

**EKSPERIMEN MODEL VISION TRANSFORMER DAN SWIN
TRANSFORMER BERBASIS FINE-TUNING UNTUK MULTI-
LABEL MULTI-KRITERIA PADA CITRA X-RAY PARU-
PARU**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 INFORMATIKA



disusun oleh

RIZKY NANDA ANGGIA

22.11.4825

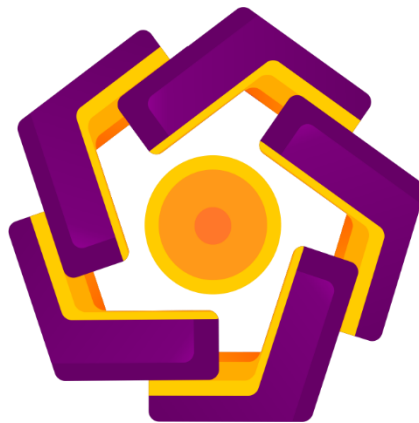
Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

**EKSPERIMEN MODEL VISION TRANSFORMER DAN SWIN
TRANSFORMER BERBASIS FINE-TUNING UNTUK MULTI-
LABEL MULTI-KRITERIA PADA CITRA X-RAY PARU-
PARU**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 INFORMATIKA



disusun oleh

RIZKY NANDA ANGGIA

22.11.4825

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**EKSPERIMEN MODEL VISION TRANSFORMER DAN SWIN
TRANSFORMER BERBASIS FINE-TUNING UNTUK MULTI-LABEL
MULTI-KRITERIA PADA CITRA X-RAY PARU-PARU**

yang disusun dan diajukan oleh

RIZKY NANDA ANGGIA

22.11.4825

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 28 Januari 2026

Dosen Pembimbing,



Theopilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375

HALAMAN PENGESAHAN

SKRIPSI

EKSPERIMEN MODEL VISION TRANSFORMER DAN SWIN
TRANSFORMER BERBASIS FINE-TUNING UNTUK MULTI-LABEL
MULTI-KRITERIA PADA CITRA X-RAY PARU-PARU

yang disusun dan diajukan oleh

RIZKY NANDA ANGGIA

22.11.4825

Telah dipertahankan di depan Dewan Penguji
pada tanggal 28 Januari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Rizqi Sukma Kharisma, S.Kom., M.Kom.
NIK. 190302215

Andika Agus Slameto, S.Kom., M.Kom.
NIK. 190302109

Theopilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 28 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : RIZKY NANDA ANGGIA
NIM : 22.11.4825

Menyatakan bahwa Skripsi dengan judul berikut:

EKSPERIMEN MODEL VISION TRANSFORMER DAN SWIN TRANSFORMER BERBASIS FINE-TUNING UNTUK MULTI-LABEL MULTI-KRITERIA PADA CITRA X-RAY PARU-PARU

Dosen Pembimbing : Theopilus Bayu Sasongko, S.Kom., M.Eng.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 28 Januari 2026

Yang Menyatakan,



METERAI TEMPEL
68AANX307629921

RIZKY NANDA ANGGIA

HALAMAN PERSEMBAHAN

Puji syukur kehadiran Allah SWT, Tuhan semesta alam, atas segala rahmat dan karunia-Nya yang tak terhingga, sehingga naskah ini dapat terselesaikan tepat waktu dan dengan segala kerendahan hati penulis persembahkan kepada:

1. Papah, Mamah, dan Adik tersayang, serta seluruh Keluarga Besar saya yang selalu memberikan dukungan, doa, semangat, dan cinta kasih tanpa henti. Kalian adalah sumber kekuatan dan inspirasi terbesar dalam hidup saya.
2. Universitas Amikom Yogyakarta, sebagai institusi pendidikan yang telah menjadi tempat saya menimba ilmu, dengan lingkungan akademis yang kondusif dan fasilitas yang mendukung proses belajar dan pengembangan diri saya.
3. Bapak/Ibu Dosen pembimbing, yang dengan sabar, telaten, dan penuh dedikasi telah memberikan bimbingan, arahan, serta ilmu pengetahuan yang sangat berharga selama proses penelitian dan penulisan skripsi ini.
4. Teman-teman seperjuangan saya selama berkuliah di Universitas Amikom Yogyakarta. Kebersamaan, bantuan, dan kenangan indah yang telah kita ukir tidak akan pernah terlupakan.

KATA PENGANTAR

Dengan puji dan syukur kehadiran Allah Subhanahu wa Ta'ala atas limpahan rahmat, taufik, dan hidayah-Nya, penulis dapat menyelesaikan skripsi yang berjudul “Eksperimen Model *Vision Transformer* dan *Swin Transformer* Berbasis *Fine-Tuning* untuk *Multi-Label Multi-Kriteria* pada Citra *X-Ray* Paru-Paru.” sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Penulis menyadari bahwa penyusunan dan penyelesaian skripsi ini tidak terlepas dari bantuan, dukungan, dan doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah Subhanahu wa Ta'ala, Tuhan semesta alam yang senantiasa memberikan nikmat iman, kesehatan, dan kemudahan dalam setiap langkah penulis hingga skripsi ini dapat terselesaikan.
2. Bapak dan Ibu tercinta, adek, serta seluruh keluarga besar, yang dengan penuh kasih sayang, doa, dan dukungan tanpa henti telah menjadi sumber semangat dan motivasi dalam menempuh pendidikan hingga penyusunan skripsi ini.
3. Bapak Theopilus Bayu Sasongko, S.Kom., M.Eng., selaku Dosen Pembimbing yang telah meluangkan waktu, memberikan arahan, serta bimbingan dengan sabar dan penuh perhatian dari tahap awal hingga akhir penelitian ini.
4. Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta yang telah memberikan kesempatan, fasilitas, dan dukungan dalam proses menempuh pendidikan tinggi di Universitas Amikom Yogyakarta.
5. Seluruh Dosen Program Studi Informatika Universitas Amikom Yogyakarta, atas ilmu, bimbingan, serta motivasi yang telah diberikan selama masa perkuliahan hingga penyusunan tugas akhir ini.

6. Rekan-rekan mahasiswa Informatika angkatan 2021, yang selalu memberikan dukungan, semangat, dan kebersamaan selama proses perkuliahan serta penyusunan skripsi.
7. Seluruh pihak lainnya yang tidak dapat penulis sebutkan satu per satu, namun telah memberikan bantuan, doa, serta dukungan baik secara langsung maupun tidak langsung dalam penyelesaian skripsi ini.

Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis dengan lapang dada menerima segala bentuk kritik dan saran yang membangun demi penyempurnaan karya ini di masa mendatang.

