

Optimasi Resolusi Citra pada MobileNetV3: Analisis Trade-Off Latensi dan Akurasi Deteksi Sampah

^{1*}Dwika Putra Adinata, ²Umi Mahdiah, ³Wahyu Cahyo Utomo

^{1,2,3}Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: *¹dwikaputraadinata@gmail.com, ²umimahdiah@gmail.com,

³wahyu.utomo@unpkdr.ac.id

Penulis Korespondens : Dwika Putra Adinata

Abstrak— Pengelolaan sampah di lingkungan sekolah memerlukan sistem pemilahan otomatis yang mampu beroperasi secara seketika. Namun, penerapan teknologi visi komputer pada perangkat standar sering terkendala oleh tingginya beban komputasi. Penelitian ini bertujuan mengimplementasikan arsitektur MobileNetV3 untuk mengklasifikasikan citra ke dalam kelas organik, anorganik, dan bukan sampah. Eksperimen difokuskan pada modifikasi dimensi citra masukan, yaitu ukuran 160, 224, dan 256 piksel, guna mengevaluasi keseimbangan antara ketepatan prediksi dan durasi komputasi. Hasil pengujian menunjukkan bahwa ukuran 160 piksel merupakan konfigurasi paling efisien yang mampu mencapai ketepatan tertinggi sebesar 0,9910 dengan waktu inferensi paling singkat, yakni 14,99 detik, tanpa dukungan pengolah grafis tambahan. Temuan empiris ini menegaskan kelayakan optimasi resolusi citra dalam mewujudkan infrastruktur deteksi sampah cerdas yang cepat, responsif, dan siap diterapkan pada perangkat berkapasitas rendah.

Kata Kunci— dimensi citra, klasifikasi sampah, MobileNetV3, visi komputer, waktu inferensi

Abstract— Waste management in school environments requires an automated sorting system capable of operating in real-time. However, the application of computer vision technology on standard devices is often hindered by high computational loads. This research aims to implement the MobileNetV3 architecture to classify images into organic, inorganic, and non-waste classes. The experiment focuses on modifying input image dimensions, specifically 160, 224, and 256 pixels, to evaluate the trade-off between prediction accuracy and computational time. The test results indicate that the 160-pixel size is the most efficient configuration, capable of achieving the highest accuracy of 0.9910 with the shortest inference time of 14.99 seconds, without the support of additional graphics processors. These empirical findings confirm the feasibility of image resolution optimization in realizing a smart waste detection infrastructure that is fast, responsive, and ready to be deployed on low-capacity devices.

Keywords— computer vision, image dimensions, inference time, MobileNetV3, waste classification

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Pengelolaan sampah yang cerdas (*smart waste management*) kini menjadi pilar penting dalam mewujudkan ekosistem lingkungan yang berkelanjutan. Salah satu kendala utama di lapangan adalah rendahnya efisiensi pemilahan material organik dan anorganik yang masih mengandalkan tenaga manusia. Penggunaan teknologi kecerdasan buatan, khususnya

Convolutional Neural Network (CNN), telah terbukti mampu memberikan solusi klasifikasi visual secara otomatis dengan tingkat akurasi yang tinggi [1]. Namun, implementasi teknologi ini pada skala praktis seringkali terhambat oleh kebutuhan daya komputasi yang masif [2].

Kebutuhan daya komputasi yang masif tersebut umumnya bersumber dari penggunaan model jaringan konvensional yang memiliki puluhan juta parameter, sehingga secara alami menghasilkan waktu tunda (latensi) yang sangat tinggi [3]. Tingginya latensi ini menjadi kendala fatal bagi perangkat standar yang dituntut untuk mengeksekusi aliran video secara seketika (*real-time*). Untuk memecahkan kebuntuan komputasi tersebut tanpa mengorbankan kualitas deteksi, pendekatan yang paling rasional saat ini adalah beralih menggunakan arsitektur jaringan yang secara arsitektural dirancang jauh lebih ringan [4].

