International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

A Comprehensive Survey of Facial Expression Recognition Methods for Basic and Compound Emotions Under Occluded and Non-Occluded Conditions

Aparna V B[1], Ajitha R S[2]

[1, 2]Computer Science, NSS College Rajakumari, Rajakumari, India

Abstract: ***Facial Expression Recognition (FER) is a technology that enables the automatic identification and interpretation of human emotions from facial images or videos. FER plays a major role in various fields such as neuro science, bio medical engineering, health care, education human computer interaction and public safety. Most of the existing methods provides solution to the problem of facial expression recognition either by considering the local patches of face or by using the global image. The field of FER faces many challenges such as pose variation, low resolution images, aging effects, variation in lighting, intra class variations, inter class similarities and partial occlusion. Among these, partial occlusion is very challenging since it hides important facial information leading to poor emotion recognition accuracy of the system. This survey provides a detailed analysis of various FER methods and focusing on occlusion and non-occluded conditions. The proposed study analyses different methods for recognizing facial expressions. Also, it portrays the current trends in technologies, datasets and various tools used for FER.***

Categories: AI applications, Image Processing and Analysis, Computer Vision

Keywords: facial expression recognition (fer), basic expression, compound expression, micro expression, convolutional neural network (cnn)

1. Introduction and Background

Facial expressions are the most honest and powerful means to communicate complex emotions [1]. Facial expressions are closely tied to automatic neurological responses and are very significant in human communication. They serve as a diagnostic indicator for detecting neurological conditions like Parkinson's disease and psychological conditions like severe depression. Facial expressions play a major role to express emotions and it is a crucial part in communication and social interactions. Facial expressions can be classified into two categories such as Basic Expressions and Compound Expressions. Basic expressions are the seven universal expressions like happiness, surprise, sadness, anger, fear, disgust, and neutral. Compound expressions are the combination of two basic expressions. For example, 'happily surprised' is a compound expression formed by combining two basic expressions happy and surprise. Human emotions are not always pure which may show a mixture of expressions. Hence compound expressions are crucial to express complex mind states of human beings.

Facial Expression Recognition is four stage problem. The first phase is Face Detection. In this phase human face is identified from an image or from a video. Second phase is Pre -Processing. Some actions like resizing, rotating, illumination correction etc. are done at this phase. Third phase is Feature Extraction. The relevant features are selected in this step. Three methods are commonly used for feature extraction [1].

1) Geometric Feature Extraction: This method focuses on shape and position of facial components like eyes, nose, mouth and eyebrows.
2) Appearance Based Feature Extraction: This method focuses on the texture and intensity changes of the face.
3) Action Unit Based Feature Extraction: This method involves decomposing facial muscle movements for recognizing expressions.

The last step of FER is Expression Classification. In this phase the relevant features are submitted to a pre-trained classifier and determines the emotional state. For this function deep learning models like CNN, SVM etc. are used Figure 1 shows the steps of a standard FER system.

From the literatures analysed, it is clear that many authors are extensively studies the problem of facial expression recognition. Most of them are focusing on basic expressions recognition and only few are focusing on the compound expression recognition.
The motive of [1-18] is to recognize basic expressions and [19-29] focuses on basic expression recognition in the presence of occlusion. Motive of [30-38] are to recognize compound expressions. The methods [39-65] are used to recognize micro expressions.

The main objectives of this study are to:
1) Provide an in-depth examination of various FER methods by analysing the workings of each method.
2) Provide a classification of different types of expressions
3) Provide a detailed analysis of datasets used in basic and compound FER techniques.
4) Address the applications and challenges involved in recent facial expression recognition research, such as occlusion, pose variation, illumination changes, and real-time constraints.

The rest of the paper is organized as follows: Section II contains a comprehensive analysis of various techniques used in the recent literatures for both basic and compound expressions. Section III describes various data structures used for basic expression recognition and compound expression recognition. Section IV analyses the evaluation metrics and tools used in the research works. Section V contains the challenges and future scope and finally section VI contains the conclusion.

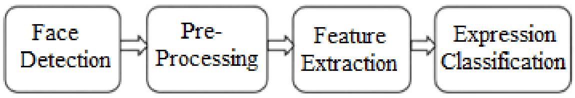

Figure 1: Steps of Facial Expression Recognition System

2. Review

Analysis of FER Methods

The basic expressions are happiness, surprise, sadness, anger, fear, disgust, and neutral. Many methods have been developed to improve accuracy and handle real-world challenges such as low-quality images, occlusion, and varied lighting. We next examine the diverse basic expression recognition methodologies in the literature. Xin Wang et al proposed a method to improve the accuracy of FER by a hierarchical deep neural network structure [1]. H Yu et al uses a Relational Convolutional Neural Network (ReCNN) in [2] and [3] uses a 3D vision Transformer with a deformable module for finding thew important regions. The methods [4-6] rely convolutional Neural Network (CNN) for FER. These methods mainly contributing to the reliability of dataset and hence improves the result. K Wang et al aims to improve facial expression recognition in real-world conditions, especially when faces are obscured, turned at different angles, or blurred in [7]. Method [8] Introduced a network called FERDERnet to address the challenges of real driving situations in emotion recognition. D Ruan, Y Yan et al introduces a Feature Decomposition and Reconstruction Learning (FDRL) method for facial expression recognition (FER) in [9]. The methods explained in [10-14] are using CNN for FER. Y Wang, Y Li et al explores the influence of activation functions in facial expression

recognition in [15]. The authors proposed a new piecewise activation function to address limitations in existing functions in [15]. Li et al introduces global convolution attention (GCA) block in [16] to rescale feature channels, and an intensity-aware loss (IAL) to guide the network in handling low-intensity expressions. Behaviour of students in the e-learning environment is analyzed and a novel pipeline based on video facial processing is proposed in [17]. Karnati Mohan et al proposed a method named as “FER-net: facial expression recognition using deep neural net” [18]. In this method, authors proposed a method to distinguish facial expressions efficiently with the help of the softmax classifier. In the field of basic expression recognition another important category is micro expression recognition. These are involuntary expressions and happen so quickly. Methods [39-59] focuses on micro expression recognition for basic expressions.

