International Journal of Science, Strategic Management and Technology
Volume **02** Issue **04** April-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Talentiq: An Intelligent Resume Intelligence & Fair Candidate Ranking System

Srivarshini I
Department of Computer Science and Engineering
CARE college of Engineering
Tamil Nadu, India
Email: geethaliyyappan771@gmail.com

Pillai Kirthika Kannan
Department of Computer Science and Engineering
CARE college of Engineering
Tamil Nadu, India
Email: kirthikapillai92@gmail.com

Tamilselvi P
Department of computer science and engineering CARE college of engineering
Tamilnadu, India
Email: tamilpandian20102004@gmail.com

Swathi.G
Department of Computer Science and Engineering
CARE college of Engineering
Tamil Nadu,India
Email: swathibalaji991@gmail.com

Ranitha R
Assistant Professor
Computer Science and Engineering
CARE College of Engineering
rranitha@care.ac.in

https://doi.org/10.55041/ijsmt.v2i4.024

Cite this Article: I, S., Kannan, P. K., P, T. & Swathi.G, (2026). Talentiq: An Intelligent Resume Intelligence & Fair Candidate Ranking System. International Journal of Science, Strategic Management and Technology, *02*(04). https://doi.org/10.55041/ijsmt.v2i4.024

License: This article is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting use, distribution, and reproduction in any medium, provided the original author(s) and source are properly credited.

Abstract - In this paper, the authors introduce a secure voice transaction system that uses deepfake technology to increase the validity of authentication in online financial transactions. Passwords and OTPs, which are commonly used as traditional authentication tools, are under increased cyber threat, such as voice spoofing attacks and AI-generated impersonation attacks. To solve these issues, the proposed system involves a two-layered security system that is based on a combination of biometric voice identification and deep learning-based fake voice recognition. In the verification of identity, Mel-Frequency Cepstral Coefficients (MFCC) are computed on the voice inputs of the user and the comparison between them is carried out by the use of cosine similarity. At the same time, a Convolutional Neural Network (CNN) is used to analyze Mel-spectrogram representations of the audio to identify it as an original or an artificial one. Approving a transaction requires fulfillment of background checks, as well as authenticity and deep fake checks. The system is installed on flask-based backend and voice capture over the browser. The experimental findings show better strength and resiliency to the contemporary voice-based fraud attacks.

Keywords - Voice Authentication, Deepfake Detection, MFCC, Cosine Similarity, Convolutional Neural Network, Mel-Spectrogram, Biometric Security, Voice Biometrics, Audio Spoofing Detection, Secure Transactions

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

I.INTRODUCTION

The high rates of development of digital financial systems have made the requirement of safe, trustful authentication tools to be highly in demand. The conventional passwords, PIN, and one-time passwords (OTP) are becoming more susceptible to hackers like phishing, credential theft, and social-engineering. Biometric authentication in reaction to this has come up as another viable substitute especially voice based systems since it is not only convenient, but it is also the only one. Voice biometrics uses vocal peculiarities to identify a person, and that is why this tool is applicable to real-time and remote systems. Recent developments in the field of artificial intelligence have thrown more security issues, in particular, deepfake voice results, which are able to successfully simulate legitimate users. Research has pointed out the increasing vulnerabilities of voice spoofing and the importance of a secure anti-spoofing system in authentication systems [1], [9]. Therefore, there has been a need to incorporate the advanced machine learning methods in the voice authentication systems to improve the level of reliability in the system to avoid fraudulent access.

To overcome these obstacles, this study suggests a safe voice-based transaction system that considers what is known as biometric authentication along with deepfake detection. The system uses Mel-Frequency Cepstral Coefficients (MFCC) and cosine similarity to check the identity of the user to enable good speech recognition. Simultaneously, a Convolutional Neural Network (CNN) is employed to process Mel-spectrograms of audio signals to identify synthetical or manipulated voices. The dual-layered model is more secure and authentic and resistant to fraud hence is a complete security solution too. Recent studies showed that deep learning models improve the overall performance of voice authentications on less training data and low false acceptance, despite the small data size [4], [10]. Moreover, privacy-friendly voice authentication and neural network speaker recognition are also developed thus enhancing the possibility of such a hybrid system [3], [5]. The proposed system will offer a scalable and effective solution to secure voice-based transactions by utilizing both conventional methods of signal processing and some recent methods of deep learning.

