

# Klasifikasi Multi-Kelas Penyakit Kulit berbasis *Transfer Learning* dan *Explainable AI*

## *Multi-Class Skin Disease Classification using Transfer Learning and Explainable AI*

<sup>1</sup>Ayutri Wahyuni\*, <sup>2</sup>Resky Ayu Dewi Talasari, <sup>3</sup>Muhammad Syawal Idil Fitrah Baharuddin

<sup>1,2,3</sup>Program Studi Informatika, Fakultas Teknik, Universitas Fajar

<sup>3</sup>Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, UIN Alauddin Makassar

<sup>1,2,3</sup>Jl. Prof. Dr. H. Abdurrahman Basalamah No. 101 (Ex. Jl. Racing Center), Panakkukang, Kota Makassar, Sulawesi Selatan 90231

<sup>3</sup>Jl. Sultan Alauddin No.63, Mangasa, Kec. Tamalate, Kota Makassar, Sulawesi Selatan 90221

\*e-mail: [ayutri@unifa.ac.id](mailto:ayutri@unifa.ac.id)

(received: 30 May 2026, revised: 12 June 2026, accepted: 13 June 2026)

### Abstrak

Penelitian ini mengkaji penerapan transfer learning dan *Explainable Artificial Intelligence* (XAI) untuk klasifikasi multi-kelas penyakit kulit. Dataset yang digunakan bersumber dari Kaggle Skin Diseases Image Dataset dan mencakup 29.153 citra asli dari sepuluh kelas penyakit kulit. Untuk mengurangi bias kelas mayoritas, data disetarakan melalui directed undersampling menjadi 12.000 citra, yaitu 1.200 citra untuk setiap kelas. Tiga arsitektur pretrained CNN, yaitu EfficientNetB0, ResNet50, dan DenseNet201, diimplementasikan dan dievaluasi menggunakan confusion matrix, akurasi, precision, recall, dan F1-score. Hasil pengujian menunjukkan bahwa DenseNet201 mencapai performa tertinggi dengan akurasi 0,8779, precision 0,8751, recall 0,8748, dan F1-score 0,8745, diikuti ResNet50 dengan akurasi 0,8629 dan EfficientNetB0 dengan akurasi 0,8269. Interpretabilitas model dianalisis menggunakan Grad-CAM, SHAP, dan LIME. Grad-CAM menunjukkan area aktivasi utama model berada pada pusat dan batas lesi; SHAP mengidentifikasi kontribusi fitur dominan pada region lesi dan variasi pigmentasi; sedangkan LIME menegaskan superpiksel lokal yang mendukung keputusan prediksi, terutama pada tepi, warna, dan tekstur lesi. Temuan penelitian mengindikasikan bahwa pendekatan transfer learning yang dikombinasikan dengan XAI berpotensi menjadi fondasi pengembangan sistem pendukung keputusan klinis pada domain klasifikasi citra dermatologis. Penelitian lanjutan yang mencakup validasi data eksternal, penyeimbangan distribusi kelas, serta penilaian interpretasi klinis oleh tenaga ahli dermatologi diperlukan untuk mempersiapkan sistem menuju penerapan dalam konteks layanan kesehatan yang nyata.

**Kata kunci:** convolutional neural networks, explainable artificial intelligence, klasifikasi multi-kelas, penyakit kulit, transfer learning.

### Abstract

*This study investigates the application of transfer learning and Explainable Artificial Intelligence (XAI) for multi-class skin disease classification. The dataset was obtained from the Kaggle Skin Diseases Image Dataset and consists of 29,153 original images spanning 10 skin disease classes. To reduce the bias introduced by class imbalance, the dataset was balanced through directed undersampling, resulting in 12,000 images, with 1,200 images per class. Three pretrained convolutional neural network (CNN) architectures—EfficientNetB0, ResNet50, and DenseNet201—were implemented and evaluated using a confusion matrix, accuracy, precision, recall, and F1-score. The experimental results demonstrate that DenseNet201 achieved the highest classification performance, with an accuracy of 0.8779, precision of 0.8751, recall of 0.8748, and F1-score of 0.8745, outperforming ResNet50 (accuracy: 0.8629) and EfficientNetB0 (accuracy: 0.8269). Model interpretability was investigated using Grad-CAM, SHAP, and LIME. Grad-CAM highlighted that the models primarily focused on the central and peripheral regions of skin lesions during prediction. SHAP identified the dominant contribution of lesion regions and pigmentation patterns to the classification process, while LIME emphasized the importance of local superpixels associated with*

<http://sistemasi.ftik.unisi.ac.id>

*lesion boundaries, color, and texture in supporting the model's predictions. The findings indicate that combining transfer learning with Explainable AI provides a promising foundation for developing clinical decision support systems for dermatological image classification. Future research should incorporate external dataset validation, more robust class balancing strategies, and clinical interpretation by dermatology experts to facilitate the deployment of such systems in real-world healthcare settings.*

Keywords: CNN, explainable artificial intelligence, multi-class classification, skin diseases, transfer learning.

## 1 Pendahuluan

Penyakit kulit mencakup gangguan infeksi, inflamasi, pigmentasi, tumor jinak, dan keganasan yang dapat dikenali melalui perubahan warna, tekstur, bentuk, serta batas lesi. Paparan sinar matahari yang berlebihan tanpa perlindungan dapat menyebabkan sunburn, gangguan pigmentasi, penuaan dini, serta meningkatkan risiko kanker kulit [1, 2]. Keragaman manifestasi klinis serta kemiripan visual antarkelas menyebabkan identifikasi penyakit kulit tidak selalu mudah. Pada penyakit kulit wajah, pemeriksaan yang terutama mengandalkan pengamatan visual dan riwayat pasien masih berpotensi menghasilkan kesalahan analisis, sehingga diperlukan dukungan klasifikasi citra yang lebih objektif [3].

Diagnosis penyakit kulit secara konvensional dilakukan melalui anamnesis, pemeriksaan fisik, dermoskopi, dan pemeriksaan histopatologi melalui biopsi pada indikasi tertentu. Citra dermoskopi mampu memperlihatkan struktur bawah permukaan lesi yang tidak selalu terlihat secara langsung, sedangkan biopsi dan histopatologi tetap digunakan untuk konfirmasi diagnosis. Namun, proses diagnostik tersebut memerlukan kompetensi klinis, waktu, fasilitas, dan pada biopsi melibatkan tindakan invasif. Penelitian pada citra dermoskopi menunjukkan bahwa analisis komputasional berbasis deep transfer learning dapat membantu membedakan lesi benign dan malignant serta mendukung deteksi dini, tetapi hasil model tetap harus diposisikan sebagai alat bantu dan diverifikasi secara klinis [4].

Perkembangan Artificial Intelligence (AI), khususnya Convolutional Neural Network (CNN), membuka peluang klasifikasi penyakit kulit secara cepat dan non-invasif melalui analisis citra. CNN dapat mempelajari fitur visual secara hierarkis, termasuk warna, tekstur, bentuk, dan batas lesi. Penerapan VGG-16 pada klasifikasi kanker kulit menunjukkan bahwa arsitektur CNN dapat membedakan citra lesi dengan performa yang kompetitif [5]. Studi lain yang membandingkan ResNet-152 dan AlexNet [6], InceptionV3 dan EfficientNetB0 [7], serta pendekatan CNN dengan hybrid preprocessing [8] juga memperlihatkan bahwa pemilihan arsitektur dan prapemrosesan memengaruhi kinerja klasifikasi. Pada klasifikasi penyakit kanker kulit, DenseNet121 dilaporkan memiliki performa lebih baik dibandingkan VGG16 [9]. Temuan-temuan tersebut menegaskan relevansi transfer learning dan CNN untuk mengembangkan klasifikasi multi-kelas pada dataset dermatologi.

Meskipun memiliki performa tinggi, model deep learning sering bersifat black box sehingga alasan di balik prediksi sulit ditelusuri. Penelitian klasifikasi mengenai kanker kulit dengan berbasis deep transfer learning menunjukkan bahwa integrasi Explainable Artificial Intelligence (XAI) diperlukan untuk meningkatkan transparansi keputusan model [10]. Grad-CAM dapat menampilkan area citra yang menjadi fokus model melalui peta aktivasi, SHAP dapat memperlihatkan bagian citra yang berkontribusi terhadap keputusan prediksi sedangkan LIME memberikan penjelasan lokal melalui pemilihan region atau superpiksel yang mendukung klasifikasi [11, 12]. Penggunaan ketiga metode secara komplementer memungkinkan evaluasi interpretabilitas dari sisi lokasi aktivasi, kontribusi fitur, dan konsistensi penjelasan lokal.

