



ARYA: A PROGENITURE INTELLIGENT ASSISTANT FOR SUPERIOR USER EXPERIENCE

Mr. Maheshwar M

MCA Scholar, School of CS & IT, Dept of MCA, Jain (Deemed-to-be) University, Bangalore

Ms. Maitrayee Mahanta

MCA Scholar, School of CS & IT, Dept of MCA, Jain (Deemed-to-be) University, Bangalore

Mr. P R Kuber Gupta

MCA Scholar, School of CS & IT, Dept of MCA, Jain (Deemed-to-be) University, Bangalore

Article DOI: <https://doi.org/10.36713/epra4335>

ABSTRACT

The model that is proposed in this article suggests ways to shorten the bridge between human and computer with less human intervention by using inclining concepts such as Machine Learning and Artificial Intelligence. The voice assistants that exists currently in market is capable to perform basic tasks whereas, ARYA on the other hand is a type of voice assistant such as google assistant, cortana etc. which has made it easier for the end users to perform and automate various tasks just using their voice. It has multiple advantages over other voice assistants such as face recognition, cross platform assistance and automate document modifications. These features enhance the Access Control System over the usage of voice assistant and also reduces time and complexity of doing many tasks manually.

KEYWORDS-machine learning, artificial intelligence, cross platform, face recognition.

I. INTRODUCTION

Voice Assistants (VA) or Personal Assistants have brought different user experiences to access gadgets. The VA basically functions based on the guidelines/commands provided by clients or end users through voice. VA is incorporated in almost every digital gadget like smart phones, laptops etc. Some of the examples of Voice Assistants are Cortana for Microsoft Windows, Google Assistant for android smart phones, Siri for iOS and Mac Book etc.

Usage of voice assistant has rapidly increased and this has opened the room for inclusion of many other features than the legacy assistants. Arya is one such VA that is developed to overcome the limitations that are present in other assistants.

Existing VAs such as Google Assistant and Siri are restricted to only one platform such as Android and iOS respectively as they are written in

programming languages such as C, C++ and Kotlin which are more complex to run in multi-platform. This makes them platform dependent and hence are limited to only certain community of users who are using that specific platform. ARYA, on the other hand is completely built in Python which generates python byte code internally from the source code that we write, which is interpreted by a virtual machine known as the Python Virtual Machine or PVM. Hence, we can run the same python code in Mac which was originally written in some other platform such as Windows with minimal complexity and easily.

Another feature that makes ARYA distinct from already present VAs is that it is integrated with a face recognition system, built using the Machine Learning technology with the help of OpenCV. This new feature that has been added to ARYA intensifies

the security of the system. It detects the face that is captured using the camera, from which different patterns are generated for each picture it takes, which is recognized and saved using various Machine Learning Algorithms. Thus, using ARYA we get a better Access Control System, which will take commands from a user only after it recognizes / identifies the person using the already generated patterns of existing users. After the face verification, the VA will further perform various tasks based on the user's voice command.

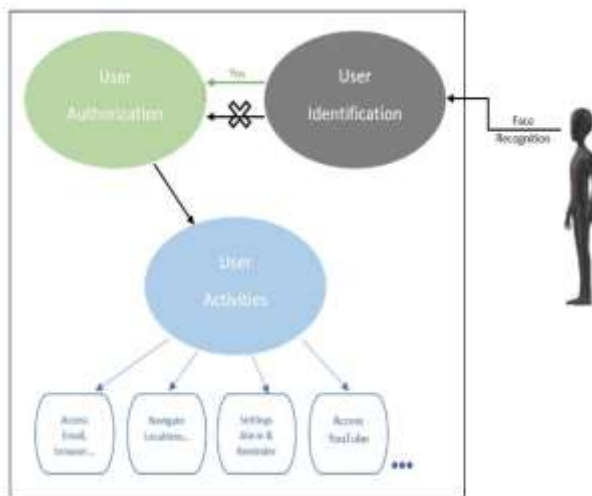


Fig 1. System model

In addition to all the features mentioned above, ARYA is also able to perform basic word operations such as creation of any word document, saving the document and can perform basic edit operations such as editing of headings, insert page breaks, insert images etc. which makes the job of document modification more easier for users.

II. LITERATURE REVIEW

[1] In this paper author quoted that Next-gen voice assistants are capable enough to recognize input in various means and process the output in the corresponding manner. Some of such input modes are speech recognition, gesture recognition. These VA's can be used in enhancement of user experience in different fields like education assistance, home automation, security control etc.

