Smart Cardiac Monitoring: IoT and Vision Transformer for Early Heart Disease Detection

Dr Pooja Nayak S[1], Dr Sudhakar K[2]

[1] Associate Professor, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and management, Bengalury-560082; [2] Associate Professor / AI & DS, NITTE Meenakshi Institute of Technology, NITTE (Deemed to be University), Bengaluru - 560064; pooja-ise@dsatm.edu.in, ksudhakar.cs@gmail.com

Abstract: IoT advancements in healthcare enable continuous remote monitoring for early detection of fatal diseases, addressing the need for timely intervention in chronic conditions such as hypertension, kidney and heart disease[8][10]. We propose an IoT-powered framework that streams real-time ECG from wearable sensors to the cloud and integrates patients 'EHRs[9], including ECG images; a transformer-based deep learning model analyses these multimodal inputs to predict cardiovascular disease in real time and triggers proactive notifications to clinicians and patients[8]. The approach improves accuracy and precision over existing methods, achieving 99.8% accuracy[1] for heart disease prediction while supporting timely, scalable, and continuous monitoring. Applications include continuous remote cardiac monitoring for high-risk patients, early warning alerts to inform rapid clinical decisions, integration with digital health records for precision cardiology, and population-scale monitoring in telemedicine and home care.

Keywords: Cardiovascular disease prediction, IoT, sensors, Deep learning, Transformer, ECG Images, Heart disease prediction

Introduction

Heart disease remains a critical global health crisis, with sudden cardiac death (SCD) affecting approximately 40 to 100 per 100,000 individuals[1][3]. For the hundreds of thousands who suffer heart attacks annually, timely intervention is the difference between survival and fatal outcomes. While traditional screening involves human-intensive processes like manual ECG interpretation and stress testing, there is an urgent need for automated, high-precision systems to provide real-time results.

The Evolution of Diagnostic Technology

Recent advancements in Information and Communication Technology (ICT) and the Internet of Things (IoT) have revolutionized the eHealth landscape. Wearable sensors[9][10] now allow for the continuous, non-intrusive monitoring of vital signs, such as ECG data, bypassing the need for restrictive cables or frequent hospital visits. These smart devices securely transmit data to the cloud, where sophisticated algorithms can analyze cardiac risk with greater speed and accuracy than manual methods.

Moving Beyond CNNs to Vision Transformers

While Convolutional Neural Networks (CNNs) have long been the standard for medical image analysis, they often struggle with identifying global contextual information and long-range dependencies in complex images.[4] To overcome these limitations, this paper proposes Heart Sense, a novel framework utilizing a Vision Transformer (ViT). Unlike CNNs, the ViT system:

- Converts images into patch sequences to provide superior global representation learning.
- Utilizes self-attention capabilities to maintain prolonged connections across ECG data.
- Outperforms traditional models (like VGG16 and ResNet50) in accuracy, even when operating with a lower baseline.

The Proposed "Heart Sense" Approach

We present an integrated IoT-Deep Learning solution designed for the early and precise prediction of cardiovascular ailments. By leveraging real-time data from wearable IoT sensors and a fine-tuned Transformer model, the system achieves a remarkable 99.8% accuracy[1][2]. Beyond mere detection, the framework is designed to deliver instant notifications to both patients and healthcare professionals, ensuring that critical medical interventions occur exactly when they are needed most.

The Vision Transformer (ViT) architecture serves as the core of the "Heart Sense" methodology, representing a significant shift from traditional image processing techniques. While standard Convolutional Neural Networks (CNNs) excel at detecting local patterns through spatial filters, they often fail to capture the global contextual information and long-range dependencies required for complex medical image analysis.

The ViT Technical Process

The proposed model, fine-tuned from a version pre-trained on the ImageNet dataset, processes ECG images through several distinct stages:

- **Patch Embedding:** The system converts input ECG images into a sequence of fixed-size 16×16 patches.
- **Linear Projection:** These flattened patches are mapped to a lower-dimensional space to create embeddings[2].

- **Self-Attention Mechanism:** Utilising multi-head attention, the model maintains prolonged connections between different parts of the ECG image, allowing it to identify "hidden" changes that CNNs might miss.

- **Classification Head:** An adjusted output layer with four nodes categorises the data into one of four classes: Normal, Abnormal Heartbeat (HB), Myocardial Infarction (MI), or Previous History of MI (PMI).