Facial expression recognition becomes a challenging task when there is partial occlusion. Many studies in facial expression recognition addressed the problem of occlusion. A significant amount of research is currently dedicated to mitigate partial occlusion in facial expression recognition. The study in [19] aims to detect facial expressions in the presence of Virtual Reality headsets. It is a challenging task because the upper half of face is fully covered. This method uses SVM based methods for object detection. D Poux et al uses generative algorithms in [20] to reconstruct occluded facial expressions and a de noising auto encoder to calculate optical flow between two occluded images. S. Selvi and M. Parvathy introduced method for improving FER in autism patients under the condition of partial occlusion in [21]. Ning Sun et al attempt to reduce the challenges of FER like face posture, illumination, and partial occlusion by using 3-D face feature reconstruction and learning [22]. G K Sahoo et al introduced a method to monitor driver behaviour from the front face images captured via dashboard camera and alert driver to improve their driving style in [23]. K Hu et al uses an occlusion detection module based on symmetric SURF in [24] to detect the occluded part. Qingyan Duan and Lei Zhang use a boosting GAN (BoostGAN) for FER under partial occlusion [25]. The method explained in [26] focuses on the occlusion where the user is wearing eye wear equipment that completely covers the area around their eyes such as glasses for visually impaired people or a head-mounted display in a virtual reality setting. The studies [26] and [28] uses transfer learning approach. The method in [27] addresses the need for lightweight deep learning models suitable for mobile devices with limited computing resources in facial expression recognition systems. It uses lightweight CNN models based on the Inverted Residual Bottleneck. Authors proposes a powerful deep learning approach to recognizing facial expressions in [29], even in real-world conditions where lighting, occlusion, and facial movement can make the task difficult. The two common methods to solve the problem of occlusion is either reconstruct the occluded parts or focusing on the visible region and feature learning.

Compound facial expression recognition focuses on identifying facial expressions that reflect a mix of two or more emotions. Recognizing them is a greater challenge due to the subtle and overlapping facial muscle movements involved. The scarcity of well-annotated datasets for compound emotions is another problem. To address these challenges, many researches are done in this field. In the next section, we explore various studies and approaches that have been proposed to identify and analyse compound facial expressions. The paper in [30] introduced a new model called MDFGAN (Multi Domain Fusion Generative Adversarial Network) to combine basic expressions and compound expressions and hence obtaining more diverse and realistic examples for training. R Dong and K-M Lam propose a new loss function called bicenter loss in [31]. Y Lv, Y Yan et al proposes a method in [32], in which basic expressions are learned first and new compound expressions are added over time. To address the challenge of learning new expressions without forgetting previously learned ones-known as the stability-plasticity dilemma, the authors proposed a method called Relationship Guided Knowledge Transfer (RGKT). The method explained in [33] recommend a method that combines spatial and frequency domain transformations. The four methods explained in [34-37] uses CNN. E Ryumina, M Markitantov et al uses Large Vision-Language Models (LVLMs) in [38]. A transfer learning approach for compound expression recognition using a fine-tuned VGG-19 pre-trained model is used in [39]. Y Zhang et al uses a deep convolutional generative adversarial network (DCGAN) and convolutional neural network (CNN) for compound expression recognition in [40]. Figure 2 demonstrates the general workflow of Facial Expression Recognition (FER) systems under two different conditions: FER without occlusion and FER with occlusion. In FER without occlusion, the entire facial region is clearly visible, allowing the system to perform standard processes such as image acquisition, face detection and alignment, preprocessing, feature extraction, and expression classification. Preprocessing techniques including resizing, normalization, and illumination correction are applied to improve image quality before extracting features using deep learning models such as Convolutional Neural Networks (CNNs) and Transformers. The extracted features are then classified using methods such as Softmax or Support Vector Machine (SVM) to recognize basic and compound facial expressions.

In contrast, FER with occlusion deals with partially hidden facial regions caused by masks, glasses, hands, hair, or other objects. Therefore, additional steps such as occlusion detection and occlusion handling are required before feature extraction. Techniques including region reconstruction, visible feature learning, and attention mechanisms help the model focus on the available facial information. Robust feature extraction methods using CNNs, GANs, and Transformers are employed to improve recognition accuracy under occluded conditions. Finally, the processed features are classified to achieve occlusion-robust facial expression recognition. The figure highlights the additional complexity and specialized processing required in occluded FER compared to conventional FER systems.

