International Journal of Computer Technology and Science
Volume. 1 No. 4 October 2024
e-ISSN :3048-1961; dan p-ISSN :3048-1899, Hal. 12-26
DOI: https://doi.org/10.62951/ijcts.v1i4.113
Available Online at : https://international.arteii.or.id/index.php/IJCTS

Design GiggleGate as Desktop Virtual Assistant with Face and Speech Recognition Authentication System

Jasmine Aulia Mumtaz[1], Kinaya Khairunnisa Komariansyah*[2], Wildan Holik[3], Reza Pratama[4], Muhammad Galuh Gumelar[5], Endang Purnama Giri[6], Gema Parasti Mindara[7]
[1-7] IPB University, Indonesia

Address: Jl. Kumbang No.14, RT.02/RW.06, Bogor - West Java 16128
*Correspondence writer: kinayakhaii@gmail.com**

Abstract. *In recent years, virtual assistants have become an integral part of everyday life, simplifying routine tasks and allowing users to focus on more important matters. This research aiming to design GiggleGate, a virtual desktop assistant integrated with both face and speech recognition technology to enhance authentication security. The objective is to develop an authentication system that not only verifies user identity but also provides a more intuitive experience and seamless interaction. The research employs a development methodology to create and implement the system, which integrates face recognition via OpenCV and speech recognition via a Python library. The findings indicate that the integration of these technologies enhances security and user experience by offering dual-factor authentication. The system is expected to contribute to more secure and accessible virtual assistant applications, offering both a practical and efficient solution for users. The implications of this study suggest that the combination of face and speech recognition can provide an effective means to protect user privacy and improve the overall functionality of desktop assistants.*

Keywords *: authentication, face recognition, speech recognition, virtual assistant, security*

1. BACKGROUND

In recent years, virtual assistants have become an integral part of our daily lives, whether we realize it or not. This is due to the various features that make it easier for users to complete routine tasks, such as making calls, writing messages, saving to-do lists, and surfing the internet, so that users can focus more on the things that are important to them (Geetha et al., 2021) . Virtual assistants, developed through complex programming, allow users to execute various commands, save time, and customize functionality. in accordance need (Joshi et et al., 2023) . Wrong One development significant in virtual assistant technology is the integration of voice recognition, which allows users to give voice commands, such as opening applications or searching for information, without having to search for them one by one. This convenience is increasingly relevant, especially in increasing the efficiency of application use and providing quick access in environments that may limit mobility users. However, with the more its vastness use virtual assistants, the need for a secure authentication process becomes increasingly important to ensure that access is granted only to authorized users (DAE Saputri et al., 2023).

Received: August 21, 2024; Revised: September 20, 2024; Accepted: October 29, 2024; Published 26 November, 2024

Along with increasing need will authentication Which safe, recognition technology face And introduction voice has develop For fulfil need the. Introduction face use feature unique face every individual as verification method, give level accuracy Which tall in ensure identity users (Susim & Darujati, 2021) . This technology, which is part of computer vision and pattern recognition, give layer security Which can prevent access No valid. Besides That, introduction voice (speech recognition) Also play role important in authentication, although it differs from speaker recognition, which focuses on voice characteristics. individual. Speech recognition more emphasize on introduction phrase or say which is spoken by users, so that can functioning as component in system identity verification (Maulana & Agoes, 2019) . Integration technology face recognition And speech recognition, allow system For verify identity user with more accurate, create layer security addition Which effective in prevent access No legitimate.

Previous research by Atmawijaya & Radiyah (2024) has designed multi-factor authentication based on facial recognition and FIDO (Fast Identity Online), but has not supported voice recognition as additional authentication. Meanwhile, Anggara et al. (2023) implemented the MTCNN algorithm for facial recognition-based authentication, but only focused on facial recognition without including voice recognition. Saputri (2023) developed a biometric security system in a mobile banking application, but its implementation was limited to biometrics without voice recognition integration. Research by Santoso & Kristianto, (2020) and Firmansyah et al. (2024) also used OpenCV in face recognition for a student attendance system, but did not apply voice recognition as part of authentication. Finally, research (Susim & Darujati, 2021) emphasizes image processing for face recognition with OpenCV, but has not yet merge method introduction voice. By Because That, GiggleGate developed as an authorization system that integrates facial and voice recognition technology, offering dual authentication in one platform to improve security and overall user experience.

