Klasifikasi Penyakit Jantung Bawaan berdasarkan citra 3D *Computed Tomography* Menggunakan CNN-LSTM

Muhammad Fathan Mukhlisan
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia
famukhlisan@gmail.com

Annisa Darmawahyuni
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia
riset.annisadarmawahyuni@gmail.com

Abstrak*—Penyakit jantung bawaan (PJB) merupakan penyakit kelainan pada struktur jantung atau pembuluh darah besar sehingga dapat mengganggu fungsi jantung manusia. Penerapan tekonologi berbasis kecerdasan buatan bisa membantu para ahli medis dalam menentukan penanganan selanjutnya. Penelitian ini bertujuan untuk mengembangkan sistem deteksi PJB dengan mengklasifikasikan citra jantung ke dalam kategori PJB Sianotik, PJB Non-Sianotik, atau jantung normal dengan menggunakan teknik *deep learning* yang mengkombinasikan model *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM). Dalam sistem yang diusulkan, model CNN digunakan untuk ekstraksi fitur dari citra, sementara model LSTM digunakan untuk menganalisis pola sekuensial dari citra. Pengujian dilakukan menggunakan total 19.800 citra berbasis pasien dari 72 orang (dengan 275 citra sekuensial per pasien), yang terdiri atas 26 pasien PJB Sianotik, 26 pasien PJB Non-Sianotik, dan 20 pasien dengan jantung normal. Hasil pengujian menunjukkan akurasi pelatihan sebesar 100% akurasi validasi sebesar 71%, dan akurasi data uji sebesar 62% yang mengindikasikan terjadinya overfitting pada model. Kondisi ini mengimplikasikan bahwa sistem mampu menghafal pola PJB secara ekstensif pada data pelatihan, namun menunjukkan keterbatasan dalam generalisasi terhadap data yang belum pernah dilihat sebelumnya. Performa sistem berpotensi untuk ditingkatkan lebih lanjut dengan ketersediaan citra pasien yang lebih beragam dan dalam jumlah yang lebih besar.

Kata kunci—Klasifikasi, Penyakit Jantung Bawaan (PJB), Deep Learning, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM).

I. PENDAHULUAN

Kecerdasan Buatan sebagai teknologi revolusioner dapat berkontribusi di berbagai sektor dan aspek kehidupan manusia, termasuk sektor medis. Salah satu metode utama dalam Kecerdasan Buatan yaitu *deep learning*, yang konsepnya meniru kemampuan jaringan saraf biologis sehingga sistem bisa belajar secara mandiri. Konsep ini memungkinkan sistem untuk memproses dan menganalisis data medis, hasil dari sistem tersebut dapat memprediksi sesuatu dari data medis tersebut. Sistem yang diusulkan ini menggunakan data medis berupa citra dari pasien yang memiliki Penyakit Jantung Bawaan (PJB) dan pasien dengan jantung normal.

Kelainan PJB merupakan salah satu kelainan yang sering menjadi penyebab kematian pada fase awal kehidupan manusia, angka prevalensi PJB mencapai 9,4 bayi per 1000 kelahiran hidup berdasarkan survei epidemiologi PJB di berbagai negara [1][2][3]. Meskipun tanda-tanda kelainan PJB sudah muncul saat fase kelahiran, terdapat kemungkinan PJB dengan kategori ringan tidak terdeteksi dengan rentan waktu mingguan, bulanan, tahunan, dan bahkan terdeteksi saat pasien telah dewasa. Sebagian besar kasus PJB tidak dapat dijelaskan penyebab spesifiknya. Ada beberapa faktor yang masih diduga menjadi penyebab PJB, seperti faktor genetik, faktor lingkungan, infeksi, radiasi, atau bahkan konsumsi ibu selama kehamilan [1][2][4][5]. PJB secara klinis diklasifikasikan menjadi dua kategori utama, yaitu PJB Sianotik dan PJB Non-Sianotik. PJB Sianotik ditandai dengan manifestasi klinis berupa warna kebiruan. Sebaliknya, PJB Non-Sianotik yang tidak menunjukkan gejala kebiruan. PJB sianotik artinya terdapat penyumbatan pada alur keluar jantung atau pencampuran darah sirkulasi sistemik dan pulmonal yang tidak adekuat (kategori: TOF, PA, TGA, dan TrA). Sedangkan PJB non-sianotik terjadi akibat aliran darah vena yang bermasalah meliputi lubang pada sekat serambi jantung, bilik jantung atau pembuluh darah besar sehingga terjadi pirau dari kiri ke kanan (kategori: VSD, ASD dan PDA) [1][2][6][7][8].

