

# PENDEKATAN PENGURANGAN OVERFITTING PADA MOBILENET UNTUK KLASIFIKASI CITRA SAMPAH

Irsyad Badruddin<sup>1</sup>, Tegar Risky Nugrahanto<sup>2</sup>, Oktafian Dyah Pangesti<sup>3</sup>,  
Tinuk Agustin\*<sup>4</sup>

<sup>123</sup>Informatika, STMIK AMIKOM Surakarta

<sup>123</sup>Jawa Tengah, Indonesia

Email: <sup>1</sup>[irsyad.10503@mhs.amikomsolo.ac.id](mailto:irsyad.10503@mhs.amikomsolo.ac.id),

<sup>2</sup>[tegar.10486@mhs.amikomsolo.ac.id](mailto:tegar.10486@mhs.amikomsolo.ac.id), <sup>3</sup>[oktafian.10490@mhs.amikomsolo.ac.id](mailto:oktafian.10490@mhs.amikomsolo.ac.id),

<sup>4</sup>[agustin.amikom@gmail.com](mailto:agustin.amikom@gmail.com)

## Abstract

*The challenge faced in waste management, especially in waste image classification, is the low accuracy of the model which can cause errors in sorting waste types, thus reducing the effectiveness of the recycling process. This study aims to develop a more accurate waste image classification model using the MobileNet approach, which is enhanced with overfitting reduction techniques such as data augmentation, dropout, and L2 regularization. The methodology used is experimental research. The first model was built without overfitting reduction techniques, while the second model applied augmentation, dropout, and L2 regularization techniques. The training results of the first model showed a high training accuracy of up to 99.95%, but there was a decrease in validation accuracy of up to 80%, indicating overfitting. In contrast, the second model managed to maintain a more stable and higher validation accuracy, with a validation accuracy reaching 91.24% at the end of the epoch, indicating an increase in the model's generalization ability to unseen data. These findings contribute to improving the accuracy and efficiency of waste sorting, which supports better waste management and recycling systems, as well as the development of machine learning models in waste and environmental management.*

**Keywords:** Image Classification, MobileNet, Overfitting, Data Augmentation, L2 Regularization

## Abstraksi

*Tantangan yang dihadapi dalam pengelolaan sampah, khususnya dalam klasifikasi citra sampah, adalah rendahnya akurasi model yang dapat menyebabkan kesalahan dalam pemilahan jenis sampah, sehingga mengurangi efektivitas proses daur ulang. Penelitian ini bertujuan untuk mengembangkan model klasifikasi citra sampah yang lebih akurat dengan menggunakan pendekatan MobileNet, yang disempurnakan dengan teknik reduksi overfitting seperti augmentasi data, dropout, dan regularisasi L2. Metodologi yang digunakan adalah penelitian eksperimental. Model pertama dibangun tanpa teknik pengurangan overfitting, sementara model kedua menerapkan teknik augmentasi, dropout, dan regularisasi L2. Hasil pelatihan model pertama menunjukkan akurasi pelatihan yang tinggi hingga 99,95%, namun terjadi penurunan akurasi validasi hingga 80%, mengindikasikan overfitting. Sebaliknya, model kedua berhasil mempertahankan akurasi validasi yang lebih stabil dan lebih tinggi, dengan akurasi validasi mencapai*

91,24% pada akhir epoch, yang menunjukkan peningkatan kemampuan generalisasi model terhadap data yang tidak terlihat. Temuan ini memberikan kontribusi dalam meningkatkan akurasi dan efisiensi pemilahan sampah, yang mendukung sistem pengelolaan sampah dan daur ulang yang lebih baik, serta pengembangan model pembelajaran mesin dalam pengelolaan sampah dan lingkungan.

**Kata Kunci:** Klasifikasi Citra, MobileNet, Overfitting, Augmentasi Data, Regularisasi L2

## 1. PENDAHULUAN

Sampah merupakan masalah global yang mendesak, terutama di negara berkembang seperti Indonesia. Menurut data dari Badan Pusat Statistik, Indonesia menghasilkan lebih dari 35 juta ton sampah setiap tahun, di mana sebagian besar tidak dikelola dengan baik[1]. Untuk mengatasi masalah ini, teknologi pengenalan citra menggunakan algoritma Convolutional Neural Network (CNN) seperti MobileNet menjadi solusi yang menjanjikan. MobileNet dirancang untuk berjalan pada perangkat dengan sumber daya terbatas, menjadikannya ideal untuk aplikasi mobile dan Internet of Things (IoT)[2].

