Nama : Nuralam Paku Bumi

Kelas : A C3

Nim : 13020200039

1. Pada area riset computer vision berdasarkan gambar dibawah, bidang yang terliput dalam computer vision cukup banyak, contohnya interaksi manusia dan komputer pada bidang ini terdapat beberapa study yang menerapkan computer vision seperti psikologi, Biomeantik dan AI (meaching learning). Pada AI juga, dapat melakukan pemrosesan suara dan citra, pemrosesan citra pastinya tidak lepas dengan computer vision yang tentunya membutuhkan input kamera untuk memproses citra yang terinput. Selain itu, bidang Robotic juga sangat erat penggunaan computer vision didalamnya, selain Robotik bidang lainnya seperti multimedia, lingkungan virtual, dan grafik komputer juga menggunakan tecnologi komputer vision.

2. Algoritma face detection berdasarkan gambar meliputi beberapa tahapan proses.
   - preprocessing
   - Deteksi wajah

- Pemihihan wajah
- ekstraksi fitur, raitu fitur penting dari wajah yang terdeteksi seperti mata, hidung, dan mulut
- Klasifikasi
- Post - Processing

Research Article

Face Detection and Recognition Algorithm in Digital Image Based on Computer Vision Sensor

Di Lu [ORCID] **and Limin Yan**

Microelectronics Research and Development Center, Shanghai University, Shanghai, 200444 Shanghai, China

Correspondence should be addressed to Di Lu; di3777@shu.edu.cn

With the continuous innovation of network technology, various kinds of convenient network technologies have grown, and human dependence on network technology has gradually increased, which has resulted in the importance of network information security issues. With the continuous development of my country's industrialization, the application of sensors is becoming more and more extensive, for example, the security vulnerabilities and defects in the operating system itself. Traditional sensors can "perceive" a certain thing or signal, convert it into an electrical signal and record it, and then use a conversion circuit to output the electrical signal into a value or other display form that is conducive to observation. Nowadays, sensors have been further developed. Based on the original "perception" function, combined with computer technology, it integrates data storage, data processing, data communication, and other functions, so that it has analysis functions and can better display information. The technical level has reached a new level. Early intelligent recognition mainly used the uniqueness of finger and palm lines to scan and contrast, but due to some weather reasons or skin texture constraints caused by skin texture, these methods showed certain limitations. This paper proposes a new computer vision-based algorithm from face detection technology and face recognition technology. In the face detection technology, it is mainly introduced from the OpenCV method. Face recognition technology is improved in practical applications through the Seetaface method and YouTu method. At the same time, using the contrast experiment, the detection and recognition rates under the three different requirements of side face detection, occlusion detection, and facial exaggerated expression are compared, and the accuracy of each method is improved. The results show that each case is compared in each case. The advantages and disadvantages of the algorithm effectively verify the effectiveness of the method.

1. Introduction

With the continuous improvement of science and technology, face detection and recognition are applied in more and more fields, such as the verification of identity by each application face scanning, the monitoring system of the bank self-service cash machine, the face unlocking of the mobile phone, and the new face-brushing technology of Alipay. All need to pass the detection and recognition technology for the face. Under the prospect of the gradual diversification of the technology, face detection and recognition have become a technology closely related to our lives. Face detec-

tion and recognition technology not only make life easier and faster but also add a touch of technology fun. Through the face of a series of operations such as unlocking the phone, paying for the face, and intelligently identifying, using high-tech technology to ensure the security of our property and identity and to realize the combination of technology and life, it is a vital part of our lives. Sensors can be combined with many technologies to form smart sensors. Vision measurement technology has been developed into a new type of industrial testing technology, and its application scope is also expanding. Early vision measurement will be limited by the software and hardware resources of image

sensors and image processing systems and is expensive, has low performance indicators, and has relatively high failure rates. The processing efficiency is not high.

The research of face detection has important research value due to the variability of facial expression, skin color, and illumination. Yong and Yanru [1] studied the face detection based on skin color features and found that the difference in skin color is obvious under different illumination levels. In order to solve this problem, it uses YCbCr and HIS two skin color space lighting as the technical basis. The skin color model is used to distinguish the skin color area, which can effectively reduce the impact of care on skin color, which is more convenient for the face detection and positioning through the SNoW classifier. The comparison experiment results show that the method has high robustness during the experiment, the detection speed of the experiment is fast, and there are excellent face detection and positioning results. Guangsheng and Huarong [2] combined the deep learning technology with the convolutional neural network technology in order to control the occurrence of abnormal events in the monitoring system, so as to acquire the characteristics of the image and use the circulating neural system to effectively process the sequence and obtain the position and size of the detection window, creating a viable monitoring system that provides real-time warning of anomalies. Experimental results show that the system has strong face detection performance. Luhong et al. [3] used the face template method to effectively locate the face edge in order to break through the face detection which can only identify the binding under the condition of no background and pass the two eyes and different aspect ratios. The face template in the case is used for face detection, and finally, the face detection success rate of the experiment is more than 90%, which brings an effective solution for detecting the face in a single positive and positive environment. Yuangen et al. [4] reported that in order to solve the problem of detection complexity caused by face complexity, side face, rotation, expression, occlusion, etc., a new improved version of the AdaBoost algorithm is proposed to achieve comprehensive skin color preproduction. The function of processing and using geometric features for filtering, and the highly robust nature of the Haar algorithm, also extracts complex scenes. According to experiments, the method has excellent detection rate and latitude for side face and rotation. It has a good implementation effect on the embedded platform. Suwen and Yinwei [5] reported that in order to solve the traditional convolutional neural network for face detection (the detection speed and weight are too random due to the excessive number of sliding windows, and the training time is too long due to random initialization) and network convergence slow problem, they decided to use a new convolutional neural network method based on grouping layered strategy window to combine selective search and the Gabor optimization to extract features and determine the generalization of network structure. In the benefits of enhanced and face detection accuracy and speed of experimentation. Chengji et al. [6] designed the influence of the experimental differences caused by the complexity of the face in the real scene and the confidence of the background

frame, and the difference in the accuracy of the feature combination method in the wild. A multilayer feature fusion method effectively improves the detection accuracy between adjacent faces. Sheping et al. [7] proposed the LBP method to enhance the influence of face detection and reduce the conversion of nonlinear data to linear structure [8]. The extracted feature vector and SVM algorithm are used for classification processing. It is found through experiments that this method can effectively avoid the complex effects of illumination caused by uneven illumination and is very effective for face detection and recognition experiments.

