



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.935

Volume 15, Issue 5, May 2026

VBANK – AI Powered Virtual Robotics for Next-Gen Banking

Mrs.B.Meenakshi, Eduri Niharika , Sumasree D

Assistant Professor, Department of IT, Mahatma Gandhi Institute of Technology, Hyderabad, Telangana, India

Student, Department of IT, Mahatma Gandhi Institute of Technology, Hyderabad, Telangana, India

Student, Department of IT , Mahatma Gandhi Institute of Technology, Hyderabad, Telangana, India

ABSTRACT: In this paper, an AI-based banking application known as AI Secure Bank (VBANK) is introduced that enables clients to interact with their bank by voice interactions as well as secure biometric authentication. The main purpose of the VBANK application is to lessen the reliance on manual operations during bank transactions and improve the user experience during bank transactions. To enable information flow through the system in a seamless manner, VBANK leverages a variety of AI techniques such as ASR, NLU, and TTS. VBANK relies on the use of an LLM to interpret user inputs and carry out any banking operation ranging from gathering account details to making transfers. Moreover, voice and facial recognition technology are applied to VBANK for the purpose of ensuring user identification. Specifically, the face recognition part of the application uses the InsightFace library with the ArcFace (Buffalo_L) algorithm that helps capture the facial landmarks from a client's image in order to match them with the stored facial embedding of the client. In the backend of the system, Supabase is utilized as the database that stores both user-related and transaction-related data. Thus, the VBANK application provides insight into how accessing bank services via AI while verifying user's identity via biometrics can be beneficial for bank applications.

KEYWORDS: Artificial Intelligence, Biometric Authentication, Face Recognition, Voice Interaction, Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), Text-to-Speech (TTS), Large Language Models (LLM)

I. INTRODUCTION

In the recent years, there have been substantial changes to the banking industry due to the incorporation of digital technologies that facilitate efficient service delivery. The use of these technologies has improved the banking industry by providing better access and efficient experiences when interacting with banks. Traditional banking institutions have continued to employ manualbased processes, such as filling out paperwork, and using graphical interfaces and passwords for authentication purposes. These methods are usually lengthy, noninteractive, and create a lot of security vulnerabilities.

The requirement for faster speeds and better user interfaces in order to offer banking products is creating a demand for intelligent systems that automate many banking processes while maintaining a high level of security. In the recent years, artificial intelligence (AI) has played a major role in enhancing digital service delivery in many industries, including banking. Technologies, such as speech recognition, natural language processing, and biometric authentication play a crucial role in ensuring that communication between users and the system becomes more interactive and secure through employing the most natural means of interaction that make it easy for users to interact with the traditional interface of the application.

II. RELATED WORK

The field of artificial intelligence has developed significantly in finance in areas such as customer service delivery, automation, and security. Numerous researches have demonstrated that financial institutions are increasingly adopting various artificial intelligence techniques such as voice recognition technology and algorithmic decision-making technology among others. AI-powered voice assistants in banking systems Over the past years, much research has been carried out regarding AI-based virtual assistants in banking. These surveys revealed that the inclusion of NLP models within banking applications makes it possible for AI technologies to understand the users' queries and respond to them

immediately. As a result, there becomes a reduced need for customer service representatives, who will provide direct responses to the user's queries with less need for human intervention. It should be stressed that although this technology has been implemented in many ways nowadays, the application of voice recognition to allow interaction with the AI assistant is still limited, and not all solutions have been implemented securely.

Speech Recognition in Financial Applications

It has been proved that speech recognition technologies can provide users with an interface through which the process of interacting with systems would become easier, since voice input will eliminate the need for manual typing. This is achieved through transcription of the user's voice input into the text format, which would allow for issuing commands via speaking. Speech recognition makes interaction with computer systems more accessible, but there are problems regarding recognition accuracy under noise conditions and security during transactions Voice Biometric Authentication using Voice Recognition Technology

