Non-Linear Functional Principal Component Analysis pada Sinyal HRV untuk Deteksi Awal Penyakit Jantung Koroner

1st Dinda Dwi Handini
Teknik Biomedis
Fakultas Teknik Eleketro
Bandung, Indonesia
dindahandini@student.telkomuniversity.ac.id

2nd Tito Waluyo Purboyo
Teknik Biomedis
Fakultas Teknik Eleketro
Bandung, Indonesia
titowaluyo@telkomuniversity.ac.id

3rd Dziban Naufal
Teknik Biomedis
Fakultas Teknik Eleketro
Bandung, Indonesia
dzibannaufal@telkomuniversity.ac.id

***Abstrak* — Penyakit Jantung Koroner (PJK) menyebabkan 17,8 juta kematian global per tahun dan 14,4% kematian di Indonesia. Metode deteksi dini konvensional terbatas pada pendekatan linear yang tidak mampu menangkap pola non-linear dalam sinyal Heart Rate Variability (HRV). Penelitian ini menganalisis karakteristik fitur HRV dari rekonstruksi sinyal EKG menggunakan Functional Principal Component Analysis (FPCA) dan Principal Component Analysis (PCA), serta mengevaluasi kontribusi fitur HRV dan TSFEL dalam klasifikasi pasien normal dan PJK menggunakan Support Vector Machine (SVM) dan Logistic Regression.**

Dataset MIT-BIH dan St. Petersburg INCART digunakan dengan 7 rekaman masing-masing pada Lead VI dan MLII. Pre-processing meliputi upsampling 360 Hz, segmentasi 30 detik, dan deteksi puncak R. Rekonstruksi sinyal menggunakan PCA (5,10,15 komponen) dan FPCA dengan basis Fourier. Analisis distribusi menggunakan uji Kolmogorov-Smirnov.

Hasil menunjukkan seluruh fitur HRV memiliki distribusi berbeda signifikan ($p<0,001$) antara subjek normal dan PJK. Fitur TSFEL mencapai akurasi 100% pada PCA 15 komponen, sedangkan kombinasi HRV+TSFEL pada FPCA basis 150 dengan 15 FPC memberikan performa terbaik dengan akurasi 99,60% dan ROC AUC 99,99%. Fitur paling signifikan adalah mean_hr, cv_rr, signal_Autocorrelation, dan signal_Spectral distance, mengkonfirmasi potensi metode untuk deteksi dini PJK yang akurat.

***Kata kunci*— *Functional Principal Component Analysis*, *Heart Rate Variability*, *Logistic Regression*, Penyakit Jantung Koroner, *Support Vector Machine*, TSFEL**

I. PENDAHULUAN

Penyakit Jantung Koroner (PJK) merupakan salah satu tantangan kesehatan global terbesar dan menjadi penyebab utama kematian di seluruh dunia. Data dari *World Health Organization* (WHO) menunjukkan bahwa penyakit ini bertanggung jawab atas lebih dari 17,8 juta kematian setiap tahunnya [1].

Interpretasi EKG secara manual melalui penglihatan visual oleh dokter sangat bergantung pada keahlian subjektif individu dalam mengukur interval dan menilai morfologi gelombang, sehingga proses ini rentan terhadap variabilitas antar pemeriksa dan kurang sensitif dalam mendeteksi perubahan halus yang mungkin mengindikasikan fase awal PJK [2].

Salah satu pendekatan utama untuk deteksi dini PJK adalah melalui analisis sinyal *Electrocardiogram* (EKG) dan parameter turunannya, yaitu *Heart Rate Variability* (HRV). HRV, atau variasi interval antar detak jantung, telah lama digunakan sebagai indikator penting kesehatan jantung karena kemampuannya mencerminkan keseimbangan sistem saraf otonom [3]. Namun, metode analisis HRV yang umum digunakan saat ini sering kali terbatas pada pendekatan linear. Keterbatasan ini menjadi celah penelitian yang signifikan, karena metode linear gagal menangkap pola-pola non-linear yang kompleks dalam sinyal HRV, yang diyakini mengandung informasi diagnostik tambahan yang berharga [4].

Oleh karena itu, penelitian ini bertujuan untuk mengembangkan sebuah metode deteksi dini PJK yang baik dengan menerapkan *Princpipal Component Analysis* (PCA) dan *Functional Principal Component Analysis* (FPCA) [6]. Penelitian ini dilakukan untuk mengembangkan dan mengevaluasi metode klasifikasi antara subjek normal dan pasien Penyakit Jantung Koroner (PJK). Penggunaan FPCA dalam penelitian ini didasarkan pada kemampuannya dalam menangkap informasi temporal dari sinyal EKG secara lebih utuh, terutama ketika sinyal telah direpresentasikan dalam domain fungsi melalui pendekatan basis seperti *Fourier*.

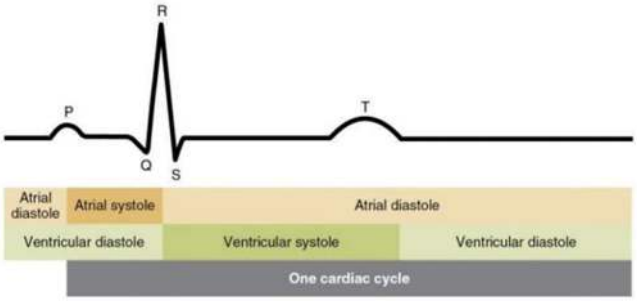
II. KAJIAN TEORI

A. Anatomi dan fisiologi Jantung

Jantung merupakan organ otot yang berongga dan berukuran sebesar kepalan tangan yang berfungsi sebagai pemompa darah ke pembuluh darah dengan kontraksi ritmik dan berulang.

Fisiologi jantung yang menyebabkan adanya gelombang EKG berasal dari depolarisasi dari sel pacemaker yang berasal dari nodus sinoatrial (SA), yang merupakan pemacu utama dalam menghasilkan impuls listrik yang dimana menjalar dan menyebabkan depolarisasi dari seluruh atrium yang menyebabkan terbentuknya gelombang P pada EKG [24], impuls listrik mengalami penundaan pada nodus

atrioventrikular (AV) yang terlihat seperti segmen PR. Penundaan ini memastikan atrium dapat berkontraksi secara penuh sebelum ventrikel melakukan aktivasi yang dimana setelah itu bergerak melalui sistem konduksi khusus untuk mendepolarisasi ventrikel secara terkoordinasi dari endokardium menuju epikardium, yang menyebabkan terbentuknya kompleks QRS. Setelah terjadinya depolarisasi, ventrikel mengalami fase repolarisasi. Proses repolarisasi miokatdium ventrikel, yang bermulai dari area yang terdepolarisasi terakhir, yang dimana menghasilkan gelombang T pada EKG [5].



