Neurocomputing 393 (2020) 214–222

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A hybrid improved kernel LDA and PNN algorithm for efficient face recognition

Aijia Ouyang[a], Yanmin Liu[b,*], Shengyu Pei[c], Xuyu Peng[a], Minghua He[d], Qian Wang[a]

[a] School of Information Engineering, Zunyi Normal University, Zunyi, Guizhou 563006, China
[b] School of Mathematics, Zunyi Normal University, Zunyi, Guizhou 563006, China
[c] Central South University, School of Automation, Changsha Hunan 410083 P.R. China
[d] School of Business, Jinggangshan University, Xueyuan Road 28, Ji'an, Jiangxi, China

ARTICLE INFO

Article history:
Received 20 November 2017
Revised 12 January 2019
Accepted 15 January 2019
Available online 25 July 2019

Keywords:
Dimension reduction
Face recognition
Feature extraction
Kernel discriminant analysis
Probabilistic neural network

ABSTRACT

This paper proposes a hybrid approach to face recognition based on a combination of probabilistic neural networks (PNNs) and improved kernel linear discriminant analysis (IKLDA). The dimensions of a sample's features are first of all reduced, whilst retaining its relevant information, A PNN method is then adopted to solve face recognition problems. The proposed IKLDA+PNN method not only improves the overall computing efficiency, but also its precision. Face recognition experiments conducted on the ORL, YALE and AR datasets, which contain a wide variety of facial expressions, facial details, and degrees of scale, were used to validate the feasibility of the IKLDA+PNN method. The results showed that it can obtain an average recognition accuracy of 97.22%, 83.8% and 99.12%, across the three datasets, respectively.

1. Introduction

Image recognition technology [1] is used across a range of domains for a variety of applications including target recognition, identity recognition and facial expression recognition. In the field of face recognition, a lot of research has been devoted to the development of effective programs for managing key problems such as image acquisition and classification [2]. One of the oldest, simplest, but also most important of these is the nearest neighbor algorithm [3,4], which is commonly employed for face recognition, etc. The nearest neighbor method involves first of all finding the closest training samples to a test sample. The latest research methods for Face Recognition include fully convolutional neural network [5], deep neural network [6] and so on.

The test sample then falls into the same class that the training samples belong to. As the nearest sample is the most similar one, a nearest neighbor method is used to analyze and explore which training sample is most similar to the test sample [7]. The distances between each training sample and the test sample are calculated successively using a traditional nearest neighbor algorithm, while the similarity between different training samples and their underlying relationship is ignored. However, if the two training samples closest to a test sample are very similar, the situation becomes difficult for a nearest neighbor algorithm to handle [8].

Various methods have been used for face recognition over the past decades. The motivation for the continuous work on face recognition is to obtain a method able to recognize different angles and poses of faces accurately and efficiently. The proposed algorithm is a complement to the traditional methods.

First of all, we propose an IKLDA algorithm, secondly the IKLDA algorithm is employed to reduce the dimensions of sample vector, and then, we propose an IKLDA-PNN hybrid algorithm, finally, the IKLDA-PNN algorithm is applied to address face recognition because PNN has high classification accuracy, the combination of IKLDA and PNN algorithms is very efficient and effective. The average accuracy achieved by the algorithm proposed in this paper for the ORL human faces database was much better than the classic classification algorithms, PCA+PNN, LDA+PNN, KPCA+PNN and KLDA+PNN. The IKLDA+PNN algorithm was faster in solving the face recognition problems for the ORL and Yale databases than the other algorithms. The classification performance of our proposed algorithm was better than some of the other algorithms for the AR database.

We aim to overcome the drawbacks of LDA and PNN, in addition to the severe degradation in performance of the PNN algorithm under degraded conditions, thus forming the underlying motivation of the work presented in this paper.

* Corresponding author.
E-mail address: lym@zync.edu.cn (Y. Liu).

https://doi.org/10.1016/j.neucom.2019.01.117

The main contributions of the study presented here are as follows:

1. We propose a new effective dimensionality reduction strategy, modify the total scatter matrix *St*, and then formulate a new equation of optimization objective.
2. We update the objective function in order to solve feature problem. $KWK = KK$, the solution of eigenvectors, $KWK\alpha = KWy = K\lambda y = \lambda Ky = \lambda KK\alpha$.
3. Find the optimal value U with the method of fuzzy c-means cluster analysis.
4. We design a hybrid model of IKLDA+PNN, where IKLDA maps multidimensional face data to the appropriate dimension space and completes the classification by PNN.

Over the remainder of the paper, Section 2 will examine related work regarding LDA and PNN, Section 3 will present the execution steps of our proposed hybrid face recognition algorithm based on IKLDA and PNN and Section 4 will provide the results of a simulation experiment and discuss in detail the performance of the IKLDA+PNN algorithm. In Section 5 we will give our concluding remarks.

2. Related work

The main idea of an LDA algorithm is to project a high-dimensional pattern sample into an optimal set of discriminant vectors so as to extract classification information and reducing the dimensions of the feature space. After the process of projection, it should be possible to guarantee that pattern sample has the maximum between-class distance and the minimum within-class distance in the new subspace. In other words, the pattern sample should have the best separability in this space [9]. This provides an effective approach to feature extraction. By using this method, the between-class scatter matrix of a pattern sample can be maximized after projection and the within-class scatter matrix kept to a minimum.

LDA is a classic statistical method that has been widely used in the fields of patient disease prognosis, brand positioning, product management and market research and across areas as diverse as cost minimization [10], face recognition, and machine learning. LDA can be employed for data classification but, first of all, it needs to be trained with pre-classified data that to establish a discrimination model. Thus, LDA needs to be considered a supervised learning algorithm. The basic idea of LDA is projection. n-dimensional data is projected into a low-dimensional space and groups are kept as far apart as possible, i.e.. they are provided with maximum separability within the space. The ideal, here, is to have a minimal within-class distance and a maximal between-class distance in any new subspace. The goal is to work out a vector that has these maximal and minimal characteristics so that it can form the basis of a discrimination model [11].