[2] In this paper author concluded researched and concluded that, Voice assistants are capable enough to perform various tasks that user instructs and commands are provided through voice. Not everyone's accent is same and different region people speak different accents of English. However, there is no much difference recorded while performing instruction test that has been done between primary and secondary English speakers.

[3] The author in this paper proposed a method on how technology can improve the mental health of elderly people. Social isolation has been a major cause for determination of life of elderly people. To minimize it, technology such as intelligent personal assistants have been introduced which can collect user's data and interact with them by using natural human voice. Hence, improving mental health condition of elderly people.

[4] Author in this paper performed thesis on how voice assistants have changed the way people interact with mobile phones or other electronic devices. Multiple assistants have been developed with best features like Google, Siri, Alexa but the choice of VAs differs with people. Survey has been conducted to determine the factor on which the choice of people for VAs depends. It has been seen that, among all the three major factors: privacy, performance and price, privacy has been the first priority of people, after which comes the performance of the VA and then the price.

[5] In this paper author stated that voice assistants interact with users based on the services that they are trained for. Using big data for training the assistant increases the accuracy in the results. The more the results are accurate, the more the users access the voice assistant.

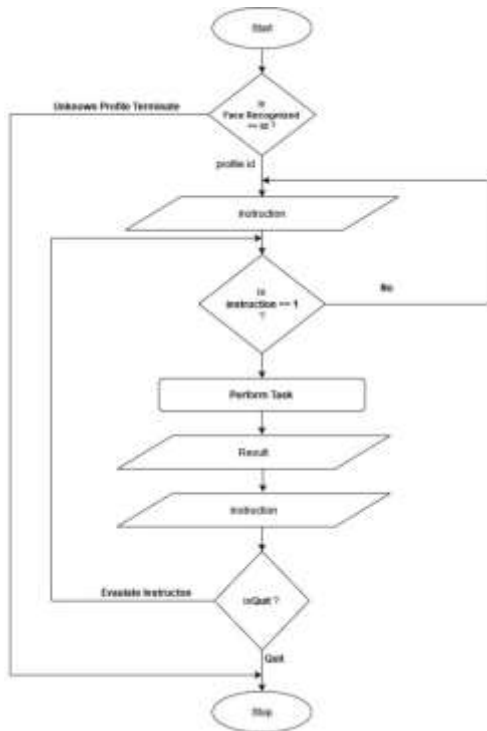
[6] In this paper author made a conclusion that privacy and performance play a vital role in the functionality of any voice assistant. Different aspects make people to access different virtual assistants. Virtual assistants have the capability to track human behavior in different environments.

III. PROBLEM STATEMENT

There are numerous voice assistants available in market which solely focuses on basic command execution that the user instructs and these are platform dependent as well. Which makes these VA's operable only by certain community of people.

IV. SOLUTION

The proposed method overcomes the snags that exists in the present-day VA's. some of the supplementary tasks that Arya performs when compared to other VA's are basic document management (creation, deletion, adding text, applying page breaks, imports images), call contacts using Your Phone application, cross platform support, face recognition for access control. The frame work that's included in the model is python and cascade files from open CV for system logic.



V. IMPLEMENTATION

5.1 Working

On detecting the voice our proposed VA ie ARYA will try to capture and recognize the face that it captures. It will check for previously saved patterns which were previously generated and saved using Machine Learning algorithms. If the patterns match then ARYA will allow its users to use its functionalities. In our project, we have given only two types of access rights to users:

- Root access
- Browse access

Once the face is recognized by ARYA, it will automatically provide the configured access rights to its user and will start operating automatically and perform task based on the commands provided by its users.

5.2 Features

Based on the voice instructions, our proposed project can perform the following task or have the following features:

- Call: can easily make call from desktop using Your Phone application automatically
- Document modification: can perform basic word operations such as creation of a word document, inserting page breaks, modifying headings etc.
- Alarm: can open and set alarms.
- Custom search: it performs all the custom search operations in Google and YouTube.
- Wikipedia: can search anything in Wikipedia.

