



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 5 No. 2 (2026) pp: 16536- 16542

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Smart Waste Berbasis IoT Menggunakan Transfer Learning MobileNetV2 dan Flask

Ayu Nursila¹, Husain², Muhammad Zulfikri³

^{1,2,3}Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Bumigora

¹2201020041@universitasbumigora.ac.id , ²husain@universitasbumigora.ac.id , ³mzulfikri@universitasbumigora.ac.id

Abstrak

Tingkat kesadaran masyarakat dalam memilah sampah secara mandiri masih rendah walaupun fasilitas tempat sampah terpisah telah tersedia di berbagai ruang publik. Kondisi ini menyebabkan sistem manajemen limbah konvensional menjadi tidak efektif karena sangat bergantung pada kepatuhan pengguna. Penelitian ini bertujuan mengimplementasikan teknologi deep learning pada perangkat IoT berbiaya rendah menggunakan modul ESP32-CAM untuk menciptakan tempat sampah pintar yang mampu memilah sampah organik dan anorganik secara otomatis. Metode yang digunakan adalah kuantitatif eksperimental dengan menerapkan transfer learning berbasis arsitektur MobileNetV2 yang ringan (lightweight), sehingga sesuai dijalankan melalui skema server-side computing berbasis Flask untuk mengatasi keterbatasan sumber daya mikrokontroler. Model dilatih menggunakan dataset gabungan 4.000 citra dari 11 label yang bersumber dari data publik Kaggle dan potret mandiri ESP32-CAM, melalui dua fase pelatihan yaitu feature extraction dan fine-tuning. Hasil penelitian menunjukkan model mencapai akurasi validasi 90% pada fase feature extraction dan meningkat menjadi 91% setelah fine-tuning, dengan weighted average F1-score 0,91. Pengujian klasifikasi real-time menghasilkan tingkat keberhasilan 90% (99 dari 110 sampel), sedangkan pengujian integrasi end-to-end membuktikan bahwa sensor ultrasonik, ESP32-CAM, server Flask, dan motor servo dapat bekerja sama secara kohesif untuk memilah sampah secara otomatis. Penelitian ini menyimpulkan bahwa pendekatan server-side computing berbasis Flask yang dikombinasikan dengan transfer learning MobileNetV2 mampu mentransformasi tempat sampah konvensional menjadi sistem cerdas yang ekonomis bagi area publik dan lingkungan pendidikan.

Kata Kunci: Smart Waste, Deep Learning, Internet of Things, Convolutional Neural Network, MobileNetV2, Flask

1. Latar Belakang

Permasalahan pengelolaan sampah di Indonesia telah mencapai titik yang kritis. Berdasarkan data Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) tahun 2025 yang mencakup 288 Kabupaten/Kota, total timbulan sampah nasional mencapai 31.461.837,19 ton per tahun, dengan persentase sampah tidak terkelola mencapai 67% atau setara 21.079.430,92 ton per tahun [1]. Kesenjangan ini sebagian besar disebabkan oleh rendahnya partisipasi masyarakat dalam memilah sampah secara mandiri akibat kebiasaan lama dan minimnya pemahaman terhadap dampak negatif tumpukan sampah [2], [3]. Kondisi ini tampak nyata pada ruang publik seperti taman kota maupun lingkungan kampus, termasuk Universitas Bumigora, di mana tempat sampah terpisah kerap diabaikan penggunaannya.

Perkembangan Smart Waste Management (SWM) berbasis Internet of Things (IoT) dan Artificial Intelligence (AI) memungkinkan proses pemilahan dilakukan secara otomatis tanpa bergantung pada kepatuhan pengguna. Metode deep learning seperti Convolutional Neural Network (CNN) terbukti efektif mengenali objek berdasarkan citra visual, dengan akurasi klasifikasi jenis sampah yang dapat mencapai di atas 90% [4], [5]. Prinsip dasar CNN memanfaatkan operasi konvolusi dan skema parameter sharing untuk mendeteksi fitur visual secara efisien [15]. Namun demikian, sebagian besar penelitian terdahulu mengandalkan perangkat keras berperforma tinggi seperti Raspberry Pi yang relatif mahal untuk diterapkan secara masif di fasilitas publik, atau menggunakan sensor sederhana yang hanya mampu mendeteksi volume tanpa mengenali jenis sampah secara spesifik.

