Designing an ECG signal measurement and analysis system using SVM algorithm for arrhythmia classification

Pham Van Nam*, Nguyen Thu Ha, Pham Van Hung,
Pham Van Minh, Tran Thi Hang

Faculty of Electrical Engineering, Hanoi University of Industry, 298 Cau Dien, Tay Tuu, Hanoi, Vietnam.
*Corresponding author: nampv@haui.edu.vn

Received 7 May 2025; Revised 3 Jun. 2025; Accepted 27 Jun. 2025; Published 25 Aug. 2025.
DOI: https://doi.org/10.54939/1859-1043.j.mst.105.2025.36-43

ABSTRACT

This paper introduces a hardware device designed for measuring electrocardiogram (ECG) signals, integrated with a remote system for automated signal analysis, facilitating health monitoring and heart disease diagnosis by doctors. The device transmits ECG signals in real-time to a server equipped with software for analyzing and classifying ECG signals using the Support Vector Machine (SVM) algorithm. Hermite basis functions are employed to generate feature vectors. The proposed solution has been validated using ECG signals from the MIT-BIH (Massachusetts Institute of Technology, Boston's Beth Israel Hospital) database, achieving a 5.07% error rate in classifying seven types of heart rhythms. The SVM algorithm demonstrates high speed and suitability for server-based data classification needs.

Keywords: Machine learning; Support Vector Machine (SVM); Hermite basis functions; Electrocardiogram (ECG) signals.

1. INTRODUCTION

In today's society, the aging population is growing. As lifestyles evolve, the demand to prioritize the health of elderly family members also rises. However, adults are often occupied with work and lack sufficient time to regularly take their grandparents and parents to hospitals for health check-ups. Currently, many hospitals have adopted IoT (Internet of Things) technology to remotely connect, manage, and monitor centralized data [4, 5, 8, 9]. Data processing software plays a crucial role in swiftly diagnosing and screening diseases to aid doctors in remote examinations, as depicted in the following illustrations:

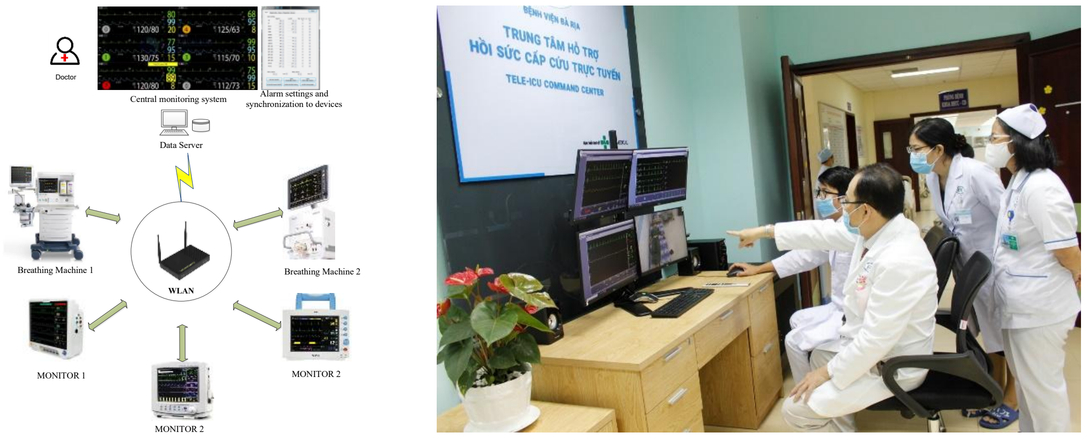


Figure 1. *A remote patient monitoring system at a hospital in Vietnam.*

Recently, advanced methods have been developed for remote ECG measurement and analysis, especially those based on deep learning. Notably, Long Short-Term Memory (LSTM) networks and hybrid models combining LSTM with 1D Convolutional Neural Networks (1D-CNNs) have shown strong performance in handling complex biomedical time-series data like ECG signals [11-

12]. These models are capable of learning temporal dependencies and extracting spatial features simultaneously, making them highly effective for arrhythmia detection and classification tasks. In addition, newer architectures that integrate Convolutional Neural Networks with Transformer models have emerged as powerful solutions [13], leveraging both local feature extraction and global attention mechanisms to enhance the robustness and accuracy of arrhythmia detection.

