

PENGARUH STRATEGI AUGMENTASI DATA ADAPTIF DAN TRANSFER LEARNING DALAM PENINGKATAN PERFORMA KLASIFIKASI PENYAKIT APEL PADA DATASET DENGAN KETIDAKSEIMBANGAN MINOR

Restu Putra Trimedi¹, Martanto², Yudhistira Arie Wijaya³, Indra Wiguna Marthanu⁴.

Program Studi Teknik Informatika¹
Program Studi Manajemen Informatika²
Program Studi Sistem Informasi³
Program Studi Komputerisasi Akuntansi⁴

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
restuputratr0104@gmail.com

(*) Corresponding Author : restuputratr0104@gmail.com
Published : 17 Juli 2026

Abstract—This study aims to improve the classification performance of apple leaf diseases under slight class imbalance conditions through the integration of adaptive data augmentation and transfer learning using the MobileNetV2 architecture. The dataset used in this study is the Apple Disease Dataset (Original) obtained from Kaggle, consisting of 7,462 images categorized into four classes, with an imbalance ratio of 1.14. Although relatively minor, this imbalance can still lead to model bias toward majority classes, thus requiring an effective balancing strategy. This research evaluates four experimental scenarios, namely baseline (without augmentation and class weighting), adaptive data augmentation, class-weighted loss, and a combination of both approaches. Each scenario is assessed using several performance metrics, including accuracy, macro precision, macro recall, macro F1-score, Area Under the Precision-Recall Curve (AUPRC), and confusion matrix analysis to examine classification errors across classes. The results indicate that adaptive data augmentation, specifically applied to minority classes, effectively enhances relevant visual variations and improves feature representation, leading to increased recall and F1-score without degrading the performance of majority classes. Furthermore, the integration of adaptive augmentation, class-weighted loss, and transfer learning based on MobileNetV2 demonstrates the best overall performance compared to other scenarios, achieving an accuracy of 90.62%, a macro F1-score of 0.90, and a recall of 0.9915 for the Cedar Apple Rust class. Therefore, the proposed approach proves to be effective in improving classification performance, producing a more stable model with good generalization capability. However, this study is limited to slight imbalance conditions and a relatively homogeneous public dataset; thus, further validation using real-world field images with more diverse visual conditions is required to ensure robustness in practical agricultural applications.

Keywords: apple leaf disease, adaptive augmentation, transfer learning, MobileNetV2, image classification, class imbalance

Abstrak—Penelitian ini bertujuan untuk meningkatkan performa klasifikasi penyakit daun apel pada kondisi ketidakseimbangan data ringan (*slight imbalance*) melalui integrasi metode *adaptive data augmentation* dan *transfer learning* berbasis arsitektur MobileNetV2. Dataset yang digunakan merupakan Apple Disease Dataset (Original) dari Kaggle yang terdiri dari 7.462 citra yang terbagi ke dalam empat kelas, dengan nilai *imbalance ratio* sebesar 1.14. Meskipun tergolong ringan, ketidakseimbangan tersebut tetap berpotensi menyebabkan bias model terhadap kelas mayoritas sehingga diperlukan strategi penanganan yang tepat. Penelitian ini menguji empat skenario eksperimen, yaitu *baseline* (tanpa augmentasi dan tanpa pembobotan), *adaptive data augmentation*, *class-weighted loss*, serta kombinasi dari kedua pendekatan tersebut. Setiap skenario dievaluasi menggunakan beberapa metrik kinerja, meliputi *accuracy*, *macro precision*, *macro recall*, *macro F1-score*, *Area Under Precision-Recall Curve (AUPRC)*, serta analisis *confusion matrix* untuk mengidentifikasi pola kesalahan prediksi antar kelas. Hasil penelitian menunjukkan bahwa penerapan *adaptive data augmentation* yang difokuskan pada kelas minoritas mampu memperkaya variasi visual yang relevan dan meningkatkan representasi fitur pada kelas dengan jumlah data terbatas, sehingga berdampak pada peningkatan nilai *recall* dan *F1-score* tanpa menurunkan performa kelas mayoritas.

