

KLASIFIKASI PENYAKIT DAUN PADI MENGGUNAKAN TRANSFER LEARNING DENGAN ANALISIS PENGARUH VARIASI DIMENSI CITRA PADA KINERJA MODEL

Akhmad Taukhid¹, Martanto², Yudhistira Arie Wijaya³, Heliyanti Susana⁴, Nana Suarna⁵

^{1,2,3,4,5} STMIK IKMI Cirebon

¹akhmadtaukhid10@gmail.com, ²martantomusijo@gmail.com, ³yudhistira010471@gmail.com, ⁴heliyanti3105@gmail.com, ⁵st_nana@yahoo.com

ABSTRACT

This study focuses on the early detection of rice leaf diseases to increase agricultural productivity and reduce misdiagnosis that often occurs in manual identification [1]. Although various studies have applied deep learning to plant disease classification, the effect of image resolution on the performance of rice leaf disease classification models, especially in limited data scenarios, has rarely been systematically studied. This study aims to analyze the performance of a transfer learning-based rice leaf disease classification model with a VGG16 architecture on images with a resolution of 224×224 pixels, while also assessing the computational efficiency of the training and testing processes. The data used consisted of 320 rice leaf images from the public dataset “Daun Padi Sultra (Southeast Sulawesi)” on Kaggle, which was divided into training, validation, and test data with a ratio of 60:20:20. The main stages of the research included exploring the characteristics and distribution of the data, pre-processing the images (resizing to 224×224, normalization, and limited augmentation), and building a transfer learning model with VGG16 as a frozen feature extractor and a custom classification head. The model was trained using the Adam optimizer with Early Stopping and ModelCheckpoint mechanisms, then evaluated using accuracy, precision, recall, F1-score, and confusion matrix. The test results show that the model achieved a test accuracy of 98.44% with a loss of 0.1815, as well as macro average and weighted average values for precision, recall, and F1-score that were close to 0.98 with only one classification error in the test data. The training and testing processes can be completed with moderate computational load on the Google Colab GPU environment, so that the VGG16 configuration with a resolution of 224×224 pixels has the potential to be an effective and efficient baseline for rice leaf disease classification in limited data scenarios.

Keywords: Transfer Learning, VGG16, rice leaf disease, image resolution.

I. PENDAHULUAN

Padi merupakan komoditas pangan utama bagi sebagian besar penduduk di kawasan Asia, termasuk Indonesia, sehingga kestabilan produksi padi menjadi salah satu faktor kunci ketahanan pangan nasional. Berbagai gangguan biotik seperti hama dan penyakit daun dapat menurunkan produktivitas secara signifikan apabila tidak terdeteksi dan ditangani sejak dini [1]. Berbagai studi menunjukkan bahwa penyakit daun padi seperti blast dan blight, termasuk di antara penyebab penurunan hasil panen yang paling sering dijumpai di lahan budidaya padi [1].

Diagnosis penyakit pada tanaman padi di lapangan umumnya masih mengandalkan pengamatan visual oleh penyuluh atau petani yang berpengalaman. Pendekatan ini bersifat subjektif, membutuhkan waktu, dan tidak selalu tersedia merata di seluruh wilayah produksi. Perbedaan kondisi pencahayaan, sudut pengambilan gambar, latar belakang, dan tingkat keparahan gejala pada daun juga dapat menyebabkan ketidakselarasan diagnosis antar pengamat. Hal ini mendorong kebutuhan akan sistem diagnosis penyakit daun padi yang lebih objektif, konsisten, dan dapat dioperasikan secara skalabel di lapangan.

Perkembangan computer vision dan deep learning membuka peluang besar untuk otomatisasi proses identifikasi penyakit tanaman berbasis citra. Berbagai

penelitian telah menunjukkan bahwa arsitektur *convolutional neural network* (CNN) mampu mencapai akurasi tinggi dalam klasifikasi penyakit daun padi menggunakan citra RGB maupun citra hiperspektral. Deng et al. [1] mengembangkan sistem diagnosis otomatis penyakit padi berbasis deep learning dan melaporkan kinerja klasifikasi yang tinggi pada beberapa jenis penyakit daun. Feng et al. [2] mengombinasikan citra hiperspektral dengan deep transfer learning untuk mendeteksi penyakit padi dengan sensitivitas yang baik. Seelwal et al. [3] melalui kajian sistematis menegaskan bahwa pendekatan deep learning dan transfer learning telah menjadi tren utama dalam otomasi diagnosis penyakit padi dalam beberapa tahun terakhir.

