

Analisis Teknik *Explainable AI Post-hoc* untuk Deteksi Infark Miokard Berbasis Sinyal EKG

1st Fathiya Nur Assyifa
Fakultas Teknik Elektro
Universitas Telkom

Kota Bandung, Indonesia
fthyna@student.telkomuniversity.ac.id

2nd Tito Waluyo Purboyo
Fakultas Teknik Elektro
Universitas Telkom

Kota Bandung, Indonesia
titowaluyo@telkomuniversity.ac.id

3rd Dziban Naufal
Fakultas Teknik Elektro
Universitas Telkom

Kota Bandung, Indonesia
dziban@telkomuniversity.ac.id

Abstrak—Model deep learning (DL) menunjukkan akurasi yang unggul dalam menganalisis sinyal EKG untuk mendeteksi infark miokard. Namun, pendekatan ini sering dianggap sebagai "black box" yang sulit dipahami oleh tenaga medis. Penelitian ini mengevaluasi penerapan explainable AI (XAI), yakni Grad-CAM dan LRP, dalam meningkatkan interpretabilitas model DL berbeda untuk deteksi infark miokard menggunakan sinyal EKG. Pada studi ini, XAI berupa Grad-CAM dan LRP diaplikasikan pada tiga jenis model konvolusional, yakni model konvolusional biasa, model residual, serta model inception. Hasil menunjukkan bahwa Grad-CAM memberikan penjelasan spasial yang konsisten dengan atribusi positif serta lebih sederhana, sementara LRP dapat memberikan atribusi positif maupun negatif, membedakan relevansi antar-lead, serta tidak tergantung pada resolusi spasial dari layer internal model. Kemudian, kombinasi Grad-CAM untuk analisis temporal dan LRP untuk analisis relevansi lead memberikan interpretasi model yang paling komprehensif dan direkomendasikan untuk evaluasi relevansi klinis model DL. Disimpulkan bahwa arsitektur InceptionTime ditemukan merupakan model terbaik, dengan akurasi tertinggi serta utilisasi lead tertinggi berdasarkan analisis XAI.

Kata kunci—deep learning, elektrokardiogram, infark miokard, XAI

I. PENDAHULUAN

Penyakit jantung iskemik merupakan penyebab kematian nomor 2 di Indonesia setelah stroke dengan angka kematian 14,38% [1]. Di antara kematian penyakit jantung iskemik, 33% merupakan akibat serangan jantung atau infark miokard [2]. Infark miokard merupakan kondisi gawat darurat yang membutuhkan penanganan cepat [3]. Penanganan terlambat dapat menyebabkan kerusakan permanen pada jantung dan menyebabkan komplikasi berbahaya yang memiliki tingkat mortalitas tinggi [4].

EKG merupakan salah satu modalitas yang penting untuk diagnosis infark miokard, dan perlu diambil secepat mungkin pada pasien dengan keluhan sakit dada [4] [5] [6]. Interpretasi EKG dini dapat menentukan ketepatan penanganan serta prognosis pasien. Belakangan, metode komputasional untuk menganalisis EKG berbasis akal imitasi (*artificial intelligence*, AI), terutama arsitektur pembelajaran dalam (*deep learning*, DL), berkembang pesat dan memperoleh tingkat keberhasilan yang tinggi [3] [7] [8], bahkan lebih baik dari analisis kardiologi [9].

Namun, model *deep learning* umumnya dianggap sebagai *black box*, terutama dalam aplikasi klinis, karena mekanisme internal dan basis keluaran model tidak eksplisit, dan sulit dipercayai oleh tenaga kesehatan [10]. Interpretabilitas algoritma *machine learning* memiliki urgensi tinggi untuk dapat menjawab permasalahan etik dan pertanggungjawaban

apabila model membuat kesalahan dalam pengambilan keputusan [11]. Dengan demikian, *explainable artificial intelligence* (XAI) telah diajukan sebagai solusi potensial untuk membuat model AI lebih transparan dan dapat dipahami [11]. Tugas akhir ini bermaksud berkontribusi pada pengembangan XAI dengan meninjau penerapannya untuk deteksi infark miokard pada sinyal EKG.

II. KAJIAN TEORI

A. Infark miokard

Infark miokard adalah kematian sel jantung akibat tersumbatnya arteri koroner yang menyuplai nutrisi ke sel tersebut. Diagnosis dan lokalisasi infark miokard dapat dilakukan melalui EKG 12-lead. Infark dinding anterior mengganggu progresi gelombang R dan menghasilkan gelombang Q pada lead lateral V1–V6. Infark lateral tinggi terlokalisasi pada lead frontal I dan aVL. Infark inferior menunjukkan elevasi ST dan gelombang Q pada lead II, III, dan aVF. Infark posterior diindikasikan oleh tinggi gelombang R yang bersifat resiprokal dan depresi ST pada lead V1–V2 (dapat juga terlihat langsung melalui lead tambahan V7–V9); serta infark ventrikel kanan muncul pada V1 dan sadapan kanan V3R–V6R [16].

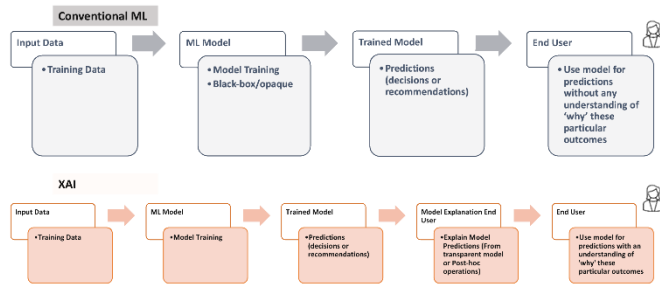
B. Deep Learning untuk Analisis EKG

Deep learning (DL) adalah metode yang *robust* dan fleksibel untuk berbagai macam analisis data, termasuk sinyal EKG. Proses penerapannya terdiri dari pra-pemrosesan, *training*, dan evaluasi [3] [18] [19] [21] [22]. Arsitektur DL yang dapat diaplikasikan umumnya jaringan satu dimensi berbasis konvolusi serta rekurensi [3], [20] [20] beserta beragam variasinya: *residual network* (ResNet) [17], jaringan *inception* [23] [24], *long short-term memory* (LSTM) [23] [20], mekanisme perhatian (*attention*) [25] [26], dan pendekatan *ensemble* [22] [27] serta *multi-branch* [28] [29].

C. Explainable AI

Explainable artificial intelligence (XAI) merupakan pendekatan yang menekankan interpretabilitas dari model DL. Survei [11] memberikan kriteria bahwa suatu studi dikatakan meninjau interpretabilitas dari AI jika terdapat usaha untuk 1) menjelaskan alasan model mengambil suatu keputusan; 2) menunjukkan pola tertentu pada mekanisme internal suatu algoritma; dan 3) mempresentasikan sistem tersebut menggunakan model atau matematika yang koheren. Dengan kata lain, XAI merupakan ekspansi dari ML di mana terdapat langkah yang memaparkan alasan model membuat prediksi tertentu (**Gambar 1**) [36]. Pada eksperimen ini, metode XAI

yang dipilih adalah Grad-CAM serta *layer-wise relevance propagation* (LRP).



Gambar 1 Perbedaan pipeline dari ML konvensional dan XAI [36].

