

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Dalam beberapa tahun terakhir, kemajuan signifikan telah dibuat dalam pemrosesan citra, seperti klasifikasi citra dan deteksi objek. Memanfaatkan kemajuan dalam bidang klasifikasi gambar dan deteksi objek, konten visual gambar dimungkinkan, yang dikenal sebagai *Image Captioning*. Membuat deskripsi gambar yang lengkap dan alami secara otomatis memiliki manfaat besar, seperti membuat teks yang dilampirkan pada gambar berita, deskripsi terkait gambar medis, pengambilan gambar berbasis teks, informasi yang dapat diakses oleh pengguna tunanetra, interaksi manusia dengan komputer (Saloni Kalra dan Alka Leeka, 2020).

Membuat deskripsi teks memerlukan penggunaan visi komputer dan teknik proses Bahasa alami. Dari perspektif visi komputer, interpretasi dapat mencakup semua aspek visual dari sebuah gambar: dapat berbicara tentang benda/objek dan fitur pemandangan, seperti di luar ruangan atau di dalam ruangan, atau mengekspresikan interaksi orang dan objek dalam sebuah pemandangan. Pemahaman gambar memang penting, tetapi tentu saja tidak cukup untuk menghasilkan deskripsi yang baik. Deskripsi yang baik harus komprehensif tetapi ringkas, kalimat yang benar secara tata bahasa (Yugandhara A. Thakare, Dr. K. H Walse, dan Dr. V. M Thakare, 2020).

Penggunaan pembelajaran mendalam untuk *image captioning* telah meledak dalam beberapa tahun terakhir. Beberapa metode yang umum digunakan adalah *Convolutional Neural Networks* (CNN) dan *Long Short-Term Memory* (LSTM). Oleh karena itu, penelitian ini dilakukan dengan menggunakan pendekatan *Convolutional Neural Networks* (CNN) dan *Recurrent Neural Networks* (RNN) (Saloni Kalra dan Alka Leeka, 2020).

Berdasarkan latar belakang diatas, penelitian ini dilaksanakan karena melihat pentingnya *image captioning* pada saat ini dan masa yang akan datang. Hasil yang diharapkan dari penelitian ini adalah pembuatan keterangan dari sebuah gambar

secara otomatis dengan hasil berupa deskripsi gambar berbahasa Indonesia dengan tata Bahasa yang benar dan senatural mungkin.

1.2. Perumusan Masalah

- a. Bagaimana implementasi metode *deep learning* menggunakan arsitektur Inception V3, algoritma *Recurrent Neural Network* (RNN) dan algoritma *Convolutional Neural Network* (CNN) dalam menghasilkan deskripsi gambar secara otomatis berbahasa Indonesia?
- b. Bagaimana tingkat akurasi yang didapatkan dari arsitektur Inception V3, algoritma *Recurrent Neural Network* (RNN) dan algoritma *Convolutional Neural Network* (CNN) untuk menghasilkan deskripsi gambar secara otomatis berbahasa Indonesia?

1.3. Tujuan Penelitian

Tujuan dari penulisan tugas akhir ini adalah sebagai berikut :

- Mengimplementasikan arsitektur Inception V3, algoritma *Recurrent Neural Network* (RNN) dan algoritma *Convolutional Neural Network* (CNN) ke dalam model untuk menghasilkan deskripsi gambar, sehingga program dapat mengidentifikasi objek di dalam gambar secara otomatis.
- Melakukan preprocessing data untuk menghasilkan kualitas data gambar yang baik saat *testing*.

1.4. Batasan Masalah

Agar dalam pengerjaan tugas akhir ini menjadi lebih terarah dan mendapatkan hasil yang lebih spesifik, maka sistem yang dirancang akan dibatasi dalam Batasan masalah sebagai berikut :

- Dataset yang di gunakan pada penelitian ini merupakan dataset dari situs *Flickr8k*.
- Deskripsi gambar (*caption*) yang digunakan sebagai dataset merupakan deskripsi dari dataset yang diperoleh.
- Deskripsi yang akan ditampilkan menggunakan bahasa Indonesia.

