



EVALUASI PERFORMA DETEKSI PUNCAK R PADA SINYAL EKG MULTI-KONDISI MENGUNAKAN ALGORITMA PAN-TOMPKINS: STUDI PADA MIT-BIH ARRHYTHMIA DATABASE

Ibrahim Tsabit Hanif Rabbiradliya¹, Fery Hadary², Redi Ratiandi Yacoub³

^{1,2,3}Program Studi Sarjana Teknik Elektro, Jurusan Teknik Elektro, Fakultas Teknik,
Universitas Tanjungpura

¹ibrahimtsabithr028@gmail.com, ²ferry.hadary@invent.untan.ac.id,

³rediyacoub@ee.untan.ac.id

Abstract

R-peak detection is a fundamental step in automatic electrocardiogram (ECG) signal analysis, yet its accuracy can be affected by variations in QRS complex morphology and signal disturbances under different clinical conditions. The Pan-Tompkins algorithm, as one of the most widely cited classical QRS detection methods, is generally evaluated under limited signal conditions, so a quantitative understanding of its performance across multiple morphological conditions simultaneously still needs to be strengthened. This study aims to evaluate and compare the performance of the Pan-Tompkins algorithm in detecting R-peaks under five different ECG signal conditions. The Pan-Tompkins algorithm was implemented and tested on five recordings from the MIT-BIH Arrhythmia Database representing Normal (record 100), Left Bundle Branch Block/LBBB (record 109), Right Bundle Branch Block/RBBB (record 118), Premature Ventricular Contraction/PVC (record 119), and baseline drift (record 222) conditions, with a sampling frequency of 360 Hz. Detection results were matched against ground truth annotations using a tolerance of ± 150 ms, then evaluated through sensitivity (Se), positive predictivity (+P), accuracy (Acc), and F1-Score. The highest performance was obtained under the Normal condition (Se 99.96%; +P 100.00%; F1 99.98%) and LBBB (F1 99.84%), followed by RBBB (F1 99.41%). Performance declined under PVC conditions (Se 79.87%; +P 79.83%; Acc 66.46%; F1 79.85%) and baseline drift (F1 89.71%). The average performance across the five conditions was Se 93.73%, +P 93.79%, Acc 89.25%, and F1 93.76%. The performance of the Pan-Tompkins algorithm was not uniform across all tested ECG signal conditions; performance was very good under relatively consistent QRS morphology but declined significantly under PVC and baseline drift conditions, indicating that multi-condition evaluation is necessary to understand the practical limits of this algorithm's use more realistically.

Keywords: Pan-Tompkins algorithm; R-peak detection; ECG; MIT-BIH Arrhythmia Database; performance evaluation

Article History

Received: September 2026

Reviewed: September 2026

Published: September 2026

Plagiarism Checker No 235

Prefix DOI :

[10.8734/Kohesi.v1i2.365](https://doi.org/10.8734/Kohesi.v1i2.365)

Copyright : Author

Publish by : Kohesi



This work is licensed under
a [Creative Commons
Attribution-
NonCommercial 4.0
International License](https://creativecommons.org/licenses/by-nc/4.0/)



Abstrak

Deteksi puncak R merupakan langkah dasar dalam analisis otomatis sinyal elektrokardiogram (EKG), namun akurasinya dapat dipengaruhi oleh variasi morfologi kompleks QRS dan gangguan sinyal pada kondisi klinis yang berbeda. Algoritma Pan-Tompkins, sebagai salah satu metode deteksi QRS klasik yang paling banyak dirujuk, umumnya dievaluasi pada kondisi sinyal yang terbatas, sehingga pemahaman kuantitatif mengenai performanya pada beberapa kondisi morfologi sekaligus masih perlu diperkuat. Penelitian ini bertujuan mengevaluasi dan membandingkan performa algoritma Pan-Tompkins dalam mendeteksi puncak R pada lima kondisi sinyal EKG yang berbeda. Algoritma Pan-Tompkins diimplementasikan dan diuji pada lima rekaman MIT-BIH Arrhythmia Database yang mewakili kondisi Normal (record 100), Left Bundle Branch Block/LBBB (record 109), Right Bundle Branch Block/RBBB (record 118), Premature Ventricular Contraction/PVC (record 119), dan baseline drift (record 222), dengan frekuensi sampling 360 Hz. Hasil deteksi dicocokkan dengan anotasi ground truth menggunakan toleransi ± 150 ms, kemudian dievaluasi melalui sensitivitas (Se), positive predictivity (+P), akurasi (Acc), dan F1-Score. Performa tertinggi diperoleh pada kondisi Normal (Se 99,96%; +P 100,00%; F1 99,98%) dan LBBB (F1 99,84%), diikuti RBBB (F1 99,41%). Performa menurun pada kondisi PVC (Se 79,87%; +P 79,83%; Acc 66,46%; F1 79,85%) dan baseline drift (F1 89,71%). Rata-rata performa pada kelima kondisi adalah Se 93,73%, +P 93,79%, Acc 89,25%, dan F1 93,76%. Performa algoritma Pan-Tompkins tidak seragam pada seluruh kondisi sinyal EKG yang diuji; performa sangat baik pada morfologi QRS yang relatif konsisten, tetapi menurun signifikan pada kondisi PVC dan baseline drift, sehingga evaluasi multi-kondisi diperlukan untuk memahami batas penggunaan algoritma ini secara lebih realistis.