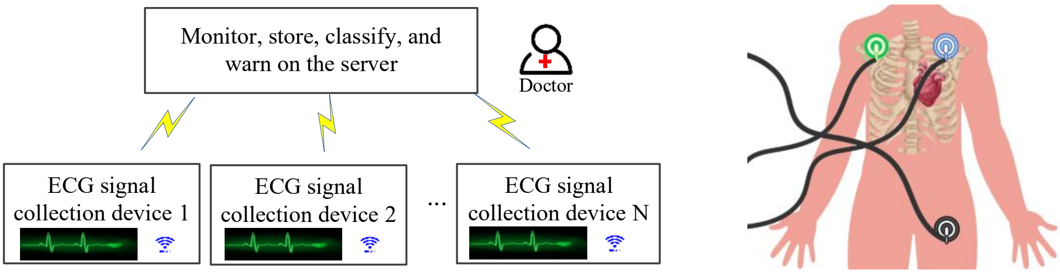
Although deep learning architectures demonstrate impressive performance, their practical deployment remains constrained by several factors, including the need for large annotated datasets, high computational demands, and lengthy training times. Moreover, their black-box nature may limit clinical interpretability, which is critical for medical decision-making. In contrast, this study adopts the Support Vector Machine (SVM) algorithm as a more practical and interpretable solution for arrhythmia classification. SVM offers a solid theoretical foundation, performs effectively in high-dimensional spaces, and generalizes well even with limited training data - conditions commonly encountered in real-world medical datasets. Additionally, SVM's computational efficiency makes it suitable for real-time applications. These attributes make SVM a reliable and transparent alternative to deep learning models, especially in scenarios requiring clarity, resource constraints, and high reliability.

Thus, developing the above models is essential as Vietnam advances into Industry 4.0. This study focuses on:

- Designing a portable ECG device with online signal transmission to support remote monitoring and early diagnosis.

- Developing SVM-based software for ECG signal classification to assist in rapid screening.

The system is validated using the standard MIT-BIH dataset [3]. The proposed model (figure 2) measures ECG and heart rate, transmits data via WiFi to a server, and processes it using server-side programs with sufficient speed and accuracy. The system enables doctors to remotely monitor multiple patients in real time, with built-in alerts for early detection of abnormalities.



a) *b)*

Figure 2. *a) Transmitting ECG signals system via IoT, automatic recognition of ECG signals, and an application for users. b) The location of the recording electrodes.*

2. SYSTEM DESIGN AND ECG SIGNAL PROCESSING

2.1. Hardware design overview

The proposed ECG device is designed for portability and remote monitoring, measuring 3-lead ECG signals (Lead I, II, III) to balance diagnostic capability with compactness. Each lead represents a distinct electrical view of the heart, derived from the potential difference between electrodes placed on the right arm (RA), left arm (LA), and left leg (LL). The device employs three channels, each processing the signal from one lead, enabling simultaneous acquisition of three distinct electrical views. The 3-lead configuration was chosen over the standard 12-lead ECG because it is sufficient for detecting common arrhythmias (e.g., premature ventricular contractions, atrial fibrillation) while enabling a lightweight, handheld design suitable for home use, reducing hardware complexity and power consumption.

The device uses the ADAS1000-3 integrated circuit [10] for high-quality signal acquisition, filtering noise in the 1–100 Hz range (eliminating motion artifacts and power-line interference), and amplifying weak ECG signals (mV range) by approximately 1000 times. Key hardware components include:

- Microcontroller: ESP32, a low-power chip with integrated WiFi for real-time data transmission to a server, enabling medical professionals to make remote diagnoses.

- Memory: 8GB SD Card for onboard storage of ECG recordings, ensuring data integrity and accessibility for offline analysis.

- Display: 3.2-inch TFT LCD - TFT LCD display (480 x 320 pixels) providing clear, detailed real-time ECG waveform display and device status information for medical professionals.

The hardware design incorporates advanced noise filtering (e.g., low-pass/high-pass filters) to improve signal clarity compared to single-lead designs [9]. The functional block diagram of the device is shown in figure 3.

