IDENTIFICATION OF EMOTIONS FROM SPEECH

Mr.M.Thirunavukkarasu
Assistant Professor of Computer Science and Engineering
SCSVMV
Kanchipuram,Enathur
mthiru@kanchiuniv.ac.in

KARTHIK BOMMARTHI
Computer Science & Engineering
SCSVMV
Kanchipuram,Enathur
11199a034@kanchiuniv.ac.in

PURUSHOTHAM REDDY ALAMURU
Computer Science & Engineering
SCSVMV
Kanchipuram,Enathur
11199a011@kanchiuniv.ac.in

Abstract— A fascinating and difficult task in human-computer interaction is recognizing emotions from speech. This topic has received a lot of interest recently. In the field of speech emotion recognition, many methods have been used to extract emotions from signals. This includes many well-known speech analysis and classification methods. Traditional methods of recognizing speech emotion extract features from the audio signal and select those features. This is a process known as the selection engine, which then identifies emotions. This is a very drawn-out process. Understand the importance of language, emotions and how they are important for emotion recognition. Identification of Emotions from our project has undergone extensive research for further research and development in the future. Automatically building speech recognition has proven effective in development. This project extends the research and development of emotion recognition through language using convolutional neural networks (CNNs). This project sheds light on using a CNN from a live video stream as input. Using various machine learning libraries such as Tensor Flow. Various fields have benefited from this development, such as medical technology, technology, and marketing.

Keywords— Emotion , Features, Speech , Techniques , Technology.

I. INTRODUCTION

In mortal- computer relations, emotional processes are nearly related to rational opinions. Emotional relations thus admit a lot of attention. thus, relating the stoner's emotional state has come an important challenge. Grounded on the proposition of psychology, it's generally accepted that we can identify her six typical feelings surprise, fear, nausea, wrathfulness, happiness and sadness. Facial expressions and speech bows play an important part in expressing these feelings. feelings can greatly change the meaning of a communication. occasionally the most important thing is how it was said, not the most important thing. Faces tend to be the most visible form of emotional communication, but compared to voice and other forms of expression, they're also the easiest to control in response to different social situations. A brief overview of the recognition system is given. Speech Recognition Research Following the long tradition of speech analysis, numerous sweats have been made to descry emotional countries from speech information. Since the 1930s, several important speech point vectors have been named for study. Abecedarian frequence, time- energy distribution vector, MFCC, LPCC portions, etc. Williams and Stevens(1972) studied spectrograms of real emotional speech and compared them with dissembled speech. They set up parallels that point to the use of traded data. Murray and Arnott(1993) presented a qualitative correlation between affective and language function. Petrushin(1998) compared mortal and machine recognition of emotion in speech and attained analogous proportions for both. The MelFrequency cepstral measure was used to exploit dynamic variation along utterances. Nwe(2001) achieved an average delicacy of 70 for her six feelings represented by two speakers using his 12 her MFCC features as inputs to a separate retired Markov model. did. Busso(2004) also argued that MFCC statistics contain emotional information. Yura.Al.(2002) used support vector machines as double 2 classifiers. They achieved 73 delicacy for four different feelings. Lee(2002) used direct discriminant, k- NN and SVM classifiers to try to distinguish between negative and positive feelings in a call center terrain and achieved an delicacy of over to 75. Batliner(2003) studied four classes of inspired affective problems in robotic language.

II. OBEJCTIVE

1. Upload speech Emotion dataset
2. Preprocess Dataset
3. Train Speech Emotion CNN Algorithm
4. Predict Speech emotion

III. LITERATURE SURVEY

S.No	Title	Authors	techniques	Disadvantages	Year
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 coincidding 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

IV. PROBLEM STATEMENT

There is a large body of literature and research on applying these algorithms to understand emotions and states of mind from human language. A wave of research is currently emerging on emotion identification and its applications, along with neural networks and applications of long-short-term memory(LSTM) network improvement, adversarial generative models, and more.

It is important to understand its application and role in emotions. Therefore, the purpose of this article is to understand machine learning techniques for speech emotion recognition, from databases to models. After further application of machine learning and feature extraction techniques, it is very difficult to achieve high model accuracy. This is because similarities between different emotions, such as happy and surprised emotions, have the same kind of frequency and tone.

Voice length is also an issue, as we all know that human emotions are not constant across sentences, but rather constantly changing. So the system needs to identify

pieces of data to fully understand the emotion of the voice.

