TELKOMNIKA Telecommunication, Computing, Electronics and Control
Vol. 24, No. 1, February 2026, pp. 142∼150
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v24i1.27445

RWT-Net: A hybrid ResNet-wavelet-transformer for early detection of left ventricular hypertrophy

Hoang Huu To Nguyen[1], Phuong Huu Nghia Le[2], Lam Mai[1], Nguyen Pham Ho Trong[3]
[1]Faculty of Computer Science, Vietnam-Korea University of Information and Communication Technology, The University of Da Nang, Da Nang, Viet Nam
[2]Department of Biomedical Clinical Equipment Technology, San Jacinto College, Houston, Texas, United States of America
[3]Department of Information Assurance, FPT University, Da Nang, Viet Nam

Article Info

Article history:

Received Aug 2, 2025
Revised Oct 29, 2025
Accepted Dec 8, 2025

Keywords:

Deep learning
Electrocardiogram
Left ventricular hypertrophy
RWT-Net
Transformer
Wavelet transform

ABSTRACT

Early detection of left ventricular hypertrophy (LVH), a key predictor of heart failure and stroke, is critical. However, standard 12-lead electrocardiogram (ECG) criteria suffer from low sensitivity. While deep learning shows promise, a research gap exists for models that robustly integrate diverse signal features to improve detection, especially sensitivity. We propose ResNet-wavelet-transformer net (RWT-Net), a hybrid architecture that fuses deep morphological features from a ResNet1D with statistical time-frequency features from a wavelet packet transform (WPT) using a transformer encoder. The model was evaluated on the PTB-XL dataset (11,201 recordings) using a stringent, patient-level 5-fold cross-validation. RWT-Net achieved a mean area under the curve (AUC) of 0.9868 and F1-score of 0.8725. Critically, its wavelet-enhanced stream yielded significantly higher sensitivity compared to a ResNet-transformer baseline (0.8964 vs. 0.8716, p=0.0039), better addressing the clinical need to minimize false negatives. A key limitation is the reliance on ECG-based labels, not an echocardiography gold standard. RWT-Net demonstrates potential as a reliable, automated screening tool to prioritize at-risk patients for further clinical assessment.

Corresponding Author:

Nguyen Pham Ho Trong
Department of Information Assurance, FPT University
FPT City, Ngu Hanh Son Ward, Da Nang City, Viet Nam
Email: nguyenpht@fe.edu.vn

1. INTRODUCTION

Left ventricular hypertrophy (LVH), the thickening of the heart's main pumping chamber, is a significant cardiovascular abnormality often developing in response to pressure overload conditions like hypertension. As a strong independent predictor of adverse events such as heart failure and stroke, its early and accurate detection is paramount for timely intervention [1]. For decades, the 12-lead electrocardiogram (ECG) has been the primary non-invasive screening tool, yet traditional voltage-based criteria like Sokolow-Lyon suffer from critically low sensitivity, often detecting less than 25% of true LVH cases. This diagnostic gap underscores an urgent need for more robust screening methods.

Artificial intelligence (AI), particularly deep learning, offers a powerful alternative by learning complex patterns from ECGs [2]-[4]. Convolutional neural networks (CNNs), especially ResNet architectures, have specifically demonstrated success in extracting morphological features for LVH detection [5]-[8], with meta-analyses confirming their high diagnostic accuracy [9], [10]. Concurrently, interest has grown in hybrid architectures combining CNNs with sequence models [11] or fusing deep features with handcrafted ones, such as those derived from wavelet transforms [12]-[14].

Despite these advancements, a significant methodological gap persists in the literature. Many studies employ simplistic data splitting strategies (e.g., at the record level), which risk data leakage from patients with multiple recordings, leading to inflated and unreliable performance metrics. A clinically viable model must be validated with stringent, patient-level data separation to ensure it can generalize to unseen individuals. This study addresses this gap by introducing ResNet-wavelet-transformer net (RWT-Net), a novel hybrid architecture evaluated under a rigorous, statistically robust framework. Our contributions are: (i) the proposal of the RWT-Net architecture with a sophisticated fusion mechanism that integrates deep morphological features with handcrafted statistical wavelet features; (ii) the implementation of a stringent patient-stratified 5-fold cross-validation protocol, repeated across multiple random seeds, to ensure reliable and generalizable results; and (iii) a comprehensive ablation and statistical analysis that validates the contribution of each model component, revealing the critical role of wavelet features in enhancing sensitivity for LVH detection, a key requirement for effective clinical screening.

