Development of a Hybrid Multimodal Biometric System

Zehir Hatem
Supervisors: Dr. Toufik Hafs & Dr. Sara Daas

Laboratory of Study and Research in Instrumentation and Communication of Annaba (LERICA)
Faculty of Technology, Badji Mokhtar-Annaba

July 9, 2025

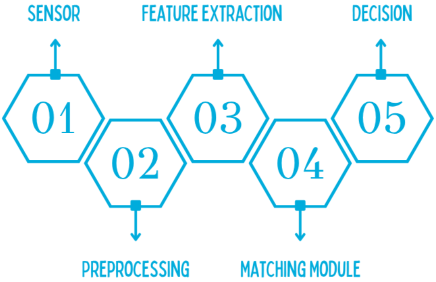
Objectives

- Single-modal biometrics face accuracy and reliability issues.
- Proposed multimodal system seeks to improve accuracy and reduce FAR/FRR.
- Explore and optimize score fusion techniques for system integration.

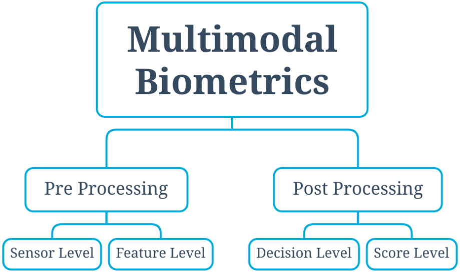
Table of Contents

Table of Contents

Introduction



Sequential representation of a biometric recognition system.

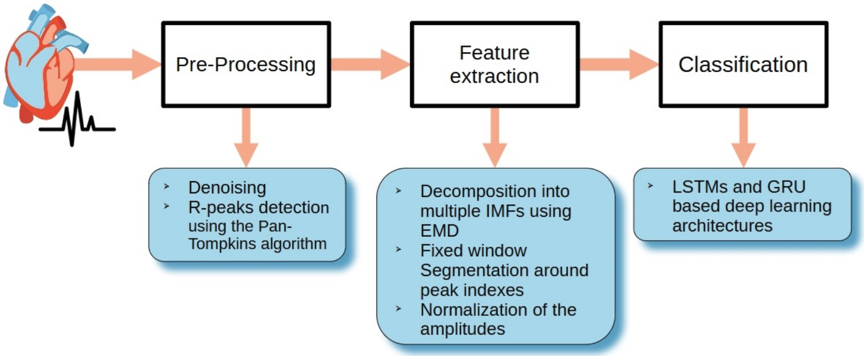


Fusion levels in multimodal biometrics: early vs. late.

Table of Contents

Subsection Overview

ECG Subsystem



MITBIH

The MITBIH database contains ECG recordings from both healthy and diseased patients for arrhythmia detection.

PTB

The PTB Diagnostic ECG Database offers ECG recordings from individuals with cardiac conditions for disease analysis.

NSRDB

The NSRDB provides ECG data from healthy subjects, serving as a reference for normal heart function studies.

ECG Subsystem

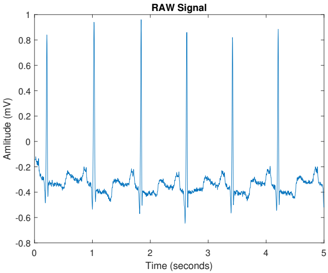
Pre-processing

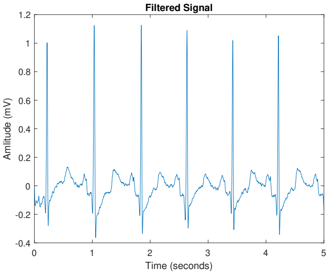
Filtering

- Denoised using a Butterworth bandpass filter (1-40 Hz).

