SPEECH EMOTION RECOGNITION WITH FACIAL EXPRESSIONS

Dr.R.Prema

Asst.Prof (CSE)

SCSVMV University

Kanchipuram

A.Vamsi Kumar

B.E(CSE)

SCSVMV University

Kanchipuram

B.Sai Eswar Reddy

B.E(CSE)

SCSVMV University

Kanchipuram

ABSTRACT:

The speech emotion recognition is a very interesting yet very challenging task of human computer interaction. In the recent years this topic has grabbed so much attention. In the field of speech emotion recognition many techniques have been utilized to extract emotions from signals , including many well-established speech analysis and classification techniques. In the traditional way of speech emotion recognition features are extracted from the speech signals and then the features are selected which is collectively know as selection module and then the emotions are recognized this is a very lengthy and time taking process.

To understand the importance of Speech emotion & facial expression and how it is important for emotion detection. Emotion detection from our project has vast study for further future research and development. To build automatically speech and facial expression recognition has an effective emerging for developments. This project extends study and development on emotion detection with speech and facial expressions using Convolution Neural network (CNN). This task is done by detecting the facial actions per every unit of face measurements as a sub part of facial action coding system. This project offers light on utilization of CNN from a live video stream as an input. Using various machine learning libraries like tensor flow and many more. With this development it's advantageous to various domains such as medical engineering, technology, marketing etc..

Key words:- Emotion , Features , Facial expressions , Speech , Techniques , Technology.

INTRODUCTION:

In mortal- computer commerce, emotional processes are inseparably connected to rational opinions; hence affective commerce has gained great attention. thus, it has come an important issue to identify the stoner emotional state. Grounded on cerebral proposition, it's extensively accepted that six archetypal feelings can be linked surprise, fear, nausea, wrathfulness, happiness and sadness. Facial emotion and tone of the speech play a major part in expressing these feelings. feelings can significantly change the communication sense occasionally it isn't what was said that's the most important, but how it was said. The face tends to be most visible form of emotion communication, but it's also most fluently controlled in response to different social situations when compared to the voice and other ways of expression. Further a brief review of being emotion recognition systems is presented. Speech Emotion Recognition Studies Following the long tradition of speech analysis, numerous sweats were taken to fete affective countries from oral information. Starting in the 1930s, some important voice point vectors have been chosen for exploration abecedarian frequency, time- energy distribution vector, MFCC, LPCC portions, etc. Williams and Stevens(1972) studied the spectrograms of real emotional speech and compared them with acted speech. They set up parallels which suggest the use of acted data. A qualitative correlation between emotion and speech point was presented by Murray and Arnott(1993). Petrushin(1998) compared mortal and machine recognition of feelings in speech and achieved analogous rates for both. To exploit the dynamic variation along an utterance, Mel Frequency Cepstral Portions were employed. Nwe(2001) achieved an average delicacy of 70 for six feelings acted by two speakers using 12 MFCC features as input to a separate retired Markov Model. Busso(2004) also argued that statistics relating to MFCCs carry emotional information. Yu et.al.(2002) used Support Vector Machines as double classifiers. On four distinct feelings, they achieved an delicacy of 73. Lee(2002) tried to distinguish between negative and positive feelings, in call center terrain, using direct demarcation, k- NN and SVM classifiers achieving a maximum delicacy rate of 75. Batliner(2003) studied a 4- class problem with inspired feelings in robotic speech. Facial Expression Recognition Studies Since 1970, Paul Ekman has performed expansive studies on mortal facial expressions. He set up that facial expressions of emotion aren't culturally determined, but universal to mortal culture and therefore natural in origin. He developed the Facial Action Coding System(FACS) where movements on the face are described by a set of action units(AUs). The studies in computer- supported recognition of facial expressions started in 1990s. The features used are generally grounded on original spatial position of specific points and regions of the face(edge of the mouth, eyes, eyebrows). Mase(1991) was one of the first experimenters who used image processing ways to fete facial expressions. With 11 windows manually located in the face, the muscle movements were uprooted by the use of optic inflow. nearest neighbour rule

was employed for the bracket task of four feelings, with an delicacy of 80. Rosenblum(1996) and Otsuka(1997) also developed an optic inflow region- grounded approach, by applying a Radial Base Function Network and a Hidden Markov Model, independently. Tian etal (2000) explored AUs recognition by using endless and flash facial features(lips, wrinkles). Geometrical models were used to detect their shapes and appearances. They achieved 96 delicacy with a Neural- Network- Grounded classifier. Cohen(2003, 2004) introduced the Bayesian Network classifiers in the static settings and a multi-level HMM classifier to automatically member an arbitrary long sequence to the corresponding facial expressions.