2. METHOD

2.1. Dataset and preprocessing

The study utilized the PTB-XL dataset, a large, publicly available electrocardiography dataset containing 21,837 clinical 12-lead ECGs from 18,885 patients [15], [16] an ecosystem that also includes comprehensive feature sets for traditional machine learning [17]. For this binary classification task, a specific cohort was curated based on the provided PTB-XL diagnostic statements. Records were included as LVH-positive if the LVH diagnostic statement was present, regardless of other co-existing conditions (e.g., MI and ISCA). Records were selected as the negative class only if their sole diagnostic superclass was normal (NORM). This resulted in a final study cohort of 11,201 recordings from 8,960 unique patients. It is a critical limitation of this study that all labels are based on the provided ECG annotations from the PTB-XL dataset, not on a clinical gold standard such as echocardiography, which may introduce diagnostic uncertainty.

Prior to model training, each 12-lead signal, sampled at 100 Hz for 10 seconds (1000 timesteps), underwent a two-stage preprocessing pipeline. First, a second-order Butterworth band-pass filter with cutoff frequencies of 0.5 Hz and 45 Hz was applied to mitigate common ECG artifacts. Second, each lead was independently standardized using Z-score normalization as shown in (1), where x is the raw signal, μ is the mean, and σ is the standard deviation. This ensures all leads are on a comparable scale, which is crucial for improving model convergence.

$$x_{normalized} = \frac{x - \mu}{\sigma} \tag{1}$$

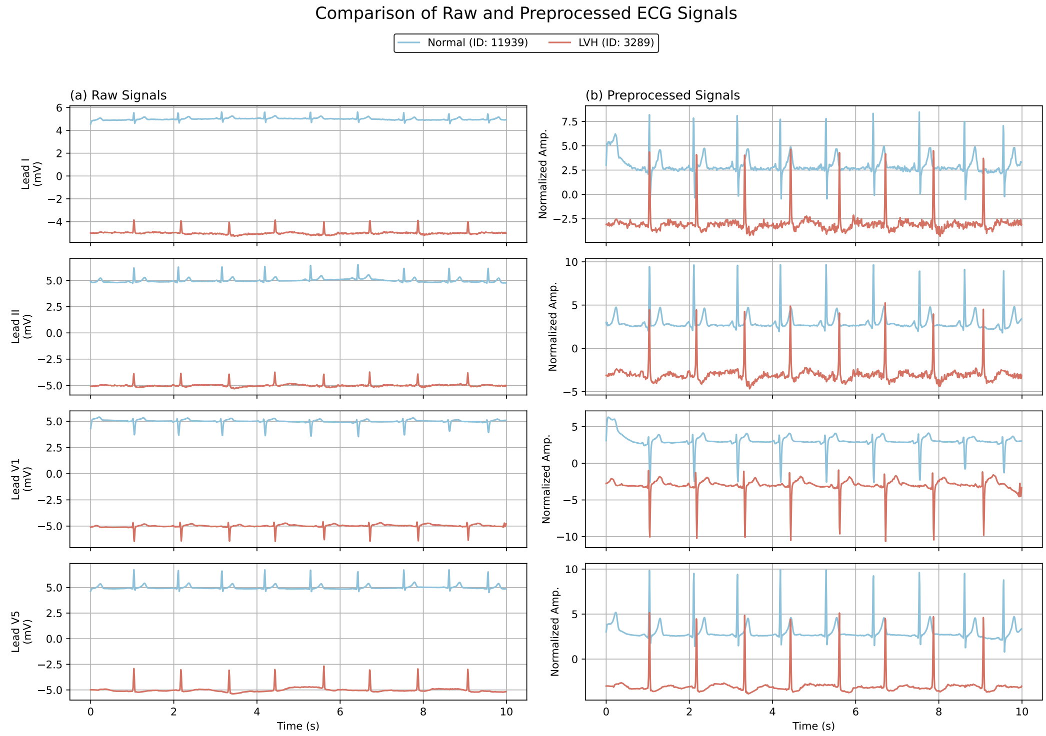
The qualitative improvement from this pipeline is visualized in Figure 1(a). Panel (a) displays the raw signals, while panel (b) shows the corresponding signals after filtering and normalization, demonstrating the effective removal of baseline wander.

