International Journal of Engineering &Technology, 7 (3.12) (2018) 1144 -1148

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



Automatic Facial Expression Detection System using Single Face Classifier

N. Durga Indira[1*], M.Venu Gopala Rao[2]

[1]Assitant Professor, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur
[2]Professor, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur
**Corresponding Author Email: [1]durgaindira@kluniversity.in*

Abstract

Several factors add to carrying emotions of a person. Posture, speech, facial expressions conduct and exercises are just some of them. Facial appearances are noteworthy in encouraging human correspondence and Associations. Facial appearances can be reflected not just as the most regular Procedure of displaying human feelings yet additionally as a significant to Non-verbal correspondence. Additionally, they are utilized as a noteworthy Device in social investigations and in medicine. In speaking with others, People can recognize feelings of included human with an impressive level of Precision. The issue of programming acknowledgment of outward appearances is yet an ebb and flow inquire about. This framework proposes a programmed outward Appearance appreciation framework, equipped for special the seven all-inclusive Feelings: disgust, anger, fear, satisfaction, trouble and amazement utilizing Profound convolution neural systems. It is intended to be a person independent. As there is a more prominent utilization of human-machine connections nowadays, it is likewise primary for machines to translate the facial ex-pressions and we Would have the capacity to accomplish exactness that is relatively practically Identical to the human mindfulness.

Keywords*: Emotions, Recognition, CNN, Haarcascade, Single face.*

1. Introduction

In existing system, the over-all, a person's emotions are known through their behavior or through some verbal communication [1]. Humans sense and understand faces and facial articulations in a scene with practically zero exertion. According to some investigation, in eye to eye correspondence, feelings are conveyed in offer of 55% through outward appearances [1] [4].

All things considered, development of a programmed framework that accomplishes this assignment is rather troublesome. Distinguishing one's feelings could be valuable by and large. This may not be that viable as just a single's perception is taken after. Empowered by advancements in Machine Learning and Computer Vision, building programmed acknowledgment frameworks has gotten a great deal of responsiveness inside the Computer Science area [7][8].

Restrictions like resolution, quality, brightening or estimate must be taken into worry for picking reasonable strategies, keeping in mind the end goal to de-liver a framework that is powerful, individual autonomous and that splendidly works continuously circumstances [1].

2. Proposed System

As the machine-human interactions are evolving day by day, it would be significantly beneficial if the machine could guess the emotions of a person by learning their facial expressions [1][4]. This is a framework talented of performing programmed acknowledgment of six feelings, caution to be all inclusive over cultures: disgust, anger, fear, happiness, sadness and surprise [1]. This consist of the following works: Face discovery [6], Facial feature extraction [3][4] and emotion grouping. Such scheme would scrutinize the duplicate of a face and produce a planned extrapolation of the expression.

Two face identification [6] calculations are connected for the face area strength organize. Eyes, mouth and eyebrows are recognized as the basic highlights and their component indicates are mined recognize the feeling [1][3]. After component reflection is accomplished a neural system approach [7][5] is utilized to distinguish the feelings confined inside the face. This system had a catholic variety of applications like in unsure finding systems etc.

2.1 Applications of Proposed System

It can be used in suspect detection systems in order to identify whether they are talking truth or lies by studying their facial expressions [4]. It is used to treat autism in children because this benefits in studying their expressions and give a good healing. It is used in the field of medicine i.e. in rehabilitations, counseling etc. It can also be used in e-learning in sensing the state of learner and these benefits in altering the performance style of tutor. This system can form a dynamic music player that plays music suitable with what you feel.

3. System Description

1. Face detection
2. Feature extraction
3. Emotion classification
4. Evaluation

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

3.1 Face Detection

Positive a picture, recognizing the presence of an individual face is a troublesome assignment because of the promising variations of the face. The disparate size, angle and postures person confront capacity inside the picture can cause this variety. Feelings which are deducible from the individual exterior and distinctive imaging conditions, for example, lighting and impediments likewise include facial appearance [1][4]. Likewise, the presence of displays, whiskers, hair and cosmetics have a significant impact in the facial appearance.

A great deal of perceptions has been done on various methodologies in confront discovery [6] can be separated into different methodologies: include invariant approach [3], layout based approach, information based approach and appearance-based approach.

