

Pendeteksian Perubahan Sinyal Dinamis ECG pada Aritmia Jantung Menggunakan Analisis Rényi Entropi dan XGBoost

* Catatan: Persyaratan Tugas Akhir S1 Teknik Biomedis

1st Fadlililla Kusuma Widi
Penulis-Teknik Biomedis
Telkom University,
Bandung Indonesia

fadliwidi@student.telkomuniversity.ac.id

2nd Tito Waluyo Purboyo
Pembimbing 1-Teknik Biomedis
Telkom University,
Bandung Indonesia

titowaluyo@telkomuniversity.ac.id

3rd Annisa Humairani
Pembimbing 2-Teknik Biomedis
Telkom University,
Bandung Indonesia

annisahumairani@telkomuniversity.ac.id

Abstrak— Deteksi aritmia jantung secara akurat dan cepat sangat penting untuk mencegah komplikasi kardiovaskular yang fatal. Penelitian ini mengusulkan metode deteksi berbasis kombinasi Discrete Wavelet Transform (DWT) multi-level, analisis Rényi Entropy, dan algoritma XGBoost. Dataset yang digunakan adalah MIT-BIH Arrhythmia Database dengan dua kanal utama, MLI dan VI. Tahapan penelitian mencakup filtering sinyal untuk peningkatan Signal-to-Noise Ratio (SNR), segmentasi beat presisi berbasis deteksi R-peak, pelabelan menggunakan anotasi MIT-BIH (.atr), serta ekstraksi fitur Rényi Entropy pada level DWT 1 hingga 4. Untuk mengatasi ketidakseimbangan kelas antara beat normal dan aritmia, dilakukan undersampling moderat dengan rasio 2:1 pada data latih. Hasil pengujian menunjukkan bahwa model XGBoost dengan balancing data mencapai akurasi 88,33% dan ROC AUC 0,9410, sedangkan tanpa balancing menghasilkan akurasi 88,61% dan ROC AUC 0,9365. Analisis SHAP mengungkap bahwa fitur entropy pada level DWT rendah memiliki kontribusi signifikan dalam klasifikasi. Peningkatan nilai SNR pasca-filtering menunjukkan keberhasilan tahap preprocessing dalam mengurangi noise dan mempertahankan morfologi PQRST. Metode yang diusulkan memberikan kinerja tinggi, sensitivitas yang baik terhadap kelas minoritas, dan potensi implementasi pada sistem pemantauan ECG waktu nyata.

Kata Kunci— ECG, Aritmia, Rényi Entropy, Discrete Wavelet Transform, XGBoost, Signal-to-Noise Ratio.

I. PENDAHULUAN

Elektrokardiogram (ECG) merupakan sinyal biomedis yang merekam aktivitas listrik jantung dan berperan penting dalam diagnosis gangguan irama jantung atau aritmia. Aritmia dapat menyebabkan komplikasi serius seperti fibrilasi atrium, takikardia ventrikel, hingga kematian mendadak apabila tidak terdeteksi secara dini. Oleh karena itu, metode analisis ECG yang cepat, akurat, dan efisien menjadi kebutuhan mendesak dalam sistem pemantauan kesehatan modern, baik di rumah sakit maupun perangkat wearable.

Namun, sinyal ECG sering kali mengandung derau (noise) akibat interferensi baseline, artefak gerakan, maupun gangguan elektromagnetik, yang dapat menurunkan akurasi deteksi. Selain itu, bentuk gelombang ECG yang bervariasi antar individu dan ketidakseimbangan jumlah data antara beat normal dan aritmia menjadi tantangan tersendiri dalam pengembangan model klasifikasi yang andal.

Berbagai studi telah mengusulkan pendekatan berbasis transformasi waktu-frekuensi, salah satunya Discrete Wavelet Transform (DWT), untuk mengekstraksi ciri dari berbagai skala frekuensi sinyal ECG. Analisis entropi, seperti Shannon Entropy dan Rényi Entropy, digunakan untuk menangkap kompleksitas nonlinier sinyal, di mana Rényi Entropy dengan parameter $\alpha < 1$ memiliki sensitivitas lebih tinggi terhadap pola langka seperti beat aritmia.

Penelitian ini mengusulkan metode deteksi aritmia yang memadukan DWT multi-level, ekstraksi fitur Rényi Entropy, dan algoritma klasifikasi XGBoost. Dataset yang digunakan adalah MIT-BIH Arrhythmia Database dengan dua kanal utama (MLI dan VI). Tahapan preprocessing meliputi filtering sinyal untuk meningkatkan Signal-to-Noise Ratio (SNR), segmentasi beat berbasis deteksi R-peak, pelabelan menggunakan anotasi (.atr), serta balancing data dengan undersampling moderat (rasio 2:1) untuk mengurangi bias terhadap kelas mayoritas.

II. KAJIAN TEORI

A. Dataset

Dataset yang digunakan dalam penelitian ini adalah MIT-BIH Arrhythmia Database yang tersedia secara terbuka di PhysioNet. Dataset ini merupakan salah satu sumber data standar internasional untuk penelitian analisis sinyal ECG, terdiri dari 48 rekaman pasien yang direkam di Beth Israel

Hospital, Boston, pada tahun 1975–1979. Setiap rekaman berdurasi 30 menit, diambil dengan sampling rate 360 Hz, resolusi 11-bit, dan amplitudo dikalibrasi sebesar 200 adu/mV.

