

KLASIFIKASI PENYAKIT *PNEUMONIA* MENGGUNAKAN METODE *CONVOLUTION NEURAL NETWORK* (CNN)

Thia Riskiyani H. Mosi¹, Yasin Aril Mustofa^{2*}, Sudirman Melangi³ Annahl Riadi⁴

^{1,3}Fakultas Ilmu Komputer, Teknik Informatika, Universitas Ichsan Gorontalo, Gorontalo, Indonesia

^{2,4}Fakultas Ilmu Komputer, Teknik Informatika, Universitas Ichsan Sidenreng Rappang, Sidrap, Indonesia

Email: thiariskiyanihm@gmail.com, arielccc@gmail.com, oedhie.lidya@gmail.com, annahlriadi@gmail.com

Abstrak - Penyakit *pneumonia* merupakan salah satu penyebab utama kematian di Indonesia dan diagnosis dini sangat penting untuk penanganan yang efektif. Penelitian ini bertujuan untuk mengembangkan model klasifikasi *pneumonia* menggunakan metode Convolution Neural Network (CNN) berbasis citra X-ray paru-paru. Metodologi penelitian ini melibatkan penggunaan dataset citra rontgen paru-paru yang diambil dari Kaggle, yang terdiri dari dua kelas yaitu *pneumonia* dan normal. Tahapan yang dilakukan meliputi pengumpulan data, pre-processing citra, pelatihan model CNN, serta evaluasi model menggunakan confusion matrix. Hasil penelitian menunjukkan bahwa model CNN yang dibangun mampu mengklasifikasikan citra X-ray paru-paru dengan akurasi mencapai 96%. Evaluasi lebih lanjut menunjukkan bahwa model ini memiliki tingkat presisi dan recall yang tinggi, yang menunjukkan kemampuannya dalam mengenali *pneumonia* dengan baik. Selain itu, confusion matrix menunjukkan bahwa model ini memiliki performa yang baik dalam mengidentifikasi kedua kelas dengan minim kesalahan. Penelitian ini memberikan kontribusi dalam bidang teknologi kesehatan, khususnya dalam pemanfaatan CNN untuk klasifikasi penyakit *pneumonia*. Temuan ini menunjukkan bahwa CNN adalah metode yang efektif untuk meningkatkan kecepatan dan akurasi diagnosis *pneumonia* melalui analisis citra X-ray. Ke depannya, penelitian lebih lanjut dengan dataset yang lebih besar dan beragam dapat lebih meningkatkan performa model ini.

Kata Kunci: *Pneumonia, Convolution Neural Network*

Abstract *Pneumonia is one of the leading causes of death in Indonesia, making early diagnosis crucial for effective treatment. This research aims to develop a pneumonia classification model based on lung X-ray images, utilizing the Convolutional Neural Network (CNN) method. The methodology involves using a dataset of lung X-ray images obtained from Kaggle, which includes two classes: pneumonia and normal. The research process consists of data collection, image pre-processing, CNN model training, and model evaluation using a confusion matrix. The results indicate that the developed CNN model can classify lung X-ray images with an accuracy of 96%. Further evaluation reveals a high level of precision and recall, demonstrating the model's strong capability to accurately identify pneumonia cases. Additionally, the confusion matrix shows that the model performs well in distinguishing between the two classes, with minimal errors. This research contributes to the field of health technology, particularly in the application of CNNs for pneumonia classification. The findings highlight that CNN is an effective method for enhancing the speed and accuracy of pneumonia diagnosis through X-ray image analysis. Future studies with larger and more diverse datasets could further improve the performance of this model.*

Keywords: *Pneumonia, Convolution Neural Network*

1. PENDAHULUAN

Penyakit paru-paru menjadi penyebab utama kematian di Indonesia. Paru-paru adalah organ vital dalam sistem pernapasan manusia yang sangat rentan terhadap berbagai penyakit, seperti asma, bronkitis, kanker, tuberculosis, infeksi saluran pernapasan akut (ISPA), dan *pneumonia* [1] *Pneumonia* merupakan salah satu penyakit yang dianggap serius. Biasanya, penyakit ini lebih rentan menyerang anak-anak di bawah usia 2 tahun serta lansia yang berusia 65 tahun ke atas [2]. *Pneumonia* merupakan penyakit yang disebabkan oleh bakteri, virus, dan jamur. Virus penyebab *pneumonia* adalah *Respiratory Syncial Virus* (RVS). Virus ini kebanyakan menyerang bagian atas saluran pernafasan manusia [3].