Although different information based and format based techniques can be produced for confront area de-end, an element invariant system in view of the skin shading is chosen as the main strategy because of its flexibility and straightforwardness. While find the veneer segment with skin shade, various calculations can be found for contrast ent shading spaces.

3.2 Feature Extraction

The face district extricated is additionally handled to extract the element focuses fundamental for the feeling arrangement organize [1][3]. In this stage a few components focuses that were perceived as critical were removed to use in the arrangement organize Feature extraction stage can be isolated into two segments: Feature region extraction, Feature point extraction.

These acquired edge pictures are then enlarged and the openings are filled. The outcome edge pictures are utilized as a part of sifting the eyebrow districts. The face distinguished in the before section was further procedure to perceive eye-temples, mouth and eye area. In last stage, a corner point location calculation is utilized to pick up the essential corner focuses from the feature regions

a) Eye Extraction:
b) Eyebrows Extraction
c) Mouth Extraction

Once the feature regions were identified [3][8], these regions are added procedure to mine the essential feature points. At that point the midpoint of left and right most focuses is picked up. This midpoint is utilized to accomplish the best and base corner focuses. In derivation following focuses are distinguished from the best, right most, base and left most focuses, the centroid of the eyes in a picture locale is figured. The point in the eyebrow which is ideal over the eye focus is achieved by handling the data of the edge picture.

3.3 Emotion Classification Stage

The mined feature indicates are taken care of secure the sources of info intended for the neural system. The neural system is taught with the goal that the feelings happiness, neutral, anger, fear, sadness, surprise and disgust are recognized are recognized.

The data sources given the neural system are as per the following.

Left eye height + Right eye height) / 2
= [(c4 - c3) + (d4 - d3)] / 2 (1)
Eye width = (Left eye width + Right eye width) / 2
= [(c2- c1) + (d1 - d2)] / 2 (2)
Mouth height = (f4 - f3), Mouth width = (f2 - f1),
Eye center to Mouth center height =
[(f5 - c5) + (f5 - d5)] / 2 (3)
Originate1at x-axis=Left eye center to Mouth top corner length
Originate2 at x-axis =Left eye center to Mouth bottom corner length
Originate3at x-axis =Right eye center to Mouth top corner length
Originate4 at x-axis=Right eye center to Mouth bottom corner length
Eye originates=Eye width/ Eye height
Mouth originates=Mouth width / Mouth height
Eyebrow originates=Eyebrow to Eye center height / (Eyebrow to Eye center height + Eye center to Mouth center height).

3.4 Evaluation

A Pre-trained prototypical will continuously capture images from the webcam and by using Haarcascade classifier we will detect the face and renovate the image into 48*48 gray scale image and feed them to the pertained convolution neural network. By using the networks, we will detect the emotion and that will be transported or released to the screen [1]. This process will continuously be going on until we close the command prompt.

Functional requirements drive the application architecture of a system. A requirement generates information to the user.

The following are the Functional requirements of our system:

1. We are providing an image as an input.
2. The inputted image is then processed for
   - Face detection
   - Feature extraction
   - Emotion classification
3. For face detection, the HAAR filter in open cv is used to automate face finding.
4. For feature extraction, corner point detection algorithm and other algorithms are used.
5. Neural networks are inputted with feature points to train the algorithms [7][3].
6. Classifiers are used to classify into any of the universal emotions by comparing with the trained dataset [1] and thus to evaluate the result.

Python includes a dynamic sort framework and programmed memory administration. Python translators are accessible for some working frameworks. CPython, the reference execution of

Python is open source programming and has a network based development show, as do about the entirety of its discretionary executions. CPy-thon is an open source which is managed by the Python Software Foundation.

4. Opencv

OpenCV (Open Source Computer Vision) is a library of professional gramming capacities primarily went for continuous PC vision. The openCV is cross-stage and complimentary for use under the open-source BSD permit. OpenCV bolsters the profound get the hang of ing frameworksTensorFlow, Torch/PyTorch and Caffe.OpenCV's application areas include:

2D and 3D feature toolkits
Emotion estimation
Facial recognition system
Gesture recognition
Human–computer interaction (HCI)

4.1 TF Learn

TF learn is a modular and it is visible as a library for deep learning supported by Tensor flow. TF Learn was intended to afford a higher-level Application Programming Interface to Tensor Flow to assist and accelerate experimentations, while remaining completely straightforward and perfect with it.

