

## Klasifikasi Penyakit Daun Sawi Menggunakan VGG19 Berbasis Citra Digital

**Wahab Sya'roni**

Universitas Nurul Jadid, Indonesia

[wahab.syaroni@gmail.com](mailto:wahab.syaroni@gmail.com)

**Yahya Auliya' Abdillah**

Universitas Nurul Jadid, Indonesia

[yahyaauliyaabdillah@unuja.ac.id](mailto:yahyaauliyaabdillah@unuja.ac.id)

**Samsul Arifin**

Universitas Nurul Jadid, Indonesia

[samsularifin434343@gmail.com](mailto:samsularifin434343@gmail.com)

**Rayhan Hibatullah**

Universitas Nurul Jadid, Indonesia

[rayhanhibat18@gmail.com](mailto:rayhanhibat18@gmail.com)

### Abstract

Agricultural productivity is greatly influenced by plant health, including mustard greens (*Brassica rapa*), which are prone to leaf diseases and have high economic value. This study aims to develop a digital image-based classification system for mustard leaf diseases using a deep learning approach, particularly the Convolutional Neural Network (CNN) VGG19 architecture, and to compare its performance with ResNet50 and VGG16 models. The dataset used consists of 999 images divided into two classes: healthy mustard leaves and diseased mustard leaves. The images were processed through preprocessing steps (resized to 224×224 and normalized), then split into training, validation, and testing sets (80:10:10). The VGG19 architecture was customized with additional layers such as Global Average Pooling and Dense layers, and trained for 50 epochs with a configuration of 32 filters, a dropout rate of 0.5, and a learning rate of 0.0003. The results showed that the VGG19 model achieved the highest validation accuracy of 96%, followed by VGG16 with 95%, and ResNet50 with 74%. Evaluation using a confusion matrix demonstrated that VGG19 exhibited the most stable and accurate classification performance in recognizing both classes. These findings reinforce the potential of VGG19 for developing automated and real-time plant disease detection systems. Furthermore, this study opens up opportunities for integration into agricultural Internet of Things (IoT) systems for continuous plant health monitoring, thereby assisting farmers in making faster and more accurate preventive decisions.

**Keywords:** CNN VGG19; Deep Learning; Disease Classification; Mustard Leaves.

## Abstrak

Produktivitas sektor pertanian sangat dipengaruhi oleh kondisi kesehatan tanaman, termasuk tanaman sawi (*Brassica rapa*) yang rentan terhadap penyakit daun dan memiliki nilai ekonomi tinggi. Penelitian ini bertujuan mengembangkan sistem klasifikasi penyakit daun sawi berbasis citra digital menggunakan pendekatan deep learning, khususnya arsitektur Convolutional Neural Network (CNN) VGG19, serta membandingkan kinerjanya dengan model ResNet50 dan VGG16. Dataset yang digunakan terdiri dari 999 gambar yang terbagi ke dalam dua kelas: daun sawi sehat dan daun sawi yang terinfeksi penyakit. Gambar diproses melalui tahapan preprocessing (resize  $224 \times 224$ , normalisasi), lalu dibagi ke dalam set pelatihan, validasi, dan pengujian (80:10:10). Arsitektur VGG19 dikustomisasi dengan layer tambahan seperti Global Average Pooling dan Dense layer, serta dilatih selama 50 epoch dengan konfigurasi filter 32, dropout 0.5, dan learning rate 0.0003. Hasil penelitian menunjukkan bahwa model VGG19 mencapai akurasi validasi tertinggi sebesar 96%, diikuti VGG16 dengan 95%, dan ResNet50 sebesar 74%. Evaluasi menggunakan confusion matrix menunjukkan bahwa VGG19 memiliki performa klasifikasi paling stabil dan akurat dalam mengenali kedua kelas. Temuan ini memperkuat potensi implementasi VGG19 dalam pengembangan sistem deteksi penyakit tanaman secara otomatis dan real-time. Selain itu, penelitian ini membuka peluang integrasi ke dalam sistem Internet of Things (IoT) pertanian untuk monitoring kesehatan tanaman secara berkelanjutan, sehingga dapat membantu petani dalam mengambil keputusan preventif dengan lebih cepat dan tepat.

**Katakunci:** Daun Sawi; Deep Learning; CNN VGG19

## 1 Pendahuluan

Peningkatan produktivitas sektor pertanian menjadi salah satu strategi utama dalam menciptakan ketahanan pangan nasional dan mendorong pertumbuhan ekonomi yang berkelanjutan (Ulfie dan Savira Putri Ayu 2024), (Lesmana, Fadhillah, dan Rozikin 2022). Di antara berbagai komoditas hortikultura yang berperan penting dalam rantai pasok pangan, tanaman sawi (*Brassica rapa*) menempati posisi strategis karena memiliki masa tanam yang relatif singkat, mudah dalam perawatannya, serta bernilai ekonomi tinggi (Palopo 2022). Selain itu, tanaman sawi dikenal memiliki kandungan gizi yang melimpah dan manfaat kesehatan yang signifikan. Salah satu bagian yang bernilai dari tanaman ini adalah bijinya yang dapat diolah menjadi minyak sawi (Sood et al. 2023), sementara daunnya kaya akan vitamin dan sangat bermanfaat sebagai bahan pangan karena kandungannya yang tinggi.

Sebagai sayuran yang banyak dikonsumsi dan memiliki nilai jual tinggi di pasar, daun sawi rentan terhadap berbagai jenis penyakit tanaman, khususnya yang menyerang bagian daunnya (Lukman Priyambodo et al. 2022). Serangan penyakit ini kerap menimbulkan kerusakan serius pada tanaman sebelum masa panen tiba, sehingga menyebabkan penurunan hasil secara kuantitas dan kualitas

(Hafidurrohman dan Kusrini 2025). Dampak dari serangan penyakit tersebut tidak hanya memengaruhi produktivitas lahan, tetapi juga berimplikasi langsung terhadap pendapatan petani serta ketersediaan komoditas di pasar. Oleh karena itu, upaya deteksi dini terhadap gejala penyakit daun sawi sangatlah krusial untuk menjaga kesehatan tanaman serta menjamin keberlangsungan dan mutu hasil panen (Abd Algani et al. 2023).