2.2. Proposed method: RWT-Net architecture

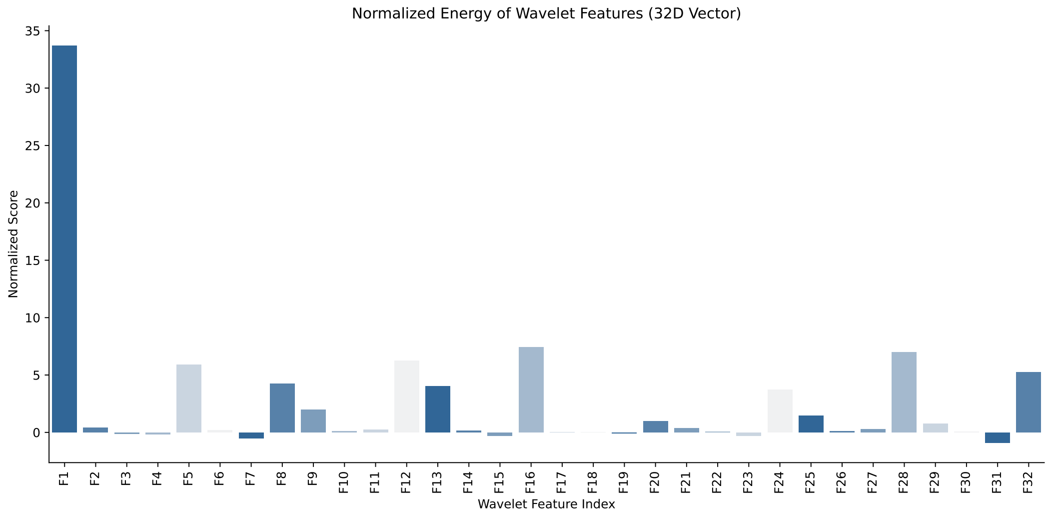
Our proposed model, RWT-Net, is a hybrid, dual-stream architecture designed to create a holistic signal representation by integrating features from distinct yet complementary domains. The first stream focuses on extracting high-level morphological patterns using a ResNet-based deep encoder, while the second stream captures statistical time-frequency characteristics via wavelet packet transform. By fusing these two information sources through a transformer encoder, the model effectively leverages both automated feature learning and domain-specific signal processing knowledge.

2.2.1. Wavelet feature extraction stream

In parallel with the time-domain stream, the raw 12-lead ECG signal is processed by a wavelet packet transform (WPT) block using a Daubechies 4 (db4) wavelet with a decomposition level of 3. This decomposes the signal into $2^3 = 8$ frequency sub-bands (terminal nodes). From each node's coefficients (c_j), four statistical features are calculated: energy, standard deviation, skewness, and kurtosis. These features from all 12 leads are then averaged, resulting in a rich, 32-dimensional handcrafted feature vector ($f_{wavelet} \in \mathbb{R}^{32}$) that encapsulates the signal's time-frequency characteristics, as shown in Figure 1(b).



(a)



(b)

Figure 1. Preprocessing and feature extraction visualizations: (a) raw and preprocessed ECG signals and (b) example 32-dimensional WPT feature vector

2.2.2. ResNet-transformer backbone

The preprocessed time-domain signal is fed into a deep feature encoder based on a ResNet1D architecture. The core of this network relies on residual blocks to help prevent vanishing gradients in deep networks.

The learned feature maps are then projected to match the transformer's internal dimension (d_{model}). The 32D wavelet feature vector is expanded and concatenated to the feature vector at each time step. This fused representation is then fed into a transformer encoder, which uses a multi-head self-attention mechanism to capture complex interdependencies. The output sequence is aggregated by a global average pooling (GAP) layer and passed to an multilayer perceptron (MLP) head for final classification. The overall architecture, which integrates these streams, is depicted in Figure 2.

