



Implementasi Arsitektur *Deep Convolutional Neural Network* (CNN) dengan *Transfer Learning* untuk Klasifikasi Penyakit Kulit

I Putu Agus¹, Khasnur Hidjah¹, Neny Sulistianingsih², Galih Hendro², Syahrir³

¹ Program Studi Ilmu Komputer, Universitas Bumigora, Indonesia

² Program Studi Ilmu Komputer, Program Pascasarjana, Universitas Bumigora, Indonesia

³ Program Studi Rekayasa Perangkat Lunak, Universitas Bumigora, Indonesia

* Korespondensi: khasnur72.h@universitasbumigora.ac.id

Situs: Agus, I. P.; Hidjah, K.; Sulistianingsih, N.; Hendro, G.; Syahrir, S. (2025). Implementasi Arsitektur Deep Convolutional Neural Network (CNN) dengan Transfer Learning untuk Klasifikasi Penyakit Kulit. *JTIM: Jurnal Teknologi Informatika Dan Multimedia*, 7(3), 461-477. <https://doi.org/10.35746/jtim.v7i3.734>

Diterima: 03-05-2025

Direvisi: 11-06-2025

Disetujui: 17-06-2025

Abstract: Skin diseases are common health problems that require early diagnosis to prevent serious complications. This study aims to develop an automatic skin disease image classification system using a *transfer learning* approach based on Convolutional Neural Networks (CNN). Image datasets were obtained from Kaggle and underwent preprocessing stages including resizing, normalization, and augmentation. Four CNN architectures were evaluated: VGG16, ResNet50, MobileNetV2, and InceptionV3, implemented using Python and the Keras library on the Google Colab platform. The dataset was split into three training and testing ratios (90:10, 80:20, and 70:30) to assess the impact of data proportion on model performance. Models were trained by modifying the output layer to match the number of classes, and evaluated using accuracy, precision, recall, F1-score, confusion matrix, and ROC curve metrics. The results show that a 70:30 ratio yielded the most optimal training performance. InceptionV3 achieved the highest validation accuracy at 80.04%, but experienced overfitting, while VGG16 demonstrated better generalization to test data. This study proves that transfer learning with CNN is effective in improving the accuracy of automatic skin disease diagnosis and has the potential to become an efficient diagnostic solution, especially in areas with limited medical infrastructure.

Keywords: classification, image, skin, transfer, learning, convolution, evaluation, overfitting



Copyright: © 2025 oleh para penulis.
Karya ini dилісанкі di bawah
Creative Commons
Attribution-ShareAlike
International License.
(<https://creativecommons.org/licenses/by-sa/4.0/>).

Abstrak: Penyakit kulit merupakan masalah kesehatan umum yang memerlukan diagnosis dini untuk mencegah komplikasi serius. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi otomatis citra penyakit kulit menggunakan metode *transfer learning* berbasis Convolutional Neural Network (CNN). Dataset citra diperoleh dari Kaggle dan melalui tahap preprocessing berupa resizing normalisasi, dan augmentasi. Empat arsitektur CNN yang diuji adalah VGG16, ResNet50, MobileNetV2, dan InceptionV3, yang diimplementasikan menggunakan Python dan pustaka Keras di Google Colab. Dataset dibagi dalam tiga rasio pelatihan dan pengujian (90:10, 80:20, dan 70:30) untuk mengevaluasi pengaruh proporsi data terhadap performa model. Model dilatih dengan mengganti layer output agar sesuai dengan jumlah kelas, dan dievaluasi menggunakan metrik akurasi, presisi, recall, F1-score, confusion matrix, dan kurva ROC. Hasil menunjukkan bahwa rasio 70:30 menghasilkan pelatihan paling optimal. InceptionV3 mencatat akurasi validasi tertinggi sebesar 80,04%, namun mengalami overfitting, sementara VGG16 menunjukkan generalisasi lebih baik terhadap data uji. Penelitian ini membuktikan bahwa transfer learning dengan CNN efektif dalam meningkatkan akurasi diagnosis otomatis penyakit kulit, khususnya di wilayah dengan keterbatasan fasilitas medis.

Kata kunci: klasifikasi, citra, kulit, transfer, pembelajaran, konvolusi, evaluasi, overfitting

1. Pendahuluan

Penyakit kulit merupakan salah satu kondisi kesehatan yang umum dan dapat menyerang individu dari berbagai usia, serta kerap bersifat kronis [1]. Penyebabnya sangat beragam, mulai dari faktor lingkungan yang tidak sehat, kebiasaan hidup yang buruk, perubahan iklim, infeksi virus atau bakteri, alergi, hingga penurunan daya tahan tubuh [2]. Gejala yang ditimbulkan pun bervariasi, seperti rasa gatal, *eritema* (kemerahan), *edema* (pembengkakan), *papul* (tonjolan padat < 5 mm), dan *vesikel* (lepuhan berisi cairan) [3]. Penyakit kulit bisa menyerang seluruh tubuh atau terbatas pada bagian tertentu, dan bila tidak ditangani dengan tepat, dapat memperburuk kondisi kesehatan penderitanya.

Di Indonesia, prevalensi penyakit kulit berkisar antara 4,60% hingga 12,95%, menjadikannya salah satu penyakit yang paling umum diderita masyarakat dan menempati urutan ketiga dalam daftar sepuluh besar penyakit terbanyak [3]. Tingginya angka kejadian ini menunjukkan bahwa penyakit kulit merupakan masalah kesehatan masyarakat yang membutuhkan penanganan serius [4]. Saat ini, diagnosis penyakit kulit umumnya dilakukan melalui pemeriksaan klinis dan prosedur *biopsi*, yaitu pengambilan sampel jaringan kulit untuk dianalisis di laboratorium menggunakan mikroskop, guna memastikan diagnosis atau mendeteksi infeksi tertentu [5]. Meskipun prosedur ini memiliki tingkat akurasi yang tinggi, terdapat sejumlah kelemahan, seperti waktu persiapan yang lama, biaya yang relatif mahal, serta ketergantungan pada fasilitas medis yang memadai. Hal ini menjadi tantangan tersendiri, terutama di daerah yang memiliki keterbatasan infrastruktur layanan kesehatan [6].

Motivasi lain dari penelitian ini muncul dari tantangan dalam proses diagnosis penyakit kulit yang sering kali dipengaruhi oleh kemiripan visual antar jenis penyakit, keterbatasan waktu konsultasi, serta minimnya alat bantu diagnosis berbasis teknologi. Beberapa studi menunjukkan bahwa kesalahan diagnosis dapat terjadi akibat keterbatasan sumber daya medis, terutama di wilayah dengan akses layanan kesehatan yang rendah. Kondisi ini mendorong perlunya solusi yang lebih akurat dan efisien melalui pemanfaatan teknologi kecerdasan buatan (*Artificial Intelligence/AI*), khususnya dalam pengolahan citra medis dan sistem klasifikasi otomatis.

Perkembangan teknologi digital, khususnya *AI*, telah membuka peluang besar dalam bidang kesehatan, terutama dalam diagnosis berbasis citra. Salah satu pendekatan *AI* yang paling menjanjikan adalah *deep learning*, khususnya arsitektur *Convolutional Neural Network (CNN)*, yang unggul dalam mengenali pola visual kompleks pada citra digital [4]. CNN mampu mengklasifikasikan objek dalam citra dengan akurasi tinggi, dan terus berkembang untuk meningkatkan performa dan efisiensinya [7]. Selain itu, pendekatan *transfer learning* juga semakin banyak digunakan, yakni dengan memanfaatkan model CNN yang telah dilatih sebelumnya pada dataset besar seperti *ImageNet* [6]. Teknik ini memungkinkan proses pelatihan menjadi lebih cepat dan efisien, serta tetap menghasilkan akurasi tinggi meskipun menggunakan *dataset* yang relatif terbatas.

Dalam konteks klasifikasi penyakit kulit, berbagai penelitian telah menunjukkan efektivitas CNN. Misalnya, penelitian yang dilakukan oleh Supirman dkk. menggunakan *VGG16* dan *MobileNet* untuk mengklasifikasikan tujuh jenis penyakit kulit dari 1.410 citra, dengan akurasi masing-masing 82% dan 80% [8]. Selanjutnya, penelitian yang dilakukan oleh Dhuhita dkk. menggunakan *MobileNetV2* pada *dataset HAM10000* dengan 10.015 gambar dan mencapai akurasi hingga 79% untuk klasifikasi kanker kulit [9]. Sementara itu, penelitian yang dilakukan oleh Oktavianus dkk. menggunakan *ResNet50* untuk klasifikasi citra cacar monyet dan memperoleh akurasi validasi 94,82% serta akurasi pengujian 76,10% dari total 3.962 gambar [10]. Penelitian lain yang dilakukan oleh Amania dkk., menerapkan *InceptionV3* untuk klasifikasi jerawat dengan akurasi 83,33%, *precision* 85,56%, *recall* 83,33%, dan *F1-score* 83,86% [11].