This research is important because security in authorization systems is increasingly becoming a major concern in the digital era, where threats to personal data and user privacy continue to increase. increase (Saputri et et al., 2023). GiggleGate designed For give solution authorization that safe And practical with merge technology introduction face And voice. The purpose of this research is to develop an authorization system that not only verifies identity user in a way accurate but Also increase experience

users through more natural and intuitive interactions. With the integration of this technology, GiggleGate is expected to be able to provide a higher level of security and present new innovations in biometric-based authentication technology.

2. STUDY THEORETICAL

Desktop Virtual Assistant

Desktop Virtual Assistant is software designed to perform tasks based on user commands, such as through voice input. This technology makes interaction with computers more efficient and accessible. By using voice commands, virtual assistants make it easier for users to perform various tasks such as opening applications and searching for information on the internet, without the need to use a keyboard (Geetha et al., 2021; Janani et al., 2021) .

This virtual assistant works by recognizing voice commands, processing them, and then providing a response, either in the form of sound or action on the device. In addition to increasing accessibility, the desktop virtual assistant also allows users to complete tasks faster and easier. This virtual assistant is here as an efficient solution to help task daily, Which flexible And can accessed When just. System This No just save time, but Also Can adapted in accordance need users, make it tool which is practical in various situations (Gupta et al., 2022; Joshi et al., 2023) .

System Authentication on Application

The authentication system in the application is used to verify the identity of the user before granting access to the service. This authentication process ensures that only authorized users can access the application, by using various authentication factors, like say password, fingerprint finger, or introduction face (Saputri et et al., 2023). There are two main authentication methods: *Single Factor Authentication* (SFA) and *Multi Factor Authentication* (MFA). SFA uses only one authentication factor, such as a password or OTP, but is considered less secure. In contrast, MFA combines multiple authentication factors, such as passwords and biometrics, to increase security (Anggara et al., 2023).

In a biometric-based authentication system, there are two main stages: *the enrollment module* , which creates a user biometric template, and *the identification module* , which compares the biometric data during login with the existing template (Anggara et al., 2023) . The use of *face recognition* technology in MFA is very popular for applications that require high level of security, such as banking and applications online (Saputra et et al.,2020). Face recognition integrated with speech recognition in MFA strengthens system authentication with

utilise various factor, so that improve user security and comfort in accessing applications (Atmawijaya & Radiyah, 2024; Saputri et al., 2023) .

Face Recognition

Face recognition is a technology used to identify and verify individuals based on unique features on their faces. This technology utilizes algorithms and image processing techniques to recognize human faces accurately and efficiently. The face recognition process can be applied in various applications, such as systems security door enter, supervision, And introduction identity in system security (Khair et al., 2024).

In general, facial recognition through the system can be divided into several stages. First, taking picture face use camera or sensors. Next, the recorded facial image is processed to extract the unique features of the face. These features are then compared with the data in the database to perform identification. Finally, verification or validation is carried out to ensure the match between the detected facial image and the stored data (Afrianto et al., 2023; Khair et al., 2024).

OpenCV

OpenCV (Open Source Computer Vision Library) is a computer vision processing library. picture And vision computer Which nature open source And can used with various programming languages, such as Python, C++, Java, and PHP (Santoso & Kristianto, 2020). OpenCV provides more than 2500 algorithms that can be used to detect, recognize, and manipulate objects in images and videos. One of the main capabilities of OpenCV is face detection, face recognition, object tracking, and feature extraction from images (Khair et al., 2024). This technology is widely used in the development of applications involving computer vision, such as face and object recognition systems.

In its implementation, OpenCV often utilizes algorithms such as Haar Cascade. Classifier For detect object, including face, in picture digital. This algorithm works by processing images using *Haar-like features* , which allows the identification of light and dark areas in the image to recognize facial patterns. The process detection face This involving steps classification sequentially in where image tested

by using integral images and certain features to produce more accurate results (Santoso & Kristianto, 2020).

Speech Recognition

Speech recognition is a technology used to recognize the voice spoken by the speaker and convert it into text. This process focuses more on recognizing words or phrases spoken by an individual, in contrast to speaker recognition which focuses on identifying who is speaking (Maulana & Agoes, 2019) . Technology This utilise algorithm Which can convert input voice become text, which can then be further analyzed by the system to take certain actions based on spoken commands (Gupta et al., 2022) .