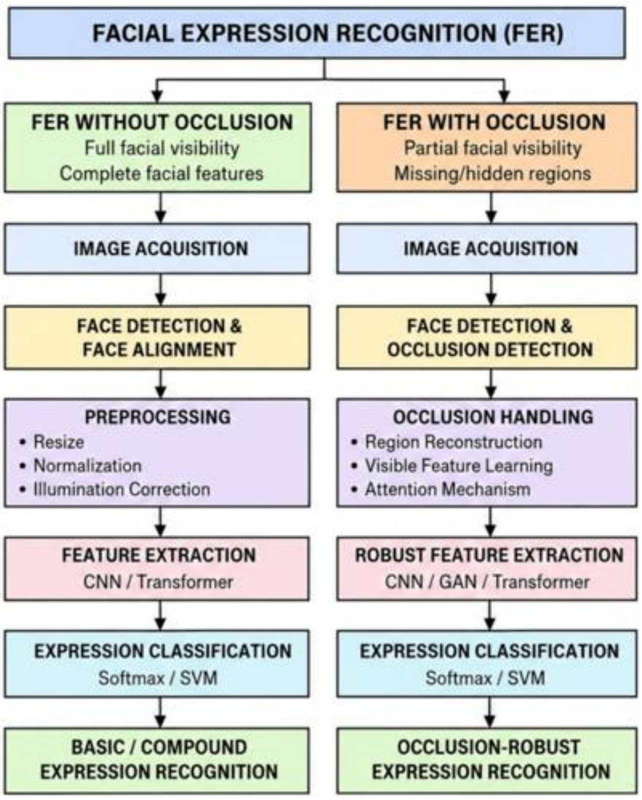

Figure 2: General workflow of FER with and without occlusion

In the next section, we explore various studies and approaches that have been proposed to identify and analyse basic and compound facial expression recognition methods. Some methods are facing the problem of Occlusion. Some approaches use convolutional neural networks (CNNs) for static images, while others use video-based methods or Transformers to capture relationships between facial regions. The commonly used techniques for FER are:

1) Convolutional Neural Network (CNN)

CNN can learn important facial features and recognize the expressions accurately. So Convolutional Neural Networks are commonly used for facial expression recognition. CNNs are capable of detecting facial patterns and muscle movements using convolutional layers. The following methods are using CNN for FER [1-6, 11-18, 21-23, 26-29, 33-37].

2) Transformer Based Methods

Transformers are able to model facial relationships using self-attention mechanism. Transformers capture global relationships of facial regions. The methods [3,7,8,10,32,38] and [39] are using Transformer based methods for FER.

3) Attention Mechanism

Attention mechanism helps to focus on more important regions of face for FER. The methods [6,12,21] and [29] are using attention mechanism.

4) Temporal or Dynamic Methods

Temporal or dynamic methods focus on how facial expressions change over time. These methods capture motion patterns and temporal relationships between consecutive facial frames and recognizes spontaneous expressions. Methods [16,14,20] uses this approach.

5) Transfer Learning

Transfer learning means using a model that is already trained on a large dataset and adapting it to recognize facial expressions. Instead of training a model from the beginning, the pre-trained model is used here and reuse previously learned facial features. Methods explained in [19,32,39] uses transfer Learning.

6) Large Vision Language Models (LVLMs)

Large Vision Language Models (LVLMs) can understand both images and text together. They learn to connect visual information with language meaning, allowing the model to describe images, answer questions about them, and understand visual concepts using words. E Ryumina, M Markitantov et al uses LVLMs in [38].

7) Generative Adversarial Networks (GANs)

GANs are capable of creating new, realistic data by learning from existing examples. In FER, GANs are mainly used for data augmentation, expression synthesis, and handling occlusion or pose variations. The research works [25,30] and [40] follows GAN model. Figure 3 shows a chart to analyse the usage of different techniques in FER. Table 1 shows the classification of different techniques used for FER and Table 2 shows the comparison of facial expression recognition methods for basic, compound, occluded, and non-occluded scenarios.

Table 1: Classification of Different Techniques Used for FER

Techniques	References
CNN Based	[1-6,11-18,21-23,26-29,33-37]
Transformer Based	[3,7,8,10,32,38,39]
Attention Mechanism	[6,12,21,29]
Temporal or Dynamic Methods	[14,16,20]
Transfer Learning	[19,32,39]
Large Vision Language Models (LVLMs)	[38]
Generative Adversarial Networks (GANs)	[25,30,40]

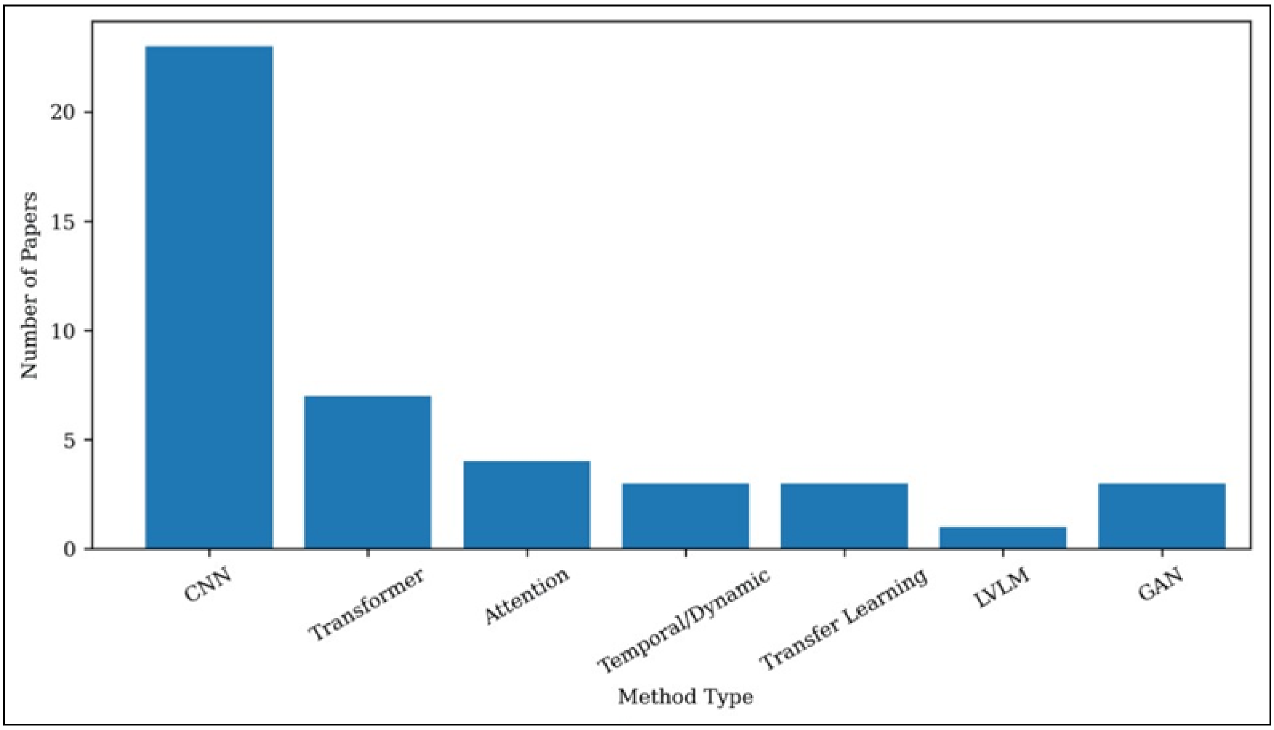

Figure 3: Distribution of methods used in FER Literatures

Table 2: Comparison of FER methods for basic, compound, occluded, and non-occluded scenarios

