

PENGARUH HYPERPARAMETER TUNING PADA KINERJA MOBILENETV2 DENGAN TRANSFER LEARNING UNTUK DETEKSI PENYAKIT KULIT

Ferry Amaludin¹, Mulki Indana Zulfa², Hari Siswanto²

¹Program Studi Teknik Elektro, Universitas Jenderal Soedirman

²Program Teknik Elektro, Universitas Jenderal Soedirman

e-mail: feryamludn@gmail.com

Abstrak

Deteksi penyakit kulit berbasis visual memerlukan model klasifikasi yang efisien dan akurat dalam membantu proses diagnosis secara otomatis. Penelitian ini bertujuan untuk menganalisis pengaruh hyperparameter tuning terhadap performa model Convolutional Neural Network (CNN) dalam melakukan klasifikasi penyakit kulit. Model yang digunakan adalah MobileNetV2 yang dikenal ringan dan efisien untuk perangkat dengan sumber daya yang terbatas. Dataset yang digunakan adalah ISIC 2019, yang terdiri dari delapan kelas penyakit kulit. Pendekatan penelitian yang dilakukan mencakup pelatihan model CNN dasar tanpa transfer learning, pelatihan MobileNetV2 tanpa menggunakan bobot pelatihan, dan pelatihan model MobileNetV2 menggunakan transfer learning. Hasil penelitian menunjukkan bahwa hyperparameter berperan penting dalam meningkatkan kemampuan model. Penggunaan transfer learning pada MobileNetV2 dengan learning rate sebesar 0.0001, batch size 16 dan 70 epoch dengan menggunakan rasio 20:80 untuk data training dan testing menghasilkan akurasi tertinggi sebesar 96.63%. Hasil ini menunjukkan bahwa kombinasi hyperparameter yang tepat serta penerapan transfer learning dapat meningkatkan kemampuan model dalam melakukan klasifikasi penyakit kulit.

Kata kunci : klasifikasi penyakit kulit; Convolutional Neural Network; transfer learning; hyperparameter tuning; MobileNetV2.

Abstract

Visual-based skin disease detection requires an efficient and accurate classification model to support the automatic diagnostic process. This study aims to analyze the effect of hyperparameter tuning on the performance of a Convolutional Neural Network model in classifying skin diseases. The model used is MobileNetV2, which is known for its lightweight architecture and efficiency on resource-constrained devices. The dataset used is ISIC 2019, consisting of eight classes of skin diseases. The experimental approach includes training the model without transfer learning, training MobileNetV2 without pretrained weights, and training MobileNetV2 using transfer learning with pretrained weights. The results indicate that hyperparameters play a crucial role in enhancing model performance. Applying transfer learning on MobileNetV2 with a learning rate of 0.0001, batch size of 16, and 70 epochs using a 20:80 train- test data split achieved the highest accuracy of 96.63%. These results show that the right combination of hyperparameters and the application of transfer learning can improve the model's ability to classify skin diseases.

Keywords : skin disease classification; Convolutional Neural Network; transfer learning; hyperparameter tuning; MobileNetV2.

1. PENDAHULUAN

Kulit merupakan organ terluas yang menutupi permukaan tubuh [1] Kulit yang sehat yaitu kulit dengan kelembaban dengan kadar air yang cukup [2]. Kulit berperan penting dalam memberikan perlindungan dan mengendalikan suhu tubuh [3]. Kulit sering terkena penyakit yang disebabkan oleh faktor lingkungan, kepadatan penduduk, dan sumber air [4]. Penyakit kulit telah meningkat drastis dalam beberapa dekade terakhir [5] Penelitian dilakukan di Irak (1991-2005) dan Singapura (1950-1990) menunjukkan faktor ekonomi, tingkat pendidikan, sanitasi, jenis pekerjaan dan sanitasi menjadi faktor utama terjadinya penyakit kulit [6]. Seiring perkembangan teknologi metode deteksi berbasis *computer vision* menggunakan deep learning telah diterapkan dalam bidang kesehatan untuk mendeteksi penyakit kulit. *Convolutional Neural Networks* (CNN) adalah salah satu arsitektur deep learning yang efektif dalam melakukan klasifikasi penyakit berdasarkan gambar visual. Namun kinerja model ini sangat dipengaruhi oleh pemilihan hyperparameter tuning yang tepat. Hyperparameter tuning merupakan arsitektur dari deep learning untuk meningkatkan performa dari model prediksi sehingga mempunyai peran yang sangat penting dalam penentuan kinerja algoritma machine learning [7].

Beberapa penelitian telah menunjukkan mengenai pentingnya hyperparameter tuning dalam meningkatkan kinerja model deteksi penyakit kulit. Misalnya, sebuah studi mengembangkan model deep learning untuk mendeteksi cacar monyet dengan menyesuaikan Hyperparameter tuning dalam menghasilkan akurasi deteksi yang tinggi [8]. Pada penelitian lain dengan menggunakan algoritma SpaSA untuk mengoptimalkan Hyperparameter tuning dalam *Fully Convolutional Encoder Decoder Network* (FCEDIN) untuk segmentasi citra kulit, menunjukkan peningkatan kinerja model [9]. Selain itu penelitian lain juga menunjukkan bahwa pemilihan Hyperparameter tuning yang tepat akan dapat meningkatkan akurasi model CNN [10].

Penelitian sebelumnya telah menunjukkan bahwa teknik hyperparameter tuning mampu meningkatkan kinerja model machine learning secara signifikan. Penelitian yang dilakukan oleh Nugraha dan Sasongko menunjukkan bahwa optimasi hyperparameter dengan menggunakan metode Grid Search mampu meningkatkan akurasi dari berbagai algoritma klasifikasi [11]. Selanjutnya, Martin juga memanfaatkan GridSearchCV untuk meningkatkan akurasi model Random Forest dalam mendeteksi malware, menunjukkan relevansi tuning dalam konteks keamanan siber [12]. Pada bidang medis, studi yang dilakukan oleh Idrove-Berrezueta et al membandingkan transfer learning dan hyperparameter tuning untuk deteksi dini pneumonia dari citra X-ray, membuktikan bahwa kombinasi keduanya dapat meningkatkan kemampuan model secara signifikan, meskipun fokus penyakit berbeda dengan penelitian ini yang berfokus pada penyakit kulit [13].