Overfitting terjadi ketika model machine learning terlalu kompleks dan belajar dari data pelatihan dengan sangat baik, tetapi gagal untuk generalisasi pada data yang belum pernah dilihat sebelumnya. Model yang memiliki terlalu banyak parameter atau lapisan dapat menyesuaikan diri dengan noise dalam data pelatihan, bukannya pola yang sebenarnya. Ini sering terjadi pada model neural network yang dalam, seperti CNN, ketika jumlah neuron atau lapisan terlalu banyak dibandingkan dengan jumlah data pelatihan yang tersedia[3].

Penelitian ini bertujuan untuk mengeksplorasi dan menerapkan pendekatan pengurangan overfitting pada model MobileNet dalam konteks klasifikasi citra sampah. Dengan memanfaatkan dataset citra sampah yang beragam, penelitian ini diharapkan dapat memberikan kontribusi signifikan terhadap pengembangan sistem klasifikasi yang lebih robust dan efektif.

Metode yang akan digunakan dalam penelitian ini mencakup pengumpulan dataset citra sampah, penerapan teknik augmentasi data untuk meningkatkan keragaman data pelatihan, serta implementasi MobileNet dengan penyesuaian parameter untuk meminimalkan overfitting[4].

Hasil dari penelitian ini diharapkan dapat memberikan wawasan baru tentang penerapan MobileNet dalam klasifikasi citra sampah dan menawarkan solusi praktis untuk pengelolaan sampah yang lebih baik. Dengan meningkatnya kesadaran akan pentingnya pengelolaan limbah yang efektif dan berkelanjutan, penelitian ini juga dapat menjadi referensi bagi peneliti lain dalam bidang machine learning dan pengolahan citra[5]. Selain itu, sistem klasifikasi otomatis berbasis MobileNet dapat membantu mengurangi beban kerja manual dalam pemisahan sampah dan mendukung upaya daur ulang secara lebih efisien.

Dengan demikian, pendekatan pengurangan overfitting pada MobileNet tidak hanya relevan secara akademis tetapi juga memiliki implikasi praktis yang signifikan dalam pengelolaan lingkungan dan keberlanjutan. Penelitian ini berpotensi memberikan kontribusi terhadap perkembangan teknologi klasifikasi limbah yang lebih baik dan lebih efisien di masa depan.

## **2. TINJAUAN PUSTAKA**

Overfitting adalah kondisi di mana model machine learning terlalu cocok dengan data pelatihan, sehingga kehilangan kemampuan untuk generalisasi pada data yang tidak terlihat. Ini terjadi ketika model memiliki kompleksitas yang tinggi, misalnya dengan banyak parameter, yang membuatnya dapat mengingat data pelatihan alih-alih belajar dari pola yang ada. Beberapa pendekatan untuk mendeteksi dan mengatasi overfitting mencakup penggunaan teknik seperti K-fold cross-validation dan penghentian awal (early stopping) selama proses pelatihan[6].

MobileNet adalah arsitektur Convolutional Neural Network (CNN) yang dirancang untuk efisiensi dalam pengolahan citra, terutama pada perangkat mobile. MobileNet menggunakan teknik depthwise separable convolutions untuk mengurangi jumlah komputasi, menjadikannya pilihan yang baik untuk aplikasi real-time dan embedded systems. Dalam penelitian sebelumnya, MobileNetV1 dan V2 telah terbukti efektif dalam mengatasi masalah overfitting, terutama saat diterapkan pada dataset kecil[7][8]. Penelitian oleh Gulzar (2023) menunjukkan bahwa modifikasi pada lapisan akhir MobileNet dengan penambahan dropout dan global average pooling dapat meningkatkan akurasi klasifikasi citra[9].

Strategi umum untuk mengurangi overfitting dalam model deep learning yaitu pelatihan sebelum model mulai mempelajari kebisingan dalam data, meningkatkan jumlah data pelatihan dengan membuat variasi dari data yang ada, seperti rotasi atau perubahan pencahayaan, mengidentifikasi dan menghilangkan fitur yang tidak relevan atau berlebihan dari dataset[6][9], memanfaatkan model yang telah dilatih sebelumnya pada dataset besar untuk meningkatkan performa pada dataset kecil[8].

