IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 8, Issue 4, Aug - Sep, 2020
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org

HUMAN EMOTION DETECTION USING MACHINE LEARNING TECHNIQUES

Rajasekeran S[1], Dr. Kousalya.G[2]
[1]Assistant Professor,Department of Information Technology, KGiSL Institute of Technology,Coimbatore
[2]Professor,Department of Computer Science and Engineering,Coimbatore Institute of Technology, Coimbatore

***Abstract*---**Picture editing is a way to digitalize a picture and execute other activities. This is done to enhance or remove valuable details from the file. Nonverbal means of contact is facial gestures. Eight basic facial expressions contain: optimistic, positive, sad, furious, disdainful, disgustful, fearful and shocked. So, identifying such feelings on the face is really necessary. An elderly surveillance program focused on technologies that detects video picture emotions. Our program proposed involves video processing technology that provides video evidence used for the purpose of real life surveillance of elders

Keywords-facial emotion recognition; Local Binary Pattern Histogram(LBPH)algorithm,Convolutional Neural Networks(CNN)

I. INTRODUCTION

The action of muscles beneath the face can be called facial speech. Facial gestures can communicate countless emotions without a phrase. The human face can communicate countless emotions. So, compared to other modes of nonverbal communication, these facial gestures are not standardized but they are common and can be interpreted by any kind of individual. The facial expressions are similar in various societies for joy, sorrow, anger, outrage, terror and disgust.

Muscle gestures relay the individual's thoughts to those who see them. They are the medium by which social communication is shared between people but often appears in most other animal species and mammals. People may freely or unwillingly follow facial expression. Unwillingness to talk is the type of language when people are sick, wounded or unhappy.

Fig. 1: Example of expression for the six basic emotions

Effective computation involves the analysis of structures and processes and technologies capable of understanding, reading, manipulating and simulating human behavior. Affective computer systems can feel the user's emotions through machines, including sensors, microphones, cameras and react through those unique, pre-defined product / service features One way to look at productive computing is through the connection between the person and the machine where a software can perceive and adapt to user-defined emotions.

The detection and assessment of emotionally based behavior and mental disorder will aid in tracking users' emotional wellbeing. Emotions are conveyed not only through interpersonal actions, but through a variety of physiological adjustments as

well. These shifts in metabolism are not regulated by humans. Thus, physiological signals can represent the actual sensations of the individuals. Several forms of physiological signs, including the electrical cardiometric (ECG), galvanic skin reaction (GSR), Electroencephalogram (EEG), respiratory suspended particles (RSP), and Blood Volume pulse (BVP), are successfully linked to emotional awareness.The most critical of these physiological measures is ECG that demonstrates how heartbeats respond to shifts in mood. Researchers have carried out comprehensive efforts on Emotional Recognition dependent on ECG.The Heart Rate Variability (HRV) derived from ECG is regarded as an significant emotional recognition parameter.

Our aim is to detect the emotions from photographs taken from live webcam in real-time. Now the camera is running a video and the faces are identified by the facial features in the frames that would involve the lips, lips, nose, mouth and corners of the forehead. Such facial features (points), which will be used to identify the facial emotions, were then removed. If the emotions are established, by imaging methods, we look for pain in the emotions.

II. BACKGROUND

A. MACHINE LEARNING

Machine Learning is a means of researching algorithms and models used in the execution of such functions by a computer system without explicit guidance. Algorithms for machine learning generate mathematical model based on some sample data named training data which can be used to forecast or make decisions without external variables that influence the output of the job.

Across many areas machine learning may be graded. The algorithm constructs a mathematical model of a data set for both the inputs and the necessary outputs in supervised learning. Algorithms for classification and regression are subject to controlled analysis. When we choose to restrict the performance to a specific range of values, classification algorithms may be used. For continuous performance regression algorithms are provided because they have meaning inside a certain range. Half-managed learning algorithms build incomplete training models that have no labeling for part of the sample application. The algorithm creates a math in unattended learningIMAGE RECOGNITION

Object recognition relates to technologies that recognizes picture positions, icons, individuals, artifacts, buildings and so on. Image recognition is a machine vision mechanism which can recognise which locate an entity in the digital picture or photograph. Photograph recognition Computer vision is the one where data from film or static images are obtained, interpreted and evaluated from the actual world. The details extracted from such a source are very broad and contain numerical or symbolic knowledge in the form of decisions. Computer vision covers error identification, listening, object recognition, camera analysis and picture r as well as visual recognition.

