

PENERAPAN *CONVOLUTIONAL NEURAL NETWORKS* (CNN) UNTUK DETEKSI PENYAKIT HEPATITIS C VIRUS (HCV)

Fathorrozi Ariyanto^{1*}, Indra Maulana², Moh. Aminollah Hamzah³, Aang Kisnu Darmawan⁴

^{1,2,3} Prodi Teknik Informatika, Universitas Islam Madura

⁴ Prodi Sistem Informasi, Universitas Islam Madura

Jln. Pondok Pesantren Miftahul Ulum Bettet, Pamekasan Madura, Kabupaten Pamekasan, Jawa Timur 69317

¹fathorroziariyanto7@gmail.com

²indralav4@gmail.com

³hamzahmni@gmail.com

⁴ak.darmawan@gmail.com

Abstract

Hepatitis C (HCV) is a serious liver infection that can progress to cirrhosis or cancer, especially if early diagnosis is neglected. While conventional diagnostic methods such as ELISA are accurate, they are often limited in terms of efficiency and accessibility. **This study introduces** an innovative approach using a one-dimensional convolutional neural network (1D-CNN) for HCV disease classification, utilizing a Kaggle dataset consisting of 615 patient samples. **The methodology** includes data preprocessing such as handling missing values and transforming categorical variables to numeric values to ensure data readiness. The 1D-CNN model was trained using the Adam optimizer with ReLU activation functions in the convolution layer and sigmoid in the output layer. Model performance was comprehensively evaluated through accuracy, precision, recall, and F1-score metrics. **The results showed** that the 1D-CNN model achieved an accuracy of 83.74% on the training data and 81.30% on the testing data after hyperparameter tuning. This improvement is significant compared to the initial accuracy of only around 52%. However, the model exhibits a strong bias towards the majority class (Blood Donor), with very poor performance on minority classes such as Hepatitis, Fibrosis, and Cirrhosis. Nonetheless, **this study contributes** to the exploration of 1D-CNN for non-image medical data, which is still rarely studied. We conclude that despite the model's potential, further developments such as data balancing are needed to improve the model's overall generalization and accuracy.

Keywords: Deep Learning, Disease Classification, Hepatitis C, 1D CNN, Medical Diagnosis

I. PENDAHULUAN

Hepatitis C adalah infeksi yang menyerang organ hati dan disebabkan oleh virus Hepatitis C (HCV) [1], [2]. Meskipun sering tidak menunjukkan gejala, infeksi kronis dapat menyebabkan jaringan parut (eskar) pada hati yang memicu sirosis setelah bertahun-tahun [2]. Dalam kasus-kasus tertentu, sirosis juga bisa menyebabkan gagal hati, kanker hati, atau pembuluh darah yang membengkak di esofagus dan lambung, yang dapat berakibat fatal [3], [4]. Dalam praktik medis, diagnosis Hepatitis C biasanya dimulai dengan pemeriksaan darah untuk mendeteksi antibodi HCV melalui uji imunoasai enzim [5].

Saat ini, metode diagnosis konvensional seperti ELISA dan PCR, meskipun akurat, memiliki keterbatasan dalam hal waktu, biaya, dan aksesibilitas, terutama di daerah dengan sumber daya terbatas [6], [7]. Keterbatasan ini menghambat upaya deteksi dini yang sangat penting. Seiring berkembangnya teknologi, kecerdasan buatan (AI) telah menarik banyak perhatian di bidang kesehatan untuk memprediksi penyakit, seperti Hepatitis C, kanker payudara, malaria, dan diabetes [7], [8], [9].

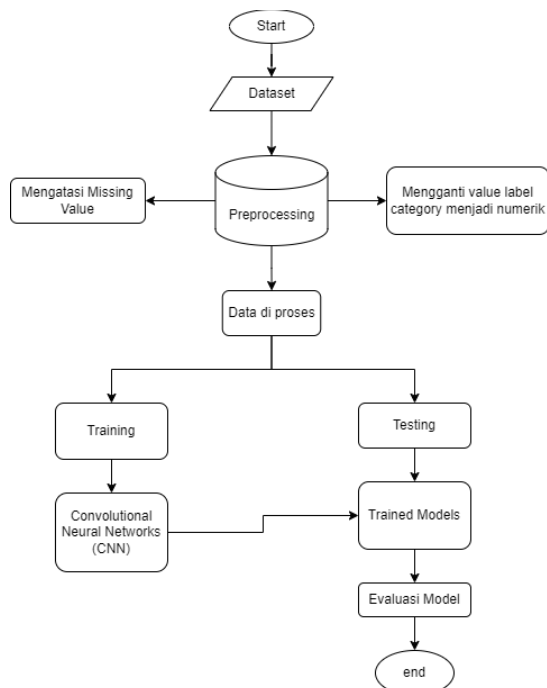
Beberapa penelitian sebelumnya telah menerapkan berbagai algoritma untuk klasifikasi Hepatitis C. Misalnya, studi tentang analisis SMOTE menggunakan algoritma **Random Forest** dan **Naive Bayes** mencapai akurasi masing-masing 98% dan 89% [10]. Pendekatan **Ensembled Learning Decision Tree** juga menghasilkan akurasi 93% [11], sementara **Artificial Neural Network (ANN)** mencapai akurasi hingga 97,78% [12]. Namun, sebagian besar penelitian sebelumnya berfokus pada citra medis, sementara analisis data satu dimensi (1D), seperti spektrum serapan molekuler atau sinyal elektroforesis, masih kurang dieksplorasi untuk HCV [13].