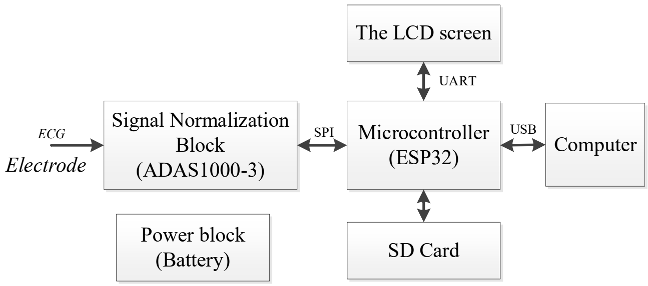


Figure 3. *Block design of the 3-lead ECG measurement device.*

The device uses the ADS1000-3 IC to simplify hardware design and capture higher-quality ECG signals compared to traditional analog circuits.

2.2. ECG signal processing and classification

The ECG signal processing pipeline analyzes the electrical activity of the heart, focusing on key waveform components such as the P wave, QRS complex, and T wave, as illustrated in figure 4. These components reflect critical cardiac events, with the QRS complex being particularly important for arrhythmia detection due to its distinct morphology.

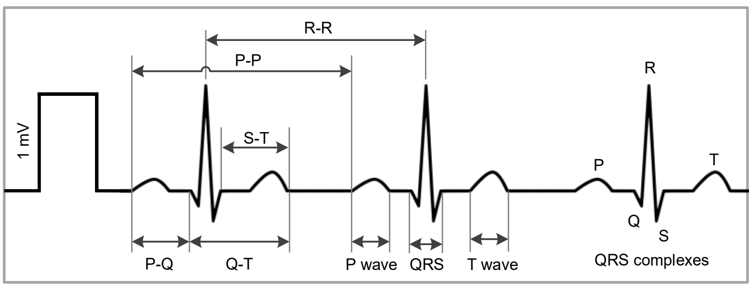


Figure 4. *Shape and main components of ECG signals.*

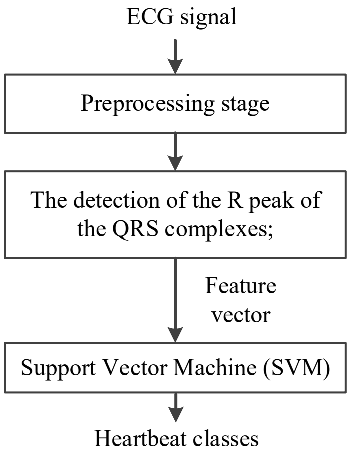


Figure 5. *General model of the ECG signal processing and classification system.*

The processing pipeline, shown in figure 5, consists of feature extraction and classification stages.

2.2.1. Feature extraction

Feature extraction simplifies the original input signal into a streamlined representation with reduced information, reducing the computational workload of the identification block. Feature selection is based on its ability to distinguish objects of interest, which is crucial for accurate and efficient identification results. In this study, Hermite functions were employed for feature extraction due to their morphological resemblance to typical ECG waveforms.

The QRS complex in the ECG signal contains valuable information for identification. The article chose a window of 250 ms around the R peak, enough to capture the QRS segments. R peak detection uses the algorithm of Hamilton and Tompkins [1, 2]; This work is performed on the measuring device. Hermite functions have the following formula for n ≥ 0:

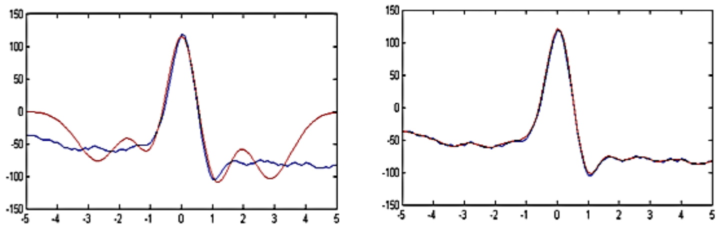
$$\psi_n(x) = \left(\sqrt{\pi} \cdot 2^n \cdot n!\right)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} H_n(x) \tag{1}$$

with $H_n(x)$ is a polynomial Hermite which is defined by:

$$H_{n+1}(x) = 2x \cdot H_n(x) - 2n \cdot H_{n-1}(x) \tag{2}$$

for $n \geq 1$, with $H_0(x) \equiv 1; H_1(x) = 2x$.

