Journal of Next-generation Convergence Information Services Technology Vol.14, Suppl. 1, November (2025) pp.29-38
ISSN 2384-101X(p) 2672-1163(e) *http://dx.doi.org/10.29056/jncist.2025.11.03*

ECG R-피크 검출을 위한 Pan-Tompkins와 SVM·RF 기법의 성능 비교 연구

A Comparative Study on the Performance of Pan-Tompkins and SVM·RF Methods for ECG R-Peak Detection

정다엘[1], 이규만[2]*

Dael Jeong[1], Kyuman Lee[2]*

요 약

본 연구에서는 MIT-BIH Arrhythmia Database의 48개 레코드(360 Hz)를 활용하여 Pan–Tompkins 알고리즘, SVM(선형 커널, C=1), Random Forest(n=100)의 R-피크 검출 성능을 동일한 전처리 및 특징 추출 조건에서 비교·분석하였다. R-피크 중심 ±50 샘플(총 100포인트)의 구간을 기준으로 총 8개의 통계적 특징(평균, 표준편차, 최댓값, 최솟값, 중앙값, Q1, Q3, 분산)을 산출하여 특징 벡터를 구성하였다. 이후 클래스 불균형 문제를 해소하기 위해 1:1 언더샘플링을 적용하고, 전체 데이터를 8:2 비율로 학습·테스트 세트로 분할하였다. Pan–Tompkins 알고리즘, SVM, Random Forest 모델에 동일한 특징과 동일한 자료 분할 방식을 적용하여 공정한 비교가 이루어지도록 하였다. 실험 결과, Random Forest는 정확도(0.9214)와 정밀도(0.8961)에서 가장 높은 성능을 보였으며, SVM은 민감도(0.9707)에서 상대적으로 우수하였다. Pan–Tompkins 알고리즘은 다른 두 모델 대비 정밀도에서 다소 낮았으나, 민감도(0.9384) 측면에서는 안정적인 검출 성능을 확인하였다. 이러한 결과는 전통적 알고리즘과 기계학습 기반 알고리즘이 각각 다른 강점을 가지며, 적용 목적에 따라 선택 기준이 달라질 수 있음을 시사한다. 특히 실시간 처리 환경에서는 민감도가 높은 알고리즘이, 오탐지 억제가 중요한 서비스 환경에서는 정밀도가 높은 알고리즘이 적합하다고 판단된다. 본 연구는 동일 조건하에서 Pan–Tompkins, SVM, Random Forest의 성능을 정량적으로 비교함으로써, 향후 ECG 기반 실시간 진단 및 임상적 활용을 위한 알고리즘 선택의 근거를 제공한다.

핵심어 : 심전도(ECG), R-피크 검출, Pan-Tompkins 알고리즘, SVM, 랜덤 포레스트

Abstract

This study compares and analyzes the R-peak detection performance of the Pan–Tompkins algorithm, a Support Vector Machine (SVM; linear kernel, C=1), and a Random Forest (n=100) using 48 records (360 Hz) from the MIT-BIH Arrhythmia Database under identical preprocessing and feature extraction conditions. A total of eight statistical features (mean, standard deviation, maximum, minimum, median, Q1, Q3, variance) were extracted from a ±50-sample (100-point) window centered on each R-peak to construct

1 Department of Bio Medical Devices, Gachon University, Seongnam, Korea [Graduate Student]
e-mail: dael0207@gachon.ac.kr
2 Department of Bio Medical Devices, Gachon University, Seongnam, Korea [Professor]
e-mail: manleekyu@gachon.ac.kr (Corresponding author)

Received(September 22, 2025), Review Result(1st: October 19, 2025), Accepted(November 12, 2025), Published(November 30, 2025)

© 2025 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc/4.0/.