Untuk mengatasi kesenjangan ini, penelitian kami berfokus pada penerapan **Convolutional Neural Network satu dimensi (1D-CNN)** untuk klasifikasi penyakit Hepatitis C. Pendekatan ini menawarkan beberapa kebaruan dan kontribusi. **Pertama**, kami mengurangi kesenjangan literatur karena mayoritas studi CNN untuk HCV menggunakan data 2D (citra), sedangkan pendekatan 1D-CNN untuk klasifikasi berbasis data spektral atau sinyal masih sangat jarang. **Kedua**, 1D-CNN menawarkan efisiensi komputasi yang lebih tinggi dan kebutuhan sumber daya yang lebih rendah dibandingkan 2D/3D-CNN, menjadikannya cocok untuk implementasi *real-time* di klinik. **Ketiga**, solusi ini berpotensi untuk diintegrasikan dengan

perangkat *point-of-care* untuk skrining HCV di daerah terpencil, mengatasi hambatan akses diagnostik. Selain itu, penelitian ini juga mencakup optimalisasi arsitektur 1D-CNN, ekstraksi fitur otomatis, dan validasi klinis menggunakan dataset pasien HCV [14]. Dengan demikian, penelitian ini bertujuan untuk mengeksplorasi apakah model 1D-CNN dapat mencapai akurasi tinggi dan menjadi alat yang efektif untuk diagnosis Hepatitis C.

II. METODOLOGI PENELITIAN

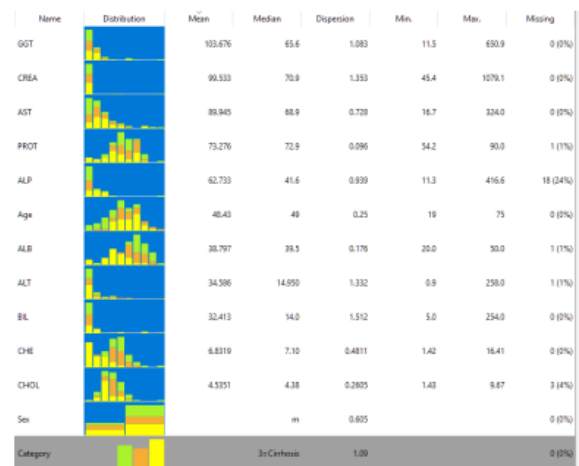
Metodologi penelitian ini menjelaskan tahapan yang dilakukan dalam mengklasifikasikan penyakit Hepatitis C menggunakan *Convolutional Neural Network (CNN) 1D*. Bagian ini merinci alat dan bahan, prosedur penelitian, teknik pengolahan data, serta evaluasi model yang diterapkan. Penjelasan ini bertujuan untuk memberikan gambaran yang jelas dan lengkap, sehingga penelitian dapat dipahami dan direplikasi oleh peneliti lain di masa mendatang. Alur pelaksanaan dalam penelitian ini mencakup serangkaian tahapan yang dilakukan secara sistematis, mulai dari pengumpulan dataset hingga evaluasi model.



Gambar 1. Flowchart Diagram untuk Metode yang Diusulkan

A. Dataset

Dataset yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari platform Kaggle dengan kata kunci "Hepatitis C". Dataset dapat diakses melalui platform Kaggle pada tautan berikut : <http://kaggle.com/datasets/fedesoriano/hepatitis-c-dataset> dataset ini terdiri dari 615 sampel pasien dengan atribut medis relevan untuk klasifikasi penyakit Hepatitis C. Atribut yang digunakan meliputi Age, ALB, ALP, ALT, AST, BIL, CHE, CHOL, CREA, GGT, dan PROT sebagai variabel input, serta Category sebagai variabel target yang mencakup *Healthy*, *HCV*, *Fibrosis*, dan *Cirrhosis*.



