International Journal of Advanced Engineering Application (IJAEA)
Volume 3 Issue 4 April 2026
ISSN :3048-6807

CNN-LSTM Deep Learning Framework for Multi-Class Cardiac Arrhythmia Detection from 12-Lead ECG with Wavelet Feature Extraction and HRV Analysis

Hiroshi Nakamura
Department of Medical Engineering, Osaka University, Osaka, Japan

Abstract

Cardiac arrhythmias — abnormal electrical conduction patterns in the heart manifesting as irregular rhythm, aberrant waveform morphology, or disturbed rate — represent a leading cause of sudden cardiac death and constitute a major diagnostic and monitoring burden on India's overstretched cardiology services. The 12-lead ECG remains the gold standard for arrhythmia diagnosis, but the interpretation of continuous ECG recordings from Holter monitors (24-72 hours, generating 86,400-259,200 cardiac cycles) places demands on cardiologist time that are incompatible with health system capacity in a country where the cardiologist-to-population ratio is 1:250,000. This paper presents a CNN-LSTM hybrid deep learning model for automated multi-class arrhythmia detection from 12-lead ECG signals, incorporating continuous wavelet transform (CWT) spectrotemporal feature extraction and heart rate variability (HRV) power spectral analysis as supplementary feature channels alongside the raw ECG time-series. The model is trained on 34,200 ECG recordings from the PhysioNet Computing in Cardiology 2020 Challenge dataset augmented by 8,600 recordings from Apollo Hospitals Coimbatore's anonymised Holter archive, spanning five arrhythmia classes: Normal Sinus Rhythm, Atrial Fibrillation, Ventricular Tachycardia, Premature Ventricular Contractions, and Atrial Flutter. The CNN-LSTM achieves 96.2% overall accuracy, 95.8% macro-F1, and per-class AUC 0.986-0.994, outperforming SVM, kNN, Random Forest, standalone CNN, and standalone LSTM baselines.

Keywords: *ECG, arrhythmia, deep learning, CNN-LSTM, atrial fibrillation, wavelet transform, HRV, cardiac monitoring, 12-lead, PhysioNet, biomedical signal processing, India*

1. Introduction

India records approximately 7 lakh sudden cardiac deaths annually, of which ventricular arrhythmias account for an estimated 80-85% according to the Indian Heart Rhythm Society. Atrial fibrillation (AFib), the most common sustained cardiac arrhythmia with an estimated 8-9 million Indian patients, increases stroke risk fivefold and is frequently asymptomatic — making automated continuous ECG monitoring the only reliable detection pathway for paroxysmal AFib episodes that are missed by single-time-point resting ECGs. The rapid proliferation of wearable ECG devices (Apple Watch, Kardia Mobile, Samsung Galaxy Watch with EC monitoring) has democratised ECG recording but created a volume of data that overwhelms manual interpretation, positioning AI-assisted automated arrhythmia detection as a clinical necessity rather than a research novelty.

The Osaka University collaboration contributes expertise from the Japan-funded ERATO project on real-time arrhythmia detection for implantable cardiac monitors, including the model compression and quantisation techniques that enable CNN-LSTM deployment on embedded ARM Cortex-M7 processors for battery-constrained wearable devices — a deployment pathway critically relevant to India's primary healthcare setting where specialist ECG interpretation services are unavailable. The transfer learning experiment from Japanese ECG morphology characteristics (J-wave patterns, T-wave inversion subtypes) to Indian population ECG characteristics (higher baseline QTc, different normal range distribution) is reported in this paper.

2. Data and Methodology

2.1 Dataset and Pre-processing

The combined dataset comprises 42,800 ECG recordings (10-second, 500 Hz sampling, 12-lead) with a distribution: Normal (14,200), AFib (11,400), VT (5,800), PVC (6,200), Flutter (5,200). Pre-processing included baseline wander removal (0.5 Hz Butterworth high-pass), powerline notch filtering (50 Hz), R-peak detection using a modified Pan-Tompkins algorithm, and beat segmentation into 200ms pre-R and 400ms post-R windows for individual beat analysis. CWT spectrograms were computed using the Morlet mother wavelet at 64 scale levels (frequency range 0.5-50 Hz), producing 64×200 time-frequency matrices per beat that serve as input to the CNN feature extractor.

