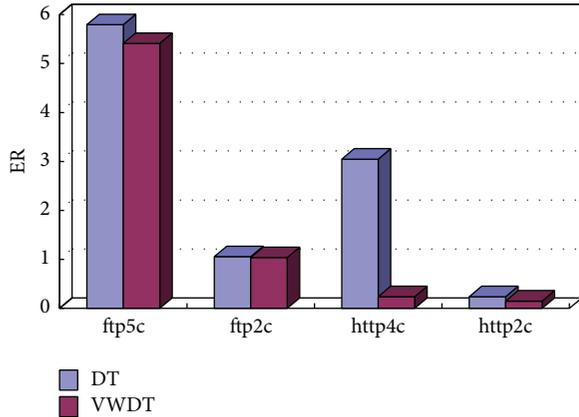


FIGURE 2: Performance comparison on the KDD data sets.

TABLE 3: Description of the UCI data sets.

Data set name	Number of samples	Number of features	Number of classes
Glass	214	9	6
Bupa	345	6	2
Pima	768	8	2
Thyroid	215	5	3
Wine	178	13	3
German	1000	24	2
Heart	270	13	2
Vehicle	846	18	4

obvious, and the complementarity is better by combining the information in the training samples around the input sample and the calculated similarity between the decision templates and the input sample's decision profile.

3.2. Using the UCI Data. A total of 8 data sets from the UCI machine learning database (<http://www.ics.uci.edu/~mllearn/MLRepository.html/>) were selected for experimental analysis. Due to the study demand for the algorithms, the selected data sets were all in digital form without errors or missing feature values. Normalization was performed on the selected data. The basic information about the selected data sets is in Table 3. The name of each data set is only the first word of its name in the original database.

For each data set, 6-fold cross-validation was used: the data set was divided into 6 subsets with similar size and distribution. Every time, a subset was selected as the test set, and the remainder comprised the training set. On each training subset, a k -NN classifier was trained. The corresponding 5 classifiers constituted the classifier set for the fusion. The test results are the average of 6 times (see Table 4 and Figure 3; AER represents the average error rate).

It can be observed that, of the 8 data sets, the VWDT algorithm is better than the DT algorithm on 7 of the data sets. On the data sets (thyroid, wine, and vehicle), which have more classes, the performances of the VWDT algorithm are improved more obviously. However, the data set of glass is

TABLE 4: Performance on the UCI data sets (in %).

Data set name	Average error rate	
	DT	VWDT
Glass	41.367	42.538
Bupa	24.751	20.526
Pima	26.042	23.828
Thyroid	14.484	3.320
Wine	6.825	2.857
German	27.564	24.154
Heart	17.407	15.926
Vehicle	33.394	26.195

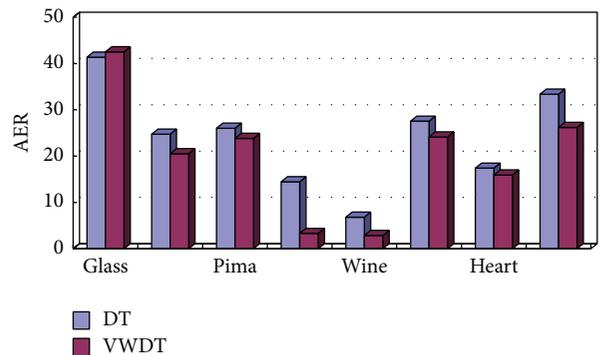


FIGURE 3: Performance comparison on the UCI data sets.

an exceptional instance, mainly because there are not many samples but lots of classes in the data set, and the distributions of the samples of the different classes are obviously varied. Under such a situation, taking the samples' decision profile average of a class as the representative of the samples of that class could result in a larger error, which is also proven by the DT algorithm's performance. For the VWDT algorithm, an additional decision template is divided, which further lowers the performance. However, it can be observed that, on the data set of glass, compared with the DT algorithm, the error rate increase of the VWDT algorithm is not obvious (2.83%). On the other data sets, the minimum error rate reduction of this algorithm is 8.5% (data set pima).

