

Image Captioning Menggunakan Metode Resnet50 Dan Long Short Term Memory

Marius Raka Satria¹ Jasman Pardede²

¹ Program Studi Informatika, Institut Teknologi Nasional

² Program Studi Informatika, Institut Teknologi Nasional

Email : rakasatria1999@mhs.itenas.ac.id

Received DD MM YYYY | *Revised* DD MM YYYY | *Accepted* DD MM YYYY

ABSTRAK

Kesalahpahaman manusia dalam mencari makna arti dari sebuah gambar menimbulkan kebingungan. Hanya karena struktur kalimat atau penggunaan kata bermakna makna lebih dari satu yang biasa disebut ambiguitas. Ambiguitas terjadi apabila arti dari kata, frasa, atau kalimat tidak pasti, maknanya lebih dari satu. Karena adanya keterkaitan dengan kecerdasan buatan dalam membantu klasifikasi gambar untuk menghindari ambiguitas, penggunaan Image Captioning dimanfaatkan pada penelitian ini. Image Captioning menghasilkan deskripsi berbahasa alami. Permasalahan yang muncul dapat diselesaikan dengan penggabungan antara kecerdasan buatan dan jaringan syaraf tiruan. Kedua metode yang digunakan dalam penelitian ini adalah Resnet50 dan Long Short Term Memory. Resnet50 berfungsi untuk klasifikasi gambar dan LSTM jaringan syaraf tiruan untuk generate caption. Penelitian ini menggunakan BLEU scoring satu gram untuk memberi nilai pada caption yang telah dibuat. Score BLEU tertinggi adalah 79,7455% dan akurasi tertinggi yang didapat adalah 85,74% pada 100 epoch.

Kata kunci: Ambiguitas, LSTM, ResNet50, Image Captioning

ABSTRACT

Human misunderstanding in finding the meaning of an image causes confusion. Just because the structure of the sentence or the use of words means more than one meaning which is usually called ambiguity. Ambiguity occurs when the meaning of a word, phrase, or sentence is uncertain, it has more than one meaning. Because of the connection with artificial intelligence in helping image classification to avoid ambiguity, the use of Image Captioning is utilized in this study. Image Captioning produces natural language descriptions. The problems that arise can be solved by combining artificial intelligence and artificial neural networks. The two methods used in this research are Resnet50 and Long Short Term Memory. Resnet50 functions for image classification and artificial neural network LSTM to generate captions. This study uses a one gram BLEU scoring to give a value to the caption that has been made. The highest BLEU score was 79.7455% and the highest accuracy obtained was 85.74% at 100 epochs.

Keywords: Ambiguity, LSTM, ResNet50, Image Captioning

1. PENDAHULUAN

Manusia ketika melihat sebuah gambar dapat terjadi kesalahpahaman dalam mencari makna yang menimbulkan kebingungan untuk menafsirkan arti dari gambar tersebut. Hal ini terjadi disebabkan karena struktur kalimat maupun karena penggunaan katanya memiliki makna yang lebih dari satu atau yang biasa disebut ambiguitas. Ambiguitas adalah saat arti dari kata, frasa, atau kalimat tidak pasti, maknanya bisa lebih dari satu. Ketika ada pernyataan yang ambigu, tentu akan membingungkan pembaca dan menghalangi makna teks. Karena arti ambigu biasanya muncul sebagai kalimat dengan makna ganda atau lebih. Arti ambigu ini berkaitan dengan perbedaan penafsiran teks, yang menyebabkan ketidakjelasan atau kebingungan. Lantas, makna ambigu sering dicantumkan dalam sastra, ide, pernyataan, hingga gambar. Bahkan kata ambigu tak lepas dalam memberikan kedalaman dan kompleksitas suatu karya atau gambar. **(PUTRI et al., 2019)**

Berdasarkan kondisi tersebut berkaitan dengan visi komputer dalam membantu klasifikasi gambar untuk menghindari ambiguitas. Dalam beberapa tahun terakhir, visi komputer di bidang pemrosesan gambar telah membuat kemajuan yang signifikan, seperti klasifikasi gambar dan deteksi objek. Manfaat dari kemajuan pada bidang klasifikasi gambar dan deteksi objek menjadi memungkinkan untuk secara otomatis menghasilkan satu kalimat atau lebih untuk menjelaskan konten visual dari suatu gambar, yang dikenal sebagai *Image Captioning*. Membuat deskripsi gambar yang lengkap dan alami secara otomatis memiliki manfaat yang besar, seperti padapembuatan judul yang dilampirkan pada gambar berita, deskripsi yang terkait dengan gambar medis, pengambilan gambar berbasis teks, informasi gambar, interaksi manusia-robot. Dengan mendapat informasi berupa deskripsi teks dari konten visual seperti gambar, maka informasi yang diperoleh akan lebih mudah diolah **(Dubey & N, 2020)**.