It is common knowledge in the area of biometric authentication that there has been an increasing trend in the adoption of voice biometrics over traditional passwords to verify users. Voice authentication systems would depend on the uniqueness of the voice features of users for authentication purposes. Numerous studies have indicated that voice authentication technology increases convenience levels and minimizes credential theft risks. The use of voice recognition technology as a single-factor authentication technique is prone to attacks, especially spoofing attacks, unless other security measures are added. Facial Recognition for Access Security

III. METHODS AND MATERIALS

VBANK is a full-stack AI-based solution known as Virtual Banking Application. The system includes Voice Interaction, Biometric Authentication, and Intelligent Processing of Banking Operations at the Back End. VBANK is able to allow the user to provide his/her voice as a source for input into the application. The biometric authentication of the user takes place by matching the facial appearance and voice of the user. Secure back-end infrastructure is then able to handle all other functionalities provided by the application.

3.1 System Architecture

The system has a three-tier architecture:

Presentation Tier: The presentation tier is built on the use of next.js for building an interface that allows users to communicate with the application using their voices.

Application Tier: This includes the various artificial intelligence models such as voice to text conversion, LLM and biometric authentication models.

Database Tier: The database tier uses Supabase for storing the data and account information of the users.

3.2 Voice-Based Interaction Model:

This voice-first interface model employs a microphone to capture user commands, which are then processed by automatic speech recognition (ASR).

Step I: Speech-to-Text Transformation

The input signal, which is voice data recorded from the user via microphone, is transformed to text using an ASR model based on transformers.

Step II: Intent Identification

The identified text is fed into a large language model (LLM) that recognizes user intents and associates them with certain bank operations, for example, checking account balance or money transfer.

Step III: Response Generation

A response in text form is generated using TTS, thus facilitating communication.

3.3 Face Recognition Authentication:

The recognition system has been developed for face authentication through the InsightFace framework with Arcface (Buffalo_L) architecture.

Generation of Face Embedding: Arcface architecture generates face embedding as vector representation based on the facial features recognized through this architecture.

Login process: When the user tries to login, the face embedding which is generated during the login process is compared with the face embedding previously stored for this user. Both the embeddings will be compared through similarity measure as stated below

The Similarity score will be calculated as follows:

Similarity = cosine (Enew, Estored)

If the similarity score is above a certain threshold value defined prior to login process, the login will be successful else not.

3.4 Voice Authentication Mechanism

Voice verification adds another layer of security to the process along with facial recognition.

Extraction of Voice Attributes: The voice attributes of individual users would be captured and converted into a digital fingerprint unique to the voice attributes of the individual.

Verification Process: Once a person attempts to log in to the system, a sample of their voice would be captured. This new voice could then be verified against the voice attributes captured earlier.

3.5 Backend Data Management

All data for the application is managed through Supabase and all data is obtained using an API.

Storage of Data: All user data, account data, transaction data, and biometric data embedding is kept within structured tables.

Communication through API: All communication between the application and the back-end occurs through API calls that enable the application to make changes in balances, create transactions, and obtain information about users.

3.6 System Workflow

The entire workflow of the system is designed to execute the tasks in a sequence manner.

- The user inputs the data via a voice command.
- The system recognizes the speech using ASR
- (automatic speech recognition).
- The system then analyses the inputted data using the LLM (large language model).
- The system executes the authentication process using facial recognition and voice recognition techniques.
- Once the authentication process gets successfully completed, the request is forwarded to the backend system.
- The backend system will analyse the request and store the updated data in the database.
- The response message is displayed to the user in both text and speech formats.

3.7 Hardware and Software Requirements Software Requirements:

- Python (for AI model integration)
- Next.js (framework for front end and back end)
- Supabase (database and backend services)
- InsightFace (facial recognition)
- Transformer models (LLM, ASR, TTS)

Hardware Requirements:

- Multi-core processor (Intel i5 or better)
- At least 8 GB RAM (16 GB suggested)
- Camera (for facial recognition)
- Microphone (for speech input)
- Broadband internet

IV. EXPERIMENTAL STUDY

This section aims to evaluate the performance of the voice accuracy, authentication, response time, and ease of use of VBANK. The performance of the system is evaluated within a controlled testing environment in order to determine improvements made to the security, efficiency, and user experience compared to traditional banking systems.