II.LITERATURE SURVEY

A. Voice-Based Authentication Systems

Voice-based authentication has acquired much popularity as an easy and secure biometric authentication method to the usual authentication methods. These systems work by using the distinctive vocal features to identify users in order to have free and non-contact access control. The initial studies aimed at creating the multi-factor authentication schemes that would utilize voice biometrics to increase the security of mobile and financial interfaces [1], [2]. It has been proven that techniques like voiceprint generation and recognition of the speaker by artificial neural network are more accurate and robust in identity verification [5]. Nevertheless, such systems are generally subject to environmental noisy issues, inter-speaker variability, and scanty training material. The new developments are intended to enhance the performance of systems by using the refined models and adaptive algorithms that help to make voice authentication more authoritative and scalable to be utilized in the real world in diverse fields.

B. Voice Authentication and Security and Privacy

As voice biometrics become more and more widely used, issues of security and privacy have been considered to be particularly critical research topics. Voice data is very sensitive and shouldn't be used by unauthorized people who may cause unimaginable consequences. As a solution to these issues, scholars have considered encryption algorithms and privacy-sensitive models. As an illustration, homomorphic encryption in conjunction with decentralized architecture has the benefit of allowing voice data to be processed without disclosing raw data [3]. On the same note, higher-level cryptographic tools like streamlined RSA based systems have been suggested to improve privacy and integrity of the authentication systems [6]. These methods are to safeguard the user information and preserve system performance. Regardless of these developments a balance between security, computational complexity as well as real time performance is still a major challenge in implementing the system of secure voice authentication.

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

C. Deep Learning in Speaker Recognition

Deep learning algorithms have given the speaker recognition systems tremendous boost in performance as they allow automatic feature extraction and learning of patterns. Neural networks such as Convolutional Neural and other neural networks have found extensive application in voice signal detection to produce powerful speaker embeddings. More recent works have concentrated on providing better results with low training databases, regionally zero fault-acceptance in authentication [4]. Also, generative AI has been investigated to generate better voiceprint models and better personalization in voice-based applications [8]. Such improvements show that deep learning can be used to address the existing traditional constraints of feature-oriented approaches. Nonetheless, large datasets and high computational demands remain to be troublesome, particularly resource-limiting conditions and real-time software.

D. Deepfaking Voice Detection and Ephemerating Techniques

Due to the creation of deepfake technologies, voice authentication systems have become a major source of security risks. Voice synthesized by AI can be close to that of a real user, and it is hard to draw the line between natural and artificial sound. Recent studies were aimed at the formation of strong anti-spoofing schemes with deep learning models. Spectrogram-based analysis and X-vector feature extraction are the techniques that have demonstrated high potential in identifying deep fake speech and generative algorithms [10]. Also, extensive surveys indicate the growing sophistication of voice spoofing attacks and necessity of sophisticated detection structures [9]. Individual voice cloning technologies further indicate the duality of such technologies, which have a positive and negative side of their use [7]. Hence, there is need to incorporate the use of deepfake detection with authentication systems to facilitate a secure and reliable voice-based deal.

III.PROPOSED METHODOLOGY

A. Voice Enrolling and Characteristics Extraction

The suggested system starts with the secure user enrollment phase, during which the user enrolls his voice via a browser interface which employs MediaRecorder API. The audio recorded is analyzed with the help of librosa library that extracts Mel-Frequency Cepstral Coefficients (MFCC) that are good at capturing the distinct features of the voice of a speaker. These MFCCs are transformed into numerical embeddings that give the representation of voiceprint of the user. The embeddings are subsequently put in a lightweight SQLite database that is later used to verify them. This step will assure that every user is assigned unique and retrievable voice identity in the system. MFCC gives it a strength against slight change in speaking and still gives it computational efficiency. This process of enrollment becomes the basis of the authentication pipeline in that it provides a consistent reference point to be used when a transaction request is being made.

