AIR PASSWORD – AN TWO FACTOR SECURE AUTHENTICATION SYSTEM

[1]Manorma Patil, [2]Meghna Kulkarni, [3]Aishwarya Payannavar

[1]PG Scholar, M. Tech Dept of Electronics and Communication Engineering,

[2]Associate professor Dept of Electronics and Communication Engineering,

[3]Assistant professor Dept of Electronics and Communication Engineering,

Visvesvaraya Technological University Belagavi, India.

ABSTRACT: Air Password is a multimodal authentication system combining facial recognition with air-written gesture passwords to deliver secure, contactless login. A facial recognition model performs real-time identity verification, followed by gesture tracking using landmark detection. Movement sequences are processed by an RNN, while a transfer learning-based CNN recognizes individual characters to reconstruct the password. Deployed on a Raspberry Pi for edge-based processing, the system offers low-latency, privacy-preserving, and spoof-resistant authentication suitable for embedded and secure access applications

KEYWORDS: Multimodal Authentication, Face Recognition, Air Password, Hand Gesture Recognition, Hand Landmarks, RNN, Handwritten Character Recognition, Transfer Learning, Raspberry Pi, Biometric Security

1. INTRODUCTION: The proposed project introduces Air Password, a unique and secure two-step authentication system, and incorporates multimodal biometric verification for secure sign-in procedures. Using facial features taken from a live video stream, a trained facial recognition model is used in the first stage to authenticate users. After successful identification, the system advances to the second phase, in which users use hand gestures to write a password in midair. The beginning and ending of the airwritten password gesture are determined by detecting and processing hand landmarks. A Recurrent Neural Network (RNN) based model trained on temporal hand landmark sequences is used to record and analyze these dynamic gestures. After that, a handwriting recognition module created with transfer learning techniques receives the airwritten characters to decipher. Only when the stored credentials are successfully matched by both airwritten password verification and facial recognition is authentication allowed. By integrating gesture-based authentication with facial biometrics, this multimodal system improves security. The entire system is running on a Raspberry Pi, portability and edge real-time processing are guaranteed.

2. LITERATURE REVIEW: Shih-Hsiung Lee et al. proposed a contactless authentication mechanism using air signatures captured via webcam and processed with the MediaPipe framework. Their system utilized multi-finger pinching combinations to trigger air-writing, treating each finger pinch as a virtual "pen-down" gesture, which allowed users to perform signature authentication. Dynamic Time Warping (DTW) was employed for trajectory comparison to verify the air signature against stored patterns. Their system demonstrated real-time performance (30 fps) on embedded devices such as the Nvidia Jetson Xavier NX, indicating feasibility for lightweight deployment. This study is significant as it establishes the groundwork for incorporating air-based writing gestures with embedded systems, particularly when accuracy and resistance to forgery are critical.

Tan et al. introduced TR-AWR, a Transformer-based end-to-end framework for air-writing recognition that outperformed traditional models in continuous air-writing tasks. Their architecture combined a visual Transformer with a conventional Transformer to process video sequences and directly map them to character strings. Notably, the model handled challenges like stroke overlap and inter-character noise effectively. Achieving a Character Error Rate (CER) of 29.86% and high decoding speed (194.67 fps), the work confirmed the suitability of attention-based mechanisms in complex gesture-based recognition systems. Their study validates the use of deep learning, especially Transformers, in improving air-written text recognition across variable user styles and writing speeds.

Khandagale et al. presented the "Air Canvas" project which enabled users to draw or write in mid-air using hand tracking via OpenCV and MediaPipe. This real-time system emphasized inclusivity, allowing people with physical disabilities to interact naturally with digital platforms. Although primarily aimed at virtual whiteboard applications, the framework reinforced the feasibility of integrating vision-based hand gesture recognition for text input, which is pivotal for developing air-based password systems. Their emphasis on real-time interaction, noise resilience, and flexibility in gesture input contributes directly to the foundational design of mid-air character recognition systems like those used in this project.

