Jurnal Ilmu Komputer dan Informatika (JIKI)
P-ISSN: 2807-6664
E-ISSN: 2807-6591
Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

Evaluation of Transfer Learning-Based Convolutional Neural Networks (InceptionV3 and MobileNetV2) for Facial Skin-Type Classification

Naufal Hafizh Muttaqin[*1], Agung Mulyo Widodo[2]

[1,2]Informatics Engineering Study Program, Faculty of Computer Science, Esa Unggul University, Indonesia

Email: [1]naufalhafizhmuttaqin@gmail.com, [2]agung.mulyo@esaunggul.ac.id

Artikel dikirim: 11 Juni 2025; Revisi-1: 23 Juli 2025; Revisi-2: 28 Juli 2025; Revisi-3: 29 Juli 2025; Diterima: 30 Juli 2025; Dipublikasikan : 03 Agustus 2025

Abstract

Manual classification of facial skin types often suffers from subjectivity and inconsistency due to reliance on human expertise. Accurate identification of skin types is crucial for selecting appropriate skincare solutions. This study evaluates the performance of two transfer-learning-based Convolutional Neural Networks (CNNs), InceptionV3 and MobileNetV2, for classifying facial skin types into four categories: normal, oily, dry, and acne-prone. A total of 1,733 facial images were collected from Kaggle and Roboflow and split into training, validation, and testing sets with a 70:20:10 ratio. Preprocessing involved normalization, augmentation, and resizing based on each model's input size. Both models were fine-tuned and evaluated using accuracy, precision, recall, and F1-score metrics. InceptionV3 achieved the highest accuracy of 90.12% and a macro F1-score of 89.47%, particularly excelling in identifying normal and acne-prone skin. MobileNetV2 reached 81.15% accuracy and performed well on dry skin types. Confusion matrices and evaluation on new, unseen data confirmed the models' generalization capabilities, though misclassifications still occurred among visually similar classes. These findings suggest that CNNs with transfer learning provide a robust foundation for developing AI-assisted facial skin-type classification systems, offering potential integration into dermatological applications.

Keywords : *Convolutional Neural Network, Facial Skin Type, InceptionV3, MobileNetV2, Transfer Learning*

This work is an open access article licensed under a Creative Commons Attribution 4.0 International License.

1. PENDAHULUAN

Kulit wajah memiliki peran penting dalam melindungi tubuh dari faktor eksternal dan menjadi indikator kesehatan secara keseluruhan, sehingga memahami jenis kulit dan cara perawatannya sangat krusial. Setiap individu memiliki tipe dan jenis kulit yang berbeda, seperti kering, normal, berminyak, atau berjerawat yang masing-masing memerlukan pendekatan perawatan yang spesifik untuk menjaga kesehatan dan penampilannya, pengetahuan mengenai jenis kulit membantu dalam pemilihan produk skincare yang tepat [1]. Oleh karena itu, edukasi tentang perawatan kulit yang sesuai dengan jenisnya dapat mengurangi risiko kerusakan kulit dan meningkatkan kualitas hidup secara keseluruhan [2]. Convolutional Neural Network telah menjadi metode yang paling populer karena kemampuan untuk mengekstraksi fitur kompleks dari gambar, yang membuatnya ideal untuk tugas-tugas klasifikasi gambar, seperti mengidentifikasi jenis kulit wajah. Metode CNN yang dioptimalkan dapat meningkatkan akurasi klasifikasi gambar wajah, termasuk pengenalan jenis kulit, dibandingkan metode konvensional. Sejauh ini, model seperti Inception dan MobileNetV2 telah menunjukkan kinerja yang baik dalam menangani kumpulan data yang kompleks. Namun, masih ada beberapa hal yang perlu diperhatikan, seperti generalisasi model dan evaluasi kinerja dalam berbagai kondisi. Tujuan dari penelitian ini adalah untuk menerapkan dan mengevaluasi model CNN untuk klasifikasi jenis kulit wajah. Diharapkan, melalui evaluasi yang menyeluruh, akan dihasilkan model yang tidak hanya akurat tetapi juga tahan terhadap berbagai kondisi pengambilan gambar [3].

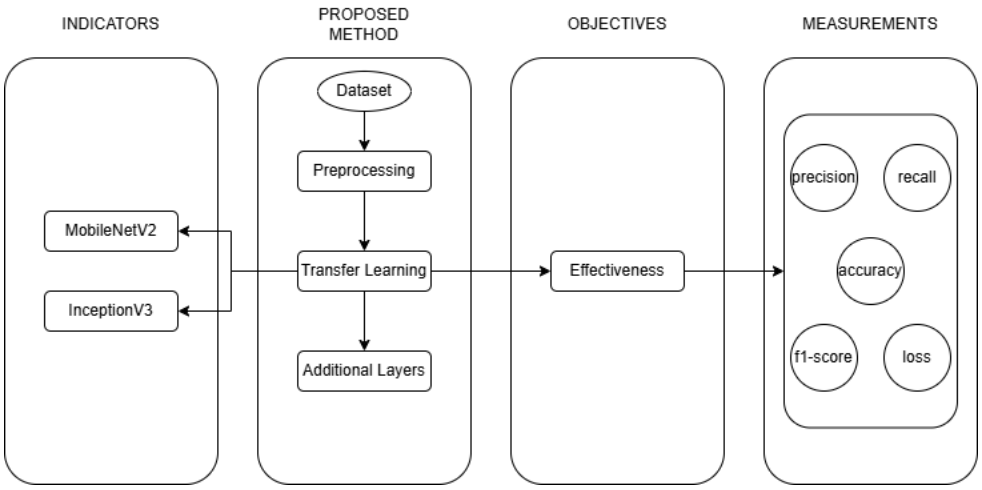
Meskipun klasifikasi jenis kulit wajah secara manual masih umum digunakan, pendekatan ini memiliki keterbatasan dalam hal objektivitas dan konsistensi, karena sangat bergantung pada persepsi

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

visual dan pengalaman individu. Sebagai alternatif, model Convolutional Neural Network (CNN) menawarkan solusi berbasis data yang lebih terukur dan andal. Namun demikian, tantangan teknis seperti sensitivitas terhadap pencahayaan masih menjadi kendala, karena variasi intensitas cahaya dapat memengaruhi representasi visual dan akurasi klasifikasi. Selain itu, pengembangan model CNN yang secara khusus ditujukan untuk aplikasi dermatologi digital dan kosmetik masih terbatas. Penelitian ini mengusulkan implementasi dan evaluasi dua arsitektur CNN InceptionV3 dan MobileNetV2 untuk tugas klasifikasi otomatis jenis kulit wajah. Kinerja model dianalisis berdasarkan akurasi, presisi, recall, dan f1-score, serta diuji dalam kondisi pencahayaan yang bervariasi guna menilai kemampuan generalisasi di lingkungan nyata. Dengan demikian, penelitian ini bertujuan untuk membandingkan performa model InceptionV3 dan MobileNetV2 berbasis transfer learning dalam klasifikasi jenis kulit wajah berdasarkan dataset 4 kelas.

2. METODE

2.1. Alur dan Kerangka Penelitian



Gambar 1. Framework of Thinking

Dalam penelitian ini, kerangka berpikir ini digunakan untuk membuat model klasifikasi jenis kulit wajah dengan menggunakan metode transfer learning berbasis Convolutional Neural Network. Dalam penelitian ini, arsitektur MobileNetV2 dan InceptionV3 dianggap sebagai indikator utama. MobileNetV2 terkenal karena kecepatan pemrosesan data yang cepat dan penggunaan daya yang rendah, sedangkan InceptionV3 unggul dalam mengekstraksi fitur yang lebih mendalam dari gambar, yang membuatnya cocok untuk klasifikasi gambar yang kompleks [4]. Penelitian dimulai dengan pengumpulan dataset foto kulit wajah yang terkait dengan pelatihan model. Preprocessing data, yang mencakup pengukuran, peningkatan, dan normalisasi gambar, kemudian dilakukan. Tujuan dari proses ini adalah untuk mempersiapkan data agar sesuai dengan format input model dan meningkatkan kualitas gambar untuk memungkinkan model untuk belajar lebih baik. Setelah preprocessing, teknik transfer learning digunakan. Dalam teknik ini, model yang telah dilatih sebelumnya pada dataset besar seperti ImageNet digunakan kembali dan disesuaikan untuk klasifikasi tertentu [5]. Ini melibatkan penambahan lapisan tambahan pada arsitektur awal, yang dapat mencakup lapisan dense dan lapisan output yang disesuaikan dengan jumlah kelas yang ada dalam dataset.

Penelitian ini mengembangkan model klasifikasi jenis kulit wajah menggunakan transfer learning berbasis Convolutional Neural Network (CNN), dengan arsitektur MobileNetV2 dan InceptionV3. MobileNetV2 dipilih karena efisien dan ringan, sedangkan InceptionV3 unggul dalam ekstraksi fitur mendalam. Proses dimulai dengan identifikasi masalah dan studi literatur, dilanjutkan dengan pengumpulan data sekunder berupa citra kulit wajah dari sumber publik. Data kemudian diproses melalui tahapan pra-pemrosesan, seperti normalisasi, pengubahan ukuran, dan augmentasi untuk meningkatkan kualitas dan keberagaman data. Model pretrained dari ImageNet digunakan kembali dengan penambahan lapisan dense dan output yang disesuaikan. Model dilatih secara iteratif

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

dan dievaluasi menggunakan lima metrik: akurasi, presisi, recall, F1-score, dan loss, guna menilai efektivitas model dalam mengklasifikasikan empat jenis kulit wajah secara otomatis dan konsisten.

2.2. Dataset dan Akuisisi

Dalam pembentukan model *deep learning*, Peneliti memilih untuk menggunakan beberapa data publik yang telah tersedia untuk umum. Peneliti menggunakan gambar dataset dari sumber universe.roboflow dan kaggle :
https://universe.roboflow.com/parin-kittipongdaja-vwmn3/skin-problem-multilabel
https://www.kaggle.com/datasets/shakyadissanayake/oily-dry-and-normal-skin-types-dataset
Dataset yang digunakan terdiri dari hampir 1750 gambar yang terdiri dari 4 kelas dengan struktur:

Tabel 1. Dataset Distribution

No	Kelas	Jumlah Gambar
1	Berjerawat	460
2	Berminyak	421
3	Kering	422
4	Normal	432

Pada dataset yang tersedia, terdapat perbedaan ukuran gambar, hal ini dapat menurunkan performa dari model *deep learning* sehingga perlu dilakukan pemrosesan data.



Gambar 2. Example images representing each skin class

2.3. Pra-pemrosesan dan Augmentasi

2.3.1. Augmentasi Data

Data augmentation adalah teknik yang digunakan untuk meningkatkan jumlah dan keberagaman data pelatihan secara artifisial dengan cara melakukan transformasi pada data yang sudah ada, seperti rotasi, pergeseran, pemotongan, atau pembalikan. Tujuannya adalah untuk membantu model belajar lebih baik, meningkatkan generalisasi, dan mengurangi risiko overfitting [6].

