

Human Emotion Detection Using DeepFace and Artificial Intelligence [†]

Ramachandran Venkatesan ¹, Sundarsingh Shirly ¹, Mariappan Selvarathi ² and Theena Jemima Jebaseeli ^{1,*}

¹ Division of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore 641114, India; rlvenkei_2000@karunya.edu (R.V.); shirly@karunya.edu (S.S.)

² Department of Mathematics, Karunya Institute of Technology and Sciences, Coimbatore 641114, India; selvarathi@karunya.edu

* Correspondence: jemima_jeba@karunya.edu

[†] Presented at the International Conference on Recent Advances in Science and Engineering, Dubai, United Arab Emirates, 4–5 October 2023.

Abstract: An emerging topic that has the potential to enhance user experience, reduce crime, and target advertising is human emotion recognition, utilizing DeepFace and Artificial Intelligence (AI). The same feeling may be expressed differently by many individuals. Accurately identifying emotions can be challenging, in light of this. It helps to understand an emotion's significance by looking at the context in which it is presented. Depending on the application, one must decide which AI technology to employ for detecting human emotions. Because of things like lighting and occlusion, using it in real-world situations can be difficult. Not every human emotion can be accurately detected by technology. Human–machine interaction technology is becoming more popular, and machines must comprehend human movements and expressions. When a machine recognizes human emotions, it gains a greater understanding of human behavior and increases the effectiveness of work. Text, audio, linguistic, and facial movements may all convey emotions. Facial expressions are important in determining a person's emotions. There has been little research undertaken on the topic of real-time emotion identification, utilizing face photos and emotions. Using an Artificial Intelligence-based DeepFace approach, the proposed method recognizes real-time feelings from facial images and live emotions of persons. The proposed module extracts the facial features from an active shape DeepFace model by identifying 26 facial points to recognize human emotions. This approach recognizes the emotions of frustration, dissatisfaction, happiness, neutrality, and wonder. The proposed technology is unique, in that it implements emotion identification in real-time, with an average accuracy of 94% acquired from actual human emotions.

Keywords: emotion detection; age prediction; gender prediction; race prediction; DeepFace; deep learning



Citation: Venkatesan, R.; Shirly, S.; Selvarathi, M.; Jebaseeli, T.J. Human Emotion Detection Using DeepFace and Artificial Intelligence. *Eng. Proc.* **2023**, *59*, 37. <https://doi.org/10.3390/engproc2023059037>

Academic Editors: Nithesh Naik, Rajiv Selvam, Pavan Hiremath, Suhas Kowshik CS and Ritesh Ramakrishna Bhat

Published: 12 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Human–machine contact is becoming more popular in modern technology and machines must comprehend human movements and emotions. If a machine recognizes human emotion, it can comprehend human behaviors and notify the person who utilizes them to identify one's feelings, thereby enhancing work efficiency. Emotions are strong sentiments that impact daily activities such as making decisions, memory, concentration, inspiration, dealing, understanding, organizing, thinking, and a lot more [1–4]. Albert Mehrabian discovered, in 1968, that in person-to-person interactions, verbal indicators account for 7% of all interactions, vocal indications account for 38%, and facial reactions account for 55% [5]. As a result, one of the most significant components of emotion identification is facial expression analysis. Although facial expression recognition from 2D photographs is a well-known issue, a real-time strategy for predicting characteristics, regardless of

poor-quality images, is lacking. More research is needed on non-frontal photographs with shifting lighting scenarios, since these global settings are not constant in real-time and visual expressions may all be utilized to recognize emotions [6,7].

The technique of detecting people's emotions is known as emotion recognition. The precision with which people identify the emotions of others differs greatly [8]. Deep learning and artificial intelligence are used to help humans with emotion identification, which is a relatively new study topic. Researchers have been interested in automatically identifying emotions since ancient times [9]. At the moment, emotion detection is accomplished via recognizing facial expressions in images and videos, evaluating speeches in audio recordings, and analyzing social media information. For emotion recognition, physiological signal measurements such as brain signals and ECG body temperature, as well as artificial intelligence algorithms, are emerging [10].

