E-ISSN: 2664-8784
P-ISSN: 2664-8776
IJRE 2025; 7(1): 59-63

www.engineeringpaper.net
Received: 20-02-2025
Accepted: 24-03-2025

Angelin Rosy M
Assistant Professor, II MCA, Department of Master of Computer Applications, Er. Perumal Manimekalai College of Engineering, Hosur, Tamil Nadu, India

Agalya K
Department of Master of Computer Applications, Er. Perumal Manimekalai College of Engineering, Hosur, Tamil Nadu, India

Correspondence:
Angelin Rosy M
Assistant Professor, II MCA, Department of Master of Computer Applications, Er. Perumal Manimekalai College of Engineering, Hosur, Tamil Nadu, India

A high-performance face anti-spoofing detection method using capsule siamese network

Angelin Rosy M and Agalya K

DOI: https://www.doi.org/10.33545/26648776.2025.v7.i1a.77

Abstract
Facial spoofing attacks—such as those involving printed photos, replayed videos, or three-dimensional masks—pose significant risks to biometric security systems by enabling unauthorized access and identity manipulation. Traditional anti-spoofing methods, including manual feature extraction and deep learning models, often face challenges in generalizing across diverse spoofing scenarios due to differences in lighting, facial angles, and domain variability. While Convolutional Neural Networks (CNNs) are commonly used in spoof detection, they frequently fall short in maintaining spatial dependencies among facial features, resulting in higher rates of incorrect acceptances and rejections. To overcome these shortcomings, this study proposes a Capsule Siamese Network (Capsule-SN) as a resilient approach to facial spoof detection. By leveraging dynamic routing between capsules, the Capsule Network maintains the spatial structure of facial features, offering a richer representation of facial texture, depth cues, and intrinsic patterns. The Siamese framework strengthens the model's capacity to distinguish between authentic and spoofed images by learning similarity metrics between image pairs, thereby reducing sensitivity to specific attack types. Furthermore, the model employs contrastive loss to shape an embedding space that clearly separates genuine from fake inputs.

Keywords: Facial spoof prevention, dynamic routing in neural networks, 3D mask-based identity fraud

Introduction
The widespread deployment of facial recognition technologies across various sectors—such as smartphone unlocking, surveillance systems, and secure access solutions—has intensified the demand for robust and secure face verification methods. A persistent threat to the reliability of these systems is face spoofing, where adversaries attempt to deceive recognition models using forged facial representations like printed images, video replays, or lifelike 3D masks. Such deceptive tactics pose significant challenges to the integrity and safety of biometric systems.

Traditional methods for detecting face spoofing often rely on handcrafted visual features, motion-based analysis, texture descriptors, or auxiliary sensor inputs such as infrared or depth information. Although these methods have achieved some success, they generally fail to perform well against novel attack types and in varied environmental conditions. In contrast, deep learning-based approaches have shown considerable promise by learning powerful feature representations from raw input data. Nonetheless, many of these models demand substantial computational resources and frequently fall short in capturing detailed spatial relationships that are key to identifying sophisticated spoofing techniques.

Proposed work
1. The Capsule-SN model retains detailed facial patterns and structural information by leveraging the representational capabilities of capsule networks.
2. Dynamic routing mechanisms between capsules strengthen the richness and precision of extracted features, allowing the model to better capture spatial dependencies.
3. The Siamese network design facilitates effective pairwise comparison, ensuring high accuracy in differentiating between genuine and spoofed face images.
4. Contrastive loss is employed to reduce both false positives and false negatives, enhancing the model's decision boundaries.
5. The approach demonstrates enhanced adaptability and resilience, making it suitable for deployment in real-world biometric security systems.

Modules

Dataset Collection

To train and evaluate the proposed face anti-spoofing model, we utilized datasets obtained from established public repositories, including well-known benchmarks such as CASIA-FASD and Replay-Attack. These datasets provide a comprehensive range of facial images captured under controlled as well as unconstrained settings, ensuring a realistic evaluation of spoof detection methods.

The collected data comprises both authentic facial images and spoofed samples, which cover a variety of attack types such as printed photographs, replayed videos, and 3D mask presentations. This mix allows the model to learn the subtle differences between real and fake inputs across multiple spoofing scenarios.

To enhance the generalization capability of the model, we ensured a broad range of visual diversity within the dataset. This includes variations in illumination, camera angles, facial expressions, and background environments. Such diversity is crucial for developing a robust system that can perform well in real-world applications.