Dataset yang digunakan pada sistem yang diusulkan adalah citra 3D *Computed Tomography* (CT) dari pasien yang hampir mencakup seluruh kategori PJB dan beberapa pasien yang tidak memiliki kelainan PJB [9]. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi PJB berbasis *deep learning* dengan menggabungkan *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM). Model CNN digunakan karena telah terbukti bahwa model CNN dapat mempelajari fitur dari citra menggunakan teknik peningkatan jaringan buatan hingga jumlah paramater yang mencapai jutaan [10][11]. Dikarenakan data citra 3D CT tersebut memiliki karakteristik sekuensial antar-irisannya (*frame*), sehingga model LSTM digunakan untuk mempelajari pola antar rangkaian citra tersebut [12][13]. Penelitian ini melibatkan data dari 72 pasien dengan masing-masing pasien memiliki 275 *frame*, sehingga total citra yang digunakan adalah 19.800 citra. Kemudian, 72 pasien tersebut dikelompokkan menjadi tiga kelas, sebanyak 26 pasien termasuk PJB sianotik, 26 pasien termasuk PJB non-sianotik, dan 20 pasien termasuk jantung normal.

II. STUDI TERKAIT

Klasifikasi PJB merupakan tantangan yang signifikan dalam bidang citra medis. Hal ini disebabkan oleh beberapa faktor, termasuk kompleksitas variasi struktur jantung dan koneksi pembuluh darah besar pada penderita PJB. Berbeda dengan deteksi lesi, PJB tidak selalu menyebabkan perubahan jaringan lokal, melainkan perubahan struktur global yang lebih sulit dideteksi. Selain itu, pengembangan sistem diagnosis otomatis untuk PJB juga terhambat oleh kurangnya

ketersediaan dataset publik yang memadai. Percobaan dalam mengatasi hal ini, Xu et al. memperkenalkan *Image Congenital Hearth Disease* (*ImageCHD*), dataset citra medis yang secara spesifik dirancang untuk klasifikasi PJB. Dataset ini terdiri dari 110 pasien dengan citra 3D CT yang telah dilabeli oleh tim radiolog kardiovaskular berpengalaman dan mencakup 16 kategori PJB yang berbeda [9]. Dalam penelitian mereka, Xu et al. juga menyajikan sebuah kerangka kerja baseline yang berbasis pada segmentasi dan analisis bentuk, yang berhasil mencapai akurasi 81,9% dengan cakupan 88,8%, namun mereka menyatakan masih terdapat ruang yang besar untuk perbaikan lebih lanjut [9].

Klasifikasi citra medis yang menggunakan arsitektur hibrida dengan menggabungkan CNN dan LSTM telah menunjukkan potensi yang besar. Penelitian yang dilakukan oleh Islam et al. menerapkan arsitektur CNN-LSTM untuk deteksi otomatis COVID-19 dari citra Sinar-X, di dalam sistem yang mereka kembangkan, CNN digunakan untuk ekstraksi fitur-fitur visual yang mendalam dari citra, sementara LSTM dimanfaatkan untuk melakukan klasifikasi berdasarkan rangkaian fitur yang telah diekstrak tersebut, sehingga Pendekatan hibrida ini terbukti sangat efektif dengan keberhasilan mencapai akurasi sebesar 99,4% dan AUC 99,9% [13]. Secara signifikan, penelitian tersebut juga menunjukkan bahwa model CNN-LSTM mereka memiliki kinerja yang lebih unggul dibandingkan dengan model CNN konvensional.

Penelitian selanjutnya menggunakan model *backbone* CNN didasari oleh implementasi 3D dari arsitektur ResNet [14]. Untuk meningkatkan performa dan efisiensi pelatihan, mereka menerapkan mekanisme *transfer learning* dengan menginisialisasi model mereka menggunakan bobot *pretrained* dari model Med3D yang dikembangkan oleh Chen et al., yang sebelumnya telah dioptimalkan untuk tugas pada citra medis 3D [15]. Dalam arsitektur hibrida tersebut, blok CNN bertugas untuk menangkap fitur-fitur pencitraan lokal dari setiap pemindaian CT, sementara blok RNN (dalam hal ini LSTM) dimanfaatkan untuk menangkap informasi deret waktu (*time-series*) global yang tertanam dalam seluruh riwayat pemindaian pasien.

