

Analisis Hyperparameter Tuning MobileNetV2 dengan Metode Sequential Search dalam Sistem Klasifikasi Penyakit Daun Kentang

Muhammad Ivan Khoirur Rizky, Akfi Rozada, Nurul Baroroh, Ricardus Anggi Pramunendar*

Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia

Email: ¹111202214778@mhs.dinus.ac.id, ²111202214724@mhs.dinus.ac.id, ³111202214802@mhs.dinus.ac.id,

^{4,*}ricardus.anggi@dsn.dinus.ac.id,

Email Penulis Korespondensi: ricardus.anggi@dsn.dinus.ac.id

Submitted: 24/11/2025; Accepted: 26/12/2025; Published: 26/12/2025

Abstrak—Produksi kentang nasional sebagai komoditas strategis menghadapi ancaman serius dari penyakit daun, sementara proses klasifikasi manual masih lambat, subjektif, dan rentan kesalahan akibat kemiripan fitur visual antar penyakit. Kondisi ini menegaskan kebutuhan akan sistem klasifikasi otomatis yang presisi dan andal. Namun, sebagian besar penelitian terdahulu belum menerapkan optimasi hyperparameter secara sistematis sehingga kemampuan arsitektur deep learning belum dimanfaatkan sepenuhnya. Berdasarkan research gap tersebut, penelitian ini bertujuan meningkatkan performa MobileNetV2 pada tugas klasifikasi penyakit daun kentang melalui strategi optimasi hyperparameter yang terstruktur. Pendekatan Sequential Search digunakan dan divalidasi melalui 3 kali Stratified Cross Validation untuk memperoleh estimasi kinerja yang stabil. Empat hyperparameter dievaluasi, yaitu learning rate dari 0.001 sampai 0.009, dropout dari 0.1 sampai 0.9, batch size dari 8 sampai 192, dan epoch dari 10 sampai 100. Konfigurasi optimal yang diperoleh terdiri dari learning rate 0.007, dropout 0.2, batch size 32, dan 60 epoch, yang menghasilkan akurasi 99 persen. Meskipun demikian, analisis evaluasi menunjukkan adanya kesalahan minor pada kelas Hawar Muda yang memiliki precision sedikit lebih rendah akibat tumpang tindih karakteristik visual. Hasil penelitian ini menetapkan tolok ukur baru bagi klasifikasi penyakit daun kentang sekaligus menyediakan kerangka optimasi yang dapat direplikasi. Penelitian ini memberikan kontribusi metodologis dan praktis bagi pengembangan sistem klasifikasi penyakit tanaman yang presisi dan efisien dalam konteks pertanian cerdas.

Kata Kunci: Klasifikasi; CNN; Hyperparameter; MobileNetV2; Kentang

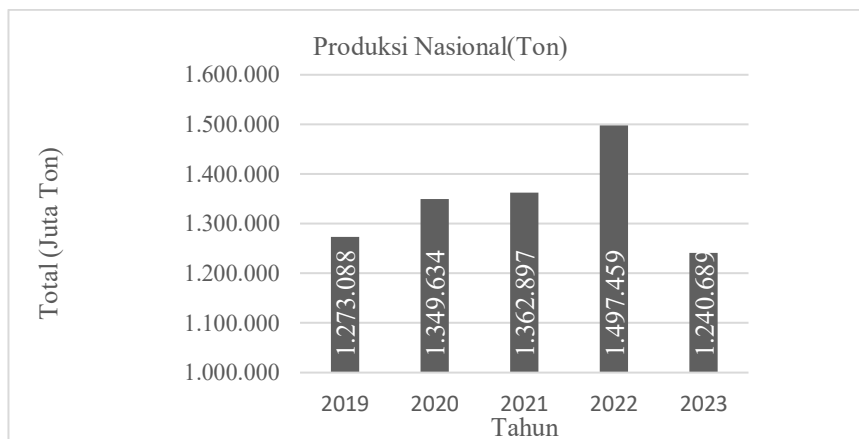
Abstract—Indonesia's national potato production faces significant threats from leaf diseases, while manual classification remains slow, subjective, and prone to error due to the high visual similarity across disease categories. This highlights the need for a precise and reliable automated classification system. However, many previous studies have not applied systematic hyperparameter optimization, leaving the capacity of deep learning architectures underutilized. Addressing this research gap, this study aims to enhance the performance of MobileNetV2 for potato leaf disease classification through a structured hyperparameter optimization process. A Sequential Search strategy validated through 3 fold Stratified Cross Validation is employed to obtain stable performance estimates. Four key hyperparameters are examined: learning rate from 0.001 to 0.009, dropout from 0.1 to 0.9, batch size from 8 to 192, and epochs from 10 to 100. The optimal configuration consists of a learning rate of 0.007, dropout of 0.2, batch size of 32, and 60 epochs, which enables MobileNetV2 to achieve an accuracy of 99 percent. Despite this strong performance, evaluation results reveal a minor limitation in the Young Blight class, where precision is slightly lower due to overlapping visual characteristics. These findings establish a new benchmark for potato leaf disease classification and provide a reproducible optimization framework for future studies. The study offers both methodological and practical contributions to the development of precise and efficient plant disease classification systems within the context of smart agriculture.

Keywords: Classification; CNN; Hyperparameter; MobileNetV2; Potato

1. PENDAHULUAN

Kentang adalah salah satu bahan pertanian terpenting di Indonesia, memiliki nilai ekonomi yang tinggi dan merupakan sumber makanan utama bagi masyarakat [1]. Kentang juga salah satu bahan makanan yang mempunyai kandungan gizi yang tinggi [2]. Namun, produktivitas tanaman kentang sering berkurang karena penyakit yang menyerang daun, seperti bintik-bintik kering, busuk daun, dan berbagai jenis penyakit lainnya [3]. Kerentanan tanaman ini terhadap penyakit dapat dilihat dari data produksi nasional yang dikeluarkan oleh Badan Pusat Statistik (BPS), sebagaimana disajikan pada Gambar 1 [4]. Data yang tertera pada Gambar 1. menunjukkan sebuah tren yang mengkhawatirkan. Setelah mengalami pertumbuhan kuat hingga tahun 2022, produksi kentang nasional mengalami penurunan secara drastis sebesar 17,15% pada tahun 2023. Penurunan tajam ini menggarisbawahi urgensi pencegahan gagal panen, dikutip dari [5]. Penyakit hawar daun pada kentang menjadi permasalahan utama, penyakit hawar daun dilaporkan dapat menurunkan hasil panen antara 20 hingga 25%, bahkan pada kondisi berat dapat menyebabkan gagal panen total. Terjadinya permasalahan tersebut menandakan adanya permasalahan serius dalam pengelolaan pertanian, khususnya pada aspek pengendalian penyakit tanaman kentang, ketergantungan petani terhadap metode konvensional dalam mendeteksi penyakit menjadikan proses pengendalian menjadi lambat dan tidak efektif. Hal tersebut menjadikan penyakit tumbuhan mudah menyebar sebelum dilakukan penanganan yang tepat.

Data yang ditampilkan pada Gambar 1 menunjukkan bahwa sistem deteksi penyakit tanaman yang digunakan saat ini belum mampu memberikan hasil diagnosis yang cepat serta akurat [4]. Sebagian besar petani dan tenaga ahli masih melakukan pengamatan secara langsung di lapangan untuk mengenali gejala penyakit [6]. Pendekatan konvensional ini memiliki keterbatasan fundamental karena membutuhkan waktu, energi, dan keahlian diagnostik yang tidak selalu tersedia di lapangan.



Gambar 1. Data Produksi Kentang Nasional.

Selain itu, metode ini sangat dipengaruhi oleh subjektivitas manusia, yang dapat menyebabkan kesalahan identifikasi[7]. Keterlambatan dalam mengenali jenis penyakit juga memberi kesempatan bagi penyakit menyebar luas yang pada akhirnya memperburuk tingkat kerusakan tanaman dan menurunkan hasil panen. Masalah identifikasi yang tidak akurat ini juga menimbulkan dampak berkelanjutan terhadap lingkungan dan perekonomian. Kesalahan identifikasi dapat menyebabkan penerapan pestisida yang salah, yang tidak hanya menjadi pemborosan sumber daya seperti biaya, waktu, dan tenaga kerja, tetapi juga berisiko merusak kesuburan tanah dan ekosistem di sekitar lahan[8]. Situasi ini menunjukkan adanya urgensi di sektor pengendalian penyakit sehingga perlunya perubahan pendekatan dalam sistem pengendalian penyakit tanaman agar proses diagnosis dapat dilakukan dengan cepat, objektif dan berkelanjutan, kebutuhan tersebut dapat dijawab oleh teknologi AI berbasis pemrosesan citra. dengan dukungan teknik pembelajaran mesin dan *deep learning* yang memungkinkan proses deteksi penyakit tanaman dilakukan secara otomatis melalui analisis gambar daun[9], [10]. Sehingga hasil diagnosis menjadi lebih konsisten dan tidak bergantung pada persepsi subjektif pengamat. Untuk mewujudkan proses analisis citra yang andal tersebut, diperlukan algoritma yang mampu mengekstraksi fitur visual secara efektif dari gambar tanaman.

