

## Research Article

# Video Face Detection Technology and Its Application in Health Information Management System

Yuxuan Liao,<sup>1</sup> Zhenyu Tang,<sup>2</sup> Jiehong Lei,<sup>1</sup> Jiajia Chen,<sup>1</sup> and Zhong Tang<sup>3</sup> 

<sup>1</sup>School of Information and Management, Guangxi Medical University, Nanning, Guangxi 530021, China

<sup>2</sup>Department of Computer Engineering, University of CA, Irvine, CA 92697, USA

<sup>3</sup>School of Humanities and Social Sciences, Guangxi Medical University, Nanning, Guangxi 530021, China

Correspondence should be addressed to Zhong Tang; tangzhong@stu.gxmu.edu.cn

Received 9 December 2021; Revised 2 January 2022; Accepted 10 January 2022; Published 4 February 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yuxuan Liao 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.

Computer face detection, as an early step and prerequisite for applications such as face recognition and face analysis, has attracted people's attention for a long time. With the popularization of computer applications, the improvement of performance, and the gradual maturity of research in the field of image processing and pattern recognition, face-related applications have become more and more a reality, so the research on face detection and positioning is also receiving more and more attention. Face detection and positioning are an important part of face analysis technology. Its goal is to search for the location of facial features (such as eyes, nose, mouth, and ears) in images or image sequences. It can be widely used in the fields of face tracking, face recognition, gesture recognition, facial expression recognition, head image compression and reconstruction, facial animation, etc. Based on the health information management system, this study mainly discusses the application of face recognition technology in video systems. Compared with other biological characteristics, such as fingerprints and eye masks, human faces are easier to obtain. In research and exploration, stable and effective face detection and face recognition algorithms have been proposed, which can achieve good recognition results even in real-time video surveillance. Aiming at the automatic face recognition technology in video surveillance, this study introduces in detail the video face detection technology in the health information management system of video image collection, image preprocessing, face detection, and face recognition. The prototype system of hygiene management is recognized.

## 1. Introduction

Face recognition is a kind of biometric identification technology based on human facial feature information. The research on face recognition (as shown in Figure 1) can be traced back to a study published by Galton in "Nature" magazine in 1888, which proposed the use of face information for personal identification, but it does not involve the problem of timely recognition of faces. In 1965, Bledsoe\_ proposed a semiautomatic face recognition system model, which is the earliest research on face detection and recognition. The model is mainly based on the geometric features of the face, such as facial features and skin color. Through the standard two-dimensional pattern recognition method, the face image is expressed as geometric feature parameters. At the end of the 1970s, a real

face recognition system began to appear [1–10]. In 1973, Dr. Kanade used the method of integration to design a backtracking recognition system based on a certain amount of artificial assistance so that the face can achieve a good matching recognition. Parke has established a high-quality face grayscale model through a computer, but the detection and recognition of human faces require prior knowledge of these human faces and cannot achieve the function of automatic recognition. By the 1990s, on the basis of a large number of scientific research results, face recognition has also received more and more attention from people around the world. Many classic algorithms are produced, among which K-L transformation achieves the purpose of reducing the dimensionality of facial image features, which becomes the basis of the later principal component analysis method. Subsequently, the eigenface method based on the



FIGURE 1: Face recognition.

template has appeared one after another, which has played a very good effect in the actual system application. The method based on the template has a better effect on face detection and recognition than the method based on geometric feature information. Fisher's linear discriminant analysis method developed on the basis of principal components uses linear judgment analysis to select the most advantageous feature points, and it is still widely used. After 1998, the system based on face detection and recognition began to enter practical research. The entire system was guaranteed to be automatically completed without human involvement. The focus of the research was also on solving the detection rate caused by illumination and complex background. On the low issue, support-vector machines and three-dimensional illumination models are applied to face detection, making the system closer to the thinking of real people, and the researched system is also more commercialized and has achieved good market response [11–15].

The main methods of face recognition technology include recognition methods based on geometric features, recognition methods based on algebraic features, recognition methods based on connection mechanisms, face recognition technologies based on three-dimensional data, and hybrid recognition methods of several technologies. A series of related technologies use cameras or cameras to capture images or video streams containing human faces, automatically detect and track human faces in the images, and then perform face recognition on the detected faces.

With the recognition method based on geometric features, the geometric feature vector represents the face, and the feature vector is constructed by extracting some feature points on the front contour line from the face profile and is used to recognize the face. Recognition methods based on algebraic features are also called eigenface recognition, including principal component analysis method, independent component analysis, and methods based on hidden Markov

models. The principal component analysis party performs orthogonal transformation according to the statistical characteristics of the image and obtains the eigenvectors with successively decreasing eigenvalues, that is, eigenfaces. The high-dimensional space representation of the image can be converted to the low-dimensional space through orthogonal transformation. The independent component analysis method performs some linear decomposition of the observed data. Hidden Markov models used for face recognition are one-dimensional hidden Markov models and two-dimensional hidden Markov models. The one-dimensional orthogonal model first transforms the two-dimensional image signal into a one-dimensional observation sequence and divides the face image into strip regions from top to bottom, and each region represents a state in the continuous hidden Markov model. Recognition methods based on connection mechanism mainly include artificial neural networks and elastic matching methods. The neural network is a complex system composed of a large number of simple processing units (neurons) interconnected to solve recognition problems. Commonly used neural networks are BP neural network, self-organizing neural network, convolutional neural network, radial basis function network, and fuzzy neural network. The elastic matching method is a method based on a dynamic link structure. Face recognition based on three-dimensional data mainly includes a method based on curvature and a method based on model synthesis. Since the curvature is the most natural and basic local feature that expresses the surface, the curvature of the face surface was first used to deal with the 3D face recognition problem [16–22].