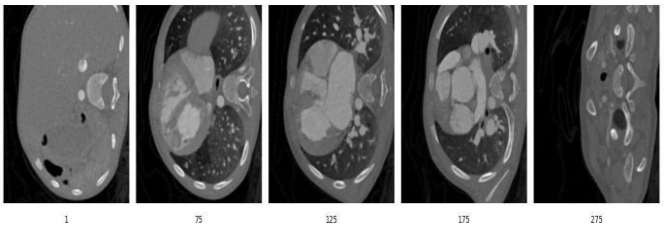
Berdasarkan tinjauan tersebut, penelitian ini bertujuan untuk menjembatani celah yang ada. Sementara Xu et al. telah menyediakan dataset PJB yang berharga dan keberhasilan arsitektur CNN-LSTM oleh Islam et al. pada domain citra medis kasus lain memberikan justifikasi kuat untuk mengadaptasi pendekatan serupa [9][13]. Oleh karena itu, penelitian ini mengusulkan penerapan model hibrida CNN-LSTM pada dataset 3D CT PJB untuk mengembangkan sistem klasifikasi dengan memanfaatkan kemampuan CNN dalam ekstraksi fitur spasial dan LSTM dalam analisis fitur sekuensial.

III. Metodologi Penelitian

Sistem klasifikasi PJB dimulai dengan mempersiapkan data agar dapat digunakan ke dalam sistem model. Sumber dataset *ImageCHD* berisi 110 total pasien. Total pasien tersebut diseleksi sehingga diambil 72 pasien yang terbagi ke dalam tiga kategori, sebanyak 26 pasien dengan PJB Sianotik (pasien dengan gejala TOF, PA, TGA, dan TrA), kemudian pasien dengan PJB non-sianotik (pasien dengan gejala VSD, ASD dan PDA), serta pasien dengan jantung normal (pasien yang tidak termasuk ke dalam kategori PJB). Seleksi ini memastikan representasi yang seimbang antar kelas.

A. Pra-pemprosesan Data

Format berkas pada sumber dataset adalah *Neuroimaging Informatics Technology Initiative* (NifTI), data tersebut akan dikonversi menjadi format PNG karena lebih familier ke beberapa pustaka yang digunakan penelitian ini (pustaka pengolahan citra: OpenCV dan kerangka deeplearning: PyTorch). Kemudian citra PNG diperkecil menjadi ukuran 256x256 piksel. Penyeragaman jumlah *frame* diperlukan agar model *deep learning* akan memiliki dimensi sekunsial yang seragam, karena jumlahnya yang sangat bervariasi mulai dari 137 hingga 551 *frame*. Jumlah target *frame* (T_F) yang akan diseragamkan yaitu 275, yang merupakan nilai modus dataset pasien.



Gambar 1. citra jantung 3D CT (indeks ke-1, 75, 125, 175, 275)

Pada Gambar 1, karakteristik sekuensial dari irisan citra 3D CT, di mana konten informatif terkait struktur jantung terkonsentrasi pada segmen tengah. Pada *frame* awal, tampak jantung belum terlihat secara signifikan. Struktur anatomis ini mulai tampak pada sekitar indeks ke-75, mencapai visualisasi jantung optimal pada indeks ke-125, dan kemudian tampak jantung mulai menghilang pada sekitar indeks ke-175. Berdasarkan observasi ini, disimpulkan bahwa kemampuan model untuk mengekstraksi fitur-fitur diskriminatif bergantung pada analisis segmen tengah dari data sekuensial tersebut.

Sekuens *frame* dari citra jantung pasien direpresentasikan sebagai (*S*) dengan jumlah *frame* awal (*F*), akan dinormalisasi dengan memprioritaskan segmen tengah. Implementasi dari pendekatan ini dilakukan dengan menerapkan salah satu dari dua metode berikut: Pemangkasan Sekuensial Simetris atau Duplikasi Sekuensial Simetris.

