IJRCS

International Journal of Robotics and Control Systems
Vol. 3, No. 2, 2023, pp. 221-232
ISSN 2775-2658
http://pubs2.ascee.org/index.php/ijrcs

ASCEE

Improving the Recognition Percentage of the Identity Check System by Applying the SVM Method on the Face Image Using Special Faces

Azita Mousavi [a,1], Amir Hossein Sadeghi [b,2,], Ali Mojarrad Ghahfarokhi [c,3], Fatemehalsadat Beheshtinejad [d,4], Mahsa Madadi Masouleh [e,5,*]

[a] School of Engineering, San Francisco Bay University, San Francisco, CA 94132, USA
[b] Department of Industrial and System Engineering, North Carolina State University, Raleigh, NC, USA
[c] Department of Computer and Information Science, University of Michigan, Dearborn, USA
[d] Department of Computer Engineering, Islamic Azad University, Najafabad Branch, Esfahan, Iran
[e] Department of Electrical and Computer Engineering, University of Victoria, BC V8P 5C2, Canada
[1]amousavi420@alumni.sfbu.edu; [2] asadegh3@ncsu.edu; [3] mojarrad@umich.edu; [4] fa_beheshti@yahoo.com; [5]Mahsa.mmadadi@gmail.com
* Corresponding Author

ARTICLE INFO

Article history
Received March 05, 2023
Revised March 26, 2023
Accepted April 01, 2023

Keywords
Face recognition;
SVM algorithm;
Special faces

ABSTRACT

Face recognition has attracted tremendous attention during the last three decades because it is considered a simple pattern recognition and image analysis method. Also, many facial recognition patterns have been introduced and used over the years. The SVM algorithm has been one of the successful models in this field. In this article, we have introduced the special faces first. In the following, we have fully explained the SVM method and its subsets, including linear and non-linear support vector machines. Suggestions for improving the recognition percentage of a person's identity check system by applying the SVM method on the face image using special faces are presented. For this test, 10 face images of 40 people (400 face images in total) have been selected from the ORL database. In this way, by choosing the optimal parameter C, determining the most suitable training samples, comparing more accurately with training images and using the distance with the closest training sample instead of the average distance, the proposed method has been implemented and tested on the famous ORL database. The obtained results are FAR=0.23% and FRR=0.48%, which shows the very high accuracy of the operation following the application of the above suggestions.

1. Introduction

Biometric systems, with the increasing development of information technology in recent years, have enjoyed high growth in the field of identity identification and verification. In these systems, identification is made by using the behavioral or physiological properties of people such as fingerprints, faces, facial expressions, voice, signature, the way of pressing the keyboard and the way of walking. Each of these methods has its own strengths and weaknesses. Identification systems or identity checks through facial images are no exception to this rule [1]. Many algorithms for feature extraction have been reviewed and introduced, specifically their application in climate change and water treatment areas [2]-[6]. Support vector machine is one of the supervised learning methods used for classification and regression [7]-[9]. This method is among the new methods that have shown good performance in recent years compared to older classification methods such as neural networks [10]. The original support vector machine algorithm was invented by Vladimir Vepnik in 1963 and was proposed by Vepnik and his group at AT&T Bell Labs in 1995 and extended to the nonlinear case [11]. This classifier is actually a binary digital classifier [12]-[14], which has the ability to classify samples of multiple classes. It is also used to solve linear and non-linear problems [15]. The working basis of the SVM classifier is the linear classification of data, and in the linear division of the data, we try to choose a line that has a higher confidence margin. The main feature of the support vector machine is to find the optimal super-plane for classifying two classes of data, and this optimization method has been used in optimization applications on yield crops [16]. The SVM algorithm is classified as a pattern recognition algorithm. SVM is used wherever there is a need for pattern classification or classification of objects in a specific class. Currently, it is used in many fields of audio and image processing and has provided very good results. The support vector machine algorithm has a relatively simple training compared to the neural network, and unlike the neural network, it does not get stuck in the local maximum, and it works well for high-dimensional data [16], [17].