Category	Objective	Main Challenges	Common Techniques	Advantages	Limitations	Example Studies
Basic Facial Expression Recognition	Recognize standard emotions such as happiness, sadness, anger, fear, surprise, disgust, and neutral	Illumination changes, pose variation, lowresolution images	CNN, Transformer, Attention Mechanism, Transfer Learning	High accuracy on standard datasets, simpler classification task	Performance decreases in real-world conditions	[1–18]
Compound Facial Expression Recognition	Recognize combinations of multiple emotions	Subtle facial muscle variations, overlapping emotions, lack of datasets	CNN, GAN, Transformer, LVLM, Transfer Learning	Better understanding of complex human emotions	Difficult labeling and lower recognition accuracy	[30–40]
FER Under Occlusion	Recognize expressions when parts of the face are hidden	Masks, glasses, VR headsets, hand occlusion	GAN, Attention Mechanism, Reconstruction Models, CNN	Robust recognition in practical environments	Missing facial information reduces accuracy	[19–29]
FER Without Occlusion	Recognize expressions from fully visible faces	Illumination, pose variation	CNN, Vision Transformer, Feature Learning Models	Higher recognition performance	Limited robustness to occlusion	[1–18], [30–40]
Micro- Expression Recognition	Detect involuntary and brief facial expressions	Short duration, subtle movements	Temporal Models, Optical Flow, CNN, Transformer	Useful in psychology and security applications	Requires high frame-rate videos	[39–59]

3. Data Sets Used in Facial Expression Recognition

Several datasets are available for facial expression recognition. In this review we can see that some authors are using publicly available datasets such as FER2013. While some authors are using some private datasets which are created by themselves. Each dataset is different in terms of number of images, size of images and videos, illumination variation, variation in head posture etc. A variety of datasets support facial expression recognition research. While many studies utilize established public benchmarks such as FER2013, others utilize proprietary datasets. The following are some of the datasets used for Facial Expression Recognition.

CK+ dataset includes 593 sequences, with 327 labelled for seven basic emotions plus contempt [66] . JAFFE consists of 213 images of seven basic facial expressions posed by 10 Japanese women [67].FER-2013 includes 35,887 images labelled with seven basic facial expressions [68]. SFEW [69] comprises 1,766 images. RAF-DB [70] contains 29,672 images, with 15,339 labelled for basic expressions and compound expressions. AffectNet [71] is collected using 1,250 emotion-related keywords in six languages, consists of approximately 1 million images. MMI [72] includes 740 images and 2,900 video sequences depicting seven basic emotions, starting and ending with neutral expressions. FERV39k [73] is a largescale, multi-scene dataset contains 38,935 video clips labelled with 7 classic expressions across 22 fine-grained scenes in 4 isolated scenarios. DFEW [74] offers 12,059 clips from 16,372 clips annotated via a voting mechanism, segmented into five sections for cross validation. FERPlus [75] includes 28,709 training images, 3589 validation images, and 3573 test grayscale images. Multi-PIE face dataset [76] contains more than 750,000 images of 337 people recorded in up to four sessions over the span of five months. OuluCASIA [77] contains videos from 80 subjects under different lighting conditions, capturing six basic emotions. Karolinska Directed Emotional Faces (KDEF) [78] is a database of pictorial emotional facial expressions for use in emotion research. It contains a total of 490 JPEG pictures (72x72 dots per inch) showing 70 individuals (35 women and 35 men) displaying 7 different emotional expressions. The Table 3 shows the different datasets used in different facial expression recognition algorithms.

Different datasets used in the field of compound expression recognition are: CFEE (Compound Facial Expressions of Emotions) [79] dataset contains 1610 images for 230 subjects.CK+ dataset and RAF DB dataset are used to recognize compound expressions also. The EmotioNet dataset [80] about 1 million facial images collected from the web in unconstrained, real-world conditions. It includes approximately 950,000 automatically annotated images and 50,000 manually labelled images. The RAF Compound (Real-world Affective Face) [81] dataset contains a total of 29,672 facial images collected from the internet under real-world conditions. Table 4 shows different datasets used for compound expression recognition and Figure 4 shows overview of different datasets for both basic and compound expression recognition.

Table 3: Summary of datasets used for basic expression recognition

Dataset	References	Dataset	References
RAF-DB	[27]	CK+, JAFFE	[1]
RAF-DB, SFEW	[9]	CK+, MMI, FERA	[14]
AffectNet, RAF-DB	[2,4,5]	CK+, OULU-CASIA, RAF-DB	[29]
RAF-DB, AffectNet, FERPlus	[6,7]	FER2013	[26]
AffectNet, RAF-DB, SFEW	[10]	FERV39K	[3,16,17]
Multi-PIE, RAF-DB, AffectNet	[22]	FERPlus, CK+	[8]
FER2013, JAFFE, CK+, KDEF	[18]	LFW	[25]
Datasets created by authors	[12,13, 21,28]		

Table 4: Overview of Compound Facial Expression Recognition Methods

Dataset	References
CFEE, CK+	[30,35,39]
RAF DB, EMO- TIONET	[32,35]
Datasets Created by Authors	[33,36-38]