Salah satu model yang terbukti unggul dalam bidang analisis citra adalah Convolutional Neural Network (CNN). Dengan kemampuan CNN dalam mengekstrak fitur visual dari gambar secara otomatis. Menjadikannya sangat sesuai untuk mendeteksi gejala penyakit pada daun tanaman kentang. Melalui lapisan-lapisan konvolusinya, CNN dapat mengenali perbedaan tekstur, warna, dan bentuk yang menjadi indikator munculnya penyakit pada tanaman. Di antara berbagai arsitektur CNN yang telah dikembangkan, Arsitektur MobileNet, khususnya, telah menunjukkan potensi besar dalam domain ini. Studi yang telah dilakukan pada [11] menunjukkan efektivitas arsitektur MobileNetV2 dalam kasus pengklasifikasian penyakit tanaman melalui analisis gambar daun dengan mencapai akurasi sebesar 85%. Penelitian lainnya yang telah dilakukan juga memperkuat pernyataan bahwa arsitektur MobileNetV2 merupakan pilihan yang efektif untuk klasifikasi penyakit daun kentang. Penelitian oleh Gunawan et al. Yang menerapkan model serupa dalam kasus klasifikasi penyakit tanaman kentang berhasil mencapai akurasi sebesar 95,31%[12]. Hasil ini menetapkan state-of-the-art (SOTA) yang solid dan membuktikan kelayakan arsitektur MobileNetV2 untuk tugas ini. Namun, analisis lebih mendalam terhadap penelitian[12] menunjukkan bahwa pencapaian ini diraih menggunakan satu set konfigurasi hyperparameter yang tetap, tanpa melalui proses pencarian atau optimalisasi sistematis. Hal ini menyisakan sebuah celah penelitian yang krusial performa yang dilaporkan, meskipun tinggi, kemungkinan besar belum mencapai potensi maksimal dari arsitektur tersebut. Oleh karena itu, kontribusi utama dari penelitian ini adalah untuk mengisi celah tersebut secara langsung. Tidak hanya bertujuan untuk mereplikasi keberhasilan sebelumnya, tetapi untuk melampauinya dengan melakukan analisis mendalam dan sistematis guna menemukan konfigurasi hyperparameter terbaik. Dengan demikian, penelitian ini memilih arsitektur MobileNet karena rekam jeaknya yang terbukti, dengan tujuan spesifik untuk memaksimalkan potensinya melalui optimalisasi proses pelatihan [13].

Upaya untuk memaksimalkan potensi MobileNetV2 bergantung pada pemilihan hyperparameter yang tepat, karena komponen seperti learning rate, dropout rate, batch size, dan jumlah epoch memiliki peran penting dalam mencegah overfitting sekaligus meningkatkan akurasi model. Penelitian oleh [14] menunjukkan bahwa pengaturan parameter pelatihan, termasuk jumlah epoch dan proses fine-tuning pada arsitektur CNN, secara signifikan memengaruhi akurasi serta kemampuan model dalam mencapai keseimbangan antara performa pelatihan dan validasi. Pemilihan hyperparameter yang tidak optimal dapat menyebabkan model gagal mencapai akurasi terbaiknya, mengalami konvergensi lambat, atau menghasilkan performa yang kurang stabil. Oleh karena itu, proses pencarian dan penyesuaian hyperparameter menjadi tahap krusial dalam pengembangan model deep learning yang andal [15], [16]. Sejalan dengan itu, tren penelitian terbaru menunjukkan bahwa hyperparameter tuning merupakan topik yang terus berkembang dan berperan besar dalam peningkatan performa model machine learning [17]. Berangkat dari temuan tersebut, penelitian ini difokuskan untuk melakukan analisis dan menentukan konfigurasi hyperparameter yang optimal pada model MobileNetV2 dalam tugas klasifikasi penyakit daun kentang sebagai salah satu komoditas

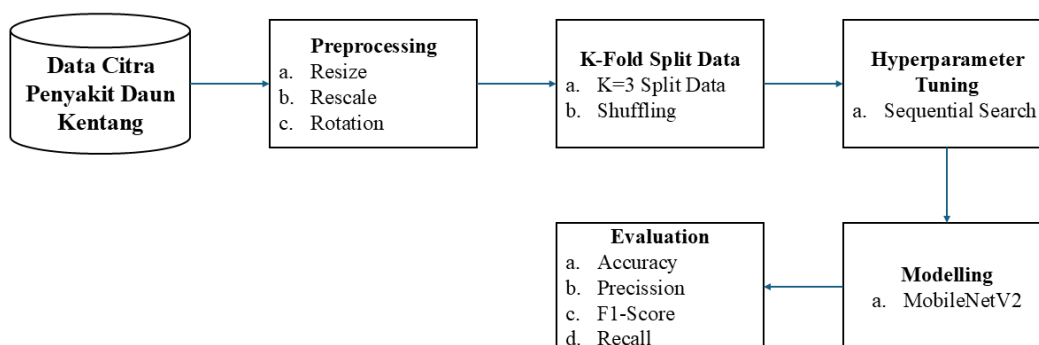
penting yang rentan terhadap serangan penyakit. Penelitian ini tidak hanya menawarkan pendekatan deteksi berbasis citra, tetapi juga memberikan kontribusi berupa konfigurasi hyperparameter yang dapat direplikasi untuk meningkatkan efisiensi dan akurasi model pada komoditas serupa [18]. Proses dilakukan secara sistematis dengan menguji pengaruh variasi nilai learning rate, dropout rate, batch size, dan jumlah epoch terhadap akurasi validasi model [19]. sehingga kombinasi hyperparameter yang paling efektif dapat ditentukan berdasarkan evaluasi metrik akurasi selama pelatihan

Secara akademik, penelitian ini bertujuan menghasilkan konfigurasi hyperparameter yang optimal pada arsitektur MobileNetV2 melalui proses tuning yang terstruktur dan terukur, sehingga dapat memberikan bukti empiris mengenai pengaruh langsung variasi hyperparameter terhadap performa model dalam tugas klasifikasi penyakit daun. Temuan ini diharapkan dapat mengisi kekosongan penelitian sebelumnya yang belum melakukan proses tuning secara sistematis, sekaligus menyediakan referensi metodologis yang dapat direplikasi untuk pengembangan model deep learning pada domain pertanian maupun domain citra lainnya. Secara praktis atau pada tingkat masyarakat, penelitian ini ditujukan untuk menghasilkan model klasifikasi penyakit daun kentang yang lebih akurat, stabil, dan mudah diimplementasikan sehingga dapat membantu petani dalam mengidentifikasi gejala penyakit secara lebih cepat dan tepat. Dengan adanya model yang lebih akurat, proses pengambilan keputusan dalam pengendalian penyakit diharapkan menjadi lebih efisien, sehingga dapat menurunkan risiko gagal panen, mengurangi kerugian ekonomi, serta mendukung penerapan teknologi pertanian cerdas di lapangan.

Untuk memastikan penyajian penelitian ini tersusun secara sistematis dan terstruktur, laporan ini dibagi ke dalam beberapa bab utama. Bagian 2, Metodologi Penelitian, menjelaskan secara rinci tahapan pelaksanaan penelitian, mulai dari persiapan dataset, perancangan arsitektur model, penyusunan skenario eksperimen untuk tuning hyperparameter, hingga metode evaluasi kinerja yang digunakan. Bagian 3, Hasil dan Pembahasan, memaparkan hasil eksperimen yang diperoleh dan memberikan analisis mendalam terkait pengaruh masing-masing hyperparameter terhadap performa model. Bagian 4, Kesimpulan, menyajikan ringkasan temuan utama, menegaskan kontribusi penelitian ini.

2. METODOLOGI PENELITIAN

Metodologi penelitian ini dirancang untuk melakukan optimalisasi hyperparameter secara sistematis pada arsitektur MobileNetV2 untuk tugas klasifikasi penyakit daun kentang. Pendekatan yang digunakan adalah optimalisasi sekuensial yang divalidasi menggunakan Stratified K-Fold Cross-Validation untuk memastikan hasil yang andal. Alur kerja penelitian ini mengikuti metodologi standar dalam pengembangan penelitian machine learning [20]. Pendekatan ini menekankan pentingnya proses sistematis agar hasil penelitian valid dan dapat direplikasi. Setiap tahap memiliki peran berurutan, di mana hasil dari satu tahap menjadi masukan bagi tahap berikutnya. Sebagaimana diilustrasikan pada Gambar 2.