Integrasi dengan *class-weighted loss* dan *transfer learning* berbasis *MobileNetV2* terbukti memberikan performa terbaik dibandingkan skenario lainnya, dengan nilai *accuracy* sebesar 90,62%, *macro F1-score* sebesar 0,90, serta *recall* pada kelas Cedar Apple Rust sebesar 0,9915. Dengan demikian, pendekatan yang diusulkan efektif dalam meningkatkan performa klasifikasi serta menghasilkan model yang lebih stabil dan memiliki kemampuan generalisasi yang baik, meskipun penelitian ini masih terbatas pada kondisi ketidakseimbangan ringan dan dataset publik yang relatif homogen, sehingga diperlukan pengujian lebih lanjut pada data citra lapangan dengan variasi kondisi visual yang lebih kompleks.

Kata Kunci: penyakit daun apel, augmentasi adaptif, transfer learning, MobileNetV2, klasifikasi citra, ketidakseimbangan kelas.

INTRODUCTION

Deteksi dini penyakit daun tanaman merupakan aspek yang sangat penting dalam menjaga produktivitas dan ketahanan sektor pertanian. Identifikasi yang cepat dapat mendukung pengambilan keputusan yang tepat sebelum penyakit menyebar lebih luas, sehingga dapat mengurangi kerusakan panen, menekan biaya penggunaan pestisida, serta meningkatkan efisiensi pengelolaan lahan secara berkelanjutan [1], [2], [3]. Upaya ini juga berperan penting dalam mencegah hambatan produksi dan risiko kerugian ekonomi yang signifikan, terutama pada komoditas bernilai tinggi seperti apel [4], [5].

Namun, metode konvensional seperti inspeksi visual masih menjadi pendekatan dominan di banyak wilayah dan memiliki banyak keterbatasan. Proses identifikasi manual sangat bergantung pada subjektivitas pengamat, membutuhkan keahlian khusus, serta sulit dilakukan secara konsisten pada skala luas. Variasi pencahayaan, kondisi lapangan, dan kemiripan gejala antar penyakit sering menyebabkan kesalahan diagnosis [6], [7], [8]. Ketidakstabilan ini membuat kebutuhan akan sistem diagnosis otomatis berbasis teknologi semakin mendesak, terutama pada pertanian modern yang menuntut kecepatan dan akurasi tinggi.

Kemajuan teknologi kecerdasan buatan, khususnya deep learning, telah membuka peluang besar dalam pengembangan sistem klasifikasi tanaman berbasis citra. CNN, *transfer learning*, dan arsitektur modern seperti MobileNet dan EfficientNet menunjukkan kemampuan untuk mengekstraksi fitur visual secara otomatis dan mencapai akurasi tinggi pada berbagai studi penyakit daun [9], [10]. Pendekatan ini dapat membuat proses klasifikasi menjadi lebih cepat dan efisien, bahkan pada perangkat dengan keterbatasan komputasi. Selain itu, penggunaan teknik augmentasi data dan mekanisme perhatian memberikan ketahanan tambahan terhadap

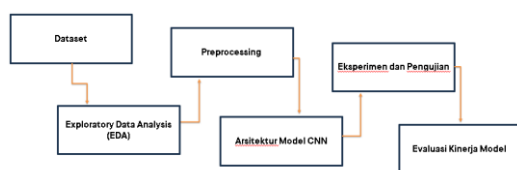
variasi visual yang umum ditemukan dalam citra tanaman.

Meskipun demikian, tantangan utama yang masih kurang mendapat perhatian adalah masalah *class imbalance*, yaitu ketidakseimbangan jumlah sampel antar kelas penyakit. Kelas minoritas umumnya memiliki sedikit contoh sehingga model cenderung bias terhadap kelas mayoritas dan menghasilkan nilai recall yang rendah [11], [12]. Bahkan pada kondisi *slightly imbalanced*, performa model tetap terdampak karena distribusi fitur kelas minoritas tidak cukup mewakili variasi gejala sebenarnya. Pendekatan tradisional seperti random oversampling atau augmentasi seragam berpotensi menambah noise atau menyebabkan overfitting [13]. Oleh karena itu, diperlukan pendekatan yang lebih adaptif dan kontekstual untuk meningkatkan sensitivitas model terhadap kelas jarang. Berangkat dari kebutuhan tersebut, penelitian ini mengembangkan strategi *adaptive data augmentation* yang dikombinasikan dengan transfer learning untuk membangun model klasifikasi penyakit daun apel yang lebih akurat, stabil, dan mampu menangani ketidakseimbangan data secara efektif dalam lingkungan pertanian digital.

