A New Framework Combining Local-Region Division and Feature Selection for Micro-Expressions Recognition

YANLIANG ZHANG[ID]1, **HANXIAO JIANG**[ID]1, **XINGWANG LI**[ID]1, **(Senior Member, IEEE),**
BING LU[ID]1, **KHALED M. RABIE**[ID]2, **(Member, IEEE),**
AND ATEEQ UR REHMAN[ID]3, **(Member, IEEE)**

[1]School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China
[2]Department of Engineering, Manchester Metropolitan University, Manchester M1 5GD, U.K.
[3]Department of Computer Science, Abdul Wali Khan University, Mardan 23200, Pakistan

Corresponding author: Xingwang Li (lixingwangbupt@gmail.com)

ABSTRACT Micro-expressions are deliberate or unconscious movements of people's psychological activities, reflecting the transient facial true expressions. Previous works focus on the whole face for micro-expressions recognition. These methods can extract a number of feature vectors which are relevant or irrelevant to the micro-expressions recognition. Besides, the high-dimension feature vectors can result in longer computational time and increased computational complexity. In order to address these problems, we propose a new framework which combines the local-region division and the feature selection. Based on the proposed framework, the original images can retain more efficient regions and filter out the invalid components of feature vectors. Specifically, with the joint efforts of the facial deformation identification model and facial action coding system, the global region is divided into seven local regions with their corresponding actions units. The ReliefF algorithm is used to select effective components of feature vectors and reduce the dimension. To evaluate the proposed framework, we conduct experiments on both the Chinese Academy of Sciences Micro-expression II Database and Spontaneous Micro-expression Database with Leave-One-Subject-Out Cross Validation method. The results show that the performance in local combined regions outperforms its counterpart in the global region, and the recognition accuracy is further improved with the combination of feature selection.

INDEX TERMS Micro-expression recognition, local region division, feature selection, ReliefF algorithm.

I. INTRODUCTION

Language is not the only way of human communication. In contrast to verbal communication, nonverbal communication plays a very important role in interpersonal communication and nurse-patient relationship accounting for 65% of all forms of communication. There are several types of nonverbal communication that can convey a person's true sentiment, thought and personality such as facial expressions, attitudes, interpersonal styles [1], etc. Facial expressions are intuitive reflection of a person's state. Currently, facial expressions can be mainly classified into macro-expressions and micro-expressions (MEs). In the early years, the automatic analysis of facial expressions focused on distinguishing the MEs and the macro-expressions [2], [3]. Macro-expressions usually last for about 2 to 5 seconds, and they are distributed throughout the face generally. However, MEs are brief, spontaneous facial movements and usually occur when a person tries to conceal the inner emotions [4], [5]. Studies showed that MEs only last one-twenty-fifth to one-third second with slight intensities of involved muscle movements [6]–[8]. Micro-expression recognition (MER) is considered as an important clue in the field of national security, medical care, psychological diagnosis, and investigative interrogation [9]–[13], etc., because it can detect the real emotions beneath the false surface.

The associate editor coordinating the review of this manuscript and approving it for publication was Venkateshkumar M[ID].

Although Ekman and his team developed Micro Expression Training Tool (METT) [14] to train people to distinguish the categories of synthetic MEs, it is not completely applicable to spontaneous expressions.

Recently, the application of computer vision in MER has becoming very popular. Pfister *et al*. applied a temporal interpolation model together with the first comprehensive spontaneous micro-expression corpus to the field of MER, achieving the first success in recognizing spontaneous facial micro-expressions [15]. The proposed framework mainly consists of five subsections, namely, face alignment, motion magnification, Temporal Interpolation Model (TIM), feature extraction, and classification [16]. In addition, several previous researches have been proposed for MER. Shreve *et al*. put forward a new approach for the automatic temporal segmentation of facial expressions in long videos which can detect and distinguish between large expressions (macro) and localized micro-expressions in [3]. Huang *et al*. proposed Spatial-Temporal performed Local Quantization Patterns (STCLQP) [17], the expansion of Completed Local Quantized Pattern (CLQP) [18] to spatial-temporal space, realizing the progress in the analysis of facial MEs. STCLQP utilized three useful difference of pixels, including sign-based, magnitude-based and orientation-based, to obtain the compact and distinctive codebook for micro-expression analysis. In [19], Wang *et al*. proposed an identification algorithm based on Discriminant Tensor Subspace Analysis (DTSA) and extreme learning machine, and extended DTSA to a high-order tensor to process micro-expression video fragments. Furthermore, Local Spatial-temporal Direction Features [20] was developed for analyzing robust principal component of micro-expressions, but it was not improved. Then, the same team proposed a novel color space model which can translate the fourth dimensional RGB containing color information into tensor independent color space for greater accuracy [21]. The aforementioned works have laid a good foundation for the latter studies on facial expression recognition.

However, to date, there are several parts which could be used to enhance the performance of micro-expression recognition. Firstly, MEs are not distributed uniformly across the whole face [22]–[25], and they occur in a combination of several facial regions. For example, the muscular movements of the cheek lift and mouth upturn show a happy emotion, while the raised eyebrows and slightly parted lips indicate that someone feels surprised; Secondly, when using different feature descriptors to extract the block features, the dimension of the feature vectors will increase exponentially with the increase of parameters. For the high-dimensional feature vector containing many components which are unrelated to the occurrence of MEs, it has a high computational cost when performing micro-expression classification. In this paper, we propose an extended framework based on the case proposed by Pfister *et al*. [15]. We divide the local regions after operating the TIM and performing feature extraction on the local regions. To address the computational complexity of the

feature vector, the feature vector with high information content is identified by feature selection for classifying the MEs. Therefore, a new method for automatically recognizing MEs is proposed. The contributions to this article are as follows:

- We propose a facial regions-of-interest (RoIs) location technique which depends on an automatic discriminant method, namely, the discriminative facial deformable model method [26], to divide the facial key-area blocks. The divided local-region blocks completely cover all Facial Action Coding System FACS) Action Units (AUs) corresponding to each micro-emotion category, and eliminate the areas which are irrelevant to MEs, effectively avoiding the influence of the irrelevant areas on MER.
- ReliefF algorithm is applied to carry out the feature selection with three spatial-temporal local texture descriptors. The combination of local regions and the feature selection [27] can adequately pick the feature subsets with good distinguishing characteristics. Feature selection can reduce the dimension of feature vectors and increase the speed of processing operation. At the same time, it can retain the feature components with high information content for the classification of MEs, which is the key to improve MER.
- Intensive experiments using three feature descriptors on the two aforementioned datasets show that for MER, the combination of local regions and feature selection performs better than the operations separately. Therefore, using a better combination of facial regions has a great influence on improving the recognition accuracy of MEs.

The remainder of the paper is organized as follows. Section II reviews the databases, the MER in global region and the spatial-temporal local texture descriptors which are used in the experiments. In Section III, we present our proposed framework for MER. Afterwards, we introduce the region division and feature selection in Sections IV and V. We design two comparisons for evaluating this framework in Section VI. Eventually, we draw the conclusion of our study in Section VII.