1. Pemangkasan Sekuensial Simetris, diterapkan pada data pasien dengan jumlah *frame* melebihi target ($F > T_F$). Tujuannya untuk mengurangi jumlah *frame* dengan cara menghapus sejumlah *frame* citra di bagian awal dan akhir sekuens secara merata sehingga simetris. Dimulai dengan menghitung jumlah *frame* yang akan dihapus (*N*).

$$N = F - T_F \tag{2}$$

Kemudian menentukan jumlah *frame* yang akan dihapus dari awal sekuens (N_{awal}) dan akhir sekuens (N_{akhir}). Untuk memastikan distribusi penghapusan yang seimbang, apabila *N* bernilai ganjil, maka satu *frame* tambahan akan dihapus di bagian awal.

$$N_{awal\ (ganjil)} = \frac{N}{2} + 1 \tag{2}$$

$$N_{awal\,(genap)} = N_{akhir} = \frac{N}{2} \tag{3}$$

Selanjutnya membentuk sekuens baru (*S'*) dengan mengambil *frame* dari sekuens asli (*S*) yang berada

di antara indeks N_{awal} dan F yang dikurangi dengan N_{akhir}.

$$S' = \{f_i \mid N_{awal} \leq i < F - N_{akhir}\} \quad (4)$$

Dimana f_i adalah *frame* ke-i dari sekuens asli. Hasilnya adalah sekuens baru dengan panjang *frame* sesuai target, dimana indeks ke-1 dimulai dari N_{awal} dari sekuens asli dan indeks ke-275 yaitu N_{akhir}.

2. Duplikasi Sekuensial Simetris, diterapkan pada data pasien dengan jumlah *frame* kurang dari target ($F < T_F$). Tujuan metode ini adalah untuk menambah jumlah *frame* hingga mencapai T_F melalui proses duplikasi *frame* persis pada bagian tengah indeks sehingga simetris. Kemudian *frame* duplikat disisipkan (*interleaved*) setelah posisi *frame* aslinya. Proses ini dibagi menjadi dua skenario berdasarkan jumlah *frame* yang perlu ditambahkan.

   Dimulai menghitung jumlah frame yang akan diduplikasi (N).

$$N = T_F - F \quad (5)$$

Selanjutnya mengidentifikasi segmen tengah dari sekuens asli (S_{tengah}) yang akan menjadi sumber duplikasi. Indeks tengah (i_{tengah}) dari sekuens asli.

$$i_{tengah} = \frac{F-1}{2} \quad (6)$$

Jumlah duplikasi (N) dibagi menjadi dua bagian, yaitu untuk sisi kiri (N_{kiri}) dan sisi kanan (N_{kanan}) dari indeks tengah. Jika N ganjil, sisi kiri akan memiliki satu *frame* lebih banyak.

$$N_{kanan}\ (ganjil) = \frac{N}{2} + 1 \quad (7)$$

$$N_{kanan}\ (genap) = N_{kiri} = \frac{N}{2} \quad (8)$$

Kemudian menentukan rentang indeks dari segmen tengah yang akan diduplikasi, yaitu dari indeks awal (i_{awal}) hingga indeks akhir (i_{akhir}).

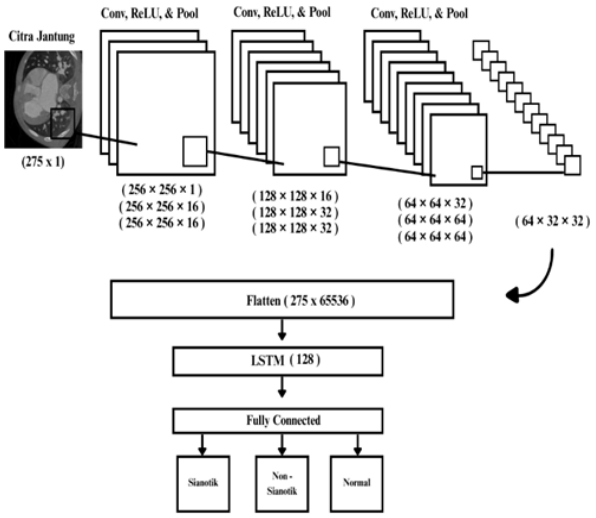
$$i_{awal} = i_{tengah} - N_{kiri} \quad (9)$$

$$i_{akhir} = i_{tengah} + N_{kanan} \quad (10)$$

Membentuk sekuens baru (S') dengan menyisipkan setiap *frame* duplikat dari rentang [i_{awal}, i_{akhir}] dengan metode penyisipan (*interleaving*). Setiap *frame* yang diduplikasi dari segmen sumber disisipkan secara langsung setelah posisi *frame* aslinya dalam sekuens. Apabila jumlah masih kurang dari target, maka akan dilakukan kembali proses duplikasi dari awal hingga akhir.