Yong [9] explores the effect of face recognition in uncontrolled lighting environment based on the influence of different illumination effects. The MSR algorithm can directly extract the illumination invariants of objects, and GF and INPS algorithms can indirectly extract illumination. The algorithm characteristics of invariants are tested. From the perspective of feature level fusion and classifier decision-level fusion, the linear discriminant analysis of illumination robustness is used to design a method to overcome illumination invariant bands in face recognition. In the complexity of the impact, according to Chenkai et al. [10], based on the application of face recognition technology, from the history of traditional face recognition algorithms to the face recognition research under the deep learning method, the application of deep learning and DCNN algorithm is analyzed. Finally, the prospects for the future development of face recognition technology are discussed, and the face recognition technology is fully discussed. According to Qianyu et al. [11], influenced by the environment at that time, deep learning has a certain degree of success in motion recognition and target detection. The research on the influence of face recognition on illumination changes and attitude changes is urgent, so I try to propose a kind of a new and effective face recognition method that is expected to build a deep network based on the foundation of deep learning and use the database of the face to preprocess the computational complexity, so that the new deep network can effectively carry out the application of the extracted features [12]. The Softmax regression model is used to judge the face category, and the experimental results are excellent. The validity of the method is verified. According to Zhouyu et al. [13], in order to improve the optimization of the recognition model and accelerate the feature extraction of face images, a particle swarm optimization algorithm is introduced based on the traditional PCA technology to optimize the SVM model and the function model. By reducing the training and recognition time of the SVM, the face features can be extracted efficiently, so that the classifier can identify the test data [14]. The final experimental results show that the PSO-optimized SVM model has better performance, better generalization ability, higher accuracy of parameter value and recognition, and excellent effect on improving face recognition efficiency and improving recognition. Feng et al. [15] conducted a more in-depth exploration of the face recognition in the field of compressed sensing theory and studied the problem of solving the L_0 norm optimization problem in the classification algorithm based on sparse representation. The image representation algorithm of first-order

information and second-order information shows that the proposed algorithm is significantly better in low-dimensional recognition rate, and the algorithm improves the accuracy of classification and reduces the difference of face recognition. Mengxi et al. [16] proposed a kind of nonlinear deformation caused by changes in illumination, posture, expression, age, etc. in face recognition, using pop-learning, and the Laplacian feature mapping face recognition algorithm based on the two-dimensional kernel principal component analysis method [17]. The feature maps the face recognition algorithm. The experiment compares other algorithms, which has the advantages of high recognition rate and low computational complexity. Huixian et al. [18] improved the recognition performance caused by the one-sample face recognition problem, explored the extraction of invariant features, and improved the local texture feature of the image accurately and quickly by using the local binary pattern. The degree information uses the gradient information to improve feature extraction and effectively combines the two information extraction feature methods to achieve the ability to enhance the recognition of facial features. Experiments show that the BGCSBP algorithm has a high recognition rate, effectively reduces the recognition time, and has good applicability.

The paper first explores the face detection and recognition algorithm [19]. The face detection technology is analyzed by the OpenCV method. Then, the face recognition technology is explored from the Seetaface method and the YouTu method [20, 21]. Finally, the data experiment is used to analyze the three aspects of the face, the face occlusion, and the face with exaggerated expression [2–23]. The effect of face detection and recognition under different conditions is compared with the accuracy of face detection and recognition in different situations according to the three methods [24]. It is found that in the face detection part, there is exaggerated expression detection. Using the Seetaface method, the side face detection module and the occlusion detection module use the YouTu method; in the face recognition part, the face recognition module uses the YouTu method to maximize the detection rate and recognition rate and to reduce the false detection rate [25].

2. Methods

2.1. OpenCV Method. To perform the face recognition function, face detection is first performed to determine the position of the face in the picture. The OpenCV method is a common method in face detection. It firstly extracts the feature images into a large sample set by extracting the face Haar features in the image and then uses the AdaBoost algorithm as the face detector. In face detection, the algorithm can effectively adapt to complex environments such as insufficient illumination and background blur, which greatly improves the accuracy of detection. For a set of training sets, different training sets are obtained for subsequent work by changing the distribution probabilities of each of the samples, and each training set is trained to obtain a weak classifier, and then these several classifiers are weighted. For example, each sample is distributed with a training class, and a new training set is obtained by changing the distribu-
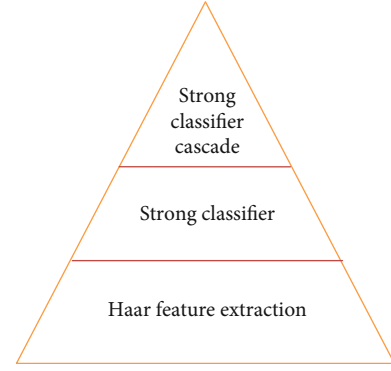

Figure 1: Human face detection model structure.

tion probability according to the correctness of the training set classification. The higher the classification accuracy rate, the lower the distribution probability. The new training set is trained to get the classifier, and it is repeated, and several classifiers are obtained, so that the weight of each classifier is increased by the classification accuracy.

Figure 1 shows the structure of the face detection model. The extraction of Haar's rectangular features and the strong classifier based on AdaBoost are an important part of face detection. The Haar feature is composed of several identical rectangles, which are distinguished by the black and white difference of colors, and the feature values of the Haar features are defined by the pixel values of the rectangle.

Regarding the number of rectangular features, Papageorgiou et al. proposed a formula:

$$\Omega^m_{[w,h]} = \left(\left[\frac{m}{w} \right] + \left[\frac{m-1}{w} \right] + \cdots + \left[\frac{w+1}{w} \right] + 1 \right) \cdot \bullet \left(\left[\frac{m}{h} \right] + \left[\frac{m-1}{h} \right] + \cdots + \left[\frac{h+1}{h} \right] + 1 \right), \tag{1}$$

where w and h represent the width and height of the rectangular features, respectively.

It can be seen from the above formula that the number of features is huge, and the OpenCV method solves this problem by introducing an integral image method. Let f be an arbitrary image and g be the integral image of the image. Then, the value of any pixel $A(x, y)$ is defined as

$$g(x, y) = \sum_{x' \leq x, y' \leq y} f(x', y'). \tag{2}$$

Calculated by the following formula, it is

$$s(x, y) = s(x, y - 1) + f(x, y),$$
$$g(x, y) = g(x - 1, y) + s(x, y), \tag{3}$$

where $s(x, y)$ is the cumulative value for each row, the initial value of $s(x, -1) = 0$, and $g(x, y)$ is the initial value of $g(-1, y) = 0$.
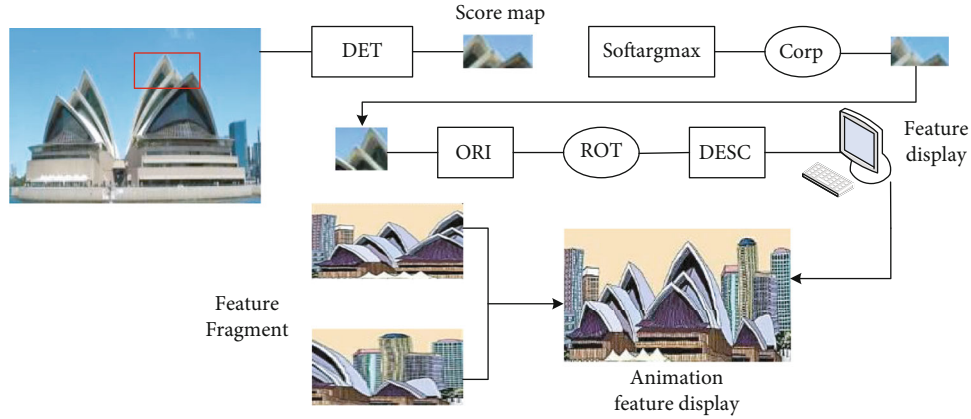
FIGURE 2: Feature extraction.

2.2. Seetaface Method. Seetaface is a new convolutional neural network structure. Face detection module Seetaface detection, feature point location module Seetaface alignment, and feature extraction and comparison module Seetaface recognition are mainly used for face detection and recognition. Firstly, the face part of the image is segmented to remove the background part. The feature points of the face are recognized and extracted to obtain the feature map, which is expressed in algebraic form and compared with the correlation to determine whether it is the same person. Feature extraction is shown in Figure 2.