feature vectors. To address the class imbalance, a 1:1 undersampling method was applied, and the dataset was divided into training and test sets at a ratio of 80:20. The same feature set and data-splitting strategy were applied across all three models to ensure fair comparison. Experimental results showed that the Random Forest achieved the highest performance in accuracy (0.9214) and precision (0.8961), while the SVM demonstrated relatively superior performance in sensitivity (0.9707). The Pan–Tompkins algorithm exhibited lower precision compared to the machine-learning models but maintained stable sensitivity (0.9384). These results indicate that traditional and machine-learning-based algorithms have different strengths, and the optimal choice may vary depending on the intended application. In particular, algorithms with higher sensitivity may be more suitable for real-time processing environments, while algorithms with higher precision are appropriate in scenarios where suppression of false positives is critical. This study provides quantitative evidence for selecting appropriate algorithms for real-time ECG-based diagnostic systems and clinical applications under uniform experimental conditions.

Keywords : electrocardiogram (ECG), R-peak detection, Pan-Tompkins algorithm, Support Vector Machine (SVM), Random Forest (RF)

1. 서론

ECG는 심장 전기 활동을 기록하는 표준 생체신호로, 부정맥 조기 탐지 및 환자 모니터링에 일상적으로 사용된다. 특히 QRS 복합파 중 R-피크는 심박수 계산, 부정맥 탐지, 심장박동 주기 분석 등에서 핵심적인 임상 지표로 간주되며, 이 R-피크를 얼마나 정확하게 검출하는지는 전체 진단의 신뢰도에 직접적인 영향을 미친다.

그동안 R-피크 검출을 위해 다양한 알고리즘이 개발되어 왔으며, 이 중에서도 Pan-Tompkins 알고리즘은 1985년 발표 이후 실시간 ECG 분석 시스템에 널리 도입되어 왔다. 해당 알고리즘은 필터링, 미분, 제곱, 적응 임곗값 설정 등 정형화된 단계를 기반으로 비교적 높은 검출 성능을 보여왔으나, 최근에는 다양한 노이즈 환경이나 비정형 심전도 신호에서의 한계가 지적되고 있다.

최근에는 지도학습 분류기(SVM, RF 등)를 R-피크 검출에 적용해 잡음·형태 변화 환경에서의 검출 오류를 줄이는 시도가 활발하다. 그러나 기존 연구들은 각 기법의 기술적 특성이나 알고리즘 성능 자체에 집중되어 있었고, 임상적 응용이나 실시간 진단 시스템에 적용 가능성을 판단할 수 있는 비교 분석 연구는 상대적으로 부족하였다. R-피크 검출이라는 동일한 목적 아래, 서로 다른 방식의 알고리즘들이 실제 환경에서 어떤 차이를 보이는지를 구조적으로 분석한 연구 또한 많지 않다.

한편, 최근 딥러닝 및 기계학습 분야에서는 단순한 정확도 향상 뿐 아니라 실제 임상 또는 서비스 적용성을 함께 고려하는 연구가 늘고 있다 [1]. 예를 들어, 딥러닝을 이용한 척추 자동 분할 및 Cobb angle 측정 시스템을 개발하며 [2], 임상 환경에서의 실시간 적용성과 데이터 품질 제약을 동시에 고찰하였다. 또한 개인정보 삭제권 실현의 기술적 한계를 분석하면서 시스템 설계 시 재현성과 투명성의 중요성을 강조하는 연구도 있다 [3]. 이러한 연구들은 본 연구가 추구하는 “기술적 정밀성과 실제 적용성의 균형”과 같은 문제의식을 공유하며, R-피크 검출 알고리즘의 임상 활용 가

능성을 고려하는 본 연구의 방향성과도 맥락을 같이한다.

본 연구의 목적은 전통적 신호처리(Pan-Tompkins)와 ML(SVM, RF)의 R-피크 검출 특성을 동일 조건에서 정량 비교하는 것이다. 단순히 알고리즘의 성능을 평가하는 데서 나아가, 전통적 기법과 머신러닝 기법 간의 검출 특성을 비교하고, 각 방법이 특정 상황에서 갖는 장단점을 정량적으로 분석하고자 한다. 특히 본 연구는 알고리즘의 일반적 효과나 구조에 대한 기술적 설명보다는, 다양한 신호 조건 속에서 각 방식이 어떻게 작동하고 어떤 상황에 더 적합한지를 밝히는 데 초점을 둔다.

