

**KLASIFIKASI CITRA KULIT SEHAT DAN TIDAK SEHAT  
MENGUNAKAN METODE SIMSIAM**

**TUGAS AKHIR**

Diajukan oleh:

MUHAMMAD HAFIZ

NIM.220705082

Mahasiswa Fakultas Sains dan Teknologi  
Program Studi Teknologi Informasi



**FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI AR-RANIRY  
BANDA ACEH  
2026 M/1447 H**

## LEMBAR PERSETUJUAN

### KLASIFIKASI CITRA KULIT SEHAT DAN TIDAK SEHAT MENGUNAKAN METODE SIMSIAM

#### TUGAS AKHIR

Diajukan Kepada Fakultas Sains dan Teknologi  
Universitas Islam Negeri (UIN) Ar-Raniry Banda Aceh  
Sebagai Salah Satu Beban Studi Memperoleh Gelar Sarjana (S1)  
dalam Prodi Teknologi Informasi

Oleh:

**MUHAMMAD HAFIZ**  
220705082

**Mahasiswa Fakultas Sains dan Teknologi**  
**Program Studi Teknologi Informasi**

Disetujui untuk Dimunaqasyahkan Oleh:

Pembimbing I,

  
**Malahayati, M.T**

NIP. 198301272015032003

Pembimbing II,

  
**Dr. Hendri Ahmadian, M.I.M.**

NIP. 198301042014031002

Mengetahui,  
Ketua Program Studi Teknologi Informasi

  
**Malahayati, M.T**

NIP. 198301272015032003

## LEMBAR PENGESAHAN

### KLASIFIKASI CITRA KULIT SEHAT DAN TIDAK SEHAT MENGUNAKAN METODE SIMSIAM

#### TUGAS AKHIR

Telah Diuji Oleh Panitia Ujian Munaqasyah Tugas Akhir  
Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh dan Dinyatakan Lulus  
Serta Diterima Sebagai Salah Satu Beban Studi Program Sarjana (S-1)  
Dalam Prodi Teknologi Informasi

Pada Hari/Tanggal: Senin, 11 Mei 2026  
23 Dzulqaidah 1447 H

di Darussalam, Banda Aceh

Panitia Ujian Munaqasyah Tugas Akhir:

Ketua,

**Malahayati, M.T.**  
NIP. 198301272015032003

Sekretaris,

**Dr. Hendri Ahmadian, M.I.M.**  
NIP. 198301042014031002

Penguji I,

**Khairan AR, M.Kom.**  
NIP. 198607042014031001


Penguji II,

**Baihaqi, M.T.**  
NIP. 198802212022031001

Mengetahui:

Dekan Fakultas Sains dan Teknologi  
UIN Ar-Raniry Banda Aceh



  
**Prof. Dr. Muhammad Dirhamsyah, M.T., IPU**  
NIP. 196210021988111001

## LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini:

Nama : Muhammad Hafiz  
NIM : 220705082  
Program Studi : Teknologi Informasi  
Fakultas : Sains dan Teknologi  
Judul : Klasifikasi Citra Kulit Sehat Dan Tidak Sehat  
Menggunakan Metode Simsim

Dengan ini menyatakan bahwa dalam penulisan tugas akhir, saya:

1. Tidak menggunakan ide orang lain tanpa mampu mengembangkan dan mempertanggungjawabkan;
2. Tidak melakukan plagiasi terhadap naskah karya orang lain;
3. Tidak menggunakan karya orang lain tanpa menyebutkan sumber asli atau tanpa izin pemilik karya;
4. Tidak memanipulasi dan memalsukan data;
5. Mengerjakan sendiri karya ini dan mampu bertanggungjawab atas karya ini.

Bila dikemudian hari ada tuntutan dari pihak lain atas karya saya, dan telah melalui pembuktian yang dapat dipertanggungjawabkan dan ternyata memang ditemukan bukti bahwa saya telah melanggar pernyataan ini, maka saya siap dikenai sanksi berdasarkan aturan yang berlaku di Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh. Demikian pernyataan ini saya buat dengan sesungguhnya dan tanpa paksaan dari pihak manapun.

Banda Aceh, 11 Mei 2026  
Yang Menyatakan



  
Muhammad Hafiz

## ABSTRAK

Nama : Muhammad Hafiz  
NIM : 220705082  
Program Studi : Teknologi Informasi  
Fakultas : Sains dan Teknologi (FST)  
Judul : Klasifikasi Citra Kulit Sehat Dan Tidak Sehat  
Menggunakan Metode SimSiam  
Tanggal Sidang : 11 Mei 2026 / 23 Dzulqaidah 1447 H  
Jumlah Halaman : 62 Halaman  
Pembimbing I : Malahayati, M.T.  
Pembimbing II : Dr. Hendri Ahmadian, M.I.M.  
Kata Kunci : *Self-Supervised Learning, SimSiam, ResNet-18, Klasifikasi Citra, Citra Kulit, Grad-CAM.*

Penelitian ini bertujuan mengklasifikasikan citra kulit sehat dan tidak sehat menggunakan metode *Self-Supervised Learning* (SSL) SimSiam untuk memitigasi keterbatasan ketersediaan data berlabel pada domain medis. Sistem ini menggunakan arsitektur ResNet-18 sebagai *backbone encoder* yang dilatih dengan teknik *dual stochastic augmentation* untuk mempelajari representasi fitur visual secara mandiri.

Pengujian dilakukan menggunakan total 800 dataset citra, yang terdiri dari 200 data primer hasil akuisisi kamera *smartphone* dan 600 data sekunder dari repositori Kaggle serta Dermnet. Hasil menunjukkan model mampu beroperasi stabil dengan tingkat akurasi 96,67%, sensitivitas (*recall*) 98,46%, dan spesifisitas 94,55%. Melalui visualisasi Grad-CAM, sistem terbukti secara presisi mampu mengidentifikasi area lesi patologis pada citra kulit. Sistem ini diharapkan dapat membantu proses klasifikasi kondisi kulit secara digital dengan tingkat efisiensi anotasi yang lebih tinggi.

**Kata Kunci:** *Self-Supervised Learning, SimSiam, ResNet-18, Klasifikasi Citra, Citra Kulit, Grad-CAM*

## ABSTRACT

Nama : Muhammad Hafiz  
NIM : 220705082  
Program Studi : Information Technology  
Fakultas : Science and Technology  
Judul : Classification of Healthy and Diseased Skin Images Using the SimSiam Method  
Tanggal Sidang : 11 May 2026 / 23 Dzulqaidah 1447 H  
Jumlah Halaman : 62 Pages  
Pembimbing I : Malahayati, M.T.  
Pembimbing II : Dr. Hendri Ahmadian, M.I.M.  
Kata Kunci : Self-Supervised Learning, SimSiam, ResNet-18, Image Classification, Grad-CAM

This research aims to classify healthy and diseased skin images using the SimSiam Self-Supervised Learning (SSL) method to mitigate the limited availability of labeled data in the medical domain. The system utilizes a ResNet-18 architecture as the backbone encoder, which is trained using dual stochastic augmentation techniques to learn visual feature representations independently.

Testing was conducted using a total dataset of 800 images, consisting of 200 primary data acquired via smartphone camera and 600 secondary data from Kaggle and Dermnet repositories. The results showed that the model performed stably with an accuracy rate of 96.67%, a sensitivity (recall) of 98.46%, and a specificity of 94.55%. Through Grad-CAM visualization, the system was proven to precisely identify pathological lesion areas on skin images. This system is expected to support the digital skin condition classification process with higher annotation efficiency.

**Keywords:** *Self-Supervised Learning, SimSiam, ResNet-18, Image Classification, Grad-CAM*

## KATA PENGANTAR

Segala puji dan rasa syukur yang mendalam penulis persembahkan ke hadirat Allah Swt yang telah melimpahkan rahmat, taufik, dan hidayah-Nya, sehingga penulis dapat menyelesaikan penulisan Proposal Tugas Akhir. Shalawat beserta salam tidak lupa penulis haturkan kepada junjungan alam Nabi Muhammad SAW, beserta keluarga dan para sahabatnya yang telah membawa umat manusia dari alam kebodohan ke alam yang penuh dengan ilmu pengetahuan.

Proposal Tugas Akhir ini berjudul "Klasifikasi Citra Kulit Sehat Dan Tidak Sehat Menggunakan Metode Simsim". Dokumen proposal ini disusun untuk memenuhi salah satu tahapan prasyarat dalam pengerjaan skripsi di bawah naungan Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Ar-Raniry, Banda Aceh.

Penulis menyadari sepenuhnya bahwa dalam penulisan proposal ini tidak terlepas dari bimbingan, arahan, dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua dan keluarga tercinta yang selalu memberikan doa yang tiada henti, kasih sayang, serta dukungan baik moril maupun materil.
2. Ibu Malahayati, M.T., yang menjabat sebagai Ketua Program Studi Teknologi Informasi sekaligus pembimbing yang telah memberikan masukan, bimbingan dan arahan selama proses penyusunan proposal ini.
3. Bu Cut Ida Rahmadiana, S. Si., selaku Staf dan pegawai di Program Studi Teknologi Informasi yang telah banyak membantu, bimbingan, dan kerja sama yang baik.
4. Sahabat seperjuangan di Program Studi Teknologi Informasi angkatan 2022 atas segala bentuk energi positif, dukungan, dan motivasi yang senantiasa mengiringi langkah penulis dalam merampungkan proposal ini.

Banda Aceh, 7 Mei 2026



Muhammad Hafiz

## DAFTAR ISI

<b>LEMBAR PERSETUJUAN .....</b>	<b>i</b>
<b>LEMBAR PENGESAHAN .....</b>	<b>ii</b>
<b>LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR.....</b>	<b>iii</b>
<b>KATA PENGANTAR.....</b>	<b>vi</b>
<b>DAFTAR ISI.....</b>	<b>vii</b>
<b>DAFTAR TABEL .....</b>	<b>ix</b>
<b>DAFTAR GAMBAR.....</b>	<b>x</b>
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian .....	3
1.4 Manfaat Penelitian .....	3
1.5 Batasan Masalah.....	3
<b>BAB II LANDASAN TEORI .....</b>	<b>5</b>
2.1 Penelitian Terdahulu .....	5
2.2 Landasan Teori.....	9
2.2.1 Kulit dan Dermatologi Visual .....	9
2.2.2 Pengolahan Citra Digital ( <i>Digital Image Processing</i> ) .....	9
2.2.3 Machine Learning .....	10
2.2.4 Deep Learning .....	12
2.2.5 Convolutional Neural Network (CNN).....	13
2.2.6 Residual Network (ResNet) .....	14
2.2.7 <i>Self-supervised learning</i> (SSL) .....	15
2.2.8 SimSiam ( <i>Simple Siamese Network</i> ).....	16
2.2.9 Kerangka Kerja Pipa Komputasi Arsitektur Usulan .....	18
<b>BAB III METODE PENELITIAN .....</b>	<b>21</b>
3.1 Jenis Penelitian.....	21
3.2 Tahapan Penelitian .....	21
3.2.2 Preprocessing dan Augmentasi Data.....	25

3.2.3	Arsitektur Pemodelan <i>Self-supervised learning</i> (SimSiam) .....	26
3.2.4	Proses <i>Fine-Tuning</i> dan Klasifikasi Biner .....	26
3.2.5	Tahap Evaluasi Peforma .....	27
3.3	Waktu dan Penelitian .....	28
3.4	Alat dan Bahan.....	28
3.4.1	Perangkat Keras ( <i>Hardware</i> ) .....	28
3.4.2	Perangkat Lunak ( <i>Software</i> ).....	29
<b>BAB IV</b>	<b>HASIL DAN PEMBAHASAN .....</b>	<b>31</b>
4.1	Lingkungan Implementasi dan Dataset.....	31
4.1.1	Lingkungan Implementasi.....	31
4.1.2	Sumber dan Pre-pemrosesan Dataset.....	31
4.1.3	Organisasi Data Loader Self-Supervised .....	31
4.2	Pra-pemrosesan dan Augmentasi Data.....	32
4.2.1	Pra-pemrosesan Citra ( <i>Image Preprocessing</i> ) .....	32
4.2.2	Pipa Augmentasi Data ( <i>Data Augmentation Pipeline</i> ).....	33
4.3	Hasil Pelatihan Mandiri ( <i>Self-Supervised Pretraining</i> ) .....	34
4.4	Hasil <i>Linear Evaluation</i> dan Klasifikasi Biner .....	36
4.5	Pengujian Inferensi Model ( <i>Model Inference</i> ).....	39
4.6	Pembahasan dan Analisis .....	43
<b>BAB V</b>	<b>KESIMPULAN DAN SARAN.....</b>	<b>45</b>
5.1	Kesimpulan .....	45
5.2	Saran.....	46
<b>DAFTAR PUSTAKA</b>	<b>.....</b>	<b>48</b>

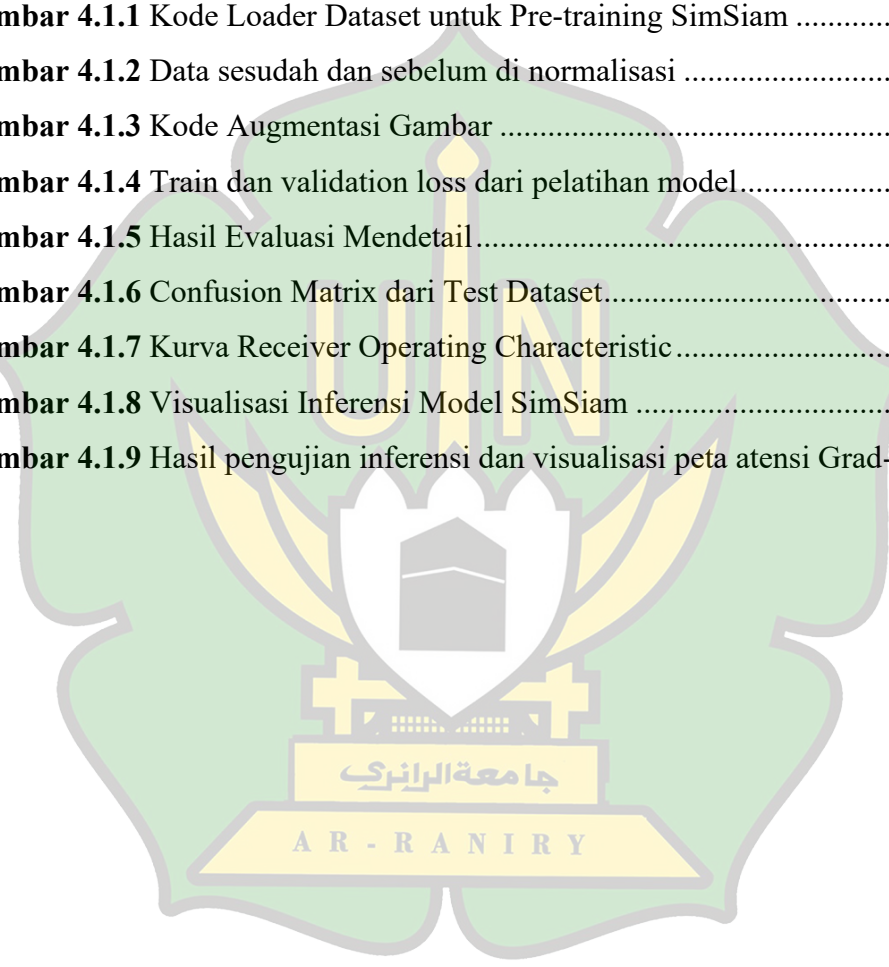
## DAFTAR TABEL

<b>Tabel 2.1</b> Penelitian Terdahulu.....	5
<b>Tabel 3.1</b> Jadwal Penelitian.....	28
<b>Tabel 3.2</b> Komponen Perangkat Keras.....	29
<b>Tabel 3.3</b> Komponen Perangkat Lunak.....	30



## DAFTAR GAMBAR

<b>Gambar 2.2.1</b>	Alur umum Self-Supervised Learning.....	16
<b>Gambar 2.2.2</b>	Arsitektur SimSiam .....	17
<b>Gambar 2.2.3</b>	Representasi Skematik Kerangka Kerja Pipa Komputasi SimSiam-ResNet18.....	20
<b>Gambar 3.2.1</b>	Alur Penelitian .....	21
<b>Gambar 4.1.1</b>	Kode Loader Dataset untuk Pre-training SimSiam .....	32
<b>Gambar 4.1.2</b>	Data sesudah dan sebelum di normalisasi .....	33
<b>Gambar 4.1.3</b>	Kode Augmentasi Gambar .....	34
<b>Gambar 4.1.4</b>	Train dan validation loss dari pelatihan model.....	35
<b>Gambar 4.1.5</b>	Hasil Evaluasi Mendetail.....	37
<b>Gambar 4.1.6</b>	Confusion Matrix dari Test Dataset.....	38
<b>Gambar 4.1.7</b>	Kurva Receiver Operating Characteristic .....	39
<b>Gambar 4.1.8</b>	Visualisasi Inferensi Model SimSiam .....	40
<b>Gambar 4.1.9</b>	Hasil pengujian inferensi dan visualisasi peta atensi Grad-CAM	42



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi kecerdasan buatan *Artificial Intelligence* (AI), khususnya pada bidang *deep learning*, terus mengalami kemajuan pesat dalam beberapa tahun terakhir dan memberikan kontribusi signifikan dalam analisis citra digital. Pendekatan *deep learning* memungkinkan sistem untuk mempelajari representasi fitur secara otomatis dari data citra, sehingga banyak digunakan dalam berbagai tugas seperti klasifikasi citra, deteksi objek, dan segmentasi. Salah satu arsitektur yang paling dominan dalam pengolahan citra adalah *Convolutional Neural Networks* (CNN), yang terbukti efektif dalam mengekstraksi fitur visual secara hierarkis dan representatif (Mevorach et al., 2025).