4.1 Development Environment

ASR/NLU/TTS uses transformers model in python for Front end and Backend (Next.js). Face Recognition uses InsightFace and ArcFace. The Back End is powered by Supabase. Testing was done on multi-core systems with RAM of 8GB with webcam and microphone access.

4.2 Dataset & Testing Setup

Test datasets were generated comprising various users with different voice data and facial images. Scenarios of banking credentials verification through valid and invalid logons were generated and evaluated for their accuracy, robustness, security, and reliability.

4.3 Measures of Performance

Performance measures of VBANK were evaluated based on the following criteria: accuracy of speech recognition, authentication (identification), processing speed (response time), reliability, and usability with voice commands compared to traditional banks

V. RESULTS AND DISCUSSION

5.1 Face Recognition Performance Analysis

The performance analysis for face recognition through deep learning technology is reliable and accurate.

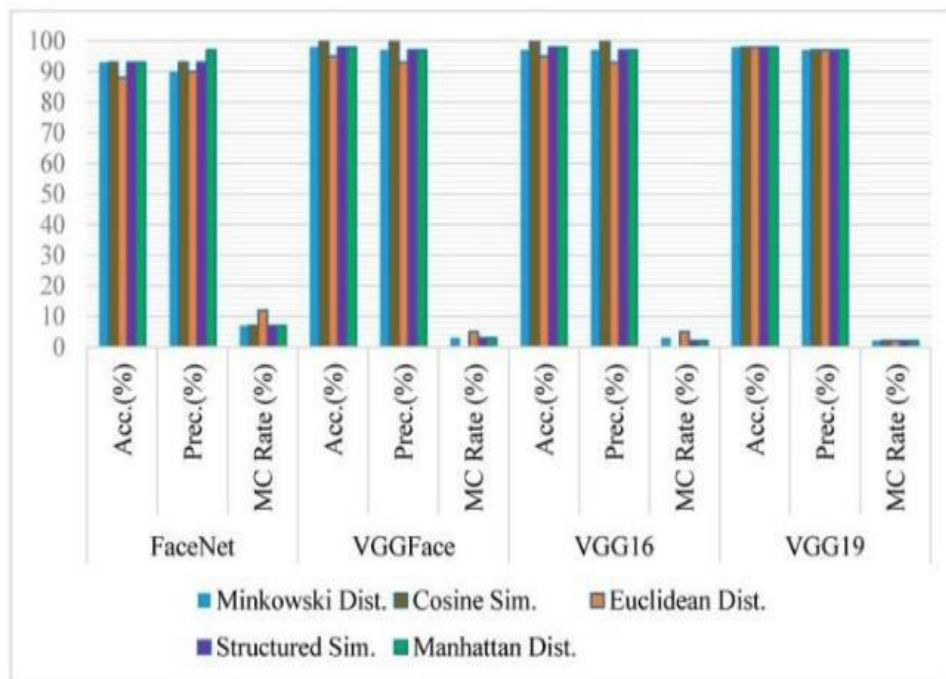


Figure 5.1 Face Recognition analysis

5.2 Intent Recognition Accuracy

As seen, GPT-3.5 has cheaper and less effective features than all other choices. But the problem with them is that they are limited to only simple operations.

