Archives available at journals.mriindia.com

International Journal on Advanced Computer Engineering and Communication Technology

ISSN: 2278-5140

Volume 14 Issue 03s, 2025

A Systematic Review of Facial Expression Recognition (FER)

[1]Barkha Sheikh, [2]Dr.Pragati Patil, [3]Ashwini Mahajan

[1,2,3] Computer Science Engineering
Tulsiramji Gaikwad-Patil College Of Engineering and technology,
Nagpur, India
Email: [1]barkhasheikh86@gmail.com, [2]Viceprincipal@tgpcet.com, [3]ashwini.cse@tgpcet.com

Peer Review Information

Submission: 05 Nov 2025

Revision: 25 Nov 2025

Acceptance: 17 Dec 2025

Keywords

Automatic Computing, Face emotion detection, Static images

Abstract

In order to improve human personification among humans, robots, and sophisticated symbols, by utilizing advances in artificial intelligence, Facial Expression Recognition (FER) seeks to identify facial expressions from still images. New problems and methods that are not given much thought in that FER mindset are confronted as the FER section moves from controlled laboratory conditions to increasingly complicated real-world scenarios. Advanced techniques have been developed quickly. This paper offers a thorough analysis of image-based static FER (SFER) methods, looking at everything from challenge-centred order to model-arranged improvement. We start with a basic analysis of current surveys, an overview of common datasets and methodologies, and a workflow using FER in order to build a solid basis.

Introduction

Emotion detection has broad implications and is important in important public domains [1]. Artificial intelligence for feeling and expression acknowledgment was named the top 50 emerging technologies that would have a big impact on the business and society in 2024 by Improve UK, the UK's advancement organization. Among the top ten boondocks scientific worries of 2024, according to the China Association for Science and Innovation, which eloquently outlined the year's key logical difficulties, was the investigation of intelligent humans and robots with emotions and sophisticated understanding. It is obvious that the advancement of technology for identifying emotions and behaviours in artificial intelligence has become a vital prerequisite for multidisciplinary investigation, improved manufacturing, and general computer-based intelligence [2].

Face emotion [3] is a basic and easy technique for humans to convey intense feelings. It is occasionally used and is extremely important in relational interactions [4][5]. Compared to other kinds of communication including voice, signs, and body postures, they use nonverbal cues to convey more sophisticated emotional information [6]. Darwin first introduced the concept of facial feelings in the book., "The Expression of the Emotions in Man and Animals (1872 It is being noted that emotion is fundamental to nature and forms the basis of both human and animal development. Six basic feelings were proposed by Ekman and Friesen [7]: joyful, sad, angry, surprised, fear, and disgust. They also found a general correlation between the types of sentiments that are consistent throughout societies and explicit facial muscle examples. Techniques for facial feeling acknowledgment (FER) have advanced quickly in recent years, keeping pace with improvements in simulated intelligence. These methods have shown several uses in clinical diagnostics [10], psychological research [8][9], and astute human-PC collaboration. Determining a person's state of residency from their look is the aim of the FER [11],[12]. The FER can be separated into two categories: picture-based static FER (SFER) along with video based dynamic FER (DFER), depending on the kind of data utilized to capture emotions.

[13] [17]. Posture impediment, cross-area inconsistency, mark vulnerability, inadequate information volume, and cross-modality are the reasons why the SFER has settling difficulties. In order to address the problems of inadequate information volume and mark vulnerability, scientists also employ various information expansion tactics and regularization procedures. Cross-modular data combining is used in expansion to improve the robustness and accuracy of emotion recognition.

Fig. 1 depicts the five FER stages. The preprocessing step uses a picture or picture succession—a sequence of pictures that are nonpartisan to an articulation—as information and proposes to complete the commotion.

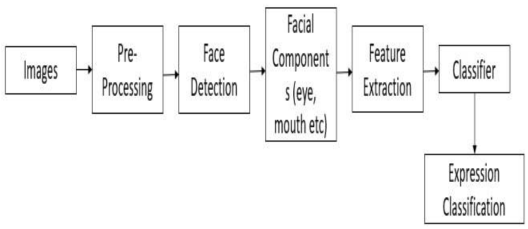

Fig.1. Classification block diagram of facial expressions

For further processing. For the nose, cheeks, lips, eye temple, eyes, ears, front head, and so forth, the facial part identification determines the return on investment. Highlights from the returns on initial capital investment are extracted by the element extraction stage. The most widely used extraction methods include, but are not limited to local binary pattern analysis (LBP), he Gabor filter, linear discriminant analysis (LDA), The local Gradient Code (LGC), principal component analysis (PCA), Local The direction Patterns (LDP). and the independent component analysis (ICA). The two prevalent order techniques are SVM (Support Vector Machine) along with NN (Nearest Neighbor). The classification step divides the pieces into distinct look classes based on arrangement tactics.