1) Grad-CAM

Grad-CAM (*Gradient-weighted Class Activation Mapping*) [31] merupakan sebuah teknik visualisasi yang memanfaatkan informasi gradien yang mengalir menuju lapisan konvolusional terakhir dari CNN untuk membangun peta lokasi kasar yang menunjukkan area-area penting pada citra masukan. Dengan menggunakan rata-rata gradien output kelas y^c terhadap *feature map* A^k pada lapisan konvolusional terakhir

Untuk suatu masukan X , CNN menghasilkan skor logit y^c untuk kelas target c dan menghitung gradien terhadap skor tersebut. Misalkan A^k adalah *feature map* ke- k dari lapisan konvolusional terakhir. Grad-CAM menghitung gradien y^c terhadap *feature map* $\frac{\partial y^c}{\partial A_{ij}^k}$, di mana merepresentasikan aktivasi pada posisi spasial (i, j) dari *feature map* ke- k . Peta gradien dari masing-masing *feature map* selanjutnya dikompresi menjadi skalar a_k^c yang menggambarkan kontribusi keseluruhan *feature map* tersebut terhadap kelas c melalui *global average pooling* (GAP) (**Persamaan (1)**).

$$a_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (1)$$

Menggunakan nilai kontribusi tersebut, *heatmap* Grad-CAM dihitung dengan memboboti *feature map* dengan kontribusi masing-masing (**Persamaan (2)**).

$$L_{\text{Grad-CAM}}^c = \text{ReLU} \left(\sum_k a_k^c A^k \right) \quad (2)$$

2) Layer-wise Relevance Propagation

Layer-Wise Relevance Propagation (LRP) [40] [46] merupakan teknik interpretabilitas untuk jaringan saraf yang mengatribusikan prediksi model kembali ke fitur-fitur input dengan mempropagasikan "skor relevansi" melalui jaringan secara berurutan dalam urutan terbalik dari komputasi maju. LRP mendekomposisi skor output akhir y^c menjadi kontribusi dari setiap dimensi input, dengan menerapkan prinsip konservasi relevansi sehingga $y^c = \sum R_i$, di mana R_i adalah relevansi yang diberikan kepada titik input x_i .

Beberapa varian aturan LRP telah diajukan untuk menangani berbagai jenis lapisan dan karakteristik aktivasi. Dalam praktiknya, kombinasi beberapa aturan ini sering

diterapkan pada lapisan-lapisan dari sebuah model tunggal untuk menghasilkan penjelasan yang lebih akurat dan dapat ditafsirkan. Ringkasan aturan umum dan rekomendasi penggunaannya disajikan pada **Tabel 1**.

Tabel 1 Aturan-aturan LRP beserta pemakaianya [46].

Aturan	Persamaan	Pemakaian
LRP-0	$R_j = \sum_k \frac{a_j w_{jk}}{\sum_{0,j} a_j w_{jk}} R_k$	Layer akhir
LRP- ϵ	$R_j = \sum_k \frac{a_j w_{jk}}{\epsilon + \sum_{0,j} a_j w_{jk}} R_k$	Layer tengah
LRP- γ	$R_j = \sum_k \frac{a_j (w_{jk} + \gamma w_{jk}^+)}{\sum_{0,j} a_j (w_{jk} + \gamma w_{jk}^+)} R_k$	Layer awal
LRP- $\alpha\beta$	$R_j = \sum_k \left(\alpha \frac{(a_j w_{jk})^+}{\sum_{0,j} (a_j w_{jk})^+} - \beta \frac{(a_j w_{jk})^-}{\sum_{0,j} (a_j w_{jk})^-} \right) R_k$	Layer awal
flat	$R_j = \sum_k \frac{1}{\sum_j 1} R_k$	Layer awal
w^2 -rule	$R_i = \sum_j \frac{w_{ij}^2}{\sum_i w_{ij}^2} R_j$	Layer pertama ($\mathbb{R}^d$)
z^b -rule	$R_i = \sum_j \frac{x_i w_{ij} - l_i w_{ij}^+ - h_i w_{ij}^-}{\sum_i x_i w_{ij} - l_i w_{ij}^+ - h_i w_{ij}^-} R_j$	Layer pertama (piksel/nilai bounded)

D. Analisis XAI

Digunakan dua strategi evaluasi untuk menganalisis hasil penjelasan yang diberikan oleh XAI. Strategi pertama berfokus pada metrik fidelitas kuantitatif dengan uji *insertion/deletion*. Strategi kedua menekankan kesesuaian terhadap literatur klinis melalui analisis atribusi kualitatif (LRP dan Grad-CAM) dan *lead relevance scoring* (LRP).

1) Tes insertion/deletion

Tes *insertion* dan *deletion* adalah metode berbasis perturbasi untuk mengevaluasi perubahan kinerja model ketika fitur input dihapus atau ditambahkan kembali berdasarkan urutan relevansinya [47], [48]. Pada tes *deletion*, fitur input dihapus secara bertahap, dan penurunan nilai output kelas target digambarkan sebagai kurva *deletion*. Penjelasan yang baik menghasilkan penurunan awal yang tajam, dengan *area under the curve* (AUC) yang lebih rendah menandakan penjelasan yang lebih baik. Tes *insertion* adalah kebalikan dari *deletion*: dimulai dari input yang sepenuhnya rusak, fitur dipulihkan bertahap, dan kenaikan output di-plot sebagai kurva *insertion*. Penjelasan yang baik menunjukkan kenaikan awal yang cepat dengan AUC yang lebih tinggi.

2) Analisis atribusi kualitatif

Analisis atribusi kualitatif dilakukan dengan memeriksa secara visual *heatmap* yang dihasilkan oleh Grad-CAM dan LRP, dengan berfokus pada segmen gelombang ECG (seperti gelombang P, kompleks QRS, dll.) yang memperoleh nilai atribusi tertinggi [49]. Analisis ini memungkinkan penilaian apakah daerah EKG yang disorot selaras dengan peristiwa infark miokard dan dapat digunakan untuk menjelaskan kasus *outcome* tertentu (*false positive* dll.). Pendekatan ini melengkapi metrik kuantitatif dengan memberikan

pemahaman intuitif mengenai bagaimana model mendistribusikan perhatian sepanjang struktur temporal ECG.

3) Lead relevance score

Lead relevance score adalah metrik interpretabilitas global yang dirancang untuk mengevaluasi bagaimana suatu metode XAI mendistribusikan relevansi pada 12 *lead* standar ECG [49]. Pendekatan ini mengagregasikan nilai atribusi sepanjang dimensi temporal setiap *lead* dan menormalkannya terhadap total atribusi absolut dari seluruh sinyal, menghasilkan 12 skalar yang mencerminkan tingkat kepentingan relatif yang diberikan model pada masing-masing *lead* ketika memprediksi suatu kelas tertentu, sehingga diperoleh profil atribusi global yang menyoroti *lead* mana yang paling berpengaruh dalam proses pengambilan keputusan model.

III. PERANCANGAN SISTEM

Studi ini menggunakan *database* PTB-XL [50], yang mencakup 21.837 rekaman EKG 12-*lead* dari 18.885 pasien. Basis data tersebut merupakan basis data multilabel yang dianotasi oleh dua orang kardiolog menggunakan kode sesuai dengan standar SCP-ECG. Tiap rekaman EKG memiliki durasi 10 detik dan tersedia dua opsi untuk laju pencuplikan, yakni 500 Hz dan 100 Hz.

Tabel 2 Distribusi data kelas induk PTB-XL.

Kelas induk	Jml. EKG	% dataset
normal (NORM)	9.528	44,46%
myocardial infarction (MI)	5.486	25,60%
ST/T change (STTC)	5.250	24,50%
conduction disturbance (CD)	4.907	22,90%
hypertrophy (HYP)	2.655	12,39%

Dataset terdiri menjadi lima kelas diagnostik induk: normal, gangguan hantaran (*conduction disturbance*), infark miokard (*myocardial infarction*), hipertropi (*hypertrophy*), serta perubahan ST/T (*ST/T change*) (**Tabel 2**). Karena patologi yang tumpang tindih cukup umum, semua model akan dilatih pada tugas klasifikasi multilabel, mencakup lima kelas induk diagnostik, supaya mencerminkan kompleksitas dan komorbiditas dalam interpretasi EKG klinis.

