

Addition of Non-Skin Classes in Skin Type Classification Using EfficientNet-B0 Architecture

Haitsam Muftin Sani^{1*}, Ajie Kusuma Wardhana^{2*}

^{*}Informatika, Universitas Amikom Yogyakarta

haitsam1012@students.amikom.ac.id¹, ajiekusuma@amikom.ac.id²

Article Info

Article history:

Received 2025-07-20

Revised 2025-08-26

Accepted 2025-09-10

Keyword:

*Skin Classification,
Nonskin Integration,
Efficientnet-B0,
Deep Learning,
Image Classification.*

ABSTRACT

Skin type classification is an essential process in dermatology and skincare, aiming to categorize skin conditions such as dry, normal, and oily. However, image-based skin classification models often struggle when confronted with non-skin objects like clothing, background, or hair that are not accounted for in standard datasets. This study proposes a novel approach by integrating a nonskin class into a skin type classification model based on the EfficientNet-B0 architecture. The dataset used consists of images categorized into four classes: dry, normal, oily, and nonskin. The model was trained using transfer learning and optimized through techniques such as data augmentation, learning rate scheduling, and early stopping. The final evaluation achieved an accuracy of 91%, with the nonskin class showing perfect precision and recall. These results demonstrate that incorporating nonskin data can significantly enhance model robustness and accuracy. This research contributes a practical method for improving the reliability of skin classification systems, especially in real-world environments.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Klasifikasi tipe kulit merupakan proses identifikasi dan pengelompokan jenis kulit manusia berdasarkan karakteristik visual maupun biologis seperti kelembaban, kandungan minyak, tekstur, serta warna permukaan kulit [1]. Dalam dunia medis dan kosmetik, klasifikasi ini sangat penting karena membantu menentukan perawatan kulit yang tepat, diagnosis penyakit dermatologis, serta mendukung personalisasi produk kecantikan. Secara umum, tipe kulit dibagi menjadi 3 kategori utama, yaitu kulit normal, kering, dan berminyak [2]. Dengan kemajuan teknologi pengolahan citra dan kecerdasan buatan, pendekatan berbasis deep learning mulai diterapkan secara luas untuk proses klasifikasi ini, sebagaimana ditunjukkan dalam studi oleh [3] yang berhasil mengklasifikasikan tipe kulit manusia menggunakan metode CNN dan teknik augmentasi data berbasis gambar.

Namun, di lapangan, penggunaan teknologi klasifikasi tipe kulit berbasis citra masih menghadapi sejumlah tantangan besar. Salah satu masalah yang sering terjadi adalah kehadiran objek-objek non-kulit dalam gambar masukan, seperti pakaian, rambut, atau bagian dari wajah namun hanya

terfokus pada 1 objek seperti mata, yang dapat menurunkan akurasi klasifikasi. Sayangnya, sebagian besar dataset yang digunakan dalam penelitian terdahulu hanya mencakup area kulit tanpa mempertimbangkan gangguan visual dari lingkungan sekitar. Ini berarti bahwa model yang dilatih tidak dibekali dengan kemampuan untuk membedakan antara fitur kulit dan non-kulit, sehingga mengalami kesalahan prediksi saat diterapkan dalam kondisi nyata. Hal ini diperkuat oleh studi [4], yang menunjukkan bahwa efisiensi CNN dalam diagnosis penyakit kulit sangat dipengaruhi oleh keberagaman dan kualitas dataset yang digunakan. Penelitian [5] juga menyoroti bahwa penyakit kulit yang diklasifikasikan saat ini umumnya terbatas pada jenis yang itu-itu saja, sehingga kurang mampu menangani kompleksitas data nyata.

Sementara itu, perkembangan arsitektur deep learning untuk klasifikasi kulit terus mengalami kemajuan. Arsitektur seperti ResNet50, MobileNet, dan EfficientNet telah banyak digunakan dalam klasifikasi penyakit kulit, termasuk lesi dan kanker kulit. Penelitian oleh [6] melakukan perbandingan antara ketiga arsitektur tersebut dan menunjukkan bahwa EfficientNet-B0 memiliki kinerja yang sangat baik dalam hal akurasi dan efisiensi. Penelitian lain [7], [8] membuktikan

bahwa versi ringan dari EfficientNet dapat diterapkan secara efektif untuk diagnosis penyakit kulit. Penelitian yang dilakukan oleh [9] menggunakan arsitektur EfficientNet-B0 juga menunjukkan hasil yang cukup tinggi dalam klasifikasi tipe kulit. Namun demikian, seiring meningkatnya kompleksitas model, kualitas dataset belum mampu mengejar [10]. Di sisi lain, industri perawatan kulit telah mengalami pertumbuhan pesat dengan munculnya berbagai produk dan obat yang menasar permasalahan kulit umum seperti kulit berminyak, kering, atau sensitif [11]. Ironisnya, teknologi klasifikasi digital masih terbatas dalam mengakomodasi keragaman tersebut akibat keterbatasan data [10].

Untuk menjawab permasalahan tersebut, penelitian ini mengusulkan pendekatan baru dengan menambahkan kelas non-kulit ke dalam dataset klasifikasi tipe kulit. Pendekatan ini bertujuan meningkatkan akurasi dan ketahanan (robustness) model dalam menghadapi kondisi gambar dunia nyata. Dengan menggabungkan data kulit dan non-kulit secara eksplisit, model dapat dilatih untuk memahami konteks visual secara lebih menyeluruh, sehingga mampu membedakan fitur kulit dari gangguan visual di sekitarnya. Hal ini berbeda dengan pendekatan konvensional yang hanya berfokus pada penambahan variasi data non-kulit. Dalam penelitian ini, digunakan arsitektur EfficientNet-B0 karena keunggulannya dalam menggabungkan efisiensi parameter dan akurasi tinggi [12].

