



## KLASIFIKASI PENYAKIT DAUN PISANG MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN)

M. Duta Pratama<sup>1</sup>, Rendra Gustriansyah<sup>2</sup>, Evi Purnamasari<sup>3</sup>

<sup>1,2,3</sup>Fakultas Ilmu Komputer dan Sains, Universitas Indo Global Mandiri  
Palembang, Sumatera Selatan, Indonesia 30129  
2020110033@students.uigm.ac.id, rendra@uigm.ac.id, evi.ps@uigm.ac.id

### Abstract

Bananas are a fruit with promising economic value in Indonesia. They are an essential commodity for farmers, but diseases affecting banana plants can harm their livelihoods. Banana diseases initially attack the leaves, and in the early stages, they are difficult to differentiate with the naked eye due to farmers' limited knowledge of pathogens. This research utilized the Convolutional Neural Network (CNN) method with transfer learning assistance using Google Colab to facilitate the classification of banana leaf diseases. The trained model experienced overfitting, so regularization was applied using dropout. The best model achieved an accuracy of 92%, precision of 92%, sensitivity of 91%, and an F1-score of 91% at a 70:20:10 ratio on epoch 80, as evaluated and validated using a confusion matrix. This study produced a reliable model for classifying banana leaf disease.

**Keywords:** Banana leaf disease, classification, CNN, dropout, transfer learning

### Abstrak

Pisang merupakan buah yang memiliki nilai ekonomi yang menjanjikan di Indonesia. Pisang merupakan komoditas yang penting bagi para petani dan penyakit yang menyerang pisang bisa merugikan petani. Penyakit pada pisang awalnya akan menyerang daun pisang, tahap awal pada penyakit susah untuk dibedakan menggunakan mata telanjang karena kurangnya pengetahuan petani akan *pathogen*. Penelitian ini menggunakan metode *Convolutional Neural Network* (CNN) dengan bantuan *transfer learning* menggunakan *Google Colab* dengan tujuan untuk memudahkan klasifikasi penyakit daun pisang. Model yang sudah dilatih mengalami *overfitting* sehingga dilakukan regularisasi menggunakan *dropout*. Hasil model terbaik didapatkan dengan rasio 70:20:10 pada *epoch* 80 dengan akurasi 92%, presisi 92%, sensitivitas 91%, dan *f1-score* 91% yang diuji dan divalidasi menggunakan *confusion matrix*. Penelitian ini menghasilkan model yang mampu mengklasifikasi penyakit daun pisang dengan baik.

**Kata kunci:** CNN, *dropout*, klasifikasi, penyakit daun pisang, *transfer learning*

### 1. PENDAHULUAN

Pisang adalah salah satu buah yang dibudidayakan secara luas di berbagai negara tropis dan subtropis [1]. Informasi yang dimuat pada Badan Pusat Statistik (BPS) di Indonesia tahun 2022 menjelaskan total produksi pisang mencapai 9,2 ton dan merupakan buah dengan produksi paling banyak di Indonesia. Pisang merupakan komoditas penting di Indonesia, namun penyakit pada pisang sering menginfeksi yang menyebabkan kerugian bagi petani. Ada beberapa jenis penyakit pada buah pisang yang sering dijumpai di Indonesia seperti bercak daun *cordana*, gugur daun *pestaliopsis*, *sigatoka*, dan layu jamur dan bakteri.

Petani di Indonesia menghadapi masalah dalam mengidentifikasi jenis penyakit pada daun pisang karena

kurangnya pengetahuan tentang jenis *pathogen* yang menyerang pisang sehingga menyebabkan kesalahan dalam penanganan penyakit. Petani di Indonesia menghadapi masalah dalam mengidentifikasi jenis penyakit pada daun pisang karena kurangnya pengetahuan tentang jenis *pathogen* yang menyerang pisang sehingga menyebabkan kesalahan dalam penanganan penyakit yang membuat pertumbuhan pisang melambat atau gagal sehingga menyebabkan kerugian pada petani akibat banyak pisang yang gagal panen [2] sehingga diperlukan sebuah sistem untuk mengklasifikasi penyakit pisang. *Convolutional Neural Network* (CNN) dipilih untuk penelitian klasifikasi penyakit pada daun pisang yang dapat membantu petani. Penelitian terkait klasifikasi daun pisang diterapkan

arsitektur VGG19 dalam CNN pada penelitian ini untuk melihat apakah hasil klasifikasi akan lebih optimal.

