International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Survey on Facial Emotion Detection

Siyona Ratheesh[1],Dr. Kala K. U.[2]

[1]Student, Jain (Deemed-to-be) University, Bangalore, rsiyona27@gmail.com
[2]Assistant Professor, Jain (Deemed-to-be) University, Bangalore, kalasarin@gmail.com

ABSTRACT:

Recognizing facial expressions is inherently challenging due to the unique and varied features of human faces. While humans can easily identify emotions from facial cues, current technologies often fall short in accurately distinguishing these emotions. This survey paper provides a thorough review of the existing models and techniques used in facial emotion recognition, focusing on key emotions such as sadness, happiness, disgust, surprise, fear, and anger. We examine both traditional methods, including feature-based approaches and classical machine learning algorithms, as well as advanced deep learning techniques like convolutional neural networks and recurrent neural networks.

The paper also explores the various datasets that are instrumental in training and evaluating FER systems, such as CK+, FER2013, and JAFFE. Applications of FER in fields such as human-computer interaction, healthcare, marketing, security, and entertainment are discussed to demonstrate the broad impact and utility of this technology. Furthermore, we address the significant technical challenges, including real-time processing, handling occlusions, and variations in lighting and facial poses, as well as ethical and privacy concerns related to the deployment of FER systems.

By comparing and contrasting different strategies, this paper aims to provide a comprehensive overview of the current state of FER technology, highlighting its achievements and limitations, and suggesting potential directions for future research and development. This survey serves as a valuable resource for researchers and practitioners seeking to understand and advance the field of facial emotion recognition.

Keywords: – Emotion Detection, FER (Facial Emotion Recognition), CNN (Convolutional Neural Network), SVM (Support Vector Machine),Machine Learning, Deep Learning

1. Introduction :

One of the various facial recognition technologies that have evolved and expanded throughout time is emotion recognition. Currently, a certain program can analyse and evaluate a person's facial emotions thanks to the use of facial emotion recognition software. This software mimics the way a human brain works by using sophisticated picture dispensation, which enables it to recognise emotions as well.

AI-based systems used for identification and verification include facial recognition and emotion detection. Thanks to artificial intelligence, these technologies are able to recognise voices, faces, and emotions. Each year, advances in facial emotion recognition technology are made. The AI utilised analyses facial expressions based on a variety of variables to determine what emotion the person is expressing.

Human emotions are recognised and examined using emotion detection. This software mimics the way a human brain works by using sophisticated picture dispensation, which enables it to recognise emotions as well. By obtaining details about someone's emotional state, it is possible to determine how they are feeling. AI-based technologies used for verification and recognition include facial recognition and emotion detection.

Factors like the position of the eyes and brows, the mouth's position, distinct changes in the features of the face

The purpose of this paper is to provide readers with a clear understanding of emotions and the various ways that a system could be able to perceive them. When computers are utilised solely as instruments for computational purposes, it is not necessary to be able to recognise emotions.[1] However, it is now crucial to teach the system how to understand the user. This can lessen the user's dissatisfaction and create resources for the development of social and emotional skills.

2. Literature Review

In the past few years, several papers have successfully solved the problem of facial expression recognition So far, several tasks have been completed for real- time emotion classification. So, some state of artwork relevant to the proposed work is been discussed in this section.