Deep learning may be used in a marketing effort, to target adverts to clients who are likely to possess a passion for the good or service that is being marketed. This may serve to enhance sales while boosting the performance of the marketing strategy. Deep learning may be used by a security system to recognize distressed clients [11]. Marketing and advertising businesses seek to know the emotional reactions of customers to adverts, designs, and products [12]. Education applications include tracking the responses of students for engagement and interest in the topic. Also, another application is the use of emotion as feedback to create customized content [13]. Real-time emotion identification can detect future terrorist behavior in a human being. Electroencephalography (EEG) and facial expressions together can improve emotion identification. The electrical activity of the brain may be measured using EEG, which can reveal clues about a person's underlying emotional state [14]. The user's emotional state can be taken into account when creating content, such as advertisements or suggestions. Apps for health and wellness can perform emotion detection to give feedback on stress levels and recommend mindfulness or relaxation activities. The extent of student interest in the classroom may be monitored in education. The systems may be used to detect aggressive, angry, or annoyed individuals. Then, before those people conduct crimes, such information might be leveraged to take action. AI systems provide offenders feedback on how they act and how they look so they may learn to regulate their emotions [15].

Challenges

- i. Due to individual variances in expression and the crucial need for context, it is difficult to correctly infer emotions from facial expressions.
- ii. The effectiveness of emotion detection systems may suffer when used on people from different cultural backgrounds.
- iii. Depending on their personalities, past events, and even their physical qualities, people display their emotions in various ways.
- iv. According to the circumstances, a single facial expression can portray a variety of emotions.
- v. Face hair, spectacles, and masks are a few examples of things that might hide facial emotions. These occlusions might make it difficult for systems to effectively identify and analyze facial signals.

The proposed research aims to enlighten the scientific community about the recent advances in emotion recognition methods using artificial intelligence and deep learning in the medical domain. From the input image, the proposed real-time emotional identification system identifies human reactions such as frustration, hatred, satisfaction, disbelief, and tolerance. When a human stands in front of a camera, the suggested approach identifies their emotion by comparing their facial expression with the reference images.

2. Dataset

The Facial Emotion Recognition (FER+) dataset is an expansion of the initial FER collection, in which the images were re-labeled as unbiased, happiness, disbelief, sorrow,

frustration, dissatisfaction, anxiety, and disapproval. Because of its tremendous scientific and business significance, FER is crucial in the domains of computational vision and artificial intelligence. FER is a technique that examines facial movements across passive images, as well as videos, to disclose details about a person's state of mind. Table 1 shows the FER 2016 dataset's test and training images [16].

Table 1. Test and train images of the FER 2013 dataset [16].

FER 2013	0	1	2	3	4	5	6
Test	3395	436	4096	7214	4830	3171	4965
Train	958	111	1024	1774	1247	831	1232

A dataset for recognizing facial expressions was made available in 2016 and is called FER 2016. Researchers from the University of Pittsburgh and the University of California, Berkeley generated the FER 2016 dataset. The dataset was gathered from a range of websites and open databases. Due to the variety of emotions and the variety of photos, it is regarded as one of the most difficult facial expression recognition datasets. The FER 2016 dataset's classes are:

1. Happiness—images of faces showing enjoyment, such as smiling or laughing, are included in this class.
2. Sadness—images of sad faces, such as those that are sobbing or frowning, are found in this class.
3. Anger—images of faces exhibiting wrath, such as scowling or staring, are included in this category.
4. Surprise—images depicting faces displaying surprises, such as enlarged eyes or an open mouth, are included in this category.
5. Fear—images depicting faces displaying fear, such as enlarged eyes or a shocked look, are included in this class.
6. Disgust—images of faces indicating disgust, such as those with a wrinkled nose or an upturned lip, are included in this category.
7. Neutral—images of faces in this category are described as neutral, since they are not showing any emotion.