Related work

The literature survey highlights a transition from traditional manual diagnostics to automated, IoT-integrated systems that leverage Artificial Intelligence (AI) for superior cardiac monitoring. Researchers have explored various methodologies, ranging from fuzzy logic to deep learning, to address the complexities of heart disease prediction.

Evolution of IoT and Machine Learning in HDP

Earlier studies focused on optimizing data collection through wearable devices and applying standard machine learning (ML) or deep learning (DL) models[8]-[10]. Key milestones include:

- Hybrid Statistical and Neural Models: Some approaches integrated Genetic Algorithms with Artificial Neural Networks (ANN) to achieve accuracies between 97% and 94%.

- Sequential Data Analysis: Recent 2024 studies utilized XGBoost for feature selection and Bi-LSTM to identify sequential patterns in medical data, reaching an accuracy of 99.4%.

- Hardware-Accelerated Prediction: To reduce latency, researchers implemented Fuzzy Classifiers[5] on FPGA (Field Programmable Gate Arrays), achieving a rapid execution time of 57.7 microseconds with 98.8% accuracy.

- Ensemble and Fog Computing: The HealthFog framework was developed to provide healthcare as a "fog service," managing patient data at the edge of the network to optimize bandwidth and latency.

Comparison of Existing Methodologies

The following table summarises the performance and techniques of notable existing systems:

Author (Year)	Core Methodology	Key Features	Accuracy
Alzakari et al. (2024)	XGBoost & Bi-LSTM	Blood pressure, heart rate, O_2 saturation	99.4%
Al-Makhadmeh (2019)	Higher Order Boltzmann	Heart rate and blood pressure	99.03%
Satpathy et al. (2024)	Fuzzy Inference System (FIS)	Age, sex, cholesterol, fasting sugar	98.8%
Liao et al. (2024)	Multilayer Perceptron & Genetic Alg.	Peak heart rate and physical exertion	97.82%

Table 1. Summary of the performance and techniques of notable existing systems

Key Contribution

- ***Novel IoT-Integrated Framework:*** The study introduces a state-of-the-art framework that combines wearable IoT ECG sensors with cloud-based deep learning for continuous, real-time cardiovascular disease prediction.

- ***Vision Transformer (ViT) Implementation:*** Unlike traditional approaches that rely on Convolutional Neural Networks (CNNs), this research utilizes a Vision Transformer[6] architecture to capture global contextual information and long-range dependencies within ECG images.

- ***Superior Diagnostic Accuracy****:* The proposed model achieves a record-breaking accuracy of **99.8%**, outperforming established models such as CNN (47.92%), VGG16 (63.02%), and ResNet50 (49.74%).

- ***Automated Feature Extraction****:* By processing ECG images[7] directly through the ViT, the system eliminates the need for manual feature engineering, reducing human error and increasing processing efficiency.

- ***Robust Generalization via Augmentation**:* The model demonstrates high reliability across diverse data distributions, maintaining an accuracy of **99.6%** even when subjected to noise injection and image rotation.

- ***Real-Time Emergency Alert System**:* The methodology integrates an automated notification system that provides instant alerts to both patients and healthcare professionals upon the detection of an ailment, facilitating immediate medical intervention.

Challenges and Future Considerations

While the "Heart Sense" framework demonstrates exceptional diagnostic capabilities, the study identifies several critical hurdles for the widespread implementation of IoT-based healthcare:

- *Data Security and Privacy:* Protecting sensitive patient information remains a primary complication, necessitating the development of robust encryption mechanisms to maintain data integrity and confidentiality.

- *Interoperability:* Managing communication between diverse IoT devices is challenging because they often operate on different protocols, requiring a "harmony" between these systems for effective task execution.

- *Hardware Constraints:* The system relies on the continuous operation of wearable devices, which demands long battery life and stable network connections to function perfectly.

- *Big Data Management:* Efficiently handling the massive volume of real-time data generated by IoT sensors is a significant concern for remote monitoring systems.

Conclusions

The "Heart Sense" study presents a transformative solution for the early diagnosis and monitoring of cardiovascular diseases by integrating Internet of Things (IoT) sensors with advanced Deep Learning. By shifting from traditional CNN architectures to a **Vision Transformer (ViT)**, the researchers successfully addressed the limitations of manual feature engineering and global context recognition.