LITERATURE SURVEY:

S.NO	TITLE	AUTHORS	TECHNIQUES	DISADVANTAGES	YEAR OF PUBLICATION
1	Speech Emotion Recognition from Spectrograms with Deep Convolutional Neural Network	. M. Badshah, J. Ahmad, N. Rahim, and S. W. Baik	Deep neural networks	If face areas are not detached, feature coinciding occurs and effects in performance degradation	2017
2	Emotion Recognition of Stressed Speech using Teager Energy and Linear Prediction features	S.B. Reddy, T. Kishore Kumar	Linear prediction features	Doesn't provide accurate results to all the facial expressions compared to other algorithms.	2018

3	Audio-visual emotion recognition	Li Zheng et al.	CNN random forest	Variations of learning rate may affect output	2018
4	Emotion Detection from Speech Signals using Voting Mechanism on Classified Frames	Ashwini Rajasekhar, et.al,	SVM	They used computerized voice datasets it is hard to find human voice.	2018
5	Evaluating deep learning architectures for Emotion Recognition	Renjith S, Manju K G.	LPCC /ANN	Accuracy is 67.18%	2017

EXISTING SYSTEM:

In the existing system affective computing is the "computing that relates to, arises from, or influence emotions", or in the other words, any form of computing that has something to do with emotions. The creation of automatic classifier involves collecting information, extracting the features which are important and finally training the data, so it classify and recognize some patterns.Emotion detection using speech, gathering emotional information from the user of a system is their voice.Any emotion from the speakers speech is represented by the large number of parameters which is contained in the speech and changes in these parameters will result in corresponding changes in emotions which is quite difficult.

DRAWBACKS:

- ➢ Creation of model in real life is difficult.
- ➢ Voice recognition software won't always put your words on the screen completely accurately.
- ➢ Programs cannot understand the context of language the way that humans can, leading to errors that are often due to misinterpretation.

PROPOSED METHOD:

To overcome the existing drawbacks, comparing the traditional machine learning approaches, deep learning based methods have shown better performance in terms of accuracy and speed of processing in speech recognition. We have used a modified Convolutional Neural Network (CNN). CNN is mostly used in image and speech recognition. CNN is a kind of artificial neural networks that employs convolutional methodology to extract features from the input data to increase the number of features. That captures each frames and test them and is trained by CNN model and later classified into different emotions. With computational power of Graphical Processing Units(GPU's), CNN has achieved remarkable cutting edge results in speech& image recognition.

ADVANTAGES:

- ➢ Automatically detects the important features without any human supervision.
- ➢ Gives good accuracy.
- ➢ Computationally Efficient

SYSTEM ARCHITECTURE:
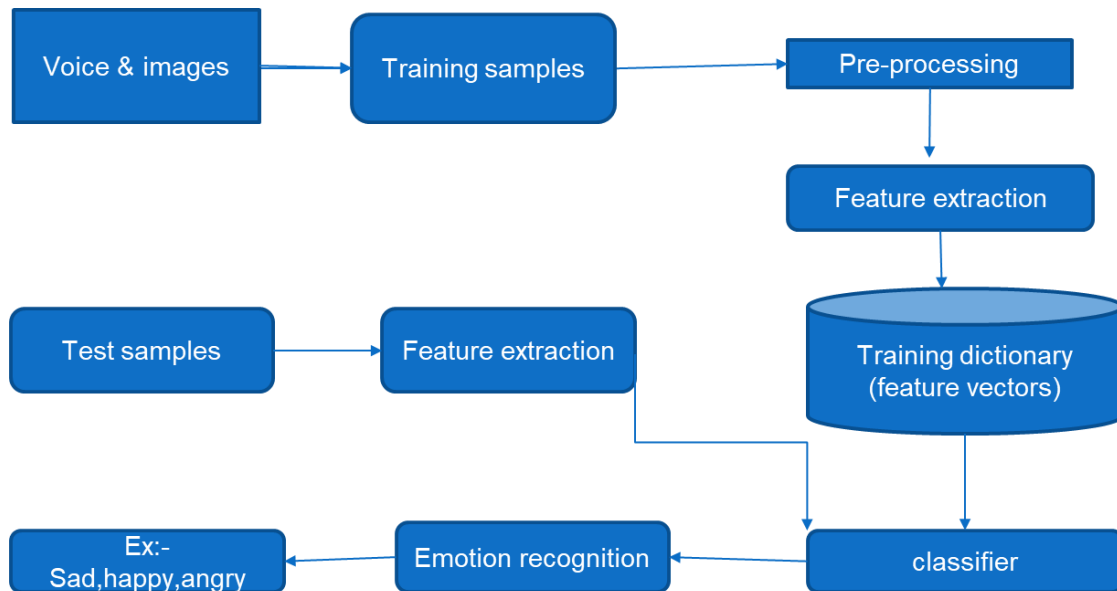
IMPLEMENTATION:

- ➢ Firstly,we have collected the dataset of speech and facial expressions.i.e ravdess dataset.
- ➢ We have to upload the dataset and the dataset will be preprocessed,after preprocessing it will give what is the accuracy that is obtained by this technique.
- ➢ By moving ahead we have to train the sppech and facial dataset with CNN algorithm that we are using.
- ➢ We can also get the accuracy comparision graph between the speech and facial expressions.
- ➢ We have to upload the image that we want to know the emotion ,then the emotion of the image will be displayed.
- ➢ We have to upload the speech sample that we want to know the emotion ,then the emotion of the speech sample will be displayed.

CONCLUSION:

Emotion recognition gives an occasion to significantly ameliorate bias like buses , phones, TVs, office outfit and indeed ménage appliances and systems by enforcing new features and interfaces, which would be much further intuitive and able of bus adaption to stoner requirements. Large commercial companies use Affective computing, who want to know at all costs whether their products, services and marketing strategy addresses the requirements and tastes of guests. similar technology can be delved and developed. Emotional expression of the subject can be used to determine their possible anti-social motives. In this we developed the speech emotion recognition with facial expressions included in it. This task is done by detecting the facial conduct per every unit of face measures as a sub part of facial action rendering system.

Emotion recognition is the subject which can be used to determine their possibleanti-social motives. To understand the significance of Speech emotion & facial expression and how it's important for emotion discovery. Emotion discovery from our design has vast study for farther unborn exploration and development. Emotion modeling has an intriguing part in the coming generation mortal machine interactive systems.

REFERENCES:

1. A. M. Badshah, J. Ahmad, N. Rahim, and S. W. Baik, "Speech Emotion Recognition from Spectrograms with Deep Convolutional Neural Network," 2017 Int. Conf. Platf. Technol. Serv., pp. 1–5, 2017.

2. Noroozi, F., Marjanovic, M., Njegus, A., Escalera, S., & Anbarjafari, G. Audio-visual emotion recognition in video clips. IEEE Transactions on Affective Computing,2017

3. Bettadapura, V. (2012). Face expression recognition and analysis: the state of the art. arXiv preprint arXiv:1203.6722.

4. Shan, C., Gong, S., &McOwan, P. W. (2005, September). Robust facial expression recognition using local binary patterns. In Image Processing, 2005. ICIP 2005. IEEE International Conference on (Vol. 2, pp. II-370).IEEE.

5. Bhatt, M., Drashti, H., Rathod, M., Kirit, R., Agravat, M., &Shardul, J. (2014). A Studyof Local Binary Pattern Method for Facial Expression Detection. arXiv preprintar Xiv: 14 05.6130.

6. Chen, J., Chen, Z., Chi, Z., &Fu,H. (2014, August). Facial expression recognition based on facial components detection and hog features. In International Workshops on Electrical and Computer Engineering Subfields (pp. 884-888).

7. Ahmed, F., Bari, H., & Hossain, E. (2014). Personindependent facial expression recognition based on compound local binary pattern (CLBP). Int. Arab J. Inf. Technol., 11(2),195-203.

8. Happy, S. L., George, A., &Routray, A. (2012, December). A real time facial expression classification system using Local Binary Patterns. In Intelligent Human Computer Interaction (IHCI), 2012 4th International Conference on (pp. 1-5). IEEE.

9. Zhang, S., Zhao, X., & Lei, B. (2012). Facial expression recognition based on local binary patterns and local fisher discriminant analysis. WSEAS Trans. Signal Process, 8(1),21- 31.

10. Chibelushi, C. C., &Bourel, F. (2003). Facial expression recognition: A brief tutorial overview. CVonline: On-Line Compendium of Computer Vision,9.

11. Sokolova, M., Japkowicz, N., &Szpakowicz, S. (2006, December). Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In Australasian Joint Conference on Artificial Intelligence (pp. 1015- 1021). Springer BerlinHeidelberg.

12. Michel, P., & El Kaliouby, R. (2005). Facial expression recognition using support vector machines. In The 10th International Conference on Human-Computer Interaction, Crete,Greece.

13. S.B. Reddy, T. Kishore Kumar , "Emotion Recognition of Stressed Speech using Teager Energy and Linear Prediction Features," in IEEE 18th International Conference on Advanced Learning Technologies, 2018

14. Zamil, Adib Ashfaq A., et al. "Emotion Detection from Speech Signals using Voting Mechanism on Classified Frames." 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST). IEEE, 2019.