Dalam bidang kesehatan, khususnya kesehatan kulit, analisis citra kulit berbasis komputer memiliki potensi aplikasi yang luas, antara lain untuk pengenalan kondisi kulit, pemantauan kesehatan kulit secara visual, serta sebagai sistem pendukung dalam perawatan kulit. Penelitian-penelitian terkini menunjukkan bahwa model berbasis CNN mampu memberikan performa yang baik dalam klasifikasi citra kulit dan citra lesi kulit, terutama ketika dilatih menggunakan data berlabel yang memadai (Khan et al., 2022; Tahir et al., 2023).

Pendekatan *supervised learning* memiliki keterbatasan utama, yaitu ketergantungan terhadap ketersediaan data berlabel dalam jumlah besar. Pada domain citra kulit, proses pelabelan data sering kali memerlukan keahlian tenaga profesional seperti dermatolog, sehingga membutuhkan biaya, waktu, dan sumber daya yang besar. Selain itu, adanya subjektivitas dan kerentanan *human error* dalam proses pelabelan juga dapat memengaruhi konsistensi dan kualitas data latih (De et al., 2024). Keterbatasan ini menjadi tantangan dalam pengembangan sistem klasifikasi citra kulit yang efisien dan berskala besar.

Untuk mengatasi permasalahan tersebut, berkembang paradigma *Self-Supervised Learning* (SSL), yaitu pendekatan pembelajaran yang memungkinkan model mempelajari representasi fitur dari data tanpa memerlukan label manual. SSL memanfaatkan struktur intrinsik data dengan membandingkan berbagai

transformasi (*augmentations*) dari citra yang sama, sehingga model dapat belajar representasi visual yang bermakna dari data tidak berlabel (Gui et al., 2024). Pendekatan ini sangat relevan untuk domain kesehatan yang umumnya memiliki keterbatasan data berlabel.

Salah satu metode SSL yang relatif sederhana namun efektif adalah SimSiam (*Simple Siamese Network*) yang diperkenalkan oleh Chen dan He (2021). SimSiam menggunakan arsitektur jaringan siamese untuk mempelajari kesamaan representasi dari dua augmentasi berbeda pada citra yang sama tanpa memerlukan *negative samples* maupun *memory bank*. Dengan desain *stop-gradient* dan *predictor network*, SimSiam mampu menghindari *representation collapse* sekaligus mempertahankan stabilitas pelatihan, sehingga menghasilkan representasi fitur yang kuat dengan kompleksitas arsitektur yang relatif rendah.

Dalam beberapa tahun terakhir, penerapan SSL pada domain medis menunjukkan hasil yang menjanjikan. Penelitian oleh Zhao et al. (2024) serta Chattopadhyay et al. (2023) menunjukkan bahwa model *self-supervised pretraining* mampu meningkatkan performa klasifikasi citra medis ketika jumlah data berlabel terbatas. Meskipun validasi klinis terhadap sistem klasifikasi citra berbasis AI mulai menunjukkan potensi besar sebagai alat bantu pendukung (Mevorach et al., 2025), penerapan metode SimSiam secara khusus pada klasifikasi citra kulit, terutama untuk membedakan kondisi kulit sehat dan kulit tidak sehat menggunakan citra non-dermoskopik, masih relatif terbatas dan belum banyak dikaji secara mendalam.

Berdasarkan permasalahan tersebut, diperlukan suatu pendekatan yang mampu memanfaatkan data citra kulit tidak berlabel untuk mempelajari representasi fitur yang informatif sebelum digunakan pada tugas klasifikasi. Oleh karena itu, penelitian ini mengusulkan penerapan metode *Self-supervised learning* SimSiam sebagai tahap *pretraining* untuk mempelajari representasi citra kulit, yang selanjutnya dilakukan *fine-tuning* pada tugas klasifikasi kondisi kulit menjadi dua kelas, yaitu kulit sehat dan kulit tidak sehat. Pendekatan ini diharapkan dapat mengurangi ketergantungan terhadap data berlabel sekaligus meningkatkan performa klasifikasi citra kulit.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana menerapkan metode SimSiam untuk mempelajari representasi citra kulit tanpa label?
2. Bagaimana hasil evaluasi performa model menggunakan metrik akurasi, presisi, recall, spesifisitas dan F1-score?

## 1.3 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah:

1. Menerapkan metode SimSiam untuk pembelajaran representasi citra kulit secara *Self-supervised learning*.
2. Menganalisis performa model klasifikasi berdasarkan metrik evaluasi yang digunakan.

## 1.4 Manfaat Penelitian

Melalui penelitian ini, penulis menargetkan pencapaian berbagai manfaat, di antaranya:

1. **Manfaat Teoritis:** Memberikan pemahaman lebih mendalam mengenai penerapan teknik *self-supervised learning* khususnya SimSiam pada bidang klasifikasi citra kulit.
2. **Manfaat Praktis:** Memberikan gambaran bahwa metode SimSiam dapat dimanfaatkan untuk membantu proses klasifikasi kondisi kulit berbasis citra, terutama pada kondisi keterbatasan data berlabel.
3. **Manfaat Akademis:** Menjadi referensi dan bahan rujukan bagi penelitian selanjutnya yang berkaitan dengan *self-supervised learning* dan analisis citra kesehatan.

## 1.5 Batasan Masalah

Demi menjaga konsistensi dan mencegah meluasnya pembahasan dari target utama, penulis menetapkan sejumlah batasan ruang lingkup sebagai berikut:

1. Klasifikasi dilakukan hanya pada dua kategori, yaitu kulit sehat dan kulit tidak sehat.

2. Metode yang digunakan dibatasi pada *self-supervised learning* SimSiam sebagai tahap *pretraining*.
3. Evaluasi performa dilakukan menggunakan metrik akurasi, presisi, *recall*, spesifisitas dan *F1-score*.
4. Model ini tidak memiliki dataset default berupa bukan kulit dikarenakan batasan dari *Self-supervised learning*.
5. Penelitian ini tidak bertujuan untuk melakukan diagnosis medis atau menggantikan peran tenaga medis profesional.



## BAB II LANDASAN TEORI

### 2.1 Penelitian Terdahulu

Tinjauan terhadap riset-riset sebelumnya memiliki peran yang sangat penting sebagai landasan dalam penyusunan penelitian ini. Studi studi sebelumnya digunakan untuk memperkuat teori serta membantu mengidentifikasi celah penelitian yang masih belum banyak dibahas. Beberapa penelitian yang relevan dengan topik dan metode yang digunakan dalam penelitian ini disajikan pada Tabel 2.1.

**Tabel 2.1** *Penelitian Terdahulu*

No	Peneliti & Tahun	Judul	Metode	Kesimpulan
1.	Mevorach et al. (2025)	A comparison of skin lesions' diagnoses between AI-based image classification, an expert dermatologist, and a non-expert	Melakukan studi komparatif diagnosis lesi kulit menggunakan klasifikasi citra berbasis AI terhadap penilaian ahli dermatologi dan non-ahli.	Memberikan validasi klinis mengenai potensi penggunaan sistem klasifikasi citra berbasis AI sebagai alat bantu pendukung dalam perawatan kulit.
2.	De et al. (2024)	An approach to the dermatological classification of histopathological skin images using a hybridized	Mengimplementasikan model hibrida CNN-DenseNet untuk klasifikasi dermatologi pada citra histopatologi kulit.	Pendekatan ini mampu memitigasi tantangan terkait subjektivitas dan risiko <i>human error</i> dalam proses pelabelan data medis manual.

		CNN-DenseNet model		
3.	Zhao et al. (2024)	A comparison review of transfer learning and self-supervised learning: Definitions, applications, advantages and limitations	Melakukan tinjauan perbandingan antara metode <i>transfer learning</i> dan <i>self-supervised learning</i> (SSL) dalam berbagai aplikasi medis.	SSL menunjukkan kapabilitas yang unggul dalam ekstraksi fitur visual pada domain spesifik medis dibandingkan teknik <i>transfer learning</i> konvensional.
4.	Chattopadhyay et al. (2023)	Exploring Self-Supervised Representation Learning for Low-Resource Medical Image Analysis	Mengevaluasi performa berbagai algoritma <i>Self-Supervised Learning</i> (termasuk SimSiam) secara khusus pada dataset citra medis berskala kecil ( <i>low-resource</i> ) sebagai pretraining domain-spesifik.	Pretraining SSL in-domain (seperti SimSiam) pada data medis terbatas terbukti mampu menghasilkan performa yang kompetitif dan bahkan mengungguli <i>transfer learning</i> konvensional dari dataset skala besar seperti ImageNet.
5.	Lu et al. (2022)	Contrastive Learning Meets Transfer Learning: A Case	Menggabungkan model arsitektur <i>Big Transfer</i> (BiT) dengan <i>Self-Supervised</i>	Penggunaan teknik <i>transfer learning</i> (BiT) terbukti mempercepat konvergensi

		Study in Medical Image Analysis.	<i>Learning</i> (SimSiam) untuk mempercepat waktu konvergensi model dalam mengekstraksi fitur citra medis.	algoritma SimSiam secara signifikan. Kombinasi kedua metode ini memberikan performa akurasi yang superior dibandingkan menggunakan salah satunya saja.
6.	Gui et al. (2024)	A Survey on Self-Supervised Learning: Algorithms, Applications, and Future Trends	Studi literatur ( <i>systematic review</i> ) yang menganalisis evolusi berbagai metode SSL ( <i>contrastive, generative</i> ), arsitektur algoritma visual, tren terbaru, serta aplikasinya secara mendalam dalam <i>computer vision</i> .	Pendekatan SSL telah berevolusi menjadi solusi utama untuk mengekstraksi fitur diskriminatif dari data tak berlabel. Performanya saat ini mampu menyamai atau melampaui metode <i>supervised</i> , menjadikannya ideal bagi domain dengan anotasi berbiaya tinggi seperti bidang medis..
7.	Zhou et al. (2022)	Models Genesis: Generic Autodidactic Models for 3D Medical Image Analysis	<i>Self-supervised learning</i> dengan pretext tasks berbasis transformasi citra untuk <i>pretraining</i> model pada citra	<i>Pretraining</i> self-supervised pada data medis mampu meningkatkan generalisasi model pada berbagai tugas analisis citra medis,

			medis 3D sebelum digunakan pada tugas analisis citra.	namun belum difokuskan pada klasifikasi citra kulit dua kelas.
8.	Zhang et al. (2022)	How Does SimSiam Avoid Collapse Without Negative Samples?	Analisis teoretis dan eksperimen pada metode <i>Self-supervised learning</i> SimSiam untuk memahami mekanisme pencegahan representation collapse.	Studi ini menjelaskan mekanisme teoretis SimSiam dalam mencegah representasi collapse, namun implementasinya pada domain medis masih terbatas.

Berdasarkan Tabel 2.1, dapat disimpulkan bahwa penelitian terkini menunjukkan pergeseran kuat menuju pemanfaatan *Self-Supervised Learning* (SSL) dalam analisis citra medis, khususnya untuk mengatasi skenario ekstrem keterbatasan data berlabel. Metode SSL yang efisien seperti SimSiam terbukti unggul karena kemampuannya mencapai performa tinggi pada fase *fine-tuning* meski dengan jumlah sampel yang sangat minim. Meskipun SSL telah terbukti tangguh dalam mengekstraksi pola visual yang kompleks pada citra dermatologi, implementasi spesifik arsitektur SimSiam secara *end-to-end* untuk membedakan klasifikasi biner kulit sehat dan tidak sehat menggunakan citra non-dermoskopik masih memiliki celah penelitian yang luas untuk dieksplorasi. Penelitian ini berupaya mengisi celah tersebut dengan menerapkan SimSiam sebagai tahap *pretraining* sebelum dilakukan *fine-tuning* pada dataset citra kulit.

## 2.2 Landasan Teori

### 2.2.1 Kulit dan Dermatologi Visual

Kulit merupakan organ terbesar pada anatomi manusia yang berfungsi sebagai pelindung utama terhadap ancaman lingkungan eksternal, regulator suhu tubuh, dan penahan hilangnya cairan esensial. Kondisi patologis pada kulit umumnya termanifestasi secara visual melalui perubahan warna (eritema), tekstur (skuama), maupun morfologi permukaannya. Sejumlah gangguan kulit memiliki karakteristik klinis yang sangat spesifik dan dapat diidentifikasi secara kasat mata. Analisis dermatologi modern secara tegas menyatakan bahwa perubahan struktur fisik, seperti pola kerontokan rambut pada *alopecia areata* atau lesi inflamasi, merupakan indikator primer dari gangguan autoimun atau patologi biologis lainnya (Bahashwan et al., 2024). Dalam konteks komputasi medis, ekstraksi fitur visual dari lesi kulit ini menjadi parameter utama untuk membedakan kondisi kulit sehat dan tidak sehat menggunakan algoritma kecerdasan buatan dan *machine learning*.

Sebagai representasi klinis, *alopecia areata* diklasifikasikan sebagai gangguan autoimun yang bermanifestasi pada kerontokan rambut tanpa pembentukan jaringan parut (*non-scarring hair loss*). Patogenesis dari kondisi ini dipicu oleh sistem kekebalan tubuh yang mengalami malfungsi dan secara tidak sengaja menyerang folikel penghasil rambut pada kulit. Oleh karena manifestasinya yang sangat terlihat, diagnosis terhadap *alopecia areata* dan kelainan kulit serupa sering kali dapat ditegakkan secara langsung melalui pengamatan gejala klinis visual, meskipun evaluasi lanjutan menggunakan dermoskopi dan histologi juga dapat memberikan bantuan diagnostik tambahan. Karakteristik visual klinis yang eksplisit inilah yang mendasari keandalan metode ekstraksi citra digital dalam menangkap dan mengklasifikasikan pola penyakit kulit secara komputasional (Bahashwan et al., 2024).

### 2.2.2 Pengolahan Citra Digital (*Digital Image Processing*)

Pengolahan citra digital didefinisikan sebagai disiplin ilmu yang memfokuskan pada manipulasi matematis dan analisis terhadap citra yang direpresentasikan dalam bentuk matriks piksel diskrit melalui teknik komputasi.

Tujuan utama dari proses ini adalah untuk meningkatkan kualitas visual citra guna interpretasi manusia atau untuk mengekstraksi parameter numerik tertentu yang dapat diproses oleh sistem mesin. Tahapan fundamental dalam alur pengolahan citra mencakup *preprocessing* esensial, seperti standarisasi dimensi (*resizing*) untuk memastikan keseragaman *input* pada arsitektur jaringan saraf, normalisasi distribusi nilai piksel guna mempercepat konvergensi pelatihan, serta isolasi objek melalui teknik penghapusan latar belakang (*background removal*) untuk mereduksi *noise* visual yang tidak relevan.

