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 DETECTION USING VOICE RECOGNITION AND FACIAL EXPRESSIONS

K. HARSHVARDHAN[1], S.RAJESH[2], M.YASWANTH[3], G.BINDHU MADHURI[4], M.GOWREESWARI[5]
A.S.C. TEJASWINI KONE[2]

1.IV YEAR B.Tech STUDENTS
Visakha Institute of Engineering & Technology
Computer Science and Engineering
Visakhapatnam, Andhra Pradesh
2.PROFESSOR & GUIDE
Visakha Institute of Engineering & Technology
HOD of Computer Science and Engineering
Visakhapatnam, Andhra Pradesh

ABSTRACT: *Using Natural Language Processing(NLP) our application detects complex and nuanced human emotions with the help of voice assistant and facial expressions provided by user and we identify the emotion class that the user belongs to at that particular instance .Facial Expression Detector can be used to know whether a person is sad, happy, angry and so on only through his/her face and it uses your Web Camera and then identifies your expression in Real Time. We develop an artificial intelligence voice-controlled emotion detector which gathers input from tone of human voice and detect emotion just the way humans express. Face expression recognition and voice assistants aid in human emotion detection by providing real-time insights into emotions. Our Project is an assortment of Natural Language Processing and Recommender System so as to sustain the user's emotions using our featured content and replenish the user to a better phase.*

1.1 INTRODUCTION

Emotion detection plays a significant role in human life by influencing various aspects of our well-being and interactions. It plays a key role in the field of Artificial Intelligence(AI). In real-time, the voice assistant can use the microphone to capture the user's voice. And simultaneously the device's camera can capture the user's facial expressions using computer vision. These data streams are then analysed by the emotion detection application.

In general , human emotions can be analysed using various methods and technologies including facial expression analysis, voice analysis, body language recognition, physiological measures, behavioural observations. The interaction between human beings and computers will be more natural if computers are able to perceive and respond to human non-verbal communication such as emotions and voice assistant.

Facial emotion recognition is the technology that analyses facial expressions from both statistic images and videos in order to reveal information on one's emotional state and the voice analyser system can identify the emotion of human by tone of human voice.

Although several approaches have been proposed to recognize human emotions based on facial expressions or speech, relatively limited work has been done to fuse these two, and other modalities to improve the accuracy and robustness of the emotion recognition system. This project analyses the strengths and the limitations of systems based only on facial expressions or acoustic information.
It discusses two approaches used to fuse these two modalities, i.e., decision level and feature level integration. Using a database recorded from a model, four emotions were classified: sadness, anger, happiness, and neutral state. By the use of markers on face, detailed facial motions were captured with motion capture, in conjunction with simultaneous speech recordings.

1.2 EMOTION DETECTION USING VOICE RECOGNITION AND FACIAL EXPRESSIONS.

Human variability in the recognition of emotions is quite significant. It is essential to note that when delving into the realm of automated emotion recognition, multiple references exist for determining the "ground truth" of an emotion – that is, what constitutes the true representation of an individual's emotional state.

Let's consider the instance of assessing Tom's emotions. One such reference is based on "how most individuals would perceive Tom's emotions." In such cases, the 'truth' might not necessarily align with Tom's genuine feelings but rather with the general consensus on how he appears to feel. For example, Tom might be experiencing sadness but chooses to project a cheerful appearance, prompting most people to interpret his emotional state as happiness.

When an automated method attains results on par with a collective of human observers, it can be deemed accurate, despite not necessarily reflecting Tom's genuine emotions. An alternative "truth" source involves directly consulting Tom about his innermost emotions. This approach is effective when Tom possesses a keen self-awareness of his emotional state, is willing to share his feelings, and can be illuminating them with precision using words or numeric representations. Nonetheless, it's essential to acknowledge that some individuals may be emotionless, lacking a profound understanding of their emotional landscape, or struggling to convey it accurately through language or numerical expressions.

In general, uncovering the authentic emotion at play demands a dedicated effort, can fluctuate based on the chosen criteria, and tends to retain an element of uncertainty.

1.3 MOTIVATION

In our everyday lives, we often lack constant companions who can fully comprehend our emotions and provide motivation. Therefore, we offer a personal space for our users to express their feelings, enabling us to offer optimal support through technology as a 24/7 intelligent system with an easier accessible voice recognition and facial expression recognition methods.