GAMBAR 1
Fase Gelombang EKG [19]

Penyakit Jantung Koroner merupakan konsekuensi progresif dari aterosklerosis arteri koroner, yaitu penumpukan plak yang tersusun atas kolesterol, lemak, kalsium, dan material seluler pada dinding pembuluh sehingga lumen menyempit dan elastisitas menurun. Pada tahap awal, penyempitan dapat belum menimbulkan gejala karena kebutuhan oksigen miokard masih tercukupi saat istirahat. Ketika kebutuhan meningkat (aktivitas fisik atau stres emosional), suplai relatif tidak memadai sehingga muncul angina berupa rasa tertekan, tertindih, atau terbakar di dada yang dapat menjalar ke lengan, leher, rahang, punggung, atau epigastrium. Jika terjadi ruptur plak diikuti pembentukan trombus yang menyumbat aliran secara mendadak, dapat terjadi infark miokard akut dengan gejala lebih berat: nyeri dada yang persisten, keringat dingin, mual, sesak, lemas, sampai henti jantung mendadak. Faktor risiko utama PJK meliputi merokok, hipertensi, dislipidemia (LDL tinggi, HDL rendah), diabetes, obesitas sentral, pola makan tinggi lemak jenuh dan gula sederhana, kurang aktivitas fisik, stres kronis, konsumsi alkohol berlebihan, usia lanjut, serta riwayat keluarga. Strategi pencegahan primer dan sekunder menitikberatkan pada modifikasi gaya hidup (diet seimbang, berhenti merokok, kontrol berat badan, aktivitas fisik rutin) ditambah terapi farmakologis sesuai indikasi (statin, antiplatelet, pengendali tekanan darah dan glukosa). Pada kasus dengan obstruksi bermakna atau gejala refrakter, intervensi revaskularisasi seperti angioplasti koroner dengan pemasangan stent atau bypass arteri koroner dipertimbangkan untuk memulihkan perfusi miokard dan menurunkan risiko kejadian kardiovaskular lanjut [6].

B. *Electrocardiogram* (EKG)

Electrocardiogram (EKG) merupakan pencatatan non-invasif aktivitas listrik jantung yang aman, tidak menimbulkan nyeri, dan memiliki diagnostik tinggi dalam menilai ritme, konduksi, serta mendeteksi kelainan miokard, termasuk akibat gangguan aliran darah koroner [7].

Pemeriksaan EKG diawali dengan penjelasan kepada pasien mengenai sifat prosedur yang non-invasif dan tidak menimbulkan nyeri. Pasien ditempatkan dalam posisi berbaring terlentang, melepaskan pakaian di area dada serta perhiasan atau aksesoris logam. Sepuluh elektroda kemudian dipasang secara tepat pada posisi anatomis standar, yaitu 6 elektroda prekordial di dada dan 4 elektroda ekstremitas di pergelangan tangan dan kaki, yang kemudian dihubungkan ke mesin EKG. Perekaman dilakukan selama beberapa detik hingga menit, bergantung pada kebutuhan klinis, untuk memperoleh kurva gelombang P, QRS, dan T yang diperlukan untuk evaluasi diagnostik.

Heart Rate Variability (HRV) adalah penanda dinamis fungsi dan keseimbangan sistem saraf otonom (simpatis–parasimpatis) yang telah dikaji luas dalam konteks risiko Penyakit Jantung Koroner (PJK). HRV merefleksikan adaptabilitas regulasi kardiovaskular; penurunannya sering dihubungkan dengan disfungsi otonom dan meningkatnya kerentanan kejadian kardiovaskular.

Dalam HRV, ada yang dikenal dengan interval R-R, yang merupakan jarak antar puncak gelombang R pada rekaman EKG itu sendiri.

C. Principal Component Analysis (PCA)

Functional Principal Component Analysis (FPCA) adalah teknik statistik yang efektif dalam menganalisis data fungsional, seperti deret waktu atau sinyal kontinu. Teknik analisis ini memperluas PCA tradisional ke data fungsional, yang biasanya dipresentasikan sebagai kurva atau permukaan, bukan sebagai titik-titik diskrit. FPCA ini berfungsi sebagai teknik pengurangan dimensi yang mengekstrak fitur-fitur penting dari data yang berdimensi tinggi yang memungkinkan pemrosesan yang lebih efisien dan meningkatkan kinerja klasifikasi.

Secara konkret pada analisis sinyal Elektrokardiogram (EKG), penerapan FPCA mengikuti langkah-langkah intuitif. Pertama, sinyal EKG yang panjang dipecah menjadi segmen-segmen individual yang merepresentasikan satu siklus detak jantung (misalnya dari satu ke berikutnya). Setiap segmen detak jantung ini diperlakukan sebagai satu 'fungsi' atau kurva. Kumpulan dari ratusan atau ribuan kurva detak jantung inilah yang menjadi data fungsional.

FPCA kemudian bekerja dengan cara:

1. Menghitung Fungsi Rata-rata (*Mean Function*): Teknik ini menghitung bentuk 'rata-rata' dari semua detak

jantung yang ada. Fungsi rata-rata ini merepresentasikan morfologi EKG yang paling umum dalam dataset.

2. Mengekstrak *Eigenfunction*: Selanjutnya, FPCA mengidentifikasi *Eigenfunction* (atau disebut juga Komponen Utama Fungsional). *Eigenfunction* ini adalah 'mode variasi' utama yang menjelaskan bagaimana setiap detak jantung individu menyimpang dari detak jantung rata-rata. Sebagai contoh:
   a. *Eigenfunction* 1 mungkin menangkap variasi pada amplitudo gelombang R.
   b. *Eigenfunction* 2 mungkin menjelaskan variasi pada durasi interval QRS.
   c. *Eigenfunction* 3 bisa jadi berhubungan dengan perubahan pada elevasi segmen ST.
3. Menghitung Skor FPC (FPC *Scores*): Untuk setiap detak jantung, dihitung skor FPC untuk masing-masing eigenfunction. Skor ini adalah sebuah angka tunggal yang mengukur sejauh mana 'mode variasi' tersebut muncul pada detak jantung tersebut.

Hasil akhirnya, setiap kurva detak jantung yang tadinya terdiri dari ratusan titik data kini dapat direpresentasikan secara efisien hanya dengan beberapa skor FPC (misalnya 3-5 skor). Skor-skor inilah yang kemudian menjadi fitur input yang jauh lebih ringkas dan informatif untuk model klasifikasi, misalnya untuk mendeteksi penyakit jantung koroner.

D. *Support Vector Machine* (SVM)

Support Vector Machine (SVM) adalah algoritma klasifikasi yang mencari *hyperplane* dengan margin maksimum untuk memisahkan kelas sehingga efektif pada data berdimensi menengah–tinggi, termasuk pola non-linier melalui pemetaan ke ruang fitur lebih tinggi menggunakan fungsi *kernel*. Kinerja SVM dipengaruhi oleh: kualitas representasi fitur (seleksi/reduksi untuk mengurangi redundansi), dan penalaan hyperparameter (misal CCC, γ, derajat polinomial). Penerapan SVM dengan seleksi fitur *information gain* menunjukkan peningkatan stabilitas dan akurasi setelah fitur tidak informatif dieliminasi [8].

E. Time Series Feature Extraction Library (TSFEL)

Time Series Feature Extraction Library (TSFEL) adalah sebuah pustaka perangkat lunak yang dirancang untuk menstandarkan dan memudahkan proses ekstraksi fitur dari beragam sinyal deret waktu. TSFEL menyediakan antarmuka yang relatif sederhana untuk memanggil puluhan fungsi fitur yang telah dikelompokkan ke dalam beberapa *domain*. Dengan demikian, peneliti dapat dengan cepat menghasilkan *feature matrix* dari kumpulan sinyal tanpa harus mengimplementasikan setiap rumus secara manual.