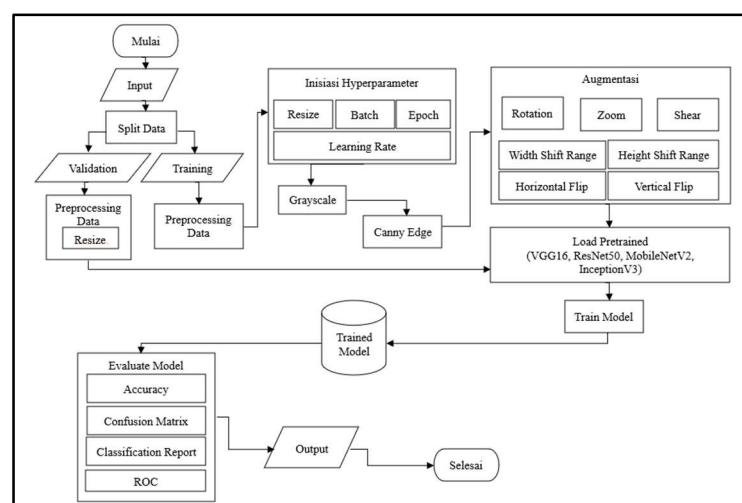
Dari studi-studi tersebut, terlihat bahwa arsitektur CNN seperti *VGG16*, *MobileNetV2*, *ResNet50*, dan *InceptionV3* memiliki potensi besar dalam tugas klasifikasi citra penyakit kulit. Oleh karena itu, penelitian ini bertujuan untuk mengimplementasikan metode *transfer learning* dengan keempat arsitektur CNN tersebut. Metode *transfer learning* dipilih karena memungkinkan pemanfaatan bobot dari model yang telah dilatih sebelumnya pada dataset berskala besar, sehingga dapat mempercepat proses pelatihan, mengurangi kebutuhan data, serta meningkatkan akurasi model. Tujuan akhir dari penelitian ini adalah membangun sistem klasifikasi penyakit kulit yang otomatis, yaitu mampu melakukan prediksi secara mandiri tanpa intervensi manual dari tenaga medis, dan efisien, yaitu meminimalkan waktu pelatihan, sumber daya komputasi, serta mempercepat proses diagnosis. Selain itu, dilakukan pula perbandingan performa antar model guna mengidentifikasi arsitektur yang paling optimal. Diharapkan, hasil penelitian ini dapat memberikan kontribusi signifikan dalam pengembangan sistem diagnosis otomatis berbasis AI di bidang dermatologi, khususnya dengan membandingkan performa empat arsitektur CNN melalui pendekatan *transfer learning* untuk klasifikasi citra penyakit kulit. Penelitian ini juga memberikan gambaran tentang rasio pembagian data yang paling optimal serta model mana yang memiliki generalisasi terbaik, sehingga dapat menjadi acuan dalam pengembangan sistem serupa di masa depan, maupun sebagai dasar penelitian lanjutan untuk meningkatkan efisiensi dan akurasi diagnosis penyakit kulit.

2. Bahan dan Metode

Bagian ini menjelaskan secara sistematis tahapan penelitian yang dilakukan, dimulai dari pemilihan dan pengumpulan dataset, proses praproses data, pembangunan model menggunakan arsitektur Convolutional Neural Network (CNN), penerapan metode *transfer learning*, pelatihan model, hingga evaluasi performa klasifikasi. Penelitian ini menggunakan empat arsitektur CNN populer, yaitu *VGG16*, *ResNet50*, *MobileNetV2*, dan *InceptionV3*. Proses pelatihan dilakukan dengan variasi pembagian data untuk mengamati pengaruhnya terhadap kinerja model. Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, F1-score, confusion matrix, dan ROC curve guna menilai efektivitas masing-masing arsitektur.

2.1. Tahapan Penelitian

Jenis penelitian yang akan dilakukan adalah untuk melihat kemampuan klasifikasi model mana yang lebih baik dalam mendekripsi penyakit kulit. Penelitian ini menggunakan metode kuantitatif tipe eksperimental dengan memanfaatkan dataset sekunder. Tahapan-tahapan penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian.

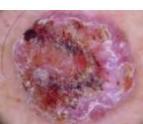
Penelitian dimulai dengan pengumpulan dataset citra penyakit kulit yang diperoleh dari situs *kaggle.com*. Dataset kemudian dibagi melalui proses splitting data menjadi tiga bagian, yaitu data pelatihan, validasi, dan pengujian. Tiga proporsi pembagian data digunakan yaitu 90:10, 80:20, dan 70:30. Proporsi pertama menunjukkan persentase data pelatihan dan validasi, sementara proporsi kedua dilakukan untuk pengujian. Tujuan variasi ini adalah untuk mengevaluasi pengaruh jumlah data pelatihan terhadap performa klasifikasi model. Data kemudian melalui tahap *preprocessing*, yang mencakup penentuan hyperparameter seperti *batch size*, jumlah *epoch*, dan *learning rate* untuk menoptimalkan proses pelatihan. *Batch size* sebesar 64 digunakan karena seimbang antara stabilitas pelatihan dan efisiensi komputasi. *Learning rate* ditetapkan sebesar 0.0001 untuk menjaga kestabilan pembaruan bobot selama pelatihan. Jumlah *epoch* sebanyak 50 dipilih untuk memberi waktu pelatihan yang cukup tanpa memicu *overfitting*, dengan dukungan *early stopping* sebagai mekanisme penghentian otomatis bila performa validasi tidak meningkat. Proses *preprocessing* gambar meliputi penyesuaian ukuran, normalisasi piksel, dan konversi format sesuai dengan input model. Untuk meningkatkan variasi dataset dan ketahanan model, *ImageDataGenerator* digunakan dengan teknik augmentasi seperti rotasi, *zooming*, dan *flipping*. Pada tahap implementasi, model CNN diterapkan dengan empat arsitektur VGG16, ResNet50, MobileNetV2, dan InceptionV3 yang dipilih karena keunggulannya dalam pengenalan citra, khususnya untuk klasifikasi berbagai jenis penyakit kulit. Jenis penyakit yang digunakan dalam penelitian ini mencakup eczema, melanoma, BCC, dan infeksi jamur, serupa dengan jenis yang digunakan dalam studi sebelumnya, seperti HAM10000 untuk kanker kulit atau studi Supirman dkk. dengan tujuh kategori penyakit. Dataset yang digunakan memiliki cakupan lebih luas dengan sembilan jenis penyakit yang telah terverifikasi secara medis. Tahap terakhir adalah evaluasi hasil, di mana model yang telah dilatih dievaluasi menggunakan *confusion matrix* dan sejumlah metrik turunan seperti *recall*, *precision*, *f1-score*, dan *accuracy*. *Confusion matrix* memberikan informasi dasar mengenai jumlah prediksi benar dan salah pada setiap kelas, yang menjadi dasar perhitungan metrik lainnya. *Recall* mengukur sensitivitas model dalam mengenali kelas positif, *precision* mengukur ketepatan prediksi, *f1-score* memberikan keseimbangan antara keduanya, dan *accuracy* menunjukkan tingkat prediksi benar secara keseluruhan. Evaluasi ini diperkuat dengan kurva ROC yang menggambarkan kemampuan model dalam membedakan antar kelas. Kombinasi metrik ini dianggap cukup representatif dan dapat memberikan gambaran yang meyakinkan terhadap performa model dalam tugas klasifikasi multikelas.

2.2. Dataset

Penelitian ini menggunakan *dataset* citra penyakit kulit yang diperoleh dari platform *Kaggle*, yang terdiri dari 9.000 gambar berwarna dengan resolusi bervariasi dan format *Joint Photographic Experts Group* (.jpg). Setiap citra telah dilabeli sesuai dengan jenis penyakit kulit berdasarkan hasil diagnosis oleh tenaga medis profesional yang terlibat dalam proses anotasi data. Karena bersumber dari repositori daring yang telah tersedia sebelumnya, dataset ini termasuk data sekunder. *Dataset* ini mencakup sembilan jenis penyakit, yaitu *eczema*, *melanoma*, *basal cell carcinoma (BCC)*, *melanocytic nevi (NV)*, *benign keratoses lesions (BKL)*, *lichen planus*, *seborrheic keratoses*, *tinea ringworm*, *candidiasis*, dan *warts molluscum*, dengan masing-masing jenis penyakit memiliki 1.000 citra.

Tabel 1. Dataset Penyakit Kulit.