Beberapa penelitian terbuka lainnya juga mendukung pendekatan ini. Studi oleh [13] memperkenalkan metode pengukuran tipe kulit secara otomatis menggunakan model Inception-v3 dengan teknik transfer learning dan segmentasi berbasis ROI, yang mampu meningkatkan presisi klasifikasi. Selain itu, studi [14] dalam tinjauan sistematisnya menyoroti kurangnya representasi tipe kulit gelap dalam dataset dermatologi dan menyarankan strategi balancing data untuk meningkatkan performa model. Penelitian [15] juga menunjukkan bahwa integrasi fitur lokal dan kontekstual secara hierarkis dapat mendorong peningkatan akurasi klasifikasi lesi kulit menggunakan pendekatan dual-track CNN. Penelitian-penelitian ini membuktikan bahwa peningkatan akurasi model klasifikasi kulit tidak hanya bergantung pada algoritma, tetapi juga pada kualitas dan keberagaman data yang digunakan.

Dengan latar belakang tersebut, penelitian ini berfokus pada integrasi kelas non-kulit sebagai bentuk augmentasi parameter dalam dataset klasifikasi tipe kulit. Tujuan utama dari pendekatan ini adalah untuk meningkatkan performa model dalam hal akurasi dan generalisasi, khususnya saat diterapkan pada data dunia nyata yang kompleks. Melalui eksperimen berbasis arsitektur EfficientNet-B0, penelitian ini diharapkan dapat memberikan kontribusi nyata terhadap pengembangan sistem klasifikasi kulit berbasis deep learning yang lebih akurat, efisien, dan tahan terhadap gangguan visual.

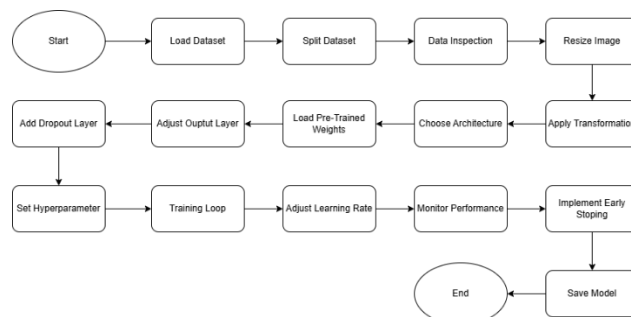
II. METODE

Penelitian ini menggunakan arsitektur EfficientNet-B0 dalam pengembangan sistem klasifikasi citra tipe kulit. Proses penelitian dilakukan dengan memanfaatkan framework PyTorch, serta pustaka pendukung seperti torchvision, numpy, dan matplotlib untuk mendukung pemrosesan data dan visualisasi hasil. Dataset yang digunakan berasal dari dua sumber, yaitu Skin Types 2 untuk citra kulit dan dataset citra acak dari Kaggle sebagai tambahan kelas *nonskin*.

Tahapan penelitian diawali dengan pemuatan dan pembagian dataset ke dalam data pelatihan, validasi, dan pengujian. Setelah itu, dilakukan proses transformasi dan augmentasi data untuk memastikan keseragaman ukuran dan meningkatkan variasi citra. Selanjutnya, model pretrained EfficientNet-B0 dimodifikasi pada bagian output untuk disesuaikan dengan jumlah kelas yang digunakan dalam penelitian ini.

Proses pelatihan dilakukan dengan menyesuaikan parameter-parameter penting seperti learning rate dan fungsi optimasi. Selama pelatihan, dilakukan pemantauan terhadap performa model berdasarkan data validasi. Untuk mencegah overfitting, digunakan strategi early stopping yang akan menghentikan pelatihan saat performa model tidak lagi meningkat. Model terbaik kemudian disimpan dan digunakan untuk evaluasi akhir menggunakan data pengujian.

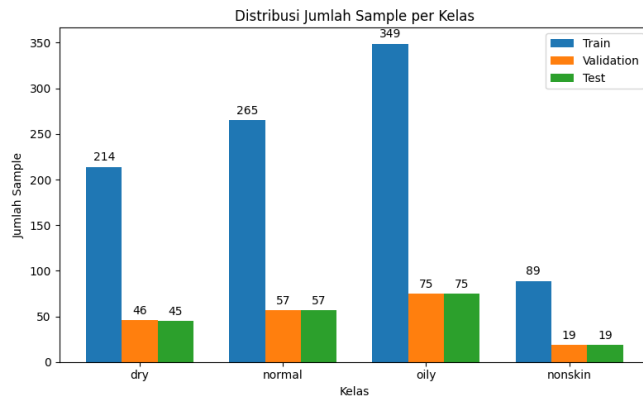
Secara keseluruhan, alur pengembangan penelitian ini mengikuti proses standar dalam pengembangan model deep learning, seperti yang ditunjukkan pada Gambar 1.



Gambar 1. gambar flowchart dari penelitian ini

A. Dataset

Penelitian ini menggunakan dataset Skin Types 2 – Computer Vision Project dari Roboflow yang berisi gambar kulit manusia dengan tiga kategori tipe kulit: *dry*, *normal*, dan *oily*. Untuk meningkatkan kemampuan model dalam mengenali gambar non-kulit, ditambahkan kelas nonskin yang berasal dari kumpulan gambar acak berbagai objek dari dataset Kaggle. Distribusi untuk kelas *dry* sebanyak 350 data, *oily* 499 data, *normal* 379 data, dan *non-skin* 127 data dengan total data sebanyak 1310 data.



