Voice Assistant Using Face Authentication

Vishnusai Bhonsle
Computer Science And Engineering (student)
REVA University
Banglore,INDIA
vishnusai.bhonsle@gmail.com

Deeksha Chilukuri
Computer Science And Engineering (student)
REVA University
Banglore,INDIA
chilukuri.deeksha@gmail.com

Bhoomika.N
Computer Science And Engineering (student)
REVA University
Banglore,INDIA
bhoomikanag@gmail.com

Chirag Girish Purohit
Computer Science And Engineering (student)
REVA University
Banglore,INDIA
2000chiragpurohit@gmail.com

sailaja Thota
Computer Science And Engineering(Assistant Professor)
REVA University
Banglore,INDIA
sailaja.thota@reva.edu.in

***Abstract*—Voice assist is becoming increasingly popular. The most popular voice assistants are Google Assistant, Amazon's Alexa, Apple's Siri, and Microsoft's Cortana. Voice virtual assistants were becoming increasingly popular, particularly among younger people. According to the report, these voice assistants have a history of giving unwanted access to the device and inflicting damage to the owners as a result, software security is crucial. We developed facial and voice recognition technologies to handle security problems and unauthorized access, and they are useful in identifying unwanted trespassers and suspects.**

Keywords-Voice Assistant, Face Authentication, CNN, NLP,SVM, Anti-face spoofing.

I. INTRODUCTION

Voice assistant technology, which was once thought to be science fiction, is now a reality. From an academic perspective, these voice assistants are also known as Intelligent Personal Assistants and Speech-based Natural User Interfaces (NUI).Using voice recognition, language processing algorithms, and voice synthesis, the voice assistant listens to specific voice commands and returns relevant information or performs specific functions as requested by the user. By listening for and filtering out specific keywords, background noise, voice assistants can revert back relevant information.

Artificial intelligence and voice recognition are the technologies behind the voice assistant, which measure accurately and efficiently to deliver the result that the user is looking for in response. Because of its flexibility and extensibility, there are several voice assistants that specialise in a specific feature set, while others prefer to be open-ended and assist with almost any situation at hand. Users may ask their assistant's questions, control home automation devices and media playback, and manage other basic chores like email, task reminders, and calendars via voice commands.

There are many voice assistants available in today's world, and due to their popularity, they have a large user base. However, there is no guarantee of security for the voice assistant application because voice can be tampered with and modified, resulting in unauthorised access. So, in order to protect our voice assistant application, we're putting a face recognition system in place.

II. LITERATURE SURVEY

The methods used in the development of a virtual assistant differ between the various products available on the market in reference[1]. One product may be better in terms of voice synthesis quality, while another may provide more accurate results. As voice assistants become more prevalent in our daily lives, we may find ourselves having to repeat ourselves from time to time in order for the VA to understand our commands.

Customers can take control of their voice interactions when they interact with a virtual assistant (VA) who combines functional intelligence, sincerity, and creativity. Voice assistants are becoming increasingly popular. However, as reference[2] shows, security issues in VA, such as unauthorised access, may be accepted at some times.

[3] Face recognition automatically detects the human face in live video images taken with a video camera. Face recognition is the best computer vision, but poor image capture can make it difficult to identify the image.

In reference [4] shows End-to-end learning for the task using a convolutional neural network is progress in this area (CNN). This work is primarily concerned with deep architectures for face recognition. Face recognition can analyse images and videos from a variety of sources. Concerns about data privacy with facial recognition make it vulnerable to hackers.

In reference [5] derives Despite decades of effort, human activity recognition remains an immature technology that has drawn a large number of people interested in computer vision. In this paper, a system framework for recognising multiple types of activities from videos using an SVM is presented. It is more efficient in high-dimensional spaces and makes good use of memory. It does not work well when target classes overlap, and it begins to lag.

A binary tree based SVM multi-class classifier. The framework is made up of three functionally linked modules: shows in reference [6], (a) detecting and locating people using a non-parameter background subtraction approach. (b) extracting various features. (c) identifying people's activities using an SVM multi-class classifier whose structure is determined by a clustering process.

This method is best suited for classes with distinct separation margins; it also works well when total dimensions exceed sample dimensions. If the number of features for each data point exceeds the number of training data samples, the SVM will underperform.