II. RELATED WORK

A. DATABASE

So far, MEs datasets have been established by some teams, but most of the MEs samples in the datasets are artificially controlled and non-spontaneous. Studies have shown that when people produce natural MEs, they themselves even do not realize that they leak their true feelings involuntarily. What's more, there are essential differences between the posed and the authentic spontaneous micro-expressions, which proved the defect of the posed MEs in practical applications. Therefore, in order to obtain more realistic and natural spontaneous data, the Spontaneous Micro-Expression (SMIC) database and the Chinese Academy of Science Micro-Expression II (CASME II)

TABLE 1. Composition of the SMIC database.

Database	Subjects	Positive	Negative	Surprise	Total
HS	16/20	51	70	43	164
VIS	8/10	28	23	20	71
NIR	8/10	28	23	20	71

TABLE 2. Composition of the CASME II database.

MEs	Happy	Disgust	Surprise	Repression	Other
MEs sequence	32	64	25	27	99
MEs number	2360	4159	1605	2187	6344

database [28], an upgraded version of the CASME, have been collected.

1) SMIC

The SMIC database is the world's first public spontaneous micro-expression database designed by a team in the University of Oulu in Finland in 2012. They carefully chose video clips to induce subjects' emotional responses. To simulate high stake situations, subjects were told that there would be a punishment if they are found any emotion in their faces.

The team collected the real emotions, including *happy, sad, disgust, fear, and surprise*. In SMIC database, they are divided into three main categories, positive (happy), negative (sad, disgust and fear), and surprise. The collection is done by cameras with different frame rates, respectively, a high speed (HS) camera of 100fps, a normal visual camera (VIS), and a near-infrared (NIR) with 25 fps. Similar experiments were conducted under VIS and NIR cameras for obtaining another two sessions of SMIC with the last eight subjects. Table 1 demonstrates the specific composition of the SMIC database.

2) CASME II

The CASME II database was established by the team of Fu Xiaolan in the Institute of psychology, Chinese academy of sciences. It was an extension of the CASME database, with higher temporal resolution (200 fps) and spatial resolution (about 280×340 pixels on facial area) in a well-controlled experimental environment. They selected 247 MEs from nearly 3000 sequences for the database with corresponding AUs and labeled emotions. Similarly, the label of CASME II database has an identical judgment-criteria with the CASME database. According to this, there are five micro-expression labels in the dataset, namely, *happy, disgust, surprise, repression, and other*. The specific number of sequence and pictures for each type of expressions is shown in the following Table 2.

B. SPATIAL-TEMPORAL LOCAL TEXTURE DESCRIPTORS

Dynamic texture analysis [29] in the field of spatial-temporal space can provide information about the dynamic process of facial appearance and movement, which all have an important impact on the recognition performance. Some descriptions about the Local Binary Pattern on Three Orthogonal Plans
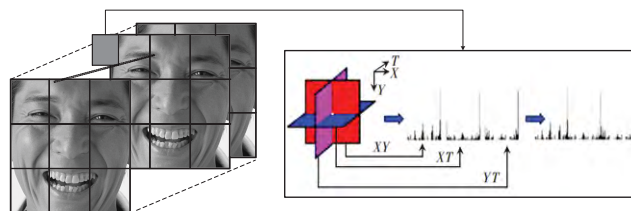

FIGURE 1. Features in each block by using LBP-TOP [30].

(LBP-TOP), the Histograms of Oriented Gradients on Three Orthogonal Plans (HOG-TOP), and the Histograms of Image Oriented Gradients on Three Orthogonal Plans (HIGO-TOP) are demonstrated in the following.

1) LBP-TOP

In the field of facial expression recognition, texture features are commonly used to reflect the distribution properties of pixel space. The Local Binary Patterns (LBP) [30], [31], proposed by the University of Oulu in Finland, is a simple, efficient, and typical texture feature extraction method which can be applied in many fields. Subsequently, Zhao and Pietikäinen put forward the LBP-TOP [29], [32], which is an extension of LBP descriptor in three-dimensional space and a measurement of dynamic image texture. The LBP-TOP operator sets three spatial-temporal axes for the dynamic image sequence, named T, X, and Y. Specifically, the XY plane contains texture information for each frame of image, the XT and YT planes include the changes of the MEs sequence in the spatial position over time. First, we can obtain the LBP values on XY, XT, and YT by using the LBP-TOP operator. Then, we form the final histogram by connecting the local LBP histograms in three orthogonal planes (XY, XT, YT) in series. The concatenated histogram represents the feature vector of LBP-TOP, including the information of the appearance and motion of the dynamic texture. Fig. 1 shows the diagram of extracting the LBP-TOP feature vector.

2) HOG-TOP

Histograms of Oriented Gradients (HOG) was proposed by Dalal and Triggs [33] in 2005. The HOG descriptor obtains the feature vectors based on the specification of image edges. Even in the case of lacking accurate information of the image direction gradient and edge position, the HOG descriptor can characterize the local appearance and shape them by calculating the gradient of image and the amplitude of gradient in different directions [34].

Firstly, we calculate the gradient amplitude and directions in the horizontal and vertical direction of the image to capture the expressions' contour information. Provided the pixel (x, y), we can obtain the gradient components G_x and G_y by the convolution of the gradient operator K and the original image in the respective directions of x and y, where $K = [-1, 0, 1]^T$. For each point of the image, its gradient amplitude G and gradient direction α of the pixel

are computed as follows:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}, \qquad (1)$$

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right). \qquad (2)$$

Secondly, we need to divide the image into cells with the same size and count the histogram of oriented gradient of each cell. Each cell unit is combined into large, spatially connected blocks. Therefore, we can calculate the histogram features of the oriented gradient on the whole block and normalize the features to reduce the influence of background color and noise.

Finally, we can construct the HOG feature by concatenating the block-based gradient histograms. Moreover, the HOG-TOP [34] operator acts on the dynamic video sequences and its algorithmic process is the same as the HOG.

3) HIGO-TOP

HIGO-TOP is a simplified feature descriptor of the HOG-TOP. This operator ignores the size of the first derivative to reduce the effects of illumination and contrast. Similarly, the HIGO-TOP [17] operator can apply to the dynamic video sequences and its operation is the same as the HIGO.

C. RECOGNIZING MICRO-EXPRESSION IN GLOBAL REGION

When calculating the description over the entire facial expression sequence, the descriptors can only encode the appearance of micro-patterns and ignore their specific positions [35]. For solving the effect, Guoying Zhao and Matti Pietikäinen proposed a new facial representation in [29]. Taking the ordinary facial expression as an example, the regularized facial image can be divided into overlapping or non-overlapping blocks. Figs. 2 and 3 are static images in which blocks of area are separated. Thereinto, Fig. 2 describes the non-overlapping 7 × 5 blocks and the overlapping 4 × 3 blocks are illustrated in Fig. 3. Extracting the feature vectors of the spatial-temporal pattern on each region block and connecting them in series, we can obtain the feature vector of the dynamic video sequences from three different directions.

