ACKNOWLEDGEMENT

The Major Project work carried out by our team in the Department of Computer Science and Engineering, Malla Reddy College of Engineering for Women, Hyderabad. ***This work is original and has not been submitted in part or full for any degree or diploma of any other university.***

We wish to acknowledge our sincere thanks to our project guide **Ms. K.Neha Nandini**, Assistant Professor, Computer Science & Engineering for formulation of the problem, analysis, guidance and her continuous supervision during the course of work.

We acknowledge our sincere thanks to **Dr. Vaka Murali Mohan,** Director, MRCEW & Professor of computer science engineering department. **Dr. Kanaka Durga Returi,** Principal & professor of computer science engineering department and **Mrs. K. Sheetal**, Professor & Head of Department and all our facilities of CSE Department for their kind cooperation in making this major project work a success.

We extend our gratitude to **Sri. Ch. Malla Reddy**, Founder Chairman and **Sri. Ch. Mahender Reddy,** Secretary for their kind cooperation in providing the infrastructure for completion of our Major Project.

We acknowledge our special thanks to the entire teaching faculty and non-teaching staff members of the Computer Science & Engineering Department for their support in making this project work a success.

GORLA VARSHITHA REDDY	**19RG1A0574**_____________
KOTHAPELLI SAIPREETHI	**19RG1A0591**_____________
NALLA RIKHITHA REDDY	**19RG1A05A3**_____________
PATHLAWATH DEEPTHI	**19RG1A05A6**_____________

INDEX

ABSTRACT

In this paper, we propose a bimodal emotion recognition system using the combination of facial expressions and speech signals. The models obtained from a bimodal corpus with six acted emotions and ten subjects were trained and tested with different classifiers, such as Support Vector Machine, Naive Bayes and K-Nearest Neighbour. In order to fuse visual and acoustic information, two different approaches were implemented: feature level fusion and match score level fusion. Comparative studies reveal that the performance and the robustness of emotion recognition systems can be improved by the use of fusion-based techniques. Further, the fusion performed at the feature level showed better results than the one performed at the score level.

LIST OF FIGURES

CHAPTER-1: SYSTEM ANALYSIS

1.1 Existing System:

- It has become an important issue to identify the user emotional state. Based on psychological theory, it is widely accepted that six archetypal emotions can be identified: surprise, fear, disgust, anger, happiness and sadness.
- Facial motion and tone of the speech play a major role in expressing these emotions.
- Emotions can significantly change the message sense: sometimes it is not what was said that is the most important, but how it was said.
- The face tends to be most visible form of emotion communication, but it is also most easily controlled in response to different social situations when compared to the voice and other ways of expression.
- Further a brief review of existing emotion recognition systems is presented.

1.1.1 Disadvantages of Existing System

- The accuracy of emotion detection using speech recognition and facial expression analysis heavily release on the quality input and background noise and low-quality images.
- There is no standardization in emotion detection algorithms, which can lead to difference in results across the system and make it challenging to compare results from different studies.
- Less security.

1.2 Proposed System:

- One of the most important tasks is to find a proper choice of feature vectors. The first challenge was to select several indicators attributable to the emotional behaviour.
- In order to fulfil this, many features have been explored and further we present the used ones: Mel Frequency Cepstral Coefficients and the statistical moments for speech, respectively wavelet coefficients and the seven moments of Hu for image There is a variety of temporal and spectral features that can be extracted from human speech.
- Prosodic features have been known to be an important indicator of emotional states, and thus they have been used in the design of many vocal emotion recognition systems.
- Several recent studies have also shown that there are variations across emotional states in the spectral features at the phoneme level (especially for vowel sounds) .
- Our purpose is to explore the spectral features by using the Mel-frequency cepstral coefficients (MFCCs).
- They have been widely employed in speech recognition also, because of superior performance when compared to other features.
- The Mel-frequency kepstrum is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear Mel scale of frequency

1.2.1 Advantages of Proposed System

- The use of speech recognition and facial expression analysis is relatively low cost compared to other approaches.
- Highly accurate in emotion detection system.
- More security.

1.3 INTRODUCTION

In human-computer interaction, emotional processes are inseparably connected to rational decisions; hence affective interaction has gained great attention. Therefore, it has become an important issue to identify the user emotional state. Based on psychological theory, it is widely accepted that six archetypal emotions can be identified: surprise, fear, disgust, anger, happiness and sadness. Facial motion and tone of the speech play a major role in expressing these emotions. Emotions can significantly change the message sense: sometimes it is not what was said that is the most important, but how it was said. The face tends to be most visible form of emotion communication, but it is also most easily controlled in response to different social situations when compared to the voice and other ways of expression. Further a brief review of existing emotion recognition systems is presented. Facial Expression Recognition Studies: Since 1970, Paul Ekman has performed extensive studies on human facial expressions. He found that facial expressions of emotion are not culturally determined, but universal to human culture and thus biological 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-assisted recognition of facial expressions started in 1990s. The features used are typically based on local spatial position of specific points and regions of the face (edge of the mouth, eyes, eyebrows). Mase (1991) was one of the first researchers who used image processing techniques to recognize facial expressions. With 11 windows manually located in the face, the muscle movements were extracted by the use of optical flow. K-nearest neighbour rule was employed for the classification task of four emotions, with an accuracy of 80%. Rosenblum (1996) and Otsuka (1997) also developed an optical flow region-based approach, by applying a Radial Basis Function Network and a Hidden Markov Model, respectively. Tian et al. (2000) explored AUs recognition by using permanent and transient facial features (lips, wrinkles). Geometrical models were used to locate their shapes and appearances. They achieved 96% accuracy with a Neural-Network-Based classifier. Cohen (2003, 2004) introduced the Bayesian Network classifiers in the static settings and a multi-level HMM classifier to automatically segment an arbitrary long sequence to the corresponding facial expressions . Vocal Emotion Recognition Studies: Following the long tradition of speech analysis, many efforts were taken to recognize affective states from vocal information. Starting in the 1930s, some important voice feature vectors have been chosen for research: fundamental frequency, time-energy distribution vector, MFCC, LPCC coefficients, etc. Williams and Stevens (1972) studied the spectrograms of real emotional speech and compared them with acted speech. They found similarities which suggest the use of acted data. A qualitative correlation between emotion and speech feature was presented by Murray and Arnott (1993). Petruchio (1998) compared human and machine recognition of emotions in speech and achieved similar rates for both. To exploit the dynamic variation along an utterance, MelFrequency Cepstral Coefficients were employed. Nwe (2001) achieved an average accuracy of 70% for six emotions acted by two speakers using 12 MFCC features as input to a discrete Hidden Markov Model. Busso (2004) also argued that statistics relating to MFCCs carry emotional information. Yu et. al. (2002) used Support Vector Machines as binary classifiers. On four distinct emotions, they achieved an accuracy of 73%. Lee (2002) tried to distinguish between negative and positive emotions, in call centre environment, using linear discrimination, k-NN and SVM classifiers achieving a maximum accuracy rate of 75%. Batliner (2003) studied a 4-class problem with elicited emotions in spontaneous speech [3]. Multimodal Emotion Recognition Studies: In order to improve the unimodal systems' recognition accuracy several studies attempted to exploit the advantage of using multimodal systems, especially by fusing audio-visual information. De Silva et al (2000) proposed a rule-based singular classification of audio-visual data recorded from two subjects into six emotions. From the audio data, they selected prosodic features, and from the video data, they chose the maximum distances and velocities between six specific facial points. Using decision-level fusion, a recognition rate of 72% was reported. A set of singular classification methods was proposed by Chen and Huang (2000), in which audio-visual data collected from five subjects was classified into the Ekman's six basic emotions [5]. In both studies, the performance of the system increased when both modalities were fused. A largescale audio-visual dataset was collected by Zeng et al. (2004), it containing five affective responses (confusion, interest, boredom and frustration) in addition to the six basic ones. They used the Naive Bayes classifier as the update rule, achieving an accuracy of almost 90%. There are only a few attempts to combine information from body movement and gestures. Kaliouby's (2005) model infers acted mental states from head movements and facial expressions. Gunes and Piccard (2006) fused at different levels facial expressions and body gestures. Another multimodal system based on facial expression, body gestures and speech were implemented by Castellano in 2007.

CHAPTER-2: LITERATURE SURVEY

1.“Multimodal emotion recognition from expressive faces, body gestures and speech”,

AUTHOR: G. Castellano, L. Kessous.

ABSTRACT: In this paper we present a multimodal approach for the recognition of eight emotions that integrates information from facial expressions, body movement and gestures and speech. We trained and tested a model with a Bayesian classifier, using a multimodal corpus with eight emotions and ten subjects. First individual classifiers were trained for each modality. Then data were fused at the feature level and the decision level. Fusing multimodal data increased very much the recognition rates in comparison with the unimodal systems: the multimodal approach gave an improvement of more than 10% with respect to the most successful unimodal system. Further, the fusion performed at the feature level showed better results than the one performed at the decision level.

2. “Multimodal Approaches for Emotion Recognition: A Survey”

AUTHOR: Sebe, I. Cohen, T. Gevers, T.S. Huang.

ABSTRACT: Recent technological advances have enabled human users to interact with computers in ways previously unimaginable. Beyond the confines of the keyboard and mouse, new modalities for human-computer interaction such as voice, gesture, and force-feedback are emerging. Despite important advances, one necessary ingredient for natural interaction is still missing-emotions. Emotions play an important role in human-to-human communication and interaction, allowing people to express themselves beyond the verbal domain. The ability to understand human emotions is desirable for the computer in several applications. This paper explores new ways of human-computer interaction that enable the computer to be more aware of the user's emotional and attentional expressions. We present the basic research in the field and the recent advances into the emotion recognition from facial, voice, and physiological signals, where the different modalities are treated independently. We then describe the challenging problem of multimodal emotion recognition and we advocate the use of probabilistic graphical models when fusing the different modalities. We also discuss the difficult issues of obtaining reliable affective data, obtaining ground truth for emotion recognition, and the use of unlabelled data.

3. "Bimodal emotion recognition using speech and physiological changes", Robust Speech Recognition and Understanding

AUTHOR: J. Kim.

ABSTRACT: In this paper, we treated all stages of emotion analysis, from data collection to classification using short-term observations, and evaluated several fusion methods as well as a hybrid decision scheme. We also compared the results from multimodal classification with the unimodal results. As in our earlier work (Kim et al. 2005) where we relied on longer observation phases and a different set of features, the best results were obtained by feature level fusion method in combination with feature selection stage. In this case, not only user dependent, but also user-independent emotion classification could be improved compared to the unimodal methods.

4. "Analysis of emotion recognition using facial expressions, speech and multimodal information"

AUTHOR: C. Busso, Z. Deng, et al.

ABSTRACT: 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. 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 paper analyses the strengths and the limitations of systems based only on facial expressions or acoustic information. It also discusses two approaches used to fuse these two modalities: decision level and feature level integration. Using a database recorded from an actress, four emotions were classified: sadness, anger, happiness, and neutral state. By the use of markers on her face, detailed facial motions were captured with motion capture, in conjunction with simultaneous speech recordings. The results reveal that the system based on facial expression gave better performance than the system based on just acoustic information for the emotions considered. Results also show the complementarily of the two modalities and that when these two modalities are fused, the performance and the robustness of the emotion recognition system improve measurably.

5. “Emotional expressions in audio visual human computer interaction “.

AUTHOR: Chen, L.S., Huang, T.S.

ABSTRACT: Visual and auditory modalities are two of the most commonly used media in interactions between humans. In the present paper, we describe a system to continuously monitor the user's voice and facial motions for recognizing emotional expressions. Such an ability is crucial for intelligent computers that take on a social role such as a tutor or a companion. We outline methods to extract audio and visual features useful for classifying emotions. Audio and visual information must be handled appropriately in single-modal and bimodal situations. We report audio-only and video-only emotion recognition on the same subjects, in person-dependent and person-independent fashions, and outline methods to handle bimodal recognition.

6. “The Nature of Statistical Learning Theory”

AUTHOR: Vatnik, V.

ABSTRACT: The aim of this book is to discuss the fundamental ideas which lie behind the statistical theory of learning and generalization. It considers learning from the general point of view of function estimation based on empirical data. Omitting proofs and technical details, the author concentrates on discussing the main results of learning theory and their connections to fundamental problems in statistics. These include: - the general setting of learning problems and the general model of minimizing the risk functional from empirical data - a comprehensive analysis of the empirical risk minimization principle and shows how this allows for the construction of necessary and sufficient conditions for consistency - non-asymptotic bounds for the risk achieved using the empirical risk minimization principle - principles for controlling the generalization ability of learning machines using small sample sizes - introducing a new type of universal learning machine that controls the generalization ability.

7. Facial Action Coding System: A Technique for Measurement of Facial Movement.

AUTHOR: Ekman, P., Friesen, W. V.

ABSTRACT: Facial expression is widely used to evaluate emotional impairment in neuropsychiatric disorders. Ekman and Friesen's Facial Action Coding System (FACS) encodes movements of individual facial muscles from distinct momentary changes in facial appearance. Unlike facial expression ratings based on categorization of expressions into prototypical emotions (happiness, sadness, anger, fear, disgust, etc.), FACS can encode ambiguous and subtle expressions, and therefore is potentially more suitable for analysing the small differences in facial affect. However, FACS rating requires extensive training, and is time consuming and subjective thus prone to bias. To overcome these limitations, we developed an automated FACS based on advanced computer science technology. The system automatically tracks faces in a video, extracts geometric and texture features, and produces temporal profiles of each facial muscle movement. These profiles are quantified to compute frequencies of single and combined Action Units (AUs) in videos, and they can facilitate a statistical study of large populations in disorders known to impact facial expression. We derived quantitative measures of flat and inappropriate facial affect automatically from temporal AU profiles. Applicability of the automated FACS was illustrated in a pilot study, by applying it to data of videos from eight schizophrenia patients and controls. We created temporal AU profiles that provided rich information on the dynamics of facial muscle movements for each subject. The quantitative measures of flatness and inappropriateness showed clear differences between patients and the controls, highlighting their potential in automatic and objective quantification of symptom severity.

8. "Emotion recognition based on phoneme classes"

AUTHOR: Lee, Chul Min et al.

ABSTRACT: Recognizing human emotions/attitudes from speech cues has gained increased attention recently. Most previous work has focused primarily on suprasegmental prosodic features calculated at the utterance level for modelling against details at the segmental phoneme level. Based on the hypothesis that different emotions have varying effects on the properties of the different speech sounds, this paper investigates the usefulness of phoneme-level modelling for the classification of emotional states from speech. Hidden Markov models (HMM) based on short-term spectral features are used for this purpose using data obtained from a recording of an actress' expressing 4 different emotional states - anger, happiness, neutral, and sadness. We designed and compared two sets of HMM classifiers: a generic set of "emotional speech" HMMs (one for each emotion) and a set of broad phonetic-class based HMMs for each emotion type considered. Five broad phonetic classes were used to explore the effect of emotional colouring on different phoneme classes, and it was found that spectral properties of vowel sounds were the best indicator of emotions in terms of the classification performance. The experiments also showed that the better performance can be obtained by using phoneme-class classifiers than generic "emotional" HMM classifier and classifiers based on global prosodic features. To see the complementary effect of the prosodic and spectral features, the two classifiers were combined at the decision level. The improvement was 0.55% in absolute (0.7% relatively) compared with the result from phoneme-class based HMM classifier.

9. "A wavelet tour of signal processing"

AUTHOR: S. Mallat.

ABSTRACT: Introduction to a Transient World. Fourier Kingdom. Discrete Revolution. Time Meets Frequency. Frames. Wavelet Zoom. Wavelet Bases. Wavelet Packet and Local Cosine Bases. An Approximation Tour. Estimations are Approximations. Transform Coding. Appendix A: Mathematical Complements. Appendix B: Software Toolboxes.

10. "Visual pattern recognition by moment invariants"

AUTHOR: M.K Hu.

ABSTRACT: In this paper a theory of two-dimensional moment invariants for planar geometric figures is presented. A fundamental theorem is established to relate such moment invariants to the well-known algebraic invariants. Complete systems of moment invariants under translation, similitude and orthogonal transformations are derived. Some moment invariants under general two-dimensional linear transformations are also included. Both theoretical formulation and practical models of visual pattern recognition based upon these moment invariants are discussed. A simple simulation program together with its performance are also presented. It is shown that recognition of geometrical patterns and alphabetical characters independently of position, size and orientation can be accomplished. It is also indicated that generalization is possible to include invariance with parallel projection.

11.“Development and validation of a facial expression database based on the dimensional and categorical model of emotions”

AUTHOR: T Fujimura, H Umemura.

ABSTRACT: The present study describes the development and validation of a facial expression database comprising five different horizontal face angles in dynamic and static presentations. The database includes twelve expression types portrayed by eight Japanese models. This database was inspired by the dimensional and categorical model of emotions: surprise, fear, sadness, anger with open mouth, anger with closed mouth, disgust with open mouth, disgust with closed mouth, excitement, happiness, relaxation, sleepiness, and neutral (static only). The expressions were validated using emotion classification and Affect Grid rating tasks [Russell, Weiss, &Mendelsohn,1989. Affect Grid: A single-item scale of pleasure and arousal. Journal of Personality and Social Psychology,57, 493–502]. The results indicate that most of the expressions were recognised as the intended emotions & could systematically represent affective valence and arousal. Furthermore, face angle and facial motion information influenced emotion classification and valence and arousal ratings.

12. “Emotion recognition in the noise applying large acoustic feature sets”

AUTHOR: Björn Schuller, Dejan Aric, Frank Wallhoff, Gerhard Rigoll.

ABSTRACT: Speech emotion recognition is considered mostly under ideal acoustic conditions: acted and elicited samples in studio quality are used besides sparse works on spontaneous field data. However, specific analysis of noise influence plays an important factor in speech processing and is practically not considered hereon, yet. We therefore discuss affect estimation under noise conditions herein. On 3 well-known public databases-DES, EMO-DB, and SUSAS-effects of post recording noise addition in diverse dB levels, and performance under noise conditions during signal capturing, are shown. To cope with this new challenge, we extend generation of functionals by extraction of a large 4k hi-level feature set out of more than 60 partially novel base contours. Such comprise among others intonation, intensity, formants, HNR, MFCC, and VOC19. Fast Information-Gain-Ratio filter-selection picks attributes according to noise conditions. Results are presented using Support Vector Machines as classifier.

13. “Multimodel face detection, head orientation and eye gaze tracking”

AUTHOR: Frank Wallhoff, Markus AblaBmeier, Gerhard Rigoll.

ABSTRACT: For several applications within the human-machine-interface domain a person's face plays a key role as an information source, such as the identification, the computation of the affective state or to predict the awareness of an user. Therefore, this paper presents a multi-modal approach for finding and tracking a face and estimating the head's gaze as well as the eyes' view direction. Throughout the paper several measurements relying on two different camera inputs are introduced which can be used to form a robust computation of the head orientation and the viewing direction of a person.