Data selection menyeleksi rekaman ECG menggunakan kode SCP untuk menyaring rekaman yang tidak memiliki label diagnostik. Kode SCP masing-masing rekaman kemudian dipetakan ke lima kelas diagnostik induk sebagai ruang label.

Tabel 3 Pembagian data latih, validasi, dan uji. Data tanpa label diagnostik tidak dimasukkan, sehingga total data lebih sedikit dari ukuran *database* PTB-XL.

Himpunan data	Jml. rekaman	%
Data latih	17.111	79,85%
Data validasi	2.156	10,06%
Data uji	2.163	10,09%
Total	21.430	100%

Data kemudian dibagi menjadi data latih, validasi, serta data uji dengan rasio sekitar 80:10:10 (**Tabel 3**). Pembagian data tersebut berdasarkan stratifikasi rekomendasi PTB-XL. Distribusi label pada masing-masing himpunan terdapat pada **Tabel 4**.

Tabel 4 Distribusi label pada masing-masing subhimpunan data.

Label	Data latih	%	Data val.	%	Data uji	%
NORM	7.607	44,46%	957	44,39%	964	44,57%
MI	4.389	25,65%	544	25,23%	553	25,57%
STTC	4.193	24,50%	534	24,76%	523	24,18%
CD	3.912	22,86%	497	23,05%	498	23,02%
HYP	2.121	12,40%	271	12,57%	263	12,16%

Tahap pra-pemrosesan sinyal terdiri dari *denoising* dan normalisasi mengikuti skema pada referensi [17]. Digunakan *low-pass filter* (LPF) dengan frekuensi *cutoff* 0,5 Hz untuk mengisolasi *baseline drift*, yang kemudian dikurangi dari sinyal untuk mendapatkan sinyal dengan *baseline* nol. Selanjutnya, LPF dengan frekuensi *cutoff* 40 Hz diaplikasikan untuk menghilangkan derau frekuensi tinggi. Normalisasi dilakukan secara global mengikuti referensi [23], dengan menggeser rata-rata sinyal ke nol dan membagi dengan deviasi standar.

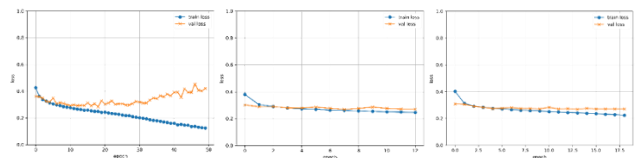
Tabel 5 merangkum model DL yang digunakan beserta referensi implementasinya. Masukan semua model berukuran (B, 12, 1000) dan keluaran berukuran (B, 5) di mana B = *batch size* = 32. Model dilatih menggunakan skema multilabel, dengan optimisasi AdamW, *learning rate* sebesar 1×10^{-3} , dan *weight decay* sebesar 0.01. *Loss function* yang digunakan adalah *binary cross-entropy*. Pelatihan dilakukan selama maksimal 50 *epoch*, serta *early stopping* untuk ResNet dan Inception apabila akurasi validasi tidak meningkat selama lima *epoch* berturut-turut. Model dengan *validation loss* terendah dipilih.

Tabel 5 Ringkasan model *deep learning* yang diinterpretasikan menggunakan XAI.

Referensi	Arsitektur	Jumlah parameter	Ukuran model
[51]	AlexNet-1D	3.414.661	15,50 MB
[23]	ResNet	473.349	18,22 MB
[23], [24]	InceptionTime	507.653	40,66 MB

IV. HASIL DAN ANALISIS

Training loss dan *validation loss* selama proses *training* dapat dilihat pada **Gambar 2**. Ketiga model mencapai hasil yang memadai setelah *training*. Tampak bahwa seiring meningkatnya jumlah *epoch*, AlexNet mengalami *overfitting* berat (*validation loss* meningkat meskipun *training loss* menurun). Namun, model terbaik dengan *validation loss* terendah merupakan versi yang disimpan untuk analisis.



Gambar 2 *Training loss* dan *validation loss* pada ketiga model selama proses *training*. Kiri: AlexNet 1D, tengah: ResNet, bawah: InceptionTime

Tabel 6 menunjukkan performansi masing-masing model pada klasifikasi semua kelas pada data latih serta uji. Terlihat bahwa ketiga model mencapai akurasi di atas 85%. Tingkat *precision* cukup rendah dan tidak mencapai 70%. Tingkat *recall* ketiga model berada pada rentang 84-86%, dan skor f1

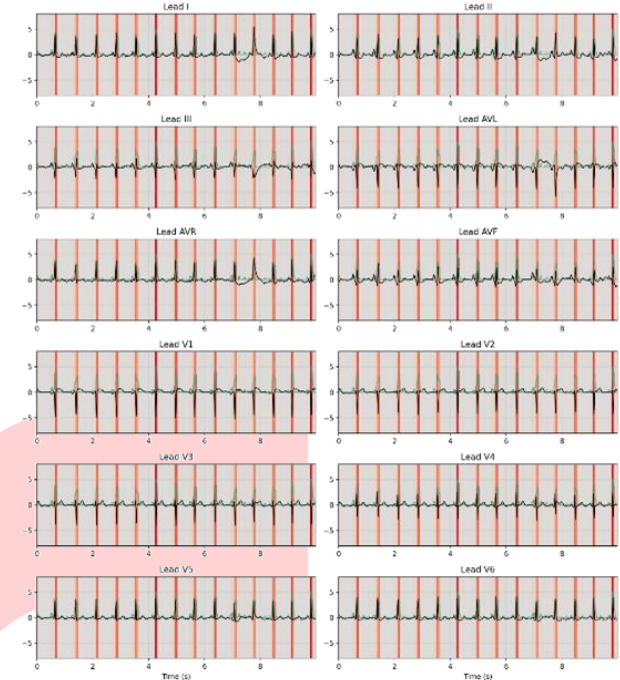
ketiga model sekitar 75%. Secara keseluruhan, arsitektur Inception memiliki akurasi, *recall*, skor f1, dan ROC AUC tertinggi.

Tabel 6 Performa ketiga model pada tugas klasifikasi semua kelas. Model terbaik dicetak tebal dan digarisbawahi. Performa model yang mendekati performa tertinggi (selisih $< 10^{-3}$) diberi cetak tebal tanpa garis bawah.

Model	AlexNet-1D		ResNet		InceptionTime	
Data	Latih	Uji	Latih	Uji	Latih	Uji
% Akurasi	88,202	85,030	86,951	85,992	89,629	86,269
% <i>Precision</i>	72,125	66,723	69,202	68,681	74,962	68,661
% <i>Recall</i>	88,927	84,184	89,668	84,398	90,199	86,433
% skor-f1	79,661	74,444	78,117	75,733	81,887	76,529
ROC AUC	0,95099	0,91019	0,94749	0,92365	0,96070	0,92781