The ROC curve of Seetaface detection on the FDDB database is shown in Figure 2.

Face detection is the top of the FuSt cascade structure consisting of several fast LAB cascade classifiers. These fast LAB cascade classifiers are mainly for face images of different poses appearing during face detection, and modules. The middle is composed of a plurality of multilayer perceptrons based on SURF features, and the same multilayer perceptual machine structure at the end of the module is responsible for processing gesture images of various faces. Among them, the feature of Seetaface face detection feature is that the upper funnel state is wide and narrow. This level of classifier makes the adopted features change gradually from top to bottom, so as to ensure that the background area is removed to the greatest extent and only the face area is retained. Effective step selection makes the next step more effective. Computer face recognition is shown in Figure 4.

In the feature point location module, the CFAN structure of the cascaded multistage stack-type self-encoder network is used to locate the feature points of the image detected by the Seetaface face. The specific steps are shown in Figure 3. It mainly adopts the classic five-point positioning method, that is, five points mark from the eyes, nose, and mouth, which ensures the accuracy of face detection recognition to the greatest extent. In the processing of the first-stage self-encoding network, a low-resolution face region is used for fast estimation, and the face shape is positioned as s0. Since the resolution of the picture at this stage is too low, the shape of the face shape is roughly contoured. The next step is to increase the resolution, so that the clarity
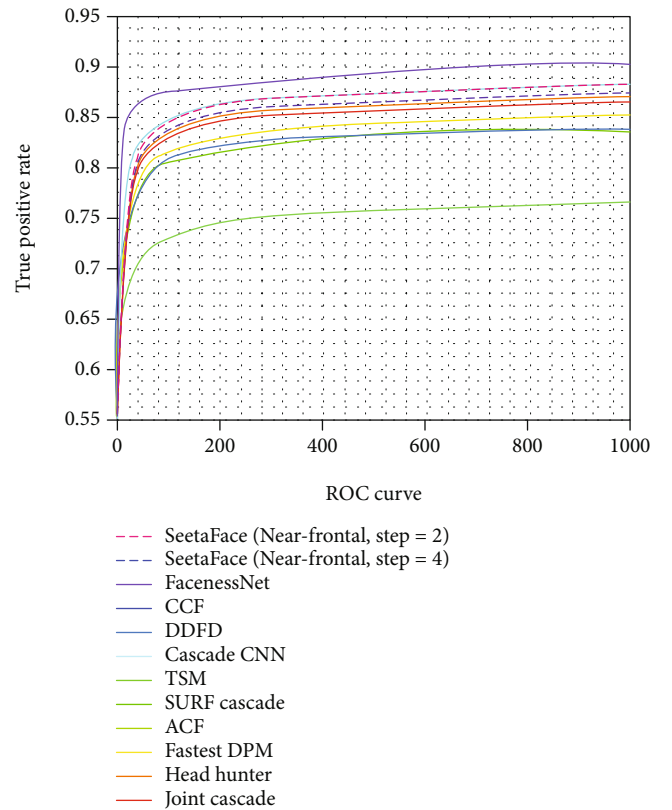

FIGURE 3: ROC curve.

of the detected pictures is continuously improved. After each step of the self-encoding network, $H_2 \longrightarrow H_3 \longrightarrow H_n$ is processed step by step, and the facial features of the face are continuously optimized more and more carefully.

According to Figure 5, the feature extraction and comparison module are mainly based on the convolutional neural network model, and the two images are similarly completed to complete the recognition step. The similarity of the same face image is more than 70%, and the similarity of different objects is less than 30%, so that it meets the needs of daily face recognition.

2.3. YouTu Method. As an emerging face recognition algorithm, the YouTu method adopts the classical boosting algorithm in analysis and facial features, and the face recognition confirmation part is completed by combining deep learning methods. Set $w_1, w_2, w_3 \cdots$ as a collection; then

$$f(x) = \sum_{t=1}^{T} \lambda_t w_t(x). \tag{4}$$

Among them, λ is the correlation coefficient of each member, and both λ and w can be learned through the boosting process.

Normally, if the input source image belongs to the input space X and the output result belongs to the output space Y, then there is a function $f : X \longrightarrow Y$ that obeys the probability distribution $P(x, y)$:

$$(x_1, y_1), (x_2, y_2), \cdots \cdots, (x_n, y_n) \in R^d \times \{-1, +1\}. \tag{5}$$

At this point, f can predict the unknown (x, y), and each input category label is given by function $f(x)$. We call this a hard classifier. However, in the process of use, we are more inclined to use the soft classifier. In the soft classifier, each input category label is given by function sign $(f(x))$; then

$$L(f) = \int \lambda(f(x), y) dp(x, y). \tag{6}$$

Among them, λ represents the loss function, and risk is called $L(f)$ as the error promotion rate. Because of the $p(x, y)$ distribution that is not known, $L(f)$ cannot be directly reduced. Therefore, an import rule is needed to reduce the generalized error rate. Then, there are

$$\widehat{L}(F) = \frac{1}{N} \sum_{n=1}^{N} \lambda(f(x_n), y_n). \tag{7}$$

According to the law of large numbers, when $N \longrightarrow \infty$, $\widehat{L}(f) \longrightarrow L(f)$.

In daily use, the amount of training data is not very large, and the actual error rate will not be large. The boosting algorithm guarantees that the complexity is not high in some cases, so there is no overadaptation.

The posterior probability of the hidden topic E can be estimated according to the existing parameters.

$$E_{r \longrightarrow G} = \min_{G \in G} |R - G| + \beta. \tag{8}$$

Maximize the complete likelihood logarithm function according to the estimated hidden topic R.

$$L(d) = \frac{1}{|M|} \sum_{r \in R} [E_{R \longrightarrow G} < D]. \tag{9}$$

We introduce the parameters $g \in G$ and use the Lagrangian coefficient method to solve the maximum complete likelihood logarithm function:

$$T_{r \longrightarrow G} = \min_{r \in R} |M - Y|. \tag{10}$$

For a two-class classification problem, define the training sample set:

$$K(d) = \frac{1}{O} \sum_{g \in G} [R_{g \longrightarrow R} < D]. \tag{11}$$

It is necessary to introduce a slack variable ξ_i to convert the solution of the optimal hyperplane problem into a quadratic optimization problem, that is, to solve the following constraints:

$$S(d) = \xi \frac{Y(d)R(d)}{Y(d) + R(d)}. \tag{12}$$

Use the Lagrange coefficient method to convert the final quadratic optimization problem into a dual problem, namely

$$H(x) = \mu h_\theta(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 +, \cdots, \beta_n x_n. \tag{13}$$

Each element in μ is a Lagrangian coefficient.

$$h(x) = \sum_{i=0}^{n} \theta_i X_i = \theta^T X + \sum_{m}^{n} (x - y)^3. \tag{14}$$

In order to evaluate whether the fitting effect of the parameter θ is optimal, a loss function is needed to express it, which is generally used in the model.

