VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

CHENNAI

FALL SEMESTER 23-24

Review on Prediction of Emotion Recognition Through Computer Vision and Deep Learning

Submitted by:

Khan Mohammed Raza (23MCA1029)

khanmohammed.raza2023@vitstudent.ac.in

Vinayak Kumar Singh (23MCA1030)

vinayakkumar.singh2023@vitstudent.ac.in

Shivanand Thakur (23MCA1012)

shivanand.thakur2023@vitstudent.ac.in

Submitted To:

Dr. Rashmi Rekha Borah

Assistant Professor

Vellore Institute of technology, Chennai

rashmirekha.borah@vit.ac.in

Abstract:

Facial expression recognition is a growing field at the crossroads of computer vision, psychology, artificial intelligence, and deep learning. This work explores how facial expression recognition analyses emotions through facial cues, highlighting its broad impact across different areas and emphasising its transformative potential. The method proposes a multi-step pipeline, including face detection, detecting facial landmarks, emotion classification, using convolutional neural networks for feature extraction, using data augmentation for noise removal, and handling various lighting conditions. To work with the model, it can utilise datasets such as FER-2013, which contains 30000 facial RGB images of different expressions; CK+, which contains many images of one hundred and twenty-three people; and AffectNet, which contains around 0.4 million images manually labelled for the presence of different facial motions. The main objective of the research is to address the complexity and variability of human emotions, making them more reliable and robust across different environmental conditions, including varying lighting, angles, and ethnic backgrounds. This research is likely to yield significantly improved accuracy in recognising and classifying human emotions from facial expressions and could enhance user experiences in technologies like virtual reality and improve the quality of human-computer interaction.

Keywords: *Facial Expression Recognition, Deep Learning, Computer Vision, Facial Landmark Detection, Feature Extraction, Human Emotions, Artificial Intelligence*

I. INTRODUCTION

Facial expression recognition is a growing field at the crossroads of computer vision, psychology, artificial intelligence, and deep learning. The field of facial expression recognition is a cutting-edge area of study that brings together computer vision, psychology, artificial intelligence, and deep learning. This emerging field promises broad implications in a variety of fields as it uses facial cue analysis to decode human emotions. This study is based on a thorough methodology that includes four important steps: convolutional neural networks for feature extraction, facial landmark identification, emotion classification, and face detection. Data augmentation helps with managing changing lighting conditions and reducing noise. Important datasets for model development are CK+, AffectNet, and FER-2013. The main objective of this study is to address the subtleties and complexity of human emotions in order to improve the accuracy and resilience of emotion recognition when faced with a variety of external stimuli. Changing lighting, different perspectives, and diverse ethnic backgrounds are some of these variables. This work is expected to lead to a major improvement in the accuracy and precision of facial expression-based human emotion recognition and classification. This advancement could improve user experiences with virtual reality and the calibre of human-computer interaction.

II. PROBLEM STATEMENT

Accurately identifying emotions across a range of cultural backgrounds and demographic groups is a challenge that faces researchers studying emotion recognition. When used on people from different age groups or cultures, current emotion recognition models may show biases and inaccuracies. Developing cross-cultural and cross-demographic emotion recognition models that take into consideration differences in expressions, norms, and cultural context is the goal of this research.

III. OBJECTIVES

To create more captivating and natural-feeling HCIs. Systems for recognising emotions, for instance, can be used to gauge a user's mood and modify the interface appropriately. For instance, the interface could offer more assistance and direction if the user is experiencing frustration to enhance client support. Systems that recognise emotions, for instance, can be used to identify customer discontent and offer proactive support. For instance, the system might suggest or offer a discount to a customer who appears hesitant to purchase a product.The system might offer more difficult or engaging content, for instance, if the student appears bored. to produce more realistic and immersive virtual worlds. For instance, virtual characters that can react to the user's emotions naturally can be made using emotion recognition systems. The virtual character might become more comforting and encouraging, for instance, if the user is experiencing fear.

IV. RESEARCH QUESTION

How can emotion recognition models be made more robust and adaptable to varying lighting conditions and image qualities?

How can emotion recognition be integrated into user-friendly applications to enhance human-computer interactions and user experiences?

Can facial expression recognition contribute to improving mental health diagnosis, and what are the potential challenges in its application in the mental health field?

V. METHODOLOGY.

Gathering a dataset of facial expressions and labelling them with the associated emotions is the first step. Numerous sources, including speech, video, and physiological recordings, can be used to gather the data. The data is then pre-processed in order to normalise it and eliminate artefacts and noise before being subjected to additional analysis. The pre-processed data is then used to extract features that are important for emotion recognition in the second step. For instance, facial muscle position and movement are often extracted by facial expression recognition systems. Pitch, loudness, and speech rate are typical features that speech emotion recognition systems extract. Heart rate, blood pressure, and skin conductance are common features that physiological emotion recognition systems extract. They choose a subset of the extracted features to train a machine learning classifier to predict the emotion of a given input sample. Support vector machines, logistic regression, and neural networks are just a few examples of the machine learning classifiers that can be used to identify emotions

VI. LITERATURE SURVEY

Kim et al (2019) suggests a hierarchical deep neural network-based facial expression recognition system that is effective. To increase FER accuracy, the suggested technique uses deep learning to integrate temporal and spatial information. The algorithm is developed to overcome the difficulties posed by shifting surroundings and individual variations in face emotions.

The suggested algorithm has the potential to be used in AI robots, emotion recognition systems, and human-computer interaction. By adding more features and refining the network design, the suggested method may be enhanced in the future.

Hussain et al (2020) provides a deep learning model-based solution for the real-time classification and recognition of facial emotions. The suggested system recognizes and categorizes facial expressions in real-time using a convolutional neural network (CNN). Using a dataset of 1,000 photos, the system's accuracy was tested and found to be 92.5%. Future research should focus on utilizing bigger datasets to increase the system's accuracy as well as investigating the use of different deep learning models. Potential uses for the technology exist in a number of industries, including entertainment, healthcare, and security. But it's also important to take into account the limitations of facial recognition technology and ethical issues.

Kong (2019) et al suggested a novel approach to facial expression recognition that makes use of enhanced LBP features and deep convolutional neural networks. With a classification accuracy of 91.28%, the suggested strategy outperforms conventional machine-based techniques. The study also demonstrates that although adding layers to the network shortens training iterations and enhances recognition performance. Further research will focus on enhancing the combination of LBP and DCNN characteristics and investigating the use of additional deep learning methods. Other kinds of picture recognition jobs can also be handled by the suggested technique.