A key use of LDA is for the reduction of dimensions in classification problems. As the mean vector and variance matrix of each class needs to be used, it can draw upon the parameters used in Finite Element (FE) approaches. As a rule, a within-class scatter matrix, between-class scatter matrix and mixture scatter matrix are used to express the standard degree of class scatter.

LDA, also known as Fisher linear discriminant (FLD), is a classic resource for solving pattern recognition problems. It was first introduced into the field of artificial intelligence and pattern recognition by Belhumeur, where it is seen to be a powerful algorithm for solving problems of face recognition [12]. LDA is particularly well-known as a high-dimensional reduction algorithm that can be applied in a wide variety of different application areas. Yu and Lu [13], for instance, have used it to develop a manifold learning-based wafer map defect detection and recognition system. Zhang et.al. [14], meanwhile, have proposed a novel model for sparse uncorrelated LDA. In [15], LDA is used to diagnose bearing damage, with this being further developed in [16] through the use of a novel MLDA [17] algorithm.

In recent years, LDA has often been adopted to solve dimension reduction problems in supervised classification [18]. Here, a mean vector and variance matrix for each class is used to express within-class [9], between-class and mixture scatter matrixes. Using the Fisher standard, the ratio between the between-class and within-class scatter matrixes is maximized in a low-dimensional space. It is on the basis of these characteristics that a hybrid algorithm for face recognition will be developed in this study. The proposed technique not only saves overall computing time, but also improves computing efficiency and precision.

PNN [19] was proposed by Specht in 1990. An exponential function used to replace the commonly used sigmoid activation function in neural network. And then the neural network used to work out nonlinear decision boundaries is created, and gradually, the decision boundaries tend to get closer to the optimum Bayesian decision surface [20]. In[21], an incremental learning technique by combing PNN algorithm and adjustable fuzzy clustering method is presented. In [22], a PNN method is adopted for estimating the health of Li-ion batteries [23]. In [24], a supervised classification method is adopted in the study. In [25], the authors propose a PNN algorithm designed on the basis of time-series DCA. In [26], a class of novel algorithms for the selection and adaptation of the smoothing parameter of the PNN algorithm is put forward.

A probabilistic neural network (PNN) is a kind of neural network that is commonly employed to solve problems of pattern classification. PNN has a statistical basis and is equivalent to a Bayes optimal classifier in terms of classification function. Indeed, it can be considered to be a kind of parallel algorithm evolved from minimal Bayes risk standards. Unlike a traditional multi-layer feedforward network, which requires a BP algorithm for back-propagation computation,it is a fully forward computational process. It can be trained very easily, is resistant to local convergence and has high classification accuracy. No matter how complicated the classification problem, an optimal solution can be obtained under the Bayes criterion as long as there is enough training data [27].

3. The proposed algorithm (IKLDA+PNN)

In this section, we proposed an improved sparse kernel linear discriminant analysis strategy. To get a proper Φ, we define an inner product function (known as a reproducing kernel Hilbert space (RKHS)) in F. This can be described as:

$$< \Phi(\mathrm{X}), \Phi(\mathrm{Y}) >= \mathrm{K}(\mathrm{X}, \mathrm{Y}), \tag{1}$$

where K(.,.) is a positive semi-definite kernel function. There are some popular frequently-used kernel functions, such as the Gaussian kernel function $\mathrm{K}(\mathrm{x}, \mathrm{y}) = \exp(-||\mathrm{x}-y||^2/2\sigma^2)$, the polynomial kernel function $\mathrm{K}(\mathrm{x}, \mathrm{y}) = (1+\mathrm{x}^T y)^d$, and the Sigmoid kernel function $\mathrm{K}(\mathrm{x}, \mathrm{y}) = \tanh(\mathrm{x}^T y + \alpha)$.

S_b^Φ, S_w^Φ and S_t^Φ are, respectively, the between-class scatter matrix, the within-class scatter matrix and the total scatter matrix in the feature space. They can be defined as:

$$S_b^\Phi = \sum_{k=1}^{c} m_k\left(\mu_\Phi^{(k)} - \mu_\Phi\right)\left(\mu_\Phi^{(k)} - \mu_\Phi\right)^T$$

$$S_w^\Phi = \sum_{k=1}^{c}\left(\sum_{i=1}^{m_k}\left(\Phi(\mathrm{X}_i^{(k)}) - \mu_\Phi^{(k)}\right)\left(\Phi(\mathrm{X}_i^{(k)}) - \mu_\Phi^{(k)}\right)^T\right)$$

$$S_t^\Phi = \sum_{i=1}^{m}(\Phi(\mathrm{X}_i) - \mu_\Phi)(\Phi(\mathrm{X}_i) - \mu_\Phi)^T, \tag{2}$$

where, $\mu_\Phi^{(k)}$ and μ_Φ are, respectively, the centroid of the kth class and the total centroid.