All samples were systematically annotated with appropriate labels indicating whether the image was genuine or spoofed. These labels were used for supervised training of the model, allowing it to learn discriminative features between different classes effectively.

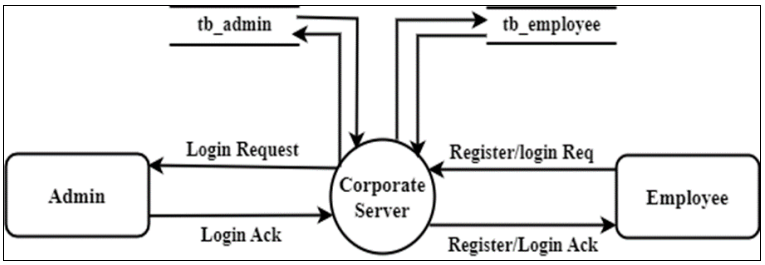
Data Cleaning and Normalization

To enhance the accuracy and efficiency of the model training process, a thorough preprocessing phase was carried out on the dataset. This involved identifying and removing defective data entries, such as images that were distorted, heavily blurred, or redundant. Filtering out these problematic instances helped ensure that the model was trained only on high-quality and informative examples.

Following this, pixel normalization techniques were employed to bring all image data to a common scale. Depending on the model's input requirements, pixel values were adjusted to fall within the range of 0 to 1 or -1 to 1, aiding in faster training convergence and reducing computational errors.

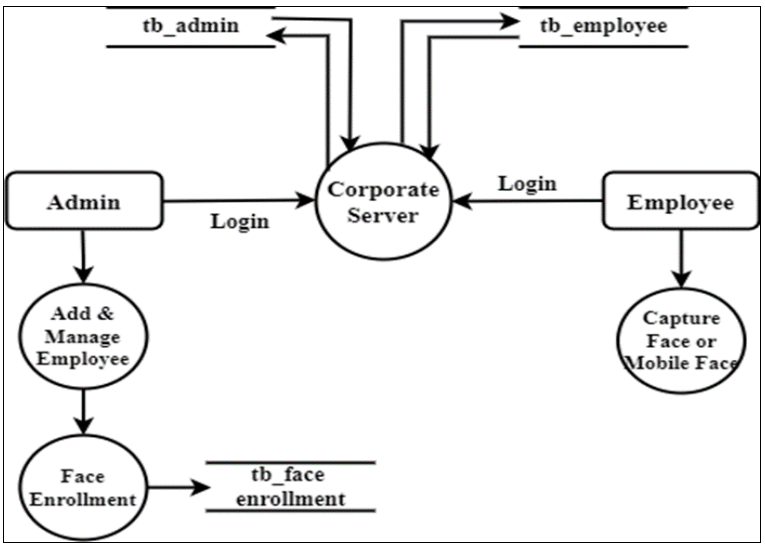
In addition, all image files were resized to a standardized dimension and reformatted uniformly. This consistency in image size and type was critical for ensuring smooth and error-free data loading during the training process.

Overall, these cleaning and normalization steps greatly improved the integrity of the dataset, resulting in more stable learning and enhanced generalization capabilities for the anti-spoofing model.



Feature Extraction Using Capsule Network

To effectively capture intricate facial details and maintain spatial coherence, a Capsule Network was employed for feature extraction. Unlike traditional convolutional networks, capsules are capable of preserving the positional and structural relationships between facial components, allowing for a more nuanced understanding of face geometry. The model learned to identify critical facial elements such as the shape of the eyes, contour of the mouth, and overall head orientation. These features are essential for distinguishing genuine facial characteristics from spoofing attempts. By leveraging the capsule architecture, the system was able to extract both fine-grained (local) and holistic (global) facial representations, enabling it to recognize subtle inconsistencies often present in fake inputs. This approach significantly improved the model's resilience to variations in head pose, lighting, and spoofing artifacts, such as printed images or screen replays.