Setiap rekaman memiliki dua kanal utama, yaitu:

- MLII (*Modified Limb Lead II*) — digunakan untuk memantau aktivitas listrik jantung dari sudut pandang inferior, yang memberikan representasi yang jelas terhadap gelombang P, QRS, dan T.
- V1 (*Lead V1*)— digunakan untuk memantau aktivitas listrik jantung dari sudut pandang anterior, yang membantu mendeteksi kelainan konduksi intraventrikular.

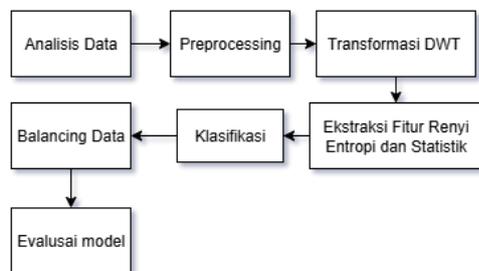
Selain data sinyal dalam format .dat dan header .hea, tersedia juga berkas anotasi .atr yang berisi posisi waktu setiap detak jantung (beat) dan kode simbol yang mewakili jenis beat (misalnya N, L, R, V, A). Pada penelitian ini, proses *labeling* mengacu pada standar AAMI (*Association for the Advancement of Medical Instrumentation*), di mana jenis beat dikelompokkan ke dalam dua kelas utama:

- Normal beat (kelas 0): meliputi simbol N, L, dan R yang merepresentasikan detak normal serta variasinya.
- Aritmia (kelas 1): meliputi PVC, APC, LBBB, RBBB, dan jenis beat lainnya yang tidak dikategorikan sebagai normal.

Hasil *labeling* menunjukkan distribusi kelas yang tidak seimbang, yaitu 12.353 beat normal dan 6.177 beat aritmia. Distribusi ini menegaskan adanya fenomena *class imbalance* yang memerlukan strategi balancing data pada tahap pelatihan model agar performa deteksi aritmia tetap optimal [6], [7].

B. Metode

Metode yang diusulkan pada penelitian ini melibatkan serangkaian tahap mulai dari akuisisi data hingga evaluasi model klasifikasi. Alur proses ditunjukkan pada bagan 1 (diagram alir). Tahapan utama yang dilakukan adalah sebagai berikut:



Bagan 1
Framework penelitian

Bagian metode penelitian ini diawali dengan analisis data, yaitu peninjauan terhadap struktur dan karakteristik dataset MIT-BIH *Arrhythmia Database* [8] yang diunduh dari PhysioNet. Tahap ini mencakup identifikasi kanal sinyal ECG yang digunakan, yaitu MLII dan V1, format file (.dat, .hea, .atr), jumlah total beat, distribusi kelas normal dan aritmia, serta pemetaan anotasi berdasarkan standar AAMI. Hasil analisis ini menjadi acuan dalam menentukan strategi preprocessing dan balancing data. Selanjutnya dilakukan preprocessing berupa penyaringan sinyal ECG menggunakan *bandpass filter* 0,5–50 Hz untuk menghilangkan baseline wander dan high-frequency noise, deteksi R-peak berbasis anotasi .atr, segmentasi sinyal per-beat dengan jendela ± 200 ms dari puncak R, serta normalisasi menggunakan metode z-score. Tahap berikutnya adalah transformasi DWT dan ekstraksi fitur *Rényi Entropy*, di mana setiap beat didekomposisi hingga level 4 menggunakan *Discrete Wavelet Transform* (DWT) dengan mother wavelet Daubechies-6 (db6), kemudian dari setiap level koefisien dihitung nilai *Rényi Entropy* dengan parameter $\alpha = 0,5$ untuk meningkatkan sensitivitas terhadap pola langka pada beat aritmia [15][16]. Proses balancing data dilakukan menggunakan teknik undersampling moderat dengan rasio 2:1 pada kelas normal untuk mengurangi ketidakseimbangan kelas tanpa kehilangan variasi morfologi beat normal secara signifikan [22]. Setelah itu, proses klasifikasi menggunakan algoritma *Extreme Gradient Boosting* (XGBoost) [23][29] dengan hyperparameter terpilih ($\text{max_depth} = 10$, $\text{learning_rate} = 0.2$, $\text{n_estimators} = 200$, $\text{sub_sample} = 0.6$, $\text{colsample_bytree} = 0.85$, $\text{use_label_encoder} = \text{False}$, $\text{random_state} = 42$, $\text{eval_metric} = \text{"logloss"}$) yang dioptimalkan untuk mendeteksi aritmia dengan akurasi tinggi. Terakhir, evaluasi model dilakukan menggunakan *confusion matrix*, *classification report*, akurasi, ROC-AUC, log loss, serta analisis interpretabilitas fitur menggunakan metode SHAP [30] untuk mengetahui kontribusi masing-masing fitur terhadap hasil prediksi model.

$$a + b = \gamma \quad (1)$$

C. Discrete Wavelet Transform

Discrete Wavelet Transform (DWT) merupakan salah satu teknik transformasi sinyal yang memproyeksikan data ke dalam domain waktu–frekuensi melalui dekomposisi bertingkat (*multi-resolution analysis*). Keunggulan DWT dibandingkan transformasi klasik seperti *Fast Fourier Transform* (FFT) atau *Short-Time Fourier Transform* (STFT) terletak pada kemampuannya dalam mempertahankan informasi lokal sinyal

di kedua domain, sehingga sangat efektif dalam menganalisis sinyal nonstasioner seperti elektrokardiogram (ECG) [9][24].