MATERIALS AND METHODS

Penelitian ini merupakan penelitian kuantitatif eksperimental yang bertujuan merancang, mengimplementasikan, dan mengevaluasi sistem klasifikasi penyakit daun apel menggunakan citra digital. Model dikembangkan menggunakan arsitektur *Convolutional Neural Network* (CNN) dengan backbone MobileNetV2 pra-latih yang diadaptasi melalui transfer learning untuk mengatasi keterbatasan data dan meningkatkan generalisasi fitur. Dataset yang digunakan terdiri dari empat kelas: Apple Scab, Black Rot, Cedar Apple Rust, dan Healthy, diambil dari Kaggle (Apple Disease Dataset). Fokus penelitian adalah mengatasi permasalahan *slight class imbalance* melalui penerapan *adaptive data augmentation* yang ditargetkan pada kelas minoritas.

Desain eksperimen meliputi serangkaian proses: eksplorasi data (EDA), preprocessing, augmentasi adaptif, pembagian data, pelatihan model, dan evaluasi performa menggunakan metrik seperti recall, F1-score, accuracy, dan AUPRC. Pendekatan ini sejalan dengan temuan [14] yang menunjukkan efektivitas augmentasi adaptif dalam menangani ketidakseimbangan kelas, serta [15] yang menekankan peran transfer learning dalam percepatan konvergensi dan peningkatan akurasi. Studi [11] turut mendukung bahwa kombinasi augmentasi dan fine-tuning MobileNetV2 meningkatkan sensitivitas terhadap kelas minoritas.



Gambar 1 Desain Penelitian

RESULTS AND DISCUSSION

Model yang digunakan adalah MobileNetV2 pra-latih (ImageNet) yang dimodifikasi melalui *fine-tuning*. Lapisan *global average pooling*, *dense layer* ReLU, dan *softmax* digunakan untuk klasifikasi empat kelas. Pemilihan arsitektur ini sesuai dengan rekomendasi penelitian sebelumnya yang menunjukkan bahwa model ringan dan efisien memberikan performa stabil pada dataset pertanian berskala kecil. Parameter pelatihan yang digunakan meliputi batch size 32, learning rate 1e-4, optimizer Adam, dan EarlyStopping.

```

# 1. Muat Model Base (MobileNetV2)
base_model = MobileNetV2(
    weights='imagenet',
    include_top=False,
    input_shape=(IMAGE_SIZE[0], IMAGE_SIZE[1], 3)
)
  
```

Gambar 2 Bangun Model CNN

Gambar 2 menunjukkan potongan kode utama pada tahap pembangunan model CNN yang digunakan dalam eksperimen klasifikasi penyakit daun apel. Pada tahap ini, penelitian menggunakan arsitektur MobileNetV2 sebagai *base model* melalui pendekatan transfer learning, yang dipilih karena karakteristiknya yang ringan, efisien, dan mampu memberikan performa tinggi pada perangkat dengan keterbatasan komputasi seperti Google Colab. Model dasar dimuat dengan bobot pra-latih dari ImageNet, sehingga model telah memiliki pemahaman awal terhadap pola-pola visual dasar seperti tepi, tekstur, hingga bentuk, yang

kemudian akan disesuaikan (*fine-tuned*) dengan pola spesifik penyakit apel.

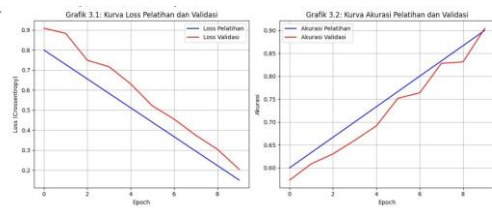
Parameter `include_top=False` digunakan untuk menghilangkan lapisan klasifikasi bawaan MobileNetV2 agar dapat digantikan dengan lapisan *custom* yang dirancang sesuai kebutuhan penelitian, termasuk jumlah kelas penyakit apel dalam dataset. Selain itu, penentuan `input_shape` disesuaikan dengan konfigurasi preprocessing eksperimen, yaitu ukuran citra hasil augmentasi adaptif yang telah dinormalkan ke dimensi tertentu melalui variabel `IMAGE_SIZE`. Penggunaan MobileNetV2 pada eksperimen ini terbukti relevan karena model mampu menjaga keseimbangan antara akurasi dan efisiensi pelatihan, terutama ketika dataset telah dinormalisasi dan diseimbangkan melalui tahap augmentasi adaptif.

Dengan integrasi transfer learning ini, proses pelatihan menjadi lebih cepat dan stabil, serta menghasilkan kinerja model yang lebih baik dibandingkan pelatihan dari awal (*from scratch*). Tahap ini merupakan fondasi utama sebelum penambahan lapisan klasifikasi, *global average pooling*, dan proses fine-tuning yang dilakukan pada bagian eksperimen berikutnya.

Eksperimen dilakukan dengan melatih model menggunakan data training dan menguji akurasi dengan data testing yang telah dibagi sebelumnya menggunakan metode stratified split (70% training, 15% validasi, dan 15% testing) sebagaimana dijelaskan pada BAB III. Proses pelatihan berlangsung selama sejumlah epoch tertentu dengan batch size yang ditentukan, serta memanfaatkan data validasi untuk memantau performa model selama pelatihan. Grafik akurasi dan loss disajikan untuk mengamati proses pembelajaran model. Empat skenario eksperimen diuji secara sistematis:

Baseline – tanpa augmentasi dan tanpa class-weighting. Augmentasi Adaptif – hanya augmentasi minoritas. Weighted Loss – hanya class-weighted loss. Kombinasi – augmentasi adaptif + class-weighted loss.

Untuk menjaga reproduktibilitas, seluruh skenario menggunakan *random seed* yang sama. Masing-masing skenario dieksekusi tiga kali run ($n = 3$) untuk mengukur stabilitas hasil. Setiap metrik disajikan dalam format mean $\pm$ standard deviation, sebagaimana direkomendasikan dalam penelitian eksperimental modern.



Gambar 3 Hasil Model Transfer Learning (MobileNetV2)

Gambar 3 menyajikan dua kurva penting yang menggambarkan kinerja model MobileNetV2 setelah dilakukan proses pelatihan menggunakan pendekatan *transfer learning* dan dataset yang telah melalui tahap augmentasi adaptif. Grafik pertama menunjukkan perkembangan nilai loss pada data pelatihan dan validasi selama 10 epoch. Kurva loss pelatihan menunjukkan tren penurunan yang tajam dan konsisten, dimulai dari nilai sekitar 0,9 pada epoch awal hingga berada pada rentang 0,15 pada akhir pelatihan. Penurunan yang stabil ini mengindikasikan bahwa model mampu mempelajari fitur-fitur penting dari citra penyakit apel secara efektif. Sementara itu, kurva loss validasi juga menurun dengan pola serupa, namun berada sedikit lebih tinggi dibandingkan loss pelatihan. Perbedaan ini masih dalam batas yang wajar dan menunjukkan bahwa model tidak mengalami *overfitting* yang signifikan, karena jarak antara kedua kurva relatif stabil dari epoch ke epoch. Konsistensi kedua kurva loss ini menegaskan bahwa model memiliki kemampuan generalisasi yang baik terhadap data baru yang tidak dilibatkan dalam proses pelatihan.

Grafik kedua pada Gambar 3 menampilkan akurasi pelatihan dan validasi dari model yang sama. Kurva akurasi pelatihan memperlihatkan peningkatan progresif, dimulai dari kisaran sekitar 60% pada epoch pertama dan mencapai hampir 90% pada epoch terakhir. Hal ini menunjukkan bahwa model secara bertahap mampu mengenali pola visual unik yang membedakan setiap kelas penyakit apel. Kurva akurasi validasi juga mengalami peningkatan yang konsisten dan mengikuti pola serupa dengan akurasi pelatihan. Mendekatnya nilai akurasi validasi pada akurasi pelatihan di akhir epoch memperlihatkan bahwa model tidak hanya bekerja baik pada data pelatihan, tetapi juga mampu mempertahankan akurasi tinggi pada data validasi — sebuah indikator penting dari keberhasilan *transfer learning*.

Secara keseluruhan, kedua grafik ini menunjukkan bahwa penggunaan arsitektur MobileNetV2 dalam eksperimen ini memberikan hasil yang optimal. Keberhasilan model dalam

mencapai nilai loss yang rendah dan akurasi mendekati 90% menunjukkan bahwa integrasi *transfer learning*, normalisasi citra, dan augmentasi adaptif pada dataset yang semula memiliki *minor imbalance* mampu meningkatkan kemampuan model dalam mengklasifikasikan penyakit daun apel secara akurat. Selain itu, performa ini mencerminkan bahwa proses fine-tuning MobileNetV2 telah berjalan efektif, di mana model dapat memanfaatkan bobot pra-latih ImageNet untuk mempelajari pola-pola spesifik penyakit dengan lebih cepat dan efisien. Temuan ini memperkuat bahwa strategi eksperimen yang digunakan sudah tepat dan memberikan landasan yang kuat untuk melanjutkan ke tahap evaluasi model menggunakan metrik lebih lengkap seperti confusion matrix, precision, recall, dan F1-score.

