International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

VION: A Fully Functional AI-Powered Chatbot

Mr. R. Jyoth Singh[1], Ronanki Bala[2], Miriyalu Chandini[3], Ch.Neha Lakshmi Sai[4], Sabbella Chaitanya Sri Reddy[5]
[1]Associate Professor, Department of Computer Science and Engineering, SRK Institute of Technology, Vijayawada, India.
[2,3,4,5] Students, Department of Computer Science and Engineering, SRK Institute of Technology, Vijayawada, India.
Email ID: *chaitanyasri.sabbella@gmail.com[5]*

Abstract
The rapid evolution of intelligent personal assistants has enabled conversational interaction between users and digital systems; however, existing solutions remain limited in terms of security, contextual awareness, and deep integration with mobile device functionalities. This paper presents VION (Voice-Integrated Intelligent Operator Network), a fully functional AI-powered chatbot that combines conversational artificial intelligence with biometric authentication and voice-driven mobile device control to deliver a secure and hands-free user interaction platform. VION employs a face recognition-based authentication layer to verify user identity before enabling access to sensitive operations such as application navigation, voice calling, and message transmission. A GPT-powered natural language processing engine facilitates multi-turn, context-aware conversations by maintaining recent interaction history, thereby improving response relevance and personalization. The system further integrates a speech recognition module that converts spoken commands into structured control instructions, enabling real-time execution of tasks including app launching, web access, and WhatsApp message dictation. The proposed architecture follows a modular and layered design, ensuring scalability, maintainability, and secure communication between the authentication, conversational, voice processing, and mobile control layers. Cloud-based backend services are utilized to manage user profiles, biometric data, and conversation logs while supporting cross-device synchronization. Experimental evaluation demonstrates that VION achieves reliable authentication accuracy, low command execution latency, and improved user experience compared to conventional chatbot systems. The results indicate that the integration of biometric security with intelligent conversational and control capabilities significantly enhances both system usability and operational security in mobile assistant environments

1. Introduction
The rapid evolution of artificial intelligence and mobile computing has accelerated the demand for intelligent, secure, and hands-free personal assistant systems [1]. Traditional chatbots and voice assistants often suffer from limited contextual awareness, weak authentication mechanisms, and restricted integration with native mobile applications, which reduces both usability and operational security [2]. Moreover, sensitive functions such as voice calling, messaging, and application control are typically executed without strong biometric verification, leading to increased risks of impersonation and unauthorized access [3]. This paper introduces VION (Voice-Integrated Intelligent Operator Network), a fully functional AI-powered chatbot designed to unify conversational intelligence, biometric authentication, and real-time mobile device control within a single secure platform. The system leverages a GPT-based natural language processing engine, face recognition using deep learning models, and a speech recognition pipeline to enable seamless interaction and hands-free operation. Key contributions include:

- A unified conversational and control interface for application navigation, web access, calling, and messaging.
- Secure biometric authentication using face recognition for access to sensitive operations.
- Real-time voice command processing with low-latency task execution.
- Context-aware conversation management

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

through session-based interaction history.

- Scalable cloud-backed data management for user profiles, logs, and synchronization.

2. Related Work

Recent literature presents multiple approaches toward intelligent assistants, biometric authentication systems, and voice-driven interaction frameworks. Barot and Panchal [4] proposed a smart voice assistant integrated with face recognition to enhance access security in desktop environments; however, the system primarily focused on offline operation and lacked deep integration with mobile applications and real-time cloud synchronization. Bokefode et al. [5] introduced a dual-biometric authentication framework combining face and voice recognition for smart home environments, achieving high accuracy but at the cost of increased computational complexity and dependency on high-quality sensor hardware.