Pada model machine learning data sangat berpengaruh terhadap proses pelatihan. Augmentasi merupakan proses manupulasi data untuk menambah banyaknya data latih sebelum digunakan sebagai input untuk model. Augmentasi diterapkan untuk meningkatkan jumlah dan keberagaman data dengan transformasi seperti rotasi dan zoom. Proses ini membantu meningkatkan generalisasi model. Pada penelitian ini, penulis akan menggunakan library python yaitu ImageDataGenerator. Dengan penggunaan ImageDataGenerator, dataset akan ditingkatkan jumlahnya untuk setiap kelas yang akan membantu membangun model yang lebih kuat dan generalisasi.

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

```
training_datagen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=15,
    width_shift_range=0.1,
    height_shift_range=0.1,
    shear_range=0.1,
    zoom_range=0.1,
    horizontal_flip=True,
    fill_mode='reflect'
)

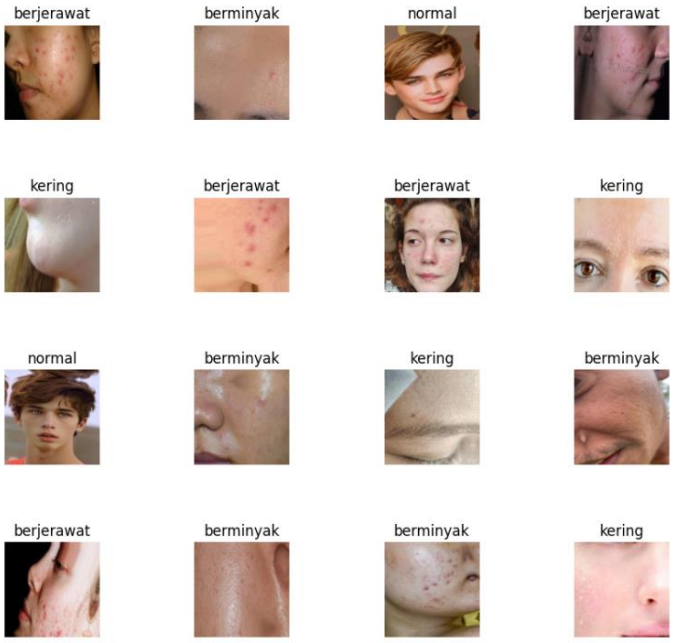
validation_datagen = ImageDataGenerator(rescale=1./255)

train_generator = training_datagen.flow_from_directory(
    training_dir,
    target_size=(224, 224),
    class_mode='categorical',
    batch_size=32,
    shuffle=True
)

validation_generator = validation_datagen.flow_from_directory(
    validation_dir,
    target_size=(224, 224),
    class_mode='categorical',
    batch_size=32,
    shuffle=True
)

testing_generator = validation_datagen.flow_from_directory(
    testing_dir,
    target_size=(224, 224),
    class_mode='categorical',
    batch_size=32,
    shuffle=True
)
```

Gambar 3. Using ImageDataGenerator for Data Augmentation



Gambar 4. Example of an augmented image

2.3.2. Mengubah Ukuran Citra

Pada tahap ini dilakukan perubahan ukuran citra agar seluruh data memiliki dimensi yang seragam sebelum digunakan sebagai input pada model. Ukuran citra sangat penting dalam proses pelatihan karena setiap arsitektur deep learning memiliki kebutuhan input yang berbeda. Penelitian ini menggunakan dua arsitektur Convolutional Neural Network berbasis transfer learning, yaitu InceptionV3 dan MobileNetV2, yang masing-masing memiliki ketentuan ukuran input sebagai berikut:

- InceptionV3 membutuhkan ukuran input sebesar 299x299 piksel, sehingga seluruh citra dalam dataset diubah ukurannya menjadi 299x299 piksel saat digunakan pada model ini.

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

- MobileNetV2 membutuhkan ukuran input sebesar 224x224 piksel, sehingga dataset juga diproses ulang menjadi ukuran tersebut saat digunakan pada model ini.

Perubahan ukuran ini dilakukan menggunakan parameter target_size pada objek ImageDataGenerator dan image_dataset_from_directory di TensorFlow. Tujuan dari proses ini adalah untuk memastikan setiap gambar sesuai dengan spesifikasi model dan agar proses pelatihan dapat berjalan dengan optimal tanpa error dimensi.

2.3.3. Distribusi Dataset

Distribusi dataset training, validation, dan test menggunakan metode pembagian data yang digunakan dalam pengembangan model. Metode ini akan membagi dataset menjadi set pelatihan dan pengujian dengan tetap mempertahankan proporsi kelas di dalam set pelatihan dan set pengujian tetap seragam [7].

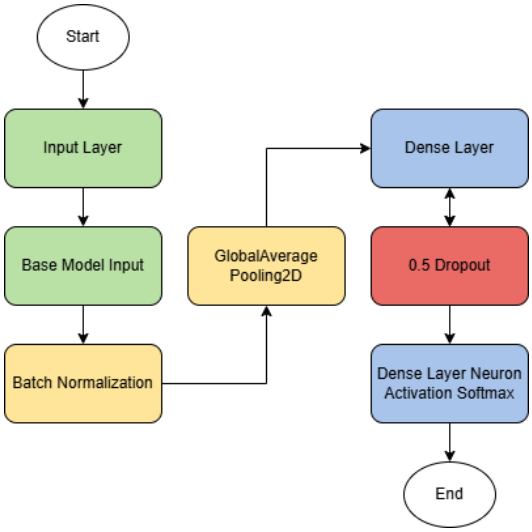
```
import splitfolders
input_folder='/content/skin2'

splitfolders.ratio(input_folder, output="skintype",
                   seed=42, ratio=(.7,.2,.1),
                  group_prefix=None)
```

Gambar 5. Dataset Distribution

2.4. Arsitektur Model

Setelah dataset telah dilakukan pemrosesan data, maka tahap selanjutnya adalah pembentukan model deep learning. Proses pemodelan dilakukan dalam beberapa kali proses percobaan dengan hyperparameter yang berbeda untuk menemukan model dengan akurasi terbaik yang bisa didapatkan. Adapun flowchart arsitektur dan rancangan hyperparameter yang telah didapatkan sebagai berikut.



Gambar 6. Model Architecture

2.4.1. Input Layer

Sebelum data citra digunakan sebagai input pada model, terlebih dahulu dilakukan proses pra-pemrosesan, salah satunya yaitu mengubah ukuran citra dan menormalkan nilai pixel. Ukuran input citra disesuaikan dengan arsitektur model yang digunakan. Pada model InceptionV3, ukuran input yang digunakan adalah 299x299 piksel dengan 3 channel warna (RGB) [8]. Sedangkan untuk MobileNetV2, digunakan ukuran input sebesar 224x224 piksel dengan 3 channel RGB. Seluruh nilai pixel pada gambar dinormalisasi ke dalam rentang [0, 1] dengan membagi setiap nilai pixel dengan 255 [9]. Hal ini dilakukan agar model dapat memproses citra dengan lebih stabil dan efisien saat pelatihan. Setelah itu, citra-citra yang sudah diproses ini akan masuk ke input layer sebagai tahap awal dari pemodelan deep learning.

P-ISSN: 2807-6664 https://jiki.jurnal-id.com
E-ISSN: 2807-6591 DOI: https://doi.org/10.54082/jiki.264

2.4.2. Menambahkan Model Transfer Learning

Pada tahap ini dilakukan inisialisasi model transfer learning sebagai model dasar untuk transfer leaning. Transfer learning memungkinkan untuk memanfaatkan pengetahuan yang sudah dimiliki oleh model berdasarkan tugas sebelumnya dan mengadaptasikannya untuk tugas identifikasi jenis kulit wajah dalam konteks penelitian ini [10]. Dengan menggunakan model yang sudah dilatih, dapat mempercepat dan meningkatkan kinerja model pada dataset yang relative kecil.

2.4.3. Batch Normalization

Batch normalization adalah teknik yang digunakan dalam jaringan saraf untuk menormalkan output dari layer sebelumnya dengan cara menyesuaikan nilai rata-rata dan varians aktivasi pada setiap mini-batch. Tujuannya adalah untuk mempercepat proses pelatihan, menjaga stabilitas distribusi data antar lapisan, serta mengurangi risiko overfitting. Teknik ini memungkinkan model menggunakan learning rate yang lebih tinggi dan membantu konvergensi lebih cepat serta stabil selama pelatihan [11].

Setelah mengintegrasikan model transfer learning sebagai dasar model, akan masuk ke dalam tahap selanjutnya yaitu batch normalization. Tahapan ini akan memanfaatkan inputan yang kita dapatkan sebelumnya yaitu output image dari model transfer learning. Batch Normalization pada tahap ini menggunakan axis=-1, momentum 0.99 dan epsilon 0.001. Pada batch normalization kali ini menggunakan axis=-1, momentum 0.99 dan epsilon 0.001. Axis merupakan sumbu atau dimensi mana yang akan dinormalisasi, dengan penerapan axis -1 maka normalisasi akan dilakukan pada sumbu terakhir seperti untuk data gambar dengan bentuk (64, 299, 299, 3) axis -1 berarti normalisasi akan dilakukan pada sumbu channels, dan momentum merupakan nilai seberapa besar kontribusi dari batch sebelumnya terhadap statistik yang diperbarui. Nilai 0.99 berarti statistik yang diperbarui akan lebih dipengaruhi oleh batch sebelumnya, membuat pembaruan lebih lambat dan lebih stabil.

2.4.4. Dense Layer

Dense layer adalah lapisan dalam jaringan saraf tiruan di mana setiap neuron terhubung secara penuh ke semua neuron di lapisan sebelumnya dan berikutnya. Lapisan ini berfungsi untuk menggabungkan informasi dari proses ekstraksi fitur dan mempelajari hubungan non-linear antar fitur. Menurut Josephine et al. (2021), dense layer memainkan peran penting dalam meningkatkan akurasi model klasifikasi, terutama ketika digunakan dalam jumlah yang lebih banyak dan dengan variasi jumlah neuron di setiap lapisan [12].

Setelah proses ekstraksi fitur selesai dilakukan oleh model dasar (InceptionV3 atau MobileNetV2), tahap selanjutnya adalah menambahkan lapisan Dense atau fully connected. Lapisan ini berfungsi untuk menghubungkan seluruh neuron dari layer sebelumnya dengan seluruh neuron pada layer ini, sehingga mampu mempelajari pola dari fitur yang telah diekstraksi. Digunakan satu buah dense layer dengan 256 neuron dan fungsi aktivasi ReLU (Rectified Linear Unit). Fungsi aktivasi ini dipilih karena dapat membantu model dalam mempelajari pola-pola non-linear dan menghindari masalah vanishing gradient. Lapisan ini bertugas untuk mengolah hasil dari fitur yang telah diringkas sebelumnya menjadi bentuk yang lebih siap untuk diklasifikasikan pada layer output.