이에 본 연구는 Pan-Tompkins, SVM, RF 세 가지 알고리즘을 중심으로, MIT-BIH 부정맥 데이터베이스에 수록된 48개의 심전도 레코드를 활용하여 평균 정확도, 민감도, 정밀도 등의 측면에서 R-피크 검출 성능을 비교한다. 이는 향후 실시간 ECG 진단 시스템 개발에 있어 어떤 알고리즘을 선택할 것인가에 대한 실증적 근거를 제공하고자 하는 데 목적이 있다.

2. 이론적 배경

2.1 Pan-Tompkins 알고리즘

Pan-Tompkins 알고리즘은 QRS 복합파의 형태학적 특성을 기반으로 실시간 ECG 신호에서 R-피크를 검출하는 방법이다. 본 기법은 밴드패스 필터링, 1차 차분, 제곱 변환, 이동 평균 및 적응 임곗값을 적용하여 노이즈 제거와 동시에 QRS 파형을 강화한다 [4]. 알고리즘의 구조적 장점은 실시간 구현 용이성에 있으나, 다양한 환자 및 노이즈 조건에서는 민감도가 저하되는 단점이 있다.

2.2 SVM (Support Vector Machine)

본 연구의 SVM 분류기는 sklearn.svm.SVC로 구현하였다. 커널은 선형(Linear)을 사용하고, 하이퍼파라미터는 C=1.0, class_weight='balanced', random_state=42로 고정했다. 특징 스케일에 따른 영향을 줄이기 위해 학습 세트의 평균과 표준편차로 z-score 표준화를 수행하고(3.2의 8개 통계 특징), 동일 변환을 테스트 세트에 적용하였다. 학습에는 3.3에서 구성한 언더샘플링 1:1 데이터와 stratified 8:2 분할을 그대로 사용하였다.

선형 커널을 채택한 이유는 다음과 같다. 입력 차원이 8로 낮고 특징 간 상관이 높지 않아 선형 결정경계로도 안정적인 분리가 가능했다. 비선형 커널(RBF 등)은 튜닝 공간(C, γ)이 커서 데이터 규모 대비 탐색 비용이 과도하며 [5][6], 목적이 방법 간 공정 비교인 점을 고려해 복잡한 모델을 피했다. 선형 모델의 가중치 벡터는 특징 기여도를 해석하기 쉽고, 실시간 적용 시 예측 비용이 작다.

모든 전처리(표준화)와 모델 학습은 훈련 세트 정보만 사용했으며, 테스트 세트에는 어떤 형태의 통계도 누수되지 않도록 했다.

2.3 Random Forest

랜덤 포레스트는 sklearn.ensemble.RandomForestClassifier로 구현하였다. 기본 설정은 n_estimators=100, max_depth=None, min_samples_leaf=1, max_features='sqrt', bootstrap=True, random_state=42이다. 입력은 SVM과 동일하게 3.2에서 정의한 8개 특징을 사용했고, 데이터 구성 역시 3.3의 언더샘플링 1:1과 stratified 8:2 분할을 그대로 따랐다(표준화 여부와 무관하게 RF는 성능 저하가 없음을 확인하여 동일 파이프라인 입력을 사용).

설정 근거는 다음과 같다. 다수 트리의 배깅으로 분산을 낮추고 과적합을 억제할 수 있어, 레코드 간 신호 변동성이 큰 ECG에서 안정적이다. max_features='sqrt'는 트리 간 상관을 줄여 앙상블 이득을 확보하는 표준 선택이다 [7]. max_depth=None과 100개 트리는 본 데이터 규모에서 예측 성능과 연산 비용의 균형을 제공했다. 클래스 불균형은 3.3의 언더샘플링으로 해소했기에 추가 가중치(class_weight)는 사용하지 않았다.

