

Klasifikasi Infark Miokard Berbasis Sinyal EKG Menggunakan *Hybrid* CNN-LSTM dan CNN BiLSTM

1st Hana Rizkia Nafisa
Penulis-Teknik Biomedis

Telkom University
Bandung, Indonesia
rizsahana@student.telkomuniversity.ac.id

2nd Tito Waluyo Purboyo
Pembimbing 1-Teknik Biomedis

Telkom University
Bandung, Indonesia
titowaluyo@telkomuniversity.ac.id

3rd Dziban Naufal
Pembimbing 2-Teknik Biomedis

Telkom University
Bandung, Indonesia
dziban@telkomuniversity.ac.id

Abstrak— Infark miokard merupakan penyebab utama kematian akibat penyakit kardiovaskular di Indonesia. Elektrokardiogram (EKG) merupakan metode yang paling umum digunakan untuk mendeteksi perubahan aktivitas listrik jantung yang berkaitan dengan kondisi ini. Penelitian ini mengembangkan sistem klasifikasi sinyal EKG guna mendeteksi infark miokard menggunakan pendekatan hybrid yang menggabungkan *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM), serta membandingkan performanya dengan arsitektur *Bidirectional LSTM* (BiLSTM). Metode yang digunakan melibatkan dua model *hybrid*, yaitu CNN-LSTM dan CNN-BiLSTM, yang berperan dalam menangkap pola temporal dari sinyal EKG. Optimasi *hyperparameter* dilakukan menggunakan *grid search* dengan 5-fold *cross-validation*. Performa model dievaluasi berdasarkan akurasi, *presisi*, *recall*, *F1-score*, dan ROC-AUC. Hasil menunjukkan bahwa model CNN-LSTM menghasilkan akurasi sebesar 0,922, *presisi* 0,921, *recall* 0,922, *F1-score* 0,922, dan ROC-AUC 0,974. Model CNN-BiLSTM memberikan sedikit peningkatan performa dengan akurasi 0,923, *presisi* 0,924, *recall* 0,923, dan *F1-score* 0,924, meskipun ROC-AUC sedikit lebih rendah, yaitu 0,973. Meskipun BiLSTM memungkinkan pemrosesan temporal dua arah, peningkatan performa relatif kecil dan tidak signifikan dibandingkan kompleksitas arsitekturnya. Oleh karena itu, CNN-LSTM dinilai lebih efisien namun tetap efektif dalam klasifikasi sinyal EKG infark miokard.

Kata kunci— infark miokard, EKG, CNN-LSTM, BiLSTM, klasifikasi sinyal, *deep learning*.

I. PENDAHULUAN

Jantung merupakan organ utama tubuh yang berfungsi memompa darah ke seluruh organ. Jika jantung berhenti bekerja, organ vital seperti otak akan berhenti berfungsi. Dua arteri koroner membawa darah beroksigen yang dibutuhkan sel otot jantung untuk tetap hidup dan berfungsi. Sel-sel jantung akan mati jika suplai darah menuju jantung terganggu akibat penyumbatan arteri atau cabangnya. Kondisi ini dikenal sebagai iskemia, yaitu infark miokard (IM), yang terjadi ketika sel-sel jantung mati karena kekurangan aliran darah dalam jangka waktu yang cukup lama [1].

Faktor terjadinya penyakit kardiovaskular secara umum meliputi predisposisi genetik seperti riwayat penyakit keluarga, faktor lingkungan seperti merokok dan penyalahgunaan obat-obatan, gaya hidup kurang aktif, serta penyakit penyerta seperti diabetes yang tidak terkontrol, hipertensi, dislipidemia, penyakit paru, penyakit mental, dan kondisi lain yang meningkatkan risiko IM [2]. Data Kementerian Kesehatan Indonesia menunjukkan bahwa penyakit kardiovaskular merupakan penyebab kematian tertinggi, dengan stroke sebesar 19,42% dan jantung iskemik

sebesar 14,38% [3]. Secara global, IM menyumbang sekitar 33% dari total kematian terkait jantung iskemik [4].

Metode diagnosis IM yang umum digunakan meliputi pemeriksaan biomarker seperti troponin, pencitraan seperti ekokardiografi dan angiografi koroner, serta elektrokardiogram (EKG) [5]. Sinyal EKG yang menunjukkan IM umumnya ditandai oleh elevasi segmen ST, inversi gelombang T, atau gelombang Q patologis [6]. Seiring kemajuan teknologi, kecerdasan buatan (AI) telah digunakan untuk mempercepat dan meningkatkan akurasi diagnosis berbasis EKG. Studi ARISE (*Artificial Intelligence Enabled Rapid Identification of ST-Elevation Myocardial Infarction With Electrocardiogram*) menunjukkan bahwa penggunaan AI-ECG dapat mengurangi waktu pengiriman pasien ke *cath lab* sebesar 9 menit dibandingkan perawatan standar [7].