Menghasilkan deskripsi dengan bahasa alami yang memiliki makna dari sebuah gambar membutuhkan tingkat pemahaman yang lebih tinggi dari klasifikasi dan deteksi gambar. Permasalahan tersebut sangat menarik karena adanya Kecerdasan Buatan. Penelitian ini, akan menggunakan dua metode yaitu RESNET50 dan *Long Short Term Memory* (LSTM). RESNET50 berfungsi untuk klasifikasi gambar seperti *Convolutional Neural Network* (CNN) dengan jaringan-jaringan yang lebih mendalam LSTM yang merupakan sebuah jaringan syaraf berulang akan digunakan untuk *generate caption* **(Hochreiter & Schmidhuber, 1997)**.

ResNet adalah jenis deep network berbasis pembelajaran residual. Pembelajaran semacam ini dapat memfasilitasi pelatihan jaringan dengan mempertimbangkan input layer sebagai referensi (Faiyaz Khan et al., 2021). ResNet-50 adalah salah satu varian ResNet yang memiliki 50 layer. Jika pada varian ResNet sebelumnya dilakukan skip connection sebanyak 2 layer, maka ResNet-50 melewati 3 layer dan terdapat 1 X 1 convolution layer **(Miranda, Novamizanti, & Rizal, 2020)**.

Long Short Term Memory Neural Network (LSTM) merupakan salah satu jenis RNN. LSTM menyimpan informasi terhadap pola-pola pada data. LSTM dapat mempelajari data mana saja yang akan disimpan dan data mana saja yang akan dibuang, karena pada setiap neuron LSTM memiliki beberapa gates yang mengatur memori pada setiap neuron itu sendiri. LSTM banyak digunakan untuk pemrosesan teks, video, dan data time series **(Wildan, Aldi, & Aditsania, 2018)**.

Dalam menguji seberapa baik deskripsi yang dihasilkan dari hasil caption ini akan dievaluasi dengan skor BLEU. BLEU (Bilingual Evaluation Understudy) adalah pengukuran perbedaan antara terjemahan otomatis dan satu atau beberapa terjemahan referensi yang dibuat manusia dari kalimat sumber yang sama **(Callison-Burch, Osborne, & Koehn, 2006)**. Sesuai dengan visi komputer dalam mempermudah dalam klasifikasi gambar, maka sangat baik menjadi solusi untuk membantu menghindari ambiguitas dalam mengklasifikasi gambar menggunakan bantuan komputer. Penggunaan metode ResNet50 dan LSTM ini merupakan salah satu langkah yang diambil untuk menjadi solusi dalam membantu menghindari ambiguitas pada klasifikasi gambar. Deskripsi gambar yang muncul dalam bahasa Inggris

tertata senatural mungkin dan selanjutnya di evaluasi menggunakan score BLEU. Maka dengan itu penulis memiliki judul penelitian "IMAGE CAPTIONING MENGGUNAKAN METODE RESNET50 DAN LONG SHORT TERM MEMORY" yang dapat dijadikan sebagai acuan penelitian selanjutnya yaitu mengenai program komputer yang dapat membuat hasil score yang optimal.

2. TINJAUAN TEORITIS

2.1 Residual Network 50 (ResNet50)

Residual Neural Network atau yang dikenal sebagai ResNet merupakan sebuah arsitektur CNN yang telah memenangkan ILSVRC (*Image Large Scale Visual Recognition Competition*) pada tahun 2015 dan mendapat error rate mencapai 3,57%. ResNet adalah jenis deep network berbasis pembelajaran residual. Pembelajaran semacam ini dapat memfasilitasi pelatihan jaringan dengan mempertimbangkan input layer sebagai referensi (Zhou, Nejadi, Do, Cheung, & Cheah, 2016). ResNet mengalahkan VGG sebagai posisi runner up, walaupun ResNet memiliki layer yang lebih banyak dari VGG, namun ResNet memiliki tingkat kompleksitas yang lebih rendah dari VGG. ResNet-50 adalah salah satu varian ResNet yang memiliki 50 layer.

2.2 Text Pre-Processing

Berdasarkan ketidak teraturan struktur data teks, maka proses sistem temu kembali informasi ataupun *text mining* memerlukan beberapa tahap awal yang pada intinya adalah mempersiapkan agar teks dapat diubah menjadi lebih terstruktur. Tahap *Text Preprocessing* adalah tahapan dimana aplikasi melakukan seleksi data yang akan diproses pada setiap dokumen. Proses *preprocessing* ini meliputi (1) *case folding*, (2) *tokenizing*, (3) *filtering*, dan (4) *stemming* (Hermawan & Ismiati, 2020).

