

PENERAPAN RESNET50 DAN SWIN TRANSFORMER PADA IDENTIFIKASI CITRA PENYAKIT DAUN KELAPA SAWIT

Siti Marwa¹⁾, Muhammad Faisal²⁾, Muhyiddin AM Hayat³⁾, Nurnawaty⁴⁾, Andi Makbul Syamsuri⁵⁾, M. Aguslim⁶⁾

1. Teknik, Universitas Muhammadiyah Makassar
email: 105841100122@student.unismuh.ac.id
2. Teknik, Universitas Muhammadiyah Makassar
email: muhfaisal@unismuh.ac.id
3. Teknik, Universitas Muhammadiyah Makassar
email: muhyiddinamhayat@gmail.com
4. Teknik Pengairan, Universitas Muhammadiyah Makassar
email: nurnawaty@unismuh.ac.id
5. Teknik Pengairan, Universitas Muhammadiyah Makassar
email: amakbulsyamsuri@unismuh.ac.id
6. Teknik Pengairan, Universitas Muhammadiyah Makassar
email: m.aguslim@unismuh.ac.id

Abstract

This study aims to implement and compare the performance of ResNet50 and Swin Transformer models in classifying palm oil leaf diseases. The decline in palm oil productivity is often caused by disease infections such as Curvularia (leaf spot) and Leaf Rust, necessitating a fast and precise automated identification system. This experimental computational research used a primary dataset of 600 digital images proportionally divided into training, validation, and testing sets. The preprocessing stage included resolution adjustment (resizing), data augmentation to prevent overfitting, and normalization. Model performance evaluation was conducted quantitatively through Confusion Matrix calculations and validated qualitatively through heatmap visualization using the Gradient-weighted Class Activation Mapping (Grad-CAM) method. The test results proved that the ResNet50 architecture outperformed the Swin Transformer with an accuracy of 98.00%, precision of 98.01%, recall of 98.00%, and F1-score of 98.00%, compared to the Swin Transformer's accuracy of 96.00%. Grad-CAM analysis also confirmed that ResNet50 is sharper in specifically localizing local infection areas. Overall, it is concluded that the ResNet50 model is more optimal, stable, and recommended for the palm oil leaf disease classification system in this dataset domain.

Kata Kunci : Deep Learning, Image Classification, Palm Oil Leaf Disease, ResNet50, Swin Transformer

A. PENDAHULUAN

Kelapa sawit (*Elaeis guineensis* Jacq) merupakan komoditas perkebunan utama

yang memiliki nilai ekonomis tinggi untuk memenuhi kebutuhan minyak nabati dunia (Yabani et al., 2023). Pada tahun 2021, minyak sawit berkontribusi sebanyak 36%

terhadap total produksi minyak nabati global, menggeser pangsa konsumsi komoditas lain seperti kedelai, rapeseed, dan bunga matahari karena menghasilkan minyak yang lebih jernih dan berkualitas (Murphy, 2025). Indonesia sendiri menjadi produsen terbesar dengan menguasai lebih dari 55% dari total produksi perkebunan dunia tersebut (Badan Pusat Statistik, 2020). Kendati demikian, realisasi produktivitas *Crude Palm Oil* (CPO) di Indonesia saat ini baru mencapai 3.732 Kg/Ha, yang menunjukkan angka 49,76% lebih rendah dibandingkan potensi maksimalnya sebesar 7,5 ton/Ha (Badan Pusat Statistik, 2020). salah satu faktor utama yang menghambat pencapaian target produksi optimal tersebut adalah serangan penyakit pada area fotosintesis (daun), seperti bercak daun (*Curvularia*) dan karat daun (*Leaf Rust*) (Priwiratama et al., 2024). Infeksi pada daun ini mengganggu kemampuan tanaman dalam mengonversi energi cahaya menjadi nutrisi, sehingga berdampak langsung pada penurunan berat tandan buah segar (TBS) dan penurunan kualitas minyak yang dihasilkan.