For scientists conducting facial expression recognition research, the FER 2016 dataset is a useful tool. Although it is a difficult dataset, face expression recognition algorithms may be trained using it. There are several issues with existing datasets, including accessibility, the absence of guidelines, safety, examination, accessing data interaction, data analysis, information sets, metadata and reconstruction, intra-class deviation from overfitting, interruption, contrast variations, spectacles, and anomalies.

3. Methodology

The following are the difficulties with emotion detecting technologies in real environments:

- i. The technology can have trouble recognizing a person's face if there is excessive or insufficient light.
- ii. Due to occlusion, the technology cannot see a person's face if it is obscured by something.
- iii. Not every facial expression has the same meaning across cultures.
- iv. The technology cannot keep up with rapid facial movements.
- v. The technology cannot see a person's face if their head is turned away from the camera.
- vi. A person's face may be hidden by facial hair.

The proposed research is used to recognize the emotions of human beings that enable the user to find whether the displayed image of a person is happy, sad, or anxious, etc. Also, it helps to monitor the psychological behaviors of the human by identifying their

facial expression. AI algorithms, as well as deep learning approaches, are used to identify human faces. The system begins by looking for the eyes of a person, then face, forehead, mouth, and nostrils. The live image flows through the deep face algorithm; it recognizes the face and detects the facial features, as shown in Figure 1.

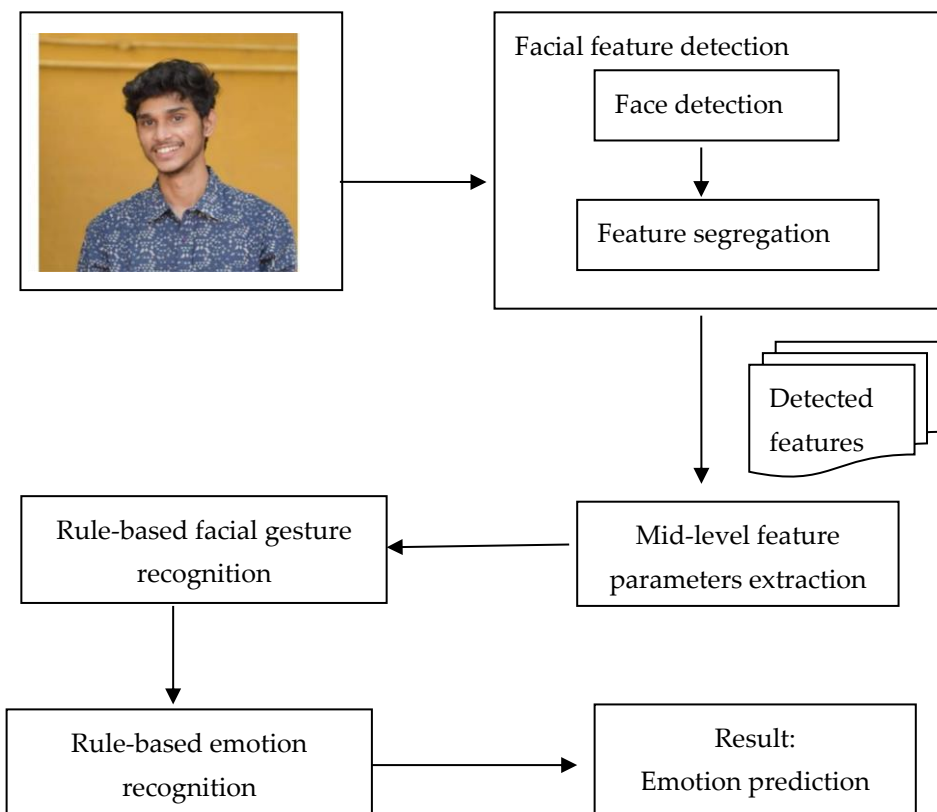


Figure 1. The proposed model for detecting human emotions.