Through this kernel matrix, the training samples can be mapped from the original input space to high-dimensional in the feature space H.

$$H = G \sum_{i=1}^{m} (H_\theta(x_i) - Y_i)^2. \tag{15}$$

In the face detection process, the YouTu method increases the number of positioning points to 90 points, which is scattered throughout the contours of the facial features, greatly increasing the accuracy of positioning. As shown in Figure 6, the 90 points used for facial features include the contours of the eyebrows, eyes, nose, mouth, and face, where the eyebrows are symmetrical at 8 points each; the left and right sides of the eyes are 9 points. There are 13 positioning points, 22 positioning points in the lips, and 21 positioning points in the face contour.

Among them, the YouTu method uses the knowledge model in face recognition to perform feature processing and calculate its feature similarity. The key step is to compare the known face image with several images in the face database, analyze the related images with high similarity, and clearly show the similarity ratio, as shown in Figure 7, namely the 1:N face search. If the photo to be retrieved contains a plurality of faces, the search result corresponding to each of the detected faces is returned. Face retrieval is applied to scenes where the user does not need to declare identity, and the identity of each
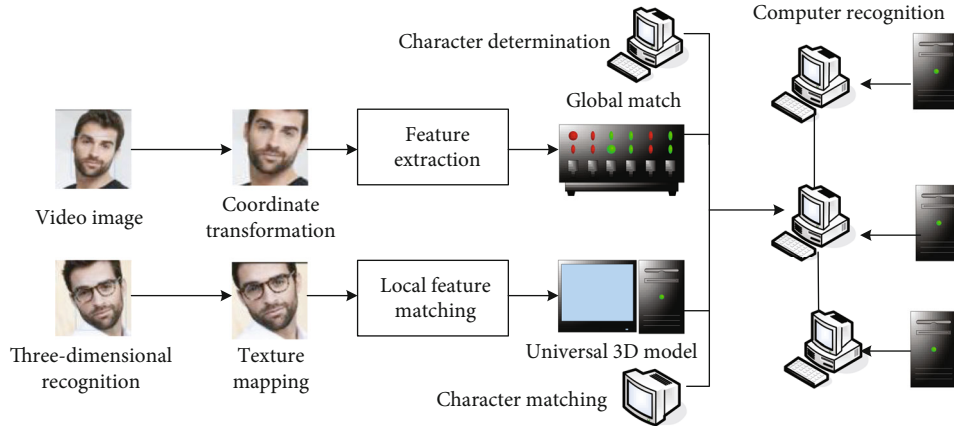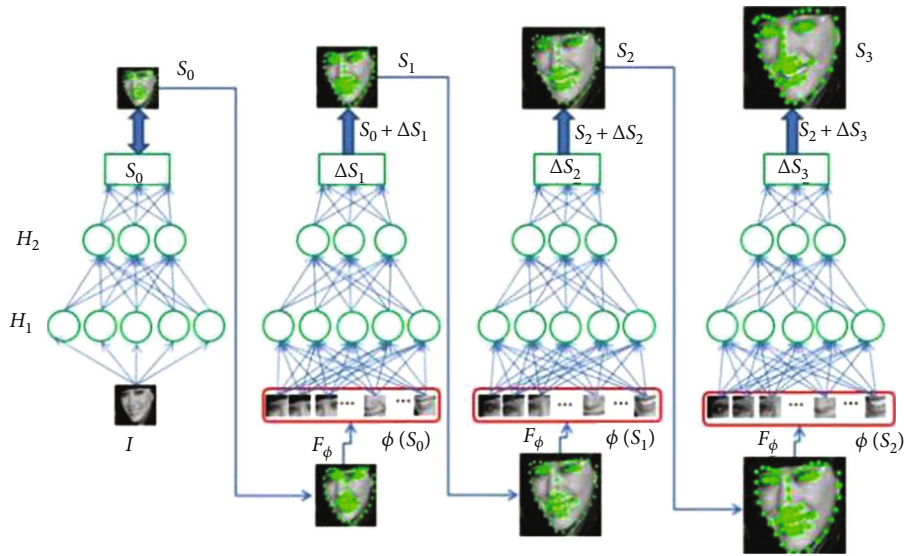
Figure 4: Computer face recognition.

Figure 5: Multilevel stack self-encoder network.

Figure 6: Five-position positioning analysis table.

person in the group is determined by performing face retrieval in the identity photo library.

This technology has high recognition rate, can search and adjust according to different scenes, and can automatically derive the face evolution model to overcome the bone differences caused by age differences. In addition, the technology also

has antiocclusion technology, which effectively reduces the impact of obstacles on face recognition, and has been applied to the public security tracking system to monitor the face position more conveniently and quickly through face recognition.

The signal-to-noise ratio (SNR) is the most important performance indicator of the ADC. The signal-to-noise ratio

Figure 7: Face recognition contrast similarity.

Table 1: The image sensor performance comparison.

Category	CCD image sensor	CMOS image sensor
Cost	High	Low
System power consumption	High	Low
Integration situation	Low (requires external chip)	High monolithic integration
Power supply	Multiple power supply	Single power supply

includes factors such as linearity, distortion, impulse, and noise, according to the quantization accuracy of the ADC.

$$\mathrm{SNR}_{hj} = \chi_{hj} \left\{ \sum_{I=1}^{M} \left[W_{ij} \left(R_{ij} + S_{ih} \right) \right]^P \right\}^{1/P}. \tag{16}$$

Harmonic distortion ratio (THD) is the ratio of the power of all harmonic distortion to the power of the fundamental wave in a certain frequency band.

$$\mathrm{THD} = k(S, t) \frac{F(S, t)}{F(t)} \left[1 - p_c \frac{\varphi(S)}{l - 1} - \lambda(S) p_m \right]. \tag{17}$$

The signal-to-noise and distortion ratio can be obtained by the fast Fourier transform (FFT) analysis β_i^3.

$$\beta_i^3 = \sum_{i=1}^{C} \frac{\partial Y}{\partial X_i^4} \times W_{ij} \frac{\sum_{k=1}^{\delta} X_k^3 - Z_j^3}{\sum_{k=1}^{\delta} S_k^3}. \tag{18}$$

Call the MATLAB function and substitute the output into the script to calculate the harmonic distortion ratio, according to the formula $\varphi_p(Z(T))$:

$$\varphi_p(Z(T)) = \frac{\prod_{m=1}^{l} R_{pm}(Z(T))}{\sum_{p=1}^{r} \prod_{m=1}^{l} R_{pm}(Z(T))}. \tag{19}$$

2.4. Image Sensor. In the past, the visual system was more complicated and expensive, usually from US dollars to more than US dollars. Generally, multiple cameras were required to complete detailed automatic detection. And due to its complexity, specialized vision experts were often required to design, integrate, and install the system. These factors naturally limit it to certain large companies, but it is obviously inappropriate for small- and medium-sized companies that require a detection system. In contrast, the visual sensor is much simpler, compact, and easier to install and operate, making it more suitable for the needs of general enterprises. The image sensor performance comparison is shown in Table 1.

The image sensor can be divided into area array type and linear array type according to the working mode. The area array image sensor uses a pixel array arranged in a two-dimensional area array to photograph objects to obtain two-dimensional image information. The area array pixel image sensor passes a complete image that can be obtained with a single exposure. Most digital cameras, mobile phone cameras, and surveillance cameras use this structure. The linear image sensor uses a pixel array arranged in a one-dimensional linear array to obtain two-dimensional image information by scanning and photographing objects. A row of pixels can be obtained at the completion of each exposure. To get an image, you need multiple rows of pixels. It is widely used in machine vision monitoring, aerial photography, spatial imaging, and medical imaging.