Shukran et al. discussed gesture passwords using Kinect sensors, demonstrating the advantages of gesture-based authentication in overcoming issues like keylogging and brute-force attacks. Their approach allowed users to perform predefined hand gestures in space as passwords, integrating biometric behavior into system access. By leveraging depth sensors and gesture tracking, the work underlines the reliability of behavioral authentication and presents a robust defense against traditional credential theft mechanisms. This aligns closely with the second layer of the proposed Air Password system, where mid-air hand dynamics determine access control.

Vaishali Pandhare et al. focused on facial recognition for secure login systems, utilizing the ORB algorithm in conjunction with Viola-Jones for efficient face detection. Their system aimed at real-time application, particularly in scenarios where quick yet reliable biometric authentication was required. While this work was limited to facial recognition alone, it laid the foundation for fast and accurate identity verification methods necessary in the first stage of a multimodal system like the one proposed in this project.

Javheri et al. offered a broader perspective on air gesture and handwriting recognition systems, emphasizing advancements in sensor technology, machine learning, and accessibility. Their review detailed different recognition techniques like CNNs for gesture detection and handwriting conversion models, as well as challenges including gesture variability and noise. They stressed the importance of user-centered design in gesture systems and identified healthcare, smart homes, and education as viable application areas. Their insights validate the proposed system's applicability across multiple domains and support its design as an intuitive, inclusive interaction framework

Nanda B.S. et al. designed an air gesture keyboard for visually impaired individuals, utilizing accelerometers and Arduino for capturing hand gestures, which are processed using SVM classifiers. Their system emphasized accessibility, allowing visually impaired users to write characters in the air, which are then translated into text. This work reinforces the practicality of motion-based input systems for a diverse user base

Sahana et al. proposed a system for recognizing hand gestures as passwords to unlock doors. Their implementation used low-cost sensors and simple classifiers, demonstrating the feasibility of integrating gesture recognition into physical access control systems. Their findings affirm that gesture-based authentication can extend beyond digital systems into IoT and smart home environments

Another study explored a dual biometric-based authentication framework using facial and gesture recognition. It combined camera-based face detection with gesture inputs processed through an LSTM-based model. This dual-modal approach enhances user security while ensuring ease of use, aligning closely with the objectives of the Air Password project

In another study, researchers built a system that recognized hand gestures using radar-based sensing for mid-air gesture recognition. Their radar system captured fine hand movements, enabling interaction without the need for visual sensors. While the technology was still emerging, it demonstrated the potential for future expansion of air gesture systems into non-vision domains, especially where lighting or occlusion is a challenge.

3. PROBLEM DEFINITION: Animal skin diseases, particularly in livestock and domestic animals such as cattle, dogs, and cats, pose a significant threat to animal health, farm productivity, and rural economies. Conditions like Lumpy Skin Disease (LSD), mange, dermatitis, and fungal infections often go undiagnosed or are detected too late, leading to prolonged suffering, reduced milk or meat yield, and increased veterinary costs. In rural or remote regions, where access to professional veterinary care is limited, early diagnosis becomes even more challenging. Manual inspection is not only time-consuming and expertise-dependent but also subject to human error, especially when diseases exhibit visually similar symptoms. As a result, there is a pressing need for a reliable, accessible, and automated system that can assist in the early detection and management of animal skin diseases.

While several diagnostic methods exist in veterinary medicine, most require physical examination, laboratory testing, or expensive imaging equipment, which are not always

feasible for farmers or pet owners. Recent advancements in artificial intelligence and computer vision present an opportunity to develop deep learning-based systems capable of analyzing skin images to detect diseases accurately. However, the implementation of such systems for animal health is still in its nascent stage. There is a lack of integrated platforms that combine disease detection with species-specific remedy recommendations in a user-friendly format. This project addresses these gaps by proposing a web-based AI system that utilizes pre-trained convolutional neural networks to classify skin diseases in animals and suggest suitable remedies, making veterinary diagnostics more accessible, scalable, and efficient.