Ringannya beban komputasi tersebut berhasil diwujudkan oleh algoritma MobileNetV3 melalui penerapan mekanisme *Depthwise Separable Convolution*. Meskipun model ini sudah terbukti efisien untuk perangkat terbatas, beban kerja operasionalnya tetap berbanding lurus dengan parameter resolusi masukan. Secara matematis dasar, pemrosesan citra dengan dimensi piksel yang lebih besar akan memperluas ukuran matriks secara kuadratik, yang pada akhirnya pasti akan menyita kapasitas memori dan memicu pelambatan kecepatan inferensi [5].

Inferensi yang melambat akibat tingginya resolusi inilah yang menjadi fokus krusial untuk dievaluasi dalam penelitian ini agar sistem pendeteksi sampah dapat berjalan mulus pada lingkungan web. Penelitian akan mengimplementasikan arsitektur MobileNetV3 dengan memanipulasi variasi dimensi citra masukan, yaitu ukuran 160, 224, dan 256 piksel. Komparasi ini ditujukan secara khusus untuk menemukan titik keseimbangan paling optimal antara durasi pemrosesan tercepat dan ketepatan metrik akurasi. Hasil temuan ini diharapkan mampu memberikan rujukan teknis yang solid dalam merancang infrastruktur pemilah sampah otomatis yang responsif, andal, dan berdaya rendah.

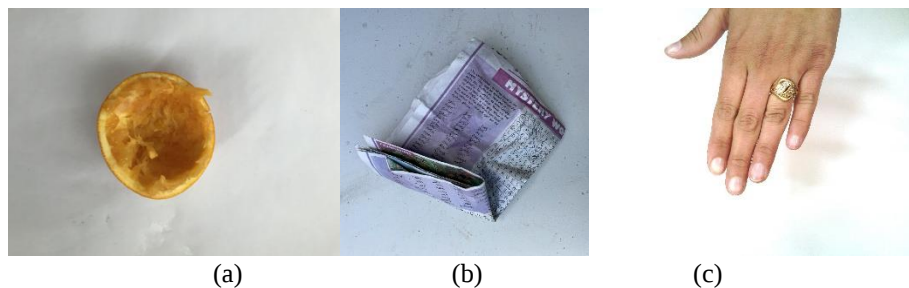
II. METODE

Penelitian ini menggunakan pendekatan eksperimental kuantitatif untuk mengukur dan mengoptimasi kinerja algoritma MobileNetV3 dalam mengklasifikasikan citra persampahan [6]. Pengolahan komputasi secara sengaja dilakukan pada lingkungan berbasis *Central Processing Unit* (CPU) menggunakan perangkat komputasi standar (Intel Core i5, RAM 8GB) tanpa menggunakan akselerator grafis eksternal. Kondisi ini dirancang untuk merepresentasikan spesifikasi perangkat keras riil yang umumnya tersedia di lingkungan sekolah dasar, sekaligus memvalidasi ketangguhan arsitektur model *lightweight* dalam mengeksekusi aliran data secara seketika (*real-time*) [7]. Seluruh tahapan eksperimen, mulai dari pra-pemrosesan hingga inferensi, dieksekusi menggunakan bahasa pemrograman Python beserta pustaka PyTorch dan OpenCV [8].

A. Pengumpulan dan Persiapan Data

Subjek penelitian berupa himpunan data (*dataset*) citra digital yang dikelompokkan menjadi tiga kategori utama: organik, anorganik, dan bukan sampah [9]. Teknik pengumpulan data dilakukan melalui dokumentasi primer di SDN Payaman menggunakan kamera gawai beresolusi tinggi

untuk menangkap objek di lingkungan sekolah. Data ini kemudian dikombinasikan dengan data sekunder dari repositori publik Kaggle guna meningkatkan variasi fitur visual [10]. Contoh visualisasi dari sampel citra untuk masing-masing kategori yang digunakan dalam penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Sampel Citra Himpunan Data: (a) Organik, (b) Anorganik, (c) Bukan Sampah

Agar model terhindar dari bias (*overfitting*) dan dapat dievaluasi secara adil, total 3.352 citra dipartisi menggunakan rasio 80:10:10 [11]. Rincian distribusi pembagian himpunan data untuk proses pelatihan, validasi, dan pengujian dapat dilihat secara detail pada Tabel 1.