3.1. Components Used in the Proposed System

The components used in this research are various libraries to process the face and to detect the emotion, age, gender, and race of the person. Face recognition and detection from the digital images and video frames are carried out using OpenCV. The deep learning face detector does not require additional libraries and Deep Neural Network (DNN) optimizes the implementation. After detecting the face, it processes the features and segregates them. Also, the algorithm detects the mid-level features, based on the input parameters. Then, the processed facial features need to be processed; the rule-based facial gestures are analyzed for subtle movement, by the facial muscle's Action Unit (AU) recognition. The plots in the face are processed, and the emotion is detected using rule-based emotion detection. Finally, the model indicates whether the individual is pleased, sad, furious, indifferent, or something else. The deep face algorithm finds the ethnicity, age, and also gender of the given face data.

As illustrated in Figure 2, numerous bits of information may be extracted from the initial image captured by the camera. The method recognizes the face of an individual from the camera image, even if the person is wearing ornaments.

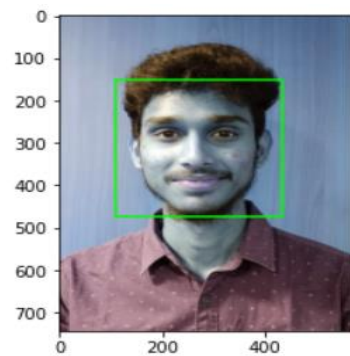


Figure 2. Live input image of a person from the camera.

The human's face is captured from the live camera is shown in Figure 3 with various expressions and is classified accurately.

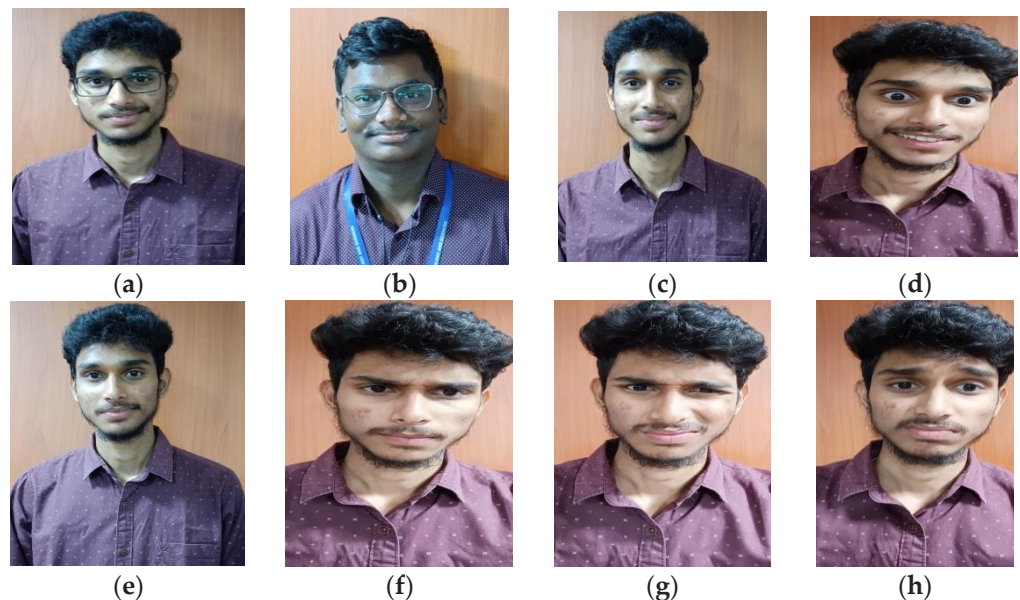


Figure 3. Live camera images: (a–c) happy; (d) surprised; (e) neutral; (f) angry; (g) disgusted; and (h) sad.

To identify the facial features, the NumPy array loads the image obtained from the camera using the `load_image_file` method, and the array of information is passed to `face_landmarks`. This will provide a Python list with the dictionary of face characteristics and their locations. Matplotlib is used in face recognition to plot and measure the dimensions of the face and facilitate its processing. It finds the face, excluding other objects, and generates the plots.