The above various face recognition methods have achieved certain success, but each has advantages and disadvantages. The processing of obstructions such as glasses is more difficult, and the robustness to large expression changes or posture changes is also relatively poor. The recognition method based on algebraic features extracts principal components through

different transformation methods [23–27]. However, this method has insurmountable defects, especially when the ambient light changes, and the recognition effect will sharply drop, which cannot meet the needs of the actual system. The advantage of the recognition method based on the connection mechanism is that it saves the material information in the face image while avoiding the more complicated feature extraction work. However, due to the huge amount of original image data and the time-consuming process, when the number of samples greatly increases, its performance will sharply drop. Face recognition methods based on three-dimensional data are characterized by the use of three-dimensional data, which opens up new ideas for face recognition, but it is difficult to obtain information sources, and the amount of data storage and calculation is also huge [28–33].

## 2. Image Preprocessing

The traditional face recognition technology is mainly based on the face recognition of visible light images, which is also a familiar recognition method with more than 30 years of research and development history. Face image preprocessing is an important link in the research and development of face recognition in video images, as shown in Figure 2. The video image quality is poor, the background is complex, the face is unpredictable, and the system must be guaranteed under a large amount of video image data. Those are the real-time requirements. Therefore, the quality of image preprocessing before face detection and recognition will directly affect the performance of the video system. The process is shown in Figure 3. Due to the changeable environment of video image capture, affected by the equipment itself, the captured video images often have various image quality problems. Through the continuous efforts of scientists, image preprocessing technology has also been greatly developed. A system has been formed for static digital image preprocessing, and it can achieve good results. Image enhancement, image grayscale transformation, and image normalization process the image, which will be separately described below. In face image preprocessing, image preprocessing for faces is a process of processing images based on the results of face detection and finally serving feature extraction.

*2.1. Eliminate Image Noise.* There is no specific definition of image noise, and there are different classifications in different fields. If it is defined according to the shape of the amplitude distribution generated by the noise, the amplitude distribution is called Gaussian noise according to the Gaussian distribution, and its probability density function is as follows:

$$p(z) = \frac{1}{\sqrt{2\pi}\delta} e^{-((z-\mu)^2/2\delta^2)}. \quad (1)$$

$Z$  represents the gray value of the image pixel,  $Z$  represents the gray value of the image pixel,  $\mu$  represents the average or expected value of  $Z$ , and  $\delta$  represents the standard deviation of  $Z$ . The square of the standard deviation  $\delta^2$  is called the variance of  $Z$ . Salt and pepper noise is usually generated in the image sensor, transmission channel,

decoding process, etc. The human eye can observe the bright and dark spots in black and white. The original image obtained by the system cannot be directly used due to various limitations and random interference, and it must be pre-processed such as grayscale correction and noise filtering in the early stage of image processing.

A large number of experimental studies have found that the digital images acquired by video surveillance cameras are more affected by salt and pepper noise and zero-mean Gaussian noise. Therefore, this study will also make the corresponding image preprocessing for these two types of noise. The salt and pepper noise uses median filtering, and the Gaussian noise uses wavelet analysis for denoising.

The median filter is a nonsmooth linear filter, which can eliminate noise while preventing the image from becoming blurred, protecting the details and edge information of the image. The median filter method has a good effect on filtering out salt and pepper noise.

Median filtering is to take the value of each pixel in the image in its square area and sort the pixel values in the area from small to large. If there are an odd number of pixel values, the middle one is taken as the pixel value of the corresponding position in the new image. If it contains an even number of pixel values, the average value of the middle two is taken as the pixel value of the corresponding position of the new image. This can eliminate isolated noise points. Time vs normalized frequency is shown in Figure 4.

In fact, the median filter is to select the closest value of the square pixel value of the area as the gray value of the pixel so that special points, namely, noise points, can be filtered out and that isolated noise points, namely, salt and pepper noise, can be eliminated by this method. Compared with the mean filter, the median filter does not change the value of the pixel point in the field but selects a representative point, and the resulting ambiguity is smaller. The solutions to the lighting problem include 3D image face recognition and thermal imaging face recognition. However, these two technologies are far from mature, and the recognition effect is not satisfactory.

Wavelet transform is an image analysis method based on image resolution. Through the multiscale refinement analysis of image signals, it can extract information on the tiny details of the image. It has been attracting attention in recent years. When denoising the image, it is necessary to analyze the affinity between the noise and the image information. When the overlap between the noise signal and the image signal is small, the traditional Fourier transform can be used for the denoising process, and the noise can effectively remove pollution, but when the noise signal and the image signal overlap a large part, the removal of noise by using this method will affect the image quality and damage the detailed information of the image. While transforming the noise signal, it will also transform the image signal. The feature extraction effect of the image is affected. Using wavelet transform to remove image noise can solve this problem. Wavelet analysis can remove image noise in the high-frequency region while retaining the image detail information in the high-frequency region. Power vs frequency is shown in Figure 5.

