

3D Face Feature Processing and Recognition Technology

Hejing Wu

East University of Heilongjiang

E-mail: 499917928@qq.com,

Abstract—Nowadays, we frequently use a lot of biometric identification technology means, such as: iris recognition, voice recognition and face recognition, but compared with other human biometric features, face information is easier to collect and more easily accepted by the public. According to the survey, in interpersonal contact, most people get information based on the first impression, even if there are changes in external conditions such as age, expression, makeup, etc., but the brain can still recognize these information, this phenomenon shows that face does have more feature information. From the information of a person's facial features, we can effectively distinguish each person. Face recognition is a technology that analyzes the face data and then uses it as an authentication mechanism. Generally, cameras are used to track and collect human faces, model theory and then analyze the collected binary data stream to extract relevant feature data. The research of human face recognition technology began in the middle of 1960s. After 1980s, due to the development of computer technology and the improvement of optical image processing technology, it gradually matured. The real use stage of face recognition is in the late 1990s, with Germany, Japan and the United States as the main technologies.

Keywords-Face Recognition; Model Theory; Image Processing

I. INTRODUCTION

The realization of face recognition technology has three steps, the first step, collect the image of

the face as a sample, the larger scientific research institutions have their own face database, the second step, extract the need to process the key information of the face, the last step is to compare with their own face database. Face recognition technology has made important contributions in the fields of psychology, criminology, medicine and so on. China began to put into the research of this technology late, but still has made outstanding results. In the later stage, another face recognition processing technology was proposed by Blanz et al. This method simulates the 3D face image space, and then matches the 3D face model with the image space, so as to estimate the shape and texture information of the face. This algorithm has a high recognition capability, the recognition rate is up to 88 % [1]The research methods that need to be adopted for face information usually rely on these four specific operation steps to achieve: The first step is face localization and detection: For a given image information, first of all, whether it is belong to face images or contain face, if contain face, defining the face area in the image, then extracted face part, because there is much easier to change the face itself factors, and is strongly influenced by the external environment, so in order to realize the recognition of human faces and detect need quite mature technology, Otherwise, it is difficult to achieve efficient face recognition. The second step is facial feature extraction: usually after the

machine determines the image as a face, it will use some algorithm to extract the facial feature data, and then save it to the face database. The most important algorithms are surface and algebraic features. The third step is the real stage of recognition: will obtain the face image and the database exists in the template to match, by comparing the same and different places to get the correct recognition results. This stage is almost not affected by external factors, mainly to see whether the algorithm is appropriate, can effectively identify, and obtain the correct face feature information. The fourth step is external interference project analysis: summarize and analyze the external environmental impact. The algorithm is improved to overcome the influence of these objective factors [2].

II. RESEARCH METHODS OF FACE RECOGNITION

Template method is the most classic pattern recognition method to solve the problem of face recognition, this method repeatedly uses the characteristics of face texture and different gray values, the main principle of this method is to identify the face image and all the templates in the database to find the best matching effect of the face. Because this method has great requirements on the image template, the two templates need to have the same proportion, the same face orientation and the same lighting conditions, so the ratio normalization and gray value normalization of the two templates are needed in the preprocessing stage. The most easy to implement method is to face image approximation as oval, then, in recognition of the image to find elliptic, another method is a set of templates to represent the whole face template, this template including small template features of various organs of the human face, but this approach is more demanding, must use each feature of different edge profile, The traditional method is

based on edge extraction, so it is difficult to obtain continuous edge data. Even if relatively continuous and reliable data are obtained, it is difficult to automatically extract the eigenvalues of different small templates that we need. Later, variable templates were used to extract the edge information of different templates in template group. The variable template consists of setting a set of variable characteristic parameters according to the existing knowledge, which depends on the special value of the energy function. Firstly, the prior knowledge of boundary value, peak value, trough value, trough information and image shape information is used to design the function, and then the parameters are modified in the direction of energy function reduction. At a minimum, the template shape corresponding to this set of parameter locks is most consistent with the feature shape. Although the recognition method based on feature setting has faster recognition speed and smaller memory consumption, the method based on elastic template is better than the traditional template method in recognition rate and other factors. Therefore, some people specially compare the recognition method based on facial geometric features and the recognition method based on elastic template. The result shows that elastic template method belongs to geometric feature method.