Kata kunci: Algoritma Pan-Tompkins; deteksi puncak-R; EKG; Basis Data Aritmia MIT-BIH; evaluasi kinerja

1. PENDAHULUAN

Deteksi kompleks QRS, khususnya posisi puncak R, merupakan tahap fundamental dalam analisis sinyal elektrokardiogram (EKG) karena menjadi dasar perhitungan laju denyut jantung serta berbagai parameter turunan seperti heart rate variability dan deteksi aritmia. Akurasi deteksi puncak R menjadi tantangan tersendiri ketika sinyal EKG menunjukkan variasi morfologi kompleks QRS – misalnya akibat gangguan konduksi, denyut ektopik – maupun gangguan non-fisiologis seperti pergeseran garis dasar (baseline drift) yang umum dijumpai pada perekaman jangka panjang.

Algoritma Pan-Tompkins, yang diperkenalkan oleh Pan dan Tompkins pada tahun 1985, merupakan salah satu metode deteksi QRS klasik yang paling luas diadopsi hingga saat ini [1]. Secara umum, algoritma ini bekerja melalui rangkaian tahap pemrosesan sinyal digital penapisan pita frekuensi, penguatan kemiringan sinyal melalui operasi turunan, penguatan



amplitudo melalui operasi kuadrat, penghalusan melalui integrasi jendela bergerak, dan penentuan puncak melalui mekanisme ambang adaptif tanpa memerlukan proses pelatihan model seperti pada pendekatan berbasis pembelajaran mesin.

Berbagai penelitian telah mengevaluasi performa algoritma Pan-Tompkins maupun variannya pada basis data EKG standar. Sebagian menunjukkan bahwa performa deteksi cenderung menurun pada sinyal yang mengandung derau atau gangguan tertentu [3], sementara penelitian lain mengembangkan modifikasi ambang adaptif untuk meningkatkan ketahanan deteksi pada kondisi yang lebih menantang [4], [9]. Studi benchmarking pada beberapa detektor QRS juga menunjukkan bahwa performa suatu algoritma dapat bervariasi cukup besar tergantung pada karakteristik morfologi sinyal yang diuji, bukan bersifat tetap untuk seluruh kondisi klinis [6], [12].

Meskipun demikian, laporan performa algoritma Pan-Tompkins pada beberapa kondisi morfologi EKG yang berbeda secara berdampingan menggunakan parameter frekuensi sampling dan toleransi evaluasi yang konsisten masih belum banyak disajikan secara eksplisit dan terstruktur. Kesenjangan ini menjadi relevan mengingat pemilihan algoritma deteksi QRS untuk suatu aplikasi seyogianya mempertimbangkan performanya pada rentang kondisi sinyal yang akan dihadapi, bukan hanya pada kondisi ideal.

Berdasarkan kesenjangan tersebut, artikel ini bertujuan menyajikan evaluasi kuantitatif performa deteksi puncak R menggunakan algoritma Pan-Tompkins pada lima kondisi sinyal EKG dari MIT-BIH Arrhythmia Database, yaitu Normal, LBBB, RBBB, PVC, dan baseline drift, dengan membandingkan nilai sensitivitas, positive predictivity, akurasi, dan F1-Score antar kondisi. Kontribusi artikel ini adalah penyajian perbandingan performa yang terukur dan dapat direproduksi pada kelima kondisi tersebut, sebagai acuan awal mengenai batas kemampuan algoritma Pan-Tompkins pada variasi morfologi sinyal EKG yang berbeda.