Selain pendekatan berbasis CNN konvensional, sejumlah penelitian terbaru mengusulkan model yang semakin beragam, mulai dari skema transfer learning untuk klasifikasi citra daun [4], rancangan fitur bertingkat secara hand-crafted [5], arsitektur jaringan ringan yang menyeimbangkan akurasi dan kecepatan [6], hingga model deteksi berbasis transformer dan varian YOLO yang dioptimalkan untuk perangkat bergerak [7], [8]. Ragam pendekatan ini menunjukkan bahwa penelitian deteksi penyakit daun padi telah berkembang pesat baik dari sisi arsitektur model maupun skenario implementasi di lapangan.



Meskipun demikian, masih terdapat beberapa tantangan yang relevan bagi penerapan sistem diagnosis berbasis deep learning pada konteks dataset yang terbatas. Pertama, banyak penelitian berfokus pada pengembangan arsitektur model yang semakin kompleks, tetapi belum secara eksplisit menganalisis dampak konfigurasi teknis seperti resolusi masukan citra terhadap performa dan efisiensi komputasi, khususnya pada dataset berukuran kecil. Seelwal et al. [3] mencatat bahwa masih dibutuhkan studi yang membahas trade-off antara akurasi, kompleksitas model, dan kebutuhan komputasi untuk memberikan panduan praktis bagi implementasi di perangkat dengan sumber daya terbatas. Kedua, sebagian penelitian melakukan pelatihan pada dataset besar atau citra berkualitas tinggi, sehingga belum sepenuhnya mewakili skenario praktis di mana data terbatas dan kualitas citra bervariasi.

Dalam konteks tersebut, transfer learning dengan memanfaatkan model pralatih seperti VGG16 menjadi salah satu pendekatan yang menarik untuk dieksplorasi pada dataset citra penyakit daun padi berukuran terbatas. Dengan membekukan sebagian besar bobot model pralatih dan hanya melatih lapisan klasifikasi di bagian akhir, proses pelatihan dapat dilakukan lebih efisien tanpa memerlukan jumlah data yang sangat besar. Pada saat yang sama, pemilihan resolusi masukan citra, misalnya 224×224 piksel, berperan penting sebagai kompromi antara kelengkapan informasi visual yang dipertahankan dan biaya komputasi yang harus dikeluarkan selama proses pelatihan dan inferensi.

Berdasarkan kondisi tersebut, penelitian ini berfokus pada klasifikasi penyakit daun padi menggunakan pendekatan transfer learning dengan arsitektur VGG16 dan resolusi citra 224×224 piksel pada dataset citra daun padi berukuran terbatas. Penelitian ini tidak hanya bertujuan untuk memperoleh model klasifikasi dengan performa yang baik, tetapi juga untuk mengevaluasi kelayakan konfigurasi tersebut sebagai baseline yang efisien bagi pengembangan sistem diagnosis penyakit daun padi pada skala yang lebih luas. Secara khusus, penelitian ini diharapkan memberikan kontribusi berupa: (1) analisis kinerja model transfer learning berbasis VGG16 pada dataset penyakit daun padi dengan jumlah sampel yang terbatas; (2) pemahaman yang lebih jelas mengenai implikasi penggunaan resolusi citra 224×224 piksel terhadap akurasi dan stabilitas model; dan (3) rekomendasi konfigurasi awal yang dapat dijadikan acuan bagi penelitian lanjutan maupun implementasi praktis di lingkungan komputasi dengan sumber daya terbatas.