The method based on the global region can extract the motion features from the facial blocks. However, there are some differences between MEs and the conventional expressions. MEs often occur in some local areas, such as eyes, nose, and mouth [23], [24]. Furthermore, there are two shortcomings in extracting feature vectors from global regions. On one hand, there will be an effect on MER because of extracting the unrelated features in global region. On the other hand, the same organ is repeatedly divided owing to the irregular distribution of the main parts, resulting an incomplete extraction of motion features.
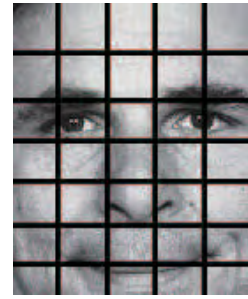

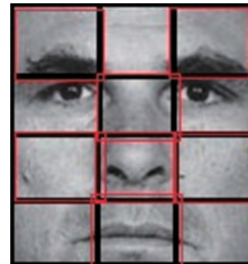**FIGURE 2.** Non-overlapping blocks (7 × 5) for facial segmentation.

FIGURE 3. Overlapping blocks (4 × 3) for facial segmentation.

FIGURE 4. Framework diagram for the proposed ME recognition method.

III. PROPOSED FRAMEWORK OF THE MEs RECOGNITION

In this paper, we propose a new framework for improving the accuracy of MER. A specific MER process of our method is shown in Fig. 4.

Firstly, according to a large number of observations, there will also be some changes in facial motions which are unrelated to the MEs. Therefore, dividing the original image into local areas of the face (i.e., eyes, mouth, and so on) and removing the unrelated facial regions are the key to study the internal structure information of the local module deeply. Then, for comparing the usability and superiority clearly, we perform the feature extraction and recognize the regions of interest (RoIs) based on three feature descriptors (i.e., LBP-TOP, HOG-TOP, and HIGO-TOP) in CASME II and SMIC databases.

Next, we use the ReliefF algorithm for feature selection. Feature selection is a process for selecting the most effective features. Therefore, the combination of feature extraction and feature selection can improve the model accuracy and reduce the running time.

There are three advantages of the proposed framework. Firstly, RoIs can exclude the unnecessary and retain the useful parts. Secondly, different combinations of the local regions are beneficial for finding a better information distribution combination. Thirdly, reducing the dimension of feature vectors can effectively shorten the running time and improve the recognition accuracy.
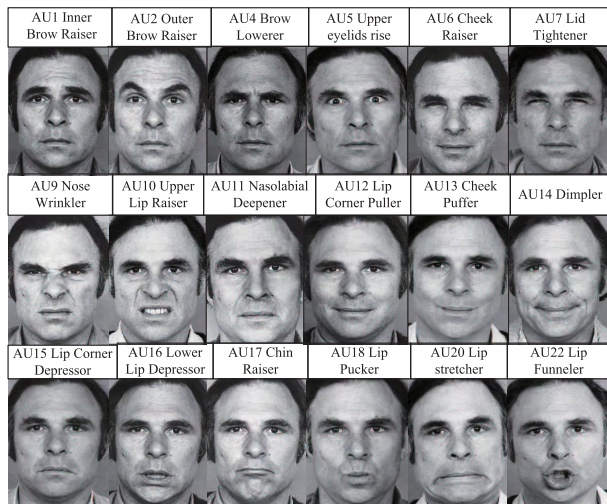
FIGURE 5. Facial expressions of some AUs.

FIGURE 6. Decomposition of Happiness with AUs.

IV. REGION DIVISION

The global method of MEs treats every divided region uniformly, while, the presence of unrelated local regions will reduce the accuracy of recognition. Hence, we need to seek out the involved facial movement regions when MEs occur and use them as the division standard. Extracting the feature vectors of spatial-temporal pattern can eliminate the redundant facial information effectively.

A. FACIAL ACTION UNITS

In our daily life, our facial expressions are varied whether it is a regular expression or a voluntary micro-expression. However, there will be something in common when the same type of expressions appear. American psychologists Ekman and Friesen developed and revised the FACS in [36]–[38] for identifying facial expressions objectively. This system assigns the muscle actions to the face, of which thirty-two are named AUs and another fourteen are Action Descriptors (ADs) [39]. Each AU can describe the specific movement of a single local muscle of the face. Facial expressions are often composed by the synergy of several specific AUs.

The excitement of the muscles can arouse many AUs, which make up different emotions. In addition, an AU can also exist in several different expressions. For example, a person may be thinking when the facial muscles drive the AU4 movement. Likewise, AU4 also appears when someone feels anger, anxiety, pain or ache. Fig. 5 exemplifies some specific AUs. For example, AU1 indicates the inner eyebrows raiser and AU5 shows the rising upper eyelids. Fig. 6 presents the detailed decomposition of AUs with the example of "Happiness". We can see that a facial expression can activate one or several AUs. For instance, there are two situations when generating a surprised expression. The first one is that people realized their surprise and concealed this emotion subconsciously when the AU5 fleeting on the face. The second is that people pose a surprised expression deliberately. In this situation, AU5 is mostly appearing with the combination of
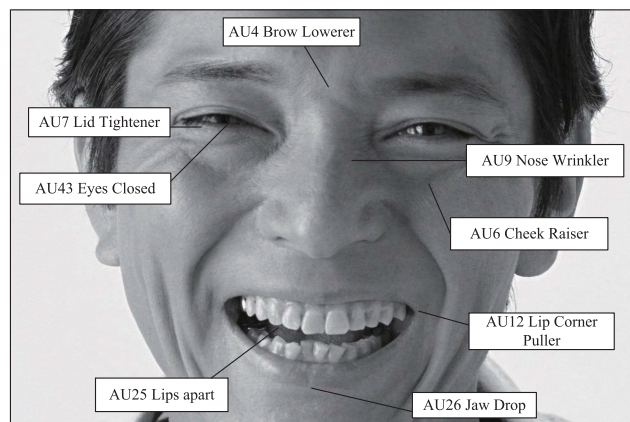
TABLE 3. The corresponding AUs involved in the conventional expressions [39].

Expression category	FACS AU
Happiness	AU6, AU12, AU25
Disgust	AU9, AU10, AU16, AU17, AU25, AU26
Surprise	AU1, AU2, AU5, AU26, AU27
Sadness	AU1, AU4, AU6, AU11, AU15, AU17
Fear	AU1, AU2, AU4, AU5, AU20, AU25, AU26, AU27
Anger	AU4, AU5, AU7, AU17, AU22 to AU26

TABLE 4. The corresponding relationship between MEs and corresponding FACS AUs [28].

MEs category	FACS AU
Happiness	AU6 or AU12
Disgust	AU9 or AU10 or AU4, AU7
Surprise	AU1, AU2, AU25 or AU2
Repression	AU15, AU17 alone or in combination
Other	AU4 or AU14 or AU17

AUs (1, 2, 5, 25, 26) or AU27. Moreover, AUs (4, 7) indicate the confused or concentrated emotion state, while AUs (5, 7) represent slightly fear.