Berdasarkan latar belakang tersebut, penelitian ini memiliki tujuan untuk mengembangkan sistem klasifikasi multi-kelas penyakit pada kulit berbasis transfer learning dengan membandingkan tiga arsitektur CNN, yaitu EfficientNetB0, ResNet50 dan DenseNet201. Penelitian ini memperluas cakupan klasifikasi dibandingkan studi sebelumnya yang umumnya terbatas pada klasifikasi biner atau pada sebagian kecil kelas penyakit. Selain itu, penelitian ini memanfaatkan kombinasi XAI, yaitu

Grad-CAM, SHAP, dan LIME, untuk meningkatkan interpretabilitas model serta memvisualisasikan area citra yang paling berpengaruh terhadap prediksi.

Kontribusi penelitian ini meliputi empat aspek utama. Pertama, penelitian menangani ketidakseimbangan data melalui *directed undersampling* sehingga setiap kelas memiliki jumlah sampel yang sama dan evaluasi tidak didominasi kelas mayoritas. Kedua, penelitian membandingkan tiga arsitektur pretrained CNN dengan head klasifikasi kustom yang memanfaatkan *Global Average Pooling* dan *Dropout 0,4* untuk mengurangi risiko *overfitting*. Ketiga, penelitian menerapkan skenario pelatihan multi-fase melalui *progressive unfreezing* dan penjadwalan *Cosine Decay* agar adaptasi bobot pretrained terhadap citra penyakit kulit berlangsung lebih stabil. Keempat, penelitian mengintegrasikan Grad-CAM, SHAP, dan LIME sebagai kerangka Tri-XAI untuk menilai apakah keputusan model didasarkan pada region lesi yang relevan secara visual dan klinis.

## 2 Tinjauan Literatur

Beberapa penelitian terdahulu yang relevan menjadi landasan bagi penelitian untuk berkembang. Klasifikasi citra ISIC2020 dilakukan menggunakan model CNN pra-latih ResNet101 dan DenseNet201 dengan dua kategori, yaitu benign dan malignant. Evaluasi kinerja sistem menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, *confusion matrix*, serta *ROC curve*. Hasil eksperimen menunjukkan bahwa ResNet101 memberikan performa terbaik dengan akurasi 87,00%, AUC sebesar 0,943, serta *recall* 89,6% pada deteksi kelas malignant. [13].

Penelitian yang dilakukan oleh Indra Raditya menunjukkan bahwa DenseNet121 memberikan performa terbaik dengan akurasi sebesar 95% dan *macro F1-score* 96%. Sebagai perbandingan, ResNet50 menghasilkan akurasi 91% dengan *F1-score* 94%. Hasil evaluasi menggunakan *confusion matrix* dan visualisasi Grad-CAM juga memperlihatkan bahwa DenseNet121 lebih konsisten dalam mengenali area lesi utama, terutama pada kelas-kelas yang memiliki kemiripan visual tinggi. Temuan ini menunjukkan bahwa model tidak hanya memiliki akurasi tinggi, tetapi juga bersifat interpretatif, sehingga berpotensi digunakan sebagai sistem pendukung diagnosis penyakit kulit berbasis citra [14].

Disisi lain Penelitian Saputra et al membandingkan performa ResNet-152 dan AlexNet untuk klasifikasi kanker kulit pada dataset ISIC 2019, yang terbagi menjadi kelas benign dan malignant. Penelitian ini menggunakan optimizer Adam dan Mini-Batch Gradient Descent (MBGD). Hasil evaluasi menunjukkan bahwa ResNet-152 memberikan kinerja terbaik dengan akurasi 87,85%, *precision* 87,25%, *recall* 88,65%, dan *F1-score* 87,95%, menegaskan efektivitas transfer learning dalam meningkatkan akurasi klasifikasi pada dataset medis meskipun model awal dilatih untuk tugas umum [15].

Dalam penelitian ini, Arlingga et al sistem klasifikasi kanker kulit dikembangkan menggunakan The-HAM10000 Dataverse dengan algoritma CNN, dan performanya dievaluasi melalui model InceptionV3 serta EfficientNetB0. Hasil pengujian menunjukkan bahwa InceptionV3 mencapai akurasi 0,7681, sedangkan EfficientNetB0 mencapai akurasi 0,9809, menandakan bahwa EfficientNetB0 memiliki kinerja lebih unggul dalam deteksi kanker kulit. Dengan tingkat akurasi tersebut, model EfficientNetB0 mampu mengklasifikasikan 327 data secara tepat, sehingga model ini dipilih sebagai metode CNN utama untuk klasifikasi karena memberikan keseimbangan optimal antara akurasi dan error validasi [16].

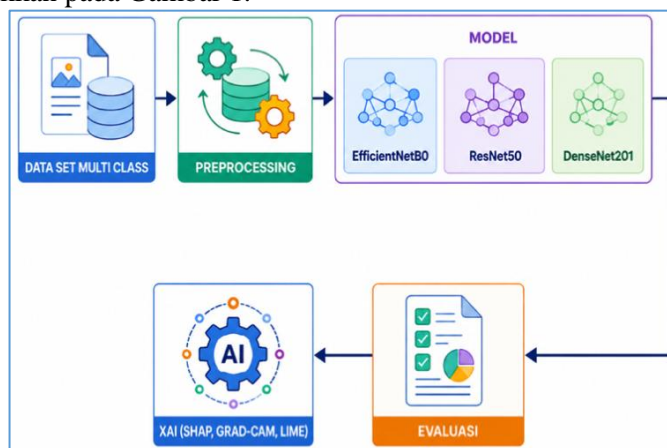
Penelitian oleh Septhya menunjukkan bahwa arsitektur DenseNet121 memberikan akurasi yang lebih tinggi dibandingkan arsitektur VGG16, dengan akurasi 93% untuk data pelatihan dan 79% untuk data pengujian, sedangkan arsitektur VGG16 mencapai akurasi 80% untuk data pelatihan dan 74% untuk data pengujian [17]. Hasil ini menunjukkan bahwa arsitektur DenseNet121 lebih unggul dalam klasifikasi kanker kulit, memberikan informasi penting untuk diagnosis yang lebih akurat.

Selain performa klasifikasi, interpretabilitas menjadi aspek krusial dalam pengembangan sistem diagnosis berbasis AI, karena akurasi tinggi saja tidak cukup untuk menjamin adopsi klinis. Alkhalaf et al. (2023) menekankan pentingnya Explainable Artificial Intelligence (XAI) dalam konteks ini. XAI memungkinkan model untuk memberikan penjelasan yang dapat dipahami pengguna mengenai faktor-faktor yang memengaruhi keputusan prediksi, misalnya area citra kulit yang paling menentukan klasifikasi. Dengan demikian, tenaga medis dapat menilai dan memverifikasi alasan di balik setiap prediksi model, meningkatkan kepercayaan dan transparansi sistem, serta meminimalkan risiko kesalahan diagnosis. Implementasi XAI juga memungkinkan deteksi potensi bias model,

evaluasi konsistensi prediksi antar kasus, dan mendukung penggunaan AI sebagai alat bantu keputusan klinis yang dapat dipertanggungjawabkan [18].

### 3 Metode Penelitian

Penelitian ini menggunakan pendekatan deep transfer learning dengan membandingkan tiga model pretrained, yaitu EfficientNetB0, ResNet50, dan DenseNet201, yang telah dilatih pada dataset ImageNet. Ketiga model tersebut kemudian dilakukan fine-tuning untuk klasifikasi multi-kelas penyakit kulit berdasarkan citra lesi kulit dari Kaggle Skin Diseases Image Dataset. Performa model dievaluasi menggunakan accuracy, precision, recall, F1-score, dan confusion matrix. Model terbaik selanjutnya dianalisis menggunakan SHAP, Grad-CAM dan LIME untuk memberikan penjelasan visual pada area-area citra yang paling berpengaruh dalam proses prediksi. Alur penelitian secara keseluruhan ditunjukkan pada Gambar 1.