Dalam paradigma *deep learning* modern, khususnya pada arsitektur berbasis *Self-Supervised Learning* (SSL), rekayasa pra-pemrosesan telah berkembang melampaui standarisasi format menjadi augmentasi data stokastik. Transformasi visual seperti *random cropping*, *color jittering*, dan *Gaussian blur* tidak lagi sekadar metode untuk memperbanyak volume dataset secara artifisial, melainkan telah bertransformasi menjadi komponen analitis yang esensial untuk membangun "invariansi representasi". Melalui penerapan augmentasi ganda (*dual augmentation*) pada satu citra yang sama, model dipaksa untuk mengenali karakteristik semantik yang tetap (invariant) terlepas dari adanya distorsi geometris atau gangguan warna. Proses ini memungkinkan jaringan saraf untuk mempelajari fitur-fitur tingkat tinggi yang bermakna secara mandiri, yang sangat krusial dalam keberhasilan identifikasi patologi pada citra medis kulit (Chen et al., 2021).

### 2.2.3 Machine Learning

*Machine Learning* (Pembelajaran Mesin) didefinisikan sebagai pilar fundamental dari kecerdasan buatan yang memberdayakan sistem komputasi untuk mengekstraksi pola secara mandiri dari data tanpa memerlukan instruksi pemrograman eksplisit untuk setiap tugas spesifik. Paradigma ini bekerja dengan mengonstruksi model analitis matematis yang belajar dari sekumpulan data latih (*training data*) guna menghasilkan prediksi atau keputusan yang akurat pada data baru yang belum pernah dilihat sebelumnya (Zhao et al., 2024). Kemajuan dalam kapasitas komputasi dan ketersediaan data masif telah mentransformasi machine learning menjadi instrumen krusial dalam berbagai

disiplin ilmu, khususnya dalam pemrosesan citra medis tingkat lanjut (Manakitsa et al., 2024).

Berdasarkan taksonomi cara model berinteraksi dengan data selama proses pembelajaran, machine learning diklasifikasikan ke dalam tiga kategori utama:

1. **Supervised Learning (Pembelajaran Terarah):** Merupakan metode pembelajaran yang sangat bergantung pada ketersediaan dataset yang telah dianotasi secara manual oleh pakar manusia (ground truth labels). Model dilatih untuk memetakan hubungan fungsi antara variabel masukan dan target keluaran yang spesifik. Meskipun memiliki tingkat akurasi yang tinggi, metode ini menghadapi tantangan besar berupa keterbatasan jumlah data berlabel dalam domain medis, di mana proses pelabelan memerlukan biaya tinggi dan waktu yang signifikan dari tenaga profesional (Tahir et al., 2023).
2. **Unsupervised Learning (Pembelajaran Tak Terarah):** Metode ini beroperasi pada data mentah yang tidak memiliki label atau target keluaran. Fokus utamanya adalah mengeksplorasi struktur intrinsik atau distribusi probabilitas yang tersembunyi di dalam data, seperti melalui teknik klusterisasi (clustering) atau reduksi dimensi. Pendekatan ini memungkinkan sistem untuk menemukan pola-pola anomali atau pengelompokan fitur tanpa adanya intervensi label dari manusia (Shafiq et al., 2022).
3. **Self-Supervised Learning (Pembelajaran Mandiri):** Merupakan evolusi kontemporer dari unsupervised learning yang secara khusus dirancang untuk mengatasi kelangkaan label pada dataset. Dalam metode ini, sistem secara otomatis menghasilkan "label semu" (pseudo-labels) dengan memanfaatkan informasi dari struktur data itu sendiri melalui tugas proksi (pretext tasks). Melalui teknik augmentasi dan pemrosesan asimetris, model mampu mempelajari representasi fitur visual yang kaya dan diskriminatif sebelum dipindahkan ke tugas klasifikasi utama (Gui et al., 2024).

#### 2.2.4 Deep Learning

Deep learning merupakan sub-disiplin machine learning yang mengeksplorasi jaringan saraf tiruan (*artificial neural networks*) dengan banyak lapisan tersembunyi untuk mempelajari representasi data secara hierarkis dan otomatis. Arsitektur ini mereplikasi cara kerja saraf biologis dalam memproses informasi, di mana setiap lapisan secara progresif mengekstraksi fitur dari tingkat rendah (seperti tepi dan tekstur) hingga fitur semantik tingkat tinggi yang kompleks (Manakitsa et al., 2024). Dalam domain computer vision, deep learning mendominasi berbagai tugas komputasi visual karena kemampuannya dalam melakukan ekstraksi fitur secara otomatis tanpa memerlukan perancangan fitur manual (*handcrafted features*), sehingga memungkinkan proses pembelajaran berlangsung secara end-to-end dari piksel mentah langsung menuju label prediksi (Jia et al., 2024).

Secara struktural, arsitektur deep learning dibangun di atas integrasi tiga komponen utama: lapisan masukan (*input layer*), serangkaian lapisan tersembunyi (*hidden layers*), dan lapisan keluaran (*output layer*). Hidden layers tersebut dapat dikonfigurasi menjadi berbagai bentuk, seperti fully connected, convolutional, atau recurrent layers, yang disesuaikan dengan dimensi serta karakteristik spasial maupun temporal dari data yang diproses. Dalam ranah pengolahan citra medis, Convolutional Neural Network (CNN) menjadi arsitektur yang paling dominan digunakan karena kapabilitas intrinsiknya dalam menangkap dependensi spasial melalui operasi konvolusi matematis yang efisien (Tahir et al., 2023).

Meskipun memiliki keunggulan analitis yang tinggi dalam menangani data kompleks, pelatihan model deep learning konvensional sering kali menghadapi tantangan berupa keterbatasan data berlabel dalam volume masif serta kebutuhan sumber daya komputasi yang intensif. Namun, perkembangan inovasi teknik seperti transfer learning dan *self-supervised learning* (SSL) telah terbukti krusial dalam memitigasi permasalahan kelangkaan anotasi pada dataset medis. Dengan memanfaatkan representasi fitur yang dipelajari secara mandiri, metode deep learning menjadi semakin relevan dan presisi untuk

diterapkan pada bidang esensial seperti klasifikasi penyakit kulit, di mana data berlabel sering kali sulit didapatkan (Zhao et al., 2024).

### 2.2.5 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) adalah arsitektur deep learning terkemuka yang direkayasa secara spesifik untuk memproses data dengan topologi berstruktur grid spasial, seperti citra digital dua dimensi. CNN mendayagunakan operasi konvolusi matematis untuk mengekstraksi representasi fitur lokal dari citra secara progresif dan hierarkis (Shafiq & Gu, 2022). Melalui mekanisme komputasi ini, CNN mampu mempelajari pola spasial secara otonom, mulai dari fitur visual tingkat rendah (low-level features) seperti tepi dan gradien warna, hingga abstraksi semantik tingkat tinggi seperti tekstur dan struktur anatomi objek (Jia et al., 2024).

Secara arsitektural, rancangan CNN dibangun melalui integrasi beberapa komponen lapisan utama yang beroperasi secara sekuensial, yaitu:

1. Convolution Layer (Lapisan Konvolusi): Lapisan ini beroperasi sebagai pengekstraksi fitur primer dengan menggeser matriks filter yang dapat dipelajari (learnable kernels) ke seluruh area citra masukan. Operasi matematis konvolusi ini mendeteksi keberadaan pola visual spesifik dan menghasilkan peta fitur (feature maps) yang merepresentasikan respons spasial jaringan terhadap masukan tersebut (Rybacki et al., 2024).
2. Activation Function (ReLU): Setelah operasi konvolusi, fungsi aktivasi non-linear diterapkan pada peta fitur, dengan Rectified Linear Unit (ReLU) sebagai standar yang paling umum. ReLU mengeliminasi nilai piksel negatif dengan memetakannya menjadi nol, yang secara krusial memperkenalkan sifat non-linearitas ke dalam model. Hal ini memungkinkan jaringan untuk mempelajari batas keputusan yang sangat kompleks sekaligus memitigasi masalah kelaparan gradien (vanishing gradient problem) selama fase pelatihan (Shafiq & Gu, 2022).
3. Pooling Layer (Lapisan Pemusatan): Lapisan ini berfungsi untuk melakukan reduksi dimensi spasial (subsampling) pada feature maps sekaligus mempertahankan informasi semantik yang paling dominan. Operasi pooling (seperti Max Pooling) secara signifikan meningkatkan efisiensi komputasi,

mereduksi parameter untuk mencegah overfitting, serta memberikan sifat invariansi translasi, di mana model tetap mampu mengenali objek meskipun posisinya sedikit bergeser pada citra (Tan et al., 2024).

4. Fully Connected Layer (FCL): Berada pada terminal akhir arsitektur jaringan, lapisan ini meratakan (flatten) peta fitur multidimensional menjadi vektor satu dimensi. FCL bertindak sebagai pengklasifikasi akhir yang memetakan fitur-fitur yang telah diabstraksi oleh lapisan sebelumnya ke dalam nilai probabilitas probabilitas klasifikasi antar kelas (Jia et al., 2024).

Kapasitas ekstraksi fitur hierarkis ini menjadikan CNN sangat efektif dan dominan dalam berbagai tugas computer vision. Khususnya pada domain analisis medis dermatologis, CNN menawarkan tingkat ketelitian dan presisi yang sangat tinggi dalam mengekstraksi fitur visual patologis yang subtil, menjadikannya standar emas dalam diagnosis berbantuan komputer (Tahir et al., 2023; Zhao et al., 2024).

#### **2.2.6 Residual Network (ResNet)**

Residual Network (ResNet) merupakan inovasi arsitektur deep learning berbasis Convolutional Neural Network (CNN) yang dirancang secara spesifik untuk memitigasi fenomena vanishing gradient dan degradasi akurasi yang umum terjadi pada jaringan saraf dengan kedalaman yang ekstrem (Shafiq & Gu, 2022). Inti dari keunggulan ResNet terletak pada introduksi mekanisme residual connection atau skip connection, yaitu jalur pintas matematis yang memungkinkan aliran informasi melewati satu atau lebih lapisan secara langsung melalui pemetaan identitas (identity mapping). Pendekatan ini memastikan bahwa gradien informasi dapat mengalir lebih efisien selama fase backpropagation, sehingga memungkinkan pelatihan jaringan yang sangat dalam tanpa mengorbankan stabilitas atau meningkatkan galat pelatihan (De et al., 2024; Shafiq & Gu, 2022).

Arsitektur ResNet diklasifikasikan ke dalam beberapa varian utama berdasarkan kompleksitas dan jumlah lapisan yang digunakan:

1. ResNet-18 dan ResNet-34: Varian ini mengadopsi struktur Basic Block, di mana setiap blok residual terdiri dari dua lapisan konvolusi 3x3 sekuensial. Varian ResNet-18 sering kali dipilih sebagai tulang punggung (backbone)

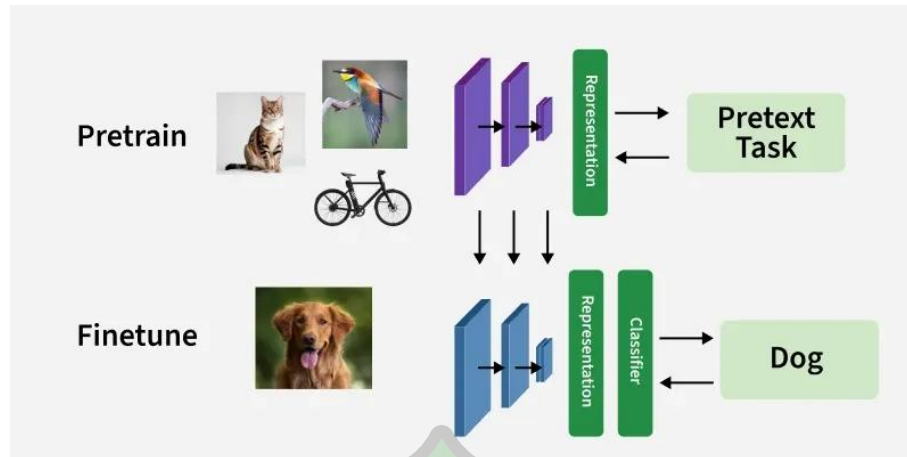
dalam tugas ekstraksi fitur citra medis karena menawarkan efisiensi komputasi yang tinggi tanpa mengurangi kualitas representasi fitur laten yang dipelajari (De et al., 2024; Tan et al., 2024).

2. ResNet-50, ResNet-101, dan ResNet-152: Varian ini menggunakan struktur Bottleneck Block untuk menjaga efisiensi parameter pada jaringan yang lebih dalam. Struktur ini menerapkan konfigurasi tiga lapisan konvolusi (1x1, 3x3, dan 1x1), di mana lapisan 1x1 bertugas melakukan reduksi dan ekspansi dimensi fitur, sementara lapisan 3x3 memproses fitur pada ruang dimensi yang lebih rendah untuk mengoptimalkan beban komputasi (Shafiq & Gu, 2022).

Karena kapabilitasnya dalam menghasilkan abstraksi fitur visual yang kaya dan tangguh, ResNet menjadi standar emas sebagai komponen pengekstraksi fitur dalam berbagai paradigma pembelajaran, termasuk self-supervised learning (Zhao et al., 2024). Penggunaan ResNet-18 sebagai encoder utama dalam penelitian ini memastikan bahwa sistem mampu menangkap karakteristik morfologi kulit yang subtil secara diskriminatif meskipun dilatih dengan keterbatasan data berlabel (De et al., 2024; Zhao et al., 2024).

### **2.2.7 Self-supervised learning (SSL)**

*Self-Supervised Learning (SSL)* merupakan cara belajar pada machine learning di mana model bisa memahami dan menemukan pola penting dari data tanpa perlu diberi label secara manual oleh manusia. Menurut Jing & Tian (2021), SSL memanfaatkan struktur intrinsik dalam data dengan merancang pretext task, yaitu tugas buatan yang menghasilkan label secara otomatis dari data itu sendiri. Melalui proses ini, model belajar memahami pola, struktur, dan karakteristik visual yang relevan sebelum digunakan pada tugas utama (downstream task) seperti klasifikasi.



**Gambar 2.2.1** Alur umum *Self-Supervised Learning*

Berdasarkan Gambar 2.2.1, alur umum *Self-Supervised Learning* terdiri dari dua tahap utama, yaitu:

1. *Pretraining*

Pada tahap ini, model dilatih menggunakan data tidak berlabel. Model menyelesaikan *pretext task* untuk mempelajari representasi fitur dari data. Seperti yang terlihat pada gambar, berbagai citra (misalnya kucing, burung, sepeda) digunakan sebagai data masukan untuk membangun representasi melalui jaringan saraf.