14. “Automatic Facial Feature Detection for Facial Expression Recognition”

AUTHOR: Taner Danis man, Loan Marius Bilasco, Nacim Ihaddadence, Chaabane Djerba

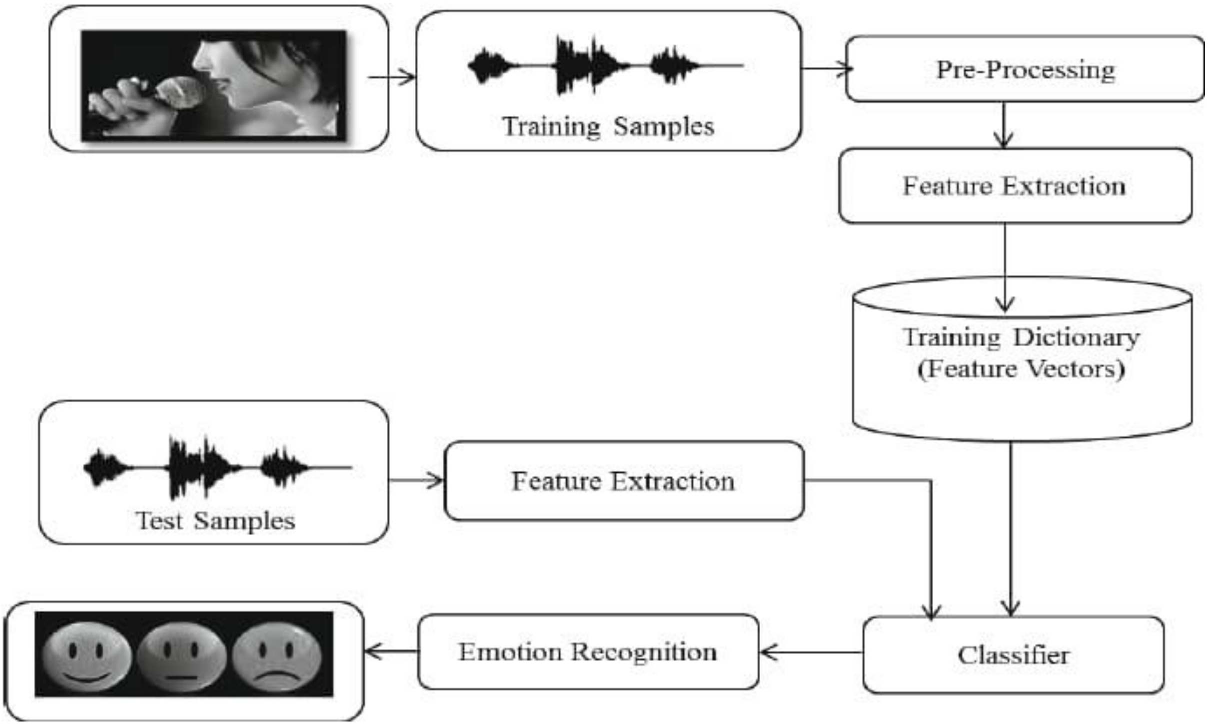
ABSTRACT: This paper presents a real-time automatic facial feature point detection method for facial expression recognition. The system is capable of detecting seven facial feature points (eyebrows, pupils, nose, and corners of mouth) in grayscale images extracted from a given video. Extracted feature points then used for facial expression recognition. Neutral, happiness and surprise emotions have been studied on the Bosphorus dataset and tested on FG-NET video dataset using OpenCV. We compared our results with previous studies on this dataset. Our experiments showed that proposed method has the advantage of locating facial feature points automatically and accurately in real-time.

15. “Real-Time Facial Expression Recognition for Natural Interaction”

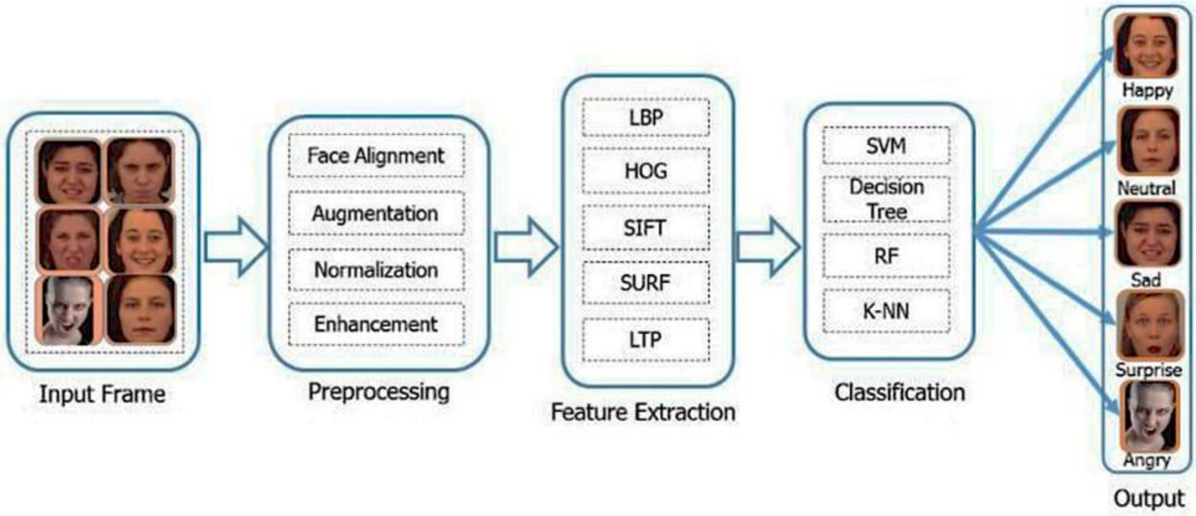
AUTHOR: Eva Cerezo, Isabelle Hupont, Cristina Manresa-Yee, Javier Varona, Sandra Baldassarri, Francisco J. Perales & Francisco J. Seron

ABSTRACT: The recognition of emotional information is a key step toward giving computers the ability to interact more naturally and intelligently with people. This paper presents a completely automated real-time system for facial expression’s recognition based on facial features’ tracking and a simple emotional classification method. Facial features’ tracking uses a standard webcam and requires no specific illumination or background conditions. Emotional classification is based on the variation of certain distances and angles from the neutral face and manages the six basic universal emotions of Ekman. The system has been integrated in a 3D engine for managing virtual characters, allowing the exploration of new forms of natural interaction.

3.1 System Architecture



3.1.1 Fig: Speech Emotion Recognition



3.1.2 Fig: Facial Emotion Expression

3.2 SYSTEM REQUIREMENTS:

3.2.1 HARDWARE REQUIREMENTS:

- System : Pentium IV 2.4 GHz.
- Hard Disk : 40 GB.
- Floppy Drive : 1.44 Mb.
- Monitor : 15 VGA Colour.
- Mouse : Logitech.
- Ram : 512 Mb.

3.2.2 SOFTWARE REQUIREMENTS:

- **Operating System:** Windows
- **Coding Language**: Python 3.7

3.3 UML DIAGRAMS

UML stands for Unified Modelling Language. UML is a standardized general-purpose modelling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group. The goal is for UML to become a common language for creating models of object-oriented computer software.

In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML. The Unified Modelling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modelling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modelling of large and complex systems. The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

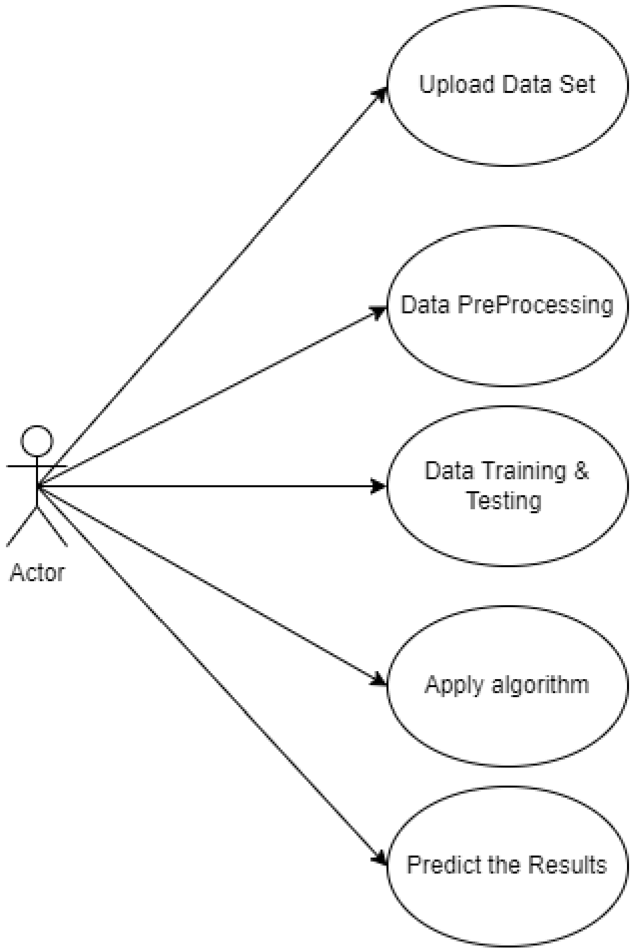
GOALS:

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modelling Language so that they can develop and exchange meaningful models.
2. Provide extendibility and specialization mechanisms to extend the core concepts.
3. Be independent of particular programming languages and development process.
4. Provide a formal basis for understanding the modelling language.
5. Encourage the growth of OO tools market.
6. Support higher level development concepts such as collaborations, frameworks, patterns and components.
7. Integrate best practices.

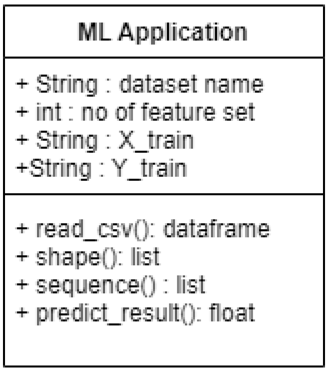
3.3.1 USE CASE DIAGRAM:

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.



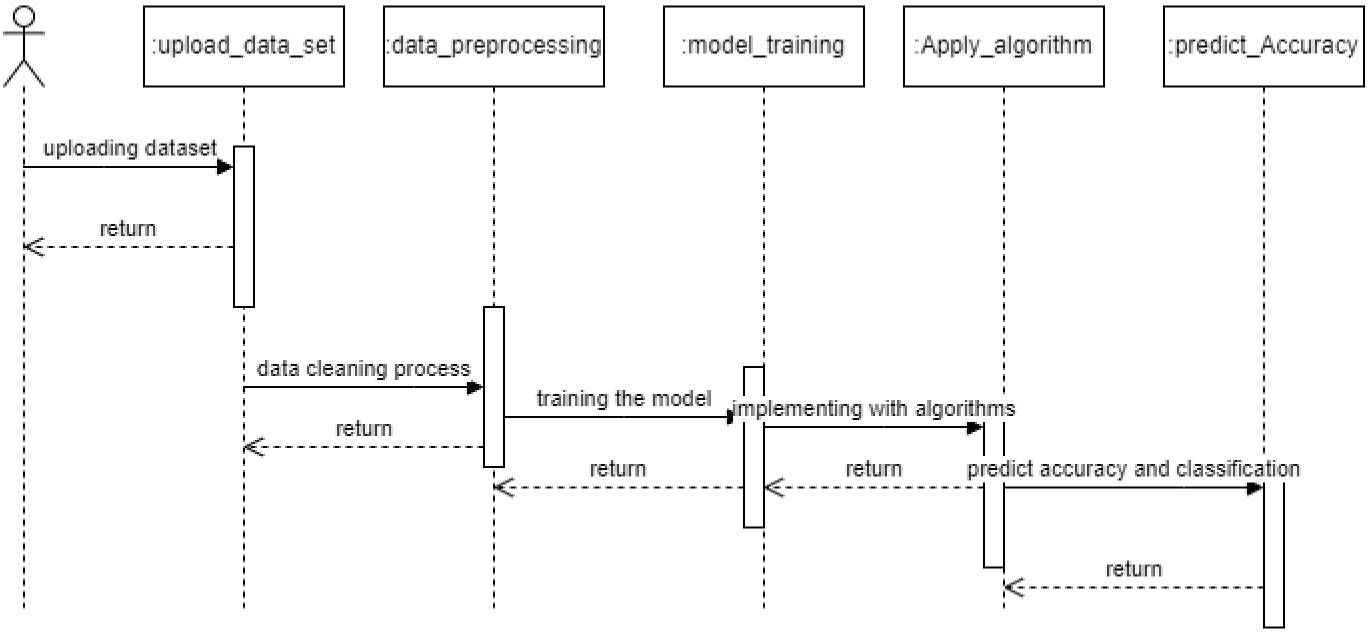
3.3.2 CLASS DIAGRAM:

In software engineering, a class diagram in the Unified Modelling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.



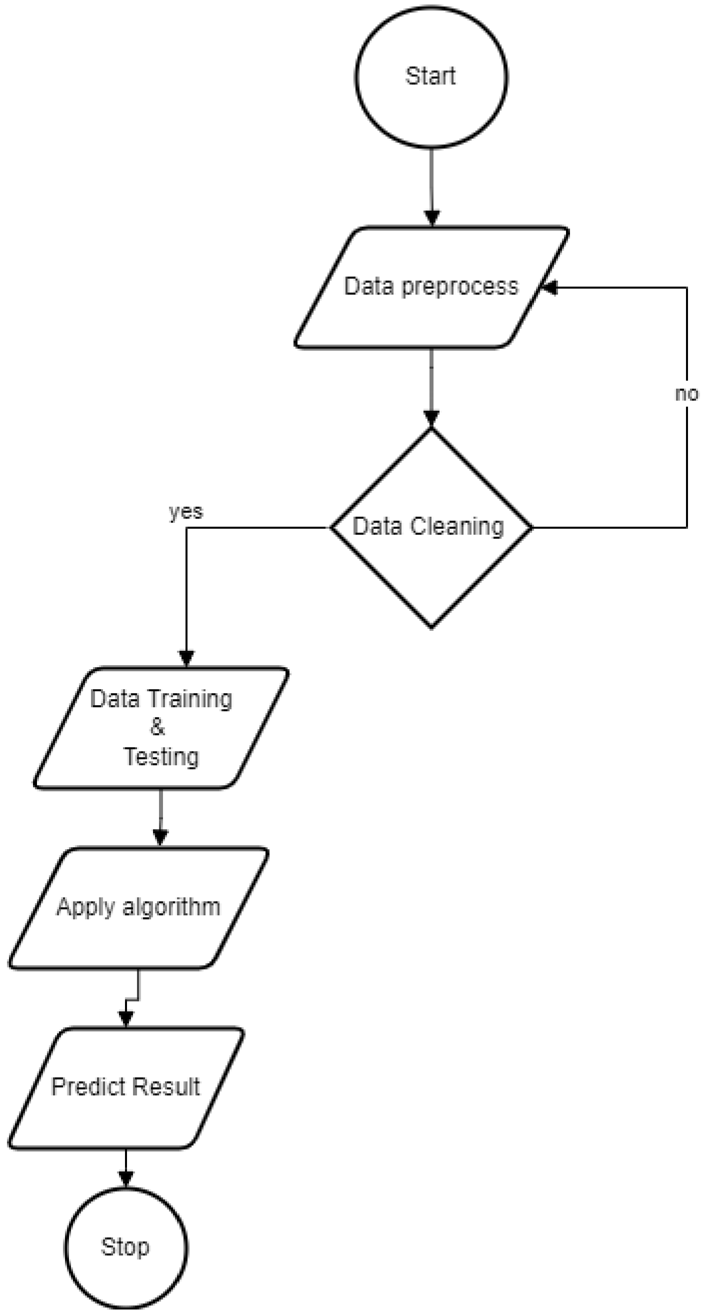
3.3.3 SEQUENCE DIAGRAM:

A sequence diagram in Unified Modelling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.



3.3.4 ACTIVITY DIAGRAM:

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modelling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.



CHAPTER-4: INPUT AND OUTPUT DESIGN

4.1 Input Design

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The input is designed in such a way so that it provides security and ease of use with retaining the privacy.

Input Design considered the following things:

¬ What data should be given as input?

¬ How the data should be arranged or coded?

¬ The dialog to guide the operating personnel in providing input.

¬ Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES

1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

2.It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors It also provides record viewing facilities.

3. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus, the objective of input design is to create an input layout that is easy to follow.

4.2 Output Design

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

2. Select methods for presenting information.

3. Create document, report, or other formats that contain information produced by the system.

The output form of an information system should accomplish one or more of the following objectives.

- Convey information about past activities, current status or projections of the Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action

CHAPTER-5: SYSTEM ENVIRONMENT

5.1 What is Python:-

Below are some facts about Python.

Python is currently the most widely used multi-purpose, high-level programming language.

Python allows programming in Object-Oriented and Procedural paradigms. Python programs generally are smaller than other programming languages like Java.

Programmers have to type relatively less and indentation requirement of the language, makes them readable all the time.

Python language is being used by almost all tech-giant companies like – Google, Amazon, Facebook, Instagram, Dropbox, Uber… etc.

The biggest strength of Python is huge collection of standard library which can be used for the following –

- Machine Learning
- GUI Applications (like Kyiv, Tkinter, PyQt etc.)
- Web frameworks like Django (used by YouTube, Instagram, Dropbox)
- Image processing (like OpenCV, Pillow)
- Web scraping (like Scrapy, BeautifulSoup, Selenium)
- Test frameworks
- Multimedia

Advantages of Python:-

Let's see how Python dominates over other languages.

1. Extensive Libraries

Python downloads with an extensive library and it contain code for various purposes like regular expressions, documentation-generation, unit-testing, web browsers, threading, databases, CGI, email, image manipulation, and more. So, we don't have to write the complete code for that manually.

2.Extensible

As we have seen earlier, Python can be extended to other languages. You can write some of your code in languages like C++ or C. This comes in handy, especially in projects.

3. Embeddable

Complimentary to extensibility, Python is embeddable as well. You can put your Python code in your source code of a different language, like C++. This lets us add scripting capabilities to our code in the other language.

4. Improved Productivity

The language's simplicity and extensive libraries render programmers more productive than languages like Java and C++ do. Also, the fact that you need to write less and get more things done.

5. IOT Opportunities

Since Python forms the basis of new platforms like Raspberry Pi, it finds the future bright for the Internet Of Things. This is a way to connect the language with the real world.

6. Simple and Easy

When working with Java, you may have to create a class to print 'Hello World'. But in Python, just a print statement will do. It is also quite easy to learn**,** understand, and code. This is why when people pick up Python, they have a hard time adjusting to other more verbose languages like Java.

7. Readable

Because it is not such a verbose language, reading Python is much like reading English. This is the reason why it is so easy to learn, understand, and code. It also does not need curly braces to define blocks, and indentation is mandatory**.** These further aids the readability of the code.

8. Object-Oriented

This language supports both the procedural and object-oriented programming paradigms. While functions help us with code reusability, classes and objects let us model the real world. A class allows the encapsulation of data and functions into one.