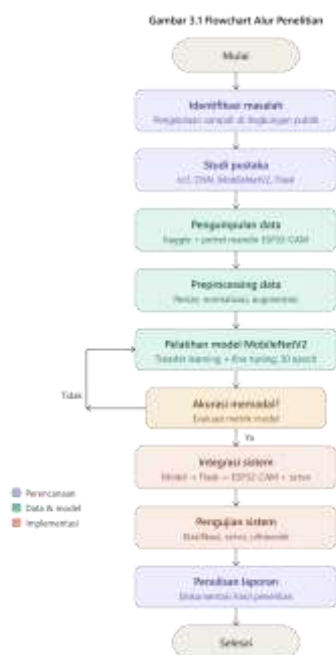
Beberapa penelitian terdahulu telah mengeksplorasi topik serupa. Ismail dkk. membangun tempat sampah IoT berbasis Raspberry Pi dengan sensor HC-SR04 untuk monitoring level sampah secara real-time [6]. Latifah mengembangkan robot pemilah sampah berbasis CNN dan SVM dengan akurasi 86% [4]. Wahid dkk. menerapkan SSD-MobileNet dengan akurasi 88-96% untuk empat kategori limbah [7]. Andriani dkk. menunjukkan implementasi transfer learning MobileNetV2 yang efektif untuk klasifikasi sampah [11]. Firnando dkk. mencatat akurasi 86,73% menggunakan MobileNetV2 pada CPU standar dengan kecepatan inferensi 69 FPS [12]. Nugroho

dkk. mengintegrasikan YOLO berbasis CNN dengan IoT untuk klasifikasi sampah pada skala kota berkelanjutan [13], sementara Roman merancang sistem pemilah sampah berbasis visi komputer menggunakan YOLO [14]. Anjasmara dkk. mengimplementasikan MobileNetV2 langsung pada ESP32-CAM, namun akurasi turun drastis dari 92,5% saat pelatihan menjadi 65,7% pada implementasi nyata akibat real-world environment gap [8]. Syahra menunjukkan bahwa pendekatan server-side computing berbasis Flask mampu mencapai akurasi 92% tanpa membebani mikrokontroler [9].

Berbeda dari penelitian-penelitian tersebut, penelitian ini menggabungkan MobileNetV2, skema server-side computing berbasis Flask, dan ESP32-CAM sekaligus dengan cakupan 11 label klasifikasi serta proporsi data primer hasil potret mandiri mencapai 85% dari total dataset, sehingga memberikan kebaruan pada cakupan label dan proporsi data primer yang lebih besar dibandingkan penelitian acuan. MobileNetV2 dipilih karena arsitektur depthwise separable convolution dan inverted residual block yang digunakan terbukti unggul dalam tugas klasifikasi citra pada perangkat edge tanpa mengorbankan akurasi secara signifikan [10]. Penelitian ini bertujuan untuk: (1) menerapkan transfer learning MobileNetV2 pada perangkat IoT berbiaya rendah untuk klasifikasi sampah organik dan anorganik, (2) membuktikan efektivitas arsitektur server-side computing berbasis Flask dalam mengatasi keterbatasan memori ESP32-CAM, (3) mengevaluasi performa model melalui metrik akurasi, precision, recall, dan F1-score, serta (4) menghasilkan prototipe smart waste yang fungsional untuk area publik dan lingkungan pendidikan.

2. Metode Penelitian

Penelitian ini menggunakan metode kuantitatif eksperimental yang dilaksanakan melalui tujuh tahapan sistematis, yaitu identifikasi masalah, studi pustaka, pengumpulan data, preprocessing data, pelatihan model, akurasi model, integrasi sistem, pengujian sistem, dan penulisan laporan. Alur penelitian secara keseluruhan disajikan pada Gambar 1.