Hingga saat ini, sebagian besar perkebunan kelapa sawit masih mengandalkan inspeksi manual secara konvensional oleh pakar atau buruh tani yang memiliki keterbatasan dalam hal efisiensi. Sebagai solusinya, perkembangan teknologi pengolahan citra dan *Computer Vision* berbasis *Deep Learning* hadir menawarkan metode otomatisasi identifikasi penyakit secara presisi, objektif, dan *real-time* (Riati, Yuhandri, and Nurcahyo 2024). Dalam perkembangannya, pendekatan klasifikasi penyakit tanaman berbasis citra digital masih didominasi oleh arsitektur *Convolutional Neural Network* (CNN). Salah satu arsitektur CNN yang telah mencapai tahap kematangan teknik adalah ResNet50. Melalui inovasi residual

learning (*skip connections*), ResNet50 berhasil mengatasi kendala *vanishing gradient* pada jaringan saraf yang dalam (Joseph Alberto, 2023). Dalam tugas klasifikasi penyakit sawit, arsitektur ini terbukti sangat stabil dan konsisten untuk mengekstraksi fitur visual lokal, seperti tepi bercak daun dan tekstur permukaan daun yang terinfeksi. Namun demikian, filter konvensional pada CNN memiliki kelemahan inheren, yaitu keterbatasan dalam menangkap hubungan spasial jarak jauh (*global context*), terutama ketika dihadapkan pada karakteristik visual lingkungan perkebunan yang kompleks dengan intensitas pencahayaan tidak seragam.

Untuk mengatasi batasan tersebut, *Swin Transformer* hadir sebagai paradigma baru dalam visi komputer dengan mengadopsi mekanisme *Self-Attention* dari dunia pemrosesan bahasa alami atau *Natural Language Processing* (NLP). Berbeda dengan CNN yang memproses citra secara lokal piksel demi piksel di area terbatas, *Swin Transformer* memanfaatkan mekanisme *shifted windows* untuk menangkap konektivitas antar-jendela dan konteks global spasial jarak jauh (Styorini et al., 2022). Kemampuan ini sangat krusial dalam mendeteksi penyakit daun kelapa sawit, di mana pola sebaran bercak atau degradasi warna daun seringkali memiliki keterkaitan spasial yang luas. Meskipun arsitektur ini menawarkan performa unggul, sebagian besar penerapannya sejauh ini masih terbatas pada dataset umum atau tanaman selain kelapa sawit. Karakteristik penyakit daun kelapa sawit yang memiliki pola bercak tidak beraturan dan *noise* visual yang tinggi akibat kondisi lapangan membuat domain ini memerlukan kajian komparatif yang lebih mendalam.

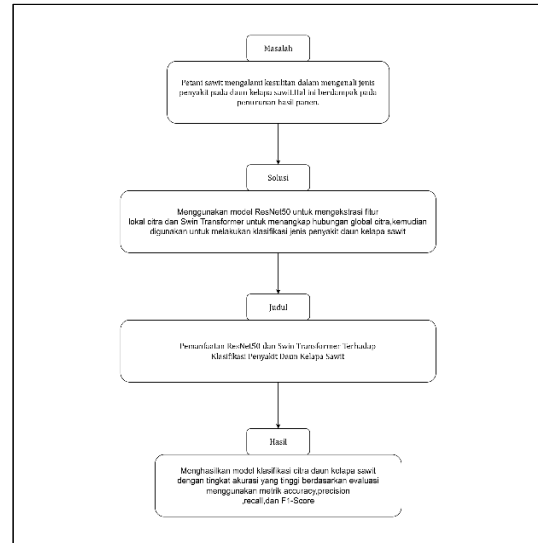
Berdasarkan kesenjangan penelitian (*research gap*) tersebut, penelitian ini diusulkan untuk mengimplementasikan

serta membandingkan kinerja arsitektur ResNet50 (sebagai representasi CNN untuk fitur lokal) dan *Swin Transformer* (sebagai representasi *Vision Transformer* untuk fitur global) dalam mengklasifikasikan jenis penyakit daun kelapa sawit. Melalui evaluasi kuantitatif menggunakan *metrik accuracy, precision, recall*, dan *F1-score*, penelitian ini diharapkan mampu menentukan model terbaik yang paling andal dan efisien, sekaligus memberikan kontribusi nyata bagi pengembangan pertanian presisi di sektor perkebunan kelapa sawit.