A. Tujuan khusus penelitian

1. Menganalisis karakteristik dan distribusi data citra daun padi pada dataset publik sebelum dilakukan proses pra-pemrosesan.
2. Mengevaluasi kinerja model transfer learning berbasis VGG16 dengan input 224×224 piksel dalam mengklasifikasikan empat kelas penyakit daun padi berdasarkan metrik akurasi, precision, recall, dan F1-score.
3. Menilai efisiensi proses pelatihan dan validasi model berdasarkan metrik performa dan waktu komputasi.

B. Kontribusi utama penelitian

1. Menyajikan implementasi terstruktur transfer learning berbasis VGG16 untuk klasifikasi penyakit daun padi dengan dataset publik yang seimbang.
2. Memberikan analisis empiris mengenai kelayakan penggunaan dimensi 224×224 piksel sebagai kompromi antara kinerja klasifikasi dan efisiensi komputasi pada kasus klasifikasi penyakit daun padi.

II. KAJIAN TERKAIT

A. Deep Learning dan Transfer

Learning untuk Penyakit Tanaman

CNN merupakan arsitektur utama dalam klasifikasi citra dan banyak digunakan di bidang pertanian, kedokteran, dan industri [9]. Arsitektur modern seperti VGG, ResNet, dan Inception memanfaatkan lapisan konvolusi dalam untuk mengekstrak fitur dari tekstur sederhana hingga pola kompleks.

Dalam konteks penyakit daun padi, Deng et al. [1] menggunakan model deep learning untuk diagnosis otomatis penyakit padi dan menunjukkan akurasi yang lebih tinggi dibandingkan dengan diagnosis manual. Feng et al. [2] menggabungkan citra hyperspectral dengan pembelajaran transfer mendalam untuk mendeteksi penyakit daun padi pada berbagai varietas. Studi lain menggunakan arsitektur seperti InceptionV3, EfficientNet, dan MobileNet yang disesuaikan pada dataset daun padi dan secara konsisten melaporkan akurasi tinggi dan F1-score [3], [10], [11].

Pendekatan federated transfer learning memungkinkan pelatihan model klasifikasi penyakit daun padi di berbagai silo data tanpa perlu mengonsolidasikan data, sehingga meningkatkan privasi dan keragaman sumber data [12]. Selain itu, beberapa studi telah mengembangkan platform pengenalan fenotipe penyakit padi berbasis transfer learning dan mikroservis untuk memfasilitasi integrasi ke dalam aplikasi praktis [13].

B. Pengaruh Resolusi Citra terhadap Kinerja Model

Beberapa studi menunjukkan bahwa resolusi citra secara langsung memengaruhi kemampuan model deep learning untuk mempertahankan detail diagnostik. Meningkatkan resolusi hingga batas tertentu dapat meningkatkan akurasi dan F1-score, tetapi setelah melampaui ambang batas tertentu, manfaat tambahan cenderung kecil sementara beban komputasi meningkat secara signifikan [14], [15]. Rajaraman et al. [16] menunjukkan bahwa konfigurasi resolusi memengaruhi kinerja segmentasi lesi pada citra rontgen dada.

Dalam kasus penyakit daun padi, Zhou et al. [17] mengevaluasi berbagai konfigurasi dataset gambar untuk akurasi dan waktu pelatihan CNN, dan melaporkan bahwa ada kombinasi ukuran citra dan konfigurasi data yang menghasilkan keseimbangan terbaik antara akurasi dan efisiensi. Secara umum, dimensi citra 224×224 piksel sering digunakan karena kompatibel dengan banyak arsitektur yang telah dilatih sebelumnya seperti VGG16 dan telah terbukti memberikan kompromi yang baik antara kemampuan representasi fitur dan persyaratan komputasi [9], [18].

Temuan ini menjadi dasar penting dalam pemilihan dimensi citra dalam studi ini, di mana resolusi 224×224 piksel dipilih untuk memanfaatkan bobot VGG16 yang



telah dilatih sebelumnya sambil mempertahankan efisiensi komputasi.