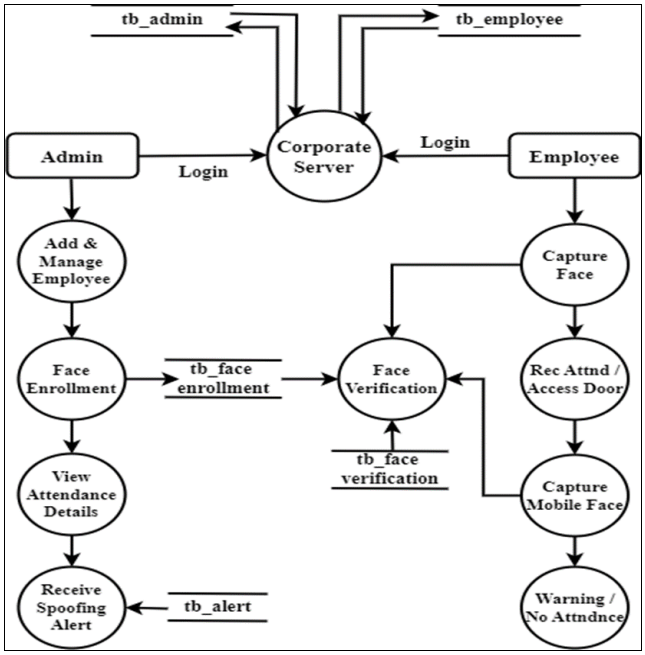


Siamese Network Implementation

A Siamese Network architecture was adopted to effectively measure the similarity between pairs of facial images. This design allows the system to learn how to distinguish between genuine and spoofed face representations by comparing their underlying feature embeddings.

The model was trained to identify matching and non-matching pairs, enabling it to recognize subtle differences between real and fake inputs. To ensure consistent and efficient learning, the twin branches of the network shared identical parameters, allowing both sides to process image pairs using the same learned transformations.

This architecture also supports one-shot and few-shot learning scenarios, making it suitable for detecting spoofing attempts even with limited training examples. As a result, the model demonstrated strong generalization capabilities in recognizing spoofed faces across a wide range of conditions and attack types.

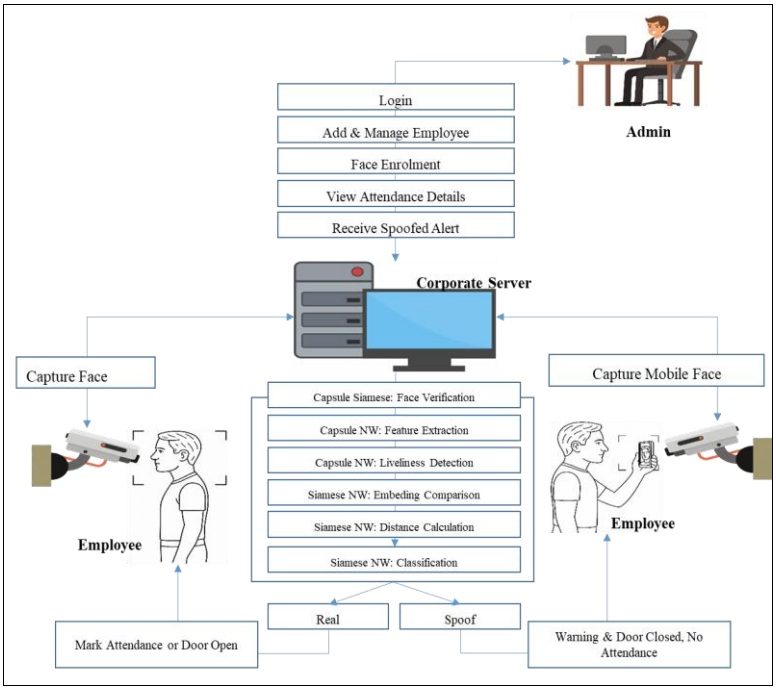


Face Identification

For identity verification, the system compared incoming facial inputs with pre-registered feature embeddings to establish a match. By analyzing the similarity between these embeddings, the model determined whether the presented face corresponded to a known identity.

This process allowed the network to simultaneously perform identity recognition and spoof detection, ensuring that access was granted only to authentic users. The learned embedding space was structured such that genuine facial representations were grouped closely together, while spoofed or fraudulent inputs were positioned at a greater distance, making them easier to distinguish.

This dual-purpose approach enhanced both the accuracy of user authentication and the robustness of spoof prevention, making the system effective for real-world biometric security applications.



Contrastive Loss Optimization

To fine-tune the model's ability to distinguish between genuine and spoofed faces, contrastive loss was employed during training. This loss function encouraged the network to minimize the distance between embeddings of authentic face pairs, reinforcing their similarity in the feature space.

