

KLASIFIKASI BUNGA FINE-GRAINED MENGGUNAKAN EFFICIENTNET-B0 DENGAN STRATEGI AUGMENTASI DATA SINTETIS BERBASIS DIFFUSION MODELS

Hekal Anfalana¹, Martanto², Yudhistira Arie Wijaya³, Heliyanti Susana⁴, Puji Pramudya Marta⁵

Program Studi Teknik Informatika¹⁴
Program Studi Manajemen Informatika²
Program Studi Sistem Informasi³⁵
STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
Email : hekalanfalana99@gmail.com

(*) Corresponding Author : hekalanfalana99@gmail.com
Published : 30 Maret 2026

Abstract— *Fine-grained flower recognition represents a significant challenge in computer vision due to the high visual similarity between classes and the large intra-class variations caused by lighting, background, and perspective differences. This study proposes a flower image classification model using the EfficientNet-B0 architecture, designed to achieve high accuracy while maintaining computational efficiency. The research utilizes the 5 Flower Types dataset, consisting of five visually similar classes. The methodological pipeline includes image preprocessing, normalization, and basic augmentation, followed by transfer-learning-based training. Although diffusion-based synthetic augmentation was conceptually planned to enrich data diversity, it was not executed fully due to computational constraints. Experimental results show that EfficientNet-B0 achieved 95.87% accuracy and a 95.8% F1-score, demonstrating consistent performance across all classes. These findings confirm that EfficientNet-B0 is an effective backbone for fine-grained classification tasks and provide a strong foundation for future research involving diffusion-based synthetic augmentation to enhance visual diversity in botanical datasets.*

Keywords: *EfficientNet, flower classification, fine-grained recognition, data augmentation, diffusion models.*

Abstract—Perkembangan teknologi visi komputer memberikan peluang besar dalam pengembangan sistem identifikasi otomatis pada objek dengan kemiripan visual tinggi, termasuk bunga. Tantangan utama pada klasifikasi citra bunga terletak pada karakteristiknya yang bersifat *fine-grained*, di mana perbedaan antarkelas sangat tipis namun variasi intrakelas justru cukup besar. Penelitian ini bertujuan mengembangkan model klasifikasi bunga menggunakan arsitektur EfficientNet-B0 yang dikenal efisien dan memiliki kemampuan ekstraksi fitur yang kuat pada skala dataset menengah. Dataset 5 Flower Types digunakan sebagai data utama, yang terdiri dari lima kelas bunga dengan kesamaan struktur morfologi dan pola warna. Proses penelitian meliputi pengumpulan dataset, preprocessing citra, augmentasi dasar, dan pelatihan model menggunakan pendekatan transfer learning. Meskipun augmentasi sintesis berbasis diffusion models direncanakan sebagai bagian dari kerangka penelitian, implementasinya belum terealisasi sepenuhnya karena keterbatasan komputasi. Hasil penelitian menunjukkan bahwa EfficientNet-B0 mampu mencapai akurasi 95,87% dan nilai F1-score 95,8%, dengan performa stabil pada seluruh kelas. Temuan ini menegaskan bahwa arsitektur efisien seperti EfficientNet tetap mampu memberikan performa tinggi pada tugas klasifikasi fine-grained, sekaligus membuka peluang penelitian lanjutan dalam pemanfaatan citra sintesis generatif untuk meningkatkan keragaman data pelatihan.

Keywords: klasifikasi bunga, EfficientNet, fine-grained, augmentasi data, diffusion models.

INTRODUCTION

Perkembangan teknologi kecerdasan buatan (AI) dan *computer vision* dalam beberapa tahun terakhir telah mendorong peningkatan

signifikan dalam berbagai aplikasi berbasis citra, termasuk botani, agrikultur presisi, konservasi biodiversitas, hingga sistem edukasi berbasis visual. Identifikasi visual otomatis menjadi kebutuhan penting karena mampu meningkatkan efisiensi dan

akurasi dalam proses klasifikasi spesies flora. Penelitian terkini menunjukkan bahwa sistem klasifikasi berbasis deep learning berperan besar dalam dokumentasi biodiversitas, otomatis herbarium digital, serta pemetaan spesies tanaman pada skala besar (Campos Leal et al., 2024; Zhang, 2021). Dengan semakin kompleksnya data visual, dibutuhkan model klasifikasi yang mampu bekerja secara akurat dan efisien dalam kondisi pencahayaan, latar belakang, dan perspektif yang sangat bervariasi (Wang et al., 2022; Wei, 2021).