No.	Gambar	Jenis Penyakit Kulit	Jumlah Citra
1		Eczema	1.000

No.	Gambar	Jenis Penyakit Kulit	Jumlah Citra
2		Melanoma	1.000
3		Basal Cell Carcinoma (BCC)	1.000
4		Melanocytic Nevi (NV)	1.000
5		Benign Keratotic Lesions (BKL)	1.000
6		Lichen Planus	1.000
7		Seborrheic Keratoses	1.000
8		Tinea Ringworm	1.000
9		Candidiasis dan Warts Molluscum	1.000
Total			9.000

Untuk keperluan pelatihan dan pengujian model, *dataset* dibagi menjadi dua kategori, yaitu data pelatihan (*training*) dan data pengujian (*testing*), dengan tiga rasio pembagian yang berbeda, yaitu 90:10, 80:20, dan 70:30. Tiga skenario ini digunakan untuk membandingkan pengaruh proporsi data terhadap performa model dan menemukan konfigurasi terbaik. Dataset terdiri dari citra berwarna berformat .jpg dengan resolusi bervariasi yang telah disesuaikan dengan kebutuhan masing-masing arsitektur CNN.

2.3. Preprocessing Data

Untuk memastikan konsistensi dan performa model, dilakukan beberapa tahap *preprocessing*, antara lain:

- 1) Inisialisasi *Hyperparameter*: Penetapan *hyperparameter* yang tepat sangat penting untuk mencapai kinerja optimal model. *Hyperparameter* yang ditentukan meliputi ukuran citra, *batch size*, jumlah *epoch*, dan *learning rate*. Pada penelitian ini, seluruh citra diubah ukurannya menjadi 224×224 piksel untuk memastikan keseragaman input dan efisiensi proses pelatihan. Meskipun InceptionV3 secara default menggunakan ukuran 299×299 piksel, penyesuaian ke 224×224 tetap memungkinkan karena arsitektur ini mendukung input fleksibel dengan sedikit penyesuaian, sehingga seluruh model dapat dilatih dengan ukuran citra yang seragam. *Batch size* ditetapkan sebesar 64, jumlah *epoch* sebanyak 50, dan *learning rate* 0.0001.

- 2) Resize: Seluruh citra yang memiliki ukuran bervariasi diubah ukurannya menjadi 224×224 piksel untuk memastikan keseragaman dimensi input pada semua arsitektur model CNN yang digunakan. Meskipun InceptionV3 secara default menggunakan ukuran 299×299 piksel, dalam penelitian ini dilakukan penyesuaian ke 224×224 piksel agar seluruh model dapat dilatih secara konsisten dan efisien tanpa mengorbankan performa secara signifikan.
- 3) Konversi ke *Grayscale*: Setelah citra diubah ukurannya, citra dikonversi dari format RGB ke *grayscale*. Konversi ini bertujuan untuk mengurangi kompleksitas data dan memudahkan penerapan teknik deteksi tepi dalam identifikasi batas atau kontur objek, yang bermanfaat dalam ekstraksi fitur dan segmentasi. Proses ini dilakukan dengan menghitung rata-rata bobot dari tiga kanal warna (R, G, B) berdasarkan kontribusi intensitas cahaya. Berikut pseudocode sederhana dari proses konversi:

For each pixel in image:

R, G, B = pixel values

$$\text{Gray} = 0.2989 * R + 0.5870 * G + 0.1140 * B$$

Replace pixel with Gray

- 4) Augmentasi Data: Untuk memperkaya variasi data pelatihan dan mengatasi ketidakseimbangan kelas, *augmentasi data* diterapkan. Teknik-teknik *augmentasi* yang digunakan antara lain *rotasi*, *flipping* horizontal dan vertikal, *pergeseran (shifting)*, serta *zooming*. *Augmentasi* ini bertujuan untuk meningkatkan kemampuan generalisasi model dengan memberikan lebih banyak variasi citra yang dapat dipelajari oleh model.

2.4. Arsitektur Model CNN

Convolutional Neural Network (CNN) adalah salah satu jenis jaringan syaraf tiruan yang banyak digunakan dalam analisis data gambar, khususnya untuk mendeteksi dan mengenali objek pada citra. CNN terinspirasi oleh cara kerja sistem saraf biologis yang terdiri dari tiga lapisan utama, yaitu *convolutional layers*, *pooling layers*, dan *fully-connected layers* [12]

Pada *convolutional layers*, *filter* digunakan untuk mendeteksi fitur visual seperti tepi, pola, dan tekstur dari gambar. *Activation function* seperti *ReLU* diaplikasikan untuk meningkatkan non-linearitas, sementara *pooling layers* berfungsi untuk mengurangi dimensi fitur guna mempercepat komputasi dan mencegah *overfitting*. Terakhir, *fully-connected layers* bertugas untuk melakukan klasifikasi berdasarkan fitur yang telah diekstraksi dari lapisan sebelumnya [13].

Saat ini, CNN dianggap sebagai model yang sangat efektif dalam berbagai tugas berbasis citra seperti *klasifikasi*, *segmentasi*, *deteksi*, dan *pengenalan objek*, serta telah diterapkan secara luas dalam bidang medis, pengenalan wajah, dan analisis video [14]. Dalam penelitian ini, CNN digunakan untuk mengklasifikasikan citra penyakit kulit berdasarkan pola visual yang terekam dalam gambar.

Dalam penelitian ini, digunakan empat arsitektur CNN populer yang telah terlatih pada dataset ImageNet sebagai bagian dari pendekatan transfer learning, yaitu: VGG16, ResNet50, MobileNetV2, dan InceptionV3. Tabel berikut merangkum karakteristik utama dari masing-masing arsitektur:

No.	Arsitektur	Tahun	Ukuran Input	Jumlah Layer	Ciri Khas
1	VGG16	2014	$224 \times 224 \times 3$	16	Desain sederhana dengan susunan layer konvolusi berurutan
2	ResNet50	2015	$224 \times 224 \times 3$	50	Memiliki residual

No.	Arsitektur	Tahun	Ukuran Input	Jumlah Layer	Ciri Khas
3	MobileNetV2	2018	224×224×3	53 (approx)	connection untuk mengatasi degradasi akurasi Ringan dan efisien, cocok untuk perangkat dengan sumber daya terbatas Menggunakan Inception modules untuk ekstraksi fitur multi-skala
4	InceptionV3	2015	224×224×3*	> 40	

*Catatan: InceptionV3 umumnya menggunakan ukuran input 299×299, namun dalam penelitian ini disesuaikan menjadi 224×224 untuk konsistensi antar model.

Setiap arsitektur dipilih karena keunggulan masing-masing dalam klasifikasi citra dan efektivitasnya dalam konteks transfer learning, terutama pada dataset berukuran sedang seperti yang digunakan dalam penelitian ini.

2.4.1. VGG16

VGG16 merupakan arsitektur CNN dari Visual Geometry Group yang terdiri dari 16 lapisan berisi bobot. Arsitektur ini dipilih karena desainnya yang sederhana namun mendalam, serta terbukti efektif dalam klasifikasi citra melalui penggunaan filter 3×3 secara berurutan. Keunggulan VGG16 terletak pada kemampuannya mengekstraksi fitur visual secara bertahap dan stabil, menjadikannya sangat cocok untuk transfer learning. Dalam penelitian ini, VGG16 diimplementasikan dengan ukuran input citra 224×224 piksel. Parameter pelatihan yang digunakan meliputi optimizer Adam, learning rate 0.0001, batch size 64, dan 50 epoch. Lapisan akhir (fully-connected) dimodifikasi agar sesuai dengan jumlah kelas dalam dataset penyakit kulit[15].

2.4.2. ResNet50

ResNet50 adalah model CNN dengan 50 lapisan yang telah dilatih pada lebih dari satu juta gambar dari database *ImageNet*. Model ini dikenal sangat andal dalam mengenali berbagai objek dan kategori secara akurat[16]. Keunggulan utama ResNet50 terletak pada penggunaan *skip connections* atau *residual connections*, yang menjaga kestabilan *gradien* selama pelatihan dan memungkinkan model mempelajari fitur kompleks secara lebih efektif [17].

Dalam penelitian ini, ResNet50 digunakan dengan pendekatan *transfer learning* dan citra berukuran 224×224 piksel. Parameter pelatihan yang digunakan meliputi optimizer Adam, learning rate 0.0001, batch size 64, dan 50 epoch. Lapisan klasifikasi akhir dimodifikasi untuk menyesuaikan jumlah kelas penyakit kulit pada dataset.

2.4.3. MobileNetV2

MobileNetV2 adalah arsitektur *neural network* yang dikembangkan oleh Google dan dirancang khusus untuk perangkat seluler atau sistem dengan keterbatasan sumber daya komputasi. Arsitektur ini merupakan penyempurnaan dari MobileNet sebelumnya, dengan memanfaatkan *depthwise separable convolutions*, *linear bottlenecks*, dan *shortcut connections* untuk mengurangi parameter dan mempercepat pelatihan tanpa mengorbankan akurasi [18]; [19].

Dalam penelitian ini, MobileNetV2 diimplementasikan dengan *input* citra 224×224 piksel dan pendekatan *transfer learning*. Parameter pelatihan meliputi optimizer Adam,

learning rate 0.0001, batch size 64, dan 50 epoch, dengan lapisan klasifikasi akhir yang disesuaikan dengan jumlah kelas penyakit kulit.