Recently, many works have been implemented using the SVM classifier in different fields. Whereas Kumar *et al.* [18] Median filter is used for noise reduction in a specific image by focusing on several facial recognition techniques used for facial identification. A facial vector is used to identify input faces as single or several, then the SVM classifier. The suggested algorithm was tested, and the precision reached up to 90 percent. From the experiment, it is observed that the SVM classifier more correctly identified the facial images of the previous classifiers. An average precision of identification of 89 percent was obtained for SVM. Dino and Abdulrazzaq [19] implemented another automated framework for the identification of facial expressions, which can identify all eight essential facial expressions. The experimental findings revealed that the presented approach offers 93.53% of the identification rate in the use of the SVM classification, 82.97% for the use of the MLP classification, and 79.97% for the KNN classification. The paper concludes that the approach provides stronger outcomes by utilizing the SVM as a classifier.

Shi *et al.* [20], Dadi & Pillutla [21] and Chen & Haoyu [22] proposed face recognition methods based on SVM. Were, Dadi & Pillutla, for both test and training images, extracted the Histogram of Oriented Gradient (HOG) features and gave them to the SVM classifier. The eigen feature-based face recognition algorithm was compared with the algorithm. Using 8 different datasets, the suggested algorithm and PCA was verified. The results showed that the proposed algorithm showed a higher face recognition rate in all face datasets compared to the traditional face recognition algorithm based on the eigenfeature. In comparison to the PCA-based face recognition algorithm, there is an 8.75 percent improvement in the face recognition rate.

The contribution of this article is to present new techniques to improve the performance of the SVM method in checking the identity based on the face image. The proposed method of "Face Special" are used as features and then by applying solutions such as optimum selection of parameter C, determination of the most suitable samples training, more accurate comparison with training images and using the distance to the closest example. Instead of the distance from the average, there have been improvements in the recognition results. In this article, we first briefly introduce the faces special and SVM, and then we will introduce our proposed algorithm. Next, the results of the system implementation and testing and the conclusion are given at the end.

2. Special Faces

In mathematical language, the main components of the distribution of faces, or the feature vector of the covariance matrix of the set of face images, behave in such a way that an image as a point or the vector is considered in a space with a very high number of dimensions. An eigenvector can be used as a set of features that are supposed to determine the difference between the images of the faces. Here, because the special vector of pictures is facing, they are called special faces [23]. In the method of special faces, the images of the reference set are considered as a linear combination of special faces (basic faces), where the coefficients of special faces are the components of the feature vectors of the above images [24]. In this method, first, the average of the faces of the reference set is obtained and then each face is subtracted from the average, and the covariance of the obtained set is calculated. The eigenvectors of the covariance matrix are the eigen figures of the reference set [25].To calculate special faces, if we have training images with size W×H, first we consider each face image as a vector with length WH, then we calculate their average. If I1, I2, I3, …, IM are our images, we get their average as:

$$\Psi = \frac{1}{M}\sum_{n=1}^{M} I_n \tag{1}$$

and calculate the difference between each face from the average. That is, we have:

$$\phi_i = I_i - \Psi \tag{2}$$

$$Cov = \frac{1}{M}\sum_{n=1}^{M} \phi_n \phi_n^t = AA^t \tag{3}$$

where $A = [\phi_1 \phi_2 \dots \phi_M]$. If we call the eigenvector of the covariance matrix (i.e., $\mathrm{L} = AA^t$), v_k, u_k obtained as:

$$u_l = \sum_{k=1}^{M} v_{lk}\, \phi_k, \qquad l = 1, \dots, M \tag{4}$$

are called special figures.

In order to move a face image, like I to the space of special faces, the operation is performed as:

$$W_k = u_k^t (I - \Psi) \tag{5}$$

and instead of using the image I, the feature vector W_K is used. Examples of special faces are shown in Fig. 1.

Fig. 1. The first eight special faces of the FERET database

3. Support Vector Machine

3.1. Linear Support Vector Machine

Consider a linear classification problem for separating two classes of training data. Let Xi, $i = 1, 2, \dots, N$ be the set of feature vectors[1] related to training data that are linearly separable and classified into two classes, 1 and 2, as shown in Fig. 2. The goal is to obtain the hyperplane $g(x)= W^T X + b$ =0 in such a way that all the training data are correctly classified. In general, this super-plane is not unique, and it is possible to obtain different coefficients for w and b. Fig. 3 shows some examples of

[1] A feature vector is a vector containing multiple elements about an object.

hyperplanes[2] that can be considered for the correct linear classification of two-class data. All the hyperplanes shown in Fig. 3. perform the separation process correctly, but there is only one hyperplane that has the largest distance to the data of both classes, which is called the optimal separator hyperplane and is expected can extend the obtained boundary to all possible ranges [26].