Penelitian terdahulu terkait klasifikasi pernah dilakukan seperti klasifikasi kematangan buah pisang menggunakan CNN sederhana [3] dengan akurasi 92%, penelitian lain untuk klasifikasi penyakit daun pisang menggunakan CNN *mobile* dengan akurasi 95,13% [4]. Penelitian tentang klasifikasi juga dilakukan dengan objek berbeda seperti tomat, makanan tradisional dan masker [5], [6], [7]. Penelitian untuk klasifikasi menggunakan metode lain juga pernah dilakukan seperti metode *Support Vector Machine* [8], *Naïve Bayes* [9], *Decision tree* [10], *K-means* [11], dan KNN [12]. Penelitian dilakukan menggunakan CNN karena metode CNN memiliki kedalaman jaringan yang tinggi dan banyak diaplikasikan pada data citra [13], [14]. Penelitian juga dilakukan menggunakan *transfer learning* dan arsitektur *Visual Geometry Group* (VGG) lebih tepatnya VGG19 karena memiliki akurasi tinggi untuk masalah identifikasi citra [15]. Penelitian ini menggunakan CNN dengan VGG19 sebagai arsitektur karena CNN cocok untuk mengolah data citra dan VGG19 memiliki akurasi yang tinggi untuk identifikasi citra [15]. Penelitian ini juga menambahkan *transfer learning* untuk menaikkan lagi akurasi pada model CNN.

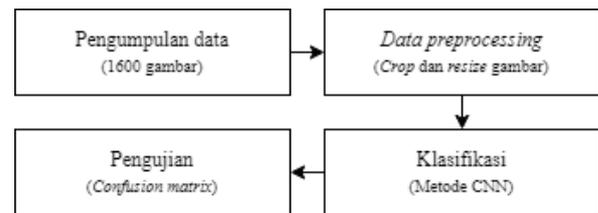
Klasifikasi merupakan kegiatan untuk menyusun sesuatu berdasarkan ciri dan sifat-sifatnya [16]. Klasifikasi menggunakan komputer dilakukan menggunakan *deep learning* dengan mencari kesamaan ciri pada suatu gambar, *deep learning* dapat mengekstraksi fitur dan pola yang akan dikelompokkan. Klasifikasi pada penyakit daun pisang dapat diambil dari ciri bentuk, warna dan tekstur pada daun pisang.

Tanaman buah pisang memiliki potensi serangan hama yang tinggi pada daerah tropis. Petani di Indonesia saat ini masih belum *familiar* dengan gejala-gejala penyakit yang menyerang pisang. Penyakit yang menyerang pisang ini menyebabkan kerugian karena produksi pisang yang menurun bahkan berhenti [2]. Penyakit berbahaya yang sering menyerang adalah *cordana*, *pestalotiopsis*, dan *sigatoka*.

CNN merupakan salah satu algoritma pada *deep learning* yang merupakan versi lebih baik dari jaringan saraf tiruan. CNN dikembangkan dengan tujuan untuk mengolah data dua dimensi seperti data citra [13]. CNN populer digunakan karena menggunakan kedalaman jaringan yang tinggi untuk mengelola data citra.

## 2. METODE PENELITIAN

Tahap-tahap penelitian dijelaskan pada Gambar 1 sebagai panduan penelitian yang akan dilakukan.