III. METODOLOGI PENELITIAN

A. Desain Penelitian

Penelitian ini menggunakan pendekatan eksperimental kuantitatif berbasis citra digital untuk mengevaluasi efektivitas transfer learning dengan arsitektur VGG16 dalam mengklasifikasikan penyakit daun padi. Data diperoleh dari dataset publik Kaggle yang berisi empat kelas, yaitu blast, blight, normal, dan tungro. Dataset tersebut dibagi menjadi data pelatihan, validasi, dan pengujian dengan rasio 60%:20%:20% menggunakan pembagian acak tanpa kebocoran data.

Kinerja model dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score yang dihasilkan dari confusion matrix untuk memberikan gambaran seimbang tentang kemampuan model dalam membedakan setiap kelas [19], [20].

B. Dataset dan Eksplorasi Data

Dataset yang digunakan adalah <https://www.kaggle.com/datasets/adesulastri/daun-padi-sultra-sulawesi-tenggara> berasal dari platform Kaggle, yang berisi 320 citra RGB daun padi. Citra tersebut dibagi secara merata ke dalam empat kelas: blast, blight, normal, dan tungro, masing-masing dengan 80 citra. Dataset publik serupa sering digunakan dalam penelitian klasifikasi penyakit daun padi karena menyediakan label yang terstruktur dan representatif [10], [11].

Eksplorasi data (EDA) dilakukan menggunakan Python di Google Colab dengan perpustakaan seperti Pandas, Matplotlib, Seaborn, dan Pillow. Analisis distribusi kelas menunjukkan bahwa dataset tersebut seimbang, sehingga risiko bias kelas mayoritas relatif kecil. Analisis dimensi citra menunjukkan bahwa ukuran citra asli bervariasi secara signifikan, sehingga memerlukan standarisasi sebelum pelatihan model.

Visualisasi citra sampel dari setiap kelas digunakan untuk memverifikasi kesesuaian label dan meninjau variasi pencahayaan dan latar belakang. Hasil eksplorasi menunjukkan bahwa data memiliki kualitas yang cukup baik dengan variasi visual yang realistis untuk kasus lapangan.

C. Pra-pemrosesan Citra

Pra-pemrosesan bertujuan untuk menstandarkan citra agar sesuai dengan arsitektur VGG16 dan meningkatkan kemampuan generalisasi model. Langkah-langkah utama yang dilakukan adalah:

1. Penyesuaian Ukuran: Semua citra disesuaikan ukurannya menjadi 224×224 piksel agar kompatibel dengan lapisan input VGG16 dan memanfaatkan bobot pra-latih ImageNet tanpa mengubah strukturnya.
2. Pembagian dataset: dataset dibagi menjadi 192 gambar pelatihan, 64 gambar validasi, dan 64 citra uji dengan pembagian acak yang menjaga kesetaraan jumlah gambar per kelas.
3. Normalisasi: Nilai piksel dinormalisasi ke rentang $[0, 1]$ dan diproses menggunakan fungsi `preprocess_input` Keras untuk sesuai dengan skema normalisasi VGG16.
4. Label Encoding: Label kelas dikonversi dari

format teks ke representasi numerik.

5. Augmentasi Data: Rotasi, translasi horizontal dan vertikal, zoom, shear, dan flip horizontal diterapkan pada data pelatihan untuk meningkatkan variasi dan mengurangi risiko overfitting [9].

D. Pembangunan dan Pelatihan Model

Model ini dibangun menggunakan VGG16 yang telah dilatih sebelumnya sebagai model dasar dengan parameter `include_top = False` dan bentuk input $224 \times 224 \times 3$. Semua lapisan konvolusi dibekukan untuk mempertahankan representasi fitur yang umum, sementara bagian klasifikasi asli dari VGG16 diganti dengan kepala klasifikasi baru yang terdiri dari lapisan flatten, lapisan padat dengan 512 neuron dan aktivasi ReLU, dropout 0.5, dan lapisan padat keluaran dengan 4 neuron dan aktivasi softmax.