Diagnosis *pneumonia* umumnya dilakukan melalui pemeriksaan rontgen dada untuk mendeteksi adanya tanda infeksi. Namun analisis gambar rontgen memerlukan proses yang kompleks karena jenis infeksi pada paru-paru sering kali sulit untuk dibedakan dengan kondisi medis lain seperti edema paru atau penyakit pernapasan lainnya. Selain itu, kualitas gambar radiografi yang buruk, seperti resolusi rendah atau gangguan noise, sering menjadi tantangan dalam mengidentifikasi *pneumonia* dengan akurat. Dengan perkembangan teknologi kecerdasan buatan (AI) terutama dalam *deep learning*, muncul peluang baru untuk mengatasi tantangan tersebut. Teknologi sering digunakan oleh manusia sebagai alat bantu dalam berbagai aktivitas dan pekerjaan sehari-hari. Seiring berjalannya waktu teknologi terus

berkembang dan berdampak pada berbagai bidang, salah satunya adalah sektor medis yang semakin mengalami perkembangan digitalisasi[2].

Dengan adanya perkembangan teknologi, klasifikasi menggunakan kecerdasan buatan seperti CNN (*Convolution Neural Network*) menjadi solusi yang cukup baik. CNN adalah metode yang dapat mengidentifikasi pola-pola dalam citra secara otomatis dan lebih cepat dibandingkan manusia.)”. Dalam penelitian klasifikasi citra, khususnya dalam diagnosis penyakit berdasarkan rontgen dada (X-ray), metode *Convolution Neural Network* (CNN) dipilih karena kemampuannya yang sangat baik dalam mengenali pola dan fitur penting dari gambar. CNN (*Convolutional Neural Network*) adalah arsitektur *deep learning* yang dirancang khusus untuk memproses data gambar. Berbeda dengan metode konvensional yang memerlukan ekstraksi fitur manual, CNN secara otomatis dapat mengekstrak fitur kompleks dalam gambar seperti tepi, tekstur, bentuk, dan pola tertentu. Keunggulan utama CNN terletak pada kemampuannya untuk menangani data gambar yang besar dan kompleks dengan efisien. Melalui proses konvolusi dan pooling, CNN dapat mengurangi dimensi gambar tanpa menghilangkan informasi penting, sehingga mempercepat proses penghitungan.

2. TINJAUAN PUSTAKA

2.1 *Pneumonia*

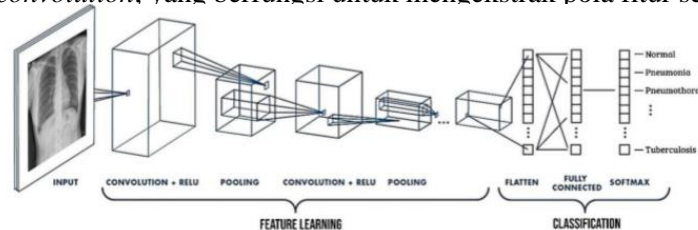
Pneumonia adalah infeksi pada saluran pernapasan yang disebabkan oleh bakteri, virus, atau jamur. Infeksi ini dapat menyerang salah satu atau kedua paru-paru yang mengakibatkan kantung udara di paru-paru kesulitan menyerap oksigen[4]. Gejala yang muncul pada penderita *pneumonia* sangat bervariasi tergantung pada penyebab infeksi, tingkat keparahannya, serta usia dan kondisi kesehatan umum pasien. Gejala ini dapat berkembang dengan cepat atau secara perlahan dalam kurun waktu 24-48 jam[5].

2.2. Klasifikasi

Klasifikasi merupakan sebuah teknik analisis data yang mengekstrak model untuk mendeskripsikannya ke dalam kelas tertentu[6]. Klasifikasi adalah proses pembelajaran untuk menghasilkan sebuah fungsi yang dapat memetakan setiap himpunan objek ke dalam salah satu label kelas yang telah ditentukan sebelumnya. Klasifikasi termasuk dalam jenis analisis yang digunakan untuk menentukan kelas label dari sejumlah sampel yang akan dikategorikan. Klasifikasi adalah bentuk dasar dari analisis data dan sering dianggap sebagai teknik untuk mengidentifikasi kategori kelas dari data yang ada. Dalam klasifikasi, terdapat berbagai teknik atau algoritma yang digunakan untuk menyelesaikan masalah atau membangun model data, seperti algoritma C4.5 (*decision tree*), *naïve bayes*, *neural network*, *k-nearest neighbor*, dan *support vector machine*[7].

