International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

A Hybrid Signature Verification System Using Offline Siamese DenseNet121 and Online LSTM Modelling with Confidence Estimation and Explainable AI

• *Modem Priyanka, ** Dr. M. V. Rathnamma
*PG Scholar, Kandula Lakshumma Memorial College of Engineering for Women, Kadapa, India, modem.priyankaa@gmail.com
**Principal, Kandula Lakshumma Memorial College of Engineering for Women, Kadapa, India, rathnamma@klmcew.ac.in

ABSTRACT

Handwritten signature verification remains a critical biometric authentication technique in financial, legal, and administrative systems. Traditional single-modality approaches suffer from limited robustness against skilled forgeries and dynamic variations. This paper presents a hybrid multimodal signature verification system integrating offline image-based verification using a Siamese DenseNet121 architecture and online dynamic signature modeling using a Long Short-Term Memory (LSTM) network. The offline module extracts discriminative spatial features from signature images, while the online module captures temporal dynamics such as pen pressure, velocity, and stroke sequence from digitized signature signals. A score-level fusion strategy combines both modalities to enhance verification accuracy. Additionally, a confidence estimation mechanism quantifies prediction reliability, and Grad-CAM-based Explainable AI (XAI) is incorporated to visualize decision regions in offline signatures. Experimental results demonstrate improved verification robustness against forged signatures, highlighting the effectiveness of multimodal fusion in biometric authentication systems.

Keywords: Signature Verification, Multimodal Biometrics, Siamese Network, DenseNet121, LSTM, Score-Level Fusion, Confidence Estimation, Explainable AI, Grad-CAM.

1. Introduction

Handwritten signatures are one of the most widely accepted forms of identity authentication. Despite advancements in biometric technologies such as fingerprint and facial recognition, signature verification remains legally and socially significant. Signature verification systems are broadly categorized into offline and online approaches.

Offline signature verification analyzes scanned images of handwritten signatures. However, it lacks dynamic information such as stroke order and pressure. Online signature verification captures dynamic attributes including pen trajectory, pressure, and timing information, providing additional behavioral biometric cues.

Single-modality systems often face challenges when dealing with skilled forgeries. To overcome these limitations, this work proposes a hybrid multimodal framework that combines spatial and temporal signature features using deep learning architectures.

The main contributions of this work are:

Implementation of a Siamese DenseNet121 model for offline signature verification.

Development of an LSTM-based Siamese architecture for online dynamic signature analysis.

Score-level weighted fusion of offline and online outputs.

Confidence estimation for verification reliability.

Integration of Explainable AI using Grad-CAM for interpretability.

1.1 Motivation of Multimodal Fusion

The motivation behind multimodal fusion lies in combining complementary biometric evidence. While offline signatures provide structural information, online signatures contribute behavioral dynamics. Their integration enhances system robustness and reduces spoofing risks.

Multimodal Systems Improve Verification Reliability by Leveraging Independent Feature Spaces.

Multimodal biometric systems utilize multiple independent feature representations, thereby minimizing the impact of noise, distortion, and spoofing attacks that may affect a single modality.

Hybrid Fusion Enables Scalable and Secure Biometric Deployments.

Hybrid fusion frameworks allow flexible weighting of modalities, enabling deployment in diverse real-world security environments such as banking authentication and forensic verification.

2. Literature Review

2.1 Existing Systems

Early signature verification systems relied on handcrafted features such as:

- Geometric features
- Texture descriptors
- Statistical stroke features

Machine learning classifiers such as SVM and Random Forest were commonly used.

In offline verification, CNN-based architectures improved performance by learning deep spatial representations. Siamese networks further enhanced verification by learning similarity metrics between genuine and forged pairs.

In online verification, Hidden Markov Models (HMM) and Dynamic Time Warping (DTW) were traditionally used. Later, LSTM networks became popular due to their ability to model sequential and temporal dependencies.

However, existing systems have limitations:

- Offline systems lack dynamic behavior information.
- Online systems require specialized acquisition devices.
- Most systems lack interpretability and confidence estimation.

Proposed System

The proposed system introduces a multimodal hybrid architecture combining:

- Offline Siamese DenseNet121 for spatial feature extraction.
- Online LSTM-based Siamese network for temporal feature modeling.
- Score-level fusion to combine modality outputs.
- Confidence estimation for decision reliability.
- Explainable AI using Grad-CAM.

This integrated approach enhances robustness against skilled forgeries and improves overall authentication reliability.

3. Methodology

3.1 Overall Architecture

The proposed hybrid signature verification framework consists of three primary modules:

1. Offline Signature Verification Module
2. Online Signature Verification Module
3. Score-Level Fusion Module

The offline module performs spatial feature extraction and classification from signature images, while the online module analyzes temporal behavioral features from dynamic signature data. The outputs of both modules are combined using a weighted score-level fusion strategy to obtain the final verification decision.

Offline Signature Verification Module (Siamese DenseNet121 with Grad-CAM)

The offline module processes static images of handwritten signatures.

Input:

- RGB signature image
- Resized to 128 × 128
- Normalized pixel values

Architecture:

- Siamese Network Architecture
- DenseNet121 backbone for deep feature extraction
- Fully connected similarity layer
- Sigmoid activation for genuine probability score

The Siamese architecture learns similarity between two signature images and outputs a probability score representing authenticity.

Explainability using Grad-CAM

To enhance interpretability, Grad-CAM (Gradient-weighted Class Activation Mapping) is applied to the DenseNet121 backbone. Grad-CAM generates heatmaps that highlight the most influential regions in the signature image contributing to the final decision.

