International Journal of Research in Computer & Information Technology (IJRCIT),
Vol.6, Issue 3, June 2021
E-ISSN: 2455-3743
Available online at www.ircit.co.in

Study on Emotion Detection Using Machine Learning

[1]Ms. Ranjana V. Gawai, [2]Dr. V. R. Raut, [3]Dr. S. M. Deshmukh

Abstract- Researchers in psychology, computer science, linguistics, neurology, and allied disciplines are becoming interested in a human-computer interface system for autonomous face recognition or facial expression recognition. An Automatic Facial Expression Recognition System (AFERS) is proposed in this paper. Face detection, feature extraction, and facial expression identification are the three stages of the proposed method. The initial phase of face detection entails detecting skin colour using the YCbCr colour model, lighting correction to achieve homogeneity on the face, and morphological procedures to keep the required face area. Using the AAM (Active Appearance Model) approach, the output of the first phase is used to extract face features such as eyes, nose, and mouth. Automatic facial expression recognition, the third stage, is straightforward. The Euclidean Distance technique is used to calculate the distance between two points. The Euclidean distance between the feature points of the training and query images is compared in this method. The output picture expression is determined using the minimum Euclidean distance. This approach has a true recognition rate of 90 percent to 95 percent. The Artificial Neuro-Fuzzy Inference System is used to further improve this method (ANFIS). When compared to previous approaches, this non-linear recognition system achieves a recognition rate of roughly 100%.

Keywords- Facial expression recognition (FER), multimodal sensor data, emotional expression recognition, spontaneous expression, real-world conditions

I. INTRODUCTION

The advances in allied domains such as machine learning, image processing, and human cognition, facial expression recognition (FER) has advanced substantially in recent years. As a result, automatic FER's effect and prospective use in a wide range of applications, such as human-computer interaction, robot control, and driver status surveillance, has been rising. However, due to the difficulties in reliably extracting the useful emotional elements, strong detection of facial emotions from photos and videos has remained a difficult task to date (1). These characteristics are frequently portrayed in various ways, including static, dynamic, point-based geometric, and region-based appearances. The movements of facial parts and muscles during emotional expression are responsible for facial movement aspects such as feature position and shape changes. When participants are exhibiting emotions, the facial components, particularly crucial elements, will constantly shift positions. As a result, the same feature appears in distinct photos at different times. Due to small facial muscle movements, the contour of the feature may be affected in rare circumstances. The mouths in the first two photographs, for example, have distinct forms than those in the third image. As a result, the geometric-based location and appearance-based form of any component indicating a specific emotion in image databases and videos typically shift from one image to the next. This type of movement features represents a large pool of both static and dynamic expression qualities that are important for FER.

The great bulk of previous FER research has ignored the dynamics of facial expressions. Attempts have been made to capture and use face movement features, almost all of which have been video-based. These efforts aim to incorporate either geometric features of the tracked facial points (such as shape vectors, facial animation parameters, distance and angular, and trajectories), or appearance differences between holistic facial regions in subsequent frames, or texture and motion changes in local facial regions. Although these methods have yielded encouraging results, they frequently necessitate precise placement and tracking of face locations, which remains a challenge.

In conventional FER approaches, the FER is composed of three major steps, as shown in Figure1:

(1) face and facial component detection, (2) feature extraction, and (3) expression classification. First, a face image is extracted from an input image, followed by the detection of facial components (e.g., eyes and nose) or landmarks in the face region. Second, from the facial components, various spatial and temporal features are derived. Third, utilising the retrieved features, pre-trained FE classifiers such as a support vector machine (SVM), AdaBoost, and random forest give recognition results.

Manuscript received May 12, 2021; Revised 28 May, 2021 and published on June 30, 2021
Ms. Ranjana V. Gawai PG Scholar, Department of Electronics & TeleCommunication, Prof. Ram Meghe Institute of Technology & Research Badnera, Amravati, India
Dr. V. R. Raut Professor, Department of Electronics & Telecommunication, Prof. Ram Meghe Institute of Technology & Research Badnera, Amravati, India
.Dr. S. M. Deshmukh Professor, Department of Electronics & Telecommunication, Prof. Ram Meghe Institute of Technology & Research Badnera, Amravati, India

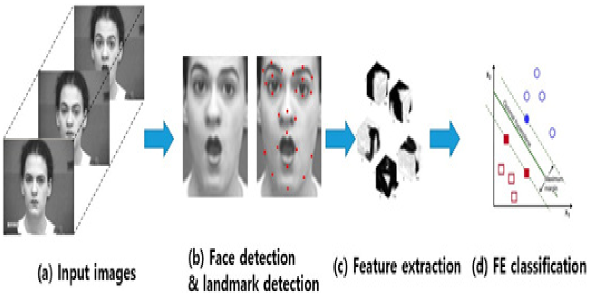

Figure 1. Procedure used in conventional FER approaches: From input images (a), face region and facial landmarks are detected (b), spatial and temporal features are extracted from the face components and landmarks (c), and the facial expression is determined based on one of facial categories using pre-trained pattern classifiers (face images are taken from CK+ dataset (d)

II. LITERATURE REVIEW

One of the factors of the strength of interpersonal relationships is the ability to read emotional facial expressions effectively. The greater one's ability to correctly interpret another's emotions, the more one is included in such encounters. Some psychopathological diseases display difficulty in social relationships, which may be linked to impairments in recognising facial expressions. Such deficits have been demonstrated in various clinical populations. Nonetheless, with respect to facial expressions, there have been discrepant findings of the studies so far. The goal of this paper is to explore the topic of emotion(3) and emotional facial expressions from ancient times, to highlight the strengths and flaws of related studies, to compare their findings, and to draw attention to this innovative issue.