3.3. Using the ELENA Data. A total of 3 data sets from the ELENA database (<https://www.elena.ucl.ac.be/neural-nets/Research/Projects/ELENA/elena.htm>) were selected for experimental analysis. Normalization was also made for the selected data. The basic information of the selected data sets is in Table 5. For each data set, 6-fold cross-validation was also used, and the test results are in Table 6 and Figure 4.

It can also be observed that the VWDT algorithm is better than the DT algorithm. On the data set of satimage, which has more classes, the performance of the VWDT algorithm is improved more obviously.

The experiment results show that the VWDT algorithm is effective for improving the classification performance. It must be noted that k -NN classifiers are adopted in the experiments of this paper. When the classifier set is composed

TABLE 5: Description of the ELENA data sets.

Data set name	Number of samples	Number of features	Number of classes
Clouds	5000	2	2
Satimage	6435	36	6
Phoneme	5404	5	2

TABLE 6: Performance on the ELENA data sets (in %).

Data set name	Average error rate	
	DT	VWDT
Clouds	10.808	7.205
Satimage	9.925	4.659
Phoneme	8.315	5.387

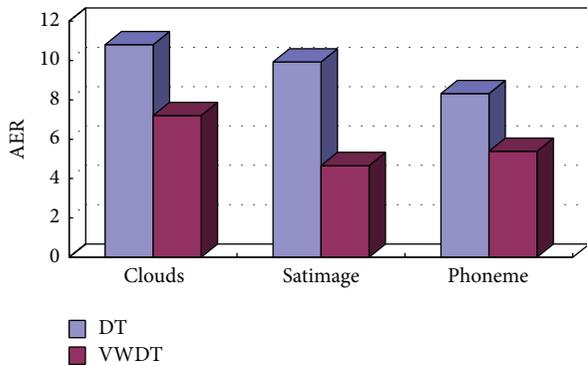


FIGURE 4: Performance comparison on the ELENA data sets.

of classifiers of a different type, because the physical meanings and measures of the classifiers' outputs are different, the performances of both the DT and VWDT algorithms could be obviously impacted. At this time, the outputs of all of the classifiers should be transformed into unified reliability.

4. Practical Application

4.1. Application Background. Age-related cataract is the leading cause of blindness in the world. Now, there are 25 million people with cataract-induced blindness, and, among them, almost 2.5 million are in China and account for 10 percent. Because the population is aging, there would be more than one million cataract patients yearly in China. Until now, an operation is the only effective method to cure cataract blindness. However, a cataract operation could not fully treat the newly increased patients, and, thus, the future appears to be rather gloomy for China's cataract blindness. The ultrasonic phacoemulsification operation is considered to be the main treatment for cataracts and is widely accepted because of its small cut, short operation time, including no suture, no need of being in hospital, and no limitation for any activities, and fast recovery of eyesight.

Before performing the operation, strict training shall be conducted for the doctor in charge of a case, to make the

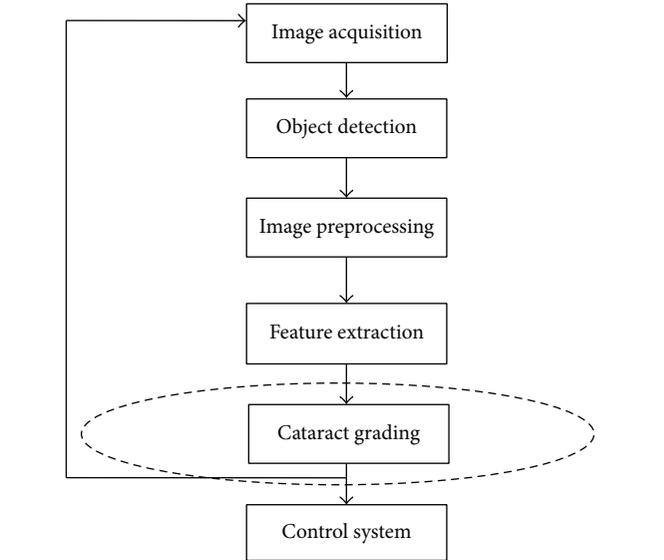


FIGURE 5: Module structure of the cataract recognition system.