Klasifikasi citra bunga termasuk dalam kategori *fine-grained image classification*, yaitu tugas yang melibatkan objek dengan kemiripan visual yang sangat tinggi dan sering kali hanya dibedakan oleh fitur morfologis yang subtil. Perbedaan halus seperti struktur kelopak, pola warna, atau variasi tekstur sering kali sulit dibedakan baik oleh pengamat manusia maupun model deep learning. Tingginya variasi intrakelas yang dipicu oleh perubahan pencahayaan, orientasi kamera, dan latar belakang menambah tingkat kesulitan dalam proses ekstraksi fitur (Li et al., 2023; Wei, 2021). Tantangan ini menempatkan *fine-grained classification* sebagai salah satu domain visual yang membutuhkan representasi fitur sangat mendalam dan model neural yang sensitif terhadap detail visual mikro.

Sejalan dengan tantangan tersebut, berbagai arsitektur *convolutional neural networks* modern telah dikembangkan untuk meningkatkan akurasi dan efisiensi pemrosesan citra. EfficientNet menawarkan pendekatan *compound scaling* yang memungkinkan peningkatan kedalaman, lebar, dan resolusi jaringan dilakukan secara proporsional. Hasil studi menunjukkan bahwa EfficientNet memberikan performa unggul pada berbagai tugas klasifikasi visual dengan jumlah parameter yang lebih sedikit dibandingkan ResNet atau VGG (Campos Leal et al., 2024; Khan et al., 2022). Keunggulan ini menjadikan EfficientNet-B0 pilihan ideal untuk klasifikasi bunga yang membutuhkan sensitivitas tinggi terhadap detail morfologi mikro.

Selain arsitektur model, kualitas dan variasi dataset sangat memengaruhi performa klasifikasi. Banyak dataset flora publik memiliki ukuran terbatas dan tidak merepresentasikan kondisi dunia nyata, sehingga meningkatkan risiko overfitting. Teknik augmentasi konvensional seperti rotasi, flipping, dan *color jitter* membantu meningkatkan generalisasi, tetapi transformasi tersebut tidak menambah keragaman semantik baru (Wang et al., 2022). Kemajuan terbaru dalam *diffusion models* seperti DDPM dan Stable Diffusion memungkinkan generasi citra sintetis berkualitas tinggi yang mampu meningkatkan keragaman dataset secara signifikan. Studi terkini melaporkan

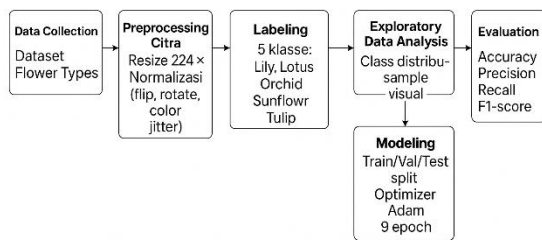
bahwa augmentasi berbasis diffusion dapat memperbaiki performa model dalam domain klasifikasi detail halus dengan menambah variasi fitur visual yang tidak dapat dihasilkan oleh augmentasi konvensional (Alimisis et al., 2025; Betzalel et al., 2024; Chen et al., 2024). Meskipun penelitian ini belum mengimplementasikan augmentasi generatif secara penuh, konsep tersebut menjadi arah pengembangan lanjutan yang memiliki potensi kuat.

Penelitian ini menggunakan dataset *5 Flower Types* yang terdiri dari lima kelas bunga dengan kemiripan morfologi yang sering saling tumpang tindih Lily, Lotus, Orchid, Sunflower, dan Tulip. Dengan kompleksitas visual yang tinggi, dataset ini menjadi benchmark yang relevan untuk mengevaluasi performa arsitektur modern seperti EfficientNet-B0. Tahap preprocessing dilakukan melalui *resize*, normalisasi, dan augmentasi ringan, diikuti pelatihan model menggunakan transfer learning berbasis bobot ImageNet untuk meningkatkan stabilitas ekstraksi fitur.