Akhir kata, penulis berharap semoga skripsi ini dapat memberikan manfaat bagi pembaca, menjadi tambahan wawasan dalam bidang Computer Vision dan Deep Learning, serta menjadi inspirasi bagi penelitian selanjutnya dalam pengembangan teknologi kecerdasan buatan di bidang kesehatan.

Yogyakarta, 28 Januari 2026

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERSEMBAHAN.....	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMBANG DAN SINGKATAN	xvi
DAFTAR ISTILAH	xxi
INTISARI.....	xxvi
<i>ABSTRACT</i>	xxvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Studi Literatur	7
2.2 Dasar Teori.....	14
2.2.1 Chest X-ray Multi-label Multi-criteria.....	14
2.2.2 Deep Learning.....	15

2.2.3	Vision Transformer (ViT).....	16
2.2.4	Swin Transformer	18
2.2.5	Matriks Evaluasi	20
BAB III METODE PENELITIAN		26
3.1	Objek Penelitian.....	26
3.2	Alur Penelitian	27
3.2.1	Identifikasi Masalah.....	28
3.2.2	Studi Literatur	28
3.2.3	Pengumpulan Data	29
3.2.4	Exploratory Data Analysis (EDA).....	30
3.2.5	Label Selection (Pemilihan Label Penyakit).....	30
3.2.6	Pembersihan dan Pra-pemrosesan Label.....	31
3.2.7	Pembagian Data (60:20:20)	31
3.2.8	Pra-pemrosesan Citra & Augmentasi Data	32
3.2.9	Perancangan Model ViT dan Swin Transformer	32
3.2.10	Konfigurasi Pelatihan Baseline (Non Fine-tuning).....	33
3.2.11	Pelatihan Model (Non Fine-Tuning / Pre-Train)	34
3.2.12	Konfigurasi Pelatihan Fine-Tuning.....	34
3.2.13	Fine-Tuning.....	35
3.2.14	Evaluasi Model Baseline dan Fine-Tuning pada Data Uji.....	35
3.2.15	Visualisasi Hasil Evaluasi.....	36
3.2.16	Analisis Hasil dan Evaluasi Kinerja Model	37
3.2.17	Kesimpulan & Saran	39
3.2.18	Deployment.....	40
3.3	Alat dan Bahan.....	42

3.3.1	Data Penelitian	42
3.3.2	Alat dan Instrumen Penelitian.....	43
BAB IV HASIL DAN PEMBAHASAN		46
4.1	Pengumpulan Data	46
4.2	Exploratory Data Analysis (EDA).....	47
4.2.1	Distribusi Label Penyakit.....	47
4.2.2	Korelasi Antar Label.....	49
4.2.3	Visualisasi Contoh Citra X-ray Dada.....	51
4.3	Label Selection (Pemilihan Label Penyakit).....	52
4.4	Pembersihan dan Pra-pemrosesan Label.....	53
4.5	Pembagian Data (60:20:20)	56
4.6	Pra-pemrosesan Citra dan Augmentasi Data	58
4.6.1	Penyiapan Dataset (Dataset Class).....	61
4.6.2	Penyiapan DataLoader	63
4.7	Perancangan Model Vision Transformer (ViT) dan Swin Transformer.....	65
4.7.1	Arsitektur Vision Transformer (ViT).....	65
4.7.2	Arsitektur Swin Transformer	66
4.7.3	Penyesuaian Model untuk Klasifikasi Multi-Label	66
4.7.4	Implementasi Model Vision Transformer (ViT).....	66
4.7.5	Implementasi Model Swin Transformer	68
4.8	Konfigurasi Pelatihan Baseline (Non Fine-Tuning)	69
4.9	Pelatihan Model Baseline (Non Fine-Tuning / Pre-Train).....	72
4.9.1	Pelatihan Model Vision Transformer (ViT).....	72
4.9.2	Pelatihan Model Swin Transformer	80
4.10	Konfigurasi Pelatihan Fine-Tuning.....	89

4.11	Fine-Tuning.....	92
4.11.1	Fine-Tuning Model Vision Transformer (ViT)	93
4.11.2	Fine-Tuning Model Swin Transformer	100
4.12	Analisis Hasil dan Perbandingan Evaluasi Kinerja Model	106
4.12.1	Analisis Kinerja Model Vision Transformer (ViT) Baseline.....	107
4.12.2	Analisis Kinerja Model Swin Transformer Baseline	108
4.12.3	Analisis Kinerja Model Vision Transformer (ViT) Fine-Tuning	109
4.12.4	Analisis Kinerja Model Swin Transformer Fine-Tuning	110
4.12.5	Perbandingan Kinerja Model Baseline dan Fine-Tuning.....	112
4.12.6	Uji Signifikansi Statistik Peningkatan Kinerja Model	114
4.12.6.1	Uji Signifikansi Statistik Kinerja Model ViT.....	115
4.12.6.2	Uji Signifikansi Statistik Kinerja Model Swin Transformer	116
4.13	Deployment Model Terbaik	118
BAB V	PENUTUP.....	123
5.1	Kesimpulan	123
5.2	Saran.....	124
REFERENSI		125