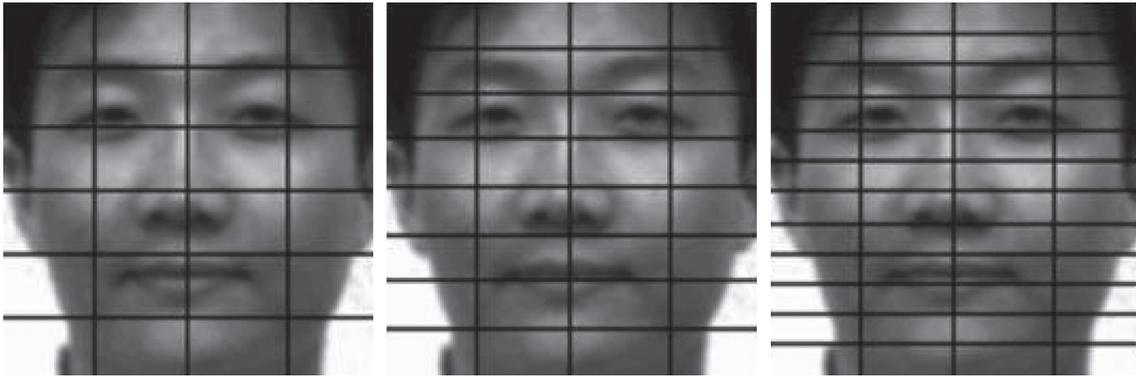


FIGURE 2: Image preprocessing.

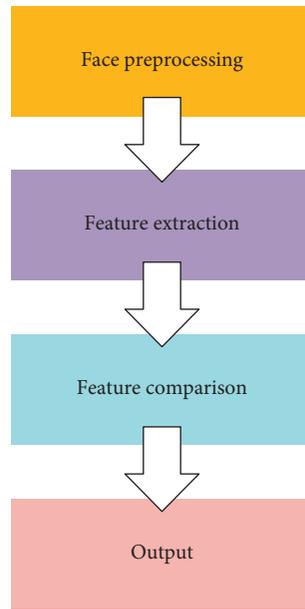


FIGURE 3: Process.

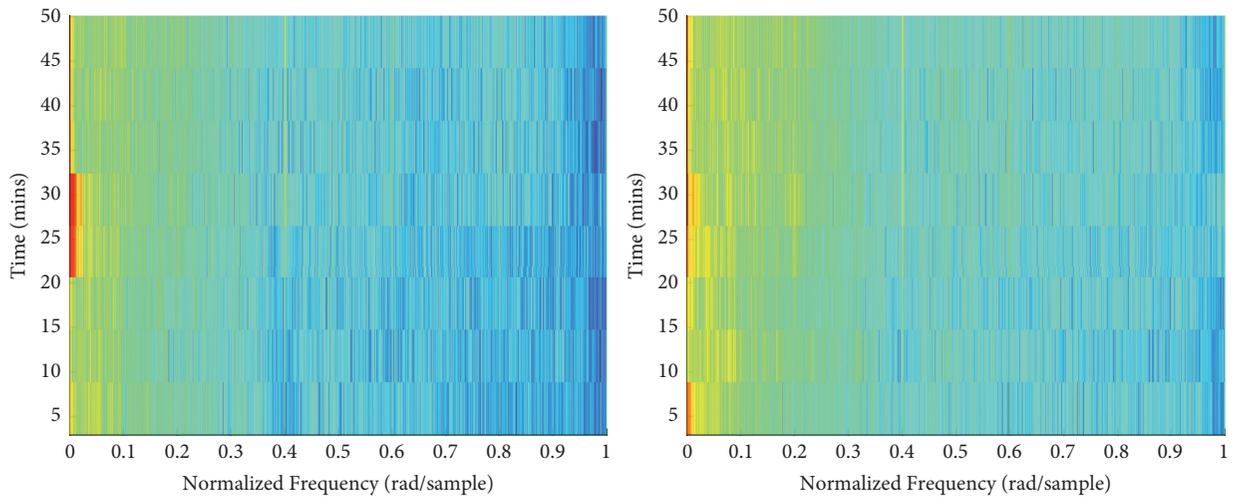


FIGURE 4: Time vs normalized frequency.

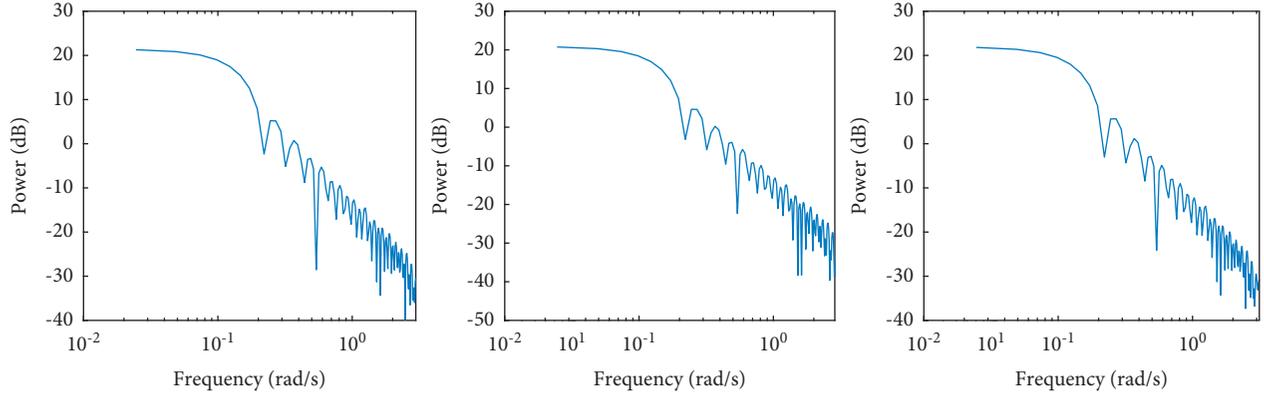


FIGURE 5: Power vs frequency.

The basic idea of wavelet analysis is as follows: a basic wavelet function is defined, operations such as scaling or translation are performed on this function, and multiscale refinement is performed to obtain a series of function clusters, which are called wavelet bases. When using wavelet transform to process an image, a set of wavelet bases is used to approximate an image signal, and then, the signal is projected onto the wavelet base to realize wavelet transform.