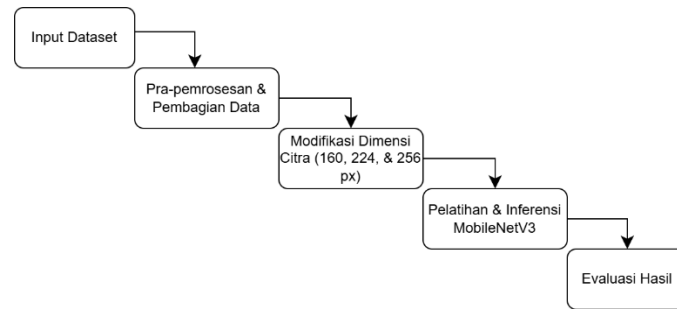
Tabel 1. Distribusi Himpunan Data Citra Sampah

Kategori Kelas	Data Pelatihan	Data Validasi	Data Pengujian
Organik	1.128	141	125
Anorganik	1.128	141	155
Bukan Sampah	427	55	52
Total	2.683	337	332

Tabel 1 menunjukkan struktur alokasi himpunan data di mana porsi terbesar digunakan model untuk mempelajari karakteristik visual objek, sedangkan dua set tersisa diisolasi murni untuk memantau performa selama pelatihan dan menguji ketepatan klasifikasi pada tahap akhir.

B. Rancangan Eksperimen

Eksperimen dirancang untuk menguji dua variabel dependen utama: tingkat keseimbangan prediksi dan kecepatan waktu inferensi (*inference time*) [12]. Alur tahapan penelitian secara keseluruhan direpresentasikan melalui diagram alir pada Gambar 2.



Gambar 2 . Diagram Alir Tahapan Eksperimen

Merujuk pada Gambar 2, prosedur penelitian diawali dengan tahap pra-pemrosesan di mana seluruh matriks nilai piksel citra dinormalisasi ke rentang 0 hingga 1 agar konvergensi pembaruan bobot model berjalan stabil [13]. Pada tahap pemrosesan utama, variabel independen yang dimanipulasi adalah ukuran dimensi citra masukan (*image resizing*). Pengujian difokuskan secara komparatif pada tiga variasi resolusi piksel, yaitu 160, 224, dan 256 piksel, guna mengevaluasi batas efisiensi performa (*trade-off*) di mana model menghasilkan latensi paling rendah namun tetap mempertahankan ketepatan analisis yang tinggi [14].

C. Analisis dan Persamaan Evaluasi Model

Kualitas kinerja model dianalisis menggunakan format *Confusion Matrix* guna mendapatkan nilai metrik statistik [15] dari tebakan sistem yang benar maupun keliru. Struktur dasar matriks evaluasi tersebut dapat dilihat pada Tabel 2.

Tabel 2. Matriks Evaluasi Kinerja (Confusion Matrix) [16]

Kategori Prediksi	Aktual: Positif	Aktual: Negatif
Prediksi: Positif	True Positive (TP)	False Positive (FP)
Prediksi: Negatif	False Negative (FN)	True Negative (TN)

