Research of Micro-expression Recognition Model based on Feature Unit

Fei Yin
School of Computer Science and Software Engineering East China Normal University
Shanghai 200062, China
51174500147@stu.ecnu.edu.cn

Jinyi Xu
MOE Enginering Research Center for Software/Hardware Co-design Technology and Application East China Normal University
Shanghai 200062, China
52184501010@stu.ecnu.edu.cn

Yixiang Chen
MOE Enginering Research Center for Software/Hardware Co-design Technology and Application East China Normal University
Shanghai 200062, China
yxchen@sei.ecnu.edu.cn

***Abstract*—Micro-expression which is the transient expression will be disclosed when people try to hide some kind of real inner emotions. Micro-expression changes so fast that few people detect its existence. As an effective behavioral clue, it is of great significance to understand the change of peoples true feelings. Based on this, image feature extraction in machine learning has made remarkable progress in the past two years.In order to improve the accuracy and practicability of micro-expression recognition, this paper deeply analyzed the practical micro-expression recognition method. Based on landmark detection. We construct a quantitative model and established a machine learning model of micro-expression judgment based on the comprehensive changes of eyes,, mouth and eyebrows. This paper combine the feature unit and model application scene for judgement of micro-expression. This method reflect the design idea that the changes of face locality contributed to the overall micro-expression judgment, and helped to complete the high autonomous construction of the micro expression judgment model.**

***Index Terms*—Face Detection, Micro-expression Recognition Machine Learning Model, Facial Landmark Detection, Feature Unit**

I. Introduction

Micro-expression can be regarded as rapid and subtle facial expression change. The duration of it start to end from 1/25 seconds to 1/5 seconds [1], which is hard to be perceived. Ekman believes that micro-expressions express the real emotions that humans try to hide [2], which often appear when people lie and conceal. It is a spontaneous and uncontrollable change in expression and movement. Micro-expressions are ubiquitous in social life and existing in various fields of life. It is a kind of reliable non-verbal clue, which reflects people's true inner feelings and potential motivations in specific situations. Therefore, the research on micro-expression recognition has important theoretical value and practical significance.

In 1969, Ekman and Friesen independently firstly named the transient movement of facial expression as Micro-expressions. Since then, researchers achieve some outstanding work in such field, such as micro-expression theory and micro-expression recognition. Polikovskys team in Japan used a 3D gradient direction descriptor on a self-built Micro-expression database to automatically identify Micro-expressions in the video stream [3]. They used a camera with a high frame rate of 200 fps to capture video clips for Micro-expression recognition, dividing the faces in the video into 12 regions, to calculate the 3D gradient direction histogram information HOG [5] about each facial region. They also applied K-means clustering method to realize the classification of micro-expressions.

Wang's team [6]used the sparse part of the robust principal component analysis method to extract the imperceptible motion information of the micro expression movement in 2014. For the local texture features of the face, they used the local spatio-temporal direction feature (LSTD) algorithm to extract. Experiments on two public micro-expression datasets, SMIC [7]and CASME II [8], show that their method achieves better performance. Guo [9]proposed a micro-expression recognition method based on CBP-TOP and ELM in 2015. CBP-TOP is different from the traditional CBP feature. It is a three-dimensional feature and can effectively extract the facial micro-expression information in the space-time dimension, and then put the extracted face CBP-TOP feature into the extreme learning machine. (ELM) [10]for multi-category learning of micro-expression recognizers. Compared with the traditional recognition method, this method can more effectively extract the motion feature information and dynamic texture of micro-expressions, and greatly improve the recognition rate of micro-expressions.

Based on the researches of Micro-expression, it is widely used in all aspects of human society. However, such kind of facial movements are too short and subtle to be recognized exactly using the existing methods. Simple facial expression classification cannot classify the micro facial expression recognition accurately. Existing algorithms for automatic recognition of micro-expressions have low recognition accuracy on the public data sets and poor applicability in actual physical scenes. Immature recognition technology influences the development of Micro-expression research. In order to solve such problems in the micro-expression recognition field, this paper combines automatic feature point change detection model with autonomous micro-expression judgment model. The main research work of this paper is as follows: Construct a change detection model based on landmarks: We locate the feature

978-1-7281-8915-4/20/$31.00 ©2020 IEEE
DOI 10.1109/QRS-C51114.2020.00080

points of each key part of the face. We design different analysis methods based on local facial feature changes to quantify the classification results of local features according to different parts of face. Design a micro-expression judgment machine learning model based on the comprehensive changes of eyes, mouth and eyebrows. We define the concept of Feature Unit (FU), and formulate the description and rules of FU. We design the model of micro-expression expression, which provides a quantified FU combination structure for micro-expression recognition. From face detection, analysis and determination of local feature changes to full-time FU results, we combine automatic detection and artificial experience recognition.