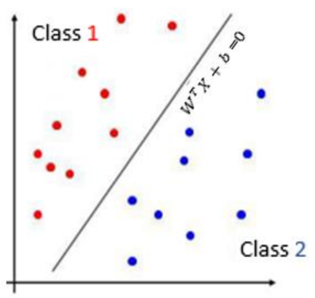

Fig. 2. Separator hyperplanes

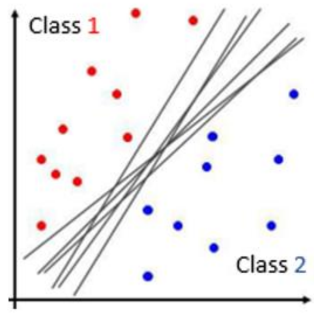

Fig. 3. Training data in two linearly separable classes

The optimal separator super-plane with maximum margin[3] value is shown in Fig. 4. This super page provides the best classification with the least amount of error due to the fact that it creates the most space for the closest data of each class on its two sides. Therefore, this type of classification shows better results when running on test samples. This result is an important issue in the design of classifiers, which is called the generalization power of the classifier.

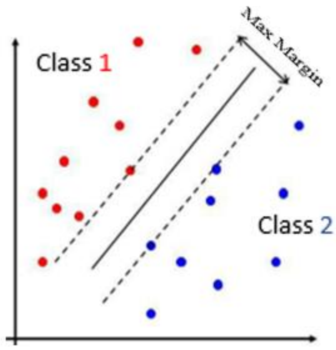

Fig. 4. Optimal separator super-plane with a maximum margin value

It is assumed that there is a problem with separating the set of educational samples that belong to two separate classes. According to (6), we define the index vector yi including the values of $+1$ and -1, in such a way that it has a value of $+1$ for class 1 and a value of -1 for class 2.

$$y_i = \begin{cases} 1 & if \quad xi\ in\ class\ 1 \\ -1 & if \quad xi\ in\ class\ 2 \end{cases} \tag{6}$$

Each hyperplane is precisely determined by its direction and location in space, where w is the direction of the hyperplane and b is its location in space. The decision function can be defined as (7) and (8).

$$d(x) = sign(W^T X + b) \tag{7}$$

$$\begin{cases} (W^T X_i + b) > 0 & if \quad y_i = +1 \\ (W^T X_i + b) < 0 & if \quad y_i = -1 \end{cases} \tag{8}$$

[2] Hyperplane is a decision boundary that divides the input space into two or more regions.
[3] Margin is the distance between the hyperplane and the observations closet to the hyperplane

The cloud page that is used in this type of problem for data classification must have two special features: First, it has the lowest possible error rate, and on the other hand, it has the largest possible distance from the data of each class. In this case, if the (8) is considered for the separating page, the training data will be placed on the top and bottom of this page, which will be, respectively, for $y_i = +1$ and $(W^T X_i + b) > 0$ and for $y_i = -1$ and $(W^T X_i + b) < 0$. Based on the stated conditions, when sets of points are optimally separated by a plane when:

1. They have been placed in their respective class without mistakes.
2. The distance between the closest points of each data class to the separating plane is maximum.

Based on this, parameters w and b should be calculated in such a way that the two mentioned conditions are met. Therefore, in linear support vector machine classification, the intention is to obtain a super-plane that, in addition to correctly classifying the training data, also maximizes the margin between two classes. Fig. 4 shows the difference between the two separator super-planes. The super-plane shown with a thicker line in Fig. 5 is the desired super-plane to separate two classes because, in addition to separating the two classes, it also maximizes the margin between the super-plane and the two classes.