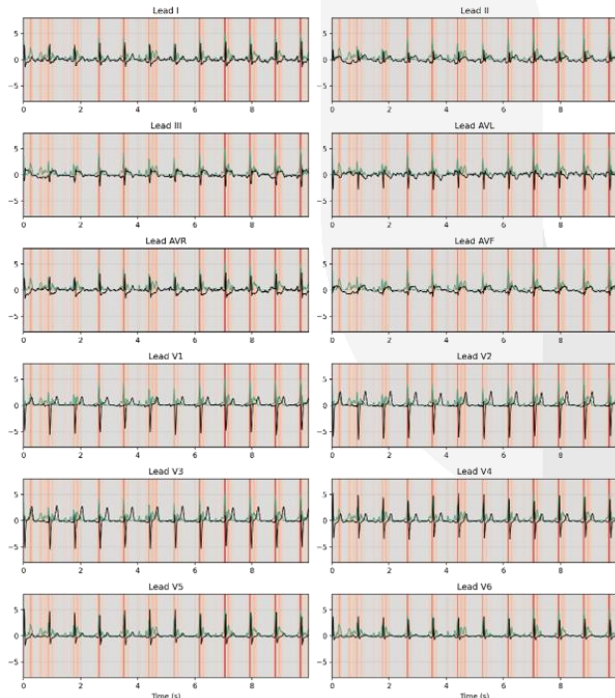
Sampel hasil output Grad-CAM pada masing-masing model dapat ditinjau pada Gambar 4 – Gambar 3. Seperti ditunjukkan pada Gambar 5, Grad-CAM tidak terlalu efektif pada AlexNet karena resolusi spasial rendah dari peta fitur akhirnya, yang tidak mampu menangkap dan membedakan fitur EKG yang lebih halus. ResNet menghasilkan heatmap dengan resolusi yang cukup tinggi untuk memungkinkan diferensiasi pola spesifik EKG secara bermakna, sementara InceptionTime mencapai resolusi tertinggi di antara ketiga arsitektur tersebut karena mempertahankan jumlah timestep EKG asli pada peta fitur akhirnya.

```
ecg_id:11605
label:tensor([1, 0, 1, 0, 0, 1])
prediction: [1.50 4.66 2.89 4.4 -1.69]
heatmap graph scale:17534.78
normalized=True
```



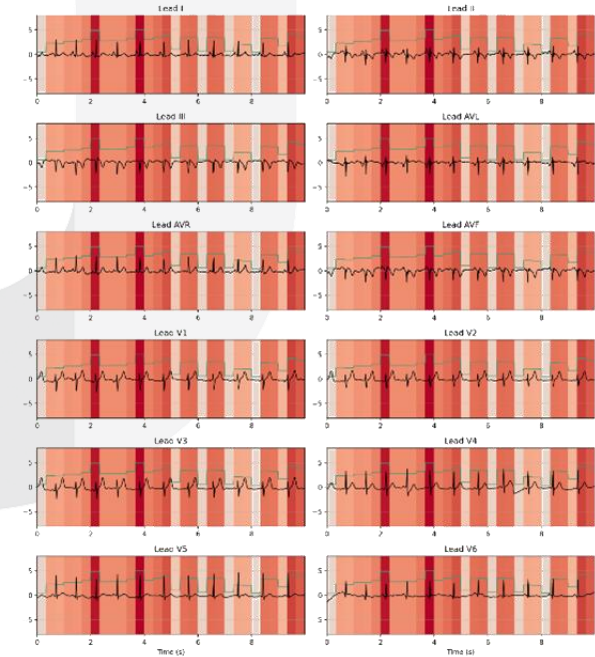
Gambar 4 Sampel Grad-CAM pada ResNet, kasus *true positive MI*

```
ecg_id:17404
label:tensor([0, 0, 1, 0, 1, 1])
prediction: [1.76 4.38 1.84 2.85 2.26]
heatmap graph scale:81022.14
normalized=True
```



Gambar 3 Sampel Grad-CAM pada InceptionTime, kasus *true positive MI*

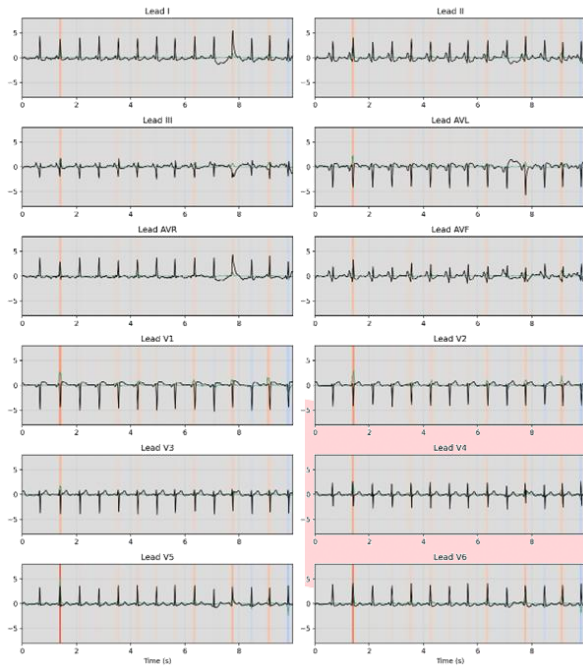
```
ecg_id:11285
label:tensor([0, 0, 1, 0, 0, 1])
prediction: [2.71 4.67 6.7 4.55 7.71]
heatmap graph scale:4511.86
normalized=True
```



Gambar 5 Sampel Grad-CAM pada AlexNet 1D, kasus *true positive MI*

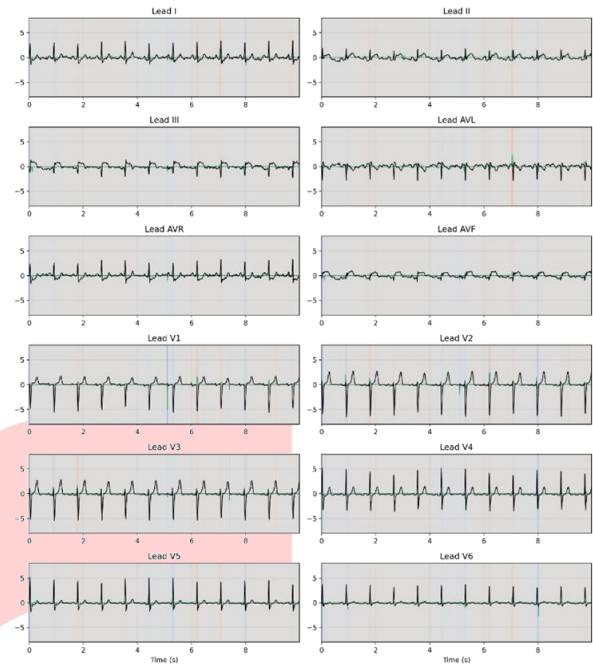
Sampel hasil output LRP dapat ditinjau pada Gambar 6 s.d. **Gambar 8**. LRP menghasilkan atribusi yang sangat jarang pada ResNet dan Inception, sedangkan atribusinya pada AlexNet cenderung berderau dan terkonsentrasi pada sebagian EKG. LRP secara eksplisit membedakan relevansi pada berbagai *lead* dan menyoroti relevansi negatif (menentang kelas target).

ecg_id:11605
label:tensor([1, 0, 1, 0, 0])
prediction:[3.36 4.66 2.89 4.4 -1.69]
heatmap_graph:scalex94.05
normalized=True



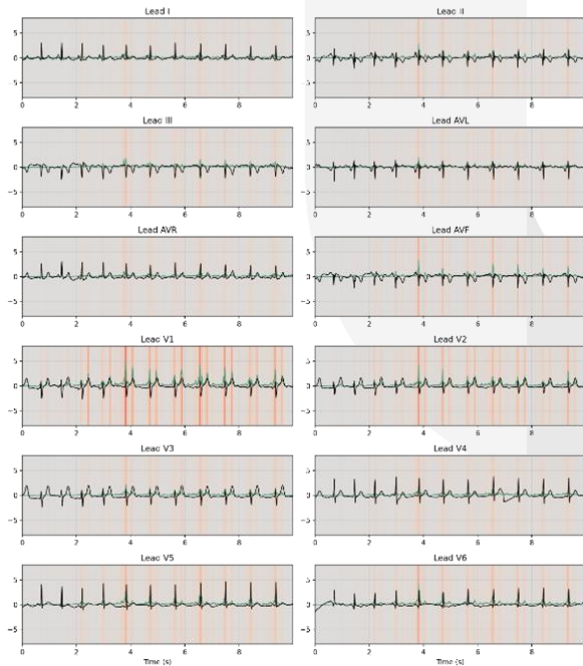
Gambar 6 Sampel LRP pada ResNet, kasus *true positive* MI

ecg_id:17404
label:tensor([0, 0, 1, 0, 1])
prediction:[1.76 4.38 1.84 2.85 2.26]
heatmap_graph:scalex24.27
normalized=True