DAFTAR TABEL

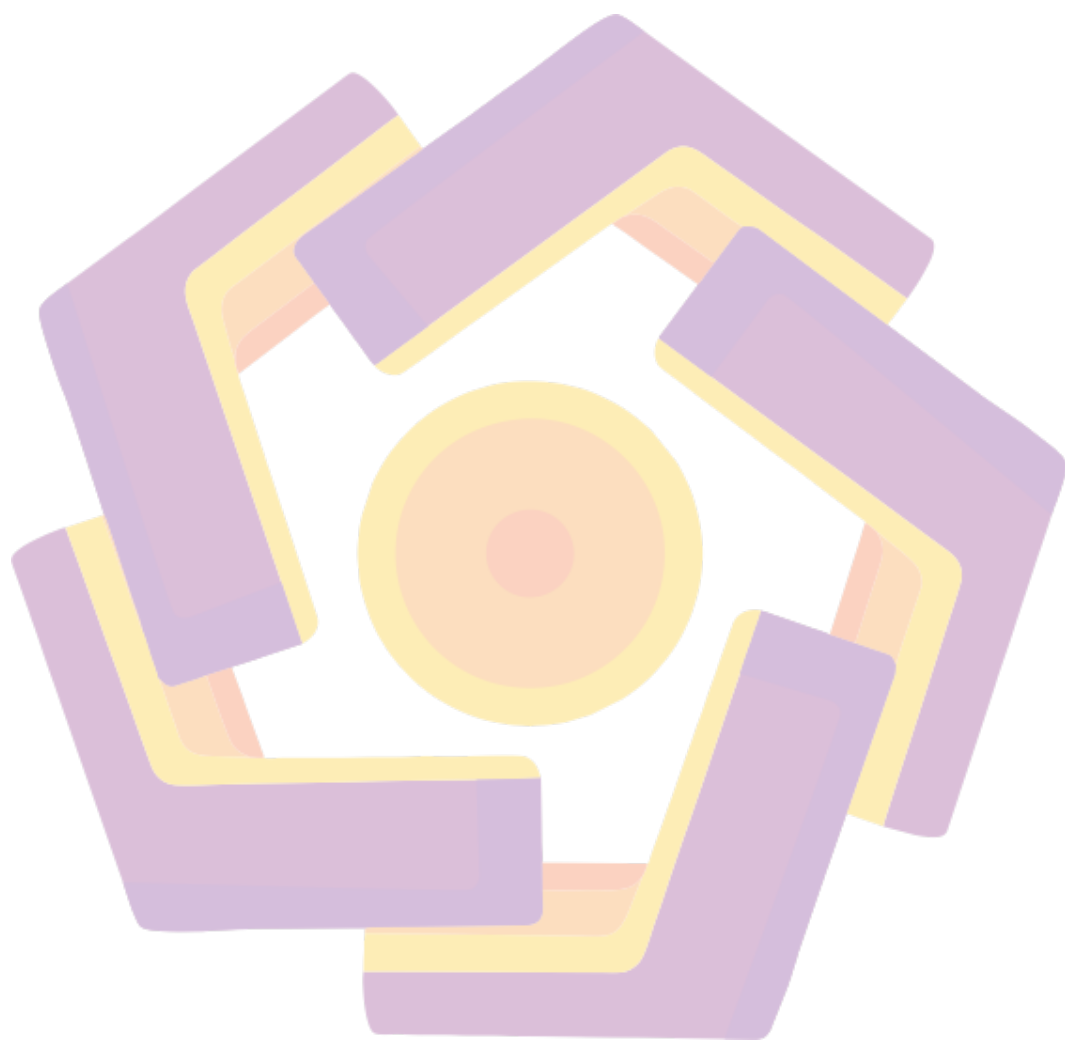
Tabel 2. 1 Keaslian Penelitian.....	10
Tabel 3. 1 Data Penelitian	42
Tabel 3. 2 Spesifikasi PC Server.....	43
Tabel 3. 3 Spesifikasi PC Sendiri.....	43
Tabel 3. 4 Daftar Perangkat Lunak dan Platform Penelitian	44
Tabel 4. 1 Jumlah Data Dataset CheXpert-v1.0-small.....	46
Tabel 4. 2 Distribusi Label Penyakit.....	48
Tabel 4. 3 Daftar Label Penyakit yang Digunakan.....	53
Tabel 4. 4 Ringkasan Hasil Pembersihan Label	55
Tabel 4. 5 Pembagian Data Dataset CheXpert.....	57
Tabel 4. 6 Konfigurasi Pra-pemrosesan, Augmentasi, dan Normalisasi Data Citra	60
Tabel 4. 7 Konfigurasi Pelatihan Model Baseline	71
Tabel 4. 8 Konfigurasi Pelatihan Model Fine-Tuning	91
Tabel 4. 9 Ringkasan Perbandingan Nilai AUC Model ViT dan Swin Transformer pada Data Uji	112

DAFTAR GAMBAR

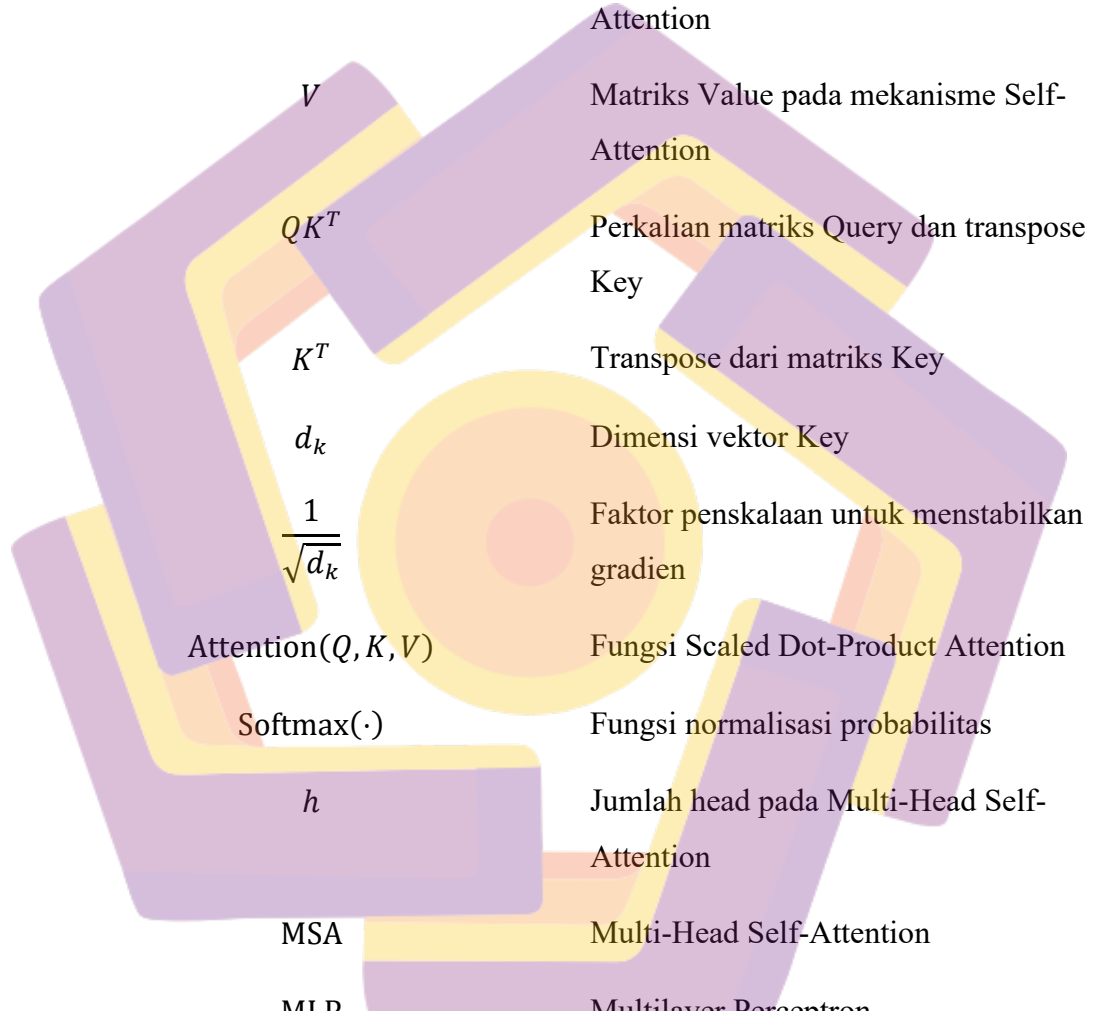
Gambar 2. 1 Contoh Chest X-ray dengan Multi-label Temuan Patologis [21]	
.....	Error! Bookmark not defined.
Gambar 2. 2 Arsitektur Dasar Vision Transformer [24]	Error! Bookmark not defined.
Gambar 2. 3 Arsitektur Swin Transformer [26]	Error! Bookmark not defined.
Gambar 2. 4 Kurva Receiver Operating Characteristic (ROC) dan Area Under the Curve (AUC) [28]	25
Gambar 3. 1 Alur Penelitian	27
Gambar 4. 1 Distribusi Label	48
Gambar 4. 2 Matriks Korelasi Antar Label Penyakit Paru	50
Gambar 4. 3 Contoh Citra X-ray Dada pada Dataset CheXpert	51
Gambar 4. 4 Daftar Label Penyakit Paru yang Digunakan dalam Penelitian	52
Gambar 4. 5 Implementasi Pembersihan dan Pra-pemrosesan Label Dataset CheXpert	54
Gambar 4. 6 Proses Pembagian Data Dataset CheXpert (Rasio 60:20:20)	56
Gambar 4. 7 Implementasi Pra-pemrosesan, Augmentasi, dan Normalisasi Data Citra.....	59
Gambar 4. 8 Implementasi Kelas Dataset untuk Pra-pemrosesan dan Augmentasi Data Citra Dada.....	62
Gambar 4. 9 Implementasi Penyiapan DataLoader untuk Data Latih, Validasi, dan Uji.....	64
Gambar 4. 10 Hasil Pemuatan Data	65
Gambar 4. 11 Penentuan Nama Model dan Model Checkpoint Vision Transformer (ViT).....	67
Gambar 4. 12 Implementasi Model Vision Transformer (ViT).....	68
Gambar 4. 13 Penentuan Nama Model dan Model Checkpoint Swin Transformer	68
Gambar 4. 14 Implementasi model Swin Transformer.....	69