Gambar 1. Alur Penelitian

2.1. Kebutuhan Perangkat Keras dan Perangkat Lunak

Perangkat keras utama terdiri atas ESP32-CAM (AI-Thinker) dengan kamera OV2640 sebagai unit akuisisi citra dan mikrokontroler utama, sensor ultrasonik HC-SR04 sebagai pemicu (trigger) otomatisasi, motor servo sebagai aktuator pemilah dengan mekanisme jungkat-jungkit, modul USB to TTL (FTDI) untuk pemrograman firmware, serta adaptor 5V/2A sebagai catu daya. Pada sisi perangkat lunak, Python beserta framework TensorFlow/Keras digunakan untuk membangun dan melatih model MobileNetV2 pada platform Google Colab dengan akselerasi GPU, sedangkan Flask digunakan sebagai web server API yang menjembatani komunikasi antara ESP32-CAM dan model klasifikasi. Arduino IDE digunakan untuk memprogram firmware ESP32-CAM, sementara OpenCV [17] dan NumPy [18] digunakan untuk decoding dan preprocessing citra pada sisi server.

2.2. Dataset dan Preprocessing

Dataset penelitian menggabungkan dua sumber, yaitu dataset publik dari platform Kaggle untuk memperluas generalisasi model terhadap variasi pencahayaan dan latar belakang, serta data primer berupa potret mandiri menggunakan kamera OV2640 pada ESP32-CAM melalui antarmuka web berbasis Flask. Dataset dikelompokkan ke dalam 11 label yang dipetakan ke tiga kategori keluaran, yaitu organik, anorganik, dan netral/background, sebagaimana dirangkum pada Tabel 1. Total populasi data sebanyak 4.000 citra dibagi menggunakan stratified random split dengan rasio 80:20, menghasilkan 3.200 citra data latih dan 800 citra data validasi.

Tabel 1. Deskripsi Label Dataset dan Pemetaan Kategori

Label	Jumlah	Kategori
Bungkus plastik	250	Anorganik
Daun/ranting	500	Organik
Gelas plastik	250	Anorganik
Kaca	250	Anorganik
Kertas tisu	250	Anorganik
Kresek	250	Anorganik
Logam	250	Anorganik
Plastik recycle	250	Anorganik
Puntung rokok	250	Anorganik
Sisa makanan	500	Organik
Background	1000	Netral

Seluruh proses preprocessing dan augmentasi diimplementasikan melalui kelas ImageDataGenerator dari TensorFlow/Keras secara on-the-fly, mencakup rescaling nilai piksel ke rentang [0,1], resize ke dimensi 224×224 piksel, serta augmentasi berupa rotasi hingga 25°, pergeseran, flipping horizontal, dan zoom. Pada fase fine-tuning, augmentasi brightness_range [0,8, 1,2] ditambahkan untuk meningkatkan ketahanan model terhadap variasi pencahayaan, dengan generator validasi dipisahkan secara mandiri agar evaluasi tetap bersih dari transformasi manipulatif.

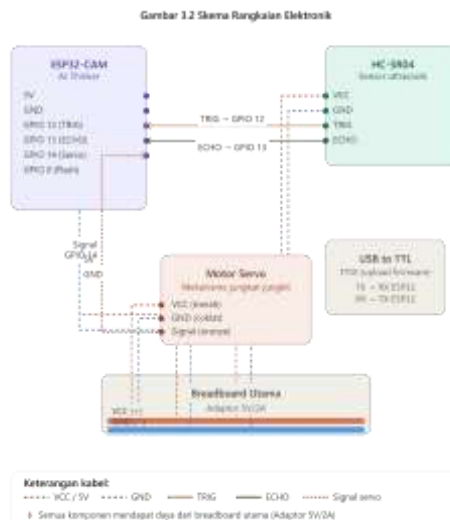
2.3. Arsitektur Model dan Strategi Pelatihan