2.4.5. Dropout Layer

Dropout adalah teknik regularisasi dalam jaringan saraf tiruan yang bekerja dengan cara menghapus secara acak sejumlah unit neuron beserta koneksinya selama pelatihan. Tujuannya adalah untuk mencegah *overfitting* dengan membuat model tidak terlalu bergantung pada neuron tertentu, sehingga model menjadi lebih robust dan mampu melakukan generalisasi yang lebih baik pada data baru [13].

Lapisan setelah dense layer yang diterapkan adalah dropout layer. Layer dropout digunakan untuk mencegah overfitting dengan secara acak menonaktifkan neuron selama pelatihan. Dropout layer dengan rate 0.5 diterapkan yang berarti bahwa 50% dari unit neuron akan secara acak dinonaktifkan selama pelatihan, pendekatan ini akan membantu meningkatkan kinerja model dengan mengurangi overfitting dan mempercepat konvergensi selama pelatihan.

2.4.6. Softmax Classifier

Softmax classifier adalah fungsi aktivasi yang digunakan pada lapisan output model klasifikasi multi-kelas untuk mengubah nilai keluaran menjadi probabilitas. Fungsi ini memastikan bahwa total seluruh output berjumlah 1, sehingga memungkinkan model memilih kelas dengan probabilitas tertinggi sebagai hasil prediksi [14].

Layer selanjutnya mengimplementasikan layer fully connected dengan 4 neuron dan menggunakan fungsi aktivasi softmax. Layer ini menggunakan 4 neuron yang berarti output dari layer yang menjadi probabilitas, fungsi aktivasi ini akan memastikan bahwa jumlah dari semua probabilitas adalah 1. Model akan memprediksi kelas dengan probabilitas tertinggi sebagai hasil klasifikasinya.

2.4.7. Setup Pelatihan

Setelah model terbentuk, langkah selanjutnya adalah melatih model dengan menggunakan dataset yang tersedia. Proses pelatihan ini dilakukan dengan memanfaatkan data latih (training data) dan memperbarui bobot model menggunakan algoritma optimasi. Tujuan dari pelatihan ini adalah agar model dapat mengenali pola dan fitur dari gambar yang ada. Selain itu, teknik augmentasi data dapat diterapkan untuk meningkatkan variasi dataset dan mengurangi kemungkinan terjadinya overfitting.

Pada saat pelatihan model menggunakan callback berupa myCallback yang berfungsi untuk menghentikan pelatihan secara otomatis apabila model telah mencapai akurasi yang diinginkan. Callback ini akan memantau nilai metrik accuracy dan val_accuracy di akhir setiap epoch. Jika kedua metrik tersebut mencapai 95%, maka pelatihan akan dihentikan menggunakan self.model.stop_training = True. Hal ini berguna untuk menghemat waktu dan sumber daya komputasi karena pelatihan tidak perlu dilanjutkan lagi jika model sudah menunjukkan performa yang optimal. Dengan menerapkan callback ini, proses pelatihan menjadi lebih efisien, karena pelatihan akan berhenti jika model sudah mencapai kinerja yang cukup baik pada data validasi. Callback ini juga membantu mencegah overfitting dan memastikan model tidak dilatih lebih lama dari yang diperlukan.

```
class myCallback(tf.keras.callbacks.Callback):
    def on_epoch_end(self, epoch, logs={}):
        if(logs.get('accuracy')>=0.95 and logs.get('val_accuracy')>=0.95):
            print("\nReached 95% accuracy so cancelling training!")
            self.model.stop_training = True
```

Gambar 7. Implementation of callbacks in training

2.5. Metrik Evaluasi

Penelitian ini akan menganalisis kerja model klasifikasi dengan menggunakan matriks konfusi sebagai instrumen evaluasi inti. Dalam metodologi ini, matriks konfusi memberikan pemahaman yang rinci tentang hasil prediksi algoritma klasifikasi. Analisis yang mendalam terhadap matriks konfusi dan metrik kinerja ini membantu mendapatkan wawasan yang lebih baik tentang sejauh mana algoritma dapat mengklasifikasikan data dengan akurat dan efektif.

Confusion matrix adalah sebuah tabel berukuran $N \times N$ yang digunakan untuk mengevaluasi kinerja model klasifikasi, di mana N adalah jumlah kelas. Setiap baris menunjukkan prediksi model, sementara setiap kolom menunjukkan label sebenarnya. Matriks ini memberikan gambaran rinci tentang jumlah prediksi yang benar dan salah untuk setiap kelas, baik dalam klasifikasi biner maupun multikelas. Dari confusion matrix dapat dihitung berbagai metrik evaluasi seperti akurasi, presisi, recall, dan F1-score, sehingga sangat berguna dalam metode *supervised learning* untuk memahami jenis kesalahan dan kekuatan model klasifikasi [15]

2.5.1. Akurasi (Accuracy)

Rasio akurasi digunakan untuk mengevaluasi kinerja algoritma pada saat klasifikasi. Akurasi akan mengukur sejauh mana model dapat melakukan prediksi dengan benar seluruh kelas. Metrik ini cocok digunakan jika distribusi data antar kelas seimbang, namun bisa menyesatkan jika data tidak seimbang.

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

2.5.2. Presisi (Precision)

Menunjukkan seberapa akurat model dalam memprediksi kelas positif. Nilai presisi tinggi berarti sebagian besar prediksi positif memang benar. Metrik ini penting jika kesalahan prediksi positif harus dihindari.

2.5.3. Recall

Mengukur kemampuan model dalam menemukan seluruh data positif yang sebenarnya. Metrik ini penting ketika melewatkan data positif lebih berdampak dibanding kesalahan prediksi lainnya.

2.5.4. F1-Score

F1-Score akan menunjukkan indikasi sejauh mana model dapat mencapai keseimbangan antara presisi dan recall. Digunakan untuk mengukur keseimbangan antara ketepatan model dalam memprediksi data positif dan kemampuannya dalam menemukan seluruh data positif yang sebenarnya.

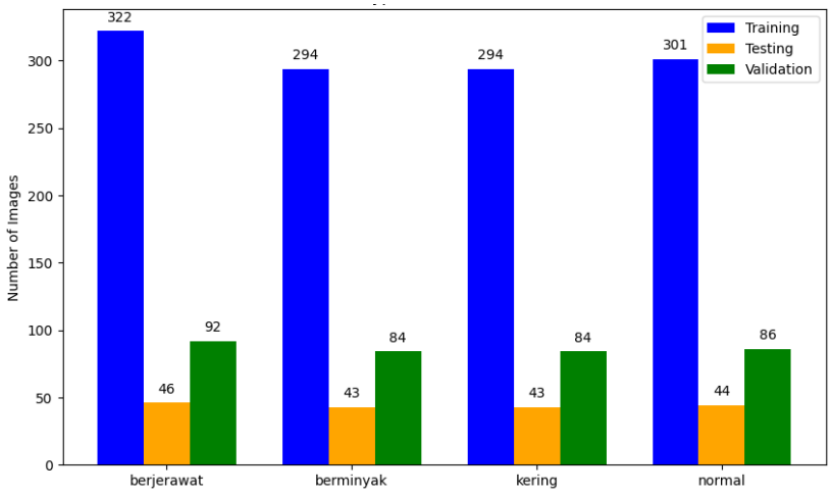
3. RESULT

3.1. Dataset

Dataset yang digunakan berasal dari platform Kaggle dengan nama "face-skin-type" yang berisi gambar wajah manusia dengan klasifikasi empat jenis kulit, yaitu Kering, Normal, Berminyak, dan Berjerawat. Dataset ini kemudian dibagi menjadi tiga bagian dengan rasio 70:20:10 untuk data *training*, *validation*, dan *testing* menggunakan library splitfolders.

Tabel 2. Distribution of Number of Images per Class

Kategori	Training	Validation	Testing
Berjerawat	322	92	46
Berminyak	294	84	43
Kering	294	84	43
Normal	301	86	44



Gambar 8. Data Distribution Visualization

3.2. Pra-Pemrosesan dan Augmentasi Data

Sebelum digunakan dalam pelatihan, citra wajah terlebih dahulu diproses agar sesuai dengan format input dari masing-masing arsitektur CNN. Ukuran gambar disesuaikan menjadi 299×299 piksel untuk InceptionV3 dan 224×224 piksel untuk MobileNetV2. Selain itu, seluruh piksel dinormalisasi ke dalam rentang 0 hingga 1 menggunakan teknik rescaling (1/255), guna memastikan stabilitas dan efisiensi selama proses pelatihan, serta menghindari permasalahan seperti *exploding gradient*. Untuk membantu model belajar dari beragam kondisi dan mencegah overfitting, diterapkan teknik augmentasi pada data latih menggunakan ImageDataGenerator. Proses augmentasi ini mencakup rotasi gambar

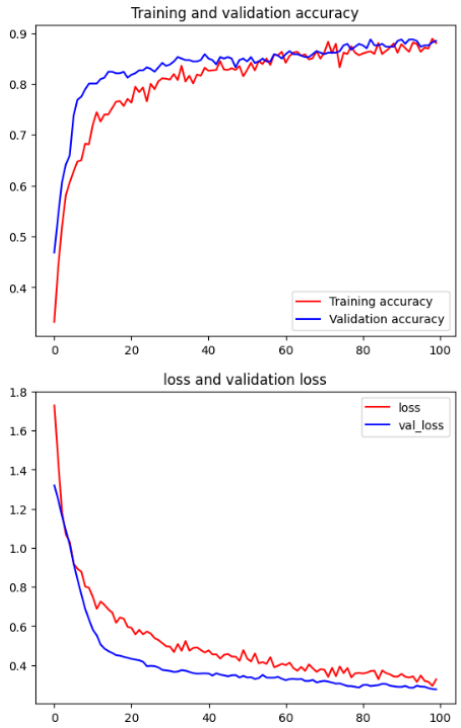
Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

hingga 45 derajat, pergeseran horizontal dan vertikal sebesar 25%, pembesaran (zoom) hingga 50%, flipping horizontal, dan shear transformation. Seluruh augmentasi dilakukan secara acak setiap epoch agar model terbiasa dengan variasi tampilan yang lebih luas. Sementara itu, data validasi dan data uji hanya melalui proses normalisasi tanpa augmentasi tambahan. Tujuannya adalah memastikan proses evaluasi berlangsung secara adil dan objektif, menggunakan data yang tidak dimodifikasi. Pendekatan ini diharapkan mampu menghasilkan model yang tidak hanya akurat saat dilatih, tetapi juga tetap andal saat dihadapkan pada data baru yang belum pernah dilihat sebelumnya.