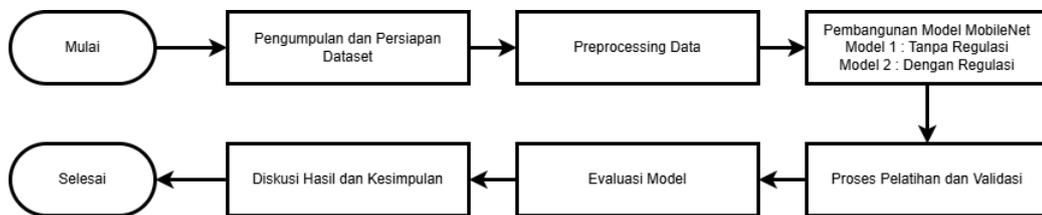
Penelitian oleh Kanjanawattana et al. (2023) menunjukkan bahwa penggunaan MobileNetV1 dan V2 dalam klasifikasi citra dapat mencapai akurasi tinggi, dengan MobileNetV1 mencapai 89,93% dan MobileNetV2 89,78%. Hal ini menunjukkan efektivitas arsitektur ini dalam konteks klasifikasi citra dengan tantangan overfitting[7].

## **3. METODE PENELITIAN**

Penelitian ini merupakan penelitian eksperimental yang bertujuan untuk menilai efektivitas strategi pengurangan overfitting yang diterapkan pada arsitektur MobileNet dalam konteks klasifikasi citra gambar. Eksperimen tersebut melibatkan analisis komparatif dari dua model MobileNet: satu yang beroperasi tanpa teknik pengurangan overfitting dan kedua menggabungkan metode untuk mengurangi overfitting (augmentasi data, dropout, dan regularisasi L2).

### 3.1. Alur Penelitian

Penelitian ini dimulai dengan pengumpulan dan penyiapan dataset, diikuti oleh preprocessing data yang mencakup augmentasi dan normalisasi. Dua model MobileNet dibangun: satu tanpa teknik regularisasi untuk menilai potensi overfitting, dan satu lagi yang menggabungkan metode regularisasi (dropout, augmentasi, L2) yang ditujukan untuk mengurangi overfitting. Selanjutnya, model menjalani pelatihan dan validasi, yang selama pelatihan tersebut akurasi dan kerugiannya dipantau. Evaluasi model dilakukan menggunakan set pengujian, dan hasilnya dibandingkan untuk memeriksa dampak regularisasi pada kinerja model. Agar lebih jelas, alur penelitian dapat divisualisasikan dalam diagram alir yang disajikan pada Gambar 1.