Gambar 2. Distribusi Data Statistik

Data berisi 615 catatan, dengan rincian 75 orang terkena infeksi HCV (Hepatitis 24 Orang, 24 Orang terkena Fibrosis dan 30 orang terkena sirosis), dan sisanya adalah pendonor darah dengan seperti biasa, Gambar 2 menunjukkan hasil data statistik dari kumpulan data Hepatitis C Virus ini.

B. Data Processing

Data preprocessing secara sederhana merupakan prosedur dalam menyiapkan data [11] hingga siap digunakan untuk proses selanjutnya. Ada beberapa langkah yang terlibat dalam memproses data, sebagai berikut:

1. Mengatasi Missing Value

Missing value pada model algoritma CNN merupakan data yang hilang atau tidak lengkap sehingga dapat mempengaruhi kinerja model karena model tidak dapat memproses data yang tidak lengkap seperti pada Gambar 3. Oleh karena itu, perlu dilakukan penghilangan missing value dengan cara menghapus kolom yang berisi missing value.

Unnamed: 0		Category	Age	Sex	ALB	ALP	ALT	AST	BIL	CHE	CHOL	CREA	GGT	PROT
0	1	0=Blood Donor	32	m	38.5	52.5	7.7	22.1	7.5	6.93	3.23	186.0	12.1	69.0
1	2	0=Blood Donor	32	m	38.5	70.3	18.0	24.7	3.9	11.17	4.80	74.0	15.6	76.5
2	3	0=Blood Donor	32	m	46.9	74.7	36.2	52.6	6.1	8.84	5.20	86.0	33.2	79.3
3	4	0=Blood Donor	32	m	43.2	52.0	30.6	22.6	18.9	7.33	4.74	80.0	33.8	75.7
4	5	0=Blood Donor	32	m	39.2	74.1	32.6	24.8	9.6	9.15	4.32	76.0	29.9	68.7

Gambar 3. (a) Dataset Sebelum Diatasi

	Category	Age	Sex	AST	BIL	CHE	CREA	GGT
0	0=Blood Donor	32	m	22.1	7.5	6.93	186.0	12.1
1	0=Blood Donor	32	m	24.7	3.9	11.17	74.0	15.6
2	0=Blood Donor	32	m	52.6	6.1	8.84	86.0	33.2
3	0=Blood Donor	32	m	22.6	18.9	7.33	80.0	33.8
4	0=Blood Donor	32	m	24.8	9.6	9.15	76.0	29.9

Gambar 3. (b) Dataset setelah diatasi Missing Value

Tujuan dari menghilangkan *Missing Value* adalah untuk memudahkan model CNN dalam melakukan proses klasifikasi dan data yang tidak diperlukan akan dibuang atau dihapus, sehingga lebih efisien dan optimal. Tabel dengan nama "Unnamed" dihapus dan beberapa missing value pada fitur dihapus agar lebih optimal dalam memproses dataset yang diberikan.

2. Mengganti nilai Kategori menjadi Angka

Mengganti nilai Kategori menjadi angka adalah proses dalam algoritma CNN yang mengubah nilai kategorikal dalam data menjadi nilai angka untuk memudahkan

pemrosesan oleh model CNN. Hal ini dikarenakan model CNN membutuhkan data numerik untuk diproses dan dipelajari. Namun, pada data medis, terdapat banyak variabel kategorikal seperti jenis kelamin, jenis obat, dan lain sebagainya. Oleh karena itu, sebelum data medis dapat diproses oleh model CNN, nilai kategorikal dari variabel tersebut perlu diubah menjadi nilai numerik. Sebagai contoh, pada Gambar 3, variabel kategorikal “donor darah”, “Hepatitis”, “Fibrosis”, dan “Chirrosis” diubah menjadi nilai numerik “0”, “1”, “2”, dan “3”.

	Category	Age	Sex	AST	BIL	CHE	CREA	GGT
0	0=Blood Donor	32	m	22.1	7.5	6.93	106.0	12.1
1	0=Blood Donor	32	m	24.7	3.9	11.17	74.0	15.6
2	0=Blood Donor	32	m	52.6	6.1	8.84	86.0	33.2
3	0=Blood Donor	32	m	22.6	18.9	7.33	80.0	33.8
4	0=Blood Donor	32	m	24.8	9.6	9.15	76.0	29.9

Gambar 4. (a) Dataset Sebelum Mengganti Category

	Category	Age	Sex	AST	BIL	CHE	CREA	GGT
0	0	32	1	22.1	7.5	6.93	106.0	12.1
1	0	32	1	24.7	3.9	11.17	74.0	15.6
2	0	32	1	52.6	6.1	8.84	86.0	33.2
3	0	32	1	22.6	18.9	7.33	80.0	33.8
4	0	32	1	24.8	9.6	9.15	76.0	29.9