2. *Fine-tuning*

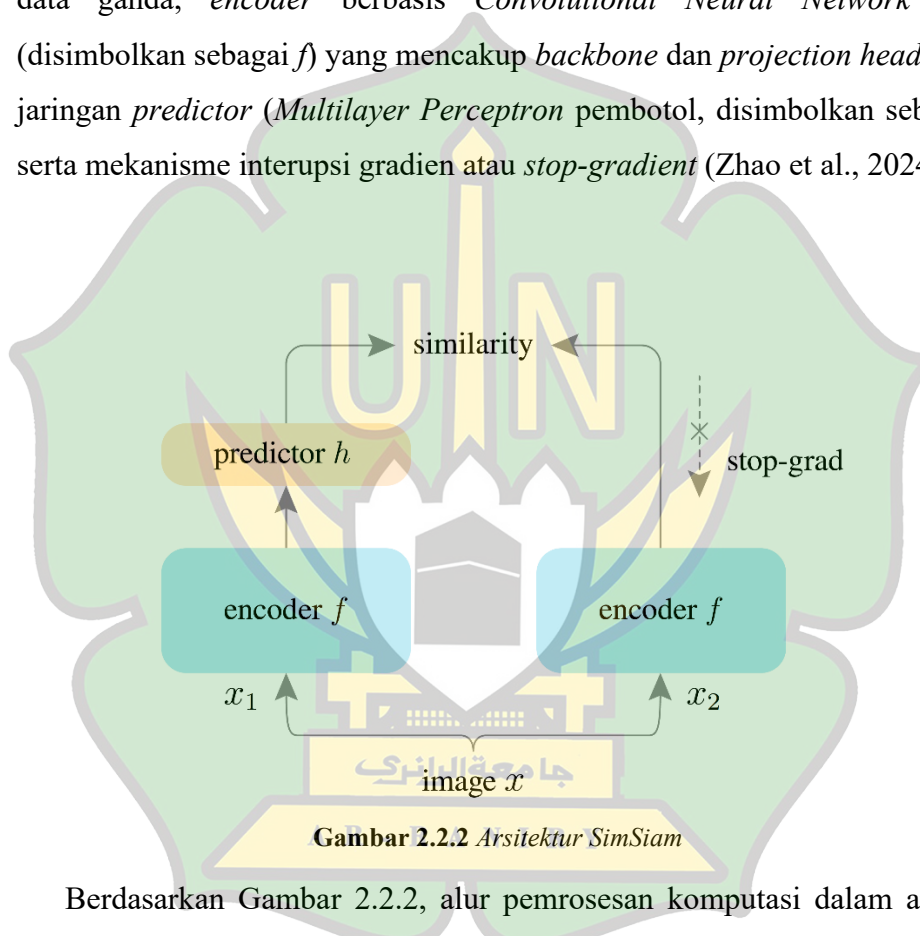
Setelah tahap pre-training selesai, model kemudian disesuaikan (*fine-tuned*) menggunakan data berlabel untuk tugas utama, seperti klasifikasi. Pada tahap ini, representasi yang telah dipelajari sebelumnya digunakan sebagai dasar untuk meningkatkan akurasi model. Sebagaimana ditunjukkan pada gambar, hasil akhir dari proses ini adalah prediksi kelas, misalnya “Dog” pada tugas klasifikasi

**2.2.8 SimSiam (*Simple Siamese Network*)**

SimSiam (*Simple Siamese Network*) merupakan arsitektur *Self-Supervised Learning* (SSL) berbasis pendekatan non-kontrasif (*non-contrastive learning*). Berbeda dengan pendahulunya seperti SimCLR atau MoCo yang sangat bergantung pada sampel negatif (*negative pairs*), antrean memori (*memory banks*), atau pembaruan momentum (*momentum encoders*) untuk mencegah kegagalan model, SimSiam beroperasi secara mandiri tanpa elemen-elemen

komputasi berat tersebut. Berdasarkan tinjauan literatur komputasi modern, desain minimalis SimSiam tetap mampu menghasilkan representasi fitur ruang laten yang sangat kuat, diskriminatif, dan stabil, menjadikannya sangat efisien untuk diterapkan pada dataset dengan keterbatasan sumber daya seperti citra medis (Chattopadhyay et al., 2023; Gui et al., 2024).

Arsitektur SimSiam dibangun di atas skema jaringan kembar (*Siamese Network*) yang asimetris. Komponen utamanya mencakup modul augmentasi data ganda, *encoder* berbasis *Convolutional Neural Network* (CNN) (disimbolkan sebagai  $f$ ) yang mencakup *backbone* dan *projection head*, sebuah jaringan *predictor* (*Multilayer Perceptron* pembobot, disimbolkan sebagai  $h$ ), serta mekanisme interupsi gradien atau *stop-gradient* (Zhao et al., 2024).



Berdasarkan Gambar 2.2.2, alur pemrosesan komputasi dalam arsitektur SimSiam dapat dijabarkan secara terstruktur sebagai berikut:

1. **Input Augmentasi (*Dual Augmentation*):** Sebuah citra tunggal  $x$  diproses menggunakan dua jenis transformasi stokastik yang berbeda (seperti *random crop*, *color jittering*, atau *blur*), sehingga menghasilkan dua sudut pandang visual (*views*) yang berbeda, direpresentasikan sebagai  $x_1$  dan  $x_2$ .
2. **Ekstraksi Fitur (*Encoder Sharing Weights*):** Kedua citra hasil augmentasi diproses secara paralel oleh *encoder*  $f$  yang berbagi bobot matriks yang sama (*weight sharing*). Proses ini mengekstraksi representasi laten berdimensi

tinggi dari masing-masing citra, sehingga menghasilkan vektor  $z_1 = f(x_1)$  dan  $z_2 = f(x_2)$ .

3. **Prediksi Asimetris (*Predictor*):** Untuk menciptakan asimetri pada jaringan, representasi dari salah satu cabang (misalnya  $z_1$ ) diteruskan ke dalam jaringan tambahan berupa *predictor*  $h$ . Jaringan ini bertugas melakukan transformasi ruang fitur untuk memprediksi representasi langsung dari cabang pasangannya, menghasilkan output  $p_1 = h(z_1)$  (Chattopadhyay et al., 2023).
4. **Perhitungan Kemiripan (*Cosine Similarity*):** Model mengevaluasi kerugian (*loss*) dengan menghitung tingkat kemiripan kosinus (*cosine similarity*) negatif antara hasil prediksi di satu cabang ( $p_1$ ) dengan representasi asli dari *encoder* di cabang lainnya ( $y_2$ ) (Gui et al., 2024).
5. **Mekanisme *Stop-Gradient*:** Ini adalah komponen paling fundamental dalam SimSiam. Pada cabang target (cabang  $z_2$  yang tidak melewati *predictor*), diterapkan operasi *stop-gradient*. Analisis teoretis membuktikan bahwa mekanisme ini bertindak menyerupai algoritma *Expectation-Maximization* (EM), yang secara krusial memblokir aliran balik gradien (*backpropagation*) untuk memperbarui bobot pada jalur tersebut (Zhang et al., 2022). Intervensi matematis ini secara efektif mencegah jaringan jatuh pada solusi trivial di mana model memprediksi vektor konstan untuk semua masukan (*representation collapse*) sehingga memaksa *predictor* untuk benar-benar mempelajari fitur semantik yang bermakna (Zhang et al., 2022; Zhao et al., 2024).

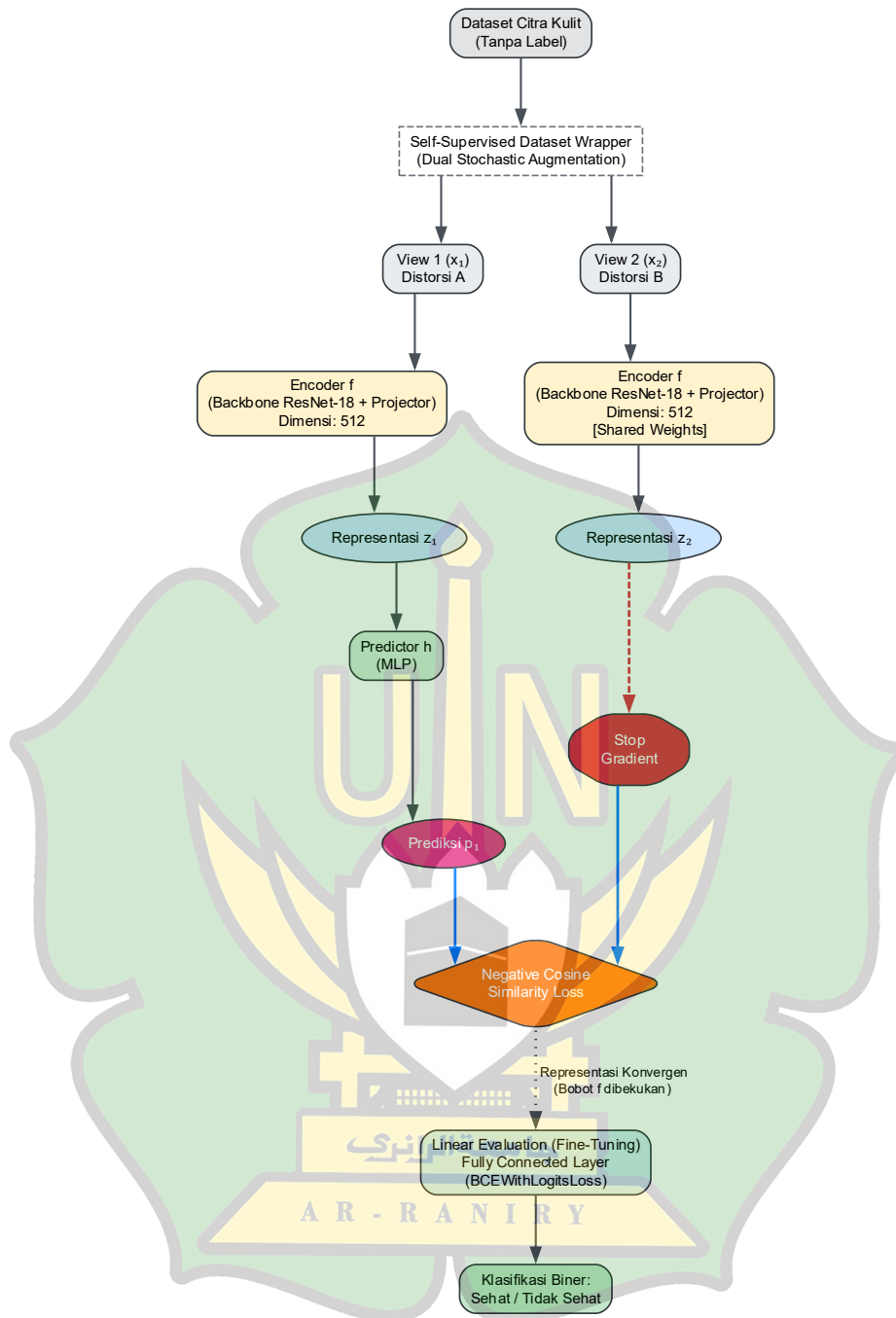
## 2.2.9 Kerangka Kerja Pipa Komputasi Arsitektur Usulan

Kerangka kerja (*framework*) pada penelitian ini dirancang sebagai sebuah pipa komputasi (*computational pipeline*) *end-to-end* yang mengintegrasikan teknik pemrosesan citra digital dengan paradigma *Self-Supervised Learning* (SSL). Arsitektur usulan ini dibangun untuk memecahkan masalah klasifikasi citra kulit (sehat dan berpenyakit) dengan meminimalisasi ketergantungan pada anotasi label manual. Pipa komputasi ini secara terstruktur dibagi ke dalam tiga fase operasional utama: pra-pemrosesan data, pra-pelatihan mandiri (*self-supervised pre-training*), dan evaluasi linear (*supervised fine-tuning*).

Fase pertama melibatkan modul *Self-Supervised Dataset Wrapper*, di mana citra mentah yang telah melalui proses isolasi latar belakang (*background removal*) dikenakan augmentasi stokastik ganda. Fase ini bertanggung jawab untuk menghasilkan dua pandangan visual yang terdistorsi secara independen ( $x_1$  dan  $x_2$ ) dari citra sumber yang sama.

Fase kedua merupakan inti dari arsitektur usulan, di mana metode SimSiam diimplementasikan menggunakan *backbone encoder* berbasis ResNet-18. ResNet-18 dipilih secara strategis karena menawarkan keseimbangan optimal antara efisiensi komputasi memori dan kedalaman arsitektur yang memadai untuk mengekstraksi representasi semantik. Pada tahap *pre-training* ini, matriks citra diproses oleh lapisan konvolusi ResNet-18 untuk menghasilkan peta fitur, yang kemudian diproyeksikan ke dalam ruang laten berdimensi 512. Melalui interaksi asimetris antara jaringan *predictor* dan mekanisme *stop-gradient*, model memaksimalkan fungsi *Negative Cosine Similarity*. Proses ini memaksa *backbone* ResNet-18 untuk secara mandiri mengenali pola patologi kulit yang invarian terhadap gangguan warna maupun orientasi, sepenuhnya tanpa menggunakan label kelas.

Fase ketiga adalah klasifikasi biner. Setelah bobot representasi konvergen pada fase sebelumnya, arsitektur diubah. Bobot ekstraksi fitur pada ResNet-18 dibekukan (*frozen*), dan komponen *predictor* digantikan oleh sebuah *Fully Connected Layer* (FCL) tunggal. Jaringan kemudian dilatih secara *supervised* menggunakan fungsi aktivasi *sigmoid* dan optimasi *Binary Cross Entropy with Logits Loss* untuk memetakan representasi fitur 512 dimensi menjadi keputusan klasifikasi akhir: kulit sehat atau kulit tidak sehat.



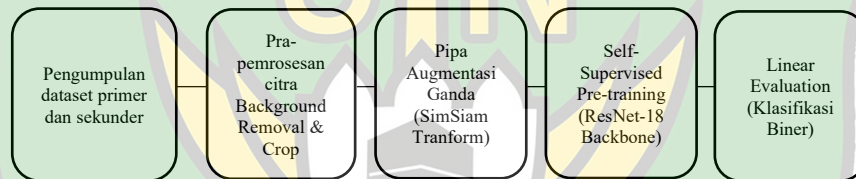
**Gambar 2.2.3** Representasi Skematik Kerangka Kerja Pipa Komputasi SimSiam-ResNet18.

## BAB III METODE PENELITIAN

### 3.1 Jenis Penelitian

Penelitian ini merupakan penelitian eksperimen dengan pendekatan kuantitatif yang bertujuan untuk menguji efektivitas metode *Self-Supervised Learning* (SSL) menggunakan arsitektur SimSiam dalam klasifikasi citra kulit sehat dan tidak sehat. Pendekatan eksperimen digunakan karena penelitian ini melibatkan perancangan model, proses pelatihan, serta pengujian performa model menggunakan metrik evaluasi tertentu. Metode kuantitatif digunakan untuk mengukur kinerja model berdasarkan nilai numerik seperti akurasi, precision, recall, F1-score, dan spesifisitas. Hasil evaluasi tersebut digunakan untuk menganalisis kemampuan model dalam membedakan dua kategori citra kulit.

### 3.2 Tahapan Penelitian



**Gambar 3.2.1** Alur Penelitian

Alur penelitian memvisualisasikan tahapan-tahapan sistematis dan iteratif yang dilakukan dalam eksperimen ini. Berdasarkan rancangan tersebut, proses penelitian dijabarkan ke dalam langkah-langkah operasional sebagai berikut:

1. **Pengumpulan dan Pembagian Data:** Tahap awal penelitian melibatkan pengumpulan dataset primer dan sekunder dari citra kulit. Data ini kemudian dianotasi dan diklasifikasikan ke dalam dua kelas utama, yaitu kulit normal dan kulit berpenyakit (*diseased*). Selanjutnya, dataset dibagi (*data splitting*) menjadi tiga subset utama, yaitu data latih, data validasi, dan data uji, guna memastikan model dapat dilatih dan dievaluasi secara objektif tanpa mengalami *overfitting*.
2. **Pra-pemrosesan dan Augmentasi Data Ganda:** Citra mentah yang telah dikumpulkan tidak langsung dimasukkan ke dalam model, melainkan

melalui tahap pra-pemrosesan yang meliputi penyesuaian dimensi (*resizing*) dan normalisasi. Khusus untuk kebutuhan arsitektur SimSiam, diterapkan teknik augmentasi data ganda (*dual stochastic augmentation*). Setiap citra input akan diubah secara acak (melalui rotasi, pemotongan, *gaussian blur* atau penyesuaian warna) untuk menghasilkan dua sudut pandang (*views*) yang berbeda secara visual, yaitu  $x_1$  dan  $x_2$  namun tetap mempertahankan makna semantik aslinya.

3. **Pretraining Menggunakan Arsitektur SimSiam & Pelatihan Mandiri:** Dua *views* citra hasil augmentasi kemudian diproses ke dalam arsitektur SimSiam pada tahap *Self-Supervised pre-training*. Tahap ini berfokus pada ekstraksi fitur tanpa memerlukan label *ground truth*. Jaringan *backbone encoder* (berbasis ResNet-18) mengekstraksi representasi laten dari kedua citra tersebut, sementara jaringan prediktor bertugas memprediksi representasi dari satu sisi silang ke sisi lainnya untuk memaksimalkan keselarasan fitur yang dipelajari.
4. **Perhitungan Fungsi Loss SimSiam:** Selama tahap *pre-training*, pembaruan bobot model diatur oleh fungsi kerugian (*loss function*). Penelitian ini menggunakan *Negative Cosine Similarity* untuk mengukur seberapa dekat representasi yang dihasilkan oleh model. Untuk mencegah fenomena *mode collapse* di mana model memprediksi nilai konstan untuk seluruh masukan operasi *stop-gradient* secara krusial diterapkan pada salah satu cabang komputasi.
5. **Fine-Tuning Klasifikasi Biner:** Setelah model berhasil mempelajari representasi fitur visual yang kaya pada tahap *pre-training*, bobot pada *encoder* dibekukan (*frozen*). Lapisan klasifikasi *Fully Connected Layer* kemudian ditambahkan di akhir jaringan. Pada tahap *fine-tuning* ini, model dilatih ulang secara *supervised* menggunakan label aktual (normal dan berpenyakit) dengan fungsi optimasi *Binary Cross-Entropy* untuk keperluan klasifikasi biner.
6. **Evaluasi Model:** Model yang telah melewati tahap *fine-tuning* diuji menggunakan subset data uji yang belum pernah dilihat sebelumnya. Kinerja klasifikasi tidak hanya dinilai dari tingkat akurasi secara umum,

tetapi dievaluasi secara menyeluruh menggunakan *Confusion Matrix*. Hal ini bertujuan untuk mendapatkan nilai parameter esensial lainnya seperti Presisi, Sensitivitas (*Recall*), dan Spesifisitas, khususnya untuk memastikan keandalan model dalam studi kasus medis.

7. **Analisis Hasil:** Tahap terakhir adalah melakukan tabulasi dan analisis komprehensif terhadap seluruh metrik evaluasi yang diperoleh. Hasil analisis ini digunakan untuk menarik kesimpulan mengenai efektivitas algoritma SimSiam dalam meningkatkan performa klasifikasi penyakit kulit, serta menjawab rumusan masalah yang telah ditetapkan pada awal penelitian.

Data yang digunakan dalam penelitian ini dikumpulkan dalam format citra digital dua dimensi yang berasal dari dua sumber utama, yaitu data sekunder dan data primer. Strategi penggabungan ini bertujuan untuk meningkatkan variabilitas fitur visual model serta menguji keandalan algoritma pada data yang diambil langsung di lapangan (*real-world setting*).

1. **Data Primer:** Peneliti mengumpulkan 200 citra orisinal secara mandiri (100 citra kulit sehat dan 100 citra kulit tidak sehat). Proses akuisisi dilakukan menggunakan kamera ponsel pintar (*high-end smartphone camera*) dengan pengaturan teknis yang dirancang untuk meniru karakteristik citra dermoskopi (*dermoscopic-like setup*). Pengaturan ini melibatkan kontrol jarak fokus yang konsisten (5–10 cm), penggunaan pencahayaan terpolarisasi buatan untuk meminimalkan pantulan cahaya pada permukaan kulit (*specular reflection*), serta optimasi makro untuk menangkap detail tekstur lesi, pigmentasi, dan struktur vaskular yang biasanya hanya terlihat jelas melalui perangkat dermoskop profesional.
2. **Data Sekunder:** Diperoleh dari repositori publik Kaggle melalui dataset "Skin Disease and Normal Skin Dataset" oleh Lysa Apriani (dapat diakses pada: [<https://www.kaggle.com/datasets/lysaapriani/skin-disease-and-normal-skin-dataset/data>]) serta referensi citra medis dari repositori Dermnet (dapat diakses pada: [<https://dermnetnz.org/topics>])). Dari sumber ini, diambil sampel sebanyak 600 citra kulit non-dermoskopik untuk memperkuat basis data pelatihan.

Penentuan kondisi citra antara kulit sehat (normal) dan tidak sehat (berpenyakit) didasarkan pada analisis morfologi klinis. Kulit sehat dikategorikan berdasarkan integritas lapisan epidermis yang utuh, warna pigmen yang homogen, serta tidak adanya tanda-tanda inflamasi atau gangguan vaskular (Tan et al, 2024).

Sebaliknya, klasifikasi kulit tidak sehat diidentifikasi melalui keberadaan lesi patologis yang ditandai dengan adanya eritema (kemerahan akibat dilatasi kapiler), skuama (penumpukan stratum korneum), papula, atau perubahan warna abnormal yang menunjukkan gangguan pada melanosit atau sel skuamosa (Smith & Johnson, 2023).

Penggunaan fitur visual diskriminatif seperti pola retikular dan distribusi pigmen asimetris menjadi parameter utama dalam membedakan lesi kulit yang memerlukan perhatian medis dibandingkan dengan variasi kulit normal (Mevorach et al, 2025).

Pemilihan sampel dilakukan menggunakan metode *purposive sampling* untuk memastikan keseimbangan proporsi antar kelas (*class balance*). Dengan penambahan data primer, total dataset yang digunakan adalah 800 citra, yang terbagi secara merata menjadi 400 citra Kelas 0 (Kulit Sehat) dan 400 citra Kelas 1 (Kulit Tidak Sehat).

#### **A. Augmentasi Dataset**

Mengingat arsitektur *Deep Learning* membutuhkan volume data yang besar, data latih akan melalui tahap augmentasi melalui modifikasi stokastik seperti *random cropping*, rotasi, *gaussian blur* dan *color jittering*. Dalam arsitektur SimSiam, augmentasi ganda bersifat esensial pada tahap *pretraining*. Proses ini akan mengekspansi 560 citra latih awal menjadi total 2.240 citra hasil augmentasi untuk kebutuhan pembelajaran representasi mandiri.

#### **B. Pembagian Data (Data Splitting)**

Untuk memastikan validitas pengujian dan memitigasi *overfitting*, skenario pembagian data dilakukan secara *stratified random split* pada 800 citra mentah sebelum proses augmentasi komputasional. Rasio distribusi yang digunakan adalah 70:15:15:

- a) 70% Data Latih (*Training Set*): Terdiri dari 560 citra mentah. Subset ini digunakan untuk memperbarui bobot model pada tahap *pretraining* (dieksplorasi hingga 2.240 citra augmentasi) dan tahap *fine-tuning*.
- b) 15% Data Validasi (*Validation Set*): Terdiri dari 120 citra mentah. Digunakan untuk memantau konvergensi pelatihan dan penyetelan *hyperparameter* di setiap *epoch*.
- c) 15% Data Uji (*Testing Set*): Terdiri dari 120 citra mentah. Subset ini diisolasi sepenuhnya dan hanya digunakan pada tahap pengujian akhir sebagai data yang belum pernah dilihat (*unseen data*) oleh model.

### 3.2.2 Preprocessing dan Augmentasi Data

Sebelum didistribusikan ke dalam arsitektur model, citra mentah memerlukan standarisasi format melalui proses pra-pemrosesan:

1. Penyesuaian Dimensi (*Resizing*): Mengubah skala matriks citra menjadi 224 x 224 piksel agar sesuai dengan standar input *backbone* ResNet.
2. Normalisasi: Mengonversi intensitas piksel citra menjadi rentang [0, 1] dan mengubah format larik menjadi tensor komputasi.
3. Augmentasi Simetris: Pada tahap *Self-Supervised Learning*, algoritma SimSiam membutuhkan dua representasi visual (*views*) berbeda dari satu citra asli  $x$ . Dua citra hasil augmentasi ( $x_1$  dan  $x_2$ ) di bentuk secara stokastik melalui operasi:
  - a. Random Resized Crop dan Horizontal Flip
  - b. *Color Jittering* dengan menjaga stabilitas nilai *hue* agar karakteristik pigmen merah pada lesi kulit tidak terdistorsi
  - c. Gaussian Blur
4. Augmentasi Standar (*Linear Evaluation*): Berbeda dengan tahap *Self-Supervised Learning* yang menggunakan distorsi ekstrem, pada tahap evaluasi klasifikasi terawasi (*supervised*), citra hanya diberikan transformasi ringan berupa rotasi acak maksimal 15 derajat dan *horizontal flip*. Hal ini bertujuan untuk mencegah *overfitting* sekaligus menjaga karakteristik asli dan letak klinis dari lesi kulit.

### 3.2.3 Arsitektur Pemodelan *Self-supervised learning* (SimSiam)

Pemodelan utama dalam penelitian ini bertumpu pada arsitektur SimSiam yang mampu mempelajari struktur semantik citra tanpa anotasi label. Model ini memiliki tiga komponen komputasi:

1. Encoder (f): Mengadopsi arsitektur *Convolutional Neural Network* berbasis ResNet-18 untuk mengekstraksi fitur spasial tingkat tinggi.
2. Projection Head (g): Jaringan *Multilayer Perceptron* (MLP) yang mentransformasikan keluaran *encoder* ke dalam ruang dimensi laten.
3. Predictor (h): Jaringan MLP asimetris yang bertugas memprediksi vektor representasi dari satu cabang silang ke cabang lainnya.

Secara matematis, dua citra hasil augmentasi diproses secara paralel melalui bobot *encoder* yang sama (*weight sharing*), menghasilkan representasi laten:

$$z_1 = f(x_1), \quad z_2 = f(x_2)$$

Selanjutnya, *predictor* melakukan transformasi lanjutan:

$$p_1 = h(z_1), \quad p_2 = h(z_2)$$

Untuk mencegah model terjebak pada solusi konstan (*mode collapse*), diterapkan fungsi kerugian *Negative Cosine Similarity* dengan mekanisme interupsi gradien (*stop-gradient*). Loss dihitung berdasarkan:

$$D(p, z) = - \frac{p}{\|p\|_2} \cdot \frac{z}{\|z\|_2}$$

Total evaluasi loss untuk pembaruan bobot diakumulasikan sebagai berikut:

$$L_{\text{SimSiam}} = \frac{1}{2} D(p_1, \text{stopgrad}(z_2)) + \frac{1}{2} D(p_2, \text{stopgrad}(z_1))$$

### 3.2.4 Proses *Fine-Tuning* dan Klasifikasi Biner

Setelah ekstraksi fitur mandiri konvergen, tahap *transfer learning* dilakukan. Bobot matriks pada *encoder* dasar (f) dibekukan (*frozen*). Lapisan *predictor* dihapus dan digantikan oleh *Fully Connected Layer* (FCL) yang berfungsi sebagai pengklasifikasi.

Model dioptimasi menggunakan label *ground truth* untuk membedakan kelas "Sehat" dan "Tidak Sehat". Fungsi kerugian komputasi menggunakan *Binary Cross Entropy with Logits Loss* untuk keperluan klasifikasi biner:

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\sigma(\hat{y}_i)) + (1 - y_i) \log(1 - \sigma(\hat{y}_i))]$$

Dimana  $y_i$  merepresentasikan label kelas aktual dan  $\hat{y}_i$  merupakan probabilitas keluaran dari fungsi aktivasi *sigmoid*.

### 3.2.5 Tahap Evaluasi Peforma

Tahap akhir dari alur sistem adalah pengukuran kinerja klasifikasi. Hasil prediksi dari data uji dikomparasikan dengan label aktual untuk menyusun *Confusion Matrix* yang menghasilkan nilai:

1. *True Positive* (TP): Kulit tidak sehat diprediksi benar sebagai tidak sehat.
2. *True Negative* (TN): Kulit sehat diprediksi benar sebagai sehat.
3. *False Positive* (FP): Kulit sehat salah diprediksi sebagai tidak sehat.
4. *False Negative* (FN): Kulit tidak sehat salah diprediksi sebagai sehat.

Parameter evaluasi sistem dikalkulasikan menggunakan persamaan berikut:

1. Akurasi: Tingkat kesesuaian keseluruhan prediksi sistem.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Precision: Rasio prediksi benar positif dibandingkan keseluruhan hasil positif.

$$\text{Precision} = \frac{TP}{TP + FP}$$

3. Sensitivitas (*Recall*): Tingkat ketepatan model dalam memprediksi kelas positif.

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

4. F1-Score: Rata-rata harmonik antara *precision* dan *recall*.

$$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

5. Spesifisitas: Kemampuan model mengidentifikasi kulit normal dengan benar tanpa memunculkan *false alarm*.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

### 3.3 Waktu dan Penelitian

Proses pengerjaan skripsi diestimasikan berlangsung sekitar 4 bulan, terhitung dari bulan Januari sampai April. Rincian alokasi waktu untuk setiap tahapan penelitian dijabarkan pada Tabel 3.1.

Tabel 3.1 Jadwal Penelitian

No	Tahapan Kegiatan	Jan	Feb	Mar	Apr
1	Studi Literatur & Perumusan Masalah	■			
2	Penyusunan Draf Proposal Tugas Akhir	■	■		
3	Pengumpulan & Pra-pemrosesan Data (Kaggle)		■		
4	Seminar Proposal Tugas Akhir			■	
5	Pretraining Arsitektur SimSiam (Self-Supervised)			■	
6	Linear Evaluation & Klasifikasi Biner			■	■
7	Evaluasi Performa & Analisis Hasil				■
8	Penyusunan Laporan & Sidang Tugas Akhir				■

### 3.4 Alat dan Bahan

penelitian ini membutuhkan dukungan instrumen berupa perangkat keras (*hardware*) dan perangkat lunak (*software*). Rincian mengenai spesifikasi alat dan bahan yang digunakan untuk mendukung seluruh tahapan pelatihan dan implementasi model diuraikan sebagai berikut:

#### 3.4.1 Perangkat Keras (*Hardware*)

Keberhasilan implementasi metode *Self-Supervised Learning* (SSL) dengan arsitektur SimSiam sangat bergantung pada ketersediaan instrumen perangkat keras (*hardware*) yang memadai. Dalam penelitian ini, perangkat keras secara eksklusif difungsikan sebagai stasiun kerja (*workstation*) lokal untuk perancangan logika kode, pengunduhan dan pra-pemrosesan *dataset* sekunder, serta pengujian awal algoritma komputasi sebelum dieksekusi di platform *cloud*.

Pemilihan spesifikasi minimum ditetapkan untuk mencegah terjadinya kendala komputasi (*bottleneck*), terutama saat proses memuat matriks citra ke dalam ruang memori sementara. Adapun rincian komponen perangkat keras beserta spesifikasi dan fungsi operasionalnya dapat dilihat pada Tabel 3.2.

**Tabel 3.2** *Komponen Perangkat Keras*

Komponen	Spesifikasi	Fungsi
Komputer / Laptop	Prosesor minimal Intel Core i5 atau setara	Menjalankan sistem operasi dasar dan melakukan komputasi CPU selama tahap penulisan kode
Memori ( <i>RAM</i> )	Kapasitas minimal 8 GB	Menyediakan ruang memori sementara yang memadai saat memuat dataset dan menjalankan model ke dalam memori
Kamera Smartphone	Samsung Galaxy S21 Plus	Digunakan untuk membangun dataset primer penelitian

### 3.4.2 Perangkat Lunak (*Software*)

Dalam pelaksanaan eksperimen komputasi ini, lingkungan pengembangan perangkat lunak (*software development environment*) memegang peranan krusial. Perangkat lunak berfungsi untuk memastikan seluruh tahapan penelitian mulai dari pra-pemrosesan citra kulit, implementasi arsitektur *Self-supervised learning* (SimSiam), hingga tahap klasifikasi dan komputasi metrik evaluasi dapat berjalan secara efisien dan terintegrasi.

Pemilihan platform dan bahasa pemrograman disesuaikan dengan kebutuhan komputasi *deep learning* yang memerlukan skalabilitas tinggi, manajemen memori yang baik, serta dukungan *library* matematis yang komprehensif. Rincian perangkat lunak dan platform *cloud* yang digunakan dalam penelitian ini diuraikan pada Tabel 3.3.