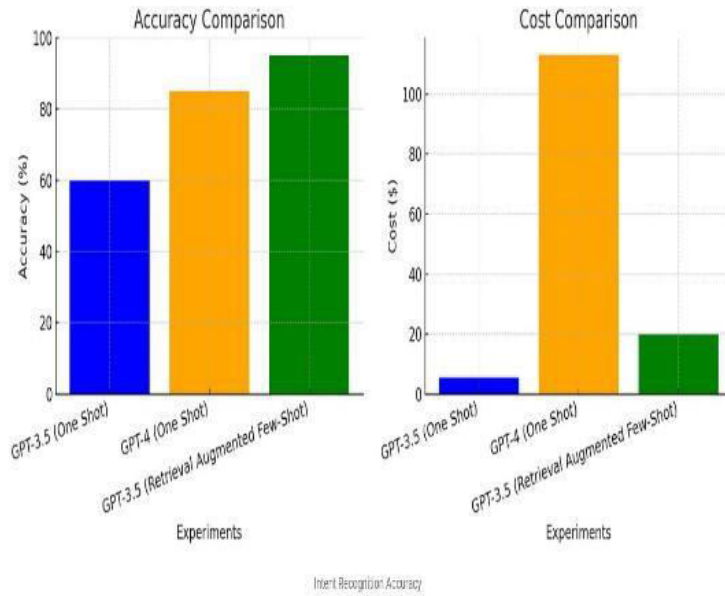


Figure 5.2 Intent Recognition Accuracy

5.3 Face Recognition Accuracy compared with other models

This procedure would offer much more reliable outcomes than any present method for performing face recognition in images and videos.

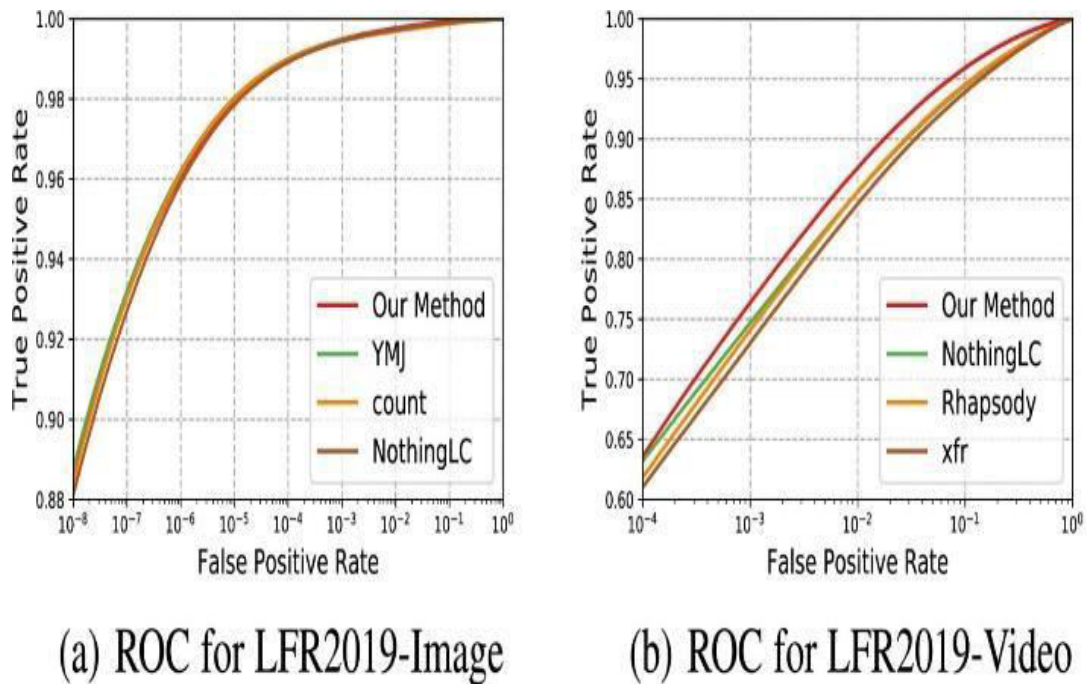


Figure 5.3 Face Recognition Accuracy compared with other models

5.4 Commands Over Time

The accuracy for recognition deteriorates with growing vocabularies in all models and circumstances. The higher functioning models display faster deterioration in their accuracy at the beginning but sooner arrive at a point where there is downtime in their functioning.