2.3 LSTM (Long Short Term Memory)

Long Short Term Memory networks (LSTM) merupakan sebuah evolusi dari arsitektur RNN, dimana pertama kali diperkenalkan oleh Hochreiter & Schmidhuber (1997). Hingga penelitian ini dilakukan banyak para peneliti yang terus mengembangkan arsitektur LSTM di berbagai bidang seperti dalam bidang *speech recognition* dan *forecasting* (Wildan, Aldi, & Aditsania, 2018).

2.4 BLEU (BILINGUAL EVALUATION UNDERSTUDY)

Merupakan sebuah metode yang digunakan untuk mengevaluasi otomatis terjemahan mesin . Cara perhitungan skor BLEU adalah sebagai berikut:

$$P_n = \frac{\sum c \text{ecorpus } n - \text{gram } \epsilon c \sum \text{count}_{clip}(n - \text{gram})}{\sum c \text{ecorpus } n - \text{gram } \epsilon c \sum \text{count}(n - \text{gram})} \quad (1)$$

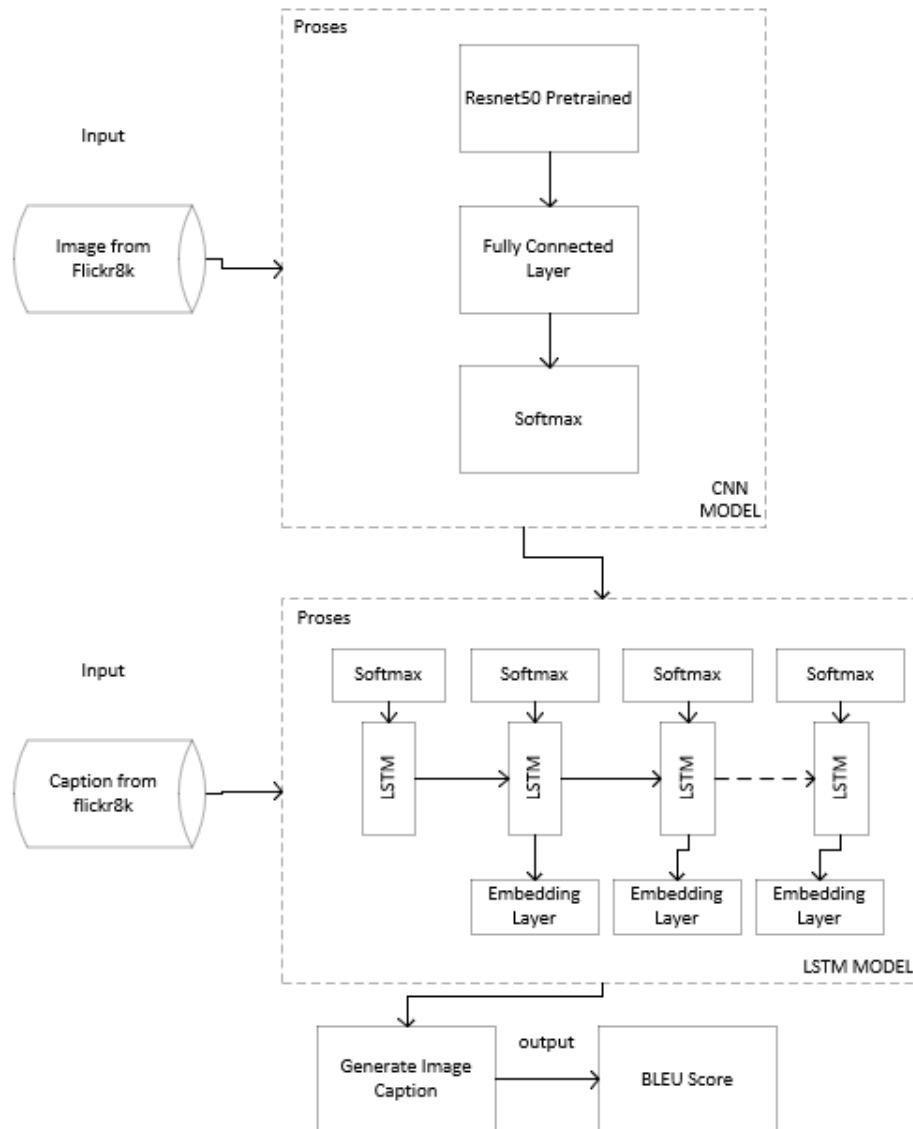
$$BP_{BLEU} = f(x) = \begin{cases} 1, & \text{if } c > r \\ 3^{(1-\frac{r}{c})}, & \text{if } c \leq r \end{cases} \quad (2)$$

$$BLEU = BP_{BLEU} \cdot e \sum_{n=1}^N w_n \log P_n \quad (3)$$

Dengan keterangan *BP* adalah *brevity penalty*. *c* adalah jumlah kata dari hasil terjemahan otomatis. *r* adalah jumlah kata rujukan. P_n adalah modified *precision score*. w_n adalah $\frac{1}{N}$ (standar nilai N untuk BLEU adalah 4). P_n adalah jumlah n-gram hasil terjemahan yang sesuai dengan rujukan dibagi jumlah n- gram hasil terjemahan.

