Facial Emotion Recognition at Micro Level Using Relief and Facial Action Unit

Ujwalla Gawande[1], Hemalatha[2,]

[1]Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia

[1]Department of IT, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

[2]Department of computer science and business systems Panimalar Engineering college Poonamallee. Chennai. Tamilnadu. India

ujwallgawande@yahoo.co.in, pithemalatha@gmail.com

Abstract: Facial micro-emotion are small and quick facial expressions that show what a person is really feeling, even if they try to hide it. Analysing such emotions proves challenging since it happens within a fraction of a second, due to minor changes in facial muscles. This paper describes an implementation for detecting micro-emotions from video or webcam stream. Initially, the videos have been converted into frames, and a sliding window will convert those frames in proper and fixed segments. Face will be identified using framework MediaPipe FaceMesh. It helps to determine 468 landmarks on face. Using face width approach the faces in the frames has been normalized to fixed size images. The relevant features /landmarks that are important for facial expression recognition has been extracted using ReliefF, to reduce the complexity of the algorithm. It works well with the noisy image, different skin tones and lesser data set. A Facial Action Unit (AU) framework has been engaged for detecting facial movements. Further those AU s have been mapped to micro level emotions and trained by convolution neural network. The experimentation has been performed on - CREMA-D (Crowd-Sourced Emotional Multimodal Actors Dataset), consisting of 7,442 original video clips. The results are promising and are able to detect the micro level emotions in noisy environment

Key words: Facial micro emotion, Facial expression, Facial action unit.

1. Introduction:

Facial micro-emotion detection is an important area of research in artificial intelligence. In today's era, understanding someone's true emotion is very important. If such emotions are detected in time, a suicide rate can be reduced. These expressions are connected to feelings. Gender and culture do not make a big difference in how micro-expressions appear, but a person's background or situation can change the way they are shown. Developing micro-expressions-based emotion recognition system, can help us understand one's emotions better and can be useful in the future for understanding psychology of a person. It helps machines understand human emotions (robots) more accurately and respond/act accordingly.

The Fig. 1 represents the human emotions in different categories. The emotions are broadly classified into two classes: Comfortable Emotions and Uncomfortable Emotions, i.e. positive and negative emotion, respectively. Each of these categories are further subdivided into macro/core level emotions such as Happy, Loved, Confident, Playful, Embarrassed, Angry, Scared and Sad. These core emotions form the foundational states. Those macro level emotions are further categorized into micro emotions which are more specific emotional expressions derived from the major states.