The essence of feature face method is realized by K-L transform, which is the best orthogonal transform method for image compression in image processing technology. Kirby and Sirovich successfully described the method of face image information through component analysis. They calculated the coordinate system with the best image compression rate for the original image information. Each coordinate is actually the eigenpoint coordinate system that they define. [4] They argue that theoretically, any face image template can be reconstructed from two sets of

images, one is the image of each face, and the other is the standard template image, also known as the eigenimage. The number of repetitions can be obtained by mapping the face image with the corresponding feature image by projection. Turk and Poggio argued that a great deal of facial information could be obtained if a set of feature images could be weighted and reconstructed. The effective recognition method is to identify the human image by analyzing the complex features of the image and comparing the feature weights needed to reconstruct the image. Thus, each face can be represented by a set of reconstructed weights. In short, vectors of higher dimensions are projected into the region of vectors of lower dimensions by matrices with specific characteristics. The eigenmatrix is characterized by a low-dimensional vector and retains the main information. That is, the corresponding high-dimensional vector can be completely reconstructed through the state eigenvector and eigenvector matrix. Compared with the image itself, this expression can completely reconstruct the corresponding high-dimensional vector. This expression is a highly compressed form relative to the image itself [5][6]. Artificial Neural Network (ANN) has a good ability to classify complex patterns, which leads to its widespread use in face recognition. There are many different models of neural network models, each of which describe and simulates different levels of biological nervous system from different perspectives. Representative neural network models include perceptron, multilayer mapping BP network, RBF network, Hopfield model and so on. At present, in

the practical application of artificial neural network, most of the neural network models are BP network and its variants. It is the most important part of artificial neural network [7].

III. CLASSIFICATION OF ALGORITHMS

There are three methods to detect facial feature points: the traditional method based on active shape model (ASM) and active appearance model (AAM), the method based on cascade shape regression, and the method based on deep learning. When the method is classified by whether there are parameters, ASM, AAM and CLM are parametric methods, while cascade regression and deep learning are non-parametric methods. Methods based on parametric shape model can be divided into local method and global method according to different shape models. Non-parametric methods can be subdivided into sample-based methods, graph-based methods and cascaded regression methods. And a deep learning-based approach. See the literature for a more detailed division. Currently, the most widely used and accurate method is based on deep learning. Therefore, in this paper, we mainly study the application of deep learning in face important point detection. Then, referring to the comprehensive review of face detail detection, two main face detection methods are proposed: parametric and non-parametric. This classification method is more acceptable and can make people better understand the meaning of the existence of parameters.

Figure 1 shows the classification figure:

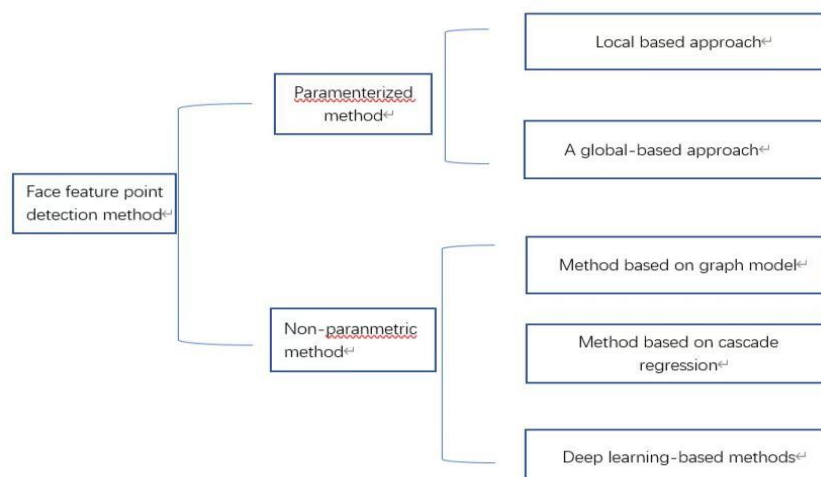


Figure 1. Classification diagram

A parametric model is a data distribution of interest to a particular target, such as a Gaussian model or Gaussian mixture Dell. Methods based on nonparametric models do not meet the requirements fundamentally, and their data should be specified to meet the requirements on the basis of distribution. Main models or non-parametric models (former) have fixed parameters, the latter numbers of parameters are Elites parameters Bakelites, Hikaru, have Hikaru electronic games with grades (Manufacture Kute STAB Silicone S Factory Yip electronic games, Hikaru of Hikaru Nad and Hikaru with Corning. Non-parametric models with video Hikaru are graph-based methods, cascaded regression based methods, deep learning and other methods [8].

The basic standard at this stage is to see the algorithm to obtain the face feature points and the actual face point between the deviations. In the quantitative evaluation of the deviation, because the face image is different, the actual size of the shooting Angle and distance will be different, in order to facilitate the comparison of the performance of the algorithm under the same conditions, the method of data normalization

should be used to maintain the constant factors. At present, the main reference method is to rely on the method of judging the distance between the two eyes for face size standardization.

A. Common database

Active Database: CMU Multi-PIE Face Database, formed by collecting multiple face data at four conferences between October 2004 and March 2005, supports the development of face recognition algorithms in a variety of conditions, such as posture, lighting conditions and expression. The database contains 337 topics and over 750,000 images of 305GB of data. Six different expressions were recorded: expressionless, smiling, surprised, leering, disgusted and screaming. Objects were recorded in 15 views and 19 different lighting conditions. A subset of the database is labeled with 68 or 39 points. The XM2VTS database collects 2360 color images, sound files and 295 people's 3D facial models. The 2360 color images have 68 key points. The AR database contains more than 4,000 color images of the faces of 126 people (70 men and 56 women). With all the uncontrollable conditions in place, Ding and Martinez manually

labeled 130 key points for each face image. The IMM database contains 240 color images of 40 people (7 women and 33 men). Each image is marked with eyebrows, eyes, nose, mouth and chin, a total of 58 markers. The MUCT database contains 3755 images of 276 people, each with 76 keypoints. The faces in the database were photographed under different lights, different ages and different races.