B. CORRESPONDENCE BETWEEN MEs AND AUs

Researchers have done a lot of studies on common expressions. Table 3 shows several conventional expressions with their corresponding AUs. However, the MEs are short and slight muscle movements in a special high-stake situation. It is different with the conventional expression in the relationship with AUs. For example, when a "surprise" expression appears on the face, it corresponds to AUs (1, 2, 5, 26) in the conventional expression. But it corresponds to AUs (1, 2), AU25 or AU2 in MEs. Therefore, the correspondence between the MEs and the related AUs have been summed up from the CASMEII database by the Fu Xiaolan group of the Chinese Academy of Sciences Psychology in Table 4.

C. BLOCK-BASED FACIAL DIVISION

In Fig. 7, we use the facial deformation identification model method [28], which is different from the classical Active Appearance Model (AAM) [40], to automatically identify
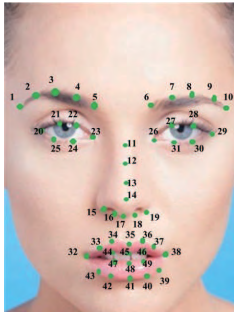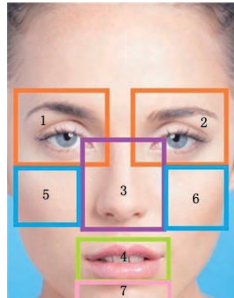
FIGURE 7. Location of the 49 facial key points.

FIGURE 8. Distribution of the divided local blocks.

TABLE 5. Coordinate Points of the boundary in each local region.

Rectangular frame	Left border	Right border	Upper border	Lower border
Left eye	1	5	3	24
Right eye	6	10	8	31
Nose	22	27	11	Midpoint between 17 and 45
Mouth	32	38	34	41
Left cheek	1	4	24	15
Right cheek	7	10	31	19
Chin	32	38	41	Lower boundary of image

the 49 key points of the face. This algorithm can realize the incremental learning to the facial model for a higher accuracy.

According to the division method of the facial region in [41], the FACS and a method of automatically locating the landmark points are used to divide the RoIs of the face. Fig. 8 demonstrates the distribution of the divided local blocks (i.e., "eye", "nose", "mouth", "cheek", and "chin"). Studying the MEs on the divided seven local areas can further eliminate the redundant facial information and retain more meaningful facial details. The coordinate points corresponding to the boundaries of the seven region blocks demonstrated in Table 5 and Table 6 explain the corresponding AUs in each local area. Comparing Table 4 and Table 6, the local-region blocks divided in this paper mostly cover all AUs, avoiding the effect of other irrelevant AUs.

V. FEATURE SELECTION
A. DIMENSION STRAIN
The feature dimension is increased due to the precise parameter setting after the feature extraction, which can easily lead to a "dimensional disaster". The excessive feature vector dimension not only increases the computational complexity and the running time, it also reduces the classification

TABLE 6. The divided facial regions and corresponding FACS AUs.

Region number	Regional part	FACS AUs
1,2	Eyes, Eyebrows	AU1, AU2, AU4, AU7
3	Nose	AU9
4	Mouth	AU10, AU12, AU14, AU15, AU25
5,6	Cheek	AU6
7	Chin	AU17

accuracy. So, it is useful to reduce the dimension of the extracted feature vector for having a great efficient feature subset. At the same time, the selected features are more discriminative to understanding the model and analyzing the data. Therefore, feature selection [42] is well applied to select the most effective features for reducing the dimension of the feature space. We can see that some proposed feature-selection methods are shown in [43]–[46].

B. RELIEFF ALGORITHM
The Relief, a supervised feature selection algorithm, was first proposed by Kira and Rendell in [47]. It judges the correlation between features and categories based on the feature distinction ability of the nearest neighbor samples, and assigns the weight value for each feature. When the weight value is greater than the predetermined threshold, the feature is retained; otherwise, it will be rejected. ReliefF algorithm is an improved extension method based on the Relief algorithm by Kononeko [48]. This algorithm can overcome the problem of limitation on two-classifications. In addition, it can also solve both the problem of incomplete data and regression problem. ReliefF algorithm makes a great contribution to the multi-label learning when comparing with the most of the multi-label feature selection methods [49].

Similar to the Relief algorithm, we input the training set $X = (x_i, x_i, \cdots, x_N)$ and $x_i \in R^m$. We set m as the number of features for each sample, the number of nearest neighbor samples is K, the number of iterations is n, and the feature weight threshold is δ. First, we need to initially set the preselected feature subset S to an empty set, and denote the weight values of all features as $w(A_f) = 0$, where $A_f(f = 1, 2, \cdots, m)$ represents the fth feature of the sample.

Second, the scope of the iterative operation is selected from 1 to n. ReliefF randomly obtains a sample R from the training sample set. Unlike the Relief algorithm, the ReliefF can find K of the nearest neighbors from the same class, named nearest hits $H_j(j = 1, 2, \cdots, K)$. Also, the K of nearest neighbors from the different classes with sample R is named near misses $M_j(C)(j = 1, 2, \cdots, K)$, where C is a different category from R [46]. Then, the update calculation for the weight values for each feature is shown as follows:

$$w(A_f) = w(A_f) - \frac{\sum_{j=1}^{K} diff(A_f, R, H_j)}{nK} + \frac{\sum_{C \notin Class(R)} \left[\frac{P(C) \sum_{j=1}^{K} diff(A_f, R, M_j(C))}{1 - P(Class(R))} \right]}{nK}. \quad (3)$$

Note that *Class*(*R*) represents the category of sample *R*, and *P*(*C*) expresses the ratio of the number of samples of the category *C* to the total number of samples. Furthermore, the function *diff* (*A_f*, *R*, *B*) for calculating the discrimination of sample *R* and *B* in the *f* th feature *A_f* is defined as:

$$diff\,(A_f, R, B) = \begin{cases} 0, & R(A_f) = B(A_f) \\ 1, & R(A_f) \neq B(A_f) \end{cases}, \quad (4)$$

where $R(A_f)$ and $B(A_f)$ represent the value of the *f* th feature in sample instances *R* and *B*, respectively.

ReliefF algorithm performs mainly by comparing the feature weight values $w(A_f)$ and thresholds δ. If the feature weight value $w(A_f)$ is greater than the threshold δ, this feature will be added into the preselected feature subset *S*. So, it is beneficial for improving the recognition accuracy by using the new-building feature subset.

VI. EXPERIMENTS

In this section, we report some results of MER experiments on the databases of SMIC and CASMEII by using three feature extraction methods, LBP-TOP, HOG-TOP, and HIGO-TOP. From these results, we will make further analysis of the effect on different combination of regions and their improvement with feature selection.

A. EXPERIMENTAL SETTING

In the experiments, Leave-One-Subject-Out Cross Validation (LOSO-CV) [51] protocol with Chi square (chi2) kernel [52] of Support Vector Machine (SVM) [53] is used as the classifier for the MER.

For the LOSO-CV, take the CASME II as an example, it collected the MEs video sequences from 26 subjects. Each time the micro-expression video sequence of 25 subjects is gathered as a training set for machine learning, the remaining subject will be used as a test. Then, the classifier SVM can build a model based on the training set. Applying this model to the target dataset can obtain the recognition accuracy. The classification accuracy (Acc) is the ratio of the number of correctly classified samples to the total number of samples (247) in the experiment.