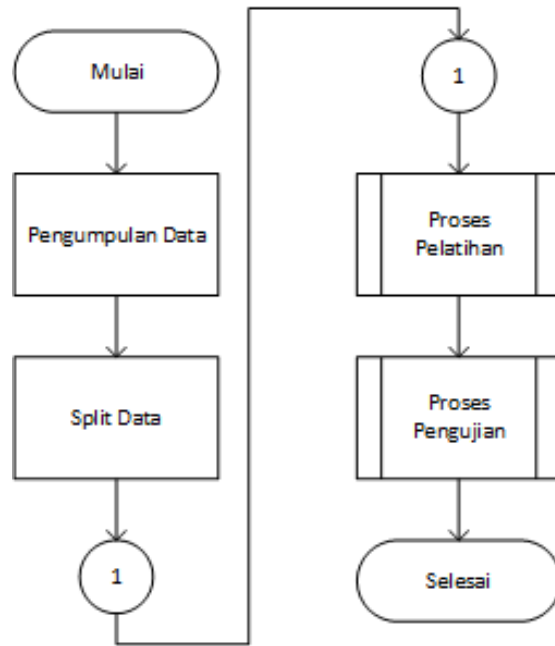
3. METODE PENELITIAN

Pada penelitian ini, metode penelitian membahas mengenai metode pengumpulan data dan metode simulasi lengkap dengan blok diagram keseluruhan dari penelitian yang dilakukan, *flowchart* program, blok diagram, dan studi kasus dari penelitian ini.



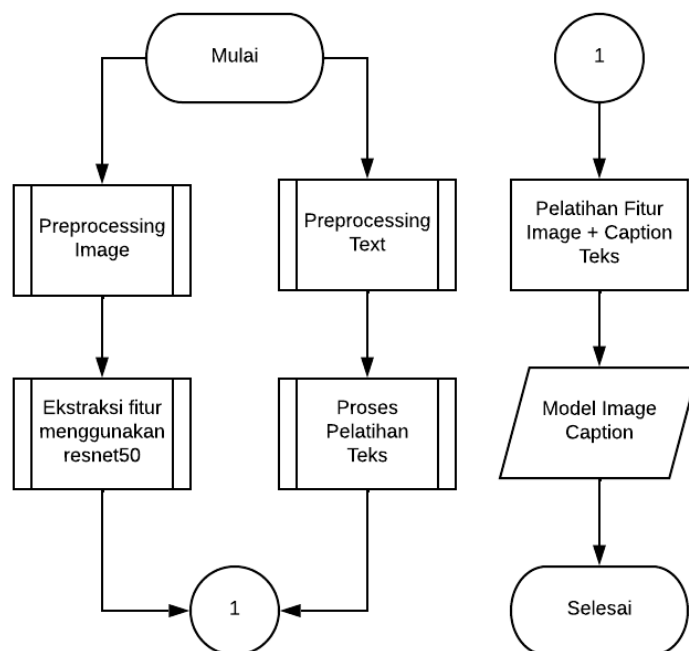
Gambar 1. Blok Diagram Penelitian

sistem mulai dari input gambar, CNN Model, kemudian *input caption* lalu LSTM Model yang selanjutnya fungsi *generator* dari gambar dan kemudian prediksinya melihat nilai BLEU dan menjadi *Output* sistem yang dianalisis. Tahapan-tahapan penjelasan dari sebuah adalah studi kasus, dimana studi kasus dalam penelitian ini dimulai dari memasukkan data gambar dari *drive*, lalu *me-running* ResNet 50 dimana data gambar diambil dari Flickr 8K, RESNET50 yang mensimulasi berfungsi untuk klasifikasi 8000 gambar dengan metode CNN method, dan lengkap juga dengan disediakan *caption* dari Flickr 8K *dataset* yang kemudian diolah dengan LSTM model, LSTM yang merupakan sebuah jaringan syaraf berulang akan digunakan untuk *generate caption* yang selanjutnya akan dihitung nilai *score* BLEU yang akan dianalisis *Output*-nya berdasarkan *Epoch* yang beragam pada gambar yang berbeda-beda maka akan menghasilkan nilai BLEU yang beragam.