V. PROPOSED SYSTEM

To overcome existing shortcomings, deep learning-based methods have shown superior performance in terms of speech recognition accuracy and processing speed compared to traditional machine learning approaches. We used a modified convolutional neural network (CNN). CNN is mainly used in speech recognition. A CNN is a type of artificial neural network that uses convolutional methods to extract features from input data and increase the number of features. It captures and tests every frame, trained by a CNN model, and later classified into different emotions. Using the computational power of graphical processing units (GPUs), CNNs have achieved remarkable excellence in identifying emotion in speech.

V.ARCHITECTURE

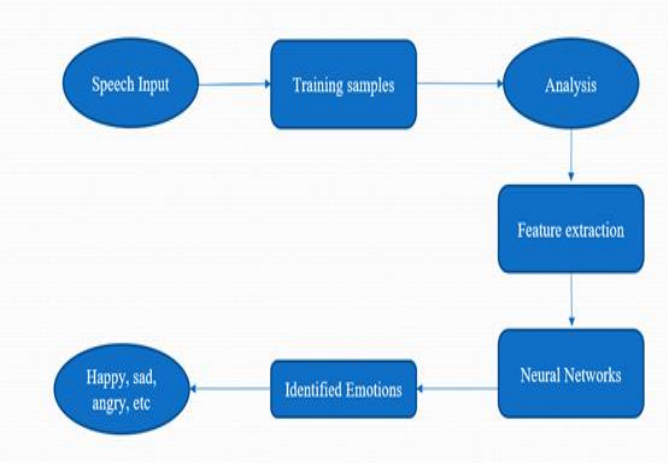

Fig 1 – System Architecture

UML DIAGRAMS

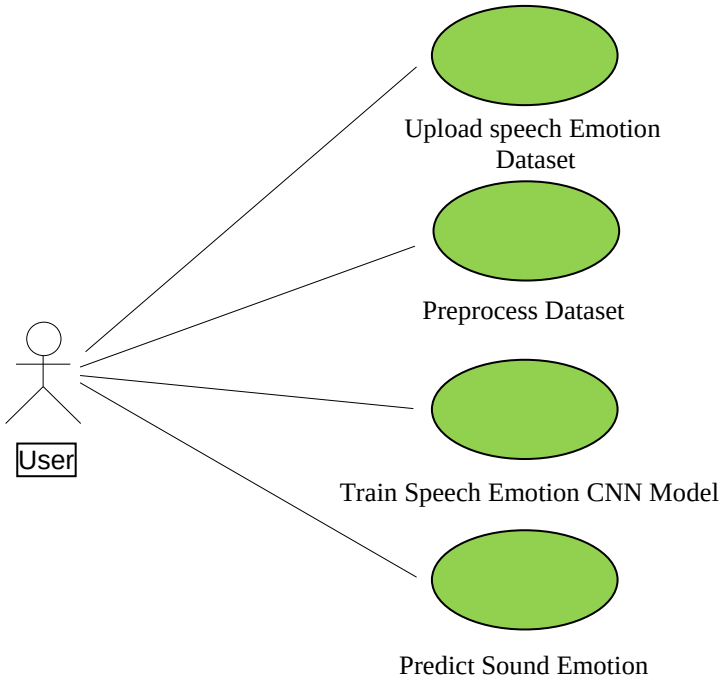

Fig 2 – Use Case Diagram

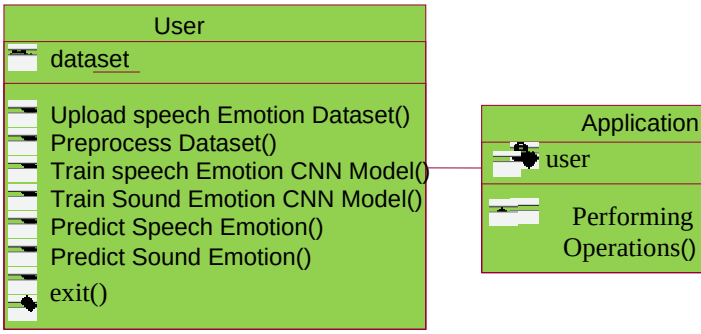

Fig 3 – Class Diagram

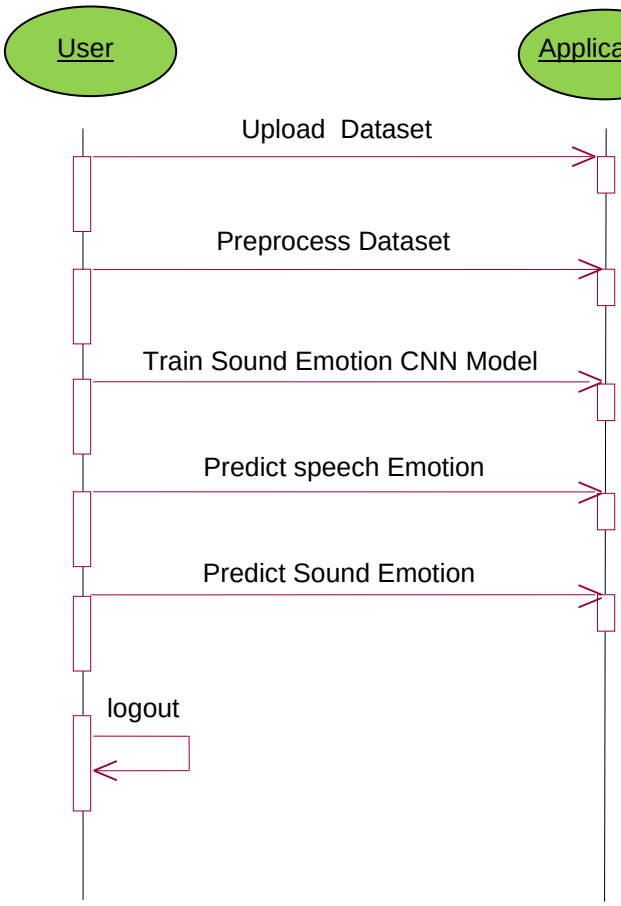

Fig 4 – Sequence Diagram

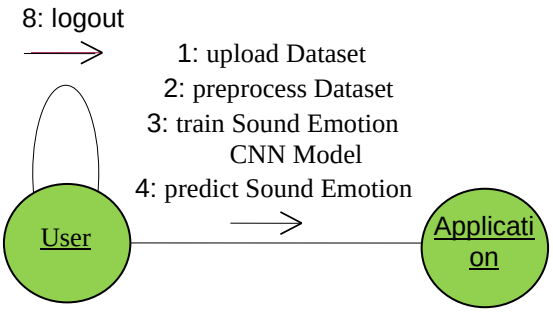

Fig 5 – Collaboration Diagram

IMPLEMENTATION

- First, we collected speech datasets. H. Love Death Records.
- The dataset must be uploaded and the dataset will be preprocessed. After preprocessing, the accuracy obtained with this technique is given.
- Next, we need to train the voice datasets using the CNN algorithm we want to use.
- You need to upload the voice sample you want to know the emotion, and the emotion of the voice sample will be displayed.

ALGORITHM

CONVOLUTION NEURAL NETWORK(CNN) :
Artificial neural networks are used in a variety of classification tasks such as sounds, and words. Different types of neural networks are used for different purposes. For example, word order prediction uses recurrent neural networks, or more precisely LSTMs, and similarly image classification uses convolutional neural networks.

1.Input Layers: The layer that provides the input to the model. The number of neurons in this layer corresponds to the total number of features in your data.

2.Hidden Layer: Inputs from the input layer are then sent to the hidden layer. Depending on the model and data size, there may be many hidden layers. Each hidden layer can have a different number of neurons, but this is typically larger than the number of features. Each layer's output is computed by a matrix that multiplies the previous layer's output by the learnable weights for that layer, plus the learnable biases, followed by an activation function that makes the network nonlinear.

3.Output Layer: The output of the hidden layer is fed into a logistic function such as Sigmoid or Softmax, which transforms each class output into a probability value for each class.