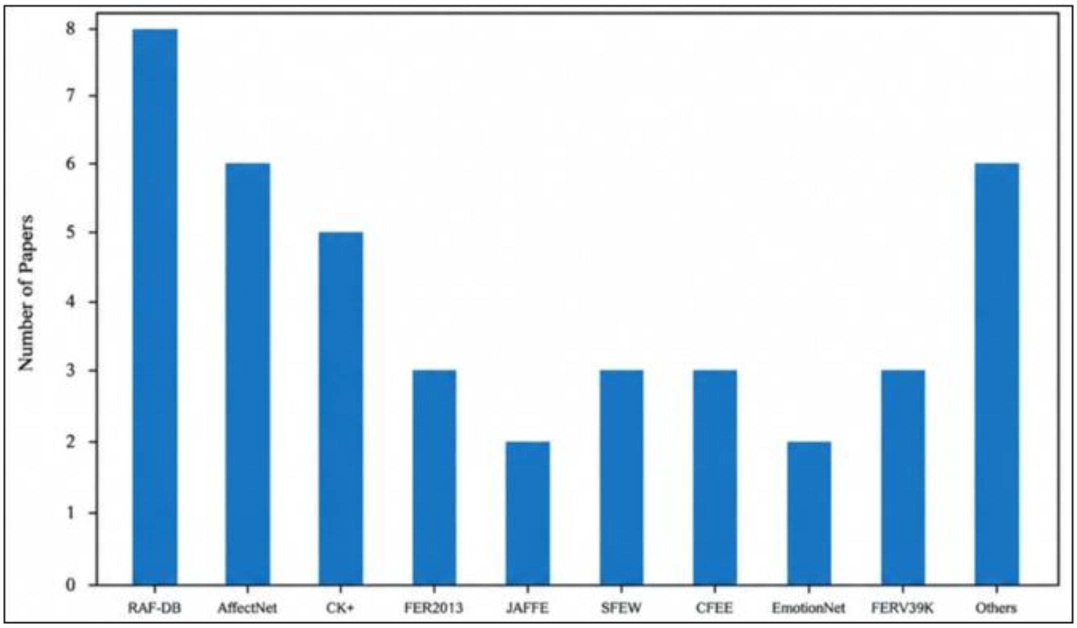

Figure 4: Overview of datasets for basic and compound expression recognition

4. Evaluation Metrics

Evaluation metrics are used to determine how accurately a model predicts outcomes by comparing its predicted results with the ground truth labels. Common metrics are explained below.

1) False Match Rate (FMR)
False Match Rate means how often a system incorrectly matches two different facial expressions as the same. Meaning of lower FMR is fewer mismatches.

FMR= False Matches/ Total Imposter Attempts

2) False Non-Match Rate (FNMR)
False Non-Match Rate measures how often the system fails to recognize the same facial expression. A high FNMR reduces system usability.

FNMR= (False Non-Matches)/ Total Genuine Attempts

3) Accuracy
Accuracy is the overall correctness of the FER system by measuring the proportion of correctly classified facial expressions out of all test samples.

Accuracy= (TP+ TN)/ (TP+ TN+FP+FN)

4) Genuine Accept Rate (GAR)
GAR measures the percentage of genuine expressions that are correctly accepted by the system. A higher GAR indicates better recognition.

GAR= 1- FNMR

5) Equal Error Rate (EER)
EER is the point where the False Match Rate equals the False Non-Match Rate. A lower EER indicates a more accurate and balanced system.

EER occurs When FMR= FNMR

6) Receiver Operating Characteristics (ROC):
The ROC curve is a graphical representation that shows the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR) at different decision thresholds.

TPR= TP/ (TP+ FN)

FPR= FP/ (FP+ TN)

7) Area Under the ROC Curve (AUC):
AUC indicates the probability that the system ranks a genuine expression higher than an imposter one. An AUC value closer to 1 indicates excellent performance, while 0.5 indicates random guessing.

$$AUC = \int_0^1 TPR\ (FPR) d(FPR)$$

5. Challenges and Future Scope

Facial Expression Recognition systems face several challenges due to the complex and dynamic nature of human emotions. Variations in pose, illumination, and occlusions such as masks, glasses, and hair significantly affect recognition accuracy. FER systems also struggle to recognize micro expressions and compound expressions, which are difficult to recognize in the presence of challenges like occlusion, varying lighting conditions. In addition, age, gender, ethnicity, and facial structure affects FER. Real-world FER further faces challenges from cross-resolution and cross-sensor data, as expressions captured using different cameras and conditions introduce noise and inconsistencies. Lack of datasets are also a significant challenge.

Future research in FER can focus on developing robust, pose-invariant, and occlusion-aware models, especially using 3D FER and multimodal approaches. Integrating generative AI and large vision-language models can help in realistic data augmentation, expression synthesis, and improved recognition of subtle and compound emotions. Exploring cross-age, cross-resolution, and cross-sensor FER will improve system reliability in real-world environments. Additionally, combining facial cues with speech, body gestures, and physiological signals, along with addressing bias, privacy, and ethical concerns, will lead to more accurate, fair, and human centred FER systems suitable for applications in healthcare, education, and human-computer interaction.

6. Conclusions

Facial Expression recognition is an important field of study due to its wide applications. This paper examined various FER techniques used for recognizing both basic and compound facial expressions, with particular focus on the challenge of facial occlusion. A comprehensive analysis of existing research revealed that occlusion remains one of the major factors affecting the accuracy and reliability of FER systems in real-world environments.

The review highlights different approaches proposed for handling occlusion, including deep learning methods, attention mechanisms, region reconstruction techniques, and robust feature extraction models such as CNNs, GANs, and Transformers. In addition, several widely used FER datasets were analysed to understand their role in evaluating model performance under occluded conditions. The study also showed that recent advancements in deep learning have significantly improved FER performance; however, challenges still exist in recognizing expressions under severe and dynamic occlusions.