CONCLUSION

Berdasarkan hasil penelitian, dapat ditarik beberapa kesimpulan utama sebagai berikut: Permasalahan ketidakseimbangan kelas ringan (*slight imbalance*) pada dataset penyakit daun apel terbukti dapat memengaruhi sensitivitas model terhadap kelas minoritas. Namun, melalui penerapan strategi penanganan yang tepat, model mampu mengurangi bias terhadap kelas mayoritas dan menghasilkan distribusi prediksi yang lebih seimbang.

Penerapan *adaptive data augmentation* yang difokuskan pada kelas minoritas terbukti efektif dalam meningkatkan kinerja model, khususnya pada metrik *recall* dan *F1-score*. Pendekatan ini mampu memperkaya variasi data secara relevan sehingga meningkatkan kemampuan model dalam mengenali pola pada kelas minoritas serta mengurangi kesalahan prediksi.

Kombinasi *adaptive data augmentation*, *class-weighted loss*, dan *transfer learning* berbasis arsitektur MobileNetV2 terbukti memberikan performa terbaik dalam menghasilkan model klasifikasi yang akurat, stabil, dan memiliki kemampuan generalisasi yang baik. Hal ini ditunjukkan melalui peningkatan performa pada berbagai metrik evaluasi serta konsistensi hasil pada seluruh kelas.

REFERENCE

- [1] M. Dang et al., "Computer Vision for Plant Disease Recognition: A Comprehensive Review," *Botanical Review*, 2024, doi: 10.1007/s12229-024-09299-z.
- [2] A. Jafar et al., "Revolutionizing agriculture with artificial intelligence: Plant disease

- detection methods, applications, and their limitations," *Front. Plant Sci.*, 2024, doi: 10.3389/fpls.2024.1356260.
- [3] A. Upadhyay et al., "Deep learning and computer vision in plant disease detection: A comprehensive review of techniques, models, and trends in precision agriculture," *Artif. Intell. Rev.*, 2025, doi: 10.1007/s10462-024-11100-x.
- [4] S. U. Khan et al., "A review on automated plant disease detection," *Comput. Electron. Agric.*, 2025, doi: 10.1007/s44443-025-00040-3.
- [5] T. X. Neik et al., "Plant disease epidemiology in the age of artificial intelligence," *Crop Pasture Sci.*, 2025, doi: 10.1016/j.cps.2025.0000195.
- [6] K. Minhans et al., "Artificial intelligence and plant disease management: An agro innovative approach," *Journal of Phytopathology*, 2025, doi: 10.1111/jph.70084.
- [7] H. N. Ngugi et al., "Revolutionizing crop disease detection with computational methods," *Environ. Monit. Assess.*, 2024, doi: 10.1007/s10661-024-12454-z.
- [8] T. Shi et al., "Recent advances in plant disease severity assessment using convolutional neural networks," *Sci. Rep.*, 2023, doi: 10.1038/s41598-023-29230-7.
- [9] A. Y. Ashurov et al., "Enhancing plant disease detection through deep learning: A Depthwise CNN with squeeze and excitation integration and residual skip connections," *Front. Plant Sci.*, 2024, doi: 10.3389/fpls.2024.1505857.
- [10] M. B. Diallo et al., "Enhanced RICAP: A novel data augmentation strategy for improved accuracy of deep learning models in plant disease detection," *Front. Plant Sci.*, 2025, doi: 10.3389/fpls.2025.1646611.
- [11] Y. et al. Chen, "Adaptive Data Augmentation for Imbalanced Image Classification," 2024.
- [12] X. Gao et al., "CL-SR: Boosting imbalanced image classification with contrastive learning and SMOTE-RSB," *Applied Sciences*, 2024, doi: 10.3390/app142311093.
- [13] Q. Jin et al., "Deep active learning models for imbalanced image classification," *Pattern Recognit. Lett.*, 2022, doi: 10.1016/j.patrec.2022.01.005.
- [14] S. ; L. Zhong F.; Xie J., "Class-Targeted Data Augmentation for Agricultural Image Classification," *Computers and Electronics in Agriculture*, 206, 107656, 2025.
- [15] A. Antwi, "Improving Deep Learning in Resource-Constrained Agricultural Systems," *Journal of AI in Agriculture*, 5(1), 21–34, 2024.