William James proposed the first physiological theory of emotion in 1884. Emotion, according to James, is anchored in bodily experience. According to him, we first notice the item, then have a physiological response, and finally, emotional arousal. When we perceive a stimulus like a bear, for example, we get a racing heart, start running, and then fear. We don't run because we're afraid; we run because we're afraid. Since 1885, when his Danish colleague Carl Lange proposed a similar view, the theory has been known as the James-Lange theory of emotions.

B. Cannon provided an alternate explanation, claiming that emotions are a cognitive state of alertness rather than a physiological one. External stimulus, cerebral processing, and physiological reactions were his perceptions of the order of events. Emotion-inducing stimuli, according to this unique hypothesis, provoke both an emotional experience, such as dread, and physiological responses, such as sweating.

The study investigated the recognition of standardized facial expressions of emotion (anger, fear, disgust, happiness, sadness, surprise) at a perceptual level (experiment 1) and at a semantic level (experiments 2 and 3) in children with autism (N = 20) and normally developing children (N = 20). Results revealed that children with autism were as able as controls to recognize all six emotions with different intensity levels, and that they made the same type of errors(4).

These negative findings are reviewed in regard to (1) earlier data indicating particular impairment in autism in detecting the belief-based expression of surprise, (2) previous research indicating that individuals with autism, like patients with amygdala damage, pass a basic emotion recognition test but fail to recognise more complex stimuli involving the perception of faces or parts of faces, and (3) the convergence of findings indicating that individuals with autism, like patients with amygdala damage, pass a basic emotion recognition test but fail to recognise more complex stimuli involving the perception of faces or parts of faces.

Psychologists have been assessing the social and affective abnormalities in autism since Kanner's original clinical account of children with autism first noted their remarkable absence of affective contact with other people. Because the empirical study on affective impairment in children and people with autism is so broad and varied, it's no surprise that the results are so contradictory. Howard et al., 2019) investigated the hypotheses of a general affective deficiency and a specific emotion perception impairment. Furthermore, by contrasting recognising tests that do and do not require the ability to describe mental states, the theory of mind (ToM) deficit hypothesis of autism enables research of selective emotion processing impairment. The current research aims to replicate and expand on these findings in autistic youngsters.

The multiclass object classification problem is tackled using a biologically inspired model of visual object recognition. Serre, Wolf, and Poggio's models are modified by ours. We use Gabor filters at all positions and scales, just like in that paper, and then alternate template matching and max pooling procedures to build up feature complexity and position/scale invariance. We refine the approach in several biologically plausible ways, using simple versions of scarification and lateral inhibitionWe show how important it is to save some position and scale data above the intermediate feature level. We arrive to a model that performs better with less characteristics using feature selection. Our final model is put to the test on the Caltech 101 item categories and the UIUC automobile localisation problem, and it excels in both. The results strengthen the case for using this class of model in computer vision.

Recognition of several item classes in natural photos has proven to be a difficult task for computer vision. Given human vision's significantly greater performance on this task, it's reasonable to look to biology for guidance. In fact, Serre, Wolf, and Poggio recently demonstrated that a computational model based on our understanding of visual cortex can compete with the best extant computer vision systems on some of the most common recognition datasets. Our research expands on their approach by introducing biologically motivated qualities such as feature sparsification, lateral inhibition, and feature localisation. We show that these tweaks increase recognition even more, adding to our knowledge of the computational limits that both biological and computer vision systems face.

III. Analysis And Design Of The Application

A. Existing Work

AFERS has three main steps

1. To detect a face from a given input image or video,

2. From the detected face, extract facial features such as the eyes, nose, and mouth.

3. Sort facial expressions into categories like happiness, rage, sadness, fear, contempt, and surprise. Detecting faces is a subset of object detection. To keep the face of the supplied image, it additionally uses light compensation techniques and morphological processes.

B. Drawbacks

Because they can reveal a person's affective state, cumulative activity, personality, intention, and psychological condition, the system plays a communication role in interpersonal relationships. Three modules make up the proposed system. The face detection module uses an image segmentation technique to turn a given image into a binary image, which is then utilised to detect faces.