Li et al (2020) offered a cutting-edge method for recognizing facial expressions utilizing a convolutional neural network with an attention mechanism. To avoid overfitting, the technique employs data augmentation together with LBP and convolution features. Additionally, the authors present a brand-new dataset known as NCUFE, which consists of 26950 photos with seven expression annotations. The experimental results demonstrate that the suggested method outperforms several existing methods, and it is assessed on four more datasets in addition to NCUFE. Future research aims to increase the accuracy of the suggested approach and investigate the use of attention mechanisms in different computer vision tasks.

Wang et al (2020) suggests a framework to enable real-time monitoring of students' emotions in online courses by integrating facial expression recognition based on convolutional neural network architecture with online education platforms. The CNN-based deep learning model and online course platforms make up the two components of the framework. The framework was put to the test by the authors with twelve participants in a real online meeting, and they made available to readers a dataset.

Ahmed et al (2019) proposed work focuses on employing a convolutional neural network (CNN) and data augmentation to increase the accuracy of facial expression recognition. Outperforming other models in use, the suggested model reached a validation accuracy of 96.24% after 120 epochs. The study also emphasizes how crucial data augmentation is to the creation of models, particularly in cases where the dataset is small. Compound emotion classes and advanced uncertainty analysis techniques such as Belief Rule Based Expert Systems (BRBES) are areas of future research.

Zadeh et al (2019) suggested a deep learning-based framework that makes use of convolutional neural networks (CNNs) and Gabor filters to quickly identify facial emotions. In comparison to alternative methods, the suggested methodology achieves both faster training and higher recognition accuracy. CNN is used for classification, and Gabor filters are used for feature extraction. The outcomes of the experiment demonstrate that the suggested features improve the neural network's training speed and accuracy. Future research will examine the potential of this methodology for real-time applications as well as the use of additional deep learning techniques and datasets for emotion recognition.

Bin Li A et al (2021) revealed a deep residual network-based enhanced face emotion recognition model. Convolutional neural networks (CNNs) are used by the model to extract features using the ResNet-50 architecture. In terms of recognition performance, the suggested method performs better than other cutting-edge techniques. The dataset used for the experiments is also covered in the paper, along with future work that will test the efficacy of various ResNet layers and variations for facial emotion recognition.

Zhang et al (2020) explains how to use image edge computing and a convolutional neural network (CNN) to recognize facial emotions. The Kirsch edge operator is used to extract edge information from the image, and the Haar classifier is used for human detection. After that, the CNN model is used to categorize the picture using the features that were extracted. The outcomes demonstrate that the suggested method recognizes facial expressions with a high degree of accuracy. The primary goal of the research is to increase face emotion recognition accuracy by combining CNN and image edge computing methods.

Sharma et al (2022) provides a study on the use of fading point detection and localized color histograms for building recognition. Using this methodology, images are divided into sub-regions, and GIST features are extracted from each one. The study assesses and contrasts the performance of the suggested biologically-plausible building recognition method with alternative approaches. The long processing time and the requirement for a simpler background are just two of the suggested method's drawbacks that are highlighted in the conclusion. Nonetheless, the analysis demonstrates that the suggested approach performs at an accuracy level of 85.3%, which is on par with alternative approaches.

Wu et al (2018) explains how to translate emotional Mandarin and Tibetan speech from sign language. Pre-processing, feature extraction, and SVM recognition are used in the methodology to recognize signs. Pre-processing, feature extraction, and SVM recognition are also used in facial expression recognition. Speaker adaptive training and a mixed language average acoustic model are used to train the emotional speech models. The three-dimensional pleasure-arousal-dominance (PAD) emotional model and the EMOS scoring standard are used to assess the synthesized emotional speech.

Mehendale (2022) outlines the current state of Emotion Recognition in Conversation (ERC), along with its challenges, benchmarks, and advancements. The technique entails examining the three elements that affect the ability to identify emotions in a conversation: the statement and its context, the speaker's emotional state, and the emotions that have been expressed in previous statements. In order to address ERC, the paper discusses several strategies, one of which is the Conversational Memory Network (CMN). The conclusion emphasizes the need for additional study to address the issues with ERC, including the dearth of annotated datasets and the requirement for more advanced models. The paper's overall goal is to present a thorough understanding of ERC and its possible uses in a range of industries.

Hua et al (2019) provide a special section on the information-centric Internet of things' smart caching, communications, computing, and cybersecurity. It focuses on the significance of the HERO (Human Emotions Recognition) technology in achieving the intelligent Internet of Things. The most recent findings on CNN applications in computer vision and facial expression recognition are covered in this paper. In order to increase the system's stability, test accuracy, and capacity for generalization, it suggests an algorithm that combines the benefits of ensemble learning and neural networks.

Poria et al (2019) presents a two-level Convolutional Neural Network (CNN) framework-based approach for facial expression recognition. To extract the facial expression vector (EV), which is directly related to changes in expression, the first step entails background removal. A CNN is used in the second level to identify facial features and categorize emotions. The algorithm outperforms other techniques that combine expression detection and background removal in a single CNN, achieving an accuracy of 96%. The impact of orientation and the existence of multiple faces in an image are two examples of the method's limitations that are covered in the study.

Guanghui et al (2021) suggests a multi-modal emotion recognition technique that combines speech and visual modalities' correlation features to enhance emotion recognition performance. The three primary components of the method are multi-modal fusion, feature learning, and pre-processing. It also makes effective use of the class information of speech and visual features to fuse the features and enhance recognition. Based on three datasets, the experimental results demonstrate that the proposed method outperforms other state-of-the-art methods in terms of recognition rates.

Frith et al (2009) suggested the various pathways of perception and action, emphasizing the processing that occurs both subconsciously and consciously. It investigates how various types of information are extracted and how these processes happen concurrently. The significance of consciousness in behavior and the intricate interplay between conscious and unconscious processes are also discussed in the paper. It also covers the impact of an audience on emotional expressions and the use of facial expressions as communication signals.

Dzedzickis et al (2020) suggested a system for evaluating and classifying emotions, with an emphasis on automated emotion recognition hardware and techniques. It draws attention to the difficulties in evaluating feelings and the multidisciplinary nature of emotion recognition. The significance of multimodal analysis and the demand for a cohesive understanding of datasets are emphasized in the paper. The importance of signal processing and analysis methods for emotion recognition is also mentioned. The potential for combining the Skin Conductance Response (SKT) technique with other approaches is discussed, as well as its limitations.

Naseem (2010) et al suggests a unique method for face identification based on nearest subspace classification and linear regression. The suggested approach produces the best outcomes for the difficult scarf occlusion problem that have ever been documented. We introduce the Linear Regression Classification (LRC) algorithm, which estimates the occluded portion of the face using a linear regression model. Standard databases are used to assess the methodology under various evaluation protocols. The findings demonstrate that the suggested approach performs better than state-of-the-art techniques. The potential for practical applications of the suggested methodology is highlighted in the paper's conclusion.