- Multimedia: can capture and record any picture and video respectively.
- Navigation: can open google maps
- Email: can send and receive emails

5.3 Software required

- Python 3.6
- OpenCV 4.1
- SQL lite3

5.4 Modules

- Speech Recognition: Provides all the libraries required to recognize voice and convert it to text by computers
- pyAudio: It provides audio I/O library which can used to play and record audio on multiple platforms such as Linux, Microsoft, mac etc. using python.
- Tkinter: It is a toolkit provided to develop Graphical User Interface.
- Wikipedia: allows Arya to search anything inside Wikipedia.
- Google-text-to-speech: Using which Arya can respond to users in a “human like voice”.
- CV2: this module is used specifically for face recognition feature.
- Pillow: it is a python library that allows to open and manipulate images of different file format.
- pysqlite: used to connect the sqlite database using python.

VI. EXPERIMENT

- It uses face recognition to unlock the Arya & provides authorization.
- Using autogui we can make calls & email using voice commands.
- Performs basic word operations on voice commands like adding images and text to the document.
- Allows you to navigate the locations using the browser.
- Greets the users based on the time.
- Allows you to search in google, YouTube and Wikipedia for specific query.

VII. RESULT ANALYSIS

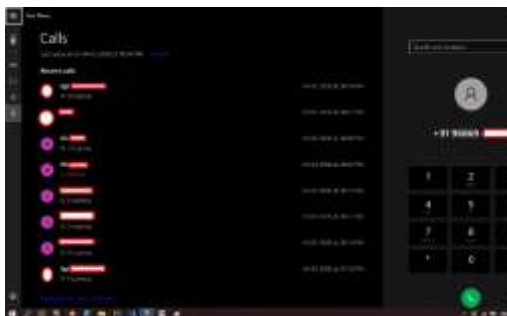
- It captures 100 images of user as a dataset and stores the user details in the database and trains the system using the dataset, after recognizing the face if the user attributes are matched up to 70% then it provides the authorization to the user.



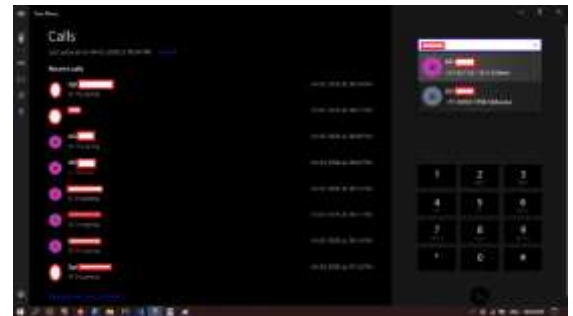
If the face attributes are doesn't match the condition then a unknowns face value is returned and the project terminates.



- On hearing the command create document, it creates a document and allow the use to perform the basic operations like adding text and images.
- On hearing the commands **call <number>**, it recognizes the phone number and makes a call using **your phone application**.



If the command is **call <name> from contacts** it searches the contact number from the phonebook and makes a call. (Note this feature solely works for windows).



- Performs basic camera operations using the cv2 modules.
- Allows us to find the location of the using google maps for browser.



- Based on the time it greets the user
- Stores the URL of the different websites to directly open the homepage. E.g. Open Facebook it uses the URL(www.facebook.com) & directly open the homepage of the Facebook in browser.
- Used autogui tools and sets the alarm on hearing the **set alarm <time>**.

VIII. CONCLUSION & FUTURE ENHANCEMENT

In this article we have used different frameworks to make a superior voice assistant. Some of the features of Arya makes it standalone when compared to other voice assistants that are available in market currently. Features that make Arya a superior among others are document management, calling contacts, face recognition etc. we can reduce the latency that is involved in face recognition and improve the document management by adding further features.

IX. REFERENCES

1. **Author:** Veton Kepuska, Gamal Bohouta. *Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home)*, 2018
2. **Author:** Debajyoti Pal, Chonlameth Arpnikanondt, Vijayakumar Varadarajan,



- ,Suree Funilkul. User Experience with Smart Voice Assistants: The Accent Perspective,2019*
3. **Author:** *Arsénio Reis, Dennis Paulino, Hugo Paredes, Isabel Barroso, Maria Joao Monteiro, Vitor Rodrigues, Joao Barroso.Using intelligent personal assistants to assist the elderlies, 2018.*
 4. **Author:** *Laura Burbach, Patrick Halbach, Nils Plettenberg, Johannes Nakayama, Martina Ziefle, Andre Calero Valdez. “Hey, Siri”, “Ok, Google”, “Alexa”. Acceptance-Relevant Factors of Virtual Voice-Assistants, 2019.*
 5. **Author:** *Kosuke Tsujino, Yusuke Nakashima, Shinya Iizuka, Yoshinori Isoda. Speech Recognition and Spoken Language Understanding for Mobile Personal Assistants,2013*