The human eye views an picture as a series of signals stored in our brain by the visual cortex. Face detection aims to replicate this method precisely. Computer treats an image as a raster or vector image. Raster images have pixel series of discrete numerical values for picture colors, whereas vector images are a collection of color-annotated polygons. The geometric encoding is translated into frameworks that reflect mechanics and artifacts in order to interpret pictures. The machines then evaluate these structures logically. Classification and extraction functions involve the arrangement of results.

The second step is to create a statistical model to use a classification algorithm. We have to practice it by showing thousands of topic and non-subject pictures in relation to the project until the classification algorithm operates. We use neural networks to construct a predictive model. The neural network is a device close to our brain and it measures functions dependent on an immense amount of uncertain inputs. In image recognition, there are several algorithms, such as vector supporters (SVM), face markers evaluation, K-nearest neighbors (KNN), technical regression, etc.

Eye recognition is the third stage. The picture data were structured with both instruction and evaluation results. Training data varies from evaluation data in which duplicates are excluded. This data was applied to the model, which identifies photos in effect. Now we need to train a classifier that measures a new test image and tells us about the closest correspondence with the subject. This requires only milliseconds to this grouping. Either topic or non-subject is the product of the assignment.

B. FEATURE EXTRACTION

Extraction of features is a method of dimensional reduction that reduces an original raw data collection for any form of processing. Characteristics describe an image's actions. Functions generally correspond to the pattern of a point or edge in an picture. The feature abstraction method is helpful if you choose to reduce the amount of data you need to handle while maintaining the relevant knowledge. Through removing features, the sum of repetitive knowledge may be that. The sampled picture utilizes methods such as baseline, resizing, normalization, binarization, etc. and attributes are omitted later. Techniques of extraction are implemented

Many of the feature identification algorithms include ORB and Color Gradient Histogram. The ORT (Oriented FAST) algorithm (Oriented and Rotated BRIEF) is essentially an picture corner combination of the Quick and the BRIEF.ORB method. The FAST part defines characteristics and is known as picture areas with a strong brightness contrast. If more than 8 pixels are higher or darker than a given pixel, it is classified as a bonus. Through translating the derived points as binary vectors, BRIEF demonstrates this. Rgb Gradient Histogram approach essentially calculates the red, green and blue picture values and chooses a picture of identical hue proportions.

Corner detection is a tool used to remove features from computer systems. The derived characteristics are used to deduce an image's meaning. Such tools for the identification of corners include motion recognition, camera registry, video capture, camera mosaicing, 3D modeling and target reconnaissance. Harris corner and shi-Tomasi corner detectors are commonly utilized for the identification of the corner. The Harris Corner Detector is a system in which it decides which windows generate significant differences in pressure, if shifted in the direction of X and Y (i.e. gradients). A threshold and then an significant cor are added.R is calculated by the formula:

$R = \min(\lambda_1, \lambda_2)$, If R is greater than the threshold, it is classified as a corner.

For face feature detection, the Viola Jones algorithm is used since it takes less time and thus provides greater precision[1]. Using basic characteristics such as hair-like attributes, the Viola Jones identifying system distinguishes the patterns or characteristics of a person. This method includes moving feature boxes around the image and measuring the discrepancy between the combined pixel values of neighboring regions. The discrepancy is then applied to a standard showing whether or not an item is observed. It calls for criteria for various function boxes and apps that were learned in advance[2].

Take the photo below into account. There are two nice features in the top half. The first feature selected reflects on the idea that the eye area is deeper than the nose and cheeks. The chosen second feature focuses on the influence of darker eyes than the nose bridge. However, the same windows on the buttocks or somewhere else are meaningless.

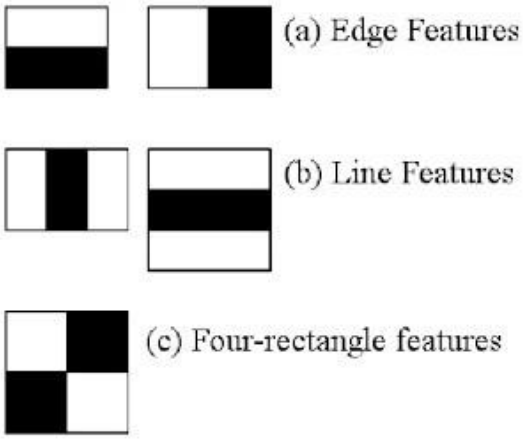
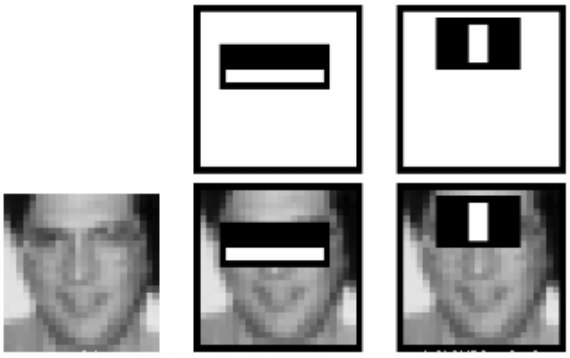


