International Journal of Scientific & Engineering Research, Volume 10, Issue 12,December 2019
ISSN 2229-5518

AN INNOVATIVE TECHNIQUE FOR CONTRAST ENHANCEMENT OF IMAGES USING NORMALIZED GAMMA-CORRECTED CONTRAST-LIMITED ADAPTIVE HISTOGRAM EQUALIZATION FOR FACIAL EMOTION RECOGNITION

NIZVIN R MOHAMED[1] , DR PRIYA S [2], MRS JAYSREE K[3]

[1] Mtech Student in Image Processing, Model Engineering College, Thrikkakara, Kochi, India
[2]Professor, Dept. of Computer Engineering, Model Engineering College, Thrikkakara, Kochi, India
[3] Professor , Dept of Computer Engineering, Model Engineering College, Thrikkakara, Kochi,India

ABSTRACT—Facial Emotion Recognition is an emerging field in research. Emotion recognition system discussed here is based on the movements of muscles beneath the face called action units. But a proper pre-processing is required to attain a high accuracy rate to classify emotions. A well engineered Convolution Neural Network is proposed in this paper which recognises emotions through action units(AUs). The various pre-processing stages discussed includes:face-detection,gray-scale conversion,cropping,resizing,normalization using adaptive histogram equalization and filtering done using a non local means filter. The dataset used is Extended Cohn–Kanade Dataset (CK+).Experiment results show a high accuracy and recognition rate.

INDEX TERMS— Emotion Recognition,FaceDetection,Rgb-gray scaleConversion,cropping resizing,CLAHE,Non local means Filtering, FeatureExtraction,Deep Learning,Convolution Network

1. INTRODUCTION

Facial expression is a major non-verbal means of expecting intentions in human communication.[1]

Emotion recognition system has applications in wide areas of image processing such as robotics(human-computer interactions. Emotion recognition is used in E-learning to analyze students approach to the faculty and changing the methodology. In monitoring the drivers behavior to ensure safety in driving. In medicine,determining the psychological behavior/emotions of patients helps in giving the best treatment.[2,3,4,5,6,7,8,9].

Various studies have been conducted to recognize emotions. But a facial emotion recognizing system is still a challenging task. The problems of lightning, illumination challenges, posing problems hinders the working of a emotion recognition system. So to overcome these problems a better preprocessing methods are used.

2. RELATED WORKS

Facial expression recognition techniques are based on appearance features or geometry features. Geometric features are extracted from the shape of the face and its components such as the eyebrows, the mouth, the nose etc. Appearance features are extracted using the texture of the face.eg:wrinkles in the face. The expressions are determined from static images or dynamic image sequences.

Ekman and Friesen invented Facial Action Coding System(FACS) to classify facial expressions using facial action units(AUs).Basic emotions include happy,sad,angry,surprise,neural ,disgust and fear emotions.[10]

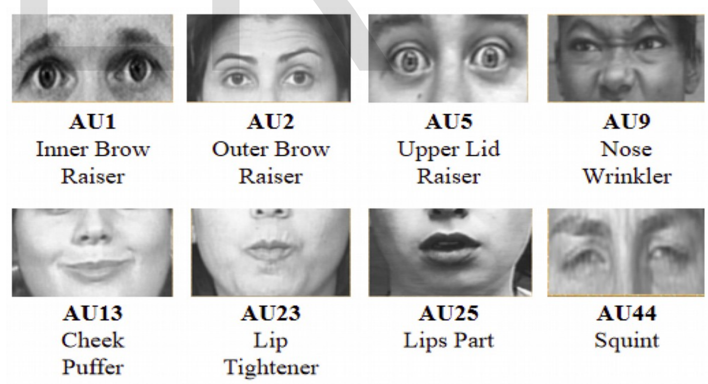

Fig 2.1 :Facial Action Coding System to categorize emotions

A convention facial emotion recognition system is shown below.The conventional FER procedure can be divided into three major steps: image preprocessing, feature extraction, and expression classification,