Deteksi dini tidak hanya memungkinkan tindakan penanganan dilakukan lebih cepat dan tepat, tetapi juga mencegah penyebaran penyakit yang lebih luas antar tanaman (Deviana dan Adinda 2024). Pendekatan ini mendukung pertanian yang berkelanjutan, dengan memberikan perlindungan terhadap tanaman secara real-time dan menjaga kepercayaan konsumen terhadap keamanan produk yang dikonsumsi (Dwianto et al. 2024). Dalam konteks tersebut, teknologi modern memiliki peranan penting, khususnya dalam mengembangkan sistem deteksi penyakit berbasis digital yang cepat dan akurat (Hanifah et al. 2020). Berbagai studi telah menekankan pentingnya pengintegrasian teknologi berbasis citra dan kecerdasan buatan dalam mendeteksi serta mengklasifikasikan penyakit tanaman secara otomatis. Hal ini dinilai dapat memberikan dampak signifikan terhadap peningkatan hasil pertanian secara keseluruhan.

Salah satu pendekatan teknologi yang menonjol dalam pengolahan citra digital adalah Convolutional Neural Network (CNN). CNN merupakan bagian dari arsitektur deep learning yang telah terbukti efektif dalam mengenali pola visual pada gambar melalui ekstraksi fitur secara hierarkis menggunakan lapisan-lapisan konvolusi (Syech Ahmad dan Sugiarto 2023). Keunggulan CNN dalam mengenali objek dan mendeteksi pola-pola visual halus menjadikannya sangat potensial dalam mengidentifikasi gejala penyakit pada daun tanaman, termasuk pada daun sawi. Dengan memanfaatkan data citra daun yang sehat dan yang terinfeksi, CNN dapat dilatih untuk membedakan karakteristik visual yang menjadi indikator adanya penyakit. Model ini tidak hanya mampu memetakan penyakit dengan akurat, tetapi juga dapat digunakan untuk pemantauan tanaman secara otomatis, baik dalam lingkungan laboratorium maupun lapangan, melalui integrasi dengan perangkat berbasis IoT (Internet of Things).

Selain arsitektur CNN murni, berbagai metode hibrid juga mulai banyak dikembangkan untuk meningkatkan akurasi deteksi. Salah satunya adalah penggabungan CNN dengan metode klasifikasi seperti Support Vector Machine (SVM). SVM dikenal sebagai salah satu algoritma pembelajaran mesin yang andal untuk klasifikasi biner dan telah banyak diaplikasikan dalam bidang pertanian. Dalam studi terbaru yang dilakukan oleh (Tian et al. 2020), perbandingan performa antara CNN dengan arsitektur ResNet dan SVM menunjukkan bahwa CNN berhasil memperoleh akurasi sebesar 95,83%, sedangkan SVM hanya mencapai 87,65%. Temuan ini menegaskan bahwa teknologi CNN lebih unggul dalam hal presisi deteksi penyakit dan hama tanaman, sekaligus memperkuat posisinya sebagai teknologi utama dalam penerapan sistem pertanian berbasis kecerdasan buatan.

Implementasi arsitektur CNN dalam sistem deteksi penyakit daun sawi diharapkan dapat memberikan kontribusi nyata dalam peningkatan efisiensi budidaya, menjaga produktivitas lahan, serta menjamin kualitas produk hortikultura yang diterima oleh pasar (Lu, Tan, dan Jiang 2021). Kendati demikian, pengembangan sistem klasifikasi penyakit tanaman menggunakan CNN masih menghadapi sejumlah tantangan teknis.

Beberapa di antaranya meliputi pemilihan metode preprocessing yang tepat, pemilihan arsitektur CNN yang optimal, pengaturan parameter pelatihan (hyperparameter tuning), serta pengolahan dataset yang bersih dan representatif.

Penelitian oleh (Nazalia et al. 2023) menyoroti pentingnya pengembangan sistem deteksi penyakit dan hama tanaman, khususnya pada tanaman sawi, mengingat besarnya dampak yang ditimbulkan terhadap sektor pertanian dan kesehatan masyarakat. Dalam studi tersebut, penerapan algoritma CNN untuk deteksi hama pada daun sawi menunjukkan hasil yang memuaskan, dengan akurasi mencapai 92%. Hal ini membuktikan bahwa metode berbasis deep learning tidak hanya efisien, tetapi juga mampu menghasilkan model klasifikasi yang andal untuk diterapkan secara langsung oleh petani, baik melalui aplikasi berbasis seluler maupun sistem deteksi otomatis di lapangan.

Meskipun CNN telah menunjukkan performa tinggi dalam mendeteksi penyakit tanaman, permasalahan teknis seperti overfitting, underfitting, dan pemilihan struktur jaringan yang optimal masih menjadi fokus dalam penelitian lanjutan (Trihardianingsih, Sunyoto, dan Hidayat 2023). Beberapa strategi mitigasi yang dapat dilakukan untuk mengatasi permasalahan tersebut antara lain: penggunaan teknik regularisasi seperti dropout dan early stopping untuk mencegah overfitting, penerapan data augmentation guna memperluas variasi data latih, pemilihan jumlah epoch dan batch size yang sesuai agar model tidak mengalami underfitting maupun overfitting, serta eksplorasi hyperparameter tuning (misalnya pengaturan learning rate, optimizer, dan fungsi aktivasi) untuk menemukan konfigurasi terbaik. Berbagai arsitektur CNN seperti ResNet-50, Xception, dan Beginning telah diuji dengan hasil akurasi tinggi (Trihardianingsih et al. 2023), namun sebagian besar penelitian masih terbatas dalam hal eksplorasi mendalam terhadap pengaruh parameter optimisasi seperti fungsi aktivasi, learning rate, optimizer, dan batch size terhadap performa akhir model.