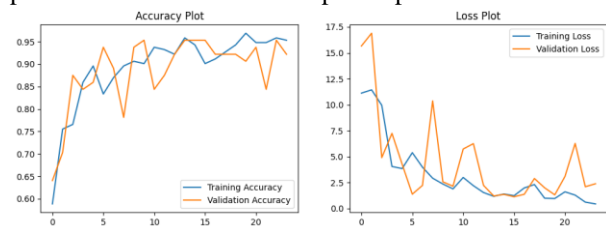
Konfigurasi pelatihan mencakup optimizer Adam dengan laju pembelajaran $1e-4$, fungsi kerugian cross-entropy kategorikal, ukuran batch 32, dan maksimum 50 epoch. Dua callback digunakan, yaitu Early Stopping (memantau `val_loss` dengan `patience = 8` dan `restore_best_weights = True`) dan ModelCheckpoint (memantau `val_accuracy` dan `save_best_only = True`) untuk menyimpan model dengan kinerja validasi terbaik.

Pelatihan dilakukan menggunakan generator data berbasis ImageDataGenerator yang mengintegrasikan normalisasi dan augmentasi. Augmentasi hanya diterapkan pada data pelatihan, sementara data validasi dan pengujian hanya dinormalisasi untuk membuat proses evaluasi lebih objektif.

IV. HASIL DAN PEMBAHASAN

A. Hasil Pelatihan dan Validasi

Sebagai dasar analisis kinerja model selama proses pelatihan, perubahan nilai akurasi dan loss pada data pelatihan serta validasi ditampilkan pada Gambar 1.



Gambar 1. Hasil Pelatihan Model

Berdasarkan Gambar 1, terlihat bahwa akurasi pelatihan meningkat secara bertahap dan mencapai nilai mendekati 1.00, sementara akurasi validasi stabil pada rentang tinggi tanpa fluktuasi ekstrem. Pola ini menunjukkan bahwa strategi pembekuan lapisan konvolusi VGG16, penggunaan augmentasi data, dan early stopping berhasil mencegah model dari overfitting, meskipun terdapat selisih kecil antara akurasi pelatihan dan validasi, yang umum terjadi pada dataset kecil.

B. Kinerja Pada Data Uji

Untuk memberikan gambaran kuantitatif mengenai kinerja model pada data uji, ringkasan nilai metrik evaluasi disajikan pada Tabel 1.



Tabel 1. Hasil Akurasi

Metrik	Nilai
Test Loss	0.1815
Test Akurasi	98.44%

Berdasarkan Tabel 1, dapat dilihat bahwa data uji memperoleh test loss sebesar 0.1815 dan akurasi uji sebesar 98.44%. Nilai test loss yang relatif rendah ini menunjukkan bahwa kesalahan prediksi model pada data yang tidak pernah dilihat selama pelatihan berada pada tingkat yang rendah, sementara akurasi mendekati 100% menunjukkan bahwa sebagian besar citra uji diklasifikasikan dengan benar. Secara keseluruhan, kombinasi antara loss rendah dan akurasi tinggi menunjukkan bahwa model mampu generalisasi dengan baik pada data uji dan memiliki kinerja prediksi yang sangat memuaskan. Selain akurasi keseluruhan, evaluasi kinerja model juga dilakukan berdasarkan metrik per kelas, yaitu precision, recall, dan F1-score. Penggunaan metrik ini dimaksudkan untuk memberikan gambaran yang lebih mendalam mengenai kemampuan model dalam mengenali setiap kelas penyakit daun padi secara seimbang. Hasil evaluasi tersebut disajikan pada tabel 2 berikut.

Tabel 2. Laporan Klasifikasi

Kelas	Precision	Recall	F1-Score	Support
Blast	1.00	0.94	0.97	16
Blight	1.00	1.00	1.00	16
Normal	1.00	1.00	1.00	16
Tungro	0.94	1.00	0.97	16
Akurasi			0.98	64
Macro avg	0.99	0.98	0.98	64
Weighted avg	0.99	0.98	0.98	64