Figure 6 illustrates that as the order of the Hermite function increases, the function exhibits more rapid fluctuations. Additionally, the shapes of these functions closely resemble the fundamental components observed in ECG signals, which form the rationale for applying Hermite functions in ECG signal analysis. To represent the ECG signal s(t) as a linear combination of the first N Hermite functions, as shown in equation (3), the Singular Value Decomposition (SVD) algorithm is employed to obtain the optimal solution for an overdetermined linear system-one that has more equations than unknowns. Further explanation can be found in [7].



(a) *(b)*

Figure 6. *ECG signals by N Hermite function: a) N = 10; d) N = 16.*

$$s(t) \approx \sum_{i=0}^{N-1} c_i \cdot \psi_i(t) \tag{3}$$

According to the process of feature extraction of ECG signals, the feature vector:

$\mathbf{X} = [c_0, \ldots, c_{15}] \in \mathbb{R}^{16}$ of each beat (QRS complex) consists of 16 components.

2.2.2. ECG signal classification algorithm

SVM are a powerful and widely used machine learning tool for pattern classification due to their ability to handle high-dimensional data and provide good generalization properties. In this study, we apply SVM to the classification of ECG signals, specifically the QRS complex, for the identification and diagnosis of cardiac issues. When classifying 07 types of ECG signals using the OVO (One-Versus-One) method: You create 21 SVM models, each trained to differentiate between a specific pair of classes (e.g., Class 1 vs. Class 2, Class 1 vs. Class 3, ..., Class 6 vs. Class 7). During classification, each sample is evaluated by all 21 models. Each model decides the class within its designated pair, and the final classification is determined by aggregating the results to select the class with the most votes. This method is chosen for its flexibility in handling multi-class scenarios and effectively managing imbalanced datasets, though it requires more computational resources due to the number of models involved. The algorithms were tested on benchmark datasets from the MIT-BIH (Massachusetts Institute of Technology, Boston's Beth Israel Hospital) [3, 9].

3. EXPERIMENTAL RESULT

3.1. Dataset and evaluation metrics

The study used the MIT-BIH arrhythmia database for the first set of ECG signals in this paper, a widely used benchmark for evaluating arrhythmia detection systems. The dataset was collected from 47 adult subjects, both male and female, with ages ranging from 22 to 89 years, all presenting various cardiac conditions. ECG signals were sampled at 360 Hz (360 data points per second). Seven types of arrhythmia were classified: Left Bundle Branch Block Beat (L), Right Bundle Branch Block Beat (R), Premature Atrial Contractions (A), Premature Ventricular Contraction (V), Ventricular Flutter Wave (I), Ventricular Escape Beat (E), and Normal beat (N).

Table 1. *Training and testing samples for 7 arrhythmia types (MIT-BIH database).*

Beat type	Total samples	Training samples	Test samples
N	2000	1400	600
L	600	420	180
R	500	350	150
A	400	280	120
V	400	280	120
I	200	140	60
E	80	56	24
Total	**4180**	**2926**	**1254**

SVM: Employing a selection structure as in [6], the 7-class classification SVM model utilizes the OVO method, resulting in 21 sets of binary SVM components. Test results, summarized in Table 2, show a total error rate of 5.07% (64 errors out of 1254 samples), with 14 False Negatives (FN) and 25 False Positives (FP). The low FN rate (1.12%) is critical, as it minimizes missed diagnoses, a significant concern in medical applications. The model achieved a sensitivity of 98% (True Positive Rate, Equation 5) and specificity of 96% (True Negative Rate, Equation 6), indicating high reliability for clinical use.

Table 2. *Matrix of identification results of 7 types of samples by SVM.*

Sample / Results	N	L	R	A	V	I	E
N	578	6	2	6	0	0	0
L	1	168	0	0	3	0	1
R	2	0	145	2	0	0	0
A	18	4	2	112	3	0	0
V	2	2	1	1	113	2	0
I	2	2	1	0	1	58	0
E	0	0	0	0	0	0	23
Total error	**25**	**14**	**6**	**9**	**7**	**2**	**1**

The results of identifying ECG signal software by using an SVM network are as follows:

- Number of error samples: 64 (samples).

$$\%\ \text{Error} = \frac{\text{Number of error samples}}{\text{Number of test samples}} \times 100\% = \frac{64}{1254} \times 100\% = 5.07\% \tag{4}$$

- **FN** (*False Negative*): A total of 14 false negative cases were observed, in which patients with the actual disease were incorrectly classified as normal.