Under the control of the timing sequence, the switch cooperates to complete the addition and subtraction of the next level signal, the sampling phase ends, and the output is completed.

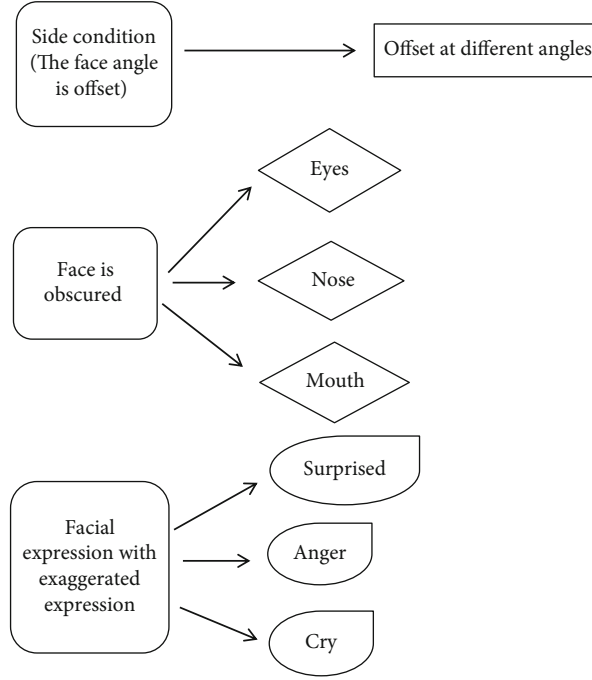$$O = \frac{C_S}{C_F} Q_{\mathrm{in}}. \tag{20}$$

FIGURE 8: Schematic diagram of the experimental process.

TABLE 2: The register function description.

On-chip registers	Symbol mark	Function description
0x001	Row start	
0x002	Column start	Set the acquisition window size of the CMOS image sensor
0x003	Row size	
0x004	Column start	

TABLE 3: The research data.

The first set of data		The second set of data		The third set of data	
x	y	x	y	x	y
6.49	1.33	6.40	1.31	6.11	1. 46
6.95	1.44	6.85	1.41	6.22	1. 56
6.41	1.54	6.31	1.52	6.68	1.68
6.86	1.64	6.66	1.64	8.14	1.64

TABLE 4: Performed data processing.

Verification test	The first set of data	The second set of data	The third set of data
Feature point 1	9.15	9.07	9.97
Feature point 2	14.97	15.05	15.88
Width information	5.82	5.98	5.91

C_S, C_F are a capacitor. The coordinates of the measurement data obtained by the sensor are based on the coordinate system N1, and the equation of the projection curve can be expressed as

$$X_1 = 0, Z = F(X, Y). \tag{21}$$

Then, the conversion relationship between coordinate systems N1 and N2 is as follows:

$$V = HV_I + HV_L + HV_R,$$
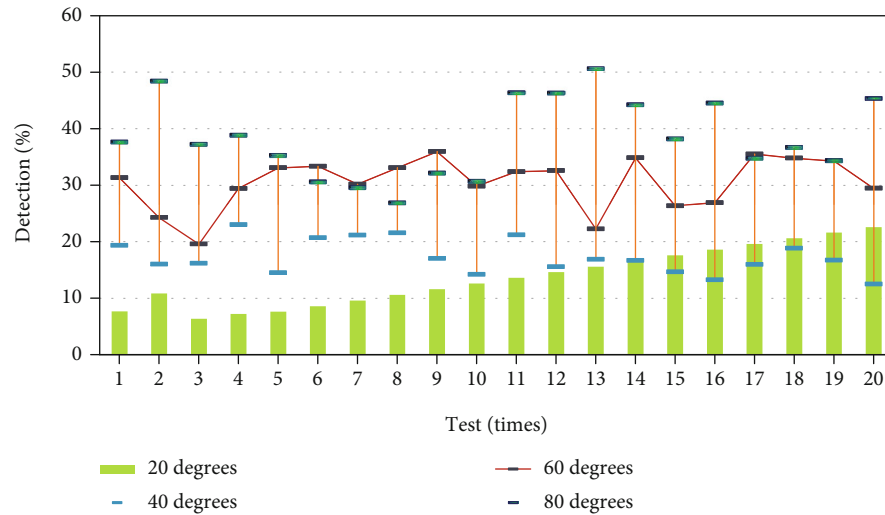$$H = P \cdot R(\alpha) \cdot R(XG) \sum_{I=1}^{N} V_I, \tag{22}$$

FIGURE 9: Face detection accuracy when the Seetaface angle shifts.

TABLE 5: The measured object is tilted.

The first set of data		The second set of data		The third set of data	
x	y	x	y	x	y
6.76	1.38	6.73	5.09	7. 08	5.13
7.22	1.46	7.18	5.14	7.53	5.21
7.68	1.58	7.63	5.21	7.98	5.29
8.14	1.67	8.08	5.29	8.44	1.51

TABLE 6: The verification test results.

Verification test	The first set of data	The second set of data	The third set of data
Feature point 1	9.15	9.07	9.97
Feature point 2	14.97	15.05	15.88
Width information	5.82	5.98	5.91

TABLE 7: The combination of each layer of the algorithm in this research.

Number of layers	Number of weak classifiers	Average detection rate	The average false detection rate of the classifier
1	820	50.15%	45.71%
2	1564	54.62%	42.45%
3	1522	64.52%	36.80%

where P is the perspective transformation matrix, expressed as

$$P = \begin{bmatrix} F & 0 & 0 \\ 0 & F & 0 \\ 0 & 1 & F \end{bmatrix}. \tag{23}$$

