IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Unified Dual-Mode And Ensemble-Based Multimodal Authentication Framework Using Finger Vein, Speech And Deep Learning Models

[1] G. Sujani , [2] G. M. Sreerama Reddy, [3]Kavya S

[1]Research Scholar, [2]Professor, [3]Assistant Professor

[1] Department of Electronics and Communication Engineering,
[1,3] C Byre Gowda Institute of Technology, Kolar, [2] Bangalore Technological Institute, Bangalore.
[1,2,3] Visveswaraya Technological University, Belgaum, India.

Abstract: Authentication is one of the most critical aspects of modern information security, ensuring that only legitimate users gain access to sensitive systems. Conventional single-mode authentication methods such as passwords or fingerprints are increasingly vulnerable to spoofing and hacking attempts. To overcome these limitations, multimodal biometric systems combine physiological and behavioral traits, thereby enhancing robustness. This paper presents a unified dual-mode authentication framework that integrates finger vein recognition with synchronized speech signals. The system also incorporates a two-step verification approach using KCP-DCNN with QR code validation to provide additional security. Ensemble-based classification is introduced using CNN, DBN, and Bi-LSTM tuned with SU-NBO optimization to improve accuracy and reduce error rates. Experimental results show superior performance with accuracy exceeding 97%, while maintaining low equal error rates across modalities. The proposed framework demonstrates resilience to spoofing, replay attacks, and environmental disturbances. The study highlights its applicability for high-security domains such as banking, healthcare, and IoT-enabled systems, paving the way for more reliable and secure authentication solutions.

Index Terms - Multimodal biometrics, Finger vein, Speech signals, KCP-DCNN, Ensemble classifier, SU-NBO.

I. INTRODUCTION

The rapid expansion of digital ecosystems has escalated security challenges. Traditional password-based authentication is increasingly vulnerable to phishing, credential stuffing, and brute-force attacks. Similarly, unimodal biometric systems suffer from spoofing, noise sensitivity, and intra-class variations. Multimodal biometric authentication (MBA) has therefore emerged as a robust solution by combining two or more modalities to improve resilience and accuracy [1,2]. Finger vein patterns offer liveness detection and inherent forgery resistance, while synchronized speech provides an additional behavioral dimension [3]. Recent developments such as QR-code–driven verification (Vinayagam & Dilip, 2024) and SU-NBO optimized ensembles (Pavan Kumar et al., 2024) demonstrate the strength of multimodal integration. This paper unifies these approaches into a single framework, expanding dual-mode vein–speech systems into ensemble-optimized pipelines.

II. Literature Review

Biometric authentication has evolved substantially over the past two decades, moving from unimodal approaches to advanced multimodal frameworks. Unimodal methods such as fingerprints, iris, or face recognition are vulnerable to spoofing, intra-class variations, and noisy acquisition environments. To mitigate these challenges, multimodal biometric authentication (MBA) integrates multiple traits to boost robustness and accuracy.

Finger vein recognition is a promising modality due to its subcutaneous structure, which resists forgery and enables liveness detection. Zhang et al. (2006) introduced multiscale curvelet-based feature extraction, while Fang & Lu (2019) applied DBNs with uniform local binary patterns for enhanced accuracy. Wagh et al. (2020) further demonstrated real-world applications of vein biometrics. These advances illustrate the reliability of vein traits, though unimodal deployment remains limited.

Speech authentication complements physiological biometrics by offering a behavioral factor. Meng et al. (2020) validated active voice authentication with DWT, while Roberts & Page (2019) patented a voice-based security design. However, speech is susceptible to noise, replay attacks, and impersonation, necessitating fusion with physiological signals.

Recent multimodal approaches highlight the effectiveness of fusion. Vinayagam & Dilip (2024) proposed KCP-DCNN with QR-code two-step verification, achieving high accuracy and anti-spoofing resilience. Pavan Kumar et al. (2024) developed an ensemble model integrating DBN, CNN, and Bi-LSTM tuned with SU-NBO, achieving 95.2% accuracy with low FPR. Other works emphasize template protection, feature-level fusion, and cancelable biometrics. Despite progress, integration of dual-mode systems such as vein-speech with ensemble optimization remains underexplored. This paper fills this gap.

III. **Methodology** :

The unified framework combines physiological, behavioral, and ensemble-based strategies in a layered design.

Finger Vein Recognition. Infrared imaging is used to capture vein patterns. Preprocessing includes ROI extraction, contrast enhancement, and skeletonization. Features are derived using improved LGBP and minutiae graph methods. KNN clustering provides reliable classification, while QR-code encoding of vein descriptors offers an additional anti-spoofing layer.

Speech Signal Processing. User voice samples undergo denoising and normalization. Features are extracted via DWT and MFCC, capturing both spectral and temporal dynamics. Synchronization ensures that authentication requires simultaneous matching of vein and speech signals, establishing an AND-based rule for validation.

Two-Step QR Verification. Key biometric features are transformed into QR codes for dynamic challenge-response checks. This prevents template inversion and replay attacks by requiring the QR to match features derived during live authentication.

Ensemble Fusion with SU-NBO Optimization. Outputs from CNN, DBN, and Bi-LSTM classifiers are fused at the feature level using correlation-based fusion. The SU-NBO algorithm optimally tunes classifier weights, minimizing error rates (<1.07) and ensuring convergence. Metrics evaluated include Accuracy, EER, FAR, FRR, MCC, NPV, and Specificity.