9. Free and Open-Source

Like we said earlier, Python is freely available. But not only can you download Python for free, but you can also download its source code, make changes to it, and even distribute it. It downloads with an extensive collection of libraries to help you with your tasks.

10. Portable

When you code your project in a language like C++, you may need to make some changes to it if you want to run it on another platform. But it isn't the same with Python. Here, you need to code only once, and you can run it anywhere. This is called Write Once Run Anywhere (WORA). However, you need to be careful enough not to include any system-dependent features.

11. Interpreted

Lastly, we will say that it is an interpreted language. Since statements are executed one by one, debugging is easier than in compiled languages.

Advantages of Python Over Other Languages:

1. Less Coding

Almost all of the tasks done in Python requires less coding when the same task is done in other languages. Python also has an awesome standard library support, so you don't have to search for any third-party libraries to get your job done. This is the reason that many people suggest learning Python to beginners.

2.Affordable

Python is free therefore individuals, small companies or big organizations can leverage the free available resources to build applications. Python is popular and widely used so it gives you better community support.

The 2019 GitHub annual survey showed us that Python has overtaken Java in the most popular programming language category.

3. Python is for Everyone

Python code can run on any machine whether it is Linux, Mac or Windows. Programmers need to learn different languages for different jobs but with Python, you can professionally build web apps, perform data analysis and machine learning, automate things, do web scraping and also build games and powerful visualizations. It is an all-rounder programming language.

Disadvantages of Python:

So far, we've seen why Python is a great choice for your project. But if you choose it, you should be aware of its consequences as well. Let's now see the downsides of choosing Python over another language.

1. Speed Limitations

We have seen that Python code is executed line by line. But since Python is interpreted, it often results in slow execution. This, however, isn't a problem unless speed is a focal point for the project. In other words, unless high speed is a requirement, the benefits offered by Python are enough to distract us from its speed limitations.

2. Weak in Mobile Computing and Browsers

While it serves as an excellent server-side language, Python is much rarely seen on the client-side. Besides that, it is rarely ever used to implement smartphone-based applications. One such application is called Carbon Nelle. The reason it is not so famous despite the existence of Bryton is that it isn't that secure.

3. Design Restrictions

As you know, Python is dynamically-typed. This means that you don't need to declare the type of variable while writing the code. It uses duck-typing. But wait, what's that? Well, it just means that if it looks like a duck, it must be a duck. While this is easy on the programmers during coding, it can raise run-time errors.

4. Underdeveloped Database Access Layers

Compared to more widely used technologies like JDBC (Java Database Connectivity) and ODBC (Open Database Connectivity), Python's database access layers are a bit underdeveloped. Consequently, it is less often applied in huge enterprises.

5. Simple

No, we're not kidding. Python's simplicity can indeed be a problem. Take my example. I don't do Java, I'm more of a Python person. To me, its syntax is so simple that the verbosity of Java code seems unnecessary.

This was all about the Advantages and Disadvantages of Python Programming Language.

5.1.1 History of Python : -

What do the alphabet and the programming language Python have in common? Right, both start with ABC. If we are talking about ABC in the Python context, it's clear that the programming language ABC is meant. ABC is a general-purpose programming language and programming environment, which had been developed in the Netherlands, Amsterdam, at the CWI (Centrum Wickenden &Informatica). The greatest achievement of ABC was to influence the design of Python. Python was conceptualized in the late 1980s. Guido van Rossum worked that time in a project at the CWI, called Amoeba, a distributed operating system. In an interview with Bill Venners[1], Guido van Rossum said: "In the early 1980s, I worked as an implementer on a team building a language called ABC at Centrum voor Wickenden Informatica (CWI). I don't know how well people know ABC's influence on Python. I try to mention ABC's influence because I'm indebted to everything I learned during that project and to the people who worked on it."Later on in the same Interview, Guido van Rossum continued: "I remembered all my experience and some of my frustration with ABC. I decided to try to design a simple scripting language that possessed some of ABC's better properties, but without its problems. So, I started typing. I created a simple virtual machine, a simple parser, and a simple runtime. I made my own version of the various ABC parts that I liked. I created a basic syntax, used indentation for statement grouping instead of curly braces or begin-end blocks, and developed a small number of powerful data types: a hash table (or dictionary, as we call it), a list, strings, and numbers."

5.2 What is Machine Learning: -

Before we take a look at the details of various machine learning methods, let's start by looking at what machine learning is, and what it isn't. Machine learning is often categorized as a subfield of artificial intelligence, but I find that categorization can often be misleading at first brush. The study of machine learning certainly arose from research in this context, but in the data science application of machine learning methods, it's more helpful to think of machine learning as a means of *building models of data.*

Fundamentally, machine learning involves building mathematical models to help understand data. "Learning" enters the fray when we give these models *tunable parameters* that can be adapted to observed data; in this way the program can be considered to be "learning" from the data. Once these models have been fit to previously seen data, they can be used to predict and understand aspects of newly observed data. I'll leave to the reader the more philosophical digression regarding the extent to which this type of mathematical, model-based "learning" is similar to the "learning" exhibited by the human brain. Understanding the problem setting in machine learning is essential to using these tools effectively, and so we will start with some broad categorizations of the types of approaches we'll discuss here.

5.2.1 Categories Of Machine Leaning:-

At the most fundamental level, machine learning can be categorized into two main types: supervised learning and unsupervised learning.

Supervised learning

Supervised learning involves somehow modeling the relationship between measured features of data and some label associated with the data; once this model is determined, it can be used to apply labels to new, unknown data. This is further subdivided into classification tasks and regression tasks: in classification, the labels are discrete categories, while in regression, the labels are continuous quantities. We will see examples of both types of supervised learning in the following section.

Unsupervised learning:

Unsupervised learning involves modeling the features of a dataset without reference to any label, and is often described as "letting the dataset speak for itself." These models include tasks such as clustering and dimensionality reduction. Clustering algorithms identify distinct groups of data, while dimensionality reduction algorithms search for more succinct representations of the data. We will see examples of both types of unsupervised learning in the following section.

5.2.2 Need for Machine Learning

Human beings, at this moment, are the most intelligent and advanced species on earth because they can think, evaluate and solve complex problems. On the other side, AI is still in its initial stage and haven't surpassed human intelligence in many aspects. Then the question is that what is the need to make machine learn? The most suitable reason for doing this is, "to make decisions, based on data, with efficiency and scale".

Lately, organizations are investing heavily in newer technologies like Artificial Intelligence, Machine Learning and Deep Learning to get the key information from data to perform several real-world tasks and solve problems. We can call it data-driven decisions taken by machines, particularly to automate the process. These data-driven decisions can be used, instead of using programing logic, in the problems that cannot be programmed inherently. The fact is that we can't do without human intelligence, but other aspect is that we all need to solve real-world problems with efficiency at a huge scale. That is why the need for machine learning arises.

5.2.3 Challenges in Machines Learning:-

While Machine Learning is rapidly evolving, making significant strides with cybersecurity and autonomous cars, this segment of AI as whole still has a long way to go. The reason behind is that ML has not been able to overcome number of challenges. The challenges that ML is facing currently are −

Quality of data − Having good-quality data for ML algorithms is one of the biggest challenges. Use of low-quality data leads to the problems related to data preprocessing and feature extraction.

Time-Consuming task − Another challenge faced by ML models is the consumption of time especially for data acquisition, feature extraction and retrieval.

Lack of specialist persons − As ML technology is still in its infancy stage, availability of expert resources is a tough job.

No clear objective for formulating business problems − Having no clear objective and well-defined goal for business problems is another key challenge for ML because this technology is not that mature yet.

Issue of overfitting & underfitting − If the model is overfitting or underfitting, it cannot be represented well for the problem.

Curse of dimensionality − Another challenge ML model faces is too many features of data points. This can be a real hindrance.

Difficulty in deployment − Complexity of the ML model makes it quite difficult to be deployed in real life.

5.2.4 Applications of Machines Learning:-

Machine Learning is the most rapidly growing technology and according to researchers we are in the golden year of AI and ML. It is used to solve many real-world complex problems which cannot be solved with traditional approach. Following are some real-world applications of ML −

- Emotion analysis
- Sentiment analysis
- Error detection and prevention
- Weather forecasting and prediction
- Stock market analysis and forecasting
- Speech synthesis
- Speech recognition
- Customer segmentation
- Object recognition
- Fraud detection
- Fraud prevention
- Recommendation of products to customer in online shopping

How to Start Learning Machine Learning?

Arthur Samuel coined the term "Machine Learning" in 1959 and defined it as a "Field of study that gives computers the capability to learn without being explicitly programmed "And that was the beginning of Machine Learning! In modern times, Machine Learning is one of the most popular (if not the most!) career choices. According to Indeed, Machine Learning

Engineer Is the Best Job of 2019 with a *344%* growth and an average base salary of $146,085 per year.

how to start learning it? So, this article deals with the Basics of Machine Learning and also the path you can follow to eventually become a full-fledged Machine Learning Engineer. Now let's get started!!!

How to start learning ML?

This is a rough roadmap you can follow on your way to becoming an insanely talented Machine Learning Engineer. Of course, you can always modify the steps according to your needs to reach your desired end-goal!

Step 1 – Understand the Prerequisites

In case you are a genius, you could start ML directly but normally, there are some prerequisites that you need to know which include Linear Algebra, Multivariate Calculus, Statistics, and Python. And if you don't know these, never fear! You don't need a Ph.D. degree in these topics to get started but you do need a basic understanding.

(a) Learn Linear Algebra and Multivariate Calculus

Both Linear Algebra and Multivariate Calculus are important in Machine Learning. However, the extent to which you need them depends on your role as a data scientist. If you are more focused on application heavy machine learning, then you will not be that heavily focused on maths as there are many common libraries available. But if you want to focus on R&D in Machine Learning, then mastery of Linear Algebra and Multivariate Calculus is very important as you will have to implement many ML algorithms from scratch.

(b) Learn Statistics

Data plays a huge role in Machine Learning. In fact, around 80% of your time as an ML expert will be spent collecting and cleaning data. And statistics is a field that handles the collection, analysis, and presentation of data. So, it is no surprise that you need to learn it!!! Some of the key concepts in statistics that are important are Statistical Significance, Probability Distributions, Hypothesis Testing, Regression, etc. Also, Bayesian Thinking is also a very important part of ML which deals with various concepts like Conditional Probability, Priors, and Posteriors, Maximum Likelihood, etc.

(c) Learn Python

Some people prefer to skip Linear Algebra, Multivariate Calculus and Statistics and learn them as they go along with trial and error. But the one thing that you absolutely cannot skip is Python! While there are other languages you can use for Machine Learning like R, Scala, etc. Python is currently the most popular language for ML. In fact, there are many Python libraries that are specifically useful for Artificial Intelligence and Machine Learning such as Kera's, TensorFlow, Scikit-learn, etc.

So if you want to learn ML, it's best if you learn Python! You can do that using various online resources and courses such as Fork Python available Free on GeeksforGeeks.

Step 2 – Learn Various ML Concepts

Now that you are done with the prerequisites, you can move on to actually learning ML (Which is the fun part!!!) It's best to start with the basics and then move on to the more complicated stuff. Some of the basic concepts in ML are:

(a) Terminologies of Machine Learning

- **Model –** A model is a specific representation learned from data by applying some machine learning algorithm. A model is also called a hypothesis.
- **Feature –** A feature is an individual measurable property of the data. A set of numeric features can be conveniently described by a feature vector. Feature vectors are fed as input to the model. For example, in order to predict a fruit, there may be features like colour, smell, taste, etc.

- **Target (Label) –** A target variable or label is the value to be predicted by our model. For the fruit example discussed in the feature section, the label with each set of input would be the name of the fruit like apple, orange, banana, etc.
- **Training –** The idea is to give a set of inputs(features) and it's expected outputs(labels), so after training, we will have a model (hypothesis) that will then map new data to one of the categories trained on.
- **Prediction –** Once our model is ready, it can be fed a set of inputs to which it will provide a predicted output(label).

(b) **Types of Machine Learning**

- **Supervised Learning –** This involves learning from a training dataset with labelled data using classification and regression models. This learning process continues until the required level of performance is achieved.
- **Unsupervised Learning –** This involves using unlabelled data and then finding the underlying structure in the data in order to learn more and more about the data itself using factor and cluster analysis models.
- **Semi-supervised Learning –** This involves using unlabelled data like Unsupervised Learning with a small amount of labelled data. Using labelled data vastly increases the learning accuracy and is also more cost-effective than Supervised Learning.
- **Reinforcement Learning –** This involves learning optimal actions through trial and error. So, the next action is decided by learning behaviours that are based on the current state and that will maximize the reward in the future.

Advantages of Machine learning: -

1.Easily identifies trends and patterns

Machine Learning can review large volumes of data and discover specific trends and patterns that would not be apparent to humans. For instance, for an e-commerce website like Amazon, it serves to understand the browsing behaviours and purchase histories of its users to help cater to the right products, deals, and reminders relevant to them. It uses the results to reveal relevant advertisements to them.

2.No human intervention needed (automation)

With ML, you don't need to babysit your project every step of the way. Since it means giving machines the ability to learn, it lets them make predictions and also improve the algorithms on their own. A common example of this is anti-virus software's; they learn to filter new threats as they are recognized. ML is also good at recognizing spam.

3.Continuous Improvement

As ML algorithms gain experience, they keep improving in accuracy and efficiency. This lets them make better decisions. Say you need to make a weather forecast model. As the amount of data, you have keeps growing, your algorithms learn to make more accurate predictions faster.

4.Handling multi-dimensional and multi-variety data

Machine Learning algorithms are good at handling data that are multi-dimensional and multi-variety, and they can do this in dynamic or uncertain environments.

5.Wide Applications

You could be an e-tailer or a healthcare provider and make ML work for you. Where it does apply, it holds the capability to help deliver a much more personal experience to customers while also targeting the right customers.

Disadvantages of Machine Learning: -

1.Data Acquisition

Machine Learning requires massive data sets to train on, and these should be inclusive/unbiased, and of good quality. There can also be times where they must wait for new data to be generated.

2.Time and Resources

ML needs enough time to let the algorithms learn and develop enough to fulfill their purpose with a considerable amount of accuracy and relevancy. It also needs massive resources to function. This can mean additional requirements of computer power for you.

2.Interpretation of Results

Another major challenge is the ability to accurately interpret results generated by the algorithms. You must also carefully choose the algorithms for your purpose.

4. High error-susceptibility

Machine Learning is autonomous but highly susceptible to errors. Suppose you train an algorithm with data sets small enough to not be inclusive. You end up with biased predictions coming from a biased training set. This leads to irrelevant advertisements being displayed to customers. In the case of ML, such blunders can set off a chain of errors that can go undetected for long periods of time. And when they do get noticed, it takes quite some time to recognize the source of the issue, and even longer to correct it.

Python Development Steps : -

Guido Van Rossum published the first version of Python code (version 0.9.0) at outsources in February 1991. This release included already exception handling, functions, and the core data types of list, dict, str and others. It was also object oriented and had a module system. Python version 1.0 was released in January 1994. The major new features included in this release were the functional programming tools lambda, map, filter and reduce, which Guido Van Rossum never liked. Six and a half years later in October 2000, Python 2.0 was introduced. This release included list comprehensions, a full garbage collector and it was supporting Unicode. Python flourished for another 8 years in the versions 2.x before the next major release as Python 3.0 (also known as "Python 3000" and "Py3K") was released. Python 3 is not backwards compatible with Python 2.x. The emphasis in Python 3 had been on the removal of duplicate programming constructs and modules, thus fulfilling or coming close to fulfilling the 13th law of the Zen of Python: "There should be one -- and preferably only one -- obvious way to do it."Some changes in Python 7.3:

- Print is now a function
- Views and iterators instead of lists
- The rules for ordering comparisons have been simplified. E.g. a heterogeneous list cannot be sorted, because all the elements of a list must be comparable to each other.
- There is only one integer type left, i.e. int. long is int as well.

- The division of two integers returns a float instead of an integer. "//" can be used to have the "old" behaviour.
- Text Vs. Data Instead of Unicode Vs. 8-bit
- **Purpose:-**

We demonstrated that our approach enables successful segmentation of intra-retinal layers—even with low-quality images containing speckle noise, low contrast, and different intensity ranges throughout—with the assistance of the ANIS feature.

Python

Python is an interpreted high-level programming language for general-purpose programming. Created by Guido van Rossum and first released in 1991, Python has a design philosophy that emphasizes code readability, notably using significant whitespace. Python features a dynamic type system and automatic memory management. It supports multiple programming paradigms, including object-oriented, imperative, functional and procedural, and has a large and comprehensive standard library.

- Python is Interpreted − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.
- Python is Interactive − you can actually sit at a Python prompt and interact with the interpreter directly to write your programs.

  Python also acknowledges that speed of development is important. Readable and terse code is part of this, and so is access to powerful constructs that avoid tedious repetition of code. Maintainability also ties into this may be an all but useless metric, but it does say something about how much code you have to scan, read and/or understand to troubleshoot problems or tweak behaviors. This speed of development, the ease with which a programmer of other languages can pick up basic Python skills and the huge standard library is key to another area where Python excels. All its tools have been quick to implement, saved a lot of time, and several of them have later been patched and updated by people with no Python background - without breaking.

5.3 Modules Used in Project :-

5.3.1 TensorFlow

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It is used for both research and production at Google.

TensorFlow was developed by the Google Brain team for internal Google use. It was released under the Apache 2.0 open-source license on November 9, 2015.

5.3.2 NumPy

NumPy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.

It is the fundamental package for scientific computing with Python. It contains various features including these important ones:

- A powerful N-dimensional array object
- Sophisticated (broadcasting) functions
- Tools for integrating C/C++ and Fortran code
- Useful linear algebra, Fourier transform, and random number capabilities

Besides its obvious scientific uses, NumPy can also be used as an efficient multi-dimensional container of generic data. Arbitrary data-types can be defined using NumPy which allows NumPy to seamlessly and speedily integrate with a wide variety of databases.

5.3.3 Pandas

Pandas is an open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures.

Python was majorly used for data munging and preparation. It had very little contribution towards data analysis. Pandas solved this problem. Using Pandas, we can accomplish five typical steps in the processing and analysis of data, regardless of the origin of data load, prepare, manipulate, model, and analyze. Python with Pandas is used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc.

5.3.4 Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Python and IPython shells, the Jupyter Notebook, web application servers, and four graphical user interface toolkits. Matplotlib tries to make easy things easy and hard things possible. You can generate plots, histograms, power spectra, bar charts, error charts, scatter plots, etc., with just a few lines of code. For examples, see the sample plots and thumbnail gallery.

For simple plotting the pyplot module provides a MATLAB-like interface, particularly when combined with IPython. For the power user, you have full control of line styles, font properties, axes properties, etc, via an object-oriented interface or via a set of functions familiar to MATLAB users.

5.3.5 Scikit – learn

Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python. It is licensed under a permissive simplified BSD license and is distributed under many Linux distributions, encouraging academic and commercial use.

FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- ECONOMICAL FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY

6.1 ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus, the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

6.2 TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

6.3 SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

CHAPTER-7: SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of tests. Each test type addresses a specific testing requirement.

7.1 TYPES OF TESTING

7.1.1 Unit Testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

7.1.2 Integration Testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

7.1.3 Functional Testing

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems : interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

7.1.4 System Testing

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration-oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

White Box Testing

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

Black Box Testing

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box. you cannot "see" into it. The test provides inputs and responds to outputs without considering how the software works.

7.1.5 Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

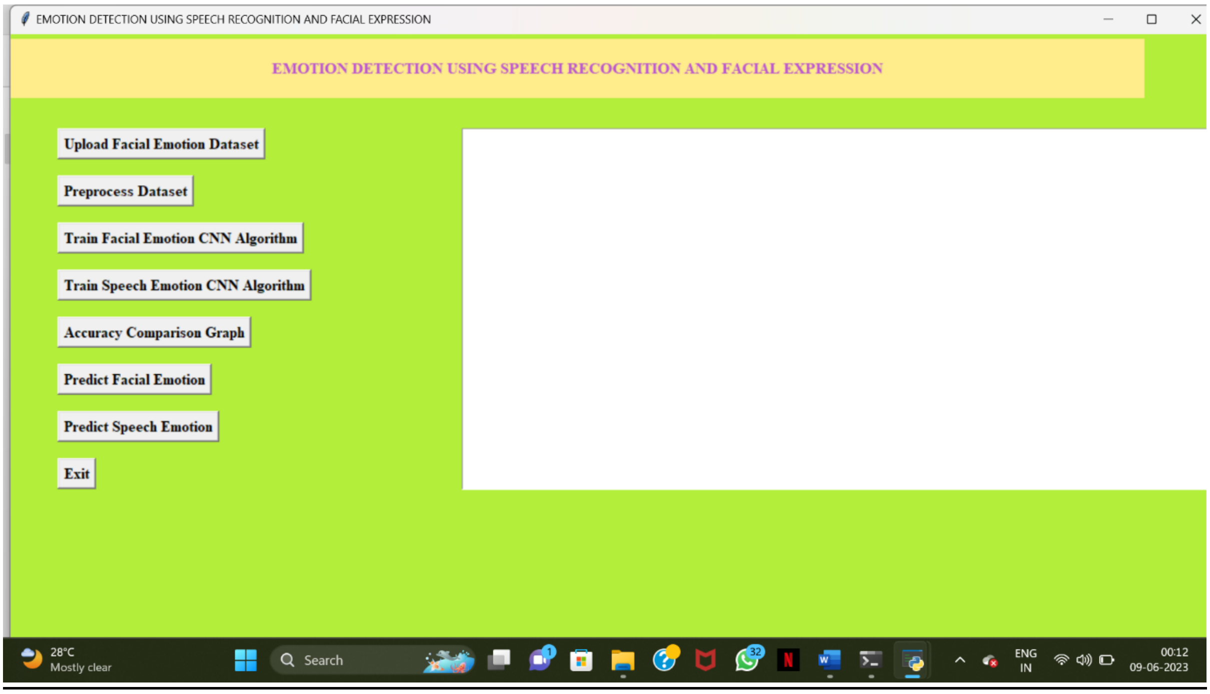


Fig8.1: Screenshot of home page

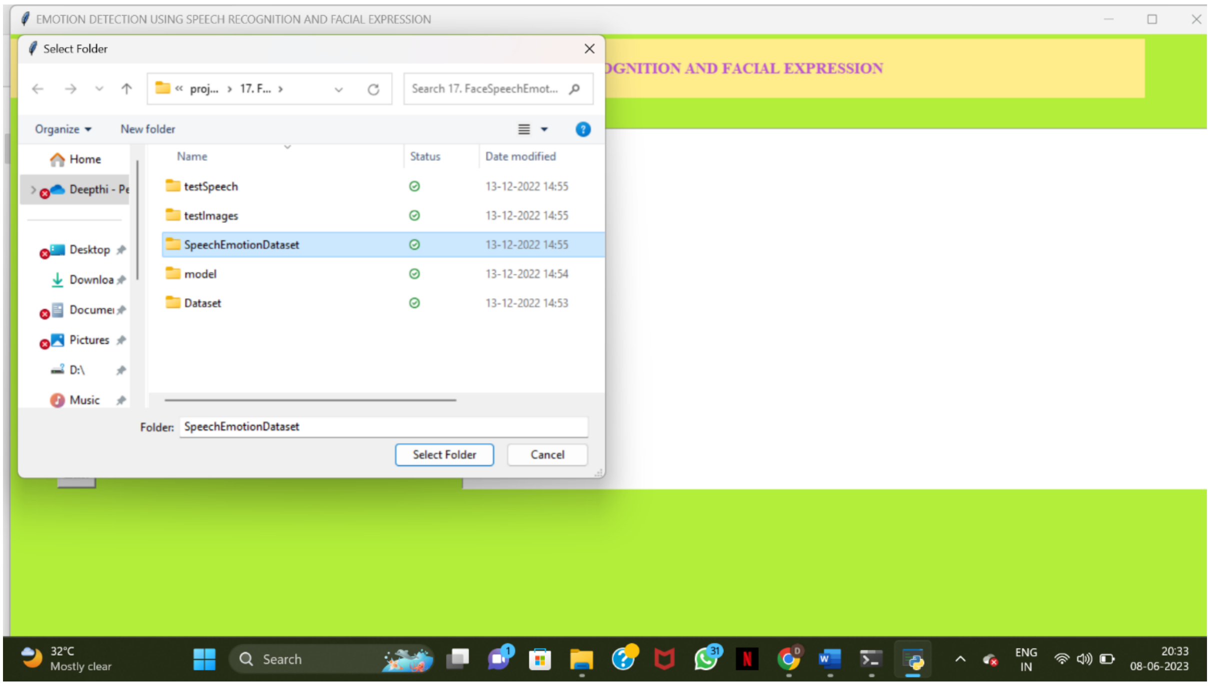


Fig.8.2 : Screenshot of Uploading of dataset

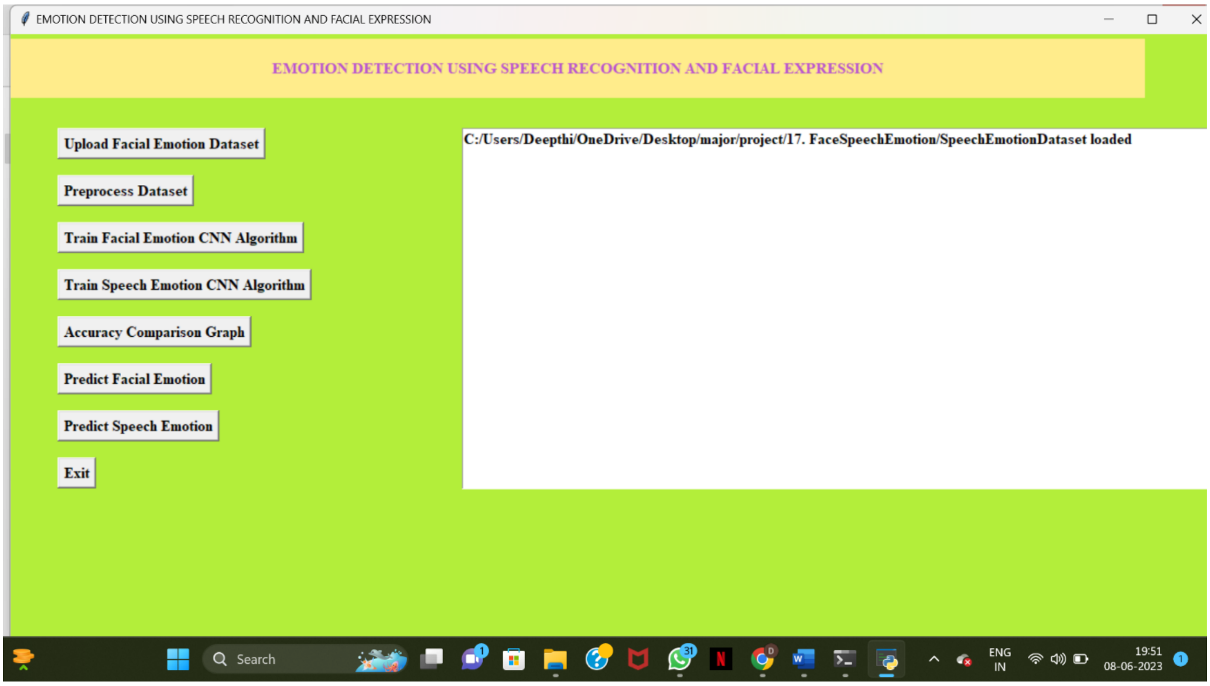


Fig 8.3: Screenshot of pre-processed data

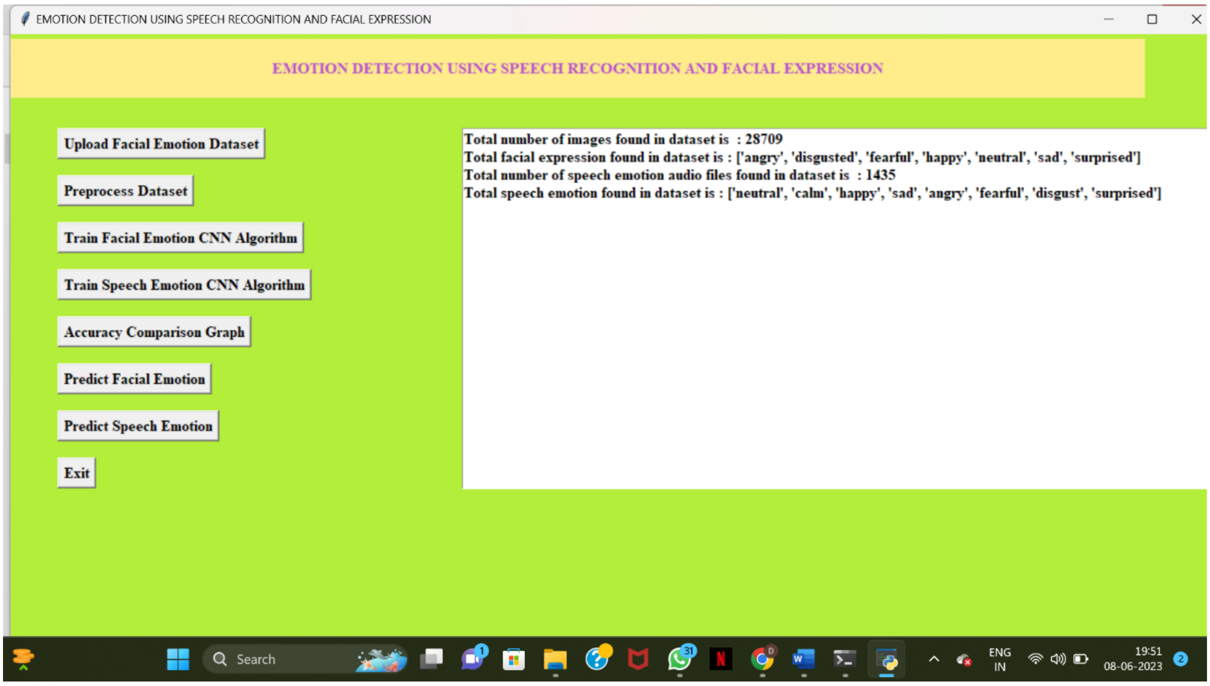


Fig 8.4: Screenshot of data uploaded

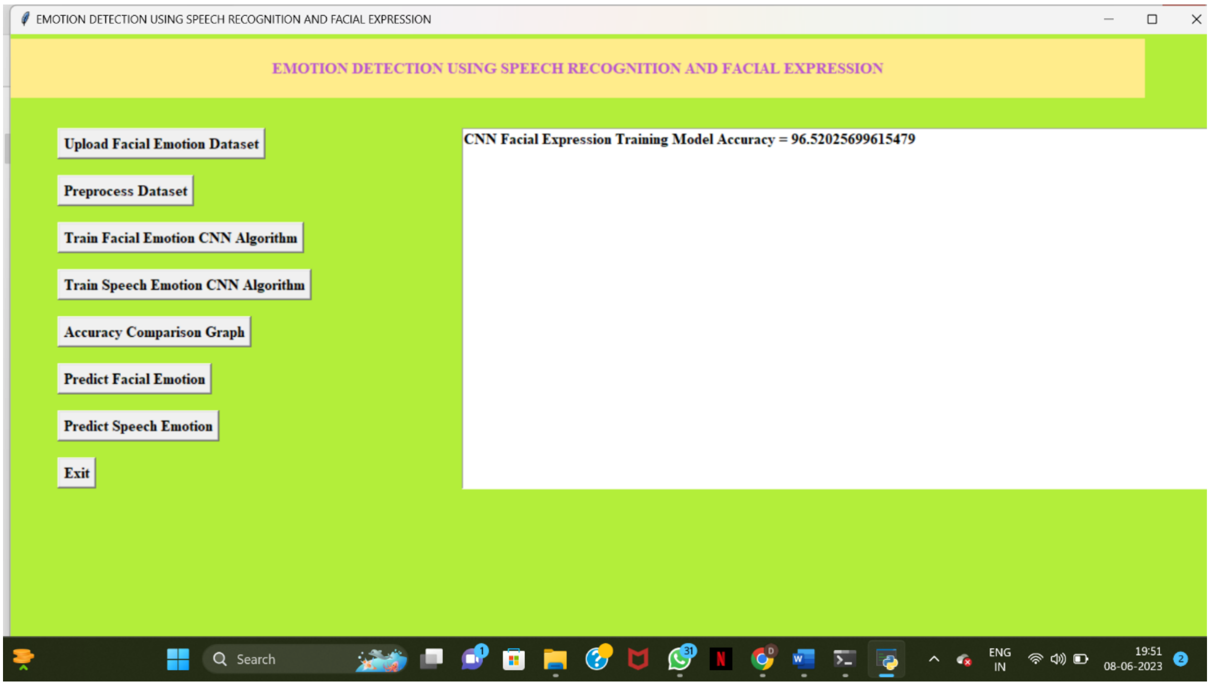


Fig 8.5: Screenshot of accuracy

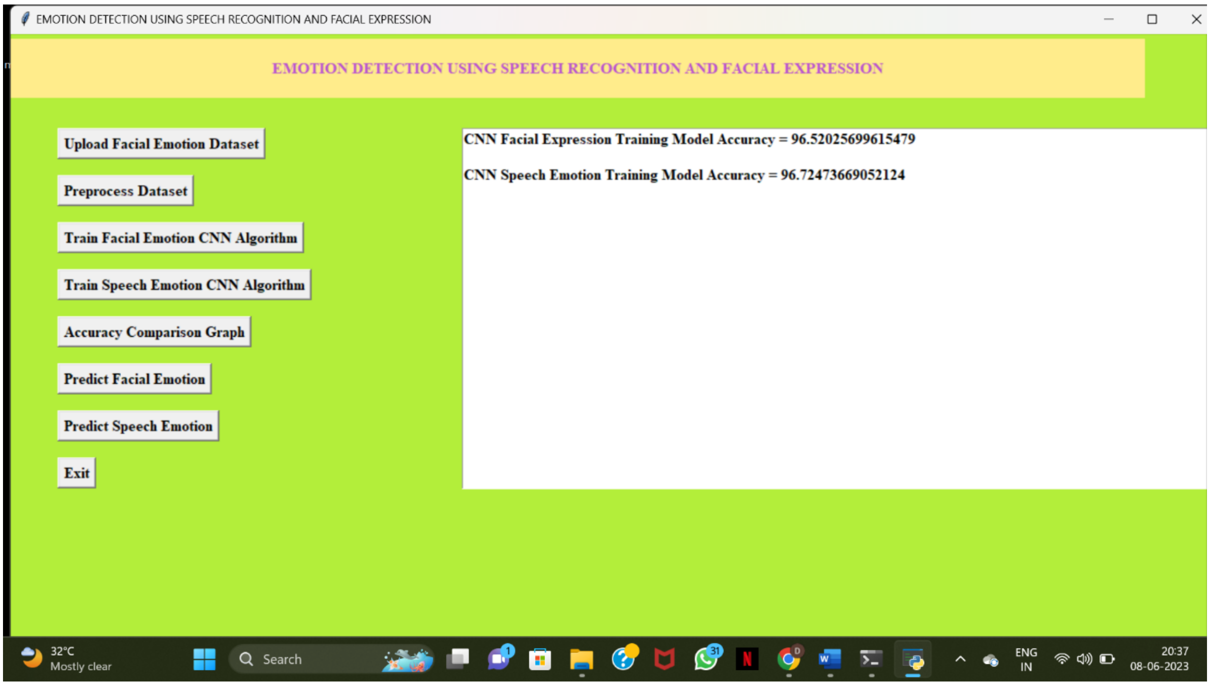


Fig 8.6: Screenshot of CNN with facial and speech accuracy

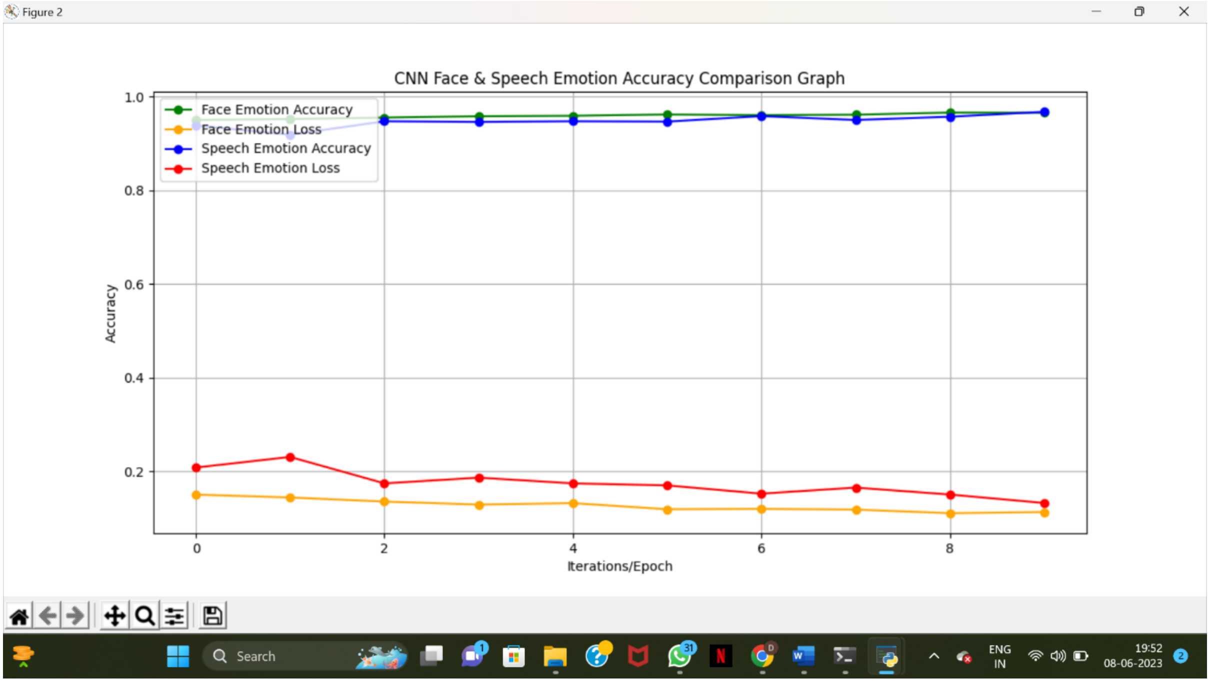


Fig 8.7: Screenshot of accuracy graph

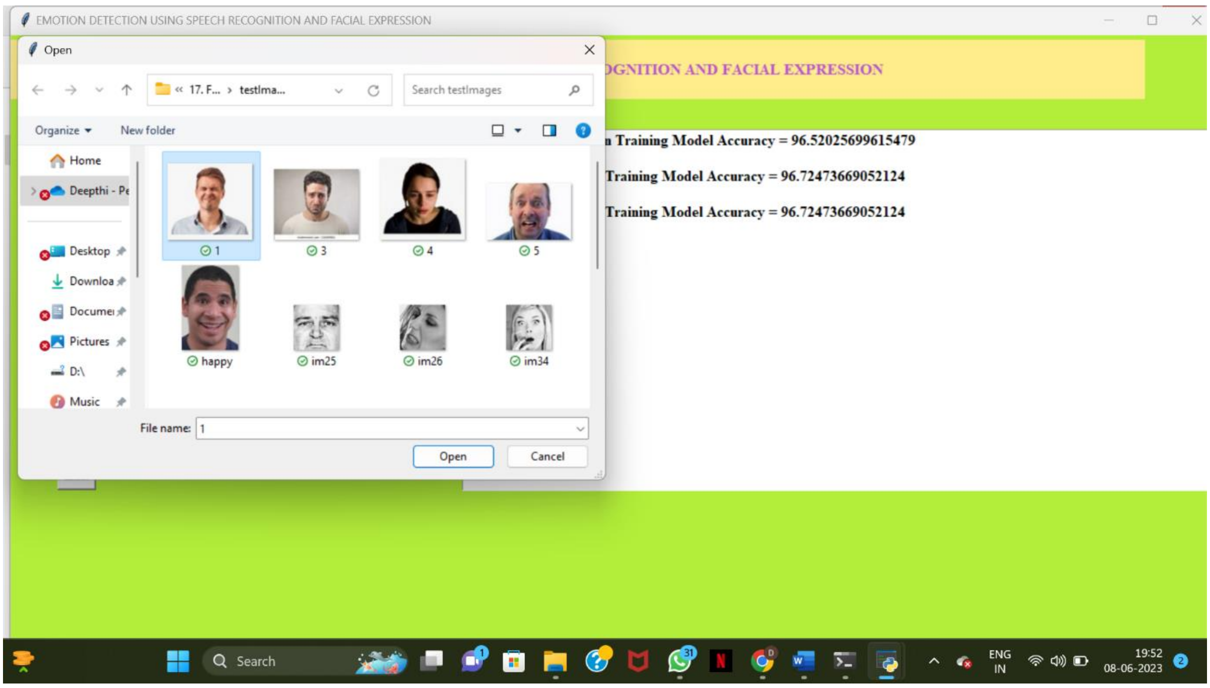


Fig 8.8: Screenshot of uploading of facial expression

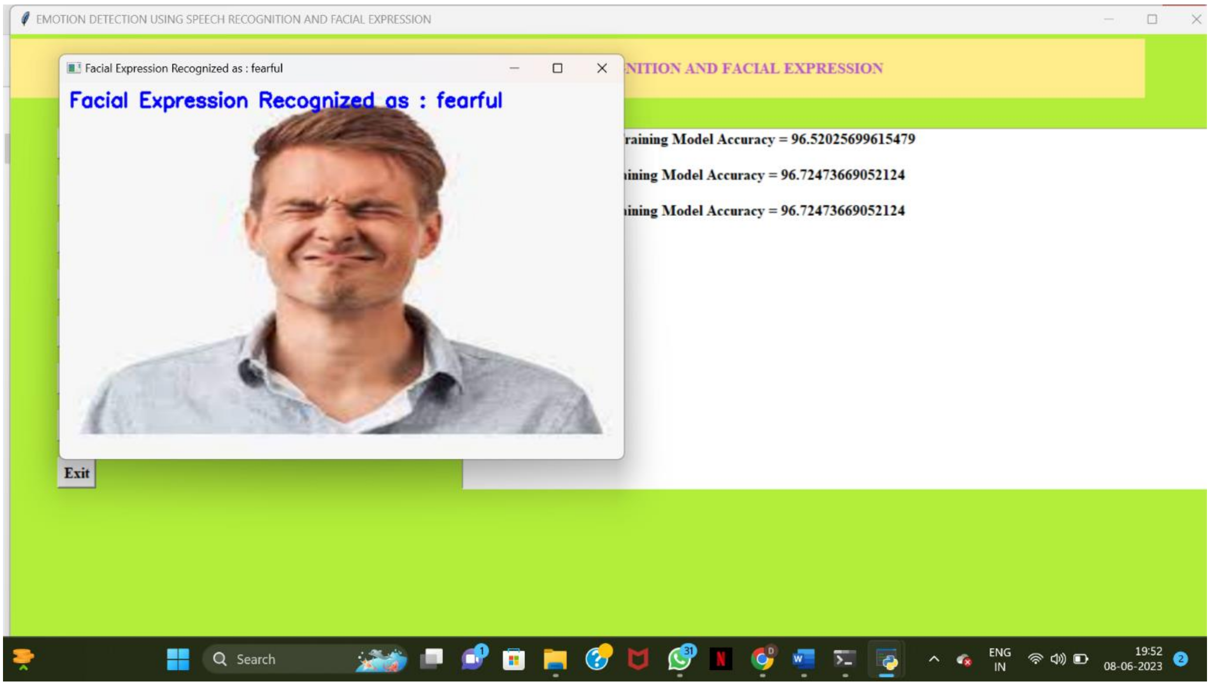


Fig 8.9: Screenshot of Recognition of facial expression

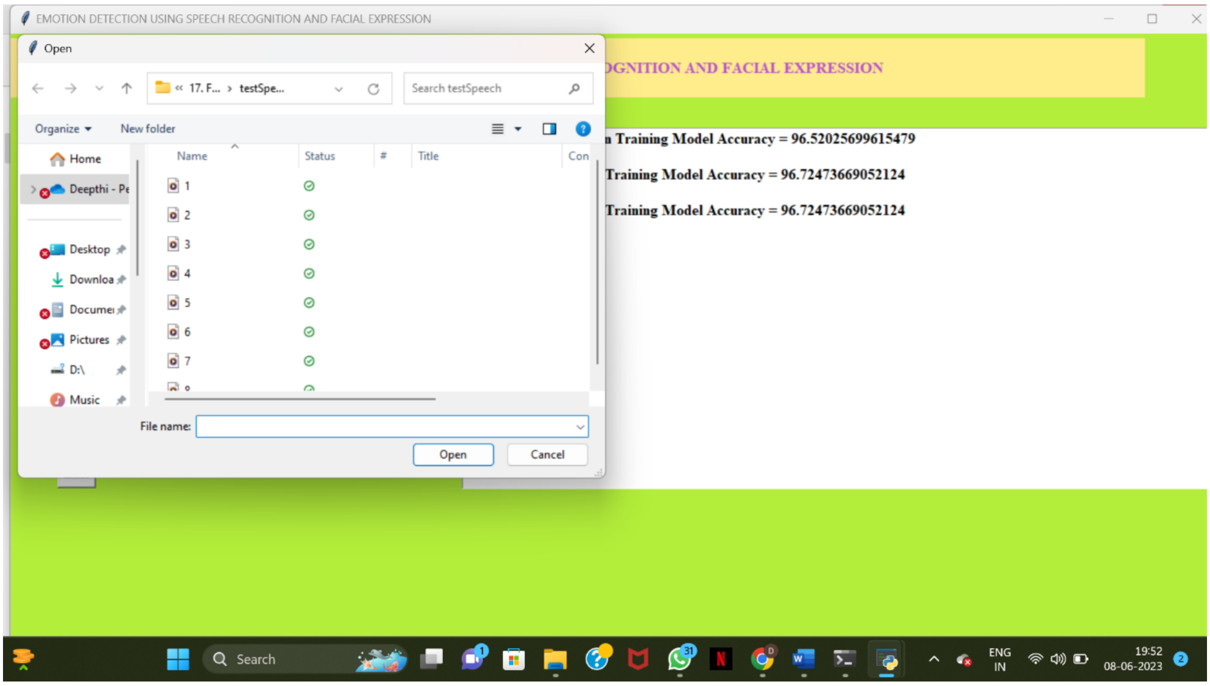


Fig 8.10: Screenshot of uploading of speech

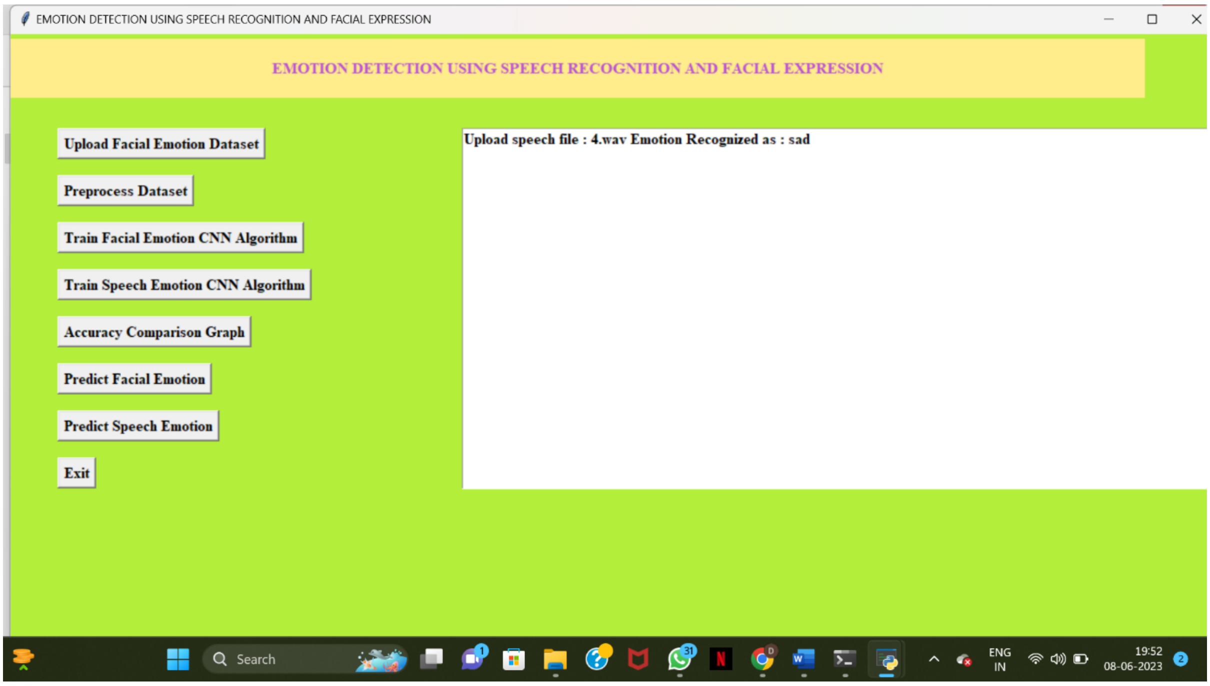


Fig 8.11: Screenshot of emotion recognized with speech

CHAPTER-9: CONCLUSION &FUTURE ENHANCEMENT

CONCLUSION

We presented a bimodal framework for emotion analysis and recognition starting from expressive faces and speech. The proposed approach tried to distinguish between six emotions (happiness, angry, fear, sadness, disgusted and neutral state) by using different classifiers. The maximum accuracy was achieved for Support Vector Machine (SVM). The first purpose was to select several suitable features for the task of emotion classification. A secondary aim was to analyse the strengths and the limitations of the unimodal emotion recognition systems based on facial expressions and speech features. The results reveal that the system based on facial expression gave better performance than the one based on speech information only for the considered emotions. It can be observed that even though the system based on audio information had poorer performance than the facial expression emotion classifier, its features have valuable information about emotions, that cannot be extracted from the visual information. Audio and visual data present complementary information. When these two modalities are fused, the performance and the robustness of the emotion recognition system are improved. Further, the fusion performed at the feature level showed better results than the one performed at the score level.

FUTURE ENHANCEMENT

For future research, the integration of physiological states such as heart beat and skin conductivity would be expected to improve the recognition rates and eventually improve the computer’s understanding of human emotional states. Gestures are widely believed to play an important role as well.

CHAPTER 10: BIBLIOGRAPHY

[1] G. Castellano, L. Kessous, “Multimodal emotion recognition from expressive faces, body gestures and speech”, 2nd International Conference on Affective Computing and Intelligent Interaction, Lisbon, September 2007N.

[2] J. Kim, “Bimodal emotion recognition using speech and physiological changes”, Robust Speech Recognition and Understanding, I-Tech Education and Publishing, Vienna, 2007.

[3] Sebe, I. Cohen, T. Gevers, T.S. Huang, “Multimodal Approaches for Emotion Recognition: A Survey” Internet Imaging VI, SPIE’05, San USA, January 2005.

[4] C. Busso, Z. Deng, et al, “Analysis of emotion recognition using facial expressions, speech and multimodal information”, Proceedings of the international conference on Multimodal interfaces, State College, USA, 2004

[5] Chen, L.S., Huang, T.S. “Emotional expressions in audiovisual human computer interaction“. Multimedia and Expo, ICME 2000.

[6] Vapnik, V. “The Nature of Statistical Learning Theory”, New York, NY: Springer-Verlag, 1995

[7] Ekman, P., Friesen, W. V. Facial Action Coding System: A Technique for Measurement of Facial Movement. Consulting Psychologists Press Palo Alto, California, 1978.

[8] Lee, Chul Min et al. "Emotion recognition based on phoneme classes" Interspeech 2004, pg. 889-892.

[9] S. Mallat, “A wavelet tour of signal processing”, New York, Academic Press, 1999

[10] M.K Hu, “Visual pattern recognition by moment invariants”, IRE Trans.Inf. Theory, It-8, 1962, pp.179-187.

[11] T Fujimura, H Umemura, “Development and validation of a facial expression database based on the dimensional and categorical model of emotions”, Cognition and Emotion, 2018 - Taylor & Francis.