Gambar 2. Alur Kerja Penelitian

2.1. Persiapan Dataset dan Pra-Pemrosesan Data

Tahap ini adalah persiapan dataset, yang merupakan fondasi penelitian. Dataset yang digunakan merupakan dataset publik yang bersumber dari Kaggle dengan nama *potato leaf disease*. Dataset ini memiliki Jumlah 900 citra daun kentang dengan tiga kelas berbeda, yaitu normal, hawar muda, dan hawar tua yang dapat dilihat pada Tabel 1. Pada tahap ini juga terjadi pengolahan data sehingga menjadi berkualitas dan representasi data secara langsung mempengaruhi kemampuan model untuk belajar serta melakukan generalisasi [21]. Sebelum diumpungkan ke arsitektur jaringan saraf, Seluruh citra dataset melalui tahapan pra-pemrosesan yang meliputi penyesuaian dimensi. Dengan demikian, Pada tahap ini dilakukan standarisasi ukuran gambar, normalisasi nilai piksel, dan augmentasi data untuk meningkatkan variasi dataset dan mencegah overfitting [22]. Dataset citra diubah ukurannya menjadi 128×128 piksel hal ini guna menstandarkan dimensi dan menjaga efisiensi komputasi tanpa mengorbankan detail visual yang penting. nilai intensitas piksel dinormalisasi dari rentang 0 hingga 255 menjadi 0, 1 langkah ini merupakan bagian penting karena model jaringan saraf sangat sensitif terhadap skala data, tanpa normalisasi ,gradien menjadi tidak stabil dan

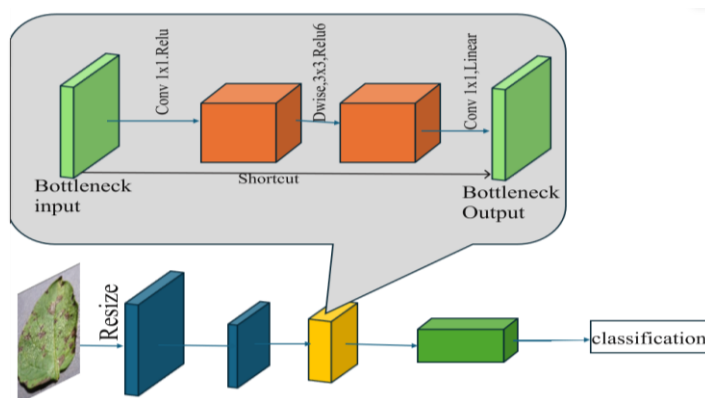
proses pelatihan menjadi kurang optimal. Penelitian ini juga menerapkan augmentasi data untuk memperluas keragaman sampel tanpa menambah jumlah data asli. Augmentasi dilakukan secara on-the-fly selama pelatihan dan mencakup rotasi acak hingga 20 derajat, translasi horizontal maupun vertikal hingga 20% dari dimensi citra, pembesaran hingga 20%, serta pembalikan horizontal. Kombinasi transformasi ini menghasilkan variasi yang kaya dari pola daun sekaligus membantu mengurangi risiko overfitting. Melalui rangkaian proses ini, dataset yang semula hanya berisi citra mentah berhasil dipersiapkan menjadi data terstandarisasi dengan variasi yang memadai, sehingga model dapat belajar dengan lebih efektif ketika memasuki tahap pelatihan

Tabel 1. Dataset penyakit Daun Kentang

| Label | Jumlah Dataset | Gambar 1 | Gambar 2 | Gambar 3 | Gambar 4 |
|------------|----------------|--|---|--|--|
| Normal | 300 |  |  |  |  |
| Hawar Muda | 300 |  |  |  |  |
| Hawar Tua | 300 |  |  |  |  |

2.2. Model MobileNetV2

Penelitian ini mengimplementasikan arsitektur MobileNetV2, sebuah jaringan saraf konvolusional yang dikenal luas karena efisiensi komputasinya yang tinggi tanpa mengorbankan performa akurasi secara signifikan. Arsitektur ini sangat relevan untuk aplikasi praktis seperti *smart farming* karena strukturnya yang ringan, yang dimungkinkan oleh blok *inverted residuals* dan *linear bottlenecks* untuk mempertahankan representasi fitur penting dengan jumlah parameter yang minimal[13]. Setiap blok utama terdiri atas lapisan ekspansi, *depthwise convolution*, dan lapisan linear, dengan fungsi aktivasi ReLU6 yang bekerja dalam rentang nilai terbatas. Gambar 3. berikut menyajikan alur kerja internal MobileNetV2 yang terdiri atas tahapan ekspansi, *depthwise convolution*, dan proyeksi linear. Struktur blok yang ditampilkan inilah yang menjadi komponen inti dalam proses ekstraksi fitur pada model.



Gambar 3. Asitektur MobileNetV2

Proses ekstraksi fitur pada setiap blok dilakukan menggunakan fungsi aktivasi ReLU6, yang membatasi rentang nilai antara 0 hingga 6 untuk menjaga stabilitas numerik, khususnya pada perangkat dengan presisi rendah. Fungsi aktivasi tersebut didefinisikan sebagai

$$f(x) = (\max(0, x), 6). \tag{1}$$

Di mana $f(x)$ merupakan keluaran fungsi terhadap masukan x . Karakter \max memilih nilai terbesar antara 0 dan hasil $\min(x, 6)$, sedangkan \min memilih nilai terkecil antara x dan 6. Dengan susunan karakter ini, nilai negatif dipetakan menjadi 0, nilai lebih besar dari 6 dipangkas menjadi 6, dan nilai di antaranya dipertahankan, sehingga stabilitas komputasi tetap terjaga.

Untuk meningkatkan efisiensi pelatihan, digunakan pendekatan transfer learning, di mana bobot awal MobileNetV2 yang telah dilatih pada dataset ImageNet digunakan sebagai inisialisasi. Pendekatan ini memungkinkan model memanfaatkan fitur dasar yang telah dipelajari sebelumnya sehingga proses konvergensi lebih cepat dan generalisasi meningkat pada dataset berukuran lebih kecil [9]. Lapisan-lapisan konvolusional dari model dasar ini dibekukan, dan sebuah kepala klasifikasi kustom dibangun untuk diadaptasikan pada tugas spesifik penelitian ini menggunakan *Global Average Pooling*, *Dropout*, dan lapisan *Dense* dengan aktivasi *Softmax*. Probabilitas keluaran kelas dihitung dengan

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \tag{2}$$

Simbol $\sigma(z_i)$ menunjukkan probabilitas kelas ke- i , sementara e^{z_i} adalah eksponensial dari logit. Garis pecahan membagi nilai eksponensial kelas i dengan total eksponensial semua kelas. Simbol \sum dan indeks $j = 1$ sampai K menandakan proses penjumlahan seluruh kelas untuk normalisasi sehingga keluaran membentuk distribusi probabilitas. Secara menyeluruh, desain model ini mengikuti alur pemrosesan berjenjang, dimulai dari tahap ekstraksi fitur oleh lapisan MobileNetV2 yang telah dibekukan, dilanjutkan dengan pereduksian fitur oleh *Global Average Pooling*, dan diikuti oleh *Dropout* untuk mengurangi risiko *overfitting*. Tahap akhir berupa lapisan *Dense Softmax* menghasilkan prediksi tiga kelas, yaitu Normal, Hawar Muda, dan Hawar Tua. Dengan struktur seperti ini, MobileNetV2 menangani proses ekstraksi fitur visual secara mendalam, sementara *classification head* berperan dalam pengambilan keputusan akhir yang sesuai dengan kebutuhan penelitian ini.

2.3. Hyperparameter Tuning

Untuk memperoleh estimasi kinerja model yang akurat serta mengurangi potensi bias dari pembagian data tunggal, penelitian ini menggunakan metode Validasi Silang Stratifikasi K-Fold dengan $K=3$. Dengan pendekatan ini, dataset akan dibagi menjadi tiga bagian (lipatan) yang menjaga proporsi kelas, model kemudian dilatih dan divalidasi dan model akan dilatih dan divalidasi dengan tiga pengulangan, dan performa akhir dihitung sebagai rata-rata nilai ketiga fold menggunakan

$$Acc_{avg} = \frac{1}{3} \sum_{i=1}^3 Acc_i. \tag{3}$$

Karakter Acc_{avg} menunjukkan akurasi keseluruhan, angka 3 pada penyebut menyatakan jumlah fold, dan simbol \sum menjumlahkan akurasi dari fold pertama hingga ketiga. Susunan ini memastikan bahwa performa model direpresentasikan melalui rata-rata dari tiga pengujian.