2. METODE PENELITIAN

2.1 Dataset

Penelitian ini menggunakan lima rekaman dari MIT-BIH Arrhythmia Database [2], [14], yaitu record 100 (kondisi Normal), record 109 (LBBB), record 118 (RBBB), record 119 (PVC), dan record 222 (baseline drift). Seluruh rekaman memiliki frekuensi sampling 360 Hz. Anotasi ground truth yang digunakan adalah anotasi bawaan basis data (file .atr), yang dihasilkan melalui proses anotasi manual oleh dua atau lebih kardiolog secara independen dan diterima luas sebagai acuan standar dalam literatur evaluasi algoritma deteksi QRS [2], [11], [14]. Tabel 1 merangkum jumlah anotasi ground truth pada tiap rekaman yang digunakan sebagai dasar evaluasi.

Tabel 1. Jumlah anotasi ground truth per record

Record	Kondisi	Fs (Hz)	Jumlah Anotasi Ground Truth
100	Normal	360	2.273
109	LBBB	360	2.532
118	RBBB	360	2.278
119	PVC	360	1.987
222	Baseline Drift	360	2.483

2.2 Implementasi Algoritma

Algoritma Pan-Tompkins diimplementasikan mengikuti spesifikasi tahap pemrosesan pada publikasi aslinya [1], yang terdiri atas enam tahap berurutan: (1) low-pass filter dan (2) high-pass filter yang bersama-sama membentuk bandpass pada rentang frekuensi karakteristik kompleks QRS; (3) derivative filter untuk menonjolkan kemiringan sinyal; (4) squaring untuk



menguatkan nilai amplitudo besar; (5) moving window integration untuk menghaluskan dan melebarkan puncak hasil deteksi awal; serta (6) adaptive thresholding untuk menentukan kandidat puncak R secara final. Parameter berbasis waktu pada tiap tahap disesuaikan terhadap frekuensi sampling basis data (360 Hz) tanpa mengubah spesifikasi rancangan filter aslinya. Detail persamaan dan koefisien tiap tahap tidak diuraikan ulang secara rinci pada artikel ini karena bukan menjadi fokus evaluasi; penjelasan pemodelan dan simulasi tahap-demi-tahap algoritma disajikan pada publikasi terpisah dari penelitian yang sama

2.3 Deteksi Puncak-R dan Pencocokan Ground Truth

Keluaran algoritma berupa daftar posisi (indeks sampel) puncak R yang terdeteksi pada tiap rekaman. Untuk mengevaluasi keakuratan deteksi, setiap posisi hasil deteksi dicocokkan terhadap posisi anotasi ground truth menggunakan skema pencocokan satu-ke-satu dengan toleransi waktu ± 150 ms, mengikuti konvensi yang umum digunakan dalam evaluasi algoritma deteksi QRS pada basis data MIT-BIH [1], [2], [11]. Proses pencocokan dilakukan secara berurutan terhadap setiap deteksi: untuk tiap posisi deteksi, dicari anotasi ground truth dengan jarak absolut terdekat; apabila jarak tersebut berada dalam batas toleransi (150 ms, dikonversi ke satuan sampel berdasarkan $F_s = 360$ Hz) dan anotasi tersebut belum dipasangkan dengan deteksi lain sebelumnya, pasangan tersebut dihitung sebagai True Positive (TP). Deteksi yang tidak memperoleh pasangan dalam batas toleransi dihitung sebagai False Positive (FP), sedangkan anotasi ground truth yang tidak memperoleh pasangan dihitung sebagai False Negative (FN). Skema pencocokan ini bersifat satu-ke-satu (one-to-one): satu anotasi hanya dapat dipasangkan dengan maksimal satu deteksi.

2.4 Perhitungan Confusion Matrix

Performa deteksi dievaluasi menggunakan empat parameter standar yang dihitung dari hasil TP, FP, dan FN pada tiap record:

$$\text{Sensitivitas: } Se = \frac{TP}{TP+FN} \times 100\%$$

$$\text{Positive Predictivity: } +P = \frac{TP}{TP+FP} \times 100\%$$

$$\text{Akurasi: } Acc = \frac{TP}{TP+FP+FN} \times 100\%$$

$$\text{F1-Score: } F1 = \frac{2 \times Se \times +P}{Se + +P} \times 100\%$$

Keempat parameter ini dipilih karena merupakan parameter standar yang umum digunakan dalam literatur evaluasi algoritma deteksi QRS berbasis MIT-BIH Arrhythmia Database [1], [2], [11], [12], sehingga memungkinkan perbandingan yang konsisten antar kondisi sinyal dalam penelitian ini.