2.4.4. InceptionV3

InceptionV3 dipilih karena merupakan arsitektur CNN yang efisien secara komputasi namun tetap mampu menghasilkan akurasi tinggi. Arsitektur ini menggabungkan beberapa ukuran filter konvolusi secara paralel melalui *Inception module*, serta menggunakan teknik *factorized convolutions* untuk mengurangi jumlah parameter tanpa kehilangan informasi penting. Dibandingkan dengan arsitektur lain seperti VGG16, InceptionV3 memiliki kompleksitas yang lebih rendah namun tetap efektif dalam mengekstraksi fitur visual dari gambar secara mendalam [20].

Dalam implementasinya, parameter utama yang digunakan meliputi ukuran input gambar 299×299 piksel, optimizer Adam, learning rate sebesar 0.001, batch size 32, jumlah epoch 50 hingga 100, dan jumlah kelas sesuai dengan kebutuhan klasifikasi. Secara umum, arsitektur InceptionV3 terdiri dari beberapa bagian utama, yaitu blok awal (*stem*), *Inception modules* (A, B, C) untuk ekstraksi fitur, *Reduction modules* untuk menurunkan dimensi spasial, serta lapisan *fully connected* yang diakhiri dengan fungsi aktivasi *softmax* untuk klasifikasi akhir [21]. Keunggulan tersebut menjadikan InceptionV3 sangat cocok digunakan dalam tugas klasifikasi citra, termasuk identifikasi penyakit kulit.

2.5. Transfer Learning

Transfer learning adalah teknik dalam *machine learning* yang memanfaatkan model yang telah dilatih pada dataset besar seperti ImageNet untuk digunakan kembali pada tugas lain yang serupa, seperti klasifikasi penyakit kulit [12]. Transfer learning dilakukan pada tahap awal sebelum pelatihan penuh model, yaitu dengan memuat bobot dari model pralatih dan kemudian menyesuaikannya (fine-tuning) terhadap dataset baru yang lebih kecil. Teknik ini sangat berguna ketika jumlah data terbatas, karena dapat mengurangi risiko overfitting dan mempercepat proses pelatihan [22].

Dalam implementasinya, parameter penting yang dibutuhkan meliputi pemilihan model pralatih (misalnya InceptionV3), penentuan layer mana yang akan dibekukan (*freeze*), learning rate (biasanya lebih kecil, seperti 0.0001 untuk fine-tuning), optimizer (seperti Adam), serta jumlah epoch dan batch size yang disesuaikan dengan dataset. Arsitektur yang digunakan tetap mengikuti struktur asli dari model pralatih (seperti InceptionV3), namun bagian akhir (*fully connected layer*) diganti atau ditambahkan sesuai jumlah kelas pada dataset target, kemudian dilatih ulang secara terbatas untuk mengadaptasi pada tugas klasifikasi baru [23]. Teknik ini sangat berguna dalam tugas klasifikasi citra medis, terutama ketika data yang tersedia terbatas, dengan menggunakan model yang dilatih pada dataset besar seperti ImageNet untuk meningkatkan akurasi [1].

2.6. Proses Pelatihan

Pelatihan model dilakukan dengan menetapkan sejumlah parameter penting untuk memperoleh hasil klasifikasi yang optimal. Pemilihan nilai-nilai parameter ini didasarkan pada praktik umum dalam transfer learning serta referensi dari penelitian sebelumnya. Berikut adalah rincian proses pelatihan yang dilakukan:

- *Optimizer*: Adam dipilih karena kemampuannya dalam menyesuaikan learning rate secara adaptif, mempercepat proses konvergensi, dan bekerja baik pada data dengan noise serta parameter yang jarang diperbarui.
- *Loss Function*: *Categorical Crossentropy* digunakan karena klasifikasi bersifat multikelas, dan fungsi ini efektif dalam mengukur selisih antara distribusi prediksi dan label target.
- *Learning Rate*: Nilai sebesar 0.0001 dipilih untuk menjaga kestabilan proses pelatihan dan mencegah lonjakan pembaruan bobot yang berlebihan.

- *Epochs*: Model dilatih selama 50 *epoch* untuk memberikan kesempatan yang cukup dalam mempelajari fitur tanpa mengalami *overfitting*.
- *Batch Size*: Setiap iterasi pelatihan menggunakan 64 sampel data, yang merupakan ukuran *batch* yang seimbang antara kecepatan pelatihan dan kestabilan pembaruan bobot.
- *Early Stopping*: Diterapkan untuk menghentikan pelatihan secara otomatis jika akurasi validasi tidak meningkat setelah sejumlah epoch tertentu, guna mencegah *overfitting*.
- *Model Checkpoint*: Digunakan untuk menyimpan bobot terbaik selama pelatihan berdasarkan performa validasi, sehingga model akhir adalah versi dengan kinerja optimal
- *Eksperimen Kombinasi Parameter*
Untuk menanggapi kemungkinan bahwa nilai parameter lain dapat menghasilkan performa yang lebih baik, dilakukan observasi tambahan dengan beberapa kombinasi parameter berikut:
 - *Learning rate*: 0.00005, 0.0001, 0.001
 - *Epochs*: 30, 50, 70
 - *Batch size*: 32, 64, 128

Tujuan dari eksperimen ini adalah untuk membandingkan pengaruh kombinasi parameter terhadap akurasi, loss, dan kestabilan pelatihan. Hasilnya dianalisis untuk menentukan konfigurasi parameter yang paling optimal pada dataset klasifikasi penyakit kulit.

2.7. Evaluasi Model

Evaluasi model merupakan tahap penting untuk mengukur seberapa baik model klasifikasi bekerja pada data yang belum pernah dilihat sebelumnya. Dalam evaluasi ini digunakan beberapa metric, antara lain *confusion matrix*, *accuracy*, *precision*, *recall*, dan *f1-score*, serta visualisasi seperti kurva *accuracy-loss* dan *ROC-AUC*.

Confusion matrix adalah tabel yang menggambarkan performa model klasifikasi dengan menampilkan jumlah prediksi benar dan salah berdasarkan kategori aktual dan prediksi. Matriks ini terdiri dari empat elemen utama: *True Positive (TP)*, *False Positive (FP)*, *True Negative (TN)*, dan *False Negative (FN)*. Metrik seperti *accuracy*, *precision*, *recall*, dan *f1-score* dihitung berdasarkan nilai-nilai tersebut.

Berikut perbedaan dan fungsi dari masing-masing metri:

- **Accuracy** : mengukur proporsi prediksi yang benar dari seluruh prediksi:

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

- **Precision**: mengukur seberapa tepat model dalam memprediksi kelas positif (berapa banyak yang benar dari semua yang diprediksi positif):

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (2)$$

- **Recall**: (atau Sensitivity) mengukur seberapa baik model dalam menemukan semua kasus positif:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (3)$$

- **F1-Score**: rata-rata harmonis dari precision dan recall, yang berguna saat data tidak seimbang:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

Evaluasi performa model tidak hanya bergantung pada accuracy, tetapi juga mencakup precision, recall, f1-score, dan ROC-AUC, terutama pada dataset yang tidak seimbang. Model yang baik adalah model yang mampu menjaga keseimbangan metrik sesuai konteks, seperti recall yang tinggi untuk diagnosis penyakit. Seluruh proses evaluasi dilakukan menggunakan pustaka Python seperti TensorFlow, Keras, Scikit-learn, Matplotlib, dan Seaborn untuk menghitung metrik dan memvisualisasikan hasil secara komprehensif.

3. Hasil

Penelitian ini menghasilkan perbandingan performa dari model arsitektur CNN (*VGG16*, *ResNet50*, *MobileNetV2*, dan *InceptionV3*) dalam tugas klasifikasi penyakit kulit. Tahapan implementasi dimulai dari preprocessing data, augmentasi, transfer learning, hingga evaluasi model, yang dijelaskan secara terstruktur sebagai berikut:

3.1. Hasil Preprocessing dan Augmentasi

Sebelum pelatihan, citra mengalami proses preprocessing yang mencakup resizing ke 224×224 piksel, normalisasi piksel ke rentang [0–1], dan konversi ke grayscale (optional). Setelah itu, dilakukan augmentasi data untuk meningkatkan variasi dan mengurangi overfitting. Teknik augmentasi yang digunakan antara lain:

- **Rotasi acak** (hingga ± 30 derajat),
- **Flipping horizontal** dan vertikal,
- **Zooming** (perbesaran acak),
- **Shifting** (pergeseran horizontal/vertikal).

Contoh hasil dari masing-masing teknik augmentasi dapat dilihat pada Gambar X (disesuaikan dengan file gambar Anda). Augmentasi ini berhasil meningkatkan jumlah variasi citra tanpa menambah jumlah data asli.