Untuk mendukung proses klasifikasi IM yang lebih komprehensif, penelitian ini menggunakan model *hybrid* CNN (*Convolutional Neural Network*) dan LSTM (*Long Short-Term Memory*). CNN digunakan untuk mengekstraksi fitur spasial sinyal EKG, sedangkan LSTM digunakan untuk menangkap pola temporal. Selain itu, digunakan pula arsitektur *Bidirectional LSTM* (BiLSTM) untuk memproses informasi temporal dari dua arah. Penelitian ini membandingkan performa CNN-LSTM dengan CNN-BiLSTM guna mengevaluasi pengaruh pemrosesan dua arah terhadap akurasi klasifikasi sinyal EKG.

Berdasarkan uraian tersebut, penelitian ini mengusulkan pengembangan sistem klasifikasi sinyal EKG menggunakan dua arsitektur gabungan—CNN-LSTM dan CNN-BiLSTM—dengan harapan dapat berkontribusi pada pengembangan sistem pendukung diagnosis otomatis berbasis EKG.

II. KAJIAN TEORI

A. Dataset

Dataset yang digunakan pada penelitian ini adalah PTB-XL (*Physikalisch-Technische Bundesanstalt*), yaitu kumpulan data elektrokardiogram (EKG) 12-lead terbesar yang terdiri dari 21.837 rekaman. Data diperoleh pada periode Oktober 1989 hingga Juni 1996, mencakup pasien dengan rentang usia mulai dari bayi hingga 95 tahun dengan median usia 62 tahun, serta distribusi gender yang seimbang, yaitu 52% laki-laki dan 48% perempuan. Seluruh data telah dianonimkan sesuai standar *Health Insurance Portability and Accountability Act* (HIPAA), dan publikasinya telah mendapatkan persetujuan dari Komite Etika Institusi (PTB-2020-1) [8].

Perekaman sinyal dilakukan menggunakan perangkat EKG dari Schiller AG dengan frekuensi sampling asli sebesar 400 Hz yang kemudian diolah ulang menjadi format 500 Hz dan 100 Hz. Setiap rekaman berdurasi 10 detik dan dilengkapi label diagnostik yang terdiri dari kelas utama serta subkelas [9]. Penelitian ini

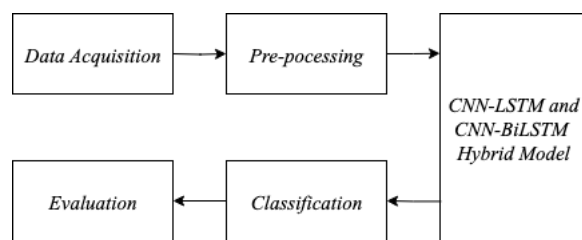
memanfaatkan seluruh data sinyal 12-lead EKG pada frekuensi 500 Hz, dengan fokus pada dua kelas diagnosis, yaitu *Myocardial Infarction* (MI) dan Normal (NORM) [10].

Sebelum digunakan dalam proses pelatihan model, dilakukan tahap pembersihan dan pelabelan data untuk memastikan hanya kelas MI dan NORM yang digunakan. Setelah itu, data dibagi menjadi *training set* dan *test set* dengan proporsi 80:20 menggunakan teknik *stratified sampling* untuk menjaga keseimbangan distribusi kelas pada kedua subset.

B. Metode

Penelitian ini menggunakan metode *hybrid deep learning* yang menggabungkan *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM), serta varian Bidirectional LSTM (BiLSTM), untuk melakukan klasifikasi sinyal elektrokardiogram (EKG) dalam mendeteksi infark miokard. Pendekatan ini dipilih karena CNN efektif dalam mengekstraksi fitur spasial dari sinyal, sedangkan LSTM dan BiLSTM mampu memodelkan pola temporal yang kompleks. Dengan memanfaatkan kedua jenis arsitektur ini secara bersama-sama, diharapkan sistem dapat mencapai kinerja klasifikasi yang lebih baik.

Metodologi penelitian mencakup beberapa tahapan, yaitu akuisisi data, *preprocessing*, perancangan dan pelatihan model, validasi, serta evaluasi kinerja model. Seluruh tahapan ini dirancang agar saling terintegrasi dan mampu menghasilkan model yang optimal untuk mendeteksi infark miokard berdasarkan sinyal EKG.