The basic idea of image denoising using wavelet analysis is as follows: through wavelet transform, most useful signals are compressed while dispersing noisy signals. Wavelet decomposition of image information is to project the image signal to each orthogonal basis and then obtain the relationship coefficient between the signal and a series of wavelet bases, that is, the wavelet transform value, to achieve the purpose of the wavelet decomposition. The wavelet decomposition of the image makes the image key information and the image noise information have different decomposition results. The wavelet transform value of the image's own signal only has a large value on some scales, while the noise signal after wavelet analysis is on most scales. Both have larger wavelet transform values. In this way, through the decomposition and reconstruction of the image signal, the purpose of denoising can be achieved.

The local singularity of the image signal can be described by the Lipschitz index. A function  $f(x)$  is defined with the following:

$$|f(x) - f(x_0)| \leq K|x - x_0|^\alpha. \quad (2)$$

There is a unique constant  $k$  that satisfies formula (2). If formula (2) holds for all  $x$  and  $x_0$ , then  $f(x)$  is consistent with Lipschitz  $\alpha$  in the interval  $(a, b)$ . If the function has a discontinuity or a certain order of derivative discontinuity at a point, then this point can be described as a singular point of the function. Then,

$$W_s f(x) \leq A s^\alpha. \quad (3)$$

Among them,  $W_s f(x)$  is the wavelet transform of  $f(x)$  on the scale  $s$ , and then, equation (3) becomes

$$W_{2^j} f(x_k) \leq A 2^{j\alpha}. \quad (4)$$

The logarithm of both sides is as follows:

$$\log_2 |W_{2^j} f(x_k)| \leq \log_2 A + ja. \quad (5)$$

It can be seen that the coefficients of the wavelet transform are affected by the Lipschitz exponent  $\alpha$  of the function  $f(x)$ . When  $\alpha > 0$ , the coefficients of the wavelet transform are proportional to the scale. If  $\alpha < 0$  of the function  $f(x)$ , the wavelet coefficient is inversely proportional to the scale. The Lipschitz exponent of a function reflects the singularity of the function at a certain point. The larger the scale, the larger the wavelet transform amplitude corresponding to the signal, but the smaller the wavelet transform amplitude corresponding to image noise. Therefore, different thresholds can be set on different decomposition scales, the wavelet transform smaller than the given threshold is regarded as the wavelet transform of noise, the pixel value is set to 0, and the wavelet transform larger than the threshold is regarded as the wavelet transform. The wavelet transform of the signal preserves the pixel. By reconstructing the image after wavelet transformation, the purpose of noise removal can be achieved. The prediction is shown in Figure 6.

### 2.2. Fractional Differentiation Realizes Image Enhancement.

The Cauchy definition of the fractional differential is as follows:

The Cauchy definition form is based on the traditional definition form of integer order calculus, which is an extension of integer order calculus:

$$D^\gamma f(t) = \frac{\Gamma(\gamma + 1)}{2\pi j} \int_c \frac{f(\tau)}{(\tau - t)^{\gamma+1}} d\tau, \quad (6)$$

where  $c$  is the smooth curve surrounding  $f(t)$  single value and analytic domain, it is the Gamma function, and its integral transformation form is as follows:

$$\Gamma(x) = \int_0^\infty y^{x-1} e^{-y} dy. \quad (7)$$

The Grunwald-Letnikov definition of the fractional differential is as follows:

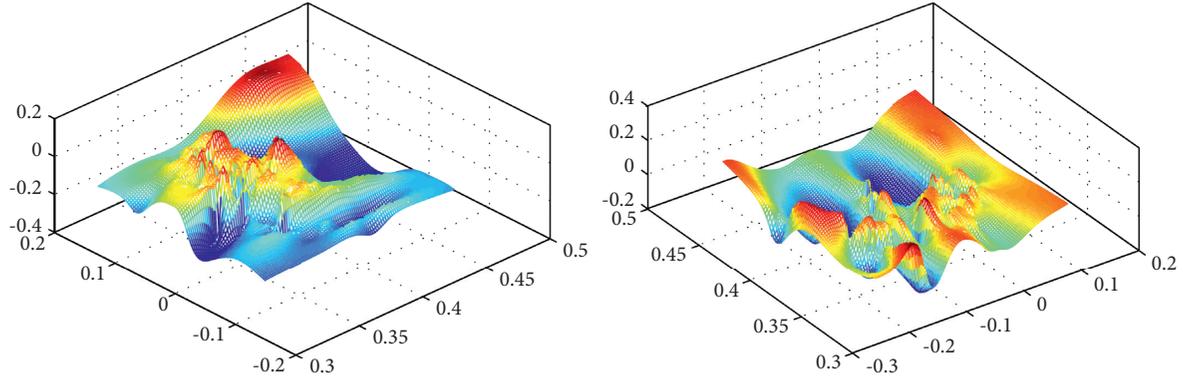


FIGURE 6: Predicted results.

$$\lim_{h \rightarrow 0} s_h^{(\nu)}(t) \approx h^{-\nu} \sum_{r=0}^{\lfloor (t-a)/h \rfloor} w_r^{(\nu)} s(t-rh), \quad (8)$$

where

$$w_r^{(\nu)} = (-1)^r \binom{\nu}{r}. \quad (9)$$

In a complete digital image, if the human eye can see a clear image effect, it means that there must be a certain correlation between adjacent pixels in this image, and the image signal is highly self-similar. The similarity is usually expressed in complex texture details. Fractional differentiation can be used to enhance the complex texture details in the image signal.