Gambar 8 Sampel LRP pada Inception, kasus *true positive* MI

ecg_id:11285
label:tensor([0, 0, 1, 0, 0])
prediction:[2.21 4.67 6.2 5.55 2.73]
heatmap_graph:scalex62.52 85
normalized=True



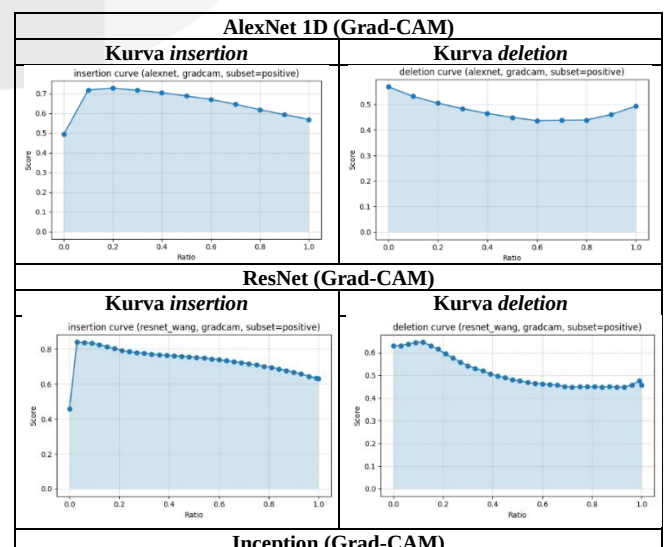
Gambar 7 Sampel LRP pada AlexNet 1D, kasus *true positive* MI

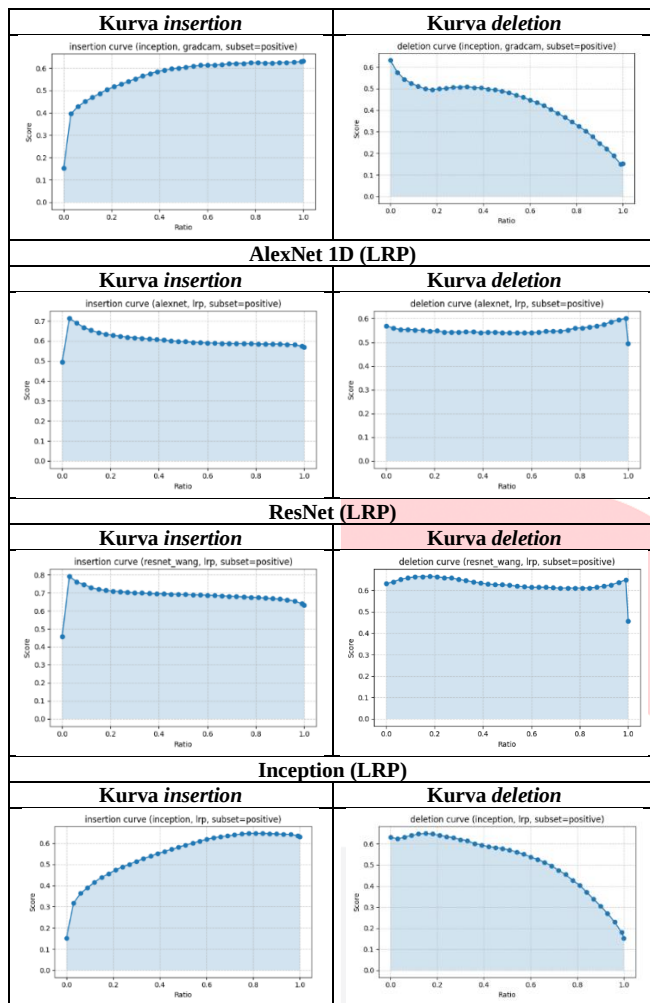
Tabel 7 menunjukkan hasil perhitungan kurva *insertion* dan *deletion* mengikuti spesifikasi yang telah dijelaskan di Bab III. Pada mayoritas kurva *insertion*, prediksi model langsung naik secara signifikan pada langkah pertama. Hal tersebut dapat mencerminkan bahwa transformasi bertahap dari *baseline* menjadi output akhir kurang tepat untuk mengevaluasi performa model pada sinyal EKG.

Satu-satunya model yang menciptakan kurva *insertion* dengan *sanity* relatif baik adalah Inception. Kurva *insertion* Inception memiliki sifat yang diinginkan yakni bentuk kurva menaik dengan relatif cepat pada awal *insertion*, sehingga dapat disimpulkan bahwa XAI yang digunakan mencerminkan atensi model dengan cukup *faithful*. Sementara itu untuk kurva *deletion*, terdapat penurunan yang relatif cepat pada awal *deletion* pada ResNet dengan Grad-CAM, serta Inception dengan Grad-CAM, menunjukkan bahwa Grad-CAM mampu memberikan relevansi secara memadai terhadap input kedua model tersebut.

Analisis atribusi kualitatif dilakukan pada ResNet dan Inception dengan Grad-CAM serta AlexNet dengan LRP. **Tabel 8 – Tabel 10** menjelaskan temuan-temuan yang diperoleh melalui analisis atribusi kualitatif.

Tabel 7 Hasil perhitungan kurva *insertion* dan *deletion*





Tabel 8 Temuan dari analisis atribusi kualitatif pada ResNet (Grad-CAM)

Kasus (MI)	Temuan pada ResNet (Grad-CAM)	Kesesuaian dengan domain klinis
True positive	Fitur paling relevan terdiri dari kompleks QRS (80%), terutama bagian akhir atau J-point (70%), serta gelombang T (60%).	Bagian akhir kompleks QRS, terutama J-point, relevan untuk diagnosis MI karena menjadi titik pengukuran elevasi- dan depresi-ST. Gelombang T juga mengalami inversi pada beberapa lead setelah terjadinya infark miokard.
True negative	Pada hampir semua kasus yang ditinjau, bagian QRS ditekankan (90%). Kecuali satu kasus CLBBB, di mana gelombang QRS memiliki bentuk yang paling berbeda, di mana bagian paling relevan terhadap MI beralih ke gelombang P (10%). Gelombang P merupakan fitur sekunder setelah QRS pada 2 kasus (20%). Gelombang T merupakan fitur sekunder pada 1 kasus (10%).	-
False positive	Kasus <i>false positive</i> terjadi pada kasus CD (40%) dan STTC (40%), serta dua kasus normal (20%). Pada kasus-kasus tersebut, model memprediksi CD dan STTC secara akurat dengan tingkat kepercayaan yang lebih tinggi dibandingkan MI. Pada satu kasus, terdapat <i>baseline drift</i> yang gagal di-filter saat pra-	Kasus CD dapat terjadi akibat penyakit jantung iskemik, sehingga komorbiditas berpotensi memengaruhi <i>output</i> model. Infark miokard juga sering menunjukkan perubahan ST-T.

	pemrosesan sehingga model gagal memprediksi STTC dengan benar (10%).	
False negative	100% kasus <i>false negative</i> memiliki komorbiditas. Paling banyak terjadi dengan kasus infark inferior (70%). Kasus lainnya terdiri dari infark anterolateral (20%) dan lateral (10%), di mana laporan PTB-XL menunjukkan <i>confidence</i> kardiolog yang relatif rendah (0,15).	Infark miokard lebih sulit untuk diidentifikasi pada EKG dengan komorbiditas.