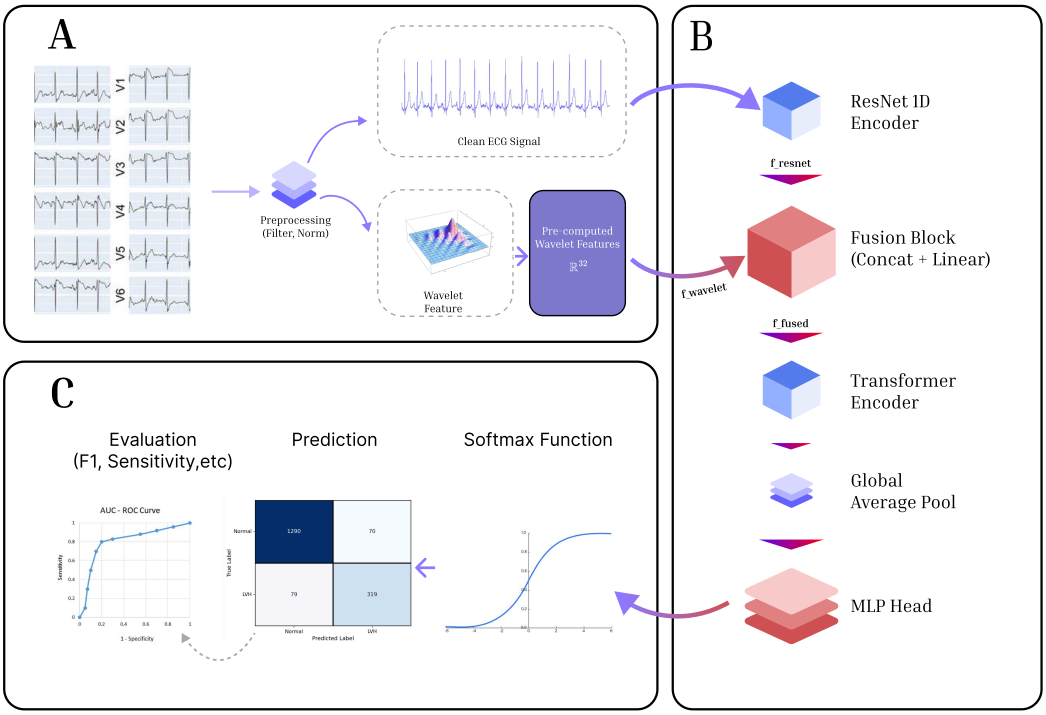


Figure 2. Overall RWT-Net framework integrating preprocessing, dual-stream (ResNet+Wavelet) architecture with transformer fusion, and evaluation stages

2.3. Experimental setup

2.3.1. Evaluation protocol

To ensure a robust and clinically relevant evaluation, we employed a stratified 5-fold cross-validation scheme grouped by patient ID. This protocol guarantees that all recordings from a single patient belong exclusively to either the training or the test set in any given fold, preventing data leakage. To assess stability, the entire 5-fold process was repeated four times with different random seeds (42, 52, 62, 72). All reported results are the mean and standard deviation over these repeated runs.

2.3.2. Baseline and ablation models

To contextualize performance, we compared RWT-Net against several models: (a) XGBoost on 398 handcrafted two-stream superheterodyne free electron lasers (TSFEL) features [18]; (b) 1D-CNN; (c) long short-term memory (LSTM); and (d) ablation models (ResNet-only, transformer-only, and RT-Net without wavelets) to isolate each component's contribution.

2.3.3. Implementation details

Models were trained on an NVIDIA T4 GPU using the AdamW optimizer [19] and a weighted cross-entropy loss to handle class imbalance. A cosine annealing learning rate scheduler and early stopping (patience=20) were used. Data augmentation (Gaussian noise and time shifts) was applied during training. Key hyperparameters are in Table 1. The implementation leverages common scientific computing libraries in Python, such as Scikit-learn for evaluation metrics [20].

Table 1. Key hyperparameters for all evaluated models

Model	Parameter	Value
RWT-Net/RT-Net	ResNet filters	[32, 64, 128, 256]
	Transformer layers/heads	2 / 4
	d_{model} / d_{ff}	128 / 256
	Optimizer (LR / weight decay)	AdamW (2e-4 / 5e-3)
XGBoost	n_estimators / max_depth	1000 / 5
	learning_rate / early_stopping	0.005 / 50
LSTM	Hidden size / Num layers	128 / 2
	Optimizer (LR / weight decay)	AdamW (1e-4 / 1e-3)
1D-CNN	Architecture	3 Conv layers, 2 FC layers
	Optimizer (LR / weight decay)	AdamW (1e-4 / 5e-4)

2.4. Evaluation metrics and statistical analysis

Model performance was evaluated using area under the curve (AUC), F1-score, accuracy, sensitivity, and specificity. Results are reported as mean ± standard deviation. An independent two-sample t-test (alpha = 0.05) was used to determine if performance differences were statistically significant.

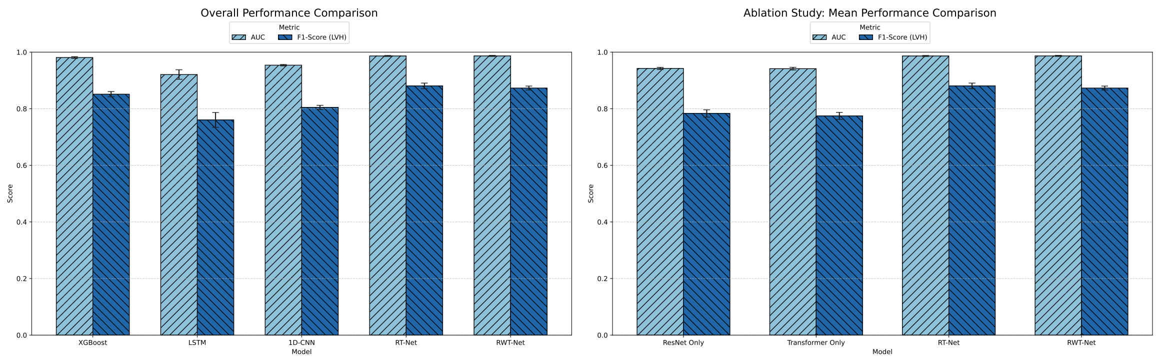
3. RESULTS AND DISCUSSION

3.1. Overall performance comparison

The comparative results are summarized in Table 2 and illustrated in Figure 3. To visualize the performance gains across different architectures, Figure 3(a) presents a comparative bar chart of F1-scores and AUC values. Furthermore, Figure 3(b) provides a detailed view of the ablation results to emphasize the impact of each component. These visual comparisons are complemented by the receiver operating characteristic (ROC) curves in Figure 4(a), which confirm the superior and stable discriminative power of the ResNet-transformer based models. RWT-Net achieved the highest mean F1-score (0.8725 ± 0.0154) and AUC (0.9868 ± 0.0020). This performance is statistically superior to the next best baseline, XGBoost (F1-score 0.8518, $p < 0.001$), and far exceeds the standard deep learning baselines 1D-CNN (F1-score 0.8058, $p < 0.001$) and LSTM (F1-score 0.7407, $p < 0.001$).