B. METODE PENELITIAN

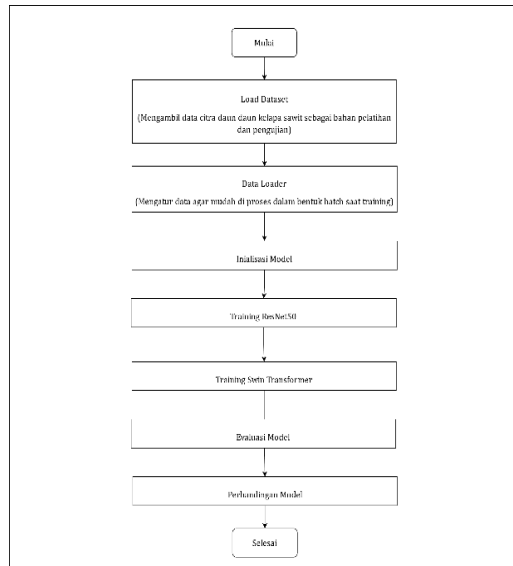
Penelitian ini menerapkan pendekatan kuantitatif dengan metode eksperimental berbasis komputasi *deep learning* untuk melakukan klasifikasi citra digital secara otomatis. Pemanfaatan teknologi ini dirancang secara terstruktur untuk menyelesaikan permasalahan konkret di lapangan, di mana pekebun kelapa sawit sering kali menghadapi kendala teknis serta memiliki pengetahuan terbatas dalam mengidentifikasi gejala infeksi secara dini dan akurat (Aditya, 2024). Hubungan konseptual antara perumusan masalah, landasan teori komparasi arsitektur, hingga target capaian akurasi tinggi dalam sistem pertanian presisi ini digambarkan secara sistematis dalam kerangka pikir penelitian berikut.

Secara teknis, pemodelan dibangun dengan mengomparasikan dua arsitektur *state-of-the-art* yang memiliki karakteristik ekstraksi fitur berbeda. Arsitektur pertama adalah ResNet50 yang merepresentasikan *Convolutional Neural Network* (CNN) matang dengan keunggulan koneksi sisa (*residual connections*) untuk mengekstrak fitur visual lokal seperti tepi bercak dan tekstur infeksi secara stabil (Joseph Alberto, 2023).



Gambar 1. Kerangka Pikir Penelitian Klasifikasi Penyakit Daun Kelapa Sawit.

Arsitektur kedua adalah *Swin Transformer* yang merepresentasikan paradigma baru berbasis *hierarchical self-attention* dengan mekanisme *shifted windows* untuk menangkap konteks spasial global jarak jauh pada daun kelapa sawit yang memiliki karakteristik tekstur tinggi dan latar belakang kompleks (Ali et al., 2025), (Mustaqim et al., 2025). Seluruh tahapan eksperimen, mulai dari pemuatan data hingga komparasi performa model akhir, dijalankan mengikuti diagram alir sistem di bawah ini.