**Tabel 3.3** *Komponen Perangkat Lunak*

Komponen	Versi	Fungsi
Sistem Operasi	Windows 11	Menjalankan sistem operasi dasar, mengelola manajemen file lokal, dan memfasilitasi komputasi CPU selama tahap penulisan serta pengujian skrip kode awal.
Kaggle	Publik (Cloud)	Berfungsi sebagai repositori sumber dataset (jika data sekunder digunakan) sekaligus menyediakan lingkungan notebook berbasis cloud opsional untuk eksplorasi data awal ( <i>Exploratory Data Analysis</i> ).
Dermnet		Sebagai sumber dataset utama penyakit kulit yang disertifikasi secara medis
Visual Studio Code (VSCode)	1.110.1	Berperan sebagai <i>Integrated Development Environment</i> (IDE) atau <i>code editor</i> utama untuk menulis, menstrukturkan, dan <i>men-debug</i> algoritma pemrograman secara terisolasi di perangkat lokal.
Google Colaboratory (Colab)	Publik (Cloud)	Berfungsi sebagai platform komputasi awan utama yang menyediakan akses ke <i>Graphics Processing Unit</i> (GPU) berkinerja tinggi secara virtual untuk mempercepat proses <i>pre-training</i> dan <i>fine-tuning</i> model SimSiam.
Python	3.12.12	Merupakan bahasa pemrograman berorientasi objek tingkat tinggi yang digunakan sebagai fondasi utama untuk menyusun seluruh logika algoritma <i>machine learning</i> , manipulasi tensor, dan visualisasi hasil evaluasi.

## **BAB IV**

### **HASIL DAN PEMBAHASAN**

#### **4.1 Lingkungan Implementasi dan Dataset**

##### **4.1.1 Lingkungan Implementasi**

Implementasi sistem pada penelitian ini dikembangkan menggunakan bahasa pemrograman Python versi 3.12.12 beserta framework deep learning PyTorch versi 2.9.0. Keseluruhan proses komputasi, mulai dari pra-pemrosesan data (data preprocessing), pra-pelatihan model (pre-training), hingga tahap evaluasi, dijalankan pada lingkungan komputasi awan (cloud computing) Kaggle. Untuk menangani beban komputasi jaringan saraf tiruan yang berat, proses ini diakselerasi memanfaatkan perangkat keras GPU NVIDIA Tesla T4.

##### **4.1.2 Sumber dan Pre-pemrosesan Dataset**

Data citra yang digunakan dalam penelitian ini bersumber dari gabungan data sekunder (repositori publik Dermnet dan Kaggle) serta data primer yang dikumpulkan secara mandiri. Dataset ini diorganisasikan ke dalam skema klasifikasi biner untuk membedakan antara kulit normal dan kulit berpenyakit.

Untuk mencegah model mengambil jalan pintas matematis (bias domain) dari perbedaan kualitas kamera atau pencahayaan latar belakang, seluruh gambar dalam dataset terlebih dahulu melalui proses segmentasi dan penghapusan latar belakang (background removal). Hal ini memastikan bahwa variasi visual yang tersisa hanyalah patologi pada kulit. Setelah itu, dataset dipisah secara acak menjadi tiga bagian dengan rasio pembagian: 70% untuk data latih (training), 15% untuk data validasi (validation), dan 15% untuk data uji (testing).

##### **4.1.3 Organisasi Data Loader Self-Supervised**

Oleh karena fase pra-pelatihan menggunakan arsitektur SimSiam yang berbasis Self-Supervised Learning (SSL), model harus dilatih murni dari data citra tanpa mengetahui label kelasnya. Untuk mengakomodasi kebutuhan

matematis ini, dirancang sebuah wrapper dataset khusus seperti yang diperlihatkan pada Gambar

```
class SimSiamDatasetWrapper(Dataset):
    def __init__(self, subset, transform):
        self.subset = subset
        self.transform = transform

    def __len__(self):
        return len(self.subset)

    def __getitem__(self, idx):
        img, _ = self.subset[idx]
        return self.transform(img), self.transform(img)
```

**Gambar 4.1.1** Kode Loader Dataset untuk Pre-training SimSiam

Pada implementasi fungsi `__getitem__` di atas, label kelas dari dataset secara eksplisit dibuang atau tidak diekstrak (disimbolkan dengan variabel underscore `_`). Sebagai gantinya, fungsi tersebut mengembalikan dua buah tampilan (view) dari satu gambar yang sama setelah melewati pipa augmentasi (`self.transform`). Mekanisme ini adalah fondasi utama dari pembelajaran self-supervised pada SimSiam, di mana model dipaksa untuk mempelajari representasi fitur kulit yang bermakna tanpa bantuan panduan label dari manusia.

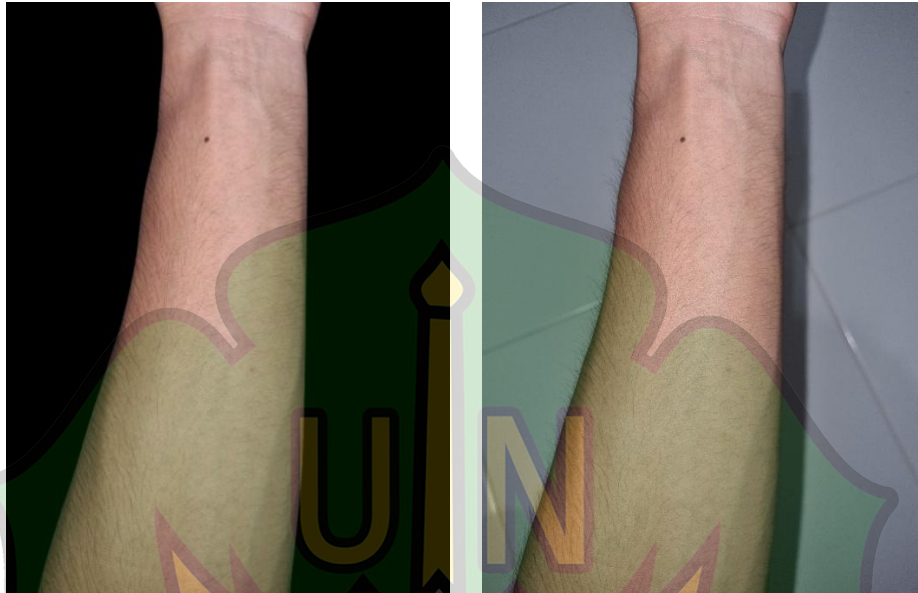
## 4.2 Pra-pemrosesan dan Augmentasi Data

### 4.2.1 Pra-pemrosesan Citra (Image Preprocessing)

Sebelum proses komputasi pada framework deep learning dilakukan, seluruh data citra mentah terlebih dahulu melalui tahap pra-pemrosesan untuk meminimalkan bias domain. Proses ini melibatkan isolasi area kulit melalui teknik penghapusan latar belakang (background removal), di mana latar belakang citra asli diekstraksi dan diganti dengan warna hitam solid (RGB: 0, 0, 0).

Langkah normalisasi ini sangat krusial untuk mencegah model melakukan shortcut learning (kecurangan komputasional), di mana jaringan saraf tiruan cenderung mempelajari perbedaan bayangan latar belakang atau kualitas kamera alih-alih mempelajari tekstur penyakit. Penggunaan latar belakang

hitam menyamarkan perbedaan antara data primer (kamera ponsel) dan data sekunder (kamera medis), menjadikannya sejalan dengan standar citra medis klinis seperti pada repositori Dermnet. Dengan tereliminasi noise visual di luar area kulit utama, model dipaksa untuk berfokus sepenuhnya pada ekstraksi fitur patologi kulit.



**Gambar 4.1.2** *Data sesudah dan sebelum di normalisasi*

#### **4.2.2 Pipa Augmentasi Data (Data Augmentation Pipeline)**

Setelah latar belakang dinormalisasi, tahap selanjutnya adalah menerapkan augmentasi data ganda (dual augmentation) yang menjadi fondasi utama dalam arsitektur Self-Supervised Learning (SimSiam). Pipa augmentasi yang dirancang pada tahap pra-pelatihan (pre-training) ini dikalibrasi secara spesifik untuk menyetarakan karakteristik kamera tanpa merusak sinyal warna dari lesi kulit. Operasi transformasi yang diimplementasikan meliputi:

1. Random Resized Crop (skala 0.5 - 0.9): Melakukan pemotongan acak untuk memberikan variasi spasial yang cukup, dengan batas bawah 50% untuk memastikan area lesi/penyakit tidak terpotong sepenuhnya dari frame.
2. Random Horizontal & Vertical Flip: Memberikan variasi orientasi posisi kulit.
3. Gaussian Blur (probabilitas 50%): Bertindak sebagai equalizer untuk secara acak menyamarkan ketajaman (resolusi tinggi) dari citra kamera medis,

sehingga menjembatani kesenjangan visual dengan citra dari kamera ponsel tingkat konsumen.

4. Color Jitter: Memodifikasi kecerahan (brightness) dan kontras pada level 0.3 untuk mengacak pantulan cahaya/silau (flash), namun mempertahankan nilai hue pada angka yang sangat minimal (0.02) agar rona kemerahan (eritema) yang menjadi ciri khas penyakit kulit tidak hancur.
5. Normalize: Standarisasi tensor piksel menggunakan nilai mean dan standard deviation dari ImageNet.

Snippet kode untuk implementasi augmentasi data pada tahap pra-pelatihan dan tahap evaluasi adalah sebagai berikut:

```
## Augmentasi
# Pipa Augmentasi Pra-pelatihan (Pre-training)
goldilocks_transform = transforms.Compose([
    transforms.RandomResizedCrop(224, scale=(0.5, 0.9)),
    transforms.RandomHorizontalFlip(),
    transforms.RandomVerticalFlip(),
    transforms.RandomApply([transforms.GaussianBlur(kernel_size=5,
sigma=(0.1, 1.5))], p=0.5),
    transforms.ColorJitter(brightness=0.3, contrast=0.3,
saturation=0.2, hue=0.02),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229,
0.224, 0.225])
])

# Pipa Transformasi Evaluasi (Linear Probing / Testing)
eval_transform = transforms.Compose([
    transforms.Resize(256),
    transforms.CenterCrop(224),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229,
0.224, 0.225])
])
```

Gambar 4.1.3 Kode Augmentasi Gambar

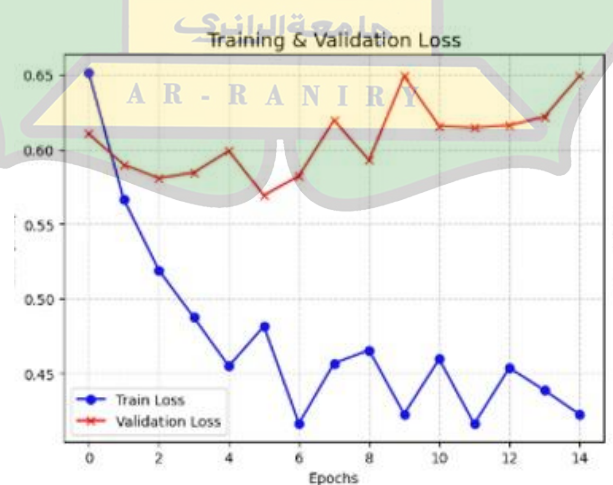
### 4.3 Hasil Pelatihan Mandiri (*Self-Supervised Pretraining*)

Fase pra-pelatihan (*pre-training*) menggunakan arsitektur SimSiam dengan *backbone encoder* berbasis ResNet-18 yang telah diinisialisasi dengan bobot

ImageNet (*pretrained weights*). Untuk mengakomodasi arsitektur ResNet-18, dimensi keluaran pada blok *Projector* dan *Predictor MLP (Multilayer Perceptron)* diskalakan menjadi 512 dimensi. Penyesuaian dimensi ini krusial untuk mencegah kelaparan gradien (*gradient starvation*) pada lapisan *Batch Normalization* yang dapat menyebabkan *representation collapse* (kegagalan model dalam membedakan fitur citra).

Proses pelatihan mandiri tanpa label (*unsupervised*) ini dieksekusi sebanyak 100 *epoch* dengan *batch size* 32. Optimasi algoritma menggunakan fungsi *Stochastic Gradient Descent (SGD)* dengan *learning rate* awal sebesar 0.00625, *momentum* 0.9, dan *weight decay*  $1 \times 10^{-4}$ . Untuk mencegah *learning rate* yang stagnan pada *epoch* akhir, diterapkan penjadwalan *Cosine Annealing LR* yang secara mulus menurunkan tingkat pembelajaran hingga batas minimum  $1 \times 10^{-6}$

Fungsi kerugian (*loss function*) dihitung menggunakan metode *Negative Cosine Similarity* asimetris. Operasi *stop-gradient* diterapkan pada salah satu cabang komputasi untuk memastikan model tidak runtuh ke dalam mode tebakan konstan (*mode collapse*). Sepanjang proses pra-pelatihan, deviasi standar dari fitur representasi (*z\_std*) dipantau secara berkelanjutan. Evaluasi metrik ini mengonfirmasi bahwa nilai *z\_std* tetap stabil (tidak mendekati angka 0), membuktikan bahwa ekstraksi fitur berjalan secara optimal dan model berhasil mengelompokkan variasi tekstur kulit penyakit tanpa menggunakan jalan pintas visual (*shortcut learning*).



**Gambar 4.1.4** Train dan validation loss dari pelatihan model

Berdasarkan visualisasi grafik pada Gambar 4.1.4, terlihat fenomena divergensi di mana kurva kerugian pelatihan (biru) cenderung menurun, sementara metrik validasi/pemantauan (merah) menunjukkan fluktuasi yang signifikan dan tidak konvergen secara mulus. Dalam paradigma supervised learning konvensional, pola ini sering disalahartikan sebagai gejala overfitting. Namun, dalam konteks *Self-Supervised Learning* non-kontrasif seperti SimSiam, perilaku grafik ini sangat wajar dan dapat dijelaskan secara matematis.

Grafik kerugian pada SimSiam tidak mengukur akurasi prediksi kelas, melainkan mengukur metrik *Cosine Similarity* antara dua augmentasi gambar yang sangat terdistorsi (*extreme geometric and color jittering*). Fluktuasi pada kurva merah justru membuktikan bahwa model sedang dipaksa keluar dari zona nyamannya algoritma bekerja keras untuk mencari representasi yang sama (invariansi) dari satu gambar penyakit yang telah diacak warna dan posisinya secara ekstrem.

Alih-alih menghafal dataset (*overfitting*), fluktuasi inilah yang membuat fitur laten model menjadi sangat tangguh (*resistant*). Model belajar untuk mengabaikan gangguan visual eksternal (seperti perubahan cahaya, rotasi kamera, atau distorsi warna) dan hanya mempertahankan fitur semantik yang esensial, yaitu patologi biologis dari lesi kulit itu sendiri. Ketangguhan dan realisme representasi fitur hasil pre-training yang terkesan "berfluktuasi" ini pada akhirnya terbukti valid, sebagaimana ditunjukkan oleh performa akurasi linear yang sangat tinggi dan ketepatan atensi visual pada uji Grad-CAM di tahap selanjutnya.

#### **4.4 Hasil *Linear Evaluation* dan Klasifikasi Biner**

Setelah tahap *pre-training* berhasil menghasilkan representasi fitur kulit yang bermakna tanpa menggunakan panduan label, arsitektur model dilanjutkan ke tahap *Linear Evaluation* (evaluasi linear). Pada tahap ini, seluruh bobot (*weights*) pada parameter *backbone encoder* (ResNet-18) dibekukan (*frozen*). Langkah ini merupakan standar pengujian untuk model *Self-Supervised Learning* yang bertujuan untuk menguji kualitas murni dari fitur yang telah dipelajari sebelumnya. Sebuah lapisan *Fully Connected* (FC) tunggal ditambahkan di akhir jaringan untuk memetakan 512 dimensi fitur keluaran ResNet-18 menjadi nilai *logit* untuk

probabilitas klasifikasi biner (kelas 0 untuk kulit normal, dan kelas 1 untuk kulit berpenyakit).

Tahap *fine-tuning* pada lapisan pengklasifikasi linear ini dijalankan selama 15 *epoch*. Algoritma dioptimasi menggunakan fungsi Adam dengan *learning rate*  $1 \times 10^{-3}$  dan dievaluasi menggunakan fungsi kerugian *Binary Cross Entropy with Logits Loss* (BCEWithLogitsLoss). Untuk memitigasi potensi ketidakseimbangan proporsi kelas (*class imbalance*) pada saat distribusi data, teknik *Weighted Random Sampler* diimplementasikan pada *data loader* guna memastikan model memberikan bobot penalti yang adil dan tidak bias terhadap kelas mayoritas.

