Facial Emotion Detection

1st Adnan Ali
Information Technology
Ajay Kumar Garg Engineering College
Ghaziabad, India
aliadnanjnv00786@gmail.com

2nd Abhishek Mohile
Information Technology
Ajay Kumar Garg Engineering College
Ghaziabad, India
abhishek2113121@akgec.ac.in

3rd Abhishek Gupta
Information Technology
Ajay Kumar Garg Engineering College
Ghaziabad, India
guptaabhishek9717@gmail.com

4th Arshad Raza
Information Technology
Ajay Kumar Garg Engineering College
Ghaziabad, India
arshadraza11199@gmail.com

***Abstract*—Facial emotions on a person face give insights on various aspects about a person, they may be their physical or mental health, or even their personalities. Facial clues also help identify a person's emotions. This article presents a comparative study of different techniques which can be used for facial emotion detection and recognition of students and office workers. Facial emotion recognition system can be a tool for educators, providing them the emotion dynamics of a class, can monitor the emotional state of employees, and identify the best suited candidates for various high stress job positions. This article presents a comparative study of the techniques which can be used for facial emotion recognition and identifies the best techniques.**

***Index Terms*—Emotion, Face Detection, Recognition**

I. Introduction

In today's world identifying the emotions of an individual plays a very important role and recognizing these emotions can lead to noticeable improvements in society i.e., education, work environments etc. Emotional intelligence plays a vital role in fostering empathy, enhancing communication, and building stronger connections between individuals and groups. There may be several problems if the society fails to understand the importance of emotions like – failing to provide guidance to children when they need it, neglecting the mental health of employees, appointing incapable individuals for certain job roles, etc.

In a class, observing the emotions of each child is important. Any extreme behavior might be indicative of nutritional deficiencies or psychological or personal problems. Often times, deficiency of essential minerals or vitamins can compromise brain function and increase symptoms of depression, irritability, anxiety, etc. A child might also need the guidance of a parent, teacher or even professional counselling to resolve personal or psychological problems. In schools or colleges, if a majority of the class shows negative emotions in a certain lecture there might be a need to re-evaluate the teacher's behavior and teaching methods. According to the report, People at Work 2023: A Global Workforce View, by ADP Research Institute, about 65 percent people say stress adversely affects their work (source: https://www.adpri.org/assets/people-at-work-2023-a-global-workforce-view/). In offices, if a majority of employees exhibit negative emotions for most of the time, it tells the organization to readjust the work pressure and help employees with their stress levels. When companies recruit for highly stressful job positions, it may be helpful to observe the candidates waiting for interviews and see their prevalent emotions in the highly stressful situation. This will help find candidates fit for the job role. It thus becomes extremely important to recognize facial emotions of individuals. In this review paper, a comparison of different techniques used for facial expression recognition is made. The best technique is identified.

II. Related Work

Gurudutt Perichetla, Akash Saravanan, Dr. K.S.Gayathri [1], in 2019 addressed the challenge of classifying images of human faces into seven basic emotions, focusing on real-time facial emotion recognition. The objective was to achieve accurate classification of universal emotions using a Convolutional Neural Network (CNN), with the exploration of various models to enhance performance in facial emotion recognition. The proposed model consisted of six convolutional layers, two max pooling layers and two fully connected layers. Upon tuning of the various hyperparameters, this model achieved a final accuracy of 60 percent which was nearly state of the art.

Marco Del Coco†, Pierluigi Carcagni, Cosimo Distante and Marco Leo [2] in 2015 demonstrated that, with an optimized set of HOG parameters, this descriptor can effectively characterize facial expression peculiarities, resulting in enhanced FER performance compared to commonly used frameworks. This technique resulted in 72.2 percent accuracy on CK+ dataset. The study contributed to overcoming the obstacle of achieving robust and accurate FER, particularly crucial for applications in human-robot interaction and other assistive technologies.

Shaogang Gong, Peter W-McOwan, Caifeng Shan McOwan [3] in 2008 explored the application of Local Binary Patterns (LBP) as a facial representation method, evaluated its performance across various databases and addressed issues related to person-independent recognition and low-resolution facial expression scenarios in real-world environments. This technique resulted in a testing accuracy of 51.1 percenton MMI dataset and 41.3 percent on JAFEE dataset.

III. LITERATURE SURVEY

Techniques used for Facial Emotion Detection.

1) Convolution Neural Networks (CNN): In this method, the convolutional layers extract facial features and identify hidden patterns in the data, an optimizer is used so that the weights can be adjusted during the backpropagation process, followed by the pooling layer which downsizes the data, without losing significant information. This is followed by fully connected layers for emotion classification (Gurudutt Perichetla, Akash Saravanan, Dr. K.S.Gayathri, 2019 [1]).