Overall, this paper provides a detailed overview of FER systems and emphasizes the importance of developing more robust, accurate, and real-time occlusion-aware FER models. The findings of this review can serve as a useful reference for future research in improving facial expression recognition under challenging real-world conditions.

References

[1] P P R J H Kim, B G Kim and D M Jeonsg: Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. IEEE Access, 2019, 07:41273-41285, 10.1109/ACCESS.2019.2907327
[2] X W M J Y Xia, H Yu and F Y Wang, Relation-aware facial expression recognition. IEEE Transactions on

Cognitive and Developmental Systems, 2022, 14:1143-1154,. 10.1109/TCDS.2021.3100131

[3] Z Zhao and Q Liu, Former-dfer: Dynamic facial expression recognition transformer. 2021, 1553-1561, 10.1145/3474085.3475292

[4] H Vo, G-S Lee, H-J Yang, and S Kim, Pyramid with super resolution for in the-wild facial expression recognition. IEEE Access, 2020, PP:1-1, 10.1109/ACCESS.2020.3010018

[5] F Ma, B Sun, and S Li: Facial expression recognition with visual transformers and attentional selective fusion. IEEE Transactions on

[6] Affective Computing. 2023, 14:1236-1248. 10.1109/TAFFC.2021.3122146

[7] D Zeng, Z Lin, X Yan, Y Liu, F Wang, and B Tang, Face2exp: Combating data biases for facial expression recognition. 2022, 2025920268, 10.1109/CVPR52688.2022.01965

[8] K Wang, X Peng, J Yang, S Lu, and Y Qiao, Suppressing uncertainties for large-scale facial expression recognition. 2020, 6896-6905, 10.1109/CVPR42600.2020.00693

[9] H Xiao, W Li, G Zeng, et al.: On-road driver emotion recognition using facial expression. Applied Sciences, 2022, 12: 807, 10.3390/app12020807

[10] D Ruan, Y Yan, S Lai, Z Chai, C Shen, and H Wang: Feature decomposition and reconstruction learning for effective facial expression recognition. 2021, 7656-7665. 10.1109/CVPR46437.2021.00757

[11] Z Wen, W Lin, T Wang, and G Xu, Distract your attention: Multi-head cross attention network for facial expression recognition. Biomimetics, 2023, 8:199, 10.3390/biomimetics8020199

[12] A H Farzaneh and X Qi,: Facial expression recognition in the wild via deep attentive center loss. 2021, 2401-2410. 10.1109/WACV48630.2021.00245

[13] L Liang, C Lang, Y Li, S Feng, and J Zhao: Fine-grained facial expression recognition in the wild. IEEE Transactions on Information Forensics and Security. 2021, 16:482-494. 10.1109/TIFS.2020.3007327

[14] B Li and D Lima: Facial expression recognition via resnet-50. International Journal of Cognitive Computing in Engineering. 2021, 2:02. 10.1016/j.ijcce.2021.02.002

[15] D Jeong, B-G Kim, and S-Y Dong, Deep joint spatiotemporal network (djstn) for efficient facial expression recognition. Sensors, 1936, 20:10.3390/s20071936

[16] Y Wang, Y Li, Y Song, and X Rong: The influence of the activation function in a convolution neural network model of facial expression recognition. Applied Sciences, 2020, 10:1897. 10.3390/app10051897

[17] H Li, H Niu, Z Zhu, and F Zhao, Intensity-aware loss for dynamic facial expression recognition in the wild. 2023, 37:67-75,. 10.1609/aaai.v37i1.25077

[18] A V Savchenko, L V Savchenko, and I Makarov, Classifying emotions and engagement in online learning based on a single facial expression recognition neural network. IEEE Transactions on Affective Computing, 2022, 13:2132-2143, 10.1109/TAFFC.2022.3188390

[19] M Karnati, A Seal, O Krejcar, and A Yazidi, Fer-net: facial expression recognition using deep neural net. Neural Computing and Applications. 2021, 33:08. 10.1007/s00521-020-05676-y

[20] B Houshmand and N Mefraz Khan, Facial expression recognition under partial occlusion from virtual reality headsets based on transfer learning. 2020, 70-75. 10.1109/BigMM50055.2020.00020

[21] D Poux, B Allaert, N Ihaddadene, I M Bilasco, C Djeraba, and M Bennamoun: Dynamic facial expression recognition under partial occlusion with optical flow reconstruction. IEEE Transactions on Image Processing, 2022, 31:446-457, 10.1109/TIP.2021.3129120

[22] S Selvi and M Parvathy, Improving facial expression recognition for autism with idensenet-rcaformer under occlusions. International journal of developmental neuroscience: the official journal of the International Society for Developmental Neuroscience, 2024, 85:11. 10.1002/jdn.10391

[23] N Sun, J Tao, J Liu, H Sun, and G Han, 3-d facial feature reconstruction and learning network for facial expression recognition in the wild. IEEE Transactions on Cognitive and Developmental Systems, 2023, 15: 298- 309, 10.1109/TCDS.2022.3157772

[24] G K Sahoo, J Ponduru, S K Das, and P Singh, Deep leaning-based facial expression recognition in fer2013 database: An in-vehicle application. 2022, 1- 6, 10.1109/INDICON56171.2022.10040121

[25] K Hu, G Huang, Y Yang, C-M Pun, W-K Ling, and L Cheng, Rapid facial expression recognition under part occlusion based on symmetric surf and heterogeneous soft partition network. Multimedia Tools and Applications, 2020, 79:11. 10.1007/s11042-020-09566-2

[26] Q Duan and L Zhang, Look more into occlusion: Realistic face frontalization and recognition with boostgan. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32:214-228, 10.1109/TNNLS.2020.2978127