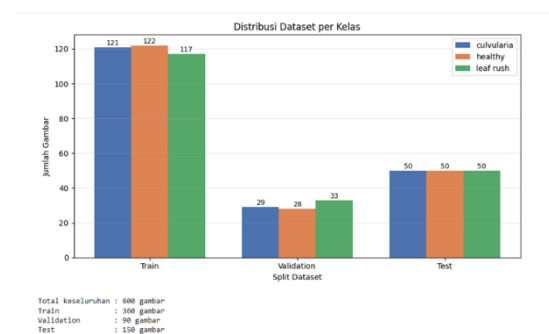
Gambar 2. Flowchart

Populasi dalam studi ini mencakup seluruh tanaman kelapa sawit yang berada di Desa Tirongan Atas, Kecamatan Bungku Utara, Kabupaten Morowali Utara, Sulawesi Tengah. Dari populasi tersebut, diambil sampel data primer berupa 600 citra digital daun kelapa sawit tunggal (single-image) yang diakuisisi menggunakan kamera gawai pada kondisi pencahayaan siang hari guna memastikan kualitas fitur visual. Proses pelabelan dataset ini sangat krusial dalam metode supervised learning, sehingga data disegmentasi ke dalam tiga kelas utama sebagaimana dijabarkan pada Tabel 1.

Tabel 1. Segmentasi Kelas Dataset Citra Daun Kelapa Sawit

Nilai Label	Kelas	Keterangan
0	<i>Curvularia</i>	Daun kelapa sawit yang terindikasi penyakit bercak daun akibat jamur <i>Curvularia</i>
1	<i>Healthy</i>	Daun kelapa sawit dalam kondisi sehat tanpa gejala penyakit
2	<i>Leaf Rust</i>	Daun kelapa sawit yang terkena Karat Daun

Untuk memastikan objektivitas pengujian, seluruh dataset diisolasi secara proporsional, yaitu data pelatihan (*training*) sebanyak 360 citra, data validasi (*validation*) sebanyak 90 citra, dan data pengujian (*testing*) sebanyak 150 citra dengan komposisi seimbang antar-kelas.



Gambar 3. Grafik Distribusi Dataset Citra Daun Kelapa Sawit per Kelas.

Sebelum dimasukkan ke dalam jaringan saraf, seluruh citra melalui tahap pra-pemrosesan (*preprocessing*) menggunakan pustaka *Torchvision*.

Langkah awal dilakukan dengan menyamakan resolusi spasial gambar menjadi 224×224 piksel guna memenuhi spesifikasi input bawaan model *pretrained* yang digunakan (Lee et al., 2023). Guna memperkaya variasi data latih sekaligus menekan risiko *overfitting*, diterapkan teknik augmentasi data geometris dan fotometrik acak yang meliputi pembalikan horizontal (*Random Horizontal Flip*), pembalikan vertikal (*Random Vertical Flip*), rotasi hingga ± 20 derajat, serta manipulasi kecerahan melalui *ColorJitter*. Citra yang telah teraugmentasi kemudian dikonversi ke dalam representasi tensor dan dinormalisasi menggunakan parameter statistik ImageNet, yaitu *mean* [0.485, 0.456, 0.406] dan standar deviasi [0.229, 0.224, 0.225] demi mempercepat konvergensi parameter selama proses pelatihan (Hastari et al., 2024).

Evaluasi kinerja klasifikasi diukur secara ketat melalui instrumen *Confusion Matrix* yang memetakan relasi antara prediksi model dan nilai observasi aktual ke dalam empat parameter fundamental: *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) (Rachman et al. 2025). Dari parameter matriks tersebut, analisis data kuantitatif dihitung berdasarkan empat formulasi metrik evaluasi matematis berikut:

Akurasi (Accuracy)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Presisi (Precision)

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F1-Score

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (4)$$

Selanjutnya, analisis kualitatif diperkuat dengan mengintegrasikan metode *Gradient-weighted Class Activation Mapping* (Grad-CAM) untuk menghasilkan visualisasi peta panas (*heatmap*) pada area piksel yang menjadi fokus utama keputusan model.

C. HASIL DAN PEMBAHASAN

Hasil

Pengujian eksperimental menggunakan citra daun kelapa sawit menghasilkan data kuantitatif dari 150 data uji dengan komposisi seimbang sebanyak 50 sampel untuk masing-masing kelas kondisi. Hasil evaluasi metrik performa berupa tingkat akurasi, presisi, *recall*, dan F1-score untuk arsitektur ResNet50 dan Swin Transformer ditabulasikan secara terperinci pada Tabel 2.