Ease of Use

This section reviewed facial expression recognition research based on its methods, techniques, datasets, results, and benefits.

In [18]/ 2021, author uses Active learning and +SVM alogorithm with Viola Jones based method generate result for female is 94% and male is 90% accuracy rate.

In [19]/ 2021, cross-connected CNN algorithm is used and MBCC-CNN method Fer2013, FERC and RAF data- sets. Author used a novel approach based on a cross- connected multiple-based CNN algorithm. Fer2013, FERC and RAF data sets recognition accuracy 71.52%, 88.10% and 87.34%.

In [20]/ 2021, KNN, radial bias, a random forest, SVM, MLP, and a decision-making board algorithm are used. Uses a minimal chi-square criteria to provide a consistency of different controlled classification approaches for facial emotion detection. Resultant 94.23 % accuracy.

In [21]/ 2021, author used Decision trees, SVM, KNN, Random Forestry, MLP, and radial basis functions algorithm. Reliever-F approach. Datasets used CK+. Effective method for gathering and classifying facial expressions Face picture sequence. It intends to examine the accuracy of six classifications that rely on the Reliever-F technique for function collection, with a focus on using the fewest possible attributes. Resultant 94.93% of accuracy.

In [22]/]2020, Two-stage neural network based on fuzzy fusion (TSFF-CNN) algorithm is used, Spectrogram and LBP-TOP used as a method, for dataset SAVEE, eNTERFACE'0 5, and AFEW. To achieve effective modality fusion. TSFFCNN performs effectively when the contribution of each modality's data to emotion recognition is significantly unbalanced. AFEW 50.28%, eNTERFACE' 05 90.82%, and SAVEE 99.79% as a resultant.

In [23]/ 2020, author utilizes the KNN, Naïve Bayes, MLP, and decision tree algorithms, for methodology information gain and correlation gains ratio, for dataset CK+. The eight essential facial expressions that people should be aware of provided FER comparison methodology to identify the most notable features of face photos using three feature selection techniques: correlation, gains ratio, and data gain.

In [24]/ 2020, (CNN) utilized to identify micro-expressions in intricate pictures that are requested. In order to enter complicated images, rank pooling is introduced, extending the facial movement magnification technique to input sequences. 0CASMEII, SMIC and SAMM dataset are used. To improve the automatic identification of face expressions using advanced manual extraction techniques possesses a thorough, effective, yet straightforward method for classifying and analyzing, in result author got SAMM 66.67%, SMIC 67.3%, and CASMEI 78.5%.

Table I. Evaluations of different approaches and strategies employed in the current system

Ref.No./Year	Algorithm	Method Used	Dataset	Approach	Merits	Result
[18]/ 2021	Active learning +SVM	Face Detection:- Viola Jones based method Feature Extraction:- Gradient Histogram	Ck+	To suggest utilizing Active learning and SVM together for the classification of facial expressions and the extraction of Aus	PCA offers a high precision rate.	Accuracy rate:- female- 94.07% Male- 90.77%
[19]/ 2021	cross-connected CNN	MBCC-CNN method	Fer2013, FERC and RAF data-sets	a novel approach based on a cross-connected multiple-based CNN	The connections and roles played by the three modules in the MBCC- CNN	Fer2013, FERC and RAF data sets recognition accuracy 71.52%, 88.10% and 87.34%,
[20]/ 2021	KNN, radial bias, a random forest, SVM, MLP, and a decision-making board.	minimum features selected by chisquare	CK+	a successful process for FER feature selection and detection from sequential facial images.	uses a minimal chi-square criteria to provide a consistency of different controlled classification approaches for facial emotion detection.	94.23 %
[21]/ 2021	Decision trees, SVM, KNN, Random Forestry, MLP, and radial basis functions	Reliever-F approach	CK+	Effective method for gathering and classifying facial expressions Face picture sequence	It intends to examine the accuracy of six classifications that rely on the Reliever-F technique for function collection, with a focus on using the fewest possible attributes.	94.93%