Hasil penelitian menunjukkan bahwa EfficientNet-B0 mampu mencapai akurasi 95.87% dan F1-score 95.8%, yang menegaskan kemampuan model dalam membedakan fitur morfologis halus antarspesies. Kesalahan prediksi sebagian besar terjadi pada kelas dengan kemiripan visual tinggi seperti Lily dan Lotus, sementara kelas lain menunjukkan stabilitas prediksi yang sangat baik. Temuan ini mengonfirmasi bahwa EfficientNet-B0 merupakan arsitektur yang efektif untuk klasifikasi bunga berbutir halus dan dapat menjadi landasan bagi penelitian selanjutnya dengan integrasi teknik augmentasi generatif berbasis diffusion untuk meningkatkan variasi visual dan memperkuat generalisasi model.

MATERIALS AND METHODS

Gambar 1 Tahapan Penelitian menyajikan alur lengkap proses penelitian klasifikasi citra bunga menggunakan arsitektur EfficientNet-B0, dimulai dari tahap pengumpulan data hingga evaluasi model. Diagram ini menggambarkan hubungan antarproses secara sistematis sehingga memudahkan pemahaman mengenai langkah-langkah utama dalam membangun model klasifikasi citra berbasis deep learning. Seluruh rangkaian tahapan mulai dari persiapan dataset, pengolahan awal citra, pelabelan, ekstraksi fitur, analisis eksploratif, pemodelan, hingga tahap evaluasi dijelaskan secara berurutan sebagaimana ditampilkan pada Gambar 3.1 Tahapan Penelitian.



Gambar 1 Tahapan Penelitian

Pengumpulan Data

Penelitian ini diawali dengan pengumpulan dataset bunga yang diperoleh dari dataset publik 5 *Flower Types* yang terdiri dari lima kelas bunga, yaitu Lily, Lotus, Orchid, Sunflower, dan Tulip. Tahap ini bertujuan menyediakan citra mentah yang merepresentasikan variasi visual bunga dari berbagai kondisi seperti sudut pengambilan gambar, pencahayaan, dan latar belakang. Sebelum digunakan, seluruh citra diperiksa untuk memastikan tidak terdapat file rusak, duplikat, atau gambar yang tidak sesuai dengan kelas. Dataset kemudian disusun dalam struktur folder per kelas untuk memudahkan proses pembacaan dan pelabelan otomatis selama pelatihan model. Persiapan data yang rapi memastikan bahwa proses pemodelan dapat dilakukan secara efisien dan akurat.

Preprocessing text

Setelah data berhasil dikumpulkan, tahap selanjutnya adalah preprocessing citra yang berfungsi untuk menyeragamkan karakteristik visual sebelum masuk ke proses pelatihan. Dalam tahap ini, setiap citra diubah ukurannya menjadi 224×224 piksel sesuai standar masukan EfficientNet-B0. Citra juga dinormalisasi menggunakan nilai *mean* dan *standard deviation* berdasarkan standar ImageNet agar distribusi pixel lebih stabil selama proses pembelajaran. Selain itu, dilakukan augmentasi dasar berupa rotasi, pembalikan horizontal, dan *color jitter* untuk menambah variasi visual pada data latih. Tahap ini sangat penting karena citra bunga sering menunjukkan perbedaan signifikan akibat kondisi lingkungan, sehingga augmentasi membantu model belajar menjadi lebih adaptif dan tidak bergantung pada pola tertentu yang hanya ditemukan pada data latih.

Labeling

Tahap labeling dilakukan untuk memberikan kelas yang benar pada setiap citra dalam dataset. Proses ini dilakukan secara otomatis berdasarkan struktur folder, di mana seluruh citra yang berada dalam folder tertentu akan diberi label

yang sama, yaitu Lily, Lotus, Orchid, Sunflower, atau Tulip. Pelabelan yang konsisten dan akurat sangat penting karena model deep learning bekerja berdasarkan pendekatan *supervised learning*, sehingga membutuhkan pasangan data dan label yang benar untuk dapat mempelajari pola visual secara efektif. Dengan penataan dataset yang terstruktur, model dapat mengidentifikasi ciri khas masing-masing bunga tanpa adanya ambiguitas kelas.