Secara umum, proses DWT melibatkan dua jenis filter utama, yaitu *low-pass filter* untuk menghasilkan koefisien aproksimasi (*approximation coefficients*) yang merepresentasikan komponen frekuensi rendah, dan *high-pass filter* untuk menghasilkan koefisien detail (*detail coefficients*) yang memuat informasi frekuensi tinggi [9]. Pada setiap tahap dekomposisi, sinyal dilewatkan melalui kedua filter ini, kemudian dilakukan *downsampling* dengan faktor dua untuk mengurangi jumlah sampel. Proses ini dapat diulang hingga mencapai level tertentu sesuai kebutuhan analisis.

Secara matematis, DWT untuk sinyal diskrit $x[n]$ dapat dinyatakan sebagai:

$$A_j(k) = \sum_n x(n) \cdot \phi_{j,k}(n) \quad (1)$$

$$D_j(k) = \sum_n x(n) \cdot \psi_{j,k}(n) \quad (2)$$

dengan $A_j(k)$ adalah koefisien aproksimasi pada level ke- j , $D_j(k)$ adalah koefisien detail, $\phi_{j,k}(n)$ merupakan fungsi skala (*scaling function*), dan $\psi_{j,k}(n)$ adalah fungsi wavelet (*wavelet function*) yang telah didilatasi dan ditranslasi [9].

Dalam konteks analisis ECG, komponen frekuensi tinggi pada level awal dekomposisi DWT (misalnya Level 1 dan Level 2) sering berkaitan dengan *QRS complex*, sedangkan komponen frekuensi rendah pada level lebih tinggi (misalnya Level 4 atau lebih) memuat informasi gelombang P dan T yang lebih lambat [25][26]. Oleh karena itu, DWT banyak digunakan untuk *denoising* [25], deteksi R-peak [21][26], dan ekstraksi fitur morfologi sinyal ECG.

Tabel 1
Wavelet

Nama Wavelet	Karakteristik	Kelebihan dalam Aplikasi ECG
Daubechies (db4, db6, db8)	Orthogonal, smooth, compact	Sangat cocok untuk deteksi QRS, karena bentuknya menyerupai gelombang QRS kompleks [25], [26]
Symlets (sym5, sym8)	Hampir Simetris	Cocok untuk mengurangi distorsi fase saat rekonstruksi sinyal
Coiflets	Banyak vanishing moments pada wavelet	Mampu menangkap halus dari gelombang P dan T

	& scaling function	
Biorthogonal (bior4.4, bior3.5)	Simetri sempurna, cocok untuk aplikasi rekonstruksi	Populer untuk kompresi dan denoising sinyal ECG
Mexican Hat & Morlet (Continuous)	Continuous wavelet, lebih cocok untuk CWT	Berguna untuk deteksi onset PQRST secara presisi (bukan DWT)

Pemilihan *mother wavelet* menjadi aspek penting dalam penerapan DWT. Beberapa penelitian menunjukkan bahwa keluarga wavelet Daubechies, khususnya db4 dan db6, memiliki kesesuaian tinggi terhadap bentuk morfologi sinyal ECG karena kemiripan bentuk basisnya dengan gelombang QRS [24][27]. Pada penelitian ini digunakan wavelet Daubechies-6 (db6) karena memberikan keseimbangan antara resolusi waktu dan frekuensi yang baik untuk mendeteksi komponen diagnostik utama ECG, sebagaimana direkomendasikan oleh studi sebelumnya [24][26][27].

Dengan kemampuan dekomposisi multi-level, DWT memungkinkan isolasi informasi pada berbagai skala, yang sangat berguna untuk perhitungan fitur nonlinier seperti Rényi Entropy pada setiap level. Hal ini memberikan representasi yang lebih kaya untuk keperluan klasifikasi sinyal ECG menjadi normal maupun aritmia.

D. Rényi Entropy

Rényi Entropy merupakan generalisasi dari Shannon Entropy yang diperkenalkan oleh Alfréd Rényi pada tahun 1961 sebagai ukuran ketidakpastian dalam sistem probabilistik yang lebih fleksibel dibandingkan entropi konvensional [10][12]. Berbeda dengan Shannon Entropy yang memiliki satu bentuk tunggal, Rényi Entropy memperkenalkan parameter orde α yang mengatur sensitivitas terhadap distribusi probabilitas, sehingga dapat menyesuaikan fokus analisis terhadap kejadian dengan probabilitas tinggi atau rendah [10][11].