Gambar 4. 15 Konfigurasi Pelatihan Baseline Model Vision Transformer dan Swin Transformer	71
Gambar 4. 16 Proses Eksekusi Pelatihan Model Vision Transformer (ViT).....	73
Gambar 4. 17 ringkasan proses training ViT selama 50 epoch (berhenti di epoch 10 karena early stopping).....	74
Gambar 4. 18 Kurva Pembelajaran (Learning Curve) Model Baseline Vision Transformer (ViT).....	75
Gambar 4. 19 Kurva ROC (Receiver Operating Characteristic) Model Baseline untuk 12 Label Penyakit Paru Vision Transformer (ViT)	76
Gambar 4. 20 Confusion Matrix Model Baseline Vision Transformer (ViT) untuk Setiap Label Penyakit Paru	77
Gambar 4. 21 Hasil Classification Report Model Baseline Vision Transformer (ViT) pada Data Uji	78
Gambar 4. 22 Visualisasi Grad-CAM dan Probabilitas Prediksi pada Citra X-ray Dada Vision Transformer (ViT)	79
Gambar 4. 23 Proses Eksekusi Pelatihan Model Vision Transformer (ViT).....	81
Gambar 4. 24 hasil training SWIN selama maksimal 50 epoch (berhenti di epoch 9 karena early stopping).....	82
Gambar 4. 25 Kurva Pembelajaran (Learning Curve) Model Baseline Swin Transformer.....	83
Gambar 4. 26 Kurva ROC (Receiver Operating Characteristic) Model Baseline untuk 12 Label Penyakit Paru Swin Transformer.....	84
Gambar 4. 27 Confusion Matrix Model Baseline Swin Transformer untuk Setiap Label Penyakit Paru	85
Gambar 4. 28 Hasil Classification Report Model Baseline Swin Transformer pada Data Uji	86
Gambar 4. 29 Visualisasi Grad-CAM dan Probabilitas Prediksi pada Citra X-ray Dada Swin Transformer	88
Gambar 4. 30 Konfigurasi Awal Pelatihan Fine-Tuning Model Vision Transformer dan Swin Transformer.....	90

Gambar 4. 31 Penyesuaian Learning Rate pada Tahap Fine-Tuning Model Vision Transformer dan Swin Transformer.....	90
Gambar 4. 32 Pemuatan Model Baseline sebagai Inisialisasi Fine-Tuning.....	92
Gambar 4. 33 Pengaktifan Kembali Lapisan Encoder pada Tahap Fine-Tuning .	93
Gambar 4. 34 Proses Eksekusi Fine-Tuning Model Vision Transformer (ViT)...	94
Gambar 4. 35 Output Proses Fine-Tuning Model Vision Transformer (ViT).....	94
Gambar 4. 36 Kurva Pembelajaran (Learning Curve) Model Fine-Tuning Vision Transformer (ViT).....	95
Gambar 4. 37 Kurva ROC (Receiver Operating Characteristic) Model Fine-Tuning untuk 12 Label Penyakit Paru Vision Transformer (ViT).....	96
Gambar 4. 38 Confusion Matrix Model Fine-Tuning Vision Transformer (ViT) untuk Setiap Label Penyakit Paru	97
Gambar 4. 39 Classification Report Per Label Penyakit Paru pada Model Fine-Tuning Vision Transformer (ViT)	99
Gambar 4. 40 Proses Eksekusi Fine-Tuning Model Swin Transformer	100
Gambar 4. 41 Output Proses Fine-Tuning Model Swin Transformer.....	101
Gambar 4. 42 Kurva Pembelajaran (Learning Curve) Model Fine-Tuning Swin Transformer.....	102
Gambar 4. 43 Kurva ROC (Receiver Operating Characteristic) Model Fine-Tuning untuk 12 Label Penyakit Paru Swin Transformer	103
Gambar 4. 44 Confusion Matrix Model Fine-Tuning Swin Transformer untuk Setiap Label Penyakit Paru	104
Gambar 4. 45 Classification Report Per Label Penyakit Paru pada Model Swin Transformer.....	105
Gambar 4. 46 Hasil Uji Wilcoxon Signed-Rank Test Model Vision Transformer (ViT) Menggunakan IBM SPSS Statistics.....	115
Gambar 4. 47 Hasil Uji Wilcoxon Signed-Rank Test Model Swin Transformer Menggunakan IBM SPSS Statistics.....	117
Gambar 4. 48 Tampilan Awal Sistem.....	119
Gambar 4. 49 Hasil Prediksi Citra Sehat / Tidak Terdeteksi Penyakit	120
Gambar 4. 50 Hasil Prediksi Citra dengan Penyakit Paru	121



