Facial Emotion Recognition And Detection

Sunil Kumar
Associate Professor,
Department of Computer Science and Engineering
Galgotias College of Engineering and Technology
Greater Noida,.
Uttar Pradesh, India
sunilkumar@galgotiacollege.edu

Neeraj Kumar Gupta
Department of Computer Science and Engineering
Galgotia College of Engineering and Technology
Greater Noida,
Uttar Pradesh, India
neerajguptabhadohi@gmail.com

Swapnil Shukla
Department of Computer Science and Engineering
Galgotia College of Engineering and Technology
Greater Noida,
Uttar Pradesh, India
shuklaswapnil009@gmail.com

Shreyash Tiwari
Department of Computer Science and Engineering
Galgotia College of Engineering and Technology
Greater Noida,
Uttar Pradesh, India
shreyashtiwari15945@gmaill.com

***Abstract* — Facial emotion detection and recognition remains a challenging and interesting in the world of computer vision. Recognizing face expression was always tough nut to crack out as the way a facial expression expressed by a person varies from one man to another. Deep learning is a most reliable technology for the study of facial expression and deep neural network is used to classify the images according to different emotions. Convolutional neural network is the technology helped to classify the images. In this paper we are introducing a new junction of these technologies for the recognition of facial emotions. In recent years, facial emotion recognition has witnessed a great surge in popularity, motivating researchers to dive deeper into this technology, getting promising outcomes in the past. There are several models which provide an overview regarding facial emotions. The main difficulty in this to identify the subtle emotions and the need of some large datasets which are also discussed in this paper. We also highlight some of the important aspect of facial emotion recognition system in many other real life fields and have lot of scope in future research. The idea is to give overview about this field and and its role and advancement in branch of artificial intelligence and computer vision.**

Keywords— CNN, Computer vision, deep learning, Tensor Flow, Keras,Machine Learning,Emotion.

I. INTRODUCTION

As we know that people communicate through their facial expression not often then most of the time.

Numerous studies have been done in past showing that 50% of the emotions through our face directly what we are feeling. The emotions are 10 times more effective than speech. In todays technologically connected world in which we are living, intelligent monitoring has emerged drastically. For example, cameras and assistive robots need to understand how human's emotions work. For humans recognising the emotion is quite a easy task but for the most intelligent AI technology it is quite hard. The main problem that come is varieties in emotions are main research area for psychologists and scientists. As of research given by the psychologist Ekman (1993) on classification of emotions developed seven face expression on the basis of movement of facial muscles, eyes. In real- life situations, facial emotions are crucial because they convey vital information. FER can be used in a variety of fields, including healthcare, human resources, law enforcement, education (via students during lectures), customer service, media (during interviews), and many .

This study uses deep convolutional neural network to identify seven basic human expressions using seven classic facial expressions (mental states such as pain, fatigue, lying, irritation, agreement, rejection, and remorse). emotions and provides a framework for understanding other mixed emotions.

II. LITERATURE REVIEW:

Facial emotion detection and Recognition is a emerging research field at the junction of computer vision, artificial intelligence and psychology. This is showing the advancement we have done to understand and interact with the emotions using the technology to decipher them. This review offers concise overview of current facial emotion detection and recognition, with the aim of summarizing and critically upraising the current research state, methods and applications in this field. Fields like Robotics, Pharmaceutical, driver assistance systems, and lie detectors uses methods like facial expression analysis. Recent facial emotion recognition system has advanced Scope in Science fields like neuroscience and cognitive science. FER has enhanced with utility and accuracy as well because of advancement in computer vision (CV) and Machine Learning (ML).

Certainly, here`s what I found from different sources like deep learning, CNN, SVM and databases like CK+, FER 2013, and professionals across diverse fields. Understanding it through facial hints has great implementation for fields as diverse as healthcare, education, security and marketing.

979-8-3503-7378-3/24/$31.00 ©2024 IEEE

A. Deep Learning:

Deep Learning has gained a recent past due to its great capability. The main focus on artificial neural networks. They are made of layers of attraction great in interconnected units called neurons, which work together to process and transform data. Deep learning is characterized by the use of deep architectures system with many interconnected layers between input and output. These deep layers allow the network to autonomously learn complex features and representations from raw data, making it ideal for complex tasks.