In implementation use Python, module speech recognition used to process input voice And convert it become text. After the sound changed become text, backend Python will analyze text the For determine whether order provided in relation to system calls, sending emails, API calls, or extraction context (Gupta et et al., 2022) . Extraction context This often involves the use of processing Language experience For interpret text Which can read by humans. One of the libraries commonly used in speech recognition is pyyttsx3,
which can convert text to speech. This library allows the system to provide voice responses in multiple languages and dialects, and works offline, making it choice Which efficient for application Which need interaction voice with users (Joshi et al., 2023). With the ability to recognize voice commands and respond to them with text or voice, this technology offers a more natural and efficient way to interact with devices.

3. METHOD STUDY

Study This use method study And development (Research and Development or R&D), which aims to produce a product in the form of GiggleGate, a Desktop Virtual Assistant with an authentication system based on Face Recognition and Speech Recognition. This method was chosen to design, develop, and test the effectiveness of the resulting product (Sati et al., 2023). The following are the stages (Santoso & Kristianto, 2020) :

1. Literature Study, studying various related theories and technologies that will be used in study This. Studies literature This involving study about technology face recognition

using OpenCV, and speech recognition using Python libraries such as SpeechRecognition.

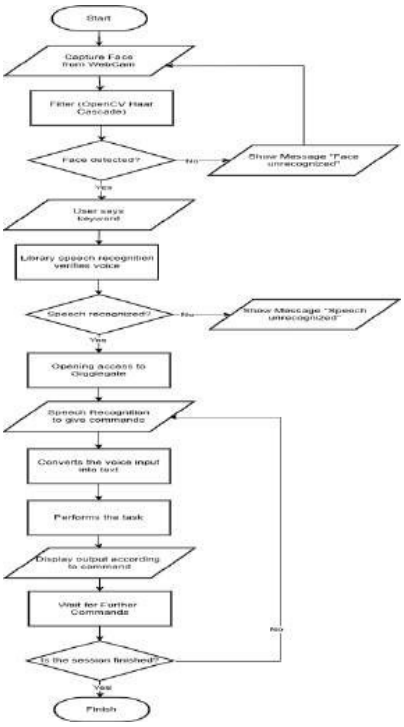
2. System Design, at this stage the workflow design for using the desktop virtual assistant system is carried out, starting from the authentication stage.
3. Implementation System, system built with use Language programming Python. OpenCV is used to perform face detection and recognition, while SpeechRecognition used For recognize voice. System This integrate both technologies the as method authentication And Speech Recognition as input for desktop virtual assistant.
4. System Testing, the system will be tested to ensure that the authentication function is working properly.
5. System Test Results Analysis, test results will be analyzed to evaluate system performance.

4. RESULTS AND DISCUSSION

Design System

In this system design stage, the flowchart illustrates the GiggleGate application process flow involving two main components: facial and voice recognition. Starting with application activation, the user activates the application to start the authentication process. Once the application is started, facial recognition is performed using a webcam. The facial image captured by the camera will be processed to detect and identify the user's face. with use technique like Hair Cascade Which implemented via OpenCV. Furthermore, after face detected, system request user For speak a password as part of voice authentication. Voice recognition is done using speech recognition, where spoken voice is converted into text to verify commands. users. If voice detected And in accordance, application will continue For open access to the system.

If the system successfully detects both the user's face and voice, then access to GiggleGate is granted, and the application proceeds to process the commands given by the user. Each voice command received will be processed by the system to perform task Which requested, for example open application or run function certain. The output of this process is then displayed according to the commands given by the user.