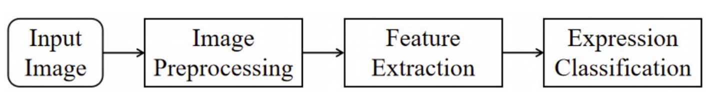

Fig 2.2 : Conventional FER System

Image preprocessing : Various image processing steps such as face detection (Viola Jonas),cropping resizing, normalization,histogram equalization and

filtering are performed. Image Preprocessing is basically done to images to remove backgrounds, to reduce noise and to elliminate light and illlumination problems.
Noise reduction/filtering uses various filters such as average Filter (AF), Gaussian Filter (GF), Median Filter (MF), Adaptive Median Filter (AMF), and Bilateral Filter (BF).
Face detection deals with localizing and extracting the face region.[11,12]
Normalization is to bring about uniformity in color and size of images for easy computation.[13,14,15]
Histogram equalisation is applied to enhance the image .

Feature Extraction: It is the process to extract useful data or information from the image, e.g., values, vectors, and symbols.
Gabor feature extraction, which uses gabor filters.[16]
Local Binary Pattern (LBP), features are represented by binary representations. [17]
Optical flow method, which uses motion vector changes.[19]
Haar-like feature extraction, feature point tracking, etc. are some of the feature extraction techniques
Another feature techniques are based on Active Appearance Model and Active Shape Model(ASM). [18]
PCA (principal component analysis) and **LDA**(linear discriminant analysis) is a dimension reduction technique for feature extraction in machine learning tasks.
In feature tracking, landmark key points are extracted. **Scale Invariant Feature Transform(SIFT**) is an example.[20]

Expression Classification:

The commonly used and widely applied classifier in FER systems include kNN (k-Nearest Neighbours) [21,22,23], SVM (Support Vector Machine)[25], Adaboost (Adaptive Boosting), Bayesian, SRC (Sparse Representation-based Classifier), and PNN (Probabilistic Neural Network).

KNN Network requires time consuming trainining alogorithm.

SVM is useful when the dataset contains less number of samples.
Adaboost is sensitive to noise but reduces overfitting problems.
Bayesian classification requires linear parameters.

The feature extraction and classification have to be designed manually and separately, which means these two phases cannot be optimised simultaneously.
Deep learning-based approaches highly reduce the reliance on image preprocessing and feature extraction and are more robust to the environments with different elements, e.g., illumination and occlusion, which means that they can greatly outperform the conventional approaches. In addition, it has potential capability to handle high volume data.

Deep convolution neural network is shown below

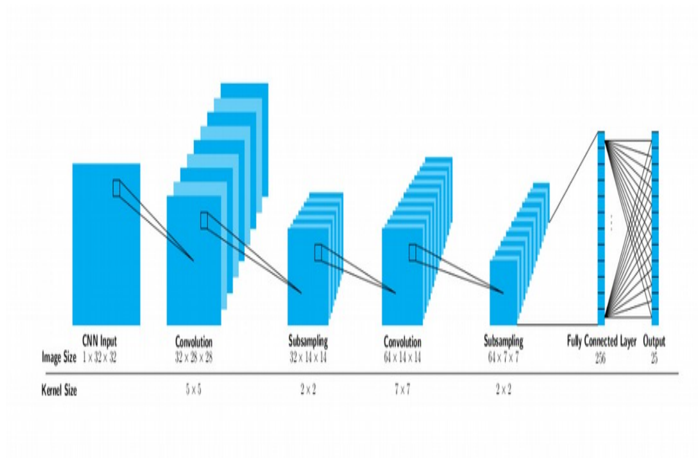

Fig 2.3 :Deep convolution Network

CNN[26] includes six components: Convolutional-layer, Sub-sampling layers, Rectified linear unit (ReLU), Fully connected layer, Output layer and Softmax layer.

Layer 0: Input layer- The network inputs are grayscale images with of size 32 by 32.

Layer 1:Convolutional layer- calculates the output of all neurons that are associated to local regions in the input layer,each calculating a dot product among their weights and a small region they are associated to .Filter used is 32.