Fig. 2: Feature boxes

The window is placed over the input picture during the detection point. The hair characteristics are determined for each sub-section of the photograph. Various applications represent various values. The gap is then contrasted with an object-separating level. -- hair trait is a "poor grouping" since it senses much more than a random devination. Many of the hair features need adequate precision to differentiate an entity from a non-item and are therefore arranged into cascades to create a good classification system.



There are many steps in the cascade system in which each step has low pupils. A technique called boosting is used to train each level. Boosting lets practice a extremely accurate assessment by taking a weighted average of poor learners' decisions.

At either point the classifier marks the area with either a positive or a negative location in the current sliding frame. Positive indicates an item is detected and negative implies there were no items identified. If the mark is negative, the area is fully labeled and the detector moves the window into the next spot. The classifier can move the region to the next level if the mark is correct. In the final stage marks the field as valid, the detector identifies the item positioned in the current window position.

III. LEARNING METHODS

A. LOCAL BINARY PATTERN HISTOGRAM

Local binary pattern (LBP) is one that tags an picture pixel and thresholds each pixel's neighborhood and calls it a binary number. LBP is seen to boost identification efficiency in conjunction with the histogram of directed gradients. The radius may be used to create a local circular binary pattern and it reflects the area around the central point, normally set to 1. LBFH uses 4 parameters:

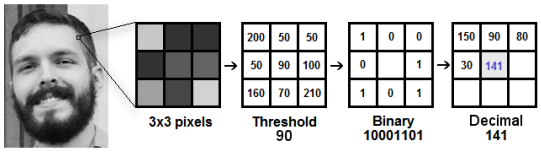
Radius: Distance may be used to create the local circular binary model which describes the distance that is usually set to 1. Radius:

Neighbors: the amount of points needed for the creation of the circular binary local pattern are the neighbours. The higher the computational expense, and typically set at 8 the more sampling points are used.

Grid X: number of horizontally aligned triangles. The smaller the atoms, the greater the scale of the resultant function vector. Normally it's set to 8.

Column Y: the vertical numbers of cells. The smaller the atoms, the greater the scale of the resultant function vector.It is usually set to 8.

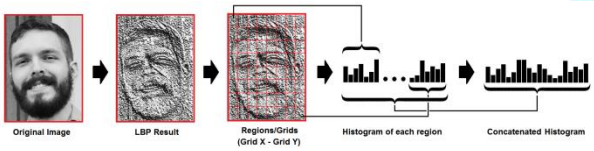
We have to use a multi-faceted dataset and create an ID in order to training the algorithm. Images must have the same ID. In order to best represent the initial picture, the first stage of LBP is to create an intermediate picture



We will take part of this picture in 3x3 pixels as can be seen in the above example because we have a grayscale setting. There are three parameters. It is the matrix which consists of the strength of each pixel (0-255). The new value.1 for values equal to or greater than the threshold, and 0 for values lower than the threshold, is required for each neighboring value of the central value.

The matrix also just includes the discrete form that lacks the main meaning. Set all the binary value to a new binary value from any place. This binary value is then translated to a decimal number, and the matrix is set to the core value. We get a fresh picture at the conclusion of the LBP process with improved picture characteristics.

We can now generate histograms using the above image. We can split the picture into different grids by using Grid X and Grid Y. Just 256 locations (0~255) reflecting the occurrences of increasing pixel strength will be included in increasing histogram from each line. Each histogram generated to reflect each picture from the training dataset will be merged into a larger histogram. Then, given an image entry, the new picture should again be performed and a histogram describing the picture generated. To locate the picture matching the data, only two histograms need to be matched and the image retrieved using the nearest histogram.