Gambar 1 Alur penelitian

Sebagaimana diperlihatkan pada Gambar 1, metode penelitian diawali dengan pengumpulan citra lesi kulit dari Kaggle Skin Diseases Image Dataset, kemudian dilanjutkan dengan tahap prapemrosesan agar data siap digunakan dalam pelatihan model. Penelitian yang diangkat ini mengevaluasi tiga arsitektur CNN berbasis transfer learning, yaitu EfficientNetB0, ResNet50 dan DenseNet201, untuk klasifikasi multi-kelas penyakit kulit. Setelah pelatihan, performa model dibandingkan menggunakan confusion matrix dan metrik evaluasi, kemudian model terbaik dianalisis menggunakan pendekatan XAI berupa SHAP, Grad-CAM, dan LIME.

#### 1. Dataset

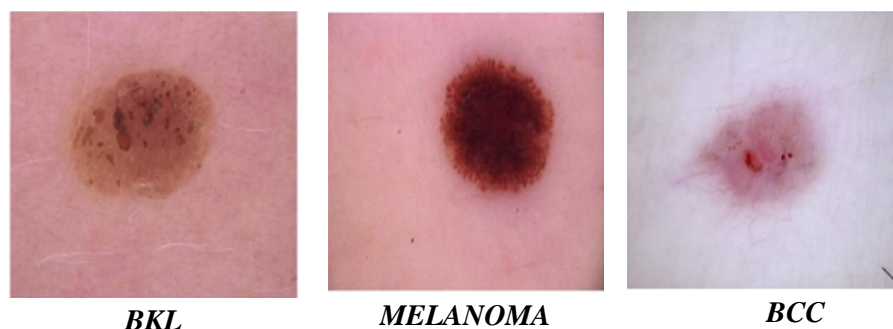
Penelitian ini menggunakan Kaggle Skin Diseases Image Dataset [17] yang terdiri atas citra lesi kulit dalam 10 kelas penyakit kulit untuk klasifikasi multi-kelas. Dataset asli memuat 29.153 citra dengan distribusi tidak seimbang antarkelas. Kelas terbanyak yaitu Melanocytic nevi (NV) sebanyak 7.970 citra, diikuti Basal cell carcinoma (BCC) sebanyak 3.323 citra dan Melanoma sebanyak 3.140 citra. Kelas lainnya meliputi BKL, Warts/viral infections, Psoriasis, Seborrheic keratoses, Tinea/fungal infections, Eczema, dan Atopic dermatitis. Ketidakseimbangan jumlah data ini menjadi aspek penting yang perlu diperhatikan dalam proses pelatihan model klasifikasi multi-kelas. Rincian distribusi dataset asli disajikan pada Tabel 1.

Tabel 1 Distribusi dataset

Nama	Jumlah
<i>Basal cell carcinoma (BCC)</i>	3323
<i>Tinea, ringworm, candidiasis, and other fungal infections</i>	1702
<i>Eczema</i>	1677
<i>Melanocytic nevi (NV)</i>	7970
<i>Warts, molluscum, and other viral infections</i>	2103
<i>Atopic Dermatitis</i>	1257
<i>Seborrheic keratoses and other benign tumors</i>	1847

<i>Melanoma</i>	3140
<i>Psoriasis, lichen planus, and related diseases</i>	2055
<i>Benign keratosis-like lesion (BKL)</i>	2079

Untuk mengatasi ketidakseimbangan tersebut, penelitian ini menggunakan directed undersampling. Dari total data asli diambil data acak sebanyak  $N=1.200$  citra untuk setiap kelas, sehingga total data yang digunakan setelah penyeimbangan adalah 12.000 citra dari 10 kelas. Pendekatan ini bertujuan mengurangi bias terhadap kelas mayoritas dan memberi peluang yang lebih seimbang bagi model untuk mempelajari fitur kelas minoritas. Contoh citra dari beberapa kelas setelah proses pemilihan data ditampilkan pada Gambar 2.



**Gambar 2 Sampel dataset**

Pada Gambar 2 menampilkan tiga contoh citra dari total 10 kelas penyakit kulit, yaitu benign keratosis-like lesion (BKL), melanoma, dan basal cell carcinoma (BCC). Ketiga contoh tersebut merepresentasikan variasi visual antarkelas pada dataset, terutama dalam aspek warna, bentuk, tekstur, dan batas lesi.

## 2. Prapemrosesan Data

ahap prapemrosesan data dilakukan untuk mempersiapkan citra penyakit kulit agar sesuai dengan kebutuhan input model deep learning. Dataset dibagi menjadi data pelatihan dan validasi dengan rasio 80:20 menggunakan teknik stratified split, sehingga distribusi setiap kelas tetap seimbang dan merepresentasikan kondisi dataset asli. Seluruh citra kemudian di-resize menjadi  $224 \times 224$  piksel agar kompatibel dengan arsitektur model pretrained yang digunakan. Setelah itu, citra dikonversi menjadi tensor dan dinormalisasi menggunakan parameter standar dari model pretrained berbasis ImageNet. Proses augmentasi data hanya diterapkan pada data pelatihan dengan berbagai transformasi, seperti rotasi, skala (scaling), translasi, dan flipping horizontal maupun vertikal, tanpa mengubah label kelas. Tahapan ini bertujuan untuk meningkatkan variasi data, memperkuat kemampuan generalisasi model, serta mengurangi risiko overfitting selama proses pelatihan.

## 3. Transfer Learning dan Fine Tuning

Penelitian ini menerapkan pendekatan transfer learning menggunakan model pretrained berbasis ImageNet untuk meningkatkan efektivitas pelatihan pada data citra medis yang terbatas. Arsitektur EfficientNetB0, ResNet50, dan DenseNet201 digunakan sebagai model pembandingan dalam klasifikasi multi-kelas penyakit kulit karena memiliki kemampuan ekstraksi fitur yang baik. Model dengan performa terbaik selanjutnya dianalisis menggunakan metode XAI untuk meningkatkan interpretabilitas hasil prediksi penyakit kulit.

### a. EfficientNetB0

EfficientNetB0 dipilih dalam penelitian ini karena memiliki arsitektur yang efisien dan mampu menghasilkan performa klasifikasi citra yang baik dengan jumlah parameter yang relatif kecil [14]. Model ini menggunakan pendekatan compound scaling yang menyeimbangkan kedalaman, lebar, dan resolusi jaringan, sehingga proses ekstraksi fitur dapat dilakukan secara optimal tanpa

<http://sistemasi.ftik.unisi.ac.id>

membutuhkan beban komputasi yang besar. Komponen seperti MBConv, depth-wise convolution, dan fungsi aktivasi Swish memungkinkan EfficientNetB0 mengenali pola visual halus pada citra lesi kulit, seperti perbedaan warna, tekstur, dan bentuk lesi. Oleh karena itu, model ini sesuai digunakan untuk klasifikasi multi-kelas penyakit kulit karena mampu menjaga keseimbangan antara akurasi, efisiensi komputasi, dan kemampuan generalisasi model.

#### b. ResNet50

ResNet50 digunakan karena memiliki kemampuan yang baik dalam melatih jaringan yang dalam melalui konsep residual learning dan shortcut connection. Mekanisme ini membantu mengatasi masalah vanishing gradient yang sering terjadi pada jaringan saraf berlapis banyak, sehingga proses pembelajaran fitur menjadi lebih stabil [15]. Dalam klasifikasi multi-kelas penyakit kulit, ResNet50 berperan penting dalam mengekstraksi karakteristik visual yang kompleks dari citra lesi, seperti pola tepi, variasi warna, asimetri, dan tekstur permukaan kulit. Model ini dipilih karena telah banyak digunakan dalam klasifikasi citra medis dan memiliki kemampuan yang kuat dalam membedakan beberapa kelas penyakit kulit dengan karakteristik visual yang saling menyerupai.

#### c. DenseNet201

DenseNet201 dipilih karena menerapkan konsep dense connectivity, yaitu setiap lapisan terhubung langsung dengan lapisan-lapisan berikutnya dalam dense block. Struktur ini memungkinkan penggunaan ulang fitur (feature reuse) secara lebih optimal dan memperlancar aliran gradien selama proses pelatihan [16]. Keunggulan tersebut membuat DenseNet201 efektif dalam menangkap detail visual yang kompleks pada citra penyakit kulit, seperti perubahan pigmentasi, pola tekstur, batas lesi, dan variasi warna yang sulit dibedakan antar kelas. Dengan kemampuan ekstraksi fitur yang kuat, DenseNet201 relevan digunakan dalam klasifikasi multi-kelas penyakit kulit, terutama ketika model harus membedakan beberapa jenis lesi yang memiliki kemiripan visual tinggi.

#### d. Skenario Pelatihan Multi-Fase (Dynamic Multi-Phase Training)

Untuk mengoptimalkan proses pemindahan pengetahuan (knowledge transfer) dari bobot pretrained ImageNet menuju dataset penyakit kulit yang baru tanpa merusak fitur generik yang telah dipelajari model, penelitian ini menerapkan strategi pelatihan tiga fase secara progresif (progressive unfreezing). Total epoch yang digunakan bukan satu nilai tunggal, melainkan maksimum 105 epoch yang terbagi menjadi 50 epoch pada head training, 25 epoch pada partial fine-tuning, dan 30 epoch pada full fine-tuning. Laju pembelajaran juga dibedakan pada setiap fase agar proses adaptasi bobot berlangsung stabil.

- Fase 1 (Pelatihan Lapisan Kustom / *Head Training*): Seluruh lapisan pada arsitektur *base model* dibekukan (*frozen*) untuk mengisolasi dan melatih lapisan klasifikasi kustom yang baru dibentuk (lapisan *Global Average Pooling*, *Dropout* 0,4, dan *Softmax*). Fase ini menggunakan laju pembelajaran awal yang relatif tinggi sebesar  $10^{-3}$  dengan pengoptimal Adam selama maksimal 50 *epoch*. Mekanisme *Early Stopping* dengan parameter *patience* sebesar 10 *epoch* diterapkan berbasis pemantauan terhadap *validation loss* guna mencegah *overfitting* dini.
- Fase 2 (Penyetelan Halus Sebagian / *Partial Fine-Tuning*): Kunci pada beberapa blok lapisan teratas dari *base model* mulai dibuka (*fine\_tune\_at* = -20 untuk EfficientNetB0, dan -30 untuk ResNet50 dan DenseNet201) agar model dapat beradaptasi secara spesifik dengan fitur-fitur visual lesi kulit. Laju pembelajaran pada fase ini diturunkan menjadi  $5 \times 10^{-5}$  dan dikombinasikan dengan fungsi penjadwalan *Cosine Decay* untuk memastikan penurunan laju pembelajaran terjadi secara halus seiring bertambahnya iterasi selama maksimal 25 *epoch*.
- Fase 3 (Penyetelan Halus Total / *Full Fine-Tuning*): Seluruh lapisan pada arsitektur jaringan dibuka secara total tanpa pengecualian. Mengingat sensitivitas lapisan konvolusi awal terhadap perubahan bobot yang drastis, fase ini menerapkan skema pelatihan yang sangat hati-hati (*ultra-gentle fine-tuning*) menggunakan laju pembelajaran mula-mula yang sangat kecil senilai  $10^{-5}$  yang dikontrol kembali lewat penjadwalan *Cosine Decay* selama 30 *epoch*.

#### 4. Visualisasi Interpretasi Model dengan Integrasi XAI

Interpretabilitas model dalam penelitian ini ditingkatkan melalui penerapan tiga metode Explainable Artificial Intelligence (XAI), yaitu SHAP (SHapley Additive exPlanations), Grad-CAM, dan LIME. Pendekatan ini diperlukan karena model deep learning seperti EfficientNetB0, ResNet50, dan DenseNet201 memiliki kompleksitas tinggi sehingga hasil prediksinya sulit dipahami secara langsung. SHAP digunakan untuk menjelaskan kontribusi setiap bagian citra terhadap keputusan model, Grad-CAM menyoroti area lesi kulit yang paling berpengaruh melalui heatmap, dan LIME menyediakan penjelasan lokal berbasis model surrogate untuk mendemonstrasikan pengaruh fitur input terhadap prediksi. Dalam klasifikasi multi-kelas penyakit kulit, kombinasi ketiga metode ini memungkinkan sistem tidak hanya menghasilkan prediksi kelas penyakit, tetapi juga memberikan visualisasi dan penjelasan yang transparan sehingga meningkatkan kepercayaan pengguna serta mendukung interpretasi hasil klasifikasi secara objektif.

#### 5. Hyperparameter

Pelatihan model menggunakan AdamW dengan batch size 32, fungsi loss categorical cross-entropy, dan metrik *accuracy*. Skenario pelatihan mengikuti dynamic multi-phase training: fase 1 head training maksimum 50 epoch dengan *learning rate*  $1 \times 10^{-3}$ ; fase 2 partial fine-tuning maksimum 25 epoch dengan *learning rate* awal  $5 \times 10^{-5}$  dan *Cosine Decay*; serta fase 3 *full fine-tuning* maksimum 30 epoch dengan *learning rate* awal  $1 \times 10^{-5}$  dan *Cosine Decay*. *Early Stopping* berbasis validation loss digunakan untuk menghentikan pelatihan lebih awal apabila performa validasi tidak membaik. Rangkuman hyperparameter dan skenario pelatihan disajikan pada Tabel 2.

**Tabel 2 Hyperparameter**

Hyperparameter	Nilai
<i>Optimizer</i>	<i>AdamW</i>
<i>Batch Size</i>	32
<i>Loss Function</i>	<i>Categorical cross-entropy</i>
<i>Metrics</i>	<i>Accuracy</i>
<i>Monitor</i>	<i>Validation Loss</i>
Strategi Pelatihan	<i>Dynamic multi-phase training</i>
Fase 1	50 epoch; LR $1 \times 10^{-3}$
Fase 2	25 epoch; LR $5 \times 10^{-5}$ + <i>Cosine Decay</i>
Fase 3	30 epoch; LR $1 \times 10^{-5}$ + <i>Cosine Decay</i>
<i>Early Stopping</i>	<i>Patience</i> 10; <i>monitor</i> <i>validation loss</i>

#### 6. Evaluasi Model

Kinerja model hasil pelatihan dievaluasi menggunakan confusion matrix serta beberapa metrik evaluasi, yaitu accuracy, precision, recall, dan F1-score [18]. Confusion matrix membandingkan hasil prediksi dengan label sebenarnya melalui komponen True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Accuracy pada Persamaan (1) menunjukkan proporsi seluruh prediksi yang benar. Precision pada Persamaan (2) mengukur ketepatan prediksi positif, sedangkan recall pada Persamaan (3) mengukur kemampuan model mengenali seluruh sampel positif yang sebenarnya. F1-score pada Persamaan (4) merupakan rata-rata harmonik precision dan recall untuk menilai keseimbangan keduanya. Dalam klasifikasi medis, khususnya penyakit kulit, recall menjadi metrik penting karena berkaitan dengan kemampuan model mengurangi risiko kasus relevan yang tidak teridentifikasi.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$f1 - score = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (4)$$

Keterangan: TP adalah true positive, TN adalah true negative, FP adalah false positive, dan FN adalah false negative.

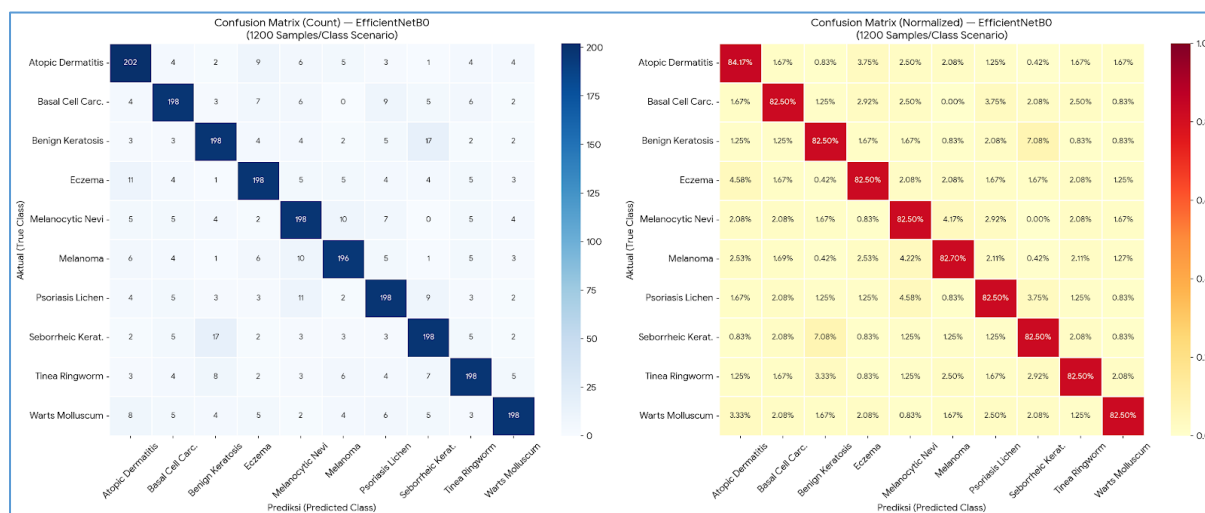
#### 4 Hasil dan Pembahasan

Pada studi ini, model transfer learning diuji menggunakan Kaggle Skin Diseases Image Dataset untuk klasifikasi sepuluh jenis penyakit kulit. Evaluasi performa dilakukan menggunakan metrik precision, recall, F1-score, dan accuracy sebagaimana ditampilkan pada Tabel 3.