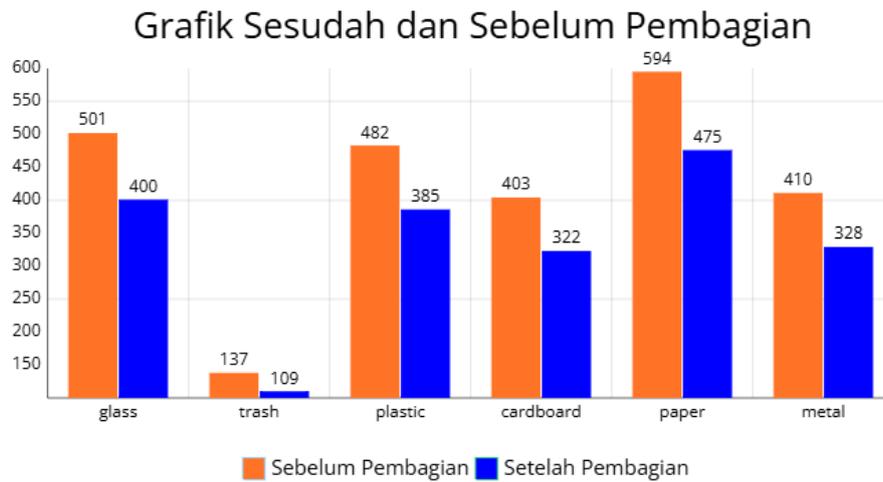


Gambar 1. Flowchart Alur Penelitian

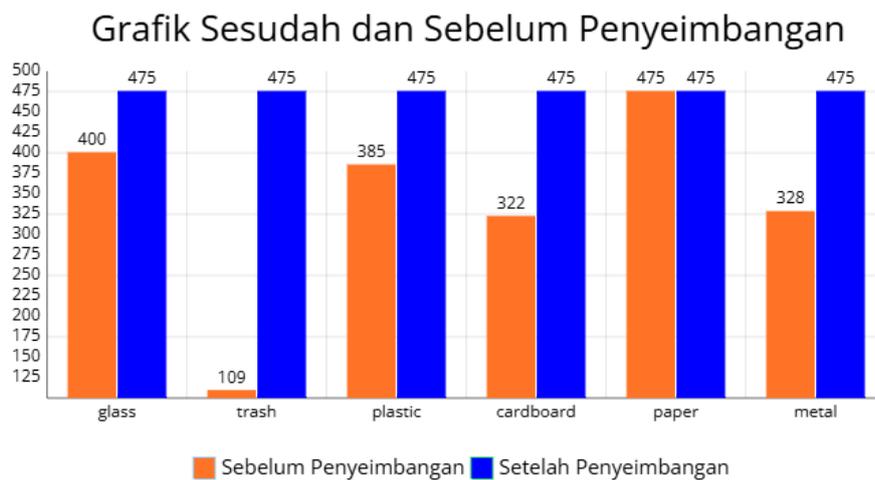
### 3.2. Dataset

Dataset digunakan untuk melatih model CNN agar dapat mengenali pola dan fitur dari data yang diberikan. Kualitas dan kuantitas dataset sangat mempengaruhi performa model. Model yang dilatih dengan dataset berkualitas tinggi cenderung menghasilkan akurasi yang lebih baik dalam klasifikasi atau deteksi objek[10].

Dataset yang digunakan dalam penelitian ini terdiri dari sekitar 2.000 gambar yang dikelompokkan ke dalam enam kelas: paper, metal, glass, plastic, cardboard, dan trash. Dataset ini pertama-tama dibagi menjadi tiga set: training set (80%), validation set (10%), dan testing set (10%) untuk tujuan pelatihan, evaluasi, dan pengujian model. Setelah pembagian, dilakukan penyeimbangan pada training set untuk memastikan tiap kelas memiliki 475 gambar, sehingga model mendapatkan representasi yang setara dari semua kelas selama proses pelatihan. Proses pembagian dan penyeimbangan dataset ini divisualisasikan pada Gambar 2 dan Gambar 3.



Gambar 2. Grafik Pembagian untuk Training set



Gambar 3. Grafik Penyeimbangan untuk Training set

Ketika dataset tidak seimbang, model cenderung bias terhadap kelas mayoritas, yang dapat menyebabkan akurasi yang tinggi tetapi rendahnya kemampuan dalam mendeteksi kelas minoritas. Penyeimbangan membantu model untuk memperhatikan semua kelas secara lebih adil, meningkatkan akurasi keseluruhan[11].

### 3.3. Skenario Pengujian

Skenario pengujian memungkinkan peneliti untuk menguji berbagai kombinasi parameter model (seperti learning rate dan jumlah lapisan) untuk menemukan konfigurasi yang memberikan hasil terbaik[12]. Dengan melakukan pengujian di berbagai skenario, peneliti dapat mengidentifikasi masalah yang mungkin muncul, seperti overfitting atau underfitting, serta memahami bagaimana model bereaksi terhadap variasi dalam data[13].

Terdapat dua skenario pengujian dalam penelitian ini. Skenario 1 (Tanpa Pengurangan Overfitting) menguji Model MobileNet tanpa menggunakan teknik

pengurangan overfitting, untuk mengamati potensi overfitting yang terjadi. Skenario 2 (Dengan Pengurangan Overfitting) menggunakan teknik augmentasi data, *dropout* sebesar 0.2, dan regularisasi L2 dengan faktor 0.0005 untuk mengurangi overfitting. Setiap skenario dilatih selama 20 epoch dengan ukuran batch 32 dan resolusi gambar 128x128.

### 3.4. Teknik Preprocessing

Praproses data adalah serangkaian teknik yang diterapkan untuk mengubah data mentah menjadi format yang lebih bersih dan terstruktur. Proses ini mencakup berbagai langkah seperti pembersihan data, integrasi data, transformasi data, dan pengurangan data. Tujuan utamanya adalah untuk meningkatkan kualitas data sehingga hasil analisis menjadi lebih akurat dan dapat diandalkan[14]. Teknik augmentasi membantu model menemukan fitur yang lebih beragam dalam data, sehingga meningkatkan kemampuannya untuk melakukan klasifikasi atau deteksi objek dengan lebih baik[12].

