

HALAMAN PENGESAHAN .....	ii
PERNYATAAN BEBAS PLAGIASI .....	iii
HALAMAN PERSEMBAHAN .....	iv
KATA PENGANTAR .....	v
DAFTAR ISI .....	vi
DAFTAR TABEL .....	ix
DAFTAR GAMBAR .....	x
DAFTAR SINGKATAN.....	xii
INTISARI.....	xiii
ABSTRACT .....	xiv
BAB I Pendahuluan .....	1
1.1 Latar Belakang .....	1
1.1.1 Perkembangan <i>Vision Foundation Models</i> yang Pesat .....	1
1.1.2 Pemrosesan Data menggunakan Pendekatan Multimodal .....	2
1.1.3 Standardisasi <i>Benchmarking Vision Foundation Models</i> pada Tugas <i>Zero-Shot Image Classification</i> .....	2
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian .....	4
1.4 Batasan Penelitian .....	4
1.5 Manfaat Penelitian .....	5
1.6 Sistematika Penulisan.....	5
BAB II Tinjauan Pustaka dan Dasar Teori .....	7
2.1 Tinjauan Pustaka .....	7
2.2 Dasar Teori .....	10
2.2.1 <i>Vision Foundation Models</i> (VFM).....	10
2.2.2 Zero-Shot Learning (ZSL).....	11
2.2.3 Konsep <i>Zero-Shot Image Classification</i> .....	12
2.2.4 <i>Convolutional Neural Networks</i> (CNN) .....	13
2.2.5 <i>Residual Networks</i> (ResNet).....	15
2.2.6 Transformer .....	16
2.2.7 CoNvNeXt.....	18
2.2.8 Large Language Models (LLM).....	19
2.2.9 <i>Prompting</i> .....	20
2.2.10 Vision Transformer (ViT).....	21
2.2.11 <i>Perception Encoder</i> .....	23
2.2.12 Contrastive Language-Image Pre-Training (CLIP).....	24

2.2.13	Sigmoid CLIP (SIGLIP) .....	26
2.2.14	<i>Baseline</i> Model .....	27
2.2.15	CIFAR100-Corrupted .....	27
2.2.16	<i>Animals with Attributes 2</i> (AwA2) .....	28
2.2.17	Oxford Pets .....	29
2.2.18	MIT-Indoor67 .....	30
2.2.19	Metrik <i>Top-1 Accuracy</i> dan <i>Top-5 Accuracy</i> .....	31
BAB III Metode Penelitian .....		33
3.1	Alat dan Bahan Tugas Akhir .....	33
3.1.1	Alat Tugas Akhir .....	33
3.1.2	Bahan Tugas akhir .....	34
3.2	Metode yang Digunakan .....	34
3.2.1	Pemilihan Dataset .....	34
3.2.2	Pemilihan Model CLIP .....	35
3.2.3	Pembuatan <i>Prompting</i> .....	36
3.2.4	Evaluasi Model .....	37
3.3	Alur Tugas Akhir .....	37
3.3.1	Identifikasi Masalah .....	38
3.3.2	Studi Literatur .....	38
3.3.3	Pemilihan Dataset .....	38
3.3.4	Pemilihan Model .....	39
3.3.5	Pengolahan Data .....	40
3.3.6	Pembuatan <i>Prompting</i> .....	40
3.3.7	Pengembangan <i>Baseline Model</i> .....	42
3.3.8	Evaluasi Model dan Perbandingan Hasil .....	42
BAB IV Hasil dan Pembahasan .....		44
4.1	<i>Setup Benchmarking Model CLIP</i> .....	44
4.2	Hasil Pembuatan <i>Prompting</i> .....	47
4.3	Hasil Eksperimen .....	49
4.3.1	Eksperimen 1: Metode <i>Prompting</i> “[kelas]” .....	49
4.3.2	Eksperimen 2: Metode <i>Prompting</i> “a photo of a [kelas]” .....	54
4.3.3	Eksperimen 3: Metode <i>Prompting</i> MPVR .....	57
4.3.4	Hasil Pengembangan <i>Baseline Model</i> .....	60
4.4	Analisis Perbandingan Performa Model .....	60
4.4.1	Performa <i>Top-1 Accuracy</i> pada Setiap Dataset .....	60
4.4.2	Performa <i>Top-5 Accuracy</i> pada Setiap Dataset .....	64
4.4.3	Performa Keseluruhan Model .....	67
4.5	Analisis Perbandingan Eksperimen dengan Uji Statistik .....	69
BAB V Kesimpulan dan Saran .....		73

5.1	Kesimpulan.....	73
5.2	Saran.....	74
	DAFTAR PUSTAKA.....	75
	LAMPIRAN .....	L-1
L.1	Source Code Penelitian .....	L-1
L.1.1	<i>Import Library</i> .....	L-1
L.1.2	Pemuatan Data dan Pengolahan Data .....	L-1
L.1.2.1	CIFAR100-Corrupted.....	L-1
L.1.2.2	Animals with Attributes 2 (AwA2) .....	L-5
L.1.2.3	MIT Indoor67 .....	L-6
L.1.2.4	Oxford Pets .....	L-7
L.1.3	Pemuatan Model CLIP.....	L-8
L.1.3.1	Perception Encoder Base/16 .....	L-8
L.1.3.2	CoNvNeXt Base .....	L-8
L.1.3.3	SIGLIP ViT-B/16 .....	L-8
L.1.3.4	SIGLIP2 ViT-B/16.....	L-8
L.1.3.5	OpenAI CLIP ViT-B/32 .....	L-9
L.1.3.6	OpenAI CLIP ViT-B/16 .....	L-9
L.1.4	Fungsi Evaluasi .....	L-9
L.1.4.1	Eksperimen 1 .....	L-9
L.1.4.2	Eksperimen 2 .....	L-14
L.1.4.3	Eksperimen 3 .....	L-18
L.1.5	Visualisasi Preprocessing Data .....	L-25
L.1.6	Visualisasi Hasil Prediksi .....	L-28
L.1.6.1	Eksperimen 1 .....	L-28
L.1.6.2	Eksperimen 2 .....	L-30
L.1.6.3	Eksperimen 3 .....	L-32