The PUT database collected 10,091 high-resolution images (2048×1536) of 10,091 individuals with partially controlled lighting conditions and rotation along pitch and yaw angles. Each image is labeled with 30 key points.

B. Face++ version DCNN

This paper uses Face++ Megvii's face recognition API. The API algorithm is improved on the traditional DCNN model, and a coarse-grained to fine-grained human face key point detection method is proposed, which mainly solves the problem of high-precision positioning of 68 human face key points. The algorithm divides the key points of the face into internal key points and external contour key points. Inside the face is not the main point eyebrows, eyes, nose, mouth a total of 51 key points, contour key points contain 17 key points. For internal keypoints and external keypoints, the algorithm uses two cascaded CNNs for keypoint detection in parallel. The network structure is shown in Figure 2 below.

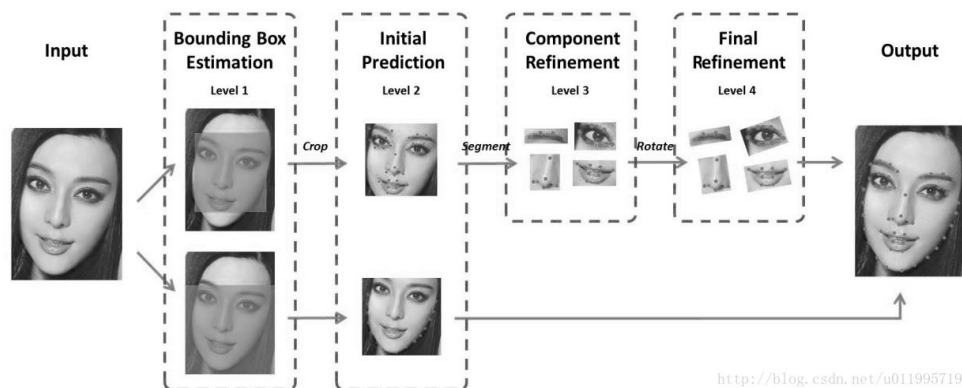


Figure 2. Network structure

For internal key points, a cascade network of four levels is adopted to continue the detection. Among them, the first layer is mainly used to locate the boundary of facial organs. The second layer is the position where the output predicts the 51 key points, which is a coarse-grained positioning for the purpose of initialization for the third layer. The third layer will be based on the features of different wonders of the face from coarse-grained to fine-grained positioning; The input of the fourth layer is to rotate the output of the third layer to a certain extent, and finally output the positions of 51 key points. For the 17

external key points, only a two-layer cascading network is used for detection. The first layer plays the same role as the internal key point detection, mainly to obtain the bounding box of the face contour; The second layer directly predicts 17 key points, cancelling the process from coarse to fine positioning, because the area of external key points is very large. If the third and fourth layers are added, a lot of time will be consumed. The final 68 key points of the face were superimposed by the output of two cascaded CNNs [15].

The main innovations of the algorithm are as follows:

1) The positioning of the face key point is divided into two parts of the internal key point and the external contour key point you to predict. This is very effective to avoid the problem of uneven feature point selection;

2) When detecting the internal points of the face, it does not use two cascaded CNNs for each key point to predict, as DCNN does, but only uses one CNN for each organ to predict, thus reducing the amount of calculation;

3) Compared with the traditional DCNN, it does not take the results of face collection as direct input, but adds a boundary localization for face border monitoring, which can effectively improve the network accuracy of coarse localization [9]. The structure and distribution area of the points are shown in Figure 3.



Figure 3. Point structure

The structure of the key points contains all the organs of the face respectively, including the contour of the eyes, nose, mouth and eyebrows. The green points in the figure are the key anchor points in the feature points. These key positioning points are the main forms of coarse precision positioning points, which are used to determine the rectangular position of the face, the distance between the outside of the eyes and the face, the distance between the two eyes, the distance between the eyes and the nose, the distance between the nose and the mouth, and the distance

between the mouth and the chin. In addition, cheek feature points are included to distinguish the gender and race of a face. At present, the latest key points are 83 and 106 points, which are the embodiment of the convolution network with real fine accuracy. Because of the increase of key points, the facial information that can be recognized becomes richer and more in line with daily needs.

IV. CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks (CNNs/ConvNets) are very similar to normal neural networks in that they are composed of neurons with learnable weights and biases. The simulation diagram of common neural network is shown in below Figure 4.