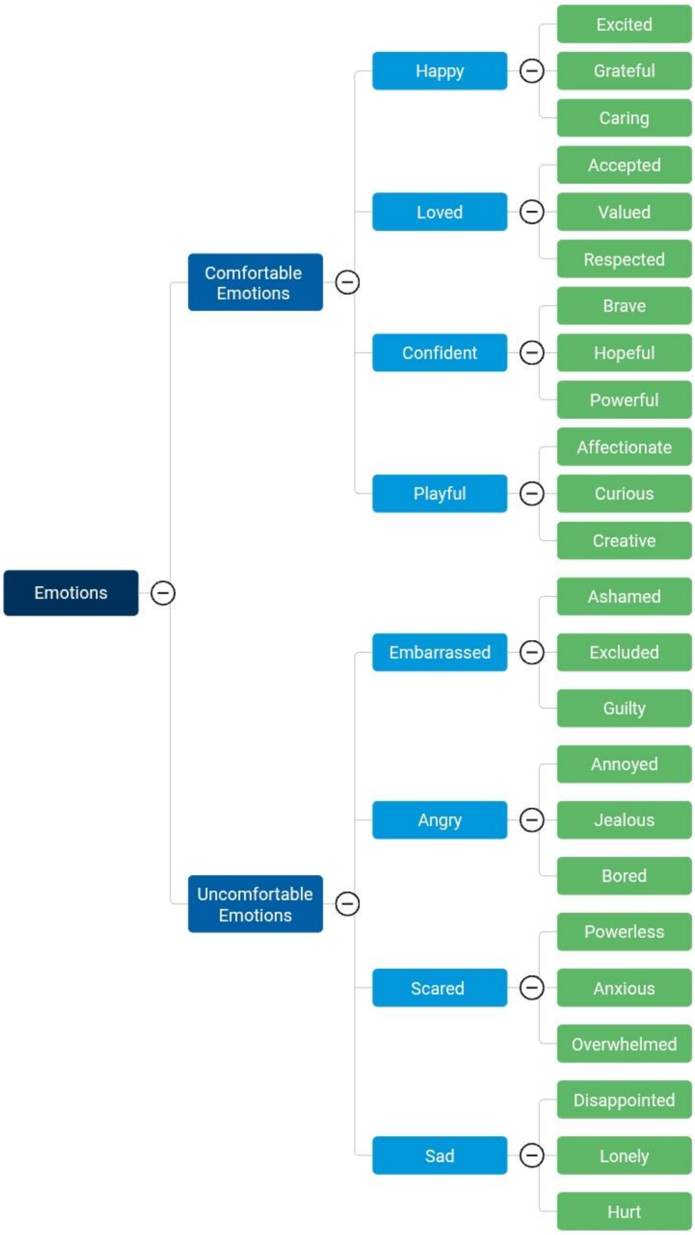

Figure 1. Categorization of emotions at macro and micro level

Most of the work done on facial emotion recognition has been performed on macro level. However very less work has been done on micro level. Hence this paper proposed a micro-expressions detection using relief and Facial Action Coding System (FACS). The FACS is a scientific framework developed to describe facial movements and relief F has been used to reduce the number of features. Those wo approaches break down facial expressions into individual Action Units (AUs), each representing specific muscle activations. Those AUs are mapped to the micro expression and are trained using CNN. The results of the proposed approach outperforms the existing approaches in literature.

2. Related Work

In today's era, humans' mental health is very important. For the same detecting an emotion at right time may be very useful for avoiding any mishaps in future. In the literature, much research has been carried out by the researchers for detecting the macro emotion from face. Very few work has been carried out on micro level emotion detection. For working on that emotion, the various datasets available are SAMM [1][4], KMU-FED [1], CASME II [3], FER-2013[11][14], Emognition Wearable Dataset [13], AffectNet [15], etc.

Recently the work introduced by Malik et al. [1] proposed the technologies such as ResNet18 for capturing visual features and Bidirectional LSTM for temporal features processing to detect AUs in facial images. They also use agglomerative clustering of AUs for increase the effectivity of system. The results claimed by them is 96.38 % accuracy on SAMM dataset and 95.96 % accuracy on n KMU-FED dataset. Though the results are high, but their work is limited to mostly macro level emotions and very few (i.e. 2) micro level emotion.

Most of the work in this area is focussed on the models like Convolutional Neural Network (CNN) which is a deep learning model. Many used hybrid models i.e. CNN along with another model or algorithm. These hybrid models help to create more powerful system and there are high chances of increase in accuracy. For example, Naidana et al. [3] work with CNN along the real-coded genetic algorithm for hyperparameter optimization. They have claimed the accuracy of 89.24% on CASME II dataset. Use of genetic algorithm is a good option but a drawback that the processing speed of CNN along GA is too slow. In the work by Bai [10], Spatio-temporal deep models like CNN + LSTM/GRU is proposed. In their work the claimed that deep models alone are insufficient for capturing extremely subtle micro-movements to detect micro emotions and therefore they have combined it with optical flow or handcrafted motion features. [10]. Similarly [4][7] captures spatio-temporal facial dynamics using a single-layer ConvLSTM. This hybrid model integrates the Convolution and LSTM model together to detect facial movements and detect micro emotions. Also, Zheng et. al. [7] proposed Hybrid_FM5 which has pipeline of CNN to ConvLSTM to transformer to fully connected layer, claiming accuracy of 65%. Further Hybrid_FM6 with same pipeline with different number of LSTM filters gives accuracy of 88%. Similarly, a hybrid model which integrates the ResNet-based CNN with micro-attention module is proposed by [5], to improve recognition of facial movements. They claimed that the accuracy has been boosted using the hybrid model.