doctor very familiar with the cataract grading. However, because the hardness grades of the cataracts are different for different individuals, the problem of “how to correctly implement a proper oscillation frequency according to a different lens nucleus” requires long-term practice for the operation, to accumulate rich operation experience toward reaching the best operation outcome. The operation is impossible to promote over a large area because of its low automatization, operation difficulty, and long learning cycle. That is the vital reason why this operation cannot be popularized.

Along with the development of information technology, computer-aided healthcare by integrating medical devices and healthcare information systems to improve healthcare quality and productivity is receiving more and more attention. In particular, a computer-aided healthcare system could provide a solution in the developing areas where medical resources are scarce [29]. Image processing and pattern recognition technologies can be applied to the computer-aided cataract recognition system of the ultrasonic emulsification apparatus, to automatically recognize the hardness grade of the cataract and, thus, to lower the operation's difficulty level, shorten the study period, and improve the safety of the operation.

Such a cataract recognition system is mainly composed of five modules, as shown in Figure 5 (the position of the research content of this paper is in the dash-dotted ellipse). The image acquisition module has a real-time operation video via camera on the operating table, achieves single-frame images by sampling at a certain frequency, and switches them into formats that can be processed by the system; the object detection module detects the emulsification pinhead from the image and locates the image area for recognition, which is the first key problem of the system; the image preprocessing module conducts the preprocessing for the image area, such as filtering and denoising, to eliminate the noise; and the feature extraction module extracts the features of the image

area. The extracted features should effectively represent the tissue and reserve the original information of the emulsification pinhead, cataract tissues, and normal tissues as much as possible; the cataract grading module adopts a certain classification algorithm according to the extracted features to recognize the tissue and submits the result to the control system. Finding a suitable recognition algorithm to ensure accurate judgment for the cataract grading is the second key problem of the system. If the recognition result is a cataract with certain hardness, the control system will emit different signals to control the ultrasonic frequency according to different hardness grades. Otherwise, no ultrasonic signal can be sent out.

4.2. Application Description. The cataract grading is mainly based on Emery and Little's grade standard, which judges the color of the lens nucleus to grade its hardness according to the examination results under the slit-lamp. Human eyes can directly distinguish normal tissue from diseased tissue and recognize different hardness of the lens nucleus based on the appearance of different colors. Considering the practical situation of lens nucleus recognition, color is selected as the image feature for the system, and the RGB color model is used in this paper.

The image of the area near the emulsification pinhead was partitioned into $m \times n$ pieces, and the color value of a piece was the average of all of the pixels in the piece. This paper selected $m = 15$ and $n = 5$ as the parameters. Thus, each image was partitioned into 75 pieces, and each piece had a similar number of pixels. The average R, G, and B value of all of the pixels in a piece were separately calculated as the feature value of that piece. In other words, every image corresponds to a 225-dimensional feature vector.

The software development environment of the recognition system was Visual C++. On the training set, three k -NN classifiers were trained separately by using R, G, and B features; that is, the input of each classifier is a feature vector of 75 dimensions. The decisions of three single feature classifiers were fused by the DT and VWDT algorithms. For a comparison, a k -NN classifier was trained by using the RGB feature; that is, the input of the classifier is a 225-dimensional feature vector. To test the effect of the parameter k on the results, the parameter is selected to be $k = 3, 5, 8$. In the VWDT algorithm, the parameter k of classifier C_{knn} is selected to be $k = 1$.

4.3. Application Effect. The cataract images were intercepted from the operation videos provided by Beijing Tongren Hospital. The size of the image had three categories: 20×20 , 30×30 , and 40×40 . The training image set had 649 images, and the test image set had 647 images. Each set had 6 types of images, normal tissues, and cataracts with I to V hardness grades. The class of each image was confirmed by the ophthalmologists from the hospital. The numeric distribution of the different classes of images is shown in Table 7. The original application results are shown in Table 8, and the average error rates (AER) are shown in Figure 6.