Operasi ini secara dinamis memperpanjang dengan menerapkan pergeseran indeks pada seluruh *frame* duplikat secara berurutan. Dengan demikian, integritas temporal relatif antar *frame* asli tetap terjaga, namun diselingi oleh duplikatnya untuk mencapai panjang sekuens target. Pendekatan ini berfungsi sebagai bentuk augmentasi temporal yang secara efektif mengutamakan segmen tengah pada sekuens citra bagi sistem model untuk proses ekstraksi fitur sekuensial.

B. Pengembangan Model



Gambar 2. Ilustrasi arsitektur CNN + LSTM

Data yang telah diproses kemudian dibagi berbasis pasien menggunakan metode *K-Fold Cross-Validation* dengan jumlah 5 *fold* (k=5) untuk mengevaluasi kinerja model [16]. Dalam setiap lipatan *fold* (k), dataset dibagi dengan perbandingan 90% untuk data pelatihan dan data validasi (total 64 pasien, 18 pasien normal, 23 pasien sianotik, 23 pasien non-sianotik) serta 10% untuk data tes uji (total 8 pasien, 2 pasien normal, 3 pasien sianotik, 3 pasien non-sianotik). Kemudian *data loader* pada penelitian dirancang untuk menangani dataset citra 3D CT yang akan disimpan sebagai sekuens citra 2D. Tujuannya adalah untuk membaca semua *frame* citra dari seorang pasien dan menyajikannya sebagai satu kesatuan data (volume) ke dalam model.

Setelah seluruh 275 *frame* dibaca, metode *stacking* digunakan untuk menumpuk semua tensor 2D tersebut menjadi satu tensor 3D tunggal dengan dimensi (275, 256, 256). Penambahan dimensi kanal membentuk tensor menjadi (1, 275, 256, 256), sehingga mempresentasi kanal, volume, tinggi, dan lebar citra.

1. *Convolutional neural network* (CNN)

   CNN telah terbukti efektif dalam berbagai aplikasi, mulai dari klasifikasi citra, deteksi objek, hingga analisis citra medis. CNN pada dasarnya mengekstrak fitur lokal *input* dari *higher layer* diubah ke *lower layer* untuk fitur yang lebih kompleks. Arsitektur CNN model yang diusulkan terdiri dari lapisan *input*, lapisan konvolusi, lapisan *pooling*, lapisan *Fully Connected* (FC) dan lapisan *output* [13].

   Dalam penelitian ini, blok CNN dirancang menggunakan tiga lapisan konvolusi. Setiap lapisan konvolusi diikuti oleh lapisan *rectified linear unit* (ReLU) untuk meningkatkan non-linearitas pada peta fitur dan dilanjutkan dengan lapisan *Max pooling* guna melakukan *downsampling* pada dimensi *input* yang diberikan untuk mengurangi jumlah parameter [13]. Kemudian ditambahkan lapisan *Dropout* dimana

merupakan teknik regularisasi untuk mengurangi *overfitting* pada model [17].

2. *Long short-term memory* (LSTM)

Fitur citra yang telah diekstrak oleh CNN, kemudian dimasukkan ke model LSTM untuk dianalisis dari segi pola sekuensial atau temporalnya. Dalam konteks ini, LSTM memperlakukan rangkaian fitur dari 275 *frame* citra sebagai sebuah sekuens, sama seperti model yang ingin memahami setiap kata dalam sebuah kalimat.

Saat memproses sekuens, LSTM akan melewati setiap vektor dari *frame* pertama hingga terakhir. Pada setiap langkahnya, LSTM akan memperbarui memori atau *hidden state* internalnya. *Hidden state* ini membawa informasi kontekstual dari *frame-frame* sebelumnya, dimana memungkinkan model untuk mempelajari bagaimana fitur pada satu *frame* berhubungan dengan fitur pada *frame* berikutnya. Kemampuan memori jangka panjang dari LSTM ini digunakan untuk mengidentifikasi perubahan atau konsistensi struktur anatomi jantung di sepanjang sumbu citra 3D CT.

