JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Emotion Recognition Using Image Processing and Deep Neural Networks

Nayan Bahirame [1]

Dr Satish Gujar. [2]

M.Tech Student, Tssm's Bscoer, Narhe , India [1]

Assistant Professor, Tssm's Bscoer, Narhe, India[2]

***ABSTRACT*:-**

Facial expressions give non-verbal information to humans. Automatic facial expression reorganization is an important role in the first of the nineties. Classical machine learning approaches typically need a fancy feature extraction method and manufacture poor results. Here we apply Mobile NET classifier model for recognizing the emotions. We tend to conjointly apply different loss functions and coaching tricks to be told CNNs with a powerful classification power. The experimental results show that our planned networks beat out progressive ways on the well-known FERC-2013 dataset provided on the Kaggle face expression recognition competition. As compared to the winning model of this competition, the quantity of parameters in our planned networks intensively decreases, which accelerates the general performance speed and makes the planned networks well appropriate for the period system

Keywords—: Image Processing, Deep Neural Networks

I. INTRODUCTION

Emotion recognition is the key feature of the social aware system. Emotion recognition is very important role in the field of education, Security, Defense Services. Speech Emotion Recognition has huge impact with speech based devices. SER models show strong performance on different conditions. The facial recognition system is trained with a supervised learning algorithm. Automatic speech recognition (ASR) training is done on large hours of data. This procedure is very costly. This drawback surely affects the performance analysis of the system.

Understanding human emotions could be a key space of analysis, since recognizing emotions could give an Embarrassment of opportunities and applications, for example, friendlier human-computer interactions with associate increased communication among humans, by refinement the emotional intelligence. Body movement Speaks louder than words. Recent analysis on psychology confirmed, the emotions area unit is most vital in deciding and rational thinking. During a day to day communications human being's categorical differing types of emotions.

With the improvement of technology, we need to analyze the emotions of people. Specialized mathematical tools are used for remodeling the qualitative amount into a quantitative one. Analysis plays a very active role in HCI Human-Computer Interaction. A lot of research is going on facial expression recognition. But the main disadvantage during this technique lies within the undeniable fact that real

feeling may be simply hidden through pretend expressions and conjointly the dependency of the visual pictures on the encircling illumination. However, recent research is currently going on emotion analysis and thermal pictures. Thermal pictures conjointly called thermograms area units primarily show of warmth distribution of AN object in kind of a picture. Thermal pictures area unit captured by a thermal infrared camera that captures the infrared emitted from an object.

II. LITERATURE REVIEW

Yahui Nan, Jianguo Ju, Qingyi Hua Haoming Zhang, Bo Wang 2021
A-Mobile Net: An approach of facial expression recognition [1]

A-MobileNet community version and attention
the module is added to beautify the version's cap potential to extract fine-grained functions of facial expressions, and the dropout era is delivered to save you overfitting. Experimental effects at the FERPlus and RAFDB display that our stepped forward version achieves higher effects than the light-weight MoibleNet collection fashions and different extremely good methods. The reputation accuracy is 84.49% and 88.11% at the RAF-DB and FER Plus, respectively.

Chenghao Zhang ,Lei Xue 2021

Autoencoder with Emotion Embedding for Speech Emotion Recognition [2]

Encoding and Decoding is used for recognizing the emotions. The emotion embedding path focuses on learning information from labels. This allows the representation from the auto encoder which deep learning is used to find out the emotions. In future work, considering the powerful capabilities of BERT in natural language processing tasks,

Hao Hu, Ming-XingXu, and Wei Wu
GMM super vector-based SVM with spectral features for speech emotion recognition [3]

In this paper, we recommend using the GMM notable vector primarily based SVM with spectral functions to speech emotion popularity. The GMM KL divergence kernel turned into proven to yield higher overall performance than different usually used kernels withinside the proposed gadget. The outcomes propose that the gender statistics ought to be taken into consideration in speech emotion popularity, and display that the GMM notable vector primarily based SVM gadget drastically outperforms general GMM gadget. For the regularly pressured emotional states, different kinds of functions, along with prosodic and voice excellent functions may be fused with our proposed technique to decorate the emotion popularity overall performance in destiny work.