One of the most remarkable capabilities of deep learning is eliminating the need for human intervention in feature engineering. In traditional machine learning, feature engineering often requires manual and time-consuming efforts to extract meaningful features from data. In contrast, deep learning models can autonomously identify and leverage relevant features, reducing the need for human intervention in this aspect of model development.

B. Convolutional Neural Network:

Convolutional neural networks (CNNs) have grown as a transformative force in area of computer vision, increasing our ability to understand and interpret visual information. Unlike traditional artificial intelligence models that depend on hand- crafted features, CNNs possess the remarkable capability of automatically extracting meaningful patterns and relationships directly from raw pixel data. CNN scan the input data and detect local patterns, edges, and textures. The pooling layer down samples the feature maps to reduce computational load and improve transformation invariance.

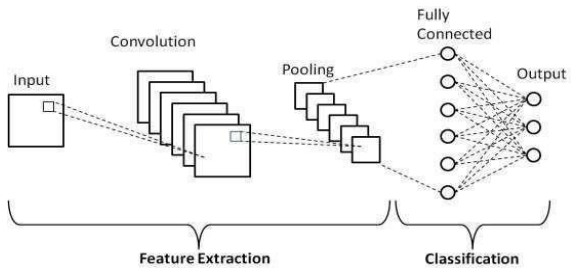

Fig.1 CNN Classification.[22]

III. RELATED WORK:

Many researchers are working on ways to improve facial emotion recognition systems. One team, Tang AI[3], has developed a system that can analyze facial expressions to evaluate the effect of something. The system has five steps: data capture, face detection, facial recognition, facial feature detection, and post- processing. It uses a classification method called KNN and a facial expression analysis technique called ULGBPHS.

Savva et al [4] research have developed systems that can automatically analyze student emotions during face-to-face instruction. Whitehill et al[5] proposed asystem that uses features and SVM model to identify facial expressions. Other researchers[7] have used ML and Computer Vision to identify facial expressions, while another system checks emotions and give real-time feedback. This system uses eye and head patterns to gather information.

Ayvaz et al [8] "created a model called FERS that detects emotions in video conferences during e- learning. The system uses many machine learning models but got good results SVM and KNN". Kim et al[9] thought of system that provides real time suggestion to improve their lectures by changing their body movement and face features. Another system[10] recognizes emotions in virtual learning environments using these models recognises different face features and classify accordingly. Chiou et al[11]uses different ML methods and network system to make high efficiency and avoid wasting of time and got great result.

IV. PROPOSED APPROACH:

Our proposed approach is to identify and analyze different aspects of face with our models and then make necessary classifications. First, it finds and identifies faces in the image. Then, it crops and resizes the detected faces to a standard size of 48x48pixels. These resized facial images are then fed into the CNN model for analysis. The CNN model then outputs the recognized facial expression, which could be Angry, Happy, Sad, Surprise, or neutral. The structural process of the system is shown in below diagram. We will first pre-process our datasets, crop and normalize as required after pre-processing step we would classify into different class with the help of CNN model and our prediction of input images is obtained.

Our approach uses a type of Mechanism called a convolutional neural network (CNN) to recognize facial expressions. CNNs are able to recognize patterns in images with very little preparation, unlike other methods. They learn to identify patterns using filters, which are like templates that are matched to the image. The CNN model uses a backpropagation algorithm to improve its ability to recognize facial expressions. This algorithm works checking difference between what has been predicted and what actual input is facial expressions and then adjusting the weights of the connections between the neurons.