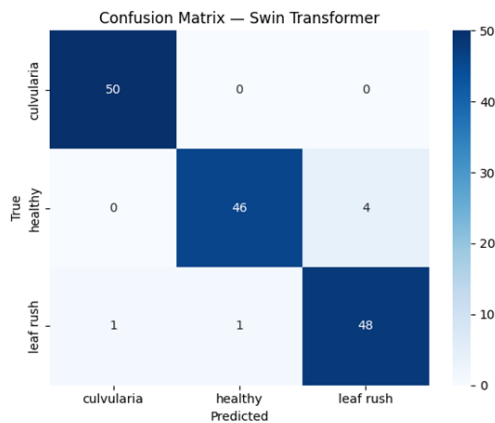
Tabel 2. Perbandingan Performa Model Klasifikasi

Model	Accuracy	Precision	Recall	F1-Score
ResNet50	98.00%	98.01%	98.00%	98.00%
Swin Transformer	96.00%	96.07%	96.00%	95.99%

Berdasarkan data pengujian kuantitatif tersebut, tingkat ketepatan prediksi dan distribusi penalaran model dipaparkan secara spasial melalui visualisasi *Confusion Matrix* yang tercantum pada Gambar 4 dan Gambar 5.



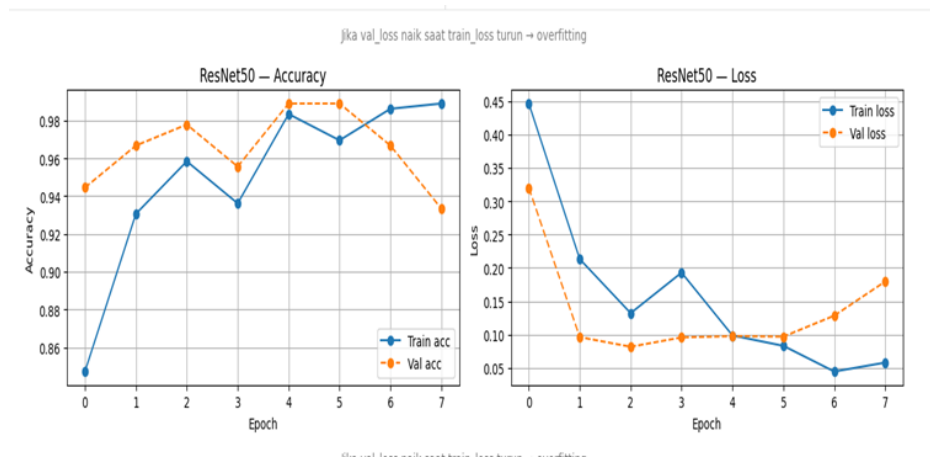
Gambar 4. Confusion Matrix ResNet50



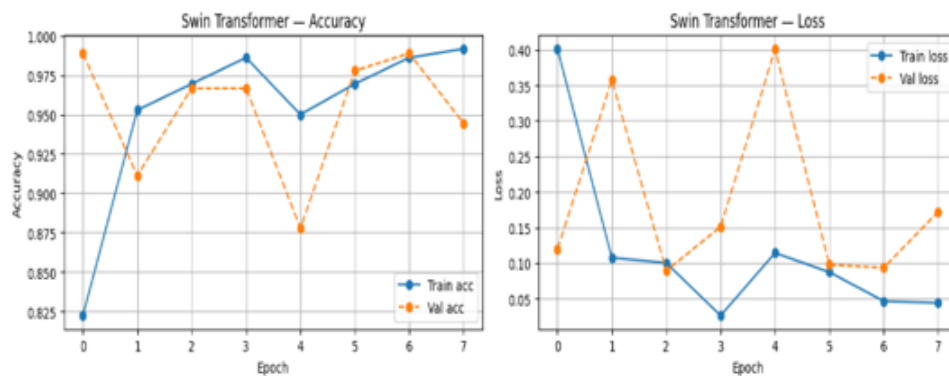
Gambar 5. Confusion Matrix Model Swin Transformer