Table 1 - Comparisons among different approaches and methods

Year	Author	Approach	Outcome	Limitation
2022 [2]	Md. Jashim Uddin et al.	Convolutional Neural Network	95%-IMDB-WIKI 66% - FER	Misclassifications, such as predicting angry instead of disgust and sad instead of fear.
2020 [3]	Akriti Jaiswal et al.	CNN	70.14 - FERC-2013 98.65 - JAFFE	large amounts of high-quality and quantitative training data FERC-2013 dataset makes it harder to interpret the images
2020 [4]	Md. Forhad Ali et al.	CNN and Viola-Jones algorithm SVM, K-means clustering, Haar features	Upto 97% accuracy. 93% accuracy on validation set.	limited quantity of training data available for developing a comprehensive framework for facial emotion detection
2020 [5]	Naveen Kumar H et al.	AFER, Viola-Jones algorithm, (MSVM)	94.2% and 93.7% on CK+ and KDEF.	Increase in training time due to block processing employed for texture feature extraction
2020 [6]	Bei Pan et al.	Genetic algorithms and Extreme Learning Machine	93.53% on CK+, 91.62% on Enterface05, and 60.77% on BAUM-1s	Genetic algorithms may risk overfitting, rely on trial and error, and require significant training data.
2023 [7]	Foo Jia Ming et al	Uses computer vision and machine learning in Python with cv2, following Rapid Application Development methodology.	The system achieved an accuracy of 60% for detecting mental stress.	Affected by cultural differences, small sample size, difficulty with subtle emotions, low-light conditions, slow response time, and low accuracy.
2021 [8]	Rupali Gill et al	Deep learning model for facial emotion recognition using CNN. Utilized three facial datasets for emotion recognition.	CNN model achieved 93% accuracy in facial emotion recognition. VGGNet16 had 67.1% accuracy, MobileNet had 47.9% accuracy.	Limited dataset size for training. Dependency on controlled environment for dataset collection.
2021 [9]	Sergio Pulido-Castro et. al.	Ensemble of machine learning models for facial emotion recognition. Improved recognition accuracy by separating emotions and selecting relevant features.	Sensitivity computed for each emotion using machine learning models. ANNs and RFs showed the best accuracy among tested models.	Reduced accuracy with fewer features. Separation of labels may affect sensitivity of similar emotions.
2021 [10]	H. Arabian et al	Localized region selection for FER modeling using Viola-Jones object detection algorithm. Utilized 3 classifiers: 2 K-Nearest Neighbor and 1 SVM.	Model accuracy reached up to 98.44% with small variations. EUC1 model showed mean accuracy of 38.92 - 1.50. Statistical data showed EUC1 model performed best with over 98% accuracy.	Limited robustness due to low accuracy in some emotional classes. Model performance not tested on subjects outside specific databases.
2023 [11]	Hien Le Nhu et al	Deep learning model with weighted face regions for emotion recognition. CNN trained on facial regions to predict emotions accurately. Utilizes FER2013 dataset to demonstrate method's accuracy.	Proposed method accuracy: 74.14% outperforms other methods. VGGNet accuracy: 73.28% on FER2013 dataset.	VGGNet model has simpler architecture, may not capture useful information. future work includes using deeper CNN architectures like VGG19.
2022 [12]	Nikolai Smirnov et al	Utilized anthropometric points for feature vectors in emotion recognition. Employed various machine learning and deep learning	LinearSVC achieved an F1 score of 0.83 in emotion recognition.	Limited discussion on real-time application feasibility. Lack of comparison with state-of-the-art emotion recognition models.