Dalam beberapa tahun terakhir, kemajuan teknologi deep learning telah mempercepat perkembangan metode identifikasi penyakit tanaman berbasis citra digital. CNN telah

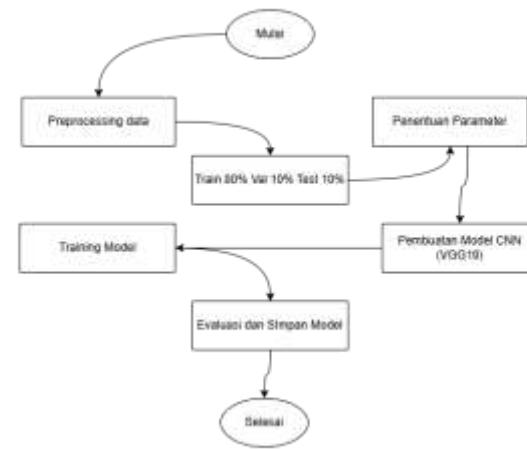
menjadi pendekatan dominan karena kemampuannya dalam mengekstraksi fitur yang kompleks dan relevan dari gambar daun. Namun, agar teknologi ini dapat diimplementasikan secara luas di sektor pertanian, diperlukan model yang tidak hanya akurat, tetapi juga efisien dalam hal kecepatan prediksi dan penggunaan sumber daya komputasi. Oleh karena itu, penelitian yang berfokus pada optimasi arsitektur dan perbandingan model menjadi sangat penting, khususnya untuk tanaman dengan nilai strategis seperti sawi.

Dalam kerangka tersebut, penelitian ini bertujuan untuk mengevaluasi dan meningkatkan performa klasifikasi penyakit pada daun sawi dengan membandingkan tiga arsitektur CNN, yaitu VGG19, VGG16, dan ResNet50. Kebaruan penelitian ini terletak pada analisis komparatif dengan konfigurasi pelatihan yang sama, sehingga dapat diidentifikasi arsitektur yang paling unggul dan stabil untuk diterapkan pada deteksi penyakit daun sawi berbasis citra digital. Dataset yang digunakan bersumber dari Kaggle, yang sebelumnya telah menjadi rujukan berbagai studi serupa. Penelitian ini juga mengusulkan pendekatan baru melalui optimasi struktur jaringan dan pengaturan parameter yang lebih presisi guna meningkatkan akurasi serta mempercepat waktu prediksi. Diharapkan hasil dari penelitian ini dapat memberikan kontribusi nyata dalam pengembangan sistem deteksi penyakit berbasis AI di bidang pertanian, sekaligus membuka peluang integrasi ke sistem IoT untuk monitoring tanaman secara real-time.

## 2 Metode

### Tahapan Penelitian

Bagian ini menjelaskan metode penelitian yang digunakan untuk membangun dan mengevaluasi kinerja model. Informasi yang diperlukan seperti metode yang dipilih untuk memperoleh kumpulan data, teknik persiapan data, teknik analisis data, dll. seperti yang ditunjukkan pada gambar 1.



**Gambar 1.** Alur Penelitian

### Dataset

Penelitian ini menggunakan Dataset yang berasal dari data kaggle yang terdiri dari 2 kelas yang terdiri dari 999 gambar. Dua kelas merupakan penyakit daun sawi dan daun sawi sehat. Dataset ini terdiri dari 499 gambar sawi ada hama dan 500 gambar daun sawi sehat. Berikut gambar 2 merupakan gambar sampel yang diambil dari dataset.



**Gambar 2.** Sample Dataset

### Preprocessing

Pada tahap preprocessing, seluruh citra diubah ukurannya menjadi  $224 \times 224$  piksel dengan 3 channel (RGB), kemudian dinormalisasi ke rentang nilai 0-1 untuk mempercepat proses pelatihan dan meningkatkan stabilitas model. Selanjutnya, data dibagi ke dalam batch berukuran 64 gambar untuk menjaga efisiensi pelatihan serta penggunaan memori GPU.

## Split Dataset

Setelah proses prapengolahan selesai dilakukan, langkah selanjutnya dalam penelitian ini adalah pembagian dataset menjadi tiga bagian utama, yaitu 80% sebagai data latih (training set), 10% data validasi (validation set), dan 10% data uji (test set). Pembagian ini bertujuan untuk memisahkan data yang digunakan untuk membangun model, mengevaluasi performanya selama proses pelatihan, serta menguji keakuratan akhir dari model yang telah dilatih.

## Pembuatan Model

Setelah data diproses dan dibagi, tahap berikutnya adalah membangun model klasifikasi menggunakan arsitektur VGG19. Model ini diimpor dari pustaka Keras dengan bobot awal dari ImageNet, namun tanpa lapisan atas (`include_top=False`), agar dapat disesuaikan dengan kebutuhan klasifikasi dua kelas. Beberapa layer tambahan ditambahkan pada model, seperti Global Average Pooling untuk meratakan output fitur, Dense layer dengan aktivasi ReLU, dan Dropout layer untuk mencegah overfitting. Lapisan output menggunakan Dense layer dengan satu neuron dan fungsi aktivasi sigmoid, karena klasifikasi bersifat biner. Model dikompilasi dengan optimizer Adam, fungsi loss `binary_crossentropy`, dan metrik akurasi. Pelatihan dilakukan dengan batch size 64 menggunakan data latih, serta divalidasi menggunakan data validasi untuk memantau kinerja model selama proses pelatihan.

## Training Model

Setelah model selesai dibangun dan dikompilasi, tahap selanjutnya adalah proses pelatihan (training) terhadap data latih yang telah disiapkan sebelumnya. Pelatihan model dilakukan dengan menggunakan data latih sebanyak 80% dari total dataset, sedangkan 10% digunakan sebagai data validasi untuk memantau kinerja model secara berkala di setiap epoch. Untuk mendapatkan hasil pelatihan yang optimal, penelitian ini menggunakan jumlah epoch, yaitu 10, 20, 50, dan 100 epoch. Tujuan dari penggunaan berbagai jumlah epoch ini adalah untuk melihat pengaruh durasi pelatihan terhadap akurasi dan ketstabilan model. Pada setiap percobaan, model dilatih menggunakan batch size 64, yang berarti setiap iterasi pelatihan

memproses 64 gambar sekaligus sebelum memperbarui bobot. Fungsi aktivasi yang digunakan pada layer output adalah sigmoid, karena output hanya terdiri dari dua kelas. Proses pelatihan dilakukan dengan fungsi loss `binary_crossentropy` dan optimizer Adam karena keduanya dinilai efektif dalam menangani klasifikasi biner dan konvergensi yang cepat. Hasil pelatihan pada setiap variasi epoch diamati melalui akurasi dan loss pada data latih dan data validasi. Dengan membandingkan hasil antar epoch, dapat dianalisis pada jumlah epoch berapa model mencapai keseimbangan terbaik antara akurasi dan generalisasi, serta menghindari gejala overfitting atau underfitting.