2) Histogram of Oriented Gradients (HOG): Histogram of Oriented Gradients (HOG) involves face detection and registration, followed by feature extraction using HOG descriptor. The extracted features are then input to Support Vector Machines (SVMs) for classification, and a temporal analysis is applied to enhance accuracy by considering consistency in facial expression predictions over a temporal window (Marco Del Coco†, Pierluigi Carcagni, Cosimo Distante and Marco Leo,2015 [2]).

3)Transfer learning: Transfer learning in the facial expression recognition system involves selecting pre-trained neural networks (VGG16, VGG19, InceptionResNetV2) and implementing two classifier architectures, one based on Global Average Pooling (GAP) and the other with a convolution layer (Fátima Rodrigues and José Almeida, 2021 [4]).

4)Attentional Convolutional Network: Attentional convolutional network for facial expression recognition incorporates less than 10 layers, including convolutional and fully connected layers, along with a spatial transformer network for attention on the crucial regions of image (Mehdi Minaei, Shervin Minaee and Amirali Abdolrashidi,2021 [7]).

5)Hybrid Approach (Transfer Learning + Attentional Convolutional Network): The Hybrid Approach (Transfer Learning + Attentional Convolutional Network) consisted of MobileNetv1, Patch Extraction Block and Attention classifier. MobileNetv1 is a pretrained model which acted as the backbone of this Hybrid model, then a custom-built Feature extractor was used which was named as Patch Extractor and finally an Attention classifier was used for better classification of emotions (Kiam Ming Kim, Jia Le Ngwe, Thian Song Ong, Chin Poo Lee 2023 [17]).

IV. DISCUSSION

Understanding and addressing the emotional well-being of individuals in various contexts is paramount for fostering healthier and more productive environments. Emotions play a significant role in shaping our daily lives, from our early years in the classroom to our careers in the corporate world. This study compares various methods for facial emotion recognition, highlighting their respective advantages and limitations. The comparisons between different techniques are shown in Table I.

The first technique is Convolution Neural Networks (Gurudutt Perichetla, Akash Saravanan, Dr. K.S.Gayathri, 2019 [1]) which employs extraction of facial features, also identifying hierarchical patterns in data. It then uses fully connected layers for emotion classification. The advantage of this Convolutional Neural Network (CNN) is that it is highly accurate at image recognition tasks as well as facial emotion recognition but it requires a large amount of training data.

The second technique is Transfer learning (Fátima Rodrigues and José Almeida, 2021 [4]) which conforms pre-trained model knowledge, fine-tunes on target data and by utilizing this learning it improves performance in new tasks. The advantage of this technique is reduced effort and time required for creating new models.

The third technique is Attentional Convolutional Network (Mehdi Minaei, Shervin Minaee and Amirali Abdolrashidi,2021 [7]) which extracts features with CNNs, focuses on relevant regions of an image and also classifies emotions based on weighted features. The advantage of ACNs is that it focuses on improving accuracy in tasks like facial emotion recognition but it is computationally expensive to train and run for large dataset.

The fourth technique is a Hybrid approach (Kiam Ming Kim, Jia Le Ngwe, Thian Song Ong, Chin Poo Lee,2023 [17]) consisting of Transfer Learning and Attentional Convolutional Network. The main advantage of this technique is that it is light weight as compared to previous state-of-the-art models but it is sensitive to incorrectly labelled or noisy data.

V. CONCLUSION

There has been significant advancement in facial emotion detection techniques in recent years. This paper focuses on leveraging Convolutional Neural Networks (CNN) and Artificial Neural Networks (ANN) in combination with Transfer Learning to improve the accuracy of emotion detection using the FER-2013 dataset.

These methods, when applied together, capitalize on the feature extraction capabilities of CNN, the flexibility of ANN for learning complex patterns, and the efficiency of Transfer Learning to enhance performance with limited data.

The availability of larger, correctly annotated datasets, coupled with advancements in data preprocessing methods, will play a pivotal role in the development of more robust emotion detection systems. These datasets, enriched with diverse and accurate labels, can significantly improve the training and testing phases of machine learning models, ensuring that they generalize well across different scenarios and demographics. Additionally, the incorporation of advanced techniques, such as Attentional Neural Networks and transformer-based architectures, introduces the ability to focus on critical features