Tabel 2 menampilkan laporan klasifikasi model pada data uji berdasarkan metrik presisi, recall, F1-score, dan support. Dapat dilihat bahwa untuk kelas blight dan normal, nilai presisi, recall, dan F1-score adalah 1.00 dengan support 16 citra, sehingga semua citra dalam kedua kelas tersebut diprediksi dengan benar. Untuk kelas blast, presisi 1.00 dan recall 0.94 dengan F1-score 0.97 tercapai, menunjukkan bahwa semua gambar yang diprediksi sebagai blast benar, tetapi masih ada sejumlah kecil gambar blast yang salah diklasifikasikan sebagai kelas lain. Sebaliknya, pada kelas tungro, presisi sebesar 0.94 dan recall sebesar 1.00 dengan nilai F1-score 0.97, artinya semua gambar tungro berhasil dikenali, tetapi terdapat beberapa citra dari kelas lain yang salah diprediksi sebagai tungro. Secara keseluruhan, akurasi pada data uji mencapai 0.98, sementara nilai rata-rata makro dan rata-rata tertimbang untuk presisi, recall, dan F1-score berada dalam rentang 0.98–0.99, menunjukkan bahwa kinerja model relatif tinggi dan seimbang di semua kelas dengan tingkat kesalahan prediksi yang sangat terbatas. Untuk menggambarkan secara lebih rinci distribusi prediksi benar dan salah pada setiap kelas, distribusi hasil klasifikasi kemudian divisualisasikan

dalam matriks kebingungan yang ditampilkan pada Gambar 2.



Gambar 2. Confusion Matrix

Interpretasi confusion matrix pada Gambar 2 menunjukkan bahwa sebagian besar citra dalam keempat kelas diklasifikasikan dengan benar, seperti yang terlihat dari nilai dominan pada elemen diagonal matriks. Citra dalam kelas blight, normal, dan tungro semuanya diprediksi dengan benar, sementara dalam kelas blast, satu gambar diprediksi secara salah sebagai milik kelas lain. Kesalahan klasifikasi yang terbatas pada sejumlah kecil gambar menunjukkan bahwa pola gejala visual utama dalam setiap kelas telah dipelajari dengan baik oleh model. Secara keseluruhan, distribusi nilai dalam confusion matrix menunjukkan bahwa kinerja model relatif seimbang di seluruh kelas dan tidak menunjukkan bias yang kuat terhadap kelas tertentu.

C. Analisis Dimensi Citra dan Efisiensi Komputasi

Pemilihan resolusi 224×224 piksel didasarkan pada dua pertimbangan utama: (1) kompatibilitas dengan arsitektur VGG16 yang telah dilatih sebelumnya, dan (2) kebutuhan untuk mempertahankan efisiensi komputasi tanpa mengorbankan detail tekstur penting pada citra daun padi. Hasil eksperimen menunjukkan bahwa pada dimensi ini, model mampu mencapai akurasi sangat tinggi dan macro F1-score tanpa memerlukan sumber daya komputasi yang besar. Waktu pelatihan per epoch berada dalam rentang yang wajar untuk lingkungan komputasi berbasis GPU di Google Colab.

Dibandingkan dengan studi yang menggunakan arsitektur lebih kompleks atau resolusi gambar lebih tinggi, konfigurasi VGG16 dengan input 224×224 piksel dalam studi ini menawarkan kombinasi menarik antara kesederhanaan implementasi, akurasi kompetitif, dan efisiensi. Kesuksesan model pada dataset relatif kecil juga menunjukkan bahwa strategi transfer learning dengan pembekuan lapisan konvolusi penuh cukup untuk domain yang karakteristik visualnya cukup dekat dengan ImageNet.

D. Kontribusi Penelitian

Penelitian ini berkontribusi pada (i) implementasi terstruktur transfer learning berbasis VGG16 dengan kinerja tinggi pada dataset publik yang seimbang, (ii) pembuktian bahwa input 224×224 memberikan



kompromi yang baik antara akurasi dan biaya komputasi pada kasus yang dikaji, serta (iii) praktik evaluasi komprehensif melalui metrik per kelas, rata-rata makro/tertimbang, dan confusion matrix sebagai acuan studi klasifikasi citra pertanian.

V. KESIMPULAN

Berdasarkan hasil penelitian, beberapa kesimpulan utama dapat dirumuskan sebagai berikut:

1. Kinerja model

Model VGG16 dengan basis konvolusi dibekukan dan kepala klasifikasi kustom mampu mengenali empat kelas daun padi dengan akurasi, precision, recall, dan F1-score tinggi pada tingkat kelas maupun rata-rata; kesalahan terbatas terutama terjadi antara kelas blast dan tungro.