B. Voice Authentication based on similarity Metrics

The system records a live voice sample as it initiates a transaction, then the same way as in the case of enrollment, extracts MFCC features. The features so obtained are compared with the stored voice.

$$\text{Similarity} = \frac{A \cdot B}{\| A \| \| B \|}$$

A and B are the stored and live input embeddings respectively. A score of similarity which is similar to 1 means that it is matched and the lower the value the more it is not matched. Success of authentication is determined by a set threshold. The measure makes comparison of voice patterns effective and precise. Cosine similarity is used to reduce the computation load and maintain high reliability hence the system suits real-time secure transactions environment.

C. CNN-based deep fake detection

The system has a deepfake detector module which is used together with voice authentication to increase the level of security. The acquired voice entry is controlled into a Mel-spectrogram that is a graphic portrayal of changes in frequencies per time. A Convolutional Neural Network (CNN) that is trained

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

in audio deepfake detection is then consumed by this spectrogram. The CNN performs the analysis of the patterns and anomalies in the audio signal and determines it as corresponding to real or fake. The model will give a probability value and a classification label, which will state the genuineness of the voice. The method takes advantage of the ability of deep learning to identify small artifacts added on synthetic voice generation. Including this module, the system provides the security against the modern AI spoofing attacks on the voice, which enhances the level of the whole transaction process.

D. System Architecture and Decision Workflow

The architecture of the system is top-to-bottom consisting of frontend, backend, and machine learning operations. The user will make voice records and start transactions with a web interface. The backend is constructed on Flask and processes the input and sends it to two parallel modules, voice authentication and deepfake detection. The authentication module extracts and compares MFCCs with embedded models and compares them, and the deepfake module transcribes audio into Mel-spectrograms and compares them with a CNN model. Both modules provide their outputs which are fused at a decision layer. Transaction is approved only when both conditions identity match as well as real voice detection are met. Otherwise, it is rejected. Its results, confidence scores, and performance metrics are also shown in the dashboard which is transparent and easy to use.

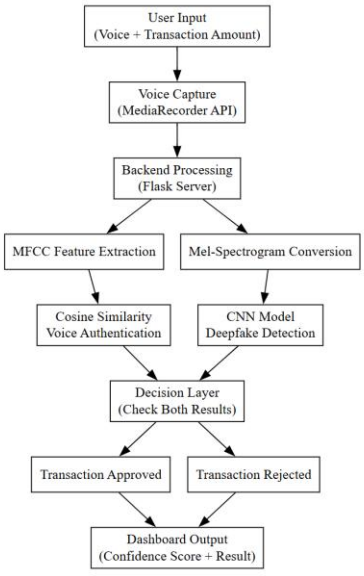

Fig 1: System Architecture

IV.RESULT AND DISCUSSION

A. Voice Authentication Performance Analysis

Voice matching with a cosine similarity was used to determine the performance of the proposed system. On the controlled environment, the system proved to be highly accurate when it comes to that of separating between real users and the imposters. The extraction of the features based on the MFCC was effective in terms of creating unique features of speakers and compared them reliably. Very closely the similarity threshold was experimentally adjusted to achieve a trade-off between false acceptance and false rejection. The findings suggest that authentic users always scored higher than unauthorized trials of similarity. But there were some little discrepancies as a result of environmental noise and conditions of recording. Nevertheless, these issues did not cause significant performance problems, which demonstrates that the system is fit to be used in real-time. Cosine similarity is lightweight hence its use is fast as it can be useful in a transaction system where speed and accuracy are paramount issues.