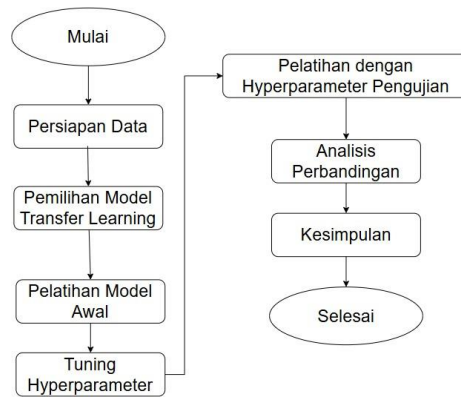
Penelitian juga dilakukan oleh, Dunia et al dalam studi simulasi terhadap berbagai prosedur hyperparameter tuning dalam model prediksi klinis berbasis data berdimensi rendah, akan menambah pemahaman tentang strategi tuning yang optimal [14]. Minarno et al menggunakan CNN yang dikombinasikan menggunakan teknik hyperparameter tuning untuk klasifikasi COVID-19 dengan menggunakan citra medis, dan menunjukkan bahwa hyperparameter tuning dapat memperbaiki performa model dalam melakukan klasifikasi [15]. Penelitian mereka dilanjutkan pada pendeteksian tumor otak yang kembali menegaskan efektivitas CNN dan hyperparameter tuning [16]. Sementara itu studi yang dilakukan oleh Alfredo dan Suharjito menerapkan CNN dan hyperparameter tuning untuk mendeteksi kematangan buah kelapa sawit menggunakan model AlexNet menunjukkan bahwa hyperparameter tuning berpengaruh dalam meningkatkan kinerja model [17].

Berdasarkan temuan dari berbagai penelitian sebelumnya, dapat disimpulkan bahwa hyperparameter tuning merupakan faktor penting dalam meningkatkan kinerja model deep learning, khususnya pada arsitektur *Convolutional Neural Network* (CNN). Oleh karena itu, penelitian ini difokuskan untuk menganalisis pengaruh hyperparameter tuning terhadap akurasi model CNN dalam klasifikasi penyakit kulit dengan menggunakan dataset ISIC 2019.

Eksperimen dilakukan dengan mencoba berbagai kombinasi hyperparameter learning rate, batch size, dan jumlah epoch. Selain itu, penelitian ini juga akan membandingkan pengaruh rasio pembagian data pelatihan dan pengujian terhadap akurasi model. Seluruh eksperimen dilakukan dengan mempertimbangkan pemilihan arsitektur CNN yang sesuai, salah satunya adalah MobileNetV2, yang dikenal ringan dan efektif untuk klasifikasi citra medis. Selain itu dataset ISIC 2019 menjadi salah satu benchmark yang paling banyak digunakan dalam penelitian klasifikasi penyakit kulit, karena menyediakan gambar dengan berkualitas tinggi dengan label yang divalidasi oleh tenaga medis profesional.

2. METODE

Pada penelitian adapun alur dan tahapan penelitian yang harus dilewati yaitu tahap persiapan dan studi literatur, tahap pembuatan model awal, tahap pengujian dan tahap akhir. Untuk tahapannya sendiri ditunjukkan pada Gambar 1. Seluruh proses dilakukan dengan menggunakan Kaggle Notebooks yang mendukung pemrosesan data untuk mempercepat pelatihan model. Model dikembangkan dengan menggunakan bahasa pemrograman python dengan bantuan pustaka Tensorflow dan Keras sebagai framework deep learning. Proses pembuatan model menggunakan GPU P100 dengan kapasitas RAM 29 GB yang disediakan oleh Kaggle Notebooks. Seluruh pipeline dikelola menggunakan notebook berbasis jupyter.



Gambar 1 Diagram Alur Penelitian

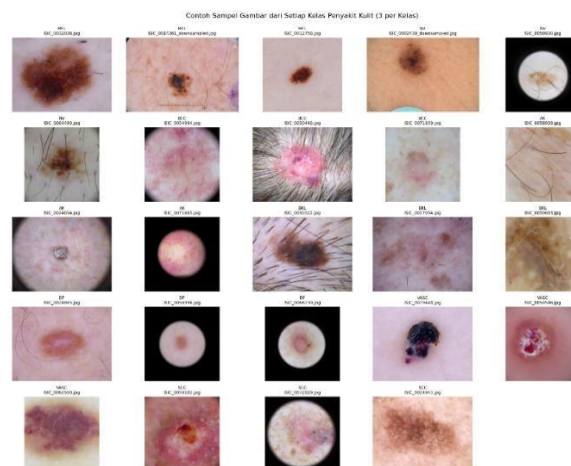
a. Tahapan Persiapan dan Studi Literatur

Tahap ini meliputi identifikasi masalah, perumusan hipotesis, serta penguatan teori melalui studi literatur dari buku, jurnal, dan artikel yang terkait dengan topik dan metode penelitian.

b. Tahap Pembuatan Model Awal

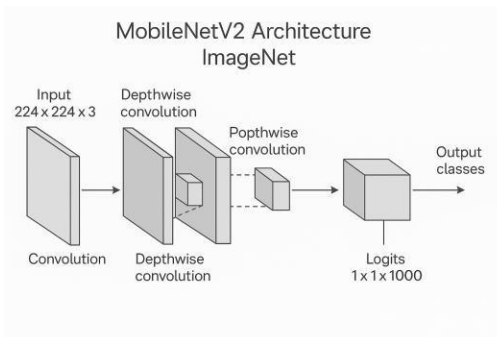
Pada tahap ini akan dilakukan pembuatan model dari sistem yang sudah direncanakan dari proses persiapan dan studi literatur. Tahapan pembuatan model ini berupa:

1. Pengumpulan data dilakukan dengan menggunakan dataset ISIC 2019 yang terdiri dari gambar penyakit kulit dan file label dalam format CSV.
2. Identifikasi data dilakukan untuk memahami struktur dan label data yang tersedia dalam format biner untuk klasifikasi multi kelas.
3. Pembersihan data mencakup penanganan missing value, penghapusan duplikasi data dan perbaikan kesalahan data.
4. Pra-pemrosesan data dilakukan melalui penyesuaian ukuran gambar 32x32 piksel dalam format RGB, encoding label, konversi label ke bentuk one-hot, serta penyeimbangan data menggunakan Teknik oversampling menggunakan Random Oversampling. Contoh hasil pra-pemrosesan data ditampilkan pada Gambar 2.