모든 난수 관련 설정은 random_state=42로 고정했고, 학습 과정에서 계산된 어떤 통계도 테스트 샘플에 선행 유입되지 않도록 처리했다.

Pan-Tompkins와 머신러닝 기반 기법(SVM·RF)의 장단점과 적용 환경을 비교한 결과는 [표 1]과 같이 정리하였다.

[표 1] Pan-Tompkins 및 머신러닝 기반 R-피크 검출 기법 비교

[Table 1] Comparison of Pan-Tompkins and ML-based R-peak Detection Methods

특징/기법	Pan-Tompkins 알고리즘	SVM·RF (머신러닝 기반)
장점	-고전적, 널리 사용: ECG 신호 처리 표준으로 자리매김 -낮은 복잡도: 구현 용이, 계산 비용 적음 -우수한 검출 성능 (고품질 신호): 깨끗한 신호에서 높은 정확도 (MIT-BIH 99.3% 이상)	-강건성: 다양한 잡음 및 비정상 파형에 강함 -높은 정확도 (복잡한 신호): 판별 어려운 R-피크, 잡음 환경에서 우수 -다양한 특징 활용: 시간-주파수, 통계적 특징 등 복합적으로 학습 가능
단점	-잡음에 취약: 저품질 신호에서 성능 저하, 오검출 발생 가능 -부정맥에 한계: 비정상 QRS 형태 검출 어려움 -고정된 임곗값 문제: 신호 진폭 변화에 취약 (개선된 버전에서 보완)	-높은 계산 복잡도: 학습 시간 길고 실시간 적용에 제약 (특히 SVM) -데이터 의존성: 학습 데이터의 품질과 양에 성능 좌우 -특징 추출 중요: 효과적인 특징 추출 과정 및 도메인 지식 요구
주요 활용 분야	깨끗하고 비교적 정상적인 ECG 신호 분석, 실시간 임베디드 시스템	잡음이 많거나 다양한 부정맥이 포함된 복잡한 ECG 신호 분석, 정교한 진단 시스템

적합성	구현 용이성과 빠른 처리가 중요한 경우	높은 정확성과 강건성이 요구되는 경우, 복잡한 패턴 분석이 필요한 경우

3. 연구 방법

3.1 데이터셋

분석에는 PhysioNet의 MIT-BIH Arrhythmia DB 중 48개 레코드를 사용했다(채널 1, 360 Hz). 각 레코드는 360Hz의 고정 샘플링 주파수로 수집되었으며, 각 레코드는 30분 길이의 2채널 ECG와 전문가 R-피크 주석을 포함하며, 본 연구는 채널 1만 사용했다 [8][9].

3.2 특징 추출

R-피크 전후 50 샘플씩 총 100포인트 윈도우 내의 신호 구간을 기준으로 R-피크 중심 구간에서 산술평균, 표준편차, 최대/최소, 중앙값, 사분위수(Q1/Q3), 분산의 8개 통계 특징을 구성했다.

R-피크 중심 ±50 샘플 윈도우 내에서 추출한 8개 통계적 특징을 [표 2]와 같이 정리하였다.

[표 2] R-피크 중심 윈도우 내 특징 추출 항목
[Table 2] Extracted Features within R-peak Centered Window

No.	Description
1	평균 (Mean)
2	표준편차 (Standard Deviation)
3	최댓값
4	최솟값
5	중앙값 (Median)
6	Q1
7	Q3
8	분산

이를 통해 각 구간을 8차원 특징 벡터로 구성하였으며, 해당 지점이 R-피크일 경우에는 양성(positive)으로, R-피크가 아닌 경우에는 음성(negative) 클래스로 라벨링하였다.