		methods for classification tasks.		
2021 [13]	Özay Ezerceli et al	CNN-based FER system with supportive techniques for high accuracy. Data quality, visualization, data augmentation, and CNN layers crucial.	State-of-the-art model achieved 93.70% accuracy on FER2013 dataset.	Subjectivity, occlusion, pose affect model accuracy. Limitations include low resolution, scale, illumination variation in images.
2023 [14]	Wanzeng Kong et al	Brain-machine coupled learning method for cognitive and visual knowledge integration. Feature-mapping-based approach to bridge cognitive and visual domains.	Visual common representation accuracy reached 96.55% using EEG signals. CMD metric used for best performance in emotion recognition.	Limited generalization due to few samples.Challenges in imitating human brain cognitive processes.
2021 [15]	Hidangmayum Bebina et al	Fusion model combines VGG-19, Densenet-121, and Sumnet for FER Regularization techniques and fusion loss enhance model optimization during training. Baseline model fine-tuned Densenet-161 using transfer learning on KDEF datasets.	Fnet model achieved 98.98% accuracy in facial emotion recognition. VGG-19 model achieved 97.75% accuracy with regularization. Densenet-121 model achieved 98.89% accuracy with regularization.	Limited discussion on model generalization. Lack of comparison with other fusion techniques.
2020 [16]	Ninad Mehendale	FERC model based on two-part CNN on Caltech faces, CMU database, and NIST database	Achieved 96% accuracy in emotion detection	Challenges with orientation, shadow detection, and multiple faces in images. Additionally, issues with facial hair and high computing power during CNN tuning were noted.
2020 [17]	Wafa Mellouk et al	CNN and CNN-LSTM architectures, to enhance performance. Databases like CK+, JAFFE, and AffectNet have been used for training models.	Significant progress has been made, with high accuracy rates exceeding 90% achieved in many studies.	Variations in head pose, lighting conditions, occlusions, and background interference. prompting the need for larger databases and more complex deep learning architectures to recognize a broader range of emotions accurately.
2021 [18]	Yousif Khaireddin et al	VGGNet model for facial emotion recognition on the FER2013 dataset	Accuracy of 73.28%, surpassing previous single-network accuracies and enhancing facial emotion recognition capabilities.	Under naturalistic conditions due to high intra-class variation and subtle differences between expressions, which may impact model performance.
2019 [19]	Milad Mohammad Taghi Zadeh et al	Gabor filters for feature extraction and a convolutional neural network	Boosted the training speed and accuracy of the convolutional neural network. 91% accuracy, outperforming non-Gabor filter systems at 82%.	Reliance on specific databases like JAFFE, which may not fully represent real-world scenarios.
2022 [20]	Mr. Rohan Appasaheb Borgalli et al	A custom CNN architecture for facial emotion recognition, training it on FER13, CK+, and JAFFE datasets using K-fold cross-validation.	accuracy rates of 86.78% for FER13, 92.27% for CK+, and 91.58% for JAFFE, showcasing the effectiveness of the custom CNN model.	Further improvement in accuracy through techniques like unsupervised pre-training, dataset balancing.
2020 [21]	Dr. Shaik Asif Hussain et al	Using deep learning algorithms and CNN models	An accuracy of 88% in facial emotion classification, showcasing the effectiveness of the proposed model.	Further refinement in handling variations in lighting conditions and diverse facial expressions to enhance the system's robustness

and accuracy.

In recent years, numerous papers have made significant advancements in solving the problem of facial expression recognition, particularly for real-time emotion classification. Researchers have extensively utilized Convolutional Neural Network (CNN) models to achieve a more accurate representation of facial data. CNNs have proven to be highly effective due to their ability to automatically and adaptively learn spatial hierarchies of features from input images, which is crucial for accurately identifying subtle variations in facial expressions.

Despite these advancements, the task of facial expression recognition is not without its challenges. One common issue faced by many researchers is the misclassification of emotions, such as confusing the emotion of fear with sadness. This is often due to the subtle and nuanced differences between certain emotional expressions, which can be difficult for models to distinguish accurately.

To address these challenges, various datasets have been employed to train and evaluate the performance of emotion recognition models. Some of the most commonly used datasets include FER (Facial Expression Recognition), JAFFE (Japanese Female Facial Expression), and CK+ (Extended Cohn-Kanade). These datasets provide a diverse range of facial expressions under different conditions, thereby helping to improve the robustness and generalization of the models.

In addition to CNNs, other machine learning techniques like Support Vector Machines (SVM) and Principal Component Analysis (PCA) have been utilized to further enhance the accuracy of emotion recognition systems. These methods often complement CNNs by providing additional feature extraction and classification capabilities. For instance, SVMs are effective in creating decision boundaries that can help distinguish between different emotional states, while PCA can reduce the dimensionality of the data, making it easier for the models to process and learn from the data.

Moreover, hybrid approaches combining CNNs with SVMs or PCA have shown promising results. These approaches leverage the strengths of each method to achieve higher accuracy and better performance in emotion classification tasks. For example, after the initial feature extraction by a CNN, an SVM can be used for the final classification, thereby improving the overall accuracy of the system.