3. Experiments

This paper analyzes the accuracy of face detection and recognition by analyzing OpenCV, Seetaface, and YouTu. In this experiment, the selected face images are randomly representative, and there is no specific case. The main work of this thesis is as follows: (1) Through the in-depth study of OpenCV, Seetaface, and YouTu, based on the convolutional neural network model, the Haar feature extraction method is used to extract the features, and the face detection of the three algorithms is performed, identify and explain. (2) For three different specific scenes, that is, the angle of the face is shifted; the face is occluded, and the eyes, nose, and mouth are, respectively, occluded; the face has an exaggerated expression, including surprise, anger, and three expressions of crying, according to research on these three different scenarios. (3) And under each specific situation, use three
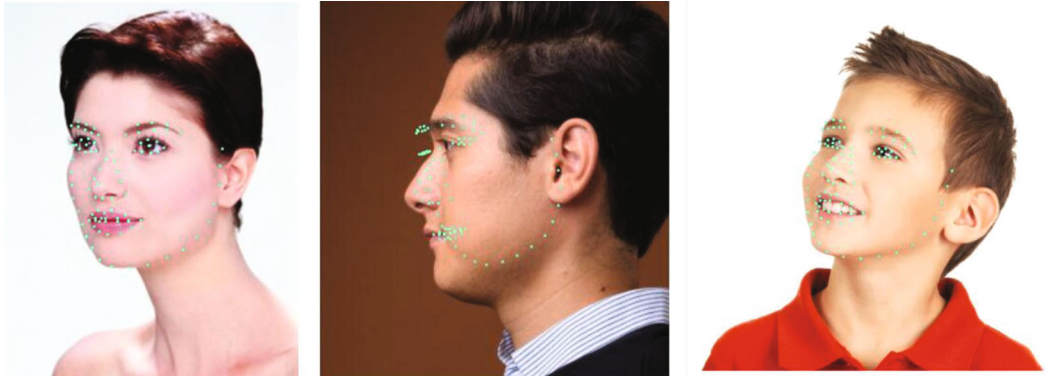
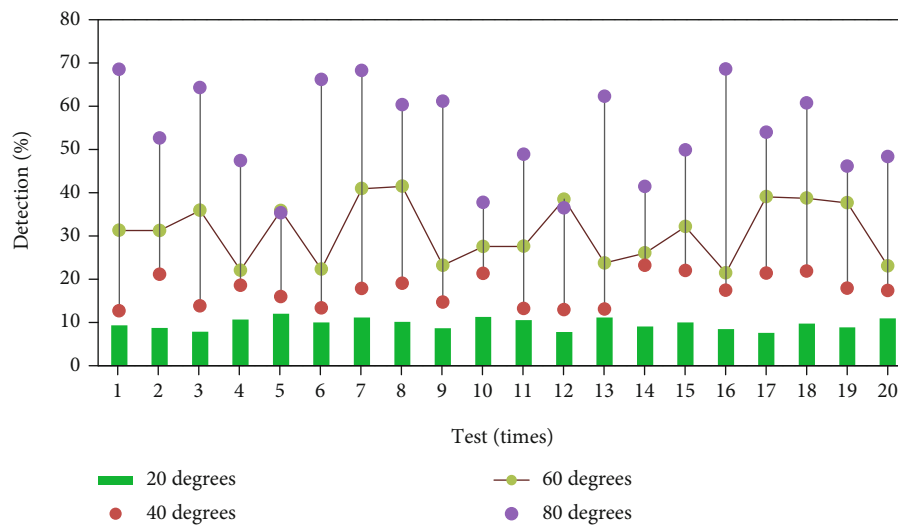Figure 10: Face detection map with three different offsets.

Figure 11: YouTu face image offset angle accuracy.

algorithms to conduct experiments, and explore the experimental results of each algorithm to compare which algorithm is applicable in this case.

The whole process of the test is similar to Figure 8. According to different aspects of face detection and recognition, it is divided into three cases for discussion. The experimental results show that this method can better control the shielding angle and shielding position. Table 2 shows the registration function description.

4. Results and Discussion

4.1. Face Detection and Recognition in Side Cases. Select three sets of experimental data, sharpen the image, extract feature points, process the characteristic information of the concave groove on the collected image, calculate the width of the concave groove, and verify the image acquisition accuracy of the sensor. The research data is shown in Table 3.

Perform data processing on the data in the above table to find the key feature points, as shown in Table 4.

Face detection and recognition are more closely applied to life, and the actual situation of the situation encountered

is more the deterioration of the face angle, that is, the side situation. Figure 9 shows the face detection accuracy of the Seetaface method when the face is angularly offset.

It can be clearly seen from the Figure 9 that as the degree of face angle shift increases, the accuracy of face detection and recognition gradually decreases. When the angular deviation is within 20°, the angular offset decreases greatly, and the recognition accuracy is above 30%, which can detect and identify the face, but there is a case where the error is too large. When the side angle shift is too large and reaches 40 degrees or more, the accuracy of face detection and recognition is only 20% or less, and it is basically in an undetectable state.

The measured object is tilted, or the incident laser light is not perpendicular to the surface during hand-held measurement (the relationship between the light bar and the gap is still vertical), as shown in Table 5.

Image processing is also performed on the above-mentioned collected images, and the verification test results are shown in Table 6.

The number of weak classifiers, usefulness, and detection rate and false detection rate of weak classifiers generated by

TABLE 8: The algorithm accuracy comparison.

Features	Detection rate	False detection rate
Best Haar features	50.10%	43.84%
The best HOG features based on different scales	82.30%	21.81%
The best HOG feature based on feature combination	90.72%	13.55%

TABLE 9: The test results of different features.

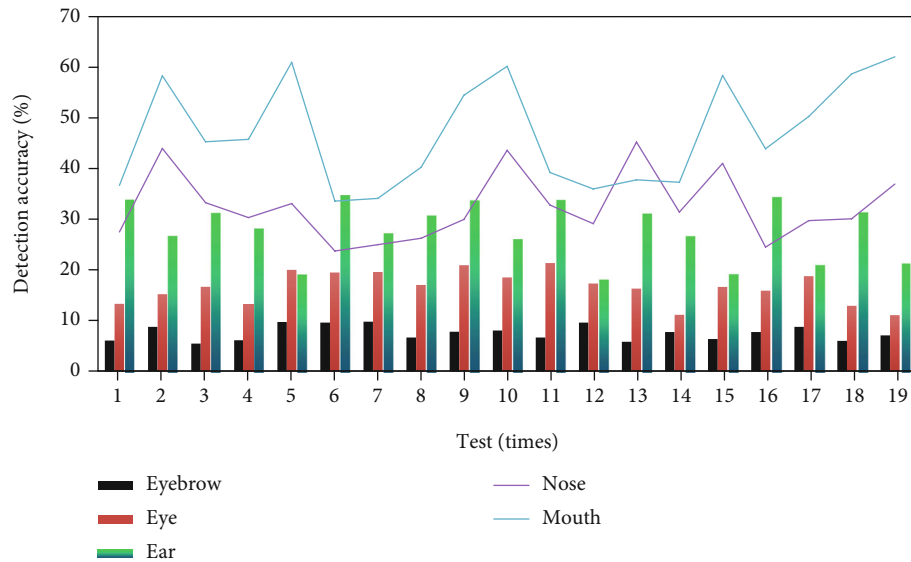| Feature	Detection rate	False detection rate	The number of layers it belongs to when it is generated
The best feature of ranking	90.72%	13.55%	10
Ranked second best feature	89 95%	13.7 1%	10
Ranked third best feature	90.24%	14.45%	7
Ranked fourth best feature	89.43%	15.21%	10

FIGURE 12: OpenCV face image offset angle accuracy rate.

the combination of each layer of the algorithm in this research are shown in Table 7.

The face detection and recognition of the three lateral off-sets in the three different gender and age test conditions are shown in Figure 10 using the YouTu method for face detection.

It can be seen that the YouTu method can accurately locate the facial features of the face even if there is a slight offset and a slight angular offset in dealing with certain face recognition and detection problems, and there is no basic deviation in its positioning. At the same time, when the side is deflected by 90 degrees, it can be clearly seen that the method can accurately perform the five-position positioning for the five senses at the identifiable place, and even for the five senses of the unidentified part, the position judgment can be made according to the situation, and the judgment is made. The location result is more reasonable.

It can be clearly seen from Figure 11 that within the angle offset of 60°, the recognition accuracy of the face image is slightly higher, and the face can be more accurately identified and detected.

The algorithm accuracy comparison is shown in Table 8. The test results of different features are shown in Table 9.

The YouTu method performs face detection on a face with a certain offset angle. The accuracy rate is shown in Figure 12. According to the chart, it can be clearly seen that when the offset angle is small, the accuracy of the detection is extremely high, but when the offset is large, the detection result is inaccurate, and as the offset angle increases, the detection accuracy is improved. Gradually lowering, face detection is not possible. The comparison on the cascade classifier is shown in Table 10.