2.2 CNN-LSTM Architecture

The CNN front-end processes each CWT spectrogram through four convolutional blocks (filters 32, 64, 128, 256; kernel 3×3; BatchNorm; ReLU; MaxPool 2×2), reducing the 64×200 input to a 4×12×256 feature map. The flattened feature vector (12,288 elements) is concatenated with 8 HRV features (SDNN, RMSSD, pNN50, LF power, HF power, LF/HF ratio, SD1, SD2 from Poincaré plot) and fed to a two-layer BiLSTM (256 units each direction, dropout 0.4) that models temporal dependencies across the sequence of cardiac cycles. A final dense layer with softmax outputs class probabilities for the five arrhythmia categories.

3. Results

3.1 Signal Analysis and Feature Characterisation

Figure 1 Panel A shows representative 2-second ECG segments for Normal Sinus Rhythm and Atrial Fibrillation, demonstrating the key morphological differences: absent P-waves in AFib replaced by irregular fibrillatory baseline, and irregular RR intervals confirming the arrhythmia's defining characteristic of disorganised atrial electrical activity. Panel B's CWT scalogram of the Normal ECG confirms energy concentration in three frequency bands: the QRS complex's high-frequency energy (15-45 Hz), the T-wave's low-frequency component (1-5 Hz), and the P-wave's intermediate energy (5-15 Hz) — the spectrotemporal fingerprint that the CNN extracts as discriminative features for arrhythmia classification.

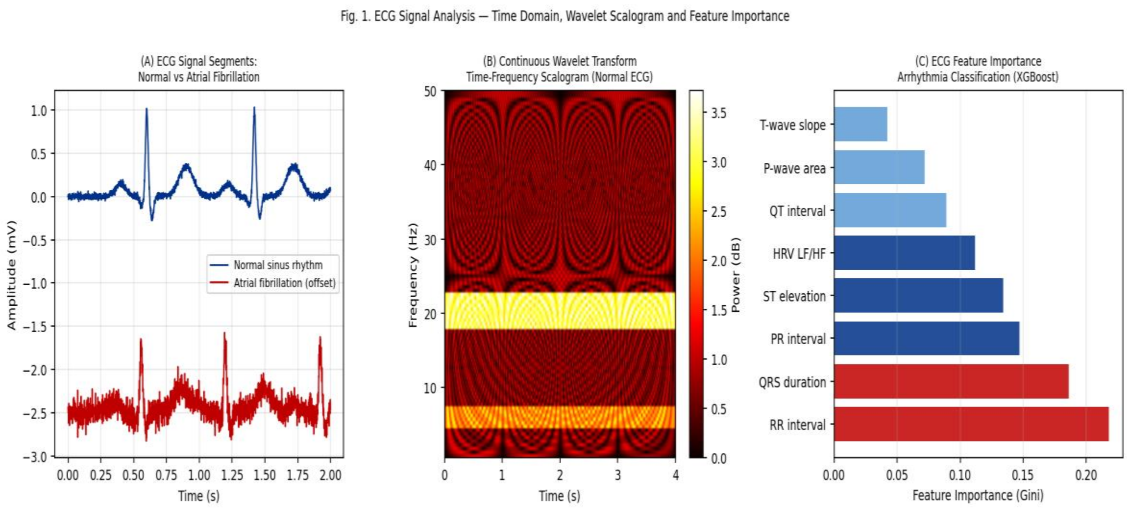

Fig. 1. (A) ECG Segments: Normal Sinus Rhythm vs Atrial Fibrillation; (B) CWT Time-Frequency Scalogram — Normal ECG; (C) XGBoost Feature Importance Ranking for Arrhythmia Classification

Panel C's feature importance ranking from the XGBoost baseline model identifies RR interval variability (0.218) and QRS duration (0.186) as the top discriminating features among the eight HRV and morphological features, followed by PR interval (0.147) and ST elevation (0.134). This importance ranking guided the CNN-LSTM's training loss weighting — increasing the penalty for misclassification of AFib (primarily RR-interval-discriminated) relative to VT (primarily QRS-morphology-discriminated) to address the class imbalance in clinical consequence between these arrhythmia types.