Jack et al (2009) investigates cultural differences in the interpretation of facial expressions using eye-tracking technology. Thirteen Western Caucasian and thirteen East Asian observers were shown images of six FACS-coded facial expressions plus neutral. The researchers established face regions and used Minimum Description Length (MDL) analysis to categorize fixations by face region. The results challenge the notion that facial expressions are universally understood and highlight the importance of considering cultural context in communication. The abstract, methodology, and conclusion all emphasize the importance of cultural context in interpreting facial expressions and the need for further research in this area.

Ayadi et al (2011) provides a comprehensive survey of speech emotion recognition systems, covering feature extraction methods, classification techniques, and emotional speech databases. The objective is to evaluate the performance of existing systems and identify limitations and challenges. The methodology includes a review of various emotional speech corpora and experimental setups used in previous studies. The paper concludes that while high classification accuracies have been achieved for high-arousal and low-arousal emotions, N-way classification remains challenging. The performance of current stress detectors needs significant improvement, and the average classification accuracy of speaker-independent speech emotion recognition systems is less than 80%.

Bartlett et al (2017) provided the challenges and advancements in machine analysis of facial expressions. It highlights the limitations of previous methods and introduces a novel, robust, fully automated facial point detector. The paper also explores the importance of temporal modeling of facial expressions and compares the effectiveness of geometric and appearance-based features. The performance of the proposed methods is evaluated using cross-validation and correlation analysis. Overall, the paper aims to improve the accuracy and reliability of facial expression analysis.

Zou et al (2012) discusses the problem of very low resolution (VLR) face images in surveillance videos and proposes a super-resolution (SR) algorithm to recover the missed details of the face image. The proposed method outperforms existing methods in terms of visual quality and recognition performance. The experiments evaluate the algorithm using subjective human visual quality and objective measurement using mean-squared error (MSE). The results show that the proposed method achieves the lowest MSE and improves recognition accuracy. The paper also discusses the use of different databases and settings for the experiments.

Lischinski et al (2013) presents a method for image-based facial modeling and animation. The main objective is to create realistic face models and perform transitions between different expressions. The methodology involves fitting a generic face model to individual faces using photogrammetric techniques. Texture maps are extracted from photographs to render photorealistic images of the face model. The paper also introduces expression synthesis techniques based on morphing and a painting interface for adding expressions. The results of experiments with the proposed techniques are presented, and directions for future research are discussed.

Szwoch (2015) et al presents an approach for recognizing facial expressions and emotions based solely on depth data from the Microsoft Kinect sensor. The proposed algorithm uses local movements detection within the face area to recognize actual facial expressions, and has been validated on Facial Expressions and Emotions Database. Although the average recognition accuracy is slightly above 50%, this approach is highly independent of illumination conditions and accepts low distance between sensor and the user. The study aims to provide a valuable tool to support other algorithms based on optical channel, as well as using skeleton or face tracking information. The conclusion suggests that this approach can be used in real-world applications, such as human-computer interaction, gaming, and healthcare

Zhong et al (2017) proposes a Temporal Information Preserving Framework (TIPF) for emotion recognition using facial expressions and physiological signals. The TIPF combines different data streams to more accurately recognize emotions, and our experiments show that it significantly improves performance, especially when combining facial expressions and physiological signals. We analyze the performances of facial expressions from different perspectives for predicting human emotional states individually and jointly. The main objective of this paper is to provide a more robust and reliable model for emotion recognition by fusing different data streams.

Pantic et al (2000) surveys the state of the art in automating facial expression analysis in facial images and image sequences. The authors identify three basic problems related to facial expression analysis: face detection, facial expression data extraction, and facial expression classification. They describe the characteristics of an ideal automated system for facial expression analysis and survey the techniques presented in the literature in the past decade for facial expression analysis by a computer. The authors selectively discuss systems which deal with each of these problems and provide possible directions for future research. The ultimate goal is to achieve a human-like interaction between man and machine.

Pighin et al (2006) provides a technique for facial modelling and animation based on images. The primary goal is to use photogrammetric techniques to fit a generic face model to a person's face and facial expressions. The procedure for texture extraction, expression synthesis, and model fitting is covered in the paper. In order to produce texture maps for rendering photorealistic facial images, the authors suggest methods for combining values from various photos. Additionally, a painting interface for adding expressions and a morphing technique for animating the face model are introduced in this paper. Experiments conducted using the suggested techniques are reported, and future research directions are discussed.

Cohen et al (2000) proposes a new architecture of Hidden Markov Models (HMMs) for automatic segmentation and recognition of human facial expressions from video sequences. The methodology explores person-dependent and person-independent recognition of expressions and increases the discrimination power between different classes. The authors suggest that recognizing emotions from facial expressions alone may not be accurate enough, and other measurements such as voice and gestures, as well as context, should also be considered. The paper concludes that this work is a step towards building more effective computers that can serve us better in various applications, including education.

Goldman et al (2005) explores the simulationist approach and theorizing approach to emotion mindreading, specifically in the context of face-based emotion recognition. The authors review existing neuropsychological research and argue that the simulation approach offers the best explanation of the data, while the theorizing approach does not fit with the evidence. They also propose and evaluate four specific models of how normal mindreaders could use simulation to arrive at emotion classifications. The paper concludes that while the simulation approach is supported by the evidence for face-based emotion recognition, it cannot be assumed that this style of mindreading is the same as that used in other subdomains. The authors call for further research and theory into this area of mental state ascriptions.

Jack et al (2009) investigates cultural differences in the interpretation of facial expressions using eye-tracking technology. Thirteen Western Caucasian and thirteen East Asian observers were shown images of faces displaying six different expressions, and their eye movements were recorded. The results showed that East Asian observers focused more on the eyes, while Western Caucasian observers focused more on the mouth. This suggests that cultural background influences the way people interpret facial expressions. The study highlights the importance of considering cultural context in communication and challenges the notion that facial expressions are universally understood.

Kalsum et al (2018) proposes a method for automatic facial emotion recognition using hybrid feature descriptors that combine spatial bag of features with spatial scale-invariant feature transform and spatial speeded up robust transform. The proposed method is evaluated using K-nearest neighbor and support vector machines with linear, polynomial, and radial basis function kernels for classification of emotions. The results show that the hybrid SBoF-SSIFT feature descriptor is most effective for emotion recognition using facial images, achieving high accuracies of 98.33% and 98.5% on JAFFE and CK+ datasets, respectively. The authors conclude that their method improves upon existing feature-based approaches and is well-suited for datasets with varying images.