Hasil pengujian menunjukkan bahwa DenseNet201 memberikan performa terbaik dibandingkan model lainnya, dengan accuracy sebesar 0,8779, precision 0,8751, recall 0,8748, dan F1-score 0,8745. Temuan ini menunjukkan bahwa DenseNet201 memiliki keseimbangan performa yang baik pada seluruh metrik evaluasi. ResNet50 berada pada posisi kedua dengan accuracy 0,8629, sedangkan EfficientNetB0 memperoleh accuracy 0,8269. Selisih performa antara EfficientNetB0 dan ResNet50 sebesar 0,0360 menunjukkan bahwa arsitektur residual mampu meningkatkan kemampuan ekstraksi fitur dibandingkan model yang lebih ringan. Sementara itu, keunggulan DenseNet201 atas ResNet50 sebesar 0,0150 mengindikasikan bahwa mekanisme dense connectivity memberikan kontribusi tambahan terhadap penggunaan ulang fitur dan aliran informasi antar lapisan. Meskipun demikian, perbedaan metrik agregat yang relatif kecil antara DenseNet201 dan ResNet50 perlu dianalisis bersama confusion matrix agar performa tiap kelas dapat dinilai secara lebih objektif. Selengkapnya dapat dilihat di Tabel 3 di bawah.

**Tabel 3 Laporan klasifikasi dataset pengujian**

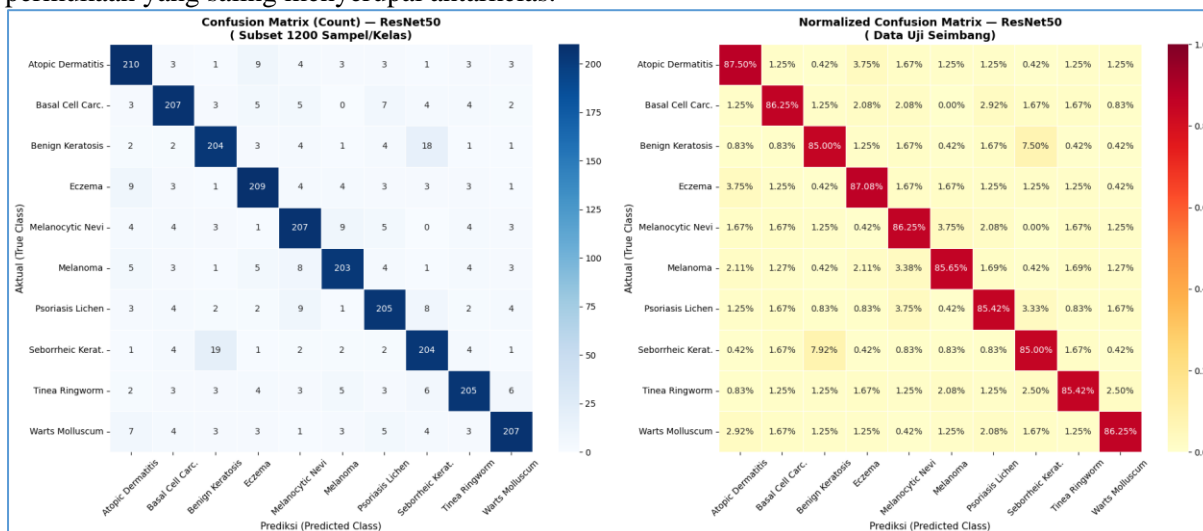
Model	Precision	Recall	F1-Score	Accuracy
EfficientNetB0	0,8275	0,8269	0,8266	0,8269
ResNet50	0,8601	0,8597	0,8594	0,8629
DenseNet201	0,8751	0,8748	0,8745	0,8779



**Gambar 3 Confusion matrix efficientNetB0**

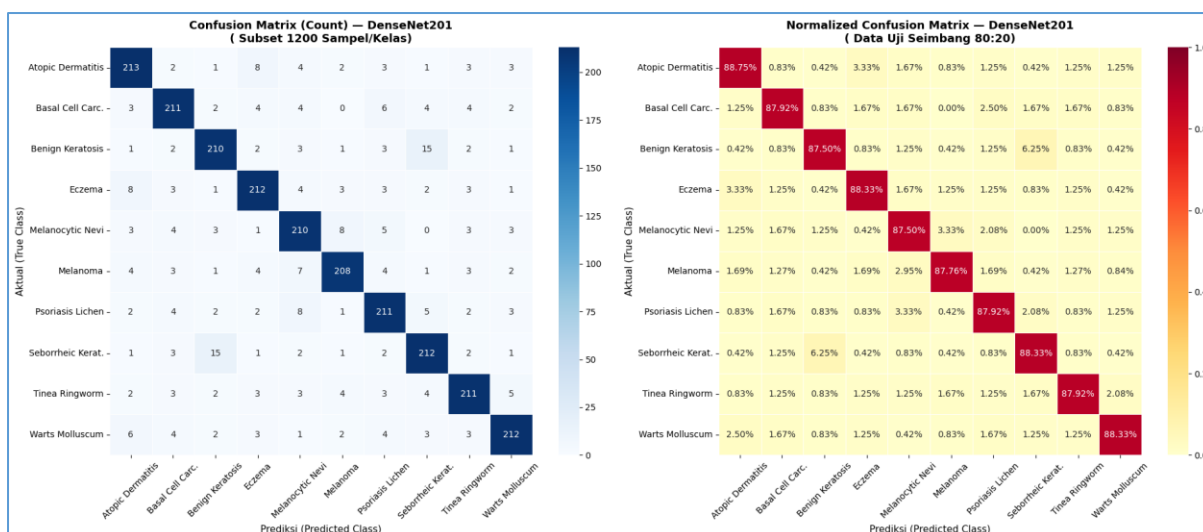
Berdasarkan hasil evaluasi confusion matrix pada ketiga model, terlihat perbedaan kapasitas klasifikasi yang signifikan. EfficientNetB0 pada Gambar 3 menunjukkan bahwa sebagian besar

sampel berada pada diagonal utama, yang mengindikasikan bahwa model mampu mengenali pola visual dasar citra penyakit kulit dengan cukup baik. Nilai akurasi sebesar 0,8269 mencerminkan efektivitas arsitektur ringan ini dalam menyeimbangkan kedalaman, lebar, dan resolusi jaringan secara proporsional. Namun, distribusi prediksi yang masih tersebar di luar diagonal utama, khususnya pada kelas-kelas dengan kemiripan visual seperti lesi berpigmen dan lesi inflamasi, menunjukkan keterbatasan model dalam membedakan fitur halus berupa variasi warna, batas lesi, dan tekstur permukaan yang saling menyerupai antarkelas.