Optimasi hyperparameter dilakukan menggunakan *sequential search*, yaitu proses pencarian yang memfokuskan evaluasi pada satu hyperparameter pada satu tahap, sementara parameter lainnya dipertahankan pada nilai terbaik sebelumnya. Ruang pencarian mencakup variasi pada *learning rate*, *dropout rate*, *batch size*, dan jumlah epoch. Dengan rentang parameter yang ditunjukkan pada Tabel 2.

Tabel 2. Variasi Nilai Parameter

| Parameter | Variasi Nilai Parameter |
|---------------|--------------------------------|
| Learning Rate | 0,001-0,009 |
| Dropout | 0,1-0,9 |
| Batch Size | 8,16,24,32,48,64,80,96,128,192 |
| Epoch | 10,20,30,40,50,60,70,80,90,100 |

Pada tahap pelatihan model, lapisan *Dropout* juga berperan penting sebagai mekanisme regularisasi yang dirancang untuk mengurangi risiko *overfitting*. Melalui pendekatan ini, sebagian neuron dinonaktifkan secara acak pada setiap iterasi pelatihan sehingga model tidak hanya bergantung pada pola tertentu, tetapi mampu mempelajari representasi fitur yang lebih umum dan robust terhadap variasi data. Secara matematis dropout dinyatakan sebagai

$$y_i = r_i \cdot f(x_i), \tag{4}$$

Rumus *Dropout* $y_i = r_i \cdot f(x_i)$ menunjukkan keluaran neuron ke- i setelah regularisasi. Simbol r_i adalah variabel biner acak, bernilai 0 atau 1, dan tanda “.” menyatakan perkalian antara nilai acak tersebut dengan aktivasi neuron $f(x_i)$. Jika $r_i = 0$, neuron dinonaktifkan; jika $r_i = 1$, neuron tetap aktif, sehingga secara keseluruhan mekanisme ini mengurangi *overfitting*. Dengan mematikan sebagian neuron secara acak pada setiap iterasi pelatihan, dropout membantu mengurangi *overfitting* dan mendorong model mempelajari representasi yang lebih robust.

2.4. Evaluasi

Proses pelatihan menggunakan algoritma optimalisasi Adam, sebuah metode adaptif yang tetap menjadi pilihan utama banyak penelitian terkini karena kemampuan stabil dalam mengombinasikan momentum dan penyesuaian laju pembelajaran secara otomatis. Berbagai studi terbaru, seperti yang disampaikan oleh [23], menegaskan bahwa Adam masih menjadi algoritma yang paling efisien dan konsisten untuk tugas klasifikasi citra berbasis deep learning,

terutama pada dataset berukuran kecil hingga menengah. Pendekatan ini memungkinkan proses pembaruan bobot berlangsung secara lebih halus dan terarah selama pelatihan, dengan persamaan matematika sebagai berikut.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \tag{5}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \tag{6}$$

$$\widehat{m}_t = \frac{m_t}{(1 - \beta_1^t)} \tag{7}$$

$$\widehat{v}_t = \frac{v_t}{(1 - \beta_2^t)} \tag{8}$$

$$\theta_t + 1 = \theta_t - \alpha \frac{\widehat{m}_t}{\sqrt{\widehat{v}_t + \epsilon}} \tag{9}$$

Rumus (5) digunakan untuk menghitung momen pertama yang berfungsi menghaluskan arah gradien sehingga proses pembaruan bobot lebih stabil. Rumus (6) menghitung momen kedua yang merekam besarnya gradien agar langkah pembaruan tidak melonjak terlalu besar. Selanjutnya, Rumus (7) dan Rumus (8) melakukan koreksi bias pada kedua momen tersebut sehingga nilainya lebih akurat pada tahap awal pelatihan. Tahap akhir menggunakan kedua momen yang telah dikoreksi untuk memperbarui bobot secara adaptif melalui Rumus (9), sehingga optimisasi berlangsung lebih efisien dan stabil.

Model dilatih menggunakan fungsi kerugian Categorical Cross-Entropy, yang mengukur divergensi antara distribusi prediksi dan label sebenarnya. Dengan K=3kelas, yaitu Normal, Hawar Daun Muda, dan Hawar Daun Tua. Mekanisme Early Stopping diterapkan untuk menghentikan pelatihan apabila tidak terjadi peningkatan akurasi validasi selama delapan epoch berturut-turut. kinerja model dievaluasi menggunakan confusion matrix. Metode ini memberikan gambaran detail mengenai bagaimana model mengenali setiap kategori daun. Untuk kelas Normal, terlihat berapa banyak citra yang diprediksi dengan benar sebagai Normal dan berapa yang salah dikenali sebagai Hawar Muda atau Hawar Tua. Hal yang sama juga tampak untuk kelas Hawar Muda, yaitu jumlah citra yang sesuai dengan kondisi hawar tahap awal serta jumlah yang salah diarahkan ke kelas Normal atau Hawar Tua. Pada kelas Hawar Tua, matrix menunjukkan citra yang berhasil diidentifikasi secara tepat maupun yang masih keliru diprediksi sebagai kategori lain. Berdasarkan informasi tersebut dihitung tiga metrik evaluasi utama, yaitu precision, recall, dan F1 score. Precision menunjukkan ketepatan model dalam memberikan label Normal, Hawar Muda, atau Hawar Tua. Recall menggambarkan kemampuan model dalam menemukan seluruh citra yang benar pada tiap kategori. F1 score digunakan untuk menilai keseimbangan antara ketepatan dan kelengkapan prediksi. Penelitian ini menggunakan pendekatan Macro Average, yaitu menghitung precision, recall, dan F1 score pada masing masing kelas kemudian merata-ratakannya. Dengan pendekatan ini setiap kategori daun memperoleh penilaian yang setara sehingga hasil evaluasi tidak dipengaruhi oleh jumlah sampel pada setiap kelas. Pendekatan ini memastikan model dinilai secara adil pada kategori Normal, Hawar Muda, dan Hawar Tua.

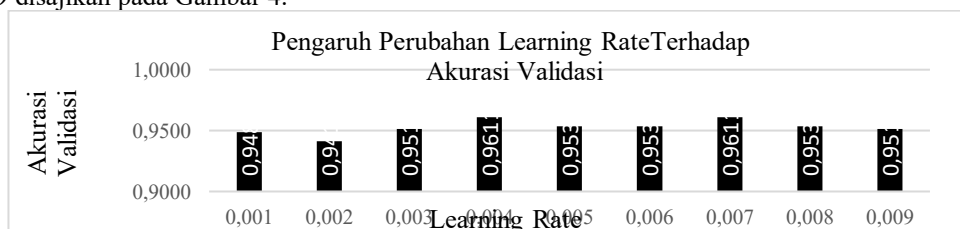
3. HASIL DAN PEMBAHASAN

Dari serangkaian eksperimen optimalisasi hyperparameter yang telah dilakukan pada arsitektur MobileNetV2, akan dilakukan analisis hasil dan Pembahasan akan difokuskan pada analisis komparatif dan kualitatif terhadap pengaruh setiap hyperparameter, evaluasi performa model final secara mendalam, serta diskusi mengenai implikasi, kontribusi, dan keterbatasan dari temuan penelitian ini.

3.1. Hasil Eksperimen Tuning Hyperparameter

Proses pencarian hyperparameter optimal dilakukan secara sekuensial dan sistematis menggunakan metode 3-Fold Stratified Cross-Validation. Metodologi ini memastikan bahwa hasil yang diperoleh bersifat robust dan tidak bias terhadap satu pembagian data tertentu. Pengaruh setiap hyperparameter learning rate, dropout rate, batch size, dan jumlah epoch erhadap rata-rata akurasi validasi dianalisis secara terpisah dan divisualisasikan untuk interpretasi yang lebih jelas.

Eksperimen pertama bertujuan untuk menemukan learning rate yang paling efektif, yang merupakan salah satu hyperparameter paling krusial dalam pelatihan model deep learning. Hasil pengujian pada rentang nilai dari 0.001 hingga 0.009 disajikan pada Gambar 4.