It can be observed that the classifiers that use the R feature have the highest error rates. The classifiers that use the G

TABLE 7: Number distribution of the different classes of images.

Class of the tissue	Number of training images	Number of test images
Normal	95	91
I	293	296
II	184	183
III	37	33
IV	28	30
V	12	14

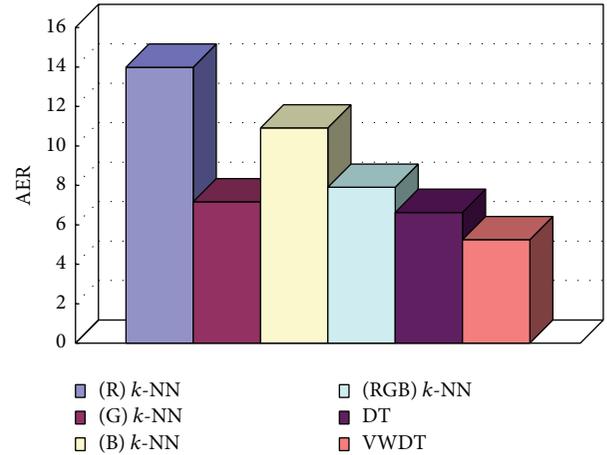


FIGURE 6: Comparison of the average error rates.

feature have the lowest error rate among the single color feature classifiers, which are slightly lower than the classifiers that use the RGB feature. The error rate of the VWDT algorithm is lower than that of all of the single classifiers and the DT algorithm, which shows the effectiveness of the algorithm again. Among the images that were misclassified by all of the algorithms, there is no normal tissue, and more than 80% of them are divided into lower grades. This result is ideal, and it will not cause adverse effects on the patients. The doctor must only properly increase the output signal intensity according to experience. The VWDT algorithm has the highest average recognition rate, 94.75%, which reaches the actual demand of the cataract ultrasonic phacoemulsification operation.

When the parameter k takes on different values, the error rates change accordingly but not very obviously (as shown in Figure 7). When $k = 5$, the error rates of all of the algorithms (except for the VWDT algorithm, when $k = 3$) are slightly lower. Thus, in practical applications of the VWDT algorithm, the parameter k can be determined by experience (a fifteenth of the number of training samples is a suggestion).

5. Conclusions

In this paper, a multiple classifier fusion algorithm is proposed and applied for recognition of the hardness of a cataract lens nucleus. The algorithm considers the difference in the classifiers' performances and makes full use of the training

TABLE 8: Application results (in %).

Feature	Method	Error rate	Method	Error rate	Method	Error rate
R	3-NN	13.75	5-NN	13.75	8-NN	14.47
G	3-NN	7.48	5-NN	6.88	8-NN	7.15
B	3-NN	10.96	5-NN	10.12	8-NN	11.65
RGB	3-NN	8.36	5-NN	6.98	8-NN	8.36
	DT	6.72	DT	6.44	DT	6.72
	VWDT	4.95	VWDT	5.26	VWDT	5.54

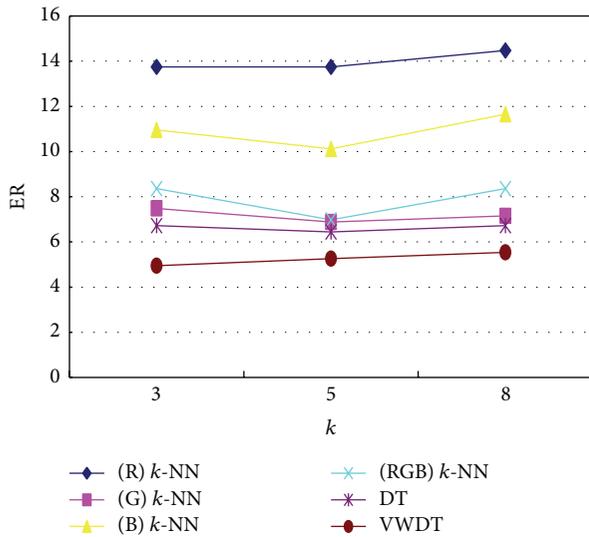


FIGURE 7: Comparison of the error rates.