Furthermore, advanced techniques such as data augmentation, transfer learning, and ensemble learning have been employed to enhance the performance of emotion recognition models. Data augmentation helps in increasing the diversity of the training data, thus making the models more robust to variations in facial expressions. Transfer learning allows models to leverage pre-trained weights from large-scale datasets, improving their performance even with limited labeled data. Ensemble learning combines predictions from multiple models to achieve better generalization and accuracy.

Overall, the continuous evolution of these techniques and the integration of new methodologies have significantly improved the capability of facial expression recognition systems, making them more accurate and reliable for real-time emotion classification. The ongoing research and development in this field hold great promise for future applications in various domains, including human-computer interaction, psychological analysis, and security systems.

3. Facial Emotions

The simplest and oldest method of reading emotions is probably based on facial expressions. It has a lengthy history since it is thought that they allude to the crucial features of emotions. Therefore, a variety of methods for classifying emotions based on facial expressions have been developed. The fundamental tenet of each of these methods is that certain facial points and areas can be spatially positioned to reveal emotional cues. [22]. There are numerous distinguishable emotions, for classifying the facial emotions a major six emotions are taken. It is Happiness, Sadness, Anger, Disgust, Fear, Surprise.

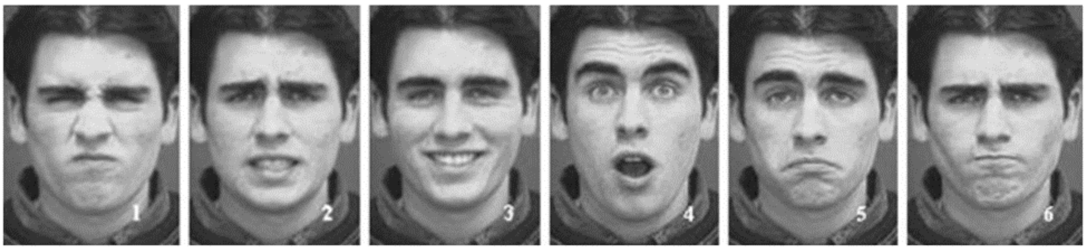

Fig. 1 - Facial Emotions [23]

Machine learning techniques such as Support Vector Machines (SVM) and Principal Component Analysis (PCA) have been integrated with CNNs to enhance the accuracy of emotion classification. SVMs are used for creating decision boundaries to distinguish between different emotional states, while PCA is employed for dimensionality reduction, making the models more efficient and effective.

Datasets like FER (Facial Expression Recognition), JAFFE (Japanese Female Facial Expression), and CK+ (Extended Cohn-Kanade) have been instrumental in training and validating these models. These datasets contain a wide variety of facial expressions captured under different conditions, which helps improve the robustness and generalization of the emotion recognition systems.

4. Dataset

Multiple datasets have been created by experts from various organisations to assess reported facial expression classification algorithms. Some of these datasets which are commonly used to analyse emotions includes:

1. FER 2013

Grayscale portraits of faces measuring 48 × 48 pixels make up the data. The faces have been automatically registered such that each one is roughly in the same location and takes up a similar amount of space. Each face must be assigned to one of seven categories, with 0 denoting anger, 1 disgust, 2 fear, 3 happiness, 4 sadness, 5 surprise, and 6 neutrality. The public test set has 3,589 cases, whereas the training set has 28,709 examples.

Fig. 2 - FER 2013 Dataset [24]

2. CK+ Cohen Kanade Dataset

The Cohen-Kanade (CK+) dataset consists of 593 video clips from 123 distinct people who range in age from 18 to 50 and are of various genders and ethnic backgrounds. Each video depicts a change in face expression from neutral to a specified peak expression. It was shot at 30 frames per second at either 640x490 or 640x480 resolution.

One of the seven expression classes—anger, contempt, disgust, fear, pleasure, sorrow, and surprise—is assigned to 327 of these movies. The majority of facial expression classification algorithms employ the CK+ database, which is widely recognised as the most frequently used laboratory-controlled facial expression classification database currently available.