Durga et al. [2] has work on Deep learning based facial emotion recognition using an adaptive tiefes FCNN. They proposed adaptive feature extraction and ignored temporal micro expressions dynamics to enhances facial muscle movements detection. Their model claims the accuracy of 99% but they have worked on majorly the macro level emotions and very few micro level emotions.

From this literature [6][8][9], it has been observed that most of the work has been performed on macro level and very few on micro level. Hence there is a need of micro level emotion detection to carefully observe the person to understand emotions such as lonely, annoyed, guilty, etc. This will help in taking necessary/corrective action in time and improve the life style.

3. Methodology

As from the literature it has been observed that major work has been done on macro level emotion detection. Also, the accuracy claimed by the researcher is around 90%. In many cases the Data set, Noise, occlusion, uncertainty, Misclassification of emotions (anger or excitement), etc. are the concerns. To address this issue this paper proposed micro level emotion detection. The proposed approach of micro-expression detection system is shown in figure 2. Here the videos from the CREMA-D data set have been given as an input to the proposed system.

3.1 Frame Extraction

Normally videos run at 24 to 30 frames per second. This video will be converted into frames using OpenCV. As a result of this operation, we shall have images for the facial expression. To get proper micro emotions, it is required to convert the sequential consecutive frames into small segments. The sliding window segmentation technique has been used to convert frames into segments for short and fixed duration. This will help in the detection of sudden facial muscle movements within short time intervals for recognizing the micro emotions efficiently.

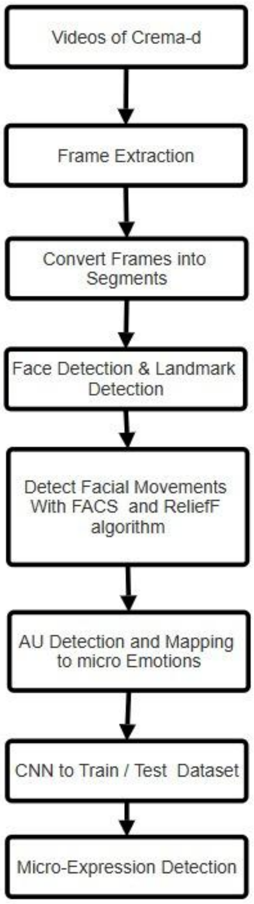

Figure 2: Block diagram for micro level emotion detection

3.2 Face Detection & Extraction

In the segments of frames, the next step in the proposed approach is to detect the face in the frames. This has been done using Media Pipe FaceMesh framework. It uses BlazeFace Detector to detect faces. This framework can detect 468 landmarks on face as shown in figure 3. Further to remove the scale variations in different frames, the FaceWidth has been used as given in equation 1.

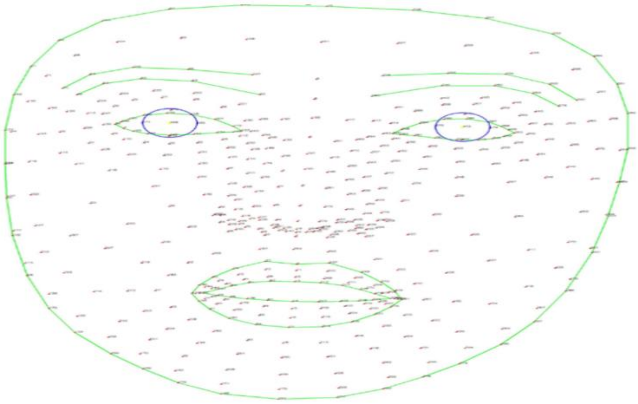

Figure 3. MediaPipe Face Mesh: A 3D Facial Landmark Detector with 468 Landmarks.[12]