Conversely, it increased the gap between embeddings of mismatched pairs, especially between real and spoofed inputs. By doing so, the model learned to clearly separate legitimate users from fraudulent attempts.

This optimization strategy significantly enhanced the model's capacity to form well-defined clusters for real faces, while ensuring spoofed faces remained distinctly isolated, resulting in more reliable face verification and spoof detection performance.

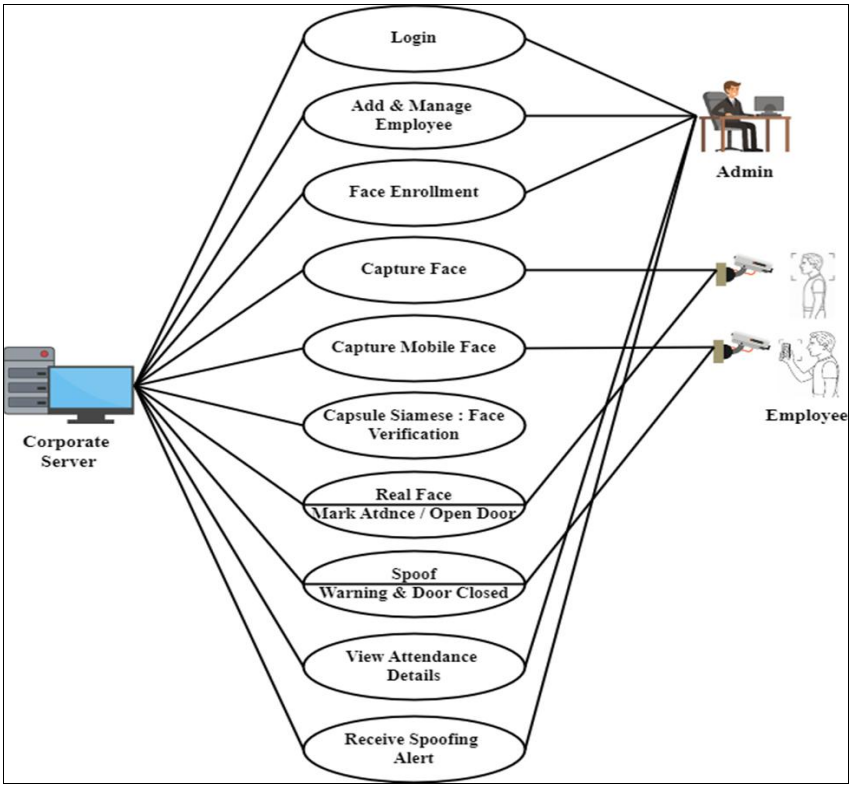
Spoof Detection and Evaluation:

The proposed Capsule-SN model was used to determine whether an input face image was authentic or fake, based on the similarity of extracted feature representations. The model's predictions were utilized to classify each input as either genuine or spoofed.

To assess the effectiveness of the detection system, several key performance indicators were measured, including Accuracy, False Acceptance Rate (FAR), False Rejection Rate (FRR), and F1-score. These metrics provided a comprehensive view of the system's reliability in real-world scenarios.

Performance results were further illustrated using confusion matrices and Receiver Operating Characteristic (ROC) curves, offering insights into classification quality and trade-offs between sensitivity and specificity.

The final system demonstrated fast, real-time detection capabilities, delivering spoof classification results with high confidence and low latency, making it suitable for practical deployment in biometric authentication systems.



Results

The Capsule Siamese Network (Capsule-SN) proposed in this study has shown outstanding effectiveness in identifying various types of facial spoofing attempts, such as photo-based forgeries, replay video intrusions, and 3D mask impersonations. By integrating the spatial retention capabilities of capsule networks with the pairwise comparison strengths of a Siamese framework, the model is able to reliably differentiate between authentic and fraudulent face inputs.