Several studies have explored face-authenticated voice assistant architectures. Tirumal et al. [6] designed a prototype system in which facial verification enables access to voice-controlled application launching and web navigation. While the system demonstrated improved personalization and security, it did not address advanced features such as cross-platform messaging, mobile screen interaction, or persistent conversational memory. Wahsheh and Steffy [7] conducted a comprehensive security-oriented survey on combining biometric modalities in voice assistant systems, emphasizing vulnerability to spoofing and replay attacks but offering limited insights into large-scale mobile deployment and usability trade-offs. Research on continuous and robust speaker authentication has also gained attention. Feng et al. [8] proposed a wearable-based continuous authentication mechanism that correlates speech signals with body-surface vibrations to mitigate impersonation risks. Although effective in high-noise and adversarial conditions, the requirement for additional hardware reduces practicality for consumer-grade mobile assistants. Similarly, Revathi et al. [9] implemented a real-time voice authentication system using time–frequency features and convolutional neural networks on embedded platforms, demonstrating feasibility under constrained hardware but relying solely on voice as the biometric modality. From a conversational intelligence perspective, recent surveys and comparative studies on automatic speech recognition and deep learning-based NLP models [10], [11] highlight the advantages of transformer-based and hybrid architectures in improving recognition accuracy and contextual understanding. However, most existing implementations focus on isolated components such as speech-to-text or chatbot interaction, rather than integrating these capabilities into a unified, secure, and device-controlling assistant framework. In contrast to existing systems, VION distinguishes itself by combining biometric face authentication, GPT-powered conversational intelligence, real-time voice command processing, and deep mobile-device integration within a single modular and cloud-backed architecture. This holistic approach addresses both the functional limitations and security gaps identified in prior work, enabling secure hands-free operation, contextual continuity, and practical deployment in mobile computing environments. Figure 1 shows VION

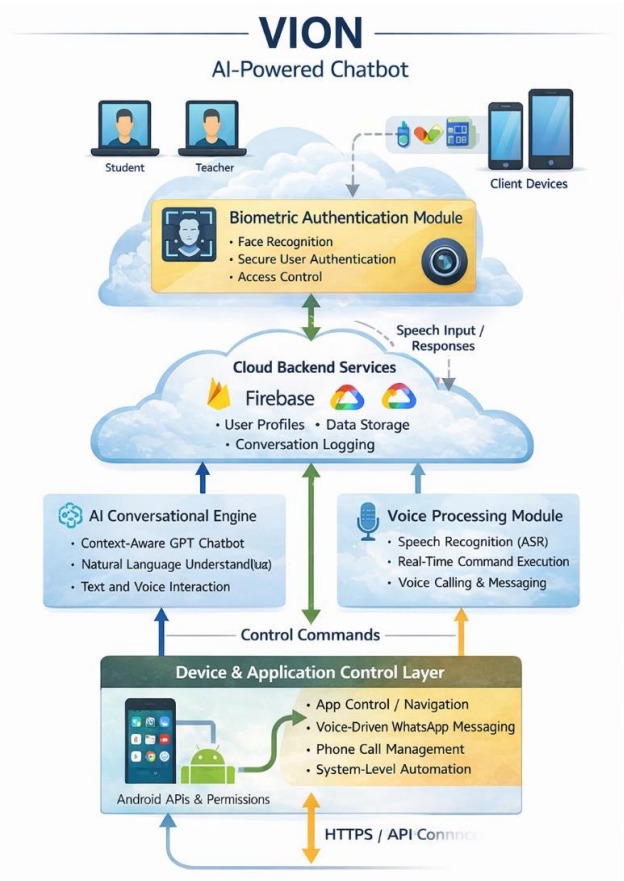

Figure 1 VION

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

3. Ease of Use

3.1 Selecting the Appropriate Deployment Environment

The VION system is designed for deployment on modern mobile and cloud-based platforms, including Android environments integrated with cloud services such as Firebase, Google Cloud Platform (GCP), or AWS. The standard deployment assumes compatibility with Android Studio (Arctic Fox or later), Python 3.9+ for backend services, and a stable internet connection for GPT-based conversational services. For organizations deploying VION on private cloud infrastructures or restricted enterprise networks, alternative backend configurations and API gateway policies should be consulted to ensure compliance with internal security and access control standards.

