International Journal of Humanities Social Science and Management (IJHSSM)
Volume 3, Issue 1, Jan.-Feb. 2023, pp: 243-247 www.ijhssm.org

Facial Emotion Recognition

Vikas Maurya , Manorma Dwivedi , Rajat Singh , Deepak Maurya

Date of Submission: 08-01-2023 Date of Acceptance: 22-01-2023

Abstract: Facial emotion recognition (FER) is an emerging and significant research area in the pattern recognition domain. In daily life, the role of non-verbal communication is significant, and in overall communication, its involvement is around 55% to 93%. Facial emotion analysis is efficiently used in surveillance videos, expression analysis, gesture recognition, smart homes, computer games, depression treatment, patient monitoring, anxiety, detecting lies, psychoanalysis, paralinguistic communication, detecting operator fatigue and robotics. In this paper, we present a detailed review on FER. The literature is collected from different reputable research published during the current decade. This review is based on conventional machine learning (ML) and various deep learning (DL) approaches. Further, different FER datasets for evaluation metrics that are publicly available are discussed and compared with benchmark results. This paper provides a holistic review of FER using traditional ML and DL methods to highlight the future gap in this domain for new researchers. Finally, this review work is a guidebook and very helpful for young researchers in the FER area, providing a general understating and basic knowledge of the current state-of-the-art methods, and to experienced researchers looking for productive directions for future work.

Keyword- Facial Emotion recognition , Facial Expressions , Intelligence augment reality

I. Introduction

Facial emotions and their analysis play a vital role in non-verbal communication. It makes oral communication more efficient and conducive to understanding the concepts. It is also conducive to detecting human attention, such as behavior, mental state, personality, crime tendency, lies, etc. Regardless of gender, nationality, culture and race, most people can recognize facial emotions easily. However, a challenging task is the automation of facial emotion detection and classification. The research community uses a few basic feelings, such as fear, aggression, upset and pleasure. However, differentiating between many feelings is very challenging for machines . In addition, machines have to be trained well enough to understand the surrounding environment—specifically, an individual's intentions. When machines are mentioned, this term includes robots and computers. A difference is that robots involve communication abilities to a more innovative extent since their design consists of some degree of autonomy . The main problem is classifying people's emotions is variations in gender, age, race, ethnicity and image quality or videos. It is necessary to provide a system capable of recognizing facial emotions with similar knowledge as possessed by humans. Recently, FER has become an emerging field of research, particularly for the last few decades. Computer vision techniques, AI, image processing and ML are widely used to expand effective automated facial recognition systems for security and healthcare applications . Face detection is the first step of locating or detecting face(s) in a video or single image in the FER process.

Indeed, human beings can easily predict facial emotions and other facial features of an image, but these are difficult tasks for machines without excellent training . The primary purpose of face detection is to separate face images from the background (non-faces). Some face detection domains are gesture recognition, video surveillance systems, automated cameras, gender recognition, facial feature recognition, face recognition, tagging and teleconferencing . These systems need first to detect faces as inputs. There is a color sensor for image acquisition that captures color images everywhere. Hence, current face recognition techniques depend heavily on grayscale, and there are only a few techniques that are able operate with color images. To achieve better performance, these systems either implement window-based or pixel-based techniques, which are the key categories of strategies for facial recognition. The pixel-based method lags in separating the face from the hands of another skin region of the person In contrast, the window-based method loses the capacity to view faces from various perspectives. Model matching approaches are the most frequently used techniques

International Journal of Humanities Social Science and Management (IJHSSM)
Volume 3, Issue 1, Jan.-Feb. 2023, pp: 243-247 www.ijhssm.org

for FER, including face detection. In contrast, the window-based approach cannot view faces from different angles. Currently, several state-of-the-art classifiers, such as ANN, CNN, SVM, KNN, and random forest (RF), are employed for different features' extraction and in the recognition of tumors in healthcare, in biometrics, in handwriting studies and in detecting faces for security measures. Main Contributions Over the past three decades, there has been a lot of research reported in the literature on facial emotion recognition (FER). Nevertheless, despite the long history of FER-related works, there are no systematic comparisons between traditional ML and DL approaches. Kołakowska presented a review whose main focus was on conventional ML approaches. Ghayoumi proposed a short review of DL in FER, and Ko et al. presented a stateof-the-art review of facial emotion recognition techniques using visual data. However, they only provided the differences between conventional ML techniques and DL techniques. This paper presents a detailed investigation and comparisons of traditional ML and deep learning (DL) techniques with the following major contributions:

• The main focus is to provide a general understanding of the recent research and help newcomers to understand the essential modules and trends in the FER field.

• We present the use of several standard datasets consisting of video sequences and images with different characteristics and purposes.