2.LITERATURE SURVEY

2.1 PROBLEM DEFINITION

Emotional fluctuations are a natural part of the human experience. Relying solely on Emotion Detection is insufficient for helping individuals maintain emotional equilibrium; what we truly require is a skilled and intelligent healer to manage one's emotions effectively.

Emotion detection using voice recognition faces challenges due to the ambiguity of vocal cues, cultural and individual differences, the masking of true emotions, and the lack of a universal emotional fingerprint. It is also limited by the need for extensive training data and ethical concerns regarding privacy and consent.

Emotion detection using facial expressions can be problematic due to cultural variations, individual differences, and the ability of some individuals to mask their true emotions. It also relies on visual cues, making it less effective in situations where facial expressions are not visible, and it raises concerns about privacy and consent in surveillance applications.

The challenge in emotion recognition using facial expressions and voice assistance lies in effectively integrating and interpreting two distinct modalities. Ambiguities, individual differences, and contextual variations can make it difficult to accurately assess emotions. Privacy concerns, data requirements, and ethical considerations also pose significant hurdles, especially in real-time applications and cross-cultural contexts.

2.2 UTILISING NLP

- By utilising Natural Language Processing, we can promptly distinguish an individual's emotional state at any given moment and classify it into specific emotion categories through the application of diverse classification algorithms.
- Emotion Detection primarily provides insights into the dynamics of emotions and assists in understanding an individual's mental state.
- It's important to recognize that the current system, based on emotion detection, is not designed to rectify emotional imbalances or facilitate an individual's healing process. Its primary function is to acknowledge and assess detected emotions.
- NLP analyses spoken language to identify emotional cues, while facial recognition uses computer vision to assess emotions based on facial expressions. These technologies employ machine learning algorithms to classify emotions into predefined categories, such as happiness, sadness, anger, etc.
- The systems work by comparing the input data, whether text or facial images, to a database of training examples. Based on patterns and characteristics in the input data, the system assigns it to one or more emotion categories. While these systems have made significant advancements, they still face challenges related to context, individual differences, and privacy concerns.

2.3 SOFTWARE ENVIRONMENT

INTRODUCTION TO PYTHON

Python is a versatile, object-oriented, high-level programming language known for its dynamic semantics. Its rich set of built-in data structures, coupled with dynamic typing and binding, make it exceptionally suitable for swift application development. Additionally, it serves as an ideal scripting or integration language to link various pre-existing components. Python's clear and easy-to-learn syntax prioritizes readability, thus minimizing the expenses associated with program maintenance. The language also supports modules and packages, fostering program modularity and code reusability. Furthermore, the Python interpreter, along with its extensive standard library, is available in both source and binary formats, free of charge, for all major platforms, and can be freely distributed.

Python, is developed by Guido van Rossum in 1991, stands as a programming language renowned for its effectiveness in statistical analysis, data manipulation, and machine learning. With Python, users can effortlessly craft objects, functions, and packages, offering a wide range of flexibility. This language knows no bounds and finds utility across diverse domains. It thrives on platform independence, seamlessly working on various operating systems. What's more, Python is an open-source tool, ensuring cost-effective accessibility, enabling its installation in any organization without the need for costly licenses.

Features of python

- Easy to understand
- Very Flexible
- Scalability
- Readability
- Robust Library, etc.,

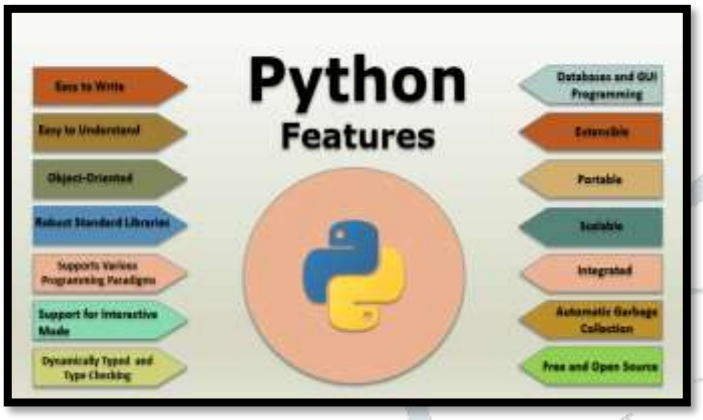


Fig 2.3 Features of Python

2.3.2 Installation of python:

- Download python from

  https://www.python.org/downloads/release/python-372/

- Click on Download python for Windows" > "base" > "Download python3.7.2 for Windows")
- Install python. Leave all default settings in the installation options.
- Set the path.