Model dibangun menggunakan pendekatan transfer learning dengan memuat arsitektur MobileNetV2 beserta bobot pretrained ImageNet (include_top=False, input_shape 224×224×3). Di atas backbone, ditambahkan lapisan GlobalAveragePooling2D, Dense (128, ReLU), Dropout (0,3), dan lapisan keluaran Dense (11, Softmax). Pelatihan dilakukan dalam dua fase menggunakan callback ReduceLROnPlateau dan EarlyStopping. Fase 1 (feature extraction) membekukan seluruh lapisan backbone dan hanya melatih classifier head baru menggunakan optimizer Adam dengan learning rate 1e-4 selama maksimum 30 epoch. Fase 2 (fine-tuning) memuat kembali model hasil Fase 1, membuka (unfreeze) 10 lapisan konvolusi teratas kecuali lapisan Batch Normalization yang tetap dibekukan untuk menjaga stabilitas statistik, menambahkan regularisasi L2 (0,0001) dan meningkatkan dropout menjadi 0,5, serta menurunkan learning rate menjadi 1e-5 guna mencegah catastrophic forgetting. Fungsi loss yang digunakan adalah categorical crossentropy dengan regularisasi L2 mengikuti persamaan (1).

$$L_{total} = L_{crossentropy} + \lambda \Sigma(w^2) \quad (1)$$

dengan L_{total} adalah fungsi kerugian total, $L_{crossentropy}$ adalah nilai categorical crossentropy, λ adalah koefisien regularisasi sebesar 0,0001, dan w adalah bobot pada lapisan Dense yang dibatasi pertumbuhannya untuk mencegah overfitting.

2.4. Perancangan Sistem dan Skema Komunikasi



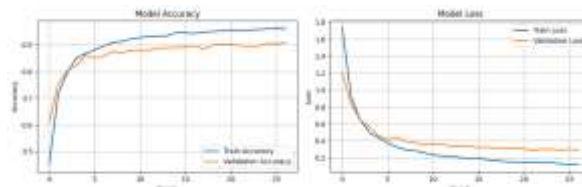
Gambar 2. Skema Rangkaian Elektronik Prototipe

Model terbaik hasil pelatihan (format.keras) ditanamkan ke back-end server Flask yang berjalan pada laptop lokal dalam satu jaringan Wi-Fi dengan ESP32-CAM. Ketika sensor ultrasonik HC-SR04 mendeteksi objek pada jarak kurang dari 10 cm, ESP32-CAM mengambil citra dan mengirimkannya ke server melalui protokol HTTP POST. Server melakukan decoding citra JPEG menggunakan OpenCV, resize dan normalisasi menggunakan NumPy, inferensi model, pemetaan label ke kategori keluaran, lalu mengembalikan respons JSON berisi label, kategori, dan confidence. ESP32-CAM mem-parsing respons tersebut menggunakan pustaka ArduinoJson dan menggerakkan motor servo ke arah kompartemen organik atau anorganik sesuai kategori yang diterima.

3. Hasil dan Diskusi

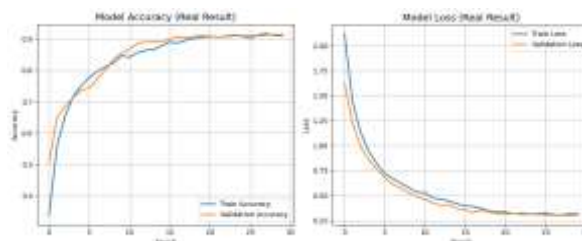
3.1. Hasil Pelatihan Model

Pelatihan Fase 1 (feature extraction) berhenti secara otomatis pada epoch ke-26 melalui mekanisme EarlyStopping, menghasilkan akurasi validasi $\pm 90\%$ dengan training accuracy $\pm 95\%$ (Gambar 3). Selisih yang kecil antara kedua kurva serta pola penurunan loss yang sejajar (dari $\pm 1,7/1,2$ menuju di bawah 0,3) mengindikasikan model tidak mengalami overfitting yang signifikan.