2. Efektivitas konfigurasi & pelatihan

Pra-pemrosesan (resizing 224×224, normalisasi), pembagian data, dan augmentasi pada data latih, dikombinasikan dengan strategi fully frozen VGG16 serta Early Stopping dan ModelCheckpoint, menghasilkan pelatihan yang stabil dan mengurangi overfitting, sehingga transfer learning tetap kompetitif pada skenario data terbatas.

3. Evaluasi komprehensif

Penggunaan metrik yang beragam termasuk macro/weighted average dan confusion matrix menunjukkan performa yang tinggi sekaligus seimbang lintas kelas, sehingga memberikan contoh pipeline evaluasi yang lebih lengkap untuk klasifikasi citra penyakit tanaman.

4. Akurasi vs efisiensi

Kombinasi transfer learning, resolusi citra moderat, dan strategi pelatihan terkontrol menghasilkan model yang efisien secara komputasi tanpa mengorbankan kinerja, serta layak dikembangkan menuju prototipe aplikasi pendukung keputusan di perlindungan tanaman.

Secara keseluruhan, kombinasi antara dataset publik yang terstruktur, pra-pemrosesan yang tepat, dan penggunaan model VGG16 yang telah dilatih sebelumnya terbukti efektif untuk pengembangan sistem deteksi penyakit daun padi berbasis citra digital.

Beberapa saran untuk pengembangan penelitian selanjutnya adalah:

1. Pengembangan & perluasan data

Dataset perlu diperbesar dan diperkaya (varietas, kondisi lapangan, pencahayaan/latar), termasuk penambahan kelas penyakit dan tingkat keparahan, serta pengambilan citra langsung di lahan untuk menguji generalisasi.

2. Eksplorasi arsitektur & kompresi

Perlu dieksplorasi arsitektur lain (ResNet, EfficientNet, MobileNet) dengan skema fine-tuning berbeda, disertai teknik kompresi (pruning, quantization, distillation) guna menyeimbangkan akurasi, ukuran model, dan kebutuhan komputasi pada perangkat terbatas.

3. Penguatan desain eksperimen & evaluasi

Disarankan penerapan k-fold cross-validation/repeated holdout, penambahan metrik/analisis kesalahan terperinci, serta uji signifikansi antarmodel agar estimasi performa lebih robust dan argumen empiris lebih kuat.

4. Prototipe & XAI

Perlu dikembangkan prototipe (web/mobile) yang dilengkapi explainable AI (mis. Grad-CAM) untuk meningkatkan keterpahaman dan kepercayaan pengguna akhir, sehingga hasil penelitian lebih mudah dialihkan ke penggunaan nyata.