Koolagudi et al (2012) provides a comprehensive review of recent research on emotion recognition from speech, with a focus on the Indian context. The authors discuss the challenges of collecting natural emotional speech corpora and the importance of using hybrid models to enhance performance. They also highlight the need to evaluate established features on different Indian languages for emotion recognition and to develop techniques for removing speaker-specific information from speech utterances. The paper concludes by identifying important research gaps and calling for further work in this area. Overall, this review serves as a valuable resource for researchers and practitioners interested in emotion recognition from speech.

Sebe et al (2005) is about Multimodal Emotion Recognition and covers various aspects of the topic. The abstract introduces the importance of understanding human emotions and the challenges in developing a system for recognizing them. The methodology section describes the different modalities used for emotion recognition, including facial expressions, physiological signals, and speech. The authors also discuss the use of probabilistic graphical models for integrating these modalities. The conclusion emphasizes the potential of multimodal emotion recognition for improving human-computer interaction and highlights the need for further research in this area. Overall, this paper provides a comprehensive overview of the current state of research in multimodal emotion recognition.

Duncan et al (2016) presents a convolutional neural network that can classify human emotions from dynamic facial expressions in real time. The network was trained on three datasets, including a new home-brewed database, and achieved a training accuracy of 90.7% and test accuracy of 57.1%. The network uses a Haar-Cascade detector to crop faces and applies a particular emoji to the subjects' faces based on the resulting classification. The results demonstrate the feasibility of implementing neural networks in real time to detect human emotion. The paper concludes that the network can be used in various applications, such as human-computer interaction, gaming, and mental health diagnosis.

Ioannou et al (2005) presents a neurofuzzy network for emotion recognition through facial expression analysis, which can improve man-machine communication systems. The methodology involves facial feature extraction and the derivation of rules for emotion recognition based on facial expression analysis. The system adapts to specific users' facial expression characteristics and uses facial animation parameters (FAPs) to validate emotional cues. The conclusion highlights the potential of this system in real-life situations and its accuracy in recognizing emotions. Overall, this paper aims to contribute to the field of affective computing and improve human-computer interaction.

Adolphs et al (1996) aimed to investigate the neural systems involved in recognizing emotions in facial expressions and how they relate to specific emotions. The researchers tested their hypotheses by analyzing data from 34 subjects with focal brain damage and found that different regions of the brain are responsible for recognizing specific emotions. They also found that some individuals may have difficulty recognizing negative emotions like fear and sadness. The study concludes that the sensory cortices within the right hemisphere are essential for recognizing emotions in facial expressions, and partly different sets of such cortical regions might be important in processing different basic emotions. Overall, this study provides a deeper understanding of the neural systems involved in recognizing emotions and their relation to specific emotions.

Mao et al (2015) proposes a real-time emotion recognition approach using Kinect sensors that combines 2D and 3D facial expression features. The methodology involves capturing animation units and feature point positions to track facial deformation, and using a fusion algorithm based on improved emotional profiles and maximum confidence to recognize emotions in real-time. The approach shows superior performance on both an emotion dataset and a real-time video. The objective is to expand the model for more comprehensive training and to disengage from the limitations of Face Tracking SDK. In conclusion, the proposed approach has potential applications in various fields, including human-computer interaction, psychology, and entertainment.

Landowska et al (2017) examines the limitations of automatic emotion recognition from facial expressions in e-learning contexts. The study involved three consecutive tutorials of varying difficulty and duration, during which participants were asked to perform operations shown in the tutorial while their facial expressions were recorded. The results showed that automatic emotion recognition tools were not reliable in detecting emotions accurately, and that there were limitations to using facial expressions as a measure of emotional states in e-learning. The authors suggest that alternative methods, such as self-reporting or physiological measures, may be more effective in monitoring emotional states in e-learning. The paper concludes with a summary of the results and a discussion of the implications for future research.

Paiva-Silva et al (2016) provides a comprehensive review of the methodologies used to assess facial emotion recognition over the past 20 years. The authors conducted a systematic search of various databases and identified 291 articles, of which 115 were included in the analysis. The articles were categorized into three groups: non-behavior-dependent methodologies (MRI and EEG), instruments used to assess facial emotion recognition, and health conditions associated with facial recognition impairments. The authors found a lack of replication and a multiplicity of instruments and methodological strategies, which may have contributed to conflicting results.

Hamann et al (1996) discusses the role of the amygdala in recognizing facial emotions. The paper presents data from two men who suffered extensive bilateral damage to the amygdaloid complex, hippocampus, and other cortical regions due to herpes simplex encephalitis. The study found that the amygdala plays a crucial role in recognizing emotions in facial expressions, and that damage to this region can impair this ability. The paper concludes that the amygdala is an important brain structure for emotional learning and recognition, and that further research is needed to fully understand its role in these processes. Overall, this paper provides valuable insights into the neural mechanisms underlying facial emotion recognition.

Sprengelmeyer et al (1998) investigate the neural substrates involved in recognizing facial expressions of basic emotions. The experiment involved presenting eight different faces depicting fear, anger, disgust, and neutral expressions to subjects while they underwent fMRI scanning. The results showed that different neural structures were involved in recognizing different emotions, with the amygdala being particularly important for recognizing fear and disgust, and the ventral parts of the frontal cortex being important for recognizing all emotions. The study concludes that the neural substrates involved in recognizing facial expressions of basic emotions are distinct and non-overlapping.

Tan et al (2021) proposes a multimodal emotion recognition method for human-robot interaction (HRI) systems using facial expressions and electroencephalography (EEG). The proposed method combines the recognition results of facial expressions and EEG using the Monte Carlo method to solve the multimodal emotion recognition problem. The HRI system developed in this work interacts with humans according to the emotion recognition results obtained through facial expressions and EEG, improving the interaction quality between humans and robots. A perceptual assessment method was proposed to evaluate the system according to human perception. The results of the experiments show that the proposed method achieved high accuracy in recognizing emotions and can be applied to real-world HRI systems.

Kolakowska et al (2014) explores the field of affective computing and its applications in various domains such as software engineering, website customization, education, and gaming. The paper provides a concise review of affect recognition methods based on different inputs such as biometrics, video channels, or behavioral data. The methodology involves proposing a set of research scenarios to evaluate the possibility of using emotion recognition methods in these areas. The paper concludes by highlighting the complexity and challenges of applying affective computing in different domains and the need for further research to address these challenges. Overall, the objective of this paper is to illustrate the diversity of possible emotion recognition applications and draw conclusions on the challenges of automatic recognition that have to be addressed by further research.