3.3 전처리 및 모델 학습

본 연구에서는 ECG R-피크 검출 모델의 효과적인 학습 및 성능 평가를 위해 데이터의 전처리

및 데이터셋 구성을 체계적으로 수행하였다. 특히, ECG 신호의 특성으로 인해 발생하는 클래스 불균형 문제를 해결하고 모델의 일반화 성능을 검증하고자, ECG 신호를 대상으로 한 데이터 분할을 수행하였다.

3.3.1 클래스 불균형 해소: 언더샘플링 기반 데이터 균형화

레코드별로 양성(R-피크) 수와 동일한 음성 구간을 무작위 선택해 1:1로 균형화했다 (random_state=42, 중복 추출 금지). 이러한 불균형한 데이터셋으로 모델을 학습시킬 경우, 이 균형화로 학습 편향을 줄였고, 양·음성의 탐지 성능을 균형 있게 평가할 수 있도록 구성하였다.

이러한 문제를 해결하고자, 본 연구에서는 각 ECG 레코드별로 언더샘플링(Undersampling) 기법을 적용하여 클래스 간의 균형을 맞추었다. 구체적으로, 다수 클래스인 R-피크가 아닌 구간에서 소수 클래스인 R-피크 구간과 동일한 수량의 샘플을 무작위로 추출하였다. 이 과정을 통해 두 클래스의 샘플 수를 균등하게 조정함으로써, 학습 과정에서 모델이 특정 클래스에 편향되지 않고 R-피크와 R-피크가 아닌 구간의 특징을 균형 있게 학습할 수 있도록 하였다.

3.3.2 데이터셋 분할: 학습 및 테스트 데이터셋 구성

클래스 균형화가 완료된 전체 샘플은 모델의 학습 및 일반화 성능 평가를 위해 학습 데이터셋(Training Dataset)과 테스트 데이터셋(Test Dataset)으로 분할되었다. 균형화 후 표본을 8:2로 분할했다. 테스트 표본은 학습/튜닝에 사용하지 않았고 레코드·구간 중복을 금지했다. 데이터 규모와 모델 단순성(SVM 선형, RF 100트리)을 고려해 8:2 비율을 적용하였다. 추가 교차검증 없이도 분산이 과도하지 않음을 제4장에서 확인했다.

3.4 R-피크 검출 모델 구현

본 연구에서는 ECG 신호의 R-피크를 효과적으로 검출하기 위해 두 가지 머신러닝 기반 분류 모델인 서포트 벡터 머신(Support Vector Machine, SVM)과 랜덤 포레스트(Random Forest, RF)를 활용하였다. 각 모델은 Python의 scikit-learn 라이브러리를 기반으로 구현되었으며, 그 세부적인 설정은 다음과 같다.