Gambar 4. (a) Dataset Sebelum Mengganti Category, (b) Dataset Setelah Mengganti Category

Model algoritma yang digunakan hanya dapat memproses data numerik. Oleh karena itu, data yang bukan berupa angka akan diubah menjadi angka. Sebagai contoh, pada Gambar 4, variabel kategori “donor darah”, “Hepatitis”, “Fibrosis”, dan “Chirrosis” diubah menjadi nilai numerik “0”, “1”, “2”, dan “3”. Dan juga mengubah variabel jenis kelamin “pria” dan “wanita” menjadi bentuk numerik untuk memudahkan pemrosesan. Variabel-variabel tersebut diubah menjadi “1” dan “0”. Serta variabel jenis kelamin “pria” dan “wanita” diubah menjadi bentuk numerik agar lebih mudah diproses. Diubah menjadi “1” dan “0”.

C. Pemodelan

Convolutional Neural Network(CNN) merupakan jaringan saraf tiruan umpan-maju yang canggih dan mendalam, yang dicirikan oleh tidak adanya siklus di antara simpul-simpulnya. Pola yang dihasilkan oleh arsitektur ini dapat bermanifestasi sebagai ekspresi, yang memungkinkan CNN mengenali pola-pola ini dalam kalimat tertentu, terlepas dari posisi leksikal kata-katanya. CNN menggunakan tiga jenis lapisan yang berbeda: lapisan konvolusional, lapisan penggabungan, dan lapisan terhubung penuh. Lapisan konvolusional terdiri dari sekumpulan filter berukuran tetap yang digunakan untuk mengonvolusi data masukan. Keluaran yang dihasilkan dari lapisan konvolusional disebut sebagai peta fitur. Persamaan berikut menggambarkan operasi konvolusional:

$$FM_{\{a,b\}} = bias + \sum_c^C Z_{c,d} + X_{\{a+c-1,b+d-1\}} \dots \dots \dots (1)$$

Pooling layer memastikan bahwa jaringan saraf berfokus secara eksklusif pada pola-pola yang paling menonjol, secara efektif memadatkan data dengan melintasi jendela di antara peta fitur dan kemudian menjalankan serangkaian transformasi linear atau non-linear pada data

yang terdapat di dalam jendela tersebut. Lebih lanjut, lapisan penggabungan mengurangi dimensionalitas peta fitur untuk memudahkan penerapannya di lapisan berikutnya.

$$f_{h(0,FM_{\{a,b\}})} = \max(0, FM_{\{a,b\}}) = \begin{cases} FM_{\{a,b\}}, & \text{jika } FM_{\{a,b\}} \geq 0, \\ 0, & \text{jika } FM_{\{a,b\}} < 0 \end{cases} \dots \dots \dots (2)$$

Layer terakhir yang digunakan adalah lapisan terhubung penuh. Lapisan ini berfungsi untuk menguraikan pola-pola yang diuraikan oleh lapisan anteseden. Neuron-neuron di dalam lapisan ini mempertahankan konektivitas penuh dengan semua aktivasi yang ada di lapisan sebelumnya. Metodologi jaringan saraf konvolusional juga menggabungkan fungsi aktivasi yang diimplementasikan antara lapisan konvolusional dan lapisan penyatuan. Fungsi aktivasi yang digunakan antara kedua lapisan ini adalah fungsi aktivasi ReLU. Fungsi aktivasi yang digunakan dalam fase keluaran adalah fungsi softmax. Representasi matematis dari fungsi aktivasi ReLU digambarkan dalam Persamaan 2.

Fungsi aktivasi softmax dirancang untuk menghasilkan hasil klasifikasi sekaligus menghasilkan nilai yang dapat diinterpretasikan sebagai probabilitas tak ternormalisasi yang sesuai dengan setiap kelas [12]. Nilai yang terkait dengan setiap kelas, yang diperoleh melalui penerapan fungsi softmax, digambarkan dalam Persamaan 3.

$$y_{\{ijk\}} = \frac{e^{x_{ijk}}}{\sum_{t=1}^D e^{x_{ijk}}} \dots \dots \dots (3)$$

Fungsi akhir yang dipertimbangkan adalah fungsi kerugian yang dirancang untuk mengkuantifikasi kerugian (yaitu, nilai kesalahan) melalui penerapan entropi silang kategoris. Persamaan 4 menggambarkan fungsi kerugian spesifik yang direferensikan.

$$L_{log}(Y, Y_{pred}) = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C 1_{y_i \in C_c} \log p_{model}[y_i \in C_c] \dots \dots \dots (4)$$

D. Evaluasi

Proses menggunakan model yang telah dilatih untuk memprediksi output untuk data input baru. Dalam CNN, model terdiri dari serangkaian lapisan yang menerima input dan menghasilkan output. Setelah model dilatih dengan data pelatihan, model tersebut dapat digunakan untuk memprediksi output untuk data yang belum pernah dilihat sebelumnya.

Hasilnya akan ditampilkan dalam bentuk hasil prediksi dari akurasi yang telah ditentukan dari data latih dan data uji. Semua metrik tersebut dihitung dengan rumus di bawah ini:

$$accuracy = (TP + TN)/(TP + TN + FP + FN) \dots \dots \dots (5)$$

$$precision = TP/(TP + FP) \dots \dots \dots (6)$$

$$recall = TP/(TP + FN) \dots \dots \dots (7)$$

$$F1\ Score = \frac{2 \times (\text{recall} \times \text{precision})}{(\text{recall} + \text{precision})} \dots\dots\dots(8)$$

Dalam konteks klasifikasi statistik, TP menandakan Positif Benar, yang menggambarkan kuantitas kejadian positif yang secara akurat diidentifikasi sebagai positif, TN menandakan Negatif Benar, yang mengacu pada jumlah kejadian negatif yang secara tepat diklasifikasikan sebagai negatif, FP menandakan Positif Palsu, yang menunjukkan jumlah kejadian positif yang secara keliru diklasifikasikan sebagai negatif, dan FN berhubungan dengan Negatif Palsu, yang menggambarkan kuantitas kejadian negatif yang secara keliru diklasifikasikan sebagai positif.

III. HASIL DAN PEMBAHASAN

A. Hasil

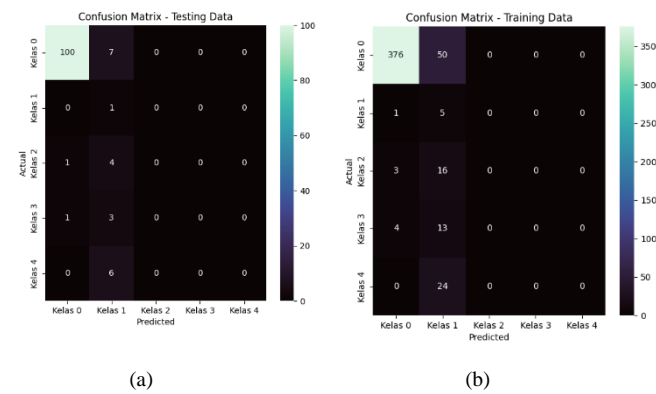
Bagian ini menyajikan temuan utama dari penelitian ini, termasuk analisis data dan interpretasi hasil untuk menjawab pertanyaan penelitian. Hasil akurasi data training dan testing yang diperoleh dari penggunaan model CNN 1D dengan struktur lapisan konvolusi yang telah ditentukan disajikan pada Tabel 1.

Tabel 1. Hasil Prediksi Pada Data Testing				
	Precision	Recall	F1-score	Support
Kelas 0	1.00	0.64	0.78	96
Kelas 1	0.05	1.00	0.09	3
Kelas 2	0.00	0.00	0.00	9
Kelas 3	0.00	0.00	0.00	6
Kelas 4	0.00	0.00	0.00	9
Accuracy			0.52	123
macro avg	0.21	0.33	0.17	123

Berdasarkan hasil evaluasi, model CNN 1D menunjukkan kemampuan yang cukup baik dalam mengklasifikasikan kelas mayoritas, yaitu Blood Donor, dengan nilai precision sebesar 1.00 dan f1-score sebesar 0.78. Namun, model gagal mengenali kelas-kelas minoritas seperti Hepatitis, Fibrosis, dan Cirrhosis, yang ditunjukkan dengan nilai precision, recall, dan f1-score sebesar 0.00 pada ketiga kelas tersebut. Akurasi keseluruhan sebesar 52% menunjukkan bahwa sebagian besar prediksi benar berasal dari dominasi kelas mayoritas, bukan dari kemampuan model dalam membedakan semua kelas secara merata. Hal ini mengindikasikan bahwa model masih cenderung bias terhadap kelas mayoritas dan memiliki kinerja yang belum optimal dalam klasifikasi multi-kelas yang tidak seimbang.

Hasil akurasi data training dan testing yang yg sudah di hypermeter tuning diperoleh dari penggunaan model CNN 1D dengan struktur lapisan konvolusi yang telah ditentukan disajikan pada Gambar 5, Tabel 2 dan 3.

Tabel 2. Hasil Prediksi Pada Data Testing				
	Precision	Recall	F1-score	Support
Kelas 0	0.97	0.93	0.95	426
Kelas 1	0.06	0.83	0.11	6
Kelas 2	0.00	0.00	0.00	19
Kelas 3	0.00	0.00	0.00	17
Kelas 4	0.00	0.00	0.00	24
Accuracy			0.81	492
macro avg	0.21	0.35	0.21	492
weighted avg	0.84	0.81	0.82	492



Gambar 5 . Hasil Confusion Matrix prediksi pada Data Test dan Train

Tabel 3. Hasil Prediksi Pada Data Train				
	Precision	Recall	F1-score	Support
Kelas 0	0.98	0.95	0.97	107
Kelas 1	0.05	1.00	0.10	1
Kelas 2	0.00	0.00	0.00	5
Kelas 3	0.00	0.00	0.00	4
Kelas 4	0.00	0.00	0.00	6
Accuracy			0.84	123
macro avg	0.21	0.39	0.21	123
weighted avg	0.85	0.84	0.84	123

Pada penelitian ini, model CNN 1D menunjukkan performa yang masih belum optimal dalam melakukan klasifikasi penyakit hepatitis C. Berdasarkan hasil evaluasi pada data uji (Tabel I), model memperoleh akurasi sebesar 83.74% yang didominasi oleh kinerja pada kelas 0 (Blood Donor) dengan nilai precision sebesar 0.98, recall 0.95, dan f1-score 0.97. Namun, pada kelas-kelas lainnya (kelas 1 hingga 4), nilai precision, recall, dan f1-score sangat rendah, bahkan bernilai 0.00 pada sebagian besar kelas, menunjukkan bahwa model gagal mengenali dan mengklasifikasikan sampel dari kelas minoritas.

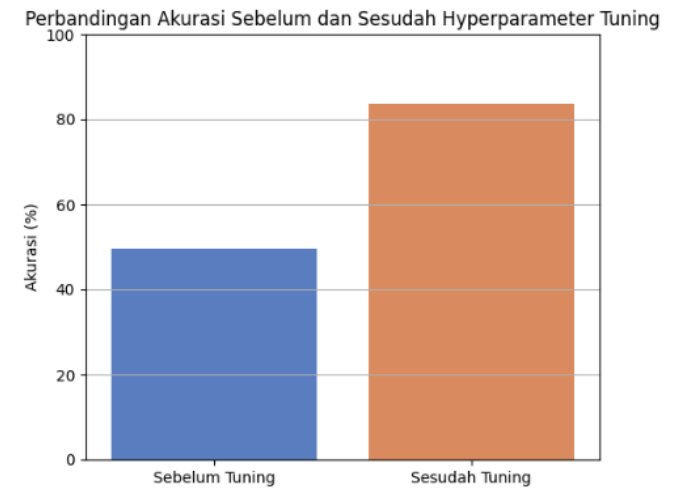
Meskipun data latih yang digunakan adalah 123 sampel atau sekitar 20% dari total dataset, hasil ini menunjukkan bahwa model CNN 1D masih mengalami kesulitan dalam membedakan kelas selain kelas 0. Ini mengindikasikan adanya kemungkinan ke tidak seimbangan kelas atau overfitting pada kelas mayoritas (Blood Donor). Oleh karena itu, diperlukan strategi peningkatan performa model, seperti penyeimbangan data atau penyesuaian parameter model agar kinerja model dapat merata pada setiap kelas.

Sementara itu, pada data latih (Tabel II), model juga menunjukkan kecenderungan yang sama, dengan akurasi hanya 81.30%, serta f1-score tertinggi tetap berada pada kelas 0. Hal ini menandakan bahwa model cenderung hanya belajar dari pola data mayoritas dan mengalami kesulitan dalam membedakan kelas-kelas lain, yang diperkuat oleh nilai macro average f1-score sebesar 0.21 pada data testing dan 0.21 pada data training.

Meskipun data latih berjumlah 492 sampel (80% dari total dataset), hasil ini mengindikasikan bahwa ketidakseimbangan kelas berperan besar terhadap kinerja model. Oleh karena itu, diperlukan upaya lebih lanjut dalam meningkatkan performa model, seperti melakukan penyeimbangan data, penyesuaian arsitektur model, atau penerapan teknik regularisasi dan augmentasi data, agar model mampu melakukan generalisasi secara lebih merata ke seluruh kelas.

B. Pembahasan

Dalam penelitian ini, dilakukan pelatihan model *Convolutional Neural Network (CNN) 1D* untuk klasifikasi penyakit Hepatitis C. Model diuji dalam dua tahap, yaitu sebelum dan sesudah dilakukan *hyperparameter tuning*. *Tuning* dilakukan menggunakan Keras Tuner dengan pendekatan *Random Search* untuk menemukan kombinasi arsitektur model yang optimal. Evaluasi dilakukan berdasarkan akurasi pada data uji (*testing*) guna menilai sejauh mana model mampu menggeneralisasi pola dari data latih.



Gambar 6. Perbandingan sebelum dan sesudah *Hyperparameter Tuning*

Berdasarkan gambar di atas, terlihat bahwa model sebelum *tuning* hanya mencapai akurasi sebesar 50%, yang menunjukkan bahwa model belum mampu belajar secara optimal dari data. Setelah dilakukan proses *tuning* terhadap beberapa parameter penting seperti jumlah filter pada *Conv1D*, ukuran kernel, jumlah neuron pada dense layer, tingkat *dropout*, dan *learning rate*, akurasi model meningkat secara signifikan menjadi 83.74%. Peningkatan ini menandakan bahwa proses *hyperparameter tuning* berhasil meningkatkan kinerja model dalam mendeteksi dan mengklasifikasikan kategori pasien berdasarkan data yang tersedia. Dengan demikian, *tuning* berperan penting dalam mengoptimalkan struktur dan parameter model untuk menghasilkan prediksi yang lebih akurat dan seimbang.

Perbandingan Hasil

Berdasarkan hasil yang telah dijelaskan sebelumnya, model *CNN 1D* yang dibangun untuk mengklasifikasikan pasien Hepatitis C menunjukkan performa yang belum merata di seluruh kelas. Model ini berhasil mencapai akurasi sebesar 83.74% pada data *testing* dan 81.30% pada data *training*, namun akurasi tinggi tersebut terutama disumbang oleh performa pada kelas mayoritas, yaitu kelas 0 (*Blood Donor*). Sementara itu, pada kelas-kelas lainnya, model gagal melakukan klasifikasi yang akurat, ditunjukkan oleh nilai *precision*, *recall*, dan *f1-score* yang sangat rendah bahkan mencapai 0.00. Hal ini menunjukkan bahwa model masih belum mampu mengenali pola dari kelas minoritas dan cenderung hanya belajar dari pola yang dominan dalam data.

Tabel 4. Perbandingan Hasil penelitian

Author	Algorithm	Accuracy
Sharfina, N., & Ramadhan, N. G [10]	Random Forest dan Naive bayes	98% dan 89%.
Charisma, R. A [11]	Decision Tree	93%
Drobo et al [12]	Artificial Neural Network	97,78%.
Metode yang diusulkan	Convolutional Neural Network	83.74%

Berdasarkan hasil perbandingan pada Tabel 3, dapat dilihat bahwa akurasi yang diperoleh dengan menggunakan metode *CNN 1D* sebagai metode yang diusulkan masih lebih rendah dibandingkan dengan metode-metode sebelumnya seperti *Random Forest*, *Naive Bayes*, *Decision Tree*, dan *Artificial Neural Network*. Hal ini menunjukkan bahwa performa model *CNN 1D* perlu ditingkatkan, baik dari segi arsitektur, parameter, maupun penanganan data untuk menghasilkan klasifikasi yang lebih akurat.

IV. KESIMPULAN

Berdasarkan hasil penelitian, model *CNN 1D* yang dikembangkan mampu melakukan klasifikasi terhadap data pasien Hepatitis C dengan akurasi sebesar 83.74% pada data pelatihan dan 81.30% pada data pengujian. Hasil ini menunjukkan bahwa proses *hyperparameter tuning* berhasil meningkatkan performa model secara signifikan dibandingkan sebelum *tuning*, yang hanya mencapai akurasi sekitar 52%. Meskipun demikian, model masih cenderung bias terhadap kelas mayoritas seperti *Blood Donor*, dengan *precision* dan *f1-score* tinggi, sementara kelas minoritas seperti Hepatitis, Fibrosis, dan *Cirrhosis* tetap sulit dikenali dengan baik. Hal ini disebabkan oleh distribusi data yang tidak seimbang antar kelas. Oleh karena itu, meskipun pendekatan *CNN 1D* terbukti potensial untuk klasifikasi medis berbasis data tabular, dibutuhkan pengembangan lebih lanjut seperti penyeimbangan kelas atau penggunaan arsitektur yang lebih kompleks untuk meningkatkan akurasi dan generalisasi model secara menyeluruh

V. SARAN

Meskipun hasil yang diperoleh cukup baik, masih ada ruang untuk perbaikan melalui eksplorasi arsitektur yang lebih kompleks dan penggunaan dataset yang lebih besar. Dengan demikian, penelitian ini diharapkan dapat berkontribusi dalam pengembangan sistem diagnosis berbasis AI yang lebih efisien dan akurat.

REFERENSI

- [1] M. Martinello, S. S. Solomon, N. A. Terrault, and G. J. Dore, "Hepatitis C," *The Lancet*, vol. 402, no. 10407, pp. 1085–1096, Sep. 2023, doi: 10.1016/S0140-6736(23)01320-X.
- [2] M. Fasano, F. Ieva, M. Ciarallo, B. Caccianotti, and T. A. Santantonio, "Acute Hepatitis C: Current Status and Future Perspectives," *Viruses*, vol. 16, no. 11, p. 1739, Nov. 2024, doi: 10.3390/v16111739.
- [3] V. Lo Re, J. C. Price, S. Schmitt, N. Terrault, D. Bhattacharya, and A. Aronsohn, "The obstacle is the way: Finding a path to hepatitis C elimination," *Hepatology*, vol. 80, no. 1, pp. 3–7, Jul. 2024, doi: 10.1097/HEP.0000000000000807.
- [4] A. S. Baber, B. Suganthan, and R. P. Ramasamy, "Current advances in Hepatitis C diagnostics," *J Biol Eng*, vol. 18, no. 1, p. 48, Sep. 2024, doi: 10.1186/s13036-024-00443-2.
- [5] I. Argirion *et al.*, "Association between Immunologic Markers and Cirrhosis in Individuals from a Prospective Chronic Hepatitis C Cohort," *Cancers*, vol. 14, no. 21, p. 5280, Oct. 2022, doi: 10.3390/cancers14215280.
- [6] C. R. Maturana *et al.*, "Advances and challenges in automated malaria diagnosis using digital microscopy imaging with artificial intelligence tools: A review," *Front. Microbiol.*, vol. 13, p. 1006659, Nov. 2022, doi: 10.3389/fmicb.2022.1006659.
- [7] M. Darvishi *et al.*, "Nanodiagnosics in global eradication of hepatitis C virus," *Clinica Chimica Acta*, vol. 565, p. 120013, Jan. 2025, doi: 10.1016/j.cca.2024.120013.
- [8] G. Bulusu, K. E. C. Vidyasagar, M. Mudigonda, and M. J. Saikia, "Cancer Detection Using Artificial Intelligence: A Paradigm in Early Diagnosis," *Arch Computat Methods Eng*, vol. 32, no. 4, pp. 2365–2403, May 2025, doi: 10.1007/s11831-024-10209-0.
- [9] Y. Kumar, S. Gupta, R. Singla, and Y.-C. Hu, "A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis," *Arch Computat Methods Eng*, vol. 29, no. 4, pp. 2043–2070, Jun. 2022, doi: 10.1007/s11831-021-09648-w.
- [10] N. Sharfina and N. G. Ramadhan, "Analisis SMOTE Pada Klasifikasi Hepatitis C Berbasis Random Forest dan Naïve Bayes," *JOINTECS*, vol. 8, no. 1, p. 33, Jun. 2023, doi: 10.31328/jointecs.v8i1.4456.
- [11] R. A. Charisma, S. Pamungkas, R. A. Saputra, N. G. Ramadhan, and F. D. Adhinata, "Analisis Penerapan Metode Ensembled Learning Decision Tree Pada Klasifikasi Virus Hepatitis C," *JoSYC*, vol. 3, no. 4, pp. 405–409, Sep. 2022, doi: 10.47065/josyc.v3i4.2064.
- [12] A. Drobo *et al.*, "Application of artificial neural networks in diagnosis of Hepatitis C," in *2022 XXVIII International Conference on Information, Communication and Automation Technologies (ICAT)*, Sarajevo, Bosnia and Herzegovina: IEEE, Jun. 2022, pp. 1–5. doi: 10.1109/ICAT54566.2022.9811126.
- [13] Arya Satya Pratama, Suci Maela Sari, Maila Faiza Hj, Moh Badwi, and Mochammad Isa Anshori, "Pengaruh Artificial Intelligence, Big Data Dan Otomatisasi Terhadap Kinerja SDM Di Era Digital," *jupiman*, vol. 2, no. 4, pp. 108–123, Oct. 2023, doi: 10.55606/jupiman.v2i4.2739.
- [14] K. Surendro, M. I. C. Rachmatullah, and J. Santoso, "Improving 1d Convolutional Neural Network (1d Cnn) Performance in Processing Tabular Datasets Using Principal Component Analysis," Oct. 17, 2022. doi: 10.21203/rs.3.rs-2149472/v1.