Layer 2: RELU layer(Rectified linear unit)- will apply on elementwise activation function, such as the max (0, x).Subsampling(Max pooling) is performed. By applying maximum function in the Max-Pooling, input value is reduced.

Layer 3:Convolution with 64 filter.Relu performed.

Layer 4:Sub Sampling performed.

Layer 5:Fully connected layer with 256 hidden units are placed.

Layer 6:Softmax layer with 25 neurons which indicate numbers of AUs and fully connected to prior layer is considered in order to classification of AUs.

3. PROPOSED METHOD

In the proposed method, in order to classify seven basic emotion states, a psychological framework which is called Facial Action Coding System (FACS) is used for increasing accuracy of recognizing system, and facial movements by AUs are used for classifier output to determine final emotion state by combination of Aus(Auction Units).

Emotion	Action Unites
Anger	4+5+7+23
Contempt	R12A+R14A
Disgust	9+15+16
Fear	1+2+4+5+7+20+26
Happiness	6+12
Sadness	1+4+15
Surprise	1+2+5B+26

Fig 3: Action Units

A overview of the system is shown below

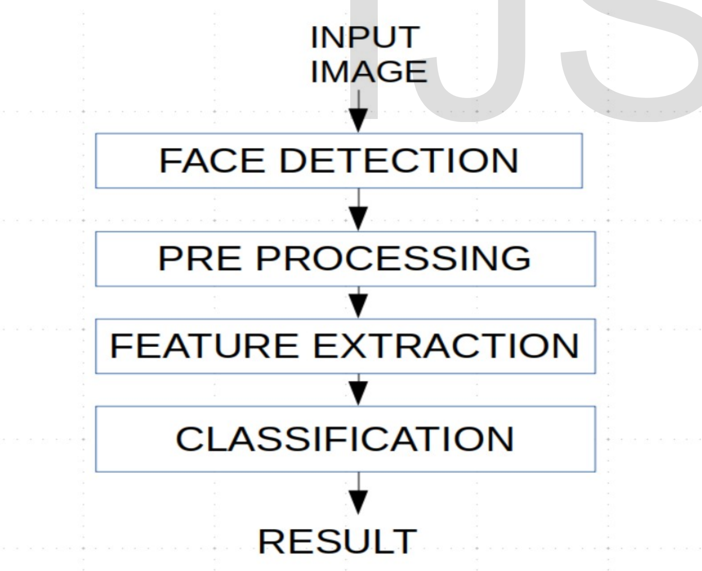

Fig 3.1: System Design

The overview of the system includes the following steps: Face Detection, pre processing, Feature Extraction,Classification.

4. IMPLEMENTATION

The input image is the original image. Face detection is performed and is followed by pre processing , feature extraction and classification stages.

Fig 4: Original Image

4.1 Face Detection

Face detection is done using viola jonas algorithm. The algorithm has four stages:

1. Haar Feature Selection
2. Creating an Integral Image
3. Adaboost Training
4. Cascading Classifiers

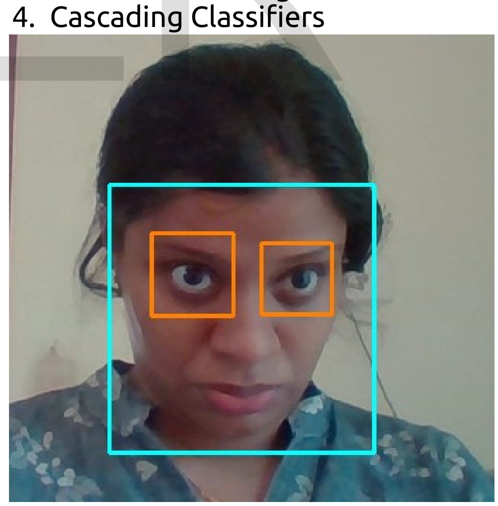

Fig 4.1: Face and Eyes Detection

4.2 Pre processing

The pre processing consists of the following steps:
1) RGB-GRAY scale conversion'
2)Cropping non face area
3)Resizing
4)Normalization
5)Filtering