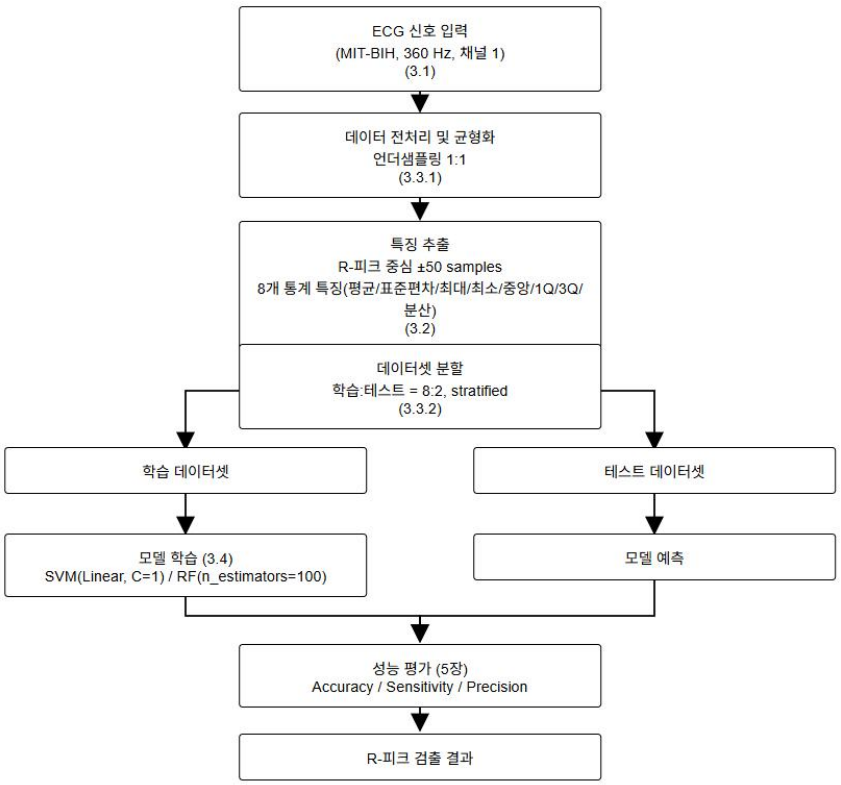
3.4.1 서포트 벡터 머신(SVM) 모델

SVM은 scikit-learn SVC로 구현했고, 선형 커널과 C=1.0, class_weight='balanced', random_state=42를 사용했다. 학습세트 통계로 z-score 표준화 후 동일 변환을 테스트에 적용했다. 본 연구에서는 'R-피크'와 'R-피크가 아닌 구간'을 분류하는 데 사용되었다. 모델 구현에는 sklearn.svm.SVC 클래스를 활용하였다. SVM의 핵심 구성 요소인 커널 함수(Kernel Function)로는 선형 커널(Linear Kernel)

을 채택하였다. 입력 차원이 8이고 특징 간 상관이 크지 않아 선형 결정경계로도 안정적인 분리가 가능하다고 판단해 선형 커널을 선택했다. 비선형(RBF) 대비 튜닝 부담이 작고 해석이 쉽다. 이는 모델의 복잡도를 낮추면서도, ECG 신호의 특징이 선형적으로 분리될 수 있는 특성을 가진다는 가설하에 선택되었다. 선형 커널의 사용은 모델의 해석 가능성을 높이고 계산 비용을 효율적으로 관리하는 데 이점을 제공한다.

3.4.2 랜덤 포레스트(Random Forest) 모델

RF는 트리 100개, 깊이 제한 없음, 특성 선택은 특성 수, 부트스트랩 사용, 난수는 42로 고정했다. 동일 특징(3.2)·동일 분할(3.3)로 SVM과 공정 비교했다. 본 연구에서는 sklearn.ensemble.RandomForestClassifier 클래스를 이용하여 모델을 구현하였다. 모델의 주요 파라미터로 100개의 개별 결정 트리를 사용하도록 설정하였다.



[그림 1] R-피크 검출 파이프라인(전처리·특징 추출·데이터 분할·모델 학습/예측·성능 평가; MIT-BIH 48 레코드)

[Fig. 1] End-to-End Pipeline for R-Peak Detection (Preprocessing, Feature Extraction, Data Split, Model Training / Inference, and Evaluation; MIT-BIH, 48 Records)

랜덤 포레스트는 각 트리를 학습 데이터의 무작위 부분집합과 특징의 무작위 부분집합을 사용

하여 독립적으로 학습시킨다. RF는 RandomForestClassifier(n_estimators=100, max_depth=None, max_features='sqrt', bootstrap=True, random_state=42)로 학습했다. 추가 가중치 없이도 3.3의 1:1 균형화로 클래스 편향을 억제했다. 이러한 앙상블 방식은 단일 결정 트리가 가질 수 있는 과적합(Overfitting) 문제를 효과적으로 줄이고, 모델의 안정성 및 예측 정확도를 전반적으로 향상시키는 데 기여한다.

본 연구의 전체 파이프라인을 도식화하여 전처리, 특징 추출, 데이터 분할, 학습 및 성능 평가 과정을 [그림 1]과 같이 정리하였다.