Teknik preprocessing data dalam penelitian ini meliputi augmentasi data untuk meningkatkan variasi dan mengatasi ketidakseimbangan kelas. Augmentasi ini dilakukan melalui beberapa teknik, termasuk *rotation\_range* hingga 40 derajat untuk menghadapi gambar dari berbagai perspektif, *width\_shift\_range* dan *height\_shift\_range* sebesar 0.2 untuk mensimulasikan translasi kecil, *shear\_range* dan *zoom\_range* masing-masing sebesar 0.2 untuk memberikan variasi bentuk dan ukuran, serta *horizontal\_flip* dan *fill\_mode 'nearest'* untuk menambah keragaman posisi objek. Teknik-teknik ini diharapkan membantu model dalam mengenali fitur dari berbagai kondisi dan meningkatkan kemampuan generalisasi model.

### 3.5. Model CNN yang Dibangun

Penelitian ini membangun dua arsitektur model untuk menguji dampak pengurangan overfitting pada model MobileNet. Arsitektur pertama (tanpa pengurangan overfitting) dan arsitektur kedua (dengan pengurangan overfitting) memiliki struktur dasar yang sama, dengan beberapa perbedaan pada penerapan dropout, augmentasi, dan regularisasi. Perbandingan antara kedua arsitektur model tersebut dapat dilihat pada Tabel 1 berikut.

Tabel 1. Perbandingan antara kedua arsitektur model

Aspek	Model Tanpa Pengurangan Overfitting	Model Dengan Pengurangan Overfitting
Base Model	MobileNet (include_top=False)	MobileNet (include_top=False)
Layer Setelah Base Model	GlobalAveragePooling2D	GlobalAveragePooling2D
Layer Dense	512 neuron, ReLU	512 neuron, ReLU + L2 Regularization (0.0005)
Dropout	Tidak ada	Dropout 0.2
Output Layer	Dense layer dengan softmax	Dense layer dengan softmax
Optimizer	Adam (learning rate 0.0001)	Adam (learning rate 0.00005)
Loss Function	Categorical Crossentropy	Categorical Crossentropy
Metrics	Accuracy	Accuracy

### 3.6. Evaluasi Model

Evaluasi model dilakukan dengan membandingkan akurasi dan validasi loss dari kedua skenario selama pelatihan dan pada data pengujian. Metrik ini digunakan untuk mengukur generalisasi model dan efektivitas teknik pengurangan overfitting. Hasil evaluasi ditampilkan dalam bentuk grafik yang menunjukkan tren akurasi dan loss dari kedua model, serta ringkasan akurasi pada data pengujian.

## 4. HASIL DAN PEMBAHASAN

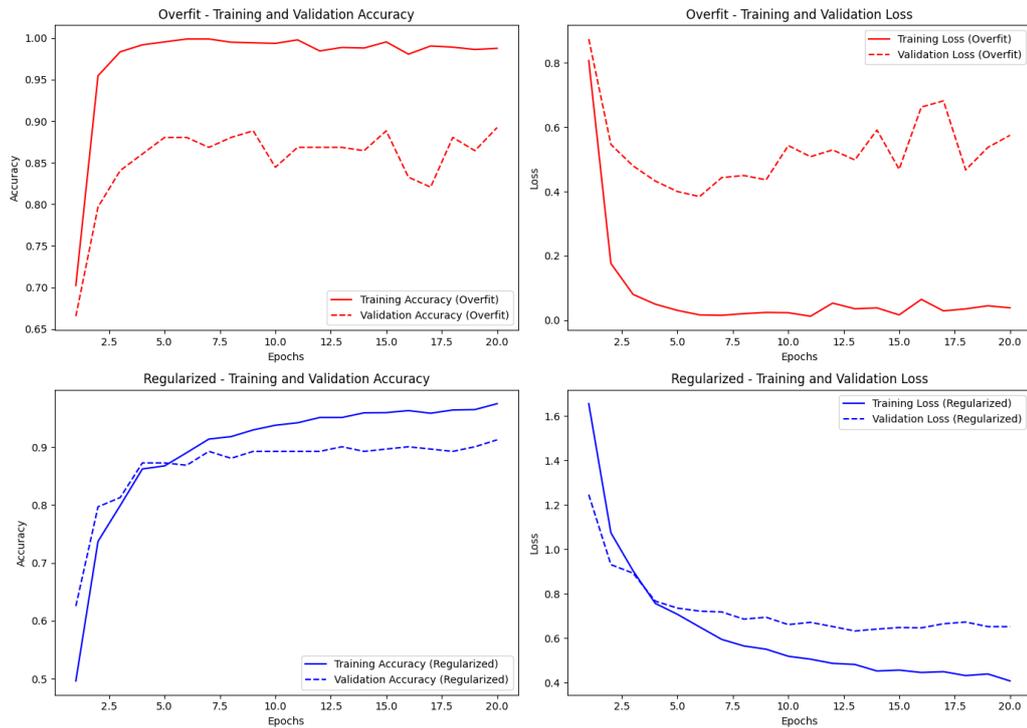
Hasil dan pembahasan dari penelitian yang menggunakan metode MobileNet untuk klasifikasi citra sampah dengan pengurangan overfitting menunjukkan efektivitas teknik ini dalam meningkatkan kemampuan generalisasi model. Pengurangan overfitting dilakukan dengan menggunakan teknik seperti augmentasi data, dropout, dan regularisasi L2. Teknik-teknik ini terbukti membantu model untuk belajar pola yang lebih umum dan tidak terjebak pada pola yang hanya ada dalam data pelatihan. Berikut adalah ringkasan hasil dan pembahasan berdasarkan eksperimen yang dilakukan pada model dengan dan tanpa pengurangan overfitting.