Melalui visualisasi matriks kesalahan tersebut, tercatat bahwa dari total 150 data uji, model ResNet50 menghasilkan 147 prediksi benar dengan 3 kesalahan klasifikasi, sedangkan *Swin Transformer* menghasilkan 144 prediksi benar dengan 6 kesalahan klasifikasi. Kedua model mampu mengenali seluruh sampel pada kelas penyakit *Curvularia* secara sempurna tanpa galat (50 dari 50 citra). Namun, penyimpangan prediksi ditemukan pada kedua arsitektur berupa pencampuran klasifikasi antara kelas *Healthy* (sehat) dan *Leaf Rust* (karat daun).

Selanjutnya, karakteristik stabilitas proses optimasi parameter dari kedua arsitektur selama tahapan iterasi (*epoch*) terekam pada grafik kurva pembelajaran (*learning curves*) yang disajikan pada Gambar 6 and Gambar 7.

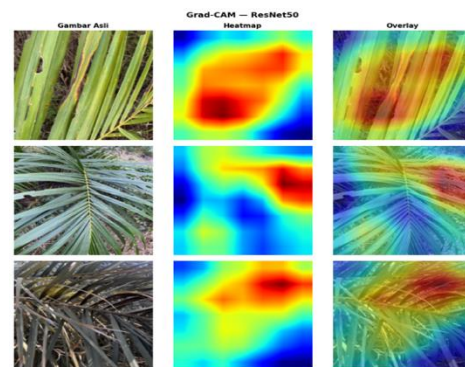


Gambar 6. Kurva Pembelajaran (Learning Curve) Model ResNet50.

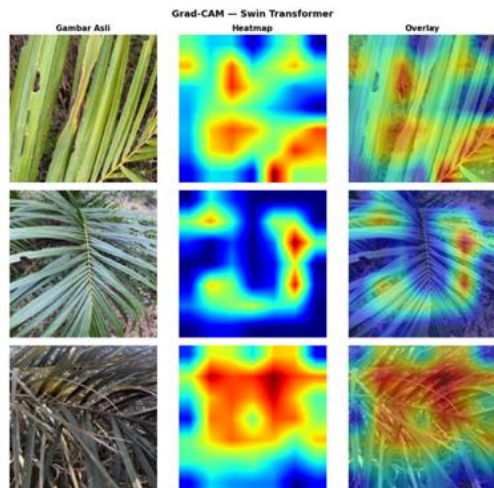


Gambar 7. Kurva Pembelajaran (Learning Curve) Swin Transformer

Untuk memvalidasi area piksel yang menjadi fokus perhatian utama model secara visual, peta panas (*heatmap*) hasil ekstraksi menggunakan metode interpretasi kualitatif Grad-CAM ditampilkan pada Gambar 8 dan Gambar 9.



Gambar 8. Visualisasi Peta Panas (Heatmap) Grad-CAM ResNet50.



Gambar 9. Visualisasi Peta Panas (Heatmap) Grad-CAM Swin Transformer.

Pembahasan

Pencapaian metrik kuantitatif pada Tabel 2 memperlihatkan bahwa model ResNet50 memiliki keunggulan performa kuantitatif dibandingkan Swin Transformer di seluruh parameter pengujian. Tingginya nilai presisi (98,01%) dan recall (98,00%) mengonfirmasi kemampuan ResNet50 dalam menekan galat *False Positive* sekaligus mengenali data infeksi secara menyeluruh. Keunggulan performa ResNet50 ini dipengaruhi secara signifikan oleh integrasi residual connection atau koneksi sisa pada arsitekturnya, yang mempermudah aliran gradien pada jaringan saraf yang dalam sehingga terhindar dari masalah *vanishing gradient*. Meskipun demikian, arsitektur Swin Transformer tetap memberikan hasil yang bersaing sebesar 96,00%, membuktikan bahwa mekanisme perhatian mandiri (*self-attention*) juga sangat efektif dalam mengenali pola penyakit tanaman.