TF Learn highlights include:

Quick prototyping through exceedingly measured implicit neural system layers, regularizes, optimizers, measurements. Full straightforwardness more than Tensor stream. All the accessible capacities are executed utilizing tensors and tensor stream utilized independently of TF Learn.

Great partner capacities is utilized to help the various data sources, yields and streamlining agents which is utilized to teach any Tensor Flow chart, Simple and fine-looking diagram picture, with insights about angles, weights and initiations
Easy gadget position for utilizing various CPU/GPU.The abnormal state API as of now bolsters the majority of ongoing profound learning models, for example, Convolutions, LSTM, BiRNN, BatchNorm, PReLU, Residual systems, Generative systems... Later on, TF Learn is additionally planned to remain fully informed regarding most recent profound learning procedures.

4.2 Tensor Flow

Is a deep learning library which was introduced by google and it was developed by using CUDA and python? Tensor flow has a fast compilation time than other deep learning libraries like keras& torch which supports for two types of environments GPU (Graphical Processing Unit) and CPU.
Tensor flow was developed to run large numerical computations. Tensors will accept data in the form of multi dimensional array of high dimensions called Tensors. Data is stored in tensors and they typically measured by using a parameter called Rank. The System with pre- defined settings may perfectly install the tensor flow.
The Instructional commands are also simple to create tensor flow environment which leads to dataflow programming over an assortment of assignments. It is a representative math library, and furthermore utilized for machine learning applications, for example, neural networks. [3] It is utilized for both research and creation at Google, frequently supplanting its closed-source predecessor, Disbelief.

4.3 Convolution Neural Network

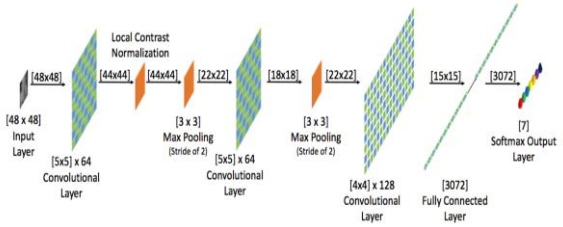

Fig. 1: Basic structure of a convent

A CNN comprises of a considerable measure of layers. These layers when utilized over and again, prompt a development of a Deep Neural Network. Three principle kinds of layers used to fabricate a CNN are:
Input: This layer holds the crude pixel estimations of picture.
Convolution Layer: This layer gets the consequences of the neuron layer that is associated with the info locales. We characterize the quantity of channels to be utilized as a part of this layer. Each channel might be a 5x5 window that slider over the information and gets the pixel with the greatest intensity as the output.
Rectified Linear Unit [ReLU] Layer: This layer applies a component astute enactment work on the picture information. We realize that a CNN utilizes back proliferation [7]. So keeping in mind the end goal to hold similar estimations of the pixels and not being changed by the back proliferation, we apply the ReLU work.
Pooling Layer: This layer play out a down-examining activity along the spatial measurements (width, stature), bringing about volume.

Fully Connected Layer: This layer is utilized to process the score classicize which class has the most extreme score relating to the information digits.

Dataset

The choice of images used for training is responsible for a big part of the performance of the eventual model. This implies the need for a both high subjective and quantitative dataset. For feeling recogni-tion, a few datasets are accessible for examine, changing from a couple of hundred high determination photographs to several thousand littler pictures. The three we will talk about are the Facial Expression Recognition Challenge (FERC-2013), Extended Cohn-Kanade (CK+), and Rad-boud Faces Database (RaFD).