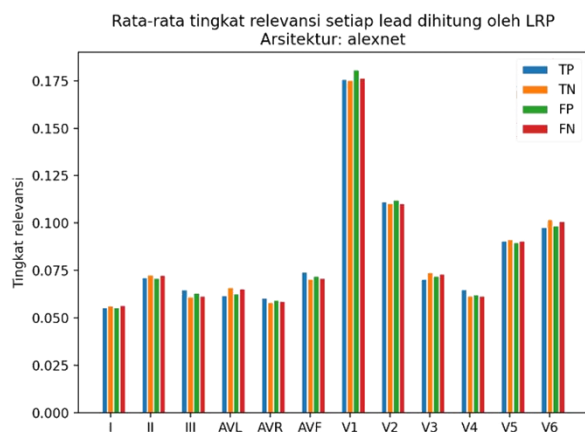
Tabel 9 Temuan dari analisis atribusi kualitatif pada Inception (Grad-CAM)

Kasus (MI)	Temuan pada Inception (Grad-CAM)	Kesesuaian dengan domain klinis
True positive	Fitur paling relevan adalah titik awal kompleks QRS atau segmen PQ (80%). Segmen ST adalah fitur sekunder pada 4 kasus (40%) dan gelombang T pada 3 kasus (30%).	Mirip dengan ResNet. Elevasi- dan depresi-ST dapat diinterpretasikan sebagai depresi- dan elevasi-PQ.
True negative	Banyak terfokus pada segmen PQ.	-
False positive	100% kasus berpotensi memiliki patologi. Mirip dengan sebelumnya, model sering memprediksi MI sebagai prediksi sekunder setelah CD. Bagian yang ditinjau terbanyak yakni segmen PQ (80%), selanjutnya segmen TP (30%). Perubahan transien ditinjau di dua kasus (20%).	Model lebih fokus kepada daerah <i>baseline</i> , yang signifikan dalam peninjauan elevasi- dan depresi-ST.
False negative	Model meninjau antara segmen PQ, kompleks QRS, dan segmen ST pada 90% kasus. Model paling umum memprediksi normal.	Seperti sebelumnya, model fokus pada daerah <i>baseline</i> , yang relatif konsisten pada klasifikasi normal.

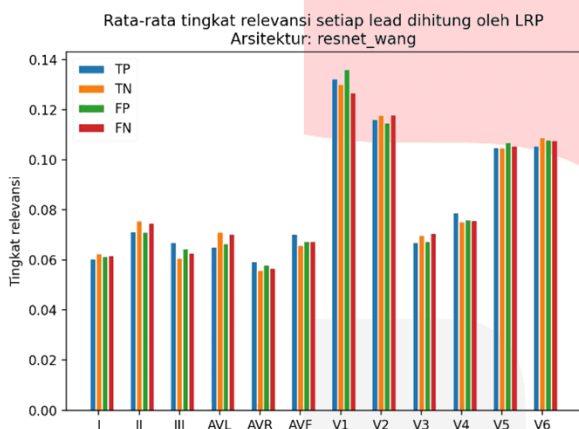
Tabel 10 Temuan dari analisis atribusi kualitatif pada AlexNet (LRP)

Kasus (MI)	Temuan pada AlexNet (LRP)	Kesesuaian dengan domain klinis
True positive	Relevansi tertinggi pada <i>lead</i> V1 (100%). J-point, segmen ST, titik mulai QRS, serta gelombang T menjadi fokus utama model.	Mirip dengan kedua model sebelumnya.
True negative	Relevansi tertinggi pada <i>lead</i> V1 (100%). Titik paling negatif terletak pada J-point serta gelombang T. Namun, titik positif tetap cukup banyak.	-
False positive	Relevansi tertinggi pada <i>lead</i> V1 (100%). Gelombang T dan kompleks QRS ditekankan secara bergantian, tidak konsisten dalam satu EKG, sulit untuk diinterpretasikan. Kasus komorbiditas dengan CD tetap muncul.	Kasus CD dapat terjadi akibat penyakit jantung iskemik, sehingga komorbiditas berpotensi memengaruhi <i>output</i> model.
False negative	Relevansi tertinggi pada <i>lead</i> V1 (100%). Model paling umum memprediksi normal dan STTC. Bagian gelombang T serta akhir kompleks QRS umumnya relevan, namun juga kurang konsisten.	Fitur yang ditinjau model kesulitan membedakan antara STTC, normal dan MI.

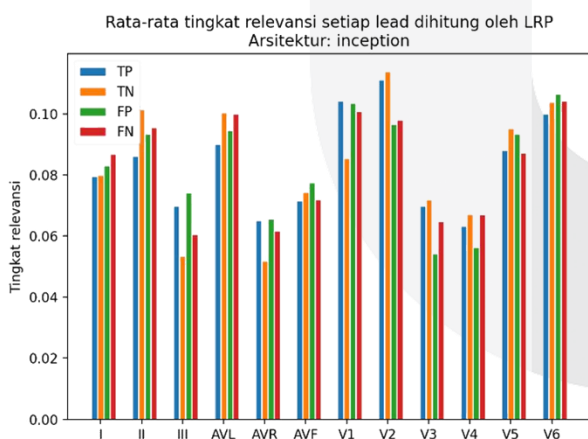
Gambar 9 – Gambar 11 menunjukkan grafik tingkat relevansi masing-masing *lead* dengan diukur oleh LRP. Distribusi relevansi tersebut berpotensi mengungkapkan bagaimana model menggunakan informasi dari *lead* berbeda.



Gambar 9 Lead relevance score AlexNet 1D



Gambar 10 Lead relevance score ResNet



Gambar 11 Lead relevance score Inception

Dapat disimpulkan bahwa:

- Lead V1, V2, V5 dan V6 memiliki relevansi lebih pada ketiga model. Secara klinis, keempat lead tersebut menunjukkan jantung bagian anterior serta lateral, sehingga menandakan model fokus mengklasifikasikan MI pada bagian jantung tersebut.
- Lead III, aVR, V3 dan V4 tidak terlalu relevan pada ketiga model. Lead III memiliki peran penting dalam klasifikasi MI inferior, sementara V3 dan V4 bermanfaat dalam klasifikasi MI anterior. Dapat disimpulkan bahwa ketiga model lebih memilih untuk menggunakan lead lain seperti

V1-V2 dan V5-V6 yang juga berguna dalam klasifikasi infark selain anterior.

- AlexNet bergantung secara ekstrim pada lead V1 untuk membuat prediksi MI.
- Inception menggunakan lebih banyak lead dibandingkan dua model lainnya, yakni lead I, II dan aVL masih sangat relevan. Lead I berperan dalam klasifikasi infark lateral, sementara lead II dan aVL berperan dalam klasifikasi MI inferior.
- Tidak seperti model lainnya, pada Inception, beberapa lead lebih penting dalam membuat outcome tertentu.

Secara keseluruhan, dari ketiga jenis arsitektur DL tersebut, arsitektur Inception merupakan yang terkuat karena melingkupi paling banyak lead, termasuk lead untuk klasifikasi infark miokard yang lebih jarang seperti infark inferior.