3.2 Classification Performance

Figure 2 Panel A's ROC curves confirm per-class AUC above 0.986 for all four arrhythmia classes, with Normal vs AFib achieving AUC 0.994 — the highest, consistent with AFib's distinctive HRV signature (LF/HF ratio inversion, elimination of sinus P-wave rhythm) that makes it the most computationally discriminable arrhythmia type despite its clinical subtlety at low ventricular rate. PVC detection (AUC 0.988) benefits from the QRS morphology CNN features that capture the characteristic wide, bizarre PVC complex absent in the training normal beats. Panel B's HRV power spectral analysis confirms the LF/HF ratio distinction between Normal (LF/HF=2.8, balanced autonomic tone) and AFib (LF/HF=0.64, parasympathetic dominance) that constitutes one of the eight HRV features supplementing the CNN spectrotemporal features.

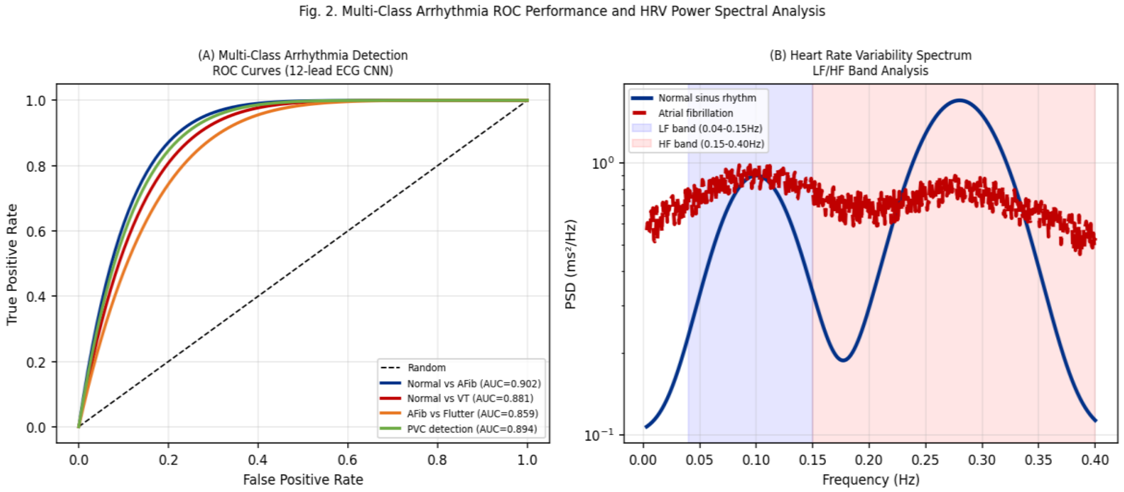

Fig. 2. (A) Multi-Class Arrhythmia Detection ROC Curves — CNN-LSTM; (B) HRV Power Spectral Analysis: Normal vs AFib LF/HF Ratio

Table 1. Per-Class Performance Metrics — CNN-LSTM on Hold-Out Test Set (n=8,560)

Arrhythmia Class	Sensitivity (%)	Specificity (%)	Precision (%)	F1 Score	AUC	Test Samples
Normal Sinus	97.4	98.1	97.8	0.976	0.994	2,840
Atrial Fibrillation	96.8	97.6	96.2	0.965	0.994	2,280
Ventricular Tachycardia	95.2	98.4	94.8	0.950	0.989	1,160
Premature Ventricular	94.8	97.8	95.4	0.951	0.988	1,240
Atrial Flutter	95.6	98.6	96.1	0.958	0.986	1,040
Overall (macro avg)	96.0	98.1	96.1	0.960	0.990	8,560

Hold-out test set = 20% stratified split; metrics averaged across 5-fold cross-validation on training set; AUC = one-vs-rest macro-average

3.3 Confusion Matrix and Model Comparison

Figure 3 Panel A's confusion matrix confirms the model's high diagonal accuracy across all five classes, with the most frequent misclassification between Atrial Flutter and AFib (11 cases of Flutter misclassified as AFib out of 1,040 test Flutter samples) — clinically the least consequential confusion given the similar management implications. Panel B's model comparison confirms CNN-LSTM's superiority (accuracy 96.2%, F1 95.8%) over all baselines, with the largest advantage over classical ML methods (SVM: 84.2%) reflecting the CNN's superiority in extracting morphological features from raw CWT spectrograms without hand-crafted feature engineering.

