

## DAFTAR ISI

<b>LEMBAR PENGESAHAN.....</b>	<b>i</b>
<b>PERNYATAAN BEBAS PLAGIASI.....</b>	<b>ii</b>
<b>INTISARI.....</b>	<b>iii</b>
<b>ABSTRACT.....</b>	<b>iv</b>
<b>DAFTAR ISI.....</b>	<b>v</b>
<b>DAFTAR GAMBAR.....</b>	<b>vi</b>
<b>DAFTAR TABEL.....</b>	<b>ix</b>
<b>BAB I.....</b>	<b>1</b>
<b>PENDAHULUAN.....</b>	<b>1</b>
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Tujuan Penelitian.....	4
1.4. Batasan Masalah.....	4
1.5. Manfaat Penelitian.....	4
1.6. Sistematika Penulisan.....	4
<b>BAB II.....</b>	<b>6</b>
<b>KAJIAN PUSTAKA.....</b>	<b>6</b>
<b>BAB III.....</b>	<b>12</b>
<b>LANDASAN TEORI.....</b>	<b>12</b>
3.1. Image Captioning.....	12
3.2. Transformer.....	15
3.2.1. Encoder.....	16
3.2.2. Decoder.....	17
3.2.3. Attention.....	18
3.2.4. Positional Encoding.....	21
3.2.4.1. Sinusoidal Positional Encoding.....	21
3.2.4.2. Learnable Sinusoidal Positional Encoding(LSPE).....	23
3.3. Convolutional Neural Network.....	25
3.4. InceptionV3.....	26
3.5. ResNet50.....	27
3.6. Matriks Evaluasi.....	29
3.6.1. BLEU.....	29
3.7. Beam Search.....	32
<b>BAB IV.....</b>	<b>33</b>
<b>METODE PENELITIAN.....</b>	<b>33</b>
4.1. Deskripsi Penelitian.....	33
4.2. Akuisisi Data.....	33
4.3. Arsitektur Model.....	35
4.3.1. Arsitektur Ekstraksi Fitur Gambar menggunakan InceptionV3.....	36

4.3.2	Arsitektur Transformer.....	37
4.3.2.1.	Transformer Decoder.....	37
4.3.2.2.	Learnable Sinusoidal Positional Encoding(LSPE).....	39
4.4.	Pra Pemrosesan Data.....	40
4.4.1.	Pra Pemrosesan Data Gambar.....	41
4.4.2.	Pra Pemrosesan Data Teks.....	41
4.5.	Pelatihan Model.....	43
4.5.1.	Optimizer.....	44
4.5.2.	Regularization.....	44
4.5.3.	Early Stopping.....	44
4.6.	Menghasilkan Caption.....	44
4.7.	Evaluasi Model.....	45
<b>BAB V.....</b>		<b>46</b>
<b>IMPLEMENTASI.....</b>		<b>46</b>
5.1.	Environment Implementasi.....	46
5.2.	Akuisisi Data.....	46
5.3.	Pra Pemrosesan Data.....	47
5.3.1.	Pembuatan Dataframe.....	47
5.3.2.	Pemisahan Data.....	47
5.3.3.	Memformat Image Path Gambar.....	49
5.3.4.	Pra Pemrosesan Data Teks.....	50
5.3.4	Pra Pemrosesan Data Gambar.....	51
5.4.	Pembuatan Tokenizer.....	52
5.6.	Pembuatan Model.....	54
5.6.1.	Ekstraksi Fitur Menggunakan CNN.....	54
5.6.2.	Embeddings dan Positional Encoding.....	54
5.6.2.1.	Fixed Sinusoidal Embeddings.....	54
5.6.2.2.	Learnable Sinusoidal Embeddings untuk LSPE.....	56
5.6.3.	Lapisan Pointwise Feed Forward.....	58
5.6.4.	Lapisan Reshape untuk Output Encoder.....	58
5.6.5.	Decoder Transformer.....	59
5.6.6.	Model Image Captioning.....	62
5.7.	Pelatihan Model.....	63
5.8.	Generate Caption dengan Beam Search.....	68
5.9.	Evaluasi Model.....	73
<b>BAB VI.....</b>		<b>76</b>
<b>HASIL DAN PEMBAHASAN.....</b>		<b>76</b>
6.1.	Pengumpulan Data.....	76
6.2.	Pra Pemrosesan Data.....	78
6.2.1.	Pra Pemrosesan Data Teks.....	78
6.2.2.	Prapemrosesan Gambar.....	80
6.3.	Hasil Hyperparameter Tuning Model.....	86
6.3.1.	Hasil Hyperparameter Tuning Model Menggunakan LSPE dan Fix Sinusoidal	

dengan Beam Width 1.....	86
6.3.2. Hasil Hyperparameter Tuning Model Menggunakan LSPE dan Fix Sinusoidal dengan Beam Width 3.....	87
6.3.3. Hasil Hyperparameter Tuning Model Menggunakan LSPE dan Fix Sinusoidal dengan Beam Width 5.....	87
6.3.4. Hasil Hyperparameter Tuning Model Menggunakan LSPE dan Fix Sinusoidal dengan Beam Width 7.....	88
6.3.5. Hasil Terbaik Hyperparameter Tuning Model Menggunakan LSPE dan Fix Sinusoidal pada Setiap Konfigurasi.....	89
6.4. Hasil Evaluasi Model pada Data Uji.....	90
6.5. Hasil Inferensi Model.....	90
6.6. Jumlah Parameter Terhadap Probabilitas Beam Search.....	94
6.7. Analisis Perhitungan dan Hasil Skor BLEU.....	95
6.8. Analisis Dataset Flickr8k Bahasa.....	96
<b>BAB VII.....</b>	<b>99</b>
<b>PENUTUP.....</b>	<b>99</b>
7.1. Kesimpulan.....	99
7.2. Saran.....	99
<b>DAFTAR PUSTAKA.....</b>	<b>101</b>