Gambar 2 Contoh Dataset Penyakit Kulit

5. Pemilihan model menggunakan pendekatan transfer learning dengan arsitektur MobileNetV2 yang telah dilatih sebelumnya menggunakan dataset ImageNet. Visualisasi arsitektur MobileNetV2 ditampilkan pada Gambar 3.



Gambar 3 Arsitektur MobileNetV2

6. Tahap Pengujian dilakukan untuk mengetahui pengaruh variasi hyperparameter epoch, learning rate, dan batch size, serta pengaruh rasio data pelatihan dan pengujian terhadap performa model.

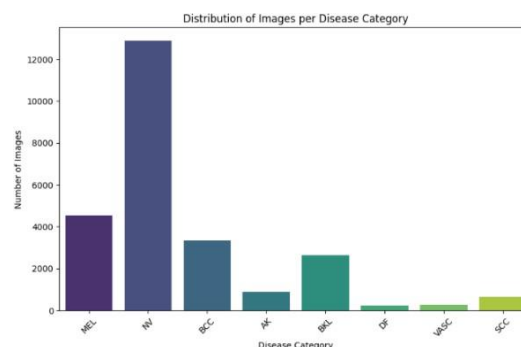
3. HASIL PENELITIAN

A. Deskripsi Dataset

Tahapan pemrosesan dan pembagian data menjadi penting karena akan mempengaruhi kualitas serta performa model yang dibangun.

1. Pembacaan Data dan Pemisahan Data Gambar

Dataset yang digunakan pada penelitian ini memiliki delapan jenis klasifikasi kelas yang terdiri dari MEL, NV, BCC, AK, BKL, DF, VASC, dan SCC. Distribusi data terlampir pada Gambar 1, data yang tidak seimbang memiliki resiko tinggi akan terjadinya overfitting atau terlalu fokus ke kelas tertentu. Sehingga diperlukan pemrosesan selanjutnya.

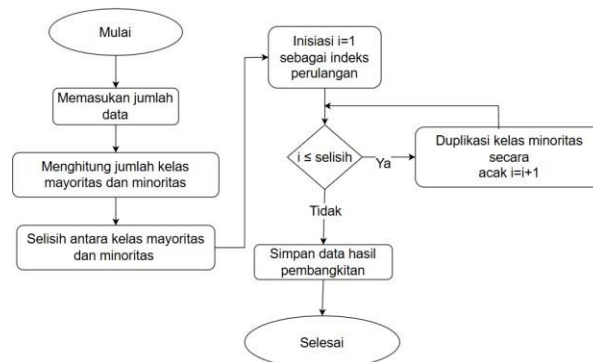


Gambar 4 Distribusi Dataset

2. Penyeimbangan Data menggunakan Random Oversampling

Random oversampling adalah teknik penyeimbangan data yang dilakukan dengan cara menggandakan secara acak kelas minoritas dalam dataset. Tujuannya adalah untuk mengurangi ketidakseimbangan data antara kelas minoritas dan kelas mayoritas. Oversampling dilakukan dengan menggandakan data pada kelas minoritas sehingga seluruh kelas memiliki jumlah data yang sama.

Dalam penelitian yang dilakukan oleh Fitriani et al (2021) dengan judul "PENANGANAN KLASIFIKASI KELAS DATA TIDAK SEIMBANG DENGAN RANDOM OVERSAMPLING PADA NAIVE BAYES (Studi Kasus: Status Peserta KB IUD di Kabupaten Kendal)" disebutkan bahwa penerapan Random Oversampling pada algoritma Naive Bayes dapat meningkatkan nilai G-mean dari kurang dari 60% menjadi 96,6% yang menunjukkan peningkatan signifikan pada model dalam mengklasifikasi data minoritas lebih akurat, selain itu penggunaan random oversampling akan membantu model belajar dari data minoritas karena frekuensi kemunculannya meningkat. Pada penelitian yang dilakukan oleh Fitriani et al.(2021) dapat disimpulkan bahwa penggunaan random sampling dinilai efektif dalam memperbaiki model dalam melakukan klasifikasi [18].



Gambar 5 Proses Random Oversampling

Berdasarkan Gambar 5 menunjukkan proses yang dilakukan dari random oversampling. Langkah pertama yang dilakukan yaitu dengan memasukkan jumlah data dari masing-masing kelas. Kemudian sistem akan menghitung jumlah data pada kelas mayoritas dan minoritas untuk menentukan selisih data antara kedua kelas tersebut. Selisih ini akan digunakan untuk mengetahui berapa banyak data yang ada pada kelas minoritas yang perlu diduplikasi agar jumlahnya seimbang dengan kelas mayoritas. Setelah mendapatkan nilai selisih, proses iterasi dimulai dengan memberikan indeks $i = 1$. Pada setiap iterasi, data dari kelas minoritas akan diduplikasi secara acak, lalu nilai indeks i ditambah satu. Proses ini akan terus berlangsung selama nilai i masih lebih kecil atau sama dengan nilai selisih. Jika sudah terpenuhi, maka sistem akan menyimpan data hasil. Dengan demikian jumlah data antara kelas mayoritas dan minoritas akan menjadi seimbang.