DeepFace is a lightweight face identification framework for analyzing facial characteristics [17]. It is a composite facial recognition framework that encapsulates cutting-edge models to recognize human emotional attributes [18]. To train and categorize the faces in the picture dataset, the DeepFace system employs a deep CNN (Convolutional Neural Network) [19]. DeepFace is composed of four modules: two-dimensional (2D) coordination, three-dimensional (3D) alignment, formalization, and a neural network. A face image cycles through these in turn, generating a 4096-dimensional characteristic vector describing the face. The matrix of features may then be utilized to carry out a range of tasks. To identify the face, the collections of feature vectors of faces are compared, to find the face with the most comparable feature vector. It accomplishes this through the use of a 3D depiction of a face [20]. The 2D alignment unit detects six fiducial locations on the observed face. Still, 2D translation fails to correct rotational motions that are out of

position. DeepFace aligns faces using a 3D model, in which 2D photographs have been reduced to 3D equivalents. The 3D image has 67 fiducial points. Following the distortion of the image, 67 anchoring points are individually placed on visualization. Because entire viewpoint perspectives are not modeled, the fitted camera is a rough representation of the individual's real face. DeepFace attempts to reduce errors by warping 2D pictures with subtle deviations. Furthermore, the camera may substitute areas of a photograph and blend them into their symmetrical counterparts. CNN's deep neural network architecture includes maximum pooling, a convolutional layer, three directly linked layers, and a layer that is fully connected. The input data are an RGB image of the human face, sized to fit the display format 152 times, whereas the result is a real vector of size 4096 that represents the facial image's characteristic vector.

3.2. Pseudocode for Human Emotion Feature Prediction Using DeepFace

The following pseudocode outlines a basic process for emotion detection using DeepFace.

```
def predict_emotion_features(image):
    # Load the DeepFace model.
    model = load_model("deepface_model.h5")
    # Extract the features of the face in the image.
    features = extract_features(image)
    # Predict the emotion features of the face.
    emotion_features = model.predict(features)
    # Return the emotion features.
    return emotion_features
```

The DeepFace model to identify emotion on the face is loaded using the `load_model()` method. The distinctive features of the human face in the image are extracted via the `extract_features()` method. These features may include the placement of the eyebrows, the contour of the lips, and the appearance of forehead wrinkles. Based on the data that are retrieved, the `predict_emotion_features()` algorithm forecasts the facial features associated with emotions. The predicted emotion features are returned by the `return emotion_features`. An array of values representing the likelihood that every emotion is related is stored in the `emotion_features` variables in this scenario. For instance, the face in the image is probably related with pleasure (0.5), then sorrow (0.3), and anger (0.2), if the `emotion_features` variables are [0.2, 0.5, 0.3].

4. Results and Discussions

The proposed method used digital identifiers, with an optical flow-based approach, to construct a real-time emotional recognition system with minimum computational demands in terms of implementation and memory. The following are the criteria for choosing the best AI system for human emotion detection.

- i. Accuracy in appropriately detecting emotions.
- ii. Robustness to function in many circumstances, such as varying illumination and movements of the face.
- iii. Scalability for large-scale data analysis.
- iv. The cost of AI technology should be affordable.

The proposed approach works effectively under irregular illumination, human head tilting up to 25°, and a variety of backgrounds and complexions. Figure 4 depicts the facial expressions and feelings of the live-captured person's face. The proposed approach recognized all of the actual user's emotions.

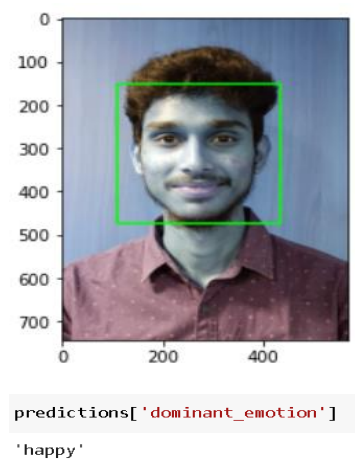


Figure 4. ROI region extraction and emotion prediction.

