An Approach to Human Face Expression Recognition Using LDA With Machine Learning Algorithm

[1]A. Ramakrishna Prasad, [2]I. Sasi Kiranmai, [3]P. Tharun Kumar,
[4]U. Brahma Raja, [5]SK. Abdul Raqib

[1]Assistant Professor, [2,3,4,5]Students
Electronics and Communication Engineering
N.B.K.R Institute of Science and Technology, Andhra Pradesh, India

Abstract- **Due to its tremendous academic and business potential, facial emotion recognition (FER) is a crucial topic in the domains of computer vision and artificial intelligence. Although FER can be carried out with a variety of sensors, this review concentrates on research that only employs facial images because facial expressions are one of the primary information routes in interpersonal communication. This project offers a succinct overview of FER research that has been done over the years. First, a synopsis of the representative categories of FER systems and their primary algorithms is given together with a description of typical FER approaches. Then, FER neural network methods that support "end-to-end" learning are introduced. The focus of this work is also on a modern hybrid learning strategy that combines a neural network (NN) for the spatial elements of a single frame and long short-term memory (LSTM) for the temporal features of subsequent frames. A brief overview of publicly accessible evaluation metrics is provided in the later section of this study, and a comparison with benchmark results—a standard for a quantitative comparison of FER researches—is explained. This review can act as a quick reference for both seasoned researchers seeking for fruitful avenues for future study and newbies to the field of FER, giving fundamental knowledge and a general overview of the most recent state-of-the-art investigations.**

1.Introduction

Human communication relies heavily on facial expressions to grasp one another's intentions. Generally speaking, people use their speech tones and facial expressions to infer other people's emotional states, such as happiness, sadness, and rage. Various surveys [1] indicate that verbal components make up one-third of human communication while nonverbal components make up the other two-thirds. Facial expressions are one of the primary information carriers in interpersonal communication among a few nonverbal cues because they convey emotional significance. So, it makes sense that research on facial expression has become more popular over the past few decades, with applications not just in the perceptual and cognitive sciences but also in affective computing and computer animations [2]. There is interest in automatic facial emotion recognition (FER), also known as facial emotion recognition and facial expression recognition (expanded form of the acronym FER). As this study works with the broad characteristics of facial emotion recognition, the term "FER" in this paper refers to facial emotion recognition.) has also accelerated recently because to advances in artificial intelligence technologies, such as those used in human-computer interaction (HCI) [3,4], virtual reality (VR) [5,6], augment reality (AR), advanced driving assistance systems (ADASs) [7], and entertainment [8,9].This study first categorizes research on automatic FER into two groups based on whether the features are created manually or as a result of a deep neural network's output.

In order to identify facial landmarks or facial components (such as the eyes and nose), a face image must first be extracted from an input image. The facial components are then used to extract a variety of spatial and temporal data. Third, using the extracted features, the pre-trained FE classifiers, including a support vector machine (SVM), AdaBoost, and random forest, provide the recognition results. The local texture of a landmark or the pairwise positions of each of two landmark points are both employed as FER feature vectors. According to the development of the models, FL detection methods may generally be divided into three categories: CNN-based methods, regression-based models with a combination of local and global models, and active shape- and appearance-based models (ASM and AAM). Seven fundamental human feelings are referred to as basic emotions (BEs): joy, surprise, anger, sadness, fear, disgust, and neutral.

2.Literature Survey:

Zhao et al. employed LBP-TOP (Local Binary Patterns from Three Orthogonal Planes) feature descriptors and near-infrared (NIR) video sequences. This study combines geometric and aesthetic information about the face using component-based facial traits. A SVM and sparse representation classifiers are employed for FER. Shen et al. extracted the horizontal and vertical temperature difference from several facial sub-regions using infrared thermal videography. The Adaboost method with k-Nearest Neighbor's weak classifiers is used for FER. Szwoch and Pieniek were able to distinguish facial expression and emotion without the aid of a camera by merely employing the depth channel from a Microsoft Kinect sensor. This study recognises facial expressions using relationships between specific emotions and local movements inside the face area as the feature.