B. Arsitektur Model CNN

1. Pelatihan Model Tanpa Transfer Learning

Pelatihan model awal tanpa Transfer Learning dilakukan dengan membangun arsitektur CNN secara manual dari awal tanpa menggunakan pre-trained model. Meskipun struktur CNN dapat mengenali pola sederhana dalam data gambar, namun terkadang hasil akurasi yang dicapai belum mencapai titik optimal. Arsitektur yang digunakan pada model ini disusun secara bertingkat, dimulai dengan tiga lapisan konvolusi dengan filter 32, 64, dan 128 masing-masing berukuran 3×3 , aktivasi ReLu, padding 'same', regularization L2, inialisasi He Normal. Setiap lapisan diikuti oleh Max Pooling dan Batch Normalization untuk mereduksi dimensi dan menstabilkan proses pelatihan. Setelah ekstraksi fitur, Dropout sebesar 50% diterapkan untuk mencegah overfitting dilanjutkan dengan tiga lapisan dense berturut-turut dengan 256, 128, dan 64 neuron, aktivasi ReLu, regulasi L2 dan batch normalization untuk menjaga kestabilan distribusi nilai aktivasi. Langkah terakhir yaitu menggunakan lapisan dense output dengan delapan neuron dan fungsi aktivasi softmax, untuk mengklasifikasikan gambar ke dalam delapan kelas penyakit kulit.

2. Pelatihan Model MobileNetV2 Tanpa Pembobotan

Pada percobaan berikutnya dengan menggunakan arsitektur MobileNetV2 tanpa bobot pelatihan. Secara desain arsitektur ini lebih efisien, namun dengan tidak menggunakan bobot dari dataset besar dari ImageNet membuat model tidak jauh berbeda dengan melatih CNN dari awal.

3. Pelatihan Model MobileNetV2 Menggunakan Transfer Learning

Berdasarkan hasil dari pendekatan sebelumnya, diterapkan transfer learning menggunakan MobileNetV2 dengan bobot pelatihan dari ImageNet dan input berupa gambar RGB berukuran 32×32 piksel. Arsitektur model dilengkapi dengan custom classifier yang terdiri dari Global Average Pooling2D untuk mereduksi dimensi feature map secara efisien, diikuti Dense layer 256 unit (ReLu), Batch Normalization dan Dropout 0.5 untuk mencegah overfitting. Selanjutnya Dense layer 128 unit (Relu), Batch Normalization, dan Dropout 0.3 untuk meningkatkan generalisasi model., serta Dense layer 8 unit dengan aktivasi Softmax untuk mengklasifikasikan delapan jenis penyakit kulit. Untuk mengoptimalkan performa model, digunakan optimizer Adam karena kemampuannya dalam menyesuaikan learning rate secara adaptif, serta callback ReduceLROnPlateau untuk menghentikan pelatihan jika tidak ada peningkatan performa model, sehingga mencegah overfitting dan menjaga bobot terbaik.

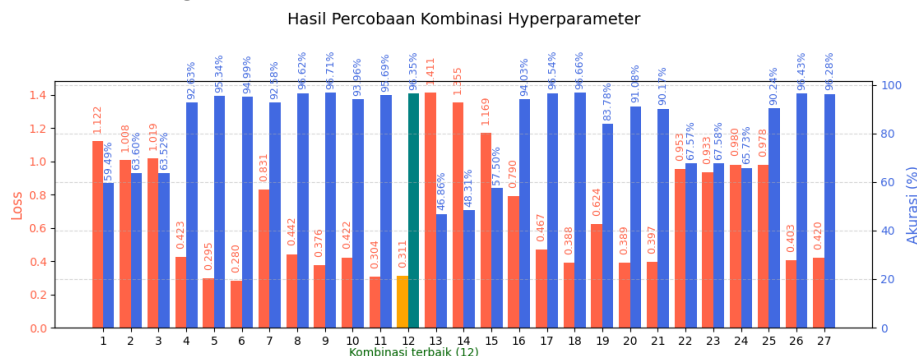
C. Pengujian dan Analisis

Analisis data hasil pengujian dilakukan terhadap beberapa pendekatan model yang dikembangkan yaitu model CNN dasar yang dibangun dari awal tanpa menggunakan transfer learning, MobileNetV2 tanpa bobot pelatihan dan MobileNetV2 dengan menggunakan bobot pelatihan dari ImageNet menggunakan transfer

learning. Analisis dilakukan berdasarkan nilai akurasi dan *loss* pada data testing. Setiap pengujian akan dilakukan proses pembagian dataset dalam rasio 20:80, 25:75, 30:70, 35:65 dan 40:60 pada data training dan testing. Setiap pembagian data akan dilakukan pengujian dengan menggunakan 27 jenis kombinasi hyperparameter yang terdiri dari variasi learning rate, batch size dan epoch. Perbandingan hasil pengujian ditampilkan dalam bentuk diagram batang untuk memudahkan dalam melakukan proses analisis. Pada diagram tersebut nilai akurasi akan ditampilkan menggunakan diagram batang warna biru sedangkan nilai *loss* warna merah. Selain itu untuk memudahkan identifikasi parameter terbaik, diberikan pembeda visual diagram dengan warna khusus yaitu grafik warna hijau untuk kombinasi nilai akurasi dan warna oren digunakan untuk nilai *loss*.

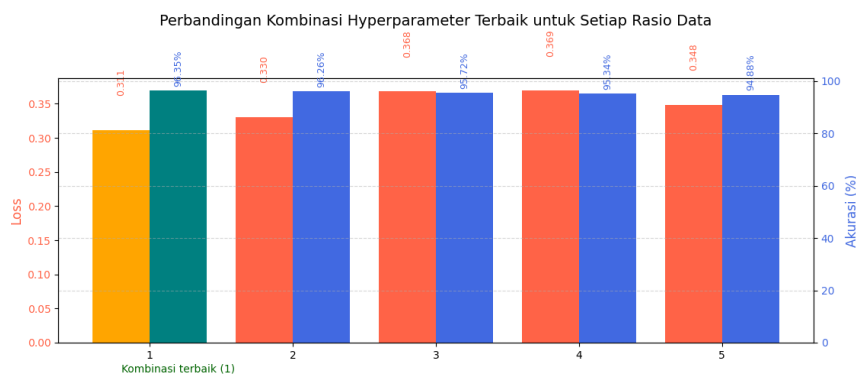
1. Hasil Model Tanpa Transfer Learning

Pengujian ini dilakukan dengan menggunakan arsitektur model CNN dasar yang dibangun dari awal tanpa menggunakan transfer learning.



Gambar 6 Diagram Hasil Pelatihan Model CNN Dasar Tanpa Transfer Learning Rasio 20:80

Berdasarkan Gambar 6 menunjukkan bahwa kombinasi hyperparameter terbaik diperoleh pada learning rate 0.001, batch size 64 dan epoch 100, yang menghasilkan akurasi sebesar 96,35% dengan nilai *loss* 0,3113.

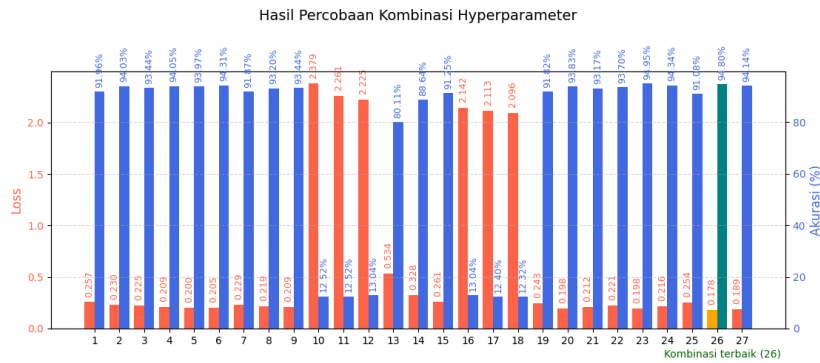


Gambar 7 Perbandingan Hasil Terbaik Setiap Rasio Model CNN Dasar Tanpa Transfer Learning

Pada arsitektur ini, diagram yang ditunjukkan pada Gambar 7 memperlihatkan bahwa kombinasi yang terbaik yaitu pada percobaan pertama dengan menggunakan rasio dataset 20:80, menggunakan parameter learning rate 0.001, batch size 64 dan epoch 100, dengan menghasilkan akurasi sebesar 96.35% dengan nilai *loss* 0.3113.

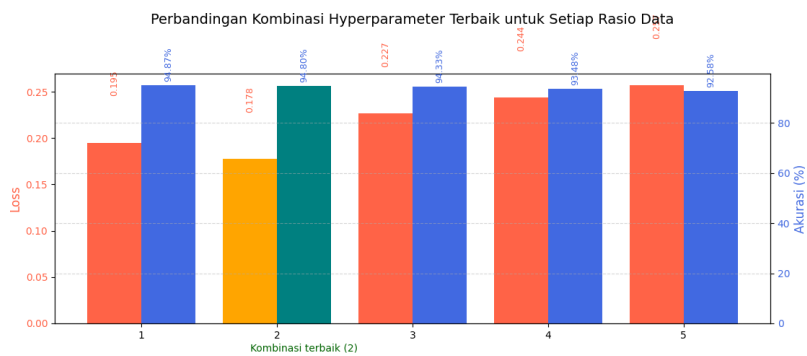
2. Hasil Model MobileNetV2 Tanpa Pembobotan

Pengujian ini dilakukan dengan memanfaatkan arsitektur MobileNetV2 yang dibangun tanpa menggunakan bobot pelatihan (*pre-trained weight*) dari ImageNet. Seluruh bobot dalam model diinisialisasi secara acak dan dilatih sepenuhnya dengan menggunakan dataset ISIC 2019.



Gambar 8 Diagram Hasil Pelatihan MobileNetV2 Tanpa Bobot Pelatihan Rasio 25:75

Berdasarkan Gambar 8 menunjukkan bahwa kombinasi hyperparameter terbaik diperoleh pada learning rate 0.001, batch size 16 dan epoch 70, yang menghasilkan akurasi sebesar 94.80% dengan nilai *loss* 0.1778.

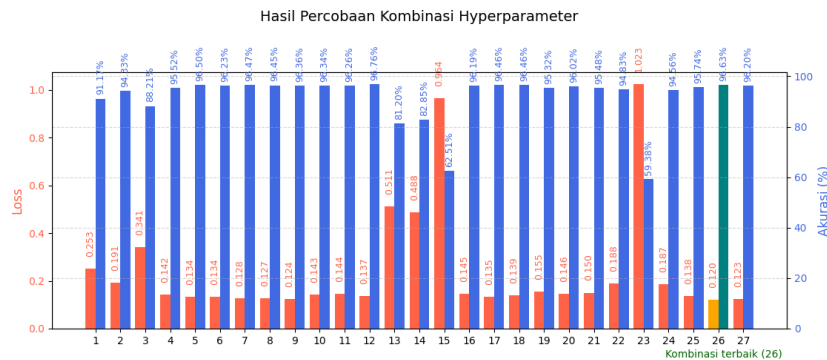


Gambar 9 Perbandingan Hasil Terbaik Setiap Rasio MobileNetV2 Tanpa Bobot Pelatihan

Pada arsitektur ini, diagram yang ditunjukkan pada Gambar 9 memperlihatkan bahwa kombinasi yang terbaik yaitu pada percobaan dengan menggunakan rasio dataset 25:75, menggunakan kombinasi hyperparameter terbaik diperoleh pada learning rate 0.001, batch size 16 dan epoch 70, yang menghasilkan akurasi sebesar 94.80% dengan nilai *loss* 0.1778.

