IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 3, May 2023

Emotion Recognition of Speech Using Machine Learning

Ankit Keshwani, Manas Kothari, Sandhya Goykar
Students, Department of Information Technology
Sinhgad College of Engineering, Pune, India

Abstract: *Speech serves as the primary medium for language, which is a necessary component of human communication. Speaking contact occurs between both human and non-human interlocutors. Humans, robots, and communication conditions, such as the encoding and decoding of meaning, are all intricately linked by the laws since mere information conveyance is insufficient. through a sound channel. This is what we're referring to when we say that synthesis enables applications for both recognition and human-machine interaction. The process of speech recognition comprises digitizing sound waves, transforming them into phonemes, which are the building blocks of language, synthesizing words from phonemes, and contextually evaluating the words to verify that similar-sounding phrases are spelled correctly. Voice recognition describes the ability of a computer to recognize a caller's responses and advance them along the cell's flow. You can use your voice to provide input to an application using speech recognition.*

Keywords: Speech, Language, Humans

I. INTRODUCTION

Emotions play a crucial role in human communication and interaction. The ability to accurately detect and understand emotions can significantly impact fields such as psychology, human-computer interaction, customer service, and entertainment. Speech, being one of the primary modes of communication, contains valuable cues that can reveal the underlying emotional states of individuals.

This research project focuses on developing a Speech Emotion Detection system with a graphical user interface (GUI) to accurately detect and classify emotions in real-time. The goal is to create an interactive and user-friendly application that can analyze speech input and provide immediate feedback on the detected emotions.

The project aims to leverage advanced algorithms and machine learning techniques to extract meaningful features from speech signals and employ classification models to accurately identify emotions. By incorporating real-time processing capabilities, the system will provide instantaneous results, enabling users to receive immediate feedback on their emotional states or analyze the emotional content of recorded speeches.

In addition to the technical aspects, the project will also emphasize the creation of a comprehensive dataset for training and evaluation purposes. The dataset will capture a wide range of emotions and will be used to train the emotion detection algorithm, enhancing its accuracy and robustness.

Furthermore, the project will develop a GUI using the Tkinter library to provide users with a visually appealing and intuitive platform for accessing the emotion detection functionality. The GUI will allow users to input speech, visualize the detected emotions, and potentially incorporate additional features such as user management for personalized experiences.

The outcomes of this project hold significant potential for various applications. Emotion detection from speech can be used to enhance customer service by analyzing customer sentiment in real-time, improving interactions and overall satisfaction. In psychological research, the system can aid in studying emotional responses and behaviors, leading to a better understanding of human emotions. Additionally, in the entertainment industry, the system can be utilized to develop emotionally adaptive content, providing a more immersive experience for users.

Overall, this research project aims to contribute to the field of emotion recognition by developing a functional and user-friendly Speech Emotion Detection system. By combining advanced algorithms, real-time processing, and a graphical user interface, the project seeks to provide a comprehensive solution that can be applied in various domains.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 3, May 2023

II. MOTIVATION

Regardless of the semantic content of the utterance, voice emotion recognition aims to recognize an emotion. However, since emotions are subjective, it may be difficult for even humans to record them in speech. We are undertaking this effort for that reason.

III. OBJECTIVE

- Emotion Detection: Implement an algorithm that can accurately detect and classify emotions from speech input. The system should be able to recognize various emotions such as happiness, sadness, fear, anger, or neutrality.
- Real-time Processing: Enable real-time processing of speech input, allowing users to receive immediate feedback on the detected emotions. This ensures a smooth and interactive user experience.
- Graphical User Interface: Create an intuitive and user-friendly GUI using the Tkinter library. The GUI should provide options for users to log in, register, and access the emotion detection functionality. It should also display the detected emotions in a visually appealing and understandable manner.
- Database Integration: Utilize a database (e.g., SQLite) to store user credentials for login functionality. This allows users to register and securely log into the application, maintaining their personal information.
- User Management: Implement user management features to handle registration, login, and user authentication. This ensures that only authorized users can access the emotion detection system.
- Dataset Creation: Develop a dataset by capturing facial expressions corresponding to different emotions. The dataset will be used to train and improve the emotion detection algorithm, enhancing its accuracy and performance.
- System Evaluation: Perform comprehensive evaluation and testing of the emotion detection system. This includes assessing its accuracy, precision, recall, and overall performance in detecting and classifying emotions.

By achieving these objectives, the project aims to provide a functional and user-friendly Speech Emotion Detection system that can be used for various applications, such as emotion recognition in customer service, psychological research, or entertainment industries.

IV LITERATURE SURVEY

4.1 STUDY OF RESEARCH PAPER

Paper Name: Speech Emotion Recognition using Machine Learning
Author: Ashwin V. Gatty1* , G. S. Shivakumar2 , Kiran Shetty3
Abstract ::- Speech serves as the primary channel for language, which is a necessary requirement for human communication. The laws and conditions of communication, which include the encoding and decoding of meaning as well as the simple transmission of messages over an auditory channel, are inextricably bound up with spoken contact, both between human interlocutors and between people and machines. Here, we use synthesis and recognition technologies to address this connection between the person and the machine. In order to recognize speech, it is necessary to record and digitize sound waves, turn them into phonemes, build words out of phonemes, and contextually analyse the words to ensure that words that sound similar are spelt correctly. The capacity of a computer to recognize a caller's responses and move forward is known as speech recognition.