• We compare DL and conventional ML approaches for FER in terms of resource utilization and accuracy. The DL-based approaches provide a high degree of accuracy but consume more time in training and require substantial processing capabilities, i.e., CPU and GPU. Thus, recently, several FER approaches have been used in an embedded system, e.g., Raspberry Pi, Jetson Nano, smart phone, etc.

This paper provides a holistic review of facial emotion recognition using the traditional ML and DL methods to highlight the future gap in this domain for new researchers. Further, Section 1 presents the related background on facial emotion recognition; Sections 2 and 3 contain a brief review of traditional ML and DL.

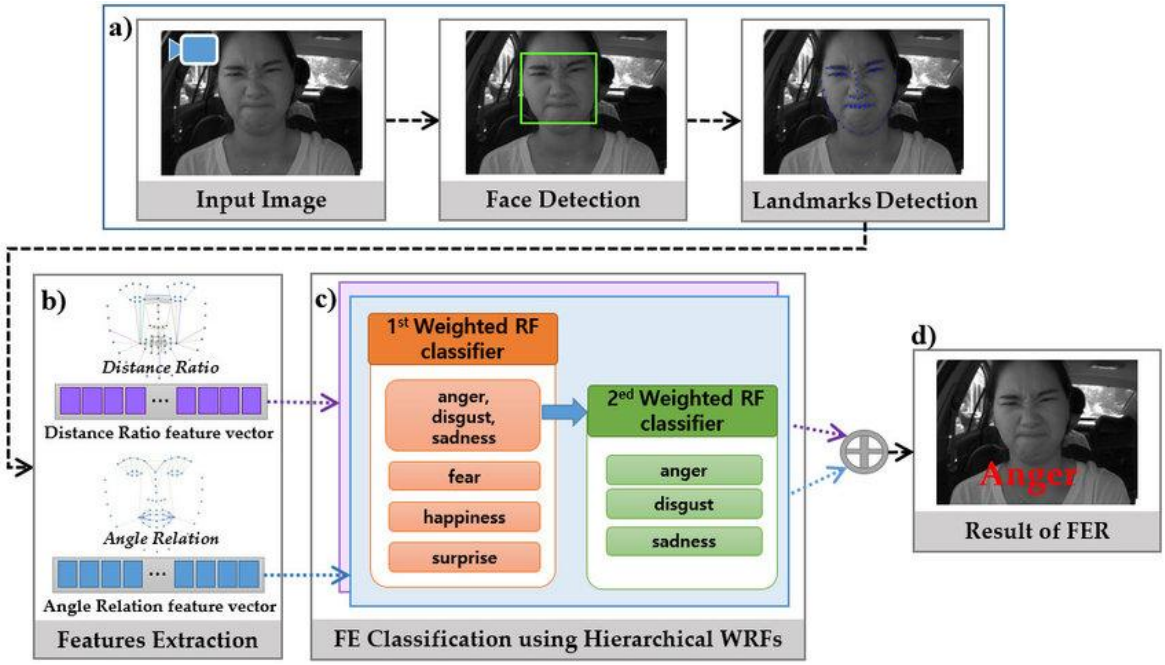
II. Facial Emotion Recognition Using Traditional Machine Learning Approaches

Facial emotions are beneficial for investigating human behavior as exhibited. Psychologically, it is proven that the facial emotion recognition process measures the eyes, nose, mouth and their locations. The earliest approach used for facial emotion intensity estimation was based on distance urged. This approach uses high-dimensional rate transformation and regional volumetric distinction maps to categorize and quantify facial expressions. In videos, most systems use Principal Component Analysis (PCA) to represent facial expression Information 2022, 13, 268 3 of 17 features . PCA has been used to recognize the action unit to express and establish different facial expressions. Other facial expressions are structured and recognized by mistreatment PCA for providing a facial action. Information 2022, 13, x FOR PEER REVIEW 3 of 19 2. Facial Emotion Recognition Using Traditional Machine Learning Approaches Facial emotions are beneficial for investigating human behavior as exhibited. Psychologically, it is proven that the facial emotion recognition process measures the eyes, nose, mouth and their locations. Facial emotion recognition (FER) process. The earliest approach used for facial emotion intensity estimation was based on distance urged. This approach uses high-dimensional rate transformation and regional volumetric distinction maps to categorize and quantify facial expressions. In videos, most systems use Principal Component Analysis (PCA) to represent facial expression features PCA has been used to recognize the action unit to express and establish different facial expressions. Other facial expressions are structured and recognized by mistreatment PCA for providing a facial action. detected and extracted the face portion via the active contour model. The researchers used Chan–Vese and Bhattacharyya's energy functions to optimize the distance.