2.3.3 Installing Packages

- Packages can be installed with the **pip install package_name** function in python.
- To import any package in the program we should use **import package**_name

2.3.4 Libraries of python:

1. tkinter

Python offers multiple options for developing GUI (Graphical User Interface). Out of all the GUI methods, tkinter is the most commonly used method. It is a standard python interface to the Tk GUI toolkit shipped in python. Python with tkinter is the fastest and easiest way to create GUI applications

How can we use tkinter:

Simply by importing the tkinter module, you can access its functionalities.

How to create a tkinter app?

Importing the module - tkinter

Create the main window (container)

Add any number of widgets to the main window Apply the eventtrigger on widgets

2.Numpy

NumPy can be used to perform a wide variety of mathematical operations on arrays. It adds powerful data structures to Python that guarantee efficientcalculations with arrays and matrices and it supplies an enormous library of high-level mathematical functions that operate on these arrays and matrices.

3.Pandas

Pandas is an open-source, BSD-licensed Python library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language. Python with Pandas is used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc.

Features of Pandas:

- **Efficient DataFrame**: A fast and efficient DataFrame object with support for both default and customized indexing.
- **Data Loading Tools**: Tools to seamlessly load data from various file formats into in-memory data structures.
- **Missing Data Handling**: Built-in capabilities for data alignment and efficient handling of missing data.
- **Data Reshaping**:

  Functions for reshaping and pivoting datasets.

- **Label-Based Selection**:

  Intuitive label-based slicing, indexing, and subsetting for large datasets.

- **Column Manipulation**:

  Ability to easily delete or insert columns in data structures.

- **Grouping and Aggregation**:

  Built-in functions for grouping data, enabling aggregation and transformations.

- **High-Performance Merging**:

  High-performance merging and joining of data from different sources.

- **Time Series Support**:

  Comprehensive Time Series functionality for working with time-based data.

2.4 Hardware and Software Requirements:

Hardware Requirements

System : pentium 4

Hard disk : 120GB

Monitor : 14 inches

Input Devices : keyboard , mouse

RAM : 8GB

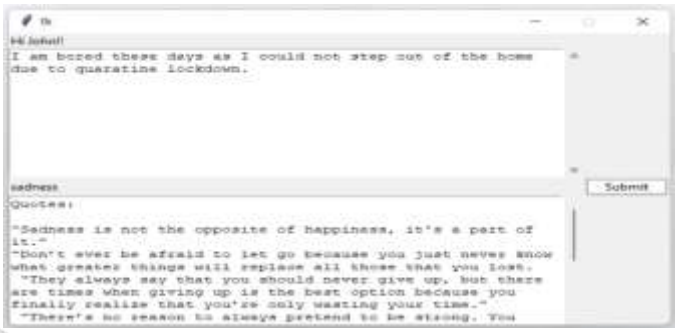
Software Requirements

Operating System : Windows 11Coding

Language : Python

3.EXISTING SYSTEM

- We can detect the emotion of a person at any instance using Natural Language Processing and identify the emotion class using various classification algorithms.
- Emotion Detection can only help us to know how emotions fluctuate and to identify the state of mind of a person .
- We cannot resolve the imbalance in emotions and heal a person by responding to the detected emotions using the existing system .
- By utilising Natural Language Processing, we can promptly distinguish an individual's emotional state at any given moment and classify it into specific emotion categories through the application of diverse classification algorithms.
- Emotion Detection primarily provides insights into the dynamics of emotions and assists in understanding an individual's mental state.
- It's important to recognize that the current system, based on emotion detection, is not designed to rectify emotional imbalances or facilitate an individual's healing process. Its primary function is to acknowledge and assess detected emotions.
- NLP analyses spoken language to identify emotional cues, while facial recognition uses computer vision to assess emotions based on facial expressions. These technologies employ machine learning algorithms to classify emotions into predefined categories, such as happiness, sadness, anger, etc.
- The systems work by comparing the input data, whether text or facial images, to a database of training examples. Based on patterns and characteristics in the input data, the system assigns it to one or more emotion categories. While these systems have made significant advancements, they still face challenges related to context, individual differences, and privacy concerns.