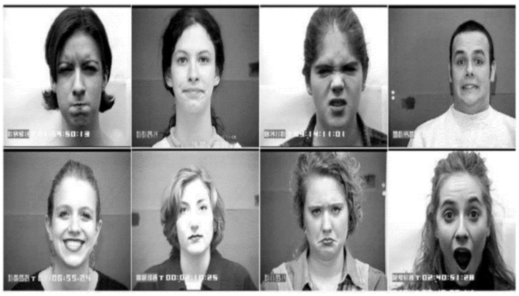

Fig. 3 - CK+ Dataset [25]

3. IMDB-WIKI Dataset

This is the largest publicly accessible dataset of face photos with gender and age labels that we are aware of. It is mainly used for age and gender prediction.

Fig. 4 - IMBD-WIKI Dataset [26]

5. Approach

There are various approaches in recognizing facial emotions from an image or a video. Some of the important techniques are

1. **Convolutional Neural Networks:**

Convolutional Neural Networks (CNNs) are among the most widely used approaches due to their ability to automatically learn and extract spatial hierarchies of features from facial images. CNNs consist of multiple layers, including convolutional layers that detect features such as edges and textures, pooling layers that reduce dimensionality, and fully connected layers that perform classification. This architecture allows CNNs to accurately capture the complex patterns associated with different emotions in facial expressions.

2. **Support Vector Machine:**

Support Vector Machines (SVMs) are another powerful technique used in facial emotion detection. SVMs work by finding the optimal hyperplane that separates different classes in a high-dimensional space. When combined with CNNs, SVMs can be used as the final classifier to improve the accuracy of emotion detection. For instance, after feature extraction by CNN layers, SVMs can classify the emotions based on these features, effectively handling the variations and subtleties in facial expressions.

3. **Principal Component Analysis:**

Principal Component Analysis (PCA) is often employed in preprocessing stages to reduce the dimensionality of facial image data. By transforming the data into a set of orthogonal components that capture the most variance, PCA helps in eliminating noise and focusing on the most critical features necessary for emotion detection. This reduction in dimensionality not only speeds up the training process but also enhances the performance of classifiers like SVMs by focusing on the most relevant features.

4. **Recurrent Neural Networks (RNNs) and Long Short-Term Memory Networks (LSTMs)**

RNNs and LSTMs are particularly effective in handling sequential data. While CNNs excel at capturing spatial features, RNNs and LSTMs can capture temporal dynamics, making them suitable for analyzing sequences of facial expressions over time. This is particularly useful for video-based emotion recognition where understanding the progression of emotions is crucial.

5. **Haar Cascades and Viola-Jones Algorithm**

The Viola-Jones algorithm, which uses Haar-like features, is a classical method for face detection. This method involves training a cascade function that is capable of detecting faces in real-time. Once the face is detected, other machine learning techniques can be used to analyze facial expressions. This method is fast and efficient but may not be as accurate as deep learning methods.

REFERENCES :

[1] Sarprasatham, Matilda. (2015). “Emotion Recognition: A Survey. International Journal of Advanced Research in Computer Science.” 3. 14-19.

[2] Md. Jashim Uddin, Dr. Paresh Chandra Barman, Khandaker Takdir Ahmed S.M. Abdur Rahim, Abu Rumman Refat, Md Abdullah-AlImran - A Convolutional Neural Network for Real-time Face Detection and Emotion & Gender Classification” IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) (2020).

[3] Jaiswal, Akriti, A. Krishnama Raju, and Suman Deb. "Facial emotion detection using deep learning." 2020 international conference for emerging technology (INCET). IEEE, 2020.

[4] Ali, Md Forhad, Mehenag Khatun, and Nakib Aman Turzo. "Facial emotion detection using neural network." the international journal of scientific and engineering research (2020).

[5] Kumar HN, Naveen, et al. "Automatic facial expression recognition combining texture and shape features from prominent facial regions." IET Image Processing 17.4 (2023): 1111-1125.

[6] Pan, Bei, et al. "A review of multimodal emotion recognition from datasets, preprocessing, features, and fusion methods." Neurocomputing (2023): 126866.