sample information. It uses a statistical vector to measure the classifier's performance, and an appropriate weight is assigned to each classifier according to the reliability of its output. Different color components of cataract images were used to train the corresponding classifiers, and the decisions of the classifiers using a single color feature were fused to judge the hardness. The experimental results show that the algorithm can achieve a high classification performance that meets the actual demands of the cataract ultrasonic phacoemulsification operation.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was financially supported by the Key Scientific Research Project of Colleges and Universities of Henan Province (no. 15A520017), Soft Science Research Project of Henan Province (no. 152400410513), the Doctoral Fund of Henan Polytechnic University (nos. B2010-046 and B2013-040), the Basic Research Development Program of Henan Province (132300410333), and Key Science and Technology Project of Henan Province (152102210102).

References

- [1] S.-Y. Liang, D.-Q. Han, and C.-Z. Han, "A novel diversity measure based on geometric relationship and its application to design of multiple classifier systems," *Acta Automatica Sinica*, vol. 40, no. 3, pp. 449–458, 2014.
- [2] D.-H. Zhou and Y.-Y. Hu, "Fault diagnosis techniques for dynamic systems," *Acta Automatica Sinica*, vol. 35, no. 6, pp. 748–758, 2009.
- [3] J.-T. Huang, M.-H. Wang, W.-J. Li, and B. Gu, "Multiple classifier fault diagnosis system based on dynamic weight," *Acta Electronica Sinica*, vol. 40, no. 4, pp. 734–738, 2012.
- [4] L. Li and H. Ji, "Radar emitter recognition based on multiple classifier fusion," *Journal of Data Acquisition and Processing*, vol. 25, no. 3, pp. 396–400, 2010.
- [5] L. Sun, C.-Z. Han, J.-J. Shen, and N. Dai, "Generalized rough set method for ensemble feature selection and multiple classifier fusion," *Acta Automatica Sinica*, vol. 34, no. 3, pp. 298–304, 2008.
- [6] A. A. López-Caloca, "Data fusion approach for employing multiple classifiers to improve lake shoreline analysis," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, E. Bayro-Corrochano and E. Hancock, Eds., vol. 8827 of *Lecture Notes in Computer Science*, pp. 1022–1029, Springer, Berlin, Germany, 2014.
- [7] M. Woźniak, M. Graña, and E. Corchado, "A survey of multiple classifier systems as hybrid systems," *Information Fusion*, vol. 16, no. 1, pp. 3–17, 2014.
- [8] V. G. Giannoglou and J. B. Theocharis, "Decision fusion of multiple classifiers for coronary plaque characterization from IVUS images," *International Journal on Artificial Intelligence Tools*, vol. 23, no. 3, Article ID 1460005, 2014.
- [9] M. De-La-Torre, E. Granger, P. V. W. Radtke, R. Sabourin, and D. O. Gorodnichy, "Partially-supervised learning from facial trajectories for face recognition in video surveillance," *Information Fusion*, vol. 24, pp. 31–53, 2015.
- [10] M. Panda, A. Abraham, and M. R. Patra, "Hybrid intelligent systems for detecting network intrusions," *Security and Communication Networks*, vol. 8, no. 16, pp. 2741–2749, 2015.
- [11] H. Altınçay, "On naive Bayesian fusion of dependent classifiers," *Pattern Recognition Letters*, vol. 26, no. 15, pp. 2463–2473, 2005.
- [12] A. Dainotti, A. Pescapé, C. Sansone, and A. Quintavalle, "Using a behaviour knowledge space approach for detecting unknown IP traffic flows," in *Multiple Classifier Systems*, vol. 6713 of *Lecture Notes in Computer Science*, pp. 360–369, Springer, Berlin, Germany, 2011.
- [13] J. Zou, Q. Ji, and G. Nagy, "A comparative study of local matching approach for face recognition," *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2617–2628, 2007.