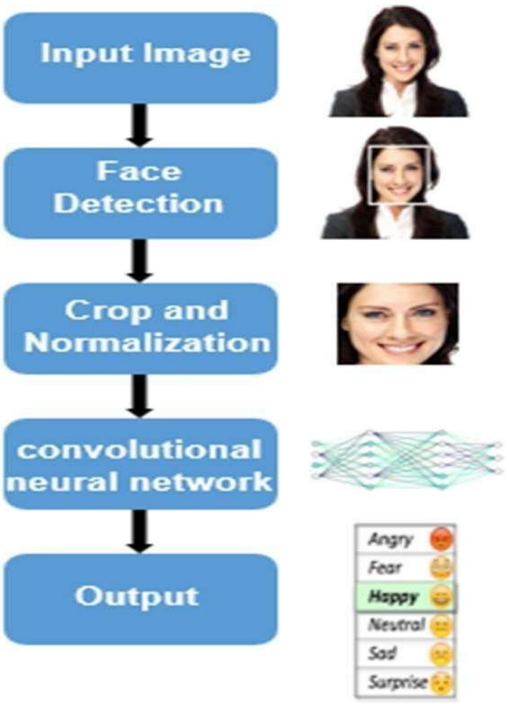

Fig.2 Methods for our system.[22]

CONVOLUTIONAL NEURAL NETWORK:

The convolutional layer is the most important part of a convolutional neural network (CNN) because it's the layer that extracts features from an image. It does this by using small squares of pixels called kernels to identify patterns in the image[21]. This operation helps the CNN to learn the important features of the image.

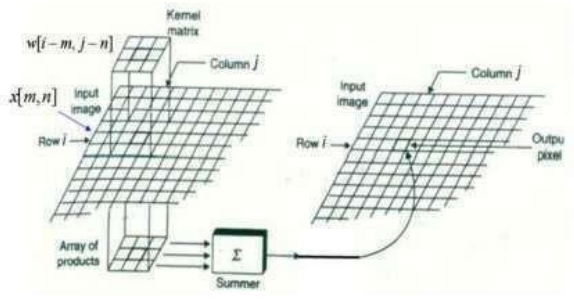

Fig 3: Layers in CNN[22]

POOLING LAYER:

A convolutional neural network (CNN) may analyse data of a smaller volume by using pooling layers. This is significant because it can enhance the CNN's capacity to generalise to new data and help prevent it from overfitting to the training set[21]. Max pooling, average pooling, and total pooling are the three primary categories of pooling layers. The result obtained pooling layer in max pooling is the maximum value of a tiny block of pixels. The result obtained in pooling layer in average pooling is the average value of a tiny block of pixels. The result obtained pooling layer in total pooling is the total of all the values within a tiny block of pixels.

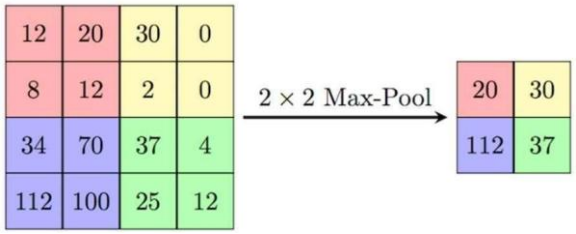

Fig 4: Example of pooling[22]

The CNN haver several layers that perform quality image processing. With the use of different layers and some activation function image features like edges, shapes and others are detected. Further unwanted dimensions and details are removed and necessary details are retained. Following these different process in different layers final result is produced in final layer and giving network prediction. These structures in CNNs slowly understand complex visual pattern and features so that predictions like image recognition and classification are done accurately[21].

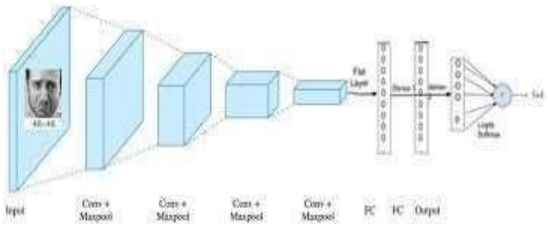

Fig 5: CNN layers for features.[22]

In the last part of CNN, it combines all what it is been trained and learned and make prediction. This step converts complex details in form of desired result, making it smart and more accurate.

V. RESULT AND DISCUSSION:

We take or extract pictures of people using web camera after this processing on captured image is done, colored photograph is converted into grayscale. Using LBPH face is detected and trained datasets and CNN is used with database for the classification into different classes.