Gambar 2. gambar distribusi data dari dataset

Sebagian besar dataset terdiri dari jenis kulit berminyak, diikuti oleh kulit normal, dan terakhir kulit kering. Penelitian ini bertujuan untuk menganalisis jenis kulit pengguna dan memberikan rekomendasi perawatan kulit yang disesuaikan dengan kebutuhan spesifik mereka. Contoh gambar untuk setiap kelas ditampilkan di bawah ini.



Gambar 3. Gambar dari setiap label dalam dataset

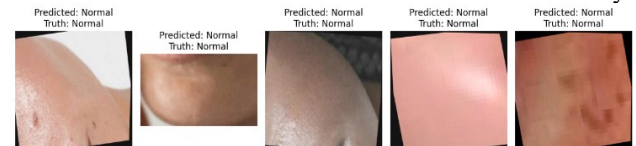
B. Preprocessing Data

Tahapan preprocessing data dalam penelitian ini dilakukan untuk memastikan bahwa seluruh citra memiliki format, dimensi, dan kualitas yang seragam sebelum digunakan dalam proses pelatihan model. Proses diawali dengan inspeksi awal terhadap dataset yang diperoleh dari dua sumber utama, yaitu dataset Skin Types 2 dari Roboflow dan dataset citra yang diambil secara acak dari Kaggle digunakan sebagai kelas tambahan nonskin. Pada tahap ini, dilakukan pengecekan terhadap format file, kesesuaian label, serta validitas citra secara umum.

Langkah selanjutnya adalah proses resize, yaitu mengubah dimensi seluruh citra menjadi ukuran tetap 224×224 piksel. Ukuran ini dipilih karena merupakan standar input pada arsitektur EfficientNet-B0 yang digunakan dalam penelitian ini. Dengan menyeragamkan resolusi citra, proses ekstraksi fitur visual oleh model dapat dilakukan secara lebih konsisten dan efisien. Resize dilakukan secara otomatis menggunakan fungsi dari pustaka torchvision.transforms.

Setelah proses resize, dilakukan penerapan transformasi pada data pelatihan untuk meningkatkan variasi dan kompleksitas data guna mencegah overfitting. Transformasi yang digunakan meliputi pembalikan horizontal secara acak (random horizontal flip), rotasi acak sebesar ± 10 derajat

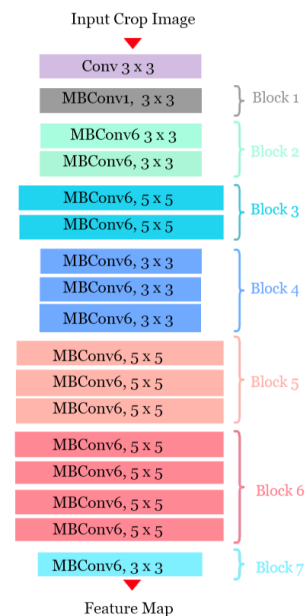
(random rotation), serta konversi citra ke dalam format tensor numerik (ToTensor). Transformasi ini diterapkan khusus pada data pelatihan agar model dapat belajar dari citra dengan variasi sudut dan orientasi yang berbeda, sementara data validasi dan pengujian hanya dikenai resize dan konversi ke tensor agar hasil evaluasi tetap konsisten. Secara keseluruhan, tahapan preprocessing ini bertujuan untuk mempersiapkan citra dalam bentuk yang optimal bagi proses pelatihan model deep learning, dengan mempertahankan fitur penting sekaligus meningkatkan generalisasi model terhadap data dunia nyata.



Gambar 4. Hasil dari preprocessing dan transform data

C. EfficientNet-B0

Mengacu pada studi-studi terdahulu, penelitian ini menggunakan model dasar EfficientNet-B0 sebagai arsitektur utama dalam klasifikasi citra tipe kulit. EfficientNet-B0 merupakan arsitektur CNN modern yang dirancang untuk menghasilkan akurasi tinggi dengan efisiensi komputasi yang optimal. Model ini diperkenalkan oleh [12] dan menjadi bagian dari keluarga EfficientNet yang dikembangkan melalui pendekatan Neural Architecture Search (NAS). Keunggulan utama EfficientNet terletak pada strategi compound scaling, yaitu metode penskalaan model yang secara proporsional memperbesar tiga aspek utama: kedalaman (*depth*), lebar (*width*), dan resolusi gambar (*input resolution*) [12], [16].



Gambar 5. Gambar dari arsitektur EfficientNet-B0 [16]

Gambar arsitektur EfficientNet-B0 terdiri dari tujuh blok

utama (Block 1 hingga Block 7), yang membentuk jalur pemrosesan fitur dari input hingga pembentukan *feature map* akhir. Setiap blok terdiri atas satu atau lebih unit MBConv (Mobile Inverted Bottleneck Convolution), yang merupakan varian dari depthwise separable convolution. MBConv memiliki struktur yang efisien, terdiri dari 1×1 expansion convolution, 3×3 atau 5×5 depthwise convolution, dan 1×1 projection convolution. Beberapa blok juga dilengkapi dengan squeeze-and-excitation (SE) module untuk meningkatkan sensitivitas terhadap fitur penting [12].