Gambar 3. Kurva Akurasi dan Loss Fase 1 (Feature Extraction)

Pelatihan Fase 2 (fine-tuning) menunjukkan kurva training dan validation accuracy yang bergerak berhimpit dari $\pm 65\%$ pada epoch pertama hingga mencapai $\pm 91\%$ pada epoch ke-29, sebagaimana ditunjukkan pada Gambar 4. Strategi unfreeze selektif 10 lapisan teratas dengan pembekuan lapisan Batch Normalization, regularisasi L2 dan dropout 0,5, serta learning rate $1e-5$ terbukti efektif mencegah overfitting sekaligus catastrophic forgetting, sehingga akurasi berhasil ditingkatkan tanpa merusak pengetahuan yang telah dipelajari pada Fase 1.



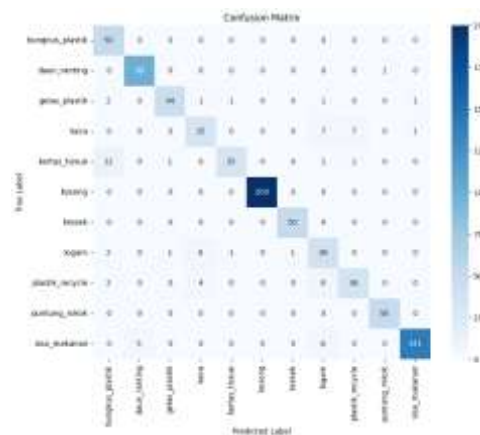
Gambar 4. Kurva Akurasi dan Loss Fase 2 (Fine-Tuning)

3.2. Evaluasi Model

Evaluasi dilakukan menggunakan confusion matrix (Gambar 5 dan Gambar 6) serta classification report yang mencakup precision, recall, dan F1-score untuk setiap label (Tabel 2).



Gambar 5. Confusion Matrix Fase 1



Gambar 6. Confusion Matrix Fase 2

Tabel 2. Perbandingan Metrik Evaluasi Fase 1 dan Fase 2 (F1-Score)

Label	F1 Fase 1	F1 Fase 2
Bungkus plastik	0.80	0.85
Daun/ranting	0.89	0.97
Gelas plastik	0.88	0.92
Kaca	0.76	0.73
Kertas tisu	0.80	0.80
Kosong/background	1.00	1.00
Kresek	0.98	0.99
Logam	0.74	0.70
Plastik recycle	0.79	0.77
Puntung rokok	1.00	0.99
Sisa makanan	0.91	0.96
Akurasi keseluruhan	0.90	0.91

Proses fine-tuning meningkatkan akurasi keseluruhan dari 90% menjadi 91%, dengan peningkatan signifikan pada label daun_ranting (0,89→0,97), bungkus_plastik (0,80→0,85), gelas_plastik (0,88→0,92), dan sisa_makanan (0,91→0,96). Label logam konsisten menjadi label dengan performa terendah pada kedua fase karena variasi visual material logam yang tinggi, sementara label kaca mengalami sedikit penurunan akibat karakteristik transparansi yang sangat dipengaruhi latar belakang dan pencahayaan. Kelas background/kosong konsisten mencapai F1-score sempurna 1,00 pada kedua fase, membuktikan efektivitasnya sebagai kelas negatif yang mencegah aktuasi servo yang tidak perlu.

3.3. Pengujian Klasifikasi Real-Time

Pengujian real-time dilakukan dengan meletakkan 10 sampel objek nyata per label di depan kamera ESP32-CAM, menghasilkan tingkat keberhasilan keseluruhan 90% (99 dari 110 prediksi benar), konsisten dengan akurasi validasi model sebesar 91%. Sebanyak 6 dari 11 label, yaitu bungkus_plastik, daun_ranting, kertas_tissue, kosong, kresek, dan sisa_makanan, mencapai keberhasilan sempurna 100%. Analisis kesalahan menunjukkan bahwa kemiringan posisi objek terhadap kamera mempengaruhi hampir seluruh kelas yang mengalami kesalahan klasifikasi, terutama gelas_plastik (70%), kaca (80%), logam (80%), dan plastik_recycle (80%), sementara pantulan cahaya pada material kaca, logam, dan plastik transparan menciptakan ambiguitas visual antar-kelas.