Hence, we can revise the second formula S_b^Φ as follows:

$$S_b^\Phi = \sum_{k=1}^{c} \frac{\sum_{i=1}^{c} m_i}{c} \left(\mu_\Phi^{(k)} - \mu_\Phi\right)\left(\mu_\Phi^{(k)} - \mu_\Phi\right)^T. \tag{3}$$

Here we set v to be the projective function in feature space. The corresponding objective function is as follows:

$$\nu_{opt} = \arg\max \frac{\nu^T S_b^\Phi \nu}{\nu^T S_t^\Phi \nu}. \tag{4}$$

This can be solved as the eigenproblem $[S_b^\Phi \nu = \lambda S_t^\Phi \nu]$. As the eigenvector is a linear combination of the function $\Phi(\mathrm{X}_i)$, there is a coefficient α_i, that can be made to satisfy the equation $\nu = \sum_{i=1}^{m} \alpha_i \Phi(\mathrm{X}_i)$.This can be proved using in the following:

$$\alpha_{opt} = \arg\max \frac{\alpha^T KWK\alpha}{\alpha^T KK\alpha}, \tag{5}$$

in which the corresponding eigen-problem is $KWK\alpha = \lambda KK\alpha$. Here, K is the eigen-vector $(K_{ij} = \mathrm{K}(\mathrm{X}_i, \mathrm{X}_j))$ and W is defined as follows:

$$W_{ij} = \begin{cases} 1/m_k, & \text{if } \mathrm{x}_i \text{ and } \mathrm{x}_j \text{ both belong to the } k\text{−th class} \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Each eigen-vector α corresponds to a projective function V in feature space. For every datapoint x, it can satisfy the following formula:

$$\begin{aligned} < \nu, \Phi(\mathrm{X}) > &= \sum_{i=1}^{m} \alpha_i < \Phi(\mathrm{X}_i), \Phi(\mathrm{X}) > \\ &= \sum_{i=1}^{m} \alpha_i \mathrm{K}(\mathrm{X}_i, \mathrm{X}) \\ &= \alpha^T K(:, \mathrm{X}), \end{aligned} \tag{7}$$

where, $K(:, \mathrm{X}) = [K(\mathrm{X}_1, \mathrm{X}), \ldots, K(\mathrm{X}_m, \mathrm{X})]^T$ and $\{\alpha_1, \ldots, \alpha_{c-1}\}$ are $c-1$ eigen-vectors of the eigen-problem with a corresponding nonzero eigen-value. The transformation matrix $\theta = [\alpha_1, \ldots, \alpha_{c-1}]$ is an $m \times (c-1)$ matrix and the datapoint x can be embedded into the $c-1$ dimension subspace thus, $\mathrm{x} \to \mathrm{z} = \theta^T K(:, \mathrm{x})$.

Here, we can set y as an eigenvector of the eigenproblem $Wy = \lambda y$ with a corresponding eigenvalue. In order to solve the eigenproblem $KWK\alpha = \lambda KK\alpha$, if it is possible to satisfy $K\alpha = y$, y will be an eigenvector and we can import the eigen-value $K\alpha = y$. The equation can now be solved in the following way:$KWK\alpha = KWy = K\lambda y = \lambda Ky = \lambda KK\alpha$; with α being an eigen-vector of the eigenproblem with a corresponding eigen-value of λ. In order to solve the above-mentioned problems, we first need to solve the eigenproblem $Wy = \lambda y$ and get y. We can initially solve any α that satisfies $K\alpha = y$ by saying that $\alpha = K^{-1}y$. However, because K is a singular matrix, the original formula now needs to be amended as follows: $(K + \delta I)\alpha = y$; where I is an identity matrix and δ $(\delta \geq 0)$ is a regularization parameter.

In this study, we are putting forward a method for solving the problem of face recognition by combining an improved sparse KLDA and PNN algorithm. The IKLDA+PNN algorithm is shown in detail in Algorithm 1 (RBF is the abbreviation of Radial Basis Function):

Algorithm 1 The IKLDA+PNN algorithm.

Require: random matrix $U = [uij]1 \leq i \leq L, 1 \leq j \leq N$,the transmission parameter of RBF initialization $sp = .15$,data matrix x_train, x_test, labels y_train_label

Ensure: Classifier f, $y\_test\_predict$.

$Uopt = fuzzy\_cmeans(x\_train, y\_train\_label)$;
$[eigvector, eigvalue] = SKLDA(x\_train)$;
$[x\_train\_new, x\_test\_new] = SKLDA.reduce\_dimension(x\_train, x\_test)$;
construct a classifier f using PNN;
$y\_test\_predict = f(x\_test\_new)$;
return Classifier f, $y\_test\_predict$.

4. Experiment results and discussion

Although they offer time efficiency, there are some problems with using a nearest neighbor classifier (NN for short) and nearest centroid classifier. To get around these problems, we are suggesting the use of a probabilistic neural network (PNN). In this section we present the results of an experiment testing the viability of the approach, followed by a discussion of the outcomes.

The hybrid algorithm was used to carry out a simulation experiment that drew upon several standard face databases [28].

The parameters of the experimental platform are as follows: an Intel Core i5-2410 laptop with a 2.30 GHz processor and 4 GB of RAM. The programming language was Matlab2012a. In addition to above settings, we also experimented with using a support vector machine (SVM) Matlab tool in the same environment. We conducted a series of experiments using different methods on the same device: Principal component analysis (PCA)+PNN; linear discriminant analysis (LDA)+PNN; kernel principal component analysis (KPCA)+PNN; kernel linear discriminant analysis (KLDA)+PNN; our proposed IKLDA+PNN; and the SVM.

The ORL, Yale and AR face databases were used as the basis of our experiments. The ORL database contains 400 human face images across 40 classes. The images cover all kinds of facial expressions (including smiling/not smiling, eyes open/eyes closed) and a rich mixture of facial details: some of them are leaning backwards or forwards or are rotated by 20^o. Each image is 32×32 pixels. The faces in the Yale database include normal face images as well as images expressing sadness or happiness, surprise, and where people are asleep or blinking. All of the images were captured under different lighting conditions. In some of the images, people are wearing glasses. From the large-scale AR database of human faces, we selected 1680 grayscale images covering 120 classes, with 14 images per class. For the ORL, Yale and AR databases, if we suppose that s samples are selected from a sample set in which each class contains n samples as training samples, then there could be$C_q^p = \frac{p(p-1)\ldots(p-q+1)}{q(q-1)\ldots 1}$ kinds of possible combinations. The same combination was used to judge the training samples and the test samples, so there could be C_n^s training samples. Table 1 shows how many training samples were selected for each class from the ORL, Yale and AR databases. A part of samples of two classes in ORL database is shown in Fig. 1, some of the samples for one of the classes in the AR database are shown in Fig. 2.