$$\mathrm{FaceWidth} = \mathrm{lm(x)} - \mathrm{lm(y)} \qquad \text{Eq. 1}$$

Where: lm(x), lm(y) = Landmark coordinates of left and right cheek

3.3 Facial Action Unit Detection

Further, in the detected FaceWidth, the Facial Action Coding System (FACS) mechanism along with ReliefF algorithm has been implemented to detect the set of facial muscle movements that correspond to a displayed emotion. Using FACS, firstly the Action Units (AU) has been detected, viz. Inner Brow Raiser, Lip Corner Depressor, etc. Those action units can detect the movements of the landmarks in the segments. To reduce the complexity of the feature vector, the most relevant features of facial expression have been extracted using ReliefF algorithm. The advantage of using ReliefF framework is its capability to work in noisy environment and few datasets. With the help of AUs, we have mapped the micro emotion as shown in table 1.

Table 1: Micro and core emotion mapping with AUs

Core Emotion	**Micro Emotion**	**AUs Combination**	**Facial Movement**
Sad	Hurt	AU1 + AU15	Inner Brow Raiser + Lip Corner Depressor
	Lonely	AU1 + AU4	Inner Brow Raiser + Brow Loewer
	Disappointed	AU4 + AU15	Brow Loewer + Lip Corner Depressor
Scared	Overwhelmed	AU5 + AU26	Upper Lid Raiser + Jaw Drop
	Anxious	AU5 + AU1	Upper Lid Raiser + Inner Brow Raiser
	Powerless	AU1 + AU20	Inner Brow Raiser + Lip Stretcher
Angry	Annoyed	AU4 + AU7	Brow lowered + Lid Tightener
	Jealous	AU4 + AU23	Brow lowered + Lip Tightener
	Bored	AU7 - AU12	Lid Tightener + Lip Corner Puller
Embarrassed	Ashamed	AU15 + AU7	Lip Corner Depressor + Lid Tightener
	Excluded	AU15 - AU1	Lip Corner Depressor + Inner Brow Raiser
	Guilty	AU15 + AU23	Lip Corner Depressor + Lip Tightener
Playful	Affectionate	AU12 + AU25	Lip Corner Puller + Lips Part
	Curious	AU12 + AU2	Lip Corner Puller + Outer Brow Raiser
	Creative	AU12 + AU20	Lip Corner Puller + Lip Stretcher
Confident	Brave	AU12 + AU7	Lip Corner Puller + Lid Tightener
	Hopeful	AU12 + AU1	Lip Corner Puller + Inner Brow Raiser
	Powerful	AU12 + AU23	Lip Corner Puller + Lip Tightener
Loved	Accepted	AU12 + AU1 + AU2	Lip Corner Puller + Inner Brow Raiser + Outer Brow Raiser
	Valued	AU12 + AU5	Lip Corner Puller + Upper Lid Raiser
	Respected	AU12 + AU7 + AU23	Lip Corner Puller + Lid Tightener + Lip Tightener
Happy	Excited	AU12 + AU26	Lip Corner Puller + Jaw Drop

Grateful	AU12 + AU1	Lip Corner Puller + Inner Brow Raiser
Caring	AU12+AU7	Lip Corner Puller + Lid Tightener

Now we have a result of mapping. We use AUs along with its mapping of micro level emotions to training of Convolutional Neural Network (CNN). CNN is an algorithm used to recognize patterns in data. With repeated training we get a proper class for micro emotion. During testing the CNN is able to classify the micro emotions. With this our approach achieved a higher solidify accuracy.

5. Result and discussion

As an outcome of the proposed approach, we have received an accuracy of 88% at micro level emotion detection. Further a comparison with the existing approaches have also been shown in Table 2. The same has been depicted in figure 4. From this it has been observed that, very less work has been performed on micro level emotion detection. In some cases, the mixed emotions are also detected, however they have not covered all the micro level emotions. Further the accuracy mentioned in their work is only 62.48%. The screenshot of the outcome is shown in figure 5, wherein the persons emotion is recognized as creative and ashamed based on the Aus trained.