4.2. Face Detection and Recognition under Face Occlusion. In addition, in face detection, the situation that the face is obstructed by the obstruction obscuring the true appearance of the detector exists in the market, so there is great research significance for face detection and recognition in the case of face occlusion.

As can be seen from Figure 13, when the detector's eyes are occluded, face detection is completely impossible. When

TABLE 10: Comparison on cascade classifiers.

SVM weak classifiers	Detection algorithm based on Haar features	HOG detection algorithm (L1 normalization)	HOG detection algorithm (L2 normalization)
Number of first-level SVM weak classifiers	18	4	2
Number of second-level SVM weak classifiers	21	4	2
The number of third-level SVM weak classifiers	24	4	3
Number of fourth-level SVM weak classifiers	30	4	3
Number of fifth-level SVM weak classifiers	39	6	5

FIGURE 13: Seetaface method for face occlusion face detection.

TABLE 11: The experimental comparison.

Compared	Algorithm 1: best single-channel Gabor characterization	Algorithm 2: multichannel Gabor characterization with equal weight	Algorithm 3: weighted multichannel Gabor characterization
ORL	0.86	0.915	0.95
AR	0.77333	0.89167	0.90167
Yale-B	0.6	0.6	0.66667

TABLE 12: Extract the features on the salient face area.

Compared	Algorithm 4: best single-channel Gabor characterization	Algorithm 5: multichannel Gabor characterization with equal weight	Algorithm 6: weighted multichannel Gabor characterization
ORL	0.95	0.96	0.915
AR	0.90167	0.95333	0.90167
Yale-B	0.66667	0.5	0.5

the nose and mouth are blocked, the five-point positioning can be clearly performed, and the portrait and background can be divided more accurately.

The different sample libraries and test libraries of ORL, AR, Yale-B, and CAS-PEAL-R1 are, respectively, cropped to 32×32 pixels. In addition, the posture of the training library is manually corrected, and the average value is

selected for multiple tests. Experiment 1 compares the recognition rates of Algorithms 1, 2, and 3 to verify the importance of multichannel weighted representation. The experimental comparison is shown in Table 11.

It is proved that it is more effective to extract the features on the salient face area. The experimental comparison is shown in Table 12.

Figure 14: YouTu method for face occlusion face detection.

Figure 15: OpenCV method for face occlusion face detection.

As shown in Figure 14, no matter whether the eyes, nose, or mouth are blocked, the YouTu method can accurately locate 90 feature points for the detection of facial contours. This method is highly effective for face detection.

As can be seen from Figure 15, the face can be effectively separated from the background only when the nose is blocked, and face recognition cannot be performed when the eyes and mouth are blocked.

4.3. Face Detection and Recognition under Exaggerated Facial Expressions. Compare the recognition rate of the methods in this chapter, as shown in Table 13.

Test the sensitivity of the algorithm in this research to noise. Usually, the picture may contain various noises. After adding several common noises to the test picture, compare the method in this chapter with the traditional overall Gabor characterization. The recognition rate is shown in Table 14.

TABLE 13: Compare the recognition rate of the methods in this chapter.

Compared	Algorithm 4: best single-channel Gabor characterization	Algorithm 6: multichannel Gabor characterization with equal weight	Algorithm 7: weighted multichannel Gabor characterization based on region selection and FFT
ORL	0.95	0.975	0.975
AR	0.90167	0.90333	0.95333
Yale-B	0.66667	0.5	0.5

TABLE 14: Test the sensitivity of the algorithm.

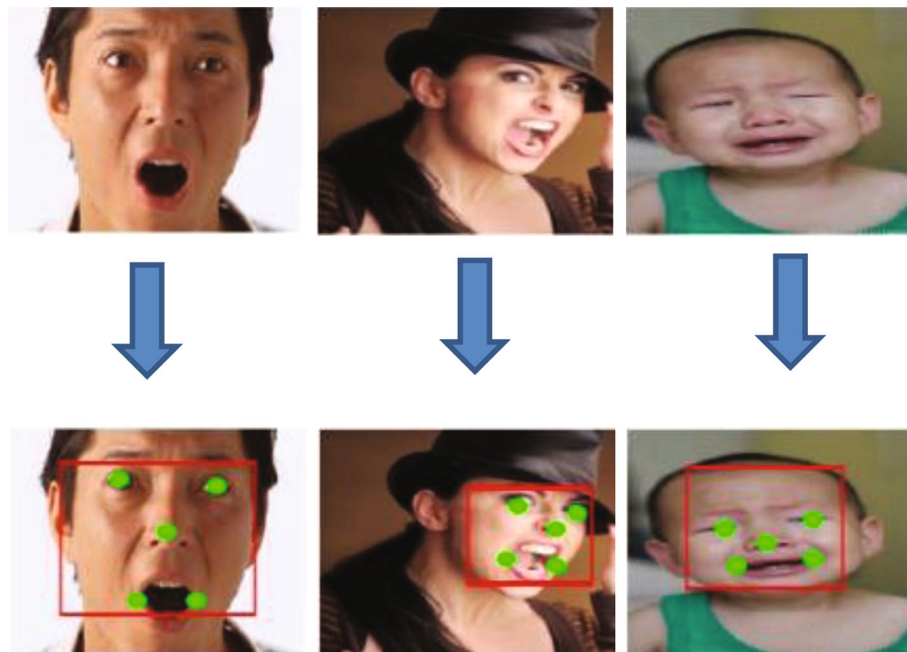| AR-600	Algorithm 5: overall Gabor characterization	Algorithm 7: weighted multichannel Gabor characterization based on region selection and FFT
Salt and pepper noise	0.69	0.95333
Gaussian white noise	0.865	0.7133
Speckle noise	0.7566	0.89833
Poisson noise	0.85833	0.77833

FIGURE 16: Seetaface method for face detection in different expressions.

The facial expression changes are extremely diverse. Each person's expression has the uniqueness and uniqueness of his own posture. This complex expression is extremely difficult in face detection and recognition. It is extremely important to find a more effective way to locate various expression changes.

It can be concluded from Figure 16 that the Seetaface method can effectively and accurately distinguish the face from the background and accurately perform the facial features in any expression state of surprise, anger, and crying.

As seen in Figure 17, the YouTu method can also perform face contour segmentation in three different expressions, but it can accurately segment the face and background. However, from a precise point of view, the facial features of the method are slightly deviated, resulting in a slight accuracy decline.

As can be seen from Figure 18, OpenCV can perform face recognition and detection for different expressions, but this method can only be used for face recognition, and accurate five-position positioning cannot be performed in face positioning.

FIGURE 17: YouTu method for face detection in different expressions.

FIGURE 18: The OpenCV method detects different expressions of faces in different expressions.

5. Conclusion

In the side case, by comparing the accuracy of the three face detection methods of Seetaface, YouTu, and OpenCV in the occurrence of angular offset, it can be clearly found that the Seetaface method can be effective when the angular offset is small. Face segmentation and detection are performed, but in the case where the offset angle is slightly larger, the accuracy is gradually reduced as the angle increases, and face detection cannot be performed even when a serious offset occurs. The YouTu method can perform more accurate face recognition detection when the angle of the face is large, and even if a serious offset occurs, it is more accurate for the unidentified part of the facial features. However, the detection result of the OpenCV method has a large deviation, and the angular offset is slightly larger, resulting in an unde-

tectable result. Ordinary photoelectric sensors have only a single light sensor element. In many applications, multiple such sensors are often required to detect the various characteristics of the components, and the vision sensor can capture an image containing millions of pixels, so as to be able to condition the inspection components shown in detail, which can prevent missed inspections and improve the accuracy of the inspection, which is especially necessary for the inspection of electronic components.