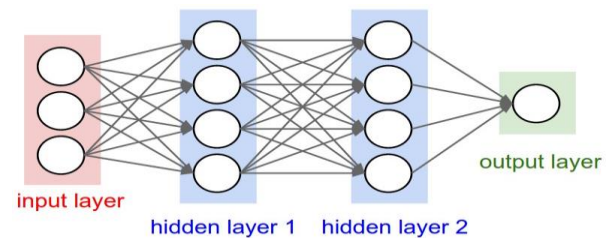


Figure 4. Traditional neural network diagram

Each neuron is involved in the calculation, and the dot product calculation is performed, and the output result is the score value of each class. The calculation method applicable in the traditional neural network is still used here. A convolutional neural network is composed of many layers, and the output is 3D. Some layers have parameters, and some layers have no parameters. General convolutional neural network contains the following layers.

Convolutional layer: Each layer in the convolutional neural network is composed of multiple convolutional units, and each parameter of the convolutional unit is optimized by the back propagation algorithm. Convolution operation is used to extract different input features. The first

layer can only extract some low-level features such as borders, lines and corners, while the multi-layer network is used to continuously screen and iterate complex features from low-level features. Pooling layer: usually, features with large dimensions will be obtained after the first layer is processed. The features will be cut into several areas to obtain the average or maximum value and new features with smaller dimensions will be obtained.

The reasonable Satisfaction Linear Units layer (ReLU layer) shall be provided to the reasonable Satisfaction Linear Units (ReLU) for nerve Activation function in this layer. Pooling layer: usually, features with large dimensions will be obtained after the first layer is processed. The features will be cut into several areas to obtain the average or maximum value and new features with smaller dimensions will be obtained. Fully Connected layer, which combines all local features to form a global feature, is used to calculate the final score of each type of feature. [10] The simulation diagram of convolutional neural network is shown in Figure 5 below:

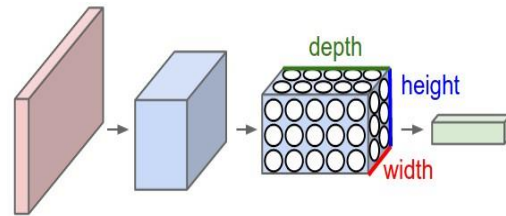


Figure 5. Convolutional neural network diagram

Modern mainstream machine learning libraries and interfaces, including TensorFlow, Keras, Thenao, Microsoft-CNTK, etc., can run convolutional neural network algorithms. In addition, some commercial numerical calculation software, such as MATLAB, also has the construction tool of convolutional neural network available.

V. CALLING PLATFORM API

The library of Python is used to send an HTTP request to the Face++ server, and upload a face picture captured by the camera. The server uses the feature point algorithm to analyze and process the picture, and returns the processed result as a object.

Figure 6 shows the execution flow of the program:

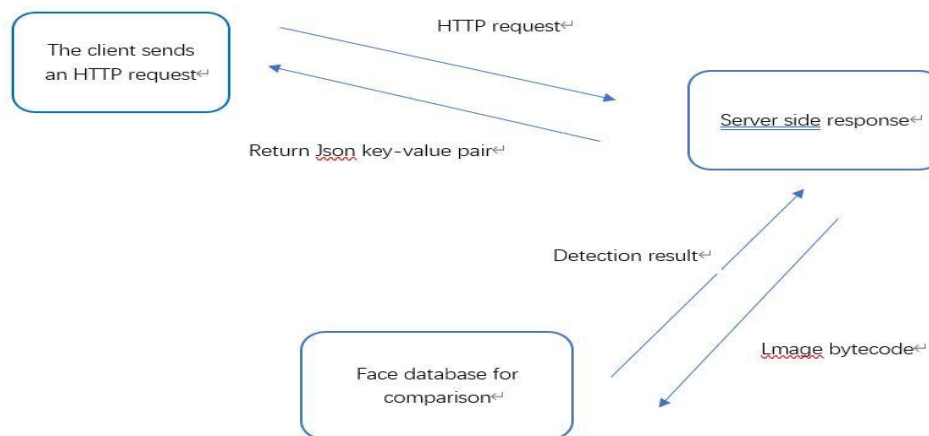


Figure 6. Flow chart of program execution

The program execution result is shown in below Figure 7.

```

Python 3.8.2 Shell
File Edit Shell Debug Options Window Help
Python 3.8.2 (tags/v3.8.2:7b3ab59, Feb 25 2020, 22:45:29) [MSC v.1916 32 bit (Intel)] on win32
Type "help", "copyright()", "credits()" or "license()" for more information.
>>>
-----RESTART: F:\PythonSpace\face.py -----
[{"request_id": "1588062839.81335679-f52c-44aa-8432-2a27b5a292d8", "time_used": 182, "faces": [{"face_token": "9aeb2323d26dc3829134025
ab41445", "face_rectangle": {"top": 425, "left": 220, "width": 346, "height": 346}, "landmark": {"contour_chin": [{"x": 396, "y": 771}, "contour_l
eft1": [{"x": 235, "y": 600}, "contour_left2": [{"x": 235, "y": 538}, "contour_left3": [{"x": 238, "y": 578}, "contour_left4": [{"x": 244, "y": 613}, "co
ntour_left5": [{"x": 253, "y": 651}, "contour_left6": [{"x": 270, "y": 685}, "contour_left7": [{"x": 293, "y": 714}, "contour_left8": [{"x": 322, "y": 7
40}, "contour_left9": [{"x": 365, "y": 762}, "contour_right1": [{"x": 560, "y": 493}, "contour_right2": [{"x": 560, "y": 535}, "contour_right3": [{"x"
: 547, "y": 573}, "contour_right4": [{"x": 541, "y": 610}, "contour_right5": [{"x": 533, "y": 648}, "contour_right6": [{"x": 517, "y": 682}, "contour_r
ight7": [{"x": 495, "y": 712}, "contour_right8": [{"x": 469, "y": 739}, "contour_right9": [{"x": 437, "y": 761}, "left_eye_bottom": [{"x": 323, "y": 503
}, "left_eye_center": [{"x": 324, "y": 496}, "left_eye_lower_left_corner": [{"x": 295, "y": 493}, "left_eye_lower_left_quarter": [{"x": 307, "y": 500}, "l
eft_eye_lower_right_quarter": [{"x": 339, "y": 502}, "left_eye_pupil": [{"x": 327, "y": 492}, "left_eye_right_corner": [{"x": 354, "y": 500}, "left
_eye_top": [{"x": 324, "y": 485}, "left_eye_upper_left_quarter": [{"x": 308, "y": 487}, "left_eye_upper_right_quarter": [{"x": 341, "y": 489}, "lef
t_eyebrow_lower_left_corner": [{"x": 271, "y": 458}, "left_eyebrow_lower_left_quarter": [{"x": 291, "y": 454}, "left_eyebrow_lower_middle": [{"x": 313
, "y": 449}, "left_eyebrow_lower_right_quarter": [{"x": 336, "y": 447}, "left_eyebrow_right_corner": [{"x": 360, "y": 443}, "left_eyebrow_upper
_left_quarter": [{"x": 285, "y": 435}, "left_eyebrow_upper_middle": [{"x": 309, "y": 425}, "left_eyebrow_upper_right_quarter": [{"x": 337, "y": 427
}, "mouth_lower_lip_bottom": [{"x": 397, "y": 703}, "mouth_lower_lip_left_contour1": [{"x": 366, "y": 670}, "mouth_lower_lip_left_contour2": [{"x": 348, "y": 685}, "mouth_lower_lip_left_contour3": [{"x": 369, "y": 698}, "mouth_lower_lip_right_co
ntour1": [{"x": 426, "y": 671}, "mouth_lower_lip_right_contour2": [{"x": 441, "y": 686}, "mouth_lower_lip_right_contour3": [{"x": 423, "y": 695}, "mou
th_lower_lip_top": [{"x": 397, "y": 672}, "mouth_right_corner": [{"x": 454, "y": 668}, "mouth_upper_lip_bottom": [{"x": 396, "y": 668}, "mouth_u
pper_lip_left_contour1": [{"x": 381, "y": 644}, "mouth_upper_lip_left_contour2": [{"x": 355, "y": 653}, "mouth_upper_lip_left_contour3": [{"x":
365, "y": 667}, "mouth_upper_lip_right_contour1": [{"x": 413, "y": 644}, "mouth_upper_lip_right_contour2": [{"x": 436, "y": 653}, "mouth_upper_l
ip_right_contour3": [{"x": 425, "y": 667}, "mouth_upper_lip_top": [{"x": 397, "y": 648}, "nose_contour_left1": [{"x": 368, "y": 500}, "nose_contour
_left2": [{"x": 358, "y": 566}, "nose_contour_left3": [{"x": 371, "y": 605}, "nose_contour_lower_middle": [{"x": 395, "y": 611}, "nose_contour_righ
t1": [{"x": 422, "y": 500}, "nose_contour_right2": [{"x": 432, "y": 566}, "nose_contour_right3": [{"x": 419, "y": 604}, "nose_left": [{"x": 350, "y": 59
4}, "nose_right": [{"x": 440, "y": 594}, "nose_tip": [{"x": 395, "y": 582}, "right_eye_bottom": [{"x": 468, "y": 502}, "right_eye_center": [{"x": 467,
y": 496}, "right_eye_lower_left_corner": [{"x": 438, "y": 502}, "right_eye_lower_left_quarter": [{"x": 452, "y": 502}, "right_eye_lower_right_quarter
": [{"x": 482, "y": 499}, "right_eye_pupil": [{"x": 465, "y": 492}, "right_eye_right_corner": [{"x": 494, "y": 495}, "right_eye_top": [{"x": 467, "y": 4
85}, "right_eye_upper_left_quarter": [{"x": 450, "y": 490}, "right_eye_upper_right_quarter": [{"x": 482, "y": 488}, "right_eyebrow_left_corner
": [{"x": 431, "y": 444}, "right_eyebrow_lower_left_quarter": [{"x": 453, "y": 449}, "right_eyebrow_lower_middle": [{"x": 476, "y": 452}, "right_ey
ebrow_lower_right_quarter": [{"x": 497, "y": 457}, "right_eyebrow_right_corner": [{"x": 518, "y": 461}, "right_eyebrow_upper_left_quarter": [{"x":
453, "y": 451}, "right_eyebrow_upper_middle": [{"x": 479, "y": 431}, "right_eyebrow_upper_right_quarter": [{"x": 503, "y": 440}], "attributes
": {"gender": {"value": "Male", "age": {"value": 17}, "smile": {"value": 0.008, "threshold": 50.000}, "headpose": {"pitch_angle": 7.5917764, "r
oll_angle": -1.0098664, "yaw_angle": -0.64553905}, "blur": {"blurriness": {"value": 0.625, "threshold": 50.000}, "motionblur": {"value": 0.625,
"threshold": 50.000}, "gaussianblur": {"value": 0.625, "threshold": 50.000}}, "eyestatus": {"left_eye_status": {"no_glass_eye_open": 99.981, "no_glass_eye_close": 0.000, "normal_glass_eye_open": 0.019, "normal_glass_eye_close": 0.000, "dark_glasses": 0.000, "occlusion": 0.000}, "right_eye_status": {"no_glass_eye_open": 99.997, "no_glass_eye_close": 0.000, "normal_glass_eye_open": 0.002, "normal_glass_eye_close": 0.000, "dark_glasses": 0.000, "occlusion": 0.000}, "emotion": {"anger": 4.654, "disgust": 0.095, "fear": 0.081, "happiness": 5.165, "neutral": 88.657, "sadness": 0.928, "surprise": 0.420}, "facequality": {"value": 92.706, "threshold": 70.100}, "ethnicity": {"value": ""}, "beauty": {"ma
le_score": 77.419, "female_score": 80.902}, "mouthstatus": {"surgical_mask_or_respirator": 0.000, "other_occlusion": 0.000, "close": 99.962, "open": 0.038}, "eyegaze": {"left_eye_gaze": {"position_x_coordinate": 0.522, "position_y_coordinate": 0.376, "vector_x_component": 0.966, "vector_y_component": -0.107, "vector_z_component": 0.992}, "right_eye_gaze": {"position_x_coordinate": 0.468, "position_y_coordinate": 0.397, "vector_x_component": -0.085, "vector_y_component": 0.017, "vector_z_component": 0.996}}, "skinstatus": {"health": 60.122, "stain": 2.372, "dark_circle": 7.832, "acne": 5.120}, "glass": {"value": "None"}]]}], "image_id": "M13qlorjhm/DbE8huAm70w==", "face_num": 1}
>>> |
    
```