4.PROPOSED SYSTEMS

4.1 PROPOSED SYSTEMS

The proposed system for detecting emotions within voice recognition and facial expressions is an integrated solution that combines advanced audio and image processing techniques with machine learning models. It captures voice data through microphones and facial expressions through cameras, preprocesses the data, and extracts relevant features. Machine learning models are employed to classify emotions based on these features, with real-time processing capabilities. The system provides users with immediate feedback on emotional states and can be applied in various domains, including mental health monitoring, customer service, and human-computer interaction, offering valuable insights into individuals' emotional well-being and enhancing user experiences. It's designed with a focus on privacy, ethical considerations, and adaptability to accommodate diverse expressions and voices.

4.2 FUNCTIONAL REQUIREMENTS

The functional requirements for a system detecting emotions within voice recognition and facial expressions would include the ability to capture and preprocess audio and facial data, employ machine learning models for emotion classification, perform real-time analysis, integrate results, offer user-friendly visualization, ensure adaptability to individual differences, prioritize security and privacy, and provide valuable feedback and reporting. These requirements are crucial for developing a comprehensive and reliable emotion detection system that caters to various applications, from mental health support to user experience enhancement.

5.SYSTEM DESIGN

5.1 SYSTEM DESIGN

To create an intelligent program that discerns emotions and suggests tailored content to enhance emotional well-being. Our recommender system identifies content in accordance with an individual's current emotional state, offering healing and the opportunity to rediscover their most positive and fulfilling life moments.

The system design for emotion detection using voice recognition and facial expressions typically involves the following components:

DATA COLLECTION:

- Gather a dataset of voice recordings and facial expression images labelled with corresponding emotions (e.g., happy, sad, angry).

PREPROCESSING:

- Clean and preprocess the audio data by removing noise, normalizing volume, and extracting relevant features (e.g., MFCCs for voice).
- Preprocess facial images, including resizing, normalization, and extracting facial landmarks.

VOICE EMOTION RECOGNITION:

- Use machine learning models (e.g., deep neural networks) to analyze voice features and classify emotions based on the voice data.

FACIAL EMOTION RECOGNITION:

- Employ computer vision techniques and deep learning models (e.g., convolutional neural networks) to detect and classify emotions from facial expressions.

FUSION:

- Combine the results from voice and facial emotion recognition for more accurate emotion classification. This can be done using various fusion techniques, such as feature-level fusion or decision-level fusion.

USER INTERFACE:

- Create a user-friendly interface for inputting voice and facial data for real-time or batch emotion detection.

POST-PROCESSING:

- Implement post-processing algorithms to smooth out emotion predictions and enhance accuracy.

EVALUATION:

- Assess the system's performance using metrics like accuracy, precision, recall, and F1-score.

INTEGRATION:

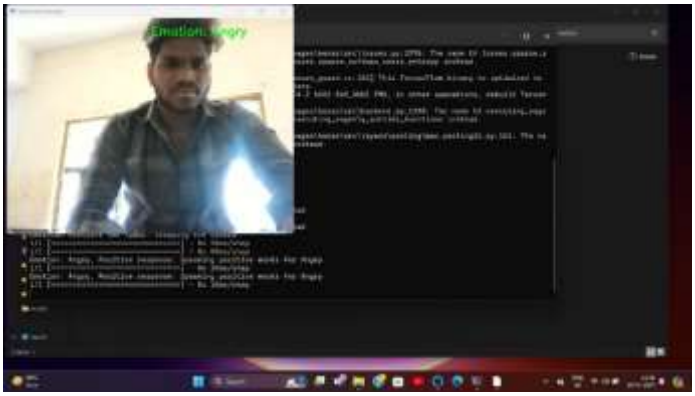
- Integrate the system into applications or devices where emotion detection can be beneficial, such as customer service, healthcare, or entertainment.