B. COMPARISON OF MER BETWEEN GLOBAL AND LOCAL REGION

The mouth and eyes ("1, 2, 4") are the main parts in expressions, so we choose them as the basic regions. Based on the basic region ("1, 2, 4") and the distribution of the divided facial local areas, we can form different region combinations by assembling the basic region and other regions.

1) PARAMETERS SETTING

According to the division code in Fig. 7, we mainly designe five local region combinations for the comparison with global region. In this sub-experiment, we set (*P*, *R*) at different values with (4, 1), (8, 2), and (8, 3). When extracting the feature vectors in each local area, the original image will be divided according to (α, β, τ), where α, β, and τ are the numbers

TABLE 7. Combination of different feature extraction regions.

Combination number	Selected regions
A	Global region
B	1, 2, 4
C	1, 2, 4, 3
D	1, 2, 4, 3, 7
E	1, 2, 4, 3, 5, 6
F	1, 2, 4, 3, 5, 6, 7

divided along the axis of X, Y, and T. For this sub-experiment, we set the number $\alpha = \beta$ referring to the setting in [16]. We vary the value of the blocking area (α, β, τ) to test the effect of classification recognition under different local area combination. Experiments are performed on CASME II by using LBP-TOP as feature descriptor. The concrete combination of the feature extraction areas and experimental results are shown in Table 7 and Table 8. The best accuracy in each combination areas are highlighted in bold in Table 8. We take the average accuracy for the comparison in Table 8.

2) DISCUSSION AND CONCLUSION

We can see that the average accuracy of the region combinations B, C, D, E, and F all performed better than that in the global region A. The increased accuracy values in the combinations are 4.05%, 3.24%, 0.49%, 5.43%, and 3.32%, respectively. It proves that the global region contains some irrelevant information for MER.

Table 8 shows that the best recognition performance of region combination C performs better than the combination B with the increased region 3. But that does not mean that increasing regions can improve the identification accuracy. For example, the best recognition performance in region combination D has decreased with the increased region 7. It suggests that the Acc of the MEs will also be reduced when extracting the features with little relevance.

C. COMPARISON OF FEATURE SELECTION IN LOCAL REGIONS

In this experiment, we compare the Acc before and after feature selection with the three spatial-temporal feature descriptors (LBP-TOP, HOG-TOP, HIGO-TOP) on two databases (CASME II and SMIC). From Table 8, we can see that region combination E has the best performance in average accuracy. To testify the validity of the framework, we conduct the following experiments under the region E. The ReliefF [54] algorithm is selected for feature selection.

1) PARAMETERS SETTING

For the LBP-TOP feature vectors, we design the same settings of (*P*, *R*) and (α, β, τ) on CASME II and SMIC databases. Similarly, we set the number of bins as 4 or 8 for the HOG-TOP and HIGO-TOP descriptors on both databases. Training and testing are all done by using Chi-Square SVM with LOSO-CV. The MER results on CASME II and SMIC databases are shown in Tables 9-14. Each table records

TABLE 8. Performance of different region combinations by using LBP-TOP descriptor on CASME II.

Parameter		A	B	C	D	E	F
(P, R)	(α, β, τ)	$Acc(\%)$	$Acc(\%)$	$Acc(\%)$	$Acc(\%)$	$Acc(\%)$	$Acc(\%)$
	$(3, 3, 2)$	44.53%	46.15%	49.80%	43.32%	49.39%	46.96%
$(4, 1)$	$(3, 3, 4)$	44.94%	46.15%	44.53%	44.53%	46.96%	44.94%
	$(5, 5, 2)$	43.32%	46.96%	42.92%	42.51%	48.99%	48.58%
$(8, 2)$	$(3, 3, 2)$	43.72%	**51.42%**	52.23%	**46.96%**	**52.23%**	51.42%
$(8, 3)$	$(3, 3, 2)$	**45.34%**	51.42%	48.58%	46.96%	51.42%	**49.39%**
Average		44.37%	48.42%	47.61%	44.86%	**49.80%**	47.69%

TABLE 9. MER results by using LBP-TOP descriptor and ReliefF algorithm on CASME II database.

Parameter		Non-selected Feature			Selected Feature			
(P, R)	(α, β, τ)	D	t	$Acc(\%)$	D	t	$Acc(\%)$	Differential
	$(3, 3, 1)$	2430	3.8258	42.51%	20	0.0359	51.82%	9.31% ↑
$(4, 1)$	$(3, 3, 2)$	4860	7.3622	49.39%	30	0.0557	53.04%	3.65% ↑
	$(3, 3, 4)$	9720	14.6234	46.96%	150	0.0996	**59.51%**	**12.55%** ↑
	$(5, 5, 2)$	13500	20.5423	47.37%	250	0.13617	53.04%	5.67% ↑
$(8, 2)$	$(3, 3, 2)$	19112	28.8231	52.23%	580	0.8497	56.28%	4.05% ↑
	$(5, 5, 3)$	79650	117.5342	48.18%	900	1.6149	57.09%	8.91% ↑
$(8, 3)$	$(3, 3, 2)$	19116	28.9987	**53.44%**	200	0.1124	55.87%	2.43% ↑

TABLE 10. MER results by using HOG-TOP descriptor and ReliefF algorithm on CASME II database.

Parameter		Non-selected Feature			Selected Feature			
Number	(α, β, τ)	D	t	$Acc(\%)$	D	t	$Acc(\%)$	Differential
	$(3, 3, 2)$	1296	2.0732	49.39%	10	0.0303	52.63%	3.24% ↑
$b = 4$	$(4, 4, 1)$	1152	1.7926	44.53%	80	0.7254	45.34%	0.81% ↑
	$(5, 5, 1)$	1800	2.8289	41.30%	140	0.1014	44.53%	3.23% ↑
	$(4, 4, 1)$	2304	3.6318	43.32%	430	0.2750	48.18%	4.86% ↑
$b = 8$	$(5, 5, 1)$	3600	5.6855	44.13%	140	0.1124	**55.47%**	**11.34%** ↑
	$(6, 6, 1)$	5184	8.3434	52.63%	450	0.3459	55.06%	2.43% ↑
	$(6, 6, 2)$	10368	16.3827	**53.04%**	340	0.2451	53.44%	0.40% ↑

TABLE 11. MER results by using HIGO-TOP descriptor and ReliefF algorithm on CASME II database.

Parameter		Non-selected Feature			Selected Feature			
Number	(α, β, τ)	D	t	$Acc(\%)$	D	t	$Acc(\%)$	Differential
	$(3, 3, 2)$	1296	2.2299	48.18%	120	0.0955	**58.30%**	10.12% ↑
$b = 4$	$(4, 4, 1)$	1152	1.9005	40.49%	10	0.0314	51.82%	**11.33%** ↑
	$(5, 5, 1)$	1800	3.0064	40.49%	100	0.0800	51.01%	10.52% ↑
	$(4, 4, 1)$	1296	3.9446	42.11%	130	0.0938	46.56%	4.45% ↑
$b = 8$	$(5, 5, 1)$	2592	6.0247	45.75%	130	0.0978	48.18%	2.43% ↑
	$(6, 6, 1)$	2304	8.6003	47.37%	260	0.1486	51.42%	4.05% ↑
	$(6, 6, 2)$	3600	16.8289	**48.58%**	220	0.1310	53.04%	4.46% ↑

TABLE 12. MER results by using LBP-TOP descriptor and ReliefF algorithm on SMIC database.