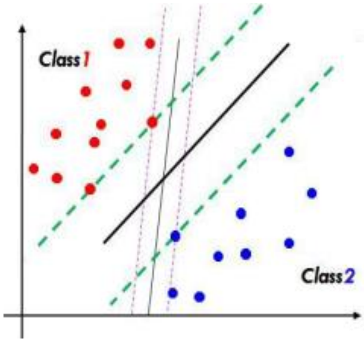

Fig. 5. Separator super-plane

Let's assume that the equation of the optimal separator hyperplane is $W^T X_i + b = 0$, so the equation of the marginal hyperplanes on both sides of the separator hyperplane will be $W^T X_i + b = +1$ and $W^T X_i + b = -1$. Fig. 6 shows the considered equations for the edges and the optimal separator plate.

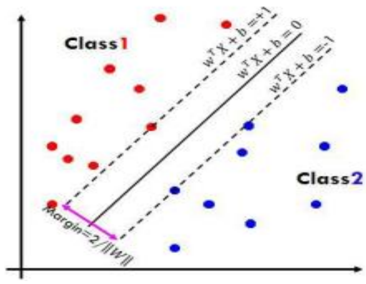

Fig. 6. The optimal separator hyperplane and its margins

Because the data is linearly separable, none of the data is placed on the $W^T X_i + b = 0$ plane. Therefore, instead of (8), you can use (9).

$$W^T X_i + b = \begin{cases} \geq \ 1 & if \quad y_i = class1 \\ \leq -1 & if \quad y_i = class2 \end{cases} \tag{9}$$

And it can be rewritten in the simplified form of (10).

$$W^T X_i + b \geq 1 \, , \mathrm{i} = \ 1, 2, \dots, \mathrm{N} \tag{10}$$

In order to introduce the best separating page, the page should have the largest distance from the classes, and this distance should be the maximum value. Now the distance between these two margins is obtained as (11).

$$\frac{|(W^TX+b-1)-(W^TX+b+1)|}{\|W\|}=\frac{2}{\|W\|} \tag{11}$$

According to (11), if the value of $\| \frac{2}{\|W\|} \|$ is maximized, the desired margin value is at its maximum value. For simplicity, instead of maximizing $\left\| \frac{2}{\|W\|} \right\|$ you can minimize the value of $\| \|W\|^2 \|$ in the denominator. Or in other words, the sentence can be written as $\frac{1}{2}WW^T$. Therefore, based on the stated conditions, the support vector machine problem becomes the following form, which is called a quadratic problem.

$$minimize\ J(w)=\frac{1}{2}WW^T \tag{12}$$

$$\text{subject to } y_iW^TX_i+b) \geq 1\ , \quad \text{i} = \ 1, 2\ , \ldots , \text{N} \tag{13}$$

Because the form of (13) is a quadratic problem with an inequality condition, the value of the objective function will be unique. That is, it has only one extremum point. This is one of the important features of support vector machines compared to multilayer neural networks, where there are many local minimum points [26]. In the above quadratic problem, Lagrange coefficients should be used to minimize and obtain the optimal values of w and b according to the existing inequality condition. Lagrange coefficients, which are sometimes called indeterminate coefficients, are used to identify specific points of a function that has several variables and constraints. Lagrange coefficients, which are sometimes called indeterminate coefficients, are used to identify specific points of a function that has several variables and constraints. In fact, if the goal is to minimize the function $f(x)$ according to the constraint $g(x) \geq 0$, then the Lagrange function defined as (14) should be minimized.

$$L(x،\alpha)=f(x)+\ \alpha g(x) \tag{14}$$

where $\alpha \geq 0$ is called the Lagrange coefficient. Considering the Lagrange function, the form with the inequality condition of the quadratic problem (13) becomes the unconditioned form of the (15), which is called the initial Lagrange function.

$$L(x،\alpha)=f(x)+\ \alpha g(x) \tag{14}$$

$$L_p(\text{w}, \text{b}, \alpha)=\frac{1}{2}W^TW-\sum_{i=1}^{N}\alpha_i\,[y_i(W^TX_i+b)-1] \tag{15}$$

In order to minimize the (15), one must obtain the stationary points of the Lagrange function. To achieve this goal, KKT conditions are applied, which is in the form of (16).

$$\begin{gathered}\frac{\partial L}{\partial W}=0\ \Rightarrow w_0=\sum_{i=1}^{N}\alpha_i x_i y_i \\ \frac{\partial L}{\partial b}=0 \Rightarrow \sum_{i=1}^{N}\alpha_i y_i\end{gathered} \tag{16}$$