Figure 7. Program execution result diagram

The returned values are presented as Json key-value pairs. These return values contain information about the success or failure of the request, the cause of the request failure, and the error code information. The error code information contains both standard HTTP error code and API-specific error code information. When the request is sent successfully, the image information will be returned, when the size of the image exceeds the standard size, or the image path is wrong, the parameter abnormal information will be reported, when the image path and size are in line with the specification, continue to return the detected face information, there are several faces. The facial feature data in the return value includes facial clarity, facial posture, whether glasses are worn, light environment, and objective information such as age, race, gender, and appearance level scores.

Due to the different face data required under different environmental requirements, the face feature information returned by the API of Face++ platform contains many aspects, which can meet the general needs of the current market from different levels. Table 1 describes the returned values.

TABLE 1. DESCRIPTION OF RETURNED VALUES

Field	Type	Instructions
Request_id	string	Distinguish each request, unique value
Faces	An array of	The array of faces that were detected
Image_id	string	The identification of the detected picture
Time_used	The integer	The time taken for the entire request, in milliseconds.
Error_message	string	Request failure information
Face_num	The integer	The number of faces recognized

The information in the returned faces array contains the face identifier, the position of the face rectangle box, the coordinates of the key points and the features of the face attributes, as shown in Table 2.

TABLE II. FACES ARRAY VALUES

Field	Type	Instructions
Face_token	string	Identification of human face
Face_rectangle	object	The position of the face box contains the following attributes.
		Top: The ordinate of the upper left corner of the face box
		left: The abscissa of the upper left corner of the face box
		width: The width of the face frame
		Height: The height of the face frame
Landmark	object	Coordinates of key points
Attributes	object	Facial feature information

After in solving the above errors, return to face feature information and standard test has great discrepancy in the document information, after the inquiry document to call API version and the latest version does not conform to, in the new version of the API, remove the glass this information, and will be integrated into the eye status information, this part of the information after the adjustment of API version, The correct facial feature information is shown below. After reviewing the technical documentation, the path to the server receiving the request was modified based on the updated version. However, the program still returns error information. After checking the error information, the initialization parameter error caused by coding problem was found. After comparing with the technical documentation of the developer platform, the coding format of the development environment was modified, and the request was successfully

sent. After the request is sent successfully, the information returned is not the desired face feature information, and the API specific error code 413 is returned. After viewing the developer's technical documentation, it is found that when selecting the face image, the size of the image is not noticed, resulting in the image corresponding to the parameter cannot be correctly parsed.