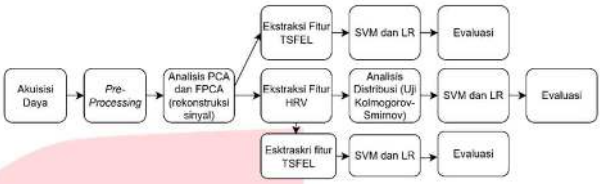
F. Logistic Regression

Logistic Regression adalah algoritma klasifikasi statistik yang menggunakan fungsi logistik untuk memodelkan probabilitas keanggotaan kelas. Algoritma ini cocok untuk deteksi dini penyakit jantung koroner karena dapat menangani fitur HRV yang sering berdistribusi non-normal dan memberikan output berupa probabilitas risiko.

Logistic regression memiliki keunggulan interpretabilitas tinggi, andal terhadap outlier, dan memberikan output probabilitas yang berguna untuk stratifikasi risiko kardiovaskular. Penelitian menunjukkan efektivitas algoritma ini dalam klasifikasi penyakit jantung berdasarkan fitur EKG dan HRV dengan performa yang kompetitif [9].

III. METODE

Desain sistem pada penelitian ini betujuan untuk membangun alur kerja dalam melakukan klasifikasi antara pasien PJK dan pasien normal. Berdasarkan PCA dan FPCA yang berasal dari fitur hasil rekonstruksi sinyal menggunakan TSFEL, ekstraksi fitur HRV, dan ekstraksi fitur HRV dan TSFEL.



GAMBAR 2
Desain Sistem Keseluruhan

Penelitian ini memanfaatkan dua sumber utama data rekaman EKG, yaitu St. Peterburg INCART *12-lead Arrythmia Database* dan MIT-BIH *Arrhhthmia Database*, yang masing-masing mewakili kelompok pasien dengan Penyakit Jantung Koroner (PJK) dan kelompok pasien dengan irama jantung normal.

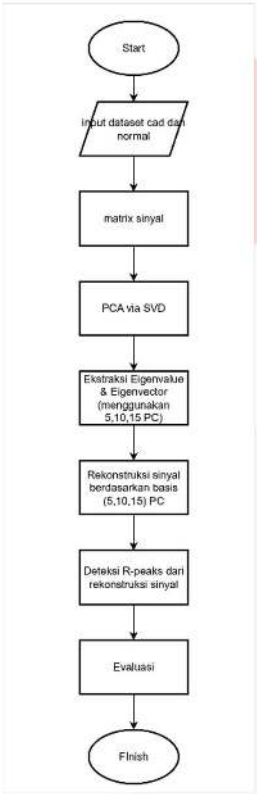
Sumber data pertama berasal dari St. *Peterburg* INCART 12-lead Arrythmia *Database*, yang terdiri atas total 75 rekaman EKG beranotasi lengkap. Sementara itu, untuk mewakili kelompok pasien dengan ritme jantung normal, digunakan *dataset* dari MIT-BIH *Arrhythmia Database*, yang merupakan satu *dataset* yang paling banyak digunakan untuk pengolahan sinyal biomedis.

Sinyal EKG yang digunakan berasal dari dua sumber yang berbeda. Masing-masing dari kedua dataset tersebut memiliki karakteristik yang berbeda. Pada tahap *pre-processing* ini, ada beberapa hal yang harus dilakukan, yaitu penyesuaian frekuensi sampling, Segmentasi dan Pendeteksian R-peaks

Berikut adalah *flowchart* dari *processing*.

GAMBAR 3
Flowchart Pre-Processing

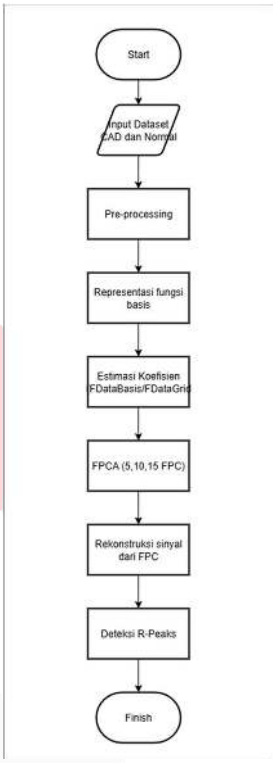
Untuk membuat analisis PCA ini, menggunakan tiga konfigurasi yang berbeda, yaitu 5,10, dan 15 *Principal Component* (PC). Untuk hasil rekonstruksi PCA dengan menggunakan 5,10, dan 15 PC, R-*peaks* juga ditandai untuk ekstraksi HRV pada tahap selanjutnya, Dalam aspek teknis sendiri, untuk analisis PCA ini menerapkan *Singular Value Decomposition* sebagai metode untuk komputasi *eigen values* dan *eigenvector* untuk menghasilkan sinyal rekonstruksi dengan 3 variasi yaitu 5,10, dan 15 PC. Di saat sinyal rekonstruksi sudah dibuat, *Mean Squared Error* (MSE) digunakan untuk mengukur akurasi numerik rekonstruksi dengan membandingkan sinyal asli dan rekonstruksi beserta R-*peaksnya* juga.



GAMBAR 4
Flowchart alur analisis PCA

Analisis *Functional Principal Component Analysis* (FPCA) dilakukan melalui beberapa tahapan untuk bisa mendapatkan representasi sinyal yang efisien dan memungkinkan untuk direkonstruksi dengan jumlah komponen 5, 10, dan 15 dengan basis *fourier* sebesar 50,150, dan 301. Sinyal yang semula berupa data diskrit di konversi menjadi fungsi kontinu menggunakan fungsi basis *fourier* Fungsi basis (*basis function*) sendiri merupakan sekumpulan fungsi yang digunakan sebagai bentuk representasi pembentuk daro sinyal kontinu.

Proses Ekstraksi Fitur *Heart Rate Variability* dalam penelitian ini dilakukan berdasarkan segmentasi sinyal EKG yang telah dibagi sebelumnya yaitu selam 30 detik per segmen. Setiap segmen 30 detik tersebut dianalisis dengan PCA dan FPCA lalu menghasilkan rekonstruksi sinyal. Hasil sinyal rekonstruksi PCA dengan 5,10, dan 15 PC dan sinyal hasil rekonstruksi FPCA dengan 5,10,15 FPC, beserta menggunakan 50,150, dan 301 basis *fourier* ini dianalisis secara *independent* untuk menghitung berbagai parameter dari HRV dari domain waktu (*time-domain*), tanoa menerapkan pendekatan *sliding window* tambahan.



GAMBAR 5
Flowchart alur analisis FPCA

Selain melakukan ekstraksi fitur HRV dari domain waktu, penelitian ini juga memanfaatkan *Time Series Feature Extraction Library* (TSFEL) untuk mengekstraksi fitur tambahan dari sinyal. Proses ekstraksi fitur TSFEL ini dilakukan terhadap dua jenis sinyal: sinyal EKG hasil rekonstruksi menggunakan metode PCA dengan 5,10, dan 15 PC dan hasil rekonstruksi menggunakan metode FPCA dengan 5,10, dan 15 FPC dengan basis 50,150, dan 301 basis *fourier*, serta sinyal HRV hasil rekonstruksi sinyal yang sebelumnya telah dihitung berdasarkan interval RR.