3.2. Penerapan Transfer Learning

Transfer learning diterapkan dengan menggunakan model pretrained dari ImageNet untuk keempat arsitektur. Hanya lapisan klasifikasi akhir (fully-connected layer) yang dimodifikasi agar sesuai dengan jumlah kelas (9 kelas penyakit kulit). Pelatihan dilakukan dengan parameter sebagai berikut:

- *Input: citra 224×224 piksel*
- *Optimizer: Adam*
- *Learning rate: 0.0001*
- *Epoch: 50*
- *Batch size: 64*
- *Loss function: Categorical Crossentropy*

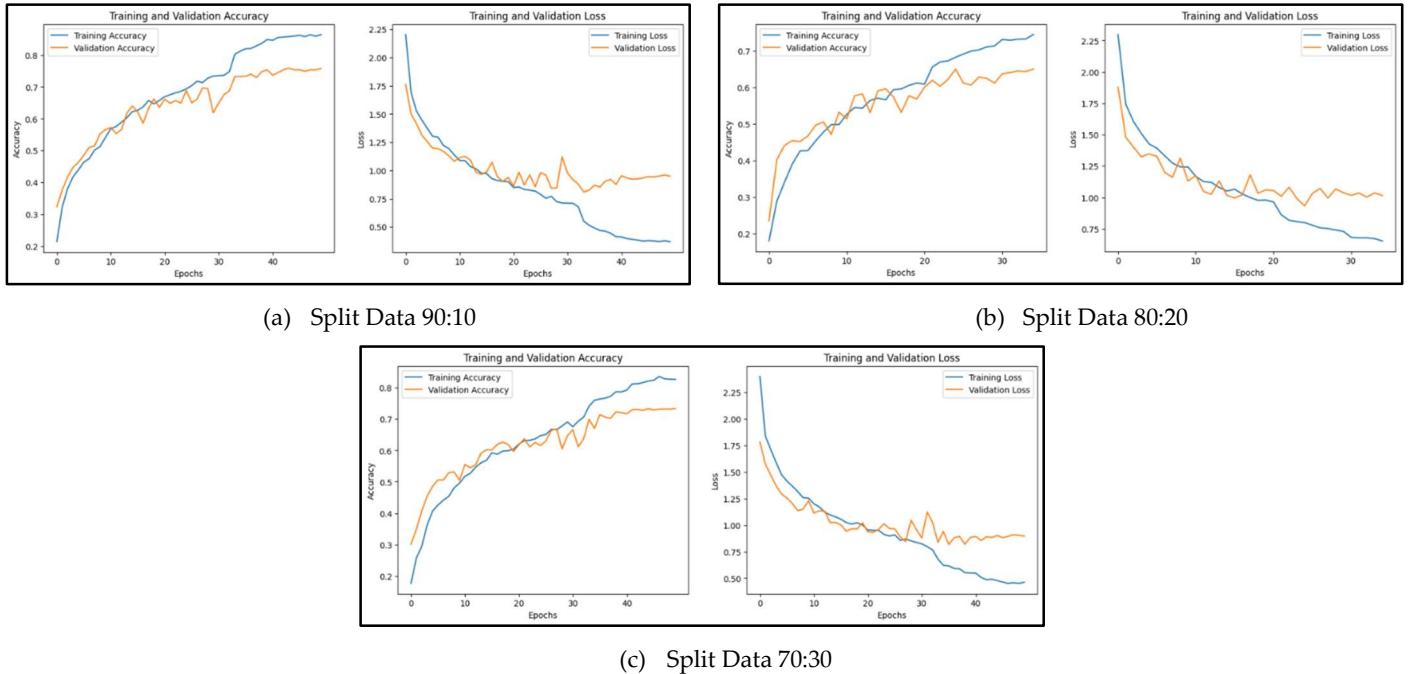
3.3. Akurasi dan Loss

Grafik akurasi data *train* dan data *test* pada masing-masing model dapat dilihat pada gambar dimana untuk masing-masing model *VGG16*, *ResNet50*, *MobileNetV2*, dan *InceptionV3* di training dengan 50 epoch.

3.3.1. VGG16

Grafik pada gambar 1 menunjukkan bahwa pada model *VGG16* pembagian data 90:10, model mencapai akurasi pelatihan sebesar 86,77% dan akurasi validasi sebesar 75,78%. Sementara itu, pembagian data 80:20 mencapai akurasi pelatihan sebesar 73,91%

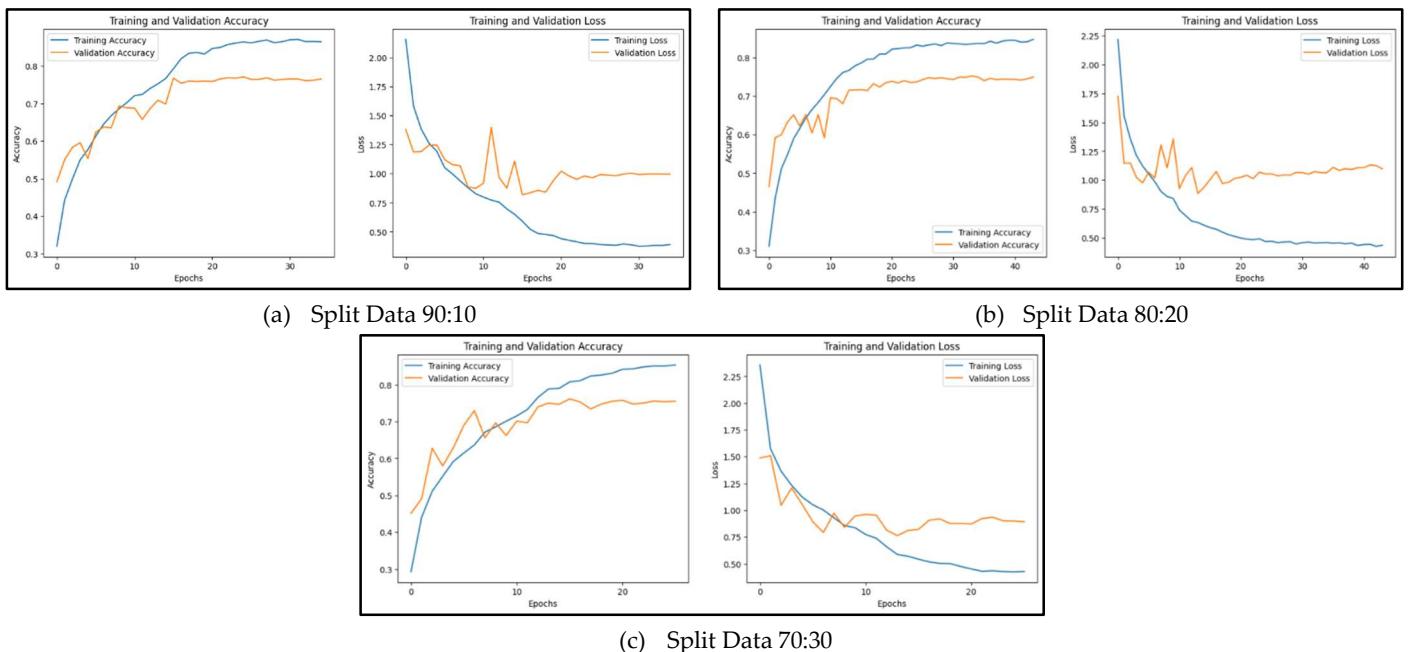
dan akurasi validasi sebesar 64,94%. Serta pembagian data 70:30 mencapai akurasi pelatihan sebesar 83,08% dan akurasi validasi sebesar 73,37%.



Gambar 2. Grafik Akurasi VGG16 Split Data 90:10, 80:20, dan 70:30.