Secara matematis, Rényi Entropy untuk orde $\alpha > 0$ dan $\alpha \neq 1$ didefinisikan sebagai:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log_2 \left(\sum_{i=1}^n P(x_i)^\alpha \right), \alpha > 0, \alpha \neq 1 \quad (3)$$

dengan $P(x_i)$ adalah probabilitas dari kejadian ke- i , n adalah jumlah total kejadian, dan α adalah parameter orde Rényi. Apabila $\alpha \rightarrow 1$, persamaan tersebut akan konvergen menjadi Shannon Entropy. Nilai $\alpha < 1$ akan memberikan bobot

lebih besar pada probabilitas kecil (kejadian langka), sedangkan $\alpha > 1$ akan menekankan pada probabilitas tinggi [10][12].

Dalam analisis sinyal fisiologis seperti ECG, Rényi Entropy terbukti memiliki keunggulan dalam menangkap perubahan kompleksitas sinyal pada berbagai skala waktu [11][15][16]. Hal ini sangat relevan untuk deteksi aritmia, karena episode aritmia sering kali muncul sebagai pola lokal yang jarang terjadi di tengah sinyal normal. Dengan memilih nilai $\alpha < 1$, seperti $\alpha = 0.5$, sensitivitas terhadap kejadian langka dapat ditingkatkan, sehingga membantu model dalam mengenali pola-pola abnormal yang tersembunyi [15][18].

Penerapan Rényi Entropy pada domain wavelet semakin meningkatkan kemampuannya dalam membedakan sinyal normal dan aritmia [16][17]. Setelah sinyal ECG diuraikan menjadi beberapa level melalui Discrete Wavelet Transform (DWT), setiap level memiliki karakteristik frekuensi yang berbeda. Perhitungan Rényi Entropy pada masing-masing level menghasilkan serangkaian fitur nonlinier yang merepresentasikan distribusi energi koefisien wavelet secara kuantitatif. Studi oleh Xin et al. [15] dan Wang et al. [18] menunjukkan bahwa kombinasi DWT dan Rényi Entropy memberikan akurasi klasifikasi yang lebih tinggi dibandingkan penggunaan fitur statistik sederhana.

Dalam penelitian ini, nilai α ditetapkan sebesar 0.5 untuk memaksimalkan sensitivitas terhadap pola-pola minoritas seperti aritmia. Fitur yang dihasilkan dari Rényi Entropy di setiap level DWT digunakan sebagai masukan bagi algoritma klasifikasi XGBoost, yang selanjutnya memutuskan apakah suatu beat termasuk ke dalam kategori normal atau aritmia. Strategi ini memadukan kekuatan analisis multi-resolusi DWT dengan kepekaan Rényi Entropy terhadap distribusi probabilitas yang jarang, sehingga mampu menghasilkan sistem deteksi aritmia yang lebih akurat dan andal [15][16][18].

E. Fitur Statistik

Fitur statistik merupakan parameter numerik yang digunakan untuk menggambarkan karakteristik distribusi data, baik dalam domain waktu maupun domain transformasi. Dalam analisis sinyal ECG, fitur statistik dapat memberikan informasi deskriptif yang melengkapi fitur nonlinier, sehingga membantu meningkatkan kinerja model klasifikasi [17], [19].

Beberapa fitur statistik yang digunakan dalam penelitian ini meliputi:

1. **Rata-rata (Mean)** — Menggambarkan nilai pusat dari distribusi data, dihitung sebagai jumlah seluruh nilai dibagi dengan jumlah sampel. Dalam konteks

koefisien DWT, mean dapat menunjukkan tren umum sinyal pada suatu level dekomposisi.

2. **Deviasi Standar (Standard Deviation)** — Mengukur seberapa besar variasi atau dispersi data terhadap rata-ratanya. Nilai deviasi standar yang tinggi mengindikasikan adanya fluktuasi besar pada sinyal.
3. **Nilai Maksimum (Max)** — Menunjukkan nilai tertinggi dari koefisien sinyal, yang dapat merepresentasikan amplitudo puncak pada suatu level dekomposisi.
4. **Nilai Minimum (Min)** — Menunjukkan nilai terendah dari koefisien sinyal, yang dapat memberikan informasi terkait amplitudo terendah atau depresi sinyal.

Penggunaan fitur statistik bersama fitur berbasis entropi seperti Rényi Entropy dapat memperkaya representasi numerik sinyal. Kombinasi ini efektif dalam mengidentifikasi pola morfologi beat yang tidak hanya berdasarkan kompleksitas sinyal, tetapi juga memperhatikan karakteristik deskriptifnya, sehingga mampu meningkatkan performa model dalam mendeteksi perbedaan antara beat normal dan aritmia [17], [22].

D. XGBoost

Extreme Gradient Boosting (XGBoost) merupakan pengembangan dari algoritma Gradient Boosting Machine (GBM) yang dirancang untuk memiliki performa tinggi, efisiensi komputasi, dan kemampuan generalisasi yang baik [23][29]. Algoritma ini bekerja dengan membangun sekumpulan pohon keputusan (decision trees) secara bertahap, di mana setiap pohon baru berusaha meminimalkan kesalahan (residual) dari gabungan pohon sebelumnya. Proses ini menggunakan fungsi objektif yang terdiri dari loss function dan regularisasi, sehingga mencegah overfitting dan meningkatkan akurasi prediksi [23].