Parameter		Non-selected Feature			Selected Feature			
(P, R)	(α, β, τ)	D	t	$Acc(\%)$	D	t	$Acc(\%)$	Differential
$(4, 1)$	$(3, 3, 1)$	2430	1.5658	36.59%	540	0.1681	45.12%	8.53% ↑
	$(3, 3, 2)$	4860	3.1147	46.34%	310	0.1054	57.93%	11.59% ↑
$(4, 2)$	$(3, 3, 1)$	2430	1.5494	39.02%	520	0.1605	54.27%	15.25% ↑
	$(3, 3, 2)$	4860	3.1321	41.46%	500	0.1524	53.05%	11.59% ↑
$(8, 1)$	$(3, 3, 1)$	9558	6.2605	40.85%	300	0.0772	56.10%	15.25% ↑
	$(3, 3, 2)$	19116	12.6503	45.12%	500	0.1139	58.54%	13.42% ↑
$(8, 2)$	$(3, 3, 1)$	9558	6.4364	42.07%	500	0.1134	57.93%	**15.85%** ↑
	$(3, 3, 2)$	19116	12.7512	**47.56%**	500	0.1280	**59.76%**	12.20% ↑

the dimension (*D*), the average time (*t*) of classification, the recognition accuracy (*Acc*) and the differentials of the selected and non-selected feature. The best accuracy values in each part are highlighted in bold. The arrow in the table intuitively reflects the tendency of the increased accuracy after feature selection.

2) DISCUSSION AND KEY FINDINGS

For the proposed framework, we used the ReliefF algorithm to operate in the combination E. Tables 9-11 and Tables 12-14 respectively demonstrate the comparison between the selected and non-selected features on both MEs databases.

TABLE 13. MER results by using HOG-TOP descriptor and ReliefF algorithm on SMIC database.

Parameter		Non-selected Feature			Selected Feature			
Number	(α, β, τ)	D	t	$Acc(\%)$	D	t	$Acc(\%)$	Differential
	$(3, 3, 2)$	1296	0.9794	45.73%	190	0.0710	54.27%	8.54% ↑
$b = 4$	$(4, 4, 1)$	1152	0.8672	32.32%	90	0.0527	38.42%	6.10% ↑
	$(4, 4, 2)$	2304	1.7166	46.34%	40	0.0345	51.83%	5.49% ↑
	$(3, 3, 1)$	1296	0.9317	42.07%	20	0.0257	48.78%	6.71% ↑
$b = 8$	$(3, 3, 2)$	2592	1.9233	**48.17%**	200	0.0754	**56.10%**	7.93% ↑
	$(4, 4, 1)$	2304	1.6907	32.02%	20	0.0262	56.10%	**17.08%** ↑
	$(4, 4, 2)$	4608	3.3916	46.95%	220	0.0923	54.27%	7.32% ↑

TABLE 14. MER results by using HIGO-TOP descriptor and ReliefF algorithm on SMIC database.

Parameter		Non-selected Feature			Selected Feature			
Number	(α, β, τ)	D	t	$Acc(\%)$	D	t	$Acc(\%)$	Differential
	$(3, 3, 2)$	1296	0.9860	39.63%	20	0.0317	52.44%	12.81% ↑
$b = 4$	$(4, 4, 1)$	1152	0.8909	34.76%	30	0.0310	42.68%	7.92% ↑
	$(4, 4, 2)$	2304	1.7441	**45.73%**	30	0.0324	52.44%	6.71% ↑
	$(3, 3, 1)$	1296	0.9652	34.76%	270	0.0928	49.39%	14.63% ↑
$b = 8$	$(3, 3, 2)$	2592	1.9588	45.73%	20	0.0264	53.05%	7.32% ↑
	$(4, 4, 1)$	2304	1.8706	36.59%	70	0.0498	52.44%	**15.85%** ↑
	$(4, 4, 2)$	4608	3.5171	41.46%	70	0.0498	**56.10%**	14.64% ↑

Firstly, with the different parameter settings, dimension of the best MER performance in these experiments are not exactly fixed. Due to varying the parameters (α, β, τ), we can divide the image into various blocks with corresponding dimensions. For most of the results, the number of the selected best dimensions will be increased with the increasing of original dimensions.

Secondly, from the time comparison between the selected and non-selected features in Tables 9-14, it can be found that the time after feature selection can reduce at least one hundred times. For example, running the MER with the parameter settings (4, 1) and (3, 3, 1) in Table 9, the average time spent in non-selected feature recognition is 3.8258s, while it only takes 0.0359s in selected feature recognition. Because of the reduction in dimension, the running speed of data is faster and the time for extracting features is also shorten up.

Thirdly, the combination E contains the effective regions for feature extraction. In addition, the ReliefF can select the subset with a high-weight value. From the results of LBP-TOP, HOG-TOP, and HIGO-TOP shown in Tables 9-14, we can find that these selected feature accuracies are all superior to the non-selected features. The largest differentials under three spatial-temporal feature descriptors are 12.55%, 11.34% and 11.33% in CASME II (15.85%, 17.08% and 15.85% in SMIC). Besides, the average values of the differential in CASME II are 6.65%, 3.76% and 6.77% (12.96%, 8.45% and 11.41% in SMIC). Even though the differentials perform less obvious in CASME II, the proposed framework still achieves contented results.

VII. CONCLUSION

In this paper, we mainly propose a novel framework combining the region division and feature selection method ReliefF for automatic MER. The FACS and the facial deformation identification model method are applied to locate the facial local areas. With the assistance of the forty-nine facial key points and the distribution of thirty-two AUs, the global region can be divided into seven regions. Then, the LBP-TOP, HOG-TOP, and HIGO-TOP descriptors are adopted to extract features from MEs respectively. Ultimately, the high-dimensional feature vectors will gain an obvious dimension reduction and higher recognition accuracy by using the novel framework.

There are two important conclusions drawn from the two groups of experiments. Firstly, the best feature subset can be obtained after the feature selection in local areas, which excludes the influence of massive redundant information on classification. Secondly, the combination of local regions can make the most of the internal structure information and significantly improve the MER accuracy when comparing with the global region.

Overall, there are still some challenges and limitations that we need to overcome in the future. The first challenge is that there is still a long way to identify the MER accurately on the basic of current methods. When developing the MER into different fields, we also need to further investigate and work out more creative methods on different databases. In addition, new spontaneous ME datasets are still in demand, which are expected to collect with a large amount of samples and more detailed emotion categories. It is a long-term and arduous task to use computers to automatic discriminate MEs more accurately.

REFERENCES

[1] P. Ekman and W. V. Friesen, ''Nonverbal leakage and clues to deception,'' *Psychiatry*, vol. 32, pp. 88–106, Mar. 1969.

[2] M. Shreve, S. Godavarthy, V. Manohar, D. Goldgof, and S. Sarkar, ''Towards macro-and micro-expression spotting in video using strain patterns,'' in *Proc. WACS*, Dec. 2009, pp. 1–6.

