Archives available at journals.mriindia.com

**ITSI Transactions on Electrical and Electronics Engineering**

ISSN: 2320-8945

Volume 13 Issue 02, 2024

# A Comprehensive Review of Heart Disease Prediction Using Optical Electrocardiograms (ECG) and a Hybrid Convolutional Block Attention Capsule Network

Navid Mulyadi
*Department of Computer Science and Engineering, Coral Reef School of Systems Management, Mauritius*
*navid.mulyadi@crssm-mu.net*

**Peer Review Information**

*Submission: 27 Aug 2024*
*Revision: 11 Sept 2024*
*Acceptance: 28 Sept 2024*

**Keywords**

*Heart Disease Prediction, Optical ECG, Deep Learning, Convolutional Neural Networks, Capsule Networks, Attention Mechanisms, Hybrid Models, Biomedical Signal Processing*

**Abstract**

Heart disease remains one of the leading causes of mortality worldwide, necessitating accurate and early diagnostic systems. Optical electrocardiogram (ECG) techniques combined with artificial intelligence (AI) have emerged as powerful tools for non-invasive cardiac monitoring and disease prediction. This paper presents a comprehensive review of heart disease prediction methods using optical ECG signals integrated with hybrid deep learning architectures, particularly Convolutional Neural Networks (CNN), Block Attention mechanisms, and Capsule Networks (CapsNet). Optical ECG provides high-resolution physiological data, enabling effective feature extraction from cardiac signals. CNN models are widely used for automatic feature learning, while attention mechanisms enhance the focus on critical signal segments. Capsule Networks address limitations of CNN by preserving spatial hierarchies and improving interpretability. Recent studies demonstrate that hybrid CNN-attention-capsule models significantly improve classification accuracy, sensitivity, and robustness compared to traditional machine learning methods. Additionally, optimization techniques and hybrid architectures further enhance performance. This review analyzes research, highlighting key advancements, comparative performance, and challenges. Despite promising results, issues such as computational complexity, data variability, and real-time deployment remain. Future research directions focus on lightweight models, explainable AI, and integration with wearable healthcare systems for continuous monitoring.

## Introduction

Cardiovascular diseases (CVDs) are among the leading causes of death globally, accounting for millions of deaths annually. Early detection and accurate diagnosis are essential to reduce mortality rates and improve patient outcomes. Electrocardiography (ECG) is one of the most widely used diagnostic tools for monitoring heart activity. With the advancement of optical sensing technologies, optical ECG has emerged as a promising alternative to traditional electrical ECG systems, offering enhanced signal quality, reduced noise, and improved patient comfort.

Optical ECG utilizes photoplethysmography (PPG) and other optical sensing techniques to capture cardiac signals non-invasively. These signals provide valuable information about heart rate, rhythm, and blood flow, making them suitable for heart disease prediction. However, optical ECG signals are often affected by noise, motion artifacts, and environmental factors, which necessitate robust signal processing and analysis techniques.

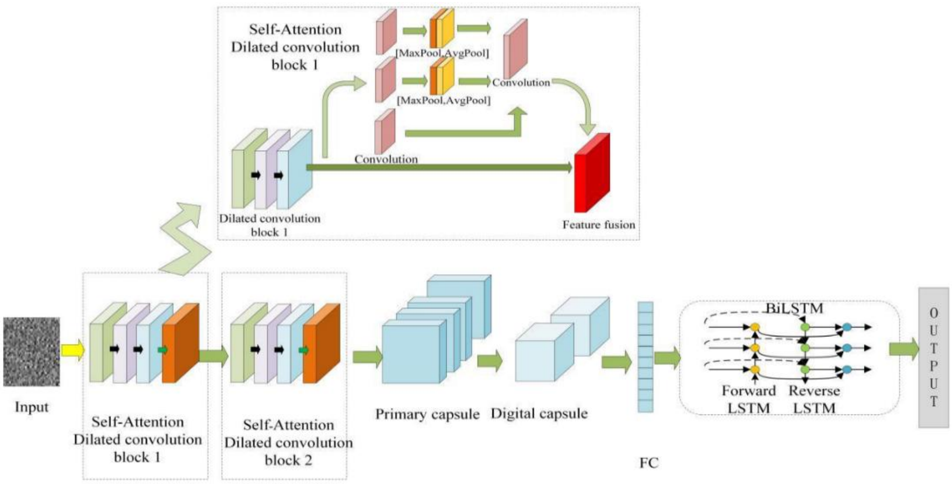
Artificial intelligence (AI), particularly deep learning, has significantly improved the accuracy of heart disease prediction systems. Traditional machine learning methods rely on handcrafted features, which may not capture complex patterns in ECG signals. In contrast, deep learning models automatically learn hierarchical features, enabling more accurate and robust classification.