Gambar 1. Alur penelitian

Penelitian dimulai dengan melakukan studi literatur terlebih dahulu untuk memahami lebih pasti dasar-dasar teori yang berkaitan pada penelitian. Masalah ditemukan dengan membaca berbagai penelitian tentang penyakit pisang di Indonesia. Tahap selanjutnya dilakukan pengumpulan data berjumlah 1600 gambar yang didapat pada *mendeley data*. Data yang telah dikumpulkan akan melalui tahap pra-proses menggunakan augmentasi data dan akan dibagi menjadi data latih dan validasi. Tahap selanjutnya adalah melakukan implementasi model CNN dengan mengatur arsitektur VGG19 melalui *Google Colab*. Tahap terakhir setelah model selesai dilatih adalah melakukan pengujian pada model dan menarik kesimpulan apakah model yang dihasilkan sudah sesuai dengan kriteria yang diinginkan.

### 2.1 Pengumpulan data

Sumber data pada penelitian ini diambil pada *platform mendeley data*. Data diambil dari *Bangadbandu Sheikh Mujibur Rahman Agricultural University* (BSMRAU) [4]. Data terdiri dari 400 gambar masing-masing 4 kelas (*cordana*, *healthy*, *pestalotiopsis*, dan *sigatoka*) dalam format jpeg.

### 2.2 Data preprocessing

*Data preprocessing* merupakan langkah yang dilakukan untuk menyiapkan data mentah agar data yang dihasilkan dapat lebih mudah untuk dilakukan *training* dan hasil model lebih baik. *Preprocessing* pada penelitian akan dilakukan *resize* dan *crop* pada gambar untuk menonjolkan fitur gambar yang ingin diklasifikasi. Pembagian data juga dilakukan pada tahap ini, data yang dihasilkan akan dibagi menjadi data latih (*train*), data validasi (*validation*) dan data *test*. Data dibagi menjadi dua rasio, rasio pertama menggunakan 70% data latih, 20% data validasi dan 10% data *test* sedangkan rasio kedua menggunakan 80% data latih, 10% data validasi dan 10% data *test*. Contoh gambar hasil data *preprocessing* bisa dilihat pada Gambar 2.

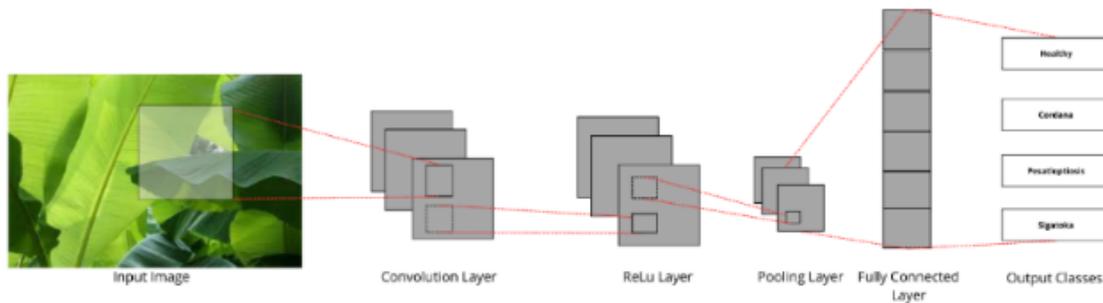


Gambar 2. Hasil *preprocessing*

### 2.3 Klasifikasi menggunakan CNN

CNN adalah salah satu algoritma dari *deep learning* dan merupakan versi lebih baik dari jaringan syaraf tiruan (JST) yang dirancang khusus untuk data gambar. Struktur dari CNN terinspirasi dari syaraf pada tubuh manusia lebih tepatnya syaraf otak pada manusia[17]. CNN secara penuh memanfaatkan *input* data dua dimensi seperti sebuah sinyal.

Proses ini memaksimalkan jumlah angka dalam parameter menjadi bagian-bagian yang lebih kecil yang berakibat pada simplifikasi proses *data training* dan mempercepat prosesnya[13], [18]. CNN biasanya menggunakan model serupa dengan *Multi Layer Perception* (MLP), yang tersusun dari banyak lapisan konvolusi dan *pooling* diakhiri dengan *fully connected layer*. Contoh arsitektur CNN untuk mengklasifikasi gambar bisa dilihat pada Gambar 3.