VI.RESULTS AND OUTPUTS

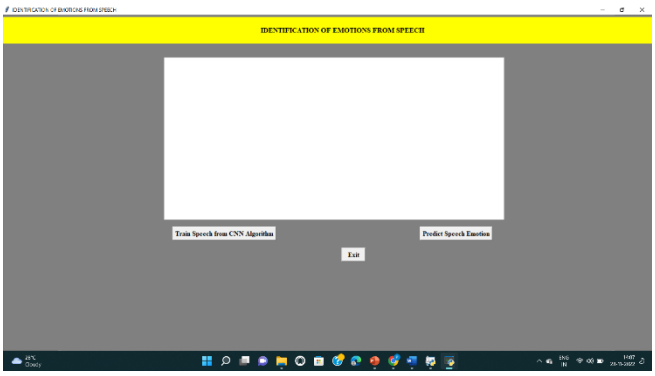

Fig 2: Interface

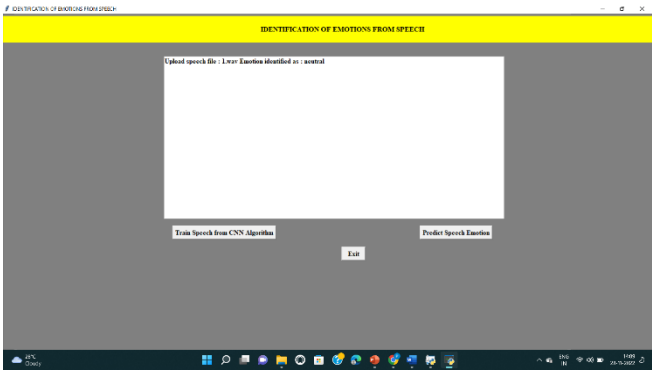

Fig 3: Identification of emotion of speech : Neutral

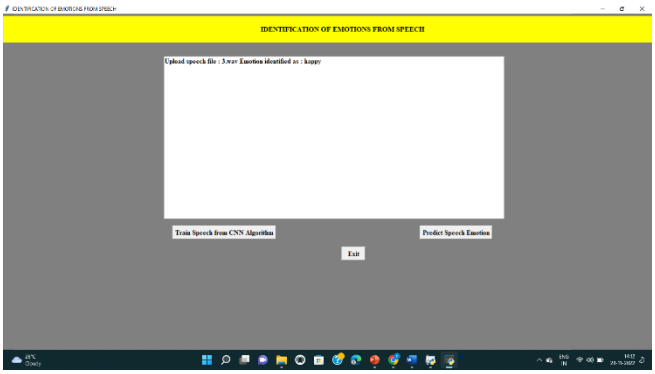

Fig 4: Identification of emotion of speech : Happy

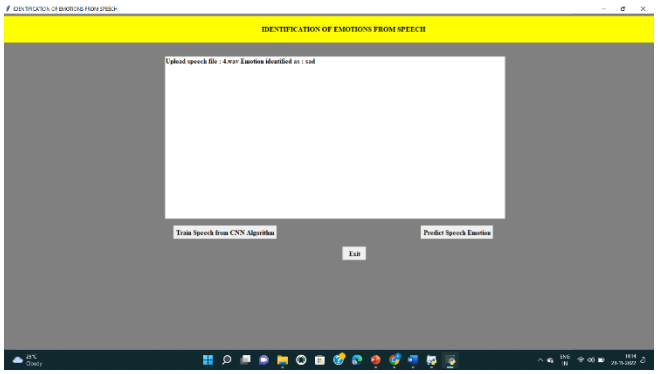

Fig 5: Identification of emotion of speech : Sad

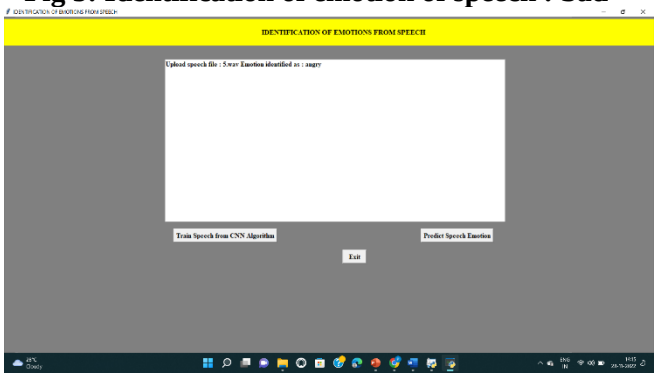

Fig 6: Identification of emotion of speech : Angry

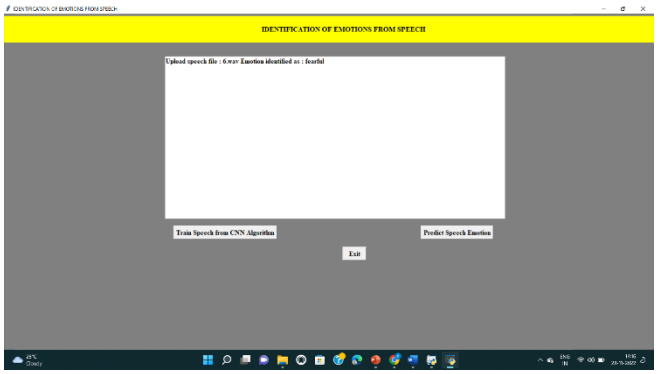

Fig 7: Identification of emotion of speech : Fearful

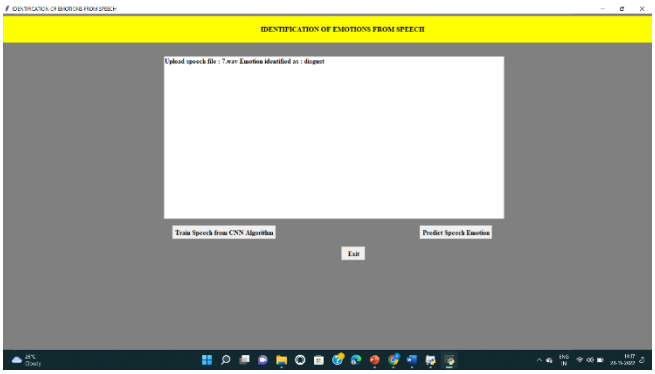

Fig 8: Identification of emotion of speech : Disgust

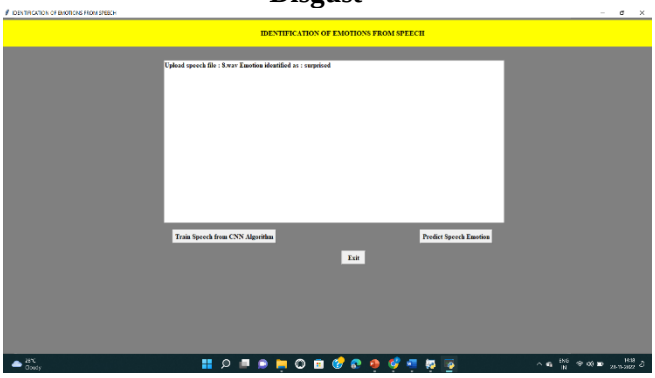

Fig 9: Identification of emotion of speech : Surprised

VII.CONCLUSION

Emotion recognition will greatly improve devices such as cars, phones, televisions, office equipment and even home appliances and systems by implementing new features and interfaces that are more intuitive and can automatically adapt to user needs. provide opportunities. Big companies are using Affective Computing because they are desperate to know if their products, services, and marketing strategies meet their customers' needs and preferences. Such technology can be researched and developed. A subject's emotional expression can be used to determine possible antisocial motives. Emotion modeling will play an interesting role in the next generation of interactive human-machine systems.

VII.REFERENCES

[1] Chen,J., Chen,Z., Chi,Z., & Fu,H.(2014, August). Facial expression recognition grounded on facial factors discovery and overeater features. In International Workshops on Electrical and Computer Engineering Subfields(pp. 884- 888).
[2] Ahmed,F., Bari,H., & Hossain,E.(2014). Personindependent facial expression recognition grounded on emulsion original double pattern(CLBP). Int. ArabJ. Inf.Technol., 11(2), 195- 203.