3. HASIL DAN PEMBAHASAN

Pada bagian ini disajikan hasil evaluasi performa deteksi puncak R menggunakan algoritma Pan-Tompkins pada lima kondisi sinyal EKG yang diuji, yaitu Normal, Left Bundle Branch Block (LBBB), Right Bundle Branch Block (RBBB), Premature Ventricular Contraction (PVC), dan baseline drift. Hasil evaluasi meliputi jumlah true positive (TP), false positive (FP), dan false negative (FN), serta nilai sensitivity (Se), positive predictivity (+P), accuracy (Acc), dan F1-Score pada masing-masing kondisi. Selanjutnya, performa algoritma dibandingkan antar kondisi untuk melihat perbedaan kemampuan deteksi pada variasi morfologi dan karakteristik sinyal EKG. Pembahasan difokuskan pada kondisi Normal, LBBB, RBBB, PVC, dan baseline drift, serta penurunan performa yang terlihat pada kondisi PVC dan baseline drift berdasarkan hasil evaluasi yang diperoleh.



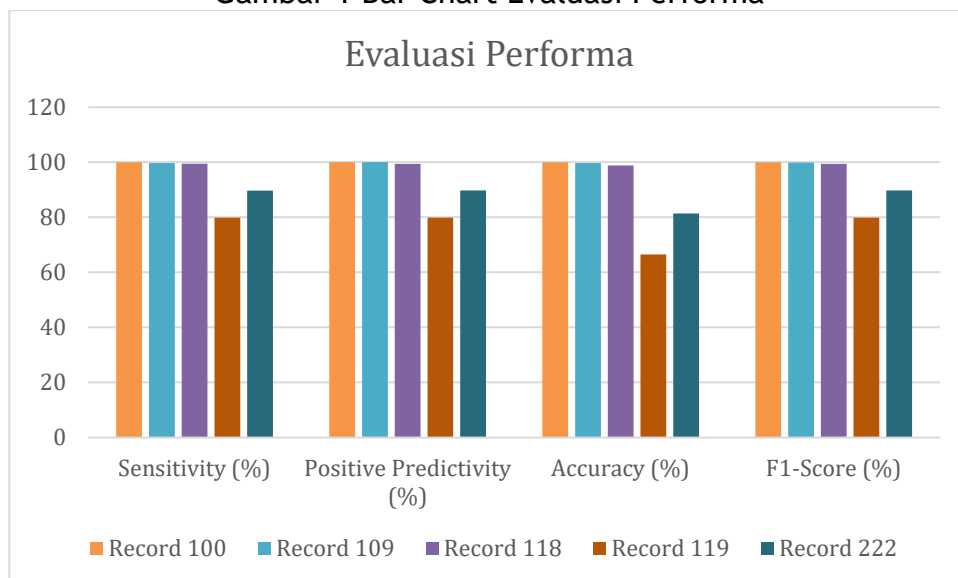
3.1 Kinerja Deteksi

Tabel 2 menyajikan hasil evaluasi performa deteksi puncak R pada kelima kondisi sinyal EKG yang diuji.

Tabel 2. Hasil evaluasi performa deteksi puncak R algoritma Pan-Tompkins

Record	Kondisi	Total Anotasi	TP	FP	FN	Sensitivity (%)	Positive Predictivity (%)	Accuracy (%)	F1-Score (%)
100	Normal	2.273	2.272	0	1	99,96	100,00	99,96	99,98
109	LBBB	2.532	2.524	0	8	99,68	100,00	99,68	99,84
118	RBBB	2.278	2.265	14	13	99,43	99,39	98,82	99,41
119	PVC	1.987	1.587	401	400	79,87	79,83	66,46	79,85
222	Baseline Drift	2.483	2.227	255	256	89,69	89,73	81,34	89,71
Rata-rata						93,73	93,79	89,25	93,76

Gambar 1 Bar Chart Evaluasi Performa



Secara umum, hasil pada Tabel 2 menunjukkan bahwa performa deteksi puncak R algoritma Pan-Tompkins tidak seragam di seluruh kondisi sinyal yang diuji, dengan rentang F1-Score dari 79,85% (PVC) hingga 99,98% (Normal).