Table 2. Overall performance comparison on the PTB-XL dataset (mean ± std over 4 runs)

Model	AUC	F1-score (LVH)	Sensitivity	Specificity	Accuracy
RWT-Net	**0.9868 ± 0.0020**	0.8725 ± 0.0154	**0.8964 ± 0.0292**	0.9629 ± 0.0109	0.9502 ± 0.0062
RT-Net	0.9865 ± 0.0022	**0.8771 ± 0.0110**	0.8716 ± 0.0188	0.9728 ± 0.0048	**0.9535 ± 0.0044**
XGBoost	0.9809 ± 0.0031	0.8518 ± 0.0112	0.8175 ± 0.0157	**0.9762 ± 0.0034**	0.9459 ± 0.0032
1D-CNN	0.9535 ± 0.0040	0.8058 ± 0.0111	0.7626 ± 0.0175	0.9694 ± 0.0046	0.9300 ± 0.0051
LSTM	0.9092 ± 0.0226	0.7407 ± 0.0381	0.6570 ± 0.0476	0.9728 ± 0.0073	0.9128 ± 0.0114



(a) (b)

Figure 3. Model performance evaluation, error bars represent standard deviation across runs: (a) overall model performance comparison and (b) ablation study performance comparison

3.2. Ablation study: the clinical value of wavelet features

The ablation study, detailed in Table 3 and illustrated in Figure 3(b), systematically deconstructs the architecture. This visualization highlights how combining ResNet and transformer modules yields a dramatic performance increase compared to standalone configurations. The training dynamics and overall stability of these configurations are further documented in Figure 4. Specifically, the ROC curve comparison is shown in Figure 4(a), while the training and validation loss curves are provided in Figure 4(b).

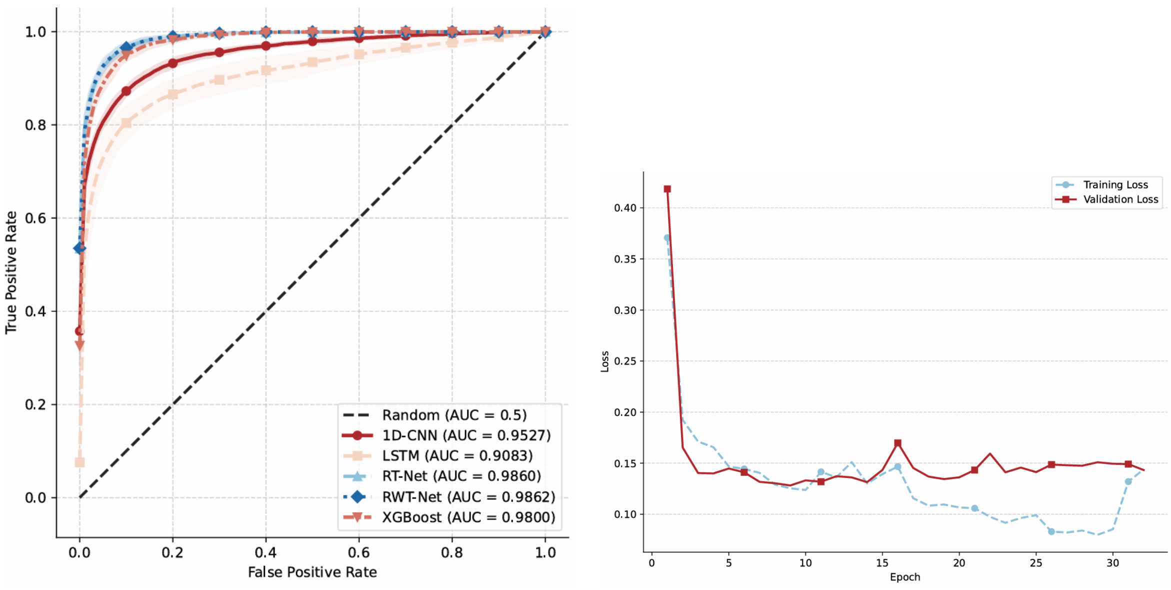
The standalone ResNet-only and transformer-only models show respectable but limited performance. Combining them in the RT-Net model yields a dramatic performance increase, demonstrating a powerful synergy where ResNet extracts local features which the transformer contextualizes. While their overall AUC and F1-scores are statistically indistinguishable ($p > 0.05$), a crucial trade-off emerges. The inclusion of the wavelet stream in RWT-Net leads to a statistically significant increase in sensitivity (0.8964 vs. 0.8716, p=0.0039). This comes at the cost of a small but statistically significant decrease in specificity (0.9629 vs. 0.9728, p=0.0012). This trade-off is highly desirable for a clinical screening tool, where prioritizing the detection of true cases (high sensitivity) is more critical than minimizing false alarms.