Setelah seluruh 275 *frame* fitur selesai diproses, model akan mengambil *output* dari *hidden state* pada bagian akhir. Vektor akhir berisi ringkasan yang kaya akan konteks dan merepresentasikan informasi dari keseluruhan volume *CT scan*, bukan hanya dari satu *frame* saja. Kemudian vektor ringkasan tersebut diteruskan ke lapisan *Fully Connected* (FC *layer*) untuk dilakukan klasifikasi akhir [18].

Lapisan terakhir FC *layer* yang berbentuk linear berfungsi sebagai kepala klasifikasi. Kepala klasifikasi bertanggung jawab untuk membuat prediksi akhir. Masukan untuk lapisan ini adalah vektor fitur ringkasan yang dihasilkan oleh *hidden state* terakhir dari lapisan LSTM. Vektor ini merepresentasikan informasi kontekstual yang telah diekstraksi dari keseluruhan sekuens citra 3D.

Melalui sebuah transformasi linear, lapisan FC ini memetakan representasi fitur tingkat tinggi tersebut ke dalam ruang prediksi yang dimensinya sesuai dengan jumlah kelas target. *Output* dari lapisan ini adalah tensor yang berisi nilai logits untuk setiap kelas—normal, sianotik, dan non-sianotik. Nilai logits inilah yang selanjutnya menjadi dasar untuk penentuan kelas prediksi oleh model.

TABLE I. PARAMETER ARSITEKTUR CNN+LSTM

No	Tipe	*Kernel Size*	*Stride*	*Kernel*	*Input Size*
1	*Convolution2D*	3 × 3	1	16	256 × 256 × 1
2	*ReLU*	-	-	-	256 × 256 × 16
3	*MaxPool2D*	2 × 2	2	-	256 × 256 × 16
4	*Convolution2D*	3 × 3	1	32	128 × 128 × 16
5	*ReLU*	-	-	-	128 × 128 × 32
6	*MaxPool2D*	2 × 2	2	-	128 × 128 × 32
7	*Convolution2D*	3 × 3	1	64	64 × 64 × 32
8	*ReLU*	-	-	-	64 × 64 × 64
9	*MaxPool2D*	2 × 2	2	-	64 × 64 × 64
10	*Flatten*	-	-	-	32 × 32 × 64
11	*Dropout (0.5)*	-	-	-	(275, 65536)
12	*LSTM*	-	-	128	(275, 65536)
13	*Linear (FC)*	-	-	3	128

Arsitektur model terdiri dari 13 lapisan seperti yang dirinci pada Tabel I. Model dilatih menggunakan beberapa *hyperparameter*. Proses optimisasi menggunakan *learning rate* sebesar 0.0001, dengan pelatihan dijalankan sebanyak 20 *epoch* untuk setiap *fold*, dan *batch size* sebesar 1.

C. Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja dan akurasi prediksi dengan menerapkan beberapa metrik evaluasi standar dalam tugas klasifikasi model, digunakan untuk mengevaluasi kinerja model yang diusulkan.

1. *Confusion Matrix*

Confusion matrix merupakan tabel sebagai tolak ukur performa dari algoritma sistem. Tabel ini menunjukkan jumlah prediksi yang tepat dan tidak tepat berdasarkan label yang sebenarnya serta label yang diprediksi oleh model. Confusion matrix terdiri dari empat komponen utama, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [19]. Pada situasi multi-kelas seperti penelitian ini yang merupakan klasifikasi untuk tiga kelas (sianotik, non-sianotik, dan normal), sehingga *confusion matrix* berupa tabel N×N, dimana N sebagai jumlah kelas.

