Emotion Detector

Dr Hariom Upadhyay[1], Mohd Azad[2], Shivam Shreevastav[3]
[1]*Professor, R.D. Engineering College*
[23]*Department of Computer Science & Engineering*
R.D. Engineering College, Uttar Pradesh, India

***Abstract*– Emotional expression is a part of face recognition, it has always been an easy task for humans, but achieving the same with a computer algorithm is challenging. With the recent and continuous advancements in computer vision and machine learning, it is possible to detect emotions in images, videos, etc. A face expression recognition method based on the Deep Neural Networks especially the convolutional neural network (CNN) and an image edge detection is proposed. The edge of each layer of the image is retrieved in the convolution process after the facial expression image is normalized. To maintain the texture picture's edge structure information, the retrieved edge information is placed on each feature image. In this research, several datasets are investigated and explored for training expression recognition models. The purpose of this paper is to make a study on emotion detection and recognition via Machine learning algorithms and deep learning. This research work will present deeper insights into emotion detection and Recognition. It will also highlight the variables that have an impact on its efficacy.**

***Keywords* - Convolutional neural network, machine learning, deep learning, computer vision, emotion recognition.**

I. INTRODUCTION

Human-computer interaction technology refers to a kind of technology that takes computer equipment as the medium, so as to realize the interaction between humans and computers. Emotion recognition system (ERS) is a mechanism that allows cameras to automatically identify people. Because of the importance of correct and effective ERS, it drives the activeness of biometric research in the race to the digital world. In recent years, with the rapid development of pattern recognition and artificial intelligence, more and more research has been conducted in the field of human- computer interaction technology.Emotion Recognition (ER) is a flourishing study topic in which many breakthroughs are being made in industries, such as automatic translation systems and machine-to-human contact. In contrast, the paper focus to survey and reviewing various facial extraction features, emotional databases, classifier algorithms, and so on. The classical ER consists of two main steps: feature extraction and emotion recognition. In addition, image pre-processing, including face detection, cropping, and resizing. Face detection crops the facial region after removing the backdrop and non-face areas. Finally, the retrieved characteristics are used to classify emotions, which is commonly done with the help of neural networks (NN) and other machine learning approaches. The challenge of emotion recognition is to automatically recognize emotion states with high accuracy. Therefore, it is challenging to find the similarity of the same emotional state between different people since they may express the same emotional state in various ways. As an example, the expression may vary in different situations such as the individual's mood, skin colour, age, and the environment surrounding. Generally, ER is separated into three major stages as shown in Figure 1: (i) Face Detection, (ii)Feature Extraction, and (iii) Emotion Classification.

Fig 1: Classification Stages

In the first, stage, which is a pre-processing stage, an image of a face is detected and facial components of the face will be detected from the region.

In the second stage, an informative feature will be extracted from different parts of the face. In the last

stage, a classifier needs to be trained before been use a classifier needs to be trained before being used to generate labels for the Emotions using the training data. Facial actions are classified into different Action Units (AUs) and emotions are categorized using collections of AU's. Deep learning is a part of machine learning approaches that can be adapted to emotion recognition and facial expression analysis. However, deep learning depends on data size which may affect its performance.

II. LITERATURE REVIEW

Facial expression is the common signal for all humans to convey a mood. There are many attempts to make an automatic facial expression analysis tool as it has applications in many fields such as robotics, medicine, driving assist systems, and lie detector. Since the twentieth century, Ekmanetal. defined seven basic emotions, irrespective of culture in which a human grows with the seven expressions (anger, fear, happiness, sad, contempt, disgust, and surprise). Discusses an extensive study on emotion identification, including the dataset's features and the facial emotion recognition study classifier. Visual features of images are examined and some of the classifier techniques are discussed in which is helpful in the further inspection of the methods of emotion recognition. This paper examined the prediction of future reactions from images based on the recognition of emotions, using different classes of classifiers. Some of the classification algorithms like Support vector machines, and Neural Networks such as Convolution Neural networks. There are many issues like excessive makeup pose and expression which are solved using convolutional networks. The development of computer vision and machine learning has made emotion recognition much more accurate and accessible to the general public. As a result, facial expression detection as a sub-field of image processing is quickly expanding. Some of the possible applications are human-computer interaction, psychiatric observations, drunk driver recognition, and the most important is a lie detector.