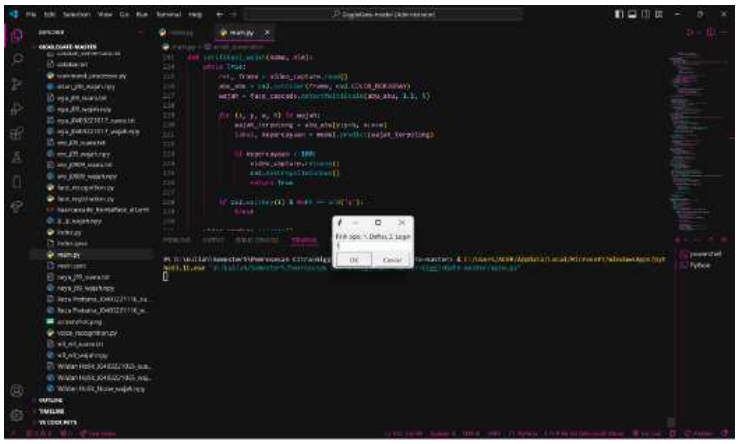


Picture 1. Flow chart System

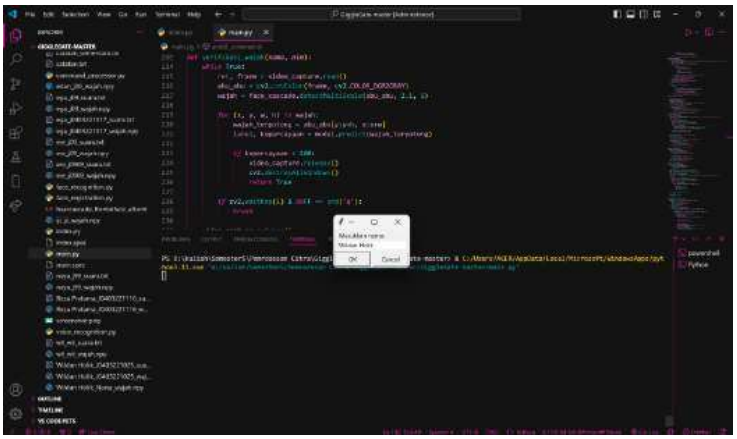
System Implementation

In the implementation phase of the GiggleGate system, the user registration process is an important initial step to ensure that each user is properly registered and has a unique identity. Users will be asked to enter their name, Student Identification Number (NIM), facial list, and keyword list for voice recognition. The name is used to record basic identity, while the NIM functions as a unique identifier, especially in the environment academic, so that system can differentiate One user from Which other. Besides In addition, users will also register their faces through scanning or uploading for facial recognition purposes. Keywords for speech recognition will also be registered, with the aim of adding a secure layer of authentication.

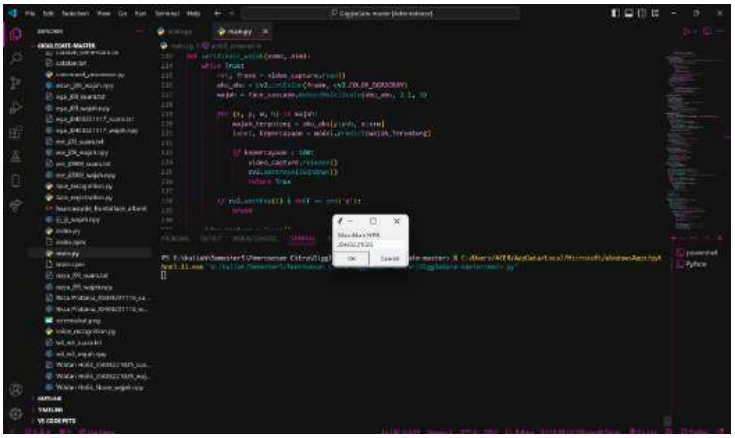
After data succeed inputted, system will do validation For ensure data format is correct, prevents duplication of NIM, and verifies the quality of facial and voice data. This validation aims to maintain data integrity and ensure that registered users have unique and verifiable identities. This stage is an important foundation for facial and voice recognition-based authentication, supporting user security and convenience in using the GiggleGate system.