Fitur TSFEL yang diekstraksi mencakup tiga domain utama, yaitu domain waktu (*time domain*), domain statistik (*statistical domain*), dan domain frekuensi (*spectral* domain), sedangkan fitur dari domain fractal tidak digunakan dalam penelitian ini. TSFEL baik pada sinyal hasil rekonstruksi FPCA dan PCA ataupun dari HRV, diperoleh Kumpulan fitur yang digunakan untuk menangkap pola tersembunyi dalam pasien dengan Penyakit Jantung Koroner (PJK) dan pasien normal. Adapun kombinasi antara fitur HRV dan fitur-fitur TSFEL ini juga akan digunakan sebagai representasi fitur gabungan (HRV+TSFEL) untuk tahap analisis selanjutnya, seperti klasifikasi.

Untuk mengetahui apakah terdapat adanya perbedaan distribusi yang signifikan antara kelompok dengan jantung normal dan kelompok dengan Penyakit Jantung Koroner. Uji dilakukan dengan *Kolmogorov-Smirnov* 2 parameter. Uji ini merupakan metode non-parametik yang digunakan untuk bisa membandingkan dua distribusi kumulatif dari kelompok

pasien PJK dan pasien normal, tanpa perlu mengasumsikan bentuk distribusi tertentu.

algoritma *Support Vector Machine* (SVM) digunakan sebagai salah satu metode klasifikasi utama untuk membedakan antara pasien dengan PJK dan pasien normal. SVM merupaka algoritma *machine learning* berbasis *supervised learning* yang efektif untuk tugas klasifikasi, baik pada data linear maupun non-linear. Untuk menangani pola non-linear dalam data fisiologis seperti sinyal jantung, digunakan *kernel Radial Basis Function* (RBF), yang memungkinkan SVM memetakan data ke dalam ruang berdimensi lebih tinggi sehingga dapat dipisahkan dengan lebih baik. Selain SVM, penelitian ini juga menggunakan algoritma *Logistic Regression* sebagai metode pembanding. *Logistic Regression* adalah model klasifikasi linear yang banyak digunakan dalam studi klinis karena interpretabilitasnya yang tinggi dan kemampuannya dalam menangani probabilitas kelas.

Evaluasi dilakukan untuk memastikan bahwa proses klasifikasi antara pasien dengan PJK dan pasien normal dilakukan secara akurat. Dalam penelitian ini, digunakan dua algoritma, yaitu *Support Vector Machine* (SVM) dan *Logistic Regression* (LR). Untuk mengevaluasi performa model, digunaan metrik evaluasi, seperti *Area Under the Curve* (AOC) dari *Receiver Operating Characteristic* (AOC-ROC) yang memberikan gambaran seberapa baik model dalam membedakan kedua kelas ini.

Akurasi klasifikasi juga dihitung berdasarkan confusion matrix, yang dihitung atas jumlah prediksi benar dan salah, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

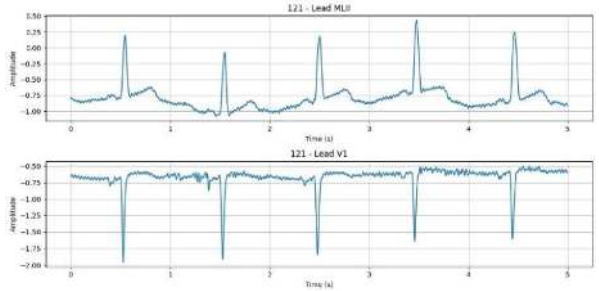
Metode pengujian digunakan untuk mengevaluasi kinerja sistem dengan membandingkan kondisi pasien normal dan PJK berdasarkan hasil rekonstruksi dengan fitur TSFEL, ekstraksi fitur HRV, dan ekstraksi fitur HRV ditambah dengan fitur TSFEL. Untuk menghasilkan data yang lebih valid, pengujian ini dilakukan dengan membagi data latih dan data uji menggunakan *cross validation*. Klasifikasi menggunakan SVM dan LR mendapatkan nilai akurasi digunakan untuk mengukur kerja dari kedua model itu. Dilakukan juga uji statistik *Kolmogrov-Smirnov* menggunakan 2 parameter untuk setiap fitur domain waktu HRV untuk mengetahui seberapa signifikan perbedaan distribusi antara kelompok pasien normal dan PJK. Kemudian, hasil pengujian ini dilakukan untuk mengetahui dan menilai kemampuan sistem untuk mengidentifikasi pola karakteristik jantung dari sinyal EKG.

IV. HASIL DAN PEMBAHASAN

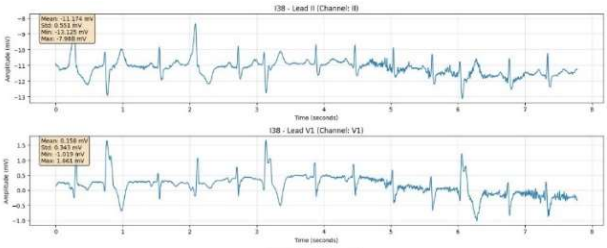
A. Pre-Processing

Tahap awal dalam alur penelitian ini adalah pre-processing sinyal EKG mentah, yang bertujuan untuk membersihkan data dari berbagai artefak dan noise, seperti baseline akibat pergerakan dan interferensi jaringan listrik, yang dapat mengganggu akurasi analisis.

Tahap awal dalam *pre-processing* adalah menyamakan frekuensi *sampling* dari *Dataset* pasien normal dan PJK. Dataset yang berasal dari MIT-BIH *Arrhythmia Database* memiliki frekuensi *sampling* sebesar 360 Hz, dan *dataset* dari St. Peterburg INCART 12-*lead Arrythmia Database* memiliki frekuensi sampling sebesar 257 Hz.

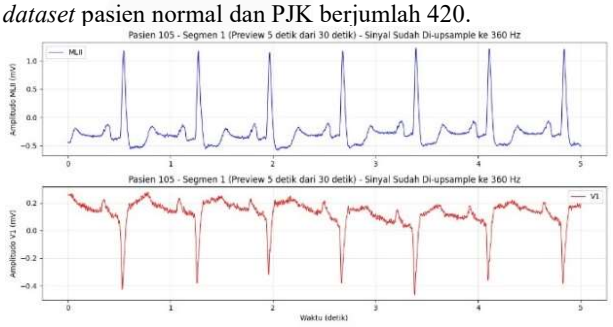


GAMBAR 6
Visualiasi Sinyal EKG normal

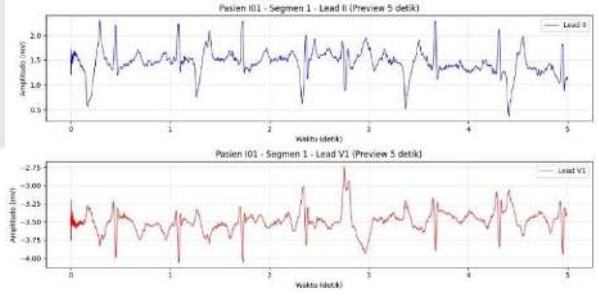


GAMBAR 7
Visualiasi Sinyal EKG PJK

Metode *upsampling* untuk *dataset* St. Peterburg INCART 12-lead *Arrythmia Database* menggunakan teknik dari *library scipy.signal.resample*. Untuk masing-masing *dataset* untuk *dataset* normal dan PJK juga di segmentasi selama 30 detik untuk 1 segmen, yang di mana total segmen untuk *dataset* pasien normal dan PJK berjumlah 420.