[27] N Petrou, G Christodoulou, K Avgerinakis, and P Kosmides: Lightweight mood estimation algorithm for faces under partial occlusion. 2023. 10.1145/3594806.3596553

[28] C Yoon and D Kim,: Mobile convolutional neural networks for facial expression recognition. 2020, 1315-1317, 10.1109/ICTC49870.2020.9289486

[29] P Lamba and D D Virmani, Dcnn-based facial expression recognition using transfer learning. 2021, 509-520, 10.1007/978-981-154992-2_48

[30] X Zhu, Z He, L Zhao, Z Dai, and Q Yang, A cascade attention based facial expression recognition network by fusing multi-scale spatio-temporal features. Sensors, 2022, 22:1350, 10.3390/s22041350

[31] S He, H Zhao, L Yu, et al.: Compound facial expression recognition with multi-domain fusion expression based on adversarial learning. 2022, 688-693, 10.1109/SMC53654.2022.9945308

[32] R Dong and K-M Lam, Bi-center loss for compound facial expression recognition. IEEE Signal Processing Letters, 2024, 31:641645, 10.1109/LSP.2024.3364055

[33] Y Lv, Y Yan, J-H Xue, S Chen, and H Wang, Relationship-guided knowledge transfer for class-incremental facial expression recognition. IEEE

Transactions on Image Processing, 2024, 33:2293-2304, 10.1109/TIP.2024.3374116 10.3390/s20174727

[34] J Hu, X Mou, B Li, and M Yan, Compound facial expression recognition based on graph convolutional multi-label learning. 2024, 433-437, 10.1109/EIECS63941.2024.10800267

[35] J. Yu, J. Zhu, W. Zhu, et al.: Multi model ensemble for compound expression recognition. 2024, 4873- 4879, 10.1109/CVPRW63382.2024.00491

[36] X Li, W Deng, S Li, and Y Li, Compound expression recognition in-the-wild with au-assisted meta multi-task learning. 2023, 57355744, 10.1109/CVPRW59228.2023.00608

[37] E Ryumina, M Markitantov, D Ryumin, H Kaya, and A Karpov: Zero-shot audio-visual compound expression recognition method based on emotion probability fusion. 2024, 4752-4760. 10.1109/CVPRW63382.2024.00478

[38] S. Ullah, J. Ou, Y. Xie, et al.: Compound facial expressions recognition approach using dcgan and cnn. Multimedia Tools and Applications, 2024, 83:85703-85723, 10.1007/s11042-024-20138-6

[39] W Kong, Z You, and X Lv,: 3d micro-expression recognition based on adaptive dynamic vision. Sensors, 2025, 25:3175, 10.3390/s25103175

[40] Y Zhang, W Lin, Y Zhang, J Xu, and Y Xu, Leveraging vision transformers and entropy-based attention for accurate microexpression recognition. Scientific Reports, 2025, 15:04. 10.1038/s41598-025-98610-y

[41] Y Zhang, X Xu, Y Zhao, Y Wen, Z Tang, and M Liu, Facial prior guided micro-expression generation. IEEE Transactions on Image Processing, 2024, 33:525-540, 10.1109/TIP.2023.3345177

[42] S. Zhao, H. Tang, S. Liu, et al: ME-PLAN: A deep prototypical learning with local attention network for dynamic micro expression recognition. Neural networks: The official journal of the International Neural Network Society, 2022, 153:427-443, 10.1016/j.neunet.2022.06.024

[43] Y Liu, Y Li, X Yi, Z Hu, H Zhang, and Y Liu,: Lightweight vit model for microexpression recognition enhanced by transfer learning. Frontiers in Neurorobotics, 2022, 16:922761, 10.3389/fnbot.2022.922761 Abstract

[44] T Xie, G Sun, H Sun, Q Lin, and X Ben, Decoupling facial motion features and identity features for micro-expression recognition. PeerJ Computer Science, 2022, 8: 1140, 10.7717/peerj-cs.1140

[45] Z Li, Y Zhang, H Xing, and K-L Chan, Facial micro-expression recognition using double-stream 3d convolutional neural network with domain adaptation. Sensors, 2023, 23:03. 10.3390/s23073577

[46] F Wu, Y Xia, T Hu, B Ma, J Yang, and H Li, Facial micro-expression recognition based on motion magnification network and graph attention mechanism. Heliyon, 2024-08, 10:35964, 10.1016/j.heliyon.2024.e35964

[47] M Yu, Z-q Guo, Y Yu, Y Wang, and S-x Cen,: Spatiotemporal feature descriptor for micro-expression recognition using local cube binary pattern. IEEE Access,. 2019, 7:1-1,. 10.1109/ACCESS.2019.2950339

[48] H Tian, W Gong, W Li, and Y Qian, Pastfnet: a paralleled attention spatiotemporal fusion network for micro-expression recognition. Medical biological engineering computing, 2024, 62:02. 10.1007/s11517-024-03041-y

[49] Z Wang, M Yang, Q Jiao, et al.: Two-level spatiotemporal feature fused two-stream network for micro-expression recognition. Sensors, 2024, 24: 1574,. 10.3390/s24051574

[50] C Lee, J Hong, and H Jung, N-step pre-training and d ́ecalcomanie data augmentation for micro-expression recognition. Sensors, 2022, 22:6671, 10.3390/s22176671

[51] Y Zhao, Z Chen, and S Luo, Micro-expression recognition based on pixel residual sum and cropped gaussian pyramid. Frontiers in Neurorobotics, 2021, 15:12. 10.3389/fnbot.2021.746985

[52] X Zhao, J Chen, T Chen, et al.: Micro-expression recognition based on nodal efficiency in the EEG functional networks. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2024, 32:887-894, 10.1109/TNSRE.2023.3347601