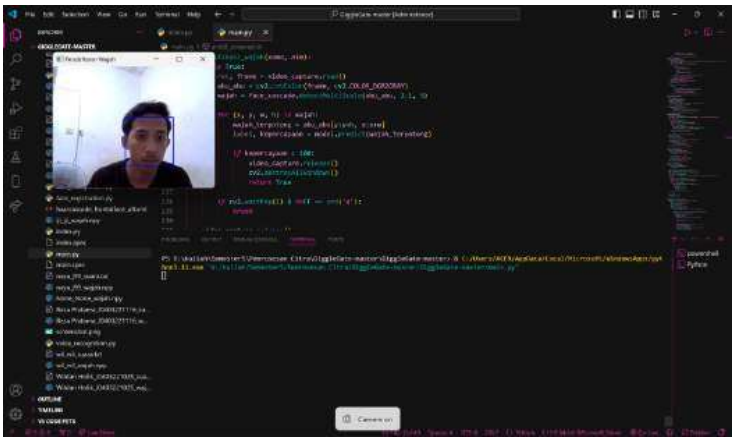
Picture 2. List Student Data



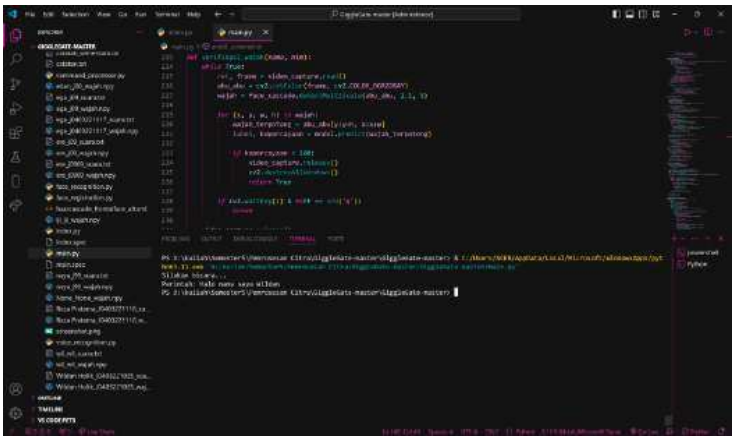
Picture 3. Input Name Student



Picture 4. Input NIM Student



Picture 5. Registration Face



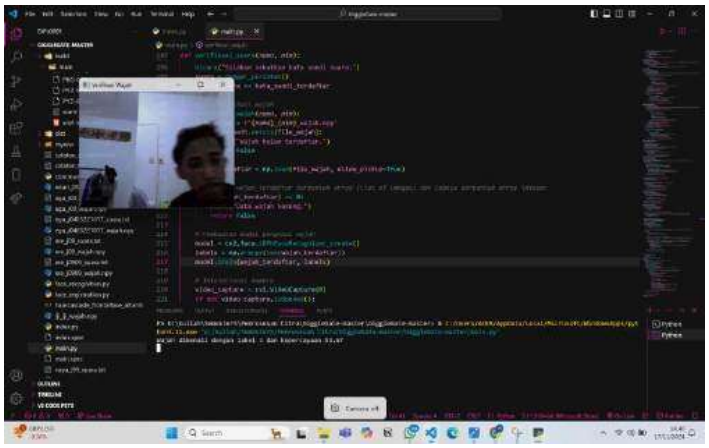
Picture 6. Input Keyword For Speech Recognition

System Testing

Stage testing system done For ensure performance Gigglegate walk as expected. This test covers several aspects, namely Face Recognition and Speech Recognition-based authentication, as well as virtual assistant functionality in supporting various tasks.

1. **Face Recognition Success**

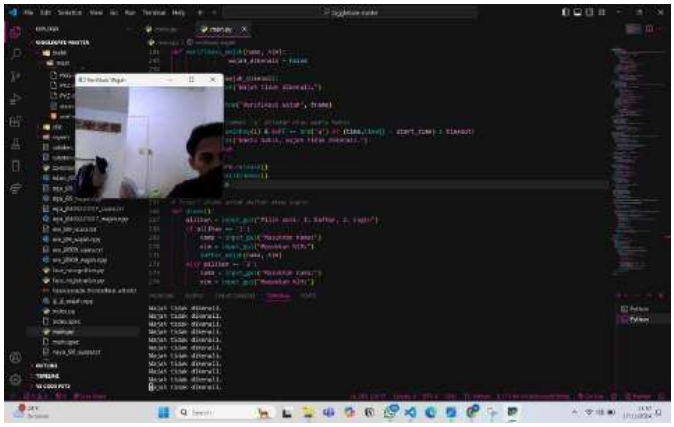
Testing This done For ensure system can recognize face student accurately. System will verify data face Which recorded with data Which stored in the database, so that the login process can be successfully carried out. This test aims to evaluate the level of accuracy and reliability of the system in detecting user faces.



Picture 7. Registered Face

2. Face Recognition Fail

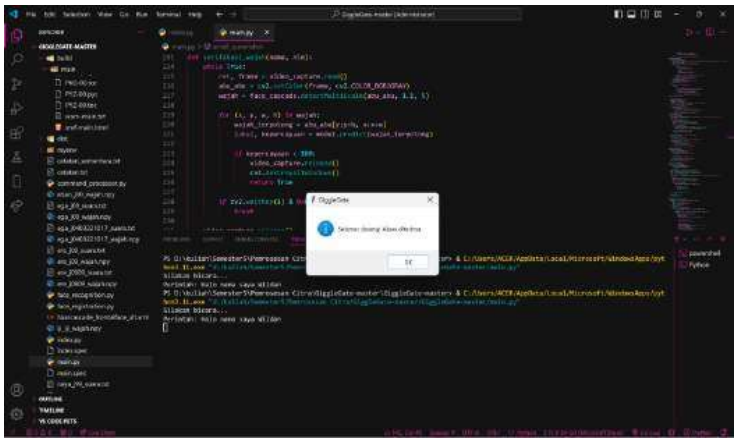
This test tests the condition when the system cannot recognize the student's face, either because the face does not match or the face position is not ideal. This test is important to identify potential weakness system And How system give response, such as displaying an error message or requesting a rescan.