Evaluasi performa prediksi dalam klasifikasi multi-kelas ini diukur secara komprehensif menggunakan metrik akurasi untuk menilai ketepatan sistem secara keseluruhan [17]. Berdasarkan komponen pada Tabel 2, instrumen ukur utama yang diformulasikan adalah nilai akurasi, yang merepresentasikan rasio total prediksi benar (baik positif maupun negatif) terhadap keseluruhan jumlah data yang diuji [18]. Perhitungan matematis evaluasi ketepatan tersebut dirumuskan pada persamaan (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Sementara itu, untuk memvalidasi kelayakan sistem deteksi *real-time* pada spesifikasi perangkat keras standar, penelitian ini juga mengukur dua metrik komputasi. Pertama, waktu inferensi dievaluasi dengan menghitung durasi selisih dari saat tensor citra diterima hingga sistem

menghasilkan probabilitas akhir, yang dirumuskan pada persamaan (2). Kedua, tingkat penggunaan memori (RAM) dicatat secara langsung selama proses evaluasi untuk menganalisis beban operasional model terhadap kapasitas perangkat.

$$InferenceTime = \frac{Total Waktu Pemrosesan}{Jumlah Frame Total} \quad (2)$$

Hasil dari perhitungan pada persamaan (1) dan persamaan (2) tersebut diakumulasikan dan dibandingkan antar setiap skenario resolusi gambar untuk menyimpulkan arsitektur sistem yang paling efisien.

III. HASIL DAN PEMBAHASAN

Bab ini menguraikan hasil pengujian model terhadap variasi modifikasi dimensi citra masukan. Temuan paling fundamental dari eksperimen ini menunjukkan bahwa resolusi 160 piksel merupakan konfigurasi paling efisien, yang secara dini telah mencapai titik jenuh akurasi maksimal sekaligus mencatatkan durasi komputasi paling cepat. Rincian komparasi metrik kinerja model pada resolusi 160, 224, dan 256 piksel disajikan secara kuantitatif pada Tabel 3.

Tabel 3. Komparasi Kinerja MobileNetV3 pada Variasi Resolusi Citra

Ukuran Gambar (Piksel)	Akurasi	Waktu Pengujian (Detik)	Penggunaan RAM
160	0,9910	14,99	1,01 GB
224	0,9910	22,32	0,97 GB
256	0,9910	24,70	1,05 GB

Tabel 3 memperlihatkan secara gamblang bahwa peningkatan dimensi masukan di atas 160 piksel memicu eskalasi perlambatan waktu inferensi tanpa memberikan kontribusi tambahan pada nilai metrik ketepatan klasifikasi. Fenomena stagnasi akurasi ini membuktikan bahwa arsitektur MobileNetV3 telah mampu mengekstraksi fitur pembeda esensial dari kelas organik, anorganik, dan bukan sampah hanya dengan matriks piksel yang minim.

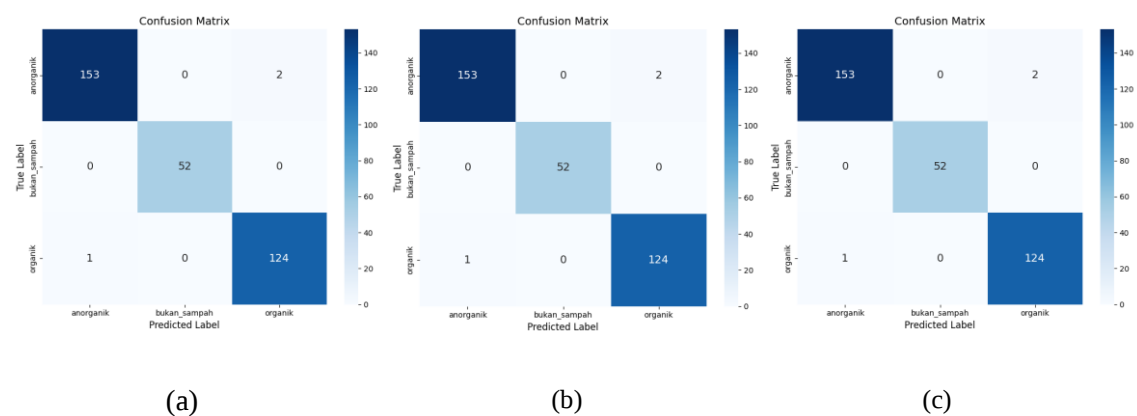
Meskipun terdapat fluktuasi minor pada aspek penggunaan RAM di mana resolusi 160 piksel mencatatkan konsumsi sebesar 1,01 GB atau sedikit lebih tinggi dibandingkan resolusi 224 piksel yang sebesar 973,66 MB akibat karakteristik pengelolaan memori dinamis (*caching allocator*) internal PyTorch hal tersebut tidak mengurangi efisiensi menyeluruh model. Kemampuan model mempertahankan akurasi tinggi pada resolusi kecil ini sangat selaras dengan konsep dasar *depthwise separable convolution*, di mana reduksi parameter dan operasi perkalian matriks berhasil menekan beban komputasi secara signifikan [19].