**2.3. Image Normalization Processing Based on Face.** The key point of a video surveillance system based on face recognition is to obtain as good a face image as possible and to accurately obtain the feature points of the face image when it is convenient to use. After the images captured from the video are processed by denoising, image enhancement, etc., a series of issues such as illumination and complex background need to be considered when performing face detection so that the obtained face images are as close to the same standard as possible. The image feature points will better reflect the facial features. The following will introduce how to normalize the face image. The data variation is shown in Figure 7.

When forming a standard face database, it is necessary to normalize the size of the face image and adjust the image to a certain size to facilitate feature extraction later and calculate the recognition rate. There are many methods for image enlargement and reduction. In fact, image enlargement can be regarded as excessive, and reduction can be regarded as undersight, but enlargement and reduction are applicable to digital images. Image scaling is not to simply obtain the pixels of the source image, and it is often necessary to find new pixels to replace, which can be approximated by the method of difference. The following describes two commonly used methods for calculating interpolation: nearest-neighbor interpolation and linear interpolation.

**2.3.1. Nearest Neighbor Interpolation.** The nearest-neighbor interpolation is to select an input pixel value closest to the desired pixel as the pixel gray value of the point. It is a simple interpolation method with intuitive thinking, but the quality of the image obtained is not high. When there are slight changes in the pixel gray level in the image, the processed image obtained by this method will be unnatural and appear as artificial traces.

**2.3.2. Bilinear Interpolation.** Bilinear interpolation is a method of equally dividing the width and height of the input image according to the width and height of the output image to determine the gray value of the output image. The interpolation function is as follows:

$$f(R_1) \approx \frac{x_2 - x}{x_2 - x_1} f(p_{11}) + \frac{x - x_1}{x_2 - x_1} f(p_{21}), \quad (10)$$

$$f(R_2) \approx \frac{x_2 - x}{x_2 - x_1} f(p_{12}) + \frac{x - x_1}{x_2 - x_1} f(p_{22}).$$

Then, linear interpolation is performed in the  $y$  direction to get

$$f(P) \approx \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2). \quad (11)$$

The result of linear interpolation has nothing to do with the order of interpolation. Regardless of whether the interpolation in the  $y$  direction is performed first, or the interpolation in the  $x$  direction is performed first, the result obtained will not change.

### 3. Face Recognition

Face image feature extraction is as follows: the features that can be used by a face recognition system are usually divided into visual features, pixel statistical features, face image transformation coefficient features, and face image algebraic features. With the continuous development of science and technology in recent years, face recognition technology has also achieved good developmental effects in different fields, and many practical technical methods have emerged. Under different lighting conditions and complex backgrounds, face

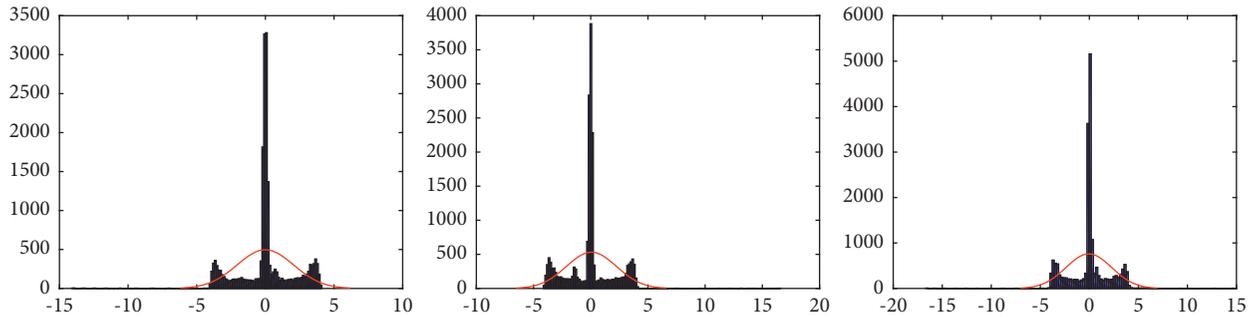


FIGURE 7: Data variation.

recognition technology has made great progress. It can achieve the expected effect in a specific application environment, but there is still a certain distance from the recognition ability of the human eye itself. Obtaining it is easier than fingerprints, eye masks, etc., but it also leads to a variety of facial features, which are more unpredictable, and it is difficult for computers to express all possible conditions through numbers. The prediction is compared in Figure 8.

At present, the common basic algorithms for face recognition can be divided into several categories: methods based on geometric features, subspace methods, neural network methods, and hidden Markov methods. It involves computer vision, pattern recognition, image processing, neural network, and other related discipline theories.

The method based on geometric features is one of the earliest face recognition methods. In fact, a face image can be regarded as composed of eyes, nose, mouth, etc. The difference between a face and a face lies in these face components and is not the same as the difference between color, size, shape, and so on. The method based on geometric features is to perform face recognition through the description of geometric features such as eyes, nose, mouth, and face shape, to obtain the geometric relationship between these geometric features, which is represented by feature vectors, angles, and curvatures, and then match these features. The similarity between vectors is usually calculated by Euclidean distance. The classic algorithms based on geometric features include the active contour model proposed by Huang to extract the contours of face parts and the deformable template model proposed by Grenander. The face recognition algorithm based on geometric features is simple and easy to understand, adapts to certain lighting changes, and has a small storage space; however, a unified feature extraction standard is not formed, the obtained features are not stable enough, and they cannot achieve good results in the case of complex expressions and side faces, such as face recognition effect.