2.3 *Convolution Neural Network* (CNN)

CNN (*Convolutional Neural Network*) adalah salah satu arsitektur jaringan saraf yang dirancang khusus untuk klasifikasi citra. CNN dapat dianggap sebagai inovasi dalam pengklasifikasian citra karena merupakan bagian dari jaringan saraf yang terinspirasi oleh cara kerja otak manusia. Salah satu keunggulan CNN adalah kemampuannya untuk mengeliminasi kebutuhan akan ekstraksi fitur pola secara manual yang biasanya memerlukan bantuan pakar. Untuk mencapai hal ini, CNN menggunakan lapisan yang disebut *convolution*, yang berfungsi untuk mengekstrak pola fitur secara otomatis[8].



Gambar 1. Ilustrasi Struktur CNN

2.4 *Confusion Matriks*

Confusion matrix merupakan salah satu alat evaluasi yang umum digunakan dalam mengukur kinerja suatu model klasifikasi, khususnya dalam konteks pembelajaran mesin dan pengolahan citra medis. Matriks ini menyajikan perbandingan antara hasil prediksi model dengan kondisi aktual data, sehingga

memberikan gambaran rinci mengenai akurasi klasifikasi terhadap masing-masing kategori. Terdapat empat komponen utama dalam confusion matrix yang merepresentasikan berbagai kemungkinan hasil klasifikasi. Pertama, *True Positive* (TP) adalah jumlah kasus di mana data yang sebenarnya positif berhasil diprediksi dengan benar sebagai positif, misalnya gambar *rontgen* yang menunjukkan pneumonia diidentifikasi dengan tepat sebagai pneumonia. Kedua, *True Negative* (TN) mencerminkan jumlah data negatif yang berhasil dikenali dengan benar sebagai negatif, yaitu gambar rontgen dada yang sehat diidentifikasi sebagai normal. Ketiga, *False Positive* (FP) terjadi ketika data negatif secara keliru diprediksi sebagai positif, contohnya gambar normal yang salah diklasifikasikan sebagai pneumonia. Keempat, *False Negative* (FN) merujuk pada kondisi ketika data yang sebenarnya positif diprediksi sebagai negatif, seperti gambar pneumonia yang salah diidentifikasi sebagai normal. Pemahaman terhadap keempat komponen ini sangat penting dalam menganalisis performa model, karena tidak hanya mencerminkan tingkat akurasi secara keseluruhan, tetapi juga mengungkap potensi kesalahan yang dapat berdampak signifikan, terutama dalam konteks diagnosis medis.

3. METODE PENELITIAN

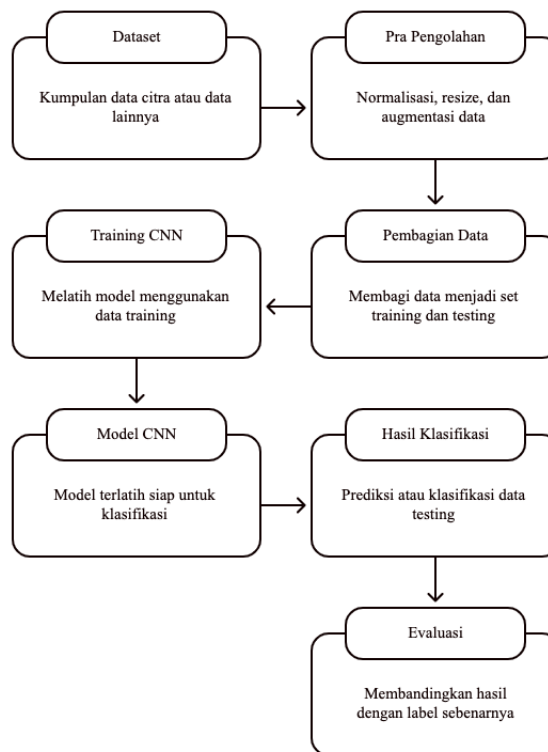
3.1 Jenis, Metode, Subjek, Objek dan Waktu Penelitian

Dilihat dari tingkat penerapannya, penelitian ini merupakan penelitian terapan. Karena sifat data atau jenis informasi yang diolah, maka penelitian ini merupakan penelitian kuantitatif. Penelitian ini menggunakan metode penelitian eksperimen. Dengan demikian, jenis penelitian ini adalah penelitian eksperimental. Subjek penelitian ini adalah klasifikasi pada objek gambar citra x-ray paru-paru manusia untuk penyakit pneumonia dan non-pneumonia. Penelitian ini dimulai dari bulan September 2024.