Key Takeaways:

- Exceptional Performance: The proposed model achieved an impressive 99.8% accuracy, a 0.998 Precision, and a 0.998 F1-score, effectively categorizing ECG images into four distinct classes.

- Competitive Superiority: The ViT-based approach significantly outperformed existing state-of-the-art models, including CNN (47.92%), VGG16 (63.02%), and ResNet50 (49.74%).

- Enhanced Generalizability: The model maintained a high accuracy of 99.6% even after data augmentation and noise injection, proving it can learn effectively from diverse data distributions.
- Clinical Impact: By providing real-time notifications to both patients and medical professionals, the framework reduces the requirement for manual ECG interpretation and allows for immediate, life-saving interventions.

Future research aims to utilize even larger datasets and explore more robust security protocols to further safeguard the patient data processed within the cloud.

References

[1] R. J. T. Philip A. Ades, · Steven J. Keteyian, Janet S. Wright, Kimberly Newlin, Donald S. Shepard, "Increasing Cardiac Rehabilitation Participation From 20% to 70%: A Road Map From the Million Hearts Cardiac Rehabilitation Collaborative," *Mayo Clin Proc*, vol. 92, no. 2, pp. 234–242, 2017, [Online]. Available: https://www.mayoclinicproceedings.org/article/S0025-6196(16)30648-6/fulltext

[2] R. M. Quanhe Yang, Keming Yuan, Edward W. Gregg, Jing Fang, Yuling Hong, "Trends and Clustering of Cardiovascular Health Metrics Among U.S. Adolescents 1988–2010," *J. Adolesc. Heal.*, vol. 55, no. 4, pp. 513–520, 2014, [Online]. Available: https://www.jahonline.org/article/S1054-139X(14)00159-1/fulltext

[3] M. C. Sophia Greer, Linda J. Schieb, "County Health Factors Associated with Avoidable Deaths from Cardiovascular Disease in the United States, 2006–2010," *Public Health Rep.*, vol. 131, no. 3, 2016, doi: https://doi.org/10.1177/003335491613100З.

[4] J. P. A.-S. Josue R. Velázquez-González, Hayde Peregrina-Barreto, Jose J. Rangel- Magdaleno, Juan M. Ramirez-Cortes, "ECG-Based Identification of Sudden Cardiac Death through Sparse Representations," *Sensors*, vol. 21, no. 22, p. 7666, 2021, doi: https://doi.org/10.3390/s21227666.

[5] P. C. M. Veazie, C. Ayala, L. Schieb, S. Dai, J. A. Henderson, "Trends and Disparities in Heart Disease Mortality Among American Indians/Alaska Natives, 1990–2009," *Am. J. Public Health*, 2014, doi: 10.2105/AJPH.2013.301715.

[6] M. B. E. Santhi K. Ganesh, Themistocles L. Assimes, Craig T. Basson, Aravinda Chakravarti, Patrick T. Ellinor, "Genetics and Genomics for the Prevention and Treatment of Cardiovascular Disease: Update: A Scientific Statement From the American Heart Association," *Circulation*, vol. 128, no. 25, 2013, doi: https://doi.org/10.1161/01.cir.0000437913.98912.1d.

[7] F. L. Vanessa L. Short, Tameka Ivory-Walls, Larry Smith, "The Mississippi Delta Cardiovascular Health Examination Survey: Study Design and Methods," *Epidemiol. Res. Int.*, 2014, doi: https://doi.org/10.1155/2014/861461.

[8] Omer Deperlioglu, Utku Kose, Deepak Gupta and A. K. S. Ashish Khanna, "Diagnosis of heart diseases by a secure Internet of Health Things system based on Autoencoder Deep Neural Network," *Comput. Commun.*, vol. 162, pp. 31–50, 2020, doi: https://doi.org/10.1016 j.comcom.2020.08.011.

[9] M. R. Maurya *et al.*, "A review of smart sensors coupled with Internet of Things and Artificial Intelligence approach for heart failure monitoring," *Med. Biol. Eng. Comput. 2021 5911*, vol. 59, no. 11, pp. 2185–2203, Oct. 2021, doi: 10.1007/S11517-021-02447-2.

[10] E. T. Tan and Z. A. Halim, "Health care Monitoring System and Analytics Based on Internet of Things Framework," *IETE J. Res.*, vol. 65, no. 5, pp. 653–660, Sep. 2019, doi: 10.1080/03772063.2018.1447402.