3.3.2. ResNet50

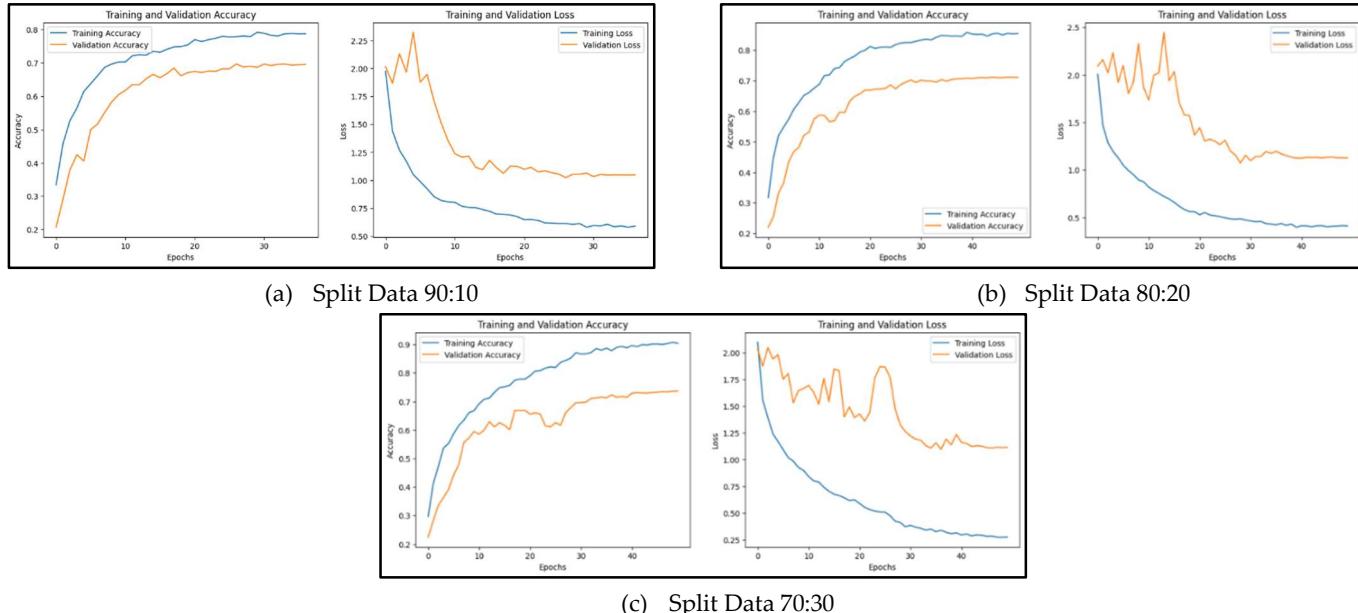
Grafik pada gambar 1 menunjukkan bahwa pada model VGG16 pembagian data 90:10, model mencapai akurasi pelatihan sebesar 86,93% dan akurasi validasi sebesar 76,56%. Sementara itu, pembagian data 80:20 mencapai akurasi pelatihan sebesar 85,03% dan akurasi validasi sebesar 74,89%. Serta pembagian data 70:30 mencapai akurasi pelatihan sebesar 85,91% dan akurasi validasi sebesar 75,52%.



Gambar 3. Grafik Akurasi ResNet50 Split Data 90:10, 80:20, dan 70:30.

3.3.3. MobileNetV2

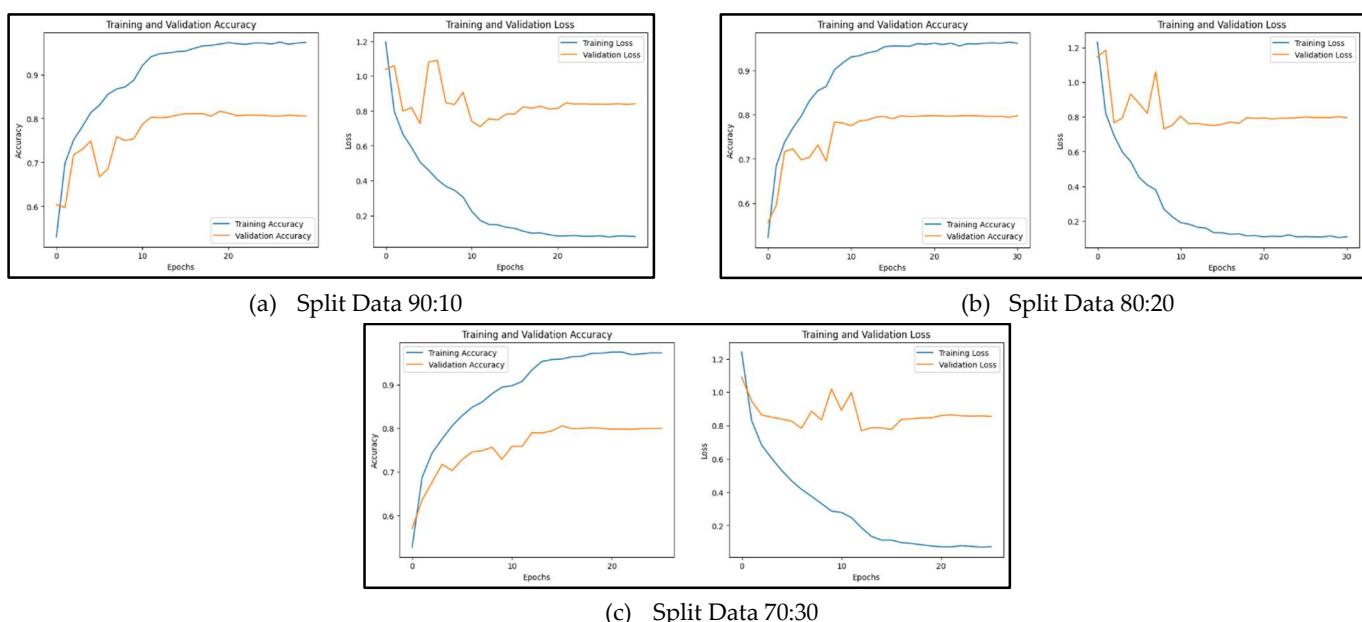
Grafik pada gambar 1 menunjukkan bahwa pada model *VGG16* pembagian data 90:10, model mencapai akurasi pelatihan sebesar 78,93% dan akurasi validasi sebesar 69,56%. Sementara itu, pembagian data 80:20 mencapai akurasi pelatihan sebesar 85,15% dan akurasi validasi sebesar 71,06%. Serta pembagian data 70:30 mencapai akurasi pelatihan sebesar 90,59% dan akurasi validasi sebesar 73,67%.



Gambar 4. Grafik Akurasi *MobileNetV2* Split Data 90:10, 80:20, dan 70:30.

3.3.4. InceptionV3

Grafik pada gambar 1 menunjukkan bahwa pada model *VGG16* pembagian data 90:10, model mencapai akurasi pelatihan sebesar 97,28% dan akurasi validasi sebesar 80,56%. Sementara itu, pembagian data 80:20 mencapai akurasi pelatihan sebesar 96,42% dan akurasi validasi sebesar 79,78%. Serta pembagian data 70:30 mencapai akurasi pelatihan sebesar 97,25% dan akurasi validasi sebesar 80,04%.



Gambar 5. Grafik Akurasi *MobileNetV2* Split Data 90:10, 80:20, dan 70:30.

3.4. Analisis Kelemahan Model

Beberapa faktor yang mempengaruhi rendahnya performa model pada kelas tertentu antara lain:

- **Kemiripan visual antar jenis penyakit** (misalnya antara eczema dan psoriasis),
- **Distribusi citra yang kurang seimbang meskipun kuantitas sama** (kualitas citra, pencahayaan, dll.),
- **Overfitting**, terutama pada model InceptionV3 yang terlalu cepat belajar dari data pelatihan,
- **Grayscale conversion** dapat menghilangkan informasi warna penting pada penyakit tertentu.

3.5. Hasil Evaluasi

Berdasarkan *confusion matrix*, berikut adalah hasil nilai *recall*, *precision*, *f1-score*, dan *accuracy* untuk model VGG16, ResNet50, MobileNet V2, dan InceptionV3.

Tabel 2. Perbandingan Hasil Evaluasi Split Data 90 : 10.

Model CNN	Precision	Recall	F1-Score	Accuracy
VGG16	76,79%	75,89%	76,15%	75,89%
ResNet50	78,58%	77,11%	77,29%	77,11%
MobileNetV2	71,30%	69,67%	69,80%	69,67%
InceptionV3	82,54%	81,67%	81,82%	81,67%

Tabel ini menunjukkan hasil evaluasi model CNN menggunakan rasio pembagian data 90% untuk pelatihan dan 10% untuk pengujian. Metrik yang ditampilkan meliputi precision, recall, f1-score, dan accuracy. InceptionV3 memperoleh hasil terbaik dengan akurasi 81,67% dan nilai f1-score tertinggi, menunjukkan kemampuannya dalam menangani data pelatihan yang besar. Sebaliknya, MobileNetV2 memiliki nilai evaluasi paling rendah pada skenario ini.

Tabel 3 . Perbandingan Hasil Evaluasi Split Data 80 : 20.

Model CNN	Precision	Recall	F1-Score	Accuracy
VGG16	66,44%	64,94%	65,04%	64,94%
ResNet50	76,72%	75,17%	75,19%	75,17%
MobileNetV2	74,16%	71,11%	71,05%	71,11%
InceptionV3	80,63%	79,83%	79,84%	79,83%

Tabel ini menampilkan performa model saat data dibagi dalam rasio 80:20. InceptionV3 tetap unggul dengan akurasi 79,83%, diikuti oleh ResNet50. Terlihat bahwa performa VGG16 menurun secara signifikan dibandingkan dengan rasio 90:10, menunjukkan sensitivitas model ini terhadap pengurangan data pelatihan.

Tabel 4. Perbandingan Hasil Evaluasi Split Data 70 : 30.