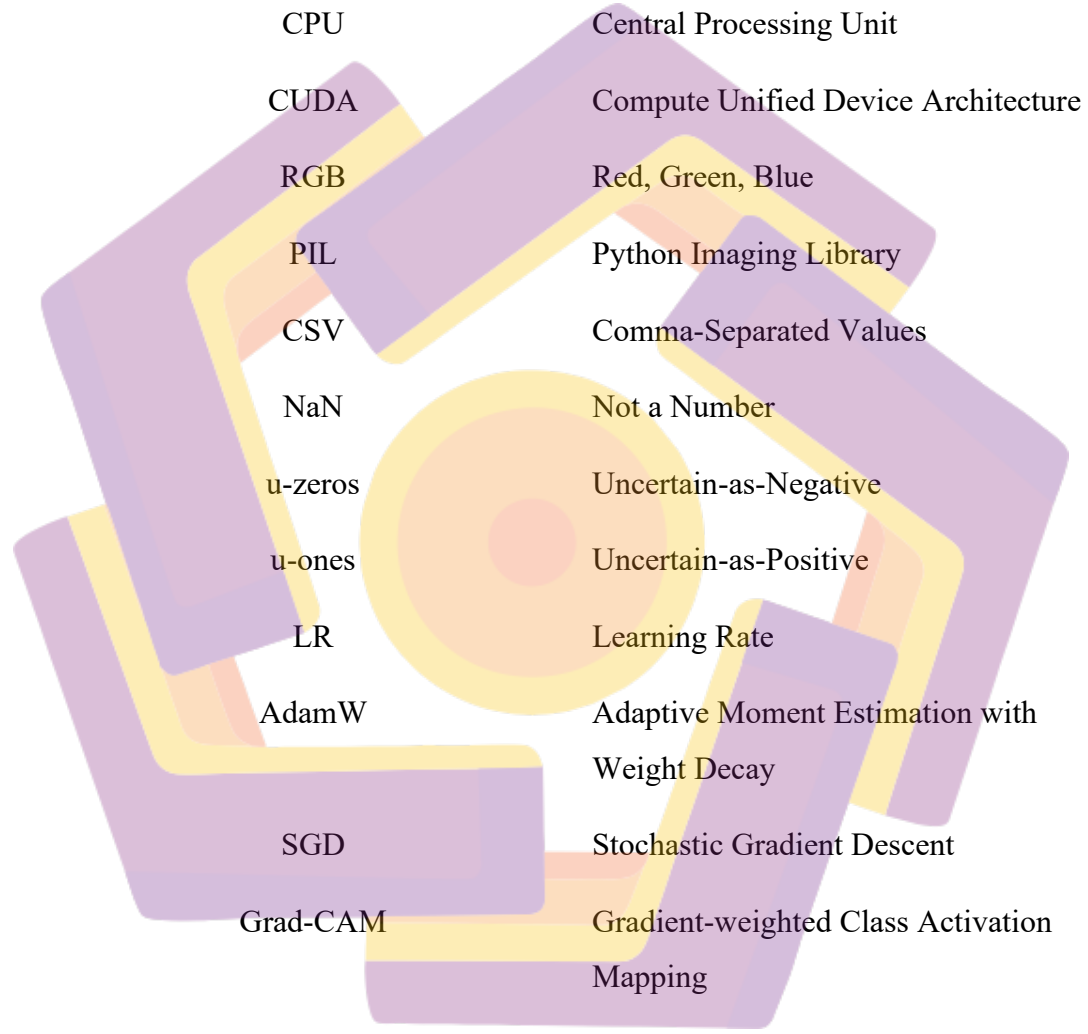
DAFTAR LAMBANG DAN SINGKATAN



Q	Matriks Query pada mekanisme Self-Attention
K	Matriks Key pada mekanisme Self-Attention
V	Matriks Value pada mekanisme Self-Attention
QK^T	Perkalian matriks Query dan transpose Key
K^T	Transpose dari matriks Key
d_k	Dimensi vektor Key
$\frac{1}{\sqrt{d_k}}$	Faktor penskalaan untuk menstabilkan gradien
$\text{Attention}(Q, K, V)$	Fungsi Scaled Dot-Product Attention
$\text{Softmax}(\cdot)$	Fungsi normalisasi probabilitas
h	Jumlah head pada Multi-Head Self-Attention
MSA	Multi-Head Self-Attention
MLP	Multilayer Perceptron
FFN	Feed-Forward Network
LN	Layer Normalization
Z^{l-1}	Representasi fitur (token) pada lapisan sebelumnya

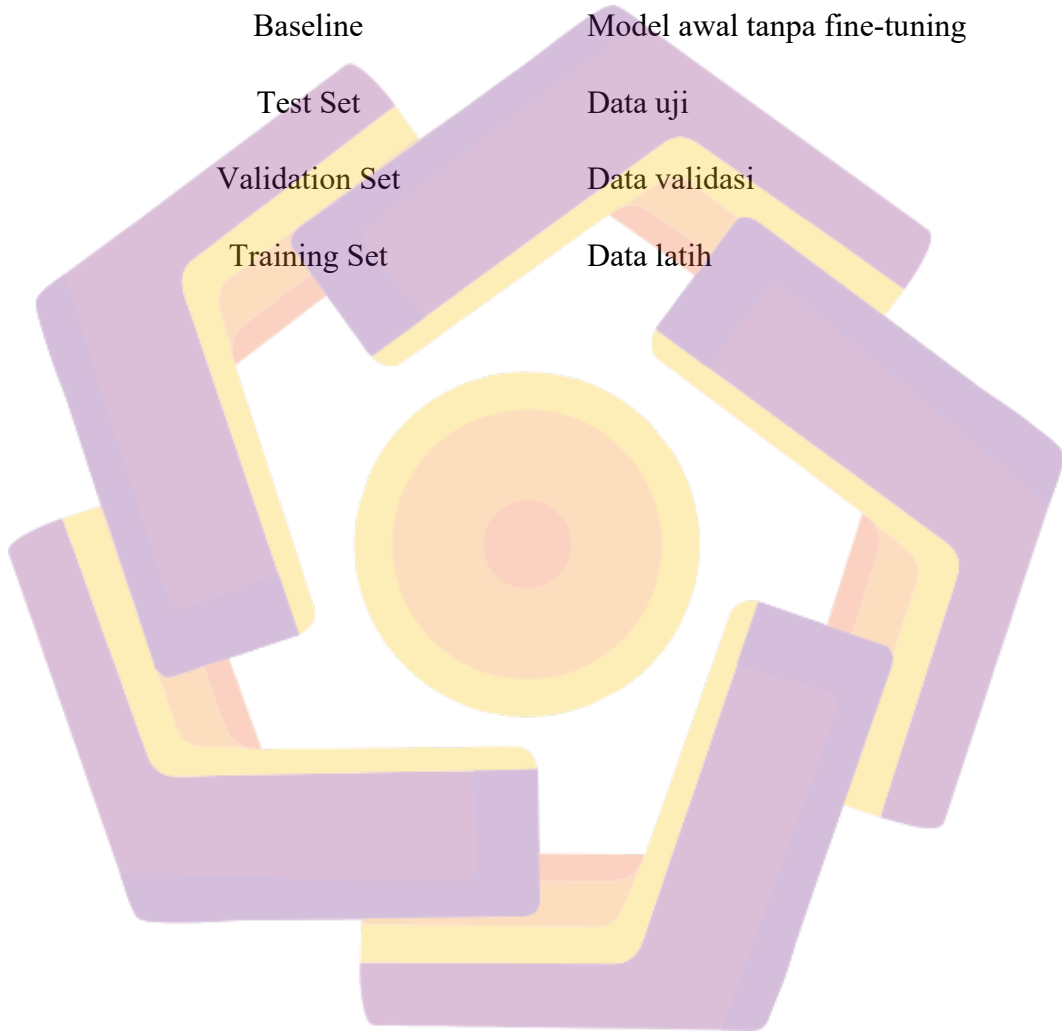
$\hat{Z}^l$	Output sementara dari sub-blok W-MSA
Z^l	Output akhir blok Swin Transformer lapisan ke- l
l	Indeks lapisan (layer)
$l + 1$	Lapisan setelah lapisan ke- l
W-MSA	Window-based Multi-Head Self-Attention
SW-MSA	Shifted Window Multi-Head Self-Attention
M	Ukuran window pada Swin Transformer
$\lfloor \frac{M}{2} \rfloor$	Besar pergeseran window
[CLS]	Token klasifikasi global pada Vision Transformer
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
TPR	True Positive Rate (Recall / Sensitivity)
FPR	False Positive Rate