The main idea of face recognition based on subspace is to map an image to a low-dimensional subspace through a certain spatial transformation, thereby reducing the complexity of classifying face images in high-level spaces and improving computational efficiency. The main subspace methods at this stage include principal component analysis (PCA), linear discriminant analysis (LDA), independent component analysis (ICA), etc. Turk first used the PCA

method for face recognition and developed it into the later eigenface method. Its main idea is to obtain eigenvalues and corresponding eigenvectors by performing KL transformation on the feature space and arranging them in descending order, before selecting  $N$  feature vectors are used as principal components to obtain a low-dimensional face vector space to achieve the purpose of dimensionality reduction. The main idea of the linear discriminant analysis (LDA) method is to find a linear transformation so that the distance between the classes is the largest and the distance within the classes is the smallest after the training sample is projected to the linear transformation. When there are many training samples, this method is better than the PCA algorithm. Face feature extraction is carried out for certain features of the face. Facial feature extraction, also known as face representation, is the process of modeling features of human faces.

The main idea of the independent component analysis (ICA) method is to find a set of mutually independent components from the training sample through linear transformation and use this to describe the sample data containing high-order statistical information. The ICA algorithm is a generalization of the PCA algorithm. It is widely used in the field of face recognition. The neural network-based face recognition method mainly uses neural network theory for facial feature extraction and recognition. The neural network simulates human thinking through computers and strives to be closer to the learning process of the human brain.

Completing complex tasks by coordinating a large number of simple arithmetic processing (neurons) has certain advantages for face processing with complex structures. The advantage of the neural network is that there is no fixed process, but it is constantly adjusted according to the changes in conditions and environment to obtain the inherent laws and characteristics of the face image. The disadvantage is that the number of neurons is large and the training time is long, which is difficult to implement in practice.

The hidden Markov model (HMM) for face recognition was first introduced by Samaria and then widely used in the field of face recognition. The hidden Markov model is based on statistical analysis and is a form of Markov chain. It consists of a Markov chain with specific state values and a set of random functions to form a bidirectional random process,

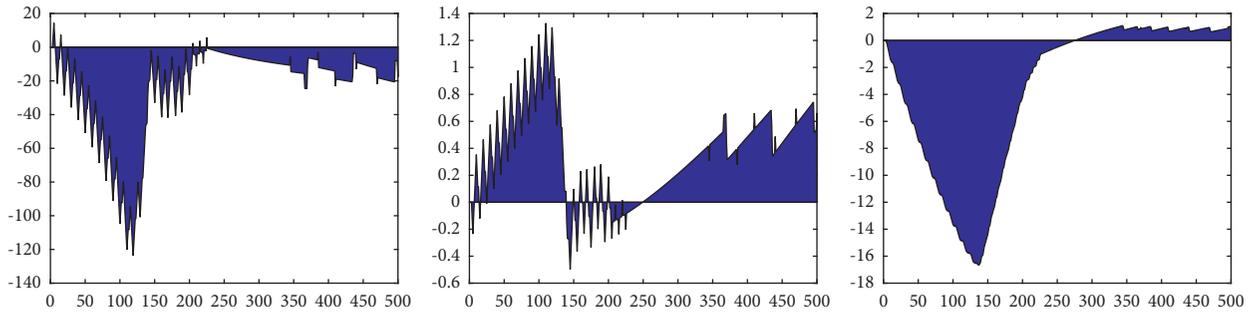


FIGURE 8: Prediction comparison.

which can extract the face image. Interrelationships between various organs are as follows. Face recognition based on the HMM model also has better recognition results when the facial expressions are complex and the face scale greatly changes, but the feature extraction and model training operations are large.

In addition, there are methods for face recognition based on elastic graph matching, support-vector machines, and other methods. In actual applications, it is often difficult to meet the needs by using only one method. Therefore, it is necessary to use the advantages and disadvantages of various methods and the scope of application. Different methods are combined to achieve better recognition results.

Face recognition has gradually entered the research of three-dimensional face recognition from the initial recognition of frontal two-dimensional images to the subsequent recognition in multiple poses and complex environments. With the deepening of research, the performance of face recognition systems has also continued. However, due to the complex environment people are in and the influence of age, face occlusion, etc., the robustness of the face recognition system still needs to be improved. The face features extracted by the SIFT algorithm have strong robustness and rich information and have good feature extraction effects under the interference of image scale transformation, noise, brightness, etc., and have been widely used in images in recent years. such as retrieval, image matching, and image tracking, and this method has been proven to have higher matching performance than other local features of the same type and has high research value and practical significance in the field of face recognition. The “eigenface” method represented by PCA is a benchmark algorithm for detecting the performance of face recognition systems. This study will analyze the specific application of the PCA algorithm and SIFT algorithm in face recognition and compare the advantages and disadvantages of the two through simulation experiments. The recognized portrait is shown in Figure 9.

This part introduces the specific implementation process of face recognition using the PCA algorithm on the ORL face database.

- (1) The face database is read, the ORL database contains a total of  $40 \times 10$  images, and each image is  $92 \times 112$  in size. Each image is converted into a row vector of  $92 \times 112$  columns, and 6 images of each person are selected as training samples, that is, 240 images. The

image forms a matrix of 240 rows and  $92 \times 112$  columns. The formula is to calculate the average image vector of the training sample to get the average face, then calculate the interpolation between each image and the average face, and then build the covariance matrix from the formula.