Feature Extraction text

Tahap selanjutnya adalah ekstraksi fitur (feature extraction), yaitu proses mengubah citra bunga menjadi representasi numerik yang dapat dipahami oleh jaringan saraf. Pada penelitian ini, digunakan arsitektur EfficientNet-B0 sebagai *backbone* untuk mengekstraksi fitur visual secara mendalam. EfficientNet-B0 memanfaatkan pendekatan *compound scaling* yang membuat jaringan lebih efisien dalam mempelajari pola dan tekstur bunga, mulai dari bentuk kelopak, warna dominan, hingga struktur tekstur halus. Lapisan dasar EfficientNet-B0 digunakan sebagai ekstraktor fitur utama, kemudian lapisan akhir dimodifikasi agar sesuai dengan jumlah kelas pada dataset. Dengan demikian, fitur yang dihasilkan mencerminkan karakteristik visual yang relevan untuk proses klasifikasi.

EDA

Sebelum melakukan pelatihan model, dilakukan tahap Exploratory Data Analysis (EDA) untuk memahami struktur dan distribusi dataset secara menyeluruh. Tahap ini mencakup analisis jumlah citra pada setiap kelas, visualisasi sampel citra dari tiap kelas, serta pemeriksaan variasi visual yang mungkin memengaruhi performa model. EDA memberikan gambaran awal mengenai apakah dataset seimbang atau tidak, serta membantu mengidentifikasi tantangan khusus seperti kemiripan morfologi antar kelas bunga. Dengan memahami karakteristik data sejak awal, peneliti dapat menentukan strategi augmentasi dan konfigurasi model yang lebih tepat untuk meningkatkan performa klasifikasi.

Pemodelan

Setelah dataset siap digunakan, proses pemodelan dilakukan menggunakan pendekatan *transfer learning* pada EfficientNet-B0. Lapisan awal jaringan dibekukan untuk mempertahankan representasi visual umum, sementara lapisan akhir dilatih ulang untuk menyesuaikan dengan lima kelas bunga. Dataset dibagi menjadi data pelatihan, validasi, dan uji dengan rasio 70:15:15. Model

dilatih menggunakan optimizer Adam selama sembilan epoch dengan pemantauan terhadap nilai loss dan akurasi. Pemodelan dilakukan secara iteratif untuk memastikan proses pembelajaran berjalan stabil tanpa indikasi *overfitting*. Tahap ini merupakan inti dari penelitian, di mana model mempelajari pola visual bunga secara bertahap.

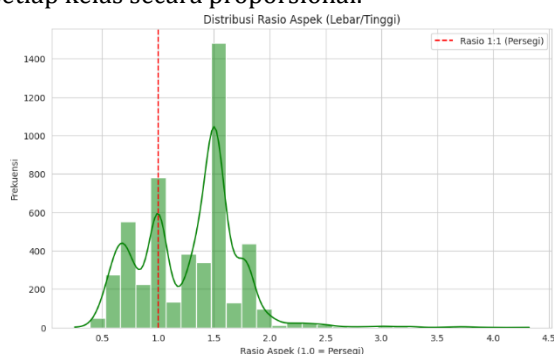
Evaluasi Model

Tahap terakhir dalam penelitian ini adalah evaluasi model untuk mengukur tingkat akurasi dan ketepatan prediksi EfficientNet-B0 dalam mengklasifikasikan citra bunga. Evaluasi dilakukan menggunakan data uji dengan metrik akurasi, presisi, recall, dan F1-score. Selain itu, confusion matrix digunakan untuk melihat pola kesalahan prediksi di antara kelas-kelas bunga, terutama pada kelas yang memiliki kemiripan visual tinggi seperti Lily dan Lotus. Evaluasi ini memberikan gambaran menyeluruh mengenai kemampuan model dalam mengenali pola visual yang halus dan membedakan kelas bunga secara akurat. Hasil evaluasi digunakan untuk menentukan efektivitas model dan potensi pengembangan lebih lanjut.

RESULTS AND DISCUSSION

Exploratory Data Analysis (EDA)

Tahap analisis eksploratif dilakukan untuk memahami karakteristik awal dataset sebelum proses pelatihan dilakukan. Dataset *5 Flower Types* terdiri atas lima kelas bunga yang memiliki jumlah citra relatif seimbang, sehingga tidak terdapat dominasi kelas tertentu yang berpotensi menimbulkan bias pada model. Kondisi ini memberikan fondasi yang baik bagi proses klasifikasi, karena model dapat mempelajari pola setiap kelas secara proporsional.



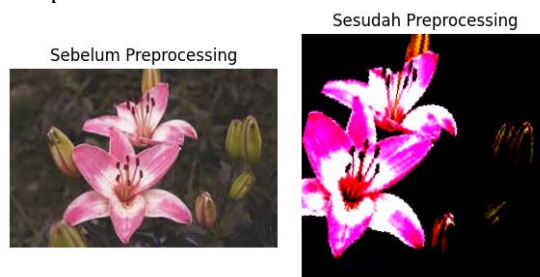
Gambar 2. Distribusi Kelas Dataset

Gambar 2 memperlihatkan jumlah citra pada masing-masing kelas, yaitu Lily, Lotus, Orchid, Sunflower, dan Tulip. Distribusi yang hampir seragam pada seluruh kelas mengindikasikan bahwa dataset berada dalam kondisi seimbang, sehingga kinerja model dapat dievaluasi secara

objektif. Dataset yang seimbang juga membantu mengurangi risiko ketidaktepatan prediksi pada kelas tertentu, khususnya pada kelas minor yang umumnya menjadi sumber ketidakseimbangan dalam klasifikasi citra.

Preprocessing Effects

Tahap preprocessing diterapkan untuk menyeragamkan karakteristik citra sebelum digunakan dalam proses pelatihan model EfficientNet-B0. Seluruh citra diubah ukurannya menjadi 224×224 piksel, dinormalisasi sesuai standar ImageNet, dan diberikan augmentasi ringan berupa rotasi, flipped horizontal, serta *color jitter*. Tujuan tahap ini adalah menghasilkan variasi visual yang cukup sehingga model mampu beradaptasi terhadap perubahan pencahayaan, orientasi, maupun tekstur.

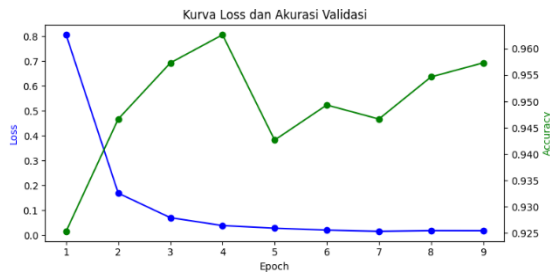


Gambar 3. Perbandingan Sebelum dan Sesudah Preprocessing

Gambar 3 menunjukkan perbedaan visual citra sebelum dan setelah preprocessing. Terlihat bahwa augmentasi menghasilkan variasi warna dan intensitas cahaya tanpa mengubah struktur morfologis utama, seperti bentuk kelopak dan pola tekstur. Hal ini menandakan bahwa preprocessing tidak hanya meningkatkan keragaman data, tetapi juga mempertahankan fitur penting yang diperlukan dalam proses klasifikasi, sehingga model dapat mengenali objek dengan lebih baik pada kondisi yang bervariasi.

Training Performance

Performa model selama proses pelatihan dievaluasi berdasarkan perubahan nilai *loss* dan akurasi validasi pada setiap epoch. Proses pelatihan berlangsung selama sembilan epoch dengan mekanisme early stopping yang menghentikan pelatihan ketika tidak terjadi peningkatan akurasi lebih lanjut. Pendekatan ini membantu mencegah model mengalami overfitting.



Gambar 4. Kurva Loss dan Akurasi Validasi per Epoch

Gambar 4 menunjukkan bahwa nilai *loss* menurun secara signifikan pada epoch awal dan kemudian stabil mendekati nol, yang mengindikasikan bahwa model mampu mempelajari pola visual dataset dengan efektif. Pada saat yang sama, akurasi validasi meningkat secara konsisten hingga mencapai nilai optimum sebesar 0.9573. Keselarasan antara penurunan *loss* dan peningkatan akurasi menggambarkan bahwa model tidak mengalami gejala overfitting. Aktivasi early stopping memastikan bahwa proses pelatihan berhenti pada titik performa terbaik tanpa berlatih secara berlebihan.