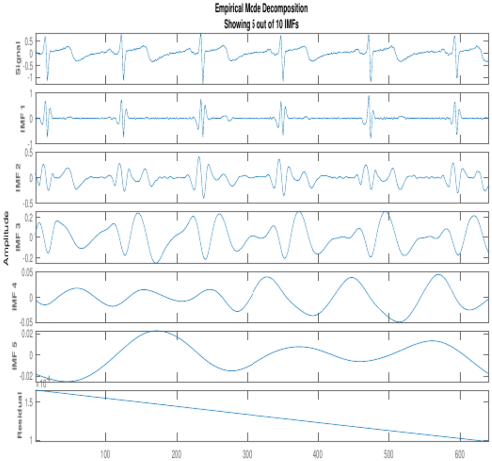
Pan-Tompkins Algorithm

- R-peaks are detected via the Pan-Tompkins algorithm.

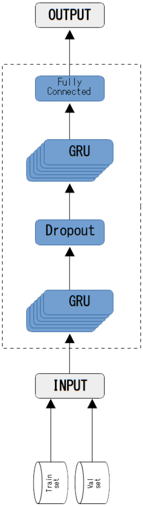




- Decomposing ECG signals using EMD.
- Normalizing and keeping the first 2 IMFs only.
- Segmenting these IMFs into 100 ms windows around R-peaks.



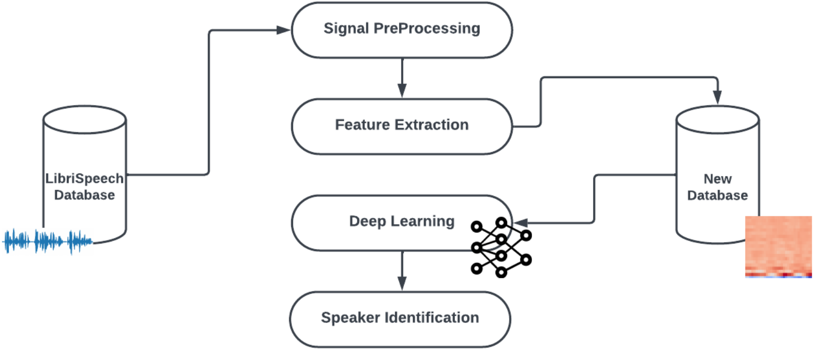
The first 5 IMFs and residual signal resulting from the application of EMD to a single-lead ECG signal from subject 16795 in the NSRDB database.



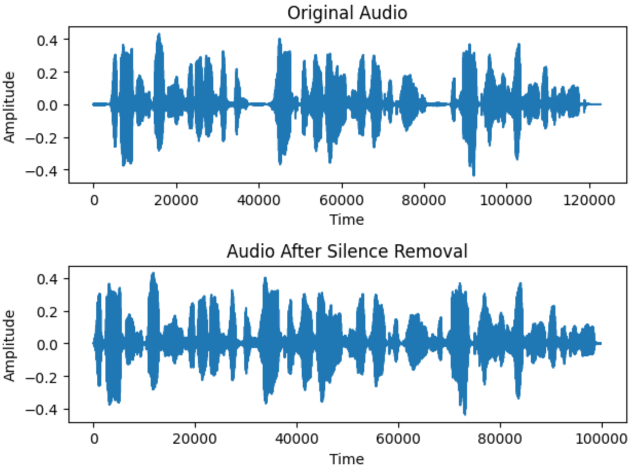
Layer #	Type	Description
1	Sequence Input	2D sequence input
2	LSTM	100 neurons
3	Dropout	Dropout at 20%
4	LSTM	100 neurons
5	Fully Connected	N hidden units
6	Softmax	Softmax activation
7	Classification Output	Cross Entropy

