

BAB I

PENDAHULUAN

1.1. Latar Belakang

Semenjak berkembangnya internet berkecepatan tinggi di Indonesia, semakin banyak pula pengguna internet yang berselancar di dunia maya. Hal ini menyebabkan mudahnya dalam mendapatkan informasi dan konten. Mulai dari informasi atau konten yang bersifat positif hingga yang bersifat negative. Salah satu konten yang bersifat negative adalah konten porno. Saat ini konten porno sangat mudah sekali untuk di akses. Meskipun provider internet melakukan bloking ke situs terkait namun pengguna internet bias menggunakan langkah khusus untuk melewati proses blokