

**Klasifikasi MIT-BIH Arrhythmia Database Metode Random
Forest dan CNN dengan Model ResNet-50: *A Systematic
Literature Review***

PROPOSAL PROGRAM INTERNSHIP I

Diajukan untuk memenuhi kelulusan mata kuliah Internship

Tahun akademik **2023/2024**

Disusun Oleh:

Nama : **M. RIZKY**

NPM : 1194021



Universitas Logistik & Bisnis Internasional

**PROGRAM DIPLOMA IV TEKNIK INFORMATIKA
UNIVERSITAS LOGISTIK & BISNIS INTERNASIONAL
BANDUNG**

2023

***MIT-BIH Classification of Arrhythmia Database Random Forest and CNN
Methods with the ResNet-50 Model: A Systematic Literature Review***

INTERNSHIP PROGRAM REPORT I

Submitted to allow Internship course graduation
Academic year **2023/2024**

Created By:

Name : **M. RIZKY**
NPM : 1194021



Universitas Logistik & Bisnis Internasional

***APPLIED BACHELOR PROGRAM OF INFORMATICS ENGINEERING
UNIVERSITY OF INTERNATIONAL LOGISTICS & BUSINESS
BANDUNG
2023***

SURAT PERNYATAAN TIDAK PLAGIAT

Bandung, 02 April 2023

Saudara yang bertanda tangan dibawah ini:

Nama : **M. RIZKY**

NPM : 1.19.4.021

Judul Laporan : PERBANDINGAN METODE RANDOM FOREST DAN
CNN DENGAN MODEL RESNET-50 PADA
KLASIFIKASI MIT-BIH ARRHYTHMIA DATABASE

Menyatakan bahwa saya tidak melakukan tindakan meniru, menjiplak ataupun menyalin skripsi karya ilmiah yang telah ada. Apabila kita terbukti melakukan tindakan tersebut, maka kita bersedia menerima sanksi yang diberikan sesuai dengan ketentuan yang ditetapkan dan berlaku di Program Studi Diploma IV Teknik Informatika Universitas Logistik & Bisnis Internasional (ULBI).

Bandung, 02 April 2023

Yang membuat pernyataan

M. RIZKY

NPM. 1.19.4.021

ABSTRAK

Machine Learning merupakan teknologi yang berkembang sangat pesat saat ini. Banyak sekali produk - produk digital yang memanfaatkan teknologi *Machine Learning* mulai dari bidang Transportasi, Kesehatan, dan lainnya. Seiring perkembangannya, sudah banyak perusahaan yang mengimplementasikan *Machine Learning* pada produk digitalnya. Pemanfaatan *Machine Learning* dibidang kesehatan memang sangat perlu dilakukan terlebih lagi pentingnya penanganan dini kepada pasien yang mengidap penyakit khususnya pada penyakit jantung. Dengan melakukan pendeteksian dini terhadap suatu penyakit, dapat meminimalisir perkembangan penyakit tersebut. Maka dari itu, pada penelitian ini akan membahas tentang bagaimana proses ekstraksi pada sinyal EKG dan melakukan klasifikasi terhadap sinyal tersebut berdasarkan kelas yang ditentukan dengan beberapa model yang dipakai. Setelah dilakukan klasifikasi, model yang digunakan untuk proses klasifikasi tersebut dilakukan komparasi dan evaluasi untuk menentukan model mana yang memiliki akurasi yang tinggi dan baik.

Kata Kunci: *Machine Learning*, Ekstraksi, EKG, Klasifikasi

ABSTRACT

Machine Learning is a technology that is developing very rapidly at this time. Lots of digital products that utilize Machine Learning technology, starting from the fields of Transportation, Health, and others. Along with its development, many companies have implemented Machine Learning in their digital products. The use of Machine Learning in the health sector really needs to be done, especially the importance of early treatment for patients with diseases, especially heart disease. By doing early detection of a disease, it can minimize the development of the disease. Therefore, in this study will discuss how the process of extracting ECG signals and classifying these signals based on the specified class with several models used. After classification, the models used for the classification process are compared and evaluated to determine which model has high accuracy and is good.

Keywords: *Machine Learning, Extraction, ECG, Classification*

KATA PENGANTAR

Dalam rangka menyelesaikan program studi yang akan saya jalani, saya ingin mengajukan proposal internship 2 dengan judul “Perbandingan Metode Random Forest dan CNN dengan model ResNet-50 pada Klasifikasi MIT-BIH Arrhythmia Database”.

Penelitian ini bertujuan untuk melakukan perbandingan antara model tradisional atau *Machine Learning* dengan model *Deep Learning* atau *Convolutional Neural Network (CNNs)* dalam melakukan klasifikasi Arrhythmia MIT-BIH Database.

Saya yakin bahwa hasil penelitian ini dapat memberikan kontribusi yang bermanfaat dalam mengembangkan kebijakan publik di Indonesia, khususnya dalam bidang kesehatan. Selain itu, melalui program internship ini, saya berharap dapat memperoleh pengalaman dan pengetahuan yang lebih luas dalam bidang penelitian dan pengembangan teknologi informasi.

Saya mengucapkan terima kasih atas perhatian dan dukungan yang diberikan, dan saya juga berharap proposal ini dapat diterima untuk dilaksanakan.

Bandung, 12 Mei 2023

M. RIZKY

DAFTAR ISI

ABSTRAK	i
<i>ABSTRACT</i>	ii
KATA PENGANTAR	iii
DAFTAR ISI	iv
BAB I	1
PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	2
1.3. Tujuan	2
1.4. Manfaat	3
1.5. Penelitian Sebelumnya	3
1.6. Sistematika Penulisan	4
BAB II	6
TINJAUAN PUSTAKA	6
2.1. State of the Art	6
2.2. MIT-BIH Arrhythmia Database	10
2.3. Elektrokardiogram (EKG)	10
2.4. Machine Learning	11
2.4.1. Supervised Learning	12
2.4.1.1. Random Forest	12
2.5. Deep Learning	13
2.5. Convolutional Neural Network (CNN)	13
BAB III	14
METODOLOGI PENELITIAN	14
3.1. Ruang Lingkup	14
3.2. Alur Metodologi Penelitian	14
3.3. Indikator Capaian Penelitian	15
3.3.1. Studi Literatur	16
3.3.2. Pengumpulan Data	16
3.3.3. Pra-Pemrosesan Data	17
3.3.4. Pemodelan	17
3.3.5. Evaluasi Model	17
BAB IV	18
PENUTUP	18
4.1. Kesimpulan dan Saran	18

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Artificial Intelligence saat ini menjadi salah satu perhatian dunia dikarenakan perkembangannya yang begitu sangat cepat. Seiring berkembangnya teknologi yang sangat cepat, banyak sekali jenis - jenis teknologi yang bermunculan untuk membantu bahkan menggantikan pekerjaan manusia salah satunya di bidang kesehatan. Contoh Artificial Intelligence yang sangat populer saat ini adalah ChatGPT atau Generative Pre-training Transformer yang mampu berinteraksi dengan manusia melalui chat. ChatGPT merupakan *Large Language Model (LLM)* yang menggunakan metode *Deep Learning* untuk dapat menghasilkan teks yang mirip manusia sebagai responnya. ChatGPT dirilis pada tanggal 30 November 2022 sebagai pengembangan teknologi terbaru dari OpenAI dari *Large Language Model* yang mampu melakukan percakapan yang cerdas [1]. Selain itu, salah satu cabang Artificial Intelligence sederhana dan banyak sekali dijadikan bahan pembelajaran adalah Machine Learning dan sampai saat ini masih terus berkembang pesat di kalangan programmer atau khususnya di dunia IT. Machine Learning sendiri terdiri atas 2 bagian yaitu Machine Learning Supervised dan Machine Learning Unsupervised. Sedangkan *Deep Learning*, merupakan kemajuan terbaru dari kecerdasan buatan yang telah digunakan mengatasi banyak masalah yang melibatkan kumpulan data besar atau big data.

Machine Learning Supervised adalah struktur dari suatu data yang hendak dianalisis telah ditentukan dahulu dan Machine Learning mencari data di struktur tersebut, sedangkan Machine Learning Unsupervised struktur dari suatu data dicari oleh Machine Learning itu sendiri [2]. Salah satu algoritma Supervised Learning yang sering digunakan pada proses klasifikasi adalah algoritma Random Forest (RF). RF adalah teknik bagging yang memiliki karakteristik signifikan

yang berjalan efisien pada dataset besar. Random forest dapat menangani ribuan variabel masukan tanpa penghapusan variabel dan memperkirakan fitur penting untuk klasifikasi.

Di dalam dunia medis, teknologi - teknologi banyak sekali diterapkan untuk memenuhi kebutuhan medis itu sendiri seperti AI pendeteksi pasien positif Covid atau tidak dengan memanfaatkan hembusan nafas dari pasien tersebut. Dan masih banyak lagi hal - hal yang bisa kita manfaatkan untuk mengembangkan teknologi di bidang kesehatan salah satunya adalah dengan memanfaatkan sinyal EKG atau Elektrokardiogram untuk mengklasifikasi penyakit gagal jantung pada pasien. EKG merupakan sebuah informasi sinyal yang digambarkan dalam bentuk diagram yang menampilkan informasi penting mengenai keadaan jantung manusia. Elektrokardiogram atau EKG adalah rekaman aktivitas listrik yang dihasilkan melalui siklus detak jantung [3].

1.2. Rumusan Masalah

1. Bagaimana proses ekstraksi fitur sinyal EKG pada dengan metode CNN?
2. Bagaimana proses implementasi metode yang akan digunakan untuk mengklasifikasi data pada sinyal EKG?
3. Metode yang memiliki akurasi paling tinggi pada klasifikasi sinyal EKG?

1.3. Tujuan

Dari penjelasan pada bagian latar belakang di atas, dapat kita pahami bahwa penulisan laporan ini memiliki tujuan antara lain:

1. Dapat memahami bagaimana proses ekstraksi fitur pada sinyal EKG dengan metode CNN.
2. Dapat memahami bagaimana proses implementasi metode yang dipilih dalam proses klasifikasi.
3. Dapat menentukan metode yang terbaik untuk proses klasifikasi sinyal EKG.

1.4. Manfaat

Manfaat yang bisa didapatkan pada penelitian ini adalah kita dapat memahami bagaimana proses dari kita melakukan ekstraksi ciri dari sinyal EKG lalu kemudian melakukan preprocessing data yang dimana proses tersebut harus kita lakukan agar data yang diproses klasifikasikan menghasilkan output yang akurat dengan hasil yang seharusnya.

1.5. Penelitian Sebelumnya

Beberapa topik penelitian sebelumnya yang berkaitan dengan penelitian sekarang:

Tabel 1.1 Penelitian sebelumnya

Topik Penelitian	Negara	Tahun	Model	Hasil	Skor
Klasifikasi sinyal EKG <i>sleep apnoea</i> [4]	Other	2021	CNN & LSTM	Akurasi Tertinggi	90.92%
Klasifikasi sinyal EKG <i>Arrhythmia</i> [5]	Other	2020	Time-series & DL	Akurasi Tertinggi	99.12%
Klasifikasi dan optimasi sinyal EKG <i>Arrhythmia</i> [6]	Other	2020	CNN & BaROA	Akurasi Tertinggi	93.19%
Deteksi <i>Ventricular Arrhythmia</i> dengan sinyal EKG [7]	Other	2021	Random Forest	Akurasi Tertinggi	94.81%
Deteksi <i>Congestive Heart Failure</i> dengan sinyal EKG [8]	Other	2019	HCMM	Akurasi Tertinggi	98.78%
Klasifikasi sinyal EKG <i>Arrhythmia</i> [9]	Other	2021	SVM	Akurasi Tertinggi	99.51%
Klasifikasi sinyal EKG <i>Arrhythmia</i> [10]	Other	2022	CNN & LSTM	Akurasi Tertinggi	99.7%
Klasifikasi sinyal EKG <i>Arrhythmia</i> [11]	Other	2021	XDM	Akurasi Tertinggi	93.6%
Klasifikasi sinyal EKG <i>Arrhythmia</i> [12]	Other	2021	CNN & SVM	Akurasi Tertinggi	96%
Klasifikasi sinyal EKG <i>Arrhythmia</i> [13]	Other	2022	CWT & CNN	Akurasi Tertinggi	99.65%

Klasifikasi sinyal EKG Covid-19 [14]	Other	2022	CNN & OWAE	Akurasi Tertinggi	95.29%
Deteksi <i>Coronary Artery, Myocardial Infarction & Congestive Heart Failure</i> dengan sinyal EKG [15]	Other	2021	Gaborn & CNN	Akurasi Tertinggi	98.74%
Deteksi <i>Congestive Heart Failure & Arrhythmia</i> dengan sinyal EKG [16]	Other	2020	<i>CNN & CQ-NSGT</i>	Akurasi Tertinggi	98.82%
Deteksi <i>Congestive Heart Failure</i> dengan sinyal EKG [17]	Other	2019	<i>QSVM</i>	Akurasi Tertinggi	99.66%
Deteksi <i>Congestive Heart Failure</i> dengan sinyal EKG [18]	Other	2020	<i>CNN</i>	Akurasi Tertinggi	100%
Deteksi <i>Heart Disease</i> dengan sinyal EKG [19]	Other	2021	<i>GAN-LSTM</i>	Akurasi Tertinggi	99.4%
Klasifikasi sinyal EKG <i>Arrhythmia</i> [20]	Other	2021	<i>CNN</i>	Akurasi Tertinggi	98.92%

1.6. Sistematika Penulisan

Penyajian laporan program Internship II dibagi menjadi beberapa bab. Pembagian bab tersebut adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini meliputi Latar Belakang, Identifikasi Masalah, Tujuan dan Manfaat, Ruang Lingkup Penelitian dan Sistematika Penulisan.

BAB II LANDASAN TEORI

Pada bab ini meliputi uraian dan sumber tentang penjelasan mengenai sistem dalam penelitian yang akan dibuat.

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan tentang metode yang akan digunakan pada proses penyelesaian program internship dan juga tahapan serta alur metode penelitian.

BAB IV HASIL PENELITIAN

Pada bab ini menjelaskan tentang hasil penelitian berdasarkan capaian yang ingin dicapai pada metodologi penelitian.

BAB V PENUTUP

Pada bab ini menjelaskan tentang hasil penelitian berdasarkan capaian yang ingin dicapai pada metodologi penelitian.

BAB II

TINJAUAN PUSTAKA

2.1. *State of the Art*

Mendeteksi awal gejala gagal jantung merupakan tindakan yang sangat penting dilakukan untuk menghindari terjadinya gejala yang lebih parah. Salah satu faktor yang mempengaruhi gagal jantung adalah *Arrhythmia*. Penyakit ini dapat menyebabkan kematian mendadak jika tidak segera ditangani dengan awal. Untuk itu, mendeteksi gejala awal gagal jantung dapat membantu menyelamatkan nyawa dari serangan gagal jantung [6]. Mendeteksi penyakit *Arrhythmia* menggunakan sinyal EKG merupakan salah satu tindakan awal untuk mendeteksi gejala awal gagal jantung. Perlu diketahui bahwa kelainan pada sinyal EKG merupakan gejala yang terputus - putus pada denyut yang berhubungan dengan *Arrhythmia* dalam waktu singkat. Untuk memantau sinyal EKG terdapat alat tradisional yang dihubungkan melalui kabel pada manusia itu sendiri dan juga tanpa menggunakan kabel untuk menghubungkannya [21].

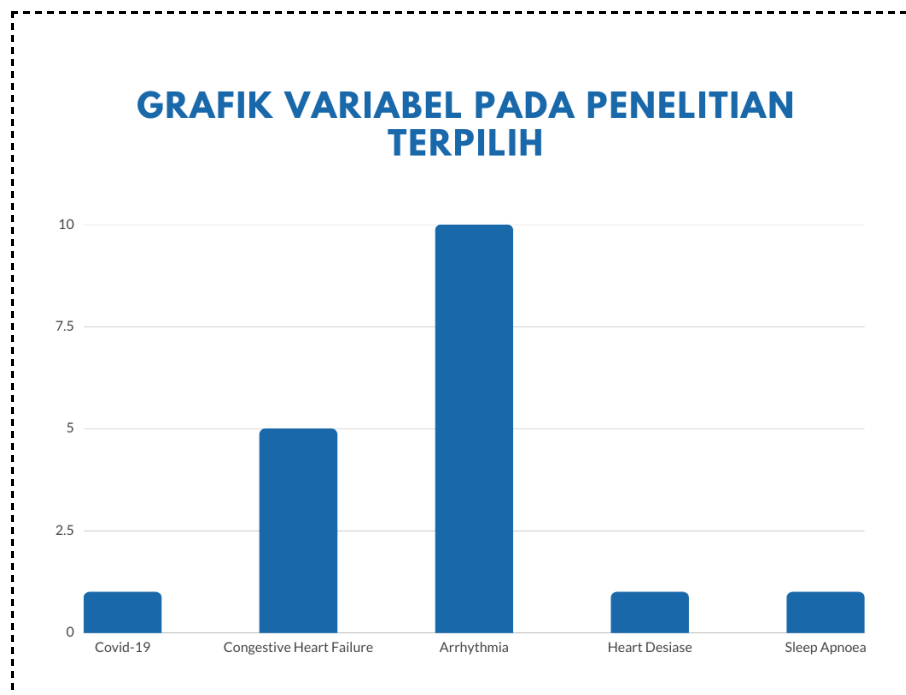
Berdasarkan studi literatur, kami mengumpulkan paper penelitian yang berhubungan dengan penelitian ini. Pada penelitian [4] menggunakan model *Hybrid* yaitu *Convolutional Neural Network* (CNN) dan *Long Short Term Memory* (LSTM) untuk mengklasifikasi sinyal EKG *Sleep Apnoea*. Pada penelitian ini menggunakan 2 model klasifikasi *Output Layer* dan keduanya bekerja dengan baik serta melakukan ekstraksi fitur otomatis. Selanjutnya adalah penelitian yang dilakukan oleh [5] dengan melakukan klasifikasi sinyal EKG *Arrhythmia*. Pada penelitian tersebut menggunakan model *Time-series Augmented* dan *Deep Learning* untuk klasifikasinya. Penggunaan *Augmented* pada proses pengklasifikasian bukan hanya membuat pelatihan model menjadi akurat, akan tetapi dapat menstabilkan pada akurasi yang lebih tinggi.

Tabel 2. Evaluasi Model Penelitian Terpilih

Paper	Variabel	Model	Pendekatan	Akurasi	Presisi	Recall	F1
H. Almutairi et al. [4]	<i>Sleep Apnoea</i>	CNN & LSTM	<i>Hybrid</i>	94.27%	94.57	93.93%	95.41%
P. Kanani & M. Padole [5]	<i>Arrhythmia</i>	Time-series & DL	<i>Hybrid</i>	99.12%	-	-	98%
D. K. Atal & M. Singh [6]	<i>Arrhythmia</i>	CNN & BaROA	<i>Hybrid</i>	93.19%	95%	93.98%	-
S. Mandal et al. [7]	<i>Ventricular Arrhythmia</i>	Random Forest	<i>Machine Learning</i>	97.02%	90.97%	97.89%	-
R. K. Tripathy et al. [8]	<i>Congestive Heart Failure</i>	HCMM	<i>Machine Learning</i>	98.78%	98.48%	99.09%	-
J. Rahul et al. [9]	<i>Arrhythmia</i>	SVM	<i>Machine Learning</i>	99.51%	99.28%	99.63%	-
S. Sowmya & D. Jose [10]	<i>Arrhythmia</i>	CNN & LSTM	<i>Hybrid</i>	99.7%	96.6%	95.6%	-
Y.-Y. Jo et al. [11]	<i>Arrhythmia</i>	XDM	<i>Deep Learning</i>	-	93.6%	79.3%	85.2%
L. Zheng et al. [12]	<i>Arrhythmia</i>	CNN & SVM	<i>Hybrid</i>	96%	-	-	-
S. C. Mohonta et al. [13]	<i>Arrhythmia</i>	CWT & CNN	<i>Hybrid</i>	99.65%	98.87%	99.85%	-
K. Prashant et al. [14]	<i>Covid-19</i>	CNN & OWAE	<i>Hybrid</i>	95.29%	95.5%	95.3%	95.4%
V. Jahmunah et al. [15]	<i>Coronary Artery, Myocardial Infarction & Congestive</i>	Gaborn & CNN	<i>Deep Learning</i>	98.74%	98.74%	99.46%	-

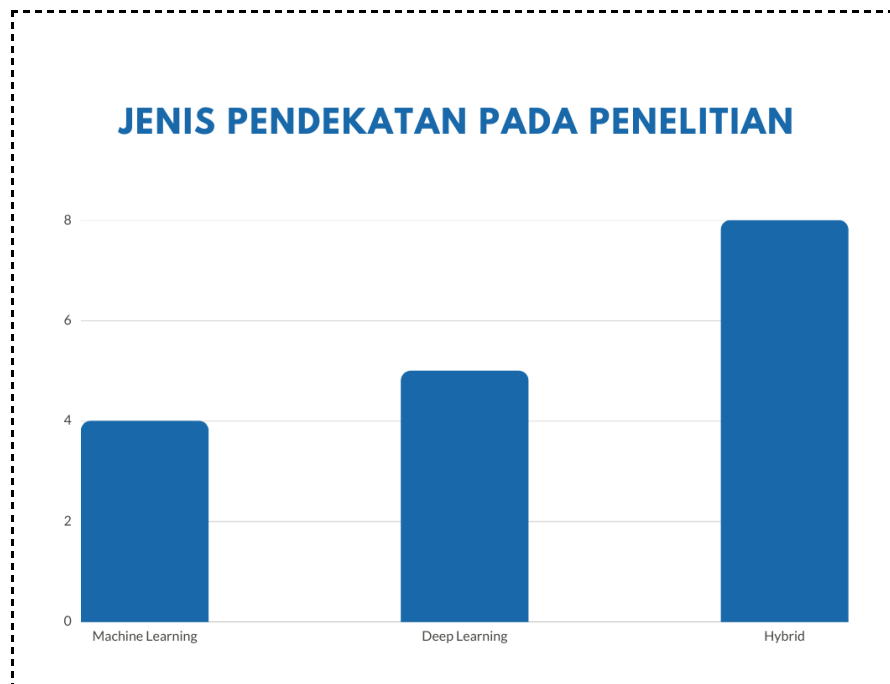
	<i>Heart Failure</i>						
A. S. Eltrass et al. [16]	<i>Congestive Heart Failure & Arrhythmia</i>	<i>CNN & CQ-NSG T</i>	<i>Hybrid</i>	98.82%	99.20%	99.21%	-
A. A. Bhurane et al. [17]	<i>Congestive Heart Failure</i>	<i>QSVM</i>	<i>Machine Learning</i>	99.66%	99.82%	99.28%	-
M. Porumb et al. [18]	<i>Congestive Heart Failure</i>	<i>CNN</i>	<i>Deep Learning</i>	100%	-	-	-
A. Rath et al. [19]	<i>Heart Disease</i>	<i>GAN-LSTM</i>	<i>Deep Learning</i>	99.4%	-	-	98.7%
A. Pal et al. [20]	<i>Arrhythmia</i>	<i>CNN</i>	<i>Deep Learning</i>	98.92%	-	-	-

Setelah dilakukan pengelompokan penelitian terkait, peneliti telah menggambarkan melalui grafik mengenai *variable* dalam penelitian terkait tersebut pada Gambar 2.1, untuk pendekatan pada penelitian terkait akan digambarkan pada Gambar 2.2.



Gambar 2.1 Grafik Variabel Penelitian Terpilih

Pengelompokan variabel pada penelitian terkait untuk menentukan variabel mana yang banyak digunakan untuk proses klasifikasi tersebut. Selanjutnya adalah dengan menentukan jenis pendekatan yang digunakan dalam proses klasifikasi dengan masing - masing model yang akan digunakan. Gambar 2.2 akan menampilkan grafik jenis pendekatan pada penelitian terkait.



Gambar 2.2. Grafik Jenis Pendekatan Penelitian

Berdasarkan jenis pendekatan yang digambarkan pada grafik diatas, pendekatan yang banyak digunakan adalah *Hybrid* dengan total 8 penelitian terkait. Pada Tabel 2, untuk menjawab pertanyaan di awal model yang memiliki performa paling baik dan stabil adalah *Convolutional Neural Network (CNN)* dengan mengevaluasi model yang paling banyak digunakan adalah Akurasi model tersebut.

Adapun teori - teori yang akan digunakan dalam penelitian ini untuk membangun model dan dilakukannya suatu perbandingan antara metode tradisional atau *Machine Learning* menggunakan *Random Forest* dengan *Convolutional Neural Network (CNN)* menggunakan arsitektur ResNet-50 adalah sebagai berikut:

2.2. MIT-BIH Arrhythmia Database

MIT-BIH Arrhythmia Database adalah rangkaian uji standar yang umumnya tersedia untuk mengevaluasi aritmia deteksi. Sejak 1980, basis data ini telah digunakan untuk dasar penelitian untuk dinamika jantung di sekitar 500 lokasi di seluruh dunia. Basis data ini sebagian besar digunakan untuk tujuan medis dan penelitian dari deteksi dan analisis aritmia jantung yang berbeda. Basis data ini mencoba menyediakan informasi yang tepat untuk mendeteksi aritmia ventrikel [22].

Aritmia adalah perubahan detak jantung yang tidak normal karena detak jantung yang tidak tepat yang menyebabkan kegagalan dalam pemompaan darah. Aritmia dapat menyebabkan kematian jantung mendadak. Gejala aritmia yang umum adalah denyut prematur, jantung berdebar, pusing, kelelahan, dan pingsan. Aritmia lebih sering terjadi pada orang yang menderita tekanan darah tinggi, diabetes dan arteri koroner penyakit [23]. Sinyal Elektrokardiogram yang akan digunakan pada penelitian ini diambil dari *MIT-BIH Arrhythmia Database*.

2.3. Elektrokardiogram (EKG)

Elektrokardiogram (EKG) adalah tes medis yang mengukur aktivitas listrik jantung. EKG digunakan untuk mendiagnosis dan memantau berbagai kondisi jantung, seperti serangan jantung, aritmia, dan gagal jantung. EKG akan merekam aktivitas listrik kecil yang dihasilkan oleh jantung selama periode waktu tertentu dengan menempatkan elektroda pada tubuh pasien [24]. Rekaman EKG berisi *noise* dan amplitudo yang bervariasi dari setiap orang sehingga sulit dalam proses mendiagnosis [25].

Elektrokardiogram (EKG) memberikan informasi penting tentang berbagai kondisi manusia [26]. Untuk melakukan EKG, petugas kesehatan akan menempelkan tambalan kecil dan lengket yang disebut elektroda ke dada, lengan, dan kaki pasien. Elektroda terhubung ke mesin EKG, yang merekam sinyal listrik yang dihasilkan oleh jantung saat bergerak ke seluruh tubuh. Mesin

tersebut menghasilkan jejak aktivitas listrik jantung, yang disebut strip EKG, yang kemudian diinterpretasikan oleh petugas kesehatan.

Machine Learning dapat diterapkan pada analisis data EKG untuk meningkatkan akurasi dan efisiensi diagnosis dan pengobatan. Misalnya, algoritma pembelajaran mesin dapat dilatih untuk mengenali pola dalam data EKG yang menunjukkan kondisi jantung tertentu. Algoritma ini kemudian dapat digunakan untuk menganalisis data EKG secara otomatis dan memberikan rekomendasi diagnosis atau pengobatan. Algoritma pembelajaran mesin dapat dilatih untuk mengklasifikasikan data EKG ke dalam kategori yang berbeda, seperti normal atau abnormal, atau untuk mengidentifikasi kondisi jantung tertentu.

Biasanya, klasifikasi sinyal EKG memiliki empat fase: *preprocessing*, segmentasi, ekstraksi fitur, dan klasifikasi. Fase *preprocessing* terutama ditujukan untuk mendeteksi dan melemahkan frekuensi sinyal EKG yang terkait dengan artefak, yang juga biasanya melakukan normalisasi dan peningkatan sinyal. Setelah *preprocessing*, segmentasi akan membagi sinyal menjadi segmen yang lebih kecil, yang dapat mengekspresikan aktivitas listrik jantung dengan lebih baik [25].

2.4. Machine Learning

Pembelajaran Mesin atau *Machine Learning* merupakan kemajuan teknologi yang penting karena dapat membantu dalam mengambil keputusan dengan mekanisme prediksi dan klasifikasi berdasarkan data yang ada [27]. Berfokus pada *performance* yang tinggi, teknik pembelajaran mesin atau *machine learning* diterapkan pada bisnis dengan data yang berkembang pesat. Karena pendekatan desain cocok untuk komunikasi komputasi paralel dan terdistribusi yang berevolusi atau data bisnis yang dinamis dan berkembang kedalam model *Machine Learning* [28].

Teknologi berbasis komputer modern banyak yang telah menggunakan Pembelajaran Mesin atau *Machine Learning*. *Machine Learning* merupakan

cabang dari Kecerdasan Buatan atau *Artificial Intelligence* yang luas dan sudah berkembang pesat saat ini yang memungkinkan komputer untuk belajar dan berkembang secara otomatis tanpa harus diprogram secara eksplisit. Teknologi ini berasal dari mempelajari pengenalan pola dan teori pembelajaran komputasi. Secara umum, metode pembelajaran yang umum digunakan oleh *Machine Learning* dapat diklasifikasikan menjadi *Supervised*, *Unsupervised*, dan *Reinforcement Learning* [29].

2.4.1. *Supervised Learning*

Supervised Learning merupakan metode *Machine Learning* untuk menyimpulkan fungsi dari data *train*. Algoritma *Supervised Learning* biasanya berisi kumpulan sampel *input (feature)* dan label yang berkaitan dengan kumpulan data tersebut. Tujuan dari pengklasifikasian adalah untuk menemukan batas yang sesuai yang dapat memprediksi label yang benar pada data *test*. Secara singkat, dalam *Supervised Learning* memiliki setiap contoh data yang berpasangan yang terdiri dari objek masukan (*input*) dan objek keluaran (*output*) yang diinginkan. Algoritma *Supervised Learning* menganalisis data *train* dan menghasilkan fungsi (*model*) [30]. Salah satu contoh metode algoritma pada *Supervised Learning* adalah *Random Forest*.

2.4.1.1. *Random Forest*

Random Forest adalah metode *Machine Learning* yang diperkenalkan pada tahun 2001 oleh Leo Breiman. Metode ini menggunakan serangkaian besar dari *Decision Tree* dengan korelasi timbal balik yang rendah dan fitur yang dipilih secara acak menggunakan metode *bagging (Bootstrap AGGREGatING)* [31].

Random Forest merupakan salah satu metode pengklasifikasian terbaik dan banyak digunakan untuk regresi dan juga klasifikasi. *Random Forest* memiliki algoritma yang sederhana sehingga menjadi salah satu pilihan yang menarik untuk mengklasifikasi teks. Selain itu, *Random Forest* juga memiliki

kemampuan untuk mengolah data berdimensi tinggi dan memiliki performa yang tinggi walaupun menggunakan data yang banyak sehingga menjadi salah satu keuntungan menggunakan model ini dibandingkan dengan model *Machine Learning* lainnya [31].

Pemilihan model ini didasarkan karena pada faktanya bahwa *Random Forest* secara luas dianggap sebagai salah satu metode *Machine Learning* yang paling sukses dan banyak digunakan hingga saat ini [32].

2.5. *Deep Learning*

Deep Learning merupakan kemajuan terbaru dari kecerdasan buatan yang telah digunakan mengatasi banyak masalah yang melibatkan kumpulan data besar atau big data, Deep Learning juga muncul sebagai alat yang menjanjikan untuk analisis data scRNA-seq, karena memiliki kapasitas untuk mengekstrak fitur informatif dan ringkas data scRNA-seq yang berisik, heterogen, dan berdimensi tinggi untuk meningkatkan analisis hilir [33].

2.5. *Convolutional Neural Network (CNN)*

Convolutional Neural Networks (CNNs), telah berhasil digunakan untuk memproses dan menganalisa gambar digital, termasuk berbagai jenis gambar medis, termasuk CT. Meskipun beberapa karya telah mempresentasikan hasil menggunakan gambar sinar-X, CT scan cenderung lebih efisien karena dua alasan. Pertama, CT scan memberikan tampilan 3 dimensi yang mendetail dari organ yang didiagnosis, sedangkan sinar-X memberikan tampilan 2-D. Kedua, CT scan tidak tumpang tindih dengan organ, sedangkan pada sinar-X, tulang rusuk tumpang tindih dengan paru-paru dan jantung [34].

BAB III

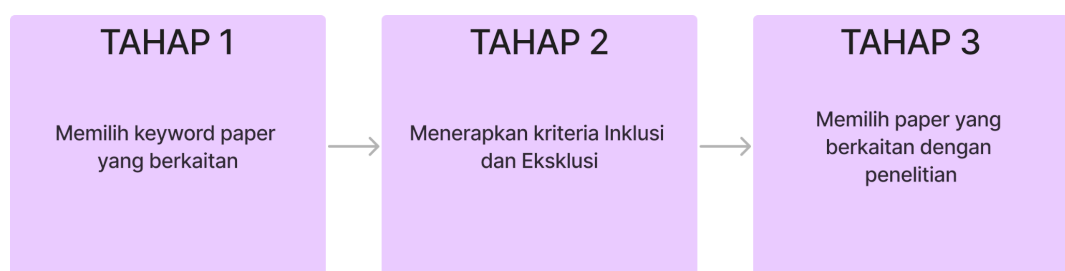
METODOLOGI PENELITIAN

3.1. Ruang Lingkup

Penelitian ini akan membahas tentang Perbandingan antara metode Random Forest dan CNN dengan model ResNet-50 pada klasifikasi arrhythmia database. Sumber data yang akan digunakan berasal dari situs PhysioNet yang merupakan Database Complex Physiologic Signals. Data yang akan diambil pada penelitian ini adalah MIT-BIH Arrhythmia Database dengan 48 record dan masing-masing durasi yang tersedia. Dari data-data tersebut akan digabungkan lalu dilakukan dengan masing-masing model.

3.2. Alur Metodologi Penelitian

Pada penelitian ini, metode akan menjadi hal penting dalam melakukan ekstraksi hingga klasifikasi sinyal karena akan mempengaruhi hasil klasifikasi atau output dari model tersebut. Oleh karena itu, alur metodologi yang tepat untuk mendapatkan hasil yang terbaik. Alur penelitian akan ditunjukkan pada gambar 3.1.



Gambar 3.1 Flow Diagram Metodologi Penelitian

3.2.1. Penentuan Database

Paper yang akan digunakan dapat dikumpulkan melalui beberapa website yang menyediakan *database* paper antara lain *Sciencedirect*, *Google Scholar*, dan *Scopus*.

3.2.2. Penentuan Keyword atau Kata Kunci

Setelah ditetapkannya beberapa website yang akan digunakan untuk mencari paper, maka peneliti perlu menentukan *keyword* atau kata kunci yang akan digunakan dalam mencari paper yang berkaitan dengan penelitian. Salah satu kata kunci yang digunakan dalam mencari paper adalah “*MIT-BIH Arrhythmia*” dan “*Classification ECG*”.

3.2.3. Penentuan Kriteria Inklusi dan Eksklusi

Pada saat mencari paper yang sesuai dengan penelitian menggunakan website yang dipilih, peneliti harus membatasi pencarian paper dengan ketentuan batasan 100 paper dengan menerapkan kriteria Inklusi dan Eksklusi pada kumpulan paper tersebut. Kriteria Inklusi merupakan karakteristik umum suatu penelitian yang akan diteliti berdasarkan tujuan dari penelitian tersebut, sedangkan untuk Eksklusi adalah suatu kriteria yang menghilangkan subjek yang tidak memenuhi kriteria dari Inklusi itu sendiri oleh peneliti. Berikut merupakan kriteria dari Inklusi dan Eksklusi adalah:

- Paper yang akan dijadikan referensi atau sitasi dalam rentang waktu publish 5 tahun terakhir atau dimulai dari tahun 2018.
- Paper memiliki hubungan atau kaitan dengan apa yang akan dilakukan oleh peneliti.
- Memiliki *SJR value* lebih dari 1.

Kemudian untuk kriteria dari Eksklusi adalah sebagai berikut:

- Tidak dapat diunduh.

BAB IV

PENUTUP

4.1. Kesimpulan dan Saran

Berdasarkan hasil kajian yang dibahas di atas, pada penelitian ini dapat ditarik kesimpulan bahwa pada penelitian ini memiliki kesimpulan adalah sebagai berikut:

1. Melakukan perbandingan antara Model Tradisional atau *Machine Learning* dengan *Deep Learning* atau *Convolutional Neural Networks (CNNs)*.
2. Melakukan evaluasi model menggunakan *confusion matrix*.

DAFTAR PUSTAKA

- [1] D. O. Eke, "Chatgpt and the rise of Generative AI: Threat to academic integrity?," *Journal of Responsible Technology*, vol. 13, p. 100060, 2023. doi:10.1016/j.jrt.2023.100060
- [2] Pojon, M. (2017). *Using machine learning to predict student performance* (Master's thesis).
- [3] R. Ganeshkumar and Dr. YSKumaraswamy, "Investigating cardiac arrhythmia in ECG using Random Forest Classification," *International Journal of Computer Applications*, vol. 37, no. 4, pp. 31–34, 2012. doi:10.5120/4599-6557
- [4] H. Almutairi, G. M. Hassan, and A. Datta, "Classification of obstructive sleep apnoea from single-lead ECG signals using convolutional neural and long short term memory networks," *Biomedical Signal Processing and Control*, vol. 69, p. 102906, 2021. doi:10.1016/j.bspc.2021.102906
- [5] P. Kanani and M. Padole, "ECG Heartbeat arrhythmia classification using time-series augmented signals and Deep Learning Approach," *Procedia Computer Science*, vol. 171, pp. 524–531, 2020. doi:10.1016/j.procs.2020.04.056
- [6] D. K. Atal and M. Singh, "Arrhythmia classification with ECG signals based on the optimization-enabled deep convolutional neural network," *Computer Methods and Programs in Biomedicine*, vol. 196, p. 105607, 2020. doi:10.1016/j.cmpb.2020.105607
- [7] S. Mandal, P. Mondal, and A. H. Roy, "Detection of ventricular arrhythmia by using heart rate variability signal and ECG Beat Image," *Biomedical Signal Processing and Control*, vol. 68, p. 102692, 2021. doi:10.1016/j.bspc.2021.102692
- [8] R. K. Tripathy, M. R. A. Paternina, J. G. Arrieta, A. Zamora-Méndez, and G. R. Naik, "Automated detection of congestive heart failure from Electrocardiogram Signal using Stockwell transform and hybrid classification scheme," *Computer Methods and Programs in Biomedicine*, vol. 173, pp. 53–65, 2019. doi:10.1016/j.cmpb.2019.03.008
- [9] J. Rahul, M. Sora, L. D. Sharma, and V. K. Bohat, "An improved cardiac arrhythmia classification using an RR interval-based approach," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 2, pp. 656–666, 2021. doi:10.1016/j.bbe.2021.04.004
- [10] S. Sowmya and D. Jose, "Contemplate on ECG signals and classification of arrhythmia signals using CNN-LSTM deep learning model," *Measurement: Sensors*, vol. 24, p. 100558, 2022. doi:10.1016/j.measen.2022.100558
- [11] Y.-Y. Jo *et al.*, "Detection and classification of arrhythmia using an explainable deep learning model," *Journal of Electrocardiology*, vol. 67, pp. 124–132, 2021. doi:10.1016/j.jelectrocard.2021.06.006
- [12] L. Zheng, Z. Wang, J. Liang, S. Luo, and S. Tian, "Effective compression and classification of ECG arrhythmia by Singular Value Decomposition," *Biomedical Engineering Advances*, vol. 2, p. 100013, 2021. doi:10.1016/j.bea.2021.100013
- [13] S. C. Mohonta, M. A. Motin, and D. K. Kumar, "Electrocardiogram based arrhythmia classification using wavelet transform with Deep Learning Model," *SSRN Electronic Journal*, 2022. doi:10.2139/ssrn.4088025
- [14] K. Prashant, P. Choudhary, T. Agrawal, and E. Kaushik, "Owae-net: COVID-19 detection from ECG images using deep learning and optimized weighted