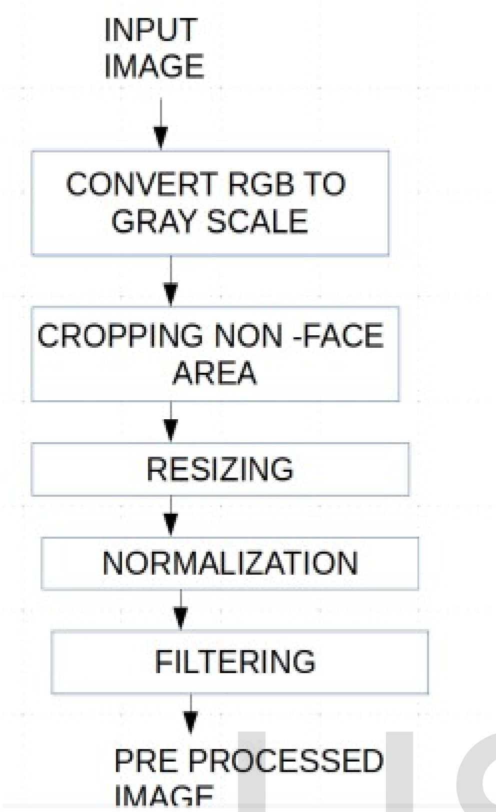

Fig 4.2.1: Pre Processing Stages

RGB – Gray scale conversion

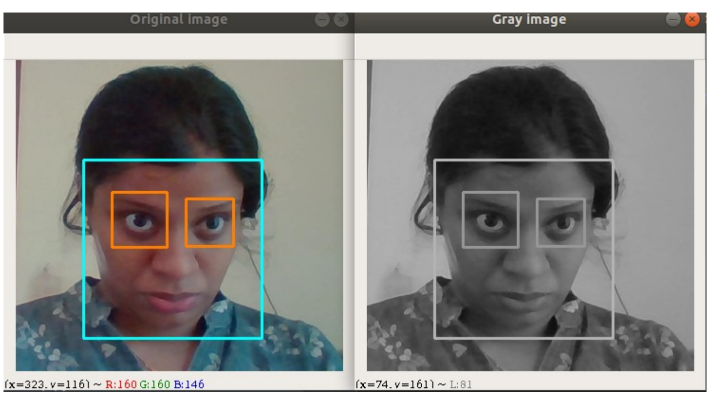

Fig 4.2.2: RGB-GRAY Scale conversion

Cropping non face area

Fig 4.2.3: Cropping

The non face area is cropped off to obtain the face region to perform further operations on the image.

Resizing

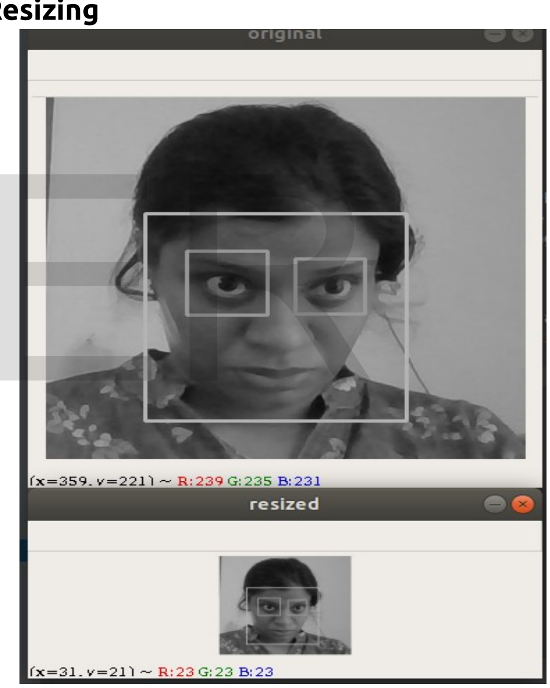

Fig 4.2.3: Resizing

Resizing image to 32* 32 which will be fed to the Convolution Neural Network.