5. Process of Operating an Application

Step 1: Get images from a webcam. Build the convnet. Input is 48x48--3072 nodes in fully connected layer [9].Target classes = ['angry', 'disgusted', 'fearful', 'happy', 'sad', 'surprised', 'neutral'].Self.network = conv_2d (self. Net work, 64, 5, activation = 'relu').self.network = max_pool_2d (self.network, 3, strides = 2)
Step 2: Detect Facial Landmarks
Step 3: Initialize the cascade.
Step 4: Initialize object of EMR class.
Step 5: Append the list with the emoji images.
Step 6: Again, find haar cascade to draw bounding box around face do nothing if no face is detected.
Step 7: Draw box around faces
Step 8: Initialize the cascade
Step 9: Image read from buffer
Step 10: Initialize the first face as having maximum area, then find the one with max_area
Step 11: Extract ROI of face
Step 12: Resize the image so that it can be passed to the neural network.

5.1 Use a Pre-Trained Neural Network Model

Step 1: Initialize object of EMR class
Step 2: Append the list with the emoji images
Step 3: Again, find haar cascade to draw bounding box around face
Step 4: Compute soft max probabilities
Step 5: Write the different emotions and have a bar to indicate probabilities for each class for index, emotion in enumerate (EMOTIONS):
Step 6: Find the emotion with maximum probability and display it. The shape of face_image is (x,y,4)the fourth channel is 0 or 1in most cases it is 0, so, we assign the roi to the emoji

```
if not len(faces) > 0:
        do nothing if no face is detected
        a = 1
else:
        draw box around face with maximum area
        max_area_face = faces[0].
```

5.2 Prediction of Emotions

1. **If image is recognized**

```
    if image is None:
return None
```

2. **If image is not recognized**

```
    image = image.reshape([-1, 48, 48, 1])
    return self.model.predict(image)
```

3. **If pretrained model is loaded**

```
ifisfile("model_1_nimish.tflearn.meta"):
self.model.load("model_1_nimish.tflearn")
print('\n---> Pre-trained model loaded')
```

4. **If pretrained model is not loaded**

```
else:
```

```
self.model.load("model_1_nimish.tflearn")
print("---> Couldn't find model")
```

5. If face is not detected

```
if not len(faces) > 0:
# do nothing if no face is detected
a = 1
else:
```

6. If face is not detected

```
# draw box around face with maximum area
max_area_face = faces[0]

for face in faces:
if face[2] * face[3] >max_area_face[2]* max_area_face[3]:
max_area_face = face
face = max_area_face
(x,y,w,h) = max_area_face
frame = cv2.rectangle(frame,(x,y-50),(x+w,y+h+10),(255,0,0),2)
```

Most of the recognition algorithms focus on two main categories. The featured-based deals with geometric parameters like distance between the eye and size of the eye and image based recognition. The input face database is matched with Eigen faces databases. The top matched Eigen faces databases are constructed for comparison tests.

The test image is the image for future reference to get the image matched with the database. ANxN picture of a face, represented as a vector involves a solitary point in N2 dimensional picture space. Images of face being comparative in general design won't be haphazardly circulated in this large picture space. Discover the vectors that best record for variation of images in entire image space. Construct a face space and project the image into this face space. The geometric Interpretation projects the information along the headings where the information differs and these directions are determined by the vectors of the covariance.

Degree of Confidence: Interested in acquiring a form of confidence score for each detection by the individual cascade classifiers. With such a confidence score I could compare the overlapping regions of interest as proposed by each classifier and remove the regions with the lowest confidence score.