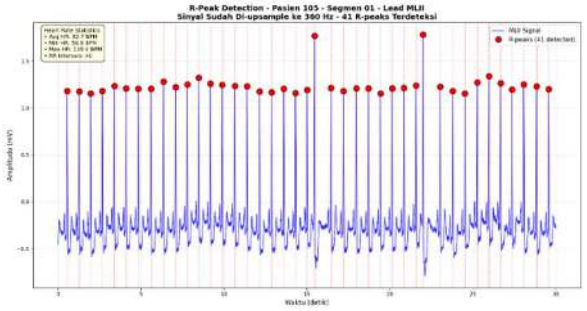


GAMBAR 8
Visualisasi sinyal EKG normal upsampling dan segmentasi

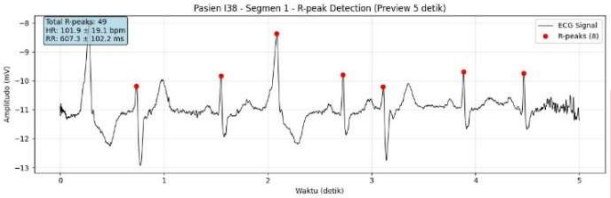


GAMBAR 9
Visualisasi sinyal EKG PJK upsampling dan segmentasi

Setelah data di *upsampling*, dapat dilihat bahwa detail sinyal jantung QRS masih terlihat jelas, *amplitude* terlihat lebih relatif besar. Setelah di *upsampling* dan segmentasi, lalu dideteksi untuk R-*peaksnya* untuk masing -masing *lead* V1 dan MLII.



GAMBAR 10
Visualisasi R-peaks pada sinyal EKG pasien Normal

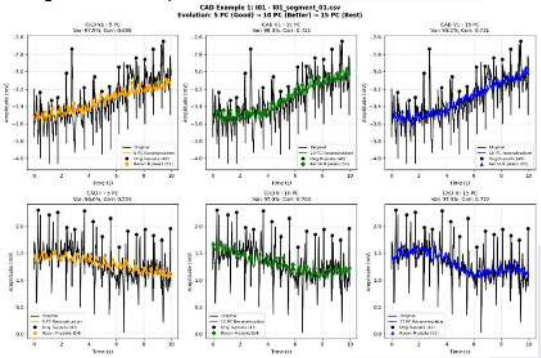


GAMBAR 11
Visualisasi s pada sinyal EKG pasien PJK

Pada hasil visualisasi R-*peaks* pada sinyal sebelum di rekonstruksi dengan menggunakan PCA dan FPCA, banyak R-*peaks* yang terdeteksi untuk satu segmen baik untuk *dataset* pasien PJK dan pasien normal berkisar di antara 39-49 R-*peaks*.

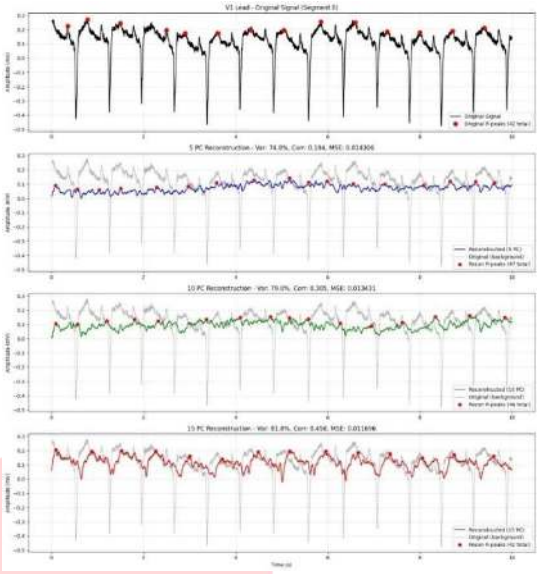
B. Rekonstruksi Sinyal dengan PCA dan FPCA

1. Hasil Rekonstruksi Sinyal Menggunakan *Principal Component Analysis* (PCA)



GAMBAR 12 Hasil Gambar Rekonstruksi PCA dengan 5,10, dan 15 PC PJK

Analisis PCA dilakukan untuk mengevaluasi kemampuan reduksi dimensi dalam mempertahankan informasi penting dari sinyal EKG. Pendekatan ini menggunakan StandardScaler untuk normalisasi yang digunakan untuk mempertahankan karakteristik variansi natural sinyal. Evaluasi dan rekonstruksi sinyal dilakukan pada 15 komponen utama dengan fokus pada konfigurasi 5, 10, dan 15 PC untuk menganalisis efisiensi komputasi dan kualitas rekonstruksi.

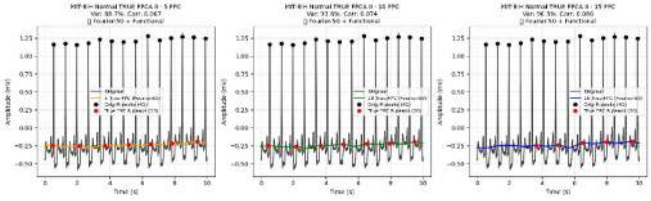


GAMBAR 13
Hasil Rekonstruksi PCA dengan 5,10, dan 15 PC Normal

Hasil dekomposisi sinyal EKG menggunakan PCA menunjukkan perbedaan karakteristik antara dataset PJK dan normal. Pada dataset PJK (gambar pertama), sinyal dapat direkonstruksi dengan sangat baik hanya dengan 5 hingga 15 komponen utama, dengan variansi mencapai di atas 97% dan korelasi yang meningkat seiring penambahan PC. Jumlah R-peak hasil rekonstruksi juga konsisten dan mendekati jumlah asli, mencerminkan sinyal yang lebih sederhana dan terpusat pada komponen dominan. Sebaliknya, pada dataset normal (gambar kedua), rekonstruksi sinyal memerlukan lebih banyak komponen untuk mencapai kualitas serupa. Meskipun variansi kumulatif meningkat dari 74,8% (5 PC) menjadi 81,8% (15 PC), korelasi terhadap sinyal asli tetap rendah, dan pola R-peak tidak setepat pada dataset PJK. Hal ini mengindikasikan bahwa sinyal EKG pasien normal memiliki struktur yang lebih kompleks dan variatif dibandingkan pasien penyakit jantung koroner.

2. Hasil Rekonstruksi Sinyal Menggunakan *Principal Component Analysis* (PCA)

Implementasi FPCA menggunakan basis *Fourier* dengan tiga konfigurasi berbeda bertujuan untuk mengevaluasi pengaruh resolusi frekuensi terhadap kualitas rekonstruksi sinyal. Pendekatan *functional analysis* ini memberikan keunggulan dalam menangani karakteristik temporal dan frekuensi sinyal EKG secara simultan.