[7] Ming, Foo Jia, et al. "Facial Emotion Recognition System for Mental Stress Detection among University Students." 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). IEEE, 2023.
[8] Gill, Rupali, and Jaiteg Singh. "A deep learning approach for real time facial emotion recognition." 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART). IEEE, 2021.

[9] Pulido-Castro, Sergio, et al. "Ensemble of Machine Learning Models for an Improved Facial Emotion Recognition." 2021 IEEE URUCON. IEEE, 2021.

[10] Arabian, Herag, et al. "Facial Emotion Recognition Focused on Descriptive Region Segmentation." 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2021.

[11] Le Nhu, Hien, Huy Vu Dang, and Huan Hoang Xuan. "Facial Emotion Recognition by Combining Deep Learning and Averaged Weight of Face-Regions." 2023 15th International Conference on Knowledge and Systems Engineering (KSE). IEEE, 2023.

[12] Smirnov, Nikolai V., and Alexander S. Chernyshov. "Emotion recognition from facial images." 2022 International Russian Automation Conference (RusAutoCon). IEEE, 2022.

[13] Ezerceli, Özay, and M. Taner Eskil. "Convolutional neural network (cnn) algorithm based facial emotion recognition (fer) system for fer-2013 dataset." 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). IEEE, 2022.

[14] Kong, Wanzeng, et al. "Causal graph convolutional neural network for emotion recognition." IEEE Transactions on Cognitive and Developmental Systems (2022).

[15] Devi, Hidangmayum Bebina, and Tushar Sandhan. "Regularized Fusion of CNN with Transfer Learning for Facial Emotion Recognition." 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2023.

[16] Mehendale, Ninad. "Facial emotion recognition using convolutional neural networks (FERC)." SN Applied Sciences 2.3 (2020): 446.

[17] Mellouk, Wafa, and Wahida Handouzi. "Facial emotion recognition using deep learning: review and insights." Procedia Computer Science 175 (2020): 689-694.

[18] Khaireddin, Yousif, and Zhuofa Chen. "Facial emotion recognition: State of the art performance on FER2013." arXiv preprint arXiv:2105.03588 (2021).

[19] Zadeh, Milad Mohammad Taghi, Maryam Imani, and Babak Majidi. "Fast facial emotion recognition using convolutional neural networks and Gabor filters." 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI). IEEE, 2019.

[20] Borgalli, Mr Rohan Appasaheb, and Sunil Surve. "Deep learning for facial emotion recognition using custom CNN architecture." Journal of Physics: Conference Series. Vol. 2236. No. 1. IOP Publishing, 2022.

[21] Hussain, Shaik Asif, and Ahlam Salim Abdallah Al Balushi. "A real time face emotion classification and recognition using deep learning model." In Journal of physics: Conference series, vol. 1432, no. 1, p. 012087. IOP Publishing, 2020.

[22] Ralph Adolphs, ‗"Recognizing Emotion From Facial Expressions: Psychological and Neurological Mechanisms" ‘, University of Iowa College of Medicine

[23] Kanade, T., Cohn, J., Tian, Y.-L.: "Comprehensive database for facial expression analysis. In: Proceedings of International Conference on Face and Gesture Recognition", pp. 46–53 (2000)

[24] Zafar, Sahar & Ali, Fayyaz & Guriro, Subhash & Ali, Irfan & Khan, Asif & Zaidi, Adnan. (2019). Facial Expression Recognition with Histogram of Oriented Gradients using CNN. Indian Journal of Science and Technology.

[25] Wafi, Muhammad & Bachtiar, Fitra & Utaminingrum, Fitri. (2023). Feature extraction comparison for facial expression recognition using adaptive extreme learning machine. International Journal of Electrical and Computer Engineering (IJECE).

[26] Rasmus Rothe, Radu Timofte, and Luc Van Gool. 2015. DEX: Deep expectation of apparent age from a single image.In IEEE International Conference on Computer Vision Workshops (ICCVW).