Sarode et al (2010) presents a method for recognizing four facial expressions using a 2D appearance-based local approach. The algorithm implements Radial Symmetry Transform and edge projection analysis for feature extraction and creates a dynamic spatio-temporal representation of the face, followed by classification into one of the expression classes. The accuracy achieved by the algorithm for facial expression recognition from grayscale images is 81.0%. The paper concludes that the proposed method is an efficient, local image-based approach for extraction of intransient facial features and recognition of four facial expressions, and does not require any manual intervention.

Mellouk et al (2020) provides a comprehensive review of recent research on facial emotion recognition (FER) using deep learning techniques. The abstract highlights the importance of FER in human-machine interaction and the need for more accurate recognition of complex emotions. The methodology section describes different architectures of convolutional neural networks (CNN) and CNN-long short-term memory (LSTM) proposed by various researchers, as well as databases used for training and testing. The conclusion emphasizes the potential of deep learning to improve FER accuracy and the need for larger databases and more powerful architectures to recognize all basic and secondary emotions. Overall, the paper provides valuable insights into the current state and future directions of FER research.

Susskind et al (2007) compares the recognition performance of humans and support vector machines (SVMs) in identifying six basic emotional expressions. The methodology involved testing both human and computer performance on a set of stimuli, and comparing computer performance with human norms. The results reveal high accuracy in recognizing expression prototypes, providing insights into the potential compatibility of different views on facial expression recognition. The study's findings suggest that SVMs can be trained to recognize emotional expressions with high accuracy, and that this has implications for the development of facial recognition technology.

Ivanova et al (2020) discusses the optimization of machine learning algorithms for emotion recognition in terms of human facial expressions. The paper explores the challenges faced in developing biologically inspired cognitive architectures for emotion recognition and proposes a methodology for optimizing machine learning algorithms. The study concludes that the use of artificial intelligence-based biometric methods can help identify, measure, and analyze not only physical and facial features but also specific human behavior. The paper highlights the potential applications of emotion recognition technology in various industries, including security, healthcare, and marketing. Overall, the study provides valuable insights into the field of emotion recognition and its potential impact on society.

Tarnowski et al (2017) aimed to recognize seven emotional states based on facial expressions using a three-dimensional face model. The methodology involved presenting participants with sample images of actors displaying neutral, joy, surprise, anger, sadness, fear, and disgust, and asking them to mimic the expressions. The Kinect device was used to calculate six action units as features for classification. The k-NN classifier and MLP neural network were used for classification, with the MLP achieving higher accuracy. The study concluded that recognizing emotions based on facial expressions is feasible using a three-dimensional face model and machine learning techniques.

Brisimi et al (2018) discusses the use of federated learning to develop predictive models for hospitalizations due to cardiac events using Electronic Health Records (EHRs) from multiple institutions. The objective is to develop a decentralized optimization framework that enables multiple data holders to collaborate and converge to a common predictive model without explicitly exchanging raw data. The methodology involves using a soft-margin l1-regularized sparse Support Vector Machine (sSVM) classifier and an iterative cluster Primal Dual Splitting (cPDS) algorithm for solving the large-scale sSVM problem in a decentralized fashion. The conclusion is that federated learning can be an efficient and secure way to process large amounts of healthcare data stored in different locations, while ensuring privacy and security of sensitive healthcare data.

Young et al (2016) aimed to test the dimensional and category accounts of emotion recognition using photographic quality continua of interpolated facial expressions. Four experiments were conducted, with six emotions tested in each experiment. The methodology involved morphing images of facial expressions to create continua, which were then presented to subjects for recognition. The results showed that both dimensional and category accounts of emotion recognition were supported, with some evidence for a hybrid model. The study concluded that facial expressions are recognized based on both categorical and dimensional information, and that the use of continua can provide a more nuanced understanding of the underlying processes involved in emotion recognition.

Busso et al (2004) explores the potential for multimodal emotion recognition using facial expressions and speech, with the goal of improving human-computer interaction. The researchers analyze the strengths and weaknesses of unimodal systems and compare two fusion approaches: feature-level and decision-level fusion. They find that the bimodal emotion classifier outperforms each of the unimodal systems, and that the best fusion technique depends on the application. The primary objective of this research is to identify the advantages and limitations of unimodal systems and to show which fusion approaches are more suitable for emotion recognition.

Emerich et al (2009) presented at the 17th European Signal Processing Conference (EUSIPCO 2009) in Glasgow, Scotland. The authors, Simina Emerich, Eugen Lupu, and Anca Apatean from the Technical University of Cluj-Napoca in Romania, discuss different methods to improve the accuracy of emotion identification in the presence of environmental noise. They use bimodal systems which utilize both voice characteristics and facial expressions. Feature extraction and fusion techniques are implemented at several levels to simplify data and improve results. The information from voice and facial images are combined to form a single feature vector. The study also uses concepts like two-dimensional wavelets and moments for pattern recognition and image understanding.

Kamachi et al (2013) investigates the role of dynamic properties in recognizing facial expressions. The study used a within-subjects design with 26 Japanese adults to examine the effects of expression type and speed on recognition accuracy. The results showed that dynamic properties, such as speed, significantly influenced recognition accuracy, suggesting that expressive sequences convey dynamic information that is useful for decoding emotions portrayed. The study concludes that dynamic properties should be considered an important factor in recognizing facial expressions, and that future research should continue to explore the role of dynamic information in emotion perception.

Samal et al (1992) provides a comprehensive survey of the challenges and solutions related to automatic recognition and analysis of human faces and facial expressions. The abstract highlights the potential benefits of a face recognition system and the paper's goal to explore the design issues and solutions related to this technology. The methodology involves identifying the five basic problems that must be addressed, discussing the capabilities of the human visual system, surveying past work in these problems, and discussing possible directions for future research. The conclusion emphasizes the ongoing interest and importance of these problems to researchers and the need to consider the speed of recognition in many applications. Overall, this paper provides a valuable resource for understanding the challenges and solutions related to face recognition technology.

Ko et al (2018) provides a brief literature review of facial emotion recognition (FER) approaches, from conventional methods to recent advanced deep-learning-based approaches. The objective of this review is to provide a general understanding of the state-of-the-art FER approaches and help new researchers understand the essential components and trends in the FER field. The paper describes various standard databases for FER use, introduces representative categories of FER systems and their main algorithms, and compares key aspects between conventional FER and deep-learning-based FER in terms of accuracy and resource requirements. The paper also presents a hybrid deep-learning approach combining a convolutional neural network (CNN) and long short-term memory (LSTM) for temporal features of consecutive frames.