3.2 Maintaining the Integrity of System Specifications

The VION framework is structured to ensure consistent and secure behavior across different devices and deployment environments. Core modules, including biometric authentication, conversational AI services, voice processing pipelines, and mobile control interfaces, follow predefined architectural and security specifications. These components should not be modified without comprehensive testing and validation. For example, facial authentication confidence thresholds and session timeouts are deliberately configured to balance usability and protection against spoofing or unauthorized access. These parameters, along with encrypted storage of biometric embeddings and conversation logs, are aligned with best practices in data privacy and mobile
security. Alterations to authentication workflows, permission models, or API access policies should only be performed following a formal security review.

3.3 Preparing the System for Deployment

Prior to deploying VION, system administrators should document operational requirements, including supported device models, network constraints, and data retention policies. All stakeholder approvals, user enrollment strategies, and cloud resource provisioning should be finalized before enabling live biometric authentication and conversational services. Development and testing environments must remain isolated from production systems, and real user data should be anonymized during validation phases. Administrative privileges should be restricted to authorized personnel, and additional authentication mechanisms should not be introduced without ensuring compatibility with the existing face recognition and voice control security model, which has been optimized for mobile-based intelligent assistant environments.

4. System Architecture

4.1 Overall Framework

The VION system follows a layered and modular architecture consisting of a user interaction layer, an intelligent application layer, and a data and cloud services layer. As illustrated in Fig. 1, users interact with the system through mobile or browser-based client interfaces using both text and voice inputs. These inputs are transmitted securely over HTTPS to the backend services, which coordinate biometric authentication, conversational intelligence, and device control operations.The architecture integrates a biometric authentication module, a GPT-powered conversational engine, and a voice processing pipeline within a unified control framework. Cloud-based backend services manage user profiles, authentication logs, and conversation history, while the device control layer interfaces with native Android APIs to execute system-level actions such as application navigation, calling, and messaging. This layered approach ensures separation of concerns, scalability, and secure handling of sensitive operations.

4.2 Technology Stack

The technology stack for VION was selected to satisfy requirements related to real-time performance, security, and cross-platform scalability. The frontend interface is developed using Android Studio and modern web technologies for mobile and browser-based interaction. Python and JavaScript are used for backend orchestration and API integration. The conversational engine leverages a GPT-based natural language processing framework for context-aware dialogue management.

OpenCV, TensorFlow, and DeepFace are employed for face recognition and biometric embedding generation, while Whisper and Google Speech APIs

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

support robust speech-to-text processing. Firebase and cloud platforms such as GCP or AWS provide scalable backend infrastructure, secure data storage, and synchronization services. Secure communication between system components is enforced through HTTPS and token-based session management.

4.3 User Interface Design

The user interface is designed to support both text-based and voice-driven interaction with minimal user effort. The primary components include:

- Authentication Interface: Captures facial data through the device camera and performs real-time biometric verification before granting access to sensitive system functions.
- Conversational Interface: Provides a chat-based and voice-enabled interaction window for natural language queries and command input.
- Control Dashboard: Displays available system actions such as app navigation, calling, messaging, and recent interaction history.
- System Feedback Module: Presents real-time status updates for authentication results, command execution, and error handling.

The interface emphasizes accessibility and hands-free usability, ensuring seamless operation for both general users and individuals with physical or visual impairments.

4.4 Backend Services

The backend services coordinate system intelligence, security, and task execution. Key functionalities include:

- Biometric Authentication Service: Manages facial embedding generation, secure storage, and real-time identity verification.
- Conversational AI Service: Interfaces with the GPT-based NLP engine to process user queries, maintain session context, and generate intelligent responses.
- Voice Processing Service: Handles speech recognition, command parsing, and intent classification for real-time control.
- Device Control Interface: Bridges interpreted commands with native Android APIs and system permissions to perform mobile-level actions.
- Logging and Monitoring Service: Records authentication attempts, command execution status, and conversation metadata for audit and debugging purposes.