3.3. Transfer Learning dengan InceptionV3

Transfer learning pada penelitian ini diterapkan menggunakan arsitektur InceptionV3, yang dikenal andal dalam klasifikasi citra karena kemampuannya mengekstraksi fitur dari berbagai skala melalui struktur Inception Modules. Model ini sebelumnya telah dilatih pada dataset ImageNet dan terbukti efektif mengenali pola visual yang kompleks, sehingga dinilai cocok untuk tugas klasifikasi jenis kulit wajah. Dalam implementasinya, InceptionV3 difungsikan sebagai feature extractor dengan menghilangkan bagian atas model menggunakan include_top=False. Output diambil dari layer *mixed4*, yang merupakan layer menengah dengan representasi fitur yang cukup informatif namun belum terlalu spesifik terhadap kelas ImageNet.

Di atas layer ini, ditambahkan beberapa lapisan klasifikasi: *GlobalAveragePooling2D*, *BatchNormalization*, satu layer Dense dengan 256 unit dan aktivasi ReLU, *Dropout* sebesar 50%, serta output layer dengan empat neuron dan fungsi aktivasi *softmax*. Untuk tahap awal pelatihan, seluruh bobot InceptionV3 dibekukan agar tetap mempertahankan pengetahuan dari pelatihan awal di ImageNet. Model dikompilasi dengan optimizer Adam (learning rate 0.0001) dan fungsi *categorical_crossentropy*. Pelatihan dijalankan hingga 100 epoch, namun akan dihentikan lebih awal jika akurasi pelatihan dan validasi keduanya mencapai ≥ 95%, dengan bantuan *callback*. Hasil pelatihan menunjukkan bahwa penambahan *BatchNormalization* berkontribusi dalam meningkatkan stabilitas model dan mempercepat proses konvergensi, tanpa indikasi overfitting.



Gambar 9. Accuracy and Loss Curve for InceptionV3

Gambar di atas memperlihatkan grafik akurasi dan loss dari model InceptionV3 selama pelatihan selama 100 epoch. Bagian atas grafik menampilkan perkembangan *training accuracy* dan *validation accuracy*, sedangkan bagian bawah menunjukkan *training loss* dan *validation loss*. Akurasi pelatihan terlihat meningkat secara konsisten dari awal pelatihan, dimulai sekitar 50% dan naik mendekati 90%.

Akurasi validasi bahkan sejak awal sudah relatif tinggi dan terus stabil di kisaran 88–90% pada akhir pelatihan. Pola ini mengindikasikan bahwa model mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat sebelumnya.

Pada grafik loss, terlihat penurunan tajam pada *training loss* dan *validation loss* di awal pelatihan, kemudian secara bertahap menurun dan stabil mendekati nilai 0.3. Menariknya, *validation loss* cenderung sedikit lebih rendah dari *training loss*, yang menunjukkan bahwa model tidak mengalami overfitting. Secara keseluruhan, hasil visualisasi ini menunjukkan bahwa proses pelatihan berjalan dengan stabil dan efektif. Model tidak hanya belajar dengan baik dari data latih, tetapi juga mempertahankan performa yang konsisten terhadap data validasi.

3.3.1. Fine-Tuning InceptionV3

Setelah *transfer learning* selesai, dilakukan *fine-tuning* dengan membuka seluruh layer pada model InceptionV3 agar dapat belajar lebih dalam dari data kulit wajah. Learning rate diturunkan menjadi 1e-5 untuk mencegah perubahan bobot yang terlalu drastis, dan fungsi loss disesuaikan dengan from_logits=False karena output menggunakan *softmax*. Model kemudian dilatih ulang selama 25 *epoch* dengan data yang sama seperti sebelumnya. Tahap ini bertujuan untuk menyempurnakan hasil transfer learning dan meningkatkan kemampuan generalisasi model.

```
Epoch 1/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 43s 955ms/step - accuracy: 0.9074 - loss: 0.2840 - val_accuracy: 0.8815 - val_loss: 0.2754
Epoch 2/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 29s 767ms/step - accuracy: 0.9062 - loss: 0.2822 - val_accuracy: 0.8815 - val_loss: 0.2751
Epoch 3/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 741ms/step - accuracy: 0.8860 - loss: 0.3256 - val_accuracy: 0.8786 - val_loss: 0.2740
Epoch 4/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 744ms/step - accuracy: 0.8601 - loss: 0.3599 - val_accuracy: 0.8786 - val_loss: 0.2755
Epoch 5/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 734ms/step - accuracy: 0.8706 - loss: 0.3182 - val_accuracy: 0.8786 - val_loss: 0.2746
Epoch 6/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 42s 749ms/step - accuracy: 0.8966 - loss: 0.3264 - val_accuracy: 0.8815 - val_loss: 0.2745
Epoch 7/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 749ms/step - accuracy: 0.8853 - loss: 0.2985 - val_accuracy: 0.8815 - val_loss: 0.2746
Epoch 8/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 29s 772ms/step - accuracy: 0.8794 - loss: 0.3216 - val_accuracy: 0.8815 - val_loss: 0.2752
Epoch 9/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 737ms/step - accuracy: 0.8840 - loss: 0.2884 - val_accuracy: 0.8815 - val_loss: 0.2742
Epoch 10/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 752ms/step - accuracy: 0.8738 - loss: 0.3302 - val_accuracy: 0.8815 - val_loss: 0.2744
Epoch 11/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 29s 761ms/step - accuracy: 0.8820 - loss: 0.3145 - val_accuracy: 0.8815 - val_loss: 0.2758
Epoch 12/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 748ms/step - accuracy: 0.8902 - loss: 0.2843 - val_accuracy: 0.8786 - val_loss: 0.2746
Epoch 13/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 738ms/step - accuracy: 0.8866 - loss: 0.3027 - val_accuracy: 0.8815 - val_loss: 0.2743
Epoch 14/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 29s 772ms/step - accuracy: 0.9007 - loss: 0.3052 - val_accuracy: 0.8815 - val_loss: 0.2742
Epoch 15/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 741ms/step - accuracy: 0.8831 - loss: 0.2983 - val_accuracy: 0.8815 - val_loss: 0.2741
Epoch 16/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 738ms/step - accuracy: 0.8656 - loss: 0.3458 - val_accuracy: 0.8844 - val_loss: 0.2742
Epoch 17/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 30s 796ms/step - accuracy: 0.8764 - loss: 0.2983 - val_accuracy: 0.8786 - val_loss: 0.2731
Epoch 18/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 750ms/step - accuracy: 0.8610 - loss: 0.3402 - val_accuracy: 0.8815 - val_loss: 0.2755
Epoch 19/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 749ms/step - accuracy: 0.8739 - loss: 0.3340 - val_accuracy: 0.8786 - val_loss: 0.2763
Epoch 20/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 29s 743ms/step - accuracy: 0.8633 - loss: 0.3224 - val_accuracy: 0.8786 - val_loss: 0.2773
Epoch 21/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 742ms/step - accuracy: 0.8967 - loss: 0.2935 - val_accuracy: 0.8786 - val_loss: 0.2756
Epoch 22/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 744ms/step - accuracy: 0.8952 - loss: 0.2713 - val_accuracy: 0.8815 - val_loss: 0.2752
Epoch 23/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 41s 733ms/step - accuracy: 0.8834 - loss: 0.3304 - val_accuracy: 0.8815 - val_loss: 0.2738
Epoch 24/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 28s 741ms/step - accuracy: 0.8707 - loss: 0.3319 - val_accuracy: 0.8786 - val_loss: 0.2731
Epoch 25/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 29s 762ms/step - accuracy: 0.8839 - loss: 0.3143 - val_accuracy: 0.8815 - val_loss: 0.2738
<keras.src.callbacks.history.History at 0x7e5ae469d610>
```

Gambar 10. Fine-Tuning InceptionV3

Gambar di atas menampilkan hasil fine-tuning model InceptionV3 selama 25 epoch. Setiap baris mencerminkan metrik pelatihan seperti *accuracy*, *loss*, *val_accuracy*, dan *val_loss* pada akhir masing-masing epoch. Selama proses ini, akurasi pelatihan terus meningkat secara stabil dan berada di kisaran 86% hingga 90%. Akurasi validasi juga menunjukkan performa yang konsisten, yakni antara 88.1% hingga 88.8%, tanpa fluktuasi besar meskipun model terus dilatih. Nilai *validation loss* juga stabil di sekitar 0.238 hingga 0.276, menandakan tidak terjadi overfitting selama proses fine-tuning. Secara keseluruhan, model mampu meningkatkan akurasi pelatihan sambil tetap menjaga performa validasi. Ini menunjukkan bahwa proses fine-tuning membantu model memahami fitur data dengan lebih baik tanpa mengorbankan kemampuan generalisasi.

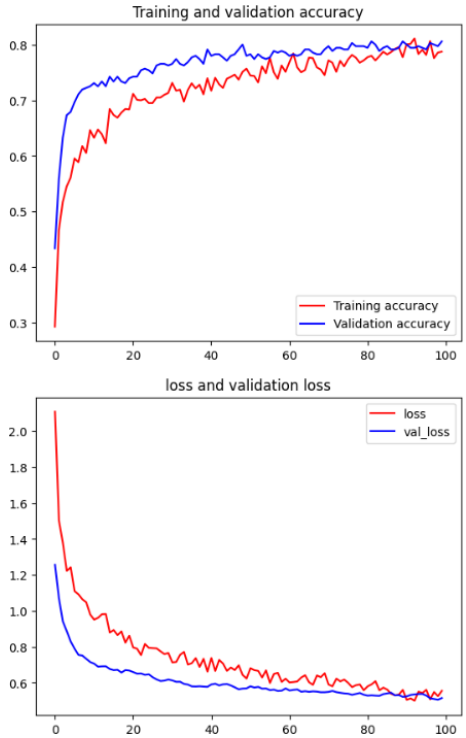
3.4. Transfer Learning MobileNetV2

Transfer learning pada penelitian ini juga diterapkan menggunakan arsitektur MobileNetV2, model deep learning ringan yang dirancang untuk efisiensi tinggi, terutama pada perangkat dengan

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

keterbatasan komputasi. Meski ukurannya lebih kecil dibandingkan model seperti InceptionV3, MobileNetV2 tetap mampu menghasilkan performa klasifikasi yang kompetitif berkat kemampuannya mengekstraksi fitur penting dengan jumlah parameter yang lebih sedikit. Model ini sebelumnya telah dilatih pada dataset ImageNet dan digunakan kembali di sini sebagai feature extractor untuk klasifikasi jenis kulit wajah. Dalam implementasinya, bagian atas model dihapus (include_top=False), dan output dari layer terakhir diproses melalui sejumlah layer tambahan. Dimulai dengan *GlobalAveragePooling2D* untuk meratakan fitur spasial, dilanjutkan dengan *BatchNormalization* untuk menstabilkan proses pelatihan.

Kemudian ditambahkan satu layer Dense dengan 256 unit dan aktivasi ReLU, serta Dropout sebesar 50% untuk mengurangi overfitting. Output layer terdiri dari empat neuron dengan aktivasi softmax, sesuai jumlah kelas yang ingin diklasifikasikan. Seluruh bobot dari MobileNetV2 dibekukan di tahap awal agar tetap mempertahankan pengetahuan dari pretraining-nya. Model dikompilasi menggunakan optimizer Adam (learning rate 0.0001) dan fungsi loss *categorical_crossentropy*. Proses pelatihan berlangsung hingga maksimal 100 epoch, dengan *callback* yang menghentikan pelatihan otomatis jika akurasi pelatihan dan validasi sama-sama mencapai ≥ 95%. Hasil pelatihan menunjukkan bahwa *BatchNormalization* membantu menjaga kestabilan dan efisiensi proses pelatihan, dengan performa validasi yang stabil dan tanpa tanda overfitting. Ini membuktikan bahwa MobileNetV2, meskipun ringan, mampu mengenali pola visual pada kulit wajah dengan cukup baik.



Gambar 11. Accuracy and Loss Curve for MobileNetV2

Gambar di atas menampilkan grafik akurasi dan loss dari model MobileNetV2 selama pelatihan selama 100 epoch. Bagian atas grafik menunjukkan perkembangan *training* dan *validation accuracy*, sementara bagian bawah menampilkan *training* dan *validation loss*. Akurasi pelatihan meningkat secara bertahap dari sekitar 30% di awal hingga melampaui 78% pada akhir pelatihan. Akurasi validasi juga mengalami peningkatan pesat di awal dan kemudian stabil di kisaran 78–80%, menunjukkan bahwa model mampu mengenali pola data dengan cukup baik, meskipun performanya sedikit di bawah model InceptionV3.

Pada grafik loss, *training loss* awalnya cukup tinggi (di atas 2.0), namun terus menurun secara konsisten hingga mendekati 0.6 di akhir pelatihan. *Validation loss* menunjukkan tren serupa, dengan selisih yang relatif kecil dari *training loss*, dan tidak ada kenaikan tajam yang mengindikasikan overfitting. Secara keseluruhan, hasil grafik menunjukkan bahwa pelatihan model MobileNetV2

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

berlangsung stabil. Meskipun akurasinya tidak setinggi InceptionV3, model tetap mampu mencapai performa yang cukup baik dan menunjukkan kemampuan generalisasi yang solid terhadap data baru.

3.4.1. Fine-Tuning MobileNetV2

Setelah *transfer learning* selesai, dilakukan *fine-tuning* dengan membuka seluruh layer pada model MobileNetV2 agar dapat belajar lebih dalam dari data kulit wajah. Learning rate diturunkan menjadi 1e-5 untuk mencegah perubahan bobot yang terlalu drastis, dan fungsi loss disesuaikan dengan from_logits=False karena output menggunakan *softmax*. Model kemudian dilatih ulang selama 25 *epoch* dengan data yang sama seperti sebelumnya. Tahap ini bertujuan untuk menyempurnakan hasil transfer learning dan meningkatkan kemampuan generalisasi model.

```
Epoch 1/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 32s 622ms/step - accuracy: 0.8022 - loss: 0.4812 - val_accuracy: 0.8064 - val_loss: 0.5140
Epoch 2/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 439ms/step - accuracy: 0.7815 - loss: 0.5653 - val_accuracy: 0.8064 - val_loss: 0.5149
Epoch 3/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 18s 471ms/step - accuracy: 0.7982 - loss: 0.5113 - val_accuracy: 0.8064 - val_loss: 0.5156
Epoch 4/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 443ms/step - accuracy: 0.8071 - loss: 0.5088 - val_accuracy: 0.8064 - val_loss: 0.5166
Epoch 5/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 450ms/step - accuracy: 0.8116 - loss: 0.5001 - val_accuracy: 0.8064 - val_loss: 0.5131
Epoch 6/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 20s 442ms/step - accuracy: 0.8075 - loss: 0.4871 - val_accuracy: 0.8064 - val_loss: 0.5132
Epoch 7/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 456ms/step - accuracy: 0.7734 - loss: 0.6050 - val_accuracy: 0.8064 - val_loss: 0.5117
Epoch 8/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 446ms/step - accuracy: 0.7784 - loss: 0.5388 - val_accuracy: 0.8064 - val_loss: 0.5138
Epoch 9/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 443ms/step - accuracy: 0.7857 - loss: 0.5833 - val_accuracy: 0.8064 - val_loss: 0.5142
Epoch 10/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 18s 464ms/step - accuracy: 0.7900 - loss: 0.5290 - val_accuracy: 0.8064 - val_loss: 0.5171
Epoch 11/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 436ms/step - accuracy: 0.7792 - loss: 0.5417 - val_accuracy: 0.8064 - val_loss: 0.5188
Epoch 12/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 18s 483ms/step - accuracy: 0.7873 - loss: 0.4897 - val_accuracy: 0.8064 - val_loss: 0.5179
Epoch 13/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 441ms/step - accuracy: 0.7868 - loss: 0.5552 - val_accuracy: 0.8064 - val_loss: 0.5171
Epoch 14/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 16s 434ms/step - accuracy: 0.8077 - loss: 0.5130 - val_accuracy: 0.8035 - val_loss: 0.5178
Epoch 15/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 18s 454ms/step - accuracy: 0.8069 - loss: 0.5266 - val_accuracy: 0.8035 - val_loss: 0.5165
Epoch 16/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 444ms/step - accuracy: 0.7810 - loss: 0.5652 - val_accuracy: 0.8064 - val_loss: 0.5148
Epoch 17/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 458ms/step - accuracy: 0.7902 - loss: 0.5179 - val_accuracy: 0.8064 - val_loss: 0.5142
Epoch 18/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 18s 467ms/step - accuracy: 0.7770 - loss: 0.5475 - val_accuracy: 0.8064 - val_loss: 0.5151
Epoch 19/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 16s 434ms/step - accuracy: 0.8073 - loss: 0.4770 - val_accuracy: 0.8064 - val_loss: 0.5166
Epoch 20/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 451ms/step - accuracy: 0.8045 - loss: 0.4978 - val_accuracy: 0.8064 - val_loss: 0.5136
Epoch 21/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 16s 430ms/step - accuracy: 0.7845 - loss: 0.5365 - val_accuracy: 0.8064 - val_loss: 0.5122
Epoch 22/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 18s 480ms/step - accuracy: 0.7910 - loss: 0.5362 - val_accuracy: 0.8064 - val_loss: 0.5122
Epoch 23/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 16s 431ms/step - accuracy: 0.8072 - loss: 0.4835 - val_accuracy: 0.8064 - val_loss: 0.5121
Epoch 24/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 16s 430ms/step - accuracy: 0.8020 - loss: 0.4900 - val_accuracy: 0.8035 - val_loss: 0.5131
Epoch 25/25
38/38 ━━━━━━━━━━━━━━━━━━━━ 17s 453ms/step - accuracy: 0.7901 - loss: 0.5099 - val_accuracy: 0.8035 - val_loss: 0.5105
<keras.src.callbacks.history.History at 0x7f9bd830f410>
```

Gambar 12. Fine-Tuning MobileNetV2

Gambar di atas memperlihatkan hasil fine-tuning model MobileNetV2 selama 25 epoch. Setiap epoch mencatat metrik *accuracy*, *loss*, *val_accuracy*, dan *val_loss* pada akhir pelatihan. Akurasi pelatihan terlihat cukup stabil di kisaran 77.3% hingga 81.6%, meskipun sempat mengalami sedikit fluktuasi. Secara umum, tren akurasi pelatihan bertahan di sekitar angka 79–80%. Akurasi validasi juga menunjukkan konsistensi yang baik, berada di rentang 80.4% hingga 80.6%, tanpa penurunan performa seiring berjalannya pelatihan. Nilai *validation loss* pun relatif stabil antara 0.509 hingga 0.517, tanpa lonjakan tajam yang biasanya menandakan overfitting. Secara keseluruhan, hasil fine-tuning MobileNetV2 menunjukkan kinerja yang stabil dan kemampuan generalisasi yang cukup baik. Meskipun peningkatan performa tidak terlalu signifikan dibandingkan tahap awal, model tetap mampu mempertahankan akurasi tinggi pada data validasi dengan loss yang terkendali.

3.5. Evaluasi Model

Setelah seluruh model selesai dilatih, langkah selanjutnya adalah melakukan evaluasi terhadap performa model dalam mengklasifikasikan jenis kulit wajah. Evaluasi dilakukan menggunakan data uji yang tidak pernah dilibatkan dalam proses pelatihan maupun validasi sebelumnya. Dengan demikian, data ini berfungsi sebagai tolak ukur objektif untuk mengukur kemampuan generalisasi model terhadap data baru.

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

Tabel 3. Model Accuracy on Test Data

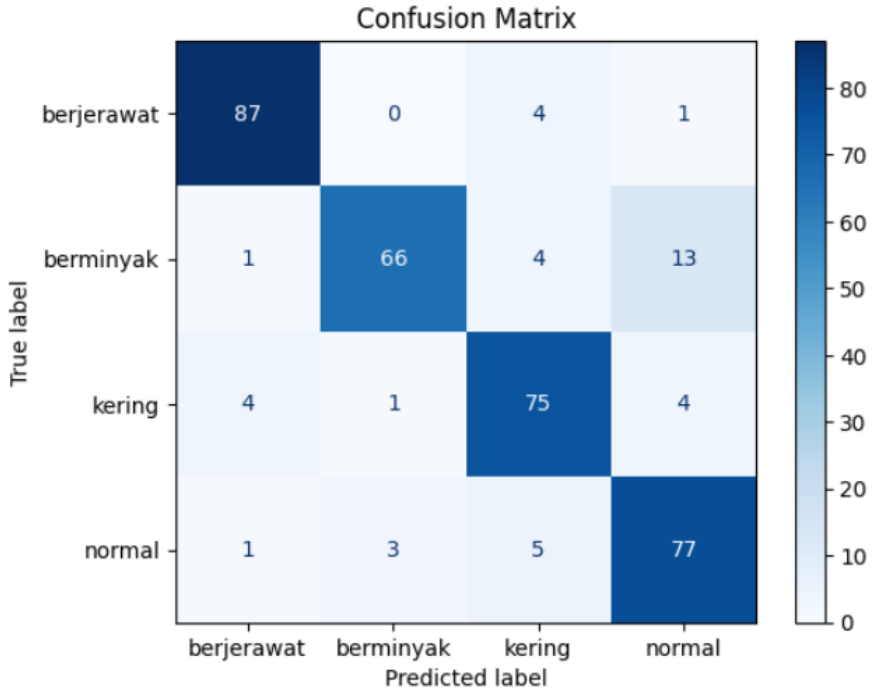
Model	Transfer Learning	Fine-Tuning
InceptionV3	88%	90%
MobileNetV2	78%	81%

Tabel 4. Classification Report of InceptionV3

Kelas	*Precision*	*Recall*	*F1-Score*	*Support*
Berjerawat	0.94	0.95	0.94	92
Berminyak	0.94	0.79	0.86	84
Kering	0.85	0.89	0.87	84
Normal	0.81	0.90	0.85	86

Tabel 5. Classification Report of MobileNetV2

Kelas	*Precision*	*Recall*	*F1-Score*	*Support*
Berjerawat	0.88	0.87	0.87	92
Berminyak	0.79	0.76	0.78	84
Kering	0.68	0.83	0.75	84
Normal	0.90	0.74	0.82	86

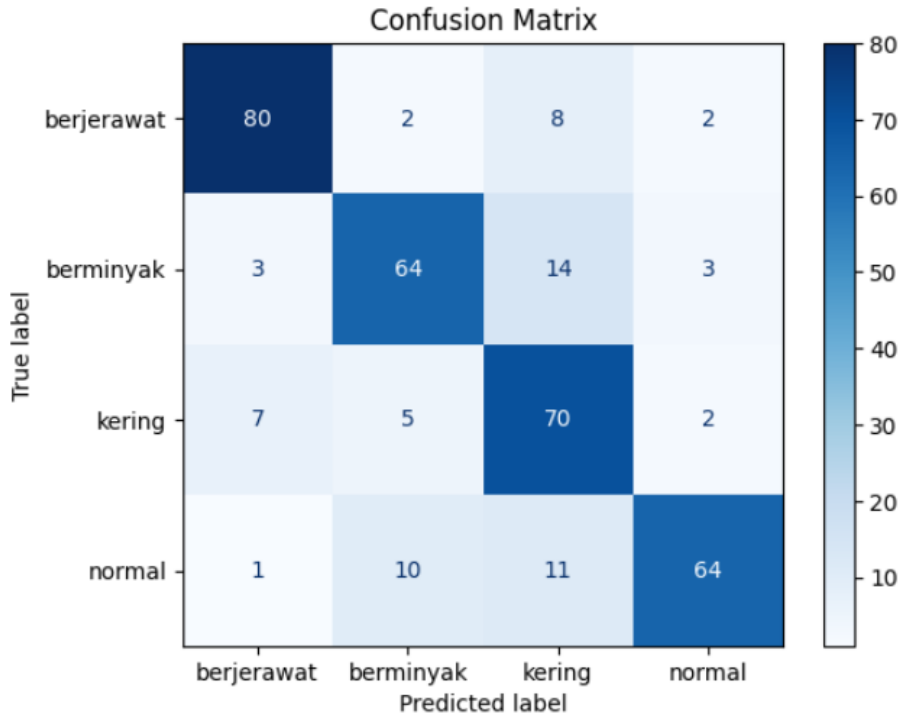


Gambar 13. Confusion Matrix of InceptionV3

Gambar di atas memperlihatkan confusion matrix hasil klasifikasi model InceptionV3 terhadap empat jenis kulit wajah: berjerawat, berminyak, kering, dan normal. Mayoritas prediksi berada pada diagonal utama, yang menunjukkan bahwa model mampu mengklasifikasikan data dengan tingkat akurasi yang cukup baik secara keseluruhan. Pada kelas berjerawat, model menunjukkan performa terbaik dengan 87 dari 92 data berhasil diklasifikasikan secara tepat. Hanya 4 data diklasifikasikan sebagai kering, dan 1 sebagai normal, tanpa kesalahan ke kelas berminyak. Hal ini mengindikasikan bahwa fitur visual pada kulit berjerawat cukup khas dan mudah dikenali oleh model. Kelas berminyak memiliki performa yang paling menantang, dengan 66 dari 84 data diklasifikasikan secara benar. Kesalahan paling dominan terjadi ke kelas normal sebanyak 13 data, serta masing-masing 4 data salah masuk ke kelas kering dan berjerawat. Ini menunjukkan bahwa model kesulitan membedakan kulit berminyak dari kelas lain, terutama karena kemiripan visualnya dengan kulit normal.

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

Pada kelas kering, model berhasil mengklasifikasikan 75 dari 84 data dengan benar. Sisanya terdiri dari 4 data yang salah diklasifikasikan sebagai berjerawat, 4 sebagai normal, dan 1 sebagai berminyak. Meski tidak seakurat kelas berjerawat, performanya tetap tergolong baik dan stabil. Sementara itu, kelas normal diklasifikasikan benar sebanyak 77 dari 86 data. Terdapat beberapa kesalahan klasifikasi, yaitu 5 data salah sebagai kering, 3 sebagai berminyak, dan 1 sebagai berjerawat. Kesalahan ini menunjukkan bahwa kulit normal cukup sering tertukar dengan kelas lain, khususnya kulit kering, yang memiliki tampilan visual yang serupa dalam beberapa kondisi pencahayaan. Secara keseluruhan, model InceptionV3 menunjukkan performa yang cukup baik dalam membedakan keempat jenis kulit wajah. Akurasi tertinggi tercapai pada kelas berjerawat dan kering, sedangkan kelas berminyak dan normal masih menunjukkan tantangan tersendiri akibat kemiripan fitur visual antar kelas. Meskipun begitu, model tetap mampu mengenali pola utama dari masing-masing kategori dengan konsistensi yang cukup tinggi.



Gambar 14. Confusion Matrix of MobileNetV2

Gambar di atas menampilkan confusion matrix hasil klasifikasi model MobileNetV2 terhadap empat jenis kulit wajah: berjerawat, berminyak, kering, dan normal. Secara umum, sebagian besar prediksi berada pada diagonal utama, yang menandakan bahwa model memiliki performa klasifikasi yang cukup baik secara keseluruhan. Pada kelas berjerawat, model berhasil mengklasifikasikan 80 dari 92 data dengan tepat. Sisanya terdiri dari 8 data yang salah diklasifikasikan sebagai kering, serta masing-masing 2 data yang salah sebagai berminyak dan normal. Hal ini menunjukkan bahwa meskipun performanya tidak seakurat InceptionV3, model masih cukup andal dalam mengenali ciri khas kulit berjerawat. Untuk kelas berminyak, terdapat 64 prediksi benar dari 84 data. Kesalahan terbesar terjadi ke kelas kering, dengan 14 data salah klasifikasi, serta masing-masing 3 data yang keliru masuk ke kelas berjerawat dan normal. Ini menandakan bahwa model masih mengalami kesulitan membedakan kulit berminyak dari kering, yang kemungkinan memiliki tekstur atau pantulan cahaya yang serupa dalam citra.

Pada kelas kering, model mencatat 70 prediksi benar dari 84 data. Namun terdapat 7 data yang salah diklasifikasikan sebagai berjerawat, 5 sebagai berminyak, dan 2 sebagai normal. Kesalahan ini relatif tersebar dan mengindikasikan bahwa karakteristik kulit kering dalam gambar bisa tumpang tindih dengan kelas lainnya, terutama dalam kondisi pencahayaan tertentu. Sementara itu, kelas normal memiliki 64 prediksi benar dari 86 data. Kesalahan terbanyak terjadi ke kelas kering (11 data), diikuti oleh berminyak (10 data), dan satu data salah sebagai berjerawat. Ini menunjukkan bahwa kulit normal

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

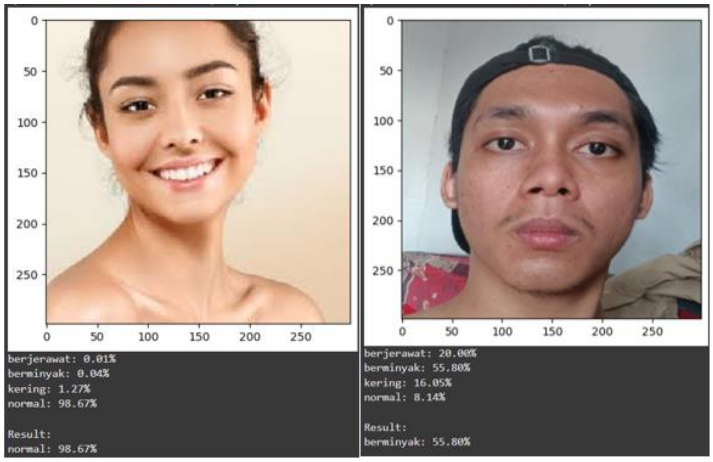
merupakan salah satu kelas yang paling sering tertukar, terutama dengan kelas kering yang secara visual memang cukup mirip. Secara keseluruhan, model MobileNetV2 menunjukkan performa klasifikasi multi-kelas yang stabil dan cukup akurat. Meskipun terdapat beberapa kesalahan antar kelas yang memiliki kemiripan visual, mayoritas data berhasil diklasifikasikan dengan benar. Hasil ini menunjukkan bahwa meskipun arsitekturnya lebih ringan, MobileNetV2 mampu mengenali pola visual utama dari masing-masing jenis kulit wajah dengan tingkat generalisasi yang layak.

3.6. Uji Validitas Model

Sebagai bagian dari evaluasi akhir, dilakukan pengujian tambahan menggunakan citra yang tidak termasuk dalam data pelatihan, validasi, maupun pengujian. Tujuan dari tahap ini adalah untuk menguji kemampuan model dalam mengenali jenis kulit wajah pada data yang benar-benar baru, guna menilai sejauh mana model mampu melakukan generalisasi di luar data yang telah dipelajari sebelumnya. Hasil prediksi ditampilkan dalam bentuk probabilitas untuk setiap kelas, dan label dengan probabilitas tertinggi ditetapkan sebagai hasil klasifikasi. Visualisasi ditampilkan secara langsung berupa gambar beserta label prediksi. Jika probabilitas tertinggi melebihi ambang 50%, maka model menampilkan nama kelas beserta persentasenya. Namun jika seluruh nilai probabilitas rendah (di bawah 50%), sistem akan memberikan peringatan bahwa prediksi tidak dapat ditentukan secara meyakinkan.

3.6.1. InceptionV3

Citra yang diuji diunggah secara manual melalui antarmuka Google Colab, kemudian diproses terlebih dahulu agar sesuai dengan format input model. Gambar diubah ukurannya menjadi 299x299 piksel dan dilakukan preprocessing menggunakan fungsi preprocess_input yang sesuai dengan arsitektur InceptionV3. Setelah itu, gambar diberikan kepada model untuk diprediksi.