- **TN** (*True Negative*): Number of cases which is diagnosed as true negatives: 578 samples.
- **FP** (*False Positive*): Number of cases which is diagnosed falsely positive, it means patients are diagnosed as normal, but being misdiagnosed is diseased: 25 samples.
- **TP** (*True Positive*): Number of cases which is diagnosed as true positive: 644 samples.
- **Sensitivity** (*True Positive Rate*): Ratio of true positive diagnoses.

$$\text{Sensitivity } = \frac{TP}{TP+FN} \cdot 100\% = \frac{644}{644+14} \cdot 100\% = 97{,}87\% \quad (5)$$

- **Specificity** (*True Negative Rate*): Ratio of true negative diagnoses.

$$\text{Specificity } = \frac{TN}{TN+FP} \cdot 100\% = \frac{578}{578+25} \cdot 100\% = 95{,}85\% \quad (6)$$

In evaluating model quality based on the criteria mentioned above, FN (14 samples) is critical because a high FN error could pose significant risks, as the model may fail to detect diseases. Thus, a lower number of FN cases indicates higher model quality. Therefore, the SVM-based ECG signal identification model performs effectively.

3.2. Hardware performance

As illustrated in figure 7, we see that the electrocardiogram signal of the measuring device using IC ADAS1000-3 gives results equivalent to the electrocardiogram device NT Cardio ECG-1100 in terms of measured heart rate, and ECG waveforms on all 3 leads. Some published works on the design of ECG measurement modules only measure 01 lead, and the noise filtering is not thorough. The device is designed to measure 3 leads simultaneously and give clearer signals.

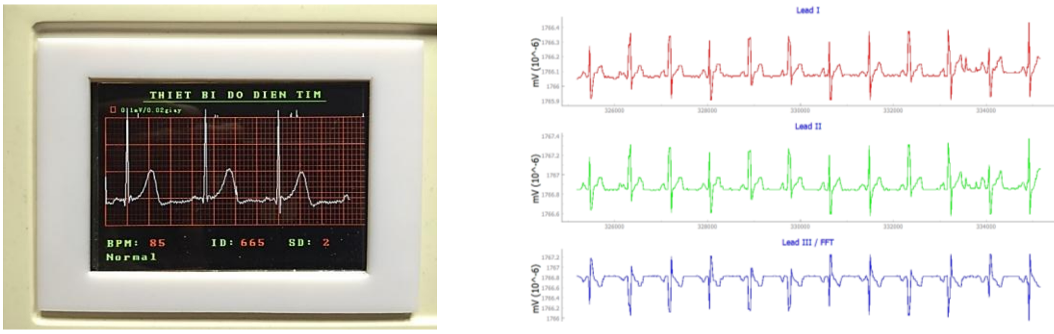


Figure 7. *ECG signal collection module and displayed results on computer interface.*

A web server was developed to collect and store device data, with a simple interface consisting of two sections: Home and Patient Details, as shown below.

Figure 8. *Main website screen, detailed screen when clicking "View" on the main website.*

5. CONCLUSIONS

The paper proposes an effective IoT solution for remote health monitoring using the ADAS1000-3 sensor to collect ECG signals, which are transmitted to the ESP32 microcontroller. This enables doctors and medical centers to remotely monitor patients' health, with potential for commercial development. Additionally, the paper outlines plans to utilize SVM for ECG signal identification software. While not the optimal solution, SVM offers simplicity, speed, and suitability for handling large datasets. The algorithms were validated using MIT-BIH datasets and by international research groups, showing high recognition accuracy. Future plans include device optimization, collaboration with medical centers for ECG data collection, and gradual testing on patients.

Acknowledgement: *This work was supported by the Project Granted number [33-2022-RD/HĐ – ĐHCN], Hanoi University of Industry.*

REFERENCES

[1]. P.S. Hamilton and W.J. Tompkins, *"Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database",* IEEE Trans. Biomed Eng., vol. BME-33, pp. 1157-1165, (1986).

[2]. Thakor, N. V., Webster, J. G., and Tompkins, W. J, *"Optimal QRS detector".* Medical and Biological Engineering, 343–50, (1983).