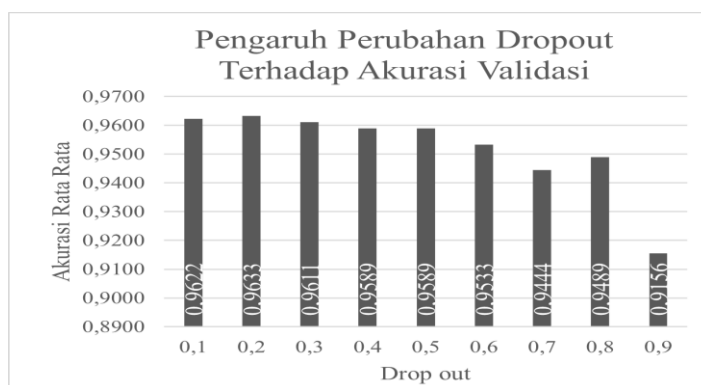


Gambar 4. Pengaruh Perubahan Learning Rate terhadap Akurasi Validasi

Hasil analisis grafik pengaruh perubahan Learning Rate yang ditampilkan pada Gambar 4. menunjukkan bahwa performa model relatif stabil pada rentang LR 0,001 hingga 0,009, dengan akurasi validasi pada kisaran 0,94 hingga 0,96. Dari seluruh nilai yang diuji, performa terbaik dicapai pada LR 0,007, yang menghasilkan akurasi validasi tertinggi sebesar 0,9611. Berdasarkan analisis statistik stabilitas ini terjadi karena perubahan nilai learning rate tidak menimbulkan variasi gradien yang signifikan selama proses pembaruan bobot, sehingga akurasi model cenderung berada dalam rentang performa yang serupa. Berdasarkan analisis statistik deskriptif, nilai rata-rata akurasi sebesar 0,9527 mengindikasikan bahwa sebagian besar konfigurasi menghasilkan performa yang mendekati nilai pusat tersebut. Hal ini terjadi karena akurasi hasil pelatihan tidak mengalami penyimpangan besar akibat perubahan parameter, sehingga nilai pusat distribusi menjadi representatif untuk keseluruhan pengamatan. Keselarasan antara nilai median dan modus, masing-masing sebesar 0,9533, menunjukkan bahwa sebagian besar nilai akurasi berkumpul pada titik tersebut. Kondisi ini muncul karena distribusi akurasi memiliki konsistensi internal yang tinggi, sehingga tidak terdapat nilai ekstrem yang dapat menyebabkan pergeseran pusat distribusi.

Standar deviasi sebesar 0,0061 mencerminkan bahwa penyebaran akurasi antar percobaan sangat rendah. Penyebaran yang rendah ini disebabkan oleh tingkat sensitivitas model terhadap perubahan learning rate yang minimal pada rentang nilai yang diuji, sehingga variasi parameter tidak menghasilkan perubahan performa yang mencolok. Rentang akurasi sebesar 0,0200, dengan nilai minimum 0,9411 dan maksimum 0,9611, memperlihatkan bahwa jarak antara performa terendah dan tertinggi sangat kecil. Hal ini terjadi karena batas bawah dan batas atas performa sama-sama berada dalam rentang yang sempit, yang menandakan bahwa performa model relatif terkontrol dan tidak terpengaruh secara drastis oleh variasi nilai learning rate tersebut. Nilai skewness sebesar $-0,3156$ menunjukkan bahwa distribusi akurasi sedikit miring ke kiri. Kemiringan ini terjadi karena sebagian besar nilai akurasi berkumpul di sisi kanan atau sedikit lebih tinggi daripada rata-rata, sehingga sisi kiri distribusi menjadi lebih panjang. Fenomena tersebut menunjukkan bahwa mayoritas percobaan cenderung menghasilkan akurasi yang optimal dalam rentang pengujian. Sementara itu, nilai kurtosis sebesar 0,9443 menunjukkan karakteristik distribusi platykurtic, yaitu distribusi dengan puncak yang lebih datar dan penyebaran yang lebih merata dibandingkan distribusi normal. Kondisi ini terjadi karena nilai akurasi tidak terkonsentrasi secara tajam pada satu titik, tetapi tersebar secara relatif merata dalam rentang sempit. Jumlah observasi sebanyak sembilan sesuai dengan jumlah variasi learning rate yang diuji, yang memberikan gambaran kuantitatif yang memadai mengenai pola distribusi performa model dalam eksperimen ini.

Dengan menetapkan learning rate 0,007 sebagai nilai paling optimal berdasarkan stabilitas performa dan pola distribusi statistik yang diperoleh, tahap eksperimen selanjutnya difokuskan pada pencarian dropout rate terbaik sebagai mekanisme regularisasi dalam mencegah overfitting. Eksperimen ini dirancang untuk mengidentifikasi nilai dropout yang mampu menjaga keseimbangan antara kemampuan model dalam menangkap fitur penting dari data dan kemampuan generalisasi pada data yang belum pernah dilihat. Melalui variasi dropout yang berbeda, dapat diamati bagaimana tingkat penghapusan neuron secara acak memengaruhi akurasi pada data validasi, termasuk apakah model menjadi lebih stabil atau justru rentan mengalami fluktuasi performa. Proses evaluasi ini juga memungkinkan peneliti untuk memahami sejauh mana dropout dapat mengendalikan kompleksitas model ketika berhadapan dengan pola visual penyakit daun yang memiliki kemiripan antar kelas. Selain itu, pengamatan terhadap perubahan tren akurasi pada tiap variasi memberikan gambaran lebih jelas mengenai sensitivitas model terhadap tingkat regularisasi yang diterapkan. Hasil pengujian ini memberikan dasar yang kuat dalam menentukan konfigurasi dropout yang optimal, sehingga model mampu mempertahankan akurasi validasi yang tinggi sekaligus menghindari kecenderungan untuk menghafal data pelatihan. Ringkasan hasil eksperimen dropout tersebut disajikan pada Gambar 5.

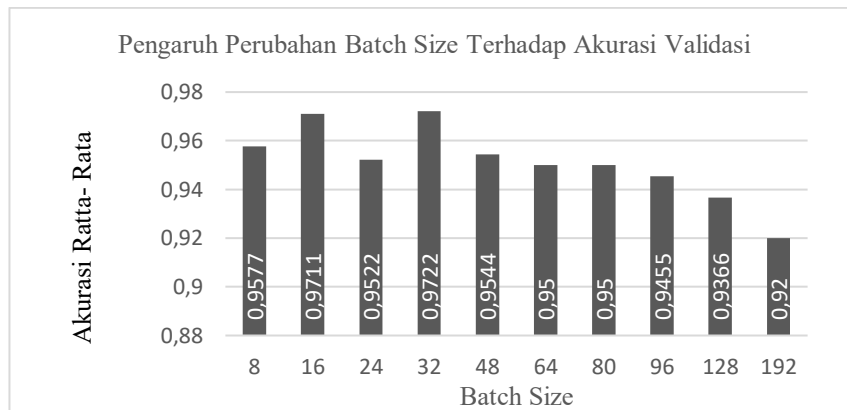


Gambar 5. Pengaruh Perubahan Dropout Rate terhadap Akurasi Validasi

Gambar 5. Menunjukkan hasil pengujian terhadap sembilan tingkat dropout menunjukkan variasi performa yang lebih lebar dibandingkan eksperimen pada learning rate, dengan akurasi berada pada rentang 0,9156 hingga 0,9633. Rentang ini mencerminkan bahwa mekanisme dropout yang menonaktifkan sejumlah neuron pada setiap iterasi secara langsung memengaruhi stabilitas representasi fitur, terutama pada nilai dropout yang lebih tinggi. Nilai rata-rata akurasi sebesar 0,9519 dan median 0,9589 menunjukkan bahwa meskipun performa umum model tetap

tinggi, terdapat penurunan akurasi pada dropout besar yang menekan nilai rata-rata. Standar deviasi sebesar 0,0150 dan rentang akurasi 0,0478 mengindikasikan penyebaran performa yang lebih lebar akibat degradasi akurasi yang cukup signifikan pada tingkat dropout tinggi. Karakteristik distribusi memperkuat temuan ini. Nilai skewness $-2,0884$ menunjukkan kemiringan yang kuat ke kiri akibat munculnya beberapa nilai akurasi rendah yang menarik distribusi menjauhi pusat. Sementara itu, nilai kurtosis $4,7689$ menandakan distribusi leptokurtic, yaitu tingginya konsentrasi nilai akurasi di sekitar area performa menengah hingga tinggi dengan keberadaan beberapa nilai ekstrem rendah. Keseluruhan pola distribusi ini menegaskan bahwa dropout memiliki pengaruh yang lebih fluktuatif terhadap performa model dibandingkan hyperparameter lainnya, terutama karena nilai dropout besar berpotensi mengurangi kapasitas representasi jaringan secara signifikan.