S.no.	*Techniques*	*Methodology used*	*Advantages*	*Disadvantages*
1.	Convolution Neural Networks (CNN)	Extract facial features, apply convolution layers for hierarchical pattern learning, followed by fully connected layers for emotion classification.	CNNs are highly accurate at image recognition tasks, including facial emotion recognition.	Require large amount of training data.
2.	Transfer learning	Adapt pre-trained model knowledge, fine-tune on target data, utilizing transfer learning for improved performance in new tasks.	Transfer Learning can significantly reduce the time and effort required for developing new models.	Transfer Learning can lead to overfitting if the source domain and target domain are not sufficiently similar.
3.	Attentional Convolutional Network	Extract features with CNNs, apply attention mechanism to focus on relevant regions, and classify emotions based on weighted features.	ACNs can focus on relevant regions of an image, improving accuracy in tasks like facial emotion recognition.	ACNs can be computationally expensive to train and run, especially for large datasets.
4.	Hybrid Approach (Transfer Learning + Attentional Convolutional Network)	MobileNetv1 serves as feature extractor, patch extraction block extracts meaningful local features followed by an attention classifier for better classification.	This state-of-the-art model is more lightweight than previous state-of-the-art models.	Extremely sensitive to incorrectly annotated/ noisy input data.

and patterns within the data, thereby enhancing the overall precision and reliability of these models.

VI. Acknowledgment

The authors thank Mrs. Chelsi Sen for mentorship.

VII. References

[1] Bin Zhang, Jian Li, Ying Tai, Yabiao Wang, Chengjie Wang, Zhenyu Zhang, Xiaoming Huang, Jilin Li, Yili Xia, proposed a Face Detection model ,arXiv preprint arxiv:2201.10781v1[cs.CV]

[2]Oussama Abdul Hay, Mohamad Alansari, Abdulhadi Shoufan, Sajid Javed, Naoufel Werghi, Yahya Zweiri (2023) proposed a model for Face Recognition, doi:10.1109/ACCESS.2023.3266068

[3]Kiam Ming Kim, Jia Le Ngwe, Thian Song Ong, Chin Poo Lee (2023), proposed a model for Facial Emotion Recognition, arXiv preprint arXiv:2306.0962v1[cs.CV]

[4]Dumitru Erhan2, Ian J. Goodfellow1, Aaron Courville, Pierre Luc Carrier, Ruifan Li, Ben Hamner, Mehdi Mirza, Yichuan Tang, Will Cukierski, Dong-Hyun Lee, David Thaler, Chetan Ramaiah, Fangxiang Feng, Yingbo Zhou, Dimitris Athanasakis, Xiaojie Wang, Maxim Milakov, John Shawe-Taylor, Radu Ionescu, John Park, Cristian Grozea, Marius Popescu, Jingjing Xie, James Bergstra, Bing Xu, Lukasz Romaszko, Yoshua Bengio, Zhang Chuang (2013), paper titled " Challenges in Representation Learning: A report on three machine learning

[5]N. Yadav, K. Banerjee, and V. Bali, "A survey on fatigue detection of workers using machine learning," International Journal of E-Health and Medical Communications (IJEHMC), vol. 11, no. 3, pp. 1-8, 2020.

[6]T. Sharma, K. Banerjee, S. Mathur, and V. Bali, "Stress analysis using machine learning techniques," International Journal of Advanced Science and Technology, vol. 29, no. 3, pp. 14654-14665, 2020.

[7]K. Banerjee, M. S. Kumar, and L. N. Tilak, "Delineation of potential groundwater zones using Analytical hierarchy process (AHP) for Gautham Buddh Nagar District, Uttar Pradesh, India," Materials Today: Proceedings, vol. 44, pp. 4976-4983, 2021.

[8]K. Banerjee and R. A. Prasad, "Reference based inter chromosomal similarity based DNA sequence compression algorithm," in 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017, pp. 234-238. doi: 10.1109/CCAA.2017.8229806.

[9]K. Banerjee and R. A. Prasad, "A new technique in reference based DNA sequence compression algorithm: Enabling partial decompression," in AIP Conference Proceedings, vol. 1618, no. 1, pp. 799-802, 2014. doi: 10.1063/1.4897025.

[10]K. Banerjee and V. Bali, "Design and Development of Bioinformatics Feature Based DNA Sequence Data

Compression Algorithm," EAI Endorsed Trans. Pervasive Health Technol., vol. 5, no. 20, p. e5, 2020.

[11]K. Banerjee, M. B. Santhosh Kumar, L. N. Tilak, and S. Vashistha, "Analysis of Groundwater Quality Using GIS-Based Water Quality Index in Noida, Gautam Buddh Nagar, Uttar Pradesh (UP), India," in Applications of Artificial Intelligence and Machine Learning, A. Choudhary, A. P. Agrawal, R. Logeswaran, and B. Unhelkar, Eds. Singapore: Springer, 2021, pp. 205-212. doi: 10.1007/978-981-16-3067-$5_1$4.