Bagan 1.
Diagram Blok Sistem Klasifikasi Infark Miokard

Diagram blok pada Bagan 1 menggambarkan tahapan penelitian yang diawali dengan data acquisition, yaitu pengambilan data sinyal EKG dari dataset PTB-XL [8]–[10] yang berisi rekaman 12-lead berdurasi 10 detik dengan frekuensi sampling 500 Hz. Pada tahap ini hanya digunakan dua kelas, yaitu *Myocardial Infarction* (MI) dan Normal (NORM). Data yang telah terkumpul kemudian diproses pada tahap preprocessing, yang meliputi pembersihan dan pelabelan data untuk memastikan hanya kelas yang relevan, pembagian data menggunakan teknik *stratified sampling* untuk menjaga proporsi kelas, penurunan frekuensi *sampling* (*downsampling*) dari 500 Hz menjadi 250 Hz untuk mengurangi beban komputasi, serta penyaringan sinyal (*filtering*) menggunakan *wavelet transform* dan *Butterworth low pass filter* guna menghilangkan *baseline wander* dan *high frequency noise*. Setelah itu dilakukan normalisasi amplitudo menggunakan metode *z-score* agar skala data seragam, dan penyeimbangan kelas (*balancing*) dengan kombinasi *Random Under Sampling* (RUS) dan *Synthetic Minority Oversampling Technique* (SMOTE) pada data latih.

Tahap berikutnya adalah model *training & validation*, di mana data hasil *preprocessing* menjadi masukan bagi arsitektur CNN-LSTM dan CNN-BiLSTM. CNN digunakan untuk mengekstraksi fitur spasial dari sinyal EKG, sedangkan LSTM atau BiLSTM mengolah urutan waktu untuk menangkap pola temporal. Pelatihan

model dilakukan dengan optimasi hyperparameter menggunakan *grid search*, dan evaluasi selama pelatihan menggunakan *5-fold cross validation* untuk mengurangi bias dan meningkatkan reliabilitas.

Tahap terakhir adalah model evaluation, di mana model yang telah dilatih diuji menggunakan data testing dan dievaluasi berdasarkan metrik akurasi, presisi, *recall*, *F1-score*, dan ROC-AUC. Tahapan ini memastikan model tidak hanya memiliki performa tinggi pada data latih, tetapi juga mampu memberikan prediksi yang akurat pada data yang belum pernah dilihat sebelumnya.

C. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan jenis jaringan saraf tiruan yang tersusun secara hierarkis dan terdiri dari beberapa *convolutional layer* serta *subsampling layer*. CNN dirancang untuk mengekstraksi fitur dari data masukan dengan memanfaatkan operasi konvolusi dan pooling secara berulang, sehingga mampu mengenali pola-pola penting pada data. Dalam penelitian ini, CNN digunakan sebagai modul ekstraksi fitur awal dari sinyal elektrokardiogram (EKG) sebelum diproses lebih lanjut oleh LSTM atau BiLSTM.

Arsitektur CNN bekerja dengan cara melakukan operasi konvolusi antara data masukan dan kernel atau filter untuk menghasilkan *feature map* yang menyoroti pola tertentu. Pemilihan ukuran kernel dan jumlah filter menjadi aspek penting karena sangat memengaruhi kemampuan jaringan dalam menangkap pola sinyal. Penyesuaian parameter ini umumnya dilakukan melalui eksperimen berulang atau metode optimasi seperti *grid search* [23].

Untuk data berdimensi satu seperti sinyal EKG, digunakan 1D CNN yang lebih efisien dibandingkan CNN dua dimensi. 1D CNN memproses data secara linier mengikuti urutan waktu, sehingga sangat sesuai untuk data sekuensial. Arsitektur 1D CNN terdiri dari beberapa komponen utama. *Convolutional layer* berfungsi mengekstraksi pola lokal dari sinyal. Setiap hasil konvolusi kemudian dilewatkan ke *activation function* seperti *Rectified Linear Unit* (ReLU) untuk memperkenalkan sifat *non-linear* pada model. Selanjutnya, *pooling layer* digunakan untuk mengurangi dimensi data dan mencegah *overfitting* dengan mempertahankan fitur-fitur penting saja. Hasil dari *pooling layer* kemudian diratakan melalui *flatten layer*, yang selanjutnya dihubungkan ke *fully connected layer* untuk melakukan klasifikasi [24][25].

Dengan kemampuannya dalam mengekstraksi representasi fitur yang informatif dari sinyal mentah, CNN berperan penting dalam tahap awal pemrosesan data pada model *hybrid* CNN-LSTM maupun CNN-BiLSTM. Kombinasi ini memungkinkan sistem untuk menangkap pola spasial dari sinyal EKG sebelum memproses dependensi temporalnya menggunakan LSTM atau BiLSTM.

D. Long Short-Term Memory (LSTM) dan Bidirectional LSTM (BiLSTM)

Long Short-Term Memory (LSTM) merupakan jenis jaringan saraf tiruan yang termasuk dalam keluarga *Recurrent Neural Network* (RNN) dan dirancang khusus untuk mengatasi masalah *long-term* dependensi pada data sekuensial. LSTM memiliki struktur internal yang terdiri dari *cell state* dan tiga gerbang pengendali aliran informasi, yaitu *forget gate*, *input gate*, dan *output gate*. *Forget gate* berfungsi untuk memutuskan informasi mana yang harus dibuang dari *cell state*, *input gate* menentukan informasi baru

yang akan ditambahkan, sedangkan output gate mengatur informasi yang akan dikeluarkan pada setiap langkah waktu [26][27].