[3] M. Shreve, S. Godavarthy, D. Goldgof, and S. Sarkar, ''Macro-and micro-expression spotting in long videos using spatio-temporal strain,'' in *Proc. Int. Conf. Autom. Face Gesture Recognit*, Mar. 2011, pp. 51–56.

[4] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *J. Personality Social Psychol.*, vol. 17, no. 2, pp. 124–129, 1971.

[5] S. Porter and L. T. Brinke, "Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions," *J. Physiol. Sci.*, vol. 19, no. 5, pp. 508–514, May 2008.

[6] W.-J. Yan, Q. Wu, J. Liang, Y.-H. Chen, and X. Fu, "How fast are the leaked facial expressions: The duration of micro-expressions," *J. Nonverbal Behav.*, vol. 37, no. 4, pp. 217–230, Jul. 2013.

[7] D. Matsumoto and H. S. Hwang, "Evidence for training the ability to read microexpressions of emotion," *Motivat. Emotion*, vol. 35, no. 2, pp. 181–191, Apr. 2011.

[8] X. Li, J. Li, and L. Li, "Performance analysis of impaired SWIPT NOMA relaying networks over imperfect Weibull channels," *IEEE Syst. J.*, vol. 14, no. 1, pp. 669–672, 2020.

[9] Q. Wu, X.-B. Shen, and X.-L. Fu, "Micro-expression and its applications," *Adv. Psychol. Sci.*, vol. 18, no. 9, pp. 1359–1368, 2010.

[10] M. O'Sullivan, M. G. Frank, C. M. Hurley, and J. Tiwana, "Police lie detection accuracy: The effect of lie scenario," *Law Human Behavior*, vol. 33, no. 6, pp. 530–538, Nov. 2009.

[11] J. J. Weiss, "Ekman, P. (2009) telling lies: Clues to deceit in the marketplace, politics, and marriage. New York: Norton," *Amer. J. Clin. Hypn.*, vol. 53, no. 4, pp. 287–288, Oct. 2011.

[12] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 680–695, Jan. 2020.

[13] X. Li, M. Liu, C. Deng, P. T. Mathiopoulos, Z. Ding, and Y. Liu, "Full-duplex cooperative NOMA relaying systems with I/Q imbalance and imperfect SIC," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 17–20, Jan. 2020.

[14] P. Ekman, *MicroExpression Training Tool (METT)*. San Francisco, CA, USA: Univ. California, 2002.

[15] T. Pfister, X. Li, G. Zhao, and M. Pietikäinen, "Recognising spontaneous facial micro-expressions," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2012, pp. 1449–1456.

[16] X. Li, X. Hong, A. Moilanen, X. Huang, T. Pfister, G. Zhao, and M. Pietikäinen, "Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods," *IEEE Trans. Affect. Comput.*, vol. 9, no. 4, pp. 563–577, Oct. 2018.

[17] X. Huang, G. Zhao, X. Hong, W. Zheng, and M. Pietikäinen, "Spontaneous facial micro-expression analysis using spatiotemporal completed local quantized patterns," *Neurocomputing*, vol. 175, pp. 564–578, Jan. 2016.

[18] X. Huang, G. Zhao, X. Hong, M. Pietikainen, and W. Zheng, "Texture description with completed local quantized patterns," in *Proc. Scand. Conf. Image Anal. (SCIA)*, 2013, pp. 1–10.

[19] S.-J. Wang, H.-L. Chen, W.-J. Yan, Y.-H. Chen, and X. Fu, "Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine," *Neural Process. Lett.*, vol. 39, no. 1, pp. 25–43, Feb. 2014.

[20] S. Wang, W. Yan, G. Zhao, X. Fu, and G. Zhou, "Micro-expression recognition using robust principal component analysis and local spatiotemporal directional features," in *Proc. Eur. Conf. Comput. Vis.*, vol. 8925, Mar. 2015, pp. 325–338.

[21] S.-J. Wang, W.-J. Yan, X. Li, G. Zhao, and X. Fu, "Micro-expression recognition using dynamic textures on tensor independent color space," in *Proc. 22nd Int. Conf. Pattern Recognit.*, Aug. 2014, pp. 4678–4683.

[22] X. Duan, Q. Dai, X. Wang, Y. Wang, and Z. Hua, "Recognizing spontaneous micro-expression from eye region," *Neurocomputing*, vol. 217, pp. 27–36, Dec. 2016.

[23] W.-J. Yan, S.-J. Wang, Y.-J. Liu, Q. Wu, and X. Fu, "For micro-expression recognition: Database and suggestions," *Neurocomputing*, vol. 136, pp. 82–87, Jul. 2014.

[24] S.-J. Wang, W.-J. Yan, X. Li, G. Zhao, C.-G. Zhou, X. Fu, M. Yang, and J. Tao, "Micro-expression recognition using color spaces," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6034–6047, Dec. 2015.

[25] X. Li, Q. Wang, H. Peng, H. Zhang, D.-T. Do, K. M. Rabie, R. Kharel, and C. C. Cavalcante, "A unified framework for HS-UAV NOMA networks: Performance analysis and location optimization," *IEEE Access*, vol. 8, pp. 13329–13340, 2020.

[26] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Incremental face alignment in the wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1859–1866.

[27] N. Kwak and C.-H. Choi, "Input feature selection for classification problems," *IEEE Trans. Neural Netw.*, vol. 13, no. 1, pp. 143–159, 2002.

[28] W. Yan, X. Li, S. Wang, G. Zhao, Y. Liu, Y. Chen, and X. Fu, "CASME II: An improved spontaneous micro-expression database and the baseline evaluation," *PLoS ONE*, vol. 9, no. 1, pp. 1–8, Jan. 2014.

[29] G. Zhao and M. Pietikäinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, Jun. 2007.

[30] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, Jan. 1996.

[31] T. Ojala, M. Pietikäinen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[32] G. Zhao and M. Pietikäinen, "Boosted multi-resolution spatiotemporal descriptors for facial expression recognition," *Pattern Recognit. Lett.*, vol. 30, no. 12, pp. 1117–1127, Sep. 2009.

[33] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 886–893.

[34] M. R. Guedira, A. E. Qadi, M. R. Lrit, and M. E. Hassouni, "A novel method for image categorization based on histogram oriented gradient and support vector machine," in *Proc. Int. Conf. Electr. Inf. Technol. (ICEIT)*, Nov. 2017, pp. 1–5.

[35] P. Yang, Q. Liu, and D. N. Metaxas, "Boosting encoded dynamic features for facial expression recognition," *Pattern Recognit. Lett.*, vol. 30, no. 2, pp. 132–139, Jan. 2009.

[36] P. Ekman and W. Friesen, "Facial action coding system (FACS): A technique for the measurement of facial action," *Riv. Psichiatr.*, vol. 47, pp. 126–138, Jan. 1978.