Secara umum, fungsi objektif XGBoost dapat dinyatakan sebagai:

$$\mathcal{L}(\phi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (4)$$

Keunggulan XGBoost dibandingkan GBM konvensional meliputi optimisasi berbasis gradient kedua (second-order gradient optimization), dukungan terhadap parallel processing, serta teknik shrinkage dan column subsampling yang mempercepat proses pelatihan tanpa mengorbankan akurasi [23][29]. XGBoost juga mendukung penanganan data tidak seimbang (imbalanced data) melalui parameter *scale_pos_weight*, yang dapat meningkatkan sensitivitas terhadap kelas minoritas, seperti pada kasus deteksi aritmia [22].

Dalam konteks analisis sinyal ECG, XGBoost telah digunakan secara luas karena kemampuannya dalam menangani dataset berdimensi tinggi yang dihasilkan dari ekstraksi fitur kompleks seperti Discrete Wavelet Transform (DWT) dan Rényi Entropy [17][19]. Studi oleh Kumar et al. [17] menunjukkan bahwa XGBoost mampu mencapai akurasi di atas 90% pada klasifikasi aritmia ketika digunakan bersama dengan fitur entropi hybrid. Selain itu, interpretabilitas model dapat ditingkatkan dengan metode *SHapley Additive exPlanations* (SHAP), yang memberikan wawasan mengenai kontribusi setiap fitur terhadap prediksi [30].

Pada penelitian ini, XGBoost digunakan untuk mengklasifikasikan beat ECG menjadi dua kelas utama, yaitu normal dan aritmia, berdasarkan fitur Rényi Entropy hasil dekomposisi DWT. Parameter-parameter seperti *max_depth*, *learning_rate*, *n_estimators*, *subsample*, dan *colsample_bytree* diatur secara empiris untuk mendapatkan keseimbangan antara akurasi, kompleksitas model, dan waktu inferensi. Selain itu, evaluasi performa dilakukan menggunakan metrik *accuracy*, *ROC AUC*, *log loss*, serta analisis kurva ROC untuk menilai kemampuan model dalam membedakan kedua kelas secara efektif [23][29][30].

Metrik akurasi digunakan untuk mengukur persentase prediksi yang benar dari keseluruhan data uji. Namun, pada dataset yang tidak seimbang, akurasi saja kurang mencerminkan kemampuan model dalam mengenali kelas minoritas. Oleh karena itu, digunakan pula metrik Recall dan Precision yang masing-masing mengukur kemampuan model dalam mendeteksi beat aritmia secara benar serta proporsi prediksi aritmia yang benar-benar tepat.

Metrik ROC AUC (*Receiver Operating Characteristic – Area Under Curve*) memberikan evaluasi menyeluruh terhadap performa klasifikasi biner dengan menghitung luas area di bawah kurva ROC. Nilai ROC AUC yang mendekati 1 menandakan bahwa model mampu membedakan kedua kelas secara konsisten di berbagai ambang keputusan [23].

Selain itu, log loss digunakan untuk menilai kualitas probabilitas prediksi yang dihasilkan model. Log loss yang lebih kecil menunjukkan bahwa probabilitas prediksi lebih selaras dengan label aktual, yang penting untuk aplikasi medis karena memberikan indikasi tingkat keyakinan model terhadap setiap prediksi.

Kombinasi metrik-metrik ini memungkinkan evaluasi model yang lebih komprehensif, tidak hanya dari segi ketepatan klasifikasi, tetapi juga dari sisi kemampuan model mendeteksi

kelas minoritas (sensitivitas) dan keandalan probabilitas prediksi yang dihasilkan [30].

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$Presisi = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$F1 - Score = 2 \cdot \frac{Presisi \cdot Recall}{Presisi + Recall} \quad (8)$$

III. HASIL PENELITIAN

A. Perangkat Penelitian

Penelitian ini dilakukan menggunakan perangkat keras berupa komputer/laptop dengan spesifikasi prosesor Intel® Core™ i7-4710HQ CPU @ 2.50 GHz dengan 8 GB RAM dan sistem operasi Windows 10 Pro (versi 22H2) arsitektur 64-bit. Spesifikasi tersebut dipilih untuk memastikan proses komputasi seperti *filtering* sinyal ECG, dekomposisi Discrete Wavelet Transform (DWT), perhitungan fitur Rényi Entropy, dan pelatihan model klasifikasi XGBoost dapat berjalan dengan waktu eksekusi yang optimal.

Perangkat lunak yang digunakan dalam penelitian ini meliputi:

Python 3.10 sebagai bahasa pemrograman utama.

Google Colaboratory sebagai lingkungan komputasi berbasis cloud untuk mempercepat pemrosesan dataset yang besar.

Jupyter Notebook untuk dokumentasi dan eksekusi kode secara interaktif.

Pustaka Python seperti NumPy, Pandas, SciPy, PyWavelets, scikit-learn, Matplotlib, dan XGBoost.

PhysioNet WFDB Toolbox untuk membaca, memproses, dan mengekstrak anotasi dari MIT-BIH Arrhythmia Database.