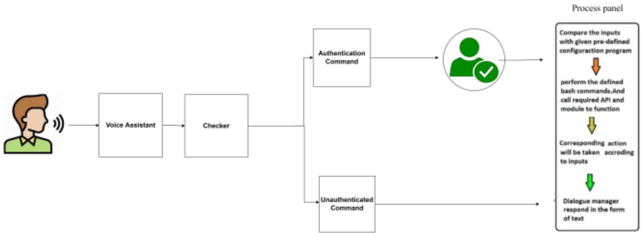
Natural Language Processing (NLP)is defined from reference [7]as the study of language issues in human-to-human and human-to-machine communication. We get an exact response in a matter of seconds because NLP is very time efficient. The NLP system is designed for a single, specific task.

In this case, [8] This technical briefing gives an overview of NLP tasks, available techniques, supporting tools, and NLP technologies. The amount of relevant information provided by the NLP system increases the accuracy of the answer. NLP is unable to adapt to new domains due to its limited functions.

Virtual assistance in the market is lacking with the privacy issues and following command to execute process where the command will perform without any authorization of user.

III. METHODOLOGY

In this paper, we propose a light-weight voice Assistant that uses face authentication to process requests and responds to the user with a synthesised voice and the required action. the voice assistant in our paper was implemented with an adaptive search engine where users can search things not only on the internet but also on a specific platform. we built voice assistance such that to support multiple platform devices like house automation tools, tv, computers, etc. the assistance is equipped with smart face authentication which help's the software to support security aspects. the command checker of our software will verify the input into authenticated or unauthenticated. If the user command is not authenticated, respond with a synthesised voice and the required action. However, if the user command is authenticated, the user must provide face authentication in order to complete the required action. Figure 1 depicts the high-level architecture of voice assistance.



Process panel
Voice Assistant
Checker
Authentication Command
Unauthenticated Command
Compare the inputs with given pre-defined configuraction program
perform the defined bash commands.And call required API and module to function
Corresponding action will be taken accroding to inputs
Dialogue manager respond in the form of text

Fig 1: Architecture of voice assistant using face authentication

The architecture is divided into two-parts

1) voice assistant: Voice assistants can return relevant information after listening for specific keywords and filtering out background noise.

The architecture was divided into three parts:

speech Recognition: The voice assistant should be able to understand and respond to the vocal request. There are numerous speech-to-text converter APIs on the market. we used to google API to convert speech to text. the text will be processed to the command checker module.

command checker module: The command checker module will check the processed text received from the speech recognition module and verify the command is authorized or unauthorized and execute to process panel with or without authentication depending on the command panel process:

The process panel will maintain a knowledge base of prepared commands that are linked to the relevant APIs for the required activity

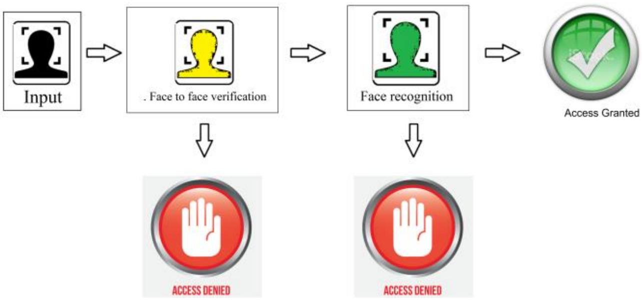


Fig 2: Functional Architecture of Face authentication

2) Face authentication: The face authentication system automatically identifies faces in images and videos from a live stream. It is divided into two modes:

1. Face to face verification (or authentication)

2. Face recognition (or recognition)

Face Verification: the module was designed to counter the face spoofing attacks using dual-stream convolutional neural networks. this module will be able to determine the input face received from Livestream is real or fake.

Face identification: the module is used to compare the face received from the face verification module with the stored in our database and if it matches more than 80% it will allow user to get access