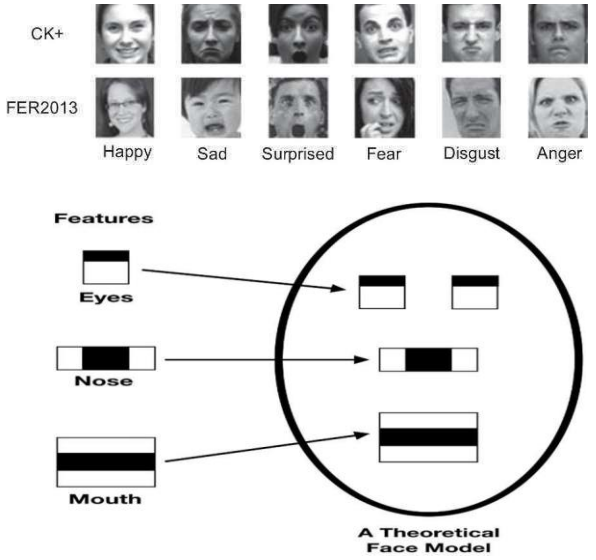

Fig 6: Face Detection LBPH[22]

80-20 is percentage in dataset used as training and testing. Datasets from FER2013 which have pre- cropped grayscale classified faced photos. Webcam photo as input processed for output. The output is in categories of expression like Happy, sad, surprised, angry and neutral.

To evaluate the effectiveness of our approach, we shown in Figures 7 and 8. Our model is particularly accurate at predicting expressions of happiness and surprise. However, it decent with angry faces, often good for sad faces.

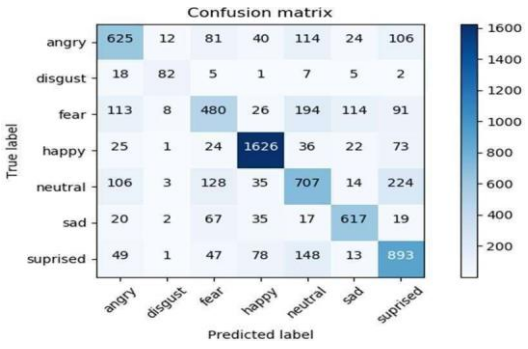

Fig 7. The confusion matrix on the FER 2013[22].

CONCLUSION AND FUTURE WORK:

Our study gives effective method for face emotion recognition with CNN and Deep learning algorithms. Our system combines many facial features, reducing data as per necessity and make noises clear using specific filters. These steps are responsible for improve performance of CNN.

Further, we can increase classes of expression and also consider the ethnicity of people belonging to different geographic location and culture. This system can further modified and can be used for application like robotics, security, easy communication with machines,and also can able to give remark on level of expression and extent of human feeling.

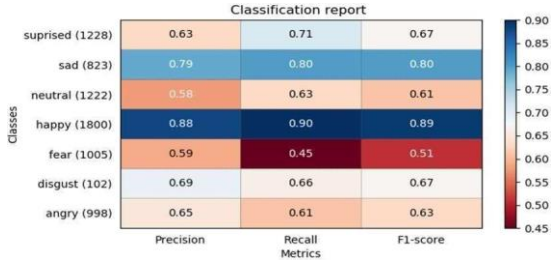

Fig 8. The classification report for method we introduced on the FER 2013 database[22].

REFERENCES:

[1] R. G. Harper, A. N. Wiens, and J. D.Matarazzo, Nonverbal communication: the state of the art. New York: Wiley, 1978.

[2] P. Ekman and W. V. Friesen, “Constants acrosscultures in the face and emotion,” Journal of Personality and Social Psychology, vol. 17, no 2, p. 124-129, 1971.

[3] C. Tang, P. Xu, Z. Luo, G. Zhao, and T. Zou, “Automatic Facial Expression Analysis of Students in Teaching Environments,” in Biometric Recognition, vol. 9428, J. Yang, J. Yang, Z. Sun, S. Shan, W. Zheng, et J. Feng, Éd. Cham: Springer International Publishing, 2015, p. 439-447.