- (2) The singular value decomposition method is used to calculate the eigenvalues and eigenvectors of the covariance matrix, arrange the eigenvalues and their corresponding eigenvectors in descending order, and select the top P largest eigenvalues and their corresponding eigenvectors. These P feature vectors (principal components) are constructed into a feature face space.
- (3) Six images of each person are selected as training samples, and the remaining 4 images are selected as test samples, and the difference between the training sample face and the average face to the eigenface space is mapped so that the matrix can be used as the basis for face recognition. Each image of the test sample is mapped to the eigenface space, and the obtained vector can be used as a comparison with the eigenface space of the training sample.
- (4) There are many ways to compare the test sample and the training sample. To find out which sample the face image to be recognized belongs to, you can first simplify the 6 images of the training sample to form a feature that can approximately represent this class vector, and the simplest way is to do the arithmetic average of 6 vectors, compare the test sample with 40 feature vectors representing 40 classes, and divide it into the closest class; you can also compare the sample to be identified with the training sample, compare each feature vector of, and classify it into the category of the image closest to it. The disadvantage is that if there is a special situation, it happens that there is an image of another category that is closest to it, which will cause and avoid misclassification. In this case, you can find the closest few pictures and divide them into the category with the largest number of similarities. In this experiment, the calculation of the distance between the vector and the vector uses the Euclidean distance metric.

The PCA algorithm and the SIFT algorithm are used to perform face recognition on the ORL database and the



FIGURE 9: Recognized portrait.

laboratory database under the condition of selecting different training samples and test samples. The overall face recognition efficiency is not as high as the recognition rate on the ORL database; as the number of samples increases, the recognition rate on both databases accordingly increases. The overall quality of the face images of the laboratory database is not as high as that of the ORL database and is affected by illumination and posture, so the recognition rate is not as high as the recognition rate on ORL.

The PCA algorithm is a characteristic face method. The entire recognition process is for the entire face image, not for the characteristics of the face. Therefore, the recognition result is greatly affected by the light, facial expressions, etc., and the difference of the characteristic face is obtained under different brightness of the big image. The PCA algorithm is only able to describe the difference between images not human faces. Different from the PCA algorithm, the SIFT algorithm realizes face recognition by extracting local feature points of the image, especially the extracted feature points are concentrated around the eyes, mouth, nose, etc. The algorithm has good invariance under different conditions such as rotation, scale scaling, and illumination changes. The feature information extracted by the SIFT algorithm is rich, and a large number of feature vectors can be extracted from a small amount of image samples. Mikolajczyk obtained through experiments that the matching performance of SIFT descriptors is higher than that of other local feature descriptors.

Through experiments, it is found that, compared to the PCA algorithm, the face recognition based on the SIFT algorithm can extract feature points for recognition without image normalization; for most algorithms, it is necessary to establish a training library first and then recognize the test samples after training. The SIFT algorithm does not need to perform sample training in advance, as long as the video image is collected, and the extracted facial features are matched with the sample features of the sample library. In summary, the SIFT algorithm has certain robustness for face recognition in the case of poses, expressions, and simple illumination changes. Although the feature vector extracted by the SIFT algorithm has 128 dimensions, it takes up a lot of

space in the matching operations of a large number of databases, which affects the speed of face recognition, but is constantly improving and expanding, such as the SURF algorithm that reduces the dimensionality of the feature vector, the PCA-SIFT algorithm that combines with the PCA algorithm, and so on.

#### 4. Conclusions

Based on the health information management system, this study mainly discusses the application of face recognition technology in video systems. Compared with other biological characteristics, such as fingerprints and eye masks, human faces are easier to obtain. In research and exploration, stable and effective face detection and face recognition algorithms have been proposed, which can achieve good recognition results even in real-time video surveillance. Aiming at the automatic face recognition technology in video surveillance, this study introduces in detail the video face detection technology in the health information management system of video image collection, image pre-processing, face detection, and face recognition. The prototype system of hygiene management is recognized.

#### Data Availability

The dataset can be accessed upon request.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

This work was supported by the Guangxi Health and Economic and Social Development Research Center Project (Research on the Practice Mechanism of Social Work in Public Health Emergencies in Guangxi; no. 2021RWB03) and 2021 Guangxi Philosophy and Social Science Planning Research Project (Research on the Construction of Urban Smart Home-Based Elderly Care Service System under the Healthy Guangxi Strategy; no. 21FSH017).