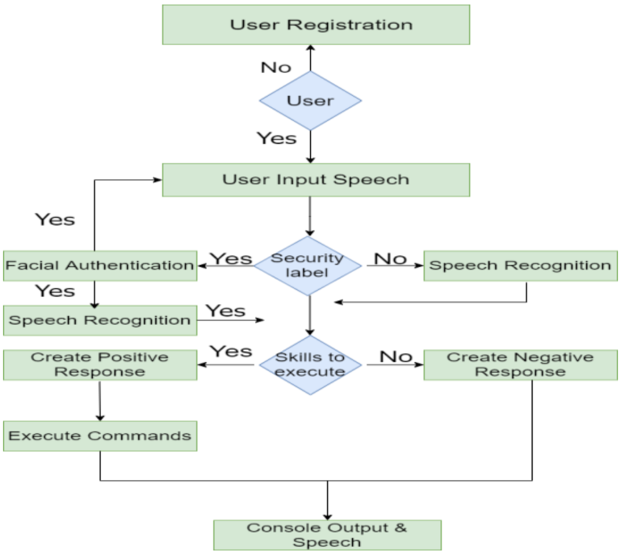


Fig 3: Functional Architecture of voice assistant using face   authentication

IV. IMPLEMENTATION

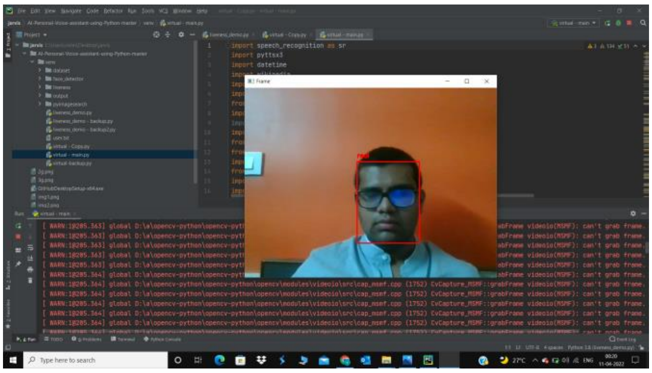
Face Authentication Implementation

In our system, we designed the authentication process such that whenever the user gives the input to the system it will check with the face verification process, and then if the process is verified then it will be checked with the face identification module for 15 times and then it will be looped with face verification for 5 times shown in the Fig 2. the above-stated process will help the software to prevent unauthorized access and to improvise the accurate verification process in the system.

Architecture implementation

In the implementation of our software, we designed the system such that the user needs to register his/her face id with the username shown in Fig 3. After registration, the system will save the face IDs and usernames in the database. the authorized user can be able to access all the features of the system with full potential.

V. RESULTS

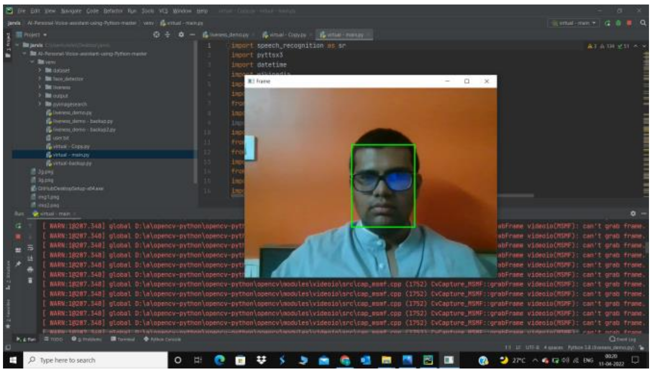


```
import speech_recognition as sr
import pyttsx3
import datetime
```

```
[ WARN:1@285.364] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
[ WARN:1@285.364] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
[ WARN:1@285.364] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
[ WARN:1@285.364] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
[ WARN:1@285.364] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
```

Fig 4: Result image.

we showed that when the user gave authorized command the software was able to authorize the user by undertaking the process of face verification

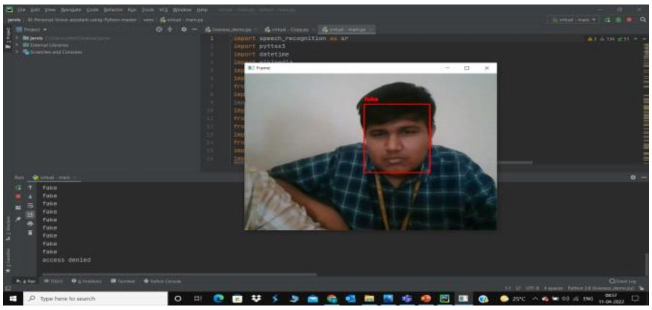


```
import speech_recognition as sr
import pyttsx3
import datetime
```

```
[ WARN:1@287.348] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
[ WARN:1@287.348] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
[ WARN:1@287.348] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
[ WARN:1@287.348] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
[ WARN:1@287.348] global D:\a\opencv-python\opencv-python\opencv\modules\videoio\src\cap_msmf.cpp (1752) CvCapture_MSMF::grabFrame videoio(MSMF): can't grab frame.
```

Fig 5: Result image.

we showed that when the software was successfully able to complete the process of face verification it will undergo the process of face recognition.