## Evaluasi dan Simpan Model

Setelah proses pelatihan selesai dilakukan pada berbagai variasi epoch, tahap selanjutnya adalah evaluasi model untuk mengetahui seberapa baik performa model dalam mengklasifikasikan gambar daun sawi ke dalam dua kategori: sehat dan terinfeksi penyakit atau hama. Evaluasi dilakukan dengan menggunakan data uji sebesar 10% dari total dataset yang sebelumnya tidak pernah dilihat oleh model selama proses pelatihan maupun validasi. Setelah proses evaluasi selesai dan model menunjukkan hasil yang optimal terutama pada jumlah epoch yang memberikan keseimbangan antara akurasi tinggi dan ketstabilan loss maka model disimpan untuk digunakan kembali di masa mendatang. Proses penyimpanan model dilakukan menggunakan format standar dari Keras, yaitu dengan menyimpan model dalam bentuk file .h5. Penyimpanan model dilakukan dengan perintah `model.save('model_vgg19_sawi.h5')`.

## 3 Hasil dan Diskusi

### Pengumpulan data

Penelitian ini mendapatkan data dari metode Secondary data analysis, yang dimana metode ini melibatkan penggunaan data yang telah dikumpulkan oleh peneliti-peneliti sebelumnya. Data yang didapatkan dapat berupa arsip yang tersedia secara publik dan diperbolehkan untuk digunakan, biasanya terdapat pada repository GITHUB, Driver, Kaggle dan lain sebagainya.

Untuk melakukan eksperimen pada penelitian kali ini, peneliti menggunakan

dataset yang berasal dari data kaggle yang terdiri dari 2 kelas yang terdiri dari 999 gambar. Dua kelas merupakan penyakit daun sawi dan daun sawi sehat. Dataset ini terdiri dari 499 gambar sawi ada hama dan 500 gambar daun sawi sehat. Berikut gambar 2 merupakan gambar sampel yang diambil dari dataset.



**Gambar 3.** Sample Dataset

Gambar 3 menampilkan sampel gambar dari dataset yang digunakan dalam penelitian ini, yaitu dataset penyakit daun sawi yang diperoleh dari Kaggle. Dengan menggunakan dataset ini, penelitian bertujuan untuk mengembangkan model deteksi penyakit jagung yang lebih akurat.

**Table 1.** Jumlah Dataset

No	Jenis Daun	Jumlah
1	Daun Sawi ada Hama	499
2	Daun Sawi Sehat	500
<b>Total Keseluruhan</b>		<b>999</b>

#### Pelatihan Model

Tahap selanjutnya yaitu pelatihan model yang dilakukan sebanyak 50 epoch untuk setiap arsitektur yang digunakan, yaitu VGG19, ResNet50, dan Xception. Pemilihan jumlah epoch ini didasarkan pada pertimbangan efisiensi pelatihan dan kestabilan kinerja model pada data validasi. Penggunaan jumlah epoch yang sama pada setiap model bertujuan untuk menjaga konsistensi dalam evaluasi performa dan membandingkan hasil akhir secara adil. Tabel 1 menunjukkan skenario pelatihan yang diterapkan oleh peneliti.

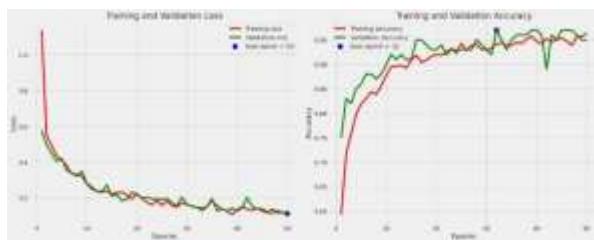
**Table 2.** Skenario-skenario yang dilakukan

No	Model	Parameter	Data	Epoch
1	VGG19	Filter 32, Do, 0.5, Lr 0.0003	Daun Sawi	50
2	Resnet50	Filter 32, Do, 0.5, Lr 0.0003	Daun Sawi	50
3	VGG16	Filter 32, Do, 0.5, Lr 0.0003	Daun Sawi	50

Pada Tabel 2 disajikan skenario pelatihan model yang dilakukan oleh peneliti. Terdapat tiga skenario utama berdasarkan arsitektur model CNN yang digunakan, yaitu VGG19, ResNet50, dan VGG16. Seluruh skenario menggunakan dataset yang sama, yaitu gambar daun sawi yang bersifat imbalance, dan pelatihan dilakukan secara konsisten selama 50 epoch untuk setiap model. Skenario pertama menggunakan model VGG19 dengan konfigurasi parameter yang meliputi jumlah filter sebanyak 32, dropout rate sebesar 0.5, dan learning rate sebesar 0.0003. Model ini dilatih selama 50 epoch tanpa variasi jumlah epoch, dengan tujuan memperoleh kestabilan dalam proses pelatihan dan menghindari overfitting akibat pelatihan yang terlalu panjang.

Skenario kedua menggunakan model ResNet50 dengan konfigurasi parameter yang identik, yaitu filter 32, dropout 0.5, dan learning rate 0.0003. ResNet50 dipilih karena kemampuannya dalam menangani permasalahan vanishing gradient melalui mekanisme residual connection, yang relevan untuk klasifikasi citra daun yang memiliki tekstur kompleks. Pelatihan dilakukan selama 50 epoch, mengikuti pendekatan yang konsisten agar hasil antar model dapat dibandingkan secara objektif. Skenario ketiga menerapkan model VGG16, yaitu versi yang lebih ringan dibandingkan VGG19 namun masih mempertahankan struktur berlapis khas arsitektur VGG. Konfigurasi parameter tetap dipertahankan (filter 32, dropout 0.5, dan learning rate 0.0003), serta pelatihan juga dilakukan selama 50 epoch. Tujuan dari skenario ini adalah untuk membandingkan kinerja antara VGG16 dan VGG19 dalam

konteks kompleksitas model terhadap akurasi dan efisiensi pelatihan. Gambar 4 menunjukkan hasil grafik dari proses pelatihan menggunakan model VGG19, Gambar 5 menunjukkan hasil grafik dari proses pelatihan menggunakan model Resnet50, dan Gambar 6 menunjukkan hasil grafik dari proses pelatihan menggunakan model VGG16.