Dengan kombinasi struktur yang ringan dan efisien ini, EfficientNet-B0 dapat mencapai akurasi tinggi setara model besar seperti ResNet50, namun dengan jumlah parameter dan waktu pelatihan yang jauh lebih sedikit. Keunggulan ini menjadikannya sangat ideal untuk aplikasi klasifikasi citra medis, terutama yang ditujukan untuk perangkat mobile atau edge computing [12], [16].

TABEL I
PSEUDOCODE

Step	Description
Input:	Skin image dataset divided into three folders: training, validation, and test. Each image has a size of (224, 224, 3). Class labels consist of four categories: dry, oily, normal, and non-skin.
Output:	Compiled deep learning model based on EfficientNetB0, ready for training and evaluation
Load pretrained EfficientNetB0 base model:	Set <code>include_top = False</code> to remove original classification head. Set <code>input_shape = (224, 224, 3)</code> . Set <code>weights = 'imagenet'</code> to use pretrained weights.
Freeze all layers in base model	<code>base_efficientnet.trainable ← False</code>
Define number of output classes	<code>num_classes ← number of columns in y</code>
Construct new classification head:	$x \leftarrow \text{Output of base_efficientnet} \rightarrow x \leftarrow \text{GlobalAveragePooling2D}(x) \rightarrow x \leftarrow \text{Dense}(128 \text{ units, activation} = \text{ReLU})(x) \rightarrow x \leftarrow \text{Dropout}(\text{rate} = 0.5)(x) \rightarrow \text{predictions} \leftarrow \text{Dense}(\text{num_classes, activation} = \text{softmax})(x)$
Assemble final model:	<code>model_efficientnet ← Model(inputs = base_efficientnet.input, outputs = predictions)</code>
Compile the model:	<code>optimizer ← Adam → loss ← categorical_crossentropy → metrics ← ['accuracy'] → model_efficientnet.compile(optimizer, loss, metrics)</code>

D. Pelatihan model

Model EfficientNet-B0 yang digunakan dalam bentuk pretrained, yaitu telah dilatih sebelumnya pada dataset ImageNet, sehingga mampu mengenali fitur visual umum yang berguna dalam klasifikasi citra. Strategi transfer

learning ini memungkinkan proses pelatihan dilakukan lebih efisien dan efektif, terutama ketika dataset yang digunakan tidak terlalu besar. Dalam penelitian ini, bagian akhir dari arsitektur EfficientNet-B0 dimodifikasi untuk menyesuaikan dengan tugas klasifikasi empat kelas, yaitu dry, normal, oily, dan nonskin.

Modifikasi dilakukan dengan mengganti lapisan klasifikasi bawaan model dengan fully connected layer baru yang terdiri dari dua komponen utama, yaitu Dropout layer sebesar 0.5 dan Linear layer dengan empat output neuron. Dropout digunakan sebagai metode regularisasi untuk mengurangi overfitting dengan menonaktifkan sebagian neuron secara acak selama proses pelatihan. Empat output neuron berfungsi untuk mengklasifikasikan citra ke dalam salah satu dari keempat kelas yang tersedia.

Model dikompilasi dengan menggunakan fungsi loss Cross Entropy, yang sesuai untuk klasifikasi multi-kelas. Untuk optimisasi, digunakan Adam optimizer dengan learning rate awal sebesar 0.0001. Penelitian ini juga menggunakan learning rate scheduler berbasis validasi loss, yaitu akan menurunkan learning rate apabila model tidak menunjukkan peningkatan performa pada data validasi selama beberapa epoch. Proses pelatihan dilakukan dengan batch size sebesar 16 dan ditargetkan untuk berjalan selama maksimal 30 epoch. Namun, untuk menjaga efisiensi dan mencegah overfitting, diterapkan strategi early stopping dengan patience sebanyak 5 epoch, yang menyebabkan pelatihan berhenti lebih awal ketika tidak terjadi penurunan nilai loss pada data validasi. Dalam eksperimen ini, pelatihan model berhenti secara otomatis pada epoch ke-23, menunjukkan bahwa model telah mencapai performa optimal lebih awal. Selama pelatihan, performa model dipantau pada setiap epoch menggunakan metrik loss dan akurasi baik pada data pelatihan maupun validasi. Model dengan nilai loss validasi terendah disimpan secara otomatis sebagai model terbaik untuk proses evaluasi. Evaluasi dilakukan menggunakan data uji untuk menilai performa akhir dari model.

Seluruh tahapan ini mencerminkan proses pengembangan model yang sistematis dan terukur, dimulai dari pemanfaatan pretrained model, penyesuaian arsitektur, penentuan *hyperparameter tuning*, hingga strategi pelatihan yang adaptif dan efisien. Dengan pendekatan ini, model diharapkan dapat melakukan klasifikasi tipe kulit dengan akurasi tinggi serta mampu membedakan citra kulit dari citra yang tidak relevan.

E. Evaluasi

Model dievaluasi menggunakan teknik umum seperti akurasi (pada data pelatihan dan validasi), *confusion matrix*, presisi, recall, skor F1, dan uji data gambar dunia nyata tambahan. Model dengan kinerja terbaik kemudian akan diimplementasikan ke dalam aplikasi web untuk pengembangan lebih lanjut. Skor F1 dihitung menggunakan presisi (P) dan recall (R) dengan rumus berikut: TP mewakili True Positive, TN adalah True Negative, FP adalah False Positive, dan FN adalah False Negative.

$$(1) \text{ Precision} = \frac{TP}{(TP + FN)}$$

$$(2) \text{ Recall} = \frac{TP}{(TP + FP)}$$

$$(3) \text{ F1-Score} = 2 \times P \times \frac{R}{(P + R)}$$

Rumus Precision (1) mengukur seberapa tepat prediksi positif model dibandingkan dengan semua prediksi positif yang dihasilkan. Semakin tinggi precision, semakin sedikit kesalahan positif (*false positive*).