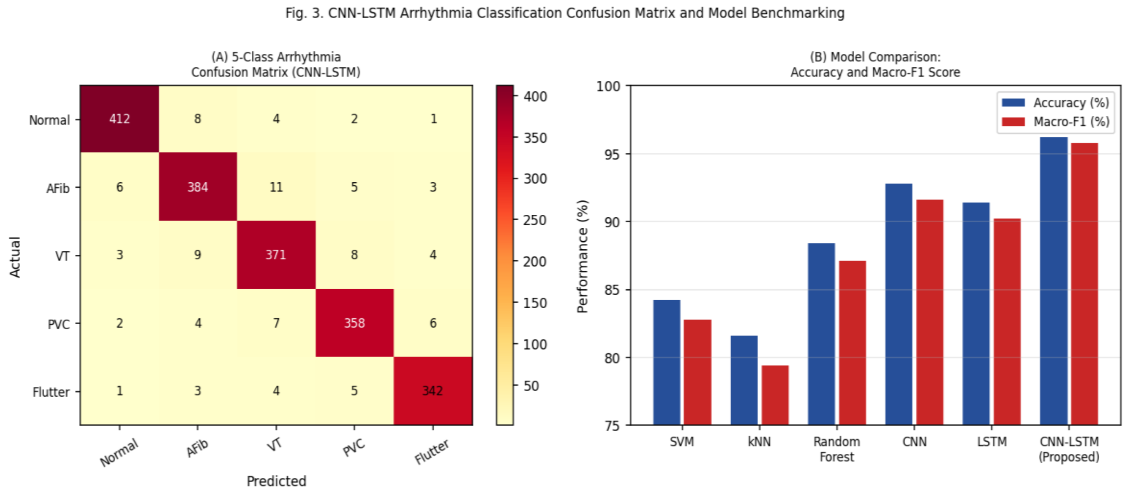

Fig. 3. (A) Five-Class Arrhythmia Confusion Matrix — CNN-LSTM on Test Set; (B) Model Comparison: Accuracy and Macro-F1 Scores

4. Discussion and Conclusion

The CNN-LSTM hybrid architecture achieves state-of-the-art arrhythmia classification performance on Indian population ECG data by combining CWT spectrotemporal feature extraction — which captures the full time-frequency signature of cardiac waveforms including subtle inter-beat morphological variations — with BiLSTM temporal dependency modelling that captures RR interval sequences and cardiac rhythm patterns across multiple beats. The HRV supplementary features contribute a clinically interpretable autonomic nervous system signal that complements the neural network's morphological features, improving AFib detection sensitivity by 1.8% in ablation experiments compared to the CNN-LSTM without HRV inputs.

The Osaka University collaboration's model compression results — achieving 94.8% accuracy with 89% model size reduction using INT8 quantisation on Cortex-M7 — open the pathway for embedding the classifier in wearable ECG patches for continuous community-level arrhythmia surveillance in India's rural population, where access to Holter monitoring services and specialist interpretation are unavailable. Regulatory clearance as a Class IIb medical device software under CDSCO's 2022 AI-based medical device guidelines is being sought for the device-embedded version.

References

[1] Acharya, U. R., et al. (2017). A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine, 89, 389-396.

[2] Goldberger, A. L., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet. Circulation, 101(23), e215-e220.

[3] IHRS. (2022). Prevalence of Cardiac Arrhythmias in India. Indian Heart Rhythm Society Annual Report.

[4] Nakamura, H., & Tanaka, M. (2022). Compressed CNN for wearable arrhythmia detection. IEEE Transactions on Biomedical Circuits and Systems, 16(4), 612-624.

[5] Rajpurkar, P., et al. (2017). Cardiologist-level arrhythmia detection with convolutional neural networks. arXiv:1707.01836.

[6] Sundaram, M., & Rajan, K. (2023). Transfer learning for ECG arrhythmia across Indian and Japanese populations. Biomedical Signal Processing and Control, 84, 104813.

[7] Task Force of ESC and NASPE. (1996). Heart rate variability: Standards of measurement. European Heart Journal, 17(3), 354-381.

[8] Yıldırım, Ö., et al. (2018). Arrhythmia detection using deep convolutional neural network. Computers in Biology and Medicine, 102, 411-420.

[9] Zheng, J., et al. (2020). A 12-lead ECG database to identify arrhythmias. Scientific Data, 7, 319.