Normalization

A normalized gamma-corrected contrast-limited adaptive histogram equalization tecnique is used.
Contrast Limited Adaptive Histogram Equalization technique(CLAHE) is followed by a NGC Function .As a final point, adding the NGC function to CLAHE can significantly improve its performance, wherein this function helps to reduce the brightness and enhance the contrast of the degraded image.
CLAHE is histogram equalisation adaptive method computes several histograms, each corresponding to a distinct section of the image, and uses them to redistribute the lightness values of the image. It is therefore suitable for improving the local contrast and enhancing the definitions of edges in each region of an image

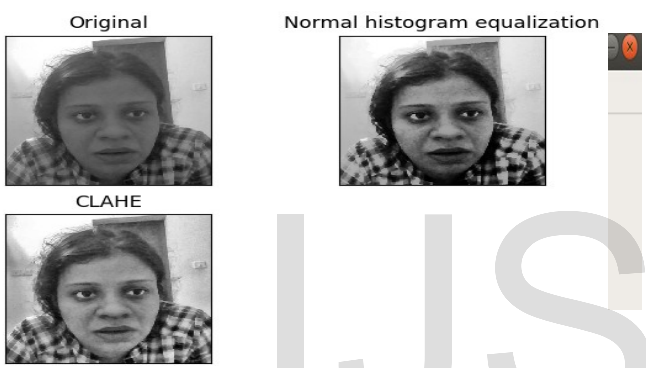

Fig 4.2.4: CLAHE

Applying gamma correction on the CLAHE image

Fig 4.2.4: Gamma Correction

Filtering
A Non- local means filter is used for noise reduction or image denoising.Non-local means filtering takes a mean of all pixels in the image, weighted by how similar these pixels are to the target pixel.

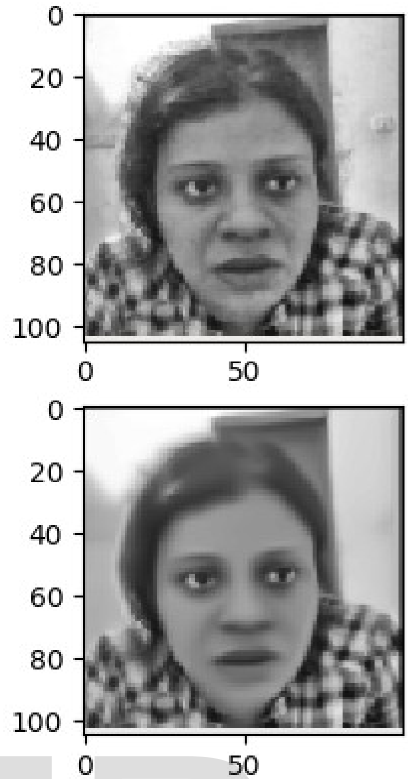

Fig
4.2.5 Non Local Means Filter

4.3.Feature Extraction

LTP(Local ternary pattern) is used for feature extraction.
In LTP, the neighborhood pixel gray values are compared with the central pixel gray values by using a threshold. Based on this comparison, the neighborhood value will be assigned one of the three values, +1 or 0 or -1.

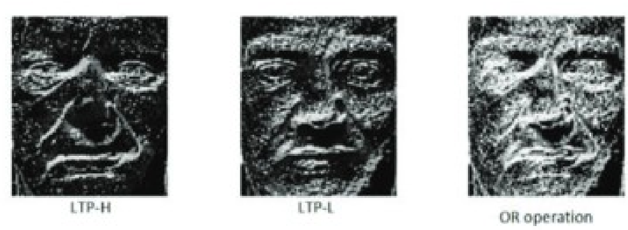

Fig 4.3: Feature Extraction

4.4 Classification

The network inputs are grayscale images with of size 32 × 32 and comprise two convolutional layers with 32 and 64 filters, and filter sizes of 5 × 5 and 7 × 7. After each convolutional layers, a Rectified Linear Unit (ReLU) activation functions is followed. Max pooling layers are also located after convolutional layers, and a fully connected layer with 256 hidden units are placed. Finally a softmax layer with 25 neurons which indicate numbers of AUs and fully connected to prior layer is considered in order to classification of Aus.