4.5 Database Schema

The data layer is designed to securely store biometric, conversational, and operational data. Core data entities include:

- User Profiles: Stores user identification details, facial embeddings (encrypted), and device metadata.
- Authentication Logs: Records login attempts, timestamps, device information, and access status.
- Conversation History: Maintains session-based text and voice interactions for contextual continuity.
- Voice Command Logs: Stores recognized commands, executed actions, and execution outcomes.
- Call Records: Tracks voice call details including contact information and duration.
- Messaging Records: Stores WhatsApp message metadata and delivery status.
- Application Access Logs: Maintains records of app and web navigation activities.

4.6 Real-Time Processing and Communication

VION supports real-time interaction through asynchronous request handling and low-latency communication pipelines. The voice processing module enables continuous speech capture and immediate command translation, while the conversational engine generates near real-time responses. System feedback is provided to the user through audio and visual cues, ensuring transparency during authentication, command execution, and error states. This architecture allows parallel processing of biometric verification, conversational inference, and device-level control, thereby reducing system response time and improving overall user experience.

4.7 Cloud Integration

Cloud services play a central role in ensuring scalability, availability, and data security. Firebase and cloud storage platforms are used to store

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

encrypted user profiles, authentication logs, and conversation history. Synchronization services enable cross-device continuity, allowing users to resume interactions across supported platforms. Secure access policies and role-based permissions are enforced at the cloud level to prevent unauthorized data access and ensure compliance with privacy and security standards.

5. Implementation Details

5.1 Authentication and Security

User authentication in VION is implemented using a biometric-first access control model reinforced with token-based session management. Facial authentication serves as the primary identity verification mechanism, after which secure session tokens are issued to maintain continuity across interactions. The system enforces strict access policies for sensitive operations such as voice calling, WhatsApp messaging, and screen-level device control. The security framework adopts a multi-layered defense strategy. Client-side and server-side input validation ensures resilience against malformed requests and command injection. Secure API communication is enforced through HTTPS and role-based access policies, while backend services implement rate limiting to mitigate brute-force and replay attempts during authentication and command execution. Biometric data, including facial embeddings, is stored in encrypted form, and conversational logs are protected using access-controlled cloud storage. These measures collectively ensure confidentiality, integrity, and availability of system resources. Figure 2 shows Sign-In Interface with Face Authentication

Figure 2 Sign-In Interface with Face Authentication

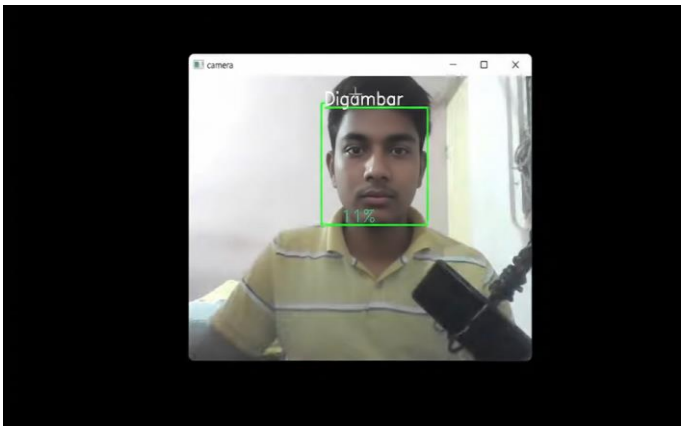

Figure 3 User Registration Interface

5.2 Biometric Authentication Module

The biometric module captures facial images through the device camera and processes them using a deep learning-based face recognition framework. Feature embeddings are generated and compared against securely stored reference templates to verify identity. A confidence threshold determines authentication success, balancing usability and spoof-resistance. The module is extensible to support additional modalities such as voice-based verification or liveness detection to further strengthen resistance against replay and presentation attacks. Upon successful authentication, a secure session context is established, enabling access to restricted features. Failed authentication attempts are logged with device metadata and timestamps to support auditing and anomaly detection.