Wani et al (2021) provides a comprehensive review of Speech Emotion Recognition (SER) systems, which have become an integral component of human-computer interaction and other high-end speech processing systems. The paper identifies and synthesizes recent relevant literature related to the varied design components and methodologies of SER systems, providing readers with a state-of-the-art understanding of this hot research topic. The paper also highlights the contrasting quantitative and qualitative ways in which humans and machines recognize and correlate emotional aspects of speech signals, presenting enormous difficulties in blending knowledge from interdisciplinary fields. The paper concludes by outlining the research gap's prominence for consideration and analysis by other related researchers, institutions, and regulatory bodies. Overall, this paper provides valuable insights and information for researchers, developers, and anyone interested in the latest advances in speech processing.

Singh et al (2012) proposes a human emotion recognition system using neural networks and facial expressions. The methodology involves analyzing existing techniques, developing a simulator using MATLAB, and conducting tests to measure accuracy. The results show that the proposed system achieved up to 97% accuracy, and the use of neural networks improved results compared to fuzzy logic. The paper concludes that facial expressions can be used to sense human emotions and that future work should focus on adding fuzzy logic membership functions to the system. Overall, this paper highlights the importance of emotion recognition in communication and interaction between people and presents a promising approach using neural networks and facial expressions.

Hess et al (1998) aimed to investigate whether facial reactions to emotional facial expressions are due to affect or cognition. The study conducted two experiments with different rating tasks and physiological measures. The results showed that depending on the nature of the rating task, facial reactions may be either affective or cognitive.

The authors concluded that facial mimicry is not solely an automatic affective response, but can also be influenced by cognitive processes. Overall, this study provides insights into the complex interplay between affect and cognition in facial reactions to emotional facial expressions.

New et al (2003) proposes a text-independent method of speech emotion recognition using short time log frequency power coefficients and a discrete hidden Markov model as the classifier. The emotions are classified into six categories: Anger, Disgust, Fear, Joy, Sadness, and Surprise. The proposed system is compared to other systems that use fundamental frequency, energy contour, duration of silence, and voice quality. The database used to train and test the proposed system consists of 720 utterances from 60 speakers. The performance of the LFPC feature parameters is found to be better than that of the LPCC and mel-frequency Cepstral coefficients. The subjective assessment of the emotional speech corpus by human subjects is carried out to compare the accuracy of classification by the proposed system with human classification performance. The proposed system achieves an overall accuracy of 70.8%, which is comparable to the human classification performance of 71.9%. The conclusion is that the proposed system is effective in recognizing emotions in speech and can be used in various applications such as human-computer interaction, speech therapy, and affective computing.

Farroni et al (2007) investigates the perception of facial expressions in newborns. The authors conducted three experiments to explore whether newborns can distinguish between different emotional facial expressions. The methodology involved presenting infants with pictures of faces displaying different emotions and measuring their eye movements. The results suggest that newborns are capable of distinguishing between different emotional expressions, with a preference for happy faces. The authors conclude that these findings provide evidence for the early development of emotional processing in newborns. Overall, this study contributes to our understanding of the cognitive abilities of newborns and highlights the importance of emotional expression in early social interactions.

Wagner et al (1986) aimed to investigate whether observers can accurately distinguish between seven affective states based on facial expressions. Six senders viewed emotionally loaded photographic slides and were instructed to match their emotional reaction to one of seven emotions listed in front of them and call out two numbers indicating how pleasant or unpleasant and how strong or weak their emotional experience was. Fifty-three receivers were then shown the videotaped senders' facial expressions and asked to identify the emotion being expressed. The results showed that overall accuracy was greater than chance, with happy, angry, and disgusted expressions being recognized at above-chance rates. Female subjects were found to be better senders than male subjects.

Wallbott et al (1986) investigates the cues and channels used in portraying various emotions and how they affect judgment accuracy. The study used actors to portray six basic emotions through facial expressions, vocal expressions, and body movements. Participants were asked to identify the emotions portrayed and rate their confidence in their judgments. Results showed that facial expressions were the most accurate channel for decoding emotions, followed by vocal expressions and body movements. The study concludes that nonverbal cues play an important role in emotion recognition and that future research should consider the interplay between different channels and the context in which emotions are expressed.

Young-Browne et al (1977) explores how infants are able to discriminate between different facial expressions. The study involved 80 infants and used a visual habituation paradigm to measure their responses to different facial expressions. The researchers found that infants as young as 2 months old were able to discriminate between happy and sad facial expressions, but not between happy and surprised expressions. The study also found that the presentation of stimuli and assessment of responses posed methodological limitations. Overall, the study suggests that infants have an early ability to discriminate between different facial expressions, which may contribute to their early social development.

Zhang et al (2020) provides a comprehensive review of emotion recognition methods based on multi-modal data and machine learning techniques. The authors surveyed over 220 papers and discussed the state-of-the-art techniques proposed in recent years, including physiological data labeling approaches, feature extraction methods, feature dimensionality reduction algorithms, and classification performance of various machine learning models.

Ambadar et al (2005) explores the importance of facial dynamics in interpreting subtle facial expressions. The study conducted two experiments to investigate the effect of motion on emotion judgments and perception of change. Experiment 1 found that motion facilitates configural processing of faces and enhances the perception of subtle emotions. Experiment 2 revealed that unique temporal characteristics of facial expressions are more important than sensitivity to change in identifying subtle emotions. The study challenges the traditional focus on static displays of intense facial expressions in previous research and highlights the significance of motion in identifying and understanding the various expressions encountered in real-life situations. Overall, the study emphasizes the importance of considering facial dynamics in interpreting subtle facial expressions.

Camras et al (1977) explores the facial expressions used by children during conflict situations. The study involved 72 pairs of unacquainted kindergarten children who were given a brief play session in which conflict was induced. The children were seated opposite each other at a table with a vertical transparent Plexiglas divider and a pair of gerbils. The study found that children used a variety of facial expressions during conflict, including anger, sadness, and surprise. The research suggests that understanding these expressions can help adults better communicate with children during conflicts.

Chen et al (2012) proposes a three-level model for speech emotion recognition that combines Ekman's six basic emotions model and Fox's multi-level emotional model. The model uses appropriate features selected from 288 candidates and Fisher rate as input parameter for SVM. The experimental results show that Fisher is better than PCA for dimension reduction, and SVM is more expansible than ANN for speaker independent speech emotion recognition. The objective of this study is to establish a more accurate and efficient speech emotion recognition system. The methodology includes describing the databases used, feature extraction, dimension reduction, and the structure of the proposed multi-level system.