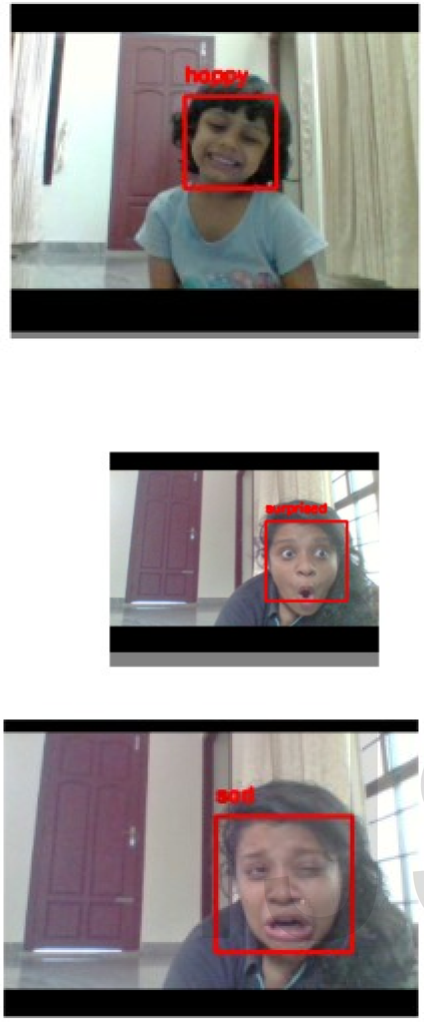

Fig 4.4: Some Emotions Classified

5. EXPERIMENTS , RESULTS AND EVALUATION

The proposed system was trained and tested on Extended Cohn-Kanade (CK+)[27] dataset which is a publicly available and used for facial expression studies. The dataset includes 123 subjects . Dataset images are grayscale with size 640 by 480 and 8-bit precision, and for each image there is a descriptor file which contain labels for denoting AUs that presented in each images. The dataset includes images for the expressions: neutral, happy, sad, surprise, fear, anger, disgust and contempt.

6.CONCLUSION

A novel approach for detecting action units (AUs) which is a coding of facial movements in psychological framework. A CNN is developed for optimal feature extraction and detecting AUs and by means of detecting seven expressed emotions. CNNs are able to learn characteristics of facial expression and increase facial emotion recognition accuracy.

REFERENCES

1]Mehrabian, A.; Russell, J.A. An Approach to Environmental Psychology; The MIT Press: Cambridge, MA, USA, 1974.

2]Cowie, R.; Douglas-Cowie, E.; Tsapatsoulis, N.; Votsis, G.; Kollias, S.; Fellenz, W.; Taylor, J.G. Emotion recognition in human-computer interaction. IEEE Signal Process. Mag. **2001**, 18, 32–80.

3]Bartlett, M.S.; Littlewort, G.; Fasel, I.; Movellan, J.R. Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction. In Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop, Madison, WI, USA, 16–22 June 2003; Volume 5, p. 53.

4]Bekele, E.; Zheng, Z.; Swanson, A.; Crittendon, J.; Warren, Z.; Sarkar, N. Understanding how adolescents with autism respond to facial expressions in virtual reality environments. IEEE Trans. Vis. Comput. Graph. **2013**, 19, 711–720.

5]Chen, C.H.; Lee, I.J.; Lin, L.Y. Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders. Res. Dev. Disabil. **2015**, 36, 396–403.

6]Assari, M.A.; Rahmati, M. Driver drowsiness detection using face expression recognition. In Proceedings of the IEEE International Conference on Signal and Image Processing Applications, Kuala Lumpur, Malaysia, 16–18 November 2011; pp. 337–341.

7]Jabon, M.; Bailenson, J.; Pontikakis, E.; Takayama, L.; Nass, C. Facial expression analysis for predicting unsafe driving behavior. IEEE Perv. Comput. **2011**, 10, 84–95.

8]Kapoor, A.; Burleson, W.; Picard, R.W. Automatic prediction of frustration. Int. J. Hum.-Comput. Stud. **2007**, 65, 724–736.