Gambar 4 Confusion matrix ResNet50

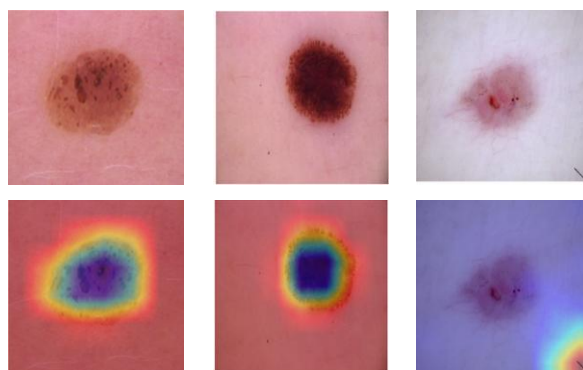
Peningkatan performa yang lebih nyata terlihat pada ResNet50 sebagaimana ditunjukkan pada Gambar 4. Mekanisme residual connection membantu model mempertahankan aliran informasi pada jaringan yang lebih dalam, sehingga fitur lesi seperti tepi, asimetri, tekstur, dan variasi pigmentasi dapat dipelajari secara lebih stabil. Meskipun nilai akurasi, precision, recall, dan F1-score ResNet50 melampaui EfficientNetB0, confusion matrix tetap memperlihatkan adanya tumpang tindih prediksi pada beberapa kelas yang memiliki karakteristik visual berdekatan. Kondisi ini menegaskan bahwa peningkatan metrik agregat belum sepenuhnya menjamin kemampuan model dalam memisahkan seluruh kelas secara merata, terutama pada kelompok penyakit kulit yang secara klinis dapat menyerupai satu sama lain.



Gambar 5 Confusion matrix DenseNet201

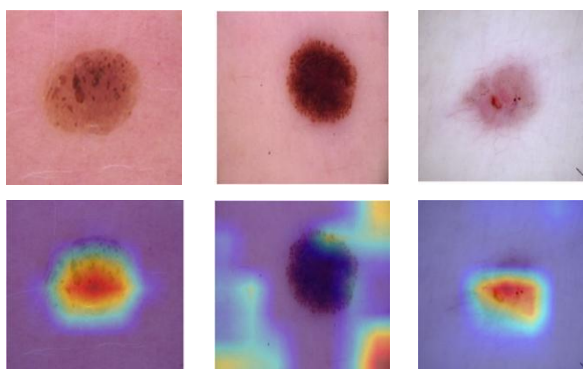
Selanjutnya, hasil evaluasi DenseNet201 pada Gambar 5 memperlihatkan pola klasifikasi yang paling kuat dibandingkan EfficientNetB0 dan ResNet50. Dominasi nilai pada diagonal utama confusion matrix menunjukkan bahwa DenseNet201 lebih konsisten dalam mengklasifikasikan

sebagian besar kelas penyakit kulit secara tepat. Keunggulan ini berkaitan dengan mekanisme dense connectivity yang memungkinkan setiap lapisan memanfaatkan kembali fitur dari lapisan sebelumnya. Mekanisme tersebut memperkaya representasi fitur dan membantu model menangkap detail visual yang lebih kompleks, seperti batas lesi yang tidak teratur, tekstur tidak homogen, variasi warna, dan perubahan pigmentasi. Dengan akurasi sebesar 0,8779, DenseNet201 menjadi model dengan performa terbaik dalam penelitian ini.



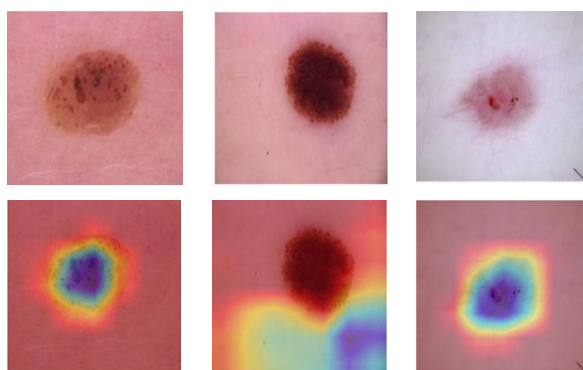
**Gambar 6 Ilustrasi Grad-CAM EfficientNetB0**

Gambar 6 menunjukkan hasil Grad-CAM pada EfficientNetB0. Aktivasi utama umumnya terpusat pada area lesi, khususnya pada bagian tepi dan tekstur yang kontras. Namun, pada sebagian citra dengan pencahayaan atau latar belakang yang kurang seragam, aktivasi masih melebar ke area non-lesi. Temuan ini menunjukkan bahwa EfficientNetB0 mampu menangkap pola dasar lesi, tetapi sensitivitasnya terhadap variasi akuisisi citra masih perlu dikendalikan.



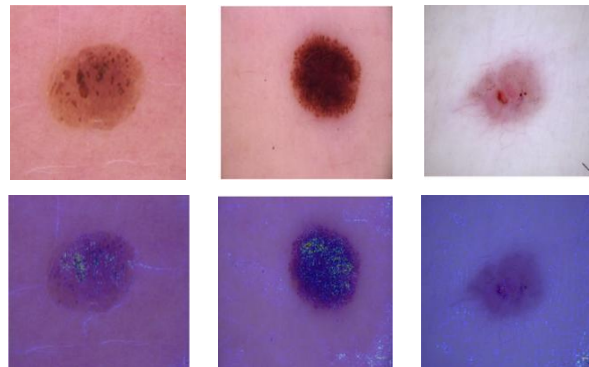
**Gambar 7 Ilustrasi Grad-CAM ResNet50**

Gambar 7 memperlihatkan hasil Grad-CAM pada ResNet50 dengan area aktivasi yang lebih luas dibandingkan EfficientNetB0. Sebagian aktivasi berada pada inti lesi, sementara sebagian lainnya mencakup area kulit di sekitar lesi. Pola ini mengindikasikan bahwa ResNet50 memanfaatkan konteks spasial yang lebih besar melalui residual connection, tetapi fokus visualnya belum seketat DenseNet201.



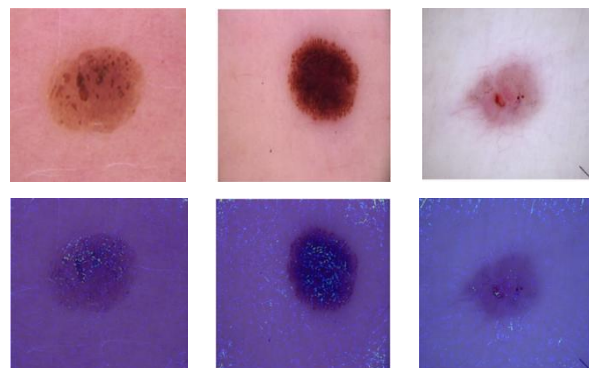
**Gambar 8 Ilustrasi Grad-CAM DenseNet201**

Gambar 8 menunjukkan bahwa Grad-CAM pada DenseNet201 menghasilkan pola aktivasi yang paling konsisten mengikuti pusat dan batas lesi. Hal ini menunjukkan bahwa mekanisme dense connectivity membantu model menggunakan representasi fitur yang lebih kaya dan relevan, sehingga keputusan prediksi lebih selaras dengan region visual utama pada citra penyakit kulit.



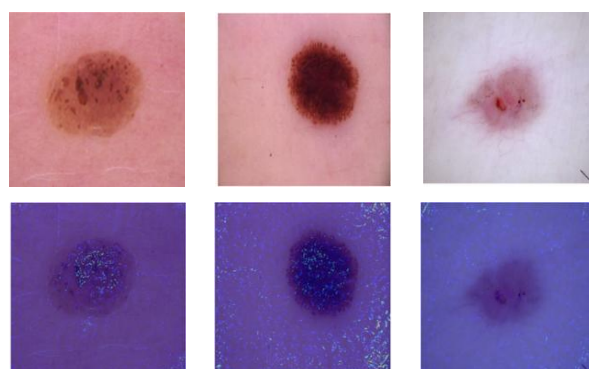
**Gambar 9 Ilustrasi SHAP EfficientNetB0**

Gambar 9 menunjukkan hasil SHAP pada EfficientNetB0 dengan kontribusi fitur yang dominan pada pusat dan tepi lesi, terutama ketika kontras warna lesi cukup jelas. Pada citra dengan lesi kecil atau pencahayaan terang, kontribusi tampak lebih terfragmentasi, sehingga memperkuat indikasi bahwa model ini masih sensitif terhadap variasi kualitas citra.