## References

- [1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [2] M. Yun, J. Zhao, J. Zhao, X. Weng, and X. Yang, "Impact of in-vehicle navigation information on lane-change behavior in urban expressway diverge segments," *Accident Analysis & Prevention*, vol. 106, no. 1, pp. 53–66, 2017.
- [3] S. Kumar Dwivedi, R. Amin, V. Satyanarayana, and R. Chaudhry, "Blockchain-based secured event-information sharing protocol in internet of vehicles for smart cities," *Computers & Electrical Engineering*, vol. 86, no. 1, pp. 1–9, 2020.
- [4] Z. Khan and S. Amin, "Bottleneck model with heterogeneous information," *Transportation Research Part B: Methodological*, vol. 112, no. 1, pp. 157–190, 2018.
- [5] J. M. Cairney, K. Rajan, D. Haley, and B. Gault, "Mining information from atom probe data," *Ultramicroscopy*, vol. 159, no. 1, pp. 324–337, 2020.
- [6] J. Yu and P. Lu, "Learning traffic signal phase and timing information from low-sampling rate taxi GPS trajectories," *Knowledge-Based Systems*, vol. 110, no. 1, pp. 275–292, 2016.
- [7] K. P. Wijayarathna, V. V. Dixit, L. Denant-Boemont, and S. T. Waller, "An experimental study of the online information paradox: does en-route information improve road network performance?" *PLoS One*, vol. 12, no. 9, pp. 184–191, 2017.
- [8] Z. Wang, H. Ren, Q. Shen, W. Sui, and X. Zhang, "Seismic performance evaluation of a steel tubular bridge pier in a five-span continuous girder bridge system," *Structures*, vol. 31, no. 1, pp. 909–920, 2021.
- [9] S. Nakayama and J. Takayama, "Traffic network equilibrium model for uncertain demands," in *Proceedings of the 82nd Transportation Research Board Annual Meeting*, Washington, DC, USA, 2021.
- [10] H. Shao, W. H. K. Lam, and M. L. Tam, "A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand," *Networks and Spatial Economics*, vol. 6, no. 3, pp. 173–204, 2019.
- [11] A. Chen, J. Kim, S. Lee, and Y. Kim, "Stochastic multi-objective models for network design problem," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1608–1619, 2020.
- [12] H. Wang, W. H. K. Lam, X. Zhang, and H. Shao, "Sustainable transportation network design with stochastic demands and chance constraints," *International Journal of Sustainable Transportation*, vol. 9, no. 2, pp. 126–144, 2015.
- [13] S.-M. Hosseininasab and S.-N. Shetab-Boushehri, "Integration of selecting and scheduling urban road construction projects as a time-dependent discrete network design problem," *European Journal of Operational Research*, vol. 246, no. 3, pp. 762–771, 2015.
- [14] S.-M. Hosseininasab, S.-N. Shetab-Boushehri, S. R. Hejazi, and H. Karimi, "A multi-objective integrated model for selecting, scheduling, and budgeting road construction projects," *European Journal of Operational Research*, vol. 271, no. 1, pp. 262–277, 2018.
- [15] M. Johnson, M. Schuster, Le et al., "Google's multilingual neural machine translation system: enabling zero-shot translation," *Transactions of the Association for Computational Linguistics*, vol. 5, no. 1, pp. 339–351, 2017.
- [16] M. D. Moreno, "Translation quality gained through the implementation of the iso en 17100:2015 and the usage of the blockchain," *Babel*, vol. 1, no. 2, pp. 1–9, 2020.
- [17] X. Wang, X. Yu, L. Guo, F. Liu, and L. Xu, "Student performance prediction with short-term sequential campus behaviors," *Information*, vol. 11, no. 4, p. 101, 2020.
- [18] Q. Guo, Z. Zhu, Q. Lu, D. Zhang, and W. Wu, "A dynamic emotional session generation model based on Seq2Seq and a dictionary-based attention mechanism," *Applied Sciences*, vol. 10, no. 6, pp. 1–10, 2020.
- [19] H. Ren, Xi Mao, W. Ma, J. Wang, and L. Wang, "An English-Chinese machine translation and evaluation method for geographical names," *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, pp. 193–201, 2020.
- [20] J. Arús-Pous, T. Blaschke, S. Ulander, J.-L. Reymond, H. Chen, and O. Engkvist, "Exploring the GDB-13 chemical space using deep generative models," *Journal of Cheminformatics*, vol. 11, no. 1, pp. 20–29, 2019.
- [21] T. Moon, T. In Ahn, and J. E. Son, "Long short-term memory for a model-free estimation of macronutrient ion concentrations of root-zone in closed-loop soilless cultures," *Plant Methods*, vol. 15, no. 1, pp. 1–12, 2019.
- [22] N. Pourdamghani and K. Knight, "Neighbors helping the poor: improving low-resource machine translation using related languages," *Machine Translation*, vol. 33, no. 3, pp. 239–258, 2019.
- [23] L. Bote-Curiel, S. Muñoz-Romero, A. Guerrero-Curieses, and J. L. Rojo-Álvarez, "Deep learning and big data in healthcare: a double review for critical beginners," *Applied Sciences*, vol. 9, no. 11, pp. 1–11, 2019.
- [24] J. Zhang and T. Matsumoto, "Corpus augmentation for neural machine translation with Chinese-Japanese parallel corpora," *Applied Sciences*, vol. 9, no. 10, pp. 1–12, 2019.
- [25] Y. Chen, Y. Ma, X. Mao, and Q. Li, "Multi-task learning for abstractive and extractive summarization," *Data Science and Engineering*, vol. 4, no. 1, pp. 14–23, 2019.
- [26] P. Zhou and Z. Jiang, "Self-organizing map neural network (SOM) downscaling method to simulate daily precipitation in the Yangtze and Huaihe River Basin," *Climatic and Environmental Research*, vol. 21, no. 5, pp. 512–524, 2016.
- [27] X. Xiao, "Analysis on the employment psychological problems and adjustment of retired athletes in the process of career transformation," *Modern Vocational Education*, vol. 5, no. 12, pp. 216–217, 2018.
- [28] S. Sahoo and M. K. Jha, "Pattern recognition in lithology classification: modeling using neural networks, self-organizing maps and genetic algorithms," *Hydrogeology Journal*, vol. 25, no. 2, pp. 311–330, 2016.
- [29] Y. Zhou and B. Yang, "Sports video athlete detection using convolutional neural network," *Journal of Natural Science of Xiangtan University*, vol. 39, no. 1, pp. 95–98, 2017.
- [30] J. Pang, "Research on the evaluation model of sports training adaptation based on self-organizing neural network," *Journal of Nanjing Institute of Physical Education*, vol. 16, no. 1, pp. 74–77, 2017.
- [31] G. Querzola, C. Lovati, C. Mariani, and L. Pantoni, "A semi-quantitative sport-specific assessment of recurrent traumatic brain injury: the TraQ questionnaire and its application in

- American football,” *Neurological Sciences*, vol. 40, no. 9, pp. 1909–1915, 2019.
- [32] J. Wang, X. Luo, and H. Yan, “Correlation analysis between injuries and functional movement screening for athletes of the National Shooting Team,” *Journal of Capital Institute of Physical Education*, vol. 5, no. 4, pp. 352–355, 2016.
- [33] G. Ma, “Research on the design of juvenile football players’ sports injury prediction model,” *Automation Technology and Application*, vol. 277, no. 7, pp. 141–144, 2018.