[22]/]2020	Two-stage neural network based on fuzzy fusion (TSFF-CNN)	Spectrogram and LBP-TOP	SAVEE, eNTERFACE'0 5, and AFEW	To achieve effective modality fusion.	TSFFCNN performs effectively when the contribution of each modality's data to emotion recognition is significantly unbalanced	AFEW 50.28%, eNTERFACE' 05 90.82%, and SAVEE 99.79%
[23]/ 2020	utilizing the KNN, Naïve Bayes, MLP, and decision tree algorithms	information gain and correlation gains ratio	CK+	The eight essential facial expressions that people should be aware of	provided FER comparison methodology to identify the most notable features of face photos using three feature selection techniques: correlation, gains ratio, and data gain.	-
[24]/ 2020	(CNN) utilized to identify micro-expressions in intricate pictures that are requested	In order to enter complicated images, rank pooling is introduced, extending the facial movement magnification technique to input sequences.	0CASMEII, SMIC and SAMM.	To improve the automatic identification of face expressions using advanced manual extraction techniques	possesses a thorough, effective, yet straightforward method for classifying and analyzing	SAMM 66.67%, SMIC 67.3%, and CASMEI 78.5%
[25]/ 2020	RNN and CNN	Bidirectional Short Memory (BiLSTM) - Standardization batch (BN) layers - VGG16 structure	CK+ OuluCASIA MMI, AffectNet	to enable a single function to understand the features (AUs) of the action units (nose, mouth, eyes, and eyebrows).	Comparing the SAANet to other specialized methods, experimental research shows that it is significantly improved. .	CK+= 99.54%AffectNet= 87.06%, Oulu CASIA MM =88.33%
[26]/ 2020	A Spatial-Temporal Focus Module and Channel (STCAM)	The network architecture of 3D-InceptionResNet	CK+ Oulu CASIA and MMI	To depict the intricate development of the face	According to the experimental results, the system outperforms state-of-the-art techniques.	CASI=89.16%, CK+=99.08%,

[27]/ 2020	Dynamic kernelbased facial expression representation	(UGMM), (MIK), (IMK)	MMI AFEW BP4D	to investigate the evolution of regional expressions captured from different facial regions.	A dynamic kernel is a crucial choice for any expression recognition algorithm that relies on measurement time or accuracy.	BP4D 74.5%
[28]/ 2019	To increase the distances between samples, a new loss function is suggested for the new facial expression DNN.	Integrates deep the pooling covariance and residual network units	-	Learning - Increase the gaps between samples from categories that are readily misunderstood in order to better achieve dynamic objectives.	The developed neural network will stop gradients from being lost.	DNN=86.47%
[29]/ 2019	FER technique for 3D Convolutional Neural Network (CNN) video	Island layer, GRU layer, Dropout layer, Stem layer, 3D LayoutResNets, and Softmax layer	(CK+, MMI, AFEW)	To improve the accuracy.	can most effectively ignore the effects of facial expression variability and adaptability, improve identification accuracy generalization	AFEW=82.36%, MMI=80.43%, and CK+=94.39%

In [25]/ 2020, RNN and CNN algorithm, Bidirectional Short Memory (BiLSTM) - Standardization batch (BN) layers - VGG16 structure used as method,dataset used is CK+ OuluCASIA MMI, AffectNet, to enable a single function to understand the features (AUs) of the action units (nose, mouth, eyes, and eyebrows). Comparing the SAANet to other specialized methods, experimental research shows that it is significantly improved. In resultant author got CK+= 99.54%AffectNet= 87.06%, Oulu CASIA MM =88.33%.

In [26]/ 2020, Author used A Spatial-Temporal Focus Module and Channel (STCAM) as n algorithm, for method the network architecture of 3D-InceptionResNet is used. CK+ as dataset, Oulu CASIA and MMI. To depict the intricate development of the face. According to the experimental results, the system outperforms state-of-the-art techniques.as a result author got CASI=89.16%, CK+=99.08%.

In [27]/ 2020, author used algorithm Dynamic kernel based facial expression representation, (UGMM), (MIK), (IMK) MMI AFEWas method. BP4D as dataset. To investigate the evolution of regional expressions captured from different facial regions. A dynamic kernel is a crucial choice for any expression recognition algorithm that relies on measurement time or accuracy. The resultant author got BP4D is 74.5%.