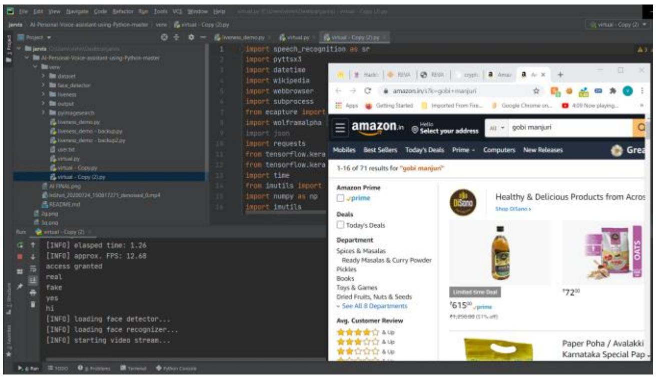


```
import speech_recognition as sr
import pyttsx3
import datetime
```

```
fake
fake
fake
fake
fake
fake
fake
fake
fake
access denied
```

Fig 6: Result image.

we showed that if the user was unable to satisfy face verification or face recognition process. Then the software won't let the user to have access



```
import speech_recognition as sr
import pyttsx3
import datetime
import wikipedia
import webbrowser
import subprocess
from ecapture import
import wolframalpha
import json
import requests
from tensorflow.kera
from tensorflow.kera
import time
from imutils import
import numpy as np
import imutils
```

```
[INFO] elasped time: 1.26
[INFO] approx. FPS: 12.68
access granted
real
fake
yes
hi
[INFO] loading face detector...
[INFO] loading face recognizer...
[INFO] starting video stream...
```

Fig 7: Result image.

we showed that if the user was able to satisfy both face verification and face recognition process. Then the software will grant access to user and execute the given task

VI. CONCLUSION

In our paper, we concluded that there is so much lack of privacy issues using voice assistance in day-to-day life. The voice-based assistants available in the market are not having an interface for face authentication or a security system to authorize the user to prevent unauthorized access. So, we focused to build a voice assistant application that can safeguard itself from unauthorized access by using face authentication where the user needs to pass a two-level authentication process. Whenever a user passes a command to voice assistance it will check the input of voice command, then categorize it into unauthorized or authorized. If the command is unauthorized (for accessing sensitive or personal information) then the application will implement a face verification and face recognition process by taking the user's face to decide whether the user is authorized or unauthorized. If the user is authorized, it will execute the user command if not it won't execute the user command.

REFERENCES

[1] Polyakov, E. V., et al. "Investigation and development of the intelligent voice assistant for the Internet of Things using machine learning." *2018 Moscow Workshop on Electronic and Networking Technologies (MWENT)*. IEEE, 2018.

[2] Poushneh, Atieh. "Humanizing voice assistant: The impact of voice assistant personality on consumers’ attitudes and behaviors." *Journal of Retailing and Consumer Services* 58 (2021): 102283.

[3] Turk, Matthew A., and Alex P. Pentland. "Face recognition using eigenfaces." *Proceedings. 1991 IEEE computer society conference on computer vision and pattern recognition*. IEEE Computer Society, 1991.

[4] Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition." (2015).

[5] Zhang, Hao, et al. "SVM-KNN: Discriminative nearest neighbor classification for visual category recognition." *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*. Vol. 2. IEEE, 2006.

[6] Qian, Huimin, et al. "Recognition of human activities using SVM multi-class classifier." *Pattern Recognition Letters* 31.2 (2010): 100-111.

[7] Jiang, Kai, and Xi Lu. "Natural Language Processing and Its Applications in Machine Translation: A Diachronic Review." *2020 IEEE 3rd International Conference of Safe Production and Informatization (IICSPI)*. IEEE, 2020.

[8] Ferrari, Alessio, Liping Zhao, and Waad Alhoshan. "NLP for Requirements Engineering: Tasks, Techniques, Tools, and Technologies." *2021 IEEE/ACM 43rd International Conference on Software Engineering: Companion Proceedings (ICSE-Companion)*. IEEE, 2021.

[9] Lewis, Patrick, et al. "Retrieval-augmented generation for knowledge-intensive nlp tasks." Advances in Neural Information Processing Systems 33 (2020): 9459-9474.

[10] Alshemali, Basemah, and Jugal Kalita. "Improving the reliability of deep neural networks in NLP: A review." Knowledge-Based Systems 191 (2020): 105210.