[3]. R. Mark Moody, *"The impact of the MIT-BIH Arrhythmia Database"*, IEEE Eng. in Medicine and Biology 20(3): 45–50, (2001).

[4]. Jalal A. Nasiri, M. Naghibzadeh, *"ECG Arrhythmia Classification with Support Vector Machines and Genetic Algorithm",* Third UKSim European Symposium on Computer Modeling and Simulation, (2009).

[5]. Pengfei Gao; Jingwei Zhao; Guijin Wang; Hengkai Guo, *"Real time ECG characteristic point detection with randomly selected signal pair difference (RSSPD) feature and random forest classifier"*, 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 732 - 735, (2016).

[6]. Tran Hoai Linh, Pham Van Nam, *"Multiple neural network integration using a binary decision tree to improve the ECG signal recognition accuracy",* International Journal of Applied Mathematics and Computer Science, vol. 24(3), pp. 647–655, (2015).

[7]. Tran Hoai Linh, Pham Van Nam, Nguyen Duc Thao, *"A hardware implementation of intelligent ECG classifier",* COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, vol. 34, Iss: 3, pp. 905 – 919, (2015).

[8]. Patel, R., Garcia, S., Lee, C. *"An IoT-based ECG Monitoring System with Real-Time QRS Detection and Alerting".* Journal of Medical Engineering & Technology, (2020).

[9]. Pham, V. N., & Tran, H. L. *"Electrocardiogram (ECG) circuit design and using the random forest to ECG arrhythmia classification".* In International Conference on Engineering Research and Applications (pp. 477-494). Cham: Springer International Publishing, (2022).

[10]. Nguyen Thanh Trung, Vuong Tri Tiep, Phạm Van Nam, *"Application of ARM IC and ADAS1000-3 for Designing a Cardiac Electrical Measurement Device with QRS Complex Detection Function"*, The International Conference and Exhibition on Control and Automation, 7th Edition -VCCA, (2024).

[11]. Alamatsaz, N., Tabatabaei, L. S., Yazdchi, M., Payan, H., Alamatsaz, N., & Nasimi, F. *"A lightweight hybrid CNN-LSTM model for ECG-based arrhythmia detection",* (2022). arXiv.org. https://arxiv.org/abs/2209.00988

[12]. Roy, M. Majumder, S., Halder, A., & Biswas, U. *"ECG-NET: A deep LSTM autoencoder for detecting anomalous ECG".* Engineering Applications of Artificial Intelligence, 124, 106484, (2023). https://doi.org/10.1016/j.engappai.2023.106484.

[13]. Shah, H., Saeed, F., Diyan, M., Almujally, N., & Kang, J. *"ECG-TransCovNet: A hybrid transformer model for accurate arrhythmia detection using Electrocardiogram signals".* CAAI Transactions on Intelligent Technology. https://doi.org/10.1049/cit2.12293.

TÓM TẮT

Thiết kế hệ thống đo và phân tích tín hiệu ECG sử dụng thuật toán SVM trong phân loại loạn nhịp tim

Bài báo này trình bày một thiết bị phần cứng đề xuất để đo tín hiệu điện tâm đồ (ECG) với hệ thống từ xa tích hợp để phân tích tín hiệu tự động, giúp bác sĩ theo dõi sức khỏe và chẩn đoán các bệnh tim mạch. Thiết bị đo lường này có thể truyền tín hiệu ECG trực tuyến tới máy chủ, máy chủ này được trang bị phần mềm để phân tích và phân loại tín hiệu ECG để phát hiện loạn nhịp tim sử dụng thuật toán Support Vector Machine (SVM). Các hàm cơ sở Hermite được sử dụng để tạo các vector đặc trưng. Giải pháp đề xuất đã được kiểm thử với các tín hiệu ECG lấy từ cơ sở dữ liệu MIT-BIH (Massachusetts Institute of Technology, Boston's Beth Israel Hospital). Bảy loại nhịp tim đã được phân loại với tỉ lệ lỗi là 5.07%. Thuật toán SVM đề xuất hoạt động rất nhanh, phù hợp với yêu cầu phân loại nhanh các bệnh tim mạch.

Từ khoá: Học máy; Máy hỗ trợ véc-tơ SVM; Hàm cơ sở Hermite; Tín hiệu điện tâm đồ (ECG).