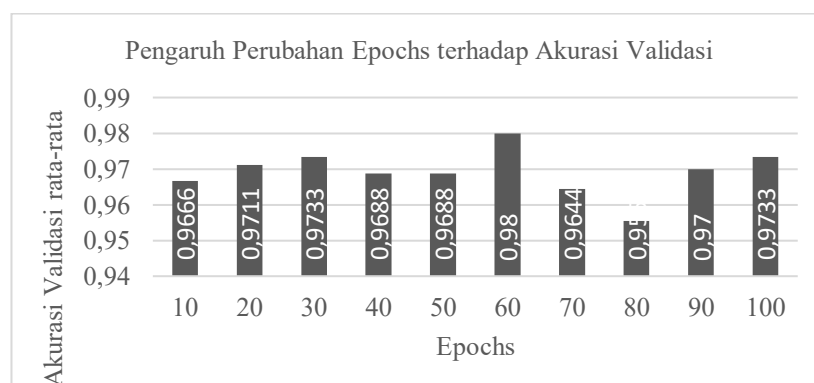
Tahap pencarian berikutnya adalah menguji pengaruh batch size terhadap performa model. Hasilnya, seperti terlihat pada Gambar 6. menunjukkan pola yang menarik. perubahan batch size tidak hanya memengaruhi kecepatan konvergensi, tetapi juga berperan dalam menentukan konsistensi performa model pada data validasi



Gambar 6. Pengaruh Perubahan Batch Size terhadap Akurasi Validasi

Analisis performa model pada berbagai konfigurasi batch size yang tertera pada Gambar 6. menunjukkan adanya variasi akurasi yang cukup jelas antar setiap pengujian. Nilai rata-rata akurasi sebesar 0,9510 dan median 0,9511 menunjukkan bahwa model mampu mempertahankan tingkat performa yang relatif stabil pada sebagian besar nilai batch size yang diuji. Perbedaan antara nilai minimum 0,9200 dan maksimum 0,9722 dengan rentang 0,0522 menggambarkan bahwa perubahan ukuran batch memberikan pengaruh nyata terhadap dinamika pembelajaran. Kondisi ini didukung oleh nilai simpangan baku sebesar 0,0153 yang menandakan fluktuasi performa yang tetap berada dalam batas yang dapat diterima dalam eksperimen pembelajaran mendalam. Nilai modus sebesar 0,9500 mengindikasikan bahwa terdapat kecenderungan nilai akurasi yang muncul lebih sering dibandingkan nilai lainnya, sehingga menjadi titik pusat kecenderungan distribusi akurasi. Karakteristik distribusi statistik akurasi memperjelas pola hubungan antara variasi batch size dan performa model. Nilai skewness sebesar $-0,5738$ menunjukkan bahwa distribusi akurasi cenderung condong ke kiri sehingga sebagian besar nilai akurasi berada sedikit di atas rata-rata. Sementara itu, nilai kurtosis sebesar 0,9518 mengindikasikan bahwa distribusi tidak memiliki puncak yang sangat tinggi maupun penyebaran ekstrem sehingga tidak terdapat nilai yang bersifat menyimpang secara signifikan dari keseluruhan pengamatan. Jumlah total akurasi sebesar 9,5100 dari sepuluh percobaan menegaskan konsistensi performa model sepanjang pengujian. Berdasarkan pola ini, dapat dijelaskan bahwa perubahan batch size memberikan pengaruh langsung terhadap stabilitas proses pembelajaran dan kualitas generalisasi model, terutama karena ukuran batch yang berbeda akan menentukan karakteristik pembaruan bobot selama proses pelatihan.

Eksperimen pencarian hyperparameter paling akhir adalah menentukan durasi pelatihan yang optimal dengan menguji variasi jumlah epochs dari 10 hingga 100, seperti ditunjukkan pada Gambar 7.



Gambar 7. Pengaruh Perubahan Epochs terhadap Akurasi Validasi












Grafik pada gambar 7. memperlihatkan hasil akurasi dari tiap epoch. Analisis performa model berdasarkan variasi jumlah epochs dari 10 hingga 100 menunjukkan bahwa akurasi yang dihasilkan tetap berada pada tingkat yang sangat tinggi dengan fluktuasi yang relatif kecil. Nilai rata-rata akurasi sebesar 0,9692 dan median 0,9694 menegaskan bahwa sebagian besar konfigurasi epochs menghasilkan performa yang konsisten. Rentang akurasi sebesar 0,0244 yang terbentuk dari nilai minimum 0,9555 dan maksimum 0,9800 memperlihatkan bahwa perubahan jumlah epochs memberikan pengaruh tertentu terhadap kualitas pembelajaran, meskipun pengaruh tersebut tidak menyebabkan perbedaan performa yang ekstrem. Nilai simpangan baku sebesar 0,0064 menunjukkan bahwa variasi antar pengamatan tetap dalam batas stabil, sehingga seluruh konfigurasi berada dalam kisaran performa yang terkontrol. Nilai modus 0,9733 mengindikasikan adanya nilai akurasi yang muncul lebih sering dibandingkan nilai lainnya, sehingga dapat dianggap sebagai kecenderungan umum dalam distribusi performa model.

Karakteristik distribusi akurasi memperjelas pola hubungan antara jumlah epochs dan stabilitas hasil pelatihan. Nilai skewness sebesar $-0,6732$ menunjukkan bahwa distribusi akurasi condong ke kiri sehingga sebagian besar nilai akurasi berada sedikit di atas rata-rata. Nilai kurtosis sebesar 2,0393 menunjukkan bahwa distribusi memiliki puncak yang lebih tinggi dibandingkan distribusi normal, yang berarti sebagian besar nilai akurasi berkumpul pada kisaran yang relatif sempit tanpa adanya penyimpangan ekstrem. Total akurasi sebesar 9,6922 dari sepuluh pengamatan menunjukkan konsistensi performa model selama seluruh rangkaian pengujian. Pola ini menunjukkan adanya hubungan sebab-akibat yang jelas, yaitu bahwa penambahan jumlah epochs memberikan kesempatan bagi model untuk menyempurnakan proses pembelajaran, namun setelah titik tertentu peningkatan tersebut tidak lagi menghasilkan kenaikan akurasi yang signifikan

3.2. Konfigurasi Optimal dan Analisis Kinerja Komparatif

Berdasarkan serangkaian eksperimen konfigurasi hyperparameter paling efektif untuk MobileNetV2 yang mencakup learning rate 0.007 dropout 0.2 batch size 32 dan 60 epoch Kombinasi ini sukses mencetak sebuah model yang sangat *robust* dan stabil. Selanjutnya, sebagai upaya untuk menguji dan membuktikan validitas keandalan model tersebut, penelitian ini melangkah pada tahap komparasi performa antara MobileNetV2 dengan arsitektur CNN konvensional. Agar perbandingan berjalan seimbang, CNN konvensional tersebut juga telah disesuaikan dengan hyperparameter learning rate 0.001, dropout 0.4, batch size 16, dan 60 epoch, serta memanfaatkan optimizer Adam dan pembobotan awal dari ImageNet. Analisis yang dilakukan dalam perbandingan ini melampaui sekadar perhitungan angka statistik; kami juga melakukan analisis visual mendalam terhadap sampel hasil klasifikasi sebagaimana ditampilkan pada Tabel 3. Pendekatan visual ini sangat krusial untuk mengamati bagaimana konsistensi prediksi dari masing-masing model dalam mendeteksi dan membedakan karakteristik pola penyakit yang terdapat pada daun kentang

Tabel 3. Hasil klasifikasi Model MobileNetV2 dan CNN

| | MobileNet True | MobileNetV2 False | CNN True | CNN False |
|---------------------|---|---|---|---|
| Normal |  |  |  |  |
| Penyakit Hawar Muda |  |  |  |  |
| Penyakit Hawar Tua |  |  |  | Tidak ada |

Analisis kinerja komparatif antara model usulan dengan CNN Hyperparameter Tuning pada Tabel 3. memperlihatkan kemampuan distingsi fitur yang dilakukan oleh kedua model. Pada kolom MobileNet True terlihat bahwa model mampu mengenali pola visual kunci dengan sangat presisi yakni tekstur halus pada daun Normal bercak konsentris khas pada Hawar Muda serta pola nekrotik yang meluas pada Hawar Tua. Namun kolom MobileNetV2 False mengungkap tantangan visual yang masih tersisa dimana terdapat kemiripan visual yang ambigu antara gejala awal hawar muda dengan lesi pada hawar tua yang menyebabkan terjadinya sedikit misklasifikasi. Meskipun demikian kesalahan ini bersifat minor dan terjadi pada citra dengan gangguan latar belakang atau kondisi pencahayaan yang kompleks. Sebagai pembanding kolom CNN True dan CNN False menunjukkan perilaku model baseline. Meskipun CNN mampu mendeteksi kelas Hawar Tua dengan sangat baik yang ditandai dengan keterangan tidak ada pada kolom

False untuk sampel tersebut model ini cenderung lebih sering mengalami kesulitan dalam membedakan gradasi halus antara daun Normal dan gejala awal Hawar Muda dibandingkan MobileNetV2. Hal ini mengindikasikan bahwa meskipun CNN sangat kompetitif dalam fitur yang jelas MobileNetV2 memiliki sensitivitas yang lebih baik dalam menangkap fitur abstrak pada keseluruhan dataset.