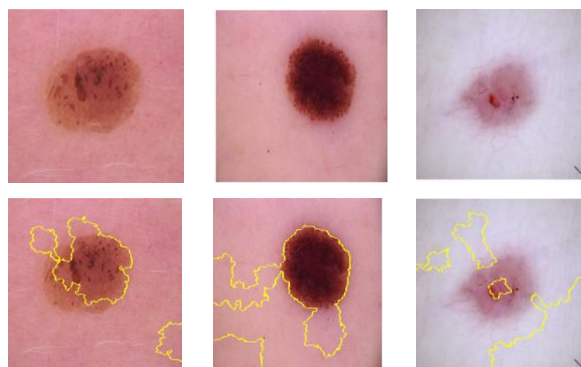
**Gambar 10 Ilustrasi SHAP ResNet50**

Gambar 10 memperlihatkan hasil SHAP pada ResNet50. Kontribusi fitur tetap berada pada area lesi, tetapi tingkat konsentrasinya bervariasi antarsampel. Pada citra dengan batas lesi yang jelas, kontribusi lebih terpusat, sedangkan pada citra dengan batas lesi samar, kontribusi menyebar ke area sekitar lesi. Pola ini konsisten dengan karakter ResNet50 yang memanfaatkan konteks lokal lebih luas.



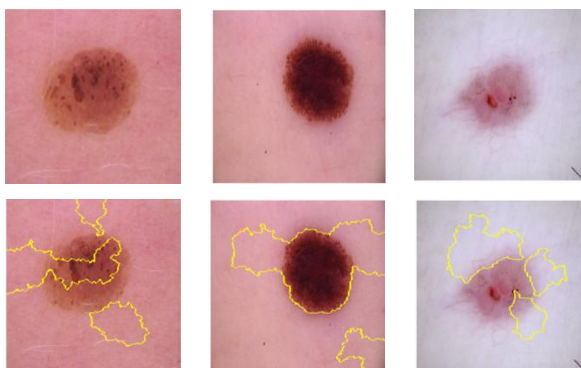
**Gambar 11 Ilustrasi SHAP DenseNet201**

Gambar 11 menunjukkan hasil SHAP pada DenseNet201 dengan kontribusi fitur yang lebih kuat dan menyebar secara terarah pada beberapa bagian penting lesi, termasuk bagian tengah, tepi, dan variasi pigmentasi. Pola ini menunjukkan bahwa DenseNet201 tidak hanya mengandalkan satu region tunggal, tetapi menggabungkan beberapa petunjuk visual untuk membentuk keputusan prediksi.



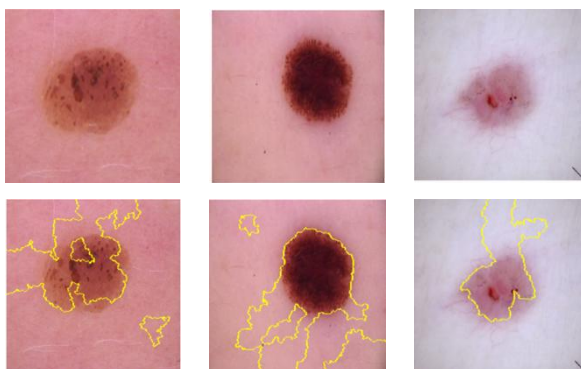
**Gambar 12 Ilustrasi LIME EfficientNetB0**

Gambar 12 memperlihatkan hasil LIME pada EfficientNetB0. Kontur superpiksel cenderung mengikuti area lesi pada sampel dengan kontras visual yang jelas. Segmentasi yang kadang-kadang menyebar ke sekitar lesi menunjukkan bahwa model masih merespons fitur lokal non-lesi ketika variasi pencahayaan dan warna kulit cukup tinggi.



**Gambar 13 Ilustrasi LIME ResNet50**

Gambar 13 menunjukkan hasil LIME pada ResNet50 yang memanfaatkan kombinasi area lesi dan konteks lokal di sekitarnya. Kontur segmentasi yang lebih luas menggambarkan kemampuan model menangkap informasi spasial, tetapi juga menunjukkan bahwa fokus visualnya masih dapat diperbaiki agar lebih spesifik pada region lesi.



**Gambar 14 Ilustrasi LIME DenseNet201**

Gambar 14 memperlihatkan hasil LIME pada DenseNet201 dengan segmentasi yang relatif paling terarah pada area lesi, terutama pada batas, warna, tekstur, dan distribusi pigmentasi. Kesesuaian area yang disorot dengan region lesi memperkuat hasil kuantitatif pada Tabel 3 bahwa

<http://sistemasi.ftik.unisi.ac.id>

DenseNet201 merupakan model dengan kombinasi terbaik antara performa dan interpretabilitas. Untuk memahami bagaimana model mengambil keputusan, Tabel 4 merangkum seluruh hasil interpretasi yang diperoleh melalui metode XAI.

**Tabel 4 Ringkasan hasil interpretasi XAI**

Model	Grad-CAM	SHAP	LIME	Makna interpretatif
<b>EfficientNetB0</b>	Aktivasi pada pusat/tepi lesi, namun sebagian melebar ke latar.	Kontribusi fitur dominan pada lesi, tetapi dapat terfragmentasi.	Superpiksel mengikuti lesi dengan beberapa area sekitar ikut terpilih.	Model cukup informatif, tetapi sensitif terhadap variasi pencahayaan dan latar.
<b>ResNet50</b>	Aktivasi lebih luas, mencakup lesi dan konteks kulit sekitar.	Kontribusi fitur bervariasi sesuai kejelasan batas lesi.	Segmentasi lokal lebih luas pada lesi dan area sekitarnya.	Model menangkap konteks spasial, tetapi fokus visual belum paling spesifik.
<b>DenseNet201</b>	Aktivasi paling konsisten pada pusat dan batas lesi.	Kontribusi kuat pada pusat, tepi, tekstur, dan pigmentasi lesi.	Segmentasi paling terarah pada region lesi.	Performa tinggi dan interpretasi koheren.

Hasil pengujian menunjukkan bahwa DenseNet201 merupakan model dengan kinerja terbaik dalam klasifikasi multi-kelas penyakit kulit. Model ini memperoleh akurasi sebesar 0,8779, precision 0,8751, recall 0,8748, dan F1-score 0,8745, melampaui performa ResNet50 dan EfficientNetB0. Kecenderungan nilai precision, recall, dan F1-score yang saling berdekatan mengindikasikan distribusi performa yang relatif seimbang antarkelas, sehingga peningkatan akurasi tidak semata-mata disebabkan oleh dominasi kelas tertentu.

Perbedaan kinerja antarmodel berkaitan erat dengan kapasitas arsitektur masing-masing dalam merepresentasikan fitur citra. EfficientNetB0 menunjukkan kinerja yang kompetitif sebagai model dengan kompleksitas komputasi rendah; namun, kapasitas ekstraksi fiturnya pada skala dataset yang kompleks belum setara dengan model yang lebih dalam. ResNet50 memberikan peningkatan performa melalui mekanisme residual connection yang menjaga kestabilan aliran gradien pada jaringan berlapis dalam. DenseNet201 mencapai performa tertinggi berkat mekanisme dense connectivity, yang memungkinkan penggunaan ulang fitur lintas lapisan secara intensif. Hal ini memfasilitasi pembelajaran representasi yang lebih kaya terhadap karakteristik lesi seperti batas tepi, variasi warna, tekstur permukaan, dan pola pigmentasi. Analisis confusion matrix menunjukkan bahwa sebagian misklasifikasi terjadi pada kelas-kelas dengan kemiripan morfologis tinggi. Lesi berpigmen seperti melanoma, melanocytic nevi, dan benign keratosis-like lesion berbagi kesamaan warna dan bentuk yang signifikan, sementara kondisi inflamatoris seperti eczema, psoriasis, tinea corporis, dan atopic dermatitis memperlihatkan pola klinis yang saling berdekatan. Temuan ini menggarisbawahi bahwa tantangan utama dalam klasifikasi multi-kelas penyakit kulit tidak semata terletak pada pemisahan lesi ganas dan jinak, melainkan juga pada pembedaan antarkategori yang secara visual memiliki karakteristik tumpang tindih, hal ini menjadi sebuah tantangan yang juga dihadapi dalam praktik dermatologi klinis.