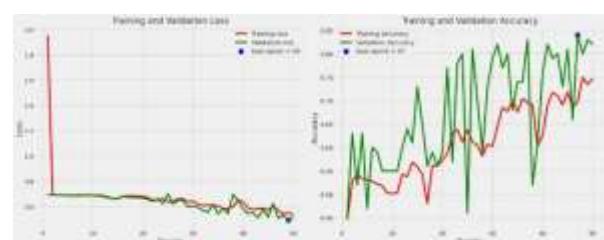
**Gambar 4.** Hasil grafik pelatihan model VGG19

Pada Gambar 4 ditunjukkan hasil pelatihan menggunakan arsitektur VGG19 yang diterapkan pada dataset citra daun sawi dengan konfigurasi parameter berupa jumlah filter sebanyak 32, dropout rate 0.5, dan learning rate sebesar 0.0003. Pelatihan dilakukan selama 50 epoch, yang dipilih secara konsisten untuk menjaga kestabilan proses pembelajaran dan menghindari risiko overfitting maupun underfitting. Model ini merupakan salah satu arsitektur CNN yang cukup dalam dan telah terbukti efektif dalam pengenalan citra dengan struktur hierarkis yang kompleks.

Hasil yang diperoleh dari pelatihan ini menunjukkan akurasi validasi sebesar 96%, yang mengindikasikan bahwa model mampu mengenali dan mengklasifikasikan pola visual dari daun sawi dengan sangat baik. Kinerja model VGG19 yang tinggi juga didukung oleh penggunaan dropout sebesar 0.5 yang berfungsi sebagai regularisasi untuk mencegah overfitting, serta pemilihan learning rate yang konservatif (0.0003) untuk memastikan proses optimisasi berjalan stabil dan tidak melompat-lompat dalam pencarian minimum loss.

Dalam proses pelatihan ini, grafik akurasi dan loss menunjukkan tren yang cukup ideal: akurasi terus meningkat dengan fluktuasi minimum, sementara loss pada data validasi menurun secara konsisten, menandakan bahwa model belajar secara efektif dan tidak

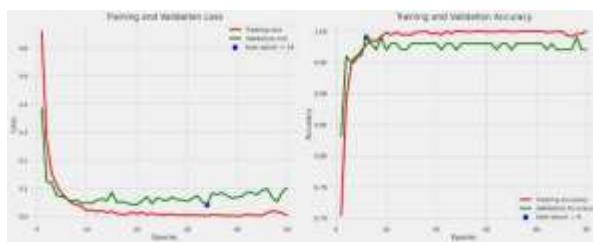
hanya menghafal data latih. Dengan demikian, eksperimen ini menunjukkan bahwa model VGG19 dengan konfigurasi tersebut sangat potensial untuk diterapkan dalam klasifikasi daun sawi, bahkan dalam kondisi dataset yang tidak seimbang (imbalance). Akurasi 96% ini menjadi indikator keberhasilan proses pelatihan dan kemampuan generalisasi model terhadap data validasi.



**Gambar 5.** Hasil grafik pelatihan model Resnet50

Pada skenario kedua, digunakan arsitektur ResNet50 yang juga diterapkan pada dataset citra daun sawi dengan konfigurasi parameter yang sama, yaitu jumlah filter sebanyak 32, dropout rate sebesar 0.5, dan learning rate sebesar 0.0003. Pelatihan dilakukan selama 50 epoch guna menjaga konsistensi antar skenario dan memastikan bahwa perbandingan performa antar model dapat dilakukan secara adil. ResNet50 merupakan model deep learning yang mengusung pendekatan residual learning, di mana lapisan-lapisan konvolusional disusun dengan shortcut connection untuk mengatasi permasalahan vanishing gradient yang umum terjadi pada jaringan yang sangat dalam. Berdasarkan hasil pelatihan, model ResNet50 mampu mencapai akurasi validasi sebesar 74%, yang menandakan bahwa model cukup mampu mengenali karakteristik citra daun sawi, meskipun belum sebaik model VGG19 dalam skenario pertama. Akurasi ini mencerminkan bahwa model dapat melakukan klasifikasi dengan tingkat ketepatan yang layak, namun mungkin belum sepenuhnya mengoptimalkan representasi fitur visual dari data yang tersedia. Hal ini bisa disebabkan oleh interaksi antara kompleksitas arsitektur ResNet50 dan kondisi dataset yang belum diseimbangkan (imbalance), sehingga model cenderung menghadapi tantangan dalam membedakan fitur minoritas secara konsisten.

Dalam grafik pelatihan, akurasi cenderung meningkat secara progresif, meskipun terdapat fluktuasi kecil pada fase akhir epoch yang mengindikasikan kemungkinan awal overfitting. Sementara itu, loss pada data validasi cenderung menurun namun tidak secepat penurunan pada model VGG19. Hal ini menunjukkan bahwa meskipun ResNet50 memiliki kemampuan yang baik dalam menyerap informasi dari data latih, efisiensi generalisasinya pada data validasi belum optimal. Dengan demikian, skenario kedua memberikan wawasan bahwa pemilihan arsitektur model memiliki dampak signifikan terhadap performa klasifikasi. Meskipun ResNet50 dikenal efisien dalam berbagai tugas klasifikasi citra, dalam konteks klasifikasi daun sawi dengan kondisi dataset imbalance, performanya masih berada di bawah model VGG19 pada skenario pertama.