III. BACKGROUND INFORMATION

A. Emotion Recognition

Facial Recognition is a branch of computer science that deals with methods and strategies for detecting emotions in facial expressions. It is expected that expressions can be the next communication medium with computers the majority of this field's research focuses on recognizing human emotions from movies or auditory data. The majority of the research has focused on recognizing and matching faces, but no convolutional neural networks have been utilized to infuse emotions into photos. Emotion Recognition is the study of identifying emotions, as well as the strategies and procedures utilized to do so. Emotions can be detected through facial expressions, verbal signals, and other indicators. Machine learning, neural networks, artificial intelligence, and emotional intelligence are just a few of the methods that have been used to infer emotions. Emotion Recognition is gaining traction in the study, which is critical to solving a variety of challenges.

B. Facial Emotion Recognition

Facial Emotion Recognition is a research area that tries to identify the emotion from the human facial expression. The surveys state that developments in emotion recognition make complex systems simpler. Emotion Recognition is a difficult process since emotions can differ depending on the environment, appearance, culture, and facial reaction, resulting in unclear data.

C. Deep Learning

Deep Learning is a machine learning technique that models the data that are designed to do a particular task. Deep learning in neural networks has wide applications in the area of image recognition, classification, decision making, pattern recognition, etc.

IV. PROPOSED METHODOLOGY

The proposed technique, the emotion database used for the study, and the Inception model are all explained in this section. This paper uses a Haar classifier for human detection. The Haar classifier is trained by Haar-like small features and also the Haar-like may be a commonly used texture descriptor, and its main features are linear, edge, canter, and diagonal. The Haar-like feature can reflect the grey level change of image, so it's very effective to explain the face because many features of external body parts have obvious contrast change characteristics. However, the calculation of eigenvalues is extremely time-

consuming. so as to enhance the calculation speed, this paper uses the integral graph method to calculate the Haar-like values.

1. Face Detection

Face detection could be a pre-processing phase to acknowledge the facial expressions of humans. a picture is segmented into two parts which have faces and other non- face regions. There are numerous methods used for face detection.

A. Haar Classifier

Haar features are often measured by expanding or reducing the dimensions of the pixel group. It used Haar-like features to detect a picture. Objects of varying sizes will be discovered using this method. Haar classifier will identify a group of features that are most contributing to the face detection problem in the training phase itself. Therefore, it's suitable for face detection in the training phase because it may indicate high detection accuracy since the computation complexity is little.

2. Feature Extraction

Feature extraction transforms pixel data from the face region into a higher-level representation of the face or its components shape, colour, texture, and spatial configuration

.Feature extraction will reduce the dimension of the input space while keeping the important information. Feature extraction is vital in formulating a stronger emotion categorization because the extracted facial feature gives inputs to the classification module which finally categorizes different emotions.

Feature extraction will be divided into two categories which are; (i) feature base and (ii) appearance base.

A. Convolutional Neural Network (CNN) Currently, CNN is one of the foremost mainstream approaches to deep learning techniques. It uses a variation of multilayer perceptron designed to want minimal pre-processing. It gets its name from the type of hidden layers it has. Convolutional layers, pooling layers, fully connected layers, and normalising layers are common components of a CNN's hidden layers. [2]