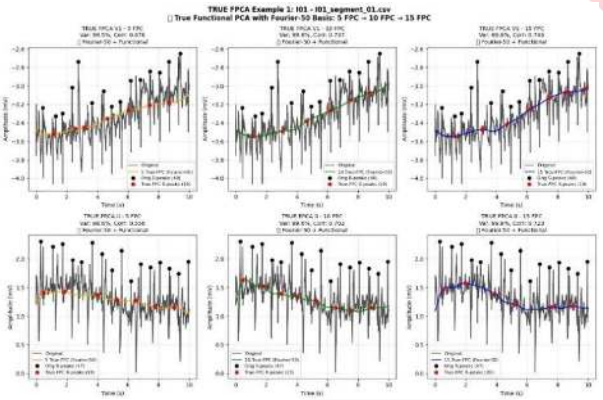
GAMBAR 14.
Hasil Gambar Rekonstruksi Sinyal FPCA dengan Basis 50

FPCA dengan basis 50 menunjukkan kemampuan yang cukup baik dalam menjelaskan *variance* sinyal dengan mencapai 96,30% pada 15 FPC. Dominasi FPC1 sebesar 71,20% menunjukkan efektivitas *functional smoothing* dalam

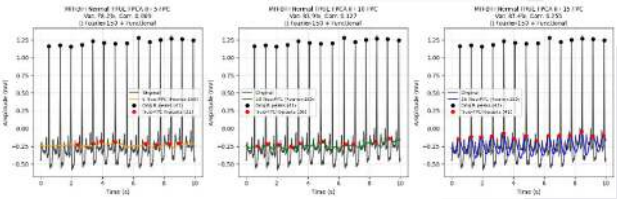
mengidentifikasi pola temporal utama. Meskipun variansi kumulatif meningkat dari 88,7% (5 FPC) menjadi 96,3% (15 FPC), bentuk QRS kompleks, yaitu puncak R serta gelombang Q dan S yang tajam dan khas, belum tergambar dengan jelas dalam sinyal hasil rekonstruksi. Sinyal tampak halus dan kurang tajam, terutama pada penggunaan 5 dan 10 FPC. Bahkan pada 15 FPC, bentuk puncak masih tampak tumpul dan tidak memiliki karakteristik lonjakan cepat khas QRS. Akibatnya, jumlah R-peak yang terdeteksi secara otomatis pada sinyal hasil rekonstruksi masih belum sepenuhnya akurat, meskipun jumlahnya bertambah seiring peningkatan FPC (dari 13 menjadi 20 *peak*).

Hasil FPCA basis 50 pada PJK menunjukkan karakteristik yang berbeda jika dibandingkan dengan pola normal. Lead V1 mencapai variance yang tinggi, dengan FPC1 sebesar 95,49%. Lead II menunjukkan pola serupa dengan FPC1 yaitu 92,1%.

Peningkatan basis menjadi 150 bertujuan untuk mencapai keseimbangan yang baik antara resolusi frekuensi dan akurasi temporal. Konfigurasi ini memberikan representasi yang lebih detail tanpa mengorbankan *precision* dalam deteksi karakteristik temporal seperti R-*peak*.

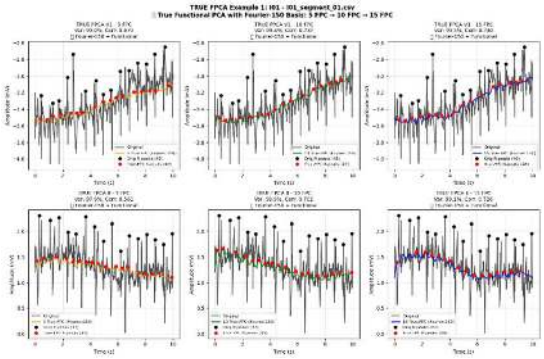


GAMBAR 15
Gambar Hasil Rekonstruksi sinyal PJK dengan basis 50



GAMBAR 16
Hasil Gambar Rekonstruksi Sinyal FPCA dengan Basis 150

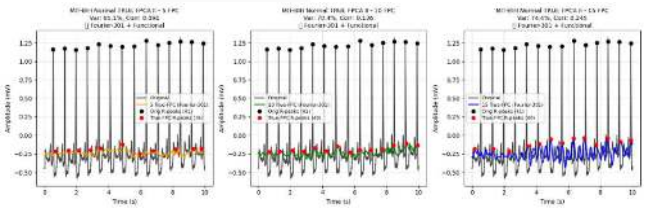
Basis 150 menunjukkan keseimbangan yang baik juga antara *variance explained* dan akurasi temporal. Meskipun *variance explained* menurun menjadi 87,43% pada 15 FPC. Penurunan FPC1 dari 71,20% menjadi 62,48% menunjukkan bahwa peningkatan resolusi basis menghasilkan dekomposisi yang lebih sedikit tajam yang memungkinkan komponen-komponen selanjutnya menangkap detail temporal yang lebih halus. Hal ini sangat penting untuk aplikasi HRV ketika analisis yang memerlukan presisi tinggi dalam temporal *feature extraction*.



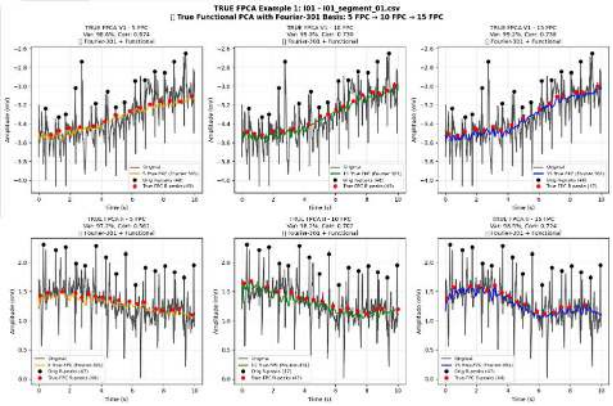
GAMBAR 17
Gambar Hasil Rekonstruksi sinyal PJK dengan basis 150

Basis 150 menunjukkan perbaikan dalam akurasi temporal untuk sinyal PJK. Konsentrasi variance pada FPC1 tetap sangat tinggi (V1: 95,04%, II: 91,37%), menunjukkan bahwa karakteristik patologis PJK konsisten dengan basis yang berbeda.

Konfigurasi basis 301 digunakan untuk dapat memberi representasi frekuensi yang jauh lebih baik jika dibandingkan dengan penggunaan 50 dab 150 basis fourier. Basis 301 menunjukkan distribusi variance yang paling merata dengan FPC1 hanya menjelaskan 51,81% variance. Meskipun total variance explained terendah (74,37% pada 15 FPC), konfigurasi ini memberikan korelasi tertinggi (0,354). Peningkatan korelasi menjadi 0,354 menunjukkan bahwa meskipun overall variance explained menurun, kualitas representasi temporal meningkat signifikan.



GAMBAR 18
Hasil Gambar Rekonstruksi Sinyal FPCA dengan Basis 301



GAMBAR 19
Gambar Hasil Rekonstruksi sinyal PJK dengan basis 150

Basis 301 memberikan representasi optimal untuk sinyal PJK, tetapi dengan pola FPC yang sedikit menurun (V1: 94,65%, II: 90,75%) mengindikasikan bahwa basis mampu menangkap detail yang tersebar dalam seluruh

komponen, memberikan representasi yang lebih jelas dari karakteristik sinyal PJK.