Dong Hwa Kim 2013
Fuzzy rule-based voice emotion control for user demand speech generation of emotion robot [4]

Speech emotion recognition is important for emotion robots and it has lots of applications. A robot can interact with humans and recognize humans with emotions. Now a day's robot service growing day by day here robots can understand the emotional situation but not only spoken command.

Yuanyuan Zhang, Jun Du, Zirui Wang, Jianshu Zhang, Yanhui Tu 2018

Effective attention mechanism in dynamic models for speech emotion recognition [5]

We verified that the CNN architectures designed for visible popularity may be at once tailored for speech emotion popularity. Besides, it's exciting to peer the switch getting to know can construct a strong bridge between herbal pictures and speech signals. Finally, we proposed an interest primarily-based FCN version. Our version is capable of coping with utterances with variable lengths and the eye mechanism empowers the community to consciousness on emotionally salient areas of a spectrogram. Our machine achieves past today's accuracy at the benchmark dataset MOCAP.

Zhaocheng Huang 2015

An investigation of emotion changes from speech [6]

To address presently unrequited issues, my analysis is anticipated to be the primary systematic investigation of feeling changes, compared with existing feeling recognition from speech analysis. This can contribute to the emotional computing analysis community in terms of the latest insights towards feeling modification issues. Conjointly it will profit a spread of analysis areas like task transition wherever psychological feature load is interfered by feeling changes, and alter quickly, feeling regulation wherever temporal order of recognizing feeling modification is important and sturdy-feeling recognition wherever temporal data facilitates recognizing emotions

Existing System

1. Facial Expression recognition is the most currently used technique for emotion recognition.
2. But in certain Cases, humans can't express emotion such as in hospitals.
3. Using Only Image processing technique we can't recognize the facial expression
4. So here we need to use Image Processing with CNN [Convolution Neural Network] LSTM Model to recognize the human emotions with image or voice.

Problem Statement

The problem statements we've got are having strong and automated face detection, analysis of the captured image and its meaningful analysis by facial expressions, creating data sets for taking a look at and coaching and so the planning and therefore the implementation of utterly fitted classifiers to be told underlying classifiers to be told the vectors of the facial descriptors.

We propose a model design that is capable of recognizing up to six models that are thought-about universal among all walks of cultures. The main are concern, happiness, sadness, surprise, disgust, and in conclusion surprise.

Algorithm Steps

Step 1: The sample image or video frame is provided as input.

Step 2: The Image Histogram is plotted.

Step 3: Mixing of Data. Dataset divides into train and test. Apply the CNN model after the divide.

Step 4: Predicting the human emotion from that trained data (sample no. - Predicted value - actual value)

II. PROPOSED SYSTEM

Any communication has two types verbal and nonverbal. Sharing mute clues or info is termed non-verbal communication. This includes visual communication like a physical look. Human feelings are often known as exploitation visual communication and posture. Posture provides info that isn't a gift in speech and face emotion. For instance, the spirit of an individual from an extended distance is often known as exploitation human posture. Thus human emotion recognition is achieved by physical look. Emotion includes like happy, anger, sad, disgust, etc Rising studies shows that folks will accurately decrypt emotional cues from others' nonverbal communications and might create logical thinking concerning the emotional states of others. A precise cluster of body actions is termed gestures. The action is often performed principally by the pinnacle, hands, and arm. These cues along and convey info of

emotional states and therefore the content within the interactions. With the support from psychological studies, distinctive emotions from material body movement have many applications. Convolution neural networks (CNN) are the preferred manner of analyzing pictures. CNN is completely different from a multi-layer perception (MLP) as they need hidden layers, known as convolution layers. The methodology works on the CNN framework. The primary level counseled is background removal, accustomed extract emotions from a picture, Here, the conventional CNN network module is employed to extract primary expressional vector (EV)... Nose, ears, etc.)