REFERENSI

- [1] R. Deng *et al.*, "Automatic diagnosis of rice diseases using deep learning," *Front. Plant Sci.*, vol. 12, p. 701038, 2021, doi: 10.3389/fpls.2021.701038.
- [2] L. Feng, B. Wu, Y. He, and C. Zhang, "Hyperspectral imaging combined with deep transfer learning for rice disease detection," *Front. Plant Sci.*, vol. 12, p. 693521, 2021, doi: 10.3389/fpls.2021.693521.
- [3] P. Seelwal, P. Dhiman, Y. Gulzar, A. Kaur, S. Wadhwa, and C. W. Onn, "A systematic review of deep learning applications for rice disease diagnosis: Current trends and future directions," *Front. Comput. Sci.*, vol. 6, p. 1452961, 2024, doi: 10.3389/fcomp.2024.1452961.
- [4] C. G. Simhadri and H. K. Kondaveeti, "Automatic recognition of rice leaf diseases using transfer learning," *Agronomy*, vol. 13, no. 4, p. 961, 2023, doi: 10.3390/agronomy13040961.
- [5] Y. M. Alsakar, N. A. Sakr, and M. Elmogy, "An enhanced classification system of various rice plant diseases based on multi-level handcrafted feature extraction technique," *Sci. Rep.*, vol. 14, p. 30601, 2024, doi: 10.1038/s41598-024-81143-1.
- [6] H. Ning, S. Liu, Q. Zhu, and T. Zhou, "Convolutional neural network in rice disease recognition: Accuracy, speed and lightweight," *Front. Plant Sci.*, vol. 14, p. 1269371, 2023, doi: 10.3389/fpls.2023.1269371.
- [7] Z. Li *et al.*, "Rice disease detection: TLI-YOLO innovative approach for enhanced detection and mobile compatibility," *Sensors*, vol. 25, no. 8, p. 2494, 2025, doi: 10.3390/s25082494.
- [8] H. Yang, X. Deng, H. Shen, Q. Lei, S. Zhang, and N. Liu, "Disease detection and identification of rice leaf based on improved detection transformer," *Agriculture*, vol. 13, no. 7, p. 1361, 2023, doi: 10.3390/agriculture13071361.
- [9] L. Alzubaidi *et al.*, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, p. 53, 2021, doi: 10.1186/s40537-021-00444-8.
- [10] M. Gogoi, V. Kumar, S. A. Begum, N. Sharma, and S. Kant, "Classification and detection of rice diseases using a 3-stage CNN architecture with transfer learning approach," *Agriculture*, vol. 13, no. 8, p. 1505, 2023, doi: 10.3390/agriculture13081505.
- [11] J. Li *et al.*, "An interpretable high-accuracy method for rice disease detection based on multisource data and transfer learning," *Plants*, vol. 12, no. 18, p. 3273, 2023, doi: 10.3390/plants12183273.
- [12] M. Aggarwal *et al.*, "Federated transfer learning for rice-leaf disease classification across multiclient cross-silo datasets," *Agronomy*, vol. 13, no. 10, p. 2483, 2023, doi: 10.3390/agronomy13102483.
- [13] P. Yuan, Y. Xia, Y. Tian, and H. Xu, "TRiP: A transfer learning-based rice disease phenotype recognition platform using SENet and microservices," *Front. Plant Sci.*, vol. 15, p. 1255015, 2024, doi: 10.3389/fpls.2023.1255015.
- [14] M. I. U. Haque, A. K. Dubey, I. Danciu, A. C. Justice, O. S. Ovchinnikova, and J. D. Hinkle, "Effect of image resolution on automated classification of chest X-rays," *J. Med. Imaging*, vol. 10, no. 4, p. 44503, 2023, doi: 10.1117/1.JMI.10.4.044503.
- [15] O. Rukundo, "Effects of image size on deep learning," *Electronics*, vol. 12, no. 4, p. 985, 2023, doi: 10.3390/electronics12040985.
- [16] S. Rajaraman, F. Yang, G. Zamzmi, Z. Xue, and S. Antani, "Assessing the impact of image resolution on deep learning for TB lesion segmentation on frontal chest X-rays," *Diagnostics*, vol. 13, no. 4, p. 747, 2023, doi: 10.3390/diagnostics13040747.
- [17] H. Zhou, J. Deng, D. Cai, X. Lv, and B. M. Wu, "Effects of image dataset configuration on the accuracy of rice disease recognition based on convolution neural



- network,” *Front. Plant Sci.*, vol. 13, p. 910878, 2022, doi: 10.3389/fpls.2022.910878.
- [18] M. Dutta *et al.*, “Rice leaf disease classification—A comparative approach using CNN, CAAR-U-Net, and MobileNet-V2 architectures,” *Technologies*, vol. 12, no. 11, p. 214, 2024, doi: 10.3390/technologies12110214.
- [19] G. M. Foody, “Challenges in the real-world use of classification accuracy metrics: From recall and precision to the Matthews correlation coefficient,” *PLoS One*, vol. 18, no. 10, p. e0291908, 2023, doi: 10.1371/journal.pone.0291908.
- [20] O. Rainio, J. Teuho, and R. Klén, “Evaluation metrics and statistical tests for machine learning,” *Sci. Rep.*, vol. 14, p. 6086, 2024, doi: 10.1038/s41598-024-56706-x.