Perbandingan antara dua gambar sebelumnya (rekonstruksi sinyal PJK dengan basis 150) dan gambar terbaru (rekonstruksi sinyal normal dengan basis 301) menunjukkan perbedaan signifikan dalam kualitas rekonstruksi, morfologi sinyal, serta jumlah *R-peak* yang berhasil dideteksi.

Pada gambar PJK (basis 150), meskipun variansi yang ditangkap cukup tinggi (sekitar 97–99%). Hal ini tercermin dari bentuk sinyal hasil rekonstruksi yang belum sepenuhnya menyerupai morfologi QRS kompleks secara konsisten, terutama pada lead V1. R-peak yang terdeteksi juga bervariasi dan terkadang tidak sepenuhnya sejajar dengan sinyal asli, menandakan rekonstruksi yang kurang presisi.

Sebaliknya, pada gambar sinyal normal (basis 301), rekonstruksi sinyal dengan 5, 10, dan 15 FPC menghasilkan variansi yang tinggi (hingga 99,5%), bahkan dengan hanya 5 FPC. Morfologi QRS kompleks tampak lebih jelas dan menyerupai sinyal asli, terutama pada lead V1. Selain itu, deteksi R-peak juga lebih akurat dan konsisten dengan sinyal asli, menunjukkan keandalan rekonstruksi.

Secara umum, sinyal dari dataset normal menunjukkan respons yang lebih baik terhadap rekonstruksi FPCA dengan basis dan jumlah FPC yang sama atau bahkan lebih rendah dibandingkan sinyal PJK. Ini mengindikasikan bahwa sinyal dari pasien normal lebih mudah direkonstruksi secara fungsional karena variabilitasnya yang lebih rendah dan bentuk QRS yang lebih konsisten, sedangkan sinyal dari pasien PJK menunjukkan kompleksitas morfologi dan variasi temporal yang lebih besar, sehingga membutuhkan basis dan FPC yang lebih tinggi untuk mencapai hasil serupa.

C. Analisis Distribusi Fitur

Analisis distribusi data HRV dilakukan menggunakan uji Kolmogorov-Smirnov (KS Test) untuk mengevaluasi perbedaan distribusi antara subjek normal dan subjek dengan PJK pada berbagai metode rekonstruksi sinyal. Uji ini membantu mengidentifikasi fitur HRV yang memiliki kemampuan diskriminasi tinggi antara kedua kelompok.

Analisis Distribusi HRV pada Rekonstruksi PCA

Hasil KS test pada Lead VI menunjukkan bahwa seluruh fitur HRV memiliki distribusi yang berbeda signifikan antara subjek normal dan PJK (18/18 fitur). Fitur dengan nilai KS Score tertinggi pada 5PC adalah std_rr_ms dan sdnn (0,829), hr_range (0,781), dan min_hr (0,767), menunjukkan kemampuan diskriminasi yang baik.

Lead MLII menunjukkan pola yang sangat baik dengan hampir seluruh fitur HRV memiliki perbedaan distribusi yang signifikan (17-18/18 fitur). Fitur dengan diskriminasi tertinggi pada 5PC adalah rr_count (0,938), mean_hr (0,902), dan mean_rr_ms (0,888). Satu-satunya fitur yang menunjukkan distribusi serupa adalah cv_rr pada 5PC (p=0,150).

Analisis Distribusi HRV pada Rekonstruksi FPCA

FPCA dengan basis 50 menunjukkan pola yang menarik dimana beberapa fitur mulai menunjukkan distribusi yang serupa (*same*), terutama fitur tinn yang konsisten menunjukkan p-value > 0,05 pada kedua lead. Lead VI umumnya menunjukkan persentase fitur signifikan yang lebih tinggi (83,3-94,4%) dibandingkan Lead MLII (77,8-83,3%).

FPCA dengan basis 150 menunjukkan peningkatan signifikan dalam kemampuan diskriminasi dengan hampir semua fitur (94,4-100%) menunjukkan perbedaan distribusi yang signifikan. Lead MLII konsisten menunjukkan performa superior dengan nilai KS Score yang tinggi terutama untuk mean_hr, mean_rr_ms, dan rr_count.

FPCA dengan basis 301 menunjukkan pola yang konsisten dengan basis yang lebih rendah, dimana Lead MLII umumnya menunjukkan nilai KS Score yang lebih tinggi dibandingkan Lead VI. Semua konfigurasi menunjukkan tingkat signifikansi yang tinggi (94,4-100%) dengan fitur tinn menjadi satu-satunya fitur yang sesekali menunjukkan distribusi serupa.

D. Pemeringkatan Fitur

Berdasarkan analisis ranking fitur ANOVA pada kedua lead (Lead VI dan Lead MLII) dengan berbagai konfigurasi PCA dan FPCA, dapat diidentifikasi beberapa fitur yang secara konsisten muncul pada posisi tinggi (peringkat 1-5) di berbagai konfigurasi.

Fitur yang Konsisten Muncul di Top 5:

1. *signal_Autocorrelation* : Fitur ini konsisten muncul di peringkat tinggi pada hampir semua konfigurasi, menunjukkan pentingnya pola korelasi temporal dalam sinyal EKG untuk membedakan kondisi normal dan PJK.
2. *signal_Spectral distance* : Fitur spectral distance menunjukkan stabilitas tinggi, mengindikasikan bahwa karakteristik spektral memberikan informasi diskriminatif yang kuat.
3. *signal_Root mean square* : Fitur ini mencerminkan energi rata-rata sinyal dan konsisten memberikan kontribusi signifikan dalam klasifikasi.
4. *signal_Area under the curve* : Menggambarkan area total di bawah kurva sinyal, fitur ini stabil muncul di berbagai konfigurasi yang menunjukkan relevansinya dalam analisis morfologi sinyal.
5. *signal_Positive turning points* : Fitur yang menghitung jumlah titik balik positif dalam sinyal, memberikan informasi tentang variabilitas dan kompleksitas sinyal jantung.

E. Klasifikasi

Evaluasi performa klasifikasi dilakukan menggunakan dua algoritma machine learning utama: *Support Vector Machine* (SVM) dan *Logistic Regression*. Analisis klasifikasi dibagi menjadi tiga kategori utama berdasarkan jenis fitur yang digunakan: TSFEL, HRV, dan kombinasi HRV+TSFEL. Setiap kategori dievaluasi pada hasil rekonstruksi PCA dan FPCA dengan berbagai konfigurasi komponen untuk mengidentifikasi kombinasi optimal dalam membedakan subjek normal dan PJK. Untuk pembagian dari train dan test dari klasifikasi ini menggunakan perbandingan 70:30, untuk masing masing *machine learning*

Implementasi fitur TSFEL pada hasil rekonstruksi PCA menunjukkan performa klasifikasi yang sangat superior dengan mencapai akurasi perfect pada hampir semua konfigurasi. Ekstraksi fitur TSFEL dilakukan pada sinyal yang telah direkonstruksi menggunakan 5, 10, dan 15 principal components untuk mengevaluasi pengaruh tingkat rekonstruksi terhadap kualitas klasifikasi.

Evaluasi komprehensif seluruh metode klasifikasi mengungkap *hierarchy* performa yang baik dan arah yang

jelas untuk analisis dalam membedakan yang normal dan PJK. Analisis ini memberikan arahan dala, pemilihan metode optimal berdasarkan trade-off antara akurasi, interpretabitas, dan yang lain.