Now, if the value of W obtained from the partial derivatives of (16) is placed in (15) itself, the basic equation of vector machines is represented as (17), which is the dual form of Lagrange's (15).

$$Max\ L_d(\alpha) = \sum_{i=1}^{N} \alpha_i - {}^1\!/_2 \sum_{ij=1}^{N} y_i y_j \alpha_i x_i^T x_j \alpha_j$$
$$\text{S.T.} \begin{cases} 0 \le \alpha_i \le C \\ \sum_{i=1}^{N} \alpha_i y_i = 0 \end{cases} \tag{17}$$

By solving the above equation, the values of Lagrange's coefficients, i.e., α, will be obtained for each of the training vectors, which will be greater than or equal to zero according to the limitation mentioned in (17). C is a positive parameter specified by the user. Among the training vectors of classes 2 and 9, those whose corresponding α value is greater than zero are called support vectors. Only these vectors are used in calculations to obtain the optimal hyperplane, so the amount of calculations is greatly reduced. The right support vectors are placed on two super-planes $W^T X + b = \pm 1$. The optimal value of w_0 was obtained from the (16). The optimal value of b is also calculated through the relationship and averaging of all the obtained values. The general equation for calculating the optimal value of b can be expressed as (18).

$$b_0 = \frac{1}{sv} \sum_{i=1}^{sv} (y_i - w^T x_i) \tag{18}$$

where SV specifies the number of support vectors. By obtaining the values of w_0 and b_0, it is possible to obtain the optimal separating hyperplane. Having these values and according to the relation of the decision function of the support vector machine, i.e., (7) for the classification of unknown data such as x, if the value of the decision function is greater than zero, the data will be classified in class 1, and if the value of the decision function is smaller than zero, the data is classified in class 2.

3.2. Non-linear Support Vector Machin

In the beginning, the support vector machine was used only for linear systems, and the optimal separator plane existed only for the linear mode. This was despite the fact that in many classification and regression problems, the linear solution did not provide a suitable answer. Extensive studies in this field presented Mercer's theory to solve this problem so that the support vector machine can also support problems that are not linearly separable [27]. The main idea of this theory was to transfer a vector such as X from a limited space (input space) to a higher space of feature space or Hilbert space by using the Hilbert transform and classifying it in the higher space. In this situation, a vector-like X is written as $\varphi(x)$ in a higher space. Fig. 7 shows how to transfer from the input space to the attribute space.

By using this theorem, by converting X to $\varphi(x)$ in equations (16) and (18), it is easy to obtain the relationships of the weight vector and the bias of the nonlinear support vector machine in the form of (19).

$$\mathrm{w}_0 = \sum_{\mathrm{i}=1}^{\mathrm{N}} \alpha_i \varphi(x_i) y_i$$
$$\mathrm{b}_0 = \frac{1}{\mathrm{SV}} \sum_{i=1}^{sv} (y_i - w^T \varphi(x_i)) \tag{19}$$

And finally, by applying the above relations, the decision function of the non-linear support vector machine will be in the form of the (20).

$$\mathrm{d(x)} = \mathrm{sign}(w^T \varphi(x_i) + b) \tag{20}$$

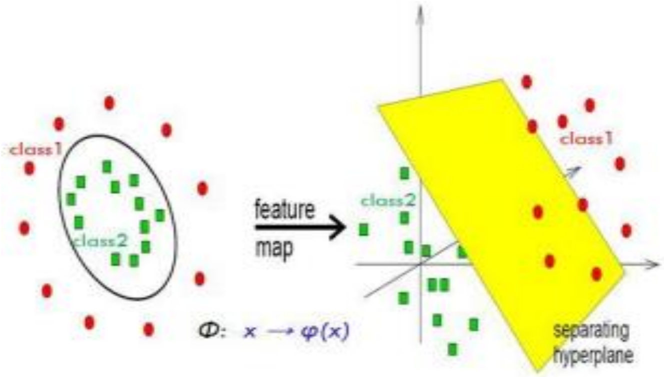

Fig. 7. Transferring the input space (left) to the feature space (right) by Hilbert transform