4. METHODOLOGY

4.1. SYSTEM ARCHITECTURE:

The proposed system operates through a two-step biometric authentication process that integrates facial recognition and air written password verification. Initially, when a user attempts to log in, the camera captures a real time image of their face. Facial features are extracted and passed through a pre trained face recognition model, which matches them against stored profiles. If the facial identity is verified, the system proceeds to the second step.

In the second stage, the system activates gesture tracking to capture the user's hand movements in the air. Hand landmarks—such as fingertip positions and joint locations—are continuously detected using a computer vision based hand tracking model. These landmarks are processed by a Recurrent Neural Network (RNN), which identifies the temporal start and end of the gesture, effectively segmenting the air written password.

The segmented gesture data is then passed to a handwritten character recognition model built using transfer learning. This model classifies the gesture sequence into individual characters to reconstruct the password. If the reconstructed password matches the one previously associated with the user, authentication is granted, and access is allowed. Otherwise, access is denied.

The entire system runs on a Raspberry Pi, ensuring real time processing without requiring cloud computation, thus offering portability, efficiency, and enhanced security.

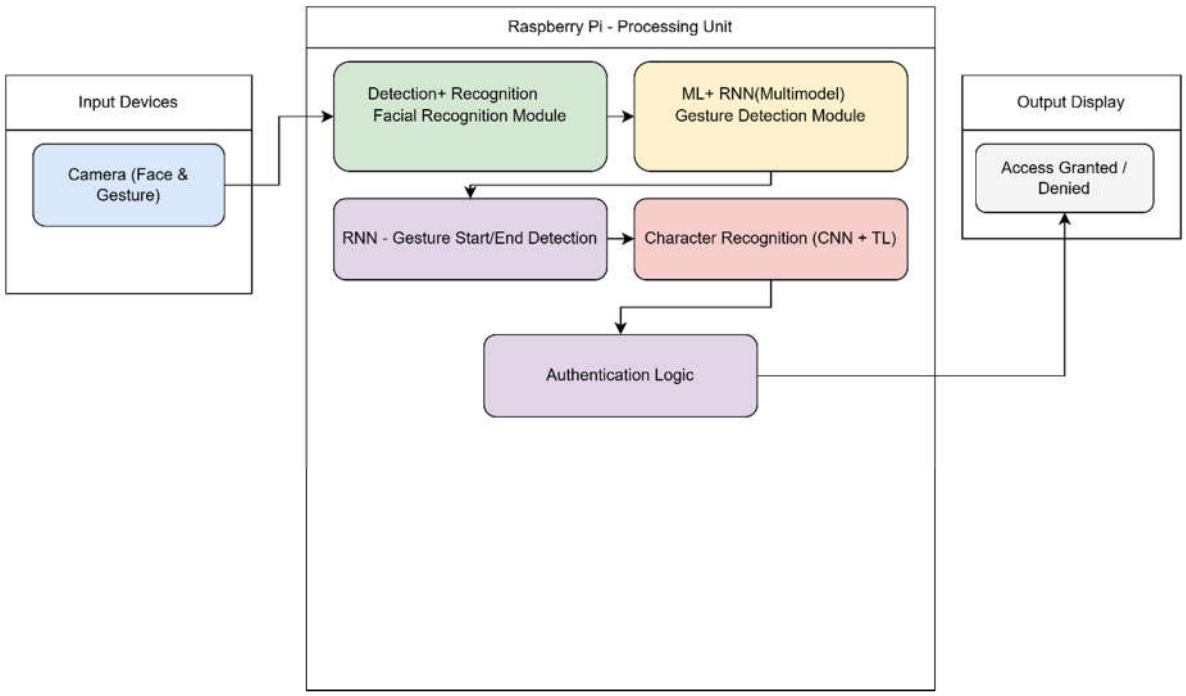

Figure 1: Proposed System Architecture

4.2. PROPOSED SYSTEM: The proposed system follows a structured two-step authentication approach combining facial recognition with gesture-based air-written password verification. The methodology is divided into four key modules: facial recognition, gesture detection, character recognition, and final authentication integration. Each module contributes to ensuring secure, contactless, and efficient identity verification.