Picture 8. Face No Registered

3. Speech Recognition Success

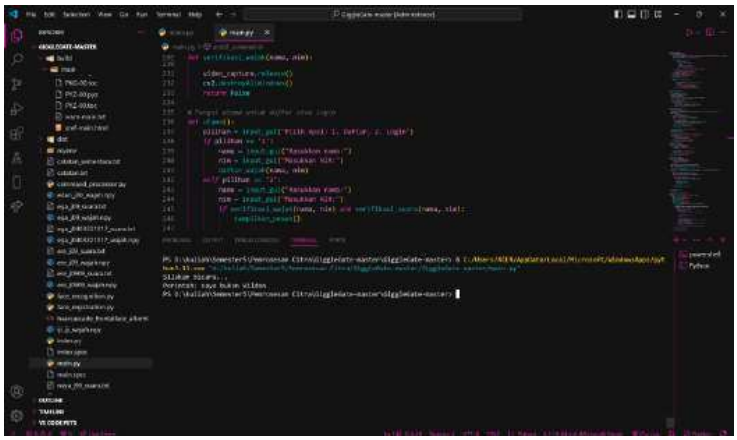
This test aims to ensure that the system can recognize the student's voice well and match it with the stored data. When the voice is recognized, the system will provide access login. Testing This evaluate effectiveness system in recognize user voice .



Picture 9. Registered Voice

4. Speech Recognition Fail

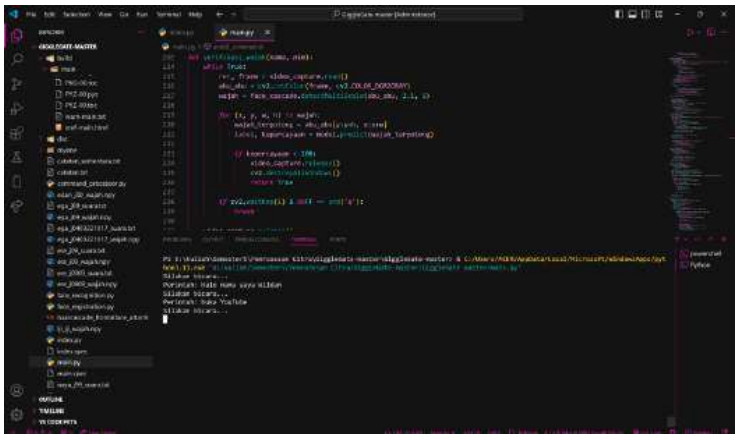
This test is conducted to identify situations when the system fails to recognize a student's voice, because the voice does not match the registered data. In this test, the system handle failure with give announcement with voice For try again.