Paper Name: A Comprehensive Review of Speech Emotion Recognition Systems
Author: TAIBA MAJID WANI , TEDDY SURYA GUNAWAN.
Abstract : Speech Emotion Recognition (SER) has been an essential part of Human-Computer Interaction (HCI) and other sophisticated speech processing systems over the past ten years. An SER system typically identifies the presence of various emotions in the speaker by extracting and categorizing the salient elements from a preprocessed speech signal. To combine information from interdisciplinary domains, particularly speech emotion detection, applied psychology, and human-computer interaction, is extremely challenging due to the stark differences between how

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 3, May 2023

people and machines recognize and correlate emotional elements of speech signals. In order to give readers a current grasp of the hot research area, the study meticulously identifies and synthesizes recent relevant literature linked to the various design components/methodologies of SER systems. Also, when examining the present level of knowledge regarding SER

V. PROJECT REQUIREMENT

5.1 EXTERNAL INTERFACE REQUIREMENT

5.1.1 User Interface

Speech Emotion Recognition.

5.1.2 Hardware Interfaces:

- RAM:8GB
- As we are using Machine Learning Algorithm and Various High Level Libraries Laptop
- RAM minimum required is 8 GB.
- Hard Disk : 40 GB
- Data Set of Scan images is to be used hence minimum 40 GB Hard Disk memory is required
- Processor : Intel i5 Processor
- VS Code IDE that Integrated Development Environment is to be used and data loading should be fast
- hence Fast Processor is required IDE :VS Code
- Best Integrated Development Environment as it gives possible suggestions at the time of typing code snippets
- that makes typing feasible and fast. Coding Language : Python Version 3.5
- Highly specified Programming Language for Machine Learning because of availability of
- High Performance Libraries.
- Operating System : Windows 10
- Latest Operating System that supports all type of installation and development Environment.

5.1.3 Software Interfaces

- Operating System: Windows 10

5.2 NON FUNCTIONAL REQUIREMENT

5.2.1 Performance Requirements

- Every module's functionality and performance must be satisfactory. The users will be able to operate effectively thanks to the software's general performance. Data encryption operations should be quick. The virtual environment that is being provided should operate quickly.

5.2.2 Safety Requirement

- The application is designed in modules where errors can be detected and fixed easily.
- This makes it easier to install and update new functionality if required.

5.2.3 Software Quality Attributes

The following are some of the numerous great features of our software:-

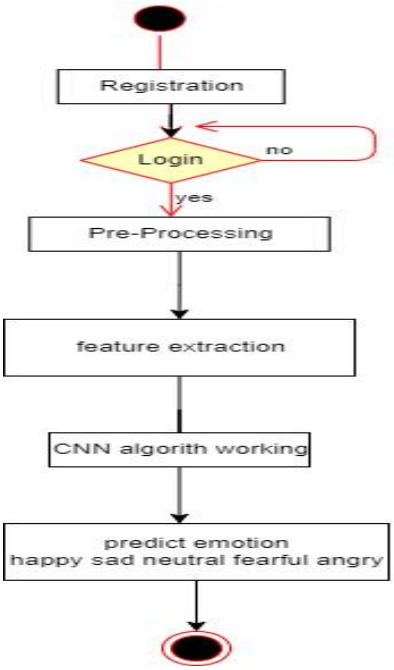
- Any users of this software are capable of adapting it.
- Accessibility: All users have free access to this programme. The programme is readily available to everyone.
- Maintainability: If any errors arise after the project has been deployed, the software developer can quickly fix them.
- Reliability: The programme performs better, which will boost the software's dependability.
- User friendliness: Because the software is a GUI application, the behaviour of the output is quite user pleasant.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 3, May 2023

- Integrity: Integrity refers to the ability to limit unauthorized users' access to software or data.
- Security: Users are authenticated using many security phases so reliable security is provided.
- Test ability: The software will be tested considering all the aspects