6. Experimental Results

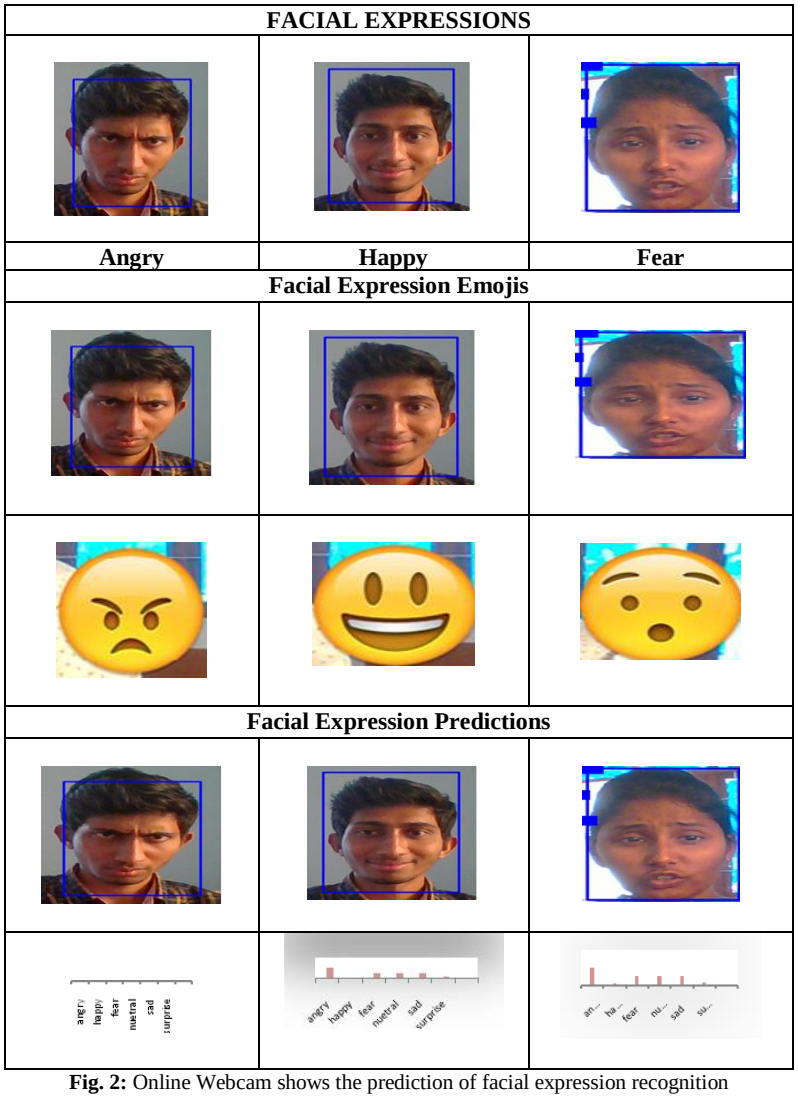

Fig. 2: Online Webcam shows the prediction of facial expression recognition

7. Scope for Further Development

Our system displays the emotion as based on the expression given by the user. As of now, it gives the type to emotion on the screen. We can extend this system to identify the eye diseases and by checking the health count of pupil and Iris. By this, we can identify and predict the eye disease at a shortest period of time and we can host the defect identified cams in social public areas. Auto-

mate the ophthalmology operations is the best way to educate the people for eye vision related issues.

8. Conclusion

The most important in deep learning is to implement the automation system using CNN. There is rapid growth in the change of hardware accordingly the same is reflected in the software and we are improving the real time non automated systems into automatic Systems.

References

[1] An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environmen tAuthor links open overlay panelD. YangaAbeerAlsadoonaP.W.C.PrasadaA.K.SinghbA. Elchouemic

[2] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," Image and Vision Computing, vol. 27, no. 6, pp. 803 – 816, 2009.

[3] Facial Emotion Recognition in Real Time Dan Duncan duncand@stanford.edu Gautam Shine gshine@stanford.edu Chris English chriseng@stanford.edu

[4] Busso, C., Deng, Z., Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., Lee, S., Neumann, U., and Narayanan, S., Analysis of emotion recognition using facial expressions, speech and multimodal information. In ICMI '04: Proceedings of the 6th international conference on Multimodal interfaces, pages 205–211, New York, NY, USA, 2004. ACM.

[5] Emotion Recognition from Facial Expressions using Multilevel HMM Ira Cohen, AshutoshGarg, Thomas S. Huang Beckman Institute for Advanced Science and Technology The University of Illinois at Urbana-Champaign iracohen@ifp.uiuc.edu, ashutosh@ifp.uiuc.edu, huang@ifp.uiuc.edu.

[6] Grimm, M., Dastidar, D. G., and Kroschel, K., Recognizing emotions in spontaneous facial

[7] Genuine and Forged Offline Signature Verification Using Back Propagation Neural Networks Authors:L Ravi Kumar, SudhirBabu, Publication date:2011,International Journal of Computer Science and Information Technologies.

[8] Fingerprint minutia match using bifurcation technique ,L Ravi Kumar, International Journal of Computer Science & Communication Networks, 2012/9,Issue:2,Volume:4,Pages478-486, Publisher: Techno park.

[9] https://github.com/GautamShine/emotion-conv net