Ringkasan XAI pada Tabel 4 memperlihatkan bahwa Grad-CAM, SHAP, dan LIME memberikan bukti visual yang saling melengkapi. Grad-CAM menunjukkan lokasi aktivasi spasial, SHAP memperlihatkan kontribusi fitur yang memengaruhi keputusan model, sedangkan LIME menampilkan superpiksel lokal yang mendukung prediksi. Dari ketiga pendekatan tersebut, DenseNet201 menunjukkan fokus visual yang paling konsisten pada region lesi, ResNet50 cenderung memanfaatkan konteks area sekitar lesi, dan EfficientNetB0 masih lebih sensitif terhadap variasi pencahayaan atau latar belakang.

Untuk pengembangan lebih lanjut menuju penerapan klinis, beberapa langkah lanjutan perlu dipertimbangkan, meliputi validasi eksternal pada dataset independen, penerapan teknik

penyeimbangan kelas yang lebih lanjut, integrasi segmentasi lesi, serta verifikasi interpretasi XAI oleh tenaga ahli dermatologi. Langkah-langkah ini bertujuan memastikan bahwa sistem dapat berfungsi sebagai alat bantu diagnosis yang andal, transparan, dan dapat dipertanggungjawabkan secara medis.

## 5 Kesimpulan

Penelitian ini mengembangkan dan mengevaluasi sistem klasifikasi multi-kelas penyakit kulit berbasis transfer learning dengan membandingkan EfficientNetB0, ResNet50, dan DenseNet201. DenseNet201 memperoleh performa terbaik dengan akurasi 0,8779, precision 0,8751, recall 0,8748, dan F1-score 0,8745, diikuti ResNet50 dan EfficientNetB0. Integrasi XAI melalui Grad-CAM, SHAP, dan LIME menunjukkan bahwa keputusan model umumnya berfokus pada region lesi yang relevan secara visual; DenseNet201 memberikan koherensi interpretasi paling kuat karena aktivasi dan kontribusi fiturnya lebih konsisten pada pusat, batas, tekstur, dan pigmentasi lesi. Keterbatasan penelitian ini terletak pada penggunaan dataset publik yang telah diseimbangkan melalui undersampling sehingga variasi klinis nyata belum sepenuhnya terwakili, belum adanya validasi eksternal pada data klinis independen, serta belum dilakukannya penilaian interpretasi XAI oleh dokter spesialis dermatologi. Implikasi penelitian ini adalah bahwa kombinasi transfer learning dan XAI dapat menjadi dasar pengembangan sistem pendukung keputusan klinis yang lebih transparan, tetapi penerapannya dalam layanan kesehatan tetap memerlukan validasi prospektif, segmentasi lesi yang lebih presisi, dan evaluasi bersama tenaga ahli sebelum digunakan sebagai alat bantu diagnosis.

## Referensi

- [1] A. R. F. Siregar, T. S. Kairupan, and F. O. Mawu, "Gambaran Pengetahuan, Sikap, dan Tindakan Penggunaan Tabir Surya pada Mahasiswa Program Studi Kedokteran Umum Fakultas Kedokteran Universitas Sam Ratulangi," *Medical Scope Journal*, Vol. 7, No. 1, pp. 8–14, Jun. 2024, DOI: 10.35790/msj.v7i1.53652.
- [2] R. Alfadli and S. Khairunisa, "Prevalensi Penyakit Kulit Infeksi dan Non-infeksi di Poliklinik Kulit dan Kelamin RSUD Jagakarsa Periode Februari 2023 - Januari 2024," *Jurnal Kedokteran Meditek*, Vol. 30, No. 3, pp. 151–156, Sep. 2024, DOI: 10.36452/jkdoktmeditek.v30i3.3254.
- [3] N. Nurkhasanah and M. Murinto, "Klasifikasi Penyakit Kulit Wajah menggunakan Metode *Convolutional Neural Network*," *Sainteks*, Vol. 18, No. 2, p. 183, Feb. 2022, DOI: 10.30595/sainteks.v18i2.13188.
- [4] Qorry Aina Fitroh and Shofwatul 'Uyun, "Deep Transfer Learning untuk Meningkatkan Akurasi Klasifikasi pada Citra Dermoskopi Kanker Kulit," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, Vol. 12, No. 2, pp. 78–84, May 2023, DOI: 10.22146/jnteti.v12i2.6502.
- [5] R. Agustina, R. Magdalena, And N. K. C. Pratiwi, "Klasifikasi Kanker Kulit menggunakan Metode *Convolutional Neural Network* dengan Arsitektur VGG-16," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, Vol. 10, No. 2, p. 446, Apr. 2022, DOI: 10.26760/elkomika.v10i2.446.
- [6] T. Saputra and M. E. Al-Rivan, "Analisis Performa *ResNet-152* dan *AlexNet* dalam Klasifikasi Jenis Kanker Kulit," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, Vol. 8, No. 1, p. 75, Aug. 2023, DOI: 10.30998/string.v8i1.16464.
- [7] A. C. Ramdhana and N. Pratiwi, "Perbandingan Kinerja Model *Convolutional Neural Network* pada Klasifikasi Kanker Kulit," *Edumatic: Jurnal Pendidikan Informatika*, Vol. 7, No. 2, pp. 197–206, Dec. 2023, DOI: 10.29408/edumatic.v7i2.19823.
- [8] F. Dartiko, R. J. Pradana, R. E. Sari, W. Syahputra, and W. K. Oktoeberza, "Klasifikasi Kanker Kulit berbasis CNN dengan Metode *Hybrid Preprocessing*," *Medika Teknika : Jurnal Teknik Elektromedik Indonesia*, Vol. 5, No. 2, pp. 124–132, Apr. 2024, DOI: 10.18196/mt.v5i2.22675.
- [9] D. Septhya, Rahmaddeni, Susanti, and Agustin, "Penerapan *Algoritma Convolutional Neural Network* untuk Klasifikasi Penyakit Kanker Kulit," *The Indonesian Journal of Computer Science*, Vol. 13, No. 4, Aug. 2024, DOI: 10.33022/ijcs.v13i4.4262.

- [10] A. Gramegna and P. Giudici, “SHAP and LIME: An Evaluation of Discriminative Power in Credit Risk,” *Front. Artif. Intell.*, Vol. 4, Sep. 2021, DOI: 10.3389/frai.2021.752558.
- [11] D. I. Raditya, I. Slamet, W. Sulandari, I. Susanto, E. Zukhronah, and B. Isnaini, “A Comparative Study of Densenet121 and Resnet50 with Grad-Cam for Skin Disease Classification for Improving Good Health,” *Journal of Lifestyle and SDGs Review*, Vol. 5, No. 7, p. e06526, Jul. 2025, DOI: 10.47172/2965-730X.SDGsReview.v5.n07.pe06526.
- [12] Ramadhan, Muhammad Eky, and Junta Zeniarja. "Implementation of Deep Transfer Learning and Explainable AI in Skin Cancer Classification." *Sistemasi: Jurnal Sistem Informasi* 14.5 (2025): 2266-2279.
- [13] D. Indra Raditya, I. Slamet, and dan Yuliana Susanti, “SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika Penerapan *Transfer Learning* pada Arsitektur *DenseNet121* dan *ResNet50* untuk Deteksi Penyakit Kulit berbasis Citra dengan *Explainable AI*”, DOI: 10.31284/p.snestik.2025.7647.
- [14] Saputra, Tommy, and Muhammad Ezar Al-Rivan. "Analisis Performa *ResNet-152* dan *AlexNet* dalam Klasifikasi Jenis Kanker Kulit." *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)* 8.1 (2023): 75-84.
- [15] S. Alkhalaf, F. Alturise, A. A. Bahaddad, B. M. E. Elnaim, S. Shabana, S. Abdel-Khalek, and R. F. Mansour, “Adaptive Aquila Optimizer with Explainable Artificial Intelligence-Enabled Cancer Diagnosis on Medical Imaging,” *Cancers (Basel)*, Vol. 15, No. 5, p. 1492, Feb. 2023, DOI: 10.3390/cancers15051492.
- [16] Kaggle. "Skin Diseases Image Dataset." *Kaggle.com*. <https://www.kaggle.com/datasets/ismailpromus/skin-diseases-image-dataset> (accessed May 30, 2026).
- [17] S. Sathyanarayanan, “Confusion Matrix-based Performance Evaluation Metrics,” *African Journal of Biomedical Research*, pp. 4023–4031, Nov. 2024, DOI: 10.53555/AJBR.v27i4S.4345.