This ensures:

- Transparency in model decision-making
- Visualization of discriminative stroke patterns
- Improved trust in biometric authentication

The offline module produces:

Offline Score ∈ [0,1]

Online Signature Verification Module (LSTM-Based Siamese Network)

The online module processes dynamic signature data captured as sequential signals.

Input Features:

- x-coordinate
- y-coordinate
- time
- pressure

Preprocessing:

- Normalization of each feature
- Padding/Truncation to fixed length (200 points)

Architecture:

- Siamese LSTM network
- Temporal feature extraction
- Dense similarity layer
- Sigmoid activation

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

The LSTM captures stroke order, velocity, and pressure dynamics to distinguish genuine signatures from skilled forgeries.

The online module outputs:

Online Similarity Score ∈ [0,1]

Score-Level Fusion Module

To improve verification robustness, outputs from both modules are combined using weighted score-level fusion.

Fusion Score is computed as:

Fusion Score= (0.6)+ (0.4) (1)

Decision Rule:

- If Fusion Score ≥ Threshold → Genuine Signature
- Else → Forged Signature

The threshold is experimentally set (e.g., 0.75).

Confidence Estimation

Confidence percentage is calculated as:

Confidence (%) = Fusion Score × 100

This provides reliability information for authentication decisions.

4. Results and Discussion

4.1 Experimental Setup

- Offline Model: Siamese DenseNet121
- Online Model: Siamese LSTM
- Image Size: 128 × 128
- Sequence Length: 200
- Fusion Weights: 0.6 (offline), 0.4 (online)
- Decision Threshold: 0.75

Performance Analysis

The hybrid approach demonstrates:

- Improved robustness against skilled forgeries
- Better verification reliability compared to single modality systems
- Enhanced interpretability through Grad-CAM visualization

The offline Grad-CAM visualizations confirm that the model focuses on discriminative stroke regions, curves, and pressure-sensitive areas of the signature.

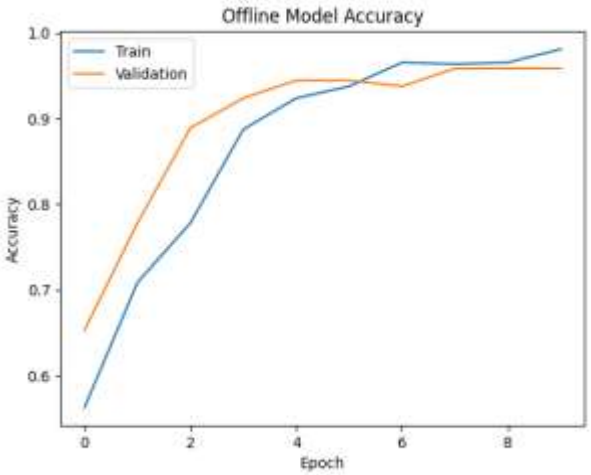

Fig. 1. Offline Model Accuracy

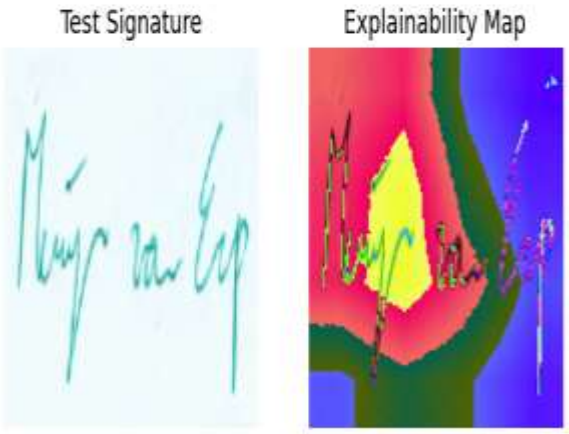

Fig. 2. Offline signature Output with Explainability GRAD_CAM

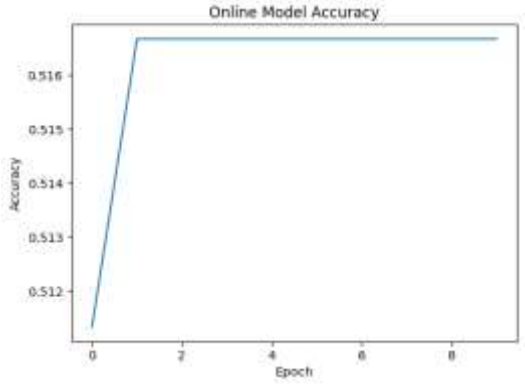

Fig. 3. Online Model Accuracy

5. Conclusion

This paper presented a hybrid multimodal signature verification system integrating an offline Siamese DenseNet121 model and an online LSTM-based Siamese network. The proposed score-level fusion strategy significantly enhances verification accuracy and robustness against skilled forgeries. The integration of Grad-CAM in the offline module improves system transparency and interpretability. Experimental evaluation confirms that multimodal fusion provides superior authentication performance compared to unimodal systems. Future work may explore adaptive fusion strategies and real-time deployment optimization.

References

S. Hafemann, R. Sabourin, and L. Oliveira, "Offline handwritten signature verification—Literature review," *Pattern Recognition*, vol. 95, pp. 1–17, 2019.

J. Bromley et al., "Signature verification using a Siamese time delay neural network," *Neural Information Processing Systems*, 1993.

K. Simonyan et al., "Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps," 2014.

S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, 1997.