So the output of the algorithm is ID with the image with the closest histogram. The algorithm should also return the calculated distance

B.CONVOLUTIONAL NEURAL NETWORKS

The neural networks in Convolution are made up in learning weight neurons. Growing neuron accepts a number of inputs, determines the quantity and generates an output. Convolutionary neural networks have an design distinct from standard neural networks. Standard neural networks transform the input to a sequence of opaque layers. Increasing layer consists of a series of neurons, in which all neurons on the previous layer are entirely connected to each layer. Finally, a layer named the output layer is the only completely connected

Neural networks are isolated from each other. The layers are divided into three dimensions: height, depth and width. In one row, the neurons will not bind to all the neurons on the next sheet. Eventually, the final output would be reduced to one variable, arranged around the depth axis of the chance ratings.

Four terms in CNN are: 1.Convolution 2.ReLu 3.Pooling 4.

Feature Extraction: Convolution

Convolution in CNN is performed on an input image using a filter or a kernel. Filtering and convolution will involves with scanning the screen which starts from top left to right and moving down a bit after covering the width of the screen and repeating the same process until we scan the whole screen. The feature from the face of the individual is lined up with the image. The image pixel is multiplied by the corresponding feature pixel. The values are added and divided by total number of pixels in the feature.

Feature Extraction: Non-Linearity

After sliding our filter over the original image the output which we get is passed through another mathematical function which is called an activation function.

Convolution in the CNN is rendered through a filter or kernel on an input file. Filtering and converting involves scanning the device, beginning from the top left to the right then going a bit down after covering the device width and doing the same cycle, before the entire screen is scanned. The individual's face is associated with the photo. The related actual pixel multiplies the picture unit. The values in the function are applied and differentiated by the total amount of pixels.

Feature Extraction: Pooling

After a convolution sheet, we have to add a pooling or a sub-sample sheet in CNN layers until you have the function charts. Compared to the Convolution system, the spatial scale of the Convolved Component is minimized by the Pooling System. The computational capacity required to process the data is diminished by a decrease in dimensionality. However the method of effectively training of the model is valuable for

removing dominant features that are invariant rotational and positional. Pooling decreases the duration of preparation and tracks health.

Classification — Fully Connected Layer (FC Layer):
Now that our input picture is transformed in the right shape, we flatten the image into a vector frame. The flattened output is fed into a neural feed network and rear propagation is implemented in each training iteration. The model will distinguish those features in photos from the prevailing ones and identify them using Softmax classification. So now we have all the parts expected for CNN building. Convolution, Pooling, and ReLU. In the classifier we talked about initially the output of max pooling, which usually consists of a multi-layer perceptron layer.

CONCLUSION

Based on Video information, the result obtained from the model proposed shows the estimated feeling of the subject. It is able to exploit its success in other cases, estimate mental illnesses and stress rates, and then function to promote, inspire and improve the emotional status of the subject in the case of "sensitive" feelings that contribute to the harmony and peace of mind of the matter. These emotional analyzes are then templates for forming culture into a place.

REFERENCES

[1] Facial Emotion Recognition System through Machine Learning approach, Renuka S. Deshmukh, ShilpaPaygude,Vandanajagtap.

[2] Emotion Detection Through Facial Feature Recognition, James Pao

[3] Renuka S. Deshmukh, VandanaJagtap, ShilpaPaygude, “Facial Emotion Recognition System through Machine Learning approach” ,2017.

[4] Dongwei Lu, Zhiwei He, Xiao Li, Mingyu Gao, Yun Li, Ke Yin, “The Research of Elderly Care System Based on Video Image Processing Technology”,2017.

[5] Shivam Gupta, ”Facial emotion recognition in real-time and static Images”,2018.

[6] Ma Xiaoxi, Lin Weisi, Huang Dongyan, Dong Minghui, Haizhou Li, “FacialEmotion Recognition”,2017.

[7] Mostafa Mohammadpour, Hossein Khaliliardali, Mohammad. M AlyanNezhadi,Seyyed Mohammad. R Hashemi, “Facial Emotion Recognition using Deep Convolutional Networks”, 2017.

[8] Dubey, M., Singh, P. L., ”Automatic Emotion Recognition Using Facial Expression: A Review,” International Research Journal of Engineering and Tech nology,2016

[9]https://stackoverflow.com/questions/42177658/how-to-switch-backend-with-keras-from- tensorflow-to-theano

[10]https://archive.ics.uci.edu/ml/datasets/Grammatical+Facial+Expressions

[11]https://www.edureka.co/blog/convolutional-neural-network/