1.5. State of The Art

Tabel 1. 1 State of The Art

Judul Jurnal	Pembahasan
<p>Klasifikasi Citra Buah Menggunakan Convolutional Neural Network</p> <p>Peneliti Febian Fitra Maulana dan Naim Rochmawati</p> <p>Lokasi Universitas Negeri Surabaya</p> <p>Tahun 2019</p> <p>Nama Jurnal Jurnal Informatika dan Ilmu Komputer</p>	<p><u>Hasil Penelitian:</u></p> <p>Pada jurnal ini membahas tentang pengembangan sistem untuk mengklasifikasi citra buah dengan menggunakan algoritma <i>Convolutional Neural Network</i> dengan mencari tingkat akurasi kecocokan pada gambar yang di input dari <i>dataset</i>. Akurasi yang didapatkan menunjukkan angka 97,97%.</p> <p><u>Alasan Menjadi Tinjauan Penelitian:</u></p> <p>Metode yang digunakan dalam penelitian ini menggunakan algoritma <i>Convolutional Neural Network</i> (CNN).</p>

<p>Generating Image Description on Indonesian Language using Convolutional Neural Network and Gated Recurrent Unit</p> <p>Peneliti</p> <p>Aditya Alif Nugraha, Anditya Sriefianto, dan Suyanto</p> <p>Lokasi</p> <p>School of Computing Telkom University Bandung, Indonesia</p> <p>Tahun</p> <p>2019</p> <p>Nama Jurnal</p> <p><i>International Conference on Information and Communication Technology (ICoICT)</i></p>	<p><u>Hasil Penelitian:</u></p> <p>Hasil yang didapatkan pada jurnal ini yaitu image caption menggunakan algoritma <i>Convolutional Neural Network</i> dan arsitektur <i>Gated Recurrent Unit</i>. Serta bagaimana perhitungannya</p> <p><u>Alasan Menjadi Tinjauan Penelitian:</u></p> <p>Dengan melihat gabungan penggunaan algoritma <i>Convolutional Neural Network</i> (CNN) dan <i>Gated Recurrent Unit</i> (GRU) dapat melakukan <i>image captioning</i> yang cukup baik dan mudah dimengerti, sehingga perpaduan metode tersebut dapat digunakan sebagai acuan untuk membuat sistem <i>Image Captioning</i>.</p>
---	--

<p>Image Caption using Deep Learning</p> <p>Peneliti</p> <p>Tarun Wadhwa, Harleen Virk, Dr. Jagannath Aghav, Savita Borole</p> <p>Lokasi</p> <p>College of Engineering Pune, Maharashtra, India</p> <p>Tahun</p> <p>2020</p> <p>Nama Jurnal</p> <p><i>International Journal of Research in Applied Science & Engineering Technology (IJRASET)</i></p>	<p><u>Hasil Penelitian:</u></p> <p>Image Caption menggunakan CNN yang digunakan untuk pemrosesan gambar dan menemukan semua detail tentang gambar. RNN untuk menghasilkan teks aktual menggunakan LSTM. <i>Dataset flickr30k</i> yang berisi 31.783 gambar dengan 5 keterangan gambar untuk setiap gambar. Model yang digunakan adalah <i>inception V3</i>.</p> <p><u>Alasan Menjadi Tinjauan Penelitian:</u></p> <p>Dengan melihat penggunaan <i>Inception V3</i> dan <i>dataset flickr</i> akan menjadi acuan untuk pembuatan sistem <i>Image Captioning</i>.</p>
--	---