**Gambar 6.** Hasil grafik pelatihan model VGG16

Pada skenario ketiga, peneliti menggunakan arsitektur VGG16 dengan konfigurasi parameter yang sama seperti pada skenario sebelumnya, yaitu jumlah filter sebanyak 32, dropout rate sebesar 0.5, dan learning rate sebesar 0.0003. Pelatihan dilakukan selama 50 epoch menggunakan dataset citra daun sawi yang bersifat imbalance. VGG16 merupakan versi yang lebih ringan dari VGG19, dengan jumlah lapisan konvolusional yang lebih sedikit, namun tetap mempertahankan pola arsitektur berlapis yang sistematis dalam proses ekstraksi fitur visual. Hasil pelatihan menunjukkan bahwa model VGG16 berhasil mencapai akurasi validasi sebesar 95%, hanya terpaut 1% dari model VGG19 yang mencapai akurasi 96% pada skenario pertama. Perbedaan ini sangat kecil dan mengindikasikan bahwa VGG16 tetap mampu melakukan klasifikasi citra daun sawi dengan sangat baik, meskipun kompleksitas arsitekturnya lebih rendah. Hal ini menjadikan

VGG16 sebagai pilihan yang efisien, khususnya dalam kondisi keterbatasan sumber daya komputasi atau kebutuhan implementasi real-time. Dari segi grafik pelatihan, model VGG16 menunjukkan tren yang stabil baik dari sisi akurasi maupun loss. Akurasi validasi meningkat secara bertahap dan konsisten sepanjang proses pelatihan, sementara loss mengalami penurunan signifikan di fase awal dan cenderung stabil di akhir epoch. Fluktuasi nilai akurasi dan loss sangat minimal, menandakan bahwa proses pembelajaran berlangsung secara optimal dan minim risiko overfitting.

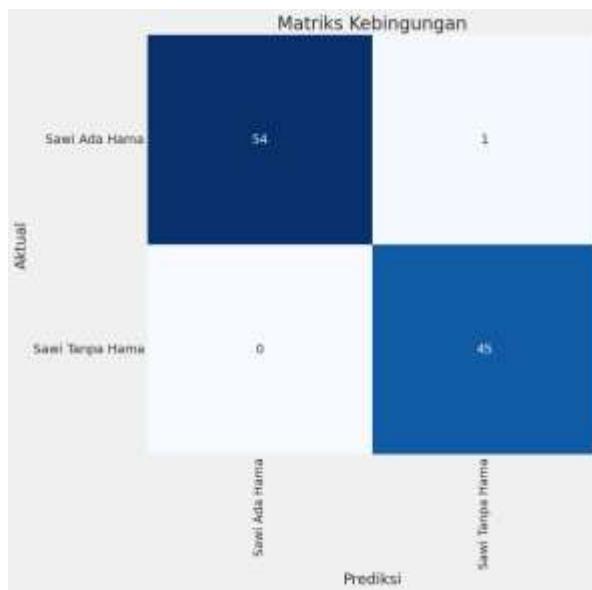
Jika dibandingkan dengan dua skenario sebelumnya, VGG19 mencatatkan performa tertinggi dengan akurasi 96%, disusul oleh VGG16 dengan 95%, dan ResNet50 dengan 74%. Perbedaan yang mencolok antara ResNet50 dan dua model VGG mengindikasikan bahwa pendekatan residual connection pada ResNet belum cukup efektif menangani karakteristik dataset yang tidak seimbang. Sementara itu, model VGG tampil lebih konsisten dan unggul dalam mengenali pola visual dari daun sawi.

Dengan mempertimbangkan hasil tersebut, model VGG19 direkomendasikan sebagai arsitektur terbaik dalam penelitian ini karena menunjukkan performa klasifikasi paling akurat secara konsisten. Meskipun VGG16 menawarkan efisiensi dalam kompleksitas model dan durasi pelatihan, VGG19 lebih unggul dalam kemampuan ekstraksi fitur yang lebih dalam, menjadikannya lebih cocok untuk kebutuhan aplikasi yang memprioritaskan ketepatan klasifikasi tinggi, seperti sistem monitoring pertanian berbasis visual atau aplikasi identifikasi tanaman berbasis AI.

### Evaluasi Model

Setelah model berhasil diimplementasi langkah terakhir pada penelitian ini adalah evaluasi model menggunakan confusion matrix untuk mendapatkan pemahaman yang lebih mendalam tentang kesalahan prediksi yang dilakukan model. Setiap elemen dalam confusion matrix menunjukkan berapa banyak prediksi yang benar atau salah dalam setiap kelas, sehingga memberikan wawasan yang jelas tentang di mana model sering salah dalam mengklasifikasikan gambar. Gambar 7 menunjukkan hasil confusion matrix dari

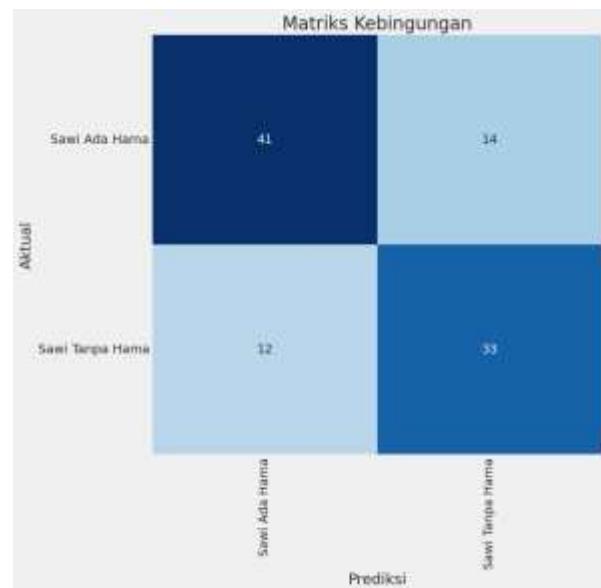
proses pelatihan menggunakan model VGG19, Gambar 8 menunjukkan hasil confusion matrix dari proses pelatihan menggunakan model Resnet50 dan Gambar 9 menunjukkan hasil confusion matrix dari proses pelatihan menggunakan model VGG16.