Gambar 3. Contoh CNN [13]

### 2.4 Pengujian

*Confusion matrix* merupakan metode yang bisa dipakai untuk mengetahui tingkat keakuratan suatu model. Pada *confusion matrix*, terdapat empat hasil dari perhitungan menggunakan metode ini yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [19]. Nilai akurasi menunjukkan kemampuan sistem untuk dapat mengklasifikasikan data dengan benar. Nilai akurasi didapat melalui perbandingan antara data yang terklarifikasi benar dan seluruh data. Nilai akurasi dapat dilihat pada persamaan 1. Nilai presisi merupakan total data positif yang diklasifikasi dengan benar lalu dibagi jumlah data positif dan dapat dilihat pada persamaan 2. Sensivitas menunjukkan ketepatan hasil positif yang diklasifikasikan oleh sistem dan dapat dilihat pada persamaan 3. *F1-score* merupakan data negatif yang diprediksi dengan benar dibagi seluruh data negatif dan dapat dilihat pada persamaan 4.

Akurasi

$$= \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (1)$$

Presisi

$$= \frac{TP}{TP + FP} \times 100\% \quad (2)$$

Sensivitas

$$= \frac{TP}{TP + FN} \times 100\% \quad (3)$$

F1 – Score

$$= 2 \times \frac{\text{Presisi} \times \text{Sensitivitas}}{\text{Presisi} + \text{Sensitivitas}} \quad (4)$$

Keterangan :

- TP adalah *True Positive* yaitu nilai data positif yang terklarifikasi benar.
- TN adalah *True Negative* yaitu nilai data negatif yang terklarifikasi benar.
- FP adalah *False Positive* yaitu nilai data positif yang terklarifikasi salah.
- FN adalah *False Negative* yaitu nilai data negatif yang terklarifikasi salah.

### 2.5 Skenario pengujian

Pada bagian ini akan diuraikan skenario pengujian yang akan dilakukan. Skenario penelitian dapat dilihat pada Tabel 1.

Tabel 1. Skenario pengujian

Skenario	Keterangan
Pembagian data	70% latih : 20% validasi : 10% uji, 80% latih : 10% validasi : 10% uji
Kelas	<i>Cordana, healthy, pestalotiopsis, sigatoka</i>
Ukuran data	224 x 224 pixel
<i>Batch size</i>	32
<i>Epochs</i>	20, 50, 80, 100

### 3. HASIL DAN PEMBAHASAN

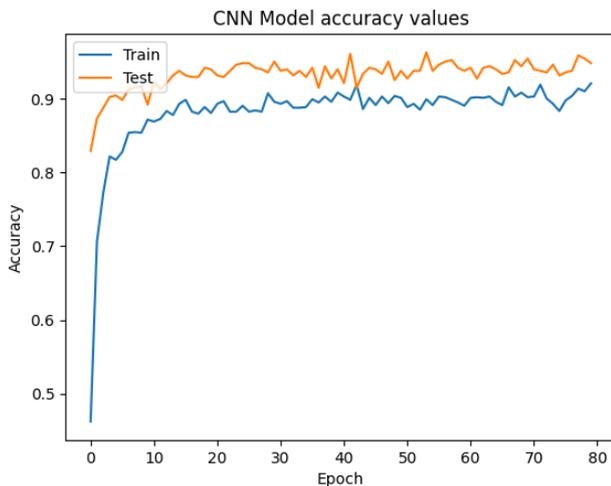
#### 3.1 Hasil Training model

Proses selanjutnya adalah melakukan *training* pada model. Sebelum melakukan *training* pada model, ditentukan terlebih dahulu *epochs* yang ingin digunakan. Penelitian terdahulu yang menjadi acuan pada penelitian ini, tidak ada ketentuan yang pasti untuk menentukan berapa *epochs* yang diperlukan karena penentuan *epochs* bisa bervariasi tergantung jumlah data dan arsitektur yang digunakan. Tabel 2 menyajikan percobaan *training* model dengan beberapa rasio pembagian data dan *epochs*.

Tabel 2 Percobaan model

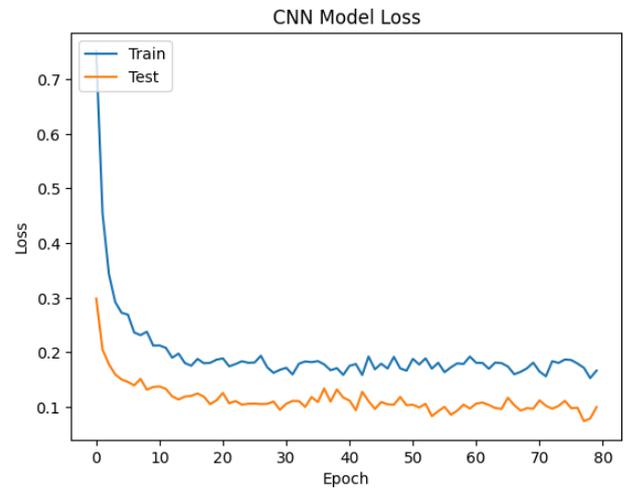
Percobaan ke	Rasio	Epochs	Latih		Validasi	
			Akurasi	Loss	Akurasi	Loss
1	70:30	20	0,86	0,19	0,91	0,11
2	70:30	50	0,90	0,15	0,93	0,11
3	70:30	80	0,92	0,16	0,94	0,10
4	70:30	100	0,90	0,18	0,93	0,11
5	80:20	20	0,87	0,18	0,94	0,09
6	80:20	50	0,90	0,17	0,93	0,10
7	80:20	80	0,89	0,18	0,91	0,12
8	80:20	100	0,90	0,18	0,93	0,07

Berbagai percobaan *training* pada model dipilih model pada percobaan ke-3 yang menghasilkan akurasi *training* dengan nilai 0,92 dan akurasi validasi bernilai 0,94. Hasil akurasi *training* dan validasi dapat dilihat pada Gambar 4.



Gambar 4. Akurasi training dan validasi

Ditampilkan juga *loss* pada *training* bernilai 0,16 dan *loss* validasi dengan nilai 0,10. Hasil *loss* pada *training* dan validasi dapat dilihat pada Gambar 5.

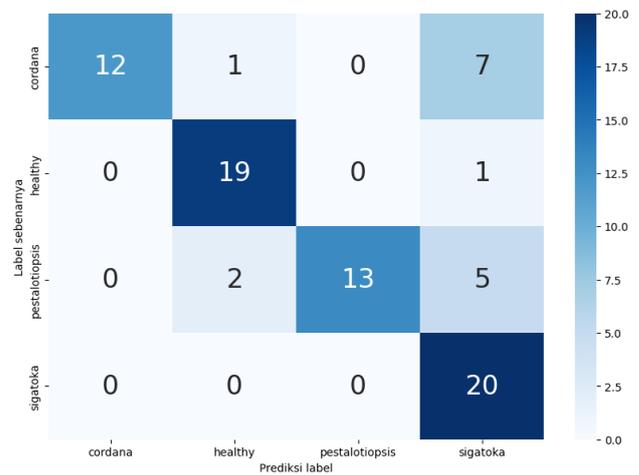


Gambar 5. Loss pada training dan validasi

Setelah mendapatkan hasil pada model. Tahap selanjutnya dalam penelitian ini adalah menguji model yang telah dibuat.

#### 3.2 Pengujian model

Pengujian adalah tahapan yang selalu dilakukan ketika selesai membuat sistem, pengujian dilakukan untuk mengukur keberhasilan yang dicapai suatu penelitian. Pengujian pada penelitian ini dilakukan dengan *confusion matrix* pada *Google Colab* untuk mengetahui akurasi, presisi, *recall*, dan *f1-score* pada model. Model diuji menggunakan dataset *test* dengan jumlah 80 gambar yang terdiri dari 20 gambar *cordana*, 20 gambar sehat, 20 gambar *pestalotiopsis*, dan 20 gambar *sigatoka*. Hasil *confusion matrix* dapat dilihat pada Gambar 6.