The rest of this paper is organized as follows: In Section II, we introduce the development of face alignment and micro-expression data sets. We provide Micro-expression Recognition Model in section III and prove the superiority of this model through experiments in section IV.

II. RELATED WORKS

In this section, we focus on the development of face alignment and current research of micro-expression data sets.

A. Face Alignment

Face Alignment is to track the distribution of landmarks of various facial organs in the image, including the feature contours of mouth, eyes, eyebrows, etc. This approach has been applied in several fields, such as facial expression recognition and target tracking.

Kazemi proposed one millisecond face alignment with an ensemble of regression trees in 2014 [11]. This algorithm cascades multiple gradient regression trees. It gradually returns the current predicted shape to the true shape of the face. The ERT algorithm takes only 1ms. It has fast extraction speed and high accuracy. This algorithm uses the features of iterative recursion with the cascade gradient boosting tree. The face shape changes from the original shape to the real shape using this algorithm. A regressor γ_t is stored on the leaf nodes of the cascade gradient tree. When the path of the feature input regression passes a certain leaf node, γ_t is added to the input to obtain the current estimated shape. As shown in the formula.

$$\hat{S}^{(t+1)} = \hat{S}^{(t)} + \gamma_t(I, \hat{S}^{(t)}) \tag{1}$$

t is the number of cascades, $\hat{S}^{(t)}$represents the coordinate shape of the landmarks of the t regressor, which stores the position information of all key points of the face as a vector. I is the image, and γ_t represents the regressor at the current level, which use the gradient improvement algorithm for training. The input is the current shape vector. And the output is the residual amount of the position. $\gamma_t(I, \hat{S}^{(t)})$ represents prediction residual, calculated by the current regressor based on the image I. The feature shape $\hat{S}^{(t)}$ represents the predicted shape of t iteration regressions.

The key of Kazemi's algorithm is the training of the feature residual γ_t. For each γ_t, this model uses a gradient-enhancing tree algorithm based on the sum of squared errors. Assuming training set $\{(I_{\pi i}, \hat{S}_i^{(t)}, \triangle S_i^{(t)})\}_{i=1}^{N}$,and training learning rate is $0 < v < 1$, we get residuals γ_t through iterative integration:

1) initialize:

$$f_0(I, \hat{S}^{(t)}) = \underset{\gamma \in R^{2p}}{\arg\min} \sum_{i=1}^{N} \left|\left|\Delta S_i^t - \gamma\right|\right|^2 \tag{2}$$

2) for k from 1 to K:
   a) set $i = 1, 2, 3, ..., N$, then

$$\gamma_{ik} = \Delta S_i^{(t)} - f_{k-1}(I_{\pi i}, \hat{S}_i^{(t)}) \tag{3}$$

   b) Fit a regression tree to the targets γ_{ik} giving a weak regression function $g_k(I, \hat{S}^{(t)})$
   c) update:

$$f_k(I, \hat{S}^{(t)}) = f_{k-1}(I, \hat{S}^{(t)}) + v g_k(I, \hat{S}^{(t)}) \tag{4}$$

3) Output:

$$\gamma_t(I, \hat{S}^{(t)}) = f_K(I, \hat{S}^{(t)}) \tag{5}$$

From this training algorithm, we can get the feature residual γ_t.