[12] F Wallhoff, M AblaBmeier, G Rigoll, “Multimodal face detection, head orientation and eye gaze tracking”, 2006 IEEE International Conference On Multisensory …, 2006

[13] Y. I. Tian, T. Kanade and J. F. Cohn, "Recognizing action units for expression analysis", IEEE Transactions on pattern analysis and machine intelligence, vol. 23, no. 2, pp. 97-115, 2001.

[14] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and "The extended Cohn Kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression", Computer Vision and Pattern Recognition Workshops (CVPRW) 2010 IEEE Computer Society Conference on, pp. 94-101, 2010.

[15] C. R. Darwin, "The expression of the emotions in man and animals" in , John Murray, London, 1872.

[16] E. Correa, A. Jonker, M. Ozo and R. Stolk, "Emotion recognition using deep convolutional neural networks", 2016.

[17] L. C. De Silva, T. Miyasato and F. Kishino, "Emotion enhanced multimedia meetings using the concept of virtual space teleconferencing", Procs. of IEEE Multimedia Systems'96 Hiroshima Japan, pp. 28-33, 1996-June.

[18] T. Sakaguchi, J. Ohya and F. Kishino, "Facial expression recognition image sequence using hidden Markov model", Transactions Engineers Japan, vol. 49, no. 8, pp. 1060-1067, August 1995.

International Journal of Gender, Science and Technology UGC Care Group I Journal
..ISSN: 2040-0748 Vol-12 Issue-01 June 2023

EMOTION DETECTION USING SPEECH RECOGNIZATION AND FACIAL EXPRESSIONS

***K.Neha Nandini*[1]., *G.Varshitha*[2]., *K.Sai Preethi*[3]., *N.Rikhitha*[4] ., *P.Deepthi*[5]**

1 Assistant Professor, Department of CSE., Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India (✉ nehanandini.kella@gmail.com)

2, 3, 4, 5 B.Tech CSE, (19RG1A05574, 19RG1A0591, 19RG1A05A3, 19RG1A05A6),

Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India

Abstract

A lot of progress has been made in the field of automated facial emotion recognition (FER) in recent years. FER has been implemented in human-centered computing and the emerging field of emotional AI (EAI), both of which aim to improve human-machine interactions. The goal of EAI research is to improve computers' ability to read and interpret human emotions and actions in a wide variety of contexts. In recent years, neural networks have undergone tremendous evolution, leading to the development of new designs to tackle ever more challenging tasks. Deep learning has been the most influential technique in this regard. In this post, we'll look at the state of the art in AI-based automatic emotion identification utilizing cutting-edge deep learning models. We demonstrate that deep learning-based FER can work in tandem with other models that employ architecture-related approaches, such as databases, to provide very precise outcomes.

Introduction

Facial expressions may be easily recognized thanks to the major and unique characteristics of the human face. Definition of FER - When a person's outward expression betrays their true feelings. It finds widespread usage in computer vision, digital image processing, and artificial intelligence, as well as HCI applications including face image processing, facial video surveillance, and facial animation. The challenging problem of automatic facial expression identification has attracted the attention of a growing number of scientists in recent years. The feature extraction process is vital in FER. A.L. and coworkers[1] that 55% of all communication is conveyed via facial expressions, whereas just 38% is through verbal and spoken communication.

There are two main approaches to FER system design. Some methods use a series of pictures, beginning with a neutral expression and progressing to the most extreme ones, as a training phase. However, systems that just employ a single face picture to make an emotion recognition determination tend to underperform when compared to state-of-the-art methods [2,3]. A FER system may be classified in more than one way, depending on the attributes it uses for recognition, in addition to the approach type it models. categories. The first group of characteristics is gleaned from the orientation of the face organs and the feel of the skin. Geometric features are the second sort of feature; they store data about the face's different locations and points and may be used to evaluate either a single picture or a series of photographs by tracking the face's motion across the frame. Extracting geometric characteristics from a face might begin with identifying key facial landmarks. Landmarks are prominent facial features that may be used to learn more about a person's identity from observation alone. Numerous more experiments have been conducted on facial landmark identification, however these are outside the purview of this paper. In order to locate these landmarks, this study makes use of the Python library dlib [4]. There are two distinct aspects of AI that are involved in the automated recognition of human emotions and psychology. Researchers in the fields of psychology and artificial intelligence are focusing on answering the challenge of how to recognize human emotions. The tone and auditory modifications [5] that are widely available, for example [6] and from which a rapid mood evaluation may be ob- tained, and other sources [7] are examples of the vocal and nonverbal sensors that create subjects like mood and accent. Mehrian's research [8] found that the senses accounted for 55% of information (emotional and linguistic) and the other 7% had some kind of physical component that was not defined. Many scientists are keen to study facial expressions since they are the first reliable indicator of an individual's emotional state.