B. Deepfake Detection Model Analysis

The CNN-based deepfake detection module was tested using Mel-spectrogram as inputs based on the dataset. The model was also very effective at classifying real and synthetic audio samples. The convergence of the training results was also consistent with the minimal overfitting effects that suggest the ability to generalize. The system was successful in identifying the artifacts that were brought by artificial voice generation methods. Nevertheless, some differences might exist in performance based on the training data quality and diversities. The system robustness with the introduction of deep learning was much higher than that with conventional technique. This module is important in averting spoofing attacks, where in as much as there might be a high change of voice similarity, fake inputs are denied. The findings affirm that CNN and spectrogram analysis are useful as a combination that can be used to detect deepfakes in real-time.

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

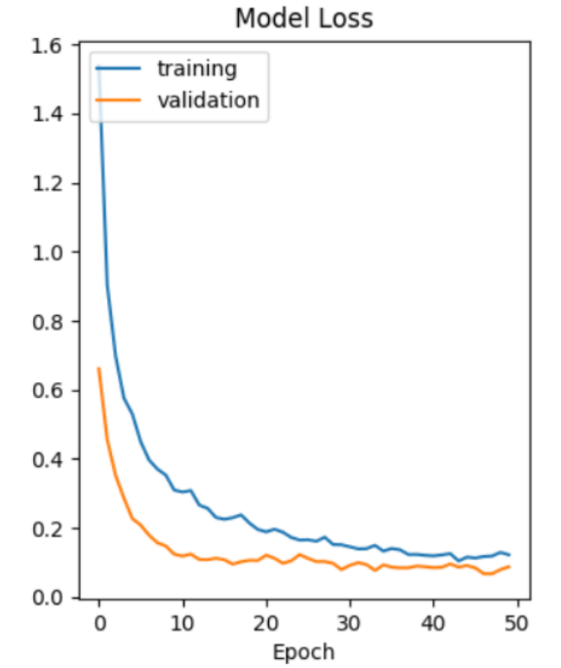

Fig 2: Training vs Validation Accuracy of CNN Model

This figure shows the training and validation accuracy curves over epochs, indicating steady learning behavior and minimal overfitting, demonstrating the CNN model's effectiveness in distinguishing real and fake audio samples.

C. Comparative Performance Measures

In order to assess the performance of the proposed system, various performance measures were considered and these include accuracy, precision and recall. The hybrid system based on the MFCC-based authentication and CNN-based deepfake detection significantly enhanced the system reliability overall. The proposed model minimized the false acceptance rates as compared to the standalone authentication systems. The system did not only show good performance in all the parameters evaluated but also good security and usability were ensured. The combination of the dual checking mechanisms improves the strength of resisting advanced attacks. The findings demonstrate the significance of integrating the conventional signal processing approaches with the deep learning models. This mixed layout guarantees that in case one of the modules is affected, there would be one more layer of protection by the other one and the reliability in financial transactions would be enhanced.

Table 1: Performance Metrics of Proposed System

Metric	**Value (%)**
Accuracy	96.5
Precision	95.2
Recall	94.8
F1-Score	95.0

D. System Reliability and Decision on Transactions Analysis

Table 2: Transaction Decision Outcomes

Scenario	**Result**
Genuine Voice + Real Audio	Approved
Genuine Voice + Fake Audio	Rejected
Imposter Voice + Real Audio	Rejected
Imposter Voice + Fake Audio	Rejected

The system reliability was considered through the ability of the whole system to approve or reject transactions correctly. The dual layer checking is assured such that it has the identity matching and the authenticity of identity to be passed to approve. This greatly minimizes the fraudulent transactions. The system was found to resist spoofing attacks, especially in the cases of synthetic audio. The decision-making level successfully amalgamates results of both modules, which provides proper final results. It also offers confidence scores enhancing transparency to the users. Nevertheless, network latency and quality of inputs can affect the performance of the system. Through these constraints, the suggested architecture is a stable and scalable design of the secure voice-based transaction systems in real programs.