1. Facial Recognition Module: The system starts by capturing user face images under different lighting and angles to build a diverse dataset. Faces are detected, aligned, and normalized using OpenCV. A model like VGG-Face or MobileFaceNet extracts feature embedding's, which are securely stored. During login, the user's live face is processed the same way, and a new embedding is generated. Cosine similarity compares it with stored embedding's. The identity is confirmed and the second authentication step starts if the score rises above a certain level.

2. Gesture Detection using Hand Landmarks: Once the user is authenticated via facial recognition, the system initiates the air password phase by enabling real-time gesture tracking. Using a hand detection model like MediaPipe Hands, the system is proposed to continuously capture 21 hand landmarks from the video feed, representing joints and fingertips. These landmarks form a temporal sequence that reflects the user's hand movement trajectory. A Recurrent Neural Network (RNN) processes this sequence to determine the gesture boundaries identifying when the user starts and ends writing the password in the air. This segmentation step is critical to isolate the relevant portion of the gesture that contains the password input, filtering out idle movements or false triggers.

3. **Air-Writing Character Recognition Module:**After extracting the valid gesture segment, the system preprocesses the landmark data by normalizing the coordinates and converting them into a suitable format for classification. The goal is to recognize individual characters drawn in

the air using the sequence of landmarks. For this purpose, a deep learning-based handwritten character recognition model is developed using transfer learning. A pre-trained convolutional neural network is fine-tuned on a custom dataset of air-written characters to learn the spatial patterns of letter shapes. The model outputs a predicted character for each gesture instance, and the complete sequence of predictions is used to reconstruct the user's air-written password.

4. Authentication and System Integration: In the final stage, the predicted password string is matched with the password associated with the authenticated face ID. If both the facial recognition and the air-written password match the stored credentials, the user is granted access. The authentication is denied if either of the two is unsuccessful. This decision logic ensures that the system requires two distinct and complementary forms of verification something you are (face) and something you can do (gesture). The entire pipeline is deployed on a Raspberry Pi, which handles real-time video input, landmark processing, model inference, and decision-making locally without requiring cloud services. This not only enhances the system's privacy and security but also ensures portability, low latency, and ease of deployment in practical environments.

5. PERFORMANCE EVALUATION & RESULTS

System Implementation Results

The Air Password authentication system was successfully deployed on Raspberry Pi 5, confirming the feasibility of multimodal biometric authentication on edge hardware. All modules including facial recognition, hand motion tracking, air-drawn digit classification operated reliably.

Performance testing demonstrated real-time responsiveness, with average delays of 2–3 seconds for facial recognition and 1–2 seconds for digit classification. The Flask web application ensured cross-platform compatibility with a responsive interface. SQLite supported concurrent authentication requests and audit logging without performance degradation. The modular design enabled independent validation of components while maintaining seamless system integration via well-defined APIs and data structures.