Sujono and Gunawan employed the Kinect motion sensor to monitor the identified face using the active appearance model (AAM) and to determine the facial region based on depth information. When there is variance in shape and texture compared to the training results, the job of AAM is to update the shape and texture model in a new face. Fuzzy logic based on past information acquired from FACS and the changing of key features in AAM are utilised to recognise face expression. Wei et al. suggested FER

employing both the colour and depth data provided by the Kinect sensor. Using sensor data that has been gathered, this study uses a face tracking algorithm to extract facial feature points and a random forest technique to recognise six different facial moods.

3.Methodology

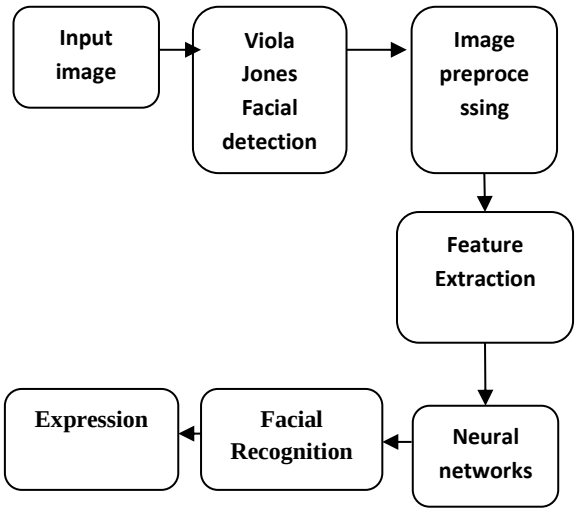

Figure: Proposed FER System

Image Acquisition

The picture acquisition stage is the first step in any image processing activity. After the image has been captured, it can be processed using a variety of techniques to carry out the various tasks needed. Although using a digital camera for digital photography is the typical method of acquiring images, alternative techniques are sometimes used. In this instance, a web camera was deployed. However, we used illustrative traffic photographs from the internet for the simulation.

RGB to Grayscale Conversion

The photos are resized after being acquired. Because every camera has a different resolution, a system that is made for one set of camera specifications won't work properly for another set of specifications, and vice versa. Therefore, image resizing is required in order to make the resolution constant for the programme. after resizing the photos. To make image processing easier, they are transformed to grey scale. A grayscale image is one that has solely intensity information, with each pixel's value being a single sample in image processing.

Image Enhancement

Now that our image is more contrasted with the backdrop, a suitable threshold level may be chosen for binary conversion. Image enhancing techniques are required here. The goal of image enhancement is to modify an image so that the final product is better suited for the intended application than the original image. The features of an image can be played with using a variety of techniques, though not always. Here are a few basic operations that are widely used to improve images.

- Power law transformations (gamma correction)
- Piecewise linear transformation functions
- Logarithmic (log and inverse log transformations)
- Linear (negative and identity transformations)
- Histogram equivalence

Viola–Jones object detection framework

Paul Viola and Michael Jones presented the Viola-Jones object detection framework as an object identification framework in 2001.[1][2] It was primarily driven by the issue of face detection, despite the fact that it may be trained to recognise a range of object classes.

Face detection in an image is the issue that needs to be solved. A person can accomplish this effortlessly, but a computer need specific guidelines and limitations. Full view frontal upright faces are necessary for Viola-Jones in order to make the task more feasible. Therefore, the full face must face the camera and cannot be slanted in any direction in order to be recognised.

Feedforward neural network

A feedforward neural network is a type of artificial neural network where there are no cycles in the connections between the nodes. It differs from its ancestor, recurrent neural networks, as a result [1].The first and most basic artificial neural network design was the feedforward neural network.[2] The information in this network only travels in one direction, forward, from the input nodes to the output nodes.

Linear Discriminant Analysis (LDA)

Data dimensions are reduced using the mathematical formula Principal Components Analysis (PCA). As a result, the PCA technique enables the recognition of standards in data and their expression in a way that highlights both their similarities and contrasts. Once patterns are identified, it is possible to compress them, or lower their dimensions while preserving a large amount of information. In conclusion, the PCA formulation can be applied as a low-loss technique for digital image compression.

LDA is a supervised classification technique that is considered a part of crafting competitive machine learning models. This category of dimensionality reduction is used in areas like image recognition and predictive analysis in marketing.