Gambar 2. Flowchart Program

Pada *encoder output*, *hidden state* diinisialisasi menjadi 0. Pada *processing training* dilakukan sebanyak 100 *epoch*. Pada tahap ini yaitu memilihkan kosakata yang terdapat pada gambar yang telah disesuaikan visualisasinya dengan *caption*. Penelitian ini, menggunakan dua metode yaitu ResNet50 dan LSTM. Metode ResNet50 digunakan untuk *generate feature* pada gambar, sedangkan LSTM digunakan untuk melakukan *generate caption*. Setelah melakukan *training*, selanjutnya adalah melakukan *generate caption* berdasar model yang telah dibuat. Tahapan ini dilakukan untuk membandingkan hasil prediksi *caption* dibandingkan dengan *caption* yang sebenarnya dari gambar. Pada analisis *output* dilakukan analisis skoring dengan BLEU (*Bilingual Evaluation Understudy*). Agar dapat mengukur seberapa tepat *caption* yang di-*generate* oleh *computer*.



Gambar 3. Proses Pelatihan

Proses pelatihan diatas menggambarkan fungsi-fungsi apa saja yang ada pada sistem mulai dari *input* gambar, CNN Model, kemudian *input caption* lalu LSTM Model yang selanjutnya fungsi *generator* dari gambar dan kemudian prediksinya melihat nilai BLEU dan menjadi *Output* sistem yang dianalisis. Tahapan-tahapan penjelasan dari sebuah proses adalah studi kasus, dimana studi kasus dalam penelitian ini dimulai dari memasukkan data gambar dari *drive*, lalu *me-running* ResNet 50 dimana data gambar diambil dari Flickr 8K, RESNET50 yang mensimulasi berfungsi untuk klasifikasi 8000 gambar dengan metode CNN method, dan lengkap juga dengan disediakan *caption* dari Flickr 8K dataset yang kemudian diolah dengan LSTM model, LSTM yang merupakan sebuah jaringan syaraf berulang akan digunakan untuk *generate caption* yang selanjutnya akan dihitung nilai score BLEU yang akan dianalisis *Output*-nya berdasarkan *Epoch* yang beragam pada gambar yang berbeda-beda maka akan menghasilkan nilai BLEU yang beragam

4. HASIL PENELITIAN DAN PEMBAHASAN

Pada step pemrosesan dan pengujian terdapat beberapa langkah diantaranya bermula dari *image processing*, *teks preprocessing*, visual gambar dengan *caption*, Membuat kosakata, membangun fungsi generator, membuat model dan terakhir adalah prediksi.

4.1 Image Preprocessing

Proses pada tahapan *image processing* berisi tentang memilih folder yang berisi dataset yang ada kemudian memilih salah satu foto untuk diproses, lalu memasukkan jenis CNN yaitu ResNet-50 yang merupakan salah satu varian *Residual Networking* yang memiliki 50 layer, dan merupakan salah satu aplikasi Keras, maka setelah model di-*import* akan muncul hasil nilai RELu, *Batch Normalization*, dan Conv2D seperti gambar dibawah.

```

avg_pool (GlobalAveragePooling (None, 2048) 0 ['conv5_block3_out[0][0]']
2D)

-----
Total params: 23,587,712
Trainable params: 23,534,592
Non-trainable params: 53,120

```

Gambar 4. Image Processing

Dimana ditahap *Image Preprocessing* gambar menggunakan ResNet50 untuk melakukan ekstraksi fitur setiap gambar dengan mengubah ukuran menjadi $224 \text{ pixels} * 224 \text{ pixels}$ dan menormalkan gambar sehingga berisi piksel dalam kisaran -1 hingga 1, yang sesuai dengan format gambar yang digunakan untuk melatih ResNet50.

4.2 Teks Processing

Pada proses teks *preprocessing* yaitu dimana proses bermula dari memasukkan *file* yang berisi *caption* kemudian melihat berapa banyak *caption* yang ada pada *file* tersebut.

```
def preprocessed(txt):  
    modified = txt.lower()  
    modified = 'startofseq ' + modified + ' endofseq'  
    return modified  
  
for k,v in captions_dict.items():  
    for vv in v:  
        captions_dict[k][v.index(vv)] = preprocessed(vv)
```

Gambar 5. Teks *Preprocessing*

Pada teks berjumlah 40.000 Kalimat dikarenakan adanya lima kalimat per gambar. Dimana dalam dataset terdapat 8000 gambar, maka $8000 * 5$ Kalimat