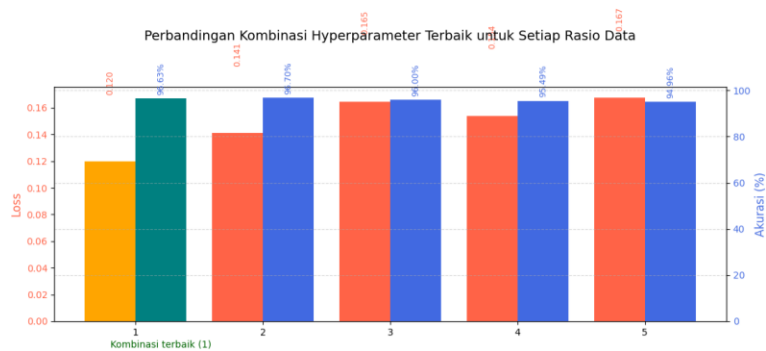
3. Hasil Model MobileNetV2 Menggunakan Transfer Learning

Pada metode ini, dilakukan pelatihan model dengan menggunakan pendekatan transfer learning menggunakan MobileNetV2 yang sudah dilatih sebelumnya dengan menggunakan dataset ImageNet. Bobot dari model pre-trained akan digunakan sebagai titik awal model dalam belajar.



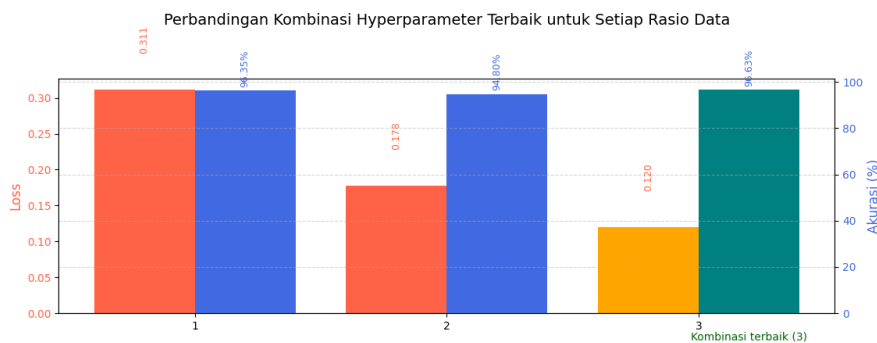
Gambar 10 Diagram Hasil Pelatihan MobileNetV2 Transfer Learning Rasio 20:80

Berdasarkan Gambar 10 menunjukkan bahwa kombinasi hyperparameter terbaik diperoleh pada learning rate 0.0001, batch size 16 dan epoch 70, yang menghasilkan akurasi sebesar 96.63% dengan nilai *loss* 0.1198.



Gambar 11 Perbandingan Hasil Terbaik Setiap Rasio MobileNetV2 Transfer Learning

Pada arsitektur ini diagram yang ditunjukkan pada Gambar 11 menunjukkan bahwa kombinasi hyperparameter terbaik ditunjukkan oleh Rasio dataset 20:80 dengan kombinasi hyperparameter terbaik yaitu learning rate 0.0001, batch size 16 dan epoch 70, yang menghasilkan akurasi sebesar 96.63% dengan nilai *loss* 0.1198, sehingga diperoleh hasil dari seluruh pengujian yang dilakukan pada semua jenis arsitektur model terbaik dengan membandingkan seluruh jenis arsitektur model pengujian yang ditunjukkan pada Gambar 12.



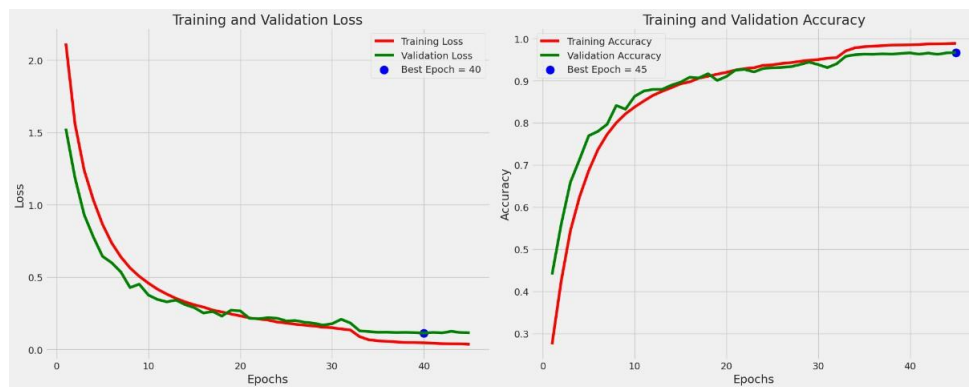
Gambar 12 Perbandingan Hasil Terbaik Setiap Jenis Arsitektur Model

Berdasarkan data kombinasi terbaik itu berasal dari Transfer learning menggunakan MobileNetV2 dengan rasio dataset 20:80 dengan pengujian nomor 26 dengan Rasio 20:80 pada learning rate 0.0001 kombinasi hyperparameter terbaik diperoleh pada learning rate 0.0001, batch size 16 dan epoch 70, yang menghasilkan akurasi sebesar 96.63% dengan nilai *loss* 0.1198. Berdasarkan parameter terbaik pada Tabel 3 adapun grafik akurasi dan validasi yang dapat dilihat pada Gambar 13.

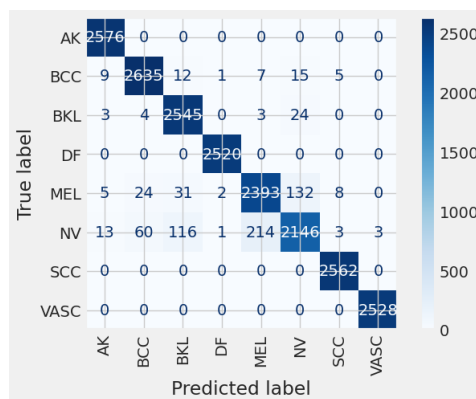
Tabel 1 Parameter Model Terbaik

Rasio	Learning Rate	Batch Size	Epoch	Akurasi	Loss
20:80	0.0001	16	70	96.63%	0.1198