Whether large Web system, Java App program, or a small Demo, have to undergo testing, only can be found in the test not appropriate, only when these problems will be improved, the program will be more introduction, strong, if too many redundant code, coupling relationship is too strong, will cause the program modification and optimization in the late very difficult.

Face recognition has many applications in people's production and life, initially as an alternative to fingerprint authentication. Compared with fingerprint authentication, face recognition authentication has the convenience of contact-free, which makes it easier for the public to accept [10].

But it's worth noting that when we authorize an app to identify us, it's likely to bundle our facial information with private identity information or even address information. This often means that when our face information is leaked, other bound information will be leaked along with the face information. An American company specializing in AI faces processing and recognition technology, has designed and released a face recognition application Clearview AI as shown in below Figure 8.

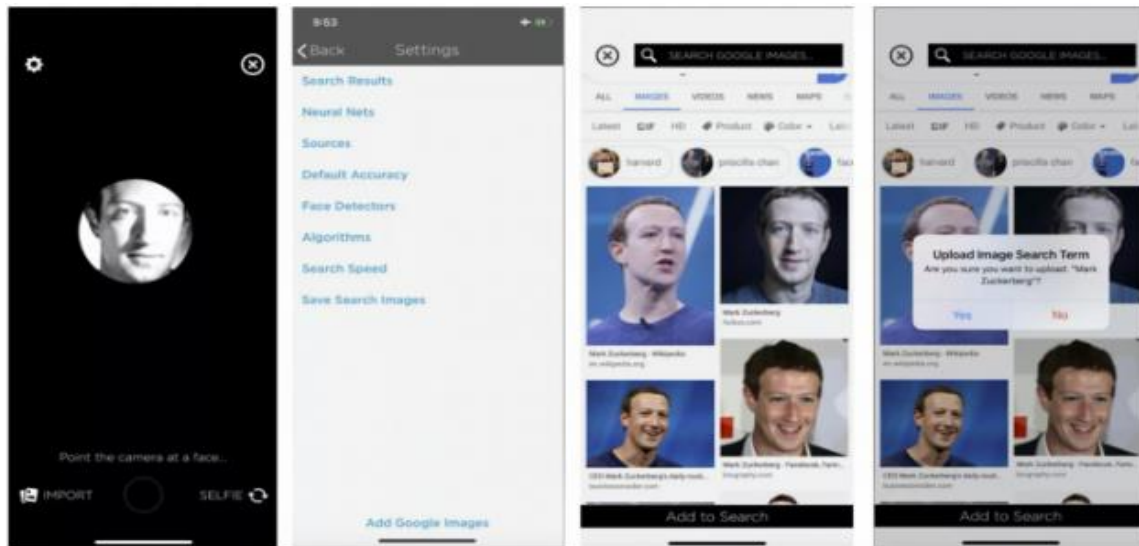


Figure 8. Facial recognition social program

Depending on the software to obtain the user's private pictures, and then upload them, you can normally browse the person's photos, as well as point to the specific information and shooting point of these photos. This photo-based technology isn't big news. Back in 2016, FindFace, a Russian startup, enabled matching a photo with a profile photo on Vkontakte, a Russian social network similar to Facebook. I can find strangers' social media accounts [13]. Although the App claims to help with "networking and intentional building". But the problem is that this is a huge database of faces. If Find Face fails to properly use the data -- school profiles, corporate customer lists, building comings and goings, mall customer movements, and so on -- harassment messages, death threats, and other incidents may occur, and facial recognition technology may violate our privacy in ways we can't imagine.

We all know that in the application of face detection identity, the information of a person's face is like a password, the password can be cracked, of course, the face can also be. As face recognition technology is widely used in applications such as face scanning authentication

and face scanning payment, news of various face recognition systems being cracked continues to emerge. 3D structured light in vivo detection for the iPhone X app has been described as the most secure facial recognition detection technology in use today, but in 2017, a tech company managed to trick Apple's Face ID into using a human face mask produced by a 3D printer, as shown in Figure 9 below [10][11].