Classification Metrics

Evaluasi performa model dilakukan melalui classification report yang memuat nilai precision, recall, dan F1-score untuk masing-masing kelas. Hasil evaluasi menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik pada sebagian besar kelas.

	precision	recall	f1-score	support
Lilly	0.921986	0.915493	0.918728	142.000000
Lotus	0.930070	0.963768	0.946619	138.000000
Orchid	0.977273	0.968894	0.969814	179.000000
Sunflower	0.985507	0.971429	0.978417	140.000000
Tulip	0.973856	0.980263	0.977049	152.000000
accuracy	0.958722	0.958722	0.958722	0.958722
macro avg	0.957738	0.958369	0.957966	751.000000
weighted avg	0.958989	0.958722	0.958770	751.000000

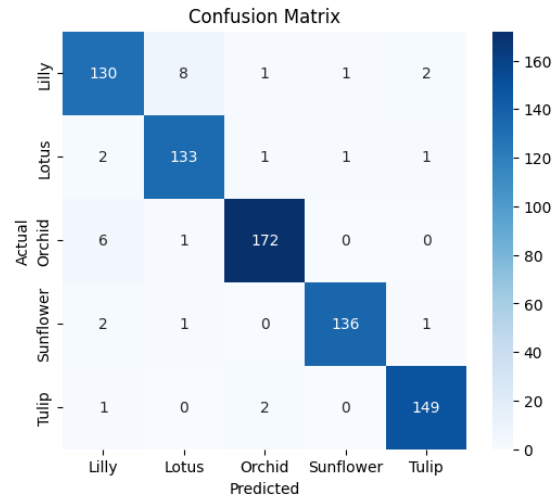
Gambar 5. Classification Report Model EfficientNet-B0

Gambar 5 menunjukkan bahwa kelas Orchid, Sunflower, dan Tulip memperoleh nilai F1-score yang tinggi, mendekati sempurna. Hal ini disebabkan oleh karakteristik visual yang lebih tegas pada ketiga kelas tersebut, sehingga lebih mudah dibedakan oleh model. Sebaliknya, nilai F1-score kelas Lily dan Lotus sedikit lebih rendah karena kesamaan bentuk kelopak dan warna dominan yang menyebabkan model lebih sulit membedakannya. Meski demikian, nilai weighted F1-score men

Confusion Matrix Analysis

Confusion matrix digunakan untuk mengidentifikasi pola kesalahan prediksi dan mengevaluasi sensitivitas model pada tiap kelas bunga. Visualisasi ini memberikan informasi penting mengenai kelas mana yang cenderung

tertukar dan sejauh mana model mampu melakukan prediksi secara tepat.



Gambar 6. Confusion Matrix Klasifikasi 5 Kelas Bunga

Gambar 6 memperlihatkan bahwa sebagian besar prediksi berada pada diagonal utama, mengindikasikan keberhasilan model dalam mengklasifikasikan citra secara benar. Kesalahan prediksi paling sering terjadi antara kelas Lilly dan Lotus, yang memiliki kemiripan morfologi sehingga berpotensi membingungkan model. Sebaliknya, prediksi untuk kelas Orchid, Sunflower, dan Tulip tampak sangat akurat, mencerminkan adanya fitur visual yang lebih khas pada ketiga kelas tersebut. Confusion matrix ini menunjukkan bahwa model memiliki sensitivitas tinggi terhadap perbedaan visual yang jelas, sementara perbedaan halus antar spesies tetap menjadi tantangan.

Overall Model Interpretation

Secara keseluruhan, model EfficientNet-B0 menunjukkan performa yang sangat kompetitif dalam tugas klasifikasi bunga berbutir halus. Nilai akurasi sebesar 95.87% dan F1-score sebesar 95.8% menggambarkan kemampuan model dalam menangkap pola visual mikro yang menjadi pembeda antar spesies bunga. Analisis dari Gambar 2 hingga Gambar 6 menegaskan bahwa preprocessing yang tepat, arsitektur model yang efisien, serta penerapan teknik transfer learning berperan penting dalam keberhasilan model. Tantangan utama terletak pada kemiripan visual antara kelas tertentu, sehingga penelitian lanjutan dapat mempertimbangkan penggunaan augmentasi generatif berbasis *diffusion models* atau teknik fine-tuning bertahap untuk meningkatkan sensitivitas model terhadap perbedaan visual yang sangat halus.