Model CNN	Precision	Recall	F1-Score	Accuracy
VGG16	74,50%	73,37%	73,58%	73,37%
ResNet50	77,45%	76,15%	76,33%	76,15%
MobileNetV2	76,12%	73,67%	73,92%	73,67%
InceptionV3	81,17%	80,59%	80,69%	80,59%

Tabel ini memperlihatkan performa model saat proporsi data pelatihan dikurangi menjadi 70%, dan pengujian ditingkatkan menjadi 30%. Hasil menunjukkan bahwa InceptionV3 masih mempertahankan akurasi tertinggi (80,59%), sementara MobileNetV2 menunjukkan peningkatan performa dibandingkan rasio sebelumnya. Performa stabil ResNet50 dan peningkatan pada MobileNetV2 menunjukkan bahwa model ini memiliki kemampuan generalisasi yang baik meskipun data pelatihan lebih sedikit.

3.6. Hasil Pengujian

Berikut ini adalah hasil pengujian model CNN yang terdiri dari VGG16, ResNet50, MobileNetV2, dan InceptionV3 pada masing-masing perbandingan data 90:10, 80:20, dan 70:30.

Tabel 5. Hasil Testing Model (90:10).

Model CNN	Waktu Proses	Kelas Asli	Kelas Prediksi	Tingkat Kepercayaan
VGG16	2 detik	5	5	55.67%
ResNet50	6 detik	5	5	16.26%
MobileNetV2	4 detik	5	5	57.30%
InceptionV3	10 detik	5	2	69.41%

Hasil pembagian data 90:10 menunjukkan kinerja model CNN dengan 90% data untuk pelatihan dan 10% untuk pengujian. VGG16 memprediksi kelas yang benar dengan tingkat kepercayaan rendah (55,67%), ResNet50 juga benar tetapi dengan kepercayaan sangat rendah (16,26%). MobileNetV2 memprediksi dengan benar dan kepercayaan lebih tinggi (57,30%), meskipun belum optimal. InceptionV3 memprediksi kelas salah (kelas 2) meskipun dengan kepercayaan tinggi (69,41%).

Tabel 6 . Hasil Testing Model (80:20).

Model CNN	Waktu Proses	Kelas Asli	Kelas Prediksi	Tingkat Kepercayaan
VGG16	2 detik	5	6	72.60%
ResNet50	6 detik	5	5	42.64%
MobileNetV2	4 detik	5	3	60.19%
InceptionV3	9 detik	5	2	37.04%

Hasil pembagian data 80:20 menunjukkan bahwa meskipun model VGG16 memprediksi kelas yang salah (kelas 6) dengan tingkat kepercayaan tinggi (72,60%), model ini cukup yakin dengan kesalahan prediksinya. ResNet50 memprediksi kelas yang benar (kelas 5) dengan kepercayaan 42,64%, sementara MobileNetV2 salah memprediksi kelas (kelas 3) dengan kepercayaan 60,19%. InceptionV3 memprediksi kelas yang salah (kelas 2) dengan kepercayaan rendah (37,04%).

Tabel 7. Hasil Testing Model (70:30).

Model CNN	Waktu Proses	Kelas Asli	Kelas Prediksi	Tingkat Kepercayaan
VGG16	4 detik	5	5	72.13%
ResNet50	6 detik	5	3	26.99%
MobileNetV2	4 detik	5	2	30.88%
InceptionV3	8 detik	5	2	41.82%

Hasil pembagian data 70:30 menunjukkan bahwa model VGG16 berhasil memprediksi label kelas yang benar, yaitu kelas 5 (jenis penyakit kulit kelima dari daftar dataset), dengan tingkat kepercayaan sebesar 72,13%. Tingkat kepercayaan ini merepresentasikan probabilitas prediksi tertinggi yang dihasilkan oleh model terhadap suatu kelas. Sebaliknya, model ResNet50 dan MobileNetV2 memprediksi kelas yang salah, masing-masing ke kelas 3 dan kelas 2, dengan tingkat kepercayaan rendah yaitu 26,99% dan 30,88%, yang menunjukkan ketidakpastian model terhadap prediksinya. Model InceptionV3 juga salah memprediksi ke kelas 2 dengan kepercayaan sebesar 41,82%.

4. Pembahasan

Penelitian ini dilakukan dengan beberapa tahapan terstruktur yang dimulai dari pengumpulan data, preprocessing, augmentasi, penerapan transfer learning, hingga evaluasi dan analisis hasil klasifikasi. Tujuan utama dari penelitian ini adalah menentukan arsitektur model CNN terbaik untuk klasifikasi penyakit kulit berdasarkan performa akurasi dan metrik evaluasi lainnya.

Dataset yang digunakan merupakan data sekunder yang diperoleh dari platform Kaggle. Dataset terdiri dari 9.000 citra berwarna dengan resolusi bervariasi dan berformat .jpg, yang telah dilabeli oleh tenaga medis profesional ke dalam sembilan jenis penyakit kulit. Data ini bersifat seimbang secara kuantitas (masing-masing kelas berisi 1.000 citra), namun tetap memiliki variasi visual yang menantang untuk klasifikasi otomatis.

Setelah pengumpulan data, tahap preprocessing dilakukan dengan mengubah ukuran seluruh citra menjadi 224×224 piksel agar sesuai dengan input model CNN. Selanjutnya, citra dinormalisasi dan dikonversi ke grayscale (opsional) untuk mengurangi kompleksitas warna.

Untuk meningkatkan kemampuan generalisasi model, dilakukan augmentasi data seperti rotasi, flipping, zooming, dan shifting. Teknik ini menambah variasi tanpa menambah jumlah data asli, sehingga membantu mengatasi potensi overfitting dan ketidakseimbangan fitur antar kelas.

Empat arsitektur CNN yaitu VGG16, ResNet50, MobileNetV2, dan InceptionV3 digunakan dengan pendekatan transfer learning. Lapisan akhir pada masing-masing model disesuaikan untuk mendeteksi sembilan kelas penyakit. Proses pelatihan dilakukan menggunakan parameter yang sama untuk semua model: ukuran input 224×224, batch size 64, learning rate 0.0001, optimizer Adam, dan 50 epoch.

Evaluasi model dilakukan dengan tiga rasio pembagian data (90:10, 80:20, dan 70:30). Hasil menunjukkan bahwa rasio 70:30 memberikan performa paling stabil dan optimal, karena menyediakan data pengujian yang cukup untuk menilai generalisasi model.

Dari keempat arsitektur yang diuji, InceptionV3 mencatatkan akurasi validasi tertinggi sebesar 80,04%. Namun, perbedaan besar antara akurasi pelatihan dan validasi menunjukkan adanya overfitting. VGG16, meskipun akurasi validasinya sedikit lebih rendah, menunjukkan generalisasi yang lebih baik dan konsisten terhadap data yang belum pernah dilihat. Model ini menjadi kandidat andal untuk implementasi nyata.

ResNet50 memberikan hasil cukup baik dengan pelatihan cepat, sedangkan MobileNetV2 efisien namun cenderung overconfident—menunjukkan kepercayaan tinggi pada prediksi yang salah.

Analisis confusion matrix dan classification report menunjukkan bahwa beberapa kelas penyakit seperti eczema, psoriasis, dan tinea sering salah diklasifikasikan. Hal ini dapat disebabkan oleh kemiripan visual antar kelas dan kualitas visual yang bervariasi.

Untuk meningkatkan akurasi di masa mendatang, disarankan untuk menambah data atau menerapkan augmentasi yang lebih selektif pada kelas-kelas dengan tingkat kesalahan tinggi.

Meskipun penelitian ini telah melibatkan proses augmentasi untuk meningkatkan variasi data, hasil klasifikasi menunjukkan bahwa misklasifikasi tetap terjadi pada beberapa kelas penyakit kulit seperti eczema, psoriasis, dan tinea. Hal ini dapat disebabkan oleh kemiripan visual yang tinggi antar kelas tersebut atau distribusi fitur yang belum sepenuhnya terwakili dalam proses augmentasi yang dilakukan. Oleh karena itu, augmentasi lanjutan yang lebih terfokus pada kelas-kelas bermasalah, atau penambahan data nyata dengan kualitas visual tinggi, dapat menjadi solusi untuk meningkatkan akurasi klasifikasi di masa mendatang.

Secara keseluruhan, implementasi CNN dalam klasifikasi citra penyakit kulit telah menunjukkan hasil yang menjanjikan. Berdasarkan evaluasi kuantitatif, model InceptionV3 memperoleh nilai evaluasi tertinggi pada rasio data 70:30 dengan precision 81,17%, recall 80,59%, f1-score 80,69%, dan akurasi 80,59%. Meskipun demikian, model ini menunjukkan kecenderungan overfitting. Sebaliknya, model VGG16 meskipun memiliki akurasi validasi sedikit lebih rendah (73,37%), menunjukkan performa yang lebih stabil dan kemampuan generalisasi yang lebih baik. Oleh karena itu, kombinasi antara arsitektur VGG16 dan rasio data 70:30 dinilai sebagai pendekatan paling efektif dalam penelitian ini.