Several traditional methods exist are used for the extraction facial features such as geometric and texture features for example local binary patterns LBP facial action local directional patterns LDA Gabor wavelet. In recent years, deep learning has been very successful and efficient approach thanks to the result obtained with its architectures which allow the automatic extraction of features and classification such as the convolutional neural network CNN and the recurrent neural network RNN; here what prompted researchers to start using this technique to recognize human emotions. Several efforts are made by researchers on the development of deep neural network architectures, which produce very satisfactory results in this area. In this paper, we provide a review of recent advances in sensing emotions by recognizing facial expressions using different deep learning architectures. We present recent results from 2016 to 2019 with an

International Journal of Humanities Social Science and Management (IJHSSM)
Volume 3, Issue 1, Jan.-Feb. 2023, pp: 243-247 www.ijhssm.org

interpretation of the problems Sand contributions. It is organized as follows: in section two, we introduce some available public databases, section three; we present a recent state of the art on the FER using deep learning and we end in section four and five with a discussion and comparisons then a general conclusion with the future works.

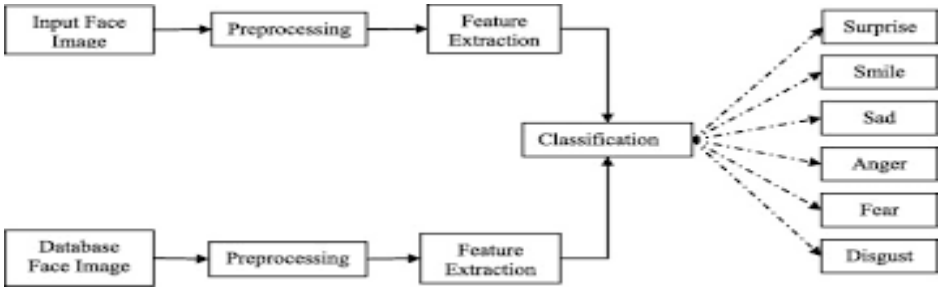


III. Facial available databases

One of the success factors of deep learning is the training the neuron network with examples, several FER databases now available to researchers to accomplish this task, each one different from the others in term of the number and size of images and videos, variations of the illumination, population and face pose. Some presented in the Table.1 in which we will note its presence in the works cited in the following section.

Experts from different institutions have generated several datasets to evaluate reported methods for facial expression classification. Accordingly, a detailed overview of some benchmark datasets is presented. • The CK+ Cohen Kanade Dataset: The Cohen Kanade database is inclusive in that it consists of subject images of all sexes and races and is open to the public. This dataset consists of seven essential emotions that often include neutral emotions. The images' resolution is 640 × 490 or 640 × 480 dimensions, with grayscale (8-bit) existing in the dataset. Approximately 81% of subjects are Euro-American, 13% are Afro American, and approximately 6% are from other races. In the CK+ dataset, the ratio of females is nearly 65%. The dataset consists of 593 images captured from 120 different people, and the age of these people varies from 18 to 30 years. • Bosphorus dataset This dataset consists of 2D and 3D faces for emotion recognition, facial action detection, 3D face reconstruction, etc. There are 105 humans with 4666 faces in different poses in the dataset. This dataset is different from other datasets in the following aspects. 1. A rich collection of facial emotions is included: i. Per person, at least 35 facial emotions are recoded; ii. FACS scoring; iii. Each third person is a professional actresses/actors. 2. Systematic head poses are included. 3. A variety of facial occlusions are included (eyeglasses, hands, hair, moustaches and beards)

International Journal of Humanities Social Science and Management (IJHSSM)
Volume 3, Issue 1, Jan.-Feb. 2023, pp: 243-247 www.ijhssm.org



IV. Literature Review

Emotion plays a significant role in human beings daily lives. Humans can easily sense a person's emotions. But in some cases devices need to sense people's emotions. Machine learning is a sub-part of artificial intelligence that produces robots handling tasks like us. Emotion recognition is a small module that can be easily achieved by machines using machine learning algorithms. This paper describes the various algorithms used to recognize the facial expressions of a person such as happy, angry, sad, disgust, neutral, fear. Gabor filters and Local Binary Pattern Operators (LBP) are discussed for the process of feature extraction. Different types of classification algorithms such as Support Vector Machines, K-Nearest Neighbors are discussed. The training of the image data is carried by comparing various neural networks including Attentional Neural Network, Convolutional neural network, shallow neural network etc.