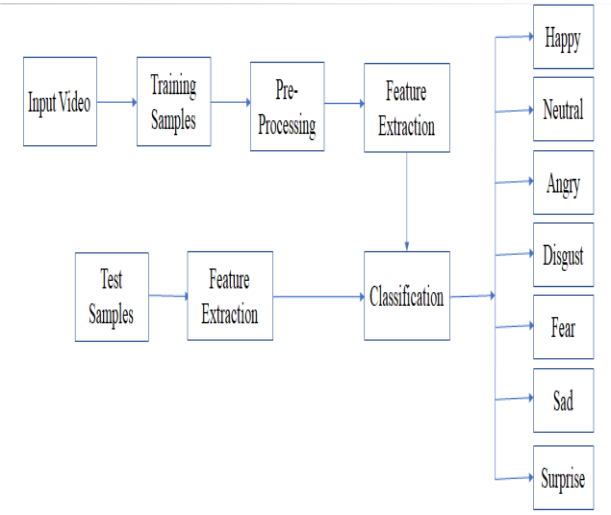

Fig1. System Architecture.

Data Sets

Neural networks and deep networks specifically, square measure familiar for his or her would like for big amounts of coaching information. Moreover, the selection of pictures used for coaching is answerable for a giant part of the performance of the ultimate model. This allows or needs an extremely qualitative and much bigger volume of a dataset. Feeling recognition has many standardized and well-acknowledged datasets on the market online that vary from many hundred photos to tens of thousands of pictures whereas additionally travel vastly in resolution. The datasets dissent principally on the amount, quality, and 'cleanliness' of the pictures. For this method, we tend to use the FER 2013 dataset which has thousands of faces with all kinds of emotions. For this method, coaching is going to be done victimization 9000 samples from the FER-2013 information with another one thousand new samples for validation.

Training Data Set: - 75 %
Testing Data Set:- 25%

III. RESULTS

1. Happy Image

Fig3. Happy Image

The Fig3. Shows Happy Image from a live video stream.

2. Sad Image

Fig4. Sad Image

The Fig4. Shows Sad Images from live Stream.

3. Neutral Image

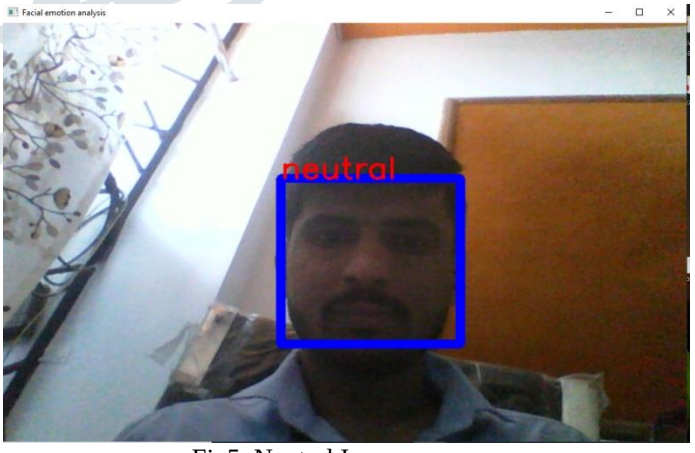

Fig5. Neutral Image

The Fig5. Shows a Neutral Image from the live stream.

4. Surprise Image

Fig6. Surprise Image

The Fig6. Shows a surprise image from the live stream.

5. Angry Image

Fig7. Angry Image

The Fig7. Shows an angry image from the live stream.

IV. EXPERIMENTAL RESULT & ANALYSIS

1. Emotion Plot Count

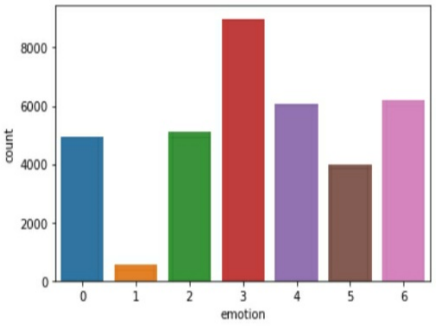

Fig8. Emotion Plot Count

The fig.8 Displays Emotion Count from the dataset.0-Angry 1- disgust 2-fear 3-happy 4-sad 5-surprise 6-neutral.