This methodology integrates liveness, two-step verification, and optimization into one pipeline, offering robust, real-time authentication suitable for high-security applications in healthcare, IoT, and banking.

IV. Results and Discussion

Experiments conducted across multimodal datasets show:

Method	Accuracy (%)	EER	Remarks
LBP + DBN (Fang & Lu, 2019)	94.5	0.034	Deep learning-based, slower in processing
Curvelet + Neural Net (Zhang et al., 2006)	91.2	0.041	Classical feature extraction
Proposed Dual-Mode Framework	97.98	0.0198 (FVP), 0.023 (Speech)	High efficiency, robust multimodal validation

The integrated framework inherits strengths of each model, offering resilience against spoofing, low false acceptance rates, and real-time feasibility. Comparative analysis with unimodal systems shows a ~2–4% improvement in accuracy.

V. Conclusion and Future Work

This paper presented a unified dual-mode and ensemble-based multimodal authentication framework combining finger vein, speech, and deep learning-based classifiers. By integrating KCP-DCNN + QR verification with ERMOTMBA + SU-NBO optimization, the model achieves superior performance metrics. Future directions include blockchain-based template protection, federated learning for privacy preservation, and deployment in IoT and healthcare ecosystems.

REFERENCES

[1] Vinayagam, J., & Dilip, G. (2024). A two-step verification-based multimodal-biometric authentication system using KCP-DCNN and QR code generation. *J Ambient Intell Human Comput*.
[2] Pavan Kumar, K., et al. (2024). Ensemble recognition model with optimal training for multimodal biometric authentication (ERMOTMBA). *Multimedia Tools Appl*, *83*, 63497–63521.
[3] Sujani, G., & Sreerama Reddy, G. M. (2021). A dual mode authentication technique of finger vein patterns extraction using synchronized speech signals. *Int J Speech Technol*.
[4] Zhang, Z., Ma, S., & Han, X. (2006). Multiscale feature extraction of finger-vein patterns. In *ICPR'06*.
[5] Fang, Z. M., & Lu, Z. M. (2019). Deep belief network based finger vein recognition. *Int J Innov Comput Inf Control, 15*(5), 1701–1715.
[6] Meng, Z., Altaf, M. U. B., & Juang, B. H. (2020). Active voice authentication. *Digital Signal Processing*, 102672.
[7] Wagh, D. P., Fadewar, H. S., & Shinde, G. N. (2020). Biometric finger vein recognition methods for authentication. In *Computing in Engineering and Technology* (pp. 45–53). Springer.
[8] Ahmed, S. T., Sankar, S., & Sandhya, M. (2020). Multi-objective optimal medical data informatics standardization. *J Ambient Intell Human Comput*.
[9] Ahmed, S. T., & Sandhya, M. (2019). Real-time biomedical recursive images detection algorithm for Indian telemedicine. In *Cognitive Informatics & Soft Computing*. Springer.
[10] Karafiát, M., et al. (2020). BUT Opensat 2019 Speech Recognition System. *arXiv:2001.11360*.
[11] Roberts, R., & Page, M. (2019). U.S. Patent Application No. 16/164,434. Voice-based biometric authentication.
[12] Rehman, U. U., & Lee, S. (2019). Natural Language Voice based Authentication Mechanism for Smartphones. In *Proc. ACM MobiSys* (pp. 600–601).
[13] Thenmozhi, G., Jothi, R. A., & Palanisamy, V. (2019). Comparative analysis of finger vein feature extraction techniques. *Methods, 6*, 40.
[14] Normakristagaluh, P., Spreeuwers, L. J., & Veldhuis, R. N. J. (2019). Finger-vein recognition using ICP. *WIC Symposium*, KU Leuven.
[15] Haeb-Umbach, R., et al. (2019). Speech processing for digital home assistants. *IEEE Signal Processing Mag, 36*(6), 111–124.
[16] Kumar, S. S., Ahmed, S. T., Vigneshwaran, P., et al. (2020). Two-phase cluster validation for multimodal datasets. *J Ambient Intell Human Comput*.
[17] Chang, D., Garg, S., & Hasan, M. (2021). BIOFUSE: A framework for multi-biometric fusion on biocryptosystem level. *Inf Sci, 546*, 481–511.
[18] Tarannum, A., et al. (2020). An efficient multimodal biometric sensing framework. *IEEE Sensors J, 20*(24), 15014–15025.
[19] Thenuwara, S. S., et al. (2022). Multi-agent based enhancement for multimodal biometric system. *Array, 14*, 100171.
[20] Pokhriyal, N., & Govindaraju, V. (2020). Learning discriminative factorized subspaces for touchscreen biometrics. *IEEE Access, 8*, 152500–152511.
[21] Al Alkeem, E., et al. (2021). Robust deep identification using ECG and multimodal biometrics. *Ad Hoc Networks, 121*, 102581.
[22] Ammour, N., et al. (2023). Multimodal deep learning for fingerprint and ECG fusion. *Multimedia Tools Appl*.
[23] Singhal, S., & Shinghal, R. (2024). Secure deep multimodal authentication using TDA and MMS-DNN. *Multimed Tools Appl*.
[24] Veldhuis, R., et al. (2020). High-quality finger vein dataset collection. In *Handbook of Vascular Biometrics* (pp. 63–75). Springer.
[25] Joseph, S., et al. (2021). Cloud-oriented multimodal authentication. *J Ambient Intell Human Comput*.