[53] Y Liu, Y Li, X Yi, Z Hu, H Zhang, and Y Liu, Micro-expression recognition model based on tv-l1 optical flow method and improved shufflenet. Scientific Reports, 2022, 12: 17522, 10.1038/s41598-022-21738-8

[54] A Muhammad Buhari, C Ooi, V Baskaran, R Phan, K Wong, and W H Tan, FACS-based graph features for real-time micro-expression recognition. Journal of Imaging, 2020, 6:130, 10.3390/jimaging6120130

[55] Y Li, X Huang, and G Zhao, Joint local and global information learning with single apex frame detection for micro-expression recognition. IEEE Transactions on Image Processing, 2021, 30:249-263, 10.1109/TIP.2020.3035042

[56] M Wei, Y Zong, X Jiang, C Lu, and J Liu, Micro-expression recognition using uncertainty-aware magnification-robust networks. Entropy. 2022-09, 24:1271. 10.3390/e24091271

[57] S Wang, S Guan, H Lin, J Huang, F Long, and J Yao, Micro-expression recognition based on optical flow and pcanet+. Sensors, 2022, 22:4296, 10.3390/s22114296

[58] H Jin, N He, Z Li, and P Yang: Micro-expression recognition based on multi-scale 3d residual convolutional neural network. Mathematical Biosciences and Engineering. 2024, 21:5007-5031. 10.3934/mbe.2024221

[59] C Zhu, M Yin, X Chen, J Zhang, and D Liu,: Ecological micro-expression recognition characteristics of young adults with subthreshold depression. PLOS ONE. 2019, 14:0216334. 10.1371/journal.pone.0216334

[60] Y Wang, Y Huang, C Liu, X Gu, D Yang, S Wang, and B Zhang, Micro expression recognition via dual-stream spatiotemporal attention network. Journal of Healthcare Engineering, 2021, 2021:1-10, 10.1155/2021/7799100

[61] H Zhou, S Huang, J Li, and S-J Wang, Dual-ATME: Dual-branch attention network for micro-expression recognition. Entropy, 2023, 25:460, 10.3390/e25030460

[62] Y Zheng and E Blasch: Facial micro-expression recognition enhanced by score fusion and a hybrid model from convolutional LSTM and vision transformer. Sensors, 2023, 23:06. 10.3390/s23125650
[63] H Pan, H Yang, L Xie, and Z Wang, Multi-scale fusion visual attention network for facial micro-expression recognition. Frontiers in Neuroscience, 2023, 17:1216181,. 10.3389/fnins.2023.1216181
[64] H Yang, L Xie, H Pan, C Li, Z Wang, and J Zhong: Multimodal attention dynamic fusion network for facial micro-expression recognition. Entropy, 2023, 25:1246. 10.3390/e25091246
[65] Y Zhang and C Zhu, The influence of face masks on micro-expression recognition. Behavioral Sciences (Basel), 2025-02, 15:200, 10.3390/bs15020200
[66] P Lucey, J F Cohn, T Kanade, J Saragih, Z Ambadar, and I Matthews: The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. 2010, 94-101, 10.1109/CVPRW.2010.5543262
[67] M Kamachi, M Lyons, and J Gyoba, The Japanese female facial expression (jaffe) database. Available: 1997,
[68] Y Khaireddin and Z Chen: Facial emotion recognition: State of the art performance on fer2013. 2021. 10.48550/arXiv.2105.03588
[69] A Dhall, R Goecke, S Lucey, and T Gedeon: Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark. 2011, 2106-2112. 10.1109/ICCVW.2011.6130508
[70] S Li, W Deng, and J Du: Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. 2017, 2852-2861. 10.1109/CVPR.2017.277
[71] A Mollahosseini, B Hasani, and M H Mahoor: AffectNet: A database for facial expression, valence, and arousal computing in the wild. IEEE Transactions on Affective Computing (TAFFC). 2017, 10:18-31. 10.1109/TAFFC.2017.2740923
[72] M Valstar, M Pantic: Induced disgust, happiness and surprise: An addition to the mmi facial expression database. 2010, 10:65.
[73] Y Wang, Y Sun, Y Huang, et al.: FERV39k: A large-scale multi-scene dataset for facial expression recognition in videos. 2022, 2089020899, 10.48550/arXiv.2203.09463
[74] X Jiang, Y Zong, W Zheng, C Tang, W Xia, C Lu, and J Liu: DFEW: A largescale database for recognizing dynamic facial expressions in the wild. 2020, 2881-2889,. 10.1145/3394171.3413620
[75] Jie Song, Mengqiao He, Jinhua Feng, et al.: Bridging the Gaps: Utilizing Unlabeled Face Recognition Datasets to Boost SemiSupervised Facial Expression Recognition. arXiv. 2024, 10.48550/arXiv.2410.17622
[76] R Gross, I Matthews, J Cohn, T Kanade, and S Baker: Multi-pie. 2008, 1-8, 10.1109/AFGR.2008.4813399
[77] G Zhao, X Huang, M Taini, S Z Li, and M Pietika¨inen: Facial expression recognition from near-infrared videos. Image and Vision Computing, 2011, 29:607-619,. 10.1109/ICPR.2008.4761697
[78] D Lundqvist, A Flykt, and A Ohman, Karolinska directed emotional faces . 1998. 10.1037/t27732-000
[79] V Mavani, S Raman, and K P Miyapuram: Facial expression recognition using visual saliency and deep learning. 2017, 2783-2788. 10.48550/arXiv.1708.08016
[80] C F Benitez-Quiroz, R Srinivasan, and A M Martinez: Emotionet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild. 2016, 5562-5570. 10.1109/CVPR.2016.600
[81] S. Li, W. Deng, J. Du,: Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition. IEEE Transactions on Image Processing. 2019, 28:356-370. 10.1109/TIP.2018.2875462.