In [28]/ 2019, To increase the distances between samples, a new loss function is suggested for the new facial expression DNN. Integrates deep the pooling covariance and residual network units, Learning - Increase the gaps between samples from categories that are readily misunderstood in order to better achieve dynamic objectives. The developed neural network will stop gradients from being lost. Author got DNN=86.47% as result.

In [29]/ 2019, FER technique for 3D Convolutional Neural Network (CNN) video algorithm is used, Island layer, GRU layer, Dropout layer, Stem layer, 3D LayoutResNets, and Softmax layer, for dataset (CK+, MMI, AFEW), to improve the accuracy. can most effectively ignore the effects of facial expression variability and adaptability, improve identification accuracy generalization AFEW=82.36%, MMI=80.43%, and CK+=94.39%

Conclusion

Facial expression recognition (FER) has drawn a lot of interest in the field of computer-based intelligence because of its prospective applications in human-machine coordination and knowledge. The foundation, datasets, traditional work method, challenge-situated scientific classification of cutting-edge approaches, ongoing advancements, applications, ethical concerns, and emerging patterns are some of the points of view that are used in this review to audit FER activities in general. We intentionally analyse and condense FER datasets, task challenges, strategies, and execution evaluations using tables and figures to give a concise summary of the most current trends in FER. Experts from a variety of disciplines benefit immensely from this in-depth investigation since it allows them to quickly recognize issues and developments in the area, encouraging cooperation for the improvement of FER as a whole.

References

[1] S. Zhao, X. Hong, J. Yang, Y. Zhao, and G. Ding, "Toward label-efficient emotion and sentiment analysis," Proceedings of the IEEE, vol. 111, no. 10, pp. 1159–1197, 2023.

[2] J. Z. Wang, S. Zhao, C. Wu, R. B. Adams, M. G.
Newman, T. Shafir, and R. Tsachor, "Unlocking the emotional world of visual media: An overview of the science, research, and impact of understanding emotion drawing insights from psychology, engineering, and the arts." Proceedings of the IEEE, vol. 111, no. 10, pp. 1236–1286, 2023.

[3] E. G. Krumhuber, L. I. Skora, H. C. Hill, and K. Lander, "The role of facial movements in emotion recognition," Nature Reviews Psychology, vol. 2, no. 5, pp. 283–296, 2023.

[4] S. Cowen, D. Keltner, F. Schroff, B. Jou, H. Adam, and G. Prasad, "Sixteen facial expressions occur in similar contexts worldwide," Nature, vol. 589, no. 7841, pp. 251–257, 2021.

[5] S. Zhang, Y. Pan, and J. Z. Wang, "Learning emotion representations from verbal and nonverbal communication," in CVPR, 2023, pp. 18 993–19 004.

[6] Y. Wang, W. Song, W. Tao, A. Liotta, D. Yang, X. Li, S. Gao, Y. Sun, W. Ge, W. Zhang, and W. Zhang, "A systematic review on affective computing: Emotion models, databases, and recent advances," Information Fusion, vol. 83, pp. 19–52, 2022.

[7] P. Ekman and H. Oster, "Facial expressions of emotion," Annual review of psychology, vol. 30, no. 1, pp. 527–554, 1979.

[8] O. Rudovic, J. Lee, M. Dai, B. Schuller, and R. W. Picard, "Personalized machine learning for robot perception of affect and engagement in autism therapy," Science Robotics, vol. 3, no. 19, p. eaao6760, 2018.

[9] Savchenko,L. V. Savchenko, and I. Makarov, "Classifying Emotions and Engagement in Online Learning Based on a Single Facial Expression Recognition Neural Network," IEEE Taffc, Vol. 13,No. 4, PP. 2132–2143, 2022.

[10] S. Zhao, H. Yao, Y. Gao, G. Ding, and T.-S. Chua, "Predicting personalized image emotion perceptions in social networks," IEEE TAFFC, vol. 9, no. 4, pp. 526– 540, 2018.

[11] S. Wang, Z. Zheng, S. Yin, J. Yang, and Q. Ji, "A novel dynamic model capturing spatial and temporal patterns for facial expression analysis," IEEE TPAMI, vol. 42, no. 9, pp. 2082–2095, 2019.