To begin, we'll be expanding our extraction features. improvements to other aspects, adding features to an existing representation might be beneficial. Each coded movement in the Facial Action Coding System (FACS) is assumed to require at least one facial muscle, as highlighted by Ekman and Friesen [9]. When it comes to coding head motions, Ekman and Friesen were the first to notice how FACS facial movement is implemented in FACS facial AUs. (among several volunteers of various sexes and/or genetic backgrounds.

Purpose of the Investigation 1.1. There are several universal and fundamental facial expressions that may be used to convey a wide range of human emotions. If an algorithm can be built to recognize, extract, and evaluate these facial expressions in real time, then it will be feasible to automatically identify emotions in still photographs and movies. Expressions on one's face may convey a wealth of information about one's state of mind and goals in a social context. They're crucial to how we communicate with one another as social beings. Facial expression processing greatly benefits from the *additional context provided by seeing faces in their natural environments. Understanding and expressing empathy in interpersonal interactions is* *essential. Automatic emotion recognition has always been a hotly debated topic in psychology. Because of this, a lot of development has occurred in this field. Words, gestures, and expressions on the face and body are only few of the ways in which we communicate our emotions.*

expressions. Therefore, it is essential for human-machine communication that both parties can read and understand expressions of emotion.

1.1. 1.1. This Study's Original Contribution

In this study, we survey the state of the art in emotion recognition using several architectures for identifying a wide variety of expressive styles from face signals alone. The most up-to-date findings from 2016-2021 are published, and analyses of the most common issues and current efforts to fixing them are provided as well. How it works is as follows. Section 2 introduces the prototypical facial expressions and other foundational kinds used to define face expressions, such as FACS. In Section 3, we see the framework of our system for recognizing people's facial expressions and moods. Practical applications of current FER findings are discussed in Section 4. The difficulties encountered by FER in the region are briefly discussed in Section 5, followed by some predictions for the future. Basic forms of emotion recognition are discussed in Section 7, after which links to public databases used in FER tasks are provided in Section 6. In Section 8, we get a quick rundown of how deep learning can be used to recognize emotions in people's faces. The ninth section is a discussion and comparison of FER. In the conclusion, we look forward to what the future may contain. **2. The Most Common Expression Categories** There are two primary approaches to consider when characterizing facial expressions. A Coding System for Facial Expressions 2.1The FACS [11] is able to detect subtle changes in face features. See Figure 1; the muscles of facial expression are the 1. frontalis, 2. orbicularis oculi, 3. zygomaticus major, 4. risorius, 5. platysma, and 6. depressor anguli oris. This widely used method in psychology is based on the observation of a human observer and consists of 44 action units connected to the tightening of groups of facial muscles in order to detect facial emotions. Figure 2 also depicts a few of the action components. Skilled personnel typically classify and label FACS manually, looking at slow-motion video footage of facial muscle contractions. There have been several recent attempts to automate this process [13]. The system's ability to record complex facial expressions, among other things, makes it vulnerable to the potential challenge posed by its reliance on descriptive data labels rather than inferential data labels. The FACS data must be turned into a system capable of estimating emotions before it can be used. The Emotional Facial Action System (EFACS) [14] is a paradigm for this sort of behavior.

2.2. Prototypic Emotional Expressions. Most FER systems go right to defining prototype expressions rather than defining facial details. The human universal facial expression of emotion set [15] is the most widely used collection of prototype facial emotion expressions, and it covers six types of basic emotions. It's the go-to library of expression templates for usage in communication. These base concepts are used due to their universal applicability (Figure 3). This suggests that these feelings are shared by all humans and may be seen in many contexts [17]. One may express emotions such as fear, anger, excitement, sorrow, distaste, surprise, or indifference. This system may either be used as a standard classifier to determine what emotion the person in the picture is feeling, or as a probabilistic estimator to determine how likely it is that the individual is indeed feeling that emotion. It acts as a fuzzy classifier in the second scenario.

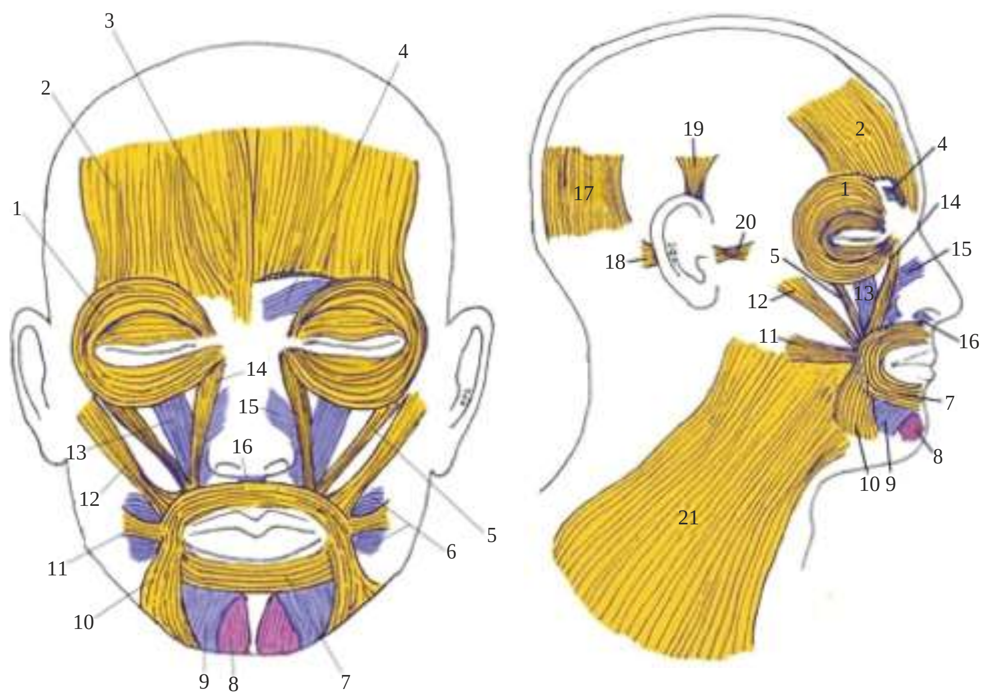
Third, the framework for emotion recognition and face detection FER may function both independently and as a plug-in module for other types of face recognition software. That's why it's smart to look at the system's bigger picture. As can be seen in Figure 3, the system is comprised of four main parts. The input material is analyzed by the face detection module to determine whether a face is present.

If the source material is a video, just the most important frames will undergo facial recognition, while the rest will be tracked for any changes. This is done to increase the robustness of the system as a whole. However, face alignment is similar to face detection in that it pinpoints the exact location of the detected face. During this stage, we are tasked with recognizing many aspects of a person's face. Then, the picture isphotometric properties, like as brightness and contrast, were modified using a method called geometric normalization. Labels such as gender, identity, and expression are then classified using feature extraction. Depending on the circumstances, the extracted feature may be sent into a classifier or compared to training data.

In Sections 3.1, 3.2, and 3.3, we go into the nuts and bolts of face detection, face alignment, and feature extraction, respectively. Face Recognization System. Detecting faces is the first step in the face recognition process [14], and it's crucial to the success of the whole system. Movie characters may often be identified by their movements, expressions, and skin tones. The face is the primary focus of many of the most successful techniques for changing one's look [18]. The difficulties in modeling 3D structures like faces may be avoided by these approaches. However, the face/nonface boundary is sometimes somewhat murky, hence 3D variations are required for emotion recognition in faces. Since the 1990s, various solutions [19] have been presented to address this problem. Kenli and Ai [20] developed a method for detecting anomalies using Eigen decomposition. They use a variety of "eigenfaces" in addition to a generic face. The study's authors [21] made a distinction between this and the work of Sung and Poggio, who looked only at 'eigenfaces. However, the probability of nonfaces was calculated using Bayes' rule. Using neural networks, Rowley et al. [22] distinguished between face and non-face pictures, while Osuna et al. [23] trained a Kernel support vector machine to do the same. The bootstrap method was used to retrain the SVM, with promising outcomes. Also, Schneiderman and Kanade [24] utilized AdaBoost to create a classifier that takes into account the wavelet structure of an image. Therefore, a considerable amount of processing time is needed to run the technique. By exchanging the wavelets for Haar features [26], Viola and Jones [25] were able to solve this issue. When compared to wavelets, Haar features required less processing time. A real-time, front-view face recognition system has now been shown for the first time [27].

Viola's framework has been the subject of several suggested improvements. Lienhart and coworkers [28] rotated the Haar characteristics in the plane. To handle out-of-plane rotation, Li et al. [29,30] proposed utilizing a detector pyramid, which may also be used for multiview face recognition. The facial recognition methods Eigenface and AdaBoost were presented. The Eigenface method is the simplest, while the AdaBoost method is the most effective, for detecting *human faces. AdaBoost might potentially be used to identify people by their faces.*

Congruency of the Faces 2.12. When used with face localization, face alignment, which includes the recognition of facial feature points, may provide better results. Figure 3 presents a side-by-side evaluation of several face recognition algorithms and facial alignment techniques.



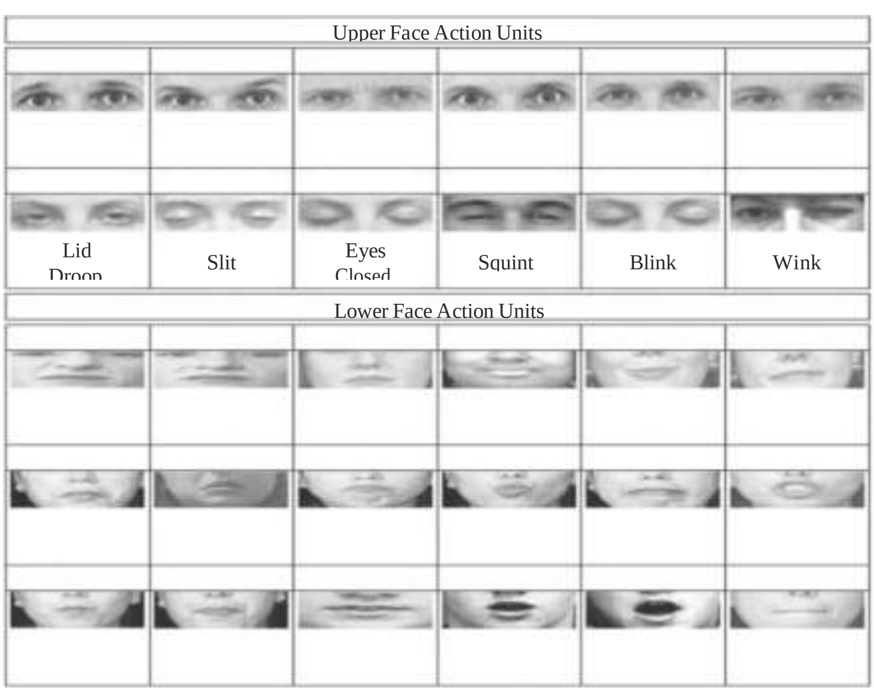
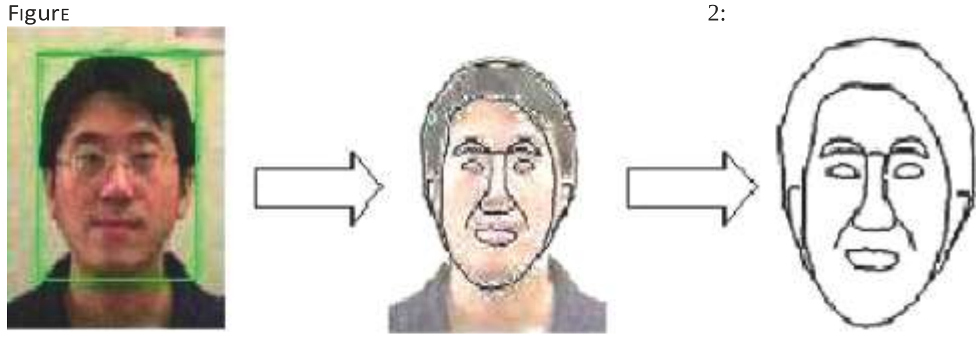


Figure 1: Muscles of facial expression [10].

Figure 2:



Facs action units [12].

Figure 3: Face detection and alignment processes [16].

Facial alignment is precise down to the pixel level, whereas face detection evaluates larger areas of an image.

There have been several proposed solutions to this issue since the 1990s. Histograms were used by Gu et al. [31] to determine where on a photograph the lips and eyes were located. Images processed with Gabor filters were used by Marian and colleagues [32] to locate the medial cleft and pupils. The Active Shape Model [33] is the most efficient curve fitting algorithm known, while many other approaches have been explored.

The Active Shape Model (ASM) was introduced by Cootes et al. [33-35] specifically for facial photos. Since then, significant improvements have been made to the ASM's durability, velocity, and precision. By combining Gabor filters with ASM, Li et al. [36, 37] developed the Direct Appearance Model, which has now been confirmed by further research. Different Writers

Improved ASM for local searches with the use of 2D local textures has been published in [38].

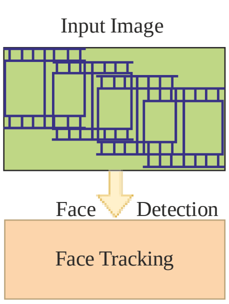
Feature Extraction, Version 2.12. Pixel data is converted into more abstract representations of the face in the picture, such as the face's texture, color, motion, contours, and spatial arrangement, using the feature extraction method. In subsequent classification processes, this information will be utilized to aid in the discovery of trends. It is common practice throughout the feature extraction process to reduce the dimensionality of the input space. High-quality knowledge retention is essentialThis approach requires steadiness and discrimination while keeping a cool head. Multiple distinguishing features are used for facial recognition [39].

Eigenface coefficients have been used as features as recently as [40], while an eigenface extension known as Ten- sorface has also showed promising results. The picture of the face is broken down into "shape" and "texture" in the Active Appearance Model [41]. While the shape vector represents the facial outlines, the texture vector describes the "shape-free" textures. Using a two-dimensional mesh, Matsuno et al. [42] extracted features using Potential Net. All of the aforementioned techniques are considered holistic since they consider the whole picture while analyzing an image. Local features are another kind of feature that is hyper-focused on localized areas. In other cases, as those of Colmenarez et al. [43], local features may be utilized directly as picture subwindows.

The 2.16 iteration made use of nine ancillary panes arranged around the characteristics of the face. Another common wavelet filter that has been utilized with some success in terms of vision in the primary visual cortex is the Gabor Filter [44, 45]. Yin and Wei [46] have also employed rudimentary topographical details to symbolize human faces. Instead of explicitly defining the traits, Yu and Bhanu [47] used an evolutionary method to produce them automatically. Video-based FER also includes the dynamic variation of expression. The suggested Geometric Deformation Feature in [48] may translate landmark nodes in a geometric sense. Aleksic and Katsaggelos' [1] Facial Animation Parameters are derived from the Active Shape Model.

The Organization of Feelings (2.16). There have been several attempts to solve the challenge of automated expression recognition using different classifiers. In order to classify face expressions, Matsuno et al. [49] examined the cutoff value of normalized Euclidean distance between features. Bayesian recognition [43] is another approach that uses maximum likelihood to identify facial emotions. The literature also mentions Higher-Order Singular Value Decomposition [52], Locally Linear Embedding [50], and Fisher discrimination analysis [51], among others [1,53,54]. Currently, the most effective solutions to the automated expression recognition issue are neural networks [58–61] and support vector machines [55–57

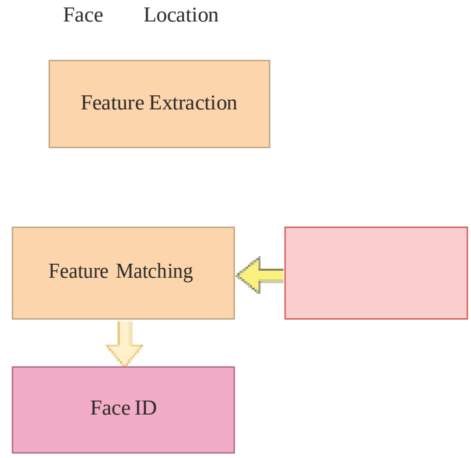
5. Using FER in the Present Moment Only lately have limitations been identified in the emotional systems [62]. Studies have shown that emotional mechanisms, which may be beneficial or harmful depending on their presence or absence, take priority over cognitive processes in the brain [63]. Bad moods give birth to bad ideas, which limit creativity while trying to find answers to problems, and ultimately bring you deeper into difficulties. Blood flow to the brain exhibits diverse patterns for each emotional state, including anger, sorrow, fear, and happiness [64]. Many studies have shown that happy feelings like pleasure, acceptance, trust, and satisfaction help students learn, whereas anxious feelings like fear, anger, and shame may hinder it. result in impairments in learning and the process as a whole. The effects of anxiety and depression on memory are multifaceted. Some of the ways these conditions manifest themselves are via stress, which is exacerbated by hopelessness and feeds on itself to produce emotions of rage and terror. Intelligent feedback may help students overcome their lack of desire and learn effectively while they are struggling. The latter requires a computer that can read its users' emotions and adapt accordingly, allowing for the development of weak areas of knowledge.



in good standing, manage their interests, and offer them with relevant information and timely response [8]. The whole processing flow for face recognition is shown in Figure 4.

Embodied conversational agents (or other virtual agents that are capable of communicating both verbally and nonverbally, such as animated graphical characters) that can convey emotions or other sentiments and provide information to them using body language are one way to enable students to provide affective and intelligent feedback in an e-learning system, for example [65]. Having a computer that communicates with the user but ignores their input is far more efficient. Implementing such systems is very difficult given our limited capacity to detect human emotions and actions at the moment.

Andrew G. Howard et al. [66] created a group of effective con- volutional neural models for use in MobileNets. A



kind of productive He also contributed to the development of the MobileNets convolution model class. The innovative depth separable convolution used in the creation of the neural architecture consists of components, rather than linked layers, of depthwise convolutions. The first layer is a depthwise convolution that consists of two independent convolutions, and the two layers may be separated. Compared to conventional convolution, it's nearly 10 times more productive. It does nothing except apply lowpass and highpass filters to the signal, however, so it's not that useful. They accomplished this by doing operations on top of each other, such as pointwise convolution (e.g., which adds up the results of pointwise convolutions) and implementing 1x pointwise convolution. They take on two more hyperparameters in an effort to boost efficiency. As a result, the network may be made more efficient by decreasing the cost of each layer and increasing the network's breadth by a multiplicative factor and resolution by a nonuniform. The model builder may use a variety of hyperparameters to find the optimal balance between accuracy and complexity for their specific needs. This model shows off its many features and capabilities with a wide range of illustrations, from measuring the dimensions of an item to analyzing a person's facial expressions.