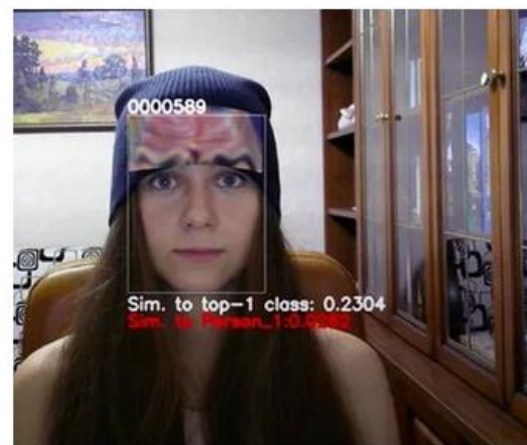


Figure 9. "Patch" cracking Face ID

Due to the technology cost is too high, we can now access to the most widely used facial recognition program is not structured light 3 d live

detection, and is only a 2 d plane detection or motion capture recognition, so we can often hear such as abundant nest "brush face take" with an inch of photo identification by criminals, intelligent security doors can use synthetic video unlock success. Considering the cost of cracking methods, the cracking methods of 3D printing materials are more flexible, but the experimental cost is too high. From the perspective of the success rate of recognition, planar photos cannot crack recognition.

Recently, a very clever method popular in machine learning technology circles is to confuse the recognition system by adding redundant image information that is difficult to recognize by the naked eye. This method is a way to trick the recognition operating system into producing errors. In 2019, researchers at Huawei's Moscow research center used a printer to print a partial image of a face and then used it to identify the face in a recognition system, an attempt that tripped up many well-known face recognition systems.

Future application prospects and a large number of benefits let face recognition technology is constantly touted by the capital

market, but also invisible to promote the exploration of technology enterprises for emerging technologies. In order to promote their own face recognition system, many enterprises develop a new way, constantly for the development of face recognition technology of multiple application scenarios.

But as we all know, the current face recognition technology still has a lot of defects, the effect of many face recognition systems is not very good, which leads to the face recognition system in some scenes that do not need to use it is abused. The wrong perception that the main app does not see the effect also exposes the risks of such blind use. Such as some schools use face recognition door card machine, its starting point is for the sake of students' safety, but the technology adopted by the belong to the false face recognition technology, is not calculated according to the feature points or vector template, but simple through photo database to compare, so not only do not implement the initial security purposes, but the way of personal information may be leaked students, There are many non-compliant face recognition machines on the market. As shown in Figure 10.



Figure 10. Campus access control system

VI. CONCLUSION

Although face recognition technology brings a lot of convenience to our life, but also let us have a broader imagination for the future, but in the history of the battle between machine and human, most people win, so we need to ensure that the new technology is more accurate application, rather than no standard of active attempt. Due to the special circumstances of the pandemic this year, masks have become a necessity and are required to be worn in any public place. However, it also brings some troubles for people to be sentenced, wearing masks, mobile phones can not be unlocked, station access can not be identified.

Earlier, Danielle Baskin, a product designer, released a mask that looks like a Face patch, allowing you to unlock your phone with Face ID. In short, patches are used to complete incomplete facial information so that the machine can recognize it.

Although, as Danielle Baskin admitted on Twitter, the recognition rate is not guaranteed at this stage. However, once the news was released, many people doubted the reliability of face recognition technology. If someone wore a mask with their face information to unlock their phone, could it be unlocked? Therefore, for the future of face recognition, we should be in a state of

experimentation, can not blindly expand the scope of use, if wrong, the impact may be world-class.

REFERENCES

- [1] Rui Gu. Overview of Occlusion Face Recognition Technology, 2018.
- [2] Li-jun Yu. Application of Face Recognition Technology in the Exam Identity Authentication System, 2017.
- [3] T.F. Cootes, C.J. Taylor, D.H. Cooper, et al. Active Shape Models-Their Training and Application [J]. *Computer Vision and Image Understanding*, 1995, 61(1):38-59.
- [4] G. J. Edwards, T. F. Cootes, C. J. Taylor. Face recognition using active appearance models [J]. *Computer Vision — Eccv'*, 1998, 1407(6):581-595.
- [5] Cootes T F, Edwards G J, Taylor C J. Active appearance models[C]// *European Conference on Computer Vision*. Springer Berlin Heidelberg, 1998:484-498.
- [6] Dollár P, Welinder P, Perona P. Cascaded pose regression [J]. *IEEE*, 2010, 238(6):1078-1085.
- [7] Sun Y, Wang X, Tang X. Deep Convolutional Network Cascade for Facial Point Detection[C]// *Computer Vision and Pattern Recognition*. IEEE, 2013:3476-3483.
- [8] Zhou E, Fan H, Cao Z, et al. Extensive Facial Landmark Localization with Coarse-to-Fine Convolutional Network Cascade[C]// *IEEE International Conference on Computer Vision Workshops*. IEEE, 2014:386-391.
- [9] Zhang Z, Luo P, Chen C L, et al. Facial Landmark Detection by Deep Multi-task Learning[C]// *European Conference on Computer Vision*. 2014:94-108.
- [10] Wu Y, Hassner T. Facial Landmark Detection with Tweaked Convolutional Neural Networks [J]. *Computer Science*, 2015.
- [11] Zhang K, Zhang Z, Li Z, et al. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks [J]. *IEEE Signal Processing Letters*, 2016, 23(10):1499-1503.