Signifikannya reduksi latensi waktu tersebut memberikan implikasi langsung terhadap kelayakan teknis implementasi sistem di lapangan. Percepatan proses komputasi memastikan

perangkat keras dengan spesifikasi unit pemroses sentral (CPU) standar di lingkungan sekolah dapat menjalankan klasifikasi dengan lancar tanpa hambatan visual (*lag*). Bukti empiris ini mengukuhkan bahwa rancangan arsitektur yang digunakan sangat ideal untuk dideploy sebagai sistem pendeteksi cerdas berbasis *real-time* [20].

Deteksi *Real-time* dalam operasionalisasinya kelak tentu masih dihadapkan pada beberapa batasan penelitian yang belum terpecahkan pada tahap pengujian ini. Kinerja model yang dilatih pada himpunan data ini mayoritas diasumsikan berjalan pada kondisi pencahayaan yang relatif stabil, sehingga fluktuasi intensitas cahaya ekstrim di area terbuka berpotensi memengaruhi kepastian prediksi visual [21]. Oleh karena itu, penelitian di masa depan perlu mengeksplorasi teknik augmentasi pencahayaan dinamis, serta menguji langsung arsitektur ini pada perangkat *Internet of Things* (IoT) seperti *Raspberry Pi* yang terintegrasi dengan tempat sampah fisik guna memvalidasi ketahanannya di dunia nyata.

Untuk membedah rincian performa klasifikasi secara lebih mendalam, distribusi ketepatan prediksi dari ketiga variasi resolusi citra masukan divisualisasikan melalui *Confusion Matrix* pada Gambar 3.



Gambar 3. *Confusion Matrix* MobileNetV3: (a) 160 Piksel, (b) 224 Piksel, dan (c) 256 Piksel

Berdasarkan Gambar 3, terlihat bahwa ketiga konfigurasi resolusi menghasilkan kemampuan pemisahan kelas yang sangat identik. Tingginya nilai pada diagonal utama di ketiga matriks tersebut menegaskan bahwa mayoritas citra uji berhasil dipetakan ke kelas yang sebenarnya secara konsisten. Mengingat ketiga resolusi mencapai batas jenuh akurasi yang sama (0,9910), pola kesalahan prediksi (misklasifikasi) minor yang terjadi pada matriks (a), (b), dan (c) tidak menunjukkan perbedaan yang berarti. Bukti visual ini semakin memperkuat argumen sebelumnya bahwa peningkatan dimensi masukan di atas 160 piksel tidak menyumbangkan fitur diskriminatif tambahan yang mampu memperbaiki misklasifikasi pada kelas persampahan, melainkan murni hanya memicu inefisiensi waktu komputasi perangkat keras.