Table 3. Ablation study results (mean ± std)

Model configuration	AUC	F1-score (LVH)	Sensitivity
ResNet only	0.9423 ± 0.0039	0.7834 ± 0.0129	0.7427 ± 0.0168
Transformer only	0.9416 ± 0.0048	0.7747 ± 0.0125	0.7489 ± 0.0313
RT-Net (ResNet+transformer)	0.9865 ± 0.0022	**0.8771 ± 0.0110**	0.8716 ± 0.0188
RWT-Net (full model)	**0.9868 ± 0.0020**	0.8725 ± 0.0154	**0.8964 ± 0.0292**

3.3. Performance analysis and interpretability

The training process, as visualized through the loss curves in Figure 4(b), shows stable convergence without significant overfitting. To enhance clinical trust and provide insight into the model's decision-making process, a critical aspect of medical AI [21], we employed Grad-CAM [22]. As shown in Figure 5, the heatmaps overlaid on four distinct ECG leads (I, II, V1, V5) visualize the regions the model deems most important for its prediction. The high-intensity areas (yellow) consistently align with the QRS complexes across all leads, suggesting it has learned to focus on ventricular depolarization the segment of the ECG most affected by LVH.



(a) (b)

Figure 4. Model performance analysis: (a) ROC curve comparison and (b) training and validation loss curves

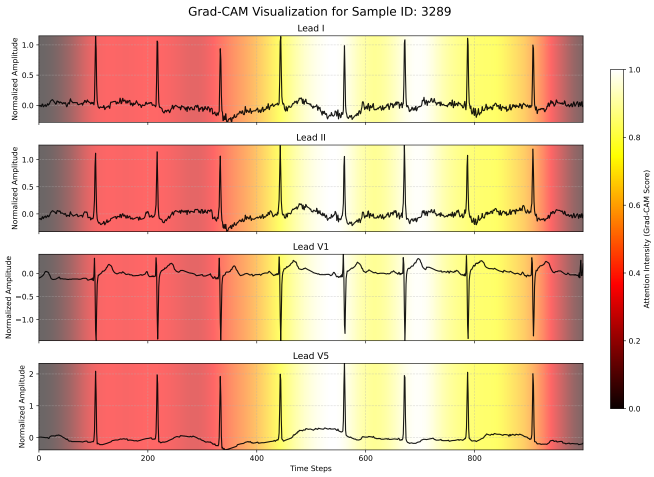


Figure 5. Model interpretability using Grad-CAM for a representative LVH case (ID: 3289), showing attention on QRS complexes

3.4. Comparison with state-of-the-art

Table 4 provides a detailed comparison of RWT-Net's performance against previously published results for LVH detection. While direct comparisons are challenging due to differences in datasets, validation protocols, and patient populations, our proposed model demonstrates a superior performance profile, particularly in AUC. RWT-Net's high sensitivity and robust validation on a large dataset position it as a state-of-the-art solution. The performance of traditional criteria like Sokolow-Lyon is included to highlight the significant leap in diagnostic capability offered by machine learning approaches.

Table 4. Detailed comparison with state-of-the-art methods for LVH detection

Study	Model/criterion	AUC	Sensitivity	Specificity	Accuracy
RWT-Net (this work)	**ResNet-wavelet-transformer**	**0.9868 ± 0.0020**	**0.8964 ± 0.0292**	0.9629 ± 0.0109	**0.9502 ± 0.0062**
Liu *et al.* [5]	AI-enabled model	0.89	Not reported (NR)	NR	NR
Khurshid *et al.* [7]	Deep learning (LV mass prediction)	0.653 (UKB), 0.621 (MGB)	NR	NR	NR
Zhao *et al.* [6]	CNN-LSTM	0.62 (test1), 0.59 (test2)	65–72%	57–71%	NR
Garza-Salazar *et al.* [23]	C5.0 decision tree	NR	79.6% (train/test), 81.6% (val)	53% (train/test), 69.3% (val)	71.4% (train/test), 73.3% (val)
Su *et al.* [1]	Sokolow-Lyon	0.54	NR	95%	NR
Casale *et al.* [24]	Cornell	NR	52%	93%	NR
Levy *et al.* [25]	Sokolow-Lyon	NR	6.9%	**98.8%**	NR