Vinola et al (2015) on Human Emotion Recognition Approaches, Databases and Applications. The abstract introduces the topic of affect detection and its importance in various fields. The methodology section discusses the different modalities and approaches used for affect detection, as well as the databases and applications that utilize these methodologies. The conclusion highlights the challenges and future directions for emotion recognition systems. The objective of the survey is to provide a comprehensive overview of the current state of research in this field. Overall, the paper provides valuable insights into the various methods and applications of human emotion recognition.

Calder et al (1996) explores the phenomenon of categorical perception of facial expressions, which refers to the tendency for people to perceive facial expressions as belonging to distinct categories rather than as continuous variations. The study uses computer-generated line-drawings and photographic-quality stimuli to investigate the effects of stimulus quality on categorical perception. The ABX task is used to measure discrimination between facial expressions, and the identification task is used to measure the ability to identify facial expressions. The results suggest that categorical perception of facial expressions is a robust phenomenon that is not affected by stimulus quality or short-term memory load.

Sullivan et al (2004) investigates emotion recognition deficits in healthy elderly adults and whether they are independent of declines in fluid ability. The study used three tasks (emotion, gender, and beaker) to test participants' recognition abilities. Results showed that elderly adults have difficulty recognizing certain emotions, such as anger and sadness, but perform similarly to younger adults in recognizing fear and happiness. The study suggests an age-related decline in emotion recognition abilities, which may have implications for healthcare and social settings.

REFERENCE

Adolphs, R., Damasio, H., Tranel, D., & Damasio, A. R. (1996). Cortical systems for the recognition of emotion in facial expressions. Journal of neuroscience, 16(23), 7678-7687.

Ahmed, T. U., Hossain, S., Hossain, M. S., ul Islam, R., & Andersson, K. (2019, May). Facial expression recognition using convolutional neural network with data augmentation. In 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR) (pp. 336-341). IEEE.

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. Psychological science, 16(5), 403-410.

Brisimi, T. S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I. C., & Shi, W. (2018). Federated learning of predictive models from federated electronic health records. International journal of medical informatics, 112, 59-67.

Busso, C., Deng, Z., Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., ... & Narayanan, S. (2004, October). Analysis of emotion recognition using facial expressions, speech and multimodal information. In Proceedings of the 6th international conference on Multimodal interfaces (pp. 205-211).

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perception of morphed facial expressions. Visual Cognition, 3(2), 81-118.

Camras, L. A. (1977). Facial expressions used by children in a conflict situation. Child development, 1431-1435.

Chen, L., Mao, X., Xue, Y., & Cheng, L. L. (2012). Speech emotion recognition: Features and classification models. Digital signal processing, 22(6), 1154-1160.

Cohen, I., Garg, A., & Huang, T. S. (2000, November). Emotion recognition from facial expressions using multilevel HMM. In Neural information processing systems (Vol. 2). State College, PA, USA: Citeseer.

Duncan, D., Shine, G., & English, C. (2016). Facial emotion recognition in real time. Computer Science, 1-7.

Dzedzickis, A., Kaklauskas, A., & Bucinskas, V. (2020). Human emotion recognition: Review of sensors and methods. Sensors, 20(3), 592.

El Ayadi, M., Kamel, M. S., & Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern recognition, 44(3), 572-587.

Emerich, S., Lupu, E., & Apatean, A. (2009, August). Emotions recognition by speechand facial expressions analysis. In 2009 17th European signal processing conference (pp. 1617-1621). IEEE.

Farroni, T., Menon, E., Rigato, S., & Johnson, M. H. (2007). The perception of facial expressions in newborns. European Journal of Developmental Psychology, 4(1), 2-13.

Frith, C. D. (2008). The social functions of consciousness. Frontiers of Consciousness: Chichele lectures, L. Weiskrantz and M. Davies, eds.(Oxford: Oxford University Press), 225-244.

Goldman, A. I., & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. Cognition, 94(3), 193-213.

Guanghui, C., & Xiaoping, Z. (2021). Multi-modal emotion recognition by fusing correlation features of speech-visual. IEEE Signal Processing Letters, 28, 533-537.

Hamann, S. B., Stefanacci, L., Squire, L. R., Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1996). Recognizing facial emotion. Nature.

Hao, N., Kilmer, M. E., Braman, K., & Hoover, R. C. (2013). Facial recognition using tensor-tensor decompositions. SIAM Journal on Imaging Sciences, 6(1), 437-463.

Hess, U., Philippot, P., & Blairy, S. (1998). Facial reactions to emotional facial expressions: Affect or cognition?. Cognition & Emotion, 12(4), 509-531.

Hua, W., Dai, F., Huang, L., Xiong, J., & Gui, G. (2019). HERO: Human emotions recognition for realizing intelligent Internet of Things. IEEE Access, 7, 24321-24332.

Hussain, S. A., & Al Balushi, A. S. A. (2020). 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.

Ioannou, S. V., Raouzaiou, A. T., Tzouvaras, V. A., Mailis, T. P., Karpouzis, K. C., & Kollias, S. D. (2005). Emotion recognition through facial expression analysis based on a neurofuzzy network. Neural Networks, 18(4), 423-435.

Ivanova, E., & Borzunov, G. (2020). Optimization of machine learning algorithm of emotion recognition in terms of human facial expressions. Procedia Computer Science, 169, 244-248.

Jack, R. E. (2010). Cultural differences in the decoding and representation of facial expression signals (Doctoral dissertation, University of Glasgow).

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. Current biology, 19(18), 1543-1548.

Kalsum, T., Anwar, S. M., Majid, M., Khan, B., & Ali, S. M. (2018). Emotion recognition from facial expressions using hybrid feature descriptors. IET Image Processing, 12(6), 1004-1012.

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2013). Dynamic properties influence the perception of facial expressions. Perception, 42(11), 1266-1278.

Kim, J. H., Kim, B. G., Roy, P. P., & Jeong, D. M. (2019). Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. IEEE access, 7, 41273-41285.

Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. sensors, 18(2), 401.

Kolakowska, A., Landowska, A., Szwoch, M., Szwoch, W., & Wrobel, M. R. (2014). Emotion recognition and its applications. Human-Computer Systems Interaction: Backgrounds and Applications 3, 51-62.

Kong, F. (2019). Facial expression recognition method based on deep convolutional neural network combined with improved LBP features. Personal and Ubiquitous Computing, 23, 531-539.

Koolagudi, S. G., & Rao, K. S. (2012). Emotion recognition from speech: a review. International journal of speech technology, 15, 99-117.

Landowska, A., Brodny, G., & Wrobel, M. R. (2017, April). Limitations of emotion recognition from facial expressions in e-learning context. In International Conference on Computer Supported Education (Vol. 2, pp. 383-389). Scitepress.

