



DAFTAR ISI

DAFTAR ISI	ii
DAFTAR TABEL	vi
DAFTAR GAMBAR	vii
DAFTAR ALGORITMA	xii
INTISARI	xiii
ABSTRACT	xiv
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.6 Metodologi Penelitian	5
1.7 Sistematika Penulisan	7
II PENELITIAN TERKAIT	8
2.1 Visual Question Answering pada Domain General	8
2.2 Visual Question Answering pada Domain Medis	9
2.3 Dataset Medical VQA Terkait	11
2.3.1 VQA-RAD	11
2.3.2 SLAKE	12
2.4 Teknik Pra-pemrosesan Data Terkait	13
2.4.1 Pra-pemrosesan Citra	13
2.4.2 Pra-pemrosesan Teks	15
2.4.3 Pra-pemrosesan Jawaban	16
2.5 Arsitektur dan Model Terkait	17
2.5.1 Mekanisme <i>Attention</i> dan Arsitektur Transformer	17
2.5.2 Vision Transformer (ViT)	20



2.5.3	BERT	21
2.5.4	BioBERT	22
2.5.5	Layer-Residual Co-Attention Network (LRCN)	22
2.6	Metrik Evaluasi Terkait	30
2.6.1	Metrik Evaluasi VQA	30
2.6.2	Metrik Analisis Distribusi <i>Attention</i>	31
2.7	Rangkuman Penelitian Terkait dan Identifikasi Celah Penelitian	32
III METODE DAN RANCANGAN PENELITIAN		38
3.1	Deskripsi Penelitian	38
3.2	Akuisisi Data	40
3.2.1	VQA-RAD	40
3.2.2	SLAKE (English subset)	41
3.3	Pemisahan Data	42
3.4	Pra-pemrosesan Data	43
3.4.1	Pra-pemrosesan Citra	43
3.4.2	Pra-pemrosesan Pertanyaan	44
3.4.3	Pra-pemrosesan Jawaban	45
3.5	Pembangunan Model dan Percobaan <i>Hyperparameter</i>	47
3.5.1	<i>Image Encoder</i>	47
3.5.2	<i>Text Encoder</i>	48
3.5.3	<i>Co-Attention Mechanism</i>	48
3.5.4	<i>Feature Fusion</i> dan Klasifikasi Jawaban	49
3.5.5	Pengaturan <i>Hyperparameter</i>	49
3.6	Evaluasi Model	51
3.6.1	Evaluasi Kinerja Jawaban	51
3.6.2	Analisis Distribusi <i>Attention</i>	51
IV IMPLEMENTASI		53
4.1	Lingkungan Implementasi	53
4.2	Akuisisi Data	53
4.2.1	Struktur Direktori Dataset (Organized Raw)	53
4.2.2	Konfigurasi Pengunduhan Dataset	54
4.2.3	Arsitektur Modul Pengunduhan	55
4.2.4	Otomasi Pengunduhan Dataset	62
4.2.5	Akuisisi Dataset SLAKE	63



4.2.6	Akuisisi Dataset VQA-RAD	65
4.2.7	Replikasi Dataset ke Kaggle	67
4.3	Konfigurasi Eksperimen	67
4.3.1	Pengaturan <i>Hyperparameter</i> dan Ruang Eksperimen	69
4.4	Data <i>Loading</i>	70
4.5	Pemisahan Data	72
4.6	Pra-pemrosesan Data	75
4.6.1	Pra-pemrosesan Citra	76
4.6.2	Pra-pemrosesan Pertanyaan	77
4.6.3	Pra-pemrosesan Jawaban	78
4.6.4	Pembentukan <i>Dataset</i> dan <i>DataLoader</i>	79
4.7	Pembangunan Model	82
4.7.1	<i>Text Encoder</i>	83
4.7.2	<i>Visual Encoder</i>	84
4.7.3	Modul <i>Self-Attention</i> , <i>Guided-Attention</i> , dan LRM	85
4.7.4	Tumpukan <i>Co-Attention</i> untuk Teks dan Visual	87
4.7.5	<i>Attention Pooling</i> dan <i>Feature Fusion</i>	89
4.7.6	Model Utama LRCN	90
4.8	Pelatihan dan Validasi Model	91
4.8.1	Perhitungan Metrik Akurasi	92
4.8.2	Fungsi Evaluasi pada <i>validation/test</i>	93
4.8.3	Pelatihan Satu <i>Epoch</i>	94
4.8.4	<i>Training Loop</i> , <i>Checkpoint</i> , dan <i>Early Stopping</i>	95
4.8.5	Orkestrasi dan Eksekusi Eksperimen	98
4.9	Evaluasi	101
4.9.1	Prosedur Evaluasi pada Data <i>test</i>	101
4.9.2	Pelaporan Hasil Evaluasi	102
4.10	Visualisasi <i>Attention</i>	102
4.10.1	Ekstraksi Bobot <i>Attention</i> untuk Visualisasi	102
4.10.2	Transformasi Bobot menjadi Artefak Visualisasi	104
4.10.3	Alur Eksekusi Visualisasi <i>Attention</i>	105
4.11	Analisis Distribusi <i>Attention</i>	106
4.11.1	Perhitungan Metrik Distribusi <i>Attention</i>	106
4.11.2	Agregasi Metrik pada Data <i>test</i>	107

V	HASIL DAN PEMBAHASAN	109
5.1	Karakteristik Data dan Implikasi Evaluasi MedVQA	109
5.1.1	Komposisi Dataset dan Pembagian Data	109
5.1.2	Standarisasi Input Model	109
5.1.3	Ruang Jawaban dan Batasan Evaluasi Berbasis Klasifikasi	110
5.2	Validasi Pengaruh LRM pada Berbagai Kedalaman <i>Co-Attention</i>	111
5.2.1	Ablation pada Dataset SLAKE: Tren Δ terhadap L	111
5.2.2	Ablation pada Dataset VQA-RAD: Stabilitas Efek LRM	112
5.2.3	Pemilihan Konfigurasi <i>base</i> dengan $L = 12$ untuk Analisis <i>Layer-wise</i>	114
5.3	Analisis <i>Information Dispersion</i>	115
5.3.1	<i>Visual self-attention</i>	115
5.3.2	<i>Guided-attention</i>	117
VI	PENUTUP	120
6.1	Kesimpulan	120
6.2	Saran	121
A	Analisis Data dan Pra-pemrosesan	126
1.1	Ringkasan Dataset dan Pembagian Data	126
1.2	Pra-pemrosesan Citra	126
1.3	Pra-pemrosesan Pertanyaan dan Penetapan L_{\max}	127
1.4	Pra-pemrosesan Jawaban, Representasi Kelas, dan Cakupan Jawaban Uji	129
1.5	Ringkasan Batch DataLoader	129
B	Rincian Hasil Evaluasi dan Kurva Konvergensi pada Dataset SLAKE	131
C	Rincian Hasil Evaluasi dan Kurva Konvergensi pada Dataset VQA-RAD	135
D	Analisis Kualitatif Visualisasi <i>Attention</i> pada Dataset SLAKE	139
E	Analisis Kualitatif Visualisasi <i>Attention</i> pada Dataset VQA-RAD	143