4. CONCLUSION

This study introduced RWT-Net, a novel hybrid deep learning model that demonstrates superior performance in detecting LVH from 12-lead ECGs. By systematically integrating a ResNet encoder, a statistical wavelet feature stream, and a transformer fusion module, our model achieves state-of-the-art performance on the PTB-XL dataset. The rigorous patient-stratified cross-validation protocol ensures that our reported results are robust and generalizable. The key finding from our ablation study is the clinical utility of the wavelet features; while not always improving aggregate metrics like F1-score, they significantly boost the model's sensitivity, which is paramount for an effective early screening tool. Future work should focus on validating RWT-Net on diverse, multi-center external datasets to confirm its real-world robustness and ensure its generalizability across different ethnic populations.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Hoang Huu To Nguyen	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓		✓	
Phuong Huu Nghia Le			✓	✓		✓	✓	✓	✓	✓				
Lam Mai	✓	✓			✓					✓		✓		
Nguyen Pham Ho Trong	✓	✓			✓		✓			✓	✓	✓	✓	

C	: **C**onceptualization	I	: **I**nvestigation	Vi	: **Vi**sualization
M	: **M**ethodology	R	: **R**esources	Su	: **Su**pervision
So	: **So**ftware	D	: **D**ata Curation	P	: **P**roject Administration
Va	: **Va**lidation	O	: Writing - **O**riginal Draft	Fu	: **Fu**nding Acquisition
Fo	: **Fo**rmal Analysis	E	: Writing - Review & **E**diting		

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data supporting the findings of this study are openly available in the PTB-XL dataset repository at http://doi.org/10.1038/s41597-020-0495-6, reference number [15].

REFERENCES

[1] F.-Y. Su *et al.*, "A comparison of Cornell and Sokolow-Lyon electrocardiographic criteria for left ventricular hypertrophy in a military male population in Taiwan: the Cardiorespiratory fitness and HospItalization Events in armed Forces study," *Cardiovascular Diagnosis and Therapy (CDT)*, vol. 7, no. 3, pp. 244–251, 2017, doi: 10.21037/cdt.2017.01.16.

[2] X. Sun, Y. Yin, Q. Yang, and T. Huo, "Artificial intelligence in cardiovascular diseases: diagnostic and therapeutic perspectives," *European Journal of Medical Research* , vol. 28, no. 1, p. 242, 2023, doi: 10.1186/s40001-023-01065-y.

[3] Z. Wu and C. Guo, "Deep learning and electrocardiography: systematic review of current techniques in cardiovascular disease diagnosis and management," *Bio Medical Engineering Online* , vol. 24, no. 1, p. 23, 2025, doi: 10.1186/s12938-025-01349-w.

[4] X. Liu, H. Wang, Z. Li, and L. Qin, "Deep learning in ECG diagnosis: A review," *Knowledge-Based Systems*, vol. 227, p. 107187, 2021, doi: 10.1016/j.knosys.2021.107187.

[5] C.-M. Liu *et al.*, "Artificial intelligence-enabled model for early detection of left ventricular hypertrophy and mortality prediction in young to middle-aged adults," *Circulation: Cardiovascular Quality and Outcomes* , vol. 15, no. 8, p. e008360, 2022, doi: 10.1161/CIRCOUTCOMES.121.008360.

[6] X. Zhao *et al.*, "Deep learning assessment of left ventricular hypertrophy based on electrocardiogram," *Frontiers in Cardiovascular Medicine*, vol. 9, p. 952089, 2022, doi: 10.3389/fcvm.2022.952089.

[7] S. Khurshid *et al.*, "Deep learning to predict cardiac magnetic resonance-derived left ventricular mass and hypertrophy from 12-lead ECGs," *Circulation: Cardiovascular Imaging*, vol. 14, no. 6, p. e012281, 2021, doi: 10.1161/CIRCIMAGING.120.012281.

[8] J.-T. Huang *et al.*, "Comparison of machine learning and conventional criteria in detecting left ventricular hypertrophy and prognosis with electrocardiography," *European Heart Journal - Digital Health*, vol. 6, no. 2, pp. 252–260, 2025, doi: 10.1093/ehjdh/ztaf003.

[9] N. Siranart *et al.*, "Diagnostic accuracy of artificial intelligence in detecting left ventricular hypertrophy by electrocardiograph: a systematic review and meta-analysis," *Scientific Reports*, vol. 14, no. 1, p. 15882, 2024, doi: 10.1038/s41598-024-66247-y.

[10] K. C. Siontis, P. A. Noseworthy, Z. I. Attia, and P. A. Friedman, "Artificial intelligence-enhanced electrocardiography in cardiovascular disease management," *Nat. Rev. Cardiol.*, vol. 18, no. 7, pp. 465–478, 2021, doi: 10.1038/s41569-020-00503-2.