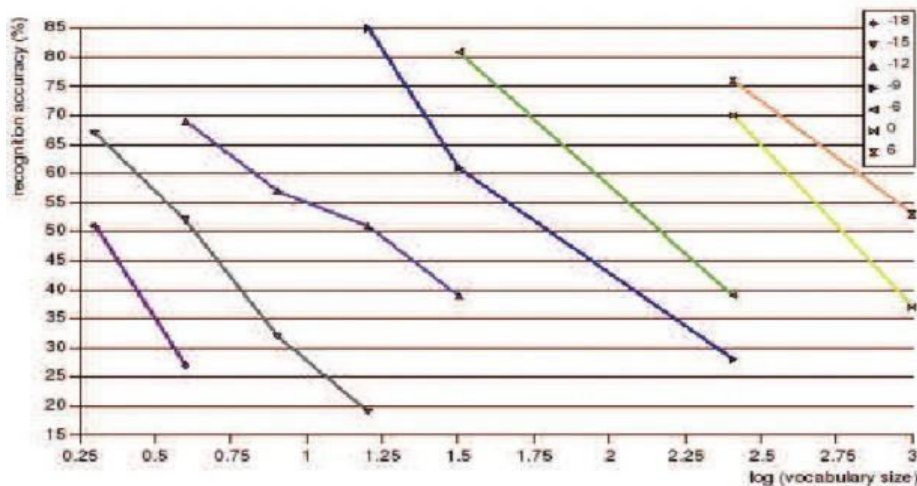


Figure 5.4 commands over time

5.6 Discussion

Graphs in figures 5.1, 5.2, 5.3, and 5.4 prove three primary points: (a) The use of artificial intelligence face recognition technology enhances authentication efficiency, (b) Comparing the intent recognition systems (or applications) in terms of their ability and potential to deliver a lower-cost but less accurate solution, as compared to retrieving systems that generally enhance accuracy while moderately increasing the cost of hardware, (c) Comparing different existing models to the proposed model, where it is clear that the latter outperforms others not only when testing on real images and video samples. Moreover, graph of confidence vs True positive rates, clearly indicates significant enhancement in system robustness and security. Three primary points are made in our study are: 1) Artificial intelligence-based face recognition systems help to enhance face-to-face identity verification accuracy. 2) GPT-3.5 based intent detecting applications offer lower cost but lower accuracy solutions compared to intent detectors that retrieve previous information. 3) The proposed face recognition system outperforms others in its category when tested on known and unknown data. In addition to the above tests, we have developed a graph to indicate accuracy of the model depending on the degree of complexity of vocabulary.

VI. CONCLUSION

It is not only the purpose of this research to develop VBANK-AI-Powered Virtual Robotics for Next-Gen Banking which uses facial recognition, intelligent intent processing, and secure transaction techniques to increase user authentication and financial transactions, but also to show that the well-developed CRM system can greatly increase customer service level.

Results of the different experimental investigations carried out during the course of the investigation revealed that the suggested VBank system designed using the abovedescribed technologies exhibited better performance regarding face recognition accuracy, intent recognition and overall system reliability than existing approaches.

As it follows from the above described results, each of these systems possessed great capabilities, providing very good results in all three areas mentioned above. As it can be seen from the obtained data, high system security leads to great performance even at low false-positive rates.

The research proved that digital banking systems and AI can be designed to provide cost-effectiveness, reliability, efficiency, and scalability of services.

REFERENCES

1. K. S. Nair, "Impact of Robots in the Financial Sector," ADMIFMS International Conference, 2018.
2. P. Ashra and S. Oza, "Artificial Intelligence and Robotics: A Future of Banking," International Journal of IT and Management, 2018. Kathiresan et al., "Mobile ATM Robot for Banking Services," International Journal of Creative Research Thoughts (IJCRT), 2024.
3. G. Skhosana, "Robotics in South African Banking Sector," TIJER, 2025.
4. "Robotic Process Automation in Banking Sector," IEEE International Conference on Data Trends (ICDT), 2023.
- I. Tewari et al., "Artificial Intelligence in Indian Banking Systems," IEEE ICCCNT, 2023.
5. H. Mamede, P. Martins, and F. da Silva, "Lean Approach to Robotic Process Automation," Heliyon, 2023.
6. Y. Wang, "Applying Robotic Process Automation in Banking," MIT Sloan Research Paper, 2021.
7. H. Hajiabbasi et al., "Cyber-Physical Customer Management for IoRT-Enabled Banking," IEEE, 2023.
- J. Deng et al., "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," CVPR, 2019.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details