V. Conclusions and Future Work

In this paper, a detailed analysis and comparison are presented on FER approaches. We categorized these approaches into two major groups: (1) conventional ML-based approaches and (2) DL-based approaches. The convention ML approach consists of face detection, feature extraction from detected faces and emotion classification based on extracted features. Several classification schemes are used in conventional ML for FER, consisting of random forest, AdaBoost, KNN and SVM. In contrast with DL-based FER methods, the dependency on face physics-based models is highly reduced. In addition, they reduce the preprocessing time to enable "end-to-end" learning in the input images. However, these methods consume more time in both the training and testing phases. Although a hybrid architecture demonstrates better performance, micro-expressions remain difficult tasks to solve due to other possible movements of the face that occur unwillingly. Additionally, different datasets related to FER are elaborated for the new researchers in this area. For example, human facial emotions have been examined in a traditional database with 2D video sequences or 2D images. However, facial emotion recognition based on 2D data is unable to handle large variations in pose and subtle facial behaviors. Information 2022, 13, 268 14 of 17 Therefore, recently, 3D facial emotion datasets have been considered to provide better results. Moreover, different FER approaches and standard evaluation metrics have been used for comparison purposes, e.g., accuracy, precision, recall, etc. FER performance has increased due to the combination of DL approaches. In this modern age, the production of sensible machines is very significant, recognizing the facial emotions of different individuals and performing actions accordingly. It has been suggested that emotion-oriented DL approaches can be designed and fused with IoT sensors. In this case, it is predicted that this will increase FER's performance to the same level as human beings, which will be very helpful in healthcare, investigation, security and surveillance.

References

[1]. E. Sariyanidi, H. Gunes, et A. Cavallaro, « Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and Recognition », IEEE Trans. Pattern Anal. Mach. Intell., oct. 2014, doi: 10.1109/TPAMI.2014.2366127.

[2]. C.-N. Anagnostopoulos, T. Iliou, et I. Giannoukos, « Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011 », Artif. Intell. Rev., vol. 43, no 2, p. 155-177, févr. 2015, doi: 10.1007/s10462-012-9368-5.

[3]. L. Shu et al., « A Review of Emotion Recognition Using Physiological Signals », Sensors, vol. 18, no 7, p. 2074, juill. 2018, doi: 10.3390/s18072074.

[4]. C. Marechal et al., « Survey on AI-Based Multimodal Methods for Emotion Detection », in High-Performance Modelling and Simulation for Big Data Applications:

Selected Results of the COST Action IC1406 cHiPSet, J. Kołodziej et H. González-Vélez, Éd. Cham: Springer International Publishing, 2019, p. 307-324.

[5]. M. H. Alkawaz, D. Mohamad, A. H. Basori, et T. Saba, « Blend Shape Interpolation and FACS for Realistic Avatar », 3D Res., vol. 6, no 1, p. 6, janv. 2015, doi: 10.1007/s13319-015-0038-7.

[6]. P. V. Rouast, M. Adam, et R. Chiong, « Deep Learning for Human Affect Recognition: Insights and New Developments », IEEE Trans. Affect. Comput., p. 1-1, 2018, doi: 10.1109/TAFFC.2018.2890471.

[7]. C. Shan, S. Gong, et P. W. McOwan, « Facial expression recognition based on Local Binary Patterns: A comprehensive study », Image Vis. Comput., vol. 27, no 6, p. 803-816, mai 2009, doi: 10.1016/j.imavis.2008.08.005.

[8]. T. Jabid, M. H. Kabir, et O. Chae, « Robust Facial Expression Recognition Based on Local Directional Pattern », ETRI J., vol. 32, no 5, p. 694 Wafa Mellouk et al. / Procedia Computer Science 175 (2020) 689–694 6 Author name / Procedia Computer Science 00 (2018) 000–000 784-794, 2010, doi: 10.4218/etrij.10.1510.0132.

[9]. S. Zhang, L. Li, et Z. Zhao, « Facial expression recognition based on Gabor wavelets and sparse representation », in 2012 IEEE 11th International Conference on Signal Processing, oct. 2012, vol. 2, p. 816-819, doi: 10.1109/ICoSP.2012.6491706.

[10]. R. Gross, I. Matthews, J. Cohn, T. Kanade, et S. Baker, « Multi-PIE », Proc. Int. Conf. Autom. Face Gesture Recognit. Int. Conf. Autom. Face Gesture Recognit., vol. 28, no 5, p. 807-813, mai 2010, doi: 10.1016/j.imavis.2009.08.002.

[11]. M. Pantic, M. Valstar, R. Rademaker, et L. Maat, « Web-based database for facial expression analysis », in 2005 IEEE International Conference on Multimedia and Expo, juill. 2005, p. 5 pp.-, doi: 10.1109/ICME.2005.1521424.

[12]. M. F. Valstar, B. Jiang, M. Mehu, M. Pantic, et K. Scherer, « The first facial expression recognition and analysis challenge », in Face and Gesture 2011, mars 2011, p. 921-926, doi: 10.1109/FG.2011.5771374.