Pada Gambar 14 terlihat grafik yang menunjukkan perkembangan akurasi dan *loss* terhadap jumlah epoch. Grafik yang berada pada sebelah kiri menunjukkan penurunan nilai *loss* yang konsisten dan baik pada data pelatihan dan baik pada data pelatihan maupun validasi yang menandakan bahwa model mampu belajar dengan efektif dan terhindar dari *overfitting*. Sedangkan pada grafik sebelah kanan menunjukkan bahwa kedua kurva akurasi meningkat dengan stabil dan cenderung konvergen. Berdasarkan kurva tersebut menunjukkan model memiliki generalization performance yang sangat baik. Puncak akurasi validasi tercapai pada epoch ke 45, sedangkan nilai *loss* terendah tercapai pada epoch ke 40 yang ditandai oleh titik biru pada masing masing grafik. Evaluasi performa model dilakukan dengan menganalisis confusion matrix yang ditampilkan pada Gambar 14, menunjukkan mayoritas prediksi berada pada diagonal utama, menandakan proses klasifikasi yang akurat untuk setiap kelas. Model berhasil mengenali AK secara sempurna dengan (2576 gambar benar) karena ciri visual yang khas sehingga mudah dikenali. Kelas SCC dan VASC juga menunjukkan akurasi tinggi (5562 dan 2528 gambar) karena memiliki fitur visual yang unik. Kalas BCC dan BKL juga memiliki klasifikasi benar yang tinggi dengan kesalahan yang minimal. Meskipun kelas MEL dan NV mengalami beberapa kesalahan klasifikasi sebanyak 132 yang terjadi karena kemiripan visual antara keduanya. Kelas DF juga memiliki performa yang baik dengan 2520 prediksi benar. Hasil ini menunjukkan bahwa model tidak hanya memiliki akurasi tinggi secara keseluruhan tetapi juga konsisten dalam membedakan antara kelas satu dengan kelas lain dan kemampuan MobileNetV2 sebagai model transfer learning memiliki stabilitas dan efektivitas model yang baik.



Gambar 13 Visualisasi Grafik Akurasi dan Loss Model Terbaik



Gambar 14 Confusion Matrix Model Terbaik

4. DISKUSI

Hasil penelitian ini menunjukkan bahwa transfer learning dengan arsitektur MobileNetV2 tidak hanya meningkatkan akurasi klasifikasi penyakit kulit, tetapi juga menunjukkan kestabilan dalam proses pelatihan dan generalisasi terhadap data baru. Keunggulan ini menguatkan posisi transfer learning sebagai solusi efektif, khususnya dalam bidang medis yang kerap menghadapi keterbatasan data terlabel. Penggunaan bobot awal dari ImageNet memberikan fondasi visual yang kaya sehingga proses pelatihan menjadi lebih efisien dan akurat, bahkan ketika ukuran input gambar direduksi menjadi 32x32 piksel. Meskipun resolusi ini tergolong rendah, akurasi tinggi yang dicapai menunjukkan bahwa MobileNetV2 mampu mengekstraksi fitur penting secara efektif. Namun, ini juga membuka diskusi mengenai potensi peningkatan akurasi apabila digunakan resolusi yang lebih tinggi, atau dengan menerapkan arsitektur yang lebih kompleks seperti EfficientNet atau DenseNet yang memiliki efisiensi parameter lebih tinggi. Lebih lanjut, hasil eksperimen membuktikan pentingnya proses hyperparameter tuning dalam meningkatkan performa model. Kombinasi optimal dari learning rate, batch size, dan epoch terbukti menghasilkan konvergensi yang stabil serta menurunkan nilai loss secara signifikan. Hal ini menegaskan bahwa tanpa tuning yang tepat, bahkan model canggih sekalipun dapat gagal mencapai performa optimal. Oleh karena itu, strategi tuning seperti Grid Search, Random Search, atau bahkan pendekatan modern seperti Bayesian Optimization dapat menjadi arah pengembangan penting dalam penelitian mendatang. Penyeimbangan data dengan Random Oversampling juga berkontribusi terhadap akurasi klasifikasi antar kelas, khususnya pada dataset tidak seimbang seperti ISIC 2019. Namun, penggunaan teknik yang lebih adaptif seperti SMOTE atau ADASYN dapat dieksplorasi untuk menangani distribusi data yang lebih kompleks.

Meskipun demikian, penelitian ini memiliki keterbatasan yang perlu dicermati. Penggunaan resolusi dengan menggunakan 32 x 32 piksel mungkin tidak mampu merepresentasikan fitur visual yang lebih halus dan kompleks, sehingga akan membatasi kemampuan pada beberapa jenis penyakit kulit. Proses tuning hyperparameter masih dilakukan secara manual dan terbatas pada pendekatan grid search sederhana yang memerlukan waktu komputasi tinggi dan belum pasti dalam menemukan konfigurasi optimal. Ketiga, validasi model terbatas pada dataset ISIC 2019 tanpa dilakukan pengujian pada dunia nyata, sehingga belum dapat dipastikan seberapa baik model akan bekerja. Keterbatasan ini menjadi landasan penting dalam penyempurnaan metode dan perluasan cakupan dalam penelitian selanjutnya.

Secara keseluruhan, keberhasilan implementasi model ini membuka peluang penerapan dalam sistem deteksi penyakit kulit berbasis aplikasi mobile atau perangkat keras ringan. Dengan integrasi model yang

efisien dan akurat, pengguna seperti dokter umum atau tenaga medis di daerah terpencil dapat terbantu dalam proses skrining awal secara cepat dan otomatis. Penelitian selanjutnya dapat diarahkan pada validasi klinis model, penerapan pada citra nyata dari berbagai jenis kamera dermatoskop, serta pengujian pada data multi-etnis untuk meningkatkan inklusivitas dan akurasi prediksi global.

5. KESIMPULAN

Berdasarkan penelitian membuktikan bahwa pemilihan hyperparameter yang tepat akan berpengaruh terhadap performa model klasifikasi penyakit kulit. Kombinasi terbaik diperoleh pada learning rate 0.0001, batch size 16 dan epoch 70, yang menghasilkan akurasi 96.63% dan loss 0.1198 yang menunjukkan peningkatan akurasi dan konvergensi model. Rasio dataset 20:80 memberikan hasil terbaik dan memiliki kemampuan generalisasi model. Penggunaan transfer learning menggunakan MobileNetV2 efektif dalam klasifikasi penyakit kulit. Penelitian selanjutnya dapat mengeksplorasi model lain seperti EfficientNet, DenseNet dan dapat menganalisis lebih lanjut mengenai pengaruh penggunaan optimizer yang berbeda.