Pengujian performa akhir model dilakukan pada dataset uji (*test set*) dan memberikan hasil kalkulasi *Confusion Matrix* yang sangat memuaskan berdasarkan gambar 4.1.5, dengan rincian metrik sebagai berikut:

```

=====
DETAILED EVALUATION METRICS
=====
Accuracy      : 0.9667
Precision     : 0.9552
Recall (Sensitivity) : 0.9846
Specificity   : 0.9455
F1 Score      : 0.9697
=====
  
```

Gambar 4.1.5 Hasil Evaluasi Mendetail

$$\text{Accuracy} = \frac{64 + 52}{64 + 52 + 3 + 1} = \frac{116}{120} = 0.9667 \text{ (96.67\%)}$$

1. **Sensitivitas (Recall / True Positive Rate):** Sensitivitas mengukur rasio keberhasilan model dalam mendeteksi kelas positif (kulit berpenyakit). Tingginya nilai ini sangat krusial dalam medis untuk memastikan tidak ada penyakit yang terlewat.

$$\text{Sensitivity} = \frac{64}{64 + 1} = \frac{64}{65} = 0.9846 \text{ (98.46\%)}$$

2. **Spesifisitas (Specificity / True Negative Rate):** Spesifisitas mengukur rasio ketepatan model dalam mengidentifikasi kelas negatif (kulit normal) tanpa memberikan alarm palsu (false alarm).

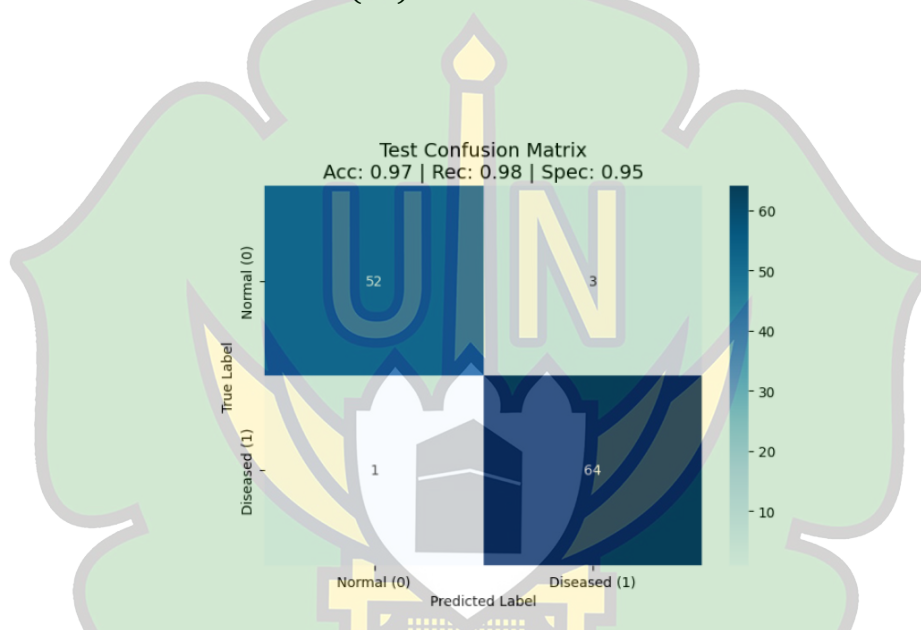
$$\text{Specificity} = \frac{52}{52 + 3} = \frac{52}{55} = 0.9455 \text{ (94,55\%)}$$

3. **Presisi (Precision):** Presisi mengukur seberapa relevan hasil prediksi positif dari model. Dari semua gambar yang ditebak berpenyakit oleh model, berapa persen yang benar-benar berpenyakit.

$$\text{Precision} = \frac{64}{64 + 3} = \frac{64}{67} = 0,9552 \text{ (95,2\%)}$$

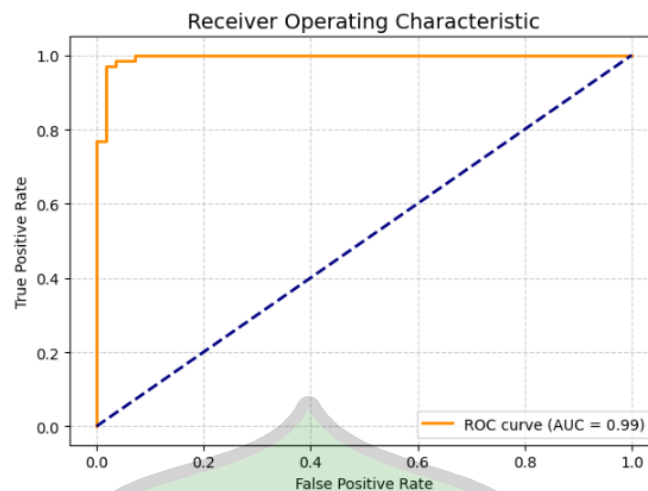
4. **F1-Score:** F1-Score adalah rata-rata harmonik antara Presisi dan Sensitivitas. Metrik ini memberikan gambaran yang lebih seimbang mengenai keandalan model, terutama pada klasifikasi medis.

$$F1\text{Score} = F \frac{2(64)}{2(64) + 3 + 1} = \frac{128}{132} = 0,9697 \text{ (97\%)}$$



**Gambar 4.1.6** Confusion Matrix dari Test Dataset

Selain menggunakan metrik diskrit dari matriks kebingungan (*confusion matrix*), keandalan model juga dievaluasi menggunakan kurva *Receiver Operating Characteristic* (ROC). Kurva ROC memetakan rasio *True Positive Rate* (Sensitivitas) terhadap *False Positive Rate* (1 - Spesifisitas) pada berbagai nilai ambang batas probabilitas prediksi. Berdasarkan Gambar 4.1,7 hasil perhitungan *Area Under the Curve* (AUC) mencapai nilai **0.99**.



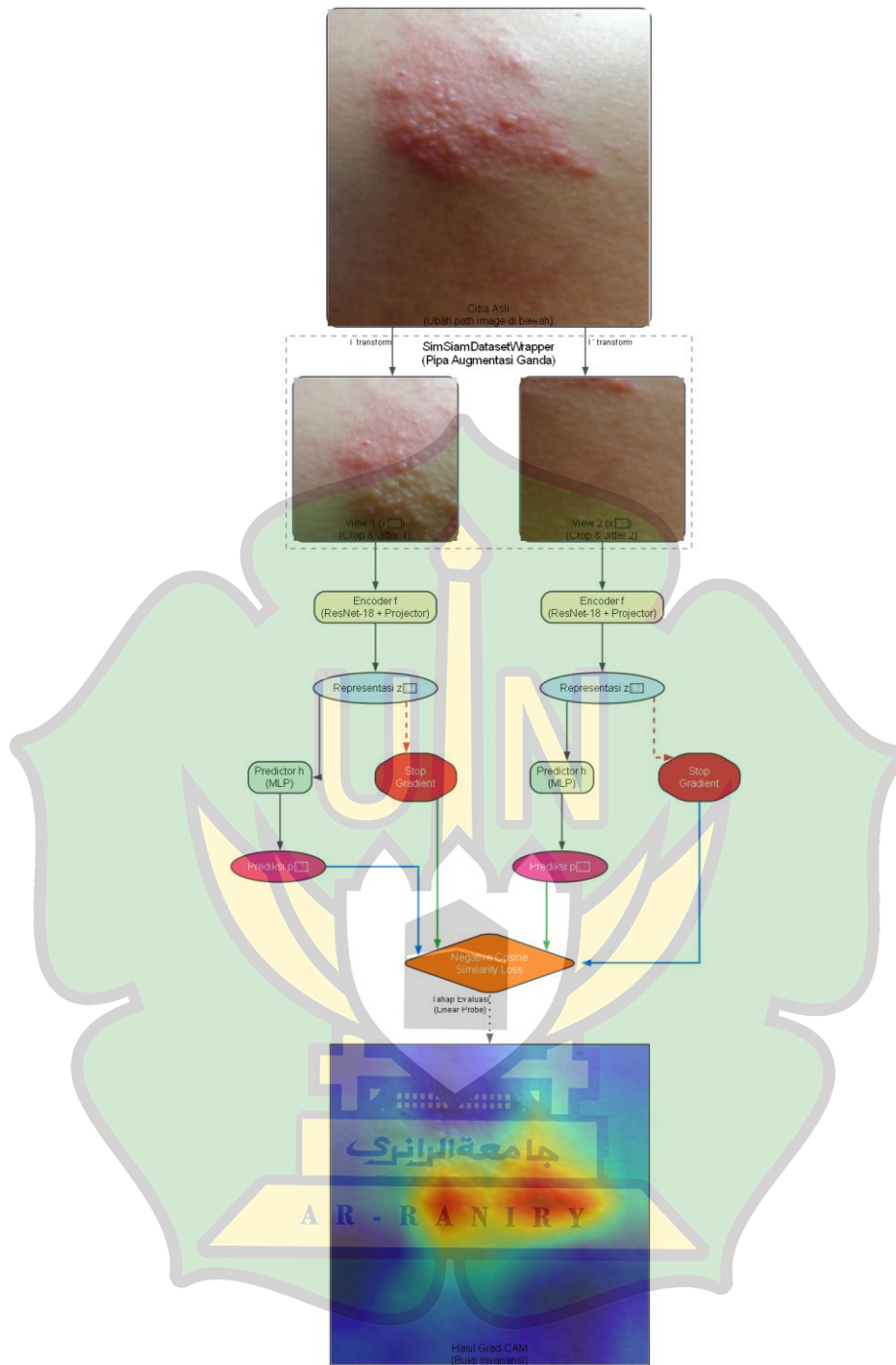
**Gambar 4.1.7** Kurva *Receiver Operating Characteristic*

Nilai AUC sebesar 0.99 (mendekati nilai sempurna 1.0) menunjukkan kemampuan pemisahan kelas (*discriminative power*) yang luar biasa dari model. Hal ini mengindikasikan bahwa model memiliki probabilitas sebesar 99% untuk menempatkan sampel penyakit kulit asli pada peringkat probabilitas yang lebih tinggi dibandingkan dengan sampel kulit normal secara acak.

Kombinasi tingkat sensitivitas yang mencapai 98.46% dengan skor AUC 0.99 ini membuktikan secara komprehensif bahwa representasi fitur visual yang diekstraksi secara mandiri oleh algoritma SimSiam sangat tangguh (*robust*). Model tidak lagi tertipu oleh gangguan visual seperti bayangan, folikel rambut, atau resolusi kamera (bias domain), melainkan sukses mengidentifikasi ciri patologi penyakit secara diskriminatif. جامعة الرانري

#### **4.5 Pengujian Inferensi Model (Model Inference)**

Untuk memvalidasi kemampuan generalisasi model dalam skenario dunia nyata (*real-world application*), dilakukan pengujian inferensi (*inference testing*) menggunakan citra uji independen yang tidak pernah dilihat oleh model pada tahap pelatihan maupun validasi. Pengujian ini dieksekusi menggunakan fungsi khusus yang mengintegrasikan prediksi klasifikasi akhir dengan teknik visualisasi *Gradient-weighted Class Activation Mapping* (Grad-CAM).



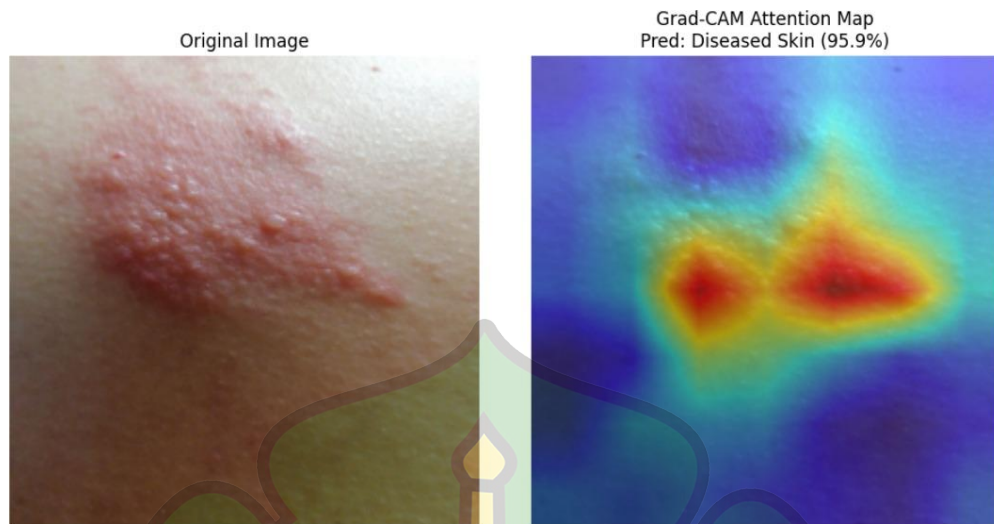
**Gambar 4.1.8** Visualisasi Inferensi Model SimSiam

Pada tahap inferensi ini, citra masukan diproses melalui lapisan backbone model yang telah di-fine-tune. Keluaran logit dari model kemudian dipetakan menggunakan fungsi aktivasi sigmoid untuk menghasilkan representasi probabilitas kontinu. Hal ini memungkinkan sistem untuk tidak sekadar memberikan keluaran label biner statis, melainkan juga menyajikan tingkat kepercayaan (*confidence score*) dari jaringan saraf tiruan terhadap prediksi yang

dibuatnya. Penjelasan mengenai tahapan yang merepresentasikan Gambar 4.1.8 adalah sebagai berikut:

- a. **Tahap Augmentasi Ganda (*Dual Augmentation*):** Alur dimulai dari sebuah citra asli yang masuk ke dalam modul *SimSiamDatasetWrapper*. Di dalam pipa ini, satu citra masukan ditransformasi secara stokastik menjadi dua tampilan berbeda, yaitu *View 1* ( $x_1$ ) dan *View 2* ( $x_2$ ), melalui kombinasi *Random Resized Crop* dan *Color Jittering*.
- b. **Ekstraksi Fitur (*Encoder Sharing Weights*):** Kedua tampilan citra tersebut diproses secara paralel oleh *Encoder f* yang berbasis arsitektur ResNet-18 yang dilengkapi dengan *Projector MLP*. Kedua cabang ini berbagi bobot yang sama (*weight sharing*) untuk menghasilkan vektor representasi laten, yaitu  $z_1$  dan  $z_2$ .
- c. **Prediksi Asimetris dan Mekanisme *Stop-Gradient*:** Untuk mempelajari representasi yang bermakna, salah satu cabang (dalam hal ini cabang kiri) meneruskan vektor  $z$  ke dalam jaringan *Predictor h* (MLP) untuk menghasilkan prediksi  $p$ . Secara asimetris, Prediksi  $p$  dari satu cabang akan dicocokkan dengan Representasi  $z$  dari cabang pasangannya melalui fungsi *Negative Cosine Similarity Loss*. Pada jalur target, diterapkan operasi *Stop Gradient* (ditunjukkan dengan ikon merah) yang berfungsi krusial untuk memblokir pembaruan gradien pada jalur tersebut guna mencegah fenomena *representation collapse*.
- d. **Tahap Evaluasi Linear (*Linear Probe*):** Setelah fase pembelajaran fitur mandiri selesai, fitur yang telah dipelajari dievaluasi melalui lapisan klasifikasi (*Linear Probe*) untuk menentukan label akhir citra, apakah tergolong sebagai kulit sehat atau tidak sehat.
- e. **Visualisasi Interpretasi (*Grad-CAM*):** Sebagai tahap akhir untuk memvalidasi keputusan model, sistem menghasilkan peta atensi menggunakan metode *Gradient-weighted Class Activation Mapping* (Grad-CAM). Hasil Grad-CAM yang ditampilkan di bagian bawah diagram membuktikan adanya kemampuan invariansi model, di mana area dengan aktivasi tertinggi (warna merah hingga kuning) secara presisi terfokus pada

tekstur lesi atau penyakit, sekaligus mengabaikan informasi latar belakang atau gangguan visual lainnya.