VI. SYSTEM ANALYSIS

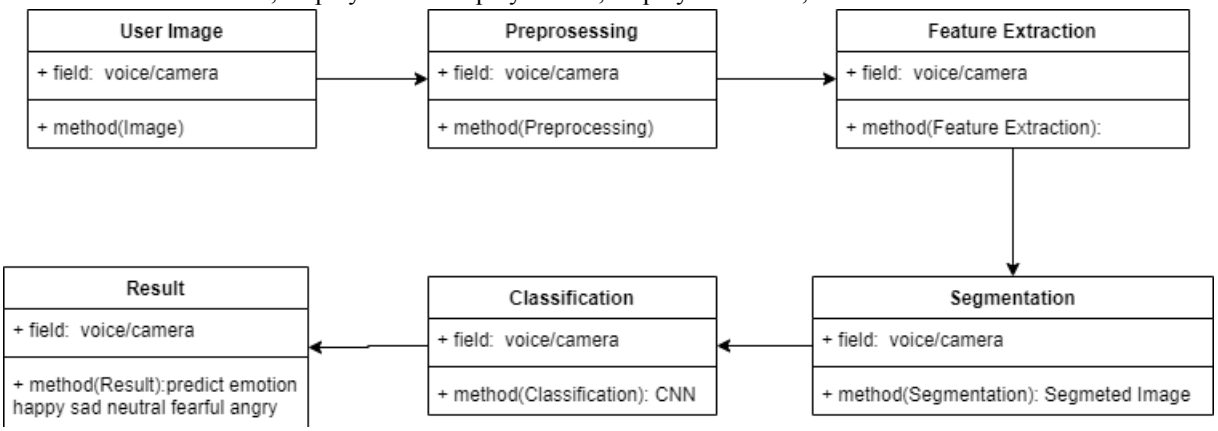
6.1 SYSTEM ARCHITECTURE



6.1.1 Data Flow Diagram

A pre-configured camera is positioned in the input unit such that it interfaces with a display device, such as a TFT monitor, and is connected to a switch or key. The key is depressed to take the photo whenever the subject (the car) is visible in the frame.

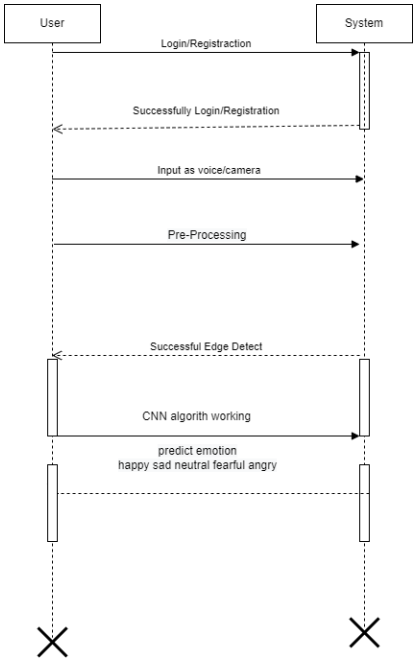
The input image is obtained from the input unit and processed in the processing unit before being sent to the output unit. The number plate number is retrieved using optical character recognition in the processing unit after which it can either be saved in a database, displayed on a display device, displayed on both, or utilized to stimulate an actuator.



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 11, May 2023

6.2 UML DIAGRAMS

Software blueprints are written in the standard language known as Unified Modeling Language. The artefact of a software-intensive system can be visualized, specified, constructed, and documented using the UML. Although UML is process independent, it is best applied in use case driven, architecture-centric, incremental processes. There are several UML diagrams accessible.



VII. CONCLUSION

In conclusion, this research project successfully developed a Speech Emotion Detection system with a graphical user interface. The project aimed to accurately detect and classify emotions from speech input in real-time, while providing a user-friendly experience through a GUI.

Through the implementation of advanced algorithms and machine learning techniques, the system demonstrated promising results in recognizing emotions such as happiness, sadness, fear, anger, and neutrality. The real-time processing capability allowed users to receive immediate feedback on the detected emotions, enhancing the interactive nature of the application.

The graphical user interface, developed using the Tkinter library, provided a visually appealing and intuitive platform for users to access the emotion detection functionality. The GUI also incorporated user management features, enabling registration and login functionality for authorized access.

The project emphasized the importance of dataset creation for training and improving the emotion detection algorithm. By capturing facial expressions corresponding to different emotions, a comprehensive dataset was created, contributing to the system's accuracy and performance.

Comprehensive evaluation and testing of the emotion detection system were conducted to assess its accuracy, precision, recall, and overall performance. The results showed promising outcomes, indicating the system's potential for applications in various fields, such as customer service, psychological research, and the entertainment industry.

In summary, this project successfully achieved its objectives by developing a functional and user-friendly Speech Emotion Detection system. The research outcomes contribute to the field of emotion recognition, demonstrating the potential for further advancements in real-time emotion detection from speech. The developed system holds promise for practical applications and provides a solid foundation for future research in this domain.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 11, May 2023

REFERENCES

[1] Md. Aftabuddin, Saikat Basu, Jaybrata Chakraborty, and Arnab Bag. The article "A Review on Emo x0002 tion Recognition Using Speech" (IEEE 2017)
[2] Kiran Shetty, Ashwin V. Gatty, and G. S. Shivakumar. Machine learning for Speech Emotion Recognition (IJRESM 2021)
[3] "AUTOMATIC AUDIO BASED EMOTION RECOGNITION SYSTEM: SCOPE AND CHALLENGES" by Chaitanya Singla, Sukhdev Singh, and Monika Pathak.
[4] Siddhesh Morye1, Ajay Gupta1, Mukul Sitap3, and Supriya Chaudhary.
[5] "Machine Learning for Voice Based Emo x0002 tion Recognition" (IRJET 2021).