[3] Happy,S.L., George,A., & Routray,A.(2012, December). A real time facial expression bracket system using Original double Patterns. In Intelligent Human Computer Interaction(IHCI), 2012 4th transnational Conference on(pp. 1- 5). IEEE.
[4] Sokolova,M., Japkowicz,N., & Szpakowicz,S.(2006, December). Beyond delicacy, F- score and ROC a family of discriminant measures for performance evaluation. In Australasian Joint Conference on Artificial Intelligence(pp. 1015- 1021). Springer BerlinHeidelberg.
[5] Bhatt,M., Drashti,H., Rathod,M., Kirit,R., Agravat,M., & Shardul,J.(2014). A Studyof Local Binary Pattern Method for Facial Expression Discovery. arXiv preprintar Xiv 1405.6130.
[6] Zamil, Adib AshfaqA., etal." Emotion Discovery from Speech Signals using Voting Medium on Classified Frames." 2019 International Conference on Robotics, Electrical and Signal Processing ways(ICREST). IEEE, 2019.
[7] Noroozi,F., Marjanovic,M., Njegus,A., Escalera,S., & Anbarjafari,G. Audio-visual emotion recognition in videotape clips. IEEE Deals on Affective Computing, 2017.
[8] Bettadapura,V.(2012). Face expression recognition and analysis the state of the art. arXiv preprint arXiv1203.6722.
[9] Zhang,S., Zhao,X., & Lei,B.(2012). Facial expression recognition grounded on original double patterns and original fisherman discriminant analysis. WSEAS Trans. gesture Process, 8(1), 21- 31.
[10] Chibelushi,C.C., & Bourel,F.(2003). Facial expression recognition A brief tutorial overview. CVonline On- Line florilegium of Computer Vision, 9.
[11] Shan,C., Gong,S., & McOwan,P.W.(2005, September). Robust facial expression recognition using original double patterns. In Image Processing, 2005. ICIP 2005. IEEE International Conference on(Vol. 2, pp. II- 370). IEEE.
[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] 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.