- average ensemble technique,” *Intelligent Systems with Applications*, vol. 16, p. 200154, 2022. doi:10.1016/j.iswa.2022.200154
- [15] V. Jahmunah, E. Y. K. Ng, T. R. San, and U. R. Acharya, “Automated detection of coronary artery disease, myocardial infarction and congestive heart failure using GABORCNN model with ECG signals,” *Computers in Biology and Medicine*, vol. 134, p. 104457, 2021. doi:10.1016/j.compbimed.2021.104457
 - [16] A. S. Eltrass, M. B. Tayel, and A. I. Ammar, “A new automated CNN deep learning approach for identification of ECG congestive heart failure and arrhythmia using constant-q non-stationary Gabor transform,” *Biomedical Signal Processing and Control*, vol. 65, p. 102326, 2021. doi:10.1016/j.bspc.2020.102326
 - [17] A. A. Bhurane, M. Sharma, R. San-Tan, and U. R. Acharya, “An efficient detection of congestive heart failure using frequency localized filter banks for the diagnosis with ECG Signals,” *Cognitive Systems Research*, vol. 55, pp. 82–94, 2019. doi:10.1016/j.cogsys.2018.12.017
 - [18] M. Porumb, E. Iadanza, S. Massaro, and L. Pecchia, “A convolutional neural network approach to detect congestive heart failure,” *Biomedical Signal Processing and Control*, vol. 55, p. 101597, 2020. doi:10.1016/j.bspc.2019.101597
 - [19] A. Rath, D. Mishra, G. Panda, and S. C. Satapathy, “Heart disease detection using deep learning methods from imbalanced ECG samples,” *Biomedical Signal Processing and Control*, vol. 68, p. 102820, 2021. doi:10.1016/j.bspc.2021.102820
 - [20] A. Pal, R. Srivastva, and Y. N. Singh, “CardioNet: An efficient ECG arrhythmia classification system using transfer learning,” *Big Data Research*, vol. 26, p. 100271, 2021. doi:10.1016/j.bdr.2021.100271
 - [21] J. A. Walsh, E. J. Topol, and S. R. Steinhubl, “Novel wireless devices for cardiac monitoring,” *Circulation*, vol. 130, no. 7, pp. 573–581, 2014. doi:10.1161/circulationaha.114.009024
 - [22] S. Kuila, N. Dhanda, and S. Joardar, “Feature extraction and classification of MIT-BiH Arrhythmia Database,” *Lecture Notes in Electrical Engineering*, pp. 417–427, 2019. doi:10.1007/978-981-15-0829-5_41
 - [23] Z. F. Apandi, R. Ikeura, and S. Hayakawa, “Arrhythmia detection using MIT-BiH Dataset: A Review,” *2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, 2018. doi:10.1109/icassda.2018.8477620
 - [24] T. Li and M. Zhou, “ECG classification using wavelet packet entropy and random forests,” *Entropy*, vol. 18, no. 8, p. 285, 2016. doi:10.3390/e18080285
 - [25] R. Ganeshkumar and Dr. YSKumaraswamy, “Investigating cardiac arrhythmia in ECG using Random Forest Classification,” *International Journal of Computer Applications*, vol. 37, no. 4, pp. 31–34, 2012. doi:10.5120/4599-6557
 - [26] M. Wasimuddin, K. Elleithy, A.-S. Abuzneid, M. Faezipour, and O. Abuzaghle, “Stages-based ECG signal analysis from traditional signal processing to machine learning approaches: A survey,” *IEEE Access*, vol. 8, pp. 177782–177803, 2020. doi:10.1109/access.2020.3026968
 - [27] N. K. Kamila *et al.*, “Machine learning model design for high performance cloud computing & load balancing resiliency: An innovative approach,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 10, pp. 9991–10009, 2022. doi:10.1016/j.jksuci.2022.10.001
 - [28] C. Renggli, S. Ashkboos, M. Aghagolzadeh, D. Alistarh, and T. Hoefler,

- “SparCML,” *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2019. doi:10.1145/3295500.3356222
- [29] H. Bhatt, V. Shah, K. Shah, R. Shah, and M. Shah, “State-of-the-art machine learning techniques for melanoma skin cancer detection and classification: A comprehensive review,” *Intelligent Medicine*, 2022. doi:10.1016/j.imed.2022.08.004
- [30] M. Khanzadeh, S. Chowdhury, M. Marufuzzaman, M. A. Tschopp, and L. Bian, “Porosity prediction: Supervised-learning of thermal history for direct laser deposition,” *Journal of Manufacturing Systems*, vol. 47, pp. 69–82, 2018. doi:10.1016/j.jmsy.2018.04.001
- [31] N. Jalal, A. Mehmood, G. S. Choi, and I. Ashraf, “A novel improved random forest for text classification using feature ranking and optimal number of trees,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 2733–2742, 2022. doi:10.1016/j.jksuci.2022.03.012
- [32] D. Makariou, P. Barrieu, and Y. Chen, “A random forest based approach for predicting spreads in the primary catastrophe bond market,” *Insurance: Mathematics and Economics*, vol. 101, pp. 140–162, 2021. doi:10.1016/j.insmatheco.2021.07.003
- [33] M. Brendel *et al.*, “Application of deep learning on single-cell RNA sequencing data analysis: A Review,” *Genomics, Proteomics & Bioinformatics*, vol. 20, no. 5, pp. 814–835, 2022. doi:10.1016/j.gpb.2022.11.011
- [34] H. Naeem and A. A. Bin-Salem, “A CNN-LSTM network with multi-level feature extraction-based approach for automated detection of coronavirus from CT scan and X-ray images,” *Applied Soft Computing*, vol. 113, p. 107918, 2021. doi:10.1016/j.asoc.2021.107918