Hasil yang diperoleh dari eksperimen ini disajikan dalam bentuk tabel dan grafik untuk memudahkan analisis dan pemahaman mengenai efektivitas teknik pengurangan overfitting pada model MobileNet. Tabel 2 menyajikan hasil perbandingan akurasi dan loss dari kedua model (tanpa dan dengan pengurangan overfitting) pada data pelatihan, validasi, dan pengujian.

Tabel 2. Perbandingan Akurasi dan Loss pada Data Pengujian

Model	Akurasi Pelatihan	Akurasi Validasi	Akurasi Pengujian	Loss Pelatihan	Loss Validasi	Loss Pengujian
Tanpa Pengurangan Overfitting	98.89%	89.24%	88.04%	0.0355	0.5748	0.5515
Dengan Pengurangan Overfitting	97.91%	91.24%	91.70%	0.4052	0.6514	0.5956

Pada tabel di atas, terlihat bahwa model tanpa pengurangan overfitting memiliki akurasi pelatihan yang lebih tinggi, tetapi akurasi validasi dan pengujian lebih rendah dibandingkan dengan model yang menggunakan teknik pengurangan overfitting. Hal ini menunjukkan adanya indikasi overfitting pada model tanpa regularisasi, yang berhasil diminimalkan pada model yang menggunakan dropout, augmentasi, dan regularisasi L2.



Gambar 4. Grafik Akurasi dan Loss selama Pelatihan

Grafik menunjukkan tren akurasi dan loss dari kedua model sepanjang 20 epoch pelatihan dapat dilihat pada Gambar 4. Model tanpa pengurangan overfitting menunjukkan akurasi yang meningkat tajam di awal pelatihan, namun kemudian mengalami penurunan pada set validasi setelah beberapa epoch. Di sisi lain, model dengan pengurangan overfitting menunjukkan tren akurasi yang stabil dengan perbedaan loss yang tidak terlalu signifikan antara data pelatihan dan validasi.

## 5. KESIMPULAN

Kesimpulan dari penelitian ini menunjukkan bahwa penggunaan teknik pengurangan overfitting pada model MobileNet untuk klasifikasi citra sampah berhasil meningkatkan kemampuan generalisasi model. Model pertama yang dibangun tanpa teknik pengurangan overfitting menunjukkan akurasi pelatihan yang sangat tinggi tetapi mengalami penurunan signifikan pada akurasi validasi, yang mengindikasikan overfitting. Sebaliknya, model kedua yang menggabungkan teknik augmentasi data, dropout, dan regularisasi L2 berhasil mempertahankan akurasi validasi yang lebih tinggi dan stabil, yang menunjukkan kemampuan model dalam mengenali pola secara lebih umum dan meningkatkan akurasi pada data yang tidak terlihat sebelumnya. Temuan ini memberikan kontribusi penting dalam pengembangan sistem klasifikasi sampah yang lebih efisien dan dapat diterapkan untuk mendukung pengelolaan sampah yang lebih baik dan lebih berkelanjutan.