Kelebihan LSTM terletak pada kemampuannya mempertahankan informasi penting dalam *cell state* untuk jangka waktu yang panjang, sehingga sangat cocok digunakan pada data deret waktu seperti sinyal EKG. Mekanisme gerbang ini membuat LSTM mampu mengenali pola temporal yang panjang tanpa mengalami degradasi memori yang signifikan, yang biasanya menjadi kelemahan pada RNN standar.

Bidirectional LSTM (BiLSTM) merupakan varian LSTM yang memproses data dalam dua arah, yaitu maju (*forward*) dan mundur (*backward*) secara bersamaan. Pada BiLSTM, dua lapisan LSTM digunakan: satu lapisan membaca urutan data dari awal ke akhir, sedangkan lapisan lainnya membaca dari akhir ke awal. Hasil dari kedua arah pemrosesan ini kemudian digabungkan untuk menghasilkan representasi yang lebih kaya. Pendekatan ini memungkinkan model memanfaatkan informasi dari masa lalu dan masa depan dalam satu urutan, sehingga seringkali menghasilkan kinerja yang lebih baik pada tugas klasifikasi [28].

Dalam konteks penelitian ini, LSTM maupun BiLSTM digunakan setelah lapisan CNN. CNN berfungsi mengekstraksi fitur spasial dari sinyal EKG, sementara LSTM atau BiLSTM mempelajari hubungan temporal antar fitur tersebut. Penggunaan BiLSTM diharapkan dapat menangkap pola temporal dari kedua arah waktu, namun tetap perlu dipertimbangkan dari segi kompleksitas model dan kebutuhan komputasi.

E. Cross Validation

Cross validation merupakan metode statistik yang digunakan untuk memperkirakan kemampuan generalisasi suatu model dengan cara membagi data menjadi beberapa *subset* atau *fold*. Setiap *fold* secara bergantian digunakan sebagai data uji, sementara *fold* lainnya digunakan sebagai data latih. Teknik ini bertujuan untuk mengurangi bias evaluasi dan memberikan estimasi performa model yang lebih akurat dibandingkan metode *holdout* sederhana [29].

Pada penelitian ini digunakan *5-fold cross validation*, di mana seluruh dataset dibagi menjadi lima bagian yang memiliki proporsi kelas seimbang. Pada setiap iterasi, empat bagian digunakan untuk pelatihan model, sementara satu bagian digunakan untuk pengujian. Proses ini diulang hingga setiap bagian pernah menjadi data uji satu kali. Nilai performa akhir dihitung sebagai rata-rata dari kelima hasil pengujian tersebut.

Penggunaan *stratified cross validation* menjadi penting pada dataset yang memiliki ketidakseimbangan kelas, seperti data EKG dengan jumlah sampel MI dan NORM yang tidak sama. Teknik ini memastikan bahwa distribusi kelas pada setiap *fold* tetap proporsional terhadap distribusi kelas pada dataset keseluruhan, sehingga hasil evaluasi lebih representatif.

Selain itu, penerapan *cross validation* dalam penelitian ini juga digunakan bersamaan dengan proses optimasi *hyperparameter* melalui *grid search*. Pendekatan ini membantu dalam memilih kombinasi parameter terbaik sambil tetap menjaga keandalan estimasi performa, serta meminimalkan risiko *overfitting* terhadap data latih.

F. Evaluasi

Evaluasi model pada penelitian ini dilakukan untuk mengukur kemampuan sistem dalam mengklasifikasikan sinyal elektrokardiogram (EKG) menjadi dua kelas, yaitu *Myocardial Infarction* (MI) dan Normal (NORM). Metode evaluasi yang

digunakan melibatkan *confusion matrix* sebagai dasar perhitungan berbagai metrik performa, yaitu akurasi, presisi, *recall*, dan *F1-score*, serta *Receiver Operating Characteristic Area Under the Curve* (ROC-AUC) [31]–[33].

Confusion matrix merupakan tabel yang merepresentasikan jumlah prediksi benar dan salah yang dibuat oleh model berdasarkan label aktual dan hasil prediksi. Tabel ini terdiri dari empat komponen: *True Positive* (TP), jumlah data positif yang diprediksi benar; *True Negative* (TN), jumlah data negatif yang diprediksi benar; *False Positive* (FP), jumlah data negatif yang salah diprediksi sebagai positif; dan *False Negative* (FN), jumlah data positif yang salah diprediksi sebagai negatif.