TABEL 1
Ranking Performa Metode Klasifikasi

Rank	Method	Configuration	Accuracy	F1 Score	ROC AUC
1	TSFEL-PCA	15 PC	100%	100%	100%
2	TSFEL-FPCA	Basis 50, 15 FPC	100%	100%	100%
3	TSFEL-FPCA	Basis 301, 15 FPC	100%	100%	100%
4	HRV+TSFEL-FPCA	Basis 150, 15 FPC	99,6%	99,6%	99,99%
5	HRV+TSFEL-FPCA	Basis 301, 10 FPC	97,96%	97,94%	99,71%
6	HRV-FPCA	Basis 301, 15 FPC	99,21%	99,21%	99,99%
7	HRV-PCA	10 PC	98,33%	98,35%	99,96%
8	HRV+TSFEL-PCA	5 PC	96,72%	96,67%	99,49%

Fitur TSFEL menunjukkan performa klasifikasi terbaik dengan PCA dan FPCA nya mencapai akurasi (100%), terutama pada rekonstruksi PCA dengan 15 PC. Hal ini mengindikasikan bahwa TSFEL mampu menangkap pola yang tidak terdeteksi oleh HRV analisis tradisional.

Kombinasi HRV+TSFEL pada FPCA basis 150 memberikan akurasi yanh hampir sempurna (99.60%) dengan tambahan berupa fitur HRV. Konfigurasi ini merepresentasikan keseimbangan antara akurasi dan penjelasan yang cukup untuk membedakan pasien normal dan PJK berdasarkan fitur fitur dari HRV dan TSFEL hasil dari rekonstruksi sinyal menggunakan PCA dan FPCA

Fitur HRV menunjukkan sensitivitas tinggi terhadap jumlah komponen (dari 78% pada 5 PC menjadi 98% pada 10-15 PC), sementara fitur TSFEL lebih robust dalam berbagai konfigurasi komponen.

FPCA basis 150-301 memberikan hasil yang baik untuk fitur HRV, sementara PCA dengan 15 PC optimal untuk fitur TSFEL karena *noise reduction*.

Untuk penggunaan HRV+TSFEL-FPCA basis 150 dengan 15 FPC direkomendasikan karena kombinasi akurasi yang hampir 100% (99,60%) dari fitur HRV, dan komputasi yang bisa membedakan pasien normal dan PJK.

Hasil ini menunjukkan bahwa *signal processing* (FPCA/PCA) yang dikombinasikam dengan ekstraksi fitur (TSFEL+HRV) dapat mencapai diagnosa *performance* yang setara atau dibandingkan dengan metode dalam menentukan pasien PJK dengan metode biasa, membuka peluang untuk membuat inovasi menggunakan FPCA/PCA yang akurasinya tinggi dalam membedakan pasien PJK dan normal, tidak berdasarkan klinis, tetapi dari hasil fiturnya dan analisis distribusi yang bisa membedakan.

V. KESIMPULAN

Berdasarkan hasil penelitian dan analisis yang telah dilakukan mengenai penerapan *Non-Linear Functional Principal Component Analysis* pada sinyal HRV untuk deteksi awal penyakit jantung koroner, dapat disimpulkan bahwa analisis distribusi menggunakan uji *Kolmogorov-Smirnov* menunjukkan bahwa seluruh fitur HRV memiliki distribusi yang berbeda secara signifikan ($p<0.001$) antara pasien normal dan pasien dengan Penyakit Jantung Koroner (PJK). *Lead* MLII memiliki nilai KS score yang lebih tinggi jika dibandingkan Lead VI, terutama dengan fitur-fitur HRV seperti *mean_rr_ms*, *mean_hr*, dan *rr_count* memiliki nilai KS *Score* tertinggi. FPCA dengan basis yang lebih tinggi (301) memberikan kemampuan pemisahan kelompok yang lebih baik dibandingkan basis rendah (50), mengindikasikan bahwa resolusi temporal yang tinggi berperan penting dalam analisis HRV.

Selain itu, evaluasi model menggunakan SVM dan *Logistic Regression* menunjukkan bahwa fitur TSFEL memberikan hasil klasifikasi yang paling baik jika dibandingkan dengan HRV ataupun HRV+TSFEL dengan akurasi sebesar 100% pada rekonstruksi PCA dengan 15 PC dan FPCA dengan berbagai konfigurasi basis. Kombinasi HRV+TSFEL pada FPCA basis 150 dengan 15 FPC mengasilkan akurasi sebanyak 96.6%, yang mencerminkan keseimbangan antara kinerja dari klasifikasi. Perfoma fitur HRV juga meningkat secara bertahap seiring peningkatan jumlah komponen, dari 77–83% (5 PC) menjadi 96–98% (10–15 PC).

Kemudia, untuk analisis dari hasil ranking fitur dengan metode ANOVA menunjukkan bahwa fitur HRV seperti mean_hr, cv_rr, mean_rr_ms, dan sdnn memiliki kontribusi signifikan. Di sisi lain, fitur TSFEL seperti signal_Autocorrelation, signal_Spectral distance, dan signal_Positive turning points menunjukkan konsistensi dalam kontribusi pemisahan kelompok. Penggabungan kedua jenis fitur memberikan informasi yang saling melengkapi dan meningkatkan akurasi klasifikasi.

REFERENSI

[1] World Health Organization, “Cardiovascular diseases (CVDs),” who.int,.

[2] P. C. A. and H. S. W. S. M. Salerno, “Competency in Interpretation of 12-Lead Electrocardiograms: A Summary and Appraisal of Published Evidence,” *Ann. Intern. Med.*, vol. 138, no. 9, pp. 751–760, 2003.

[3] P. M. L. D. S. O. I. M. and L. P. R. Castaldo, “The Assessment of the Autonomic Nervous System Functioning by Heart Rate Variability Analysis: A Methodological Review,” *Sensors*, vol. 21, no. 5, p. 1824, 2021.

[4] S. R.-A. dan J. F. G. R. S. A. Imtiaz, “A Comprehensive Survey on Heart Rate Variability (HRV) Analysis and its Application for the Detection of Cardiac Abnormalities,” *IEEE Access*, vol. 9, pp. 119859–119883, 2021.

[5] G. and H. J. E. Hall dan M. E. Hall, *Textbook of Medical Physiology, ed ke-14*. Philadelphia: Elsevier, 2020.

[6] Alodokter, “Penyakit Jantung Koroner – Gejala, Penyebab, dan Pengobatan,” https://www.alodokter.com/penyakit-jantung-koroner.

[7] American Heart Association, “Electrocardiogram (ECG or EKG),” https://www.heart.org/en/health-topics/heart-attack/diagnosing-a-heart-attack/electrocardiogram-ecg-or-ekg.

[8] P. G. A. B. A. M. H. R. T. and V. V. H. D. Jahed Armaghani, "Examining hybrid and single SVM models with different kernels to predict rock brittleness," *Sustainability*, vol. 12, no. 6, p. 2229, 2020.

[9] H. T. R. and A. S. F. T. Cao, "Functional principal component analysis and sparse-group LASSO to identify associations between biomarker trajectories and mortality among hospitalized SARS-CoV-2 infected individuals," *BMC Med Res Methodol*, vol. 23, no. 2, 2023.