Rumus Recall (2) mengukur seberapa baik model menemukan semua data positif yang sebenarnya. Semakin tinggi recall, semakin sedikit kasus positif yang terlewat (*false negative*).

Rumus F1-Score (3) merupakan rata-rata harmonis antara precision dan recall. Cocok digunakan saat ingin menyeimbangkan keduanya, terutama pada data yang tidak seimbang (*imbalanced dataset*).

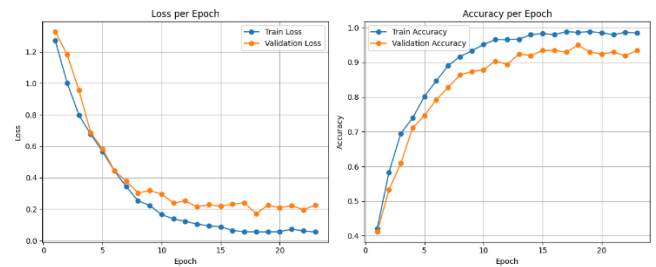
F. Test Split

Evaluasi performa model dilakukan dengan membagi dataset menjadi tiga subset, yaitu *training set*, *validation set*, dan *test set* dengan rasio 70% : 15% : 15% dengan penjabaran training set sebanyak 917 sampel, validation set sebanyak 197 sampel, dan test set sebanyak 196 sampel. Dengan menggunakan pembagian stratifikasi (*stratified split*) bertujuan agar distribusi setiap kelas tetap seimbang pada masing-masing subset. *Training set* digunakan untuk proses pelatihan model, *validation set* digunakan untuk memantau kinerja model selama pelatihan dan mencegah *overfitting* melalui mekanisme *early stopping*, sedangkan *test set* digunakan untuk mengukur performa akhir model pada data yang belum pernah dilihat sebelumnya. Selain pembagian data, evaluasi juga menggunakan beberapa metrik kinerja, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*, yang dihitung berdasarkan *confusion matrix* pada data *validation* dan *test*. Metrik *macro average* digunakan untuk memastikan performa model dievaluasi secara seimbang di seluruh kelas, sedangkan *weighted average* digunakan untuk mempertimbangkan pengaruh distribusi jumlah data per kelas.

III. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan satu model arsitektur EfficientNet-B0 dengan metode transfer learning. Model dilatih pada dataset gabungan yang mencakup tiga kelas kulit (*dry*, *normal*, *oily*) dan satu kelas tambahan *nonskin*. Selama pelatihan, diterapkan strategi *early stopping* untuk mencegah *overfitting*, dengan masa tunggu selama lima epoch. Meskipun model dirancang untuk berjalan hingga 30 epoch, pelatihan berhenti lebih awal pada epoch ke-23 karena nilai *validation loss* tidak menunjukkan penurunan yang signifikan setelahnya. Hal ini menandakan bahwa model telah mencapai

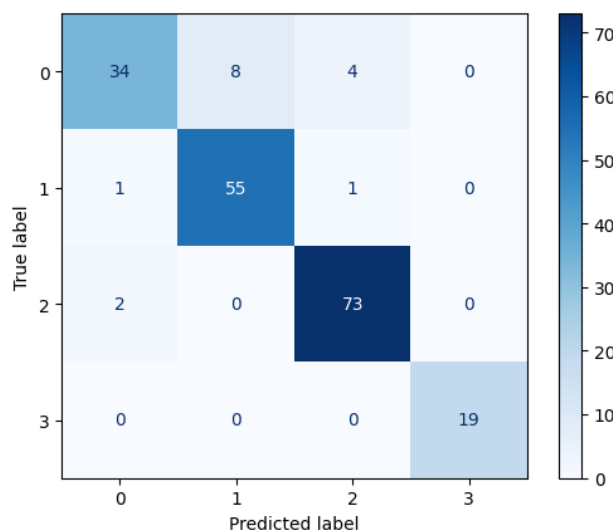
performa terbaiknya pada tahap tersebut dan dihentikan secara otomatis oleh mekanisme *early stopping*.



Gambar 6. Gambar train loss dan validation loss menggunakan kelas non-skin

Perbedaan yang terlihat antara training loss dan validation loss pada epoch pertama mencerminkan proses awal adaptasi model EfficientNet-B0 yang telah dilatih sebelumnya (*pretrained*) terhadap domain klasifikasi citra kulit dan non-kulit. Model ini sebelumnya dioptimalkan untuk klasifikasi umum pada dataset ImageNet, sehingga pada tahap awal pelatihan, bobot-bobotnya perlu menyesuaikan diri dengan karakteristik visual baru dari dataset kulit yang digunakan. Seiring bertambahnya epoch, training loss menurun secara signifikan dari sekitar 1.27 menjadi hampir 0.05, sedangkan validation loss juga turun dari sekitar 1.32 menjadi sekitar 0.18. Penurunan yang konsisten ini disertai dengan jarak yang tetap relatif kecil di antara keduanya, menunjukkan bahwa model mampu beradaptasi tanpa mengalami *overfitting* yang berarti.