IV. KESIMPULAN

Penelitian ini telah berhasil melakukan optimasi resolusi citra pada arsitektur MobileNetV3 untuk sistem deteksi sampah *real-time*. Berdasarkan analisis *trade-off* antara

latensi dan akurasi, disimpulkan bahwa modifikasi dimensi citra masukan merupakan metode yang sangat efektif untuk memangkas beban inferensi sistem tanpa mengorbankan kualitas prediksi kelas persampahan. Resolusi 160 piksel terbukti secara empiris sebagai konfigurasi paling optimal karena mampu menyentuh batas jenuh akurasi tertinggi sebesar 0,9910 dengan tingkat latensi (durasi pemrosesan) paling singkat yakni 14,99 detik, serta efisiensi penggunaan memori RAM sebesar 1,01 GB. Konfigurasi ini secara mutlak mengungguli dimensi masukan yang lebih besar yang justru murni memicu inefisiensi waktu komputasi. Temuan ini secara langsung memberikan kontribusi penting bagi perkembangan bidang ilmu visi komputer dan teknologi lingkungan, dengan membuktikan bahwa model cerdas *lightweight* dapat diimplementasikan secara lancar hanya dengan memanfaatkan *Central Processing Unit* (CPU) berbiaya rendah tanpa intervensi akselerator grafis tingkat lanjut. Implikasi praktis dari penelitian ini membuka jalan bagi terwujudnya infrastruktur pemilahan sampah otomatis yang cepat, hemat daya, dan berkelanjutan untuk diterapkan pada ekosistem lingkungan pendidikan dasar.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada Tuhan Yang Maha Esa, dosen pembimbing, serta seluruh pihak terkait yang telah membantu penulis dalam menyelesaikan penelitian ini.

DAFTAR PUSTAKA

- [1] M. Chhabra, B. Sharan, M. Elbarachi, and M. Kumar, "Intelligent waste classification approach based on improved multi-layered convolutional neural network," pp. 84095–84120, 2024, doi: 10.1007/s11042-024-18939-w.
- [2] J. Teknologi, S. Informasi, J. Nurhidayah, Z. Zulfiandri, and S. Informasi, "Tren dan Tantangan Arsitektur Komputasi Neuromorfik : Tinjauan Literatur Sistematis," vol. 6, no. December 2024, pp. 103–113, 2025, doi: 10.35957/jtsi.v6i1.10046.
- [3] M. A. Shafique and G. S. Member, "ON-CNN : Low Latency and High Throughput Online Arithmetic-Based Convolutional Neural Network Accelerator," *IEEE Access*, vol. 12, no. November, pp. 175698–175714, 2024, doi: 10.1109/ACCESS.2024.3502665.
- [4] J. Zhong, J. Chen, A. Mian, and S. Member, "DualConv : Dual Convolutional Kernels for Lightweight Deep Neural Networks," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 34, no. 11, pp. 9528–9535, 2023, doi: 10.1109/TNNLS.2022.3151138.
- [5] J. Karim, O. F. Goni, M. Ahsan, J. Haider, and M. Kowalski, "Enhancing agriculture through real - time grape leaf disease classification via an edge device with a lightweight CNN architecture and Grad - CAM," *Sci. Rep.*, pp. 1–23, 2024, doi: 10.1038/s41598-024-66989-9.
- [6] J. Patel, H. Talsania, and K. Modi, "Performance Evaluation of Indian Food Image Classification system using Transfer Learning with MobileNetV3," vol. 14, no. 3, 2022, doi: 10.18090/samriddhi.v14i03.24.
- [7] S. M. Raza, S. Murtaza, H. Abidi, and S. Y. Shin, "DSConvNet : A Lightweight Architecture for Extracting Image Features From Depthwise Separable Convolution Network for Edge Devices," *IEEE Access*, vol. 13, no. December, pp. 210102–210116, 2025, doi: 10.1109/ACCESS.2025.3639410.
- [8] S. Naveen and M. R. Kounte, "Optimized Convolutional Neural Network at the IoT edge for image detection using pruning and quantization," pp. 5435–5455, 2025, doi: 10.1007/s11042-024-20523-1.