Subsection Overview

Voice Subsystem



Silence Removal & Feature Extraction

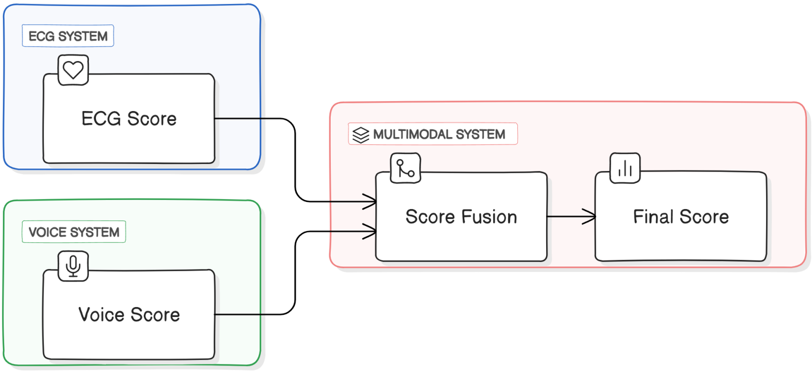


CNN Architecture for Speaker Identification

Layer #	Layer Type	Description
Block 1: Initial Convolution and Pooling		
1	Input	Input shape: $(120, 10, 1)$
2	Conv2D	96 filters, kernel size: 3×3, linear activation
3	Batch Normalization	Batch normalization
4	Leaky ReLU	Leaky ReLU with $\alpha = 0.2$
5	MaxPooling2D	Pool size: 3×1, padding: same
6	Dropout	Dropout rate: 25%
Block 2: Second Convolution and Pooling		
7	Conv2D	64 filters, kernel size: 3×3, linear activation
8	Batch Normalization	Batch normalization
9	Leaky ReLU	Leaky ReLU with $\alpha = 0.2$
10	MaxPooling2D	Pool size: 1×3, padding: same
11	Dropout	Dropout rate: 25%
Block 3: Final Convolution and Pooling		
12	Conv2D	64 filters, kernel size: 3×3, linear activation
13	Batch Normalization	Batch normalization
14	Leaky ReLU	Leaky ReLU with $\alpha = 0.2$
15	MaxPooling2D	Pool size: 1×3, padding: same
16	Dropout	Dropout rate: 50%
Block 4: Classification Layers		
17	Flatten	Flatten output
18	Dense	128 units, linear activation
19	Batch Normalization	Batch normalization
20	Leaky ReLU	Leaky ReLU with $\alpha = 0.2$
21	Dense	Number of output classes (softmax)

Subsection Overview

Multimodal System



Score Fusion

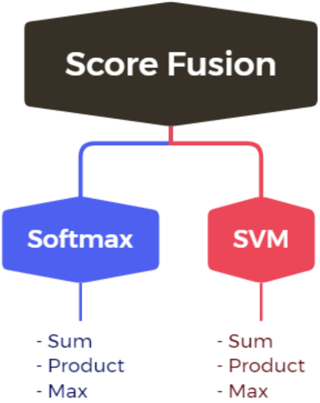
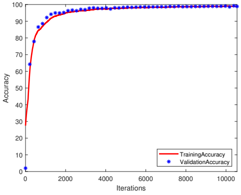


Table of Contents

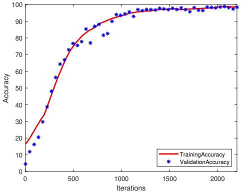
Subsection Overview

Table: The classification results of the proposed GRU model.

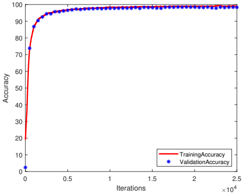
Dataset	Accuracy	Precision	Recall	Specificity	F1-score	FAR	FRR
MIT-BIH	98.57%	98.58%	98.62%	99.97%	98.60%	0.031%	1.42%
NSRDB	99.17%	99.16%	99.14%	99.95%	99.15%	0.048%	0.84%
PTB	98.26%	98.18%	98.18%	99.96%	98.14%	0.037%	1.82%



(a) NSRDB



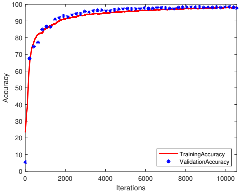
(b) PTB



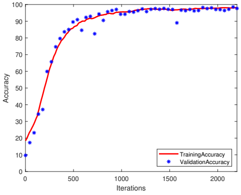
(c) MITBIH

Table: The classification metrics of the proposed LSTM model.