Pada grafik Accuracy per Epoch, terlihat peningkatan akurasi pelatihan dari sekitar 42% pada epoch pertama hingga hampir 100% pada akhir pelatihan, sedangkan akurasi validasi meningkat dari 41% menjadi sekitar 95%. Peningkatan ini berlangsung stabil, dengan gap akurasi yang kecil, memperkuat indikasi kemampuan generalisasi model. Selama proses pelatihan, performa model juga dievaluasi menggunakan *confusion matrix* dan metrik klasifikasi seperti *precision*, *recall*, dan *f1-score* untuk mendapatkan gambaran menyeluruh terkait kinerja pada tiap kelas. Visualisasi ini menunjukkan bahwa model dapat mempelajari pola citra kulit dengan baik dan mempertahankan kinerja yang tinggi hingga akhir pelatihan.



Gambar 7. Gambar confusion matrix pada setiap kelas menggunakan kelas nonskin

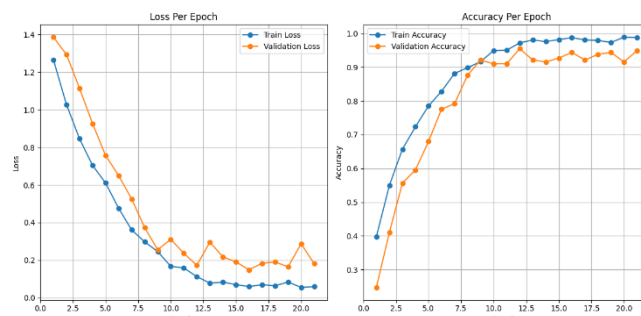
Selanjutnya, kinerja model divisualisasikan pada Gambar 7. Berdasarkan confusion matrix, model menunjukkan performa klasifikasi yang tinggi di semua kelas dengan akurasi keseluruhan sebesar 93%. Pada kelas 3 (nonskin), model mampu mengklasifikasikan semua 19 citra secara akurat tanpa kesalahan (precision dan recall masing-masing 1,00). Kelas 1 (normal) memiliki 57 citra dengan recall 0,96, hanya terdapat dua kesalahan klasifikasi silang dengan kelas 2 (oily). Kelas 2 (oily) dengan 75 citra mencapai recall 0,97, mengalami dua kesalahan prediksi ke kelas 0 (dry). Sementara itu, kelas 0 (dry) memiliki 46 citra dengan recall 0,80, yang sebagian kecil salah diprediksi sebagai kelas 1 atau kelas 2.

TABEL II
HASIL KLASIFIKASI MENGGUNAKAN KELAS NON-SKIN

Kelas	Precision	Recall	F1-Score	Support
0 (dry)	0.95	0.80	0.87	46
1 (normal)	0.90	0.96	0.93	57
2 (oily)	0.94	0.97	0.95	75
3 (nonskin)	1.00	1.00	1.00	19
Accuracy			0.93	197
Macro Avg	0.95	0.94	0.94	197
Weighted Avg	0.94	0.93	0.93	197

Tabel 2 Hasil Klasifikasi menunjukkan bahwa model memperoleh akurasi keseluruhan sebesar 93%, dengan nilai precision tertinggi pada kelas nonskin (1,00) serta precision dan recall yang juga tinggi pada kelas oily (0,94–0,97), normal (0,90–0,96), dan dry (0,95–0,80). Skor *macro average* untuk precision, recall, dan f1-score masing-masing berada pada angka 0,95, 0,94, dan 0,94, menunjukkan bahwa model memberikan performa yang merata di seluruh kelas. *Weighted average* pun mendukung hal ini dengan nilai precision 0,94,

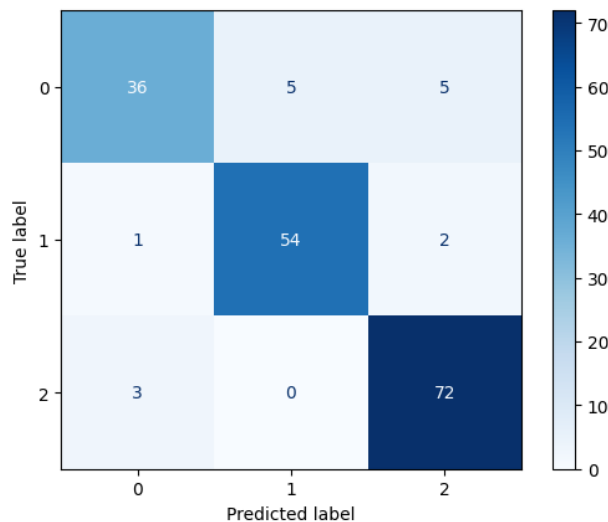
recall 0,93, dan f1-score 0,93, menandakan tidak ada bias signifikan terhadap kelas tertentu.



Gambar 8. Gambar train loss dan validation loss tanpa menggunakan kelas non-skin

Sebagai pembandingan, dilakukan eksperimen dengan menghapus kelas non-skin namun tidak mengubah arsitektur agar pembandingan ini setara. Dapat dilihat dari Gambar 8 perbedaan awal antara training loss (1.27) dan validation loss (1.39) pada epoch pertama menunjukkan bahwa model EfficientNet-B0 masih dalam tahap adaptasi terhadap domain klasifikasi citra kulit sehat dengan tiga kelas, yaitu dry, oily, dan normal. Seiring bertambahnya epoch, training loss menurun drastis hingga sekitar 0.05 pada epoch ke-20, sementara validation loss juga turun konsisten hingga sekitar 0.18. Perbedaan yang relatif kecil antara keduanya menunjukkan bahwa model mampu mempelajari pola data tanpa mengalami overfitting yang signifikan, meskipun terdapat sedikit fluktuasi pada validation loss di epoch 13 dan 20.