V. KESIMPULAN

A. Kesimpulan

Hasil implementasi Grad-CAM untuk menjelaskan model DL dalam kasus deteksi infark miokard pada sinyal EKG menunjukkan bahwa Grad-CAM menekankan fitur rekuren dengan relevansi konsisten, serta hanya memberikan atribusi positif terhadap kelas target yakni infark miokard, hanya efektif untuk model dengan lapisan konvolusi terakhir yang resolusinya tinggi, serta tidak mampu membedakan relevansi antara lead input EKG. Sementara itu, LRP memberikan atribusi positif maupun negatif, detail halus dan membedakan relevansi antara lead berbeda, namun pada implementasi tugas akhir ini, kemampuannya untuk menunjukkan fitur rekuren masih kurang konsisten dan intensitas penjelasannya dapat ditingkatkan. Inception merupakan model dengan performa terbaik, dengan akurasi tertinggi serta kombinasi Grad-CAM untuk analisis temporal, serta LRP untuk analisis relevansi lead menunjukkan bahwa arsitektur Inception menggunakan fitur dari lead terbanyak dan menghasilkan interpretasi paling lengkap dan cocok dengan pengetahuan klinis.

B. Saran

Saran untuk penelitian selanjutnya terkait XAI pada ranah sinyal EKG adalah sebagai berikut.

1. Penting untuk eksplorasi lebih lanjut variasi XAI yang beragam untuk DL klasifikasi EKG untuk mendapatkan hasil penjelasan yang lebih baik.
2. Peninjauan lebih lanjut metode evaluasi XAI yang lebih cocok untuk sinyal berulang seperti EKG sangat diperlukan, karena tes insertion/deletion memberikan input kepada model yang bukan merupakan sinyal EKG valid. Pendekatan seperti generative adversarial network (GAN) serta input optimization dapat menjadi langkah selanjutnya untuk mengevaluasi XAI dengan EKG sintetis.

DAFTAR PUSTAKA

- [1] E. Susanti, "Kemenkes : Penyakit Kardiovaskular Penyebab Kematian Tertinggi di Indonesia," Dinas Kesehatan Provinsi Aceh. Diakses: 5 November 2024. [Daring]. Tersedia pada: <https://dinkes.acehprov.go.id/detailpost/kemenkes-penyakit-kardiovaskular-penyebab-kematian-tertinggi-di-indonesia>
- [2] P. L. Laforgia, C. Auguadro, S. Bronzato, dan A. Durante, "The Reduction of Mortality in Acute Myocardial Infarction," *Int J Prev Med*, vol. 13, no. 1, hlm. 56, Jan 2022, doi: 10.4103/ijpvm.IJPVM_122_20.
- [3] E. Radwa, H. Ridha, dan B. Faycal, "Deep learning-based approaches for myocardial infarction detection: A comprehensive review recent advances and