Berdasarkan nilai-nilai tersebut, metrik evaluasi dapat dihitung menggunakan persamaan berikut:

Presisi menunjukkan proporsi prediksi positif yang benar, dihitung dengan:

$$\text{Precision (Pr\%)} = \frac{TP}{TP + FP} * 100 \quad (1)$$

Recall mengukur kemampuan model dalam mengidentifikasi data positif secara benar, dengan rumus:

$$\text{Recall (Re\%)} = \frac{TP}{TP + FN} * 100 \quad (2)$$

F1-score adalah rata-rata harmonis antara precision dan recall, yang dirumuskan sebagai:

$$\text{Score(F1\%)} = \frac{2 * \text{precision} * \text{Recall}}{2 * \text{precision} + \text{Recall}} * 100 \quad (3)$$

Akurasi menunjukkan proporsi prediksi yang benar (baik positif maupun negatif) dari seluruh data, dihitung sebagai:

$$\text{Accuracy (Acc\%)} = \frac{TP + TN}{TP + FP + TN + FN} * 100 \quad (4)$$

Selain itu, kinerja model juga dievaluasi menggunakan ROC-AUC, yang merepresentasikan kemampuan model dalam membedakan kelas positif dan negatif pada berbagai ambang batas keputusan. ROC-AUC bernilai 1 untuk model sempurna, sedangkan nilai 0,5 menunjukkan kinerja setara dengan tebakan acak. ROC-AUC dihitung dari kurva ROC yang memplot True Positive Rate (TPR) terhadap False Positive Rate (FPR), dengan FPR dirumuskan sebagai:

$$\text{FPR} = \frac{FP}{FP + TN} \quad (5)$$

Penggunaan kombinasi metrik ini memberikan gambaran menyeluruh terhadap performa model, sehingga tidak hanya menilai dari satu aspek seperti akurasi saja, tetapi juga memperhatikan keseimbangan antara kemampuan mendeteksi positif dan menghindari kesalahan prediksi negatif.

III. HASIL PENELITIAN

A. Hasil Implementasi Cross Validation

Hasil penelitian ini disajikan berdasarkan implementasi model hybrid CNN-LSTM dan CNN-BiLSTM untuk klasifikasi sinyal elektrokardiogram (EKG) infark miokard. Evaluasi dilakukan melalui *5-fold cross validation* dan pengujian menggunakan data testing untuk membandingkan performa kedua arsitektur.

1) CNN-LSTM

Model CNN-LSTM dilatih menggunakan data hasil preprocessing dengan optimasi hyperparameter melalui grid search. Kombinasi parameter terbaik yang diperoleh adalah batch size 32, learning rate 0,001, optimizer Adam, dan 64 unit LSTM. Evaluasi menggunakan 5-fold cross validation menghasilkan nilai rata-rata akurasi sebesar 0,919, presisi 0,920, recall 0,919, dan F1-score 0,919. Nilai ini menunjukkan konsistensi kinerja model di setiap fold dan mengindikasikan kemampuan generalisasi yang baik.

TABEL 1
Hasil Cross Validation CNN-LSTM

No	Optimizer	Epochs	CV Accuracy	CV F1-Score	CV Precision	CV Recall	CV ROC-AUC
1	Adam	30	0.918	0.912	0.913	0.912	0.974
2	Adam	50	0.926	0.923	0.923	0.923	0.976
3	Adam	70	0.920	0.918	0.918	0.918	0.975
4	SGD	30	0.876	0.872	0.877	0.872	0.936
5	SGD	50	0.893	0.892	0.894	0.892	0.950
6	SGD	70	0.905	0.899	0.899	0.899	0.962
7	RMSprop	30	0.923	0.923	0.923	0.923	0.977
8	RMSprop	50	0.920	0.921	0.922	0.921	0.976
9	RMSprop	70	0.922	0.921	0.921	0.921	0.976

2) CNN-BiLSTM

Model CNN-BiLSTM dilatih menggunakan parameter hasil optimasi yang sama, yaitu batch size 32, learning rate 0,001, optimizer Adam, dan 64 unit BiLSTM. Hasil 5-fold cross validation menunjukkan nilai rata-rata akurasi sebesar 0,920, presisi 0,921, recall 0,920, dan F1-score 0,920. Perbedaan nilai metrik dibandingkan CNN-LSTM relatif kecil, namun BiLSTM menunjukkan sedikit peningkatan pada presisi dan akurasi.

Tabel 2
Hasil Cross Validation CNN-BiLSTM

No	Optimizer	Epochs	CV Accuracy	CV F1-Score	CV Precision	CV Recall	CV ROC-AUC
1	Adam	30	0.922	0.921	0.921	0.921	0.977
2	Adam	50	0.923	0.920	0.921	0.920	0.976
3	Adam	70	0.922	0.922	0.923	0.922	0.954
4	SGD	30	0.886	0.888	0.892	0.888	0.954
5	SGD	50	0.895	0.894	0.898	0.894	0.956
6	SGD	70	0.889	0.890	0.893	0.890	0.955
7	RMSprop	30	0.922	0.923	0.924	0.923	0.864
8	RMSprop	50	0.925	0.925	0.925	0.925	0.977
9	RMSprop	70	0.923	0.920	0.921	0.920	0.753

B. Evaluasi Model dengan Data Uji

Pengujian akhir pada data *testing* menghasilkan nilai akurasi untuk CNN-LSTM sebesar 0,922, presisi 0,921, recall 0,922, F1-score 0,922, dan ROC-AUC 0,974. Sementara itu, CNN-BiLSTM memperoleh akurasi 0,923, presisi 0,924, recall 0,923, F1-score 0,924, dan ROC-AUC 0,973.