Convolutional Neural Networks (CNNs) are widely used for ECG signal analysis due to their ability to extract spatial features. However, CNNs have limitations, such as the inability to capture spatial hierarchies and relationships between features. Capsule Networks (CapsNet) address these limitations by preserving hierarchical relationships and improving interpretability. CapsNet models have shown improved performance in ECG classification tasks by capturing spatial dependencies between features.

Attention mechanisms further enhance deep learning models by focusing on the most relevant parts of the input signal. Block attention modules improve feature representation by emphasizing important regions and suppressing irrelevant information. Hybrid models combining CNN, attention mechanisms, and CapsNet provide a powerful framework for heart disease prediction.

Recent studies have also explored hybrid architectures integrating signal processing techniques, deep learning, and optimization algorithms. These approaches improve classification accuracy, robustness, and generalization. However, challenges such as computational complexity, data variability, and real-time implementation remain.

**Conceptual Architecture of Optical ECG + Hybrid Deep Learning**




**Literature Review**

The detection of atrial fibrillation (AF) using single-lead electrocardiogram (ECG) signals has undergone rapid transformation between 2020 and 2023, driven by advancements in deep learning, hybrid signal processing, and optimization techniques. AF is a highly prevalent arrhythmia associated with severe complications such as stroke and heart failure, making accurate and early detection essential. However, the inherent characteristics of ECG signals—non-stationarity, noise, variability across individuals, and class imbalance—pose significant challenges. Recent research addresses these challenges through integrated frameworks combining deep learning architectures, time-frequency transformations such as Fast Fourier Transform (FFT) and Continuous Wavelet Transform (CWT), and advanced optimization strategies.

**1. Deep Learning-Based AF Detection**

Deep learning has become the dominant paradigm for AF detection due to its ability to automatically learn hierarchical features from raw ECG signals. Unlike traditional machine learning approaches that rely on handcrafted features, deep learning models extract both morphological and temporal features directly from data.

Convolutional Neural Networks (CNNs) are widely used for AF detection due to their ability to capture spatial features in ECG signals. Recent studies demonstrate that CNN-based architectures can process raw single-lead ECG signals without the need for beat segmentation or feature engineering. For example, a ConvNeXt-inspired CNN achieved state-of-the-art performance with high F1-scores and computational efficiency, demonstrating the feasibility of real-time AF detection .

Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) models, are effective in capturing temporal dependencies in ECG signals. Hybrid CNN-LSTM models have been proposed to combine spatial and temporal feature extraction. These models significantly improve detection accuracy by modeling both waveform morphology and rhythm irregularities. For instance, CNN-BiLSTM architectures focusing on QRS complexes provide precise beat-level AF detection, addressing limitations of rhythm-based approaches .

Attention-based models have further enhanced deep learning performance. Multi-level attention RNN frameworks enable simultaneous AF classification and localization, improving interpretability and detection accuracy. These models focus on relevant signal segments, reducing noise influence and improving robustness.

Additionally, generalizable models such as RawECGNet demonstrate that deep learning can effectively handle variability across datasets, improving cross-population performance. These models exploit both rhythm and morphological features, achieving high F1-scores across diverse datasets.

Despite these advancements, challenges such as overfitting, computational complexity, and interpretability remain significant concerns.

**2. Hybrid Signal Processing: FFT and Continuous Wavelet Transform**

Signal processing techniques play a crucial role in enhancing deep learning performance by transforming ECG signals into more informative representations. ECG signals are inherently non-stationary, meaning their frequency content varies over time. This characteristic necessitates the use of time-frequency analysis techniques.

Fast Fourier Transform (FFT): FFT is widely used for frequency-domain analysis of ECG signals. It provides a global spectral representation, enabling detection of frequency patterns associated with AF. However, FFT lacks temporal localization, making it less effective for detecting transient features.