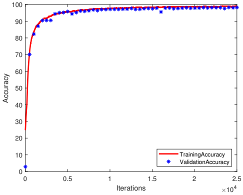
Dataset	Accuracy	Precision	Recall	Specificity	F1-score	FAR	FRR
MIT-BIH	98.33%	98.39%	98.43%	99.96%	98.40%	0.036%	1.61%
NSRDB	98.27%	98.27%	98.23%	99.90%	98.24%	0.101%	1.73%
PTB	97.89%	97.71%	97.83%	99.96%	97.70%	0.045%	2.29%



(a) NSRDB



(b) PTB



(c) MITBIH

ECG Subsystem

Advantages

- Non-invasive and easy to collect biometric data.
- Unique cardiac patterns ensure high accuracy in identification.

Disadvantages

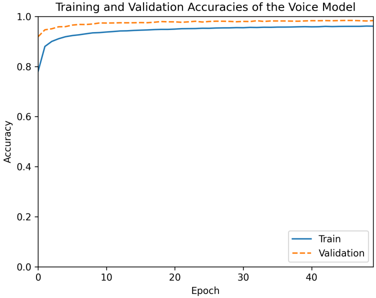
- Susceptibility to noise and artifacts in recordings.
- Variability in ECG signals among individuals and sessions.

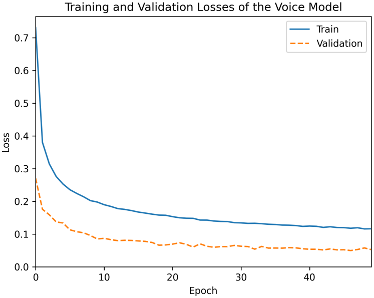
Subsection Overview

Voice Subsystem

Table: The classification results of the proposed speaker recognition system.

Accuracy	Precision	Recall	F1-score	FAR	FRR	EER
98.42%	98.46%	98.45%	98.45%	0.03%	1.55%	0.79%





Voice Subsystem

Advantages

- Convenient and user-friendly for authentication processes.
- Distinctive voice characteristics enhance accuracy in identification.

Disadvantages

- Susceptible to background noise and voice modulation changes.
- Performance affected by health conditions or emotional states.

Subsection Overview

Multimodal System

Table: Results of the proposed Voice-ECG Multimodal System on the MIT-BIH and LibriSpeech Databases.

Method	**Accuracy**	**FAR**	**FRR**	**EER**
Softmax + Sum	**99.61%**	**0.01%**	**0.43%**	**0.22%**
Softmax + Product	98.32%	0.06%	2.59%	1.32%
Softmax + Max	99.55%	**0.01%**	0.45%	0.23%
SVM + Sum	99.33%	**0.01%**	0.67%	0.67%
SVM + Product	97.53%	0.05%	2.47%	1.26%
SVM + Max	99.36%	**0.01%**	0.64%	0.33%

Multimodal System

Table: Comparaison With the Proposed Unimodal Systems.

Method	**Accuracy**	**FAR**	**FRR**	**EER**
ECG	98.57%	0.03%	1.42%	-
Voice	98.42%	0.03%	1.55%	0.79%
Multimodal	99.61%	0.01%	0.43%	0.22%

Table: Comparaison With Multimodal State-of-the-art Methods.

Method	Modalities	Accuracy	FRR	FAR	EER
Rabab A Rasool	Iris and Face	97.53%	0.24%	0.24%	-
Joshi et al.	Face, Fingerprint, Signature, and Iris	-	1.66%	0.00%	0.4%
Tharewal et al.	Ear and Face	99.25%	-	-	-
Ammour et al.	Iris and Face	99.33%	-	-	-
Proposed Method	ECG and Voice	99.61%	0.43%	0.01%	0.22%

Multimodal System

Advantages

- Enhanced accuracy through complementary biometric traits.
- More difficult for impostors to mimic both traits.
- Better performance under noisy and varying environmental conditions.
- Increased robustness to individual modality noise.
- Minimizes risk from stolen biometric templates.