2. Metrik Evaluasi

Penggunaan kombinasi metrik evaluasi digunakan untuk memberikan penilaian yang lebih adil terhadap performa klasifikasi [20]. Perhitungan metrik-metrik ini didasarkan pada empat komponen utama dari *confusion matrix* yang dalam konteks penelitian ini didefinisikan sebagai berikut: TP yang menyatakan prediksi kelas PJB (Sianotik atau Non-Sianotik) yang benar; FP yang menyatakan prediksi jantung normal yang salah diklasifikasikan sebagai PJB; TN yang menyatakan prediksi kelas jantung normal yang benar; dan FN yang menyatakan prediksi kelas PJB yang salah diklasifikasikan.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (11)$$

$$Sensitivitas = \frac{TP}{TP+FN} \quad (12)$$

$$Spesifisitas = \frac{TN}{TN+FP} \quad (13)$$

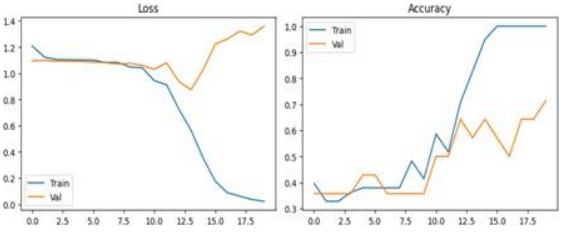
$$Presisi = \frac{TP}{TP+FP} \quad (14)$$

$$F1-Score = \frac{2*TP}{2*TP+FP+FN} \quad (15)$$

Metrik evaluasi ini akan digunakan untuk memberikan penilaian kuantitatif terhadap kinerja model yang diusulkan.

IV. HASIL PENELITIAN

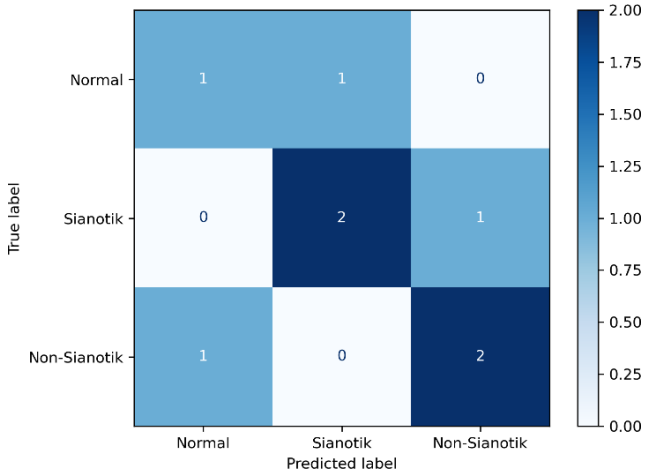
A. Hasil Pelatihan Model



Gambar 3. *loss* dan akurasi terhadap pelatihan dan validasi model

Hasil pelatihan sistem menunjukkan akurasi pelatihan mencapai 100% dan akurasi validasi mencapai 71%. Pada Gambar 3., kesenjangan yang signifikan antara kedua nilai akurasi mengindikasi bahwa sistem model mengalami *overfitting*. Hal ini mengartikan bahwa sistem model mudah menghafal data pelatihan tetapi tidak mampu melakukan generalisasi pada data validasi.

B. Evaluasi Model



Gambar 4. Tabel *confusion matrix* pada data tes

Pada Gambar 4., merupakan hasil yang merepresentasikan performa model terhadap data tes (10%) yang merupakan subset data yang diisolasi dan tidak pernah terlibat dalam fase pelatihan maupun validasi model.

TABLE II. HASIL PERFORMA MODEL

Kelas	Presisi (%)	Sensitivitas (%)	Spesifisitas (%)	F1-Score (%)
Sianotik	66,7	66,7	75,0	66,7
Non-Sianotik	66,7	66,7	75,0	66,7
Normal	50,0	50,0	80,0	50,0

Secara keseluruhan, model mencapai Akurasi sebesar 62.5%, mengindikasi bahwa model berhasil memprediksi 5 dari 8 total pasien uji. Sementara metrik per kelas, yang dirangkum pada Tabel II, menampilkan performa model untuk setiap kelas secara individual.

V. KESIMPULAN

Pendekatan arsitektur kombinasi CNN-LSTM terbukti memiliki kapasitas dalam mengekstraksi fitur spasial dan sekuensial pada klasifikasi PJB dari citra 3D CT. Kemampuan ini ditunjukkan oleh keberhasilan model mencapai akurasi pelatihan sempurna. Namun, performa model pada data validasi yang lebih rendah mengindikasikan terjadinya overfitting yang signifikan, yang kemungkinan disebabkan oleh keterbatasan jumlah dataset dimana hanya terdiri dari 72 pasien. Dengan demikian, sistem model yang diusulkan memiliki potensial untuk klasifikasi PJB, namun peningkatan performa generalisasi kedepannya bergantung pada jumlah dataset pasien yang lebih beragam untuk mengatasi masalah *overfitting* dan meningkatkan kemampuan prediksinya.