In addition, the algorithm extracts emotions from the provided input image. The result of the testing and training dataset is given in Table 2. DeepFace employs a deep learning technique to attain its high accuracy of 94%. Additionally employing a hierarchical methodology, DeepFace learns the characteristics of faces at many levels of abstraction. As a result, it is more resistant to changes in face expression.

Table 2. Result of testing and training dataset.

Data Set	Avg. Accuracy
Training Data (Not Normalized)	94.40%
Training Data (Normalized)	95.93%
Testing Data (Normalized)	92.02%

Human–machine interaction technology, including machines that can comprehend human emotions, holds immense importance in various domains, and it can significantly improve efficiency in multiple ways. Customer happiness may be measured in real-time in customer service by machines that comprehend emotions. As a result, problems may be resolved right away, decreasing customer annoyance and raising the general effectiveness of support procedures. In medical applications, motion recognition technologies can be quite useful. Medical personnel can deliver more individualized and sympathetic treatment by using machines that can recognize modifications to patients’ mental health. Machines that can understand student emotions in educational settings can modify lesson plans and instructional strategies. When kids are suffering, bored, or disinterested, they can spot these and make adjustments. Virtual assistants can modify the replies and tones based on the emotions of the user, improving interactions. In addition, the system assists physically and socially challenged people, such as those that are deaf, dumb, bedridden, or autistic, to recognize their emotions. Furthermore, it influences corporate outcomes and assesses the audience’s emotional responses. It is more useful for individualized online learning than for maximizing performance.

As shown in Table 3, the proposed system outperforms competitive methods.

Table 3. Metrics of the proposed technique on human emotion detection.

Classifiers	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
Emotion detection	80.45	80.23	89.67	90.25
Age prediction	86.55	83.27	77.02	87.67
Gender prediction	95.67	91.94	95.26	99.99
Race prediction	90.5	93.64	92.27	96.22

5. Conclusions

The same emotion may be expressed in many ways by different people, which can make it challenging for AI systems to recognize emotions with accuracy. Emotions frequently show themselves in subtly changing facial expressions or body language. This can make it challenging for AI programs to reliably recognize emotions. Despite these obstacles, human emotion recognition, utilizing DeepFace and artificial intelligence, is a promising topic with several applications. As AI technology advances, we should expect more precise and complex emotion recognition systems in the future. The proposed method differentiates emotions in 99.81% of face coordinates and 87.25% of FER datasets. The proposed technique can also be utilized to extract more characteristics from other datasets as well. In addition to refining system procedures, putting participants in real-life circumstances to communicate their true sentiments can assist to increase the performance of the system.

Author Contributions: Conceptualization, R.V. and S.S.; methodology, M.S. and T.J.J.; formal analysis, S.S.; investigation, R.V. and S.S.; resources, T.J.J.; writing—original draft preparation, S.S., M.S. and T.J.J.; writing—review and editing, R.V. and T.J.J.; visualization, S.S.; supervision, T.J.J.; project administration, T.J.J.; funding acquisition, R.V., S.S., M.S. and T.J.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article.

Acknowledgments: The authors would like to thank the Karunya Institute of Technology and Sciences for all the support in completing this research.

Conflicts of Interest: The authors do not have any conflict of interest.