The IKLDA+PNN algorithm proposed in this paper was compared with some other well-known algorithms: PNN [29,30] PCA [31–34], LDA [32,35], KLDA,KPCA, nearest neighbor classifier (NNC) [36], center-based nearest neighbor classifier (CBNNC) [36] and near-neighbor list (NNL) [37]. The comparative results are shown in the following table:

The experimental simulation program selected images from the standard human faces databases using the parameter settings proposed in [38]. The transmission parameter of initialization radial basis function was $sp = .1$. The results comparing the algorithm proposed in this paper with the other algorithms are shown in Tables 2–8. A diagram showing our proposed algorithm's recognition rate for the ORL database is shown in Fig. 3. The recognition

Table 1
Number of training samples.

Training sample number	2	3	4	5	6	7	8
ORL	45	120	210	252	210	120	45
Yale	55	165	330	462	462	330	165
AR	91	364	1001	2002	3003	3432	3003

Table 2
Comparison of the average accuracy between the proposed algorithm and some other algorithms for the ORL database.

training number	PCA+PNN AA (%)	LDA+PNN AA (%)	KPCA+PNN AA (%)	KLDA+PNN AA (%)	IKLDA+PNN AA (%)	SVM AA (%)
2	64.66	78.56	64.64	79.32	77.96	66.62
3	71.23	86.04	69.14	87.05	86.83	72.52
4	75.23	90.20	73.23	91.30	91.44	75.77
5	77.61	92.62	75.89	93.73	93.95	78.47
6	79.29	94.14	78.02	95.24	95.43	80.48
7	80.78	95.08	79.64	96.28	96.35	82.19
8	81.92	95.72	81.31	96.89	97.22	83.00

Table 3
Comparison of the standard deviation between the proposed algorithm and some other algorithms for the ORL database.

training number	PCA+PNN std	LDA+PNN std	KPCA+PNN std	KLDA+PNN std	IKLDA+PNN std	SVM std
2	3.95	2.96	3.96	2.87	2.96	3.76
3	3.80	2.22	3.76	2.28	2.47	2.96
4	3.38	2.06	3.36	2.00	2.04	3.04
5	3.15	1.93	3.17	1.78	2.00	2.96
6	3.03	1.79	2.97	1.71	1.97	3.02
7	3.36	2.10	3.43	1.76	2.01	3.28
8	3.74	2.56	3.89	2.01	2.21	3.84

Table 4
Comparison of the average accuracy between the proposed algorithm and some other algorithms for the Yale database.

training number	PCA+PNN AA (%)	LDA+PNN AA (%)	KPCA+PNN AA (%)	KLDA+PNN AA (%)	IKLDA+PNN AA (%)	SVM AA (%)
2	50.84	62.55	50.81	63.00	62.95	50.79
3	57.04	70.45	56.99	71.71	71.59	57.28
4	61.19	74.76	61.13	76.48	76.59	61.16
5	64.39	77.24	64.31	79.33	79.60	64.12
6	66.95	79.05	65.73	81.15	81.56	66.46
7	68.95	80.22	67.24	82.34	82.95	68.38
8	70.48	80.67	69.02	83.06	83.80	69.72

Fig. 1. A part of samples of two classes in ORL database.

Fig. 2. Some of the samples for one of the classes in the AR database.

Table 5
Comparison of the standard deviation between the proposed algorithm and some other algorithms for the Yale database.

training number	PCA+PNN std	LDA+PNN std	KPCA+PNN std	KLDA+PNN std	IKLDA+PNN std	SVM std
2	6.97	5.06	6.90	5.36	5.69	7.93
3	4.57	3.97	4.49	4.20	4.21	6.07
4	3.65	4.47	3.65	4.17	4.29	5.29
5	3.87	5.41	3.91	5.46	5.50	5.47
6	5.09	6.42	5.33	6.59	6.62	6.55
7	6.92	7.67	7.12	8.03	7.86	7.91
8	9.60	8.98	9.60	9.45	9.67	10.40

Table 6
Comparison of the average accuracy between the proposed algorithm and some other algorithms for the AR database.

training number	PCA+PNN AA (%)	LDA+PNN AA (%)	KPCA+PNN AA (%)	KLDA+PNN AA (%)	IKLDA+PNN AA (%)	SVM AA(%)
2	80.31	84.14	77.28	87.27	87.15	81.23
3	85.14	90.01	81.76	92.99	92.89	82.34
4	87.99	92.92	84.48	95.94	95.86	85.68
5	89.95	94.44	86.32	97.53	97.46	87.31
6	91.41	95.14	87.67	98.40	98.35	89.46
7	92.50	95.42	88.65	98.87	98.85	89.94
8	93.37	95.47	89.45	99.16	99.12	94.21

Table 7
Comparison of the standard deviation between the proposed algorithm and some other algorithms for the AR database.

training number	PCA+PNN std	LDA+PNN std	KPCA+PNN std	KLDA+PNN std	IKLDA+PNN std	SVM std
2	7.66	9.14	6.76	8.63	8.64	9.87
3	6.49	7.82	5.99	7.06	7.02	8.14
4	5.61	5.78	5.67	4.96	4.97	6.07
5	4.96	3.98	5.30	3.15	3.21	4.77
6	4.41	2.66	5.00	1.82	1.91	6.02
7	3.99	1.98	4.69	0.99	1.08	5.34
8	3.68	1.88	4.45	0.62	0.67	4.21

Table 8
Comparison of the average accuracy between the proposed algorithm some other algorithms for the AR database.

algorithm	2 AA (%)	4 AA (%)	6 AA (%)	8 AA (%)
PCA[7]	69.55	67.99	69.53	67.01
NNC[8]	61.24	58.41	62.33	59.06
CBNNC[8]	60.35	58.08	61.97	60.52
NNL[9]	59.99	58.68	62.97	59.89
IKLDA+PNN	87.15	95.86	98.35	99.12

Table 9
Comparison of the average accuracy and standard deviation between the proposed algorithm and the SVM for the ORL database.