5. Kesimpulan

Penelitian ini berhasil mengimplementasikan empat arsitektur CNN—VGG16, ResNet50, MobileNetV2, dan InceptionV3—untuk klasifikasi citra penyakit kulit. Evaluasi menunjukkan bahwa rasio data 70:30 memberikan hasil terbaik karena meningkatkan kemampuan model dalam mengenali pola. InceptionV3 mencapai akurasi validasi tertinggi (80,04%), namun menunjukkan potensi overfitting. Sebaliknya, VGG16 menunjukkan generalisasi yang lebih baik terhadap data uji.

Ke depan, penelitian dapat dikembangkan dengan augmentasi adaptif untuk kelas yang sulit diklasifikasi, serta eksplorasi arsitektur CNN yang lebih ringan dan efisien untuk implementasi pada sistem diagnosis berbasis mobile.

Referensi

- [1] R. Agarwal dan D. Godavarthi, "Skin Disease Classification Using CNN Algorithms," *EAI Endorsed Trans. Pervasive Heal. Technol.*, vol. 9, no. 1, hal. 1–8, 2023, <https://doi.org/10.4108/eetph.9.4039>.
- [2] D. D. Putri, M. T. Furqon, dan R. S. Perdana, "Klasifikasi Penyakit Kulit Pada Manusia Menggunakan Metode Binary Decision Tree Support Vector Machine (BDTSVM)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 5, hal. 1912–1920, 2019, <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/1425>.
- [3] N. S. Rahayu, A. D. Puteri, dan L. M. A. Isnaeni, "Hubungan Perilaku Masyarakat Dan Penggunaan Air Sungai Dengan Gangguan Penyakit Kulit Di Desa Kampung Pinang Wilayah Kerja Puskesmas Pantai Raja," *J. Imliah Ilmu Kesehat.*, vol. 1, no. 3, hal. 2023, 2023.
- [4] F. N. Darmawan, E. P. Silmina, dan T. Hardiani, "Sistem Klasifikasi Penyakit Kulit Menggunakan Metode Convolutional Neural Network (CNN) Berbasis Website," *Pros. Semin. Nas. LPMP Univ. 'Aisyiyah Yogyakarta*, vol. 2, hal. 871–881, 2024, <https://proceeding.unisyogya.ac.id/index.php/proseminnaslpmp/article/view/796>
- [5] F. Mahyudin, M. Edward, M. H. Basuki, dan Y. A. Barri, "Diagnosis dan Terapi Tumor Muskuloskeletal," *Sagung Seto*, hal. 119–150, 2017.
- [6] A. Ajrana, A. Lawi, dan A. M. A. Siddik, "Implementasi Arsitektur Dengan Pemilihan Model Transfer Learning Convolutional Neural Network Dalam Mengklasifikasikan Penyakit Kanker Kulit," *Pros. Semin. Nas. Tek. Elektro dan Inform.*, vol. 8, no. 1, hal. 292–297, 2023, <https://jurnal.poliupg.ac.id/index.php/snsei/article/view/3628>
- [7] T. Saputra dan M. E. Al-Rivan, "Analisis Performa ResNet-152 dan AlexNet dalam Klasifikasi Jenis Kanker Kulit," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.)*, vol. 8, no. 1, hal. 75, 2023, <https://doi.org/10.30998/string.v8i1.16464>.

- [8] S. Supirman, C. Lubis, D. Yuliarto, dan N. J. Perdana, "Klasifikasi Penyakit Kulit Menggunakan Convolutional Neural Network (Cnn) Dengan Arsitektur Vgg16," *Simtek J. Sist. Inf. dan Tek. Komput.*, vol. 8, no. 1, hal. 135–140, 2023, <https://doi.org/10.51876/simtek.v8i1.217>.
- [9] W. M. Pradnya Dhuhita, M. Y. Ubaid, dan A. Baita, "MobileNet V2 Implementation in Skin Cancer Detection," *Ilk. J. Ilm.*, vol. 15, no. 3, hal. 498–506, Des 2023, <https://doi.org/10.33096/ilkom.v15i3.1702.498-506>.
- [10] T. O. Saputra dan D. Alamsyah, "Klasifikasi Penyakit Cacar Monyet Menggunakan," *Mdp Student Conf.*, hal. 179–184, 2023, <https://doi.org/10.35957/mdp-sc.v2i1.4400>.
- [11] Amania Salwa Ahla, Sri Mulyono, dan Sam Farisa Chaerul Haviana, "Klasifikasi Jenis Jerawat Wajah Menggunakan Arsitektur Inception V3," *J. Ilm. Sultan Agung*, no. September, hal. 738–752, 2023, <https://jurnal.unissula.ac.id/index.php/JIMU/article/view/33631>.
- [12] A. E. Putra, M. F. Naufal, dan V. R. Prasetyo, "Klasifikasi Jenis Rempah Menggunakan Convolutional Neural Network dan Transfer Learning," *J. Edukasi dan Penelit. Inform.*, vol. 9, no. 1, hal. 12, 2023, <https://doi.org/10.26418/jp.v9i1.58186>.
- [13] C. A. Sanjaya, M. Waluyo, dan T. Industri, "ANALISIS PERBANDINGAN METODE TRANSFER LEARNING DENSE-NET201 DAN VGG-19 TERHADAP," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 1, 2025, <http://dx.doi.org/10.23960/jitet.v13i1.5810>.
- [14] F. Marpaung, F. Aulia, dan R. C. Nabila, *Computer Vision Dan Pengolahan Citra Digital*. 2022. [Daring]. Tersedia pada: www.pustakaaksara.co.id
- [15] A. Dhini Sephya, Rahmaddeni, Susanti, "Penerapan Algoritma Convolutional Neural Network Untuk Klasifikasi Penyakit Kanker Kulit," *Indones. J. Comput. Sci.*, vol. 13, no. 4, hal. 6590–6600, 2024, <https://doi.org/10.33022/ijcs.v13i4.4262>.
- [16] D. A. Agustina, "Klasifikasi Citra Jenis Kulit Wajah Dengan Algoritma Convolutional Neural Network (Cnn) Resnet-50," *J. Ris. Sist. Inf.*, vol. 1, no. 3, hal. 01–07, 2024, <https://doi.org/10.69714/13sbby24>.
- [17] N. I. Khani dan S. Rakasiwi, "Penerepan Convolutional Neural Network dengan ResNet-50 untuk Klasifikasi Penyakit Kulit Wajah Efektif," *J. Pendidik. Inform.*, vol. 9, no. 1, hal. 217–225, 2025, <https://doi.org/10.29408/edumatic.v9i1.29572>.
- [18] A. ANHAR dan R. A. PUTRA, "Perancangan dan Implementasi Self-Checkout System pada Toko Ritel menggunakan Convolutional Neural Network (CNN)," *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 11, no. 2, hal. 466, 2023, <https://doi.org/10.26760/elkomika.v11i2.466>.
- [19] R. FATURRAHMAN, Y. S. HARIYANI, dan S. HADIYOSO, "Klasifikasi Jajanan Tradisional Indonesia berbasis Deep Learning dan Metode Transfer Learning," *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 11, no. 4, hal. 945, 2023, <https://doi.org/10.26760/elkomika.v11i4.945>.
- [20] I. Mustikasari, "Identifikasi Kanker Kulit Melanoma Berbasis Inception V3 Menggunakan Pra-Pemrosesan dan Augmentasi Data pada Dataset Citra Kulit," *Telkom Univ.*, vol. 10, no. 5, hal. 4170–4176, 2023, https://repository.telkomuniversity.ac.id/pustaka/files/196591/jurnal_eproc/identifikasi-kanker-kulit-melanoma-berbasis-inception-v3-menggunakan-pra-pemrosesan-dan-augmentasi-data-pada-dataset-citra-kulit.pdf
- [21] A. W. Kosman, Y. Wahyuningsih, dan F. Mahendrasusila, "Pengujian Metode Inception V3 dalam Mengidentifikasi Penyakit Kanker Kulit," *J. Teknologi Inform. dan Komput.*, vol. 10, no. 1, hal. 136–146, 2024, <https://doi.org/10.37012/jtik.v10i1.1940>.
- [22] O. Sharif, M. M. Hoque, A. S. M. Kayes, R. Nowrozy, dan I. H. Sarker, "applied sciences Learning Techniques," hal. 1–23, 2020.
- [23] A. M. Ibrahim, M. Elbasheir, S. Badawi, A. Mohammed, dan A. F. M. Alalmin, "Skin Cancer Classification Using Transfer Learning by VGG16 Architecture (Case Study on Kaggle Dataset)," *J. Intell. Learn. Syst. Appl.*, vol. 15, no. 03, hal. 67–75, 2023, <https://doi.org/10.4236/jilsa.2023.153005>.