[12] D. Liu, W. Dai, H. Zhang, X. Jin, J. Cao, and W. Kong, "Brain-machine coupled learning method for facial emotion recognition," IEEE TPAMI, vol. 45, no. 9, pp. 10 703–10 717, 2023.

[13] H. Zhou, S. Huang, F. Zhang, and C. Xu, "Ceprompt: Cross-modal emotion aware prompting for facial expression recognition," IEEE TCSVT, pp. 1–1, 2024.

[14] Y. Wang, S. Yan, W. Song, A. Liotta, J. Liu, D. Yang, S. Gao, and W. Zhang, "Mgr3net: Multigranularity region relation representation network for facial expression recognition in affective robots," IEEE TII, 2024.

[15] X. Jiang, Y. Zong, W. Zheng, C. Tang, W. Xia, C. Lu, and J. Liu, "Dfew: A large-scale database for recognizing dynamic facial expressions in the wild," in ACM MM, 2020, pp. 2881–2889.

[16] Y. Wang, Y. Sun, W. Song, S. Gao, Y. Huang, Z. Chen,
W. Ge, and W. Zhang, "Dpcnet: Dual path multi-excitation collaborative network for facial expression representation learning in videos," in ACM MM, 2022,
pp. 101–110.

[17] X. Mai, J. Lin, H. Wang, Z. Tao, Y. Wang, S. Yan, X. Tong, J. Yu, B. Wang, Z. Zhou, Q. Zhao, S. Gao, and W. Zhang, "All rivers run into the sea: Unified modality brain-inspired emotional central mechanism," in ACM MM, 2024.

[18] Yao, L., Wan, Y., Ni, H., & Xu, B. (2021). Action unit classification for facial expression recognition using active learning and SVM. Multimedia Tools and Applications, 80(16), 24287-24301

[19] Shi, Cuiping & Tan, Cong & Wang, Liguo. (2021). A Facial Expression Recognition Method Based on a Multibranch Cross-Connection

Convolutional Neural Network. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3063493.
[20] M. R. Mahmood, M. B. Abdulrazzaq, S. Zeebaree, A. K. Ibrahim, R. R. Zebari, and H. I. Dino, "Classification techniques' performance evaluation for facial expression recognition," Indonesian Journal of Electrical Engineering and Computer Science, vol. 21, pp. 176-1184, 2021.
[21] M. B. Abdulrazaq, M. R. Mahmood, S. R. Zeebaree, M. H. Abdulwahab, R. R. Zebari, and A. B. Sallow, "An Analytical Appraisal for Supervised Classifiers' Performance on Facial Expression Recognition Based on Relief-F Feature Selection," in Journal of Physics: Conference Series, 2021, p. 012055.
[22] M. Wu, W. Su, L. Chen, W. Pedrycz, and K. Hirota, "Two-stage Fuzzy Fusion basedConvolution Neural Network for Dynamic Emotion Recognition," IEEE Transactions on Affective Computing, 2020.
[23] H. I. Dino and M. B. Abdulrazzaq, "A Comparison of Four Classification Algorithms for Facial Expression Recognition," Polytechnic Journal, vol. 10, pp. 74-80, 2020.
[24] T. T. Q. Le, T.-K. Tran, and M. Rege, "Dynamic image for micro-expression recognition on region-based framework," in 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI), 2020, pp. 75-81.
[25] Liu, X. Ouyang, S. Xu, P. Zhou, K. He, and S. Wen, "SAANet: Siamese action-units attention network for improving dynamic facial expression recognition," Neurocomputing, vol. 413, pp. 145-157, 2020.
[26] W. Chen, D. Zhang, M. Li, and D.-J. Lee, "STCAM: Spatial-Temporal and Channel Attention Module for Dynamic Facial Expression Recognition," IEEE Transactions on Affective Computing, 2020.
[27] N. Perveen, D. Roy, and K. M. Chalavadi, "Facial Expression Recognition in Videos Using Dynamic Kernels," IEEE Transactions on Image Processing, vol. 29, pp. 8316-8325, 2020.
[28] G. Wen, T. Chang, H. Li, and L. Jiang, "Dynamic Objectives Learning for Facial Expression Recognition," IEEE Transactions on Multimedia, 2020.
[29] L. Deng, Q. Wang, and D. Yuan, "Dynamic Facial Expression Recognition Based on Deep Learning," in 2019 14th International Conference on Computer Science & Education (ICCSE), 2019, pp. 32-37.