3.5 성능 평가 지표

각 지표는 수식 (1)~(3)에 따라 정의되며, 48개 레코드에 대해 산출된 평균값을 사용하였다.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$\mathrm{Precision} = \frac{TP}{TP + FP} \tag{3}$$

모든 지표는 레코드별 산출 후 평균을 취했다(n=48).

4. 연구 결과

총 48개 레코드에 대해 평균 성능을 산출한 결과는 다음과 같으며 Pan-Tompkins, SVM, Random Forest 세 기법의 평균 정확도, 민감도, 정밀도를 비교한 결과를 [표 3]과 같이 정리하였다.

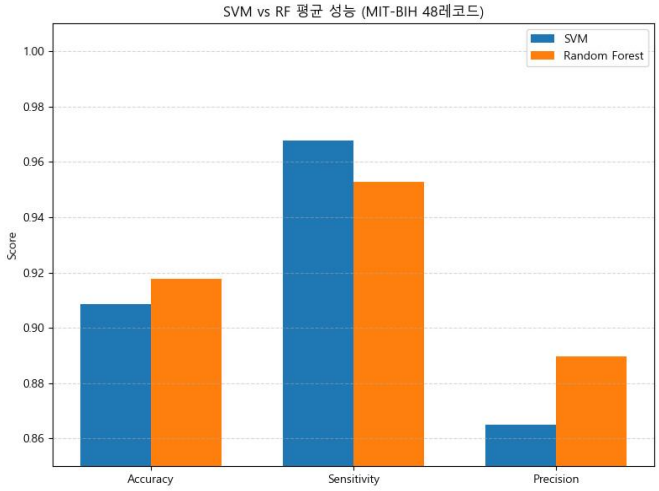
[표 3] Pan-Tompkins vs. 머신러닝 기반 R-피크 검출 성능 비교

[Table 3] Performance Comparison of Pan-Tompkins and ML-based R-peak Detection

모델 (Model)	정확도 (Accuracy)	민감도 (Sensitivity)	정밀도 (Precision)
Pan-Tompkins[3]	0.9017	0.9384	0.8653
SVM	0.9122	0.9707	0.8708
Random Forest	0.9214	0.9548	0.8961

Accuracy, Sensitivity, Precision 세 항목을 [그림 2]와 같이 정리하였다. 전체적으로 Random Forest가 모든 지표에서 우수한 성능을 보였으며, 특히 Precision에서 0.8961로 가장 높은 값을 기록하였

다.



[그림 2] SVM vs Random Forest 평균성능 비교(48레코드, MIT-BIH)

[Fig. 2] Average Performance Comparison by Model

5. 결론

이 연구는 전통적인 알고리즘과 머신러닝 기법이 심전도(ECG) 신호에서 R-피크를 검출하는 데 있어 어떠한 성능 차이를 보이는지를 실험적으로 분석하였다. Pan-Tompkins 알고리즘과 머신러닝 기반 분류기인 SVM(Support Vector Machine), Random Forest(RF)를 대상으로 MIT-BIH 부정맥 데이터베이스의 48개 전체 레코드를 활용하였으며, 정확도, 민감도, 정밀도 지표를 기준으로 비교하였다. 분석 결과, 두 머신러닝 기법 모두 기존 알고리즘보다 높은 성능을 보였으며, 특히 Random Forest는 데이터에 대한 일반화 측면에서 강점을 보이며 정확도와 정밀도 모두에서 우수한 결과를 나타냈다.

이러한 결과는 머신러닝 알고리즘이 ECG 신호 분석에서 실질적인 성능 향상을 기대할 수 있는 대안임을 시사한다. 기존의 Pan-Tompkins 알고리즘이 명확한 절차와 임곗값 조절을 통해 안정적인 성능을 유지해 온 반면, 머신러닝 기법은 데이터 특성에 맞춰 적응하는 구조적 유연성을 지닌다. 이러한 차이는 특히 다양한 환자 조건과 신호 환경이 공존하는 임상 현장에서 더욱 중요하게 작용할 수 있다.