Table 2. Comparison analysis of proposed methodology with existing approaches

Macro/Micro	**Algorithm / Model**	**Dataset**	**Size of the dataset**	**Performance / Accuracy**	**Reference**
Micro	Proposed algorithm	Crema-D	7,442 videos	88%	Proposed methodology
6 emotions Macro	CNN	FER-2013	35,887 grayscale images	92%	Kaur & Kumar [11]
9 emotions Mixed	CNN + RNN	Emognition Wearable Dataset	2,535 facial images	62.48%	Manalu & Rifai [13]
7 emotions Macro	Deep CNN	FER-2013	35,887 grayscale images	91%	Meena et al. [14]
8 emotions Macro	Knowledge Distillation CNN	AffectNet	1M images, 45K Annotated Images	73.62%	Lee et al. [15]

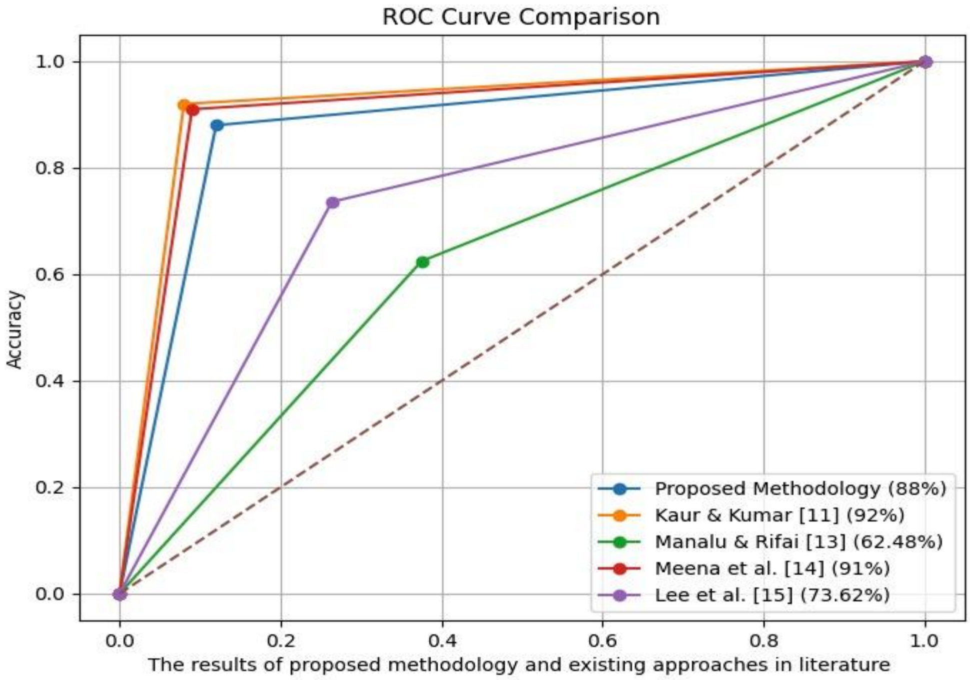

Figure 4: ROC Curve of Comparison analysis of proposed methodology with existing approaches

Figure 5. Result of Micro emotion detection

From the above results, we can claim that our proposed algorithm outperforms the existing approached. It efficiently recognized the micro level emotions with the Accuracy of 88%.

4. Conclusion

This paper presents an approach for micro level emotions detection in videos. Here firstly the videos are converted into frames, and a sliding window approach is used to segment the frames. This approach is able to detect finest change in facial muscle movement. Further the MediaPipe FaceMesh has been used to detect the face in frames and 468 landmarks in face. To normalize the face images the face width approach has been used. The relevant features are detected/segregated using ReliefF. This will help in reducing the training time and computational complexity. From the combination of Facial Action Unit (AU) framework and Relief F, we are able to obtain the Action unit, that are the most prominent features, which is further mapped to the micro level emotions. Based on this mapping the CNN model is trained. The work has been performed on CREMA-D data set consisting of 7,442 original videos. The results have shown the accuracy of 88%. This approach is very useful for real time scenarios to understand the emotions of a person. The future work lies in increasing the accuracy of the proposed framework for detecting the micro level emotions in face.