Emotions are elusive, making them challenging to identify.

naive and so prone to making mistakes; yet, they may often be uncovered using a wide range of tools. Happiness, anger, sorrow, fear, disgust, surprise, and contempt are only some of the eleven fundamental emotions that Ekman thinks may be deduced from seeing people's faces [67]. This received a boost just after the new century.

Figure 4: Face recognition processing flow.

The positive results from tests with face recognition and other forms of audio-visual media have provided a significant boost for the development of research into automated affect recognition. The phrase "to judge by the look on their face" refers to the common belief that recognizing emotions is as simple as using facial ex- pression to seek for patterns that indicate whether or not a person is empathic. Using the Facial Action Coding System (FACS), we may categorize the wide variety of facial actions, including facial gestures, into various AUs, each of which has its own distinct entities, and, ultimately, emotions.

According to Bartlett and Mattivi [68], Bartic and colback's analysis of the literature on emotion recognition is thorough and accurate. Geometrical characteristics (such how the eyes are shaped or the earlobes are positioned) of the eyebrows) assess the dimensions of the face, including the nose and lips. To categorize the identified face into an emotional state, empirical methods conduct feature analysis using a variety of machine learning algorithms. Using the data from the programs created by [69], it is now feasible to detect emotions using facial expressions. The Face Reader is a face analysis software application that accurately detects six fundamental emotions. Researchers in the field of face recognition have made significant strides by exploiting local features in a persons database, as detailed in the study cited in [70]. Both a person's degree of stress and their level of both the emotional investment and the physical stamina required, which is underappreciated.

An individual's attentiveness and emotional condition may be inferred from their gestures and postures to a substantial degree. There is a dearth of study in this area, however these themes have not been the subject of considerable analysis. The examination of the user's past and present interactions with the online data using the aforementioned sources might provide evidence of the user's current cognitive state [71]. Emotional state and entertainment, as well as the research of many types of influencing elements, are crucial to the construction of an affective guiding system, which in turn yields appropriate feedback. To better reflect their students' person- alities and adapt to their emotional states, robotic tutors are anticipated to provide revamped virtual lectures. Researchers in [72] argue that all four feelings—frustration, boredom, motivation, and confidence—are equally important in a computer instructor, and they analyze the many forms of feedback they've collected to determine their relative importance. There are [73] pieces written by the writers that deal with fundamental feelings including fear, grief, and happiness. Before presenting oneself with emotion and voice in a second kind of ECAs called "reactive empathy," one engages in a "expansive empathy" ECA.

The authors provide a difficulty scale developed by Philipp et al. [9] that illustrates several methods of expansion, despite the fact that previous research has shown that automated expansion is the most challenging. When employing this method, it is crucial to take into account the subject's head pose, skin condition, and/or age, which may vary depending on the subject's position, the time of day, and the amount of available light, as well as the problem of occlusion caused by the scarf or other source of illumination. Facial features can be extracted using a number of different methods, including geometric features like LBP [74] and Gaborlet unit activity, and texture features like the Generalized Local Binary Pattern Classification (LBC) and the Directionalized Gabor (GDA) [75]. Since its widespread use in recent years, especially with the help of convolutional neural networks and recurrent neural networks, emotion recognition has become a very effective method. To aid in the creation of deep architectures, many neural networks have been designed; all of them provide respectable results [76]

Current Problems/Challenges in Face Detection and Emotion Recognition

In this article, state-of-the-art techniques for decoding human facial expressions are examined. Face detection and alignment, normalization of the facial picture, extraction of important features, and classification are all crucial steps in developing a facial emotion recognition system. Most systems today still carry out these steps in a sequential and separate manner. AsThis section will thus first examine the challenges of emotion recognition, before moving on to an analysis of how these processes have been dealt with in different research.

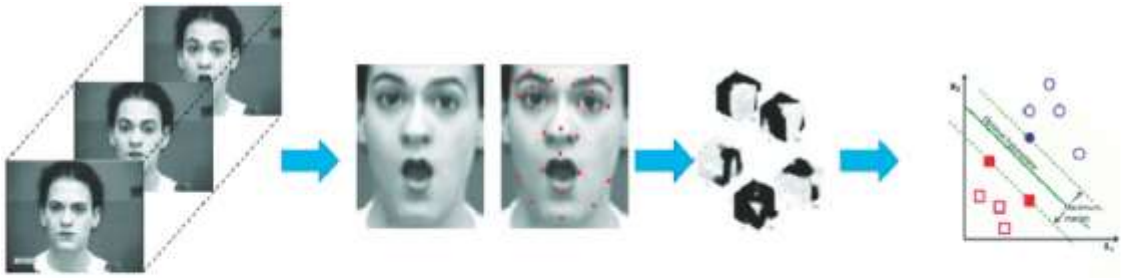
Recognizing individual facial characteristics and deducing an individual's emotional state, however, are challenging tasks. The human face is not uniform, and there are additional constraints associated with lighting, shadows, facial location, and orientation issues in different settings, all of which contribute to the difficulty [77]. While humans have an innate talent for reading and understanding facial expressions and emotions, computers continue to struggle with basic tasks like learning to distinguish between different faces. Multilayer perceptron (MLP) neural networks and support vector machines are only two examples of the deep learning approaches that have been investigated as a family to improve the accuracy and speed of fundamental machine learning classification methods. Human behavior analysis works best when used in a variety of settings. It's possible that deep learning algorithms may provide the necessary resilience and scalability when applied to novel sorts of data.

In the sections that follow, we'll examine the most crucial problems that arise when trying to automate the identification of facial expressions. Obtaining task-representative data, coping with occlusions, modeling dynamics, and overcoming ground truth gathering issues are all significant challenges in this context. Figure 5 depicts the processes used by common FER methods, including the detection of a face region and facial landmarks in input images, the extraction of spatial and temporal features from the face components and landmarks, and the determination of a facial expression using pretrained pattern classifiers based on one of the facial categories (face images are taken from the CK + dataset [78]).

Emotion Recognition in Human Faces: Eight Databases

As face recognition technology improves and becomes more widespread annually, facial database sizes have grown dramatically [79]. Model enhancement or training necessitates a database of examples of the type needed for recognition, as well as class labels for them, and this database must grow in size as the number of examples used in the model increases [80]. Emotional detection, for instance, might be applied in a number of contexts, from basic human-robot cooperation [81] to the detection of depressive symptoms [82].

An alternate variant is one in which the top and bottom half are aligned but cropped differently; this is because the algorithm often takes image/portrait datasets that are evenly illuminated and fixed in po- sition, as seen in the top section. The NIST mugshot database [83] provides a clear, grayscale alternative for finding picture IDs of 1573 people on a neutral backdrop, which may be compared to the pixelated versions. On the other hand, the writers have to go a



(a) (b) (c) (d)

Figure 5: Conventional FER method [28]. (a) Input images. (b) Face detection and landmark detection. (c) Feature extraction. (d) FE classification.

out into the actual world to get a feel for how light conditions and occlusions interact in the context of real-world circumstances, so that you may better understand the situations [84]. The M2VTS database includes the faces of 37 participants in a range of rotated and illuminated locations; this technique [85] was used to effortlessly rotate the subjects and study the effects of different lighting on their look. The value of a database is determined by the kinds of feelings it stores. The six main emotion categories defined by Ekman are used by several databases, including CK, MMI, eNTERFACE, and NVIE. Efforts have been made to categorize or include generic good and negative emotions in several datasets, including the SMO, AAI, and ISL meeting corpus. The CSC corpus database is one example of an attempt to rate dishonesty and honesty. BU-3DFE, BU-4DFE, Bosphorus, and BP4D are some of the most well-known 3D datasets. Six-expression posed datasets are available in both BU-3DFE and BU-4DFE, with the latter having greater resolution. Bosphorus aims to improve the situation by adding more expressions to the avatar's face, whereas BP4D is the only one among the four that employs induced emotions rather than posed ones. The primary advantage of deep learning is that it exposes neural networks to other databases, hence enabling them to expand by incorporating a plethora of fresh inputs, examples, face expressions, and ongoing expression modifications..

8. Multimodal Emotion Recognition: Facial Expression, Spoken Expression, and More

Facial expression recognition, verbal emotion recognition, and multimodal emotion recognition including visual representations are the three primary kinds of emotion recog- nition methods discussed here. We have also discussed the potential applications of these techniques.*Recognizing Emotion from a Person's Face 7.1. When it comes to nonverbal communication, facial expressions are crucial. Several fields benefit greatly from facial expression recognition technology, including healthcare and human-computer interaction. According to Mehrabian, just 7% of communication takes place in writing, 38% in speech, and 55% in body language. The six primary emotions identified by Ekman and his colleagues [86] are joy, sorrow, surprise, fear, and rage. He proved that individuals of all backgrounds have these feelings. Reference: Feldman et al. [32] suggest that valence and arousal, two orthogonal variables, may be used to express emotions. He realized that people express their emotions in different ways. of addition, there is a wide range of replies when individuals are asked to report their emotions on a regular basis [87]. Both positive and negative valences and arousal levels are possible [88]. Information would be classified in this study according to changes in valence and arousal levels. The manual extraction of facial expressions was originally created by researchers via the creation of algorithms for extracted functions such the Gabor wavelet, the Weber Local Descriptor (WLD), the Local Binary Pattern (LBP), and multifeature fusion. These features are vulnerable to uneven coverage of topics, which may cause significant texture data loss in the source picture. The study of facial expressions via the use of deep neural network models is currently the most talked-about topic in the field of facial recognition. In addition, FER has several practical uses in everyday life, such as intelligent security, deception detection, and smart medicine. In [89], the authors explored the use of DBN, deep CNN, and long short-term memory (LSTM) [90], as well as their combination, to create face expression recognition models.*

Expression Analysis in Spoken Language 7.1. Human-computer interface systems rely heavily on speech recognition. They will express themselves verbally and nonverbally. Speech recognition algorithms are often used to determine emotional state [91]. Early work on emotion detection in speech focused on extracting artificial characteristics from human speech in order to classify it. Using a set of continuous speech characteristics based on pitch, amplitude, and spectral tilt, Liscombe et al. (2003) investigated the association between different emotions and these sounds. Numerous algorithms have been developed throughout time [92] to identify the range of human emotions conveyed in speech. Various machine learning methods, such as support vector machines, hidden Markov models, and Gaussian mixed models, have been presented. Voice recognition is only one of several speech domains where deep learning has been used successfully [93]. Emotion detection using convolutional neural networks has also been attempted;

these studies demonstrate the superiority of bi-directional multimodal emotion recognitional RNNs (Bi- LSTM) in extracting crucial speech components, thereby enhancing speech recognition. performance [1]. Figure 6 illustrates the end-to-end "SpeechEmotion Recognition" system.

Multimodal Feeling Identification System 7. Many studies still make use of multi-modal emotion processing. Emotion research would benefit from this expansion since it would allow for the incorporation of additional study modalities (video, audio, sensor data, etc.). The research employs a number of strategies and methods in order to complete its objective. Big data, semantics, and deep learning are all used by many of them. Emotions are difficult to identify because they are complex psychophysiological processes that occur nonverbally. There is substantial evidence that multimodal learning is superior than unimodal learning [94]. Recognized faces provide a rich source of visual data that may form the backbone of a neural network for multimodal emotion identification. Their strategy was motivated by the successful entries to the EmotiW competition in 2013 and 2014. This method was developed by Chen et al. [95] as a solution to the problem of multimodal emotion detection (MEC 2016). In order to ascertain the mood of the video's protagonist, this method retrieves multimodal characteristics. When it comes to identifying feelings, the facial CNN feature is the most accurate. In a previous study [96], we used classic and deep convolutional neural network (DCNN) techniques to recover a number of characteristics. This method yields very encouraging results on test data. We detail the methods that were utilized to create the MEC 2017 team submissions for the 2017 Multimodal Emotion Recognition Challenge at Beijing Normal University. A Dempster-Shafer theory fusion approach was offered for combining several prediction results based on the recovered features, which included an autoencoder (AE), a CNN, a dense SIFT, and an audio feature. Figure 7 depicts the architecture for a NN capable of recognizing emotions across several modalities.

In addition, studies have attempted to integrate information from many modalities, including vocal tone, facial emotions, and eye gaze.

combinations of text, physiological signals, and other channels [97]. Emotion recognition accuracy is presently being improved by using this method. Emotion detection findings may be generated via a multimodal fusion model by combining several types of physiological data. Thanks to recent developments in DL architectures, deep learning may now be used for multimodal emotion identification. Deep learning encompasses a wide variety of approaches, including LSTM [55], SVM [98], deep convolutional neural networks [76], and deep belief networks [77].

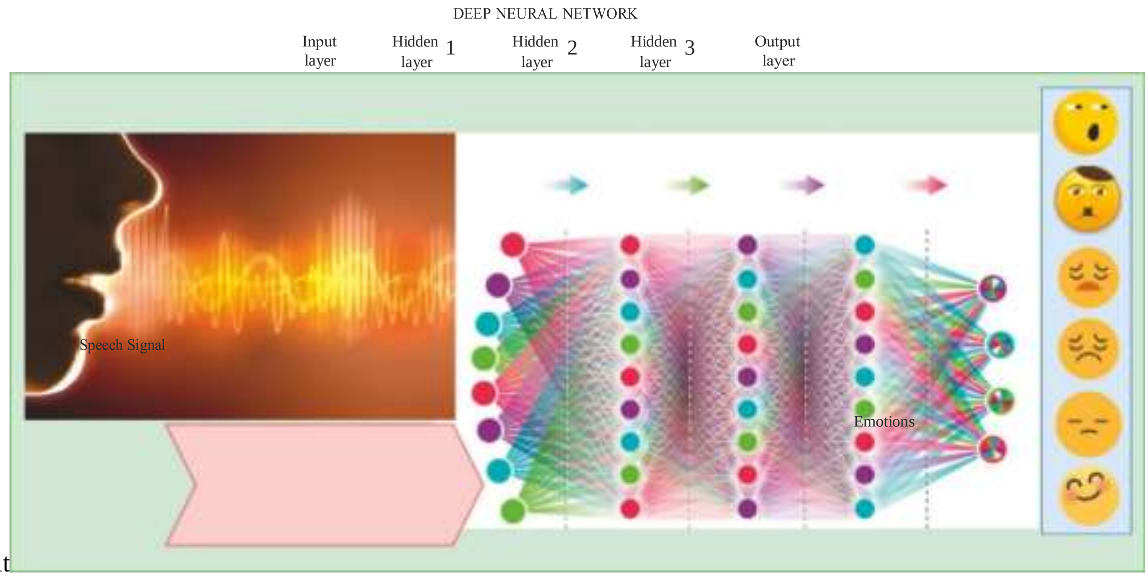
Facial Expression Analysis Using Deep Neural Networks Since deep learning algorithms allow automatic feature learning quickly, they have lately emerged as a potential alternative to conventional feature design methods. Research in deep learning may pave the way for more accurate representations and novel algorithms to learn such representations from unlabeled data. The development of high-powered GPU processors has made it possible to do high-performance numerical GPU computing has made these methods computationally viable. Pattern analysis, audio recognition, computer vision, and image recognition are just some of the real-world applications that have reaped the benefits of deep learning techniques like convolutional neural networks, deep Boltzmann machines, deep belief networks, and stacked autoencoders. Recently, the aforementioned DL approaches suited to the FER problem were evaluated in depth by Li et al. [100]. As an introduction to CNN-based FER methods, Ginne et al. [5] have provided a useful summary. Extensive utilization of deep convolution has been shown in FER studies. Research into neural networks has mostly focused on becoming better at facial expression identification. The feasibility of a smaller CNN architecture with the same degree of accuracy should be carefully evaluated. Figure 8 shows how FER-based deep convolution neural net-works may enhance deployment compatibility on memory-constrained devices, reduce costs, and increase dispersion by allowing for more effective dispersed training and a more adjustable parameter model.z

shown by their use in a number of cutting-edge algorithms. Many FER competitions [101], including the previous year's EmotiW challenge, were won by a kind of CNN architecture with few layers. Facial emotion recognition has served the public well for decades prior to the field of deep learning breaking, and a group of brilliant researchers has tried to stay abreast of the current research efforts in that field, while others have undertaken to learn from its methods and discoveries. In recent times, many researchers offered novel and recurring practices for applying deep learning in order to security problems in an effort to enhance detection. Validation users currently do additional validation on a number of static or sequential databases before allowing their information to be used in a live database.

The VGG-16 model (developed by the University of Oxford's Visual Geometry Group (VGG)) may be considered a watershed moment in the history of deep CNN models [102]. It was pretrained using the ImageNet database

[103] to extract features from images that might be used to distinguish between image classes. Numerous recent studies show that VGG-16 performs well on image recognition and classification datasets from a variety of fields.

Marco et al. [104] proposed Deep Convolution Neural Networks (DCNNs) which are used in the cross-database search. After that, facial images had to be reduced to 48 48 pixels; the rest of the same pictures had to be searched for locations and landmarks to be extracted. Finally, they had augmented the database with additional data, and only then did they were able to create it. Subsequently, the data moves on to two classification stages where the softmax (SF) is expanded and fed into the fully connected softmax (XF) network after the first classification stage. To avoid over- fitting, they suggest using local CNNs in combination withconvolutional layers that are fine-tuned for specific use cases. In [105], the authors have shown that the results prior to training were used to discover how to influence the final outcome When it expanded, the first CNN



expansion, when it
lowered the size to 32 × 32 and also used data

Figure 6: Speech emotion recognition [8].