Kombinasi perangkat keras dan perangkat lunak tersebut memungkinkan penelitian ini dilaksanakan secara efisien, baik dalam tahap analisis data maupun pelatihan model, tanpa mengalami keterbatasan sumber daya yang signifikan.

B. Evaluasi Metrik

Perbandingan antara dua skenario klasifikasi yang diuji dalam penelitian ini menunjukkan bahwa penambahan fitur statistik ke dalam fitur *Rényi Entropy* secara konsisten

meningkatkan performa klasifikasi *beat* ECG. Kombinasi kedua jenis fitur ini memperkaya representasi sinyal, karena tidak hanya menangkap kompleksitas non-linear, tetapi juga variasi amplitudo yang mencerminkan dinamika morfologi jantung.

Pada skenario pertama, model XGBoost dilatih menggunakan lima fitur *Rényi Entropy* dari level dekomposisi DWT (D1 hingga A4). Model ini menghasilkan akurasi 88,21%, precision 84%, *recall* 79%, dan F1-score 82% untuk kelas aritmia. Nilai ROC AUC sebesar 94,11% menandakan kemampuan klasifikasi yang cukup baik, namun masih terdapat keterbatasan dalam mendeteksi *beat* aritmia secara menyeluruh. Hal ini terlihat dari nilai log loss uji sebesar 29,85%, yang relatif tinggi dan menunjukkan bahwa prediksi probabilistik model masih kurang stabil, terutama untuk kelas minoritas.

Sebaliknya, pada skenario kedua, model dilatih menggunakan gabungan fitur *Rényi Entropy* dan fitur statistik (mean, standar deviasi, maksimum, dan minimum) dari setiap level koefisien DWT. Hasilnya menunjukkan peningkatan performa secara signifikan: akurasi meningkat menjadi 97,90%, dengan precision 98%, *recall* 96%, dan F1-score 97%. Peningkatan paling mencolok terlihat pada ROC AUC yang mencapai 99,56%, yang mendekati performa optimal dalam membedakan dua kelas. Nilai *log loss* uji juga menurun drastis menjadi 7,19%, mencerminkan akurasi probabilistik yang lebih baik dan generalisasi model yang lebih kuat terhadap data baru.

Tabel 2

Perbandingan 2 Model

Metrik Evaluasi	Skenario 1 (Entropi)	Skenario 2 (Entropi + Statistik)
Akurasi	88,21%	97,90%
<i>Precision</i>	84%	98%
<i>Recall</i>	79%	96%
<i>F1-score</i>	82%	97%
ROC AUC	94,11%	99,56%
<i>Train Log Loss</i>	9,96%	0,29%
<i>Test Log Loss</i>	29,85%	7,19%
CV ROC AUC (5-fold)	93,75%	99,59%
<i>Inference Time</i> (detik)	0,1484	0,1583

Kedua model menunjukkan waktu inferensi rata-rata sekitar 0,15 detik per *beat*, yang berada jauh di bawah ambang waktu untuk sistem monitoring *real-time*. Hal ini memastikan bahwa kedua pendekatan tetap layak digunakan pada aplikasi berbasis perangkat portabel atau sistem pemantauan berbasis *Internet of Things* (IoT), terutama pada lingkungan dengan keterbatasan komputasi atau sinyal yang fluktuatif.

Secara keseluruhan, skenario kedua terbukti lebih unggul secara menyeluruh, baik dari aspek akurasi, stabilitas,

sensitivitas, maupun ketahanan model terhadap variasi data. Hasil ini menegaskan bahwa integrasi fitur statistik ke dalam *pipeline* ekstraksi fitur tidak hanya memperkuat klasifikasi, tetapi juga meningkatkan interpretabilitas model dalam konteks klinis.

IV. PEMBAHASAN

Hasil penelitian ini menunjukkan bahwa kombinasi Discrete Wavelet Transform (DWT), *Rényi Entropy*, dan fitur statistik mampu meningkatkan performa klasifikasi sinyal ECG secara signifikan, terutama dalam mendeteksi aritmia. Pada Skenario 1, yang hanya menggunakan fitur *Rényi Entropy*, model XGBoost telah menunjukkan kinerja yang cukup baik dengan akurasi 88,21% dan ROC AUC sebesar 94,11%. Hal ini mengindikasikan bahwa informasi nonlinier berbasis distribusi probabilitas pada masing-masing level DWT memiliki kontribusi penting dalam membedakan pola *beat* normal dan aritmia. Namun, nilai *recall* yang relatif rendah (79%) menunjukkan adanya keterbatasan model dalam menangkap seluruh variasi *beat* aritmia, kemungkinan akibat kurangnya informasi deskriptif dalam representasi fitur.

Pada Skenario 2, penambahan fitur statistik seperti rata-rata (mean), simpangan baku (standard deviation), nilai maksimum, dan nilai minimum berhasil memperkaya representasi numerik sinyal, sehingga model memperoleh informasi tambahan terkait bentuk, dispersi, dan rentang amplitudo *beat* pada domain waktu. Dampaknya terlihat jelas dengan lonjakan akurasi menjadi 97,90% dan peningkatan signifikan pada *recall* hingga 96%, yang berarti model mampu mengidentifikasi hampir seluruh *beat* aritmia secara benar. Nilai F1-score yang mencapai 97% menandakan keseimbangan optimal antara deteksi dan ketepatan prediksi.