본 연구는 이러한 분석을 바탕으로, 향후 다음과 같은 세 가지 확장 가능성을 제안한다. 첫째, 딥러닝 기반 알고리즘과의 성능 비교이다. CNN(Convolutional Neural Networks)이나 LSTM(Long Short-Term Memory)과 같은 신경망 기반 구조는 자동화된 특징 추출을 통해 보다 정밀한 분석을 가능하게 하며, 기존 머신러닝 기법과의 비교는 알고리즘 선택의 폭을 넓히는 데 기여할 수 있다.

둘째, 다양한 노이즈 환경에서의 알고리즘 성능 평가가 필요하다. 실제 임상에서는 근전도 간섭, 베이스라인 드리프트 등 다양한 신호 왜곡 요인이 존재하며, 알고리즘이 이러한 상황에서도 안정적인 검출 능력을 유지할 수 있는지가 핵심적인 평가 기준이 된다.

셋째, 실시간 적용 가능성에 대한 검토가 요구된다. 높은 정확도 외에도 빠른 처리속도, 낮은 지연시간, 시스템 자원 효율성 등은 실제 디바이스 구현 시 필수 고려 요소이며, 본 연구의 결과가 이러한 실용적 요건과 어떻게 조응할 수 있는지를 확인할 필요가 있다.

결론적으로, 본 연구는 Pan-Tompkins 알고리즘과 머신러닝 기법 간의 정량적 비교를 통해, ECG R-피크 검출에서의 방법론적 대안을 제시하였다. 특히 Random Forest는 기술적 성능뿐 아니라 데이터 환경에 대한 유연한 대응력 면에서 강점을 보였다. 알고리즘 선택은 단순한 정확도 수치에 그치지 않고, 신호 특성, 활용 목적, 시스템 환경을 종합적으로 고려한 판단이 필요하다는 점에서, 본 연구의 결과는 향후 임상 적용 가능성을 모색하는 데 중요한 자료가 될 것이다.

References

[1] Z. Zhang, Y. Zheng, “Deep learning for ECG analysis: Bench to bedside”, Circulation Research, vol. 122, no. 5, March 2018, pp. 614-636, doi: 10.1161/CIRCRESAHA.117.311336.

[2] M. H. Baek, J. G. Jeong, Y. J. Kim, K. G. Kim, J. Y. Jeon, “Clinical application of deep learning-based automatic spinal segmentation and Cobb angle measurement”, Journal of Next-generation Convergence Information Services Technology, vol. 9, no. 1, March 2020, pp. 1-9, doi: 10.29056/jncist.2020.03.01.

[3] M. N. Shim, “A study on the barriers and implementation model for realizing the right to be forgotten in the domestic environment”, Journal of Next-generation Convergence Information Services Technology, vol. 9, no. 1, March 2020, pp. 21-34, doi: 10.29056/jncist.2020.03.03.

[4] J. Pan, W. J. Tompkins, “A real-time QRS detection algorithm”, IEEE Transactions on Biomedical Engineering, vol. BME-32, no. 3, March 1985, pp. 230-236, doi: 10.1109/TBME.1985.325532.

[5] C. Cortes, V. Vapnik, “Support-vector networks”, Machine Learning, vol. 20, September 1995, pp. 273-297, doi: 10.1007/BF00994018.

[6] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.

[7] L. Breiman, “Random forests”, Machine Learning, vol. 45, October 2001, pp. 5-32, doi: 10.1023/A:1010933404324.

[8] G. B. Moody, R. G. Mark, “The impact of the MIT-BIH arrhythmia database”, IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, June 2001, pp. 45-50, doi: 10.1109/51.932724.

[9] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley, “PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals”, Circulation, vol. 101, no. 23, June 2000, pp. e215-e220, doi: 10.1161/01.CIR.101.23.e215.