3.2 Perbandingan Antar kondisi EKG

Kondisi Normal (record 100) menunjukkan performa tertinggi di antara seluruh kondisi yang diuji, dengan +P mencapai 100,00% karena tidak ditemukan satu pun deteksi palsu (FP = 0), dan Se 99,96% karena hanya satu dari 2.273 anotasi yang tidak terdeteksi (FN = 1). Hasil ini mengindikasikan bahwa algoritma bekerja hampir optimal pada sinyal dengan morfologi QRS yang relatif ideal.

Kondisi LBBB (record 109) menunjukkan performa yang tetap sangat tinggi (Se 99,68%; +P 100,00%; F1 99,84%), dengan FP = 0 dan FN yang sedikit lebih besar (FN = 8) dibandingkan kondisi Normal. Berdasarkan hasil yang tersedia, perubahan morfologi QRS akibat LBBB tidak menunjukkan penurunan performa yang besar terhadap algoritma ini, meskipun jumlah FN yang sedikit lebih tinggi menunjukkan bahwa morfologi LBBB tetap dapat memengaruhi beberapa deteksi secara terbatas.



Kondisi RBBB (record 118) menunjukkan performa yang sedikit lebih rendah dibandingkan Normal dan LBBB (Se 99,43%; +P 99,39%; F1 99,41%), dengan munculnya FP (14) di samping FN (13) berbeda dari dua kondisi sebelumnya yang tidak memiliki FP sama sekali. Hal ini mengindikasikan bahwa morfologi RBBB berkaitan dengan sedikit peningkatan kesalahan deteksi dibandingkan kondisi Normal dan LBBB, meskipun performanya secara keseluruhan tetap tergolong sangat baik (F1 di atas 99%). Karakteristik lebih rinci mengenai pola kesalahan pada kondisi ini (termasuk kemungkinan pergeseran presisi temporal) berada di luar cakupan evaluasi performa pada artikel ini dan dibahas pada studi lanjutan yang terpisah.

Kondisi PVC (record 119) menunjukkan penurunan performa paling besar di antara seluruh kondisi yang diuji, dengan Se dan +P yang sama-sama berada pada kisaran 80% (Se 79,87%; +P 79,83%) dan akurasi turun ke 66,46%. Jumlah FP (401) dan FN (400) meningkat secara signifikan dibandingkan kondisi lainnya. Hasil ini menunjukkan bahwa karakteristik morfologi PVC dapat menjadi kondisi yang menantang bagi algoritma Pan-Tompkins berdasarkan hasil pengujian ini, meskipun mekanisme penyebab spesifik di balik penurunan performa tersebut misalnya karakteristik distribusi kesalahan secara temporal tidak diuji secara mendalam pada artikel ini dan menjadi fokus kajian terpisah.

Kondisi Baseline Drift (record 222) juga menunjukkan performa di bawah kondisi Normal, LBBB, dan RBBB, dengan Se 89,69%, +P 89,73%, dan F1 89,71%, meskipun tidak serendah kondisi PVC. Jumlah FP (255) dan FN (256) pada kondisi ini juga lebih tinggi dibandingkan tiga kondisi pertama. Hasil ini mengindikasikan bahwa gangguan garis dasar sinyal dapat berkaitan dengan penurunan performa deteksi, walaupun besaran penurunannya berdasarkan hasil yang tersedia tergolong lebih moderat dibandingkan kondisi PVC.

3.3 Penurunan Kinerja dalam Kondisi Menantang

Secara keseluruhan, dua kondisi yang menunjukkan penurunan performa paling nyata pada evaluasi ini adalah PVC dan *baseline drift*, dengan F1-Score masing-masing 79,85% dan 89,71%, jauh di bawah rata-rata keseluruhan (93,76%) maupun performa pada tiga kondisi lainnya (di atas 99%). Berdasarkan hasil pengujian ini, kedua kondisi tersebut dapat dianggap sebagai kondisi yang relatif membebani kemampuan deteksi algoritma Pan-Tompkins dibandingkan kondisi Normal, LBBB, dan RBBB.

Perlu ditekankan bahwa hasil ini tidak dapat diinterpretasikan sebagai bukti bahwa algoritma "gagal total" pada kedua kondisi tersebut, karena algoritma tetap berhasil mendeteksi sebagian besar anotasi *ground truth* pada kedua record (TP = 1.587 dari 1.987 pada PVC; TP = 2.227 dari 2.483 pada *baseline drift*). Penurunan performa yang teramati lebih tepat diinterpretasikan sebagai indikasi bahwa mekanisme *adaptive thresholding* pada algoritma ini memiliki keterbatasan tertentu ketika berhadapan dengan morfologi denyut yang menyimpang (PVC) atau gangguan amplitudo garis dasar (*baseline drift*), tanpa dapat dipastikan mekanisme penyebab spesifiknya hanya dari data performa agregat ini. Analisis lebih lanjut mengenai karakteristik kesalahan tersebut secara temporal termasuk pola distribusi kesalahan dan ketahanan algoritma dari waktu ke waktu disajikan sebagai studi lanjutan yang terpisah dari artikel ini.