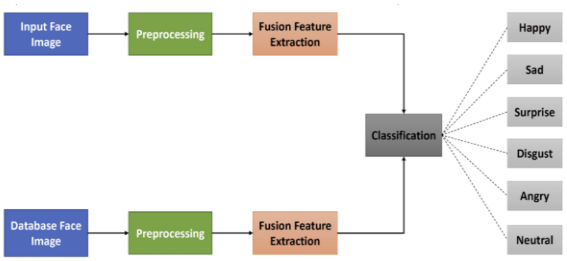


Fig 2: Emotion Detection Process

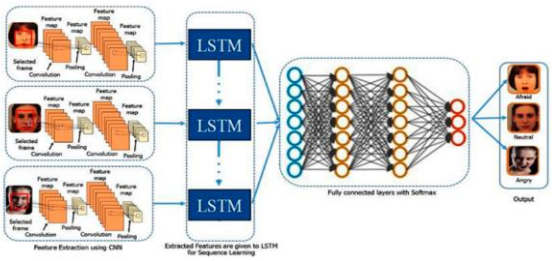


Fig 3: Image Classification

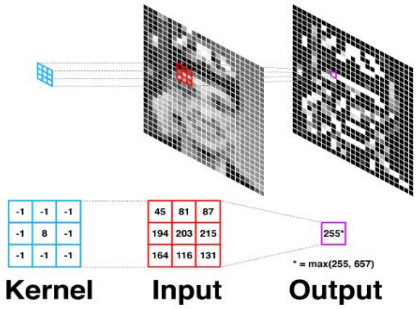


Fig 4: Convolution Filter Operation

3. Expression Classification

This stage is performed by a classifier. There are various classifications methods accustomed extract expressions.

Supervised Learning-

Supervised learning is a way of training a system using labelled data. The tagged data serves as a manager. The model is given both inputs and outputs to learn from. Following that, the model would forecast for a new data point. Classification and regression are the two types of supervised learning.

mm

A. Support Vector Machine (SVM): SVM is one of the famous statistical techniques employed in machine learning to analyse data used for classification and multivariate analysis. SVM used different kernel functions to map data in input space into high-dimensional feature space.

B. Neural Network (NN): NN executes a nonlinear reduction of the input dimensionality. It makes a statistical determination regarding the category of the observed expression. Every output unit will estimate the probability of the examined expression belonging to the associated category .

1. Inception-V1 toV3

The Inception network was a significant step forward in the evolution of CNN classifiers. It is a 22-layer design with a total of 5M parameters. It employed numerous techniques to improve performance, both in terms of speed and precision. This is frequently utilised in machine learning applications [20, 21]. Inception V2, It is the 24M parameter successor to Inception-V1. Inception-v3 is a popular image recognition model that has been shown to achieve more than 78.1 percent accuracy on the Image Net dataset. However, it is not widely utilised.

V. DATASET

To perform an experiment on FER, a regular database is required. The information will be perceived as primary or secondary. A primary dataset consumes an extended period to be completed with dataset collection. For study in FER, a spread of datasets available currently There are few datasets available for the emotion recognition problem; among those, Karolinska Directed Emotional Faces (KDEF)and Japanese Female facial features (JAFFE) datasets are well-known and regarded during this study. The dataset's images are divided into seven main emotion categories [3]. The KDEF dataset (also refer as KDEF for simplicity, henceforth) was developed by Karolinska Institute, Sweden. Specifically, the aim of the dataset was to use for perception memory emotional attention, and backward masking experiment. The dataset contains 4900 photos of 70 people, each of whom is depicted in seven different emotional states.