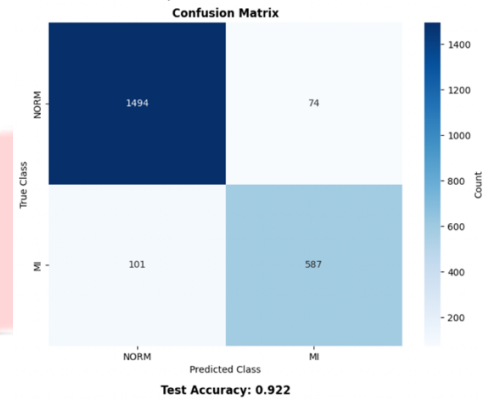
Tabel 3
Hasil Perbandingan Model dengan Data Uji

Model	Test Accuracy	Test Precision	Test Recall	Test F1-Score	Test ROC-AUC
LSTM	0.922	0.921	0.922	0.922	0.974
BiLSTM	0.923	0.924	0.923	0.924	0.973

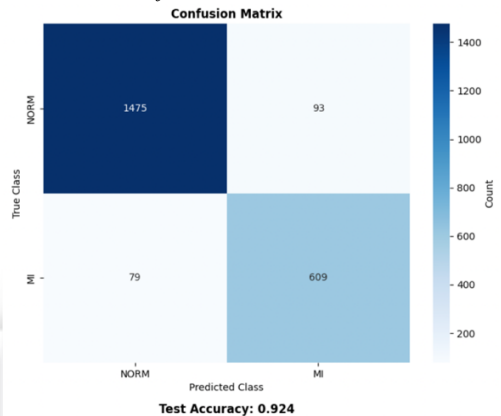
C. Analisis Perbandingan Metrik Klasifikasi

Perbandingan metrik menunjukkan bahwa CNN-LSTM unggul sedikit pada nilai ROC-AUC, sedangkan CNN-BiLSTM unggul tipis pada akurasi, presisi, dan F1-score. Analisis confusion matrix mengungkapkan bahwa sebagian besar kesalahan klasifikasi terjadi pada sinyal dengan morfologi gelombang yang mirip antara MI dan NORM.

Gambar 1
Confusion Matrix LSTM



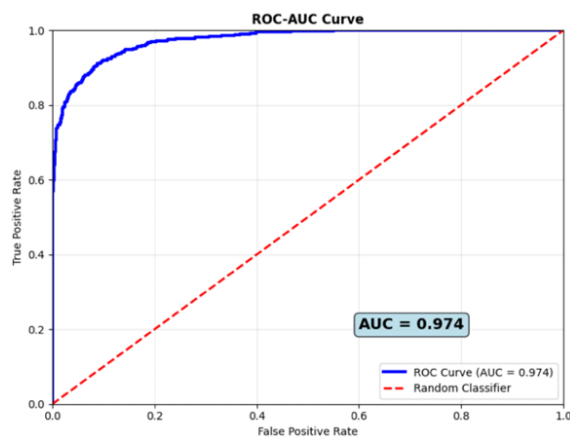
Gambar 2
Confusion Matrix BiLSTM



D. Analisis Learning Curves dan ROC-AUC

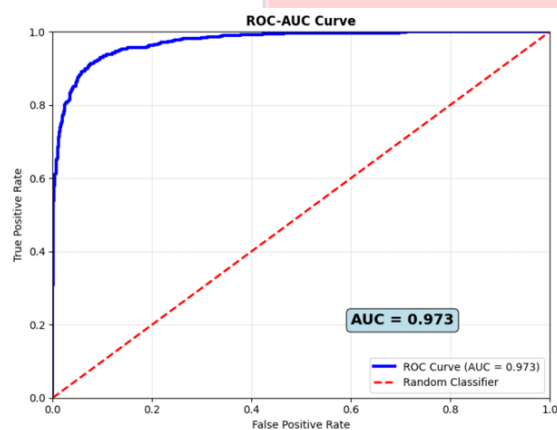
Kurva loss dan accuracy menunjukkan konvergensi stabil pada kedua model tanpa indikasi *overfitting*. Kurva ROC menunjukkan AUC mendekati 1 pada keduanya, yang menandakan kemampuan diskriminasi yang sangat baik.

Gambar 3
ROC-AUC Curve LSTM



Gambar 4

ROC-AUC Curve BiLSTM



E. Rekomendasi Model

CNN-LSTM direkomendasikan sebagai model utama karena menawarkan keseimbangan yang lebih baik antara kinerja dan efisiensi komputasi, meskipun CNN-BiLSTM memberikan sedikit peningkatan pada beberapa metrik.