Training number	SVM AA (%) ± std	IKLDA+PNN AA (%) ± std
2	66.62 ± 0.038	77.96 ± 0.030
3	72.52 ± 0.030	86.83 ± 0.025
4	75.71 ± 0.030	91.44 ± 0.020
5	78.47 ± 0.030	93.95 ± 0.020
6	80.48 ± 0.030	95.43 ± 0.020
7	82.19 ± 0.033	96.35 ± 0.020
8	83.00 ± 0.038	97.22 ± 0.022

Table 10
Comparison of the average accuracy and standard deviation between the proposed algorithm and the SVM for the Yale database.

Training number	SVM AA (%) ± std	IKLDA+PNN AA (%) ± std
2	50.79 ± 0.079	62.95 ± 0.057
3	57.28 ± 0.061	71.59 ± 0.042
4	61.16 ± 0.053	76.59 ± 0.043
5	64.12 ± 0.055	79.60 ± 0.055
6	66.46 ± 0.065	81.56 ± 0.066
7	68.38 ± 0.079	82.95 ± 0.079
8	69.72 ± 0.104	83.80 ± 0.097

rate for the SVM is better than the PCA+PNN algorithm,but its performance is significantly worse than the algorithm proposed here. Our proposed algorithm's recognition rate for the Yale database is shown in Fig. 4. Here, it can be seen that the recognition rate for the SVM was approximately the same as the PCA+PNN algorithm, but, once again, it did not perform as well as our proposed algorithm. The recognition rate results for the AR database are shown in Fig. 5. Here, the proposed algorithm and the KLDA+PNN algorithm had approximately the same recognition rate.

Table 2 shows the final recognition results. The average accuracy achieved by the algorithm proposed in this paper for the ORL human faces database was much better than the classic classification algorithms, PCA+PNN, LDA+PNN, KPCA+PNN and KLDA+PNN. Table 3 shows the standard deviation for the final recognition results. The standard deviation (std) for the ORL human faces database achieved by our proposed algorithm was much better it was for the PCA+PNN, LDA+PNN, KPCA+PNN and KLDA+PNN algorithms. Tables 4 and 5 give the final recognition results for five algorithms (including the one proposed here) for the Yale human faces database. Tables 6 and 7 provide the final recognition results for the AR human faces database. Even if the number of samples for each class is relatively small, the proposed algorithm still achieves a good classification result. It can be seen

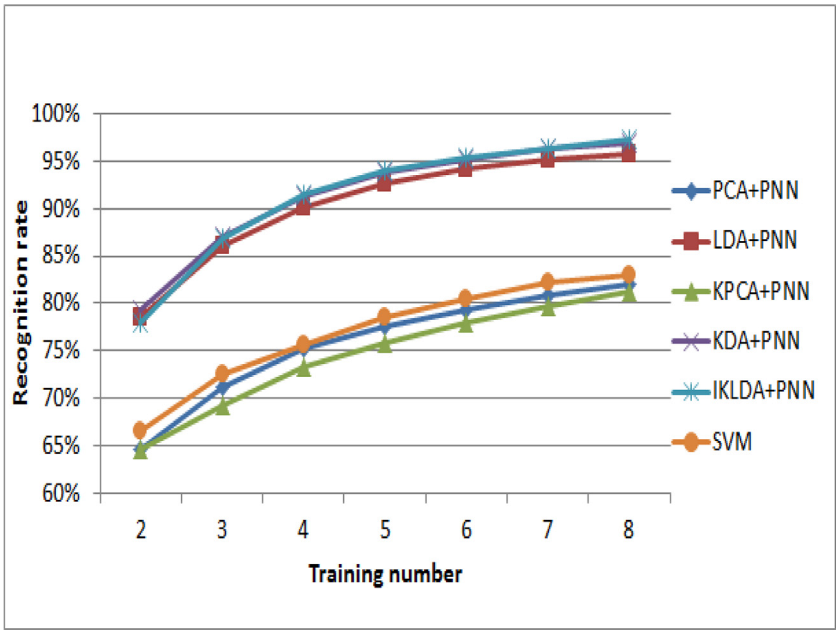

Fig. 3. Recognition rate for the six algorithms (ORL database).

