

1. PENDAHULUAN

1.1. Latar Belakang

Image captioning merupakan sebuah Teknik untuk membuat caption atau deskripsi dari sebuah gambar. Image captioning menjadi hangat akhir-akhir ini karena teknologi komputasi yang makin meningkat, sehingga pengembangan semakin meningkat. Image captioning bisa diadaptasikan dalam banyak hal, salah satunya adalah *self driving cars*. Dengan adanya image captioning maka informasi yang ada disekitar kendaraan bisa dideskripsikan, sehingga mobil bisa mengetahui bagaimana keadaan diluar.

Dalam image captioning terdapat dua langkah utama, yaitu encoder dan decoder[1]. Encoder adalah melakukan ekstraksi fitur dalam sebuah gambar, hasilnya berupa kumpulan data yang merepresentasikan objek-objek apa saja yang ada didalam gambar. Selanjutnya adalah decoder, yaitu proses rekonstruksi kalimat berdasarkan kumpulan objek yang telah diidentifikasi sebelumnya sehingga bisa menghasilkan sebuah teks deskripsi mengenai gambar.[2]

Convolutional Neural Network (CNN)[3] merupakan *state-of-the-art* dalam melakukan encoder[3]. CNN menjadi yang terbaik karena tingkat akurasi dalam mendeteksi objek-objek yang ada dalam sebuah gambar. Recurrent Neural Network (RNN)[4] adalah metode decoder yang banyak digunakan saat ini, karena RNN dapat menyimpan informasi dari kata-kata sebelumnya sehingga bisa membuat sebuah kalimat yang lebih alami.

Model CNN dengan tingkat akurasi yang tinggi dan kecepatan komputasi yang relatif cepat saat ini adalah Mask R-CNN[5]. Model ini mampu melakukan *object detection*, *object localization* dan *object segmentation*. RNN mengalami banyak perkembangan, salah satunya adalah LSTM (*Long Short Term Memory*) [6]. Model ini lebih baik dari RNN karena bisa mengatasi *vanishing gradient*[7] pada RNN.

Untuk meningkatkan detail yang ada pada kalimat, metode adaptive attention[2] mampu memilih kapan sebaiknya melihat objek pada gambar dan objek apa yang harus dilihat dan kapan sebaiknya melihat susunan kalimat yang ada.

Untuk menghasilkan kalimat yang lebih baik lagi, metode learning yang digunakan juga perlu ditingkatkan. Metode learning yang umumnya digunakan adalah

Maximum loglikelihood estimation (MLE) [8]. Metode learning ini bekerja dengan cara mencari nilai maximum dari kondisi loglikelihood dari sampel training. Dengan metode ini, kalimat yang dihasilkan cukup bagus akan tetapi beberapa ciri khas dari gambar biasanya terlewatkan [9]. Metode Contrastive Learning merupakan metode yang diusulkan oleh [9] untuk bisa meningkatkan detail pada sebuah kalimat dari image captioning.

1.2. Perumusan Masalah

Pembuatan teks deskripsi menggunakan metode CNN dan LSTM masih menghasilkan kalimat yang kurang detail.

1.3. Tujuan

Tujuan dari penelitian ini adalah mendesain model CNN dan LSTM yang mampu menghasilkan kalimat yang lebih detail untuk menjelaskan kondisi pada gambar.

1.4. Batasan Masalah

Gambar dan kalimat yang digunakan sebagai data latih dan data uji adalah gambar dan kalimat yang berasal dari MSCOCO dataset yang terbatas hanya 80 kelas objek.

1.5. Rencana Kegiatan

1. Studi literatur

Tahap ini dilakukan sebelum melakukan penelitian. Pada tahap ini pustaka yang dikumpulkan berupa paper, jurnal dan teori-teori yang berkaitan dengan image captioning.

2. Merancang model

Pada encoder model yang digunakan adalah Mask R-CNN dan untuk decoder model yang digunakan adalah LSTM dengan tambahan Adaptive Attention dan untuk fungsi objektifnya menggunakan Contrastive learning.

3. Pengujian dan analisis model

Model Mask R-CNN yang digunakan adalah model yang sudah ditrain oleh pihak ketiga, sedangkan model decodernya akan dilatih dan diuji secara pribadi yang kemudian hasil ujinya akan diujikan dengan metric pengujian B-1, B-2, B-3, B-4, ROUGE_L dan CIDEr. Kemudian hasil tersebut dibandingkan dengan metode lainya untuk mendapatkan kesimpulan apakah model yang diajukan memiliki performa yang baik untuk image captioning atau tidak.

4. Menulis dokumentasi

Hasil penelitian akan didokumentasikan ke dalam sebuah paper.

1.6. Jadwal Kegiatan

Jadwal pelaksanaan Tugas Akhir adalah sebagai berikut

Tabel 1. Jadwal kegiatan penulisan tugas akhir

Kegiatan	Bulan					
	1	2	3	4	5	6
Studi Literatur	■	■				
Merancang Model		■	■			
Pengujian dan Analisi Model			■	■	■	
Menulis Dokumentasi					■	■