4.3 Membangun Fungsi Generator

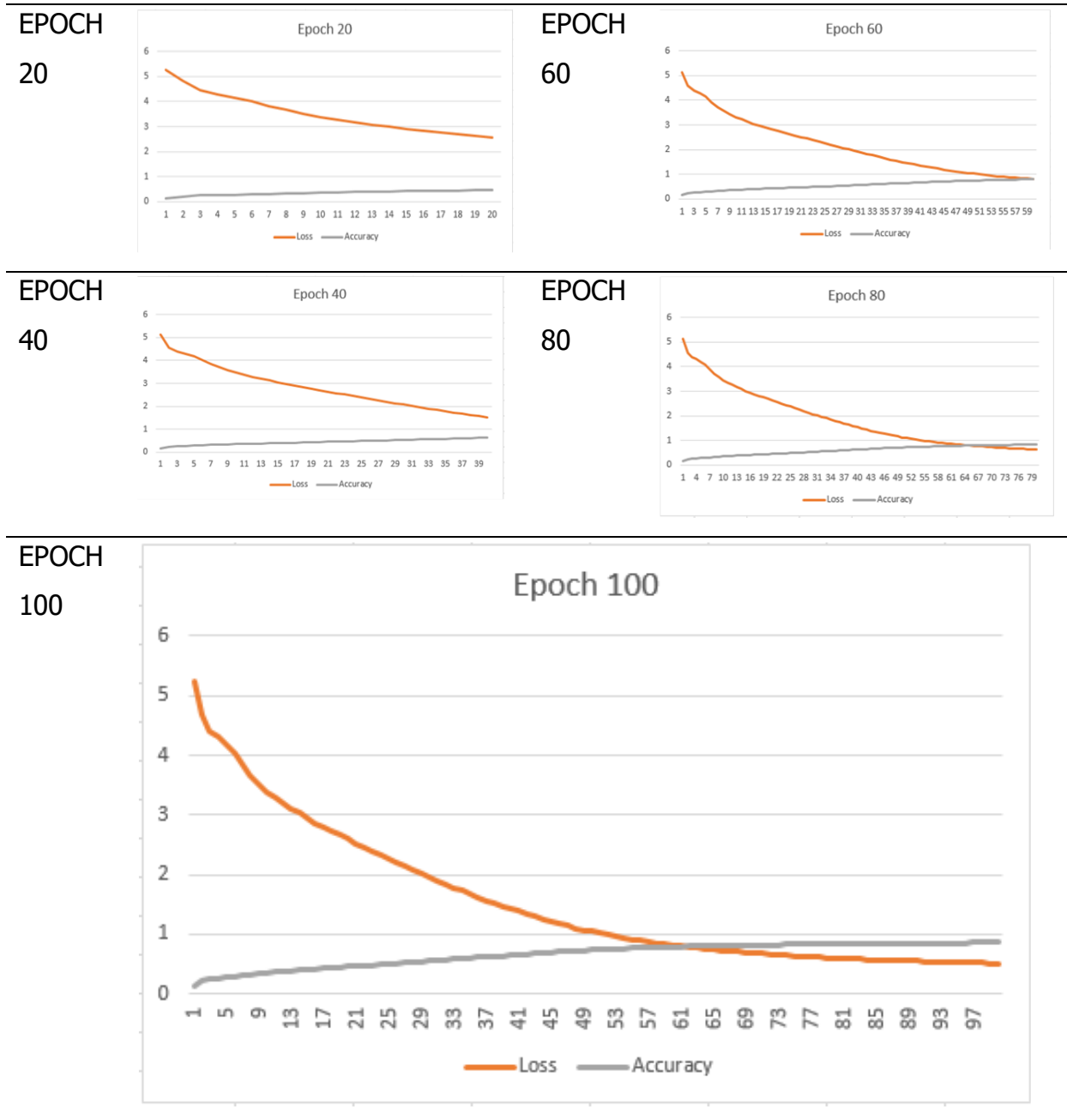
Inisialisasi ResNet50 dan memuat bobot dari *pretrained imagenet*, setelah itu *caching* fitur yang diekstrak dari ResNet50. Lalu setelah melakukan ekstraksi langkah selanjutnya adalah *preprocessing* teks dan melakukan *tokenisasi*. Tujuannya untuk mendapatkan setiap kosa kata, lalu dari kosa kata tersebut dipilih 5000 kosa kata yang tertinggi.

4.4 Membangun Model

Pada tahapan membangun model yaitu menggunakan dua metode yaitu ResNet50 dan LSTM. Metode ResNet50 digunakan untuk *generate* fitur pada gambar, sedangkan LSTM digunakan untuk melakukan *generate caption*.