4.Software Requirement:

MATLAB:

Engineers and scientists can use the programming environment MATLAB® to analyze, create, and test systems and technologies that will change the world. The MATLAB language, a matrix-based language that enables the most natural expression of computer mathematics, is the core of MATLAB.

5.Results:

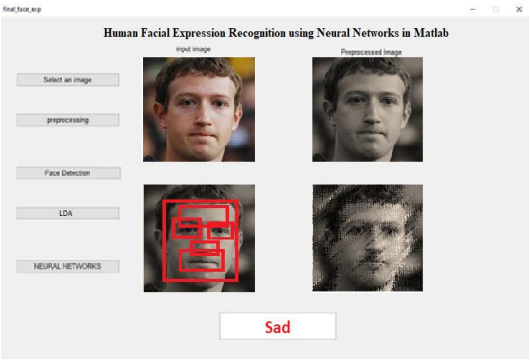

Fig: FER Using Neural Networks

The above figure shows that human emotions like happiness, sadness, joy, guilty, mightiness, cry, laughing and so on. This paper gives the results of human facial expressions observing the different movements of eyes, mouth, nose. Previously there are multiples techniques are available to find out the FER. Moreover, this paper will generate the accurate results compared to previous one.

6.Conclusion

Happiness, melancholy, joy, guilt, mightiness, crying, laughing, boring, fear, aggression, and jealousy are a few significant human emotions. Therefore, in order to create these emotions in robots, A.I. engineers must pay close attention to two areas: "facial expression" and "sound/phonic/speech recognition" to recognise human emotion and the responses from Humanoid robots to it. With the aid of "facial expression" and "speech/sound" recognition combined A.I engineering as emotion input to process and identify by Humanoid, this research article provides an idea about the core notion of how A.I engineers could build and execute human-like emotions in Humanoids. Recognizing human gestures for use in everyday life is particularly fascinating in the field of image processing.

References:

1. S. W. Lee, V. Mail, J. Jang, J. Shin, and J. Paik, “Noise-adaptive spatiotemporal filter for real-time noise removal in low light level images,” IEEE Trans. Consumer Electron., vol. 51, no. 2, pp. 648-653, May. 2005.
2. E. Bennett and L. McMillan, “Video enhancement using per-pixel virtual exposures,” ACM Trans. Graphics, vol. 24, no. 3, pp. 845-852, Jul. 2005.
3. H. Malm, M. Oskarsson, E. Warrant, P. Clarberg, J. Hasselgren, and C. Lejdfors, “Adaptive enhancement and noise reduction in very low lightlevel video,” In Proc. IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, pp. 1-8, Oct. 2007.
4. Q. Xu, H. Jiang, R. Scopigno, and M. Sbert, “A new approach for very dark video denoising and enhancement,” In Proc. IEEE International Conference on Image Processing, Hong Kong, China, pp. 1185-1188, Sept. 2010.
5. X. Dong, G. Wang, Y. Pang, W. Li, J. Wen, W. Meng, and Y. Lu, “Fast efficient algorithm for enhancement of low lighting video,” In Proc. IEEE International Conference on Multimedia and Expo, Barcelona, Spain, pp. 1-6, Jul. 2011.
6. X. Zhang, P. Shen, L. Luo, L. Zhang, and J. Song, “Enhancement and noise reduction of very low light level images,” In Proc. International Conference on Pattern Recognition, Tsukuba, Japan, pp. 2034-2037, Nov. 2012.
7. M. Kim, D. Park, D. K. Han, and H. Ko, “A novel framework for extremely low-light video enhancement,” in Proc. IEEE International Conference on Consumer Electronics, Las Vegas, USA, pp. 95-96, Jan. 2014.
8. A. Loza, D. Bull, and A. Achim, “Automatic contrast enhancement of low-light images based on local statistics of wavelet coefficients,” In Proc. IEEE International Conference on Image Processing, Hong Kong, China, pp. 3553-3556, Sept. 2010.
9. F. Drago, K. Myszkowski, T. Annen, and N. Chiba, “Adaptive logarithmic mapping for displaying high contrast scenes,” Computer Graphics Forum, vol. 22, no. 3, pp. 419-426, Sept. 2003.