Accuracy	Akurasi klasifikasi
AUC	Area Under the Curve
ROC	Receiver Operating Characteristic
TP(t)	True Positive Rate sebagai fungsi ambang batas
FP(t)	False Positive Rate sebagai fungsi ambang batas
$\int_0^1$	Operator integral
$\sum_i$	Operator penjumlahan diskrit
AUC	Area Under the Curve
ROC	Receiver Operating Characteristic
AUC-ROC	Area Under the Curve – Receiver Operating Characteristic
BCE	Binary Cross Entropy
BCEWithLogitsLoss	Binary Cross Entropy with Logits Loss
BCE Loss	Binary Cross Entropy Loss
CXR	Chest X-ray
EDA	Exploratory Data Analysis
ViT	Vision Transformer
Swin Transformer	Shifted Window Transformer
CNN	Convolutional Neural Network

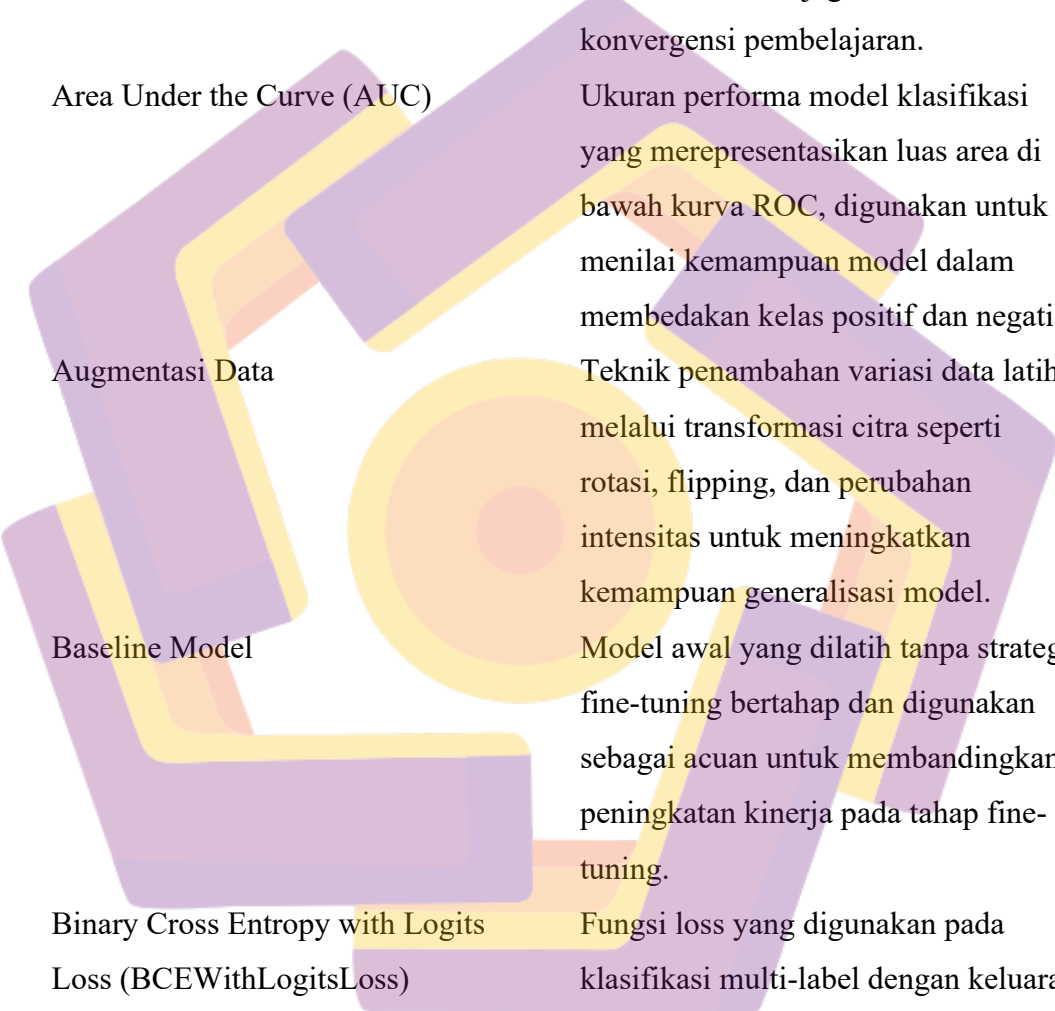


DL	Deep Learning
ML	Machine Learning
AI	Artificial Intelligence
GPU	Graphics Processing Unit
CPU	Central Processing Unit
CUDA	Compute Unified Device Architecture
RGB	Red, Green, Blue
PIL	Python Imaging Library
CSV	Comma-Separated Values
NaN	Not a Number
u-zeros	Uncertain-as-Negative
u-ones	Uncertain-as-Positive
LR	Learning Rate
AdamW	Adaptive Moment Estimation with Weight Decay
SGD	Stochastic Gradient Descent
Grad-CAM	Gradient-weighted Class Activation Mapping
F1-score	Harmonic mean dari precision dan recall
TPR	True Positive Rate
FPR	False Positive Rate
EPOCH	Satu siklus pelatihan penuh