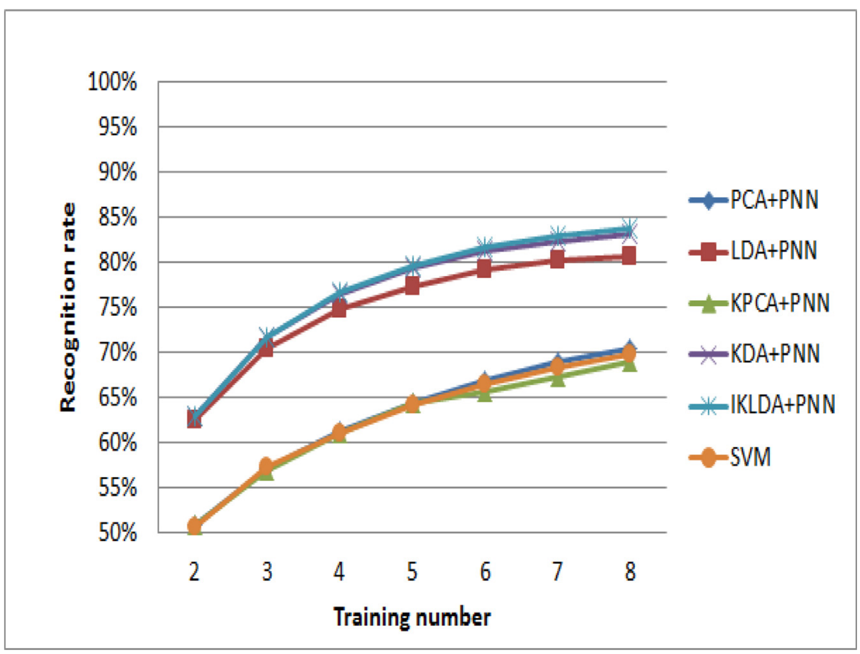

Fig. 4. Recognition rate for six algorithms (for the Yale database).

in Table 3 that the performance of the IKLDA+PNN algorithm is slightly worse than the other algorithms for the ORL database, regardless of the number of samples for each class. However, on the whole, its classification effect is much better than the PCA+PNN, LDA+PNN, KPCA+PNN and KLDA+PNN algorithms.

For the AR database, the 2nd, 4th, 6th and 8th samples were selected as training samples for each class. The remaining ones were used as test samples. Although the proposed algorithm produced a poorer result than the KLDA+PNN algorithm for the AR database for some samples in each class, its result was better with other samples. It can also be seen from Table 8 that the classification performance of our proposed algorithm was better than some of the other algorithms for the AR database. Tables 9 and 10 show the performance (in terms of recognition rate and standard deviation) of the SVM and our proposed algorithm for the ORL and Yale face data. Here it can be seen that, for both the ORL and Yale face data, the recognition rate for the IKLDA+PNN was better than the SVM, as is the standard deviation.

A time comparison over the ORL and Yale databases for each of the IKLDA+PNN and PCA+PNN, LDKPCA+PNN, KLDA+PNN, and SVM algorithms is shown in Table 11. A time comparison for the AR database in relation to the IKLDA+PNN and PCA+PNN, LDA+PNN, KPCA+PNN, KLDA+PNN algorithms is shown in Table 12. We can see that the IKLDA+PNN algorithm was faster at solving the face recognition problems for the ORL and Yale databases than the other algorithms. When the number of training samples was 2 and 3, the IKLDA+PNN algorithm was faster than the PCA+PNN and KLDA+PNN algorithms for the AR database, but

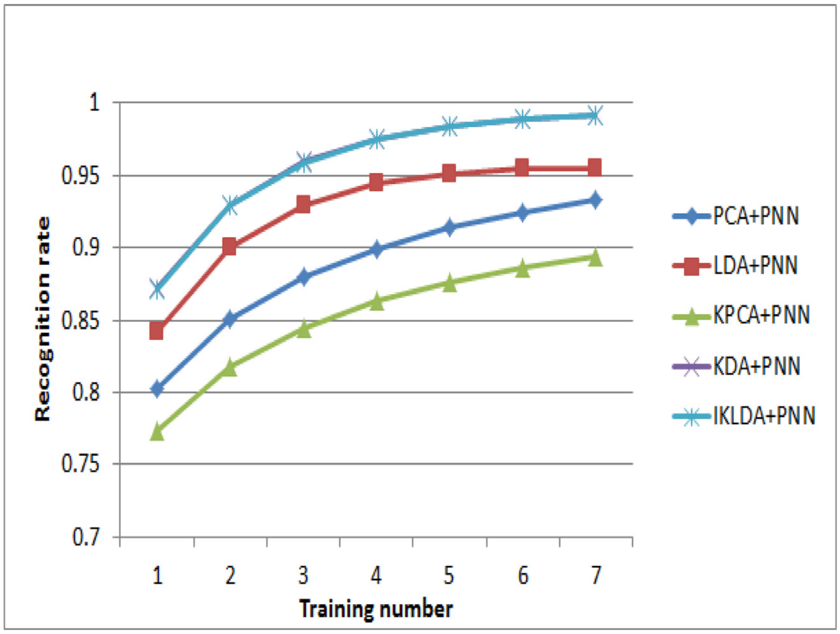

Fig. 5. Recognition rate for five algorithms (for the AR database).

Table 11
Comparison of time performance for the ORL and Yale databases for the IKLDA and PCA, LDA, KPCA, KLDA, and SVM algorithms.

Algorithm	PCA+PNN	LDA+PNN	KPCA+PNN	KLDA+PNN	IKLDA+PNN	SVM
ORL(s)	60.76	55.00	55.59	55.61	48.28	6062.70
yale(s)	78.54	77.79	78.86	78.72	76.53	1334.86

Table 12
Comparison of time performance for the AR database for the IKLDA and PCA, LDA, KPCA, and KLDA algorithms.

Training number	Algorithm	Time(s)
2	PCA+PNN	17.69
	LDA+PNN	13.08
	KPCA+PNN	13.91
	KLDA+PNN	15.57
	IKLDA+PNN	14.52
3	PCA+PNN	113.14
	LDA+PNN	66.51
	KPCA+PNN	68.61
	KLDA+PNN	79.80
	IKLDA+PNN	71.23
4	PCA+PNN	513.55
	LDA+PNN	224.24
	KPCA+PNN	243.55
	KLDA+PNN	275.04
	IKLDA+PNN	229.86

slower than the LDA+PNN and KPCA+PNN algorithms. When the number of training samples was 4, the IKLDA+PNN algorithm was faster than the PCA+PNN, KPCA+PNN, and KLDA+PNN algorithms for the AR database, but slower than the LDA+PNN algorithm.