B. Micro-expressions Dataset

At present, the public micro-expression data sets with research value are rich and diverse. As shown in Table I, the common micro-expression data sets are summarized. The duration of micro-expressions is milliseconds. The ordinary camera with low frame speed is hard to capture useful micro-expression information. The images in Micro-expression data sets are captured by high-speed cameras. As show in Table I, different data sets use cameras with different frame rates. For example, SMIC [7]capture more facial images using a camera with a frame rate of 100 frames per second. The more common CASME II [8]and Polikovsky [3]both use high-speed cameras with 200 frames per second. The data sets in Table I can be divided into two types according to whether they are classified or not. One is the non-induced micro-expression data set, such as Polikovsky, USF-HD [12], etc .; the other is the induced micro-expression data set, including SMIC, CASME II, etc.

TABLE I
EXISTING MICRO-EXPRESSION DATASET

dataset	Number of samples	frame rate	induction
SMIC	164	100	induced
CASME	195	60	induced
CASME II	255	200	induced
USF-HD	100	29.7	non-induced
Polikovsky	42	200	non-induced

1) USF-HD: USF-HD [12] is a Micro-expression data set established by the Shreve's team. A sample example is shown in Figure 1. USF-HD is an non-induced micro-expression dataset. This data set collects data under normal lighting conditions including both macro and micro expressions. For micro emoticons, the subjects are shown some sample videos containing micro emoticons before recording. They are asked to imitate these emotions randomly without out-of-plane head movements.

Fig. 1. USF-HD sample

2) CASME II: CASME II [8] is a high-acceptance induced micro-expression dataset. It is designed and improved by Fu's team of the Institute of Psychology of the Chinese Academy of Sciences. This dataset is base on the original micro-expression dataset CASME. Compared with CAME, CASME II uses a high-speed camera with a frame rate of 200 fps for data acquisition. It has higher resolution and obtains more detailed information about facial muscle movements. In addition, the collected sample data is about 280×340 in the face area Pixels. Therefore, CASME II has better spatial resolution and larger face size. An example of the CASME II dataset is shown in Figure 2.

Fig. 2. CASME II sample

III. PROPOSED METHOD

The preformance of machine learning and deep learning in facial feature extraction is excellent. It has the advantages of fast speed and high accuracy [11]. In the conditions with sufficient light, it can detect changes of local facial features accurately. In this paper, micro-expression recognition is modeled and analyzed. We use the location and recognition of feature points to achieve micro-expression detection and judgment. We construct a micro-expression recognition model. The structure of this model is shown in Figure 3.

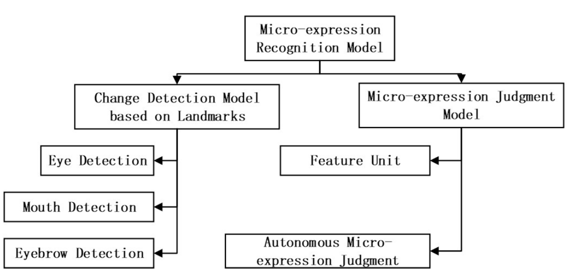

Fig. 3. The structure of micro-expression recognition model

The model consists of two parts:

- Construct a facial expression detection model based on 68 Landmarks. The identification of facial feature points is applied to micro-expression detection and recognition.
- Design an autonomous micro-expression model. We define the feature unit concept and apply it to recognize micro-expressions.

A. Change Detection Model based on Landmarks

The movement of facial muscles on the face causes the change of micro-expression. This model focuses on these muscular movement, detecting the characteristic changes of various local organs of the face, such as eyes, nose and mouth. We recognize each local part of the face, and use feature point labeling technology to detect and identify the changes of each organ. The difference of adjacent frames is applied to perform the detection of the short and small characteristics of micro-expressions. This method achieves the transformation from partical characteristics to the overall face expression.

1) Eye Detection: The method of eye changes detection mainly focus on the opening degree of eye and the frequency of blinking. The changes of micro-expressions are closely related to the opening degree of eyes. The frequency of blinking is also an important manifestation of facial emotion expression. Therefore, measuring the current state of the eyes accurately is crucial for micro-expression recognition and analysis. Karson [13]and Tsubota [14] proposed that the spontaneous resting blink rate of humans is almost 15 to 30 blinks per minute. If the frequency of the currently detected blink is lower or higher than this normal frequency, it means that the subject is expressing a particular emotion with certain micro-expression.