LR Scheduler	Learning Rate Scheduler
AutoImageProcessor	Modul otomatis pemrosesan citra
Pre-Train	Bobot hasil pelatihan awal
Fine-Tuning	Pelatihan lanjutan model
Baseline	Model awal tanpa fine-tuning
Test Set	Data uji
Validation Set	Data validasi
Training Set	Data latih



DAFTAR ISTILAH



Adaptive Learning Rate	Mekanisme pengaturan laju pembelajaran yang menyesuaikan nilai learning rate selama proses pelatihan model untuk menjaga kestabilan dan konvergensi pembelajaran.
Area Under the Curve (AUC)	Ukuran performa model klasifikasi yang merepresentasikan luas area di bawah kurva ROC, digunakan untuk menilai kemampuan model dalam membedakan kelas positif dan negatif.
Augmentasi Data	Teknik penambahan variasi data latih melalui transformasi citra seperti rotasi, flipping, dan perubahan intensitas untuk meningkatkan kemampuan generalisasi model.
Baseline Model	Model awal yang dilatih tanpa strategi fine-tuning bertahap dan digunakan sebagai acuan untuk membandingkan peningkatan kinerja pada tahap fine-tuning.
Binary Cross Entropy with Logits Loss (BCEWithLogitsLoss)	Fungsi loss yang digunakan pada klasifikasi multi-label dengan keluaran berupa nilai logits, menggabungkan sigmoid activation dan binary cross entropy secara numerik stabil.
CheXpert Dataset	Dataset citra X-ray dada berskala besar yang menyediakan anotasi multi-label penyakit paru dan

Classification Report

digunakan sebagai sumber data pada penelitian ini.

Ringkasan metrik evaluasi klasifikasi yang mencakup precision, recall, F1-score, dan support untuk setiap label penyakit.

Cosine Annealing Learning Rate Scheduler

Metode penjadwalan learning rate yang menurunkan nilai learning rate secara bertahap mengikuti kurva kosinus selama pelatihan model.

Data Uji (Test Set)

Subset data yang digunakan untuk mengevaluasi kinerja akhir model dan tidak dilibatkan dalam proses pelatihan maupun validasi.

Data Validasi (Validation Set)

Subset data yang digunakan untuk memantau kinerja model selama pelatihan dan menentukan mekanisme early stopping.

Deep Learning

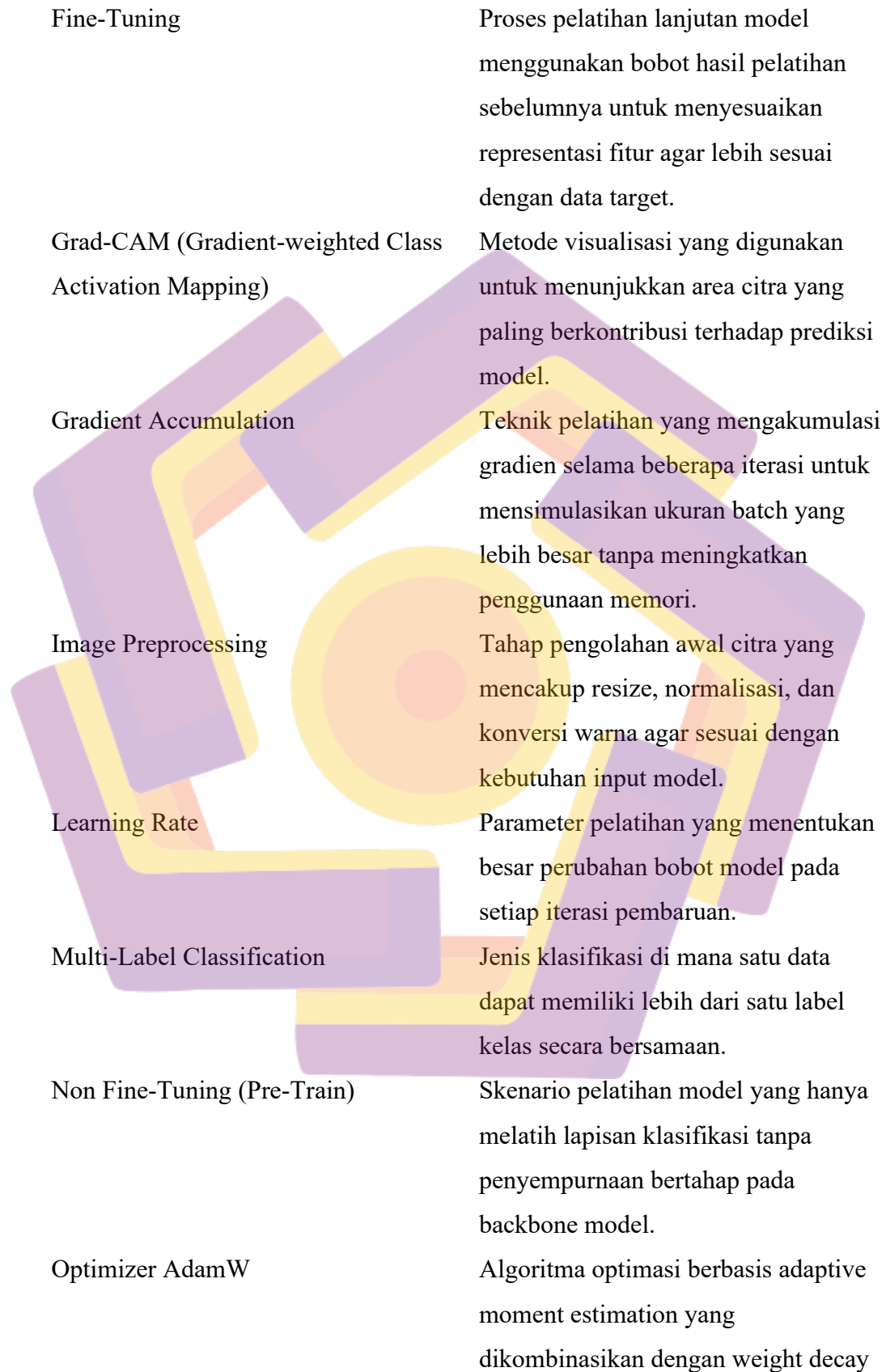
Pendekatan pembelajaran mesin berbasis jaringan saraf dalam dengan banyak lapisan untuk mempelajari representasi data secara hierarkis.

Early Stopping

Mekanisme penghentian pelatihan model secara otomatis ketika tidak terjadi perbaikan kinerja pada data validasi dalam beberapa epoch berturut-turut.