Perbedaan performa ini juga tercermin pada metrik log loss, di mana Skenario 2 mencatat penurunan yang sangat signifikan baik pada data latih (dari 9,96% menjadi 0,29%) maupun data uji (dari 29,85% menjadi 7,19%), menandakan peningkatan kepercayaan model terhadap prediksi yang dihasilkan. Peningkatan nilai CV ROC AUC (5-fold) menjadi 99,59% juga membuktikan bahwa peningkatan kinerja tidak hanya terjadi pada subset data tertentu, tetapi konsisten pada seluruh proses validasi.

Secara praktis, meskipun Skenario 2 menambah kompleksitas fitur dan sedikit meningkatkan waktu inferensi (selisih $\pm 0,01$ detik), peningkatan ini tergolong minimal dan masih sangat layak untuk implementasi pada sistem monitoring ECG waktu nyata. Dengan demikian, strategi penggabungan

fitur berbasis entropi dan statistik dapat menjadi pendekatan yang efektif untuk meningkatkan akurasi dan sensitivitas deteksi aritmia, tanpa mengorbankan efisiensi komputasi.

V. KESIMPULAN

Penelitian ini berhasil merancang dan mengimplementasikan sistem deteksi aritmia jantung berbasis analisis Rényi Entropy dan Discrete Wavelet Transform (DWT) yang dikombinasikan dengan model klasifikasi XGBoost. Hasil pengujian menunjukkan bahwa metode DWT dan Rényi Entropy mampu menangkap informasi frekuensi dan kompleksitas sinyal ECG secara efektif. Pada Skenario 1 yang hanya menggunakan fitur Rényi Entropy, model telah mencapai akurasi 88,21% dan ROC AUC 94,11%, menunjukkan potensi kuat dari fitur entropi dalam membedakan beat normal dan aritmia. Penambahan fitur statistik seperti rata-rata (*mean*), simpangan baku (*standard deviation*), nilai maksimum, dan nilai minimum pada Skenario 2 terbukti memperkaya representasi fitur sinyal, menghasilkan peningkatan signifikan pada semua metrik evaluasi, dengan akurasi mencapai 97,90%, recall 96%, dan ROC AUC 99,56%. Strategi balancing data dengan undersampling moderat rasio 2:1 (normal:aritmia) berhasil mengurangi bias terhadap kelas mayoritas tanpa mengorbankan keragaman morfologi beat normal, sehingga model tetap sensitif terhadap kelas minoritas (aritmia) dan mempertahankan stabilitas performa. Efisiensi waktu inferensi tetap terjaga, dengan rata-rata sekitar 0,15 detik per prediksi, menjadikan metode ini layak diimplementasikan pada sistem monitoring ECG real-time. Kombinasi analisis multi-level DWT, parameter Rényi Entropy dengan $\alpha = 0.5$, fitur statistik, dan XGBoost memberikan solusi yang akurat, sensitif, dan efisien untuk deteksi aritmia berbasis sinyal ECG. Penelitian selanjutnya disarankan untuk menguji pendekatan ini pada dataset ECG multilead dengan durasi perekaman lebih panjang, mengintegrasikan metode balancing adaptif, serta mengeksplorasi teknik optimasi model berbasis *ensemble learning* dan *deep learning* untuk meningkatkan generalisasi deteksi pada kondisi klinis yang lebih bervariasi.