[4] A. Savva, V. Stylianou, K. Kyriacou, and F. Domenach, “Recognizing student facial expressions: A web application,” in 2018 IEEE Global Engineering Education Conference (EDUCON),Tenerife, 2018, p. 1459-1462

[5] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J.R. Movellan, “The Faces of Engagement: Automatic Recognition of Student Engagement from Facial Expressions,” IEEE Transactions on Affective Computing, vol. 5, no 1, p. 86-98, janv. 2014.

[6] N. Bosch, S. D'Mello, R. Baker, J.Ocumpaugh, V. Shute, M. Ventura, L. Wang and W. Zhao, “Automatic Detection of Learning-Centered Affective States in the Wild,” in Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI ’15, Atlanta, Georgia, USA, 2015, p. 379-388

[7] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, “A survey of affect recognition methods: Audio, visual, and spontaneous expressions,” IEEE transactions on pattern

analysis and machine intelligence, vol. 31, no. 1, pp. 39–58, 2009.

[8] U. Ayvaz, H. Gürüler, and M. O. Devrim, “USE OF FACIAL EMOTION RECOGNITION IN E-LEARNING SYSTEMS,” Information Technologies and Learning Tools, vol. 60, no 4, p. 95, sept.2017.

[9] Y. Kim, T. Soyata, and R. F. Behnagh, “Towards Emotionally Aware AI Smart Classroom: Current Issues and Directions for Engineering and Education,” IEEE Access, vol. 6, p. 5308-5331, 2018.

[10] D. Yang, A. Alsadoon, P. W. C. Prasad, A. K. Singh, and A. Elchouemi, “An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment”.

[11] C.-K. Chiou and J. C. R. Tseng, “An intelligent classroom management system based on wireless sensor networks,” in 2015 8th International Conference on Ubi-Media Computing (UMEDIA), Colombo, Sri Lanka, 2015, p. 44-48.

[12] I. J. Goodfellow et al., “Challenges in Representation Learning: A report on three machine learning contests,” arXiv:1307.0414 [cs, stat], juill. 2013.

[13] A. Fathallah, L. Abdi, and A. Douik, “Facial Expression Recognition via Deep Learning,” in 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Hammamet, 2017, p. 745-750.

[14] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 2001, vol. 1, p. I-511-I-518.

[15] Y. Freund and R. E. Schapire, “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting,” Journal of Computer and System Sciences, vol. 55, no 1, p. 119-139, août 1997.

[16] H. Jung, S. Lee, J. Yim, S. Park, and J. Kim, “Joint fine-tuning in deep neural networks for facial expression recognition,” in Computer Vision (ICCV), 2015 IEEE International Conference on. IEEE, 2015, pp. 2983–2991.

[17] X. Zhao, X. Liang, L. Liu, T. Li, Y. Han, N. Vasconcelos, and S. Yan, “Peak-piloted deep network for facial expression recognition,” in European conference on computer vision. Springer, 2016, pp. 425–442.

[18] P. Ekman, “Facial action coding system (facs),” A human face, 2002.

[19] S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in 2017 International Conference on Engineering and Technology (ICET), Antalya, 2017, p. 1-6.

[20] B. Fasel and J. Luettin, “Automatic facial expression analysis: a survey,” Pattern recognition, vol. 36, no. 1, pp. 259–275, 2003

[21] Z. Zhang, P. Luo, C. L. Chen, and X. Tang, “From facial expression recognition to interpersonal relation prediction,” International Journal of Computer Vision, vol. 126, no. 5, pp. 1–20, 2018.

[22] Imane Lasri, Anouar Riad Solh, Mourad El Belkacemi “Facial Emotion Recognization of students using convolutional neural network” research gate 2019.

[23] Emad Barsoum, Cha Zhang, Cristian Canton Ferrer, Zhengyou Zhang. "Training deep networks for facial expression recognition with crowd-sourced label distribution" , Proceedings of the 18th ACM International Conference on Multimodal Interaction - ICMI 2016, 2016