- [9] R. Nurul *et al.*, “WARNA DAN TEKSTUR MENGGUNAKAN ARTIFICIAL NEURAL NETWORK WASTE TYPE CLASSIFICATION SYSTEM BASED ON A COMBINATION OF COLOR AND TEXTURE FEATURES USING ARTIFICIAL NEURAL NETWORK,” vol. 11, no. 2, pp. 411–420, 2024, doi: 10.25126/jtiik.20241128330.
- [10] L. Fischer-brandies, L. Müller, B. Rebholz, R. Buettner, and S. Member, “To Combine or Not to Combine ? The Influence of Combining Training Datasets on the Robustness of Deep Learning Models : An Analysis for Optical Character Recognition of Handwriting,” no. April, pp. 59039–59056, 2025, doi: 10.1109/ACCESS.2025.3556582.
- [11] U. Sumalatha, K. K. Prakasha, S. Prabhu, and V. C. Nayak, “Enhancing Finger Vein Recognition With Image Preprocessing Techniques and Deep Learning Models,” *IEEE Access*, vol. 12, no. November, pp. 173418–173440, 2024, doi: 10.1109/ACCESS.2024.3498601.
- [12] L. Papa, P. Russo, I. Amerini, L. Zhou, and S. Member, “A Survey on Efficient Vision Transformers : Algorithms , Techniques , and Performance Benchmarking,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 12, pp. 7682–7700, 2024, doi: 10.1109/TPAMI.2024.3392941.
- [13] H. Zhu, Y. Liu, X. Gao, and L. Zhang, “Combined CNN and Pixel Feature Image for Fatty Liver Ultrasound Image Classification,” vol. 2022, 2022, doi: 10.1155/2022/9385734.
- [14] M. H. Saleem, K. K. Velayudhan, and J. Potgieter, “Weed Identification by Single-Stage and Two-Stage Neural Networks : A Study on the Impact of Image Resizers and Weights Optimization Algorithms,” vol. 13, no. April, pp. 1–19, 2022, doi: 10.3389/fpls.2022.850666.
- [15] K. Riehl, M. Neunteufel, and M. Hemberg, “Hierarchical confusion matrix for classification performance evaluation,” *J. R. Stat. Soc. Ser. C Appl. Stat.*, vol. 72, no. 5, pp. 1394–1412, 2023, doi: 10.1093/jrssc/qlad057.
- [16] S. Amaliah and M. Nusrang, “Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi Di Kedai Kopi Konijiwa Bantaeng,” vol. 4, no. 2, pp. 121–127, 2022, doi: 10.35580/variasiunm31.
- [17] O. Rainio, “Evaluation metrics and statistical tests for machine learning,” *Sci. Rep.*, pp. 1–14, 2024, doi: 10.1038/s41598-024-56706-x.
- [18] M. Te Wu, “Confusion matrix and minimum cross - entropy metrics based motion recognition system in the classroom,” *Sci. Rep.*, no. 1821, pp. 1–10, 2022, doi: 10.1038/s41598-022-07137-z.
- [19] C. Y. Kim, “A novel MobileNet with selective depth multiplier to compromise complexity and accuracy,” vol. 45, no. August 2022, pp. 666–677, 2023, doi: 10.4218/etrij.2022-0103.
- [20] W. Lim and M. B. Bonab, “An Aggressively Pruned CNN Model With Visual Attention for Near Real-Time Wood Defects Detection on Embedded Processors,” *IEEE Access*, vol. 11, no. April, pp. 36834–36848, 2023, doi: 10.1109/ACCESS.2023.3266737.
- [21] R. Islam *et al.*, “Deep Learning and Computer Vision Techniques for Enhanced Quality Control in Manufacturing Processes,” *IEEE Access*, vol. 12, no. August, pp. 121449–121479, 2024, doi: 10.1109/ACCESS.2024.3453664.