[37] P. Ekman and W. Friesen, Eds., *Facial Action Coding System*. Salt Lake City, UT, USA: Consulting Psychologists Press, 2002. [Online]. Available: http://www.cs.cmu.edu/afs/cs/project/face/www/facs.htm

[38] J. F. Cohn, Z. Ambadar, and P. Ekman, "Observer-based measurement of facial expression with the facial action coding system," *Neurosci. Lett.*, vol. 394, no. 3, pp. 203–221, 2007.

[39] B. Martinez, M. Valstar, B. Jiang, and M. Pantic, "Automatic analysis of facial actions: A survey," *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 325–347, Jul./Sep. 2019.

[40] T. Cootes, G. Edwards, and C. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.

[41] S. Liong, J. See, C. Phan, K. Wong, and S. Tan, "Hybrid facial regions extraction for micro-expression recognition system," *J. Signal Process Syst.*, vol. 90, no. 4, pp. 601–607, Aug. 2017.

[42] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Jan. 2003.

[43] K. Kira and L. Rendell, "The feature selection problem: Traditional methods and a new algorithm," in *Proc. 10th Natl. Conf. Artif. Intell.*, Jul. 1992, pp. 129–134.

[44] G. John, R. Pfleger, and K. Pfleger, "Irrelevant features and the subset selection problem," in *Proc. 11th Int. Conf.*, Jul. 1994, pp. 121–129.

[45] L. Yu and H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," in *Proc. 20th Int. Conf. Mach. Learn. (ICML)*, Aug. 2003, pp. 856–863.

[46] S. Shreem, S. Abdullah, and M. Z. A. Nazri, "Hybridizing relief, mRMR filters and GA wrapper approaches for gene selection," *J. Theor. Appl. Inform. Technol.*, vol. 47, no. 3, pp. 1034–1039, Jan. 2013.

[47] K. Kira and L. A. Rendell, "A Practical Approach to Feature Selection," in *Proc. 9th Int. Workshop Mach. Learn. (ML)*, Dec. 1992, pp. 249–256.

[48] I. Kononenko, "Estimating attributes: Analysis and extensions of RELIEF," in *Proc. Mach. Learn. (ECML)*, Apr. 1994, pp. 171–182.

[49] N. Spolaor, E. A. Cherman, M. C. Monard, and H. D. Lee, "ReliefF for multi-label feature selection," in *Proc. Brazilian Conf. Intell. Syst.*, Oct. 2013, pp. 6–11.

[50] M. Wu and Y. Wang, "A feature selection algorithm of music genre classification based on ReliefF and SFS," in *Proc. IEEE/ACIS 14th Int. Conf. Comput. Inf. Sci. (ICIS)*, Jun. 2015, pp. 539–544.

[51] H. Cheng, D. J. Garrick, and R. L. Fernando, "Efficient strategies for leave-one-out cross validation for genomic best linear unbiased prediction," *J. Animal Sci. Biotechnol.*, vol. 8, no. 1, p. 38, May 2017.

[52] S. Bahassine, A. Madani, and M. Kissi, "An improved Chi-sqaure feature selection for arabic text classification using decision tree," in *Proc. 11th Int. Conf. Intell. Syst., Theories Appl. (SITA)*, Oct. 2016, pp. 1–5.

[53] C. Junli and J. Licheng, "Classification mechanism of support vector machines," in *Proc. 5th Int. Conf. Signal Process.*, Aug. 2000, pp. 1556–1559.

[54] A. A. Karizaki and M. Tavassoli, "A novel hybrid feature selection based on ReliefF and binary dragonfly for high dimensional datasets," in *Proc. 9th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, Oct. 2019, pp. 300–304.

YANLIANG ZHANG received the B.Sc. degree from Henan University, in 2001, and the Ph.D. degree from the School of Electronic Engineering, Xidian University, in 2011. He is currently an Associate Professor with the School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo, China. His research interests include machine vision and affective computing.

HANXIAO JIANG received the B.Sc. degree in electronic information engineering from the School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo, China, in 2018, where she is currently pursuing the M.Sc. degree in communication and information systems. Her current research interests include machine learning, image recognition, and transfer learning.

XINGWANG LI (Senior Member, IEEE) received the B.Sc. degree from Henan Polytechnic University, in 2007, the M.Sc. degree from the University of Electronic Science and Technology of China, in 2010, and the Ph.D. degree from the Beijing University of Posts and Telecommunications, in 2015.

From 2010 to 2012, he was worked as an Engineer with Comba Telecom Ltd., Guangzhou, China. From 2017 to 2018, he was a Visiting Scholar with Queen's University Belfast, Belfast, U.K. He is currently an Associate Professor with the School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo, China. His research interests include MIMO communication, cooperative communication, hardware constrained communication, non-orthogonal multiple access, physical layer security, unmanned aerial vehicles, and the Internet of Things. He has served as many TPC/Co-Chair, such as IEEE Globecom, IEEE/CIC ICCC, IEEE WCNC, IEEE VTC, and IEEE/IET CSNDSP. He is also an Editor on the Editorial Board of IEEE Access, *Computer Communications*, *Physical Communication*, and *KSII Transactions on Internet and Information Systems*. He is also a Lead Guest Editor of the special issue on *Recent Advances in Physical Layer Technologies for 5G Enabled Internet of Things of Wireless Communications and Mobile Computing*.

BING LU received the B.Sc. degree from the Zhongyuan University of Technology, in 2016, and the M.Sc. degree in communication and information systems from the School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo, China, in 2019. Her research interests include machine learning, artificial intelligence, and image recognition.

KHALED M. RABIE (Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical and electronic engineering from The University of Manchester, in 2011 and 2015, respectively. He is currently an Assistant Professor with the Department of Engineering, Manchester Metropolitan University (MMU), U.K. He has worked as a part of several large–scale industrial projects. He has published over 100 journal articles and conference papers (mostly IEEE). His primary research interest includes various aspects of the next–generation wireless communication systems. He is also a Fellow of the U.K. Higher Education Academy (FHEA). He serves regularly on the Technical Program Committee (TPC) of several major IEEE conferences, such as GLOBECOM, ICC, and VTC. He has received numerous awards over the past few years in recognition of his research contributions, including the Best Student Paper Award at the IEEE ISPLC (TX, USA, 2015), the MMU Outstanding Knowledge Exchange Project Award of 2016, as well as IEEE ACCESS Editor of the month award for August 2019. He serves as an Editor for the IEEE Communications Letters, an Associate Editor for IEEE Access, an Area Editor for *Physical Communications*, and an Executive Editor for *Transactions on Emerging Telecommunications Technologies*.

ATEEQ UR REHMAN (Member, IEEE) received the Ph.D. degree from the University of Southampton, in 2017. He worked with the Next Generation Wireless Research Group, University of Southampton, where he focused reliable data transmission in cognitive radio networks. He is currently working as a Lecturer with the Department of Computer Science, Abdul Wali Khan University, Mardan, Pakistan. His main research interests are in next–generation wireless communications, cognitive radio networks, the Internet of Things, the Internet of Vehicles, blockchain technology, and differential privacy. He was a recipient of the several academic awards, such as the Faculty Development Program, Islamic University of Technology (OIC) Dhaka, Bangladesh Distinction Award, and Higher Education Commission Pakistan OIC scholarship for undergraduate studies.

● ● ●