CONCLUSION

Berdasarkan keseluruhan rangkaian penelitian yang meliputi pengumpulan data citra bunga, preprocessing citra, analisis eksploratif, pelatihan model EfficientNet-B0, serta evaluasi performa melalui classification report dan confusion matrix, dapat disimpulkan bahwa model mampu memberikan performa klasifikasi yang sangat baik dengan akurasi 95.87% dan F1-score 95.8%. Dataset *5 Flower Types* yang seimbang memungkinkan proses pelatihan berlangsung tanpa bias kelas, namun kemiripan morfologi pada kelas tertentu, terutama Lily dan Lotus, tetap menjadi tantangan yang menyebabkan sebagian kesalahan prediksi. Meskipun demikian, kelas seperti Orchid, Sunflower, dan Tulip berhasil diklasifikasikan hampir sempurna berkat ciri visual yang jelas dan mudah dibedakan. Analisis confusion matrix juga memperlihatkan bahwa kesalahan utama terjadi pada kelas dengan kemiripan visual tinggi, sedangkan kelas lainnya menunjukkan stabilitas prediksi yang kuat. Temuan ini menunjukkan bahwa EfficientNet-B0 merupakan arsitektur yang efektif untuk tugas klasifikasi *fine-grained*, namun peningkatan performa masih dapat dicapai melalui pendekatan lanjutan seperti augmentasi sintesis berbasis *diffusion models*, fine-tuning bertahap, atau teknik interpretabilitas seperti Grad-CAM untuk meningkatkan sensitivitas model terhadap fitur visual yang sangat halus antar spesies bunga. Secara keseluruhan, penelitian ini membuktikan bahwa kombinasi preprocessing yang tepat, augmentasi sederhana, serta pemanfaatan transfer learning mampu menghasilkan model yang akurat dan konsisten, sekaligus memberikan landasan kuat bagi pengembangan penelitian selanjutnya.

REFERENCE

- Alimisis, P., Chatzipantazis, E., & Tefas, A. (2025). Advances in Diffusion Models for Image Data Augmentation: A Review. *Artificial Intelligence Review*. <https://doi.org/10.1007/s10462-024-10768-9>
- Betzalel, E., Penso, C., & Fetaya, E. (2024). Evaluation Metrics for Generative Models: An Empirical Study. *Machine Learning and Knowledge Extraction*, 6(3), 1531–1544. <https://doi.org/10.3390/make6030073>
- Campos Leal, J. A., Yee Rendón, A., López, I. F. V., & Arellano, J. R. L. (2024). Analysis of the EfficientNet architectures for flower classification. *Applied Sciences*, 14(1). <https://doi.org/10.3390/app14010370>
- Chen, D., Qi, X., Zheng, Y., Lu, Y., Huang, Y., & Li, Z. (2024). Synthetic data augmentation by diffusion probabilistic models to enhance weed recognition. *Computers and Electronics in Agriculture*, 216, 108486. <https://doi.org/10.1016/j.compag.2024.108486>
- Khan, S., Naseer, M., & Hayat, M. (2022). Transformers and CNNs in Visual Recognition: A Comparative Review. *IEEE Access*, 10, 110111–110131. <https://doi.org/10.1109/ACCESS.2022.3189640>
- Li, W., Duan, X., & Zhang, H. (2023). Challenges and Advances in Fine-Grained Visual Recognition: A Survey. *Pattern Recognition Letters*, 168, 68–75. <https://doi.org/10.1016/j.patrec.2022.12.011>
- Wang, Y., Chen, Y., & Li, J. (2022). Data Augmentation Strategies for Fine-Grained Classification: A Comprehensive Survey. *Pattern Recognition*, 132, 108983. <https://doi.org/10.1016/j.patcog.2022.108983>
- Wei, X.-S. (2021). Fine-grained visual classification: A review. *Information Fusion*, 79, 29–52. <https://doi.org/10.1016/j.inffus.2021.09.018>
- Zhang, M. (2021). Classification of flower images based on attention mechanism. *Image and Vision Computing*, 112, 104202. <https://doi.org/10.1016/j.imavis.2021.104202>