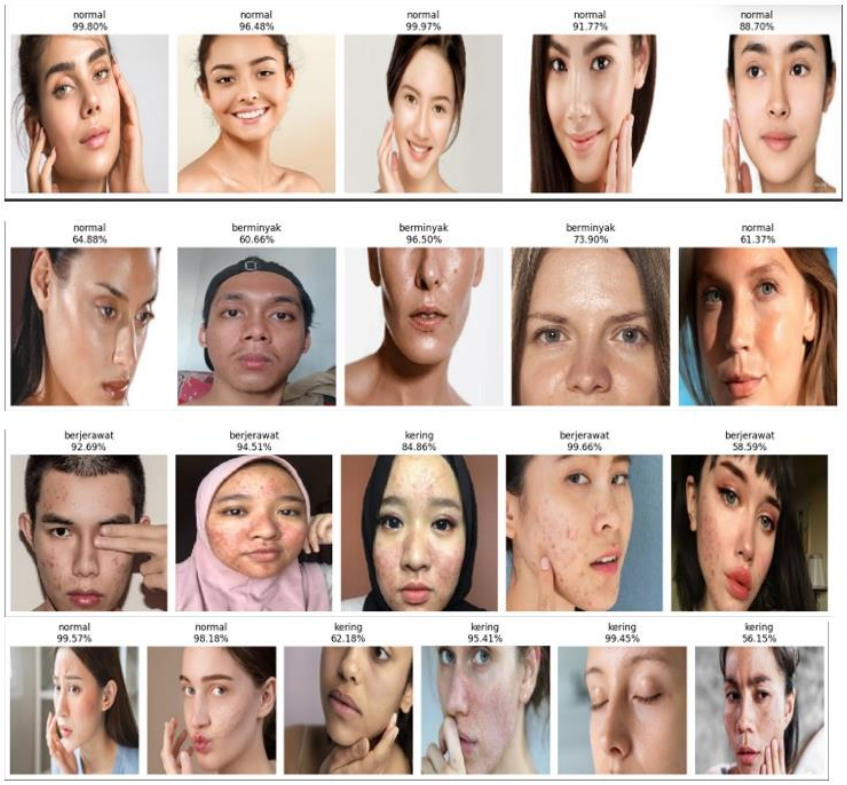
Gambar 15. InceptionV3 prediction results for normal and oily skin images



Gambar 16. InceptionV3 prediction results for acne-prone and dry skin images

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

Berdasarkan hasil pengujian terhadap empat citra uji dari luar dataset, model InceptionV3 berhasil mengklasifikasikan seluruh gambar dengan tepat sesuai label aslinya. Gambar dengan jenis kulit normal diprediksi dengan tingkat keyakinan sangat tinggi sebesar 98.67%, sedangkan gambar kulit berminyak juga diklasifikasikan dengan benar sebagai berminyak dengan probabilitas 55.80%. Untuk citra berjerawat, model menunjukkan performa yang sangat baik dengan prediksi sebesar 99.58%, dan gambar kulit kering diprediksi sebagai kering dengan probabilitas tertinggi, yaitu 60.55%. Hasil ini menunjukkan bahwa model InceptionV3 mampu mengenali ciri khas visual dari masing-masing jenis kulit dengan cukup baik, meskipun gambar yang diuji berasal dari luar data pelatihan.



Gambar 17. InceptionV3 predictions on new unseen facial images across four skin types.

Model InceptionV3 diuji kembali menggunakan gambar wajah baru yang benar-benar berbeda dari data pelatihan sebelumnya untuk mengukur kemampuannya dalam melakukan generalisasi. Hasilnya cukup menjanjikan: mayoritas gambar berhasil diklasifikasikan dengan benar. Untuk kelas normal, seluruh gambar dikenali dengan tepat, dengan tingkat keyakinan tinggi, mulai dari 88,70% hingga 99,97%. Ini menunjukkan bahwa model sangat sensitif terhadap ciri khas kulit normal, seperti tekstur halus dan tidak adanya kilap atau jerawat. Pada kelas berminyak, tiga dari lima gambar diklasifikasikan dengan akurat, salah satunya bahkan mencapai probabilitas 96,50%. Namun, dua gambar lainnya justru diprediksi sebagai normal. Kemungkinan besar, pencahayaan yang kurang optimal membuat kilap khas kulit berminyak tidak terlihat jelas, sehingga membingungkan model.

Untuk kelas berjerawat, performa model cukup baik dengan empat dari lima gambar dikenali secara tepat, seluruhnya di atas 90%. Satu gambar yang keliru diprediksi sebagai kulit kering kemungkinan memiliki jerawat yang tampak samar atau kering. Sedangkan pada kelas kering, empat dari lima gambar berhasil diklasifikasikan dengan benar. Namun, satu gambar diprediksi sebagai kulit normal, kemungkinan akibat pencahayaan dan rona wajah yang menyerupai kulit sehat. Secara keseluruhan, InceptionV3 menunjukkan kemampuan generalisasi yang kuat, terutama dalam mengenali kulit normal dan berjerawat. Namun, kesalahan masih terjadi pada kelas berminyak dan kering yang memiliki kemiripan visual, terutama ketika pencahayaan mengaburkan tekstur dan kilap kulit. Temuan ini menunjukkan bahwa model cukup andal, meskipun tetap perlu ditingkatkan dari segi keragaman data dan preprocessing citra untuk meminimalkan kebingungan antar kelas.

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

3.6.2. MobileNetV2

Citra yang diuji diunggah secara manual melalui antarmuka Google Colab, kemudian diproses terlebih dahulu agar sesuai dengan format input model. Gambar diubah ukurannya menjadi 224x224 piksel dan dilakukan preprocessing menggunakan fungsi preprocess_input yang sesuai dengan arsitektur MobileNetV2. Setelah itu, gambar diberikan kepada model untuk diprediksi.



Gambar 18. MobileNetV2 prediction results for normal and oily skin images



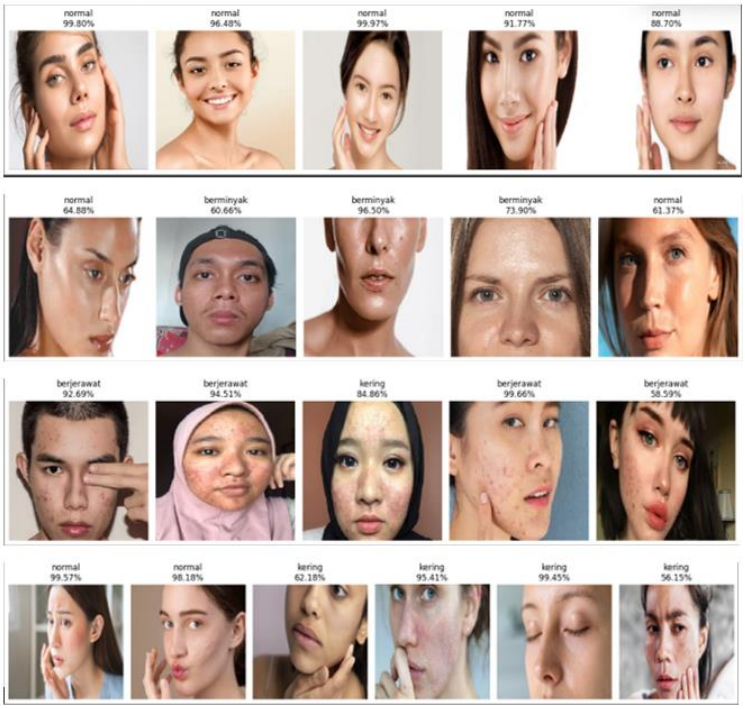
Gambar 19. MobileNetV2 prediction results for acne-prone and dry skin images

Model MobileNetV2 diuji menggunakan empat citra baru yang mewakili masing-masing jenis kulit wajah: normal, berminyak, berjerawat, dan kering. Hasil pengujian menunjukkan bahwa semua gambar berhasil diklasifikasikan dengan benar, meskipun dengan tingkat keyakinan yang bervariasi. Gambar kulit normal diprediksi dengan probabilitas 81,19%, meski masih menunjukkan sedikit keraguan terhadap kelas kering. Untuk kulit berminyak, model memberikan prediksi benar dengan tingkat keyakinan 80,98%, menandakan sensitivitas yang cukup baik terhadap ciri kulit mengilap. Gambar berjerawat juga diklasifikasikan secara tepat dengan probabilitas 75,27%, meskipun terdapat kecenderungan kecil terhadap kelas lain. Sementara itu, kulit kering dikenali dengan sangat meyakinkan, dengan prediksi mencapai 99,62%. Secara keseluruhan, meskipun tergolong ringan, MobileNetV2 menunjukkan performa klasifikasi yang mampu mengenali ciri visual utama dari masing-masing kelas.

Model MobileNetV2 juga diuji kembali menggunakan gambar wajah baru yang belum pernah dilihat sebelumnya untuk menilai kemampuannya dalam menggeneralisasi. Hasilnya menunjukkan bahwa meskipun sebagian gambar berhasil diklasifikasikan dengan benar, masih terdapat sejumlah kesalahan, terutama pada kelas yang memiliki kemiripan visual. Untuk kelas normal, satu gambar justru diklasifikasikan sebagai kulit kering dengan tingkat keyakinan tinggi (88,53%). Kemungkinan, pencahayaan yang lembut dan tekstur kulit yang datar membuat model terkecoh. Pada kelas berminyak, hanya tiga dari lima gambar berhasil dikenali dengan tepat. Dua gambar tidak dapat diklasifikasikan karena semua probabilitas di bawah ambang batas, dan satu gambar malah diprediksi sebagai kulit

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

normal dengan keyakinan tinggi (91,82%). Hal ini menunjukkan bahwa model masih sulit membedakan kulit berminyak dari normal saat kilap tidak tampak jelas.



Gambar 20. MobileNetV2 predictions on new unseen facial images across four skin types

Untuk kelas berjerawat, empat gambar dikenali sebagai berjerawat, meskipun satu diprediksi sebagai berminyak (58,13%) dan satu lainnya hanya mendapat probabilitas sedang (64,35%) untuk kelas yang benar. Kemungkinan besar, jerawat yang tidak terlalu menonjol atau efek pencahayaan memengaruhi hasil prediksi. Sementara itu, pada kelas kering, dua gambar gagal diklasifikasikan karena seluruh kelas menunjukkan probabilitas sangat rendah, dan satu gambar lainnya justru diprediksi sebagai berminyak dengan tingkat keyakinan tinggi (93,23%). Hal ini menunjukkan kebingungan model saat tekstur kasar dari kulit kering tampak seperti kilap tipis akibat cahaya. Secara keseluruhan, model MobileNetV2 menunjukkan performa yang cukup baik pada kelas dengan ciri visual yang kuat, seperti berjerawat dan kering. Namun, tantangan masih muncul pada kelas yang cenderung mirip secara visual—seperti normal dan berminyak. Hal ini menandakan perlunya peningkatan kualitas dan keragaman data, serta teknik preprocessing yang lebih optimal agar klasifikasi antar kelas menjadi lebih akurat.

4. DISCUSSIONS

4.1. Analisis Hasil Model

Model InceptionV3 menunjukkan performa klasifikasi yang lebih unggul dibandingkan MobileNetV2, baik dari segi akurasi maupun kemampuan generalisasi terhadap data baru. InceptionV3 mencatat akurasi sebesar 90%, sementara MobileNetV2 mencapai 81%. Selisih performa ini sejalan dengan perbedaan kompleksitas arsitektur: InceptionV3 dirancang dengan modul multi-path yang mampu mengekstraksi fitur spasial dan semantik secara lebih mendalam, sedangkan MobileNetV2 dioptimalkan untuk efisiensi komputasi melalui penggunaan depthwise separable convolutions dan parameter yang lebih ringan. Keunggulan InceptionV3 semakin terlihat dari nilai macro F1-score yang mencapai 0.90, mengindikasikan keseimbangan antara presisi dan recall di seluruh kelas kulit wajah. Sebaliknya, MobileNetV2 memperoleh nilai F1-score sebesar 0.81, yang masih cukup baik untuk model ringan namun menunjukkan keterbatasan dalam membedakan kelas yang memiliki kemiripan visual seperti kulit berminyak dan normal.