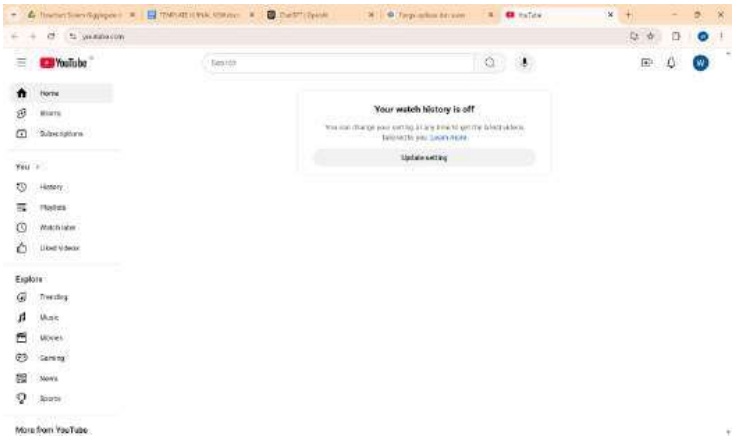
Picture 10. Sound Not listed

5. Testing Functionality Virtual Assistant

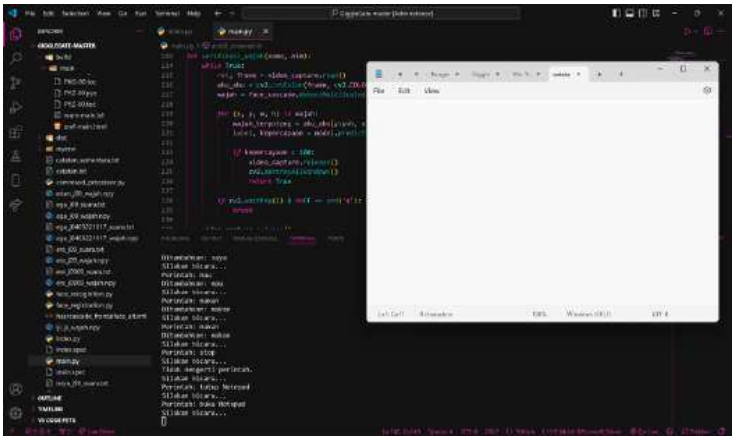
This test includes an evaluation of all Desktop Virtual Assistant features to ensure that each function runs according to user needs. This test ensures that Desktop Virtual Assistant can work optimally and provide an efficient user experience.



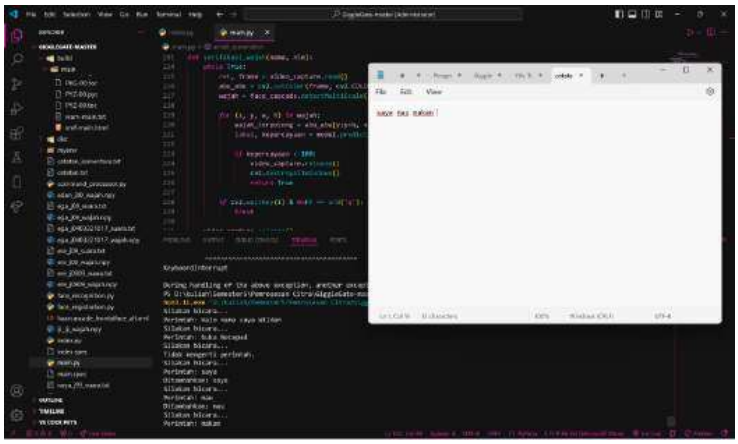
Picture 11. Users Input Voice Command to Opening Youtube



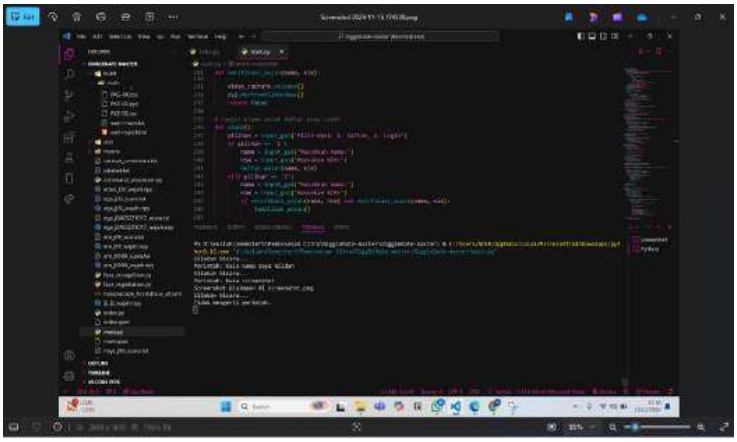
Picture 12. Assistant Opening Youtube



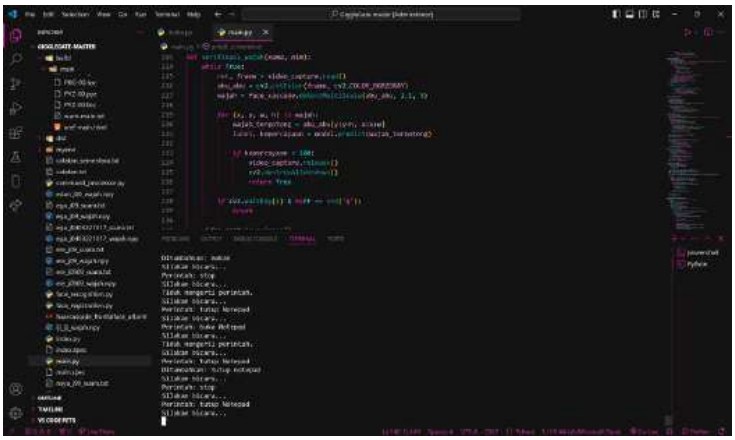
Picture 13. Asking Assistant to Open Notepad