Li, B., & Lima, D. (2021). Facial expression recognition via ResNet-50. International Journal of Cognitive Computing in Engineering, 2, 57-64.

Li, J., Jin, K., Zhou, D., Kubota, N., & Ju, Z. (2020). Attention mechanism-based CNN for facial expression recognition. Neurocomputing, 411, 340-350.

Mao, Q. R., Pan, X. Y., Zhan, Y. Z., & Shen, X. J. (2015). Using Kinect for real-time emotion recognition via facial expressions. Frontiers of Information Technology & Electronic Engineering, 16(4), 272-282.

Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). SN Applied Sciences, 2(3), 446.

Mellouk, W., & Handouzi, W. (2020). Facial emotion recognition using deep learning: review and insights. Procedia Computer Science, 175, 689-694.

Naseem, I., Togneri, R., & Bennamoun, M. (2010). Linear regression for face recognition. IEEE transactions on pattern analysis and machine intelligence, 32(11), 2106-2112.

Nwe, T. L., Foo, S. W., & De Silva, L. C. (2003). Speech emotion recognition using hidden Markov models. Speech communication, 41(4), 603-623.

Paiva-Silva, A. I. D., Pontes, M. K., Aguiar, J. S. R., & de Souza, W. C. (2016). How do we evaluate facial emotion recognition?. Psychology & neuroscience, 9(2), 153.

Pantic, M., & Bartlett, M. S. (2007). Machine analysis of facial expressions (Vol. 558). INTECH Open Access Publisher.

Pantic, M., & Rothkrantz, L. J. M. (2000). Automatic analysis of facial expressions: The state of the art. IEEE Transactions on pattern analysis and machine intelligence, 22(12), 1424-1445.

Pighin, F., Auslander, J., Lischinski, D., Salesin, D. H., & Szeliski, R. (1997). Realistic facial animation using image-based 3D morphing. Microsoft Research.

Pighin, F., Hecker, J., Lischinski, D., Szeliski, R., & Salesin, D. H. (2006). Synthesizing realistic facial expressions from photographs. In Acm siggraph 2006 courses (pp. 19-es).

Poria, S., Majumder, N., Mihalcea, R., & Hovy, E. (2019). Emotion recognition in conversation: Research challenges, datasets, and recent advances. IEEE Access, 7, 100943-100953.

Samal, A., & Iyengar, P. A. (1992). Automatic recognition and analysis of human faces and facial expressions: A survey. Pattern recognition, 25(1), 65-77.

Sarode, N., & Bhatia, S. (2010). Facial expression recognition. International Journal on computer science and Engineering, 2(5), 1552-1557.

Sebe, N., Cohen, I., & Huang, T. S. (2005). Multimodal emotion recognition. In Handbook of pattern recognition and computer vision (pp. 387-409).

Sharma, S., Tomar, P., & Sharma, P. (2022, November). Emotional Recognition Through Facial Expression Using Support Vector Machine. In 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP) (pp. 248-253). IEEE.

Singh, D. (2012). Human emotion recognition system. International Journal of Image, Graphics and Signal Processing, 4(8), 50.

Song, N., Yang, H., & Wu, P. (2018, May). A gesture-to-emotional speech conversion by combining gesture recognition and facial expression recognition. In 2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia) (pp. 1-6). IEEE.

Sprengelmeyer, R., Rausch, M., Eysel, U. T., & Przuntek, H. (1998). Neural structures associated with recognition of facial expressions of basic emotions. Proceedings of the Royal Society of London. Series B: Biological Sciences, 265(1409), 1927-1931.

Sullivan, S., & Ruffman, T. (2004). Emotion recognition deficits in the elderly. International Journal of Neuroscience, 114(3), 403-432.

Susskind, J. M., Littlewort, G., Bartlett, M. S., Movellan, J., & Anderson, A. K. (2007). Human and computer recognition of facial expressions of emotion. Neuropsychologia, 45(1), 152-162.

Szwoch, M., & Pieniążek, P. (2015, June). Facial emotion recognition using depth data. In 2015 8th International Conference on Human System Interaction (HSI) (pp. 271-277). IEEE.

Tan, Y., Sun, Z., Duan, F., Solé-Casals, J., & Caiafa, C. F. (2021). A multimodal emotion recognition method based on facial expressions and electroencephalography. Biomedical Signal Processing and Control, 70, 103029.

Tarnowski, P., Kołodziej, M., Majkowski, A., & Rak, R. J. (2017). Emotion recognition using facial expressions. Procedia Computer Science, 108, 1175-1184.

Vinola, C., & Vimaladevi, K. (2015). A survey on human emotion recognition approaches, databases and applications. ELCVIA: electronic letters on computer vision and image analysis, 00024-44.

Wagner, H. L., MacDonald, C. J., & Manstead, A. S. (1986). Communication of individual emotions by spontaneous facial expressions. Journal of personality and social psychology, 50(4), 737.

Wallbott, H. G., & Scherer, K. R. (1986). Cues and channels in emotion recognition. Journal of personality and social psychology, 51(4), 690.

Wang, W., Xu, K., Niu, H., & Miao, X. (2020). Emotion recognition of students based on facial expressions in online education based on the perspective of computer simulation. Complexity, 2020, 1-9.

Wani, T. M., Gunawan, T. S., Qadri, S. A. A., Kartiwi, M., & Ambikairajah, E. (2021). A comprehensive review of speech emotion recognition systems. IEEE access, 9, 47795-47814.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (2016). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. In Facial Expression Recognition (pp. 67-110). Psychology Press.

Young-Browne, G., Rosenfeld, H. M., & Horowitz, F. D. (1977). Infant discrimination of facial expressions. Child Development, 555-562.

Zadeh, M. M. T., Imani, M., & Majidi, B. (2019, February). Fast facial emotion recognition using convolutional neural networks and Gabor filters. In 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI) (pp. 577-581). IEEE.

Zhang, H., Jolfaei, A., & Alazab, M. (2019). A face emotion recognition method using convolutional neural network and image edge computing. IEEE Access, 7, 159081-159089.

Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. Information Fusion, 59, 103-126.

Zhong, B., Qin, Z., Yang, S., Chen, J., Mudrick, N., Taub, M., ... & Lobaton, E. (2017, November). Emotion recognition with facial expressions and physiological signals. In 2017 IEEE symposium series on computational intelligence (SSCI) (pp. 1-8). IEEE.

Zou, W. W., & Yuen, P. C. (2011). Very low resolution face recognition problem. IEEE Transactions on image processing, 21(1), 327-340.