5. Conclusion

This paper has proposed a classification method for solving face recognition problems that is based on a hybrid IKLDA and PNN algorithm. The proposed algorithm not only saves overall computational time, but also improves the effectiveness of LDA-based face recognition. The results of an experimental simulation have validated the algorithm's effectiveness for face recognition and classification and have demonstrated that the algorithm can outperform other traditional algorithms in terms of average accuracy and standard deviation when applied to the face image data held in three separate standard datasets.

Declaration of Competing Interest

None.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grants nos. 61662090, 71461027, 71471158); the science and technology project of Guizhou ([2017]1207); the training program of high level innovative talents of Guizhou ([2017]3); the Guizhou province natural science foundation in China (KY[2016]018);Guizhou province natural science foundation in China (Qian Jiao He KY [2014]295); Zhunyi innovative talent team (Zunyi KH (2015)38); Science and technology talent training object of Guizhou province outstanding youth (Qian ke he ren zi [2015]06); Guizhou science and technology cooperation plan (Qian Ke He LH zi [2016]7028); Project of teaching quality and teaching reform of higher education in Guizhou province(Qian Jiao gaofa [2015]337); 2013-2015 and 2017-2018 Zunyi 15851 talents elite project funding; Innovative talent team in Guizhou Province (Qian Ke HE Pingtai Rencai [2016]5619).

References

[1] X.-H. Han, Y.-W. Chen, Scene and object recognition with supervised nonlinear neighborhood embedding, Int. J. Innov. Comput. Inf. Control 7 (8) (2011) 4861–4870.
[2] J. Mei, K. Li, K. Li, Customer-satisfaction-aware optimal multiserver configuration for profit maximization in cloud computing, IEEE Trans. Sust. Comput. 2 (1) (2017) 17–29.

[3] K.Q. Weinberger, J. Blitzer, L.K. Saul, Distance metric learning for large margin nearest neighbor classification, in: Proceedings of the Advances in Neural Information Processing Systems, 2006, pp. 1473–1480.
[4] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inf. Theory 13 (1) (1967) 21–27.
[5] Y. Li, L. Jia, Z. Wang, Y. Qian, H. Qiao, Un-supervised and semi-supervised hand segmentation in egocentric images with noisy label learning, Neurocomputing (2018), doi:10.1016/j.neucom.2018.12.010.
[6] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, F.E. Alsaadi, A survey of deep neural network architectures and their applications, Neurocomputing 234 (2017) 11–26.
[7] L. Tan, E. Li, A. Ouyang, M. He, X. Zhou, Network intrusion clustering method based on improved LDA and CNN, Comput. Eng. Appl. 49 (2) (2013) 89–91.
[8] D. Gao, C. Wu, Y. Li, A new hybrid classifier model and its application on face recognition, Syst. Eng. Theory Pract. 36 (6) (2016) 1517–1525.
[9] K. Li, X. Tang, B. Veeravalli, K. Li, Scheduling precedence constrained stochastic tasks on heterogeneous cluster systems, IEEE Trans. Comput. 64 (1) (2015) 191–204.
[10] C. Liu, K. Li, K. Li, Minimal cost server configuration for meeting time-varying resource demands in cloud centers, IEEE Trans. Parallel Distr. Syst. 29 (11) (2018) 2503–2513.
[11] A. Ouyang, X. Peng, J. Liu, A. Sallam, Hardware/software partitioning for heterogenous MPSOC considering communication overhead, Int. J. Parallel Program. (2016) 1–24.
[12] M.H. Siddiqi, R. Ali, A.M. Khan, Y.T. Park, S. Lee, Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields, IEEE Trans. Image Process. 24 (4) (2015) 1386–1398.
[13] J. Yu, X. Lu, Wafer map defect detection and recognition using joint local and nonlocal linear discriminant analysis, IEEE Trans. Semicond. Manufact. 29 (1) (2016) 33–43.
[14] X. Zhang, D. Chu, R.C.E. Tan, Sparse uncorrelated linear discriminant analysis for undersampled problems, IEEE Trans. Neural Netw. Learn. Syst. 27 (7) (2016) 1469–1485.
[15] C.P. Mbo'o, K. Hameyer, Fault diagnosis of bearing damage by means of the linear discriminant analysis of stator current features from the frequency selection, IEEE Trans. Ind. Appl. 52 (5) (2016) 3861–3868.
[16] H. Wang, L. Yan, H. Huang, C. Ding, From protein sequence to protein function via multi-label linear discriminant analysis, IEEE/ACM Trans. Comput. Biol. Bioinf. 14 (3) (2017) 503–513.
[17] K. Li, X. Tang, K. Li, Energy-efficient stochastic task scheduling on heterogeneous computing systems, IEEE Trans. Parallel Distr. Syst. 25 (11) (2014) 2867–2876.
[18] K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press, 1990.
[19] D.F. Specht, Probabilistic neural networks, Neural Netw. 3 (1) (1990) 109–118.
[20] Y. Xu, K. Li, L. He, L. Zhang, K. Li, A hybrid chemical reaction optimization scheme for task scheduling on heterogeneous computing systems, IEEE Trans. Parallel Distr. Syst. 26 (12) (2015) 3208–3222.
[21] Z. Wang, M. Jiang, Y. Hu, H. Li, An incremental learning method based on probabilistic neural networks and adjustable fuzzy clustering for human activity recognition by using wearable sensors, IEEE Trans. Inf. Technol. Biomed. 16 (4) (2012) 691–699.
[22] H.T. Lin, T.J. Liang, S.M. Chen, Estimation of battery state of health using probabilistic neural network, IEEE Trans. Ind. Inf. 9 (2) (2013) 679–685.
[23] K. Li, W. Yang, K. Li, Performance analysis and optimization for SPMV on GPU using probabilistic modeling, IEEE Trans. Parallel Distr. Syst. 26 (1) (2015) 196–205.
[24] O. Antropov, Y. Rauste, H. Astola, J. Praks, T. Hime, M.T. Hallikainen, Land cover and soil type mapping from spaceborne polsar data at l-band with probabilistic neural network, IEEE Trans. Geosci. Remote Sens. 52 (9) (2014) 5256–5270.
[25] H. Hayashi, T. Shibanoki, K. Shima, Y. Kurita, T. Tsuji, A recurrent probabilistic neural network with dimensionality reduction based on time-series discriminant component analysis, IEEE Trans. Neural Netw. Learn. Syst. 26 (12) (2015) 3021–3033.
[26] M. Kusy, R. Zajdel, Application of reinforcement learning algorithms for the adaptive computation of the smoothing parameter for probabilistic neural network, IEEE Trans. Neural Netw. Learn. Syst. 26 (9) (2015) 2163–2175.
[27] C. Liu, K. Li, C. Xu, K. Li, Strategy configurations of multiple users competition for cloud service reservation, IEEE Trans. Parallel Distr. Syst. 27 (2) (2016) 508–520.
[28] W. Yang, K. Li, Z. Mo, K. Li, Performance optimization using partitioned SPMV on GPUS and multicore cpus, IEEE Trans. Comput. 64 (9) (2015) 2623–2636.
[29] S.-H. Lin, S.-Y. Kung, L.-J. Lin, Face recognition/detection by probabilistic decision-based neural network, IEEE Trans. Neural Netw. 8 (1) (1997) 114–132.
[30] X. Tan, S. Chen, J. Li, Z.-H. Zhou, Learning non-metric partial similarity based on maximal margin criterion, in: Proceedings of the Null, IEEE, 2006, pp. 138–145.
[31] I.T. Jolliffe, Principal component analysis, J. Market. Res. 87 (100) (1986) 513.
[32] D. Cai, X. He, Y. Hu, J. Han, T. Huang, Learning a spatially smooth subspace for face recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition CVPR, IEEE, 2007, pp. 1–7.
[33] C. Zhou, L. Wang, Q. Zhang, X. Wei, Face recognition based on PCA image reconstruction and LDA, Optik Int. J. Light Electron Opt. 124 (22) (2013) 5599–5603.
[34] T.Y. Kim, K.M. Lee, S.U. Lee, C.-H. Yim, Occlusion invariant face recognition using two-dimensional PCA, in: Proceedings of the Advances in Computer Graphics and Computer Vision, Springer, 2007, pp. 305–315.
[35] W. Dong, M. Zhou, G. Geng, Face recognition based on DWT and improved linear discriminant analysis, in: Proceedings of the Recent Advances in Computer Science and Information Engineering, Springer, 2012, pp. 699–704.
[36] Q.-B. Gao, Z.-Z. Wang, Center-based nearest neighbor classifier, Pattern Recogn. 40 (1) (2007) 346–349.
[37] W. Zheng, L. Zhao, C. Zou, Locally nearest neighbor classifiers for pattern classification, Pattern Recogn. 37 (6) (2004) 1307–1309.
[38] A. Ouyang, X. Peng, Q. Wang, Y. Wang, T.K. Truong, A parallel improved iwo algorithm on GPU for solving large scale global optimization problems, J. Intell. Fuzzy Syst. 31 (2) (2016) 1041–1051.