4. Results and Discussion

4.1. Proposed Algorithm

Our goal was to design and implement an algorithm to check the identity of people by applying the SVM method to facial images. In identity verification systems, a person's claim about her identity is checked, and finally, her claim is confirmed or rejected. Among the possible errors in this type of system, we can point to false acceptance or false rejection. The features used in the above system are special faces. After normalization, the special faces of the images are calculated with the algorithm presented in Section 2 for each person and the vectors obtained from the special faces are used for SVM training. As mentioned earlier, SVM performs classification between two classes, so if there are N people in the system, N-1 machines are trained for each person. So, we will have an N(N-1) number of SVMs in the system. In our proposed algorithm, C is calculated, and its best value is selected for each machine. The algorithm to find the best value of C is a binary search. In this way, at the time of training, first, a small value and a large value are considered for C, and after calculating the average of the two values, three SVMs are trained with them. In the next step, the input data is tested with the above machines, and the two machines with the lowest error value are selected. This work continues until the error is minimized and the best classification is performed, although, in our system, this error is zero for all training data. After finding the optimal value of C, the above machine has been tested with the obtained parameters and training inputs, and in addition to the parameters obtained from the machines, such as the values and number of support vectors, α coefficients and the optimal value of C, the output of the decision function is also stored.

In the testing phase, when a person introduces himself as a special person, his facial image is used as the input of the system, and first, the feature vector based on the person's special faces (using the average and face special features obtained in the training phase) are calculated. The resulting vector is then applied as input to all N-1 trained machines belonging to the claimant. At this stage, it can be assumed in first place that if all N-1 machines perform the classification operation correctly, the person's claim will be confirmed and rejected otherwise. In this case, the test data may be located in the area of one class, but it is very close to the area or one of the training data of another class. To avoid this error, the test algorithm was corrected as follows: if 2N or fewer machines classify the person in a class other than the claimed class, the person's claim is rejected. Otherwise, the output of the decision function is compared with the values of this function obtained from the test data, and if this distance is less than a threshold value, the obtained class is wrong, and the sample is assigned to the second class. Their distance is used to compare the output of the decision function during testing and training. But since there are several samples of the individual available at the time of training and we will have several values of the output function, one way to calculate the distance of the output of the function at the time of testing is from the average of the outputs of the function at the time of training. On the other hand, since the images of training faces are prepared from different angles and with different lighting conditions, and therefore their average is not very stable, so instead of the previous solution, it is better to compare the output of the decision function during the test with all the outputs obtained in the step of training that is calculated and the smallest distance is selected among them. In this case, if the person has turned his head slightly to the left in the test sample, he will have a very small distance from his similar image in the training samples.

After calculating the smallest distances in machines that have been misclassified, the largest distance is selected if this value is less than a threshold value (i.e., when all the distances to the test samples are sufficiently close). In this case, the person's claim is confirmed, and otherwise, it is rejected.

4.2. Implementation Results

ORL face database images that are shown in Fig. 8 were used for training and testing. This database contains 10 facial images of 40 people (400 facial images in total), and the images of each person are made at different times and with changes in lighting, different poses (open and closed eyes, smile and no smile) and other situations such as with glasses and without glasses. The above face images are in PGM format, with a size of 112×92 pixels and 256 gray levels. Out of 10 images of each person, 5 images were used in the training phase, and 5 images were used in the testing phase. The educational images are intelligently selected so that they have the maximum information that is not dependent on each other. From the 200 images used for training, a number of 200 special faces are obtained, and each image can be coded with a number of these special faces and a feature vector (which has the number of selected special faces) gained. In the training phase, 5 vectors obtained as class -1 and 5-second odd vectors as class +1 are applied to SVM with RBF internal multiplication kernel, and then the optimal C is calculated. In this way, 39 machines are trained for each person, and the parameter values are saved. In the test phase, the remaining 200 images were applied as input to the system. At this stage, 40 claims are made for each image, only one of which is true (a total of 8000 claims were checked). The obtained results include FAR=0.23% and FRR=0.48%, which gives an average recognition rate of 98.9% and an error rate of 0.4%, which is a very good result considering the variety of face images in the ORL database. The training time in this method is relatively long due to many calculations. For example, it takes 520 seconds to train the above system for 40 people with 200 images (5 images for each person) and 200 special faces. But the test time for each identity check is about 0.0625 seconds, which is completely acceptable.