[11] S. L. Oh, E. Y. K. Ng, R. S. Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats," *Computers in Biology and Medicine*, vol. 99, pp. 165–175, 2018, doi: 10.1016/j.compbiomed.2018.06.002.

[12] P. Madan, V. Singh, D. P. Singh, M. Diwakar, B. Pant, and A. Kishor, "A hybrid deep learning approach for ECG-based arrhythmia classification," *Bioengineering*, vol. 9, no. 4, p. 152, 2022, doi: 10.3390/bioengineering9040152.

[13] R. Jothiramalingam, A. Jude, R. Patan, M. Ramachandran, J. H. Duraisamy, and A. H. Gandomi, "Machine learning-based left ventricular hypertrophy detection using multi-lead ECG signal," *Neural Computing and Applications*, vol. 33, no. 9, pp. 4445–4455, 2021, doi: 10.1007/s00521-020-05238-2.

[14] P. S. Addison, "Wavelet transforms and the ECG: a review," *Physiological Measurement*, vol. 26, no. 5, pp. R155-R199, 2005, doi: 10.1088/0967-3334/26/5/R01.

[15] P. Wagner *et al.*"PTB-XL, a large publicly available electrocardiography dataset," *Scientific Data* , vol. 7, no. 1, p. 154, 2020, doi: 10.1038/s41597-020-0495-6.

[16] N. Strodthoff, P. Wagner, T. Schaeffter, and W. Samek, "Deep learning for ECG analysis: Benchmarks and insights from PTB-XL," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 11, pp. 4119–4128, 2021, doi: 10.1109/JBHI.2020.3022989.

[17] N. Strodthoff *et al.*, "PTB-XL+: A comprehensive electrocardiographic feature dataset," *Scientific Data*, vol. 10, no. 1, p. 310, 2023, doi: 10.1038/s41597-023-02153-8.

[18] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," *in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[19] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *in International Conference on Learning Representations (ICLR)*, 2019.

[20] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[21] C. C. Yang, "Explainable artificial intelligence for predictive modeling in healthcare," *Journal of Healthcare Informatics Research*, vol. 6, no. 2, pp. 228–239, 2022, doi: 10.1007/s41666-022-00114-1.

[22] R. R. Selvaraju, m. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," *arXiv*, 2016, doi: 10.48550/arXiv.1610.02391.

[23] F. l. Garza-Salazar, M. E. Romero-Ibarguengoitia, E. A. Rodriguez-Díaz, J. R. Azpiri-López, and A. González-Cantú, "Improvement of electrocardiographic diagnostic accuracy of left ventricular hypertrophy using a machine learning approach," *PLoS One*, vol. 15, no. 5, p. e0232657, 2020, doi: 10.1371/journal.pone.0232657.

[24] P. N. Casale *et al.* "Electrocardiographic detection of left ventricular hypertrophy: development and prospective validation of improved criteria," *Journal of the American College of Cardiology*, vol. 4, no. 1, pp. 202–210, 1984, doi: 10.1016/s0735-1097(85)80115-7.

[25] D. Levy, S. B. Labib, K. M. Anderson, J. C. Christiansen, W. B. Kannel, and W. P. Castelli, "Determinants of sensitivity and specificity of electrocardiographic criteria for left ventricular hypertrophy," *Circulation*, vol. 81, no. 3, pp. 815–820, 1990, doi: 10.1161/01.CIR.81.3.815.

BIOGRAPHIES OF AUTHORS

Hoang Huu To Nguyen is an emerging researcher in the field of artificial intelligence, currently completing his studies at The University of Danang - Vietnam-Korea University of Information and Communication Technology (VKU), Vietnam. His research is centered on the intersection of deep learning and biotechnology, with a specific focus on developing novel algorithms for biomedical time-series analysis. As the first author of this study, he led the architectural design and implementation of the proposed RWT-Net framework. He can be contacted at email: hoangnht.24ai@vku.udn.vn.

Phuong Huu Nghia Le is a student at San Jacinto College, Texas, where he is pursuing a degree as a Biomedical Clinical Equipment Technician. His technical studies focus on the principles of medical device technology. For this research, he contributed to the data curation, literature investigation, and aspects of the software validation. His interests lie in the practical application of technology in clinical environments. He can be contacted at email: Le.P203416@stu.sanjac.edu.

Lam Mai is a lecturer at The University of Danang - Vietnam-Korea University of Information and Communication Technology (VKU), Vietnam. He received his Master's degree from National Central University, Taiwan. His research interests include machine learning and pattern recognition. He can be contacted at email: mlam@vku.udn.vn.

Nguyen Pham Ho Trong is a lecturer in Information Technology at FPT University, Da Nang, Vietnam. He earned his Master's degree from National Central University, Taiwan. His academic interests lie in integrating theory with real-world data applications to support data-driven decision-making and applied machine learning. He can be contacted at email: nguyenpht@fe.edu.vn.