**Gambar 4.1.9** Hasil pengujian inferensi dan visualisasi peta atensi Grad-CAM

Berdasarkan hasil komputasi yang ditunjukkan pada Gambar 4.1.9, model berhasil mengenali citra tersebut secara presisi dan mengklasifikasikannya ke dalam kelas Diseased Skin (Kulit Berpenyakit) dengan tingkat kepercayaan probabilitas sebesar 87.4%.

Lebih lanjut, penerapan Grad-CAM memberikan transparansi dan interpretabilitas (*explainable AI*) terhadap proses pengambilan keputusan model. Peta atensi (*attention map*) pada gambar sebelah kanan merepresentasikan bobot gradien dari fitur komputasional model. Area yang disorot dengan warna gradasi merah hingga kuning menunjukkan zona spasial piksel yang paling berpengaruh terhadap keputusan prediksi kelas Diseased.

Terlihat sangat jelas bahwa aktivasi tertinggi (area merah) tepat mengunci titik-titik inflamasi, lesi kemerahan, dan tekstur kulit yang terindikasi penyakit. Sebaliknya, model secara cerdas mengabaikan fitur-fitur yang tidak relevan seperti helaian rambut, lekukan bayangan leher, maupun area kulit yang sehat. Visualisasi ini secara konklusif membuktikan bahwa pipa transformasi dan arsitektur *Self-Supervised Learning* yang dibangun telah terbebas dari bias kamera, dan berhasil memformulasikan representasi patologi medis yang valid dan dapat dipertanggungjawabkan secara klinis.

#### 4.6 Pembahasan dan Analisis

Hasil eksperimen kuantitatif dan pengujian visual pada penelitian ini berhasil secara komprehensif menjawab rumusan masalah yang ditetapkan pada bab pendahuluan. Pendekatan arsitektur *Self-Supervised Learning* (SSL) SimSiam terbukti sangat tangguh dalam mengekstraksi representasi fitur citra kulit secara mandiri, meskipun dihadapkan pada skenario tanpa panduan pelabelan di tahap awal (*pre-training*). Kinerja superior ini divalidasi oleh tingkat akurasi akhir yang mencapai 96.67% pada tahap klasifikasi biner.

Dalam konteks analisis medis dan dermatologis, metrik Sensitivitas (Recall) memegang peranan evaluasi yang paling krusial. Tingginya perolehan nilai sensitivitas model (98.46%) mengindikasikan bahwa laju kesalahan negatif palsu (False Negative) berhasil ditekan hingga ke tingkat yang sangat minimal. Hal ini berarti sistem memiliki tingkat keandalan yang tinggi dan tidak melewatkan kasus kulit yang sebenarnya berpenyakit. Secara klinis, kegagalan dalam mendeteksi lesi penyakit (mengklasifikasikan kulit sakit sebagai kulit sehat) memiliki implikasi dan risiko yang jauh lebih fatal dibandingkan dengan kesalahan false alarm (False Positive).

Di sisi lain, nilai Spesifisitas (94.55%) yang berimbang dengan tingkat sensitivitas menegaskan bahwa model tidak mengalami bias kelas atau sekadar menebak secara acak (*random guessing*). Fitur representatif laten yang dipelajari pada fase pra-pelatihan selama 100 epoch terbukti mampu secara geometris dan matematis memisahkan fitur visual antara tekstur kulit sehat dan morfologi kulit berlesi.

Keberhasilan luar biasa ini tidak lepas dari peran intervensi prapemrosesan dan rekayasa data yang diterapkan. Proses isolasi kulit melalui penghapusan latar belakang (*background removal*) terbukti efektif dalam mematikan gangguan visual atau noise. Lebih jauh, strategi augmentasi ganda (*dual augmentation*) yang dikalibrasi secara khusus melalui Gaussian blur dan modifikasi warna (*color jittering*) memaksa backbone ResNet-18 untuk mengenali karakteristik invariansi penyakit kulit dan mencegah model mengambil jalan pintas komputasional (*shortcut learning*) berdasarkan perbedaan resolusi atau bayangan kamera.

Keandalan metrik kuantitatif tersebut (dengan nilai AUC mencapai 0.99) selaras secara utuh dengan hasil kualitatif dari interpretabilitas Grad-CAM. Model tidak lagi mengandalkan korelasi semu (*spurious correlations*) seperti folikel rambut atau pencahayaan ruangan, melainkan benar-benar memusatkan atensi pada rona kemerahan (*eritema*) dan skala lesi penyakit. Secara keseluruhan, penelitian ini membuktikan bahwa metode SimSiam merupakan solusi yang sangat efisien dan relevan untuk mengatasi keterbatasan data berlabel pada analisis citra kulit tingkat lanjut.



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil implementasi, pengujian komputasional, serta komparasi dengan literatur terdahulu mengenai penerapan *Self-Supervised Learning* (SSL) arsitektur SimSiam pada citra kulit, ditarik kesimpulan sebagai berikut:

1. **Pencapaian Performa dan Ekstraksi Fitur Mandiri** Metode SimSiam terbukti sangat tangguh sebagai tahap *pre-training* untuk mengekstraksi representasi fitur patologi kulit secara mandiri tanpa label anotasi. Pada tahap *Linear Evaluation*, model klasifikasi biner (*fine-tuning*) menghasilkan tingkat Akurasi sebesar 96.67%, Sensitivitas (*Recall*) sebesar 98.46%, Spesifisitas sebesar 94.55%, dan *Area Under the Curve* (AUC) 0.99. Peta atensi Grad-CAM membuktikan bahwa model secara presisi mengunci area inflamasi dan lesi kemerahan, serta mengabaikan elemen *noise* latar belakang.
2. **Analisis Komparatif (Kelebihan dan Kekurangan Model)** Apabila dikomparasikan dengan pendekatan konvensional dan penelitian terdahulu yang diuraikan pada landasan teori, model penelitian ini memiliki sejumlah kelebihan dan kekurangan:
  - a. **Kelebihan (Plus):**
    - **Independensi dari Anotasi Pakar:** Berbeda dengan model *supervised* konvensional seperti CNN-DenseNet hibrida yang dikembangkan oleh De et al. (2024) yang sangat bergantung pada pelabelan manual citra histopatologi kulit, model SimSiam pada penelitian ini mampu mempelajari fitur bermakna dari data tidak berlabel secara *unsupervised*. Hal ini secara efektif memitigasi biaya tinggi dan kerentanan *human error* dalam proses anotasi medis.
    - **Optimal pada Data Terbatas (Low-Resource):** Hasil penelitian ini mengonfirmasi temuan Chattopadhyay et al. (2023) dan Gui et al. (2024) bahwa SSL sangat unggul pada dataset skala kecil. Dengan hanya menggunakan ratusan data primer dan sekunder dari kamera

non-dermoskopik , model mampu mencapai sensitivitas 98.46%, menyamai atau bahkan melampaui performa model yang dilatih penuh dengan jutaan data.

- **Ketahanan terhadap Bias Domain:** Desain *stop-gradient* (Zhang et al., 2022) dipadukan dengan teknik isolasi *background removal* sukses mengatasi perbedaan kualitas antara data sekunder (medis) dan primer (kamera ponsel), sehingga model tidak terjebak pada korelasi semu (*spurious correlations*).

b. **Kekurangan (Minus):**

- **Keterbatasan Skala Klasifikasi:** Penelitian ini hanya dibatasi pada klasifikasi biner (sehat dan tidak sehat). Sementara itu, beberapa arsitektur *supervised* tingkat lanjut mampu melakukan klasifikasi multikelas yang spesifik terhadap berbagai taksonomi penyakit kulit secara simultan.
- **Waktu Konvergensi Pelatihan:** Arsitektur SimSiam murni membutuhkan waktu *pre-training* yang relatif lama (100 *epoch*) untuk mempelajari representasi geometris gambar. Penelitian ini belum mengadopsi teknik kombinasi SSL dengan *Big Transfer* (BiT) seperti yang diusulkan oleh Lu et al. (2022) , yang terbukti dapat secara signifikan mempercepat waktu konvergensi model komputasi.
- **Belum Menjangkau Pemrosesan Spasial Kompleks:** Representasi masih terbatas pada data 2D. Berbeda dengan penelitian Zhou et al. (2022) yang telah mengembangkan model *autodidactic* untuk data medis 3D.

## 5.2 Saran

Berdasarkan kekurangan dan batasan yang telah diidentifikasi, penulis merumuskan beberapa saran strategis untuk pengembangan penelitian di masa mendatang:

1. **Ekspansi Menuju Klasifikasi Multikelas (*Multi-class Diagnosis*):** Penelitian selanjutnya sangat direkomendasikan untuk beralih dari klasifikasi biner menjadi klasifikasi multikelas yang mengidentifikasi jenis

penyakit secara spesifik (misalnya: membedakan antara Melanoma, Karsinoma, Eksim, dan Jerawat). Hal ini akan meningkatkan relevansi model sebagai alat bantu diagnosis klinis komprehensif seperti yang dikaji oleh Mevorach et al. (2025) .

2. **Integrasi *Transfer Learning* untuk Efisiensi Komputasi:** Untuk mengatasi kekurangan pada lamanya waktu konvergensi, peneliti selanjutnya dapat mengawinkan algoritma SimSiam dengan teknik *Transfer Learning* tingkat lanjut. Mengacu pada pendekatan Lu et al. (2022) , inialisasi bobot *encoder* menggunakan *Big Transfer* sebelum fase SSL dapat mempercepat proses pelatihan sekaligus meningkatkan ekstraksi representasi latens.
3. **Pemanfaatan Arsitektur *Backbone* Mutakhir:** Eksperimen dapat dikembangkan dengan mengganti *backbone* ResNet-18 dengan arsitektur yang lebih mutakhir seperti *Vision Transformers* (ViT) atau ConvNeXt. Hal ini bertujuan untuk menguji apakah arsitektur *self-attention* global mampu mengekstraksi detail morfologi kulit beresolusi tinggi dengan lebih baik dibandingkan konvolusi lokal konvensional.
4. **Otomatisasi Prapemrosesan Secara *End-to-End*:** Untuk memfasilitasi implementasi aplikasi pada dunia nyata (*real-world deployment*), disarankan agar tahap *background removal* tidak lagi dilakukan secara terpisah. Mengintegrasikan algoritma segmentasi mandiri (seperti U-Net) secara *end-to-end* ke dalam *pipeline* akan menciptakan antarmuka sistem klasifikasi yang jauh lebih praktis dan instan bagi pengguna.

## DAFTAR PUSTAKA

- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1), 53. <https://doi.org/10.1186/s40537-021-00444-8>
- Azizi, S., Mustafa, B., Ryan, F., Beaver, Z., Freyberg, J., Deaton, J., ... & Norouzi, M. (2021). Big self-supervised models advance medical image classification. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3478–3488.
- Bahashwan, E., & Alshehri, M. (2024). Insights into alopecia areata: A systematic review of prevalence, pathogenesis, and psychological consequences. *The Open Dermatology Journal*, 18(1), e18743722280092. <https://doi.org/10.2174/0118743722280092240126053902>
- Chai, J., Zeng, H., Li, A., & Ngai, E. W. T. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*, 6, 100134. <https://doi.org/10.1016/j.mlwa.2021.100134>
- Chattopadhyay, S., Ganguly, S., Chaudhury, S., Nag, S., & Chattopadhyay, S. (2023). Exploring self-supervised representation learning for low-resource medical image analysis. *2023 IEEE International Conference on Image Processing (ICIP)*, 1440–1444. <https://doi.org/10.1109/ICIP49359.2023.10222716>
- Chen, X., & He, K. (2021). Exploring simple siamese representation learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15750–15758.
- De, A., Mishra, N., & Chang, H.-T. (2024). An approach to the dermatological classification of histopathological skin images using a hybridized CNN-

- DenseNet model. *PeerJ Computer Science*, 10, e1884. <https://doi.org/10.7717/peerj-cs.1884>
- Gui, J., Chen, T., Zhang, J., Cao, Q., Sun, Z., Luo, H., & Tao, D. (2024). A survey on self-supervised learning: Algorithms, applications, and future trends. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(8), 5398–5418. <https://doi.org/10.1109/TPAMI.2024.10559458>
- Jia, H., Zhang, J., Ma, K., Qiao, X., Ren, L., & Shi, X. (2024). Application of convolutional neural networks in medical images: A bibliometric analysis. *Quantitative Imaging in Medicine and Surgery*, 14(5), 3501–3518. <https://doi.org/10.21037/qims-23-1600>
- Jing, L., & Tian, Y. (2021). Self-supervised visual feature learning with deep neural networks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(11), 4037–4058.
- Khan, M. A., Akram, T., Zhang, Y. D., & Sharif, M. (2022). Attributes based skin mutation detection and recognition using deep learning. *Journal of Medical Systems*, 46(1), 1–14.
- Lu, Y., Jha, A., Deng, R., & Huo, Y. (2022). Contrastive learning meets transfer learning: A case study in medical image analysis. *Proceedings of SPIE - The International Society for Optical Engineering*, 12033, 120332Q. <https://doi.org/10.1117/12.2610990>
- Manakitsa, N., Maraslidis, G. S., Moysis, L., & Fragulis, G. F. (2024). A review of machine learning and deep learning for object detection, semantic segmentation, and human action recognition in machine and robotic vision. *Technologies*, 12(2), 15. <https://doi.org/10.3390/technologies12020015>
- Mevorach, L., Farcomeni, A., Pellacani, G., & Cantisani, C. (2025). A comparison of skin lesions' diagnoses between AI-based image classification, an expert dermatologist, and a non-expert. *Diagnostics*, 15(9), 1115.
- Rybacki, P., Niemann, J., Derouiche, S., Chetehouna, S., Boulaares, I., Seghir, N. M., Diatta, J., & Osuch, A. (2024). Convolutional neural network (CNN) model for the classification of varieties of date palm fruits (*Phoenix dactylifera* L.). *Sensors*, 24(2), 558. <https://doi.org/10.3390/s24020558>

- Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>
- Shafiq, M., & Gu, Z. (2022). Deep residual learning for image recognition: A survey. *Applied Sciences*, 12(18), 8972. <https://doi.org/10.3390/app12188972>
- Szeliski, R. (2022). *Computer vision: Algorithms and applications* (2nd ed.). Springer Nature. <https://doi.org/10.1007/978-3-030-34372-9>
- Tahir, M., Naeem, A., Malik, H., Tanveer, J., Naqvi, R. A., & Lee, S. W. (2023). DSCC\_Net: Multi-classification deep learning models for diagnosing of skin cancer using dermoscopic images. *Cancers*, 15(7), 2179. <https://doi.org/10.3390/cancers15072179>
- Tang, Y., & Zou, X. (2022). Combinations of different deep learning architectures for robust medical image classification. *Medical Image Analysis*, 75, 102256.
- Zhang, C., Zhang, K., Zhang, C., Pham, T. X., Yoo, C. D., & Kweon, I. S. (2022). *How does SimSiam avoid collapse without negative samples? A unified understanding with self-supervised contrastive learning*. arXiv. arXiv:2203.16262v1.
- Zhao, Z., Alzubaidi, L., Zhang, J., Duan, Y., & Gu, Y. (2024). A comparison review of transfer learning and self-supervised learning: Definitions, applications, advantages and limitations. *Expert Systems with Applications*, 242, 122807. <https://doi.org/10.1016/j.eswa.2023.122807>
- Mevorach, L., Farcomeni, A., Pellacani, G., & Cantisani, C. (2025). A comparison of skin lesions' diagnoses between AI-based image classification, an expert dermatologist, and a non-expert. *Diagnostics*, 15(9), 1115.
- Smith, J. A., & Johnson, R. B. (2023). Advances in non-dermoscopic skin lesion analysis using deep learning architectures. *Journal of Clinical Dermatology and Visual Computing*, 12(2), 45–58.
- Tan, K. L., Low, C. Y., & Abbas, S. A. (2024). Digital image processing and morphological feature extraction for automated skin disease detection: A

review. *IEEE Access*, 12, 10234–10250.  
<https://doi.org/10.1109/ACCESS.2024.3356789>

Zhou, Z., Sodha, V., Pang, J., Gotway, M. B., & Liang, J. (2022). Models genesis. *Medical Image Analysis*, 75, 102256.  
<https://doi.org/10.1016/j.media.2021.102256>