Exploratory Data Analysis (EDA)

Tahap analisis awal data untuk memahami distribusi, karakteristik, dan potensi permasalahan pada dataset sebelum proses pelatihan model.



Fine-Tuning	Proses pelatihan lanjutan model menggunakan bobot hasil pelatihan sebelumnya untuk menyesuaikan representasi fitur agar lebih sesuai dengan data target.
Grad-CAM (Gradient-weighted Class Activation Mapping)	Metode visualisasi yang digunakan untuk menunjukkan area citra yang paling berkontribusi terhadap prediksi model.
Gradient Accumulation	Teknik pelatihan yang mengakumulasi gradien selama beberapa iterasi untuk mensimulasikan ukuran batch yang lebih besar tanpa meningkatkan penggunaan memori.
Image Preprocessing	Tahap pengolahan awal citra yang mencakup resize, normalisasi, dan konversi warna agar sesuai dengan kebutuhan input model.
Learning Rate	Parameter pelatihan yang menentukan besar perubahan bobot model pada setiap iterasi pembaruan.
Multi-Label Classification	Jenis klasifikasi di mana satu data dapat memiliki lebih dari satu label kelas secara bersamaan.
Non Fine-Tuning (Pre-Train)	Skenario pelatihan model yang hanya melatih lapisan klasifikasi tanpa penyempurnaan bertahap pada backbone model.
Optimizer AdamW	Algoritma optimasi berbasis adaptive moment estimation yang dikombinasikan dengan weight decay

	untuk meningkatkan stabilitas pelatihan model. Proses pembagian citra menjadi potongan kecil (patch) yang kemudian direpresentasikan sebagai vektor masukan pada Vision Transformer.
Patch Embedding	
Pneumonia	Penyakit infeksi paru-paru yang menjadi fokus utama deteksi dalam penelitian ini.
Receiver Operating Characteristic (ROC)	Kurva evaluasi yang menggambarkan hubungan antara true positive rate dan false positive rate pada berbagai ambang klasifikasi.
Self-Attention	Mekanisme pada transformer yang memungkinkan model mempelajari hubungan antar bagian data secara global.
Swin Transformer	Arsitektur transformer berbasis citra yang menggunakan window-based self-attention untuk menangkap fitur lokal dan global secara hierarkis.
Training Loss	Nilai kesalahan model yang dihitung pada data latih selama proses pelatihan.
Uncertain-as-Negative (u-zeros)	Kebijakan penanganan label tidak pasti pada dataset CheXpert dengan menganggap label uncertain sebagai negatif.
Validation AUC	Nilai AUC yang dihitung menggunakan data validasi untuk

Vision Transformer (ViT)

memantau kinerja model selama pelatihan.

Arsitektur transformer yang memproses citra sebagai urutan patch untuk memodelkan hubungan global antar bagian citra.

Weight Decay

Teknik regularisasi yang digunakan untuk mengurangi kompleksitas model dan mencegah overfitting.



INTISARI

Penyakit paru-paru merupakan salah satu penyebab utama morbiditas dan mortalitas di dunia, sehingga deteksi dini memiliki peran penting dalam meningkatkan akurasi diagnosis dan efektivitas penanganan medis. Citra rontgen dada (Chest X-ray) merupakan modalitas pencitraan yang paling umum digunakan untuk mendeteksi berbagai penyakit paru, namun proses interpretasinya masih sangat bergantung pada keahlian radiolog dan berpotensi menimbulkan subjektivitas, terutama pada kasus dengan lebih dari satu kondisi patologis dalam satu citra. Kompleksitas ini dapat berdampak pada keterlambatan diagnosis dan ketidakkonsistenan hasil interpretasi. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan dan mengevaluasi pendekatan deteksi otomatis penyakit paru berbasis deep learning menggunakan arsitektur transformer, yaitu Vision Transformer (ViT) dan Swin Transformer. Penelitian dilakukan menggunakan dataset CheXpert-v1.0-small yang terdiri dari citra X-ray dada dengan 12 label penyakit paru yang relevan. Tahapan penelitian meliputi pembersihan dan pra-pemrosesan data, penanganan label tidak pasti, pembagian data menjadi data latih, validasi, dan uji, serta pelatihan model dalam dua skenario, yaitu baseline (non fine-tuning) dan fine-tuning. Evaluasi kinerja model dilakukan menggunakan metrik Area Under the Curve – Receiver Operating Characteristic (AUC-ROC) sebagai indikator akurasi klasifikasi multi-label. Hasil penelitian menunjukkan bahwa penerapan fine-tuning mampu meningkatkan performa kedua model secara konsisten, di mana Swin Transformer hasil fine-tuning menunjukkan kinerja paling stabil dengan nilai AUC berada pada rentang 0,75–0,88 pada sebagian besar label penyakit, termasuk pneumonia. Penelitian ini diharapkan dapat dimanfaatkan sebagai dasar pengembangan sistem pendukung keputusan klinis berbasis kecerdasan buatan untuk membantu proses diagnosis penyakit paru secara lebih cepat dan objektif.

Kata kunci: Pneumonia, Chest X-ray, Vision Transformer, Swin Transformer, Fine-Tuning

ABSTRACT

Lung diseases remain one of the leading causes of morbidity and mortality worldwide, making early detection essential to improve diagnostic accuracy and the effectiveness of medical treatment. Chest X-ray imaging is the most commonly used modality for detecting lung diseases; however, its interpretation still heavily depends on radiological expertise and is prone to subjectivity, particularly when multiple pathological conditions appear within a single image. This complexity may lead to delayed diagnosis and inconsistent interpretation results. Therefore, this study aims to develop and evaluate an automated lung disease detection approach based on deep learning using transformer-based architectures, namely Vision Transformer (ViT) and Swin Transformer. The study utilizes the CheXpert-v1.0-small dataset, which consists of chest X-ray images with 12 relevant lung disease labels. The research stages include data cleaning and preprocessing, handling uncertain labels, splitting the dataset into training, validation, and testing sets, and training the models under two scenarios: baseline (non fine-tuning) and fine-tuning. Model performance is evaluated using the Area Under the Curve – Receiver Operating Characteristic (AUC-ROC) metric as an indicator of multi-label classification accuracy. The experimental results show that fine-tuning consistently improves the performance of both models, with the fine-tuned Swin Transformer demonstrating the most stable performance, achieving AUC values in the range of 0.75–0.88 across most disease labels, including pneumonia. This research is expected to serve as a foundation for developing artificial intelligence-based clinical decision support systems to assist lung disease diagnosis in a faster and more objective manner.

Keyword: Pneumonia, Chest X-ray, Vision Transformer, Swin Transformer, Fine-Tuning.