The image digital pixel sensor is a sensitive element, and the highly integrated chip makes the system avoid the use of multiple drive chips and improves the stability of the sensor. In the case of facial occlusion, the YouTu method is obviously more accurate, no matter which part is occluded, it can effectively locate the contour of the face, and the position prediction for the occlusion part is more accurate. The

effect of the OpenCV method is the lowest, the face can be detected only when the nose is occluded, and face detection cannot be performed once any of the eyes and mouth are occluded. The Seetaface method also detects that the effect is blocked during occlusion. Face detection is not possible when the eyes are blocked. Face detection can be performed slightly when the nose and mouth are blocked, but the detection accuracy is not high.

The information collected by the image sensor greatly simplifies the bandwidth of the filter in the subsequent image signal processor and enhances the signal-to-noise ratio, ensuring the quality of imaging within the available dynamic range. From the analysis of facial expressions with exaggerated expressions, Seetaface is the most effective method. This method can effectively and accurately perform face recognition and localization and is equally accurate in the facial features. The YouTu method can also perform segmentation background and facial features in face detection and recognition, but its status has certain errors, which is less accurate than Seetaface. The OpenCV method cannot perform five-point positioning in the detection, and functionally simple, and can only perform background segmentation.

Data Availability

This article does not cover data research. No data were used to support this study.

Conflicts of Interest

There are no potential competing interests in our paper. And all authors have seen the manuscript and approved the submission. We confirm that the content of the manuscript has not been published or submitted for publication elsewhere.

References

[1] D. Yong and Y. Wu, "Face detection method based on double skin model and improved SNoW algorithm," *Computer Applications and Software*, vol. 34, no. 5, pp. 135–140, 2017.

[2] G. Liang and Z. Huarong, "Design of intelligent video surveillance face detection system based on ARM," *Computer Application*, vol. 37, no. a02, pp. 301–305, 2017.

[3] L. Liang, A. Haizhou, and K. He, "Single face detection based on multi-template matching," *Journal of Image and Graphics*, vol. 4, no. 10, pp. 825–830, 2018.

[4] Q. Yuangen, C. Danchi, and J. Peng, "Face detection algorithm based on geometric features and new Haar features," *Sensors and Microsystems*, vol. 36, no. 2, pp. 154–157, 2017.

[5] S. Wu and Z. Yinwei, "Face detection based on selective search and convolutional neural networks," *Computer Application Research*, vol. 34, no. 9, pp. 2854–2857, 2017.

[6] C. Wang, Z. Luo, Z. Zhun, and L. Shaozi, "A face detection method with multi-layer feature fusion," *Journal of Intelligent Systems*, vol. 13, no. 1, pp. 138–146, 2018.

[7] Y. Sheping, L. Wei, M. Ma, and S. Gao, "Lens detection based on LBP and SVM," *Computer Technology and Development*, vol. 27, no. 9, pp. 44–47, 2017.

[8] S. Qu, W. Xu, J. Zhao, and H. Zhang, "Design and implementation of a fast sliding-mode speed controller with disturbance compensation for spmsm syste," *IEEE Transactions on Transportation Electrification*, vol. 99, pp. 1–1, 2021.

[9] Y. Cheng, "Complex illumination face recognition based on multi-feature fusion," *Computer Engineering and Applications*, vol. 53, no. 14, pp. 39–44, 2017.

[10] J. Chenkai, S. Tao, Z. Lei, L. Gang, L. Wang, and L. Kailun, "Overview of face recognition technology based on deep convolutional neural networks," *Computer Applications and Software*, vol. 35, no. 1, pp. 223–231, 2018.

[11] L. Qianyu, J. Jianguo, and Q. Meibin, "Face recognition algorithm based on improved deep network," *Chinese Journal of Electronics*, vol. 45, no. 3, pp. 619–625, 2017.

[12] N. Xiao, R. Xinyi, Z. Xiong et al., "A diversity-based selfish node detection algorithm for socially aware networking," *Journal of Signal Processing Systems*, vol. 93, no. 7, pp. 811–825, 2021.

[13] Z. Liao, Y. Wang, X. Xie, and L. Jianming, "Support vector machine face recognition based on particle swarm optimization," *Computer Engineering*, vol. 43, no. 12, pp. 248–254, 2017.

[14] Q. Jiang, F. Shao, W. Lin, K. Gu, G. Jiang, and H. Sun, "Optimizing multistage discriminative dictionaries for blind image quality assessment," *IEEE Transactions on Multimedia*, vol. 20, no. 8, pp. 2035–2048, 2018.

[15] L. Feng, M. Fanrong, and Z. Liang, "Face recognition based on first- and second-order information image representation," *Computer Application Research*, vol. 34, no. 2, pp. 603–606, 2017.

[16] M. Xu, D. Xu, and W. Mingjun, "Laplacian feature mapping based on 2d-kpca," *Computer Application Research*, vol. 34, no. 7, pp. 2212–2215, 2017.

[17] S. Dong, H. Zhu, S. Zhong, K. Shi, and Y. Liu, "New study on fixed-time synchronization control of delayed inertial memristive neural networks," *Applied Mathematics and Computation*, vol. 399, article 126035, 2021.

[18] H. Yang, D. He, L. Fan, L. Yang, and L. Zhao, "One-sample face recognition based on bidirectional gradient center symmetric local binary mode," *Journal Of Computer-Aided Design and Computer Graphics*, vol. 29, no. 1, pp. 130–136, 2017.

[19] Z. Akhtar and A. Rattani, "A face in any form: new challenges and opportunities for face recognition technology," *Computer*, vol. 50, no. 4, pp. 80–90, 2017.

[20] Z. Lv and L. Qiao, "Optimization of collaborative resource allocation for mobile edge computing," *Computer Communications*, vol. 161, pp. 19–27, 2020.

[21] Z. Lv, Y. Han, A. K. Singh, G. Manogaran, and H. Lv, "Trustworthiness in industrial iot systems based on artificial intelligence," *IEEE Transactions on Industrial Informatics*, vol. 99, 2020.

[22] K. Bong, S. Choi, C. Kim, and H. J. Yoo, "Low-power convolutional neural network processor for a face-recognition system," *IEEE Micro*, vol. 37, no. 6, pp. 30–38, 2017.

[23] C. Li, W. Wei, J. Li, and W. Song, "A cloud-based monitoring system via face recognition using Gabor and cs-lbp features," *Journal of Supercomputing*, vol. 73, no. 4, pp. 1532–1546, 2017.

[24] Z. Cai and X. Zheng, "A private and efficient mechanism for data uploading in smart cyber-physical systems," *IEEE Transactions on Network Science & Engineering*, vol. 7, no. 2, pp. 766–775, 2018.

[25] M. Zhou, H. Lin, S. S. Young, and J. Yu, "Hybrid sensing face detection and registration for low-light and unconstrained conditions," *Applied Optics*, vol. 57, no. 1, pp. 69–78, 2018.