Gambar 6. Confusion matrix

Menggunakan persamaan (1-4) diketahui akurasi *test* sebesar 82%, presisi sebesar 85%, *recall* sebesar 83% dan *f1-score* sebesar 82%. Secara keseluruhan model yang dihasilkan baik dan hasil *test* identik dengan hasil pelatihan dan validasi. Model yang telah dihasilkan bisa diterapkan untuk melakukan klasifikasi penyakit dengan cara mengambil gambar daun pisang dari kebun petani dan

melakukan *test* klasifikasi menggunakan model yang telah dibuat.

#### 4. KESIMPULAN

Model yang dihasilkan telah dianggap layak digunakan karena hasil *test* identik dengan hasil *training*. Evaluasi kinerja model dilakukan dengan menilai beberapa metrik untuk mengetahui nilai presisi, *recall*, *f1-score*, dan akurasi. Presisi mengukur seberapa akurat prediksi nilai positif, *recall* mengukur kemampuan model untuk menemukan semua nilai positif yang sebenarnya, sedangkan *f1-score* adalah gambaran keseluruhan keseimbangan antara presisi dan *recall*. Model klasifikasi yang dihasilkan, penelitian ini telah sukses mengembangkan sistem klasifikasi penyakit daun pisang menggunakan metode CNN. Model dengan kinerja terbaik bisa mencapai Tingkat akurasi sebesar 82%, presisi sebesar 85%, *recall* sebesar 83% dan *f1-score* sebesar 82%. Penelitian yang telah dilakukan berhasil membuat suatu model yang bisa membantu petani mengklasifikasi penyakit pada daun pisang dan model penelitian dapat diterapkan menggunakan *drone* ataupun kamera untuk mendapat gambar daun pisang yang akan diklasifikasi.

Penelitian menemui limitasi yaitu kurangnya data yang ada untuk pelatihan model. Untuk penelitian ini bisa semakin berkembang, beberapa saran perlu ditambahkan. Pertama, *dataset* pada model bisa ditambahkan untuk memperluas cakupan variasi situasi pada daun pisang. Kelas pada model juga bisa ditambahkan dengan memasukkan penyakit lain pada daun pisang. Pengujian model juga lebih baik dilakukan pada situasi nyata pada kebun pisang dan juga lebih baik lagi jika mencoba menggunakan arsitektur lain pada model untuk mencari tahu apakah ada arsitektur dengan kinerja lebih baik untuk melakukan klasifikasi penyakit pada daun pisang.