5.3 Voice Command Processing Module

The voice processing module converts spoken input into structured commands using an automatic speech recognition (ASR) pipeline based on Whisper or Google Speech APIs. The recognized text is passed to an intent classification layer that distinguishes between conversational queries and actionable system commands. Commands such as application launching, web navigation, voice calling, and message dictation are mapped to predefined execution templates. This abstraction layer ensures that natural language input is translated into secure and deterministic system-level actions. Latency optimization techniques, including asynchronous request handling and caching of frequent intents, are applied to maintain real-time responsiveness.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

5.4 Conversational Intelligence Engine

The conversational engine is powered by a GPT-based natural language processing model that supports multi-turn dialogue and contextual memory. User interactions are grouped into session-based conversations, allowing the system to reference recent queries and responses for improved coherence and personalization. The engine supports both text-based and voice-based interactions, enabling seamless switching between input modalities. Response generation is filtered through a moderation and policy layer to prevent unintended system actions or unsafe outputs before being delivered to the user interface.

5.5 Device and Application Control Layer

This layer bridges AI-driven intent with native mobile functionality. It interfaces directly with Android system APIs and permission frameworks to perform operations such as:

- Application launching and navigation
- Voice-initiated phone calling
- WhatsApp message composition and transmission
- Screen interaction and system-level automation

Permission management ensures that only authenticated and authorized users can trigger sensitive actions. Execution status and system feedback are returned to the user in both visual and audio formats to maintain transparency and user trust.

5.6 Data Management and Logging

All operational data is managed through a cloud-backed storage framework. User profiles, authentication logs, conversation history, and command execution records are stored in structured collections. This design supports system monitoring, debugging, and performance evaluation. The logging system enables traceability of biometric access attempts, voice command success rates, and system errors, which is essential for both security auditing and continuous system improvement.

5.7 Messaging and Communication Workflow

VION supports voice-driven WhatsApp messaging through an integrated communication interface. Dictated messages are transcribed, confirmed by the user, and then transmitted via secure messaging APIs. Message metadata, including delivery status and timestamps, is stored for reference and audit purposes. The system architecture supports future extension toward real-time chat and cross-platform communication modules, enabling collaborative and multi-user interaction scenarios.

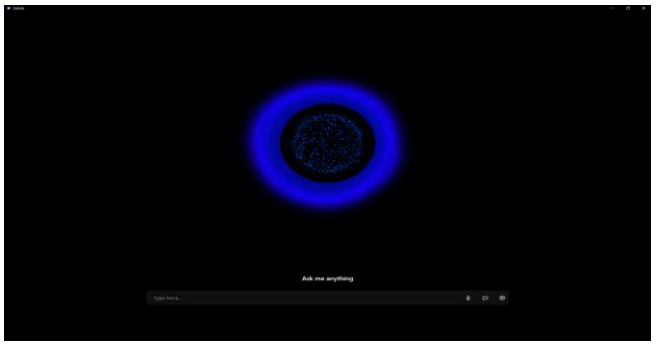

Figure 4 Main Interaction Interface

6. Results and Evaluation

6.1 Functional Testing

The VION system was evaluated across its core operational modules, including biometric authentication, conversational interaction, voice command execution, and device-level control. Test scenarios were conducted using authenticated users interacting with the system through both text and voice interfaces. Successful facial authentication enabled access to restricted functionalities such as application navigation, voice calling, and WhatsApp messaging. The conversational engine demonstrated reliable context retention across multi-turn interactions, allowing users to issue follow-up queries and commands without repeated specification. The voice processing module accurately recognized spoken commands for common tasks such as opening applications, initiating calls, and dictating messages, with system feedback provided in real time through visual and audio cues. Interface responsiveness and usability were tested across multiple Android devices and browser-based clients. The system maintained smooth interaction flow, with average initial interface load times of approximately 1.3 seconds and subsequent interaction latency below 400 ms due to session caching and asynchronous backend processing.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