Fig. 8. A number of images from the ORL database

5. Conclusion

In this article, the combination of the SVM method and special images is used to check the identity based on the face image. To improve the performance of the SVM method, the following four steps have been taken. 1: With a method inspired by binary search, the best value of C is calculated in the training phase. 2: An algorithm for the optimal comparison of the alleged person's image with other images is presented. 3: Instead of calculating the distance with the average training images in the test phase, the distance with the closest training sample is used. The educational images are intelligently selected so that they have the maximum information that is not dependent on each other. As a result, a very high recognition rate has been achieved. Compared to [28], which is a face recognition system with SVM and tested with the ORL database, we see that our proposed algorithm is more accurate. This method has a lot of calculations in the training phase, but instead, the testing phase is not time-consuming because only one feature vector is calculated for the face image (with the help of special faces calculated during the training). The next step is to use this system in real-time and optimize the above algorithm by reducing its complexity.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] J. Ashbourn, "Identification, verification and templates," in *Biometrics: Advanced Identity Verification: The Complete Guide*: Springer, pp. 65-78, 2000, https://doi.org/10.1007/978-1-4471-0747-7_4.

[2] T. Kamyab, A. Delrish, H. Daealhaq, A. M. Ghahfarokhi, and F. Beheshtinejad, "Comparison and Review of Face Recognition Methods Based on Gabor and Boosting Algorithms," *International Journal of Robotics and Control Systems*, vol. 2, no. 4, pp. 610-617, 2022, https://doi.org/10.31763/ijrcs.v3i1.849.

[3] T. Kamyab, H. Daealhaq, A. M. Ghahfarokhi, F. Beheshtinejad, and E. Salajegheh, "Combination of Genetic Algorithm and Neural Network to Select Facial Features in Face Recognition Technique," *International Journal of Robotics and Control Systems*, vol. 3, no. 1, pp. 50-58, 2023, https://doi.org/10.31763/ijrcs.v3i1.849.

[4] A. Yousefiankalareh, T. Kamyab, A. M. Ghahfarokhi, F. Beheshtinejad, H. Mirzanejad, and S. Seddighi, "Utilizing Multi-Agent Systems Approach in Firefly Algorithm," in *2021 International Conference of Modern Trends in Information and Communication Technology Industry (MTICTI)*, pp. 1-6, 2021, https://doi.org/10.1109/MTICTI53925.2021.9664757.

[5] I. Yoosefdoost, M. Basirifard, J. Álvarez-García, and M. d. l. C. del Río-Rama, "Increasing Agricultural Resilience through Combined Supply and Demand Management (Case Study: Karaj Reservoir Dam, Iran)," *Agronomy*, vol. 12, no. 9, p. 1997, 2022, https://doi.org/10.3390/agronomy12091997.

[6] I. Yoosefdoost, M. Basirifard, and J. Álvarez-García, "Reservoir Operation Management with New Multi-Objective (MOEPO) and Metaheuristic (EPO) Algorithms," *Water*, vol. 14, no. 15, p. 2329, 2022, https://doi.org/10.3390/w14152329.

[7] A. Yousefiankalareh, T. Kamyab, F. Shahabi, E. Salajegheh, H. Mirzanejad, and M. M. Masouleh, "Face recognition based on sparse coding using support vector machine classifier," in *2021 International Conference on Intelligent Technology, System and Service for Internet of Everything (ITSS-IoE)*, pp. 1-5, 2021, https://doi.org/10.1109/ITSS-IoE53029.2021.9615322.

[8] B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: Global versus component-based approach," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 2, pp. 688-694, 2001, https://doi.org/10.21236/ADA459707.

[9] P. Phillips, "Support vector machines applied to face recognition," *Advances in neural information processing systems*, vol. 11, 1998, https://doi.org/10.6028/NIST.IR.6241.

[10] S. Taghiyeh, D. C. Lengacher, A. H. Sadeghi, A. Sahebifakhrabad, and R. B. Handfield, "A multi-phase approach for product hierarchy forecasting in supply chain management: application to MonarchFx Inc," *arXiv preprint arXiv:2006.08931*, 2020, https://doi.org/10.48550/arXiv.2006.08931.