Performance Metrics

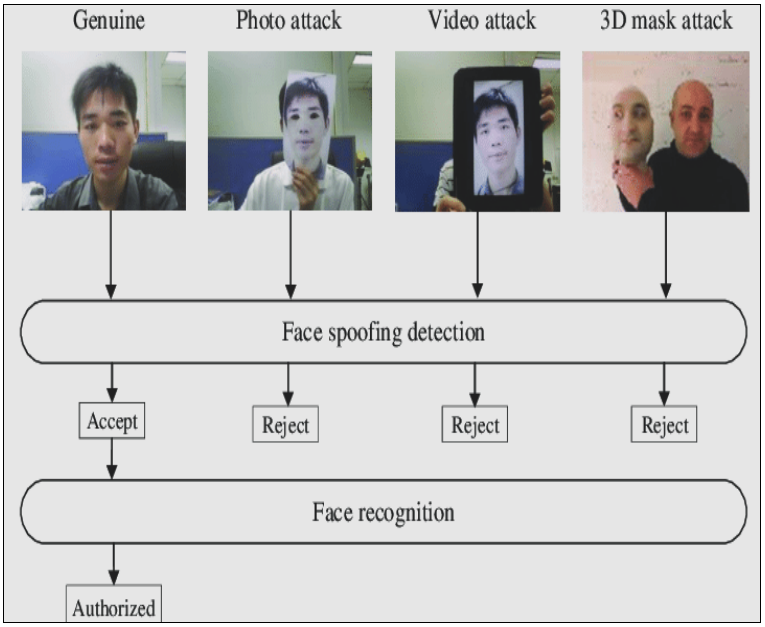
Classification Accuracy: The model achieved high accuracy scores across several standard datasets, showcasing its ability to consistently identify spoofing attempts with precision.

False Acceptance Rate (FAR)

The system maintained a very low rate of false positives, significantly reducing the likelihood of unauthorized access being granted.

False Rejection Rate (FRR)

The FRR remained within an acceptable threshold, ensuring real users are not mistakenly denied access.



Conclusion

This project successfully introduces a high-performance Face Anti-Spoofing Detection System based on the Capsule Siamese Network (Capsule-SN), offering a smart and dependable solution to combat unauthorized access in facial recognition technologies. By integrating the spatial structure retention of Capsule Networks with the pairwise comparison strength of Siamese architectures, the system excels at identifying authentic faces and detecting spoofing attempts—including photo attacks, video replays, and 3D mask impersonations.

The system was developed using technologies such as Python, Flask, MySQL, TensorFlow, and Keras, supported by a suite of contemporary libraries. Key functionalities include real-time spoof detection, an intuitive user interface, and detailed performance tracking. Its architecture effectively reduces both false acceptance and false rejection, improving trust in real-world security applications.

Hosting the solution in a web-accessible environment enhances both accessibility and flexibility, making it ideal for use in applications like secure access control and surveillance monitoring. While the current implementation addresses important challenges such as domain variability and attack generalization, future work may explore continuous learning, edge deployment, and broader coverage of spoof types.

In essence, this system marks a significant step forward in the reliability and precision of face-based authentication technologies, strengthening the foundation for secure and scalable biometric systems.

Acknowledgment

The authors confirm that there are no acknowledgments or external contributions to declare for this study.

References

1. Verkruysse W, Svaasand LO, Nelson JS. Contactless imaging of blood volume changes using natural lighting. Opt Express. 2008;16(26):21434-21445.
2. Li X, Komulainen J, Zhao G, et al. Pulse-based face spoofing detection adaptable to various scenarios via video analysis. In: Proc 23rd Int Conf Pattern Recognit (ICPR). 2016.
3. Hernandez-Ortega J, Fierrez J, Morales A, et al. Examining time-series signals in pulse-driven facial spoof detection across visible and near-infrared ranges. In: CVPR Workshops (CVPRW). 2018.
4. Nowara EM, Sabharwal A, Veeraraghavan A. Photoplethysmogram-based biometric attack detection in PPGSecure. In: IEEE Int Conf Autom Face Gesture Recognit. 2017.
5. Liu S, Yuen PC, Zhang S, et al. Combating 3D facial mask attacks using remote photoplethysmography cues. Cham: Springer International Publishing; 2016.
6. Liu S, Yuen PC, Zhang S, et al. Remote physiological signal detection for 3D mask spoof prevention. Cham: Springer International Publishing; 2016.
7. Liu S, Lan X, Yuen PC. Detection of 3D facial forgeries using rPPG correspondence features. In: Eur Conf Comput Vis (ECCV). 2018.
8. Liu Y, Jourabloo A, Liu X. Face anti-spoofing via deep models using both binary and auxiliary supervisory signals. In: IEEE Conf Comput Vis Pattern Recognit. 2018.
9. Wang W, Stuijk S, D Haan G. A spatial subspace rotation approach for extracting remote photoplethysmography signals. IEEE Trans Biomed Eng. 2015;pp.1-1.