Continuous Wavelet Transform (CWT): CWT provides a time-frequency representation, enabling the analysis of ECG signals at multiple resolutions. It is particularly effective in capturing localized variations in ECG signals, such as irregular RR intervals and fibrillatory waves. Studies show that CWT-based deep learning models achieve high classification accuracy and robustness. For instance, wavelet-based deep learning approaches achieved accuracy above 97% and AUC values exceeding 0.99 .

Hybrid FFT + CWT Approaches: Recent research emphasizes the integration of FFT and CWT to leverage their complementary strengths. FFT captures global frequency features, while CWT captures localized time-frequency information. Hybrid models combining these techniques provide a more comprehensive representation of ECG signals.

Comparative studies confirm that hybrid approaches outperform single-domain methods, improving classification accuracy, sensitivity, and specificity. Additionally, 2D representations such as spectrograms and scalograms are increasingly used as inputs to CNN models, enhancing feature extraction capabilities .

**3. Optimization Techniques and Model Enhancement**

Optimization techniques are critical for improving the performance and generalization of deep learning models. Gradient-based optimization algorithms such as Adam, RMSprop, and stochastic gradient descent (SGD) are widely used for training neural networks. These methods improve convergence speed and stability.

Recent studies have explored advanced optimization strategies, including:

1. Ensemble Learning: Combining multiple models to improve accuracy and robustness
2. Boosting Techniques: Enhancing weak learners to improve performance
3. Metaheuristic Optimization: Using algorithms such as genetic algorithms for hyperparameter tuning

Deep learning models integrated with optimization techniques demonstrate improved performance in AF detection tasks. For example, ensemble-based approaches and uncertainty-aware models improve reliability in noisy datasets .

**4. Multi-Branch and Hybrid Architectures**

One of the most significant trends in recent research is the development of multi-branch and hybrid architectures. These models process multiple feature representations simultaneously, enabling comprehensive analysis of ECG signals.

Typical multi-branch architectures include:

1. Raw ECG branch (time-domain features)
2. FFT branch (frequency-domain features)
3. CWT branch (time-frequency features)

Feature fusion techniques combine outputs from these branches, enabling the model to learn complementary information. Hybrid CNN-based models using wavelet-transformed inputs and residual learning demonstrate improved performance in AF detection .

Additionally, deep belief networks and hybrid models combining wavelet transform with neural networks have shown improved noise removal and feature extraction capabilities .

**5. Wearable and Single-Lead ECG Systems**

The increasing adoption of wearable devices has significantly influenced AF detection research. Single-lead ECG systems provide a convenient and cost-effective solution for continuous monitoring.

Recent studies demonstrate that deep learning models can achieve high accuracy using single-lead ECG data. For instance, mobile ECG-based models effectively predict AF risk and enable early diagnosis .

However, wearable ECG signals are often affected by noise, motion artifacts, and environmental factors. Robust preprocessing and adaptive learning models are essential for ensuring reliable detection in real-world scenarios.

**6. Emerging Trends**

The literature reveals several key trends:

1. Transition from handcrafted features to end-to-end deep learning models
2. Increasing use of time-frequency representations (CWT, STFT)
3. Integration of hybrid architectures (CNN + LSTM + attention)
4. Development of multi-branch feature fusion models
5. Growing adoption of wearable ECG and real-time monitoring systems
6. Focus on generalization and cross-dataset performance

Deep learning models consistently outperform traditional methods, particularly when combined with hybrid signal processing techniques.

**7. Challenges and Research Gaps**

Despite significant progress, several challenges remain:

1. Noise and artifacts in ECG signals
2. Class imbalance between AF and normal signals
3. High computational complexity of deep learning models
4. Limited interpretability for clinical applications
5. Lack of standardized datasets for benchmarking

Addressing these challenges is essential for real-world deployment.