Pada grafik akurasi, terlihat peningkatan yang signifikan. Akurasi pelatihan meningkat dari sekitar 39% pada epoch pertama hingga hampir 99% pada epoch ke-20, sedangkan akurasi validasi naik dari 25% menjadi sekitar 95%. Peningkatan ini berlangsung stabil dengan gap yang kecil antara training accuracy dan validation accuracy, menandakan kemampuan generalisasi model yang baik. Early stopping pada epoch ke-21 menunjukkan bahwa model telah mencapai performa optimal sebelum potensi overfitting terjadi.



Gambar 9. Gambar confusion matrix pada setiap kelas tanpa menggunakan kelas nonskin

Selanjutnya, kinerja model divisualisasikan pada Gambar 7. Berdasarkan confusion matrix, model menunjukkan performa klasifikasi yang tinggi di semua kelas dengan akurasi keseluruhan sebesar 94%. Pada kelas 0 (dry), dari 46 citra terdapat 36 yang berhasil diprediksi dengan benar, sementara 5 citra salah diprediksi sebagai normal dan 5 lainnya sebagai oily, sehingga recall tercatat sebesar 0,85. Pada kelas 1 (normal) dengan total 57 citra, model mampu mengklasifikasikan 54 citra secara benar, hanya terjadi 1 kesalahan ke kelas dry dan 2 kesalahan ke kelas oily, menghasilkan recall sebesar 0,96. Sementara itu, pada kelas 2 (oily) dengan 75 citra, sebanyak 72 berhasil diprediksi dengan benar, sedangkan 3 citra salah diprediksi sebagai dry, sehingga recall mencapai 0,97. Secara keseluruhan, hasil ini menunjukkan bahwa model mampu mengklasifikasikan ketiga tipe kulit dengan baik, meskipun masih terdapat kesalahan minor pada kelas dry yang lebih sering terklasifikasi silang ke kelas lain.

TABEL III
HASIL KLASIFIKASI TANPA KELAS NON-SKIN

Kelas	Precision	Recall	F1-Score	Support
0 (dry)	0.97	0.85	0.91	46
1 (normal)	0.92	0.96	0.94	57
2 (oily)	0.94	0.97	0.95	75
Accuracy			0.94	178
Macro Avg	0.94	0.93	0.93	178
Weighted Avg	0.94	0.94	0.94	178

Dari Tabel 3 dapat dilihat eksperimen ini menunjukkan hasil precision, recall, dan f1-score pada hampir tiap kelas berada di atas 0.90, dengan akurasi keseluruhan sebesar 94%.

Detail performa tiap kelas yaitu: kelas dry memiliki precision 0.97, recall 0.85, dan f1-score 0.91; kelas normal memperoleh precision 0.92, recall 0.96, dan f1-score 0.94; serta kelas oily mencapai precision 0.94, recall 0.97, dan f1-score 0.95. Hasil ini ditunjang oleh tren penurunan loss serta peningkatan akurasi per epoch pada grafik pelatihan, yang konsisten hingga tercapainya performa optimal yang menghasilkan akurasi sebesar 94%.

Sebagai pembandingan lain, penelitian [9] menerapkan arsitektur EfficientNet-B0 untuk klasifikasi penyakit kulit wajah. Dataset yang digunakan mencakup beberapa kategori penyakit kulit dengan jumlah 1158 sampel yang dibagi menjadi 8 kelas dengan 80% untuk training dan 20% untuk test, model dilatih dengan optimasi hyperparameter tuning seperti learning rate 0.001, batch size 32, dan optimizer Adam. Hasil yang dicapai menunjukkan akurasi sebesar 95,70% pada data uji.

Namun penelitian [9] hanya menggunakan dataset kulit yang terseragam dan tidak terdapat penambahan objek non-skin, sehingga dapat dikategorikan sebagai *clean dataset*. Karakteristik dataset yang seragam ini membuat citra lebih terfokus pada area kulit tanpa adanya gangguan dari latar belakang, pencahayaan yang bervariasi, maupun objek lain di luar kulit. Kondisi tersebut mempermudah proses klasifikasi karena model tidak perlu melakukan adaptasi terhadap variasi data yang kompleks. Namun, di sisi lain, penggunaan dataset yang *clean* ini juga menjadi keterbatasan karena belum sepenuhnya mencerminkan kondisi nyata, di mana citra kulit sering kali bercampur dengan objek non-skin, noise visual, atau perbedaan kualitas pengambilan gambar.

Dari beberapa penelitian pembandingan ini menunjukkan bahwa penelitian utama (Penambahan Kelas Non-Kulit dalam Klasifikasi Tipe Kulit Menggunakan Arsitektur EfficientNet-B0) mampu menghasilkan akurasi yang dapat bersaing sekaligus mampu mengklasifikasikan kelas non-skin dengan baik. Penambahan kelas non-skin ini menambah ketahanan (robustness) dari model prediksi, karena model tidak hanya belajar mengenali jenis kulit, tetapi juga mampu membedakan citra yang bukan kulit sehingga lebih relevan untuk aplikasi nyata.