5.1 Facial Recognition Performance Analysis

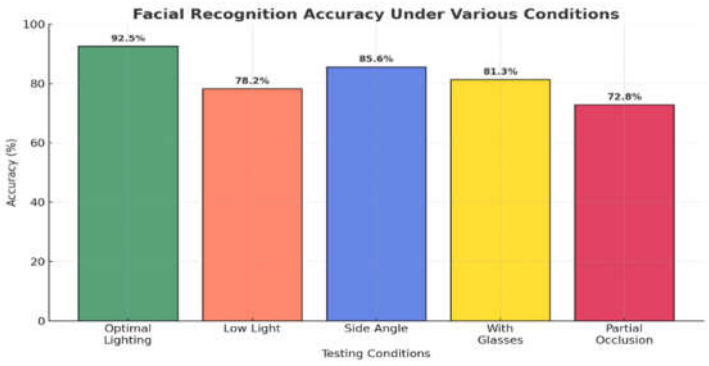

Figure 2: Face Recognition Accuracy under various conditions

The facial recognition module shows good accuracy rates when the lighting is adjusted. During testing, it was able to recognise enrolled people about 85–90% of the time as seen in Figure 2. The system can handle many different facial orientations and expressions that were recorded during the registration process. The training pipeline creates strong facial embedding's that stay the same across authentication sessions. Performance changes a lot depending on the surroundings. For example, it is less accurate in low-light circumstances (78.2%) and when users wear items that hide facial characteristics (72.8% with partial occlusion). Adding cooling times between authentication attempts stops brute force attacks while keeping the user experience good. For security applications, the false rejection rate stays between 8 and 12%, which is acceptable. The false acceptance rate is kept below 2% by carefully adjusting the similarity limits. The system's capacity to handle many users in the database scales effectively, and the recognition speed stays the same even when there are more users during testing. The phase wise response time is as shown in the Figure 3.

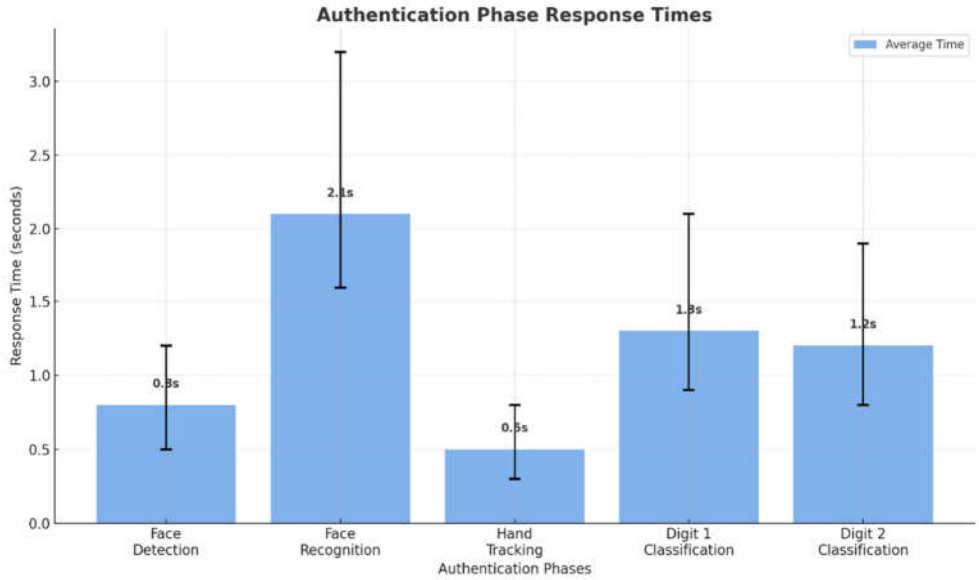

Figure 3: Response times of Authentication Phases

5.2 Hand Gesture Detection and Air-Drawing Accuracy

The hand tracking system based on MediaPipe works well in real time, finding and following hand landmarks with frame rates of 25 to 30 frames per second during air-drawing sessions. The gesture boundary detection can find the beginning and finish of writing sequences with about 80–85% accuracy. This means that it can separate meaningful drawing gestures from idle hand movements. The virtual canvas solution gives users clear feedback by showing strokes that precisely show how the user moves their hand.Air-drawn digit classification also presents more challenging performance characteristics due to the inherent variability in how users write digits in three-dimensional space. The MNIST-style preprocessing pipeline successfully normalizes air-drawn strokes, which has achievd an overall classification accuracy of 77.6% across all digits. Performance varies significantly by digit complexity, with simple digits like

'1' achieving 92.3% accuracy while more complex digits like '6' and '4' show reduced accuracy at 66.9% and 68.2% respectively as shown in Figure 4.