DAFTAR PUSTAKA

- [1] A. R. MZ, I. G. P. S. Wijaya, and F. Bimantoro, "Sistem Pakar Diagnosa Penyakit Kulit pada Manusia dengan Metode Dempster Shafer," *Journal of Computer Science and Informatics Engineering (J- Cosine)*, vol. 4, no. 2, pp. 129–138, Dec. 2020, doi: 10.29303/jcosine.v4i2.285.
- [2] A. Widiawaty, Sukasihati, K. P. Ayda, and S. A. Nasution, "Skin Moisture Profile of Normal Skin and Inflammatory Skin Disease Using Skin Analyzer," *Galaxy Science*, Dec. 2022. doi: 10.11594/nstp.2022.2825.
- [3] R. Binti Roslan, I. N. Mohd Razly, N. Sabri, and Z. Ibrahim, "Evaluation of psoriasis skin disease classification using convolutional neural network," *LAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 2, p. 349, Jun. 2020, doi: 10.11591/ijai.v9.i2.pp349-355.
- [4] Sumantri Alikhan Sains, "Relationship of Personal Hygiene and Environmental Conditions with Complaints of Skin Disease," *MIRACLE Journal Of Public Health*, vol. 4, no. 1, pp. 90–101, Jun. 2021, doi: 10.36566/mjph/Vol4.Iss1/240.
- [5] M. A. Hashmani, S. M. Jameel, S. S. H. Rizvi, and S. Shukla, "An Adaptive Federated Machine Learning-Based Intelligent System for Skin Disease Detection: A Step toward an Intelligent Dermoscopy Device," *Applied Sciences*, vol. 11, no. 5, p. 2145, Feb. 2021, doi: 10.3390/app11052145.
- [6] Q. Anum *et al.*, "Identifikasi Penyakit Kulit dan Pemeriksaan Lesi Kulit kepada Masyarakat di Daerah Binaan Universitas Andalas," *Jurnal Warta Pengabdian Andalas*, vol. 28, no. 3, pp. 334–340, Sep. 2021, doi: 10.25077/jwa.28.3.334-340.2021.
- [7] 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.
- [8] N. Dahiya *et al.*, "Hyper-parameter tuned deep learning approach for effective human monkeypox disease detection," *Sci Rep*, vol. 13, no. 1, p. 15930, Sep. 2023, doi: 10.1038/s41598-023-43236-1.
- [9] R. Ali, A. Manikandan, R. Lei, and J. Xu, "A novel SpaSA based hyper-parameter optimized FCEDN with adaptive CNN classification for skin cancer detection," *Sci Rep*, vol. 14, no. 1, p. 9336, Apr. 2024, doi: 10.1038/s41598-024-57393-4.
- [10] S. G. Malik, S. S. Jamil, A. Aziz, S. Ullah, I. Ullah, and M. Abohashrh, "High-Precision Skin Disease Diagnosis through Deep Learning on Dermoscopic Images," *Bioengineering*, vol. 11, no. 9, p. 867, Aug. 2024, doi: 10.3390/bioengineering11090867.
- [11] 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.
- [12] I. Muhamad Malik Matin, "Hyperparameter Tuning Menggunakan GridsearchCV pada Random Forest untuk Deteksi Malware," *MULTINETICS*, vol. 9, no. 1, pp. 43–50, May 2023, doi: 10.32722/multinetics.v9i1.5578.
- [13] P. Idrovo-Berrezueta, D. Dutan-Sanchez, and V. Robles-Bykbaev, "Comparison of Transfer Learning vs. Hyperparameter Tuning to Improve Neural Networks Precision in the Early Detection of Pneumonia in Chest X-Rays," 2023, pp. 263–272. doi: 10.1007/978-3-031-33258-6_24.
- [14] Z. S. Dunias, B. Van Calster, D. Timmerman, A. Boulesteix, and M. van Smeden, "A comparison of hyperparameter tuning procedures for clinical prediction models: A simulation study," *Stat Med*, vol. 43, no. 6, pp. 1119–1134, Mar. 2024, doi: 10.1002/sim.9932.
- [15] A. E. MINARNO, M. H. C. MANDIRI, and M. R. ALFARIZY, "Klasifikasi COVID-19 menggunakan Filter Gabor dan CNN dengan Hyperparameter Tuning," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 9, no. 3, p. 493, Jul. 2021, doi: 10.26760/elkomika.v9i3.493.
- [16] A. E. Minarno, M. Hazmi Cokro Mandiri, Y. Munarko, and H. Hariyady, "Convolutional Neural Network with Hyperparameter Tuning for Brain Tumor Classification," *Kinetik: Game Technology*,

Information System, Computer Network, Computing, Electronics, and Control, May 2021, doi:
10.22219/kinetik.v6i2.1219.

- [17] I. Alfredo and Suharjo, "PERBAIKAN MODEL ALEXNET UNTUK MENDETEKSI KEMATANGAN TBS KELAPA SAWIT DENGAN MENGGUNAKAN IMAGE ENHANCEMENT DAN HYPERPARAMETER TUNING," *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 27, no. 1, pp. 56–68, Apr. 2022, doi: 10.35760/tr.2022.v27i1.5973.
- [18] R. D. Fitriani, H. Yasin, and T. Tarno, "PENANGANAN KLASIFIKASI KELAS DATA TIDAK SEIMBANG DENGAN RANDOM OVERSAMPLING PADA NAIVE BAYES (Studi Kasus: Status Peserta KB IUD di Kabupaten Kendal)," *Jurnal Gaussian*, vol. 10, no. 1, pp. 11–20, Feb. 2021, doi: 10.14710/j.gauss.v10i1.30243.