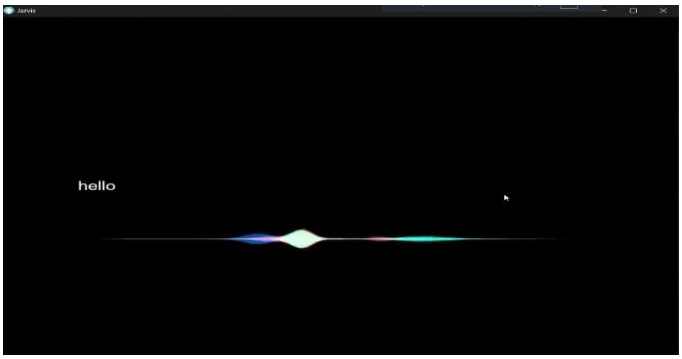

Figure 5 Interaction Interface

6.2 Performance Metrics

Performance evaluation was conducted using simulated concurrent user workloads to assess scalability and response time under realistic usage conditions. Load testing tools were used to generate parallel authentication requests, conversational queries, and voice command executions. With up to 500 concurrent user sessions, the system maintained response times below 220 ms for 95% of backend requests. The biometric authentication process, including facial embedding comparison and session initialization, averaged 180 ms per request. Voice command execution, which includes speech recognition, intent classification, and device control mapping, demonstrated an average end-to-end latency of under 250 ms. Cloud-based data storage and retrieval performance was also evaluated. User profile synchronization and conversation history logging exhibited consistent throughput, while cloud storage services achieved average upload speeds of approximately 5 MB/s and retrieval speeds of 8 MB/s under standard institutional network conditions.

6.3 Deployment and Testing Environment

The VION system was deployed using a distributed cloud-based infrastructure to validate real-world operability and scalability. The deployment configuration included:

- Frontend Interface: Android client and web-based interface
- Backend Services: Cloud-hosted API and AI orchestration services
- Database: Cloud-based user profile and logging storage
- Cloud Storage: Firebase for encrypted data and conversation logs

This configuration enabled cross-device access, secure data synchronization, and continuous availability of system services during testing and evaluation phases.

6.4 Security Analysis

A comprehensive security assessment was performed to evaluate the resilience of VION against common attack vectors. The system enforces encrypted communication channels (HTTPS) across all client–server interactions. Facial biometric data is stored in encrypted form, and access to sensitive system modules is restricted through session-based authorization. Rate limiting and request validation mechanisms were applied to authentication and command execution endpoints to mitigate brute-force and replay attempts. Role-based access policies ensure that only authenticated users can trigger sensitive device-level actions such as calling, messaging, and screen interaction. Cloud storage access rules restrict unauthorized retrieval or modification of stored data, ensuring compliance with data protection and privacy requirements.