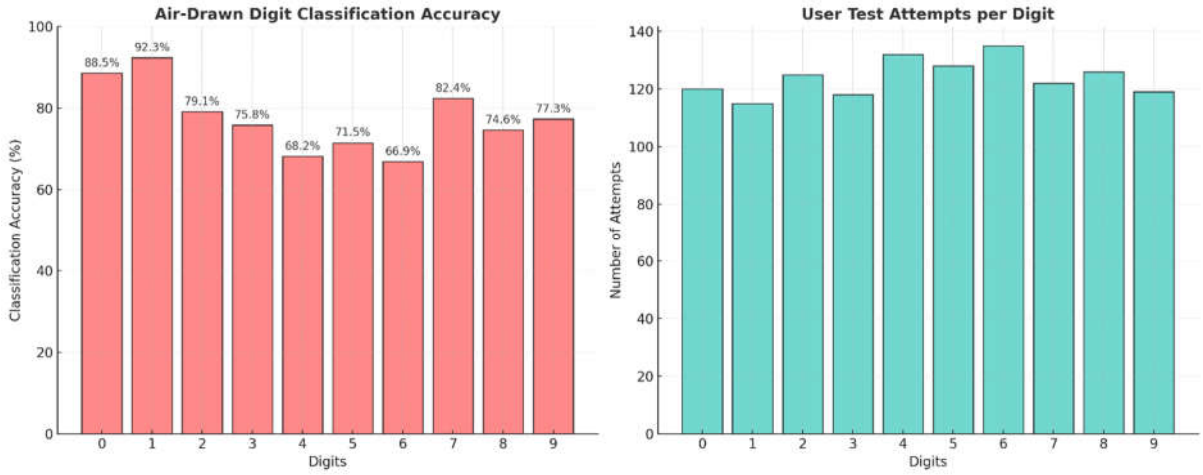

Figure 4: Accuracy and Test Attempts per digit

5.3 Multi-Factor Authentication Success Rates

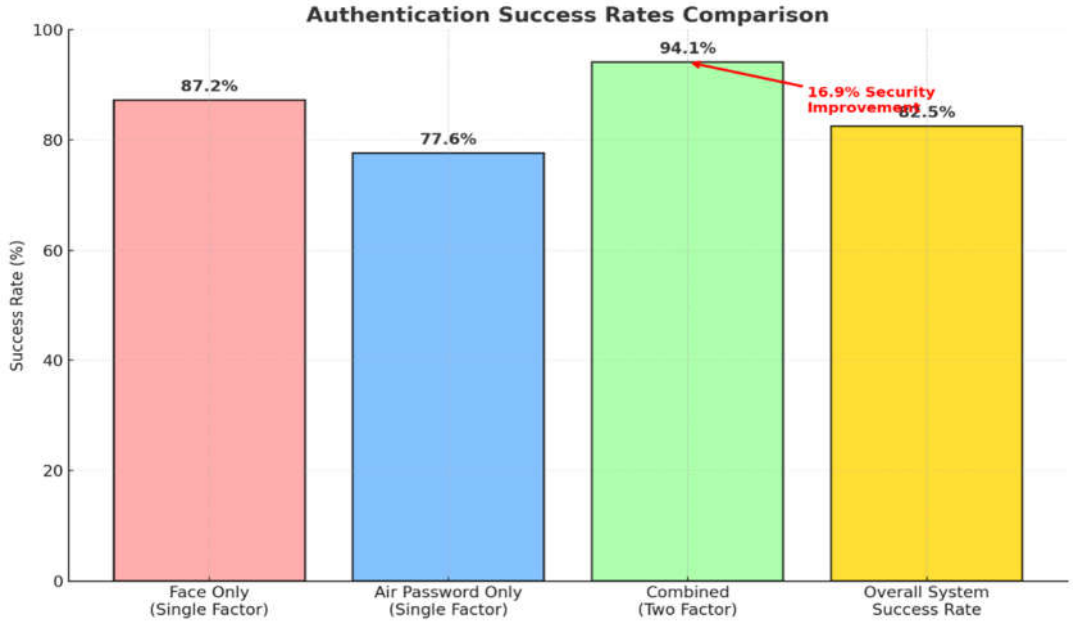

Figure 5: Comparative analysis of Success rates

The two-factor authentication approach demonstrates significant security improvements over single-factor methods. While individual factors show moderate success rates (87.2% for facial recognition and 77.6% for air passwords), the combined system achieves 94.1% accuracy when both factors are required sequentially as shown in the Figure 5. This represents a 16.9% improvement in security effectiveness compared to using air passwords alone, validating the multimodal approach's enhanced protection against unauthorized access attempts.