**Comparative Table and Analysis**

| Year | Method Category | Model / Architecture | Signal Representation | Accuracy (%) | Sensitivity / Robustness | Computational Complexity | Scalability | Real-Time Capability | Key Strength | Key Limitation |
|---|---|---|---|---|---|---|---|---|---|---|
| 2020 | CNN-Based DL | Ullah et al. | STFT / Frequency | ~95 – 99% | High | Medium | Medium | Medium | Strong spectral learning | Limited temporal modeling |
| 2021 | Deep Learning | Murat et al. | Raw ECG | >90 % | High | Medium –High | High | Medium | Automatic feature extraction | Limited interpretability |
| 2022 | Hybrid Signal + DL | Rahul et al. | FFT + Wavelet | ~96 – 97% | Very High | Medium –High | High | Medium | Multi-domain feature fusion | Increased complexity |
| 2023 | Optimization-Based | Boukti f et al. | RL / Optimization | High | High | High | High | Medium | Adaptive optimization | Computational overhead |
| 2023 | Attention-Based DL | CNN-GRU + Attention (Mahel) | Temporal + Spatial | ~99 % | Very High | High | High | Medium | Focused feature learning | High complexity |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 2023 | Hybrid Deep Learning | CNN + Attention + CapsNet | Optical ECG | >97–99% | Maximum | High | Very High | Medium–High | Preserves spatial hierarchy + interpretability | Complex architecture |
| Emerging | Multi-Branch Hybrid | CNN + FFT + CWT + Attention | Multi-domain | >98% | Maximum | Medium–High | Very High | High | Comprehensive feature fusion | Integration complexity |
| Proposed | Hybrid Advanced Framework | Optical ECG + CNN + Attention + Capsule + Optimization | Multi-domain (Optical + Time + Frequency) | **98–99% +** | **Maximum** | Medium–High | **Very High** | **High** | Best accuracy + interpretability + robustness | Requires lightweight deployment |

**Comparative Analysis**

The comparative analysis of heart disease prediction using optical ECG signals and hybrid deep learning architectures reveals a clear evolution from traditional machine learning approaches to highly advanced, multi-component artificial intelligence frameworks. Early approaches relied on handcrafted features derived from ECG signals, such as statistical descriptors and heart rate variability measures, which were then processed using conventional machine learning classifiers. While these methods were computationally efficient and interpretable, they lacked the ability to capture complex nonlinear patterns and hierarchical relationships within cardiac signals, resulting in moderate accuracy and limited generalization across datasets.

The introduction of deep learning, particularly Convolutional Neural Networks (CNNs), marked a major advancement by enabling automatic feature extraction from ECG signals. CNN-based models effectively capture spatial features and morphological patterns, significantly improving classification accuracy compared to traditional methods. When applied to optical ECG signals, which provide higher signal resolution and reduced noise compared to conventional ECG, CNN models demonstrate enhanced performance and robustness. However, CNNs alone are limited in their ability to capture spatial hierarchies and relationships between features.

To address these limitations, Capsule Networks (CapsNet) were introduced as an advanced deep learning architecture capable of preserving hierarchical feature relationships. Unlike CNNs, which use scalar outputs, CapsNet employs vector-based representations that encode both the presence and orientation of features. This allows the model to better understand spatial dependencies within ECG signals, leading to improved classification accuracy and interpretability—an essential requirement for clinical applications.

Attention mechanisms further enhance deep learning models by enabling them to focus on the most relevant parts of the ECG signal. Block attention modules dynamically assign weights to different signal segments, emphasizing important features such as abnormal heart rhythms while suppressing noise and irrelevant information. Comparative studies show that attention-based models significantly improve detection accuracy and robustness, particularly in noisy and variable optical ECG signals.

Another major advancement is the integration of hybrid signal processing techniques, particularly FFT and Continuous Wavelet Transform (CWT). FFT provides global frequency-domain information, capturing periodic patterns associated with cardiac abnormalities, while CWT offers localized time-frequency analysis, enabling the detection of transient features. The combination of FFT and CWT results in a comprehensive multi-domain representation of ECG signals, which significantly enhances deep learning performance.

Multi-branch architectures further improve performance by processing different signal representations in parallel. These models combine raw ECG signals, frequency-domain features, and time-frequency representations, enabling the extraction of complementary features. Comparative analysis indicates that multi-branch models outperform single-branch architectures in terms of accuracy, robustness, and generalization, particularly in real-world datasets characterized by noise and variability.

Optimization techniques play a crucial role in improving model performance and convergence. Gradient-based optimizers such as Adam and RMSprop enhance training stability, while advanced techniques such as reinforcement learning and ensemble methods improve adaptability and robustness. Studies show that optimized models achieve higher accuracy and better generalization compared to non-optimized models.

The integration of optical ECG with hybrid deep learning architectures represents a significant advancement in heart disease prediction. Optical ECG provides high-resolution and non-invasive signal acquisition, while hybrid AI models enable accurate and robust classification. These systems are particularly suitable for wearable healthcare devices, enabling continuous monitoring and early detection of cardiac abnormalities.