- emerging challenges,” 1 September 2024, *Elsevier B.V.* doi: 10.1016/j.medntd.2024.100322.
- [4] O. J. Mechanic, M. Gavin, dan S. A. Grossman, *Acute Myocardial Infarction*. 2024. [Daring]. Tersedia pada: <http://www.ncbi.nlm.nih.gov/pubmed/30269080>
 - [5] M. Maleki, “Evaluation of Patient With Cardiovascular Problem,” dalam *Practical Cardiology*, Elsevier, 2022, hlm. 7–16. doi: 10.1016/B978-0-323-80915-3.00015-6.
 - [6] S. Ghaffari, “Non-ST-Elevation Acute Coronary Syndromes,” dalam *Practical Cardiology*, Elsevier, 2022, 23, hlm. 413–428. doi: 10.1016/B978-0-323-80915-3.00028-4.
 - [7] J. Hassannataj Joloudari dkk., “Application of artificial intelligence techniques for automated detection of myocardial infarction: a review,” 31 Agustus 2022, *Institute of Physics*. doi: 10.1088/1361-6579/ac7fd9.
 - [8] L. Karatzia, N. Aung, dan D. Aksentijevic, “Artificial intelligence in cardiology: Hope for the future and power for the present,” 13 Oktober 2022, *Frontiers Media S.A.* doi: 10.3389/fcvm.2022.945726.
 - [9] W.-C. Liu dkk., “A deep learning algorithm for detecting acute myocardial infarction,” *EuroIntervention*, vol. 17, no. 9, hlm. 765–773, Okt 2021, doi: 10.4244/EIJ-D-20-01155.
 - [10] P. Linardatos, V. Papastefanopoulos, dan S. Kotsiantis, “Explainable ai: A review of machine learning interpretability methods,” 1 Januari 2021, *MDPI AG*. doi: 10.3390/e23010018.
 - [11] E. Tjoa dan C. Guan, “A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI,” *IEEE Trans Neural Netw Learn Syst*, vol. 32, no. 11, hlm. 4793–4813, Nov 2021, doi: 10.1109/TNNLS.2020.3027314.
 - [12] I. Rehman dan A. Rehman, *Anatomy, Thorax, Heart*. 2024. [Daring]. Tersedia pada: <http://www.ncbi.nlm.nih.gov/pubmed/0>
 - [13] F. Kusumoto, *ECG Interpretation*. Cham: Springer International Publishing, 2020. doi: 10.1007/978-3-030-40341-6.
 - [14] D. G. Strauss, D. D. Schocken, dan G. S. Wagner, *Marriott's Practical Electrocardiography*. Wolters Kluwer, 2021.
 - [15] M. J. Alemzadeh-Ansari, “Electrocardiography,” dalam *Practical Cardiology*, Elsevier, 2022, hlm. 17–59, doi: 10.1016/B978-0-323-80915-3.00005-3.
 - [16] A. L. Goldberger, Z. D. Goldberger, dan A. Shvilkin, *Goldberger's Clinical Electrocardiography: A Simplified Approach*, 9 ed. Philadelphia (PA): Elsevier, 2018.
 - [17] X. Chen dkk., “Acute Myocardial Infarction Detection Using Deep Learning-Enabled Electrocardiograms,” *Front Cardiovasc Med*, vol. 8, 2021, doi: 10.3389/fcvm.2021.654515.
 - [18] M. Ganeshkumar, V. Ravi, V. Sowmya, E. A. Gopalakrishnan, dan K. P. Soman, “Explainable Deep Learning-Based Approach for Multilabel Classification of Electrocardiogram,” *IEEE Trans Eng Manag*, vol. 70, no. 8, hlm. 2787–2799, Agu 2023, doi: 10.1109/TEM.2021.3104751.
 - [19] V. Jahmunah, E. Y. K. Ng, R. S. Tan, S. L. Oh, dan U. R. Acharya, “Explainable detection of myocardial infarction using deep learning models with Grad-CAM technique on ECG signals,” *Comput Biol Med*, vol. 146, Jul 2022, doi: 10.1016/j.combiomed.2022.105550.
 - [20] P. Xiong, S. M. Y. Lee, dan G. Chan, “Deep Learning for Detecting and Locating Myocardial Infarction by Electrocardiogram: A Literature Review,” 25 Maret 2022, *Frontiers Media S.A.* doi: 10.3389/fcvm.2022.860032.
 - [21] W. J. Tompkins, “A Real-Time QRS Detection Algorithm,” 1985.
 - [22] H. M. Rai dan K. Chatterjee, “Hybrid CNN-LSTM deep learning model and ensemble technique for automatic detection of myocardial infarction using big ECG data,” *Applied Intelligence*, vol. 52, no. 5, hlm. 5366–5384, Mar 2022, doi: 10.1007/s10489-021-02696-6.
 - [23] N. Strodthoff, P. Wagner, T. Schaeffter, dan W. Samek, “Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL,” *IEEE J Biomed Health Inform*, vol. 25, no. 5, hlm. 1519–1528, Mei 2021, doi: 10.1109/JBHI.2020.3022989.
 - [24] H. Ismail Fawaz dkk., “InceptionTime: Finding AlexNet for time series classification,” *Data Min Knowl Discov*, vol. 34, no. 6, hlm. 1936–1962, Nov 2020, doi: 10.1007/s10618-020-00710-y.
 - [25] Y. Jin dkk., “A Novel Interpretable Method Based on Dual-Level Attentional Deep Neural Network for Actual Multilabel Arrhythmia Detection,” *IEEE Trans Instrum Meas*, vol. 71, hlm. 1–11, 2022, doi: 10.1109/TIM.2021.3135330.
 - [26] J. Yoo, T. J. Jun, dan Y.-H. Kim, “xECGNet: Fine-tuning attention map within convolutional neural network to improve detection and explainability of concurrent cardiac arrhythmias,” *Comput Methods Programs Biomed*, vol. 208, hlm. 106281, Sep 2021, doi: 10.1016/j.cmpb.2021.106281.
 - [27] J. M. Kwon dkk., “Artificial intelligence assessment for early detection of heart failure with preserved ejection fraction based on electrocardiographic features,” *European Heart Journal - Digital Health*, vol. 2, no. 1, hlm. 106–116, Mar 2021, doi: 10.1093/ehjdh/ztaa015.
 - [28] C. Han, J. Sun, Y. Bian, W. Que, dan L. Shi, “Automated Detection and Localization of Myocardial Infarction With Interpretability Analysis Based on Deep Learning,” *IEEE Trans Instrum Meas*, vol. 72, hlm. 1–12, 2023, doi: 10.1109/TIM.2023.3258521.
 - [29] W. Liu, F. Wang, Q. Huang, S. Chang, H. Wang, dan J. He, “MFB-CBRNN: A Hybrid Network for MI Detection Using 12-Lead ECGs,” *IEEE J Biomed Health Inform*, vol. 24, no. 2, hlm. 503–514, Feb 2020, doi: 10.1109/JBHI.2019.2910082.
 - [30] D. Hendrycks dan K. Gimpel, “Gaussian Error Linear Units (GELUs),” *arXiv: Learning*, Jun 2016, [Daring]. Tersedia pada: <http://arxiv.org/abs/1606.08415>
 - [31] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, dan D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” dalam *2017 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Okt 2017, hlm. 618–626. doi: 10.1109/ICCV.2017.74.
 - [32] K. He, X. Zhang, S. Ren, dan J. Sun, “Deep Residual Learning for Image Recognition,” dalam *Computer Vision and Pattern Recognition (CVPR)*, 2016, hlm. 770–778. doi: 10.1109/CVPR.2016.90.
 - [33] C. Szegedy dkk., “Going Deeper with Convolutions,” *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, hlm. 1–9, Jun 2015.
 - [34] C. Han, Y. Zhou, W. Que, Z. Li, dan L. Shi, “An Overview of Algorithms for Myocardial Infarction Diagnostics Using ECG Signals: Advances and Challenges,” *IEEE Trans Instrum Meas*, vol. 73, 2024, doi: 10.1109/TIM.2024.3418105.
 - [35] C. P. Langlotz dkk., “A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop,” *Radiology*, vol. 291, no. 3, hlm. 781–791, Jun 2019, doi: 10.1148/radiol.2019190613.
 - [36] S. M. Anwar, “AIM and Explainable Methods in Medical Imaging and Diagnostics,” dalam *Artificial Intelligence in Medicine*, Springer International Publishing, 2022, hlm. 501–510. doi: 10.1007/978-3-030-64573-1_293.
 - [37] A. M. Salih dkk., “A review of evaluation approaches for explainable AI with applications in cardiology,” *Artif Intell Rev*, vol. 57, no. 9, Sep 2024, doi: 10.1007/s10462-024-10852-w.
 - [38] M. T. Ribeiro, S. Singh, dan C. Guestrin, “‘Why should i trust you?’ Explaining the predictions of any classifier,” dalam *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Agu 2016, hlm. 1135–1144. doi: 10.1145/2939672.2939778.
 - [39] S. Lundberg dan S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” Nov 2017, [Daring]. Tersedia pada: <http://arxiv.org/abs/1705.07874>
 - [40] S. Bach, A. Binder, G. Montavon, F. Klauschen, K. R. Müller, dan W. Samek, “On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation,” *PLoS One*, vol. 10, no. 7, Jul 2015, doi: 10.1371/journal.pone.0130140.
 - [41] C. Han, J. Sun, Y. Bian, W. Que, dan L. Shi, “Automated Detection and Localization of Myocardial Infarction With Interpretability Analysis Based on Deep Learning,” *IEEE Trans Instrum Meas*, vol. 72, 2023, doi: 10.1109/TIM.2023.3258521.
 - [42] P. Singh dan A. Sharma, “Interpretation and Classification of Arrhythmia Using Deep Convolutional Network,” *IEEE Trans Instrum Meas*, vol. 71, 2022, doi: 10.1109/TIM.2022.3204316.
 - [43] Z. I. Attia, G. Lerman, dan P. A. Friedman, “Deep neural networks learn by using human-selected electrocardiogram features and novel features,” *European Heart Journal - Digital Health*, vol. 2, no. 3, hlm. 446–455, Sep 2021, doi: 10.1093/ehjdh/ztab060.
 - [44] T. A. A. Abdullah, M. S. M. Zahid, W. Ali, dan S. U. Hassan, “B-LIME: An Improvement of LIME for Interpretable Deep Learning Classification of Cardiac Arrhythmia from ECG Signals,” *Processes*, vol. 11, no. 2, Feb 2023, doi: 10.3390/pr11020595.
 - [45] A. Karoui, M. Bendahmane, dan N. Zemzemi, “Cardiac Activation Maps Reconstruction: A Comparative Study Between Data-Driven and Physics-Based Methods,” *Front Physiol*, vol. 12, Agu 2021, doi: 10.3389/fphys.2021.686136.
 - [46] G. Montavon, A. Binder, S. Lapuschkin, W. Samek, dan K. R. Müller, “Layer-Wise Relevance Propagation: An Overview,” dalam *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11700 LNCS, Springer Verlag, 2019, hlm. 193–209. doi: 10.1007/978-3-030-28954-6_10.
 - [47] W. Samek, A. Binder, G. Montavon, S. Lapuschkin, dan K.-R. Müller, “Evaluating the Visualization of What a Deep Neural Network Has Learned,” *IEEE Trans Neural Netw Learn Syst*, vol. 28, no. 11, hlm. 2660–2673, Nov 2017, doi: 10.1109/TNNLS.2016.2599820.
 - [48] S. Otsuki dkk., “Layer-Wise Relevance Propagation with Conservation Property for ResNet,” Jul 2024, [Daring]. Tersedia pada: <http://arxiv.org/abs/2407.09115>
 - [49] T. Bender dkk., “Analysis of a Deep Learning Model for 12-Lead ECG Classification Reveals Learned Features Similar to Diagnostic Criteria,” *IEEE J Biomed Health Inform*, vol. 28, no. 4, hlm. 1848–1859, Apr 2024, doi: 10.1109/JBHI.2023.3271858.
 - [50] P. Wagner dkk., “PTB-XL, a large publicly available electrocardiography dataset,” *Sci Data*, vol. 7, no. 1, Des 2020, doi: 10.1038/s41597-020-0495-6.
 - [51] A. Krizhevsky, I. Sutskever, dan G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks.” [Daring]. Tersedia pada: <http://code.google.com/p/cuda-convnet/>
 - [52] Z. Wang, W. Yan, dan T. Oates, “Time series classification from scratch with deep neural networks: A strong baseline,” dalam *2017 International Joint Conference on Neural Networks (IJCNN)*, IEEE, Mei 2017, hlm. 1578–1585. doi: 10.1109/IJCNN.2017.7966039.
 - [53] C. J. Anders, D. Neumann, W. Samek, K.-R. Müller, dan S. Lapuschkin, “Software for Dataset-wide XAI: From Local Explanations to Global Insights with Zennit, CoRelAy, and ViRelAy,” Feb 2023.