Menguatkan analisis visual di atas hasil evaluasi kuantitatif menempatkan MobileNetV2 sebagai arsitektur yang superior. Model final MobileNetV2 berhasil mencapai akurasi pengujian sebesar 99 % sebuah angka yang mencerminkan keseimbangan ideal antara kemampuan learning dan generalization. Capaian ini tidak lepas dari peran kombinasi hyperparameter learning rate 0.007 dan dropout 0.2 yang mencegah overfitting meski model dilatih dengan 32 batch size dan hingga 60 epoch. Dalam uji komparatif model CNN Hyperparameter tuning mencatatkan akurasi akhir sebesar 97 %. Angka ini menjadi tolak ukur yang kuat membuktikan bahwa model konvensional pun dapat berkinerja luar biasa dengan strategi optimasi yang tepat. Namun selisih 2 % tersebut sangat krusial dalam konteks deteksi penyakit kentang dimana MobileNetV2 membuktikan efektivitas arsitekturnya yaitu blok inverted residuals dalam menangani variasi pola visual yang lebih rumit dibandingkan struktur konvolusi standar pada CNN. Hasil penelitian ini secara signifikan melampaui standar terbaik State of the Art sebelumnya. Penelitian oleh Gunawan et al [12] menetapkan standar kinerja pada angka 95,31 % menggunakan arsitektur serupa. Fakta bahwa bahkan model CNN Baseline dengan akurasi 97 % dalam penelitian ini mampu melampaui angka tersebut dan MobileNetV2 melonjak hingga 99 % menegaskan bahwa kontribusi utama penelitian ini tidak hanya terletak pada pemilihan model melainkan pada metodologi optimasi hyperparameter yang sistematis. Gabungan antara arsitektur yang efisien dan strategi pelatihan yang presisi inilah yang menghasilkan lonjakan akurasi dan menetapkan standar baru dalam klasifikasi penyakit daun kentang.

3.3 Pembahasan

Dari hasil analisis hyperparameter, dapat disimpulkan bahwa setiap hyperparameter memberikan pengaruh yang berbeda terhadap dinamika pembelajaran MobileNetV2. Learning rate cenderung memberikan dampak yang stabil karena mekanisme inverted residual menjaga perubahan bobot tetap terkontrol; akibatnya, variasi nilai LR tidak menimbulkan perbedaan kinerja yang mencolok. Dropout menjadi faktor dengan pengaruh paling jelas karena ia langsung menentukan seberapa banyak informasi fitur yang dipertahankan; dropout terlalu tinggi menyebabkan hilangnya representasi penting, sementara dropout terlalu rendah meningkatkan risiko overfitting. Batch size hanya memberikan pengaruh moderat sebagai pengatur dinamika gradien, bukan penentu kualitas generalisasi. Jumlah epoch justru berperan dalam memperkaya representasi fitur secara gradual hingga titik jenuh pembelajaran, sehingga menjadi faktor yang paling menentukan kedalaman pengetahuan model.

Hasil analisis dengan anova tanpa replikasi memperlihatkan bahwa kombinasi Learning rate dan dropout menunjukkan nilai F sebesar 0,7812 dengan p-value sebesar 0,6323, yang berarti keduanya tidak memberikan pengaruh signifikan terhadap perubahan performa. Hal ini terjadi karena MobileNetV2 memiliki stabilitas gradien bawaan, sehingga variasi learning rate tidak memengaruhi arah pembaruan bobot secara drastis, sementara dropout tidak mengubah sensitivitas pelatihan terhadap gradien. Akibatnya, performa tetap stabil meskipun kedua parameter ini divariasikan.

Pada kombinasi dropout dan batch size menunjukkan nilai F sebesar 6,1739 dengan p-value sebesar 0,0093, yang menandakan bahwa dropout memberikan pengaruh signifikan terhadap akurasi. Penyebabnya adalah dropout secara langsung menentukan jumlah neuron aktif dan mempengaruhi seberapa banyak informasi fitur yang tetap dipertahankan selama pelatihan. Ketika nilai dropout berubah, kualitas representasi fitur juga berubah sehingga performa ikut terpengaruh. Sebaliknya, batch size tidak signifikan karena perubahan ukuran batch tidak cukup memengaruhi dinamika gradien untuk menghasilkan variasi performa yang konsisten.

Hasil batch size dan Epoch menunjukkan nilai F sebesar 11,5467 dengan p-value sebesar 0,0079, yang berarti durasi pelatihan berpengaruh signifikan terhadap kualitas generalisasi model. Sebabnya, semakin banyak epoch memungkinkan model menyempurnakan representasi fitur hingga mencapai titik konvergensi yang stabil. Batch size kembali tidak signifikan karena ukuran batch tidak memberikan dampak langsung pada kedalaman pembelajaran. Akibatnya, performa akhir lebih ditentukan oleh banyaknya iterasi pelatihan, bukan oleh ukuran batch. Secara keseluruhan, ANOVA menegaskan bahwa dropout dan epoch merupakan hyperparameter yang memiliki dampak nyata, sedangkan learning rate dan batch size cenderung memberikan pengaruh yang lebih moderat.

Meskipun hasil yang dicapai sangat baik, refleksi kritis terhadap keterbatasan metodologis penelitian ini esensial. Akuisisi dataset dalam lingkungan yang terkontrol memunculkan pertanyaan mengenai robustitas model terhadap variabilitas kondisi lapangan, seperti kompleksitas latar belakang visual, diversitas varietas tanaman, serta tingkat keparahan penyakit. Validasi pada dataset di lingkungan menjadi arah pengembangan logis selanjutnya. Selain itu, adopsi arsitektur tunggal MobileNetV2 membatasi cakupan investigasi terhadap arsitektur alternatif lainnya yang berpotensi menawarkan optimalisasi yang lebih baik antara akurasi dan efisiensi komputasi. Lebih lanjut, evaluasi metrik implementasi praktis, termasuk latensi inferensi dan konsumsi daya pada perangkat *edge*, merupakan aspek yang belum dieksplorasi namun krusial untuk penentuan viabilitas implementasi di dunia nyata.

Terlepas dari batasan tersebut, penelitian ini menyajikan kontribusi ilmiah yang substantif. Pertama, penelitian ini menetapkan *benchmark* performa baru dengan akurasi 99% untuk tugas klasifikasi penyakit daun kentang. Kedua, penelitian ini mendemonstrasikan secara empiris signifikansi optimalisasi hyperparameter sebagai tahapan krusial untuk memaksimalkan kinerja model. Ketiga, dan yang paling fundamental, penelitian ini menghasilkan sebuah

konfigurasi hyperparameter yang teroptimisasi dan dapat direplikasi mencakup learning rate 0.007, dropout 0.2, batch size 32, dan 60 epoch yang berfungsi sebagai rujukan berharga bagi riset lanjutan. Dengan demikian, studi ini tidak hanya menyajikan kemajuan inkremental dalam akurasi, tetapi juga memperkuat landasan metodologis bagi pengembangan sistem deteksi penyakit tanaman berbasis deep learning di masa depan.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa pengaturan hyperparameter berperan penting dalam meningkatkan kinerja MobileNetV2 untuk klasifikasi penyakit daun kentang. Melalui 3-Fold Stratified Cross-Validation, diperoleh kombinasi hyperparameter yang mampu menghasilkan akurasi 99 %, yaitu learning rate 0.007, dropout 0.2, batch size 32, dan 60 epoch. Capaian ini lebih tinggi dibandingkan model CNN yang digunakan sebagai pembanding maupun penelitian terdahulu, sehingga membuktikan bahwa MobileNetV2 mampu mempelajari pola visual daun secara lebih mendalam, termasuk pada gejala penyakit yang sulit dibedakan dari daun normal. Meskipun demikian, hasil matriks evaluasi menunjukkan adanya kesalahan minor, terutama pada kelas Hawar Muda yang memiliki nilai precision sedikit lebih rendah dibandingkan kelas lain, menandakan bahwa sebagian kecil citra Hawar Muda masih berpotensi terklasifikasi sebagai kategori lain dengan gejala yang mirip. Hasil analisis ANOVA memperjelas bahwa tidak semua hyperparameter memberikan pengaruh yang sama. Learning rate dan batch size tidak menunjukkan perbedaan kinerja yang signifikan karena arsitektur MobileNetV2 memiliki kestabilan gradien yang baik, sehingga perubahan pada kedua parameter tersebut tidak banyak memengaruhi arah pembaruan bobot. Sebaliknya, dropout dan epoch terbukti memberikan pengaruh nyata. Dropout menentukan jumlah neuron aktif selama pelatihan sehingga memengaruhi seberapa banyak informasi fitur yang dapat dipertahankan, sedangkan epoch berperan dalam memperkuat proses pembelajaran hingga model mencapai titik konvergensi yang stabil. Penelitian ini juga membuka peluang eksplorasi lanjut, seperti pengujian ketangguhan model terhadap variabilitas kondisi lapangan untuk memastikan konsistensi performa pada perubahan cahaya maupun latar belakang yang kompleks. Selain itu, studi perbandingan dengan arsitektur modern seperti EfficientNet atau Vision Transformer dapat memberikan wawasan yang lebih mendalam mengenai kompromi antara akurasi, efisiensi, dan kebutuhan sumber daya. Pada akhirnya, prototipe ini berpotensi dikembangkan menjadi solusi aplikatif yang dapat dijalankan pada perangkat di lapangan melalui teknik optimisasi seperti kuantisasi dan pruning, sehingga menghasilkan sistem yang ringan, hemat daya, dan benar-benar bermanfaat bagi pengguna.