REFERENSI

- [1] World Health Organization, "Cardiovascular diseases (CVDs) fact sheet," WHO, 2023. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- [2] Y. Ansari, O. Mourad, K. Qaraqe, and E. Serpedin, "Deep learning for ECG arrhythmia detection and classification: An overview of progress for 2017–2023," *Frontiers in Physiology*, vol. 14, 2023, doi: 10.3389/fphys.2023.1246746.
- [3] A. Antzelevitch and M. Burashnikov, "Overview of basic mechanisms of arrhythmias," *Circulation Research*, vol. 109, no. 5, pp. 648–666, 2011.
- [4] D. P. Zipes and J. Jalife, *Cardiac Electrophysiology: From Cell to Bedside*, 6th ed. Philadelphia, PA, USA: Elsevier, 2014.
- [5] D. S. Levine, J. H. Rosenthal, and S. S. Podrid, "Cardiac arrhythmias: Diagnosis and treatment," in *Heart Disease and Arrhythmias*, 4th ed., pp. 455–470, 2019.
- [6] D. P. Goldberger, *Clinical Electrocardiography: A Simplified Approach*, 8th ed. Philadelphia, PA, USA: Elsevier, 2018.
- [7] G. D. Clifford, F. Azuaje, and P. E. McSharry, *Advanced Methods and Tools for ECG Data Analysis*. Norwood, MA, USA: Artech House, 2006.
- [8] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001, doi: 10.1109/51.932724.
- [9] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. Burlington, MA, USA: Academic Press, 2008.
- [10] P. Jizba and T. Arimitsu, "The world according to Rényi: Thermodynamics of multifractal systems," *Annals of Physics*, vol. 312, no. 1, pp. 17–59, Nov. 2004, doi: 10.1016/j.aop.2004.03.016.
- [11] M. Shi et al., "Rényi distribution entropy analysis of short-term heart rate variability signals and its application in coronary artery disease detection," *Frontiers in Physiology*, vol. 10, 2019, doi: 10.3389/fphys.2019.00809.
- [12] M. Ribeiro et al., "The entropy universe," *Entropy*, vol. 23, no. 2, p. 222, 2021, doi: 10.3390/e23020222.
- [13] Y. Sun, Y. Zhang, and M. Wang, "A novel deep entropy learning approach for ECG arrhythmia detection using hybrid wavelet-domain features," *Computers in Biology and Medicine*, vol. 157, p. 106760, 2023, doi: 10.1016/j.compbiomed.2023.106760.
- [14] S. Śmigiel, K. Pałczyński, and D. Ledziński, "ECG signal classification using deep learning techniques based on the PTB-XL dataset," *Entropy*, vol. 23, no. 9, p. 1121, 2021, doi: 10.3390/e23091121.
- [15] Y. Xin et al., "Paroxysmal atrial fibrillation recognition based on multi-scale Rényi entropy of ECG," *Technology and Health Care*, vol. 25, pp. 189–196, 2017, doi: 10.3233/THC-171321.
- [16] K. D. Shiva and V. V. R. Kondreddy, "Multiscale Rényi entropy analysis for classification of ECG signals," in *Proc. Computing in Cardiology*, vol. 50, Granada, Spain, 2023, pp. 1–4. [Online]. Available: <https://www.cinc.org/archives/2023/pdf/CinC2023-447.pdf>
- [17] A. Kumar, S. Pal, and R. Balasubramanian, "Hybrid entropy features and machine learning for ECG-based arrhythmia classification," *IEEE Access*, vol. 10, pp. 18102–18113, 2022, doi: 10.1109/ACCESS.2022.3148654.
- [18] M. Wang, C. Zhang, Q. Sun, J. Zhao, and Y. Zhang, "A multiscale Renyi entropy-based method for atrial fibrillation detection," *Frontiers in Physiology*, vol. 13, p. 846158, 2022. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/35492618/>
- [19] M. Qodri, N. A. Rahman, and D. C. E. Wirawan, "Optimasi klasifikasi aritmia jantung berbasis ECG menggunakan hybrid entropy dan CNN," in *Proc. Applied Computer and Engineering*

(ACE), 2023. [Online]. Available: <https://www.ewadirect.com/proceedings/ACE/article/view/15901>

[20] I. A. Fadhlurrahman, A. Rachmatillah, and F. A. Nugroho, "Pendeteksian aritmia jantung berdasarkan citra sinyal ECG dengan metode convolutional neural network," *J. Teknol. Inform.*, vol. 7, no. 1, pp. 1–10, 2023. [Online]. Available: <https://e-journals.dinamika.ac.id/joti/article/view/792>

[21] Y.-H. Lin, H.-Y. Hsieh, and C.-Y. Lin, "QRS complex detection using wavelet transform and neural networks," *IEEE Access*, vol. 7, pp. 60441–60449, 2019, doi: 10.1109/ACCESS.2019.2914782.

[22] L. Sharma and S. Kaushik, "ECG arrhythmia classification using ensemble learning with entropy-based features," *Biomedical Signal Processing and Control*, vol. 83, p. 104577, 2023, doi: 10.1016/j.bspc.2023.104577.

[23] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017. [Online]. Available: <https://arxiv.org/abs/1705.07874>

[24] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[25] M. Arif, M. A. Ghayour, and A. El-Rafei, "Comparative analysis of wavelet families for ECG feature extraction and

classification," *Journal of Medical Systems*, vol. 46, no. 7, pp. 1–15, 2022, doi: 10.1007/s10916-022-01822-0.

[26] A. Bouaziz, M. Deriche, and A. Kardouchi, "ECG signal denoising and features extraction using discrete wavelet transform (DWT)," *Int. J. Comput. Appl.*, vol. 96, no. 20, pp. 36–42, June 2014, doi: 10.5120/16720-6401.

[27] T. Phukpattaranont, "Comparison of wavelet functions for QRS detection in ECG signal," *BioMedical Engineering OnLine*, vol. 14, no. 1, p. 59, 2015, doi: 10.1186/s12938-015-0048-8.

[28] K. A. Basha and S. S. Kumar, "ECG signal analysis using wavelet transform," *Int. J. Eng. Trends Technol. (IJETT)*, vol. 17, no. 4, pp. 168–170, Nov. 2014. [Online]. Available: <https://ijettjournal.org/volume-17/number-4/IJETT-V17P234.pdf>

[29] M. U. Zahid et al., "Robust R-peak detection in low-quality Holter ECGs using 1D convolutional neural network," *arXiv preprint*, arXiv:2101.01666, Jan. 2021. [Online]. Available: <https://arxiv.org/abs/2101.01666>

[30] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017. [Online]. Available: <https://arxiv.org/abs/1705.07874>