Setelah mendapatkan model, langkah selanjutnya adalah melakukan pelatihan terhadap gambar dan *caption* yang telah dipersiapkan. Pada penelitian ini dibuat lima jenis *epoch* yakni *epoch 20*, *epoch 40*, *epoch 60*, *epoch 80*, dan *epoch 100*. *Epoch* ke-20 menghasilkan *loss* 2,5476 dan *accuracy* 0,4571.

Tabel 1. Tabel *loss* & *accuracy*



Analisis Loss & Accuracy Epoch 100

Tabel 2. Loss & Accuracy Epoch 100

Epoch	Waktu	Loss	Accuracy
40	35	1,426	0,644
41	35	1,3819	0,6546
42	35	1,3381	0,6633
43	35	1,2904	0,6745
44	35	1,2488	0,6834
45	34	1,2086	0,6934
46	34	1,1717	0,6994
47	35	1,1336	0,7084
48	35	1,1019	0,7159
49	34	1,0684	0,7234
50	35	1,0413	0,7322
51	35	1,0139	0,7362
52	35	0,9861	0,7435
53	35	0,962	0,7497
54	35	0,9424	0,7551
55	34	0,913	0,7612
56	35	0,8966	0,7648
57	35	0,8674	0,771
58	34	0,8463	0,7764
59	35	0,8304	0,779
60	35	0,8146	0,7841
61	34	0,7991	0,7895
62	34	0,7828	0,7931
63	35	0,7708	0,7953
64	35	0,7572	0,799
65	35	0,7443	0,8012
66	35	0,7238	0,8065
67	35	0,7148	0,8084
68	35	0,704	0,812
69	35	0,6871	0,8137
70	34	0,6823	0,8151
71	35	0,6712	0,8179
72	35	0,6585	0,8212
73	35	0,6521	0,8219
74	35	0,6427	0,8262
75	35	0,6329	0,8281

76	35	0,6216	0,8301
77	35	0,6215	0,8293
78	35	0,6118	0,8318
79	35	0,6039	0,8341
80	35	0,5961	0,8364
81	35	0,5921	0,8373
82	35	0,588	0,8378
83	35	0,5795	0,8402
84	35	0,5721	0,844
85	35	0,5683	0,8425
86	35	0,566	0,8436
87	34	0,5537	0,8467
88	34	0,5536	0,8472
89	35	0,5467	0,8499
90	35	0,5451	0,8495
91	35	0,5411	0,8502
92	35	0,5336	0,8515
93	35	0,5317	0,8527
94	35	0,5332	0,8513
95	35	0,5252	0,8535
96	35	0,5232	0,8534
97	35	0,5169	0,8562
98	35	0,5131	0,8556
99	35	0,5087	0,8569
100	35	0,5026	0,8574

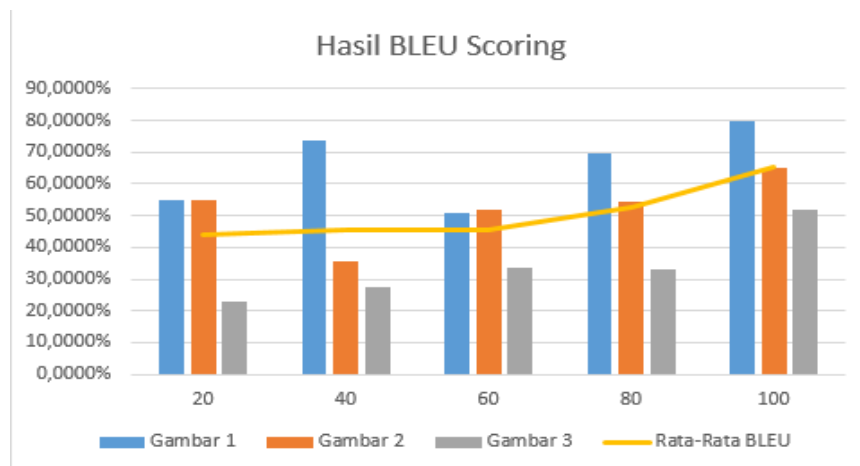
Total Waktu:	3486	Seconds
---------------------	------	---------

Pada tabel dan grafik *epoch* 100 memerlukan waktu selama 3486 detik dan menghasilkan *loss* 0,5026 *accuracy* 0,8574. *Epoch* 100 ini adalah hasil yang paling tinggi diantara *epoch epoch* lainnya, dilihat dari nilai *loss* yang paling kecil dan nilai *accuracy* yang paling besar.