9]Lankes, M.; Riegler, S.; Weiss, A.; Mirlacher, T.; Pirker, M.; Tscheligi, M. Facial expressions as game input with different emotional feedback conditions. In Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, Yokohama, Japan, 3–5 December 2008; pp. 253–256.

10]Ekman, P.; Friesen, W.V. Facial Action Coding System: Investigator's Guide; Consulting Psychologists Press: Palo Alto, CA, USA, 1978.

11]Viola, P.; Jones, M.J. Robust real-time face detection. Int. J. Comput. Vis. **2004**, 57, 137–154.

12]Hsu, R.L.; Abdel-Mottaleb, M.; Jain, A.K. Face detection in color images. IEEE Trans. Pattern Anal. Mach. Intell. **2002**, 24, 696–706.

13]Shan, S.; Gao, W.; Cao, B.; Zhao, D. Illumination normalization for robust face recognition against varying lighting conditions. In Proceedings of the 2003 IEEE International SOI Conference, Nice, France, 17 October 2003; pp. 157–164.

14]Chen, W.; Er, M.J.; Wu, S. Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. IEEE Trans. Syst. Man Cybern. Part B (Cybern.) **2006**, 36, 458–466.

15]Du, S.; Ward, R. Wavelet-based illumination normalization for face recognition. In Proceedings of the IEEE International Conference on Image Processing 2005, Genova, Italy, 14 September 2005; Volume 2, p. II-954.

16]Lyons, M.; Akamatsu, S.; Kamachi, M.; Gyoba, J. Coding facial expressions with gabor wavelets. In Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 14–16 April 1998; pp. 200–205.

17]Ahonen, T.; Hadid, A.; Pietikäinen, M. Face recognition with local binary patterns. In European Conference on Computer Vision; Springer: Berlin, Germany, 2004; pp. 469–481

18]Cootes, T.F.; Taylor, C.J.; Cooper, D.H.; Graham, J. Active shape models-their training and application. Comput. Vis. Image Underst. **1995**, 61, 38–59.

19]Yacoob, Y.; Davis, L.S. Recognizing human facial expressions from long image sequences using optical flow. IEEE Trans. Pattern Anal. Mach. Intell. **1996**, 18, 636–642.

20]Liu, Y.; Wang, J.D.; Li, P. A Feature Point Tracking Method Based on The Combination of SIFT Algorithm and KLT Matching Algorithm. J. Astronaut. **2011**, 7, 028.

21]Sohail, A.S.M.; Bhattacharya, P. Classification of facial expressions using k-nearest neighbor classifier. In Proceedings of the International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications; Springer: Berlin, Germany, 2007; pp. 555–566.

22]Wang, X.H.; Liu, A.; Zhang, S.Q. New facial expression recognition based on FSVM and KNN. Optik **2015**, 126, 3132–3134.

23]Valstar, M.; Patras, I.; Pantic, M. Facial action unit recognition using temporal templates. In Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication, Kurashiki, Okayama, Japan, 22 September 2004; pp. 253–258.

24]Michel, P.; El Kaliouby, R. Real time facial expression recognition in video using support vector machines. In Proceedings of the 5th International Conference on Multimodal Interfaces, Vancouver, BC, Canada, 5–7 November 2003; pp. 258–264.
Facial Expression Recognition Based on Improved Local Ternary Pattern and Stacked Auto-Encoder Yao Wu a) and Weigen Qiu

25]Gudipati, V.K.; Barman, O.R.; Gaffoor, M.; Abuzneid, A. Efficient facial expression recognition using adaboost and haar cascade classifiers. In Proceedings of the Annual Connecticut Conference on Industrial Electronics, Technology & Automation (CT-IETA), Bridgeport, CT, USA, 14–15 October 2016; pp. 1–4.

26]Walecki, R.; Rudovic, O.; Pavlovic, V.; Schuller, B.; Pantic, M. Deep structured learning for facial expression intensity estimation. Image Vis. Comput **2017**, 259, 143–154.

27]Lucey, P.; Cohn, J.F.; Kanade, T.; Saragih, J.; Ambadar, Z.; Matthews, I. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 13–18 June 2010; pp. 94–101