REFERENCES

- [1] B. Setiawan, P. Yudono, and S. Waluyo, "Evaluasi Tipe Pemanfaatan Lahan Pertanian dalam Upaya Mitigasi Kerusakan Lahan Di Desa Girirtira, Kecamatan Pejawaran, Kabupaten Banjarnegara," *Vegetalika*, vol. 7, no. 2, doi: <https://doi.org/10.22146/veg.35769>.
- [2] A. Kurniawan, "Uji Kandungan Flavanoid Pada Ekstrak Kentang Secara Kualitatif Dan Kuantitatif," *BENZENA Pharm. Sci. J.*, vol. 1, no. 01, June 2022, doi: 10.31941/benzena.v1i01.2024.
- [3] S. Ghandi, I. Ma'ruf Nugroho, and Y. Raymond Ramadhan, "Penerapan Metode Convolutional Neural Network (CNN) dalam Aplikasi Pendeteksi Penyakit Daun Tanaman Kentang Berbasis Android," *JATI J. Mhs. Tek. Inform.*, vol. 8, no. 5, pp. 8701–8708, Sept. 2024, doi: 10.36040/jati.v8i5.10769.
- [4] Badan Pusat Statistik Indonesia, "Produksi Tanaman Sayuran, 2023," Badan Pusat Statistik Indonesia. [Online]. Available: <https://www.bps.go.id/id/statistics-table/2/NjEjMg==/produksi-%20tanaman-sayuran.html>
- [5] S. Savary, L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts, and A. Nelson, "The global burden of pathogens and pests on major food crops," *Nat. Ecol. Evol.*, vol. 3, no. 3, pp. 430–439, Feb. 2019, doi: 10.1038/s41559-018-0793-y.
- [6] S. Somantri, A. Fergina, A. Renjani, J. A. Malik, and A. Yasa, "Pemanfaatan AgriVision Sebagai Media Pembelajaran dan Deteksi Penyakit Tanaman untuk Petani Indonesia," *J. Pengabd. Masy. Bangsa*, vol. 3, no. 5, pp. 1941–1952, July 2025, doi: 10.59837/jpmba.v3i5.2597.
- [7] M. S. Pramono and A. P. Wibowo, "Penerapan Convolutional Neural Network untuk Identifikasi Penyakit pada Tanaman Padi dari Citra Daun Menggunakan Model ResNet-101," *Djtechno J. Teknol. Inf.*, vol. 5, no. 3, pp. 415–430, Dec. 2024, doi: 10.46576/djtechno.v5i3.5098.
- [8] S. Wahyuni, Indratin, Poniman, and A. N. Ardiwinata, "Identifikasi Cemaran Insektisida Profenofos dari Lahan Bawang Merah di Kabupaten Brebes," *J. Litbang Provinsi Jawa Teng.*, vol. 17, no. 2, pp. 207–215, Dec. 2019, doi: 10.36762/jurnaljateng.v17i2.800.
- [9] T. Alami, Y. Herdiyeni, W. A. Kusuma, B. Tjahjono, and I. Z. Siregar, "Kecerdasan Buatan untuk Monitoring Hama dan Penyakit pada Tanaman Eucalyptus: Systematic Literature Review," vol. 10, no. 2, doi: <https://doi.org/10.29244/jika.10.2.224-237>.
- [10] F. Irhamna Rahman, Lukman, and Ikbal, "Klasifikasi Image Tinggi Tanaman Jagung dengan Menggunakan Algoritma Convolution Neural Network (CNN)," *J. Inform. Dan Tek. Elektro Terap.*, vol. 12, no. 3S1, Oct. 2024, doi: 10.23960/jitet.v12i3S1.5348.
- [11] S. Sodikin, Tutik Khotimah, and Ahmad Jazuli, "Penerapan Transfer Learning Menggunakan Mobile NetV2 untuk Klasifikasi Penyakit Daun Jagung Berbasis Citra," *J. Ekon. Manaj. Sist. Inf.*, vol. 6, no. 6, pp. 4276–4282, July 2025, doi: 10.38035/jemsi.v6i6.5743.
- [12] R. Gunawan, F. Salim, A. I. Wahyudhy, A. Y. Wibowo, G. Yordan, and R. Fauzan, "Klasifikasi Penyakit Daun Kentang dengan Transfer Learning Menggunakan CNN optimalisasi Arsitektur MobileNetV2," vol. 6, no. 2, 2025, doi: <https://doi.org/10.37859/coscitech.v6i2.8599>.



- [13] M. F. Wijayanto, D. Swanjaya, and R. Wulanningrum, “Penerapan MobileNet Architecture pada Identifikasi Foto Citra Makanan Indonesia,” *Digit. Transform. Technol.*, vol. 4, no. 1, pp. 652–662, Aug. 2024, doi: 10.47709/digitech.v4i1.4449.
- [14] D. P. Prabowo *et al.*, “Adaptive Inertia Weight Particle Swarm Optimization for Augmentation Selection in Coral Reef Classification with Convolutional Neural Networks,” *JOIV Int. J. Inform. Vis.*, vol. 9, no. 1, p. 216, Jan. 2025, doi: 10.62527/joiv.9.1.2726.
- [15] Joshua Agung Nurcahyo and Theopilus Bayu Sasongko, “Hyperparameter Tuning Algoritma Supervised Learning untuk Klasifikasi Keluarga Penerima Bantuan Pangan Beras,” *Indones. J. Comput. Sci.*, vol. 12, no. 3, July 2023, doi: 10.33022/ijcs.v12i3.3254.
- [16] F. Amaludin, M. I. Zulfa, and H. Siswanto, “Pengaruh Hyperparameter Tuning pada Kinerja MobileNetV2 dengan Transfer Learning untuk Deteksi Penyakit Kuli,” *J. SINTA Sist. Inf. Dan Teknol. Komputasi*, vol. 2, no. 2, May 2025, doi: 10.61124/sinta.v2i2.43.
- [17] R. Fadilatul Fajriyah and Y. Sulisty Nugroho, “Analisis Tren Penelitian Hyperparameter Tuning dalam Software Engineering melalui Systematic Literature Review dan Bibliometric Analysis,” *J. Pendidik. Dan Teknol. Indones.*, vol. 5, no. 8, pp. 2278–2294, Aug. 2025, doi: 10.52436/1.jpti.817.
- [18] Afis Julianto, Andi Sunyoto, and Ferry Wahyu Wibowo, “Optimasi Hyperparameter Convolutional Neural Network untuk Klasifikasi Penyakit Tanaman Padi,” *Tek. Teknol. Inf. Dan Multimed.*, vol. 3, no. 2, pp. 98–105, Dec. 2022, doi: 10.46764/teknimedia.v3i2.77.
- [19] W. Nugraha and A. Sasongko, “Hyperparameter Tuning on Classification Algorithm with Grid Search,” *SISTEMASI*, vol. 11, no. 2, p. 391, May 2022, doi: 10.32520/stmsi.v11i2.1750.
- [20] N. Gill, P. Hall, K. Montgomery, and N. Schmidt, “A Responsible Machine Learning Workflow with Focus on Interpretable Models, Post-hoc Explanation, and Discrimination Testing,” *Information*, vol. 11, no. 3, p. 137, Feb. 2020, doi: 10.3390/info11030137.
- [21] A. Soni, C. Arora, R. Kaushik, and V. Upadhyay, “Evaluating the Impact of Data Quality on Machine Learning Model Performance,” *J. Nonlinear Anal. Optim.*, vol. 14, no. 01, pp. 13–18, 2023, doi: 10.36893/JNAO.2023.V14I1.0013-0018.
- [22] Rachmat Santoso, “Augmentasi Data pada Prasasti Logam untuk Deteksi Aksara Kawi,” *J. FASILKOM*, vol. 14, no. 1, pp. 234–241, Apr. 2024, doi: 10.37859/jf.v14i1.6952.
- [23] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” 2014, *arXiv*. doi: 10.48550/ARXIV.1412.6980.