3.2 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini merupakan dataset publik yang bersumber dari situs Kaggle, dengan nama Gambar Rontgen Dada (Pneumonia). Dataset ini digunakan untuk keperluan klasifikasi citra guna mendeteksi keberadaan pneumonia dalam hasil rontgen dada. Secara keseluruhan, jumlah data dalam dataset ini mencapai 5.856 gambar yang dibagi menjadi tiga bagian, yaitu data training, data testing, dan data validasi. Pada data training, terdapat 1.341 gambar kategori Normal dan 3.875 gambar kategori Pneumonia. Sementara itu, untuk data testing terdiri dari 234 gambar Normal dan 390 gambar Pneumonia. Selain itu, dataset ini juga mencakup data validasi dengan jumlah yang lebih kecil, yakni 8 gambar Normal dan 8 gambar Pneumonia. Semua gambar dalam dataset ini memiliki ukuran 150 x 150 piksel dan disimpan dalam format JPG, yang sesuai untuk pemrosesan dengan model CNN. Pembagian data ini bertujuan untuk memastikan bahwa model dapat dilatih, diuji, dan divalidasi secara optimal guna menghasilkan performa klasifikasi yang andal.

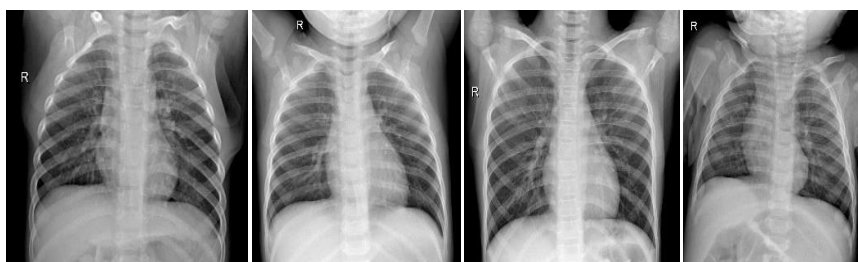
3.3 Pemodelan



Gambar 2. Pemodelan

4. HASIL DAN PEMBAHASAN

Data set yang digunakan dalam penelitian ini adalah dataset *public* dengan nama Gambar Rontgen Dada (*pneumonia*) yang berasal dari <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>. Dataset ini merupakan dataset yang pada umumnya digunakan dalam eksperimen untuk klasifikasi penyakit *pneumonia*. Ukuran dataset yang digunakan adalah 150 x 150 pixel dengan format gambar jpg. Dengan menggunakan data training 5216 dan data testing 624.



Gambar 3. Citra X-Ray Paru-paru Normal



Gambar 1. Citra X-Ray Paru-paru *Pneumonia*

4.1 Pra-Pengolahan Data

Pra-pengolahan citra (*image pre-processing*) merupakan tahapan awal yang sangat krusial dalam sistem pengolahan citra, khususnya sebelum proses klasifikasi utama dilakukan. Tujuan dari tahapan ini adalah untuk mempersiapkan data citra agar lebih sesuai dengan kebutuhan pemrosesan dan meningkatkan

efektivitas serta efisiensi kinerja model klasifikasi. Dalam konteks penelitian ini, citra hasil *X-ray* yang digunakan dikonversikan dan disesuaikan melalui beberapa langkah sistematis. Tahapan pertama adalah *resize*, yakni proses penyesuaian ukuran citra dari bentuk aslinya yang bervariasi menjadi ukuran tetap yaitu 150 x 150 piksel. Ukuran ini dipilih karena lebih kecil dan efisien, sehingga dapat mempercepat proses pelatihan model tanpa mengorbankan informasi penting dalam citra. Selanjutnya dilakukan *encoding* data, di mana label kategori yang semula berbentuk teks diubah ke dalam bentuk numerik agar dapat dikenali oleh sistem pembelajaran mesin. Dalam hal ini, label *Pneumonia* dikodekan sebagai 0 dan *Normal* sebagai 1. Tahapan terakhir adalah *augmentasi* data, yaitu teknik yang digunakan untuk memperbesar jumlah data latih secara artifisial melalui manipulasi citra tanpa mengubah karakteristik inti dari citra tersebut. Teknik *augmentasi* yang diterapkan dalam penelitian ini mencakup pengaturan *rotation range* sebesar 30 derajat, *horizontal flip*, *zoom range* sebesar 0,2, serta *width* dan *height shift range* masing-masing sebesar 0,1. Dengan melakukan pra-pengolahan yang sistematis ini, data citra menjadi lebih seragam, representatif, dan optimal untuk digunakan dalam proses pelatihan dan evaluasi model klasifikasi.