Analisis terhadap distribusi galat pada Gambar 4 dan Gambar 5 memperlihatkan adanya ambiguitas klasifikasi yang spesifik antara kelas *Healthy* dan *Leaf Rust*. Hal ini terjadi secara logis karena

beberapa citra daun kelapa sawit yang baru terinfeksi karat daun pada fase awal menunjukkan degradasi warna dan perubahan tekstur permukaan yang sekilas masih sangat menyerupai karakteristik visual daun sehat, sehingga memicu misklasifikasi pada model.

Tinjauan terhadap kurva pembelajaran pada Gambar 6 dan Gambar 7 memperlihatkan perbedaan stabilitas optimasi yang kontras. Kurva ResNet50 menunjukkan *tren* peningkatan akurasi validasi yang selaras dengan penurunan nilai *loss* secara konsisten, meskipun ada indikasi *overfitting* ringan setelah *epoch* kelima karena kenaikan kecil nilai *validation loss*. Di sisi lain, kurva Swin Transformer memperlihatkan fluktuasi yang tajam pada data validasi, ditandai dengan lonjakan *validation loss* yang signifikan pada *epoch* keempat dan ketujuh. Pola fluktuatif ini mengonfirmasi landasan teoritis bahwa arsitektur berbasis Vision Transformer sangat sensitif terhadap batasan volume data latih, sehingga rentan mengalami penurunan konsistensi generalisasi ketika dihadapkan pada jumlah dataset yang berskala relatif terbatas.

Interpretasi visual kualitatif Grad-CAM pada Gambar 8 dan Gambar 9 mengungkap perbedaan mendasar cara kedua arsitektur dalam mengekstrak fitur visual daun kelapa sawit. Peta perhatian visual ResNet50 menunjukkan pola aktivasi tertinggi (warna merah hingga kuning) yang terisolasi secara tajam langsung pada area bercak infeksi *Curvularia* atau kerusakan jaringan daun. Karakteristik ini membuktikan bahwa filter konvensional CNN pada ResNet50 sangat andal dalam menangkap detail tekstur lokal yang spesifik. Sebaliknya, Swin Transformer menunjukkan pola perhatian yang lebih tersebar secara luas dan global. Meskipun mekanisme *shifted windows* pada Swin Transformer memberikan keunggulan dalam

memahami konteks relasi spasial global dari citra daun secara menyeluruh, penyebaran atensi global ini justru mengurangi ketajaman model dalam melokalisasi tanda-tanda penyakit berukuran kecil. Dengan menggabungkan hasil pengujian numerik dan analisis visual tersebut, dapat divalidasi secara ilmiah bahwa model ResNet50 menyajikan performa yang lebih stabil, akurat, dan sesuai dengan karakteristik biologis infeksi penyakit daun kelapa sawit pada domain penelitian ini.

D. KESIMPULAN DAN SARAN

Penelitian ini menyimpulkan bahwa implementasi arsitektur *deep learning* menggunakan model ResNet50 dan *Swin Transformer* terbukti sangat efektif dan andal dalam melakukan klasifikasi citra jenis penyakit daun kelapa sawit untuk kategori *Curvularia* (bercak daun), *Healthy* (sehat), dan *Leaf Rust* (karat daun). Berdasarkan hasil evaluasi kuantitatif secara komprehensif pada data uji, arsitektur ResNet50 menunjukkan performa yang paling optimal dan unggul dengan mencapai nilai akurasi sebesar 98,00%, presisi 98,01%, *recall* 98,00%, dan *F1-score* 98,00%, mengungguli model *Swin Transformer* yang memperoleh akurasi sebesar 96,00%. Keunggulan ini secara teknis didorong oleh keandalan residual *learning* pada ResNet50 yang mampu mengekstrak fitur visual lokal daun secara mendalam dan stabil tanpa kendala *vanishing gradient*, serta didukung oleh interpretasi kualitatif Grad-CAM yang terbukti lebih presisi dalam melokalisasi area infeksi daun dibandingkan atensi global *Swin Transformer* yang masih menunjukkan fluktuasi tajam pada kurva validasi akibat keterbatasan ukuran dataset. Sebagai langkah pengembangan ke depan untuk memperkuat implementasi di sektor pertanian presisi, disarankan bagi