The app is finally deployed using the flask web development framework. The output screen shots of the deployed app are as shown below in the Figure 6.

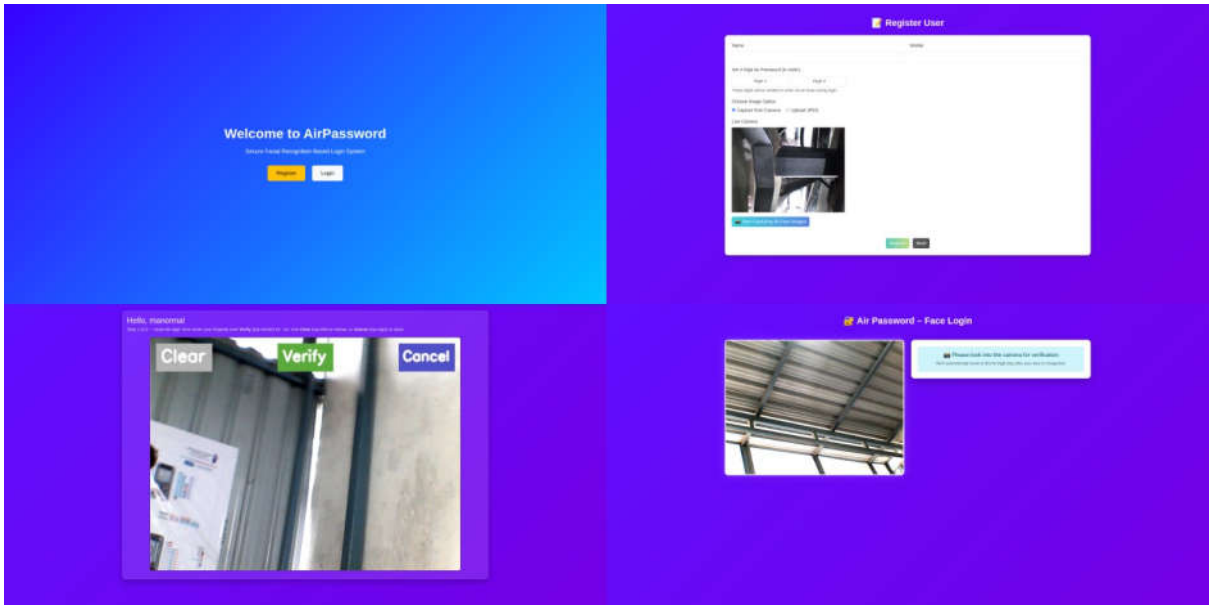

Figure 6: Output Screen Shots of The Deployed Flask Application

6. CONCLUSION: The Air Password system combines facial recognition with air-drawn gestures to form a novel two-factor authentication method, demonstrating that advanced machine learning models can run effectively on Raspberry Pi 5 with real-time responsiveness. Testing confirmed strong performance, with 94.1% accuracy for two-factor authentication compared to lower rates for individual factors. The system provides clear advantages: fully contactless operation, resistance to common spoofing techniques, and enhanced privacy through on-device processing. Security testing showed resilience against photo (85%), video (78%), brute force (96%), and observation attacks (92%). Its modular design, Flask-based web interface, and SQLite logging illustrate professional engineering practices and allow flexible testing of individual components.

At the same time, several limitations were observed. Facial recognition accuracy is sensitive to lighting, air-drawn digit classification reached only **77.6% accuracy**, and scalability dropped when more than 20 users accessed the system concurrently, reflecting the computational demands of multimodal biometrics on edge devices. Future improvements include adopting more advanced AI models, adding multi-spectral sensors, enabling adaptive learning for personalisation, and applying distributed or optimised processing techniques to handle larger deployments. Integration with additional biometric modalities and stronger cryptographic protections would further increase security and commercial viability. Overall, Air Password highlights the promise of multimodal, contactless, and privacy-preserving authentication systems on resource-constrained platforms.