4.5 Analisis Hasil BLEU Scoring

Tabel 3. Analisis Hasil Bleu Scoring

Jumlah Epoch	Gambar 1	Gambar 2	Gambar 3	Rata-Rata BLEU
20	54,9100%	54,6633%	22,8942%	44,1558%
40	73,6923%	35,4948%	27,5687%	45,5853%
60	50,8133%	51,6973%	33,7646%	45,4251%
80	69,8534%	54,4518%	33,2602%	52,5218%
100	79,7455%	65,0059%	51,9603%	65,5706%



Gambar 6. Grafik Hasil BLEU Scoring

Pada hasil BLEU Scoring ini terlihat bahwa semakin besar *epoch* maka akan semakin besar juga hasil BLEU *scoring-nya*. Hasil BLEU *scoring* juga tergantung pada gambar yang diuji, gambar yang dilatih, gambar yang berada pada *dataset*, jumlah gambar yang ada di *dataset*.

5. KESIMPULAN & SARAN

5.1 Kesimpulan

Kesimpulan pada penelitian ini dapat diketahui sebagai berikut. Implementasi dan pengujian *Image Captioning* menggunakan Metode ResNet50 dan *Long Short Term Memory* dapat dilakukan dengan hasil yang sudah dipaparkan. Dalam pengaplikasiannya *Image Captioning* memiliki langkah yang penting yaitu *preprocessing*, *split data*, pembuatan model, *training*, *caption* dan validasi, hingga analisis *output*. Pada analisis *output* dilakukan analisis hasil BLEU *scoring* pada setiap perbedaan gambar dan jumlah *epoch*. Agar dapat mengukur perbedaan hasil BLEU *scoring* dari perbedaan gambar dan perbedaan jumlah *epoch*. Hasil score BLEU menunjukkan bahwa *score* terbesar bernilai 79,7455% pada *Epoch* 100 dan Gambar kesatu.

5.2 Saran

Berikut adalah saran yang diberikan untuk melanjutkan penelitian ini. Mengganti *dataset* gambar yang lebih lengkap dan banyak gambarnya. Mengganti *dataset text* menjadi berbahasa Indonesia. Melakukan scoring lebih dari 1 jenis *scoring* agar lebih akurat untuk menilai metode. Perlu dikembangkan kembali sehingga hasil *score* optimal.

6. DAFTAR PUSTAKA

- Akbar, S. N. A. F., Hendra, & Supri Bin Hj. Amir. (2020). *PERBANDINGAN KINERJA ARSITEKTUR INCEPTION-V4 DAN RESNET-50 DALAM MENGLASIFIKASIKAN CITRA PARU-PARU TERINFEKSI COVID-19*. 2.
- Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). *YOLOv4: Optimal Speed and Accuracy of Object Detection*. <http://arxiv.org/abs/2004.10934>
- Faiyaz Khan, M., Sadiq-Ur-Rahman, S. M., & Saiful Islam, M. (2021). *Improved Bengali Image Captioning via Deep Convolutional Neural Network Based Encoder-Decoder Model*. 217–229. https://doi.org/10.1007/978-981-16-0586-4_18
- Ha, I., Kim, H., Park, S., & Kim, H. (2018). *Image retrieval using BIM and features from pretrained VGG network for indoor localization*. *Building and Environment*, 140(February), 23–31. <https://doi.org/10.1016/j.buildenv.2018.05.026>
- Hochreiter, S., & Schmidhuber, J. (1997). *Long Short-Term Memory*. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Huo, Y., Zhang, M., Liu, G., Lu, H., Gao, Y., Yang, G., Wen, J., Zhang, H., Xu, B., Zheng, W., Xi, Z., Yang, Y., Hu, A., Zhao, J., Li, R., Zhao, Y., Zhang, L., Song, Y., Hong, X., ... Wen, J.-R. (2021). *WenLan: Bridging Vision and Language by Large-Scale Multi-Modal Pre-Training*. <http://arxiv.org/abs/2103.06561>
- Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). *Focal Loss for Dense Object Detection*. *Proceedings of the IEEE International Conference on Computer Vision, 2017-October*, 2999–3007. <https://doi.org/10.1109/ICCV.2017.324>
- Miranda, N. D., Novamizanti, L., & Rizal, S. (2020). *Convolutional Neural Network Pada Klasifikasi Sidik Jari Menggunakan Resnet-50*. *Jurnal Teknik Informatika (Jutif)*, 1(2), 61–68. <https://doi.org/10.20884/1.jutif.2020.1.2.18>
- PUTRI, A., Fachrurrozi, M., & Yusliani, N. (2019). *Identifikasi Kalimat Ambigu Dalam Bahasa Indonesia Menggunakan Top Down Parsing*. <https://repository.unsri.ac.id/6222/>
- Talita, A. S., & Wiguna, A. (2019). *Implementasi Algoritma Long Short-Term Memory (LSTM) Untuk Mendeteksi Ujaran Kebencian (Hate Speech) Pada Kasus Pilpres 2019*. *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 19(1), 37–44. <https://doi.org/10.30812/matrik.v19i1.495>