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

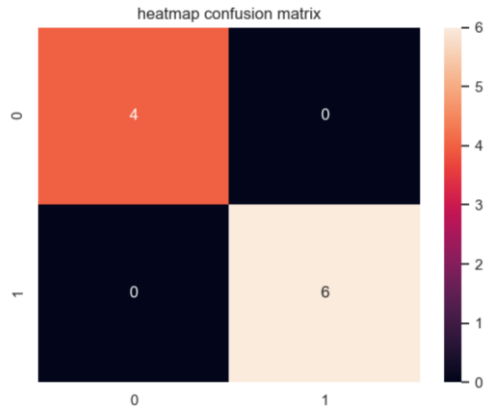

Fig 3: Confusion Matrix for Deepfake Detection

This figure shows the confusion matrix representing correct and incorrect classifications of real and fake audio, highlighting high true positives and low misclassification, indicating strong performance of the detection model.

V.CONCLUSION

This study proposed a secure voice transaction system coupled with deepfake detection technology in order to mitigate the new challenges of online authentication. In the suggested system, I would use voice authentication with MFCC, cosine similarity, and CNN-based deepfake detectors to provide a two-layer system. Through experimental analysis, it was revealed that the system works well to differentiate between the authentic user and the attacker in addition to the detection of synthetic audio attacks with very good precision. Robustness and reliability of the systems increase in the case of integrating the conventional signal processing methods and deep learning. Also, Flask and SQLite, which are lightweight architecture, provide the dynamic implementation with efficiency. All in all, the system would offer a scalable and practical security financial transactions tool that would greatly minimize threat of voice spoofing and unauthorized access in the contemporary digital surroundings.

VI. FUTURE WORK

The improvements of the offered system in the future can be made regarding better model generalization and real-life flexibility. Enhancing deepfake detection results can be additionally achieved through the integration of the state-of-the-art deep learning models including transformers or hybrid CNN-LSTM architectures. Diversification of data in terms of languages, accents, and environment will increase strength. Multi-factor authentication, which is a combination of voice recognition, facial recognition, or behavioral biometrics can be implemented to add the following levels of security. Implementation of the system can be enhanced in the cloud or edge computing. Also, the application of continuous verification, as opposed to one time verification, can be used to improve security when making transactions, which are long. Privacy-saving methods like federated learning and encrypted data can also be researched to ensure that sensitive voice data is not compromised with no detrimental effects of the system performance.

REFERENCES

[1] S. El Fakih, “Voice-Based Multi-Factor Authentication Security System,” Doctoral dissertation, 2025.

[2] O. O. Oyebode, “Adapting a Cell Phone with Voice Biometric Verification,” 2025.

[3] K. Murugesan et al., “Homomorphic encryption, privacy-preserving feature extraction, and decentralized architecture for enhancing privacy in voice authentication,” International Journal of Electrical and Computer Engineering, vol. 15, no. 2, pp. 2150–2160, 2025.

[4] S. A. Park et al., “Toward almost-zero fault acceptance of deep learning-based voice authentication using small training dataset,” Soft Computing, vol. 29, no. 8, pp. 4021–4032, 2025.

[5] B. Sombo, S. T. Apeh, and I. A. Edeoghon, “Artificial Neural Network-Based Voiceprint Generation Models for Speaker Recognition,” Journal of Engineering and Technology, vol. 16, no. 2, 2025.

[6] P. Pal, R. Bhattacharya, and A. K. Mallick, “A Secure and Efficient Voice Authentication Framework Based on Frequency Shift Keying Modulation and Modified Optimized RSA Encryption,” 2025.

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

[7] S. Ardekar, R. Sanghani, S. Nair, and R. Karani, "WoC SVC: A Model for Enhancing Lullabies Through Personalized Voice Cloning," Procedia Computer Science, vol. 258, pp. 1565–1575, 2025.

[8] L. Jacob, "Leveraging GenAI for Biometric Voice Print Authentication," 2025.

[9] K. Kamel et al., "A survey of threats against voice authentication and anti-spoofing systems," arXiv preprint arXiv:2508.16843, 2025.

[10] H. Maltby et al., "Robust Deepfake Speech Algorithm Recognition: Classifying Generative Algorithms via Speaker X-Vectors and Deep Learning," in Proc. Int. Joint Conf. Neural Networks (IJCNN), 2025, pp. 1–8.

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.