Selain itu, hasil pengujian pada citra uji baru menunjukkan bahwa InceptionV3 memiliki sensitivitas lebih baik terhadap fitur visual yang kompleks, seperti jerawat dan tekstur halus, sedangkan MobileNetV2 lebih rentan terhadap kesalahan klasifikasi akibat pencahayaan atau fitur yang kurang menonjol. Meski begitu, MobileNetV2 tetap sangat layak digunakan untuk aplikasi di perangkat terbatas seperti smartphone, karena waktu komputasi yang singkat dan kebutuhan memori yang rendah. Penelitian ini juga memiliki kontribusi kebaruan dengan membandingkan dua model CNN berarsitektur kontras pada dataset multi-kelas seimbang dalam domain klasifikasi kulit wajah, yang masih jarang dibahas secara mendalam dalam studi terdahulu. Temuan ini memberikan gambaran praktis bagi pengembang aplikasi berbasis AI dalam memilih model berdasarkan kebutuhan performa dan efisiensi sistem, serta membuka peluang eksplorasi lebih lanjut terhadap strategi hybrid, seperti distillation atau ensemble learning, guna menggabungkan keunggulan kedua arsitektur.

4.2. Kesalahan Klasifikasi dan Tantangan

Kesalahan klasifikasi yang paling umum terjadi antara kelas kulit berminyak dan normal. Hal ini kemungkinan besar disebabkan oleh kemiripan visual antara kedua kategori tersebut, khususnya dalam pencahayaan yang kurang ideal. Misalnya, pantulan cahaya pada kulit berminyak bisa saja menyerupai kulit normal, terutama jika pencahayaan lembut. Selain itu, tekstur kulit berjerawat yang tidak terlalu menonjol dapat menyebabkan kebingungan klasifikasi saat kualitas citra rendah. Tantangan ini mengindikasikan pentingnya peran preprocessing dan kualitas data dalam model deep learning berbasis citra. Namun, kelas berminyak dan normal tetap menunjukkan tingkat kesalahan tertinggi dalam confusion matrix. Ini menunjukkan tantangan dalam membedakan dua kelas dengan karakteristik visual serupa, khususnya pada pencahayaan variatif.

4.3. Perbandingan dengan Penelitian Terdahulu

Jika dibandingkan dengan penelitian terdahulu, hasil dalam penelitian ini tergolong kompetitif. Penelitian oleh Fiera Meiristika Utami et al. menggunakan arsitektur MobileNet dan berhasil mencapai akurasi 93.14% dalam klasifikasi empat jenis kulit wajah. Sementara itu, hasil MobileNetV2 dalam penelitian ini hanya mencapai 81%, yang kemungkinan dipengaruhi oleh distribusi data dan variasi pencahayaan pada dataset. Penelitian oleh Rianto dan Listianto menggunakan CNN standar untuk klasifikasi tingkat keparahan jerawat dan memperoleh akurasi 92.34%, namun fokusnya bukan pada klasifikasi jenis kulit. Sementara itu, Dian Anisa Agustina menggunakan ResNet-50 dan mencapai akurasi tinggi 99.86%, namun dilaporkan mengalami overfitting karena data tidak seimbang, masalah yang telah diminimalisir dalam penelitian ini melalui augmentasi dan pembagian data yang proporsional. Meskipun Utami et al. memperoleh akurasi 93.14% menggunakan MobileNet, perbedaan dalam kualitas dataset, jumlah data latih, dan tidak digunakannya validasi pada data baru dapat mempengaruhi generalisasi model. Sebaliknya, penelitian ini menguji performa model pada data tak terlihat untuk mengukur ketahanan klasifikasi terhadap kondisi dunia nyata.

4.4. Keterbasan dan Pengembangan

Penelitian ini masih memiliki beberapa keterbatasan, terutama dalam hal kualitas dan distribusi dataset. Gambar dengan pencahayaan ekstrem atau resolusi rendah masih mempengaruhi akurasi model. Selain itu, meskipun data telah dibagi secara proporsional, variasi tekstur dan warna kulit antar kelas masih belum merata. Untuk selanjutnya, disarankan untuk menambahkan data dari berbagai kondisi lingkungan, mempertimbangkan arsitektur alternatif seperti EfficientNet atau Xception, serta mengimplementasikan preprocessing berbasis segmentasi wajah agar klasifikasi lebih akurat dan spesifik. Untuk pengembangan selanjutnya, model dapat diperkuat melalui teknik interpretabilitas seperti Grad-CAM untuk mengidentifikasi bagian wajah yang berkontribusi terhadap keputusan klasifikasi. Selain itu, pendekatan *domain adaptation* dan eksplorasi arsitektur ringan seperti EfficientNet-Lite memungkinkan integrasi lebih optimal di aplikasi mobile, dengan tetap mempertahankan akurasi yang tinggi. Dataset multispektral juga dapat menjadi arah eksplorasi untuk mengatasi keterbatasan visual citra RGB konvensional.

5. CONCLUSION

Penelitian ini berhasil menerapkan dua arsitektur CNN berbasis transfer learning, yakni InceptionV3 dan MobileNetV2, untuk tugas klasifikasi otomatis jenis kulit wajah ke dalam empat kategori utama. Secara umum, model InceptionV3 menunjukkan performa lebih tinggi dibanding MobileNetV2, terutama dalam mendeteksi jenis kulit normal dan berjerawat. Sebaliknya, MobileNetV2 menunjukkan efisiensi yang baik dan cocok diterapkan pada perangkat dengan sumber daya terbatas. Evaluasi terhadap data baru juga menunjukkan kemampuan generalisasi yang cukup baik dari kedua model, meskipun kesalahan klasifikasi masih ditemukan pada kelas dengan kemiripan visual tinggi. Kontribusi utama terletak pada penerapan dua arsitektur CNN dengan kompleksitas berbeda pada dataset seimbang, disertai teknik augmentasi untuk meningkatkan kemampuan generalisasi. Hasil penelitian ini dapat menjadi landasan untuk pengembangan sistem klasifikasi kulit berbasis AI dalam bidang dermatologi digital maupun aplikasi mobile. Namun demikian, masih terdapat beberapa keterbatasan seperti sensitivitas terhadap pencahayaan ekstrem dan variasi etnisitas yang kurang merata pada dataset. Disarankan untuk mengeksplorasi model yang lebih efisien seperti EfficientNet-Lite, menerapkan pendekatan domain adaptation untuk peningkatan ketahanan model terhadap kondisi dunia nyata, serta teknik interpretabilitas seperti Grad-CAM untuk analisis keputusan model yang lebih transparan.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest related to the publication of this article.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to the supervising lecturers for their guidance and valuable input during the research process. Appreciation is also extended to colleagues at the Informatics Engineering Study Program, Universitas Esa Unggul, for technical assistance and insightful discussions. Special thanks to Kaggle and Roboflow for providing the datasets used in this study. Lastly, the authors sincerely thank their families and all those who offered encouragement and motivation throughout the writing of this article.

REFERENCES

[1] R. Oliveira, J. Ferreira, L. F. Azevedo, and I. F. Almeida, "An Overview of Methods to Characterize Skin Type: Focus on Visual Rating Scales and Self-Report Instruments," *Cosmetics*, vol. 10, no. 1, pp. 1–15, 2023.

[2] F. Bernerd, T. Passeron, I. Castiel, and C. Marionnet, "The Damaging Effects of Long UVA (UVA1) Rays: A Major Challenge to Preserve Skin Health and Integrity," *Int. J. Mol. Sci.*, vol. 23, no. 15, pp. 1–33, 2022.

[3] F. B. KARA, R. KARA, and S. SAKACI ÇELİK, "Skin Type Detection with Deep Learning: A Comparative Analysis," *Düzce Üniversitesi Bilim ve Teknol. Derg.*, vol. 11, no. 2, pp. 729–742, 2023.

[4] A. H. Abdulaal *et al.*, "Cutting-Edge CNN Approaches for Breast Histopathological Classification : The Impact of Spatial Attention Mechanisms CUTTING-EDGE CNN APPROACHES FOR BREAST HISTOPATHOLOGICAL CLASSIFICATION : THE IMPACT OF SPATIAL ATTENTION MECHANISMS," no. October, 2024.

[5] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer learning for medical image classification: a literature review," *BMC Med. Imaging*, vol. 22, no. 1, pp. 1–13, 2022.

[6] X. Liu, G. Karagoz, and N. Meratnia, "Analyzing the Impact of Data Augmentation on the Explainability of Deep Learning-Based Medical Image Classification," *Mach. Learn. Knowl. Extr.*, vol. 7, no. 1, pp. 1–28, 2025.

[7] J. Sadaiyandi, P. Arumugam, A. K. Sangaiah, and C. Zhang, "Stratified Sampling-Based Deep Learning Approach to Increase Prediction Accuracy of Unbalanced Dataset," *Electron.*, vol. 12, no. 21, pp. 1–16, 2023.

[8] K. Vinay, V. Surya, S. Thushar, T. Singh, and A. Sahay, "A Deep Learning Framework for Early

Detection and Diagnosis of A Deep Learning Framework for Early Detection and Diagnosis of Plant a Plant Diseases," *Procedia Comput. Sci.*, vol. 258, pp. 1435–1445, 2025.

[9] Y. M. Alsakar, N. Elazab, N. Nader, W. Mohamed, M. Ezzat, and M. Elmogy, "Multi-label dental disorder diagnosis based on MobileNetV2 and swin transformer using bagging ensemble classifier," *Sci. Rep.*, vol. 14, no. 1, pp. 1–23, 2024.

[10] J. Rashid *et al.*, "Skin Cancer Disease Detection using Transfer Learning Technique," *Appl. Sci.*, vol. 12, no. 11, 2022.

[11] X. Li *et al.*, "A Migration Learning Method Based on Adaptive Batch Normalization Improved Rotating Machinery Fault Diagnosis," *Sustain.*, vol. 15, no. 10, 2023.

[12] V. L. Helen Josephine, A. P. Nirmala, and V. L. Alluri, "Impact of Hidden Dense Layers in Convolutional Neural Network to enhance Performance of Classification Model," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1131, no. 1, p. 012007, 2021.

[13] I. Salehin and D. K. Kang, "A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain," *Electron.*, vol. 12, no. 14, 2023.

[14] I. B. Santoso, S. N. Utama, and Supriyono, "Meta-learning based softmax average of convolutional neural networks using multi-layer perceptron for brain tumour classification," *Array*, vol. 26, no. December 2024, p. 100398, 2025.

[15] S. Sathyanarayanan and B. R. Tantri, "Confusion Matrix-Based Performance Evaluation Metrics," no. November, 2024.

Vol. 5, No. 1, June 2025, Page. 11-32
https://jiki.jurnal-id.com
DOI: https://doi.org/10.54082/jiki.264

Halaman Ini Dikosongkan