**Gambar 7.** Hasil confusion matrix pelatihan model VGG19

Pada Gambar 7 ditampilkan confusion matrix yang merepresentasikan performa klasifikasi model VGG19 dalam skenario pertama pada dataset citra daun sawi. Matriks tersebut menunjukkan dua kategori utama, yaitu "Sawi Ada Hama" dan "Sawi Tanpa Hama", sebagai label aktual dan prediksi. Hasil klasifikasi menunjukkan bahwa sebanyak 54 gambar "Sawi Ada Hama" berhasil diprediksi dengan benar, sementara hanya 1 gambar salah diklasifikasikan sebagai "Sawi Tanpa Hama". Pada sisi lain, untuk kategori "Sawi Tanpa Hama", seluruh 45 gambar berhasil diprediksi dengan benar tanpa ada kesalahan klasifikasi. Tidak terdapat kasus false positive dari kategori "Sawi Tanpa Hama" yang diprediksi sebagai "Sawi Ada Hama". Secara kuantitatif, model mencatatkan jumlah klasifikasi benar (True Positives and True Negatives) sebesar 99 gambar, dan hanya 1 kasus klasifikasi salah (False Negative). Hal ini mengindikasikan bahwa model VGG19 memiliki sensitivitas (recall) yang tinggi, terutama dalam mengenali daun sawi yang tidak terkena hama, serta akurasi keseluruhan yang sangat tinggi dalam konteks dua kelas yang diuji.

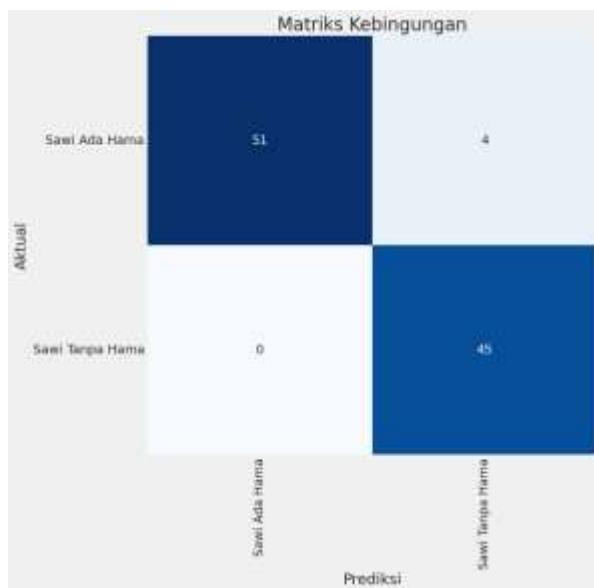
Berdasarkan distribusi confusion matrix tersebut, dapat disimpulkan bahwa model tidak hanya mampu membedakan fitur visual antar kelas secara efektif, tetapi juga menunjukkan performa yang sangat kuat dan stabil, terutama dalam kondisi dataset yang bersifat tidak seimbang. Capaian ini memperkuat posisi VGG19 sebagai model dengan performa terbaik dalam eksperimen ini.



**Gambar 8.** Hasil confusion matrix pelatihan model Resnet50

Pada Gambar 8 ditampilkan confusion matrix hasil pelatihan model ResNet50 dalam skenario kedua untuk klasifikasi citra daun sawi yang terbagi menjadi dua kategori, yaitu "Sawi Ada Hama" dan "Sawi Tanpa Hama". Matriks ini memberikan gambaran kinerja klasifikasi model dalam mengelompokkan data aktual terhadap hasil prediksi. Untuk kategori "Sawi Ada Hama", model berhasil memprediksi dengan benar sebanyak 41 gambar, namun terdapat 14 gambar yang salah diklasifikasikan sebagai "Sawi Tanpa Hama". Sementara itu, pada kategori "Sawi Tanpa Hama", sebanyak 33 gambar berhasil diklasifikasikan dengan benar, namun masih terdapat 12 gambar yang salah diklasifikasikan sebagai "Sawi Ada Hama". Dengan demikian, jumlah prediksi benar secara keseluruhan adalah 74 gambar, dan jumlah prediksi salah sebanyak 26 gambar. Hasil ini menghasilkan akurasi total sebesar 74%, sesuai dengan hasil evaluasi pada skenario kedua. Performa model ResNet50 dalam hal ini masih cukup layak, namun

terlihat adanya ketidakseimbangan dalam mengenali kedua kelas secara akurat, khususnya pada kasus false negative di mana model seringkali gagal mendeteksi keberadaan hama pada daun sawi. Hal ini dapat disebabkan oleh struktur dataset yang imbalance atau karakteristik visual antar kelas yang tidak kontras pada fitur yang ditangkap oleh ResNet50.



**Gambar 9.** Hasil confusion matrix pelatihan model VGG16

Pada Gambar 9 ditampilkan confusion matrix hasil prediksi model VGG16 dalam skenario ketiga yang digunakan untuk mengklasifikasikan citra daun sawi ke dalam dua kategori: "Sawi Ada Hama" dan "Sawi Tanpa Hama". Pelatihan dilakukan selama 50 epoch dengan konfigurasi yang sama seperti dua skenario sebelumnya, yaitu filter 32, dropout rate 0.5, dan learning rate 0.0003, menggunakan dataset yang bersifat imbalance. Hasil klasifikasi menunjukkan bahwa dari kategori "Sawi Ada Hama", sebanyak 51 gambar berhasil diprediksi dengan benar, sedangkan 4 gambar salah diklasifikasikan sebagai "Sawi Tanpa Hama". Sementara itu, pada kategori "Sawi Tanpa Hama", model menunjukkan performa sempurna dengan 45 gambar diprediksi dengan benar tanpa ada kesalahan klasifikasi (false positive = 0). Secara keseluruhan, model VGG16 menghasilkan jumlah prediksi benar sebanyak 96 gambar, dan hanya 4 gambar yang diklasifikasikan secara salah. Hal ini memberikan akurasi keseluruhan sebesar 96%, yang sangat tinggi dan hanya sedikit

lebih rendah dari hasil model VGG19 (96%) namun jauh lebih baik dibandingkan model ResNet50 (74%).