IV. PEMBAHASAN

Hasil pengujian menunjukkan bahwa kedua arsitektur hybrid yang diusulkan, yaitu CNN-LSTM dan CNN-BiLSTM, memiliki kinerja yang tinggi dalam klasifikasi sinyal EKG untuk mendeteksi infark miokard. Pada tahap *cross validation*, CNN-BiLSTM memperoleh nilai akurasi, presisi, dan F1-score sedikit lebih tinggi dibandingkan CNN-LSTM, sedangkan CNN-LSTM unggul pada nilai ROC-AUC. Perbedaan performa ini relatif kecil, mengindikasikan bahwa kemampuan memproses informasi dua arah pada BiLSTM tidak memberikan peningkatan signifikan terhadap kemampuan diskriminasi dibandingkan LSTM.

Evaluasi menggunakan data *testing* memperlihatkan hasil yang konsisten dengan tahap *cross validation*. Proses preprocessing yang meliputi downsampling, wavelet filtering, normalisasi, dan penyeimbangan kelas menggunakan SMOTE-RUS terbukti efektif dalam mempersiapkan data sehingga model dapat mempelajari pola sinyal dengan baik. Nilai ROC-AUC yang mendekati 1 pada kedua model mengindikasikan kemampuan klasifikasi yang sangat baik.

Analisis kesalahan klasifikasi menunjukkan bahwa sebagian besar misprediksi disebabkan oleh kemiripan morfologi gelombang antara kelas MI dan NORM, yang merupakan tantangan umum dalam klasifikasi sinyal medis. Hal ini membuka peluang pengembangan lebih lanjut melalui feature engineering lanjutan atau augmentasi data yang spesifik terhadap domain EKG.

Kedua model menunjukkan kurva learning yang stabil tanpa indikasi overfitting, menandakan kapasitas generalisasi yang baik. Pertimbangan kompleksitas komputasi menjadi faktor penting; CNN-BiLSTM memerlukan sumber daya lebih besar, sedangkan CNN-LSTM menawarkan kinerja yang kompetitif dengan efisiensi yang lebih tinggi. Berdasarkan hasil ini, CNN-LSTM direkomendasikan sebagai model utama untuk implementasi sistem deteksi dini berbasis EKG, sementara CNN-BiLSTM dapat digunakan pada aplikasi yang tidak memiliki keterbatasan sumber daya dan memprioritaskan presisi lebih tinggi.

V. KESIMPULAN

Penelitian ini mengusulkan dan membandingkan dua arsitektur hybrid berbasis CNN-LSTM dan CNN-BiLSTM untuk klasifikasi sinyal elektrokardiogram (EKG) dalam mendeteksi infark miokard. Hasil pengujian menggunakan 5-fold cross validation dan data testing menunjukkan bahwa kedua model memiliki performa yang tinggi, dengan nilai akurasi, presisi, *recall*, dan F1-score di atas 0,92, serta nilai ROC-AUC mendekati 1.

Perbedaan performa antara kedua model relatif kecil. CNN-BiLSTM memperoleh nilai akurasi, presisi, dan F1-score sedikit lebih tinggi, sedangkan CNN-LSTM unggul pada ROC-AUC. Hal ini menunjukkan bahwa kemampuan memproses informasi dua arah pada BiLSTM tidak memberikan peningkatan yang signifikan terhadap kemampuan diskriminasi model.

Dari segi efisiensi, CNN-LSTM memerlukan sumber daya komputasi yang lebih rendah dibandingkan CNN-BiLSTM, sehingga lebih sesuai untuk implementasi pada sistem deteksi dini berbasis EKG yang memiliki keterbatasan perangkat keras atau waktu pemrosesan. CNN-BiLSTM dapat menjadi alternatif pada aplikasi yang tidak memiliki keterbatasan sumber daya dan memprioritaskan presisi yang sedikit lebih tinggi.