<p>A Review of Convolutional Neural Networks</p> <p>Peneliti</p> <p>Arohan Ajit, Koustav Acharya, dan Abhishek Samanta</p> <p>Lokasi</p> <p>School of Electrical Engineering KIIT Deemed to be University Bhubaneswar, India</p> <p>Tahun</p> <p>2020</p> <p>Nama Jurnal</p> <p><i>International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)</i></p>	<p><u>Hasil Penelitian :</u></p> <p>Convolutional Neural Networks telah berusaha memberikan akurasi yang lebih tinggi di semua bidang yang paling populer di antaranya adalah Deteksi Objek, Digit, dan Pengenalan Gambar. Hal ini menggunakan algoritma langkah-langkah yang pasti untuk diikuti termasuk metode seperti Backpropagation, Convolutional Layers, formasi Fitur dan Pooling dan menggunakan berbagai kerangka kerja dan alat yang melibatkan model CNN.</p> <p><u>Alasan Menjadi Tinjauan Penelitian :</u></p> <p>Dengan melihat penjelasan dari penggunaan algoritma Convolutional Neural Network dapat menjadi acuan dalam pengerjaan tugas akhir ini.</p>
--	---

<p>Deep Learning Based Image Captioning: The State of The Art</p> <p>Peneliti Yugandhara A. Thakare, Dr. K. H.Walse, dan Dr. V. M. Thakare</p> <p>Lokasi Computer Science & Engineering Department SIPNA COET, Amravati, MS, India</p> <p>Tahun 2020</p> <p>Nama Jurnal <i>International Journal of Advance Research in Computer Science and Management Studies</i></p>	<p><u>Hasil Penelitian:</u></p> <p>Image Caption memberikan deskripsi singkat dan ringkas tentang gambar dalam Bahasa alami dan telah didapat dengan menggunakan Teknik kombinasi visi computer, pemrosesan Bahasa alami dan metode pembelajaran mesin serta membandingkan metrik evaluasi dan kumpulan data yang berbeda.</p> <p><u>Alasan Menjadi Tinjauan Penelitian:</u></p> <p>Dengan melihat hasil penjelasan penggunaan metrik dan kumpulan data menjadi acuan dalam pengerjaan tugas akhir ini.</p>
--	---

1.6. Sistematika Penulisan

Sistematika penulisan pada tugas akhir ini disusun untuk memberikan gambaran secara umum mengenai permasalahan dan pemecahannya. Penyusunan ini diuraikan dalam beberapa pokok permasalahan yang terbagi dalam beberapa bab. Sistematika penulisan tugas akhir ini adalah sebagai berikut :

BAB 1 PENDAHULUAN

Pembahasan pada bab ini yaitu mengenai pendahuluan penelitian yang terdiri dari latar belakang, perumusan masalah, tujuan penelitian, Batasan masalah, *state of the art* dan sistematika penulisan.

BAB 2 **LANDASAN TEORI**

Pembahasan pada bab ini yaitu mengenai landasan teori penelitian yang terdiri dari teori dasar mengenai *Image Captioning*, *Citra Digital*, *Machine Learning*, *Algoritma Machine Learning*, *Convolutional Neural Network (CNN)*, *Inception V3*, *Recurrent Neural Networks (RNN)*, *Feed forward*, aktivasi *Softmax*, *Loss Function*, *Backpropagation*, *Tensorflow*, *Generate Caption*.

BAB 3 **METODE DAN PEMBAHASAN**

Pembahasan pada bab ini yaitu analisis yang digunakan dalam membangun sistem dan merancang sistem sesuai dengan analisis yang dilakukan.

BAB 4 **IMPLEMENTASI**

Pembahasan pada bab ini yaitu implementasi algoritma *Convolutional Neural Network (CNN)* dan *Recurrent Neural Networks (RNN)* dalam mendeteksi gambar yang akan menghasilkan *captioning* berbahasa indoensia.

BAB 5 **KESIMPULAN DAN SARAN**

Pada bab ini memuat kesimpulan dan saran dari pembahasan analisis pendeteksi gambar.