Tabel 1. Hasil Data Testing

No.	Image	Klasifikasi	Hipotesis	No.	Image	Klasifikasi	Hipotesis
1.		Pneumonia	Pneumonia	16.		Normal	Pneumonia
2.		Pneumonia	Pneumonia	17.		Normal	Normal
3.		Pneumonia	Pneumonia	18.		Normal	Pneumonia
4.		Pneumonia	Pneumonia	19.		Normal	Normal
5.		Pneumonia	Pneumonia	20.		Normal	Normal

Tabel 2. Confusion Matriks

	Pneumonia	Normal
Pneumonia	15	3
Normal	0	12

Data yang diklasifikasikan dengan benar = 27

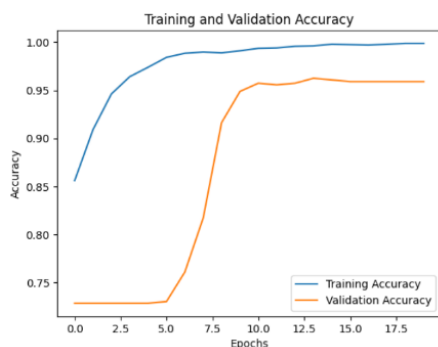
Data yang diklasifikasikan dengan salah = 3

Klasifikasi	TP	TN	FP	FN	Hasil %		
					Akurasi	Precision	Recall
Pneumonia	15	0	0	0	100	100	100
Normal	12	0	3	0	80	80	100
Rata-rata					90	90	100

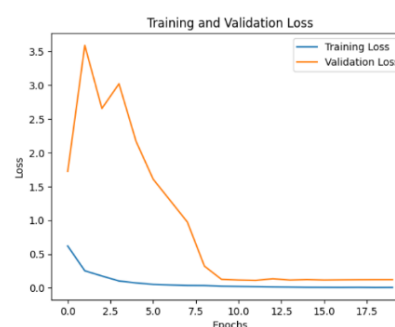
- a. Arsitektur CNN pada penelitian ini dibangun dengan 6 blok, citra berukuran 150x150 piksel yang berasal dari layer input dikonvolusi pada lapisan konvolusi dengan ukuran 8x8 dan stride 1. Konvolusi tersebut menghasilkan feature map yang kemudian dinormalisasi oleh Batch Normalization. Hasil dari normalisasi selanjutnya masuk ke layer aktivasi, di dalamnya terdapat fungsi ReLu yang digunakan untuk membuat hasil dari ekstraksi fitur non-linier. Selanjutnya nilai output dari fungsi ReLu direduksi pada lapisan MaxPooling sebelum diteruskan pada tahap konvolusi kedua.
- b. Tahap konvolusi kedua sampai dengan kelima, proses ekstraksi fitur dilakukan oleh kombinasi convolutional blok dan identity blok. Setelah proses ekstraksi fitur selesai, kemudian feature map diproses ke dalam fully connected layer untuk melakukan proses klasifikasi dibantu dengan fungsi aktivasi sigmoid.

- c. Dalam proses training data menggunakan algoritma CNN dengan nilai learning rate sebesar 0.000001, jumlah step epochs sebesar 74 dan epochsnya sebesar 50.

Berdasarkan hasil proses learning data yang dilakukan, diperoleh nilai akurasi untuk data training yang berkisar diangka 0.1 yang menunjukkan bahwa model sangat baik dalam mempelajari data pelatihan dan nilai akurasi untuk data validasi berkisar diantara 0.95. Sementara nilai loss untuk data training dan validasi mengalami penurunan yang sangat lambat dengan kedua kurva yang cenderung mendekati nilai konstan mendekati 0. Sebagaimana yang ditunjukkan pada grafik dibawah ini.