IV. KESIMPULAN

Penelitian ini berhasil menunjukkan bahwa penambahan kelas non-kulit ke dalam sistem klasifikasi tipe kulit berbasis EfficientNet-B0 dapat bersaing dan meningkatkan ketahanan model terhadap gangguan visual dari lingkungan nyata. Dengan menggunakan pendekatan *transfer learning* dan strategi pelatihan adaptif seperti *early stopping* dan *learning rate scheduling*, model berhasil mencapai akurasi 93% dan menunjukkan performa yang sangat baik terutama dalam mengidentifikasi kelas nonskin dengan nilai precision, recall, dan f1-score sempurna (1,00). Hasil evaluasi juga memperlihatkan bahwa semua kelas memperoleh nilai

precision, recall, dan f1-score yang tinggi serta seimbang, dengan *macro average* masing-masing sebesar 0,95, 0,94, dan 0,94. Temuan ini membuktikan bahwa penambahan kelas non-kulit tidak hanya efektif dalam menurunkan tingkat kesalahan prediksi, tetapi juga berkontribusi signifikan terhadap kemampuan generalisasi model. Penelitian ini membuka peluang baru dalam pengembangan sistem klasifikasi kulit yang lebih adaptif dan akurat di lingkungan nyata.

DAFTAR PUSTAKA

- [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, p. 14, Jan. 2023, doi: 10.3390/cosmetics10010014.
- [2] S. Saidah, Y. N. Fuadah, F. Alia, N. Ibrahim, R. Magdalena, and S. Rizal, "Facial Skin Type Classification Based on Microscopic Images Using Convolutional Neural Network (CNN)," 2021, pp. 75–83. doi: 10.1007/978-981-33-6926-9_7.
- [3] S. Saiwaeo, S. Arwatchananukul, L. Mungmai, W. Preedalikit, and N. Aunsri, "Human skin type classification using image processing and deep learning approaches," *Heliyon*, vol. 9, no. 11, p. e21176, 2023, doi: 10.1016/j.heliyon.2023.e21176.
- [4] O. Iparraguirre-Villanueva and M. Cabanillas-Carbonell, "Application of Convolutional Neural Networks in Skin Disease Prediction: Accuracy and Efficiency in Dermatological Image Analysis," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 21, no. 02, pp. 18–37, Feb. 2025, doi: 10.3991/ijoe.v21i02.52871.
- [5] S. Inthiyaz *et al.*, "Skin disease detection using deep learning," *Advances in Engineering Software*, vol. 175, p. 103361, Jan. 2023, doi: 10.1016/j.advengsoft.2022.103361.
- [6] M. Y. Shams *et al.*, "Skin Disease Classification: A Comparison of ResNet50, MobileNet, and Efficient-B0," *Journal of Current Multidisciplinary Research*, vol. 1, no. 1, pp. 1–7, Jan. 2025, doi: 10.21608/jcmr.2025.327880.1002.
- [7] H. Van Thanh, N. D. Quang, T. M. Phuong, K.-H. Jo, and V.-D. Hoang, "A compact version of EfficientNet for skin disease diagnosis application," *Neurocomputing*, vol. 620, p. 129166, Mar. 2025, doi: 10.1016/j.neucom.2024.129166.
- [8] M. Harahap *et al.*, "Skin cancer classification using EfficientNet architecture," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 4, pp. 2716–2728, Aug. 2024, doi: 10.11591/eei.v13i4.7159.
- [9] P. Khairunnisa, W. E. Putra, W. Yitong, A. Jufrizal, and M. N. A. Makmum, "Convolutional Neural Networks Using EfficientNetB0 Architecture and Hyperparameters on Skin Disease Classification," *Public Research Journal of Engineering, Data Technology and Computer Science*, vol. 2, no. 1, pp. 127–137, 2025.
- [10] T. Talaei Khoei, H. Ould Slimane, and N. Kaabouch, "Deep learning: systematic review, models, challenges, and research directions," *Neural Comput Appl*, vol. 35, no. 31, pp. 23103–23124, Nov. 2023, doi: 10.1007/s00521-023-08957-4.
- [11] A. Xie, X. Zhang, Q. Huang, and J. Wu, "Utilizing Untargeted Lipidomics Technology to Elucidate Differences in Lipid Compositions Among Sensitive Dry, Sensitive Oily and Healthy Skin Types," *Metabolites*, vol. 15, no. 5, p. 292, Apr. 2025, doi: 10.3390/metabo15050292.
- [12] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," 2019.
- [13] P. Zhang and D. Chaudhary, "Hybrid Deep Learning Framework for Enhanced Melanoma Detection," Jul. 2024, [Online]. Available: <http://arxiv.org/abs/2408.00772>
- [14] N. Alipour, T. Burke, and J. Courtney, "Skin Type Diversity in Skin Lesion Datasets: A Review," *Curr Dermatol Rep*, vol. 13, no. 3, pp. 198–210, Aug. 2024, doi: 10.1007/s13671-024-00440-0.
- [15] K. Ramamurthy, I. Thayumanaswamy, M. Radhakrishnan, D. Won, and S. Lingaswamy, "Integration of Localized, Contextual, and Hierarchical Features in Deep Learning for Improved Skin Lesion Classification," *Diagnostics*, vol. 14, no. 13, p. 1338, Jun. 2024, doi: 10.3390/diagnostics14131338.
- [16] J. Raihan and S. A. Ridita, "PhytoCare : A hybrid approach for identifying Rice , Potato and Corn diseases PhytoCare : A hybrid approach for identifying Rice , Potato and Corn diseases," no. February, 2024.