#### DAFTAR PUSTAKA

- [1] I. Ryan and S. Pigai, "Morfologi Tanaman Pisang Jiigikago Berdasarkan Kearifan Lokal Suku Mee Di Kampung Idayo Distrik Obano Kabupaten Paniai," 2020.
- [2] F. Riana and D. Primasari, "Diagnosis Penyakit Utama Pisang karena Jamur Patogen dengan Dempster-Shafer," *Krea-TIF*, vol. 7, no. 2, p. 72, Nov. 2020, doi: 10.32832/kreatif.v7i2.2649.
- [3] A. Upadhyay, S. Singh, and S. Kanojiya, "Segregation of Ripe and Raw Bananas Using Convolutional Neural Network," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 461–468. doi: 10.1016/j.procs.2023.01.028.
- [4] M. A. B. Bhuiyan, H. M. Abdullah, S. E. Arman, S. Saminur Rahman, and K. Al Mahmud, "BananaSqueezeNet: A very fast, lightweight convolutional neural network for the diagnosis of three prominent banana leaf diseases," *Smart Agricultural Technology*, vol. 4, pp. 1–13, Aug. 2023, doi: 10.1016/j.atech.2023.100214.
- [5] R. Soekarta, N. Nurdjan, and A. Syah, "Klasifikasi Penyakit Tanaman Tomat Menggunakan Metode Convolutional Neural Network (CNN)," *INSECT*, vol. 8, no. 2, pp. 143–151, 2023.
- [6] A. Rohim and Y. Arum Sari, "Convolution Neural Network (CNN) Untuk Pengklasifikasian Citra Makanan Tradisional," 2019. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [7] Y. Arvio, I. Sangadji, and K. Dewantara, "Penerapan Metode Convolution Neural Network (CNN) Dalam Proses Pengolahan Citra Untuk Mendeteksi Cacat Produksi Pada Produk Masker," vol. 16, no. 4, pp. 340–350, 2023, doi: 10.30998/faktorexacta.v16i4.20073.
- [8] R. Gustriansyah, N. Suhandi, S. Puspasari, and A. Sanmorino, "Machine Learning Method to Predict the Toddlers' Nutritional Status," vol. 16, no. 1, pp. 1–12, 2024, doi: 10.20895/INFOTEL.V15I4.988.
- [9] R. Gustriansyah, N. Suhandi, S. Puspasari, A. Sanmorino, and D. Sartika, "Toddlers' Nutritional Status Prediction Using the Multinomial Logistics Regression Method," *Journal of Computer Networks, Architecture and High Performance Computing*, vol. 6, no. 1, pp. 25–33, Dec. 2023, doi: 10.47709/cnahpc.v6i1.3372.
- [10] D. Sartika and R. Gustriansyah, "Comparison of naive Bayes and decision tree algorithms to assess the performance of Palembang City fire and Disaster management employees," *TEKNOSAINS: Jurnal Sains, Teknologi dan Informatika*, vol. 11, no. 1, pp. 132–138, Jan. 2024, doi: 10.37373/tekno.v11i1.843.
- [11] R. Gustriansyah, J. Alie, A. Sanmorino, R. Heriansyah, and M. M. Noor, "Machine Learning for Regencies-Cities Clustering Based on Inflation and Poverty Rates in Indonesia," 2022.
- [12] H. Azis, P. Purnawansyah, F. Fattah, and I. P. Putri, "Performa Klasifikasi K-NN dan Cross Validation Pada Data Pasien Pengidap Penyakit Jantung," *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, pp. 81–86, Aug. 2020, doi: 10.33096/ilkom.v12i2.507.81-86.
- [13] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J Big Data*, vol. 8, no. 54, pp. 1–74, Dec. 2021, doi: 10.1186/s40537-021-00444-8.
- [14] M. Zaid Romegar, C. Widya, Y. Evi, M. Dona, Theresiawati, and I. Ika Nurlaili, "Convolutional Neural Network Analysis on Handwriting Patterns

- and Its Relationship to Personality: A Systematical Review*". IEEE, 2023.
- [15] W. Setiawan, "Perbandingan Arsitektur Convolutional Neural Network Untuk Klasifikasi Fundus," *Jurnal Ilmiah SimanteC*, vol. 7, no. 2, pp. 49–54, 2019.
- [16] Ò. Lorente, I. Riera, and A. Rana, "Image Classification with Classic and Deep Learning Techniques," May 2021, [Online]. Available: <http://arxiv.org/abs/2105.04895>
- [17] Z. R. Mair and M. H. Irfani, "Permainan INGBAS (Gunting, Batu, Kertas) Menggunakan Arsitektur Convolutional Neural Network," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 10, no. 1, pp. 1019–1026, 2023.
- [18] A. Anshu, "Review Paper on Data Mining Techniques and Applications," *International Journal of Innovative Research in Computer Science & Technology*, vol. 7, no. 2, pp. 22–26, Mar. 2019, doi: 10.21276/ijrcst.2019.7.2.4.
- [19] I. Düntsch and G. Gediga, "Confusion matrices and rough set data analysis," 2019.