CONTINUOUS IMPROVEMENT:

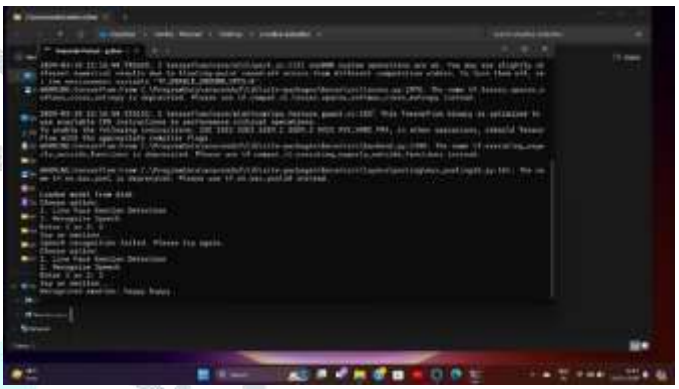
- Continuously update and fine-tune the system with new data to improve accuracy and adapt to evolving user needs.

It's important to note that the success of an emotion detection system depends on the quality of the dataset, the choice of machine learning models, and the accuracy of feature extraction and preprocessing techniques. Additionally, ethical considerations related to privacy and consent are crucial when working with voice and facial data.

6.RESULTS



FACE RECOGNIZATION OUTPUT



VOICE RECOGNIZATION OUTPUT

7.CONCLUSION

In conclusion, the merging of facial expressions and voice recognition for emotion detection presents a promising avenue for various applications, spanning psychology, human-computer interaction, and artificial intelligence. Through this project, we have explored the significant advancements and potential of utilizing both modalities to accurately understand and interpret human emotions.

8.REFERENCES

[1] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, “Emotion recognition in human-computer interaction,” in IEEE Signal Processing Magazine, vol. 18(1), Jan. 2001, pp. 32-80, doi: 10.1109/79.911197

[2] Parrott, W.G, “Emotions in Social Psychology,” in Psychology Press, Philadelphia 2001

[3] C. Maaoui, A. Pruski, and F. Abdat, “Emotion recognition for human machine communication”, Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 08), IEEE Computer Society, Sep. 2008, pp. 1210-1215, doi: 10.1109/IROS.2008.4650870

[4] Chun-Chieh Liu, Ting-Hao Yang, Chang-Tai Hsieh, Von-Wun Soo, ”Towards Text-based Emotion Detection: A Survey and Possible Improvements ”,in International Conference on Information Management and Engineering,2009.

[5] N.Fragopanagos, J.G. Taylor, “Emotion recognition in human–computer interaction”, Department of Mathematics, King‟s College, Strand, London WC2 R2LS, UK Neural Networks 18 (2005) 389–405 march 2005.

[6] Sherry Turkle, Alone Together: Why We Expect More from Technology and Less from Each Other, October 2, 2012.

[7] Breazeal, C., Regulating human-robot interaction using 'emotions', 'drives' and facial expressions. Presented at Autonomous Agents 1998 workshop 'Agents in Interaction-Acquiring Competence through Imitation', Minneapolis/St Paul, May. 1998.

[8] Breazeal, C. and Aryananda, L. Recognition of affective communicative intent in robot-directed speech. Autonomous Robots 12 1, 2002. pp. 83104.

[9] Mohammed E. H., Courgeon M., Martin J.C., Mutlu B., Picard R, MACH: My Automated Conversation coacH, in Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing, pp. 697-706, 2013

[10] Healey, J. and R.W. Picard,"Detecting Stress During Real-World Driving Tasks Using Physiological Sensors," IEEE Trans. on Intelligent Transportation Systems, Volume 6, No. 2, pp. 156166, June 2005.

[11] Global Personal Robots Market Size, Share, Development, Growth and Demand Forecast to 2022 - Industry Insights by Type (Cleaning Robot, Entertainment & Toy Robot, Education Robot, Handicap Assistance Robot, Companion Robot, Personal Transportation Robot, Security Robot, and Others) published by P&S Market Research, Feb 2017. https://www.psmarketresearch.com/marketanalysis/personal-robot-market.

[12] Wiseguy Reports, Global Pet Companion Robots Market Outlook 2024: Global Opportunity and Demand Analysis, Market Forecast, 2016-2024, 12 September, 2017.

[13] Zuzanna Wojcik, Robotics&AI, May 2016.

[14] Garay, Nestor; Idoia Cearreta; Juan Miguel Lpez; Inmaculada Fajardo "Assistive Technology and Affective Mediation". Human Technology. 2 (1): 5583, April 2006.