Gambar 5. Grafik Training dan Validasi Akurasi



Gambar 6. Gambar training dan Validasi Loss

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan dan pembahasan yang telah diuraikan sebelumnya, maka dapat disimpulkan bahwa,

- Penelitian ini berhasil menerapkan metode Convolution Neural Network (CNN) untuk mengklasifikasikan penyakit pneumonia berdasarkan citra X-ray paru-paru. Model CNN yang dibangun dapat secara efektif memproses citra dan mengidentifikasi pneumonia dengan hasil yang memadai.
- Hasil pengujian model *Convolution Neural Network* (CNN) menunjukkan akurasi yang tinggi dalam mengklasifikasikan citra rontgen paru-paru, dengan tingkat akurasi mencapai 96%. Evaluasi model menggunakan confusion matrix juga menunjukkan performa yang baik, dengan recall dan precision yang tinggi pada klasifikasi pneumonia dan normal.

Setelah melakukan penelitian, terdapat beberapa saran yang perlu diperhatikan yaitu :

- Untuk meningkatkan akurasi model, disarankan untuk menggunakan dataset yang lebih besar dan lebih bervariasi, termasuk gambar dengan kualitas lebih tinggi dan lebih banyak kelas untuk diagnosis penyakit paru-paru lainnya. Penelitian lebih lanjut bisa memanfaatkan dataset dari berbagai sumber untuk meningkatkan generalisasi model.
- Untuk meningkatkan kualitas dan akurasi model, tahapan pre-processing dapat diperluas dengan teknik yang lebih kompleks, seperti denoising untuk mengatasi masalah noise pada citra rontgen atau peningkatan kualitas citra melalui algoritma super-resolution. Penelitian lebih lanjut dapat mengeksplorasi penggunaan teknik-teknik baru dalam pre-processing citra untuk memaksimalkan hasil pengklasifikasian.

DAFTAR PUSTAKA

- [1] N. Pratiwi *et al.*, "PENERAPAN HEALING ENVIRONMENT PADA RUMAH SAKIT KHUSUS PARU-PARU DI PROVINSI GORONTALO," *JAMBURA Journal of Architecture*, vol. 5, no. 2, p. 2023.
- [2] M. A. Alghozali, J. R. Triosaputra, and A. Kayan, "Klasifikasi Penyakit Pneumonia Citra Digital X-Ray Menggunakan Metode Convolutional Neural Network dan RGB Equalization," vol. 3, pp. 229–236, 2024.

- [3] A. Ainun, D. Halim, and S. Anraeni, "Analisis Klasifikasi Dataset Citra Penyakit Pneumonia Menggunakan Metode K-Nearest Neighbor (KNN)," *Indonesian Journal of Data and Science (IJODAS)*, vol. 2, no. 1, pp. 1–12, 2021.
- [4] R. M. Diar, R. Y. N. Fu'adah, and K. Usman, "Klasifikasi Penyakit Paru-Paru Berbasis Pengolahan Citra X Ray Menggunakan Convolutional Neural Network (Classification Of The Lung Diseases Based On X Ray Image Processing Using Convolutional Neural Network)," vol. 9, pp. 476–484, Apr. 2022.
- [5] E. Zalukhu, "Analisa Perbandingan Metode Certainly Faktor-Naive Bayes Terhadap Diagnose Penyakit Pneumonia," *ADA Journal of Information System Research*, vol. 1, no. 3, pp. 103–112, Jun. 2024.
- [6] L. U. Khasanah, Y. N. Nasution, F. Deny, and T. Amijaya, "Klasifikasi Penyakit Diabetes Melitus Menggunakan Algoritma Naïve Bayes Classifier," vol. 1, no. 1, pp. 41–50, 2022, [Online]. Available: <http://jurnal.fmipa.unmul.ac.id/index.php/basis>
- [7] A. Srirahayu and L. Setya Pribadie, "JURNAL ILMIAH INFORMATIKA GLOBAL Review Paper Data Mining Klasifikasi Data Mining," *JURNAL ILMIAH INFORMATIKA GLOBAL*, vol. 14, pp. 7–12, Apr. 2023, doi: 10.36982/jiig.v13i2.2307.
- [8] I. Md. D. Maysanjaya, "Klasifikasi Pneumonia pada Citra X-rays Paru-paru dengan Convolutional Neural Network," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 9, no. 2, pp. 190–195, May 2020, doi: 10.22146/jnteti.v9i2.66.