- [14] M. A. Zahid and H. de Swart, "The borda majority count," *Information Sciences*, vol. 295, pp. 429–440, 2015.
- [15] J. P. Hwang, S. Park, and E. Kim, "A new weighted approach to imbalanced data classification problem via support vector machine with quadratic cost function," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8580–8585, 2011.
- [16] L. Zhang and Y. Dong, "Research on diagnosis of AC engine wear fault based on support vector machine and information fusion," *Journal of Computers*, vol. 7, no. 9, pp. 2292–2297, 2012.
- [17] Z. Wenbo, J. Hongbing, and W. Lei, "Adaptive weighted feature fusion classification method," *Systems Engineering and Electronics*, vol. 35, no. 6, pp. 1133–1137, 2013.
- [18] A. C. Ben Abdallah, H. Frigui, and P. Gader, "Adaptive local fusion with neural networks," in *Artificial Neural Networks-ICANN 2010*, vol. 6353 of *Lecture Notes in Computer Science*, pp. 486–491, Springer, Berlin, Germany, 2010.
- [19] Y. Yang, D. Han, and C. Han, "A novel diversity measure of multiple classifier systems based on distance of evidence," *Acta Aeronautica et Astronautica Sinica*, vol. 33, no. 6, pp. 1093–1099, 2012.
- [20] L. I. Kuncheva, J. C. Bezdek, and R. P. W. Duin, "Decision templates for multiple classifier fusion: an experimental comparison," *Pattern Recognition*, vol. 34, no. 2, pp. 299–314, 2001.
- [21] R. Ebrahimpour and S. Hamed, "Hand written digit recognition by multiple classifier fusion based on decision templates approach," *World Academy of Science, Engineering and Technology*, vol. 33, pp. 560–565, 2009.
- [22] W. Chen, S.-W. Zhang, Y.-M. Cheng, and Q. Pan, "Prediction of protein-protein interaction types using the decision templates based on multiple classifier fusion," *Mathematical and Computer Modelling*, vol. 52, no. 11-12, pp. 2075–2084, 2010.
- [23] M. Re and G. Valentini, "Integration of heterogeneous data sources for gene function prediction using decision templates and ensembles of learning machines," *Neurocomputing*, vol. 73, no. 7–9, pp. 1533–1537, 2010.
- [24] M. S. Haghghi, A. Vahedian, and H. S. Yazdi, "Extended decision template presentation for combining classifiers," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8414–8418, 2011.
- [25] R. Ebrahimpour, R. D. Vahid, and B. M. Nezhad, "Decision templates with gradient based features for Farsi handwritten word recognition," *International Journal of International Journal of Hybrid Information Technology*, vol. 4, no. 1, pp. 1–12, 2011.
- [26] B. Bigdeli, F. Samadzadegan, and P. Reinartz, "Feature grouping-based multiple fuzzy classifier system for fusion of hyperspectral and LIDAR data," *Journal of Applied Remote Sensing*, vol. 8, no. 1, Article ID 083509, 2014.
- [27] L. Valet, E. Ramasso, and S. Teyssier, "Quality evaluation of insulating parts by fusion of classifiers issued from tomographic images," *Information Fusion*, vol. 9, no. 2, pp. 211–222, 2008.
- [28] P. W. R. Duin, P. Juszczak, P. Paclik et al., *PRTools4, A Matlab Toolbox for Pattern Recognition*, Delft University of Technology, Delft, Netherlands, 2007.
- [29] L. Guo, J.-J. Yang, L. Peng, J. Li, and Q. Liang, "A computer-aided healthcare system for cataract classification and grading based on fundus image analysis," *Computers in Industry*, vol. 69, pp. 72–80, 2015.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