Aijia Ouyang received the Ph.D. degree in Computer Science, Hunan University, China, in 2015. He is an associate professor from Department of Information Engineering, Zunyi Normal University, Zunyi, Guizhou, China. His research interests include intelligence computing, parallel computing, cloud computing and parallel data mining. He has published more than 60 research papers in international conference and journals of intelligence algorithm and parallel computing.

Yanmin Liu received the B.S. degree in Applied mathematics from Harbin Institute of Technology, Haerbin, China, in 2001, the M.S. degree in Control science and engineering from Heilongjiang Bayi Agricultural University, Daqing, China, in 2006, and the Ph.D. degree in Decision theory and application from Shandong Normal University, Jinan, China, in 2011. He is presently serving as a professor in School of Mathematics, Zunyi Normal University. His main fields of research are Swarm Intelligence, Bio-inspired Computing, Multi-objective Optimization and their applications on supply chain.

Shengyu Pei received the M.E. degree in the department of computer science and technology from Guangxi University for Nationalities, Nanning, China, in 2010. He is a Lecturer from Fangchenggang College, Guangxi University of Finance and Economics, Fangchenggang, China. His research interests include artificial intelligence. He has published more than 20 research papers in international conference and journals of swarm intelligence algorithms and pattern recognition.

Xuyu Peng received her B.E. degree from the Department of Computer Science and Technology of Xiangtan University, China. She is a Lecturer from Department of Information Engineering, Zunyi Normal University, Zunyi, Guizhou, China. Her research interests include parallel computing and intelligence computing. She has published more than 10 research papers articles in international conferences and journals on parallel computing and scheduling algorithms.

Minghua He is a Ph.D. candidate in School of Economics and Management, Tongji University, China. He is an assistant researcher in Institution of Higher Education, Jinggangshan University, Ji'an, Jiangxi, China. His current research interests include operations management, sharing economy, online consumers' behaviors, and computer-assisted language learning. He has published several research papers in international conference and Chinese journals of management and computer science.

Qian Wang received his Ph.D. in pattern recognition from the Institute of Automation, Chinese Academy of Sciences, China, in 2007. His research interests include computer vision, image processing and artificial intelligence. He has published 10 research papers in international conferences and journals on pattern recognition.