REFERENSI

- [1] A. K. Dohare, V. Kumar, and R. Kumar, "Detection of myocardial infarction in 12 lead ECG using support vector machine," *Appl. Soft Comput.*, vol. 64, pp. 138–147, Mar. 2018, doi: 10.1016/j.asoc.2017.12.001.
- [2] S. Abbas, S. Ojo, M. Krichen, M. A. Alamro, A. Mihoub, and L. Vilcekova, "A Novel Deep Learning Approach for Myocardial Infarction Detection and Multi-Label Classification," *IEEE Access*, vol. 12, pp. 76003–76021, 2024, doi: 10.1109/ACCESS.2024.3401744.
- [3] Dinas Kesehatan Provinsi Aceh, "Kemenkes: Penyakit Kardiovaskular Penyebab Kematian Tertinggi di Indonesia." Accessed: Nov. 03, 2024. [Online]. Available: <https://dinkes.acehprov.go.id/detailpost/kemenkes-penyakit-kardiovaskular-penyebab-kematian-tertinggi-di-indonesia>
- [4] P. L. Laforgia, C. Auguadro, S. Bronzato, and A. Durante, "The Reduction of Mortality in Acute Myocardial Infarction: From Bed Rest to Future Directions," *Int. J. Prev. Med.*, vol. 13, p. 122, 2022, doi: 10.4103/ijpvm.IJPVM_122_20.
- [5] dr. Nuri Kurniawan, "Pentingnya Pemeriksaan Penunjang dalam Penyakit Jantung dan Pembuluh Darah." Accessed: Nov. 03, 2024. [Online]. Available: https://yankes.kemkes.go.id/view_artikel/2995/pentingnya-pemeriksaan-penunjang-dalam-penyakit-jantung-dan-pembuluh-darah
- [6] M. P. Dr Araz Rawshani, "ECG interpretation: Characteristic of the normal ECG (P-wave, QRS complex, ST

segment, T-wave.” Accessed: Nov. 03, 2024. [Online]. Available: <https://ecgwaves.com/topic/ecg-normal-p-wave-qrs-complex-st-segment-t-wave-j-point/>

[7] M. Anthony A. Bavry, “Artificial Intelligence Enabled Rapid Identification of ST-Elevation Myocardial Infarction With Electrocardiogram-ARISE,” Taiwan, Nov. 2023. [Online]. Available: <https://www.acc.org/latest-in-cardiology/clinical-trials/2023/11/10/23/02/arise>

[8] A. L. Goldberger et al., “PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals,” *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[9] P. Wagner et al., “PTB-XL: A Large Publicly Available ECG Dataset,” *Sci. Data*, vol. 7, no. 1, pp. 1–15, 2020, doi: 10.1038/s41597-020-0495-6.

[10] P. Wagner, N. Strodthoff, R. Bousseljot, W. Samek, and T. Schaeffter, “PTB-XL, a large publicly available electrocardiography dataset (version 1.0.3),” 2022.

[23] S. H. Kim, Z. W. Geem, and G. T. Han, “Hyperparameter optimization method based on harmony search algorithm to improve performance of 1D CNN human respiration pattern recognition system,” *Sensors*, vol. 20, no. 13, pp. 1–20, Jul. 2020, doi: 10.3390/s20133697.

[24] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, “1D convolutional neural networks and applications: A survey,” *Mech. Syst. Signal Process.*, vol. 151, Apr. 2021, doi: 10.1016/j.ymssp.2020.107398.

[25] F. F. Taliningsih, Y. N. Fu’adah, S. Rizal, A. Rizal, and M. A. Pramudito, “Sistem Otentikasi Biometrik Berbasis Sinyal EKG Menggunakan Convolutional Neural Network 1 Dimensi,” *MIND Journal*, vol. 7, no. 1, pp. 1–10, Jun. 2022, doi: 10.26760/mindjournal.v7i1.1-10.

[26] H. M. Rai and K. Chatterjee, “Hybrid CNN-LSTM deep learning model and ensemble technique for automatic detection of

myocardial infarction using big ECG data,” *Appl. Intell.*, vol. 52, no. 5, pp. 5366–5384, Mar. 2022, doi: 10.1007/s10489-021-02696-6.

[27] S. M. Al-Selwi et al., “RNN-LSTM: From applications to modeling techniques and beyond—Systematic review,” *J. King Saud Univ. - Comput. Inf. Sci.*, Jun. 01, 2024, doi: 10.1016/j.jksuci.2024.102068.

[28] E. Zvornicanin, “Differences Between Bidirectional and Unidirectional LSTM,” Baeldung. Accessed: Dec. 09, 2024. [Online]. Available: <https://www.baeldung.com/cs/bidirectional-vs-unidirectional-lstm>

[29] D. Wilimitis and C. G. Walsh, “Practical Considerations and Applied Examples of Cross-Validation for Model Development and Evaluation in Health Care: Tutorial,” *JMIR AI*, vol. 2, no. 1, 2023, doi: 10.2196/49023.

[30] F. Pedregosa et al., “Cross-validation: evaluating estimator performance.” Accessed: Aug. 08, 2025. [Online]. Available: https://scikit-learn.org/stable/modules/cross_validation.html

[31] I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis, and N. Doulamis, “Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem,” *Technologies*, vol. 9, no. 4, Dec. 2021, doi: 10.3390/technologies9040081.

[32] M. Gragnaniello et al., “Real-Time Myocardial Infarction Detection Approaches with a Microcontroller-Based Edge-AI Device,” *Sensors*, vol. 24, no. 3, Feb. 2024, doi: 10.3390/s24030828.

[33] Evidently AI Team, “How to explain the ROC curve and ROC AUC score?,” Evidently AI. Accessed: Dec. 11, 2024. [Online]. Available: <https://www.evidentlyai.com/classification-metrics/explain-roc-curve>