[11] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 1999, https://doi.org/10.1007/978-1-4757-3264-1.

[12] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273-297, 1995, https://doi.org/10.1007/BF00994018.

[13] X. Wang, S. M. Abtahi, M. Chahari, and T. Zhao, "An adaptive neuro-fuzzy model for attitude estimation and control of a 3 DOF system," *Mathematics*, vol. 10, no. 6, p. 976, 2022, https://doi.org/10.3390/math10060976.

[14] D. Q. Zeebaree, H. Haron, and A. M. Abdulazeez, "Gene selection and classification of microarray data using convolutional neural network," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, pp. 145-150, 2018, https://doi.org/10.1109/ICOASE.2018.8548836.

[15] M. Sajjadi, M. Chahari, H. N. Pishkenari, and G. Vossoughi, "Designing nonlinear observer for topography estimation in trolling mode atomic force microscopy," *Journal of Vibration and Control*, vol. 28, no. 23-24, pp. 3890-3905, 2022, https://doi.org/10.1177/10775463211038140.

[16] O. Mohammadrezapour, I. Yoosefdoost, and M. Ebrahimi, "Cuckoo optimization algorithm in optimal water allocation and crop planning under various weather conditions (case study: Qazvin plain, Iran)," *Neural Computing and Applications*, vol. 31, pp. 1879-1892, 2019, https://doi.org/10.1007/s00521-017-3160-z.

[17] E. Osuna, R. Freund, and F. Girosit, "Training support vector machines: an application to face detection," in *Proceedings of IEEE computer society conference on computer vision and pattern recognition*, pp. 130-136, 1997, https://doi.org/10.1109/CVPR.1997.609310.

[18] S. Kumar, S. Singh, and J. Kumar, "Multiple face detection using hybrid features with SVM classifier," in *Data and Communication Networks: Proceedings of GUCON 2018*, pp. 253-265, 2019, https://doi.org/10.1007/978-981-13-2254-9_23.

[19] H. I. Dino and M. B. Abdulrazzaq, "Facial expression classification based on SVM, KNN and MLP classifiers," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, pp. 70-75, https://doi.org/10.1109/ICOASE.2019.8723728.

[20] L. Shi, X. Wang, and Y. Shen, "Research on 3D face recognition method based on LBP and SVM," *Optik*, vol. 220, p. 165157, 2020, https://doi.org/10.1016/j.ijleo.2020.165157.

[21] H. S. Dadi and G. M. Pillutla, "Improved face recognition rate using HOG features and SVM classifier," *IOSR Journal of Electronics and Communication Engineering*, vol. 11, no. 04, pp. 34-44, 2016, https://doi.org/10.9790/2834-1104013444.

[22] H. Chen and C. Haoyu, "Face recognition algorithm based on VGG network model and SVM," in *Journal of Physics: Conference Series*, vol. 1229, no. 1, p. 012015, 2019, https://doi.org/10.1088/1742-6596/1229/1/012015.

[23] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103-108, 1990, https://doi.org/10.1109/34.41390.

[24] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991, https://doi.org/10.1162/jocn.1991.3.1.71.

[25] M. G. Ramos and S. S. Hemami, "Eigenfeatures coding of videoconferencing sequences," in *Visual Communications and Image Processing'96*, vol. 2727, pp. 100-110, 1996, https://doi.org/10.1117/12.233178.

[26] S. Abe, *Support vector machines for pattern classification*. Springer, 2005, https://doi.org/10.1007/978-1-84996-098-4.

[27] M. Martínez-Ramón and C. Christodoulou, "Support vector machines for antenna array processing and electromagnetics," *Synthesis Lectures on Computational Electromagnetics*, vol. 1, no. 1, pp. 1-120, 2005, https://doi.org/10.2200/S00020ED1V01Y200604CEM005.

[28] K. I. Kim, J. Kim, and K. Jung, "Recognition of facial images using support vector machines," in *Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing (Cat. No. 01TH8563)*, pp. 468-471, 2001, https://doi.org/10.1109/SSP.2001.955324.