Table of Contents

General Conclusion

- This thesis addresses the challenge of score-level fusion in multimodal biometric systems and its practical implementation.
- Although extensively studied in the literature, score fusion still presents challenges related to efficiency and accuracy.
- Our work introduces several contributions to develop a multimodal biometric system with high accuracy and low error rates, achieving 99.61% accuracy.
- The accuracy of the proposed system has been thoroughly evaluated and validated using the defined evaluation strategy, dataset, and performance metrics.
- The developed system is capable of facilitating real-world ECG-voice biometric applications.

Perspectives

- Hardware Implementation: Extend the research by deploying the biometric recognition system on embedded platforms.
- Custom Hardware Accelerators: Explore designing custom hardware accelerators (e.g., FPGA).
- Real-World Testing: Validate the system across diverse real-world environments.

Peer Reviewed Journal Articles

Accepted Publications

- **H. Zehir**, T. Hafs, and S. Daas, "*Unifying Heartbeats and Vocal Waves: An Approach to Multimodal Biometric Identification At the Score Level*", Arabian Journal for Science and Engineering.
- **H. Zehir**, T. Hafs, and S. Daas, "*Empirical mode decomposition-based biometric identification using GRU and LSTM deep neural networks on ECG signals*", Evolving Systems.
- **H. Zehir**, T. Hafs, and S. Daas, "*Involutional neural networks for ECG spectrogram classification and person identification*", International Journal of Signal and Imaging Systems Engineering.
- **H. Zehir**, T. Hafs, and S. Daas, "*Hardware-Optimized CNN Architecture for ECG Biometric Identification on Embedded Systems*", International Journal of Signal and Imaging Systems Engineering.
- **H. Zehir**, T. Hafs, S. Daas, , and A. Nait-ali, "*Support vector machine for human identification based on non-fiducial features of the ECG*", Journal of Engineering Studies and Research.

Peer Reviewed Conferences

Conference Presentations

- **H. Zehir**, T. Hafs, and S. Daas, "*TinyCNN: An Embedded CNN Model for Speaker Identification Using ESP32*", ICEERES'23: The 1st International Conference on Electrical Engineering & Renewable Energies Systems, Bechar, Algeria, November 2023.
- **H. Zehir**, T. Hafs, and S. Daas, "*ECG-Based Biometric System using TinyML: Implementation and Performance Evaluation on ESP32*", ICAECCT'23: The 1st International Conference on Advances in Electronics, Control and Computer Technologies, Mascara, Algeria, October 2023.
- **H. Zehir**, T. Hafs, and S. Daas, "*Healthcare Decision-Making with an ECG-Based Biometric System*", DASA23: The International Conference on Decision Aid Sciences and Applications, Annaba, Algeria, September 2023.
- **H. Zehir**, T. Hafs, S. Daas, and A. Nait-ali, "*An ECG Biometric System Based on Empirical Mode Decomposition and Hilbert-Huang Transform for Improved Feature Extraction*", BioSMART2023: 5th International Conference on Bio-engineering for Smart Technologies, Paris, France, June 2023.
- **H. Zehir**, T. Hafs, and S. Daas, "*Edge Based Online Signature Identification: A TinyML Approach with ESP32 Microcontroller*", ICTAEE23: Fourth International Conference On Technological Advances in Electrical Engineering, Skikda, Algeria, May 2023.
- **H. Zehir**, T. Hafs, and S. Daas, "*Bidirectional Long Short-term Memory Neural Networks Based Electrocardiogram Biometric System*", ICESTI'22, Annaba, Algeria, Dec. 2022.

Thank You

Questions and Discussion