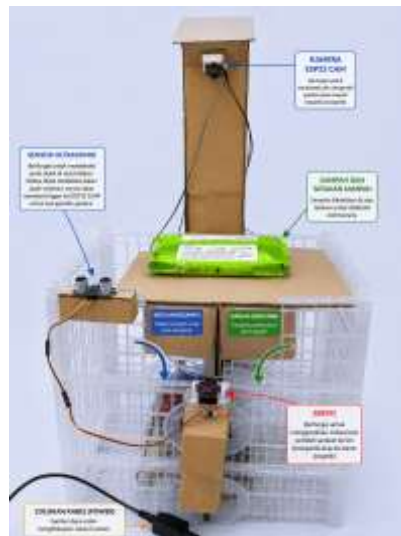
3.4. Pengujian Integrasi Sistem End-to-End

Pengujian integrasi memverifikasi sinkronisasi seluruh alur kerja mulai dari deteksi objek oleh sensor ultrasonik, pengiriman citra via HTTP POST, inferensi model, penerimaan respons JSON, hingga pergerakan servo. Hasil pengujian dirangkum pada Tabel 3.

Tabel 3. Hasil Pengujian Integrasi Sistem

Aspek Pengujian	Hasil
Deteksi objek ultrasonik (standby ± 291 cm, trigger < 10 cm)	Berhasil
Pengiriman gambar ke Flask (HTTP 200)	Berhasil
Inferensi model MobileNetV2	Berhasil
Penerimaan & parsing respons JSON	Berhasil
Pergerakan servo sesuai kategori	Berhasil
Stabilitas sistem (read timeout)	Sebagian, auto-recover
Pemilahan sampah fisik	Berhasil

Seluruh komponen, yaitu sensor ultrasonik, ESP32-CAM, server Flask, model MobileNetV2, dan motor servo, terbukti bekerja sama dalam satu alur kerja yang kohesif. Kejadian read timeout yang muncul sesekali (1-2 kali per sesi) diduga disebabkan oleh variasi ukuran file JPEG, latensi jaringan Wi-Fi lokal, dan waktu inferensi yang lebih lama pada kondisi CPU sibuk, namun tidak mengganggu fungsionalitas karena sistem mampu pulih secara otomatis (auto-recover) tanpa restart manual. Bentuk fisik prototipe yang diuji ditunjukkan pada Gambar 7.



Gambar 7. Prototipe Sistem Smart Waste Management

4. Kesimpulan

Penelitian ini membuktikan bahwa transformasi tempat sampah konvensional menjadi sistem smart waste dapat diwujudkan dengan komponen berbiaya rendah melalui integrasi ESP32-CAM, sensor ultrasonik HC-SR04, motor servo, dan server Flask. Arsitektur server-side computing efektif mengatasi keterbatasan memori ESP32-CAM dengan mendelegasikan seluruh komputasi inferensi ke server. Strategi transfer learning dua fase pada MobileNetV2 meningkatkan akurasi validasi dari 90% (feature extraction) menjadi 91% (fine-tuning) dengan weighted average F1-score 0,91, konsisten dengan tingkat keberhasilan pengujian real-time sebesar 90%. Faktor kemiringan posisi objek dan pantulan cahaya pada material reflektif teridentifikasi sebagai tantangan utama,

khususnya pada label kaca, logam, dan plastik recycle, sementara kelas background sebagai kelas negatif terbukti krusial dengan F1-score sempurna 1,00 pada seluruh pengujian. Pengujian integrasi end-to-end membuktikan seluruh komponen sistem dapat bekerja sama secara kohesif dengan ketahanan yang baik terhadap gangguan komunikasi sesaat. Penelitian lanjutan disarankan untuk memperbanyak dan memvariasikan dataset menggunakan kamera ESP32-CAM langsung dengan variasi sudut dan pencahayaan, memindahkan back-end Flask ke server cloud agar sistem lebih skalabel untuk implementasi multi-unit, serta melakukan pengujian pada lingkungan publik nyata sebagai langkah menuju ekosistem smart bin yang terintegrasi dalam kerangka smart city.