REFERENSI

[1] Eva Miranda Marwali, Yoel Purnama, and Poppy Surwianti Roebiono, "Modalitas Deteksi Dini Penyakit Jantung Bawaan di Pelayanan Kesehatan Primer," *J. Indones. Med. Assoc.*, vol. 71, no. 2, pp. 100–109, 2021, doi: 10.47830/jinma-vol.71.2-2021-241.

[2] Lilly LS, *Pathophysiology of heart disease: a collaborative project of medical students and faculty*. Lippincott Williams & Wilkins, 2012.

[3] Y. Liu *et al.*, "Global birth prevalence of congenital heart defects 1970–2017: updated systematic review and meta-analysis of 260 studies," *Int. J. Epidemiol.*, vol. 48, no. 2, pp. 455–463, Apr. 2019, doi: 10.1093/ije/dyz009.

[4] G. M. Harimurti, P. S. Roebiono, L. I. Rilantono, and W. S. Suhardiman, "Congenitally malformed hearts in newborn infants in several hospitals in Indonesia," in *Asian Congress of Cardiology*, 1997.

[5] "Penyakit Jantung Bawaan," Inaheart.org. [Online]. Available: http://www.inaheart.org/education_for_patient/2019/7/10/penyakit_jantung_bawaan

[6] Roebiono PS, "Kenali penyakit jantung bawaan pada anak yang sedang tumbuh kembang," in *Rilantono LI, ed. Penyakit Kardio-vaskular (PKV)*, Jakarta: 5 Rahasia, 2012, pp. 495–508.

[7] R. I. Koppel *et al.*, "Effectiveness of Pulse Oximetry Screening for Congenital Heart Disease in Asymptomatic Newborns," *Pediatrics*, vol. 111, no. 3, pp. 451–455, Mar. 2003, doi: 10.1542/peds.111.3.451.

[8] H. W. Sniderman, S Taeusch, "Initial evaluation: history and physical examination of the newborn," in *Avery's Diseases of the Newborn*, 8th editio., C. A. Taeusch, H William Ballard, Roberta A Gleason, Ed., Philadelphia, PA: Elsevier Saunders, 2005, pp. 301–322.

[9] X. Xu *et al.*, "ImageCHD: A 3D Computed Tomography Image Dataset for Classification of Congenital Heart Disease BT - Medical Image Computing and Computer Assisted Intervention – MICCAI 2020," A. L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M. A. Zuluaga, S. K. Zhou, D. Racoceanu, and L. Joskowicz, Eds., Cham: Springer International Publishing, 2020, pp. 77–87.

[10] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-Scale Video Classification with Convolutional Neural Networks," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1725–1732. doi: 10.1109/CVPR.2014.223.

[11] D. Matthew Zeiler and F. Rob, "Visualizing and understanding convolutional neural networks," in *Proceedings of the European Conference on Computer Vision, Zurich, Switzerland*, 2014, pp. 6–12.

[12] Y. Lu *et al.*, "A hybrid CNN-RNN approach for survival analysis in a Lung Cancer Screening study," *Heliyon*, vol. 9, no. 8, p. e18695, 2023, doi: https://doi.org/10.1016/j.heliyon.2023.e18695.

[13] Z. Islam, M. Islam, and A. Asraf, "Informatics in Medicine Unlocked A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Informatics Med. Unlocked*, vol. 20, p. 100412, 2020, doi: 10.1016/j.imu.2020.100412.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[15] S. Chen, K. Ma, and Y. Zheng, "Med3d: Transfer learning for 3d medical image analysis," *arXiv Prepr. arXiv1904.00625*, 2019.

[16] D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, and S. Ridella, "The'K'in K-fold Cross Validation.," in *ESANN*, 2012, pp. 441–446.

[17] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, vol. 25, 2012.

[19] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[20] I. M. De Diego, A. R. Redondo, R. R. Fernández, J. Navarro, and J. M. Moguerza, "General Performance Score for classification problems," *Appl. Intell.*, vol. 52, no. 10, pp. 12049–12063, 2022, doi: 10.1007/s10489-021-03041-7.