Picture 14. Assistant Taking a Short Note



Picture 15. Assistant Taking a Screenshot



Picture 16. Assistant Closing Notepad

5. CONCLUSION

This research successfully designed and implemented GiggleGate as a Desktop Virtual Assistant with an authentication system using facial and voice recognition. Results testing show that system authentication This capable Work with fine in scenario introduction face and also voice, give access Which safe And efficient. In addition, the virtual assistant functionality has also been successfully implemented, allowing users to manage various tasks such as opening and closing applications, taking screenshots, and making notes in a notepad. The presence of this system is expected to increase user convenience and productivity.

REFERENCE

Afrianto, R., Lubis, H., & Irvan. (2023). Employee attendance application design use face recognition on University Hope Medan. Journal of Education Tambusai, 7(2), 5398–5404. https://doi.org/10.31004/jptam.v7i2.6568

Anggara, K. D., Primani, D., Kartikasari, & Bakhtiar, F. A. (2023). Implementation of the algorithm MTCNN in face recognition-based authentication mechanism. Journal Development of Information Technology and Computer Science, 7(8), 3613–3621.

Atmawijaya, R., & Radiyah, U. (2024). Multi-authentication design factor with introduction face and FIDO (Fast Identity Online).

Geetha, V., Gomathy, C. K., Vardhan, K. M. S., & Kumar, N. P. (2021). The voice-enabled personal assistant for PC using Python. International Journal of Engineering and Advanced Technology, 10(4), 162–165. https://doi.org/10.35940/ijeat.D2425.0410421

Gupta, U., Jindal, U., Goel, A., & Malik, V. (2022). Desktop voice assistant. International Journal for Research in Applied Science and Engineering Technology, 10(5), 901–905. https://doi.org/10.22214/ijraset.2022.42390

Janani, N., Jessica, R. J., & Vinmathi, M. S. (2021). Desktop assistant. International Journal of Advances in Engineering and Management (IJAEM), 3(9), 528–536. https://doi.org/10.35629/5252-0309528536

Joshi, R., Kar, S., Bamud, A. W., & R, M. T. (2023). Personal AI desktop assistant. Journal Information Engineering, 9(2), 75–85.

Khair, M. A., Dear Sir, P., Angelina, P. N., Zukhrufa, M. Z., & Adrezo, M. (2024). Design student attendance system based on face recognition in UPN Veteran Environment Jakarta. Informatics: Journal of Computer Science, 20(1), 35–42. https://doi.org/10.52958/iftk.v20i1.6696

Maulana, A., & Agoes, S. (2019). Analysis application system processing voice as safety house based on microcontroller. JETRI: Journal of Electrical Technology Research and Innovation, 19(1), 46–53. https://doi.org/10.33480/inti.v19i1.5263

Santoso, B., & Christian, R. P. (2020). Implementation use OpenCV on face recognition for system presence lecture.

Saputra, S. A., Ii, J. M. K., & Indah, B. S. (2020). E-voting simulation application election of the chairman of the ISTN information systems association with authentication face recognition use method LBPHFACES. Systemation: Journal of Information Systems, 9(2), 352–361. https://doi.org/10.32520/stmsi.v9i2.822

Saputri, D. A. E., Ernawati, I. A., Rabbaanii, N. A. N., & Satriani, A. D. (2023). Implementation AAA security in application BNI mobile banking. English Journal of Innovation Multidisciplinary Research, 1(2), 63–73. https://doi.org/10.69693/ijim.v1i2.9

Saputri, V. D. (2023). Implementation of biometric-based security system on mobile banking application. Journal Computer Indonesia, 2(1), 25–32.

Sati, A. T., Aditya, D. T., Azzahra, N. L., & Djutalov, R. (2023). Design system information finance orange elevation raya (OPERA) based on desktop with Java SE & MySQL use method. Scientific Electrical Engineering, 179–196. https://doi.org/10.25105/jetri.v16i2.3610

Susim, T., & Darujati, C. (2021). Image processing for face recognition (face recognition) using OpenCV. Syntax Admiration Journal, 2(3), 534–545. https://doi.org/10.46799/jsa.v2i3.202