C. Proposed Work

The Artificial Neuro-Fuzzy Inference System is used in the third step to further increase the system's recognition rate (ANFIS). Static photos as well as video input can be used to test the expressions in this method. A neuro-fuzzy based automatic facial expression identification system has been proposed to recognise human facial expressions such as happiness, fear, sadness, anger, disgust, and surprise. Initially, a video of various expressions is framed into various images. The selected image sequence is then saved in a database folder. The features of all the photographs are found and saved in the form of. ASF files using the AAM approach. After that, a mean shape is created for all of the images in the data folder. The distance or difference (6) between Neutral and other facial emotions is measured by the change in the AAM shape model in response to changes in facial expressions. These values are saved in a.MAT file, and during ANFIS training, a specific value is assigned to each individual expression. The ANFIS is then provided these differential values as input (Artificial Neuro-Fuzzy Inference System). The system is trained for different images and their video input sequences for varied expressions using the ANFIS tool accessible in Matlab.

D. Advantages

One advantage of employing these colour spaces is that they are already used to encode most video files. The conversion from RGB to any of these spaces is a simple linear operation.

1. Face detection,

2. Feature extraction and

3. Recognition of facial expressions. Face detection starts with skin colour detection using the YCbCr colour model, lighting adjustment for uniformity on the face, and morphological procedures to keep the required face portion.

IV. System Implementation

A. Skin Color Segmentation

To segment skin colour, we first contrast the image. After that, we do skin colour segmentation.

B. Face Detection

To detect faces, we must first transform an RGB image to a binary image. To convert a binary image, we calculate the average RGB value for each pixel and replace it with a black pixel if the average value is less than 110, otherwise with a white pixel. We can convert an RGB image to a binary image using this method.

C. Eyes Detection

We transform the RGB face to the binary face for eye detection. Now we multiply W by the face width. To locate the centre position of the two eyes, we scan from W/4 to (W-W/4). The centre position of the two eyes is represented by the highest white continuous pixel along the height between the ranges.

D. Apply Bezier Curve

In the lip box, the lip, and possibly a bit of the nose. So, there is skin colour or skin around the box. As a result, we turn the skin pixel into a white pixel and the other pixel into a black pixel. We also identify and convert pixels that are comparable to skin pixels to white pixels. When the difference in RGB values between two pixels is less than or equal to 10, we term them similar pixels. The distance between the lower average RGB value and the higher average RGB value is calculated using the histogram.

E. Database and Training

There are two tables in our database. One table, "Person," stores people's names as well as their indexes for four different types of emotions, which are recorded in another table, "Position.". For each index in the "Position" table, there are 6 control points for the lip Bezier curve and 6 control points for the left eye Bezier curve. 6 control points on a Bezier curve for the right eye Lip height and width, left eye height and width, and right eye height and width are all part of the Bezier curve. As a result of this strategy, the programme learns people's emotions.

F. Emotion Detection

We must find the Bezier curves of the mouth, left eye, and right eye in order to detect emotion in an image. Then we change each Bezier curve's width to 100 and its height to its width. If the database contains information on the person's emotions, the computer will match which emotion's height is closest to the present height and output the closest emotion.

CONCLUSION & FUTURE SCOPE

This paper reviewed the efforts of many researchers, with an emphasis on including as many references from recent years as feasible. Based on reviews, the study addressed some of the concerns with facial expression identification by employing various strategies for face detection, feature extraction, analysis, and classification. The study provides thorough information on available strategies for Facial Expression Recognition (FERs) at all phases. The study is particularly valuable to both experienced and new researchers in the field of FER since it provides detailed information about existing methodologies at various phases of the field to help them strengthen their awareness of current trends and plan their future research goals. The study also discussed numerous strategies of their technology, as well as their benefits and drawbacks, in order to improve the performance of Facial Expression Recognition in image processing.

REFERENCES

[1] Eldar, C. Yonina, "Compressed sensing: theory and applications," Cambridge University, 2012.

[2] Solomon, Chris, Fundamentals of Digital Image Processing: A practical approach with examples in Matlab. John Wiley & Sons, 2011.

[3] Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition." In bmvc, vol. 1, no. 3, p. 6. 2015.

[4] Grafsgaard, Joseph, Joseph B. Wiggins. "Automatically recognizing facial expression: Predicting engagement and frustration." In Educational Data Mining 2013.

[5] Moridis, Christos N., and Anastasios. "Affective learning: Empathetic agents with emotional facial and tone of voice expressions." IEEE Transactions on Affective Computing, vol. 3, pp. 260-272,2012.

[6] Brodny, Grzegorz, AgataKołakowska, Agnieszka Landowska, MariuszSzwoch, WioletaSzwoch, and Michał R. Wróbel. "Comparison of selected off-the-shelf solutions for emotion recognition based on facial expressions." In 2016 9th ICHSI, IEEE. pp. 397-404, 2016.

[7] H. Ding, S.K. Zhou, "Facenet2expnet: Regularizing a deep face recognition net for expression recognition," In Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference pp. 118-126,2017.

[8] Y. Wu, T. Hassner, K. Kim, "Facial landmark detection with tweaked convolutional neural networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017.