References

1. Huang, D.; Guan, C.; Ang, K.K.; Zhang, H.; Pan, Y. Asymmetric spatial pattern for EEG-based emotion detection. In Proceedings of the 2012 International Joint Conference on Neural Networks (IJCNN), Brisbane, Australia, 10–15 June 2012; pp. 1–7.
2. Chowdary, M.K.; Nguyen, T.N.; Hemanth, D.J. Deep learning-based facial emotion recognition for human–computer interaction applications. *Neural Comput. Appl.* **2021**, *35*, 23311–23328. [\[CrossRef\]](#)
3. Singh, S.K.; Thakur, R.K.; Kumar, S.; Anand, R. Deep learning and machine learning based facial emotion detection using CNN. In Proceedings of the 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 23–25 March 2022; pp. 530–535.
4. Cui, Y.; Wang, S.; Zhao, R. Machine learning-based student emotion recognition for business English class. *Int. J. Emerg. Technol. Learn.* **2021**, *16*, 94–107. [\[CrossRef\]](#)
5. Kakuba, S.; Poulose, A.; Han, D.S. Deep learning-based speech emotion recognition using multi-level fusion of concurrent feature. *IEEE Access* **2022**, *30*, 125538–125551. [\[CrossRef\]](#)
6. Tripathi, S.; Kumar, A.; Ramesh, A.; Singh, C.; Yenigalla, P. Deep learning based emotion recognition system using speech features and transcriptions. *arXiv* **2019**, arXiv:1906.05681.
7. Chen, Y.; He, J. Deep learning-based emotion detection. *J. Comput. Commun.* **2022**, *10*, 57–71. [\[CrossRef\]](#)
8. Schoneveld, L.; Othmani, A.; Abdelkawy, H. Leveraging recent advances in deep learning for audio-visual emotion recognition. *Pattern Recognit. Lett.* **2021**, *146*, 1–7. [\[CrossRef\]](#)
9. Sun, Q.; Liang, L.; Dang, X.; Chen, Y. Deep learning-based dimensional emotion recognition combining the attention mechanism and global second-order feature representations. *Comput. Electr. Eng.* **2022**, *104*, 108469. [\[CrossRef\]](#)
10. Sajjad, M.; Kwon, S. Clustering-based speech emotion recognition by incorporating learned features and deep BiLSTM. *IEEE Access* **2020**, *8*, 79861–79875.
11. Jaiswal, A.; Raju, A.K.; Deb, S. Facial emotion detection using deep learning. In Proceedings of the 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 5–7 June 2020; pp. 1–5.
12. Neumann, M.; Vu, N.T. Attentive convolutional neural network based speech emotion recognition: A study on the impact of input features, signal length, and acted speech. *arXiv* **2017**, arXiv:1706.00612.
13. Imani, M.; Montazer, G.A. A survey of emotion recognition methods with emphasis on E-Learning environments. *J. Netw. Comput. Appl.* **2019**, *147*, 102423. [\[CrossRef\]](#)

14. Kamble, K.S.; Sengupta, J. Ensemble machine learning-based affective computing for emotion recognition using dual-decomposed EEG signals. *IEEE Sens. J.* **2021**, *22*, 2496–2507. [[CrossRef](#)]
15. Sahoo, G.K.; Das, S.K.; Singh, P. Deep learning-based facial emotion recognition for driver healthcare. In Proceedings of the 2022 National Conference on Communications (NCC), Mumbai, India, 24–27 May 2022; pp. 154–159.
16. FER-2013. Available online: <https://www.kaggle.com/datasets/msambare/fer2013> (accessed on 2 November 2023).
17. Chiurco, A.; Frangella, J.; Longo, F.; Nicoletti, L.; Padovano, A.; Solina, V.; Mirabelli, G.; Citraro, C. Real-time detection of worker's emotions for advanced human-robot interaction during collaborative tasks in smart factories. *Procedia Comput. Sci.* **2022**, *200*, 1875–1884. [[CrossRef](#)]
18. Sha, T.; Zhang, W.; Shen, T.; Li, Z.; Mei, T. Deep Person Generation: A Survey from the Perspective of Face, Pose, and Cloth Synthesis. *ACM Comput. Surv.* **2023**, *55*, 1–37. [[CrossRef](#)]
19. Karnati, M.; Seal, A.; Bhattacharjee, D.; Yazidi, A.; Krejcar, O. Understanding Deep Learning Techniques for Recognition of Human Emotions Using Facial Expressions: A Comprehensive Survey. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–31.
20. Mukhiddinov, M.; Djuraev, O.; Akhmedov, F.; Mukhamadiyev, A.; Cho, J. Masked Face Emotion Recognition Based on Facial Landmarks and Deep Learning Approaches for Visually Impaired People. *Sensors* **2023**, *23*, 1080. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.