In order to detect eyes movement, we use eye aspect ratio. In the facial feature point annotation, each eye is represented by 6 coordinate points, called the feature points. These feature points revolve clockwise around the eye area from the left. When the state of eyes changes, the lateral length of eyes change slightly and the vertical length of eyes changes apparently. Based on the distribution characteristics of the eye feature points, Soukupov proposed a equation to discribe the relationship between the width and height of the eye in 2016. This equation is called Eye Aspect Ratio (EAR) [15].

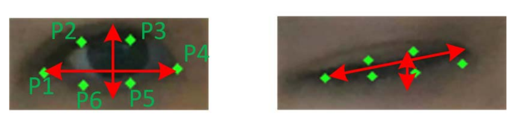

Fig. 4. Eye Aspect Ratio

As shown in Figure 4, according to the distribution of these six points, the ratio of the width and height of the eye is defined:

$$EAR = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||} \tag{6}$$

We find that the aspect ratio of the eyes is constant when the eyes open, and decreases rapidly when the eyes close. EAR can capture the fluctuation of surrounding feature points when certain facial expressions appear. So we use the EAR value to detect the changes of eyes.

In this model, EAR-SVM classifier is designed to accurately detect the opening states of the eyes. Support vector machine is a common classification method in machine learning. The basis of it is a hyperplane classifier, which find the largest positive and negative sample interval on the training sample feature space. In this model, EAR values of adjacent frames form the feature vector of the frame. A support vector machine multi-classification model is used to solve the three-classification problem of EAR values.

The training results of the model are shown in Figure 5. The blue dots in the figure are the EAR value samples of the squinting eyes. The white are the EAR value samples of the opening eyes, and the red are the normal EAR value samples. As can be seen from the above figure, EARSVM linearly divides the eyes into three categories: squint (blue area), open eyes (white area), and normal (red area) based on the distribution of EAR values. The accuracy of this classifier on the test set is about 95%.

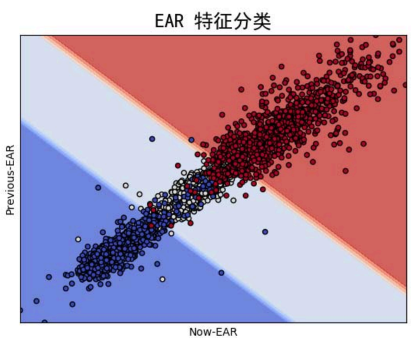

Fig. 5. EAR-SVM classifier

As the Figure 6 shows, when blinking, the EAR value in successive frames is approximately equal to 0. Therefore, when the fluctuation range of the EAR value in the detection video stream that continuously appears in 3 to 4 frames is 0 to 0.1, we think that a blinking action has occurred. In the eye detection method, the number of blinks per minute is counted in real time based on the EAR value. According to Karson et al. [13], it is normal to blink 15 to 30 times per minute, which is lower or higher than this frequency. Intervals mean a slight change in expression on the eyes.

2) Mouth Detection:

Open/Shut Mouth: In the Figure 6, there are 19 points related to the mouth. The point 63 and 67 are located below the upper lip and above the lower lip respectively. The displacement of point 63 and 67 is obvious when the mouth changes. Assuming the pixel coordinate value of point 63 is$H(x_h, y_h)$

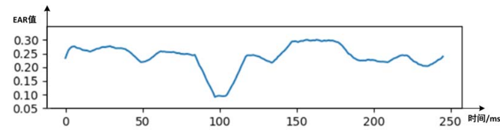

Fig. 6. EAR change curve

and the pixel coordinate value of point 67 is $L(x_l, y_l)$. The Euclidean distance D between these two points is

$$D = \sqrt{(x_h - x_l)^2 + (y_h - y_l)^2} \tag{7}$$

For each frame of image, calculate the Euclidean distance D between the two points H and L. If $D > 0$, the mouth is open, and if $D = 0$, the mouth is tightly closed. In this way, the non-zero nature of D is monitored to detect the mouth opening degree of the subject in real time.

Mouth corner: Because of the relative position stability of the nose feature point 31 among the facial feature points, we establish an angle model between the corner features of the mouth and the nose feature points. As shown in Figure 7, when the mouth corner is raised, the values of, $\angle NLR$ and $\angle NRL$ will decrease accordingly. When the mouth corner is lowered, the values of $\angle NLR$ and $\angle NRL$ will be larger.