KESIMPULAN

Penelitian ini menyajikan evaluasi kuantitatif performa deteksi puncak R menggunakan algoritma Pan-Tompkins pada lima kondisi sinyal EKG dari MIT-BIH Arrhythmia Database. Hasil evaluasi menunjukkan bahwa performa algoritma berbeda secara nyata antar kondisi sinyal yang diuji. Performa sangat tinggi diperoleh pada kondisi Normal (F1 99,98%) dan LBBB (F1 99,84%), menunjukkan bahwa algoritma bekerja sangat baik pada morfologi QRS yang relatif konsisten. Sedikit penurunan performa teramati pada kondisi RBBB (F1 99,41%), sementara penurunan performa paling besar terjadi pada kondisi PVC (F1 79,85%). Kondisi baseline drift juga menunjukkan penurunan performa (F1 89,71%) meskipun tidak seberat kondisi PVC. Rata-



rata performa pada kelima kondisi adalah Se 93,73%, +P 93,79%, Acc 89,25%, dan F1 93,76%. Temuan ini menegaskan bahwa performa algoritma Pan-Tompkins tidak dapat diasumsikan seragam pada seluruh kondisi sinyal EKG, sehingga evaluasi pada beberapa kondisi morfologi sekaligus penting untuk memahami batas penggunaan algoritma ini secara lebih realistis dalam aplikasi praktis. Artikel ini tidak mengusulkan algoritma atau metode baru, melainkan menyajikan evaluasi kuantitatif terhadap algoritma Pan-Tompkins yang sudah ada pada data yang digunakan; karakteristik kesalahan dan ketahanan algoritma secara lebih mendalam dibahas pada studi lanjutan yang terpisah.

DAFTAR PUSTAKA

- [1] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230-236, Mar. 1985, [doi: 10.1109/TBME.1985.325532](https://doi.org/10.1109/TBME.1985.325532).
- [2] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45-50, May/Jun. 2001, [doi: 10.1109/51.932724](https://doi.org/10.1109/51.932724).
- [3] M. A. Z. Fariha, R. Ikeura, S. Hayakawa, and S. Tsutsumi, "Analysis of Pan-Tompkins algorithm performance with noisy ECG signals," *J. Phys.: Conf. Ser.*, vol. 1532, no. 1, p. 012022, 2020, [doi: 10.1088/1742-6596/1532/1/012022](https://doi.org/10.1088/1742-6596/1532/1/012022).
- [4] N. Khan and M. N. Imtiaz, "Pan-Tompkins++: A robust approach to detect R-peaks in ECG signals," *arXiv:2211.03171*, 2022.
- [5] A. N. Vest et al., "Algorithm for mobile platform-based real-time QRS detection," *Sensors*, vol. 23, no. 3, p. 1625, 2023, [doi: 10.3390/s23031625](https://doi.org/10.3390/s23031625).
- [6] M. A. F. Pimentel et al., "QRS detection in single-lead telehealth ECG signals: Benchmarking open-source algorithms," *PLOS Digital Health*, vol. 2, no. 9, p. e0000538, 2023, [doi: 10.1371/journal.pdig.0000538](https://doi.org/10.1371/journal.pdig.0000538).
- [9] Anas Fouad Ahmed, "A novel improved Pan-Tompkins method for QRS complex detection," *Signal, Image and Video Processing*, 2025, [doi: 10.1007/s11760-025-04587-6](https://doi.org/10.1007/s11760-025-04587-6).
- [11] 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, no. 12, pp. 1157-1165, Dec. 1986, [doi: 10.1109/TBME.1986.325695](https://doi.org/10.1109/TBME.1986.325695).
- [12] F. Liu et al., "Performance analysis of ten common QRS detectors on different ECG application cases," *J. Healthc. Eng.*, 2018.
- [14] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215-e220, Jun. 2000, [doi: 10.1161/01.CIR.101.23.e215](https://doi.org/10.1161/01.CIR.101.23.e215).