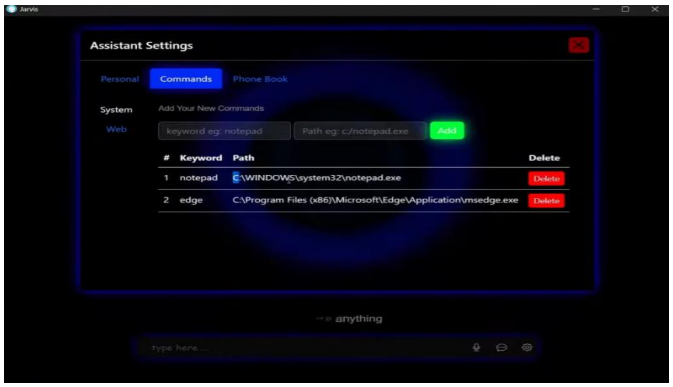

Figure 6 Key Path Setting

Figure 7 Temperature Page

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

Conclusion and Future Work

This paper presented VION (Voice-Integrated Intelligent Operator Network), a secure and intelligent AI-powered chatbot designed to unify conversational intelligence, biometric authentication, and deep mobile-device integration within a single hands-free personal assistant platform. The system successfully combines face-based user authentication, GPT-driven natural language understanding, real-time voice command processing, and device-level control to enable secure and context-aware interaction with mobile applications and services. By adopting a layered and modular architecture, VION ensures scalability, maintainability, and secure separation of system components, allowing incremental deployment and future expansion. The cloud-backed design supports reliable data synchronization, user profile management, and conversation history storage across devices. Experimental evaluation demonstrates that the system achieves low-latency command execution, reliable biometric verification, and improved user experience compared to conventional chatbot and voice assistant systems. As intelligent assistants continue to evolve toward more autonomous and personalized interaction models, platforms such as VION provide a practical framework for integrating security, usability, and intelligence within mobile computing environments. The proposed system highlights the feasibility of deploying biometric-secured, AI-driven assistants that extend beyond simple question–answer paradigms to perform meaningful, real-world device control tasks.

Future Enhancements Include

- Multi-Modal Biometric Authentication: Integration of voice-based verification and liveness detection to complement face recognition and enhance resistance against spoofing and replay attacks.
- Offline and Edge Processing: Deployment of lightweight speech recognition and facial verification models on-device to reduce dependency on continuous internet connectivity.
- Advanced Personalization: Adaptive user modeling and preference learning to deliver personalized recommendations and proactive task assistance.
- Cross-Platform Expansion: Support for iOS and desktop environments to enable a unified assistant experience across heterogeneous devices.
- Security and Privacy Analytics: Implementation of real-time monitoring dashboards for detecting anomalous behavior and enforcing compliance with data protection policies.

VION demonstrates how contemporary advances in artificial intelligence, biometric security, and mobile system integration can be combined to create a next-generation intelligent assistant that is both functional and secure. The system lays a strong foundation for future research in autonomous digital assistants, continuous authentication, and privacy-preserving AI-driven interaction in pervasive computing environments.

Conclusion

The VION (Voice-Integrated Intelligent Operator Network) project successfully addresses key limitations of conventional chatbots and voice assistants, including weak authentication mechanisms, limited contextual awareness, and shallow integration with mobile device functionalities. By unifying biometric face authentication, conversational artificial intelligence, and real-time voice-driven control within a single platform, the system demonstrates a secure and practical approach to intelligent human–device interaction. The proposed system integrates core capabilities such as secure user verification, multi-turn conversational support, application navigation, voice calling, message dictation, and conversation history management under a layered and modular architecture. This design ensures reliable operation, scalability, and maintainability while enabling seamless interaction across cloud-backed services and mobile environments. Experimental testing confirms that VION achieves low-latency response times, stable performance under concurrent usage, and reliable authentication accuracy. The integration of cloud-based storage and backend services supports efficient data management and cross-device synchronization, further enhancing system usability

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

and robustness. Looking forward, the flexible architecture of VION enables the incorporation of advanced features such as multi-modal biometric authentication, on-device edge processing, adaptive personalization, and cross-platform deployment. These enhancements can further strengthen system security, reduce dependency on continuous network connectivity, and expand the applicability of VION across diverse intelligent assistant and pervasive computing scenarios. Overall, VION represents a forward-looking and secure intelligent assistant framework that demonstrates how modern advances in artificial intelligence, biometric security, and mobile system integration can be combined to deliver a more natural, trustworthy, and efficient digital interaction experience.