References:

[1] P. Malik, J. Singh, F. Ali, and S. Sehra, "Action unit based micro-expression recognition framework for driver emotional state detection," Scientific Reports 15, July 2025. doi:10.1038/s41598-025-12245-7.

[2] B. K. Durga, V. Rajesh, S. Jagannadham, P. S. Kumar, A. N. Z. Rashed, and K. Saikumar, "Deep learning-based micro facial expression recognition using an adaptive Tiefes FCNN model," Traitement du Signal, vol. 40, no. 3, pp. 1035–1043, June 2023, doi: 10.18280/ts.400319.

[3] K. S. Naidana, Y. Yarra, and L. P. Divvela, "Facial micro-expression classification through an optimized convolutional neural network using genetic algorithm," Bulletin of Electrical Engineering and Informatics, vol. 14, no. 1, February 2025.
DOI: https://doi.org/10.11591/eei.v14i1.8048

[4] S. Shukla, P. K. Rai, and T. T. Verlekar, "Micro-expression recognition using a shallow ConvLSTM-based network," in Proceedings of the Asian Conference on Computer Vision (ACCV) Workshops, Springer, Lecture Notes in Computer Science, vol. 13848, pp. 19–30, 2022.

[5] C. Wang, M. Peng, T. Bi, and T. Chen, "Micro-attention for micro-expression recognition," arXiv:1811.02360, 2019.

[6] T. Shuai, S. Beng, F. B. Khalid, and R. W. B. O. K. Rahmat, "Advances in Facial Micro-Expression Detection and Recognition: A Comprehensive Review," Information, vol. 16, no. 10, p. 876, October 2025. doi:10.3390/info16100876.

[7] Y. Zheng and E. Blasch, "Facial micro-expression recognition enhanced by score fusion and a hybrid model from convolutional LSTM and vision transformer," Sensors, vol. 23, no. 12, p. 5650, June 2023. doi:10.3390/s23125650.

[8] R. Vats, M. Kumar, R. Rai, and G. Verma, "Comprehensive review and analysis on facial emotion recognition methods," *J. Electron. Imaging*, vol. 32, no. 4, Art. no. 040901, Aug. 2023, doi: 10.1117/1.JEI.32.4.040901.

[9] Haritha S., S. Gupta, S. Mehraj, and D. D. J., "Emotion Detection Using Facial Expression – A Comprehensive Review," Indiana Journal of Multidisciplinary Research, vol. 4, no. 3, June 2024, doi: 10.5281/zenodo.12677782.

[10] M. Bai, "Detection of micro-expression recognition based on spatio-temporal modelling and spatial attention," in Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI '20), Oct. 2020, pp. 703–707. doi: 10.1145/3382507.3421160.

[11] M. Kaur and M. Kumar, “Facial emotion recognition: A comprehensive review,” Expert Systems, vol. 41, no. 10, Jun. 2024, doi: 10.1111/exsy.13670.

[12] M Alanazi, G. Ushaw, and G. Morgan, “Improving detection of DeepFakes through facial region analysis in images,” *Electronics*, vol. 13, no. 1, p. 126, Dec. 2023, doi: 10.3390/electronics13010126.

[13] H. V. Manalu and A. P. Rifai, “Detection of human emotions through facial expressions using hybrid convolutional neural network-recurrent neural network algorithm,” Intelligent Systems with Applications, vol. 21, p. 200339, Feb. 2024, doi: 10.1016/j.iswa.2024.200339.

[14] G. Meena, K. K. Mohbey, A. Indian, M. Z. Khan, and S. Kumar, “Identifying emotions from facial expressions using a deep convolutional neural network-based approach,” Multimedia Tools and Applications, vol. 83, no. 6, pp. 15711–15732, Jul. 2023, doi: 10.1007/s11042-023-16174-3.

[15] K. Lee, S. Kim, and E. C. Lee, “Fast and accurate facial expression image classification and regression method based on knowledge distillation,” Applied Sciences, vol. 13, no. 11, p. 6409, May 2023, doi: 10.3390/app13116409.