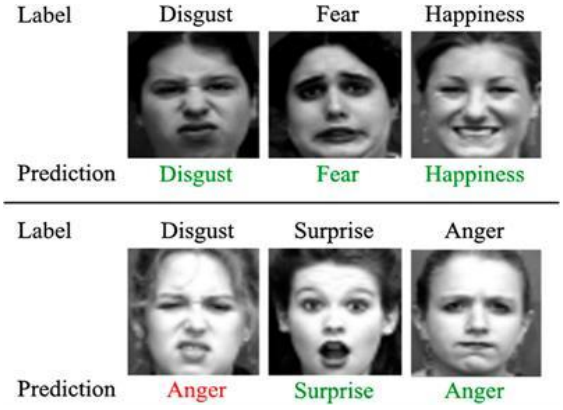
VI. RESULT AND DISCUSSION

To analyse the performance of the algorithm, the FER-2013 expression dataset was used initially. Dataset had only 7178 with 412 posers, causing accuracy to reach up to 55% maximum. To overcome the problem of low efficiency, multiple datasets were downloaded from the Internet, and also author's own pictures of different expressions were included. As the number of images in the dataset increases, the accuracy also increased. We kept 70% of 11K dataset images as training and 30% of dataset images as testing images. The number of layers and the number of filters, for background removal CNN (first-part CNN) as well as face feature extraction CNN (the second-part CNN), are the same. The number of layers in this experiment ranged from one to eight. We discovered that the greatest accuracy was around 4. We assumed that the number of layers is directly proportional to accuracy and inversely proportional to execution time, which was counterintuitive. Hence due to the maximum accuracy obtained with 4 layers, we selected the number of layers to be 4. The execution time increased as the number of layers increased, however, it did not bring much value to our research. Based on the achieved test set accuracies, the new method outperformed the existing ones.

Method	Accuracy
LBP-TOP [11]	88.99%
HOG 3D [9]	91.44%
IACNN [12]	95.37%
DTAGN [10]	97.25%
CNN (baseline)	89.50%
Ours Proposed Method	97.98%

Table 1: Accuracy Table

It's worth noting that the proposed method only misclassified a few photographs with perplexing perspectives, and overall identification accuracy remains impressive. As a result, the method suggested in this work holds promise in a real-world environment where non- frontal or angularly captured photos are the norm.



The output images of different images are show from

figure number 5 to 8. When numerous faces were present in the same image and were at the same distance from the camera, the algorithm failed. It was discovered that when the number of photons increases, accuracy decreases due to over-fitting. Also, when the number of training photos is reduced, accuracy remains low. The ideal number of images was found to be in the range of 2000–11,000 for FER to work properly.

VII. CONCLUSION

We propose a face expression identification approach based on a CNN model that effectively extracts facial features in this research. The suggested method uses training sample image data to directly input the picture pixel value. The ability to accurately determine emotions was greatly enhanced by the removal of the background. Emotion expression is important in communication, hence improving the quality of interaction between humans. Furthermore, in the near future, the study of facial expression detection may provide improved feedback to society as well as the interaction between Human-Robot interfaces (HRI). Emotion detection mostly involves the geometric part of the face (e.g.; eyes, eyebrows, and mouth). The review takes into consideration of experiments which been conducted in a controlled environment, in real-time, and in wild images. The recent research, particularly the performance with profile views, will be applicable to a greater range of real- world commercial applications, such as patient monitoring in a hospital or surveillance security. Furthermore, the concept of facial emotion recognition could be expanded to include emotion recognition from speech or body motions in order to address emerging industrial applications.

REFERENCES

[1] K. F. Azizan Illiana, "Facial Emotion Recognition: A Brief Review," in International Conference on Sustainable Engineering, Technology and Management 2018 (ICSETM-2018), 2020.

[2] R. Shyam, "Convolutional Neural Network and its Architectures.," Journal of Computer Technology & Applications, vol. 12, no. 2, pp. 6-14, 2021.

[3] R. Shyam, "Machine Learning and Its Dominant Paradigms," Journal of Advancements in Robotics, vol. 8, no. 2, pp. 1-10, 2021.

[4] R. Shyam, "Automatic Face Recognition in Digital World," Advances in Computer Science and Information Technology (ACSIT), vol. 2, no. 1, pp. 64-70, 2015.

[5] S. R. N. S. M. A. H. Akhand, "Facial Emotion Recognition Using Transfer Learning in the Deep CNN," MDPI, vol. 10, no. 9, 2021.

[6] N. Mehendale, "Facial emotion recognition using convolutional neural networks (FERC)," SN Applied Sciences, vol. 2, no. 3, 2020.

[7] N. R. S, "Emotion Recognition from Facial Expression using deep learning," International Journal of Engineering and Advanced Technology (IJEAT), vol. 8, no. 6S, 2019.

[8] R. Shyam, "Enhanced Object Detection with Deep Convolutional Neural Networks," International Journal of All Research Education and Scientific Methods (IJARESM), vol. 9, no. 7, pp. 27-36, 2021.