## 4 Kesimpulan

Penelitian ini berhasil membuktikan bahwa arsitektur VGG19 merupakan model paling optimal untuk klasifikasi penyakit daun sawi berbasis citra digital, dengan akurasi mencapai 96%. Hasil ini menunjukkan bahwa pendekatan deep learning berbasis CNN, khususnya dengan struktur berlapis dalam seperti VGG19, mampu mengekstraksi fitur visual daun secara efektif dan mengidentifikasi keberadaan penyakit secara akurat, bahkan pada dataset yang tidak seimbang. Implikasi dari penelitian ini adalah potensi pengembangan sistem deteksi penyakit tanaman yang dapat diterapkan langsung oleh petani menggunakan perangkat digital sederhana, seperti kamera smartphone. Sistem ini diharapkan mampu meningkatkan kecepatan deteksi dan mengurangi ketergantungan terhadap diagnosis manual yang subjektif. Namun demikian, penelitian ini masih memiliki keterbatasan pada jumlah kelas penyakit yang terbatas dan penggunaan dataset sekunder. Untuk penelitian selanjutnya, disarankan dilakukan pengumpulan data primer yang lebih representatif serta pengujian pada berbagai kondisi pencahayaan dan latar belakang yang lebih bervariasi. Selain itu, eksplorasi arsitektur lain dan implementasi dalam bentuk aplikasi lapangan juga menjadi rekomendasi penting sebagai pengembangan lebih lanjut.

## 5 Referensi

- Abd Algani, Yousef Methkal, Orlando Juan Marquez Caro, Liz Maribel Robladillo Bravo, Chamandeep Kaur, Mohammed Saleh Al Ansari, dan B. Kiran Bala. 2023. "Leaf disease identification and classification using optimized deep learning." *Measurement: Sensors* 25(September 2022):100643. doi: 10.1016/j.measen.2022.100643.
- Deviana, Lyla Putri, dan Puja Restu Adinda. 2024. "Sistem Monitoring Pertumbuhan Tanaman Sawi Menggunakan Artificial Intelligence Pada Aquaponik." 9(3):306–14. doi: 10.30591/jpit.v9i3.5897.
- Dwianto, Sigit, Farhan Naufal Mubarok, Dzaki

- Satriatama, Tinuk Agustin, dan Sukoharjo Indonesia. 2024. "PENERAPAN CONVOLUTIONAL NEURAL NETWORK ( CNN ) DALAM." (November):270-80.
- Hafidurrohman, M., dan K. Kusrini. 2025. "Classification Of Mustard Leaf Diseases Using Convolutional Neural Network Architecture." *Journal of Electrical Engineering and Computer (JEECOM)* 7(1):22-30. doi: 10.33650/jecom.v7i1.10779.
- Hanifah, Siti, Nur Apriliani, Eddy Tri Sucianto, dan Endang Sri Purwati. 2020. "Identifikasi Jamur Penyebab Penyakit pada Tanaman Sawi Putih ( Brassica rapa L.) dan Persentase Penyakitnya di Desa Serang Kecamatan Karangreja, Kabupaten Purbalingga." *Bio Eksaka: Jurnal Ilmiah Biologi Unsoed* 2(2020):487-501.
- Lesmana, Alang Mulya, Ronna Putri Fadhillah, dan Chaerur Rozikin. 2022. "Identifikasi Penyakit pada Citra Daun Kentang Menggunakan Convolutional Neural Network (CNN)." *Jurnal Sains dan Informatika* 8(1):21-30. doi: 10.34128/jsi.v8i1.377.
- Lu, Jinzhu, Lijuan Tan, dan Huanyu Jiang. 2021. "Review on convolutional neural network (CNN) applied to plant leaf disease classification." *Agriculture (Switzerland)* 11(8).
- Lukman Priyambodo, Hanin Latif Fuadi, Naura Nazhifah, Ibrohim Huzaimi, Angga Bagus Prawira, Tasya Enjelika Saputri, Mas Aly Afandi, Eka Setia Nugraha, Agung Wicaksono, dan Petrus Kerowe Goran. 2022. "Klasifikasi Kematangan Tanaman Hidroponik Pakcoy Menggunakan Metode SVM." *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)* 6(1):153-60. doi: 10.29207/resti.v6i1.3828.
- Nazalia, Cendekia Luthfieta, Pritasari Palupiningsih, Budi Prayitno, dan Yudhi Setyo Purwanto. 2023. "Implementation of Convolutional Neural Network Algorithm to Pest Detection in Caisim." *ICCoSITE 2023 - International Conference on Computer Science, Information Technology and Engineering: Digital Transformation Strategy in Facing the VUCA and TUNA Era* 609-14. doi: 10.1109/ICCoSITE57641.2023.1012779 2.
- Palopo, Biologi Sman. 2022. "batang dan buah . Daun tanaman yang." 12:42-50.
- Sood, Akshit, Pradeepa Kumar Sarangi, Ashok Kumar Sahoo, Lekha Rani, Karan Bajaj, dan Alok Kumar Agrawal. 2023. "AI-Driven Mustard Disease Identification: A Multiclass and Binary Classification Approach for Advanced Crop Health Monitoring." *Proceedings - International Conference on Technological Advancements in Computational Sciences, ICTACS 2023* 256-61. doi: 10.1109/ICTACS59847.2023.10390082.
- Syech Ahmad, Mochamad Taufik Ali, dan Bambang Sugiarto. 2023. "Implementasi Convolutional Neural Network (CNN) untuk Klasifikasi Ikan Cupang Berbasis Mobile." *Digital Transformation Technology* 3(2):712-23. doi: 10.47709/digitech.v3i2.3245.
- Tian, L. G., C. Liu, Y. Liu, M. Li, J. Y. Zhang, dan H. L. Duan. 2020. "Research on plant diseases and insect pests identification based on CNN." *IOP Conference Series: Earth and Environmental Science* 594(1). doi: 10.1088/1755-1315/594/1/012009.
- Trihardianingsih, Liana, Andi Sunyoto, dan Tonny Hidayat. 2023. "Classification of Tea Leaf Diseases Based on ResNet-50 and Inception V3." *Sinkron* 8(3):1564-73. doi: 10.33395/sinkron.v8i3.12604.
- Ulfi, Meitra, dan Tengku Savira Putri Ayu. 2024. *Implementasi Deep Learning dengan Convolutional Neural Network untuk Tingkat Akurasi Citra Image Hama Sawi Hijau Menggunakan Google Colab.* Vol. 10.