penelitian selanjutnya untuk memperluas volume dataset citra, menambah variasi kategori penyakit kelapa sawit lainnya, serta mengintegrasikan model terbaik ke dalam aplikasi berbasis mobile atau web agar dapat digunakan secara praktis dan real-time oleh para pekebun di lapangan.

E. REFERENSI

- Aditya, P. N. (2024). Isolasi dan Karakteristik Jamur dari *Ganoderma* di Desa Pernantian Perkebunan PT. Umada. *Jurnal Pengabdian Kepada Masyarakat*, 14(1), 167–172.
- Ali, M., Salma, M., Haji, M. El, & Jamal, B. (2025). Plant disease detection using vision transformers. *International Journal of Electrical and Computer Engineering (IJECE)*, 15(2), 2334. <https://doi.org/10.11591/ijece.v15i2.pp2334-2344>
- Badan Pusat Statistik. (2020). *Statistik Kelapa Sawit Indonesia 2020*. <https://www.bps.go.id/id/publication/2021/11/30/5a3d0448122bc6753c953533/statistik-kelapa-sawit-indonesia-2020.html>
- Hastari, D., Winanda, S., Pratama, A. R., Nurhaliza, N., & Ginting, E. S. (2024). Application of Convolutional Neural Network ResNet-50 V2 on Image Classification of Rice Plant Disease. *Public Research Journal of Engineering, Data Technology and Computer Science*, 1(2). <https://doi.org/10.57152/predatecs.v1i2.865>
- Joseph Alberto, D. H. (2023). Bird Species Classification Using CNN Method and ResNet-50 Architecture. *Jurnal Teknik Informatika Dan Sistem Informasi*, 10(3), 34–46.
- Lee, C. P., Lim, K. M., Song, Y. X., & Alqahtani, A. (2023). Plant-CNN-

- ViT: Plant Classification with Ensemble of Convolutional Neural Networks and Vision Transformer. *Plants*, 12(14), 2642. <https://doi.org/10.3390/plants12142642>
- Mustaqim, T., Safitri, P. H., & Oktavia, V. R. (2025). Peningkatan Sensitivitas Deteksi Diabetic Retinopathy melalui Mekanisme Hierarchical Self-Attention pada Swin Transformer. *Jurnal Algoritma*, 22(2), 2225–2235. <https://doi.org/10.33364/algoritma/v.22-2.2986>
- Priwiratama, H., Wiyono, S., Tondok, E. T., Hidayat, S. H., Wening, S., Wijayanti, E., & Rozziansha, T. A. P. (2024). Genus *Curvularia* and *Pestalotiopsis* as the primary pathogen of leaf spot disease on oil palm seedlings throughout Indonesia. *IOP Conference Series: Earth and Environmental Science*, 1308(1), 5–10. <https://doi.org/10.1088/1755-1315/1308/1/012018>
- Styorini, W., Putra, W. E., Khabzli, W., & Triyani, Y. (2022). Penerapan Deep Learning Pada Jenis Penyakit Tanaman Kelapa Sawit Menggunakan Algoritma Convolutional Neural Network. *Jurnal Komputer Terapan*, 8(2), 359–367. <https://jurnal.pcr.ac.id/index.php/jkt/article/view/5522>
- Yabani, Y., Kuswardani, R. A., Susanto, A., & Syah, R. (2023). Production approach model for oil palm (*Elaeis guineensis* Jacq) superior plant materials using artificial pollination studies. *Journal of Global Innovations in Agricultural Sciences*, 587–594. <https://doi.org/10.22194/JGIAS/23.1237>