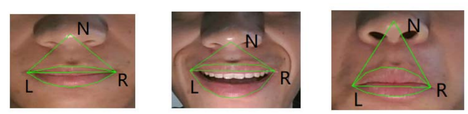

Fig. 7. Changes in mouth angle

According to the 1-minute prior video, We get the angle threshold of the mouth corner α_l and α_r.

$$\alpha_l = \frac{1}{N}\sum_{i}^{N}\alpha_{li} + b, \alpha_{li} = \arccos\frac{d_{NL}^2 + d_{LR}^2 - d_{NR}^2}{2 * d_{NL} * d_{LR}} \tag{8}$$

$$\alpha_r = \frac{1}{N}\sum_{i}^{N}\alpha_{ri} + b, \alpha_{ri} = \arccos\frac{d_{LR}^2 + d_{NR}^2 - d_{NL}^2}{2 * d_{RL} * d_{LR}} \tag{9}$$

Where N is the total number of frames. α_i represents the included angle of the mouth corner of each frame. b represents the offset of the model. And $b = 2°$,d is the distanced between each point of mouth.

3) Eyebrow Detection: When an expression change has happened in the face, there are two types of changes for the eyebrows: frowning or raising. For example, when a surprised facial micro-expression appears, the eyebrows will rise slightly. On the contrary, when in an anxious mood, eyebrows tend to frown.

As shown in Figure 8, five feature points are distributed on the left and right eyebrows respectively. These points will change their positions when the facial expression changes. The movement of point 22 and point 23 are the most obvious.

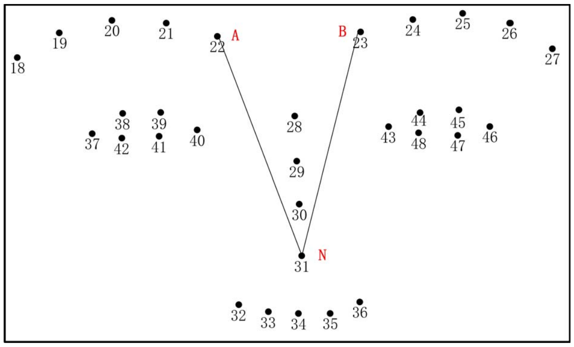

Fig. 8. Eyebrows' distribution

We take the stable feature point 31 of the nose bridge as the axis, detecting the state of the eyebrow by judging the change in the Euclidean distance D_{brow}, which is between the characteristic point of the eyebrow tip and the point 31. Ihe formula is:

$$D_{browL} = \sqrt{(x_a - x_n)^2 + (y_a - y_n)^2} \tag{10}$$

$$D_{browR} = \sqrt{(x_b - x_n)^2 + (y_b - y_n)^2} \tag{11}$$

For the detected faces in the video, we calculate D_{browL} and D_{browR} of each frame, comparing them with the average distance. A smaller value means that the eyebrow is in a tight state. And a larger value means that the eyebrow is raised. It should be noted that this distance is determined according to each person's situation, and is calculated with the prior data.

B. Micro-expression Judgment Model

1) Feature Unit: The appearance of a micro-expression is composed of one or more local areas movement of the face. In the previous section, we design a micro-expression detection model based on landmarks. This model detects the changes of the three facial organs : eyes, mouth, and eyebrows. As the definition of facial action units in FACS, the corresponding action state of facial organ, such as eyes, mouth and eyebrows, is defined as Feature Unit (FU). We can describe the changes of different facial organs, and judge different facial expressions by combining different FU. This paper defines eleven FUs, as shown in Table II.

E1-E4 describe the changes of Eye. M1-M4 represent the changes of Mouth. Eb1-Eb3 describes the changes of Eyebrow.

2) Autonomous Micro-expression Judgment: Micro-expressions are not independently. The appearance of them depend on the changes of each FU, and can be composed of different FUs. The micro-expression judgment model proposed in this section use FU. It detects and combines FUs with the model's physical scene. This model can make a highly autonomous judgment on the current micro-expression.

According to experience, we define several combinations of FU when basic micro-expressions occur. A reference case for identifying micro-expressions is shown in Table III. For these five basic micro-expressions: smile, surprise, anger, sadness and tension, all of them can be expressed and judged by the corresponding combination of FUs. There are more changes in micro-expressions than we show in the table. This model can achieve real-time detection of micro-expression with more conditions, based on the micro-expressions theory and researchers' knowledge.