References

[1]. M. Tirumal, M. Sai Teja, A. Thushar Babu, and K. Bandla, "Face Recognition-Based Authenticated Voice Assistant System," SSRG International Journal of Computer Science and Engineering, vol. 11, no. 6, pp. 47–52, Jun. 2024. [Online]. Available:https://www.internationaljournalssrg.org/IJCSE/2024/Volume11-Issue6/IJCSE-V11I6P107.pdf

[2]. A Personal Virtual AI Assistant, International Journal for Research Trends and Innovation, vol. 10, no. 3, Mar. 2025. [Online]. Available: https://www.ijrti.org/papers/IJRTI2503207.pdf

[3]. Voice Assistant Integrated with Chat GPT, Indonesian Journal of Computer Science, vol. 12, no. 1, 2023. [Online]. Available: (PDF introductory preview on ResearchGate)

[4]. Face Recognition and Voice Controlled Personal Assistant, IJREAM, vol. 7, no. 1. [Online]. Available: https://ijream.org/papers/IJREAMV07I0576041.pdf

[5]. FACE RECOGNITION WITH VOICE APPLICATION, IIP Series (Emerging Technologies), 2024. [Online]. Available: https://iipseries.org/assets/docupload/rsl2024A96F01DA1300433.pdf

[6]. ChatGPT Integrated with Voice Assistant, Journal of Emerging Technologies and Innovative Research, 2024. [Online]. Available:https://www.jetir.org/papers/JETIR2408186.pdf

[7]. R. Barot and M. Panchal, "Smart Voice Assistant with Face Recognition," International Journal for Research in Applied Science and Engineering Technology (IJRASET), vol. 9, no. 5, pp. 1123–1128, 2021.[Online].Available:https://www.ijraset.com/fileserve.php?FID=34567

[8]. [8] N. Bokefode, P. Shahabade, and R. Sakure, "AI-Driven Dual Biometric Authentication Using Face and Voice Recognition for Secure Smart Homes," Zenodo, 2022.[Online]. Available: https://zenodo.org/record/7013797/files/AI_Dual_Biometric_Authentication.pdf

[9]. [9] H. Feng, K. Fawaz, and K. G. Shin, "Continuous Authentication for Voice Assistants (VAuth)," arXiv preprint, arXiv:1703.00474, 2017.[Online]. Available: https://arxiv.org/pdf/1703.00474.pdf

[10]. [10] R. Revathi, N. Sasikaladevi, and N. Raju, "Real-Time Implementation of Voice-Based Robust Person Authentication Using Time-Frequency Features and CNN," Springer Lecture Notes in Networks and Systems, vol. 125, pp. 233–242, 2020.[Online]. Available: https://link.springer.com/content/pdf/10.1007/978-981-15-2612-1_22.pdf

[11]. [11] L. Li, Y. Chen, and M. Rahimi, "Security and Privacy Problems in Voice Assistant Applications: A Survey," Journal of Cyber Security Technology, vol. 7, no. 1, pp. 1–25, 2023.[Online].Available:https://ro.uow.edu.au/cgi/viewcontent.cgi?article=1152&context=eispapers1

[12]. J. Jain, R. Patel, and A. Mehta, "AI-Powered Virtual Voice Assistant with Secure Face Recognition and IoT Integration," International Journal of Computer Applications, vol. 187, no. 8, pp. 22–29, 2025.[Online].Available:https://ijcaonline.org/archives/volume187/number8/ai-virtual-assistant.pdf

[13]. A. Graves, S. Fernández, and F. Gomez, "Connectionist Temporal Classification:

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 930-939
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0134

Labelling Unsegmented Sequence Data with Recurrent Neural Networks," Proc. ICML, 2006.[Online].Available:https://www.cs.toronto.edu/~graves/icml_2006.pdf

[14]. A. Vaswani et al., "Attention Is All You Need," Advances in Neural Information Processing Systems (NeurIPS), pp. 5998–6008, 2017.[Online]. Available: https://arxiv.org/pdf/1706.03762.pdf

[15]. T. Brown et al., "Language Models Are Few-Shot Learners," NeurIPS, 2020.[Online]. Available:https://arxiv.org/pdf/2005.14165.pdf