2. Feature Extraction

Fig9. Feature Extraction

The Fig9. Shows feature extraction from the test set

3. Accuracy on Test Set

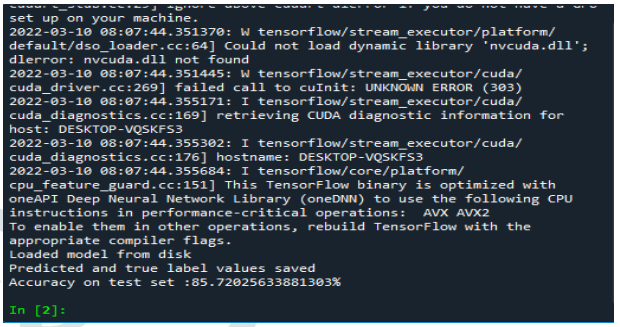

Fig10. Accuracy on Test Set

The Fig10. Shows the accuracy of the test set.

4. Confusion Matrix

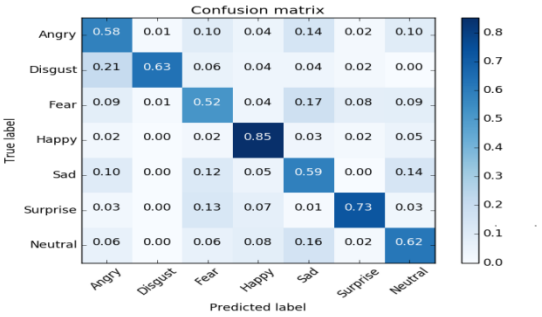

Fig11. Confusion Matrix

The Fig.11 figure shows the confusion matrix of the model with predicted label and true label.

5. Training & Validation Graph

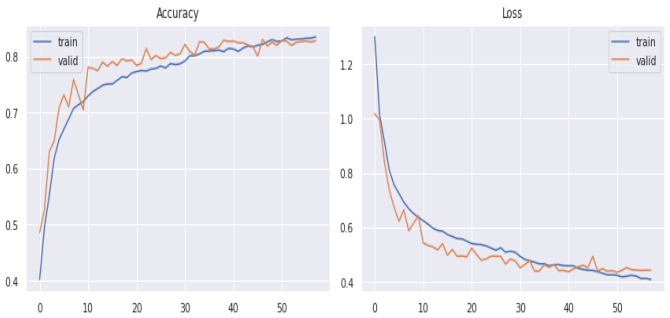

Fig12. Training & Validation graph

The Fig12 shows that accuracy gradually increased up to 80 on training & validation sets

V. CONCLUSION

In this paper, we propose a model for predicting human emotions from body movements on the sequence of frames. This model is representing deep convolution options to extract strikingness info at multiple scales.
Emotion recognition could be a novel approach to facial feeling detection that uses the advantages of CNN and supervised learning (feasible due to huge data). the most advantage of the Emotion recognition rule is that it works with different orientations (less than 30°) because of the distinctive twenty-four digit long electron volt feature matrix. The background removal more an excellent advantage in accurately deciding the emotions. Emotion recognition might be the starting step, for several of the emotion-based applications such as polygraph and additionally mood-based learning for college kids, etc.

VI. REFERENCES

1. A MobileNet: An approach of facial expression recognition Yahui Nan, Jianguo Ju, Qingyi Hua, Haoming Zhang, Bo Wang Received 26 August 2021; revised 18 September 2021; accepted 27 September 2021 Available online 19 October 2021.

2. Chenghao Zhang, and Lei Xue, " Autoencoder With Emotion Embedding for Speech Emotion Recognition.," 2021 Autoencoder With Emotion Embedding for Speech Emotion Recognition.

3. Hao Hu, Ming-XingXu, and Wei Wu, " GMM super vector-based SVM with spectral features for speech emotion recognition.," 2007 IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP '07.

4. Dong Hwa Kim, " Fuzzy rule-based voice emotion control for user demand speech generation of emotion robot," 2013 International Conference on Computer Applications Technology (ICCAT).

5. Po-Wei Hsiao and Chia-Ping Chen, “Effective attention mechanism in dynamic models for speech emotion recognition," International Journal of Speech Technology, vol. 19, no. 4, pp. 1-11, August 2018.

6. Zhaocheng Huang, "An investigation of emotion changes from speech; 2015 International Conference on Affective Computing and Intelligent Interaction (ASCII)