## DAFTAR PUSTAKA

- [1] A. R. Fahcruroji, "Implementasi Algoritma Cnn Mobilenet Untuk Klasifikasi Gambar Sampah Di Bank Sampah Program Studi Matematika," 2024.
- [2] S. Alden and B. N. Sari, "Implementasi Algoritma CNN Untuk Pemilahan Jenis Sampah Berbasis Android Dengan Metode CRISP-DM," *Jurnal Informatika*, vol. 10, no. 1, pp. 62–71, Mar. 2023, doi: 10.31294/inf.v10i1.14985.
- [3] M. Valdenegro-Toro and M. Sabatelli, "Machine Learning Students Overfit to Overfitting," Sep. 2022, [Online]. Available: <http://arxiv.org/abs/2209.03032>
- [4] H. Oktafiandi and Winarnie, "Implementasi Algoritma Convolution Neural Network pada Klasifikasi Limbah dengan Arsitektur MobileNet," 2023. [Online]. Available: <https://winco.cilacapkab.go.id>
- [5] M. F. Syahid, "IMPLEMENTASI DEEP LEARNING VGG16 DENGAN TRANSFER LEARNING PADA DETEKSI PENYAKIT TANAMAN SINGKONG SKRIPSI Oleh," 2021.
- [6] B. D. Laraswati, "Cara Menghindari Overfitting Data dalam Machine Learning," [www.algorit.ma](http://www.algorit.ma). Accessed: Nov. 09, 2024. [Online]. Available: <https://blog.algorit.ma/cara-menghindari-overfitting/>
- [7] W. Bismi, D. Novianti, and M. Qomaruddin, "Analisis Perbandingan Klasifikasi Citra Genus Panthera dengan Pendekatan Deep learning Model MobileNet," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 6, pp. 1–9, 2024.
- [8] A. G. Pradana, D. R. I. M. Setiadi, and A. R. Muslikh, "Fine tuning model Convolutional Neural Network EfficientNet-B4 dengan augmentasi data untuk klasifikasi penyakit kakao," *Journal of Information System and Application Development*, vol. 2, no. 1, pp. 01–11, Jun. 2024, doi: 10.26905/jisad.v2i1.11899.
- [9] D. Liandaputra and A. Zahra, "Klasifikasi Tingkat Kematangan Buah Nanas Menggunakan Metode Deep Learning," *G-Tech: Jurnal Teknologi Terapan*, vol. 8, no. 2, pp. 1091–1103, Apr. 2024, doi: 10.33379/gtech.v8i2.4122.
- [10] C. R. A. Widiawati, "PENGARUH DATASET TERHADAP PERFORMA CONVOLUTIONAL NEURAL NETWORK PADA KLASIFIKASI X-RAY PASIEN COVID-19," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 9, no. 6, pp. 1109–1118, 2022, doi: 10.25126/jtiik.202295645.
- [11] R. K. Hapsari and T. Indriyani, "SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika Implementasi Algoritma SMOTE Sebagai Penyelesaian Imbalance Hight Dimensional Datasets," *Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika*, pp. 427–432, 2022, doi: 10.31284/p.snestik.2022.2868.
- [12] R. Setya Nugraha and A. Hermawan, "OPTIMASI AKURASI METODE CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI KUALITAS BUAH APEL HIJAU," *Jurnal MNEMONIC*, vol. 6, no. 2, pp. 149–156, 2023.
- [13] A. E. Putra, K. Kartini, and A. P. Sari, "Metode Convolutional Neural Network dan Extreme Gradient Boost untuk Mengklasifikasi Penyakit Pneumonia," *JASIEK (Jurnal Aplikasi Sains, Informasi, Elektronika dan Komputer)*, vol. 6, no. 1, pp. 33–40, Jul. 2024, doi: 10.26905/jasiek.v6i1.11464.
- [14] M. R. B. Ulum, B. Rahmat, and H. P. S. Made, "Implementasi Metode CNN Dan K-Nearest Neighbor Untuk Klasifikasi Tingkat Kematangan Tanaman Cabai Rawit," *Jurnal Informatika dan Sains Teknologi*, vol. 1, no. 3, pp. 112–123, 2024, doi: 10.62951/modem.v1i3.131.