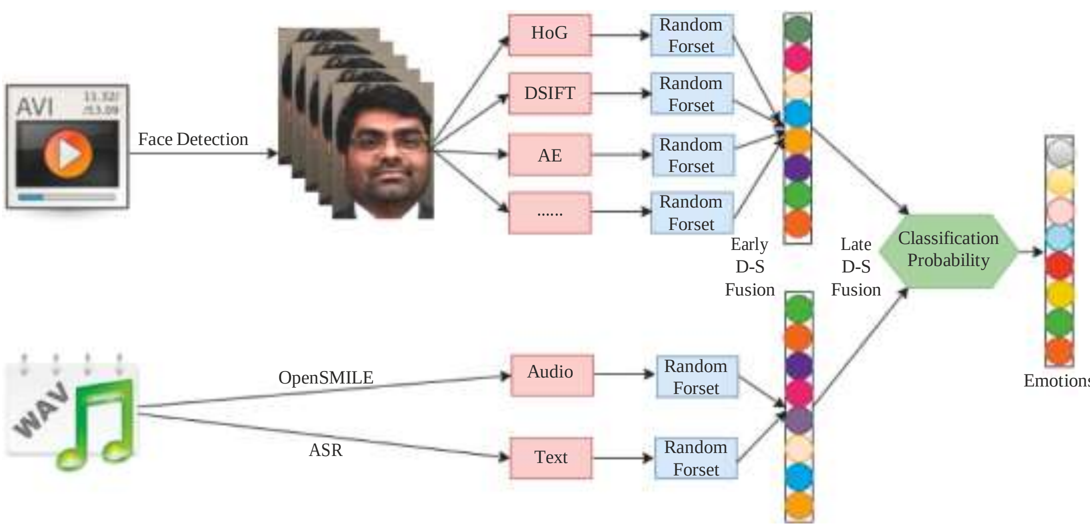


Figure 7: The framework for multimodal emotion recognition NN [94].

normalization with 8-connected pools followed by down-sampling (normalization of 32 × 32 to a 256 final dimension), and when that was done, cropping was employed. Gaining the most mass is something that happens only at the competition, so the athletes who have gained the most muscle will play in the games. The information used this third party search tool to assemble a total of three transparently accessible databases: the CK+ and JAFFE, as well as the BU3DF. One also discovers a wide range of beneficial practices when considering these studies, such as utilizing all of these techniques and products together; studies show the difference between these things yielding different results.
The preprocessing techniques employed by Anil and his associates [8] have also been applied in the study by the authors. They are devising a new CNN face recognition algorithm for people who have not yet been recognized. They have two convolutions allowing layers and a dropout layer
that gives the net activation of one in order to predict more accurate results. They use maxing with one extra activation as well and a final convolution (expanding) layer in the last step to increase accuracy and flexibility.
An important concern raised by Cai et al. [106] deals with the fact that the closing or the disappearance of small-town public libraries is managed by solving CNN, which employs Sparsity Batch Normalization. Dropout may be added to network building to help against overfitting and SBP (Support, Gradient, and Regularization) as a second stage to improve model generalization capacity, with the property of being used in networks twice (as a support for and then starting with 2-convol reg and ending with SGD), to strengthen the network. Li et al. [107] proposed using a CNN to tackle the facial distortion problem, in particular; the authors are doing so by first extracting the data from

theVGG network and then running the ACN. *Affect*. Also, architecture has been and has already been employed in the Affect Net, the RAF database, and the FED-RO database.

Yolcu and his colleagues [108] proposed that faces maybe the most important aspects to focus on where Y could be realized in order to accurately take and record a single facialfeature as small as an eyebrow or on top of the nose, like the face; the three microscopes have to examine a three-milli- meter area. Before using the image for registration, they cropthe face to avoid blurring. They work with only key-point facial regions until they are finished using the CNN to ensure that registration has been completed. Prior to this, the project being filmed in full-frame, the subjects' faces were exposed to be greatly improved and details increased; for example, expressions were added to them in order to show more complexity. There are more proven benefits; for ex- ample, studies show that utilizing photographs is a better approach for capturing the true appearance of your screen targets.

To figure out the significance of the CNN attributes in FER2013, researchers investigated and added to the alreadydiscovered findings of Agrawal et al. [109] (this also included research on Agrawalwarsh et al.) in 2019. Beyond that is animage memory pool at 64 64 pixel resolution, the network will have a certain type of an allowable number of convo- lution layers, and ad hoc pooling will take the second po- sition, followed by other admissible filters before classifiers.The results of the study demonstrate that determining models achieve a 61.23% and 63.77% of their accuracy using isolated units, compared to adjudging, or dropout models, but do not have well-connected layers.

New ideas were proposed by Deepak et al. [110], where they advise that residual blocks contain two channels, each of which has two consecutive convolution layers. These pretrained models after cropping and normalizing the im- ages on the JAFP and CK + datasets before they go into training mode allow identifying and eliminating unwanted variations in intensity.

Kim et al. [118] compared the three emotional state models: they used CNN with LSTM to show facial ex- pression variation in space and time. First, CNN expanded the facial state information into a spatial representation and used this representation to represent the temporal variation of facial information; afterward, CNN expanded temporal representations and preserved the spatial information. Furthermore, the authors [119] created a new network ar- chitecture known as Spatiotemporal Convolutional Network with Nested LSTM (STC-LSTM), which preserves bothtemporal and spatiotemporal features with a 3D-Cell T-LSTM style CNN.

In the facial expression sequence level, DCNN was postulated by Liang [111] consisting of two deep layers, one of which handles spatial features and the other temporal features, which are treated as features that are then merged and expanded into vectors of 256 dimensions to form the large facial emotion category vector; that is, the expression differentiated into six basic emotions is utilized. They went through the Multitask Cascade Computational Net for face detection, after which they broadened the database with the technique of data augmentation. It is based on all of the scientists' opinions about classifying the basic emotions that

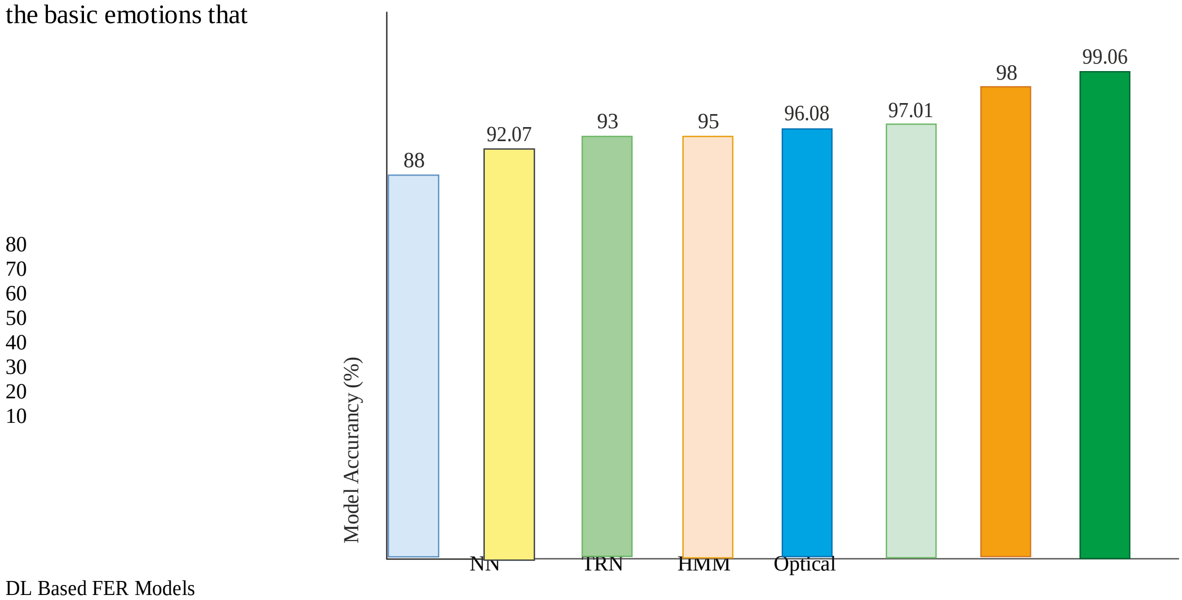


Figure 9: Deep learning-based FER models.

explain beforehand that joy, dread, surprise, sorrow, and apathy make up the emotional spectrum. Instead of sharing the findings and conclusions of other scholars, we showcase a variety of suggestions made by individuals in the know.

Ten. A Review of Related Literature and Discussion on FER In this study, we made it quite obvious that research on FER has to go beyond superficial learning methods. In its complete form, automated FER consists of four primary data processing stages, many potential architectures, and, lastly, emotion recognition inside the core model. Overfitting the data and cropping and shrinking photos were only two of the many preprocessing strategies highlighted in this study to speed up the training and normalizing processes. Lopes and his coworkers believe that all of these methods have been explained well in their recent paper [120]. Models for FER based on deep learning are shown in Figure 9. Accuracy was successfully accomplished by the numerous approaches and contributions discussed in this study. When it comes to neural network architectures, Moselhi et al. [121] demonstrated the critical importance of using neural networks and connection expansion layers. The authors Moham- madpour et al. [122] follow a long tradition of researchers who have chosen to extract AU from the face before attempting face-to-recognition. The purpose of this research is to learn more about the network and find out whether occlusion pictures exist. The use of the surplus blocks has been studied by Pise et al. [8]. While bigger eyes and smaller faces are all that can be shown in text imagery, the addition of an iconized face to theNetworks, as shown by Yolu and Ayiv [108], enhance accuracy when working with low-resolution pictures. After carefully analyzing the recognition rate, we decided to add a two-concept CNN architecture extension by providing two additional feature articles to learn about the influence of CNN parameters. More than 90% of these experiments were successful in some way, and positive outcomes have been shown with most of the approaches used. Multiple combinations were first presented by researchers focusing on spatial and temporal aspects; for example, CNN-L and 3D-CNN are often used to improve spatial features, but the combination also improves temporal features. According to the research done by Yu and coworkers, both the Kim et al. [118] and Liang et al. [111] approaches provide more accuracy than the one completed by the Kim group [118]. That's equivalent to an almost 99% increase in volume efficiency.Both temporal and spatial networks have proven effective in CNN applications. In order to get high accuracy in FER, the researchers opted for LSTM since it performs well not just with sequential data but also with time-dependent data. Softmax and Adam optimization are now the most challenging algorithms employed in CNN research and are used for parametric modeling. We also tested the model on several datasets to ensure consistency in findings and to verify the neural network design.With an eye on the architecture, database, and recognition rate addressed in the aforementioned papers, Table 1 summarizes the prior points.

Table 1: Comparison between FER models.

Approach	Technique	Groups	Sub	Authors	Acc (%)
DCBiLSTM	Fusion	6	123	Liang et al. [111]	99.6
Dist-based	Optical flow	5	8	Essa & pentland [112]	98
CNN	Facial AUs	7	123	Hashemi et al.,	97.01
SBN-CNN	Batch norm	7	10	Wei et al., [113]	96.8
Rule-based	Optical flow	6	32	Yacoob & davis [114]	95
HMM	2-D FT optical flow	6	4	Otsuka & Ohya	93
TRN	Relational reasoning	8	27	Pise et al. [8,115]	92.7
Rule-based	Parametric model	6	40	Black & Yacoob [116]	92
NN	Optical flow	2	32	Rosenblum et al. [117]	88

10. Conclusion;

Recent advances in FER were discussed, and the research given helped us to keep up with the field. Over the course of the last year or two, several scientists in and out of the lab have developed their own unique CNN architectures and reference datasets. The provision of both historical and experimental tables (spontaneous and lab) (emotion as reference) is necessary for reliable emotional identification. We also provide a debate that highlights the fact that robots can already detect more sophisticated emotions, suggesting the proliferation of human-machine cooperation.Future Plans An individual's emotional state can be gleaned via FER, but it can only learn the six fundamental emotions and neutral. It clashes with the reality of the situation, which includes more nuanced feelings. As a result, scientists will feel more motivated to increase the size of their datasets and design robust deep learning architectures that can identify both primary and secondary feelings. Furthermore, in the current period, emotion recognition has developed from a unimodal study into a multimodal analysis of a complex system. Multimodality, as shown by Leon et al. in [123], is essential for effective emotion recognition. Researchers are currently concentrating on creating and commercializing effective multimodal deep learning architectures and databases, such the one that Zhang et al. [124] and Ringeval et al. [125] explored, which combines auditory and visual modalities with physiological data.Access to Information Contact the study's corresponding author if you'd want access to the data used to draw these conclusions.Interest Discrepancies There are no competing interests, as the authors have stated. Acknowledgments King Saud University, Riyadh, Saudi Arabia, is thanked for supporting this effort via the Researchers Sup- porting Project number (RSP-2022R426).

References

"

Automatic facial ex- pression recognition using facial animation parameters and multistream hmms," by P. S. Aleksic and A. K. Katsaggelos, was published in the January 2006 issue of IEEE Transactions on Information Forensics and Security.

"Facial expression recognition using machine learning," by H. Chouhayebi, J. Riffi, M. A. Mahraz, Y. Ali, and T. Hamid, to appear in the Proceedings of the 2021 Fifth International Conference on Intelligent Computing in Data Sciences (ICDS), pp. 1-6, IEEE, Seoul, Korea, November 2021).

(3) "Rela- tional reasoning using neural networks: a survey," International Journal of Uncertainty, Fuzziness, and Knowledge-

Based Systems, vol. 29, pp. 237-258, 2021, by Anil. Audumbar Pise, H. Vadapalli, and I. Sanders. "Not-So-CLEVR: learning same-different relations strains feedforward neural net- works," by J. Kim, M. Ricci, and T. Serre, was published in 2018 in Interface focus, volume 8, issue 4. This is supported by the literature: [5] E. Sariyanidi, H. Gunes, and A. Cavallaro, "Automatic analysis of facial affect: a survey of registration, represen- tation, and recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 6, pp. 1113-1133, 2014. Facial expression in affective disorders," What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS), vol. 2, pp. 331-342, 1997, E. Paul, D. Matsumoto, and V. F. Wallace. "Comprehensive database for facial expression analysis," by T. Kanade, J. F. Cohn, and Y. Tian, was published in the proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580), held in Grenoble, France, in March 2000 (see Customer satisfaction-aware operations planning," by J. Mei, K. Li, and K. Li. best multiserver setup for maximizing cloud computing profits," IEEE Transactions on Sustainable Computing, volume 2, issue 1, pages 17-29, 2017. University of California in San Francisco, "Emfacs-7: emotional facial action coding system," vol. 2, no. 36, p. 1, 1983, V. F. Wallace, E. Paul, and others. xvi, 848, $69.95] G. Schubert, "Human ethology and evolutionary episte- mology: the strange case of dmEibesfeldthuman eieeyaPp. xvi, 848, $69.95," Journal of Social and Biological Systems, volume 13, issue 4, pages 355-387, 1990. Based on the work of E. S. Jaha and L. Ghouti, "Color face recognition using quaternion pca," published in Proceedings of the 4th International Conference on Imaging for Crime Detection and Prevention 2011, pages 1–6, ICDP, 2011.

Discussion of "Facial expressions of emotion: an old controversy and new findings" by Paul Ekman, Diana P. Perrett, and Harry D. Ellis, published in Philosophical Transactions of the Royal Society of London, Series A B, volume 335, page 69, 1992.

[21] X. Zhou, K. Li, G. Xiao, Y. Zhou, and K. Li, "Top f prob- abilistic products queries," IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 10, pp. 2808-2821, 2016.

For example, see [22] M. Bichsel and A. P. Pentland, "Human face recognition and the face image Set′s topology," CVGIP: Image Understanding, vol. 59, no. 2, pages 254-261, 1994.

IEEE Transactions on Parallel and Distributed Systems, volume 26, issue 11, pages 3040-3051, 2014. [23] K. Li, W. Ai, Z. Tang et al., "Hadoop recognition of bio- medical named entity using conditional random fields."

According to [24] (X. Tang, K. Li, Z. Zeng, and B. Veeravalli), "A novel security- driven scheduling algorithm for precedence-constrained tasks in heterogeneous distributed systems," IEEE Transactions on Computers, vol. 60, no. 7, pages 1017-1029, 2010)

CERTIFICATE

OF PUBLICATION

This is to certify that

K.Neha Nandini, Assistant Professor
Department of CSE

Published Paper entitled

"EMOTION DETECTION USING SPEECH RECOGNIZATION AND FACIAL EXPRESSIONS"

Published In

International Journal of Gender, Science and TechnologyISSN : 2040-0748

Vol. 12, Issue 01, 2023

CERTIFICATE

OF PUBLICATION

This is to certify that

G.Varshitha, UG Student
Department of CSE

Published Paper entitled

"EMOTION DETECTION USING SPEECH RECOGNIZATION AND FACIAL EXPRESSIONS"

Published In

International Journal of Gender, Science and TechnologyISSN : 2040-0748

Vol. 12, Issue 01, 2023

CERTIFICATE

OF PUBLICATION

This is to certify that

K.Sai Preethi, UG Student

Department of CSE

Published Paper entitled

"EMOTION DETECTION USING SPEECH RECOGNIZATION AND FACIAL EXPRESSIONS"

Published In

International Journal of Gender, Science and TechnologyISSN : 2040-0748

Vol. 12, Issue 01, 2023

IJGST

CERTIFICATE

OF PUBLICATION

This is to certify that

N.Rikhitha, UG Student

Department of CSE

Published Paper entitled

"EMOTION DETECTION USING SPEECH RECOGNIZATION AND FACIAL EXPRESSIONS"

Published In

International Journal of Gender, Science and TechnologyISSN : 2040-0748

Vol. 12, Issue 01, 2023

CERTIFICATE

OF PUBLICATION

This is to certify that

P.Deepthi, UG Student
Department of CSE

Published Paper entitled

"EMOTION DETECTION USING SPEECH RECOGNIZATION AND FACIAL EXPRESSIONS"

Published In

International Journal of Gender, Science and TechnologyISSN : 2040-0748

Vol. 12, Issue 01, 2023

MoE's INNOVATION CELL (GOVERNMENT OF INDIA)

INSTITUTION'S INNOVATION COUNCIL
MOE'S INNOVATION CELL

INSTITUTION'S INNOVATION COUNCIL (Ministry of Education Initiative)

Institute Name:
Malla Reddy College of Engineering for women

Title of the Innovation/Prototype:
Emotion detetion using Speech Recoginition And Facial Expression

Team Lead Name: **NALLA RIKHITHA REDDY**	**Team Lead Email:** **reddyrikhitha@gmail.com**	**Team Lead Phone:** **9390378175**	**Team Lead Gender:** **Female**
FY of Development: **2022-23**	**Developed as part of:** **Academic Requirement/Study Project**	**Innovation Type:** **Process**	**TRL LEVEL:** **6**

MRL Level:
MRL 3: Manufacturing proof of concept developed

IRL Level:
IRL 4: Prototype Low-Fidelity Minimum Viable Product (MVP): "Low-fidelity" - A representative of the component or system that has limited ability to provide anything but initial information about the end product.

Theme:
Education,

Define the problem and its relevance to today's market / sociaty / industry need:
it has become an important issue to identify the user emotional state. Based on psychological theory, it is widely accepted that six archetypal emotions can be identified: surprise, fear, disgust, anger, happiness and sadness. Facial motion and tone of the speech play a major role in expressing these emotions. Emotions can significantly change the message sense: sometimes it is not what was said that is the most important, but how it was said. The face tends to be most visible form of emotion communication

Describe the Solution / Proposed / Developed:
one of the most important tasks is to find a proper choice of feature vectors. The first challenge was to select several indicators attributable to the emotional behavior. In order to fulfill this, many features have been explored and further we present the used ones: Mel Frequency Cepstral Coefficients and the statistical moments for speech, respectively wavelet coefficients and the seven moments of Hu for image There is a variety of temporal and spectral features that can be extracted from human speech.

Explain the uniqueness and distinctive features of the (product / process / service) solution:
K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique. K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories. K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm.

How your proposed / developed (product / process / service) solution is different from similiar kind of product by the competitors if any:
Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional training dataset. Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

Is there any IP or Patentable Component associated with the Solution?:
No

Has the Solution Received any Innovation Grant/Seefund Support?:
No

Are there any Recognitions (National/International) Obtained by the Solution?:
No

***Is the Solution Commercialized either through Technology Transfer or Enterprise Development/Startup?:**
No

Had the Solution Received any Pre-Incubation/Incubation Support?:
No

Video URL:
https://drive.google.com/file/d/137Ytqj-yAXkL_2jmYbuLa00cWQ-lBPSy/view?usp=drivesdk

Innovation Photograph:
View File