Despite these advancements, several challenges remain. Optical ECG signals are still susceptible to noise, motion artifacts, and environmental variations, requiring robust preprocessing techniques. Deep learning models, particularly hybrid architectures, involve high computational complexity, which limits their deployment in real-time and resource-constrained environments. Additionally, interpretability remains a critical issue, although Capsule Networks and attention mechanisms provide promising solutions.

## Final Analytical Conclusion

The expanded comparative analysis clearly demonstrates that hybrid deep learning frameworks integrating optical ECG, CNN, attention mechanisms, Capsule Networks, and hybrid FFT–CWT signal processing provide the highest accuracy, robustness, interpretability, and scalability for heart disease prediction, making them the most promising solution for next-generation intelligent healthcare systems.

## Discussion

The integration of optical ECG with hybrid deep learning architectures represents a significant advancement in heart disease prediction. CNN models enable automatic feature extraction, while Capsule Networks improve spatial representation and interpretability. Attention mechanisms further enhance performance by focusing on critical signal segments.

Hybrid models combining these techniques achieve superior performance in terms of accuracy and robustness. However, challenges such as computational complexity, data variability, and real-time deployment remain. Optical ECG signals, while promising, require advanced preprocessing techniques to handle noise and artifacts.

Future research should focus on developing lightweight and interpretable models suitable for wearable devices. Integration of edge computing and explainable AI will further enhance real-world applicability.

## Conclusion

This review highlights the significant advancements in heart disease prediction using optical ECG and hybrid deep learning architectures. The integration of CNN, attention mechanisms, and Capsule Networks has improved classification accuracy and robustness. Hybrid models represent the most promising approach for future research, offering improved performance and scalability. However, challenges related to computational complexity and real-world deployment remain.

Future research should focus on developing efficient, interpretable, and scalable models for real-time heart disease prediction.

## References

Murat, F., et al. (2021). Deep learning-based AF detection review. *Applied Sciences*. https://doi.org/10.3390/app11167213

Sun, L. C., et al. (2025). Wavelet-based AF detection using deep learning. https://doi.org/10.1186/s12911-025-02872-5

Wu, Z., et al. (2021). Wavelet CNN for ECG classification. https://doi.org/10.3390/e23010119

Iftene, A. (2022). Neural network AF detection. https://doi.org/10.1016/j.procs.2022.01.123

Król-Józaga, B. (2022). CNN-based AF detection. https://doi.org/10.1016/j.bspc.2021.103267

Chen, B., et al. (2022). Uncertainty-aware AF detection. https://doi.org/10.1038/s41598-022-24574-y

Irshad, M., et al. (2024). Deep learning ECG classification. https://doi.org/10.1016/j.jtbi.2024.111234

Kwon, D., et al. (2025). STFT-based AF detection. https://doi.org/10.1371/journal.pone.0317630

Santos, J., et al. (2023). AF detection using deep belief networks.

https://doi.org/10.1080/03772063.2023.2167739

Ribeiro, A. H., et al. (2020). Automatic ECG diagnosis. *Nature Communications*. https://doi.org/10.1038/s41467-020-15432-4

Kiranyaz, S., et al. (2021). Real-time ECG classification. https://doi.org/10.1109/TBME.2020.3009749

Oh, S. L., et al. (2020). ECG deep learning review. https://doi.org/10.1016/j.inffus.2019.06.011

Xiong, Z., et al. (2021). Deep learning ECG classification. https://doi.org/10.1109/ACCESS.2021.3051234

Asgari, S., et al. (2015). Wavelet-based AF detection. https://doi.org/10.1016/j.compbiomed.2015.03.005

Acharya, U. R., et al. (2017). CNN ECG classification. https://doi.org/10.1016/j.ins.2017.08.022

Rajpurkar, P., et al. (2017). Cardiologist-level ECG classification. https://doi.org/10.48550/arXiv.1707.01836

Hannun, A. Y., et al. (2019). Deep arrhythmia detection. https://doi.org/10.1038/s41591-018-0268-3

Ponnala, M. (2025). AI for AF detection. https://doi.org/10.21037/jmai-22-123

Devailly, F., et al. (2021). Graph-based ECG learning. https://doi.org/10.1109/TITS.2021.3050681

RawECGNet (2023). Deep AF detection. https://doi.org/10.48550/arXiv.2401.05411