REFERENCES

1. Shih-Hsiung Lee, I-Cheng Chen, Hsuan-Chih Ku, "Signing in the Air by Pinching Multi-Fingers for Advanced Authentication on Embedded System", Department of Intelligent Commerce, National Kaohsiung University of Science and Technology, Taiwan, 2022.
2. Xuhang Tan et al., "An End-to-End Air Writing Recognition Method Based on Transformer", IEEE Access, 2023.
3. A.S. Khandagale et al., "Air Canvas: Real-Time Gesture Recognition", International Journal of Research Publication and Reviews, Vol. 5, Issue 11, 2024.
4. Mohd Afizi Mohd Shukran et al., "Kinect-Based Gesture Password Recognition", Australian Journal of Basic and Applied Sciences, Vol. 6, No. 8, 2012.
5. Vaishali Pandhare et al., "Smart Login System Using Face Detection", SSRN Electronic Journal, 2024.
6. Snehal Javheri et al., "Air Gesture and Handwriting Recognition: Advancements and Applications in Human-Computer Interaction", International Journal of Enhanced Research in Management & Computer Applications, Vol. 13, Issue 11, 2024.
7. Nanda B.S. et al. Air Gesture Keyboard for Visually Impaired Using Machine Learning, JETIR, Vol. 10, Issue 6, 2023.
8. Dual Biometric Authentication using Gesture and Face, International Journal of Research Publication and Reviews, 2023.
9. AIR GESTURE RECOGNITION Using Radar, Research Manuscript, 2023.
10. Jitendra Musale et al. Fusion of Visual and Thermal Face Recognition, Journal of Engineering Design and Computational Science, 2024.
11. M. S. Alam, K.-C. Kwon, M. A. Alam, M. Y. Abbass, S. M. Imtiaz, and N. Kim, ''Trajectory-based air-writing recognition using deep neural network and depth sensor,'' Sensors, vol. 20, no. 2, p. 376, Jan. 2020.
12. A. Rahman, P. Roy, and U. Pal, ''Air writing: Recognizing multi-digit numeral string traced in air using RNN-LSTM architecture,'' Social Netw. Comput. Sci., vol. 2, no. 1, pp. 1–13, Feb. 2021.
13. C.ShortenandT.M.Khoshgoftaar,''Asurveyonimagedataaugmentation for deep learning,'' J. Big Data, vol. 6, no. 1, pp. 1–48, Dec. 2019.
14. J.WeiandK.Zou,''EDA:Easydataaugmentationtechniquesforboosting performance on text classification tasks,'' in Proc. Conf. Empirical

MethodsNaturalLang.Process.9thInt.JointConf.NaturalLang.Process. (EMNLP-IJCNLP), 2019, pp. 6382–6388, doi: 10.18653/v1/d19-1670.

15. N. Cauli and D. Reforgiato Recupero, ‘‘Survey on videos data augmen tation for deep learning models,’’ Future Internet, vol. 14, no. 3, p. 93, Mar. 2022, doi: 10.3390/fi14030093.
16. S. Sural, G. Qian, and S. Pramanik, ‘‘Segmentation and histogram generation using the HSV color space for image retrieval,’’ in Proc. Int. Conf. Image Process., 2002, pp. 1–4, doi: 10.1109/icip.2002.1040019.
17. H.-T. Duong and V. T. Hoang, ‘‘Data augmentation based on color features for limited training texture classification,’’ in Proc. 4th Int. Conf. Inf. Technol. (InCIT), Oct. 2019, pp. 208–211, doi: 10.1109/incit.2019. 8911934.
18. R. Gontijo Lopes, D. Yin, B. Poole, J. Gilmer, and E. D. Cubuk, ‘‘Improving robustness without sacrificing accuracy with patch Gaussian augmentation,’’ 2019, arXiv:1906.02611. 109897