Referensi

- [1] Kementerian Lingkungan Hidup dan Kehutanan (KLHK), "SIPSN - Sistem Informasi Pengelolaan Sampah Nasional," 2025. [Daring]. Tersedia: <https://sampahnasional.kemenvh.go.id/portal-indikatif/>
- [2] M. Zubair Muda dan A. Nurmanina, "Perilaku masyarakat dalam membuang sampah di Kawasan Kelurahan Sungai Pinang Luar Kecamatan Samarinda Kota Samarinda," *Jurnal Ilmiah*, vol. 2025, no. 1, pp. 55-65, 2025.
- [3] R. A. Pratama, P. K. Utomo, dan R. D. Jati, "Perilaku masyarakat dalam membuang sampah di tempat penampungan sementara (TPS) di Kecamatan Pontianak Barat Kota Pontianak," 2016.
- [4] R. A. Latifah, "Pengembangan robot pemilah sampah menggunakan Convolutional Neural Network," Skripsi, UIN Syarif Hidayatullah Jakarta, 2022.
- [5] L. Stephen Pieters, "Development of automatic waste classification system using CNN based deep learning to support smart waste management," *Jurnal Inovtek Polbeng*, vol. 10, no. 1, 2025.
- [6] M. A. Ismail dkk., "Tempat sampah pintar berbasis Internet of Things (IoT) dengan sistem teknologi informasi," 2021.
- [7] Z. Wahid, M. Kallista, dan I. P. D. Wibawa, "Deep learning untuk sistem pemilah sampah otomatis dengan algoritma SSD-MobileNet," *e-Proceeding of Engineering*, 2024.
- [8] E. R. Anjasmara, F. Fattah, dan A. W. M. Gaffar, "Implementasi model MobileNetV2 pada ESP-32 CAM untuk klasifikasi botol plastik," *Jurnal Minfo Polgan*, vol. 15, no. 2, pp. 1262-1271, 2026, doi: 10.33395/jmp.v15i2.16197.
- [9] Y. Syahra, "Inovasi tempat sampah IoT dengan otomatisasi dan mekanisme keamanan modern," 2026.
- [10] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, dan L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," *arXiv preprint*, 2018.
- [11] F. Andriani, A. Yulianti, A. Yaturahmah, dan S. Mutmainah, "MobileNetV2 transfer learning implementation for waste classification," *Indonesian Applied Research Computing and Informatics*, vol. 1, no. 2, pp. 37-46, 2026, doi: 10.64479/iarci.v1i2.62.
- [12] R. Firmando, M. Ali Buchari, A. Dwi Marjusalinah, dan R. Fadli Isnanto, "Performance analysis of MobileNetV2 based automatic waste classification using transfer learning," *Jurnal Mandiri*, vol. 14, no. 1, pp. 257-268, 2025.
- [13] W. Nugroho, A. Alfattah, M. Jimmy, F. Arifianto, dan A. Hadi, "Automated waste classification for sustainable cities using YOLO based CNN integrated IoT," *Jurnal Sistem Cerdas*, 2025.
- [14] A. Roman, "Rancang bangun sistem pemilah sampah secara otomatis berbasis visi komputer menggunakan YOLO," 2024.
- [15] I. Goodfellow, Y. Bengio, dan A. Courville, *Deep Learning*. MIT Press, 2016.
- [16] D. Heryadi dan T. Wahyono, *Machine Learning*. Penerbit Andi, 2020.
- [17] G. Bradski, "The OpenCV library," *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, vol. 25, no. 11, pp. 120-123, 2000.
- [18] C. R. Harris dkk., "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357-362, 2020.