TABLE II
DEFINITION OF FEATURE UNIT

organ	FU	Description
Eye	E1	Widen
	E2	Normal
	E3	Squint
	E4	Blinking too fast
Mouth	M1	Open widely
	M2	Shut up
	M3	Open slightly
	M4	Raise mouth corner
	M5	Press mouth corner
Eyebrow	Eb1	Frown Eyebrow
	Eb2	Raise Eyebrow
	Eb3	Normal

TABLE III
FU COMBINATION RECOGNIZE MICRO-EXPRESSION

Micro-expression	FU combination
Happy	E3+M1+M3
Surprise	M1+Eb1 or E1+Eb1
Anger	E3+M2+Eb2
Sadness	E3+M5
Tension	E4+M2

IV. EXPERIMENT

The CASME II dataset contains 247 micro-expression samples from 26 participants. Each video sample consists of a starting frame, a vertex frame and an ending frame. The micro-expression action starts from the starting frame, reaches the peak of the action amplitude at the vertex frame and completes in the end frame. The video before the start frame shows the participant's normal face. Firstly, the normal face information of 26 participants in the CASME II dataset is detected by using Change Detection Model. The information obtained includes EAR-SVM, mouth corner threshold and eyebrow threshold based on Landmarks. As shown in the Table IV, we train the EAR-SVM eye classifiers for each participant in the CASME II dataset, and calculate the corners of the mouth and eyebrow thresholds.

The CASME II dataset divides the participants' micro-expressions into five categories: disgust, happiness, depression, surprise, and others. And in the reference use cases of this model smile, different categories are defined as surprise, anger, sadness and tension. Our evaluation experiment takes the union of the two kinds of categories and identify three categories of happiness (smile), depression (sad), surprise (surprise). We calculates the recognition accuracy in CASME II data set.

Using different EAR-SVM classifiers and thresholds listed in Table IV, we detect and judge 42 happy samples of different participants in the dataset.

TABLE IV
PARTICIPANT THRESHOLD

Participant	Amount	EAR-SVM	Mouth corner	Eyebrow
sub02	13	Yes	70.3	115.4
sub05	19	Yes	58.5	109.8
sub09	15	Yes	60.0	118.6
sub10	14	Yes	59.4	112.8
sub11	10	Yes	54.5	124.1
sub12	12	Yes	58.7	126.1
sub17	36	Yes	76.5	122.1
sub19	16	Yes	62.1	122.0
sub20	11	Yes	58.7	103.8
sub23	12	Yes	68.7	85.9
sub24	11	Yes	62.1	106.3
sub26	17	Yes	72.3	116.7

For example, as shown in Figure 9, we use Change Detection Model to detect a sample of the sub17 participant. Her EAR-SVM classification result is squinting, mouth corner detection value is 70.1, and eyebrow detection value is 120.1. Compared with the threshold corresponding to sub17 in the threshold table, the participant is in the state of squinting eyes, mouth corners raised and eyebrows frowning. With the definition and combination of FUs we can make micro-expression judgments: participant is in a happy mood status.

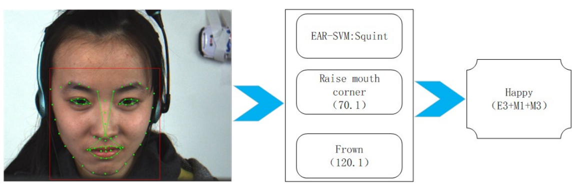

Fig. 9. Sample of sub17

This model performs micro-expression recognition on 42 happy samples, 42 depression samples and 40 surprised samples. The accuracy of each category is shown in Table VI. The recognition accuracy of this model on the CASME II dataset is 68.2%.

TABLE V
ACCURACY OF RECOGNIZING DIFFERENT EXPRESSIONS

Category	Happy	depression	Surprise
Amount	31	28	27
Accuracy	73.7%	66.7%	64.3%

There are some different recognition methods adopted by other micro-expression research teams. The accuracy comparision between them and our method are shown in Table VI. We use the CASME II dataset for this comparision. [7] adopte the LBP-TOP method to recognize micro-expressions. The recognition accuracy of this method on the CASME II dataset is 46.46%, which is about 20% less than the recognition rate of this model. [16] use EVM algorithm to process image features, and use SVM to perform micro-expression classification. The accuracy of this algorithm on the CASME II dataset is 67.2%, which is similar to the recognition accuracy of our model, but we has higher recognition-autonomy. And our model can be applied to recognition tasks in different scenarios.

TABLE VI
ACCURACY COMPARISION OF DIFFERENT RECOGNITION MODEL

Model	LBP-TOP	EVM+SVM	Our method
Accuracy	46.46%	67.2%	68.2%

V. CONCLUSION

This paper constructs a micro-expression recognition model based on feature units. It combines an automatic local detection model with a highly autonomous micro-expression judgment model. We first find the facial feature points and establishes the change model of the three characteristic organs of eyes, mouth and eyebrows. Then we define the concept of feature activity unit of eye, mouth and eyebrow. We realize the change measurement of feature activity unit finally.

The combination of feature units and specific physical scenes are used to realize the autonomous judgment of micro-expressions. The model is evaluated on the micro-expression dataset CASME II and compares with other algorithm. The experiments show that our model performs best. It has both high accuracy and high autonomy.

REFERENCES

[1] P EKMAN, FRIESEN W V. Constants across cultures in the face and emotion[J]. J Pers Soc Psychol, 1971, 17(2) : 124 129.

[2] EKMAN P. Telling lies: Clues to deceit in the marketplace, politics, and marriage (revised edition): WW Norton & Company, 2009.

[3] POLIKOVSKY S, KAMEDA Y, OHTA Y. Facial micro-expressions recognition using high speed camera and 3D-gradient descriptor[J], 2009 : 16 16(1).

[4] POLIKOVSKY S, KAMEDA Y, OHTA Y. Facial micro-expressions recognition using high speed camera and 3D-gradient descriptor[J], 2009 : 16 16(1).

[5] DALAL N, TRIGGS B. Histograms of Oriented Gradients for Human Detection[C] // International Conference on Computer Vision & Pattern Recognition (CVPR 05) : Vol 1. [S.l.] : IEEE Computer Society, 2005 : 886 893.

[6] WANG S-J, YAN W-J, ZHAO G, et al. Micro-expression recognition using robust principal component analysis and local spatiotemporal directional features[C] // European Conference on Computer Vision. 2014 : 325 338.

[7] PFISTER T, LI X, ZHAO G, et al. Recognising spontaneous facial microexpressions[C] // 2011 international conference on computer vision. 2011 : 1449 1456.

[8] YAN W-J, LI X, WANG S-J, et al. CASME II: An improved spontaneous micro-expression database and the baseline evaluation[J]. PloS one, 2014, 9(1) :1 8.

[9] GUO Y, XUE C, WANG Y, et al. Micro-expression recognition based on CBPTOP feature with ELM[J]. Optik - International Journal for Light and Electron Optics, 2015, 126(23) : 4446 4451.

[10] HUANG G B, ZHOU H, DING X, et al. Extreme learning machine for regression and multiclass classifcation[J]. IEEE Transactions on Systems Man & Cybernetics Part B, 2012, 42(2) : 513 529.

[11] KAZEMI V, SULLIVAN J. One millisecond face alignment with an ensemble of regression trees[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2014 : 1867 1874.

[12] SHREVE M, GODAVARTHY S, GOLDGOF D, et al. Macro-and microexpression spotting in long videos using spatio-temporal strain[C] // Face and Gesture 2011. 2011 : 51 56.

[13] KARSON C N. Spontaneous eye-blink rates and dopaminergic systems[J]. Brain, 1983, 106(3) : 643 653.

[14] TSUBOTA K. Tear dynamics and dry eye[J]. Progress in retinal and eye research, 1998, 17(4) : 565 596.

[15] SOUKUPOVA T, CECH J. Eye blink detection using facial landmarks[C] // 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia. 2016.
[16] LI X, HONG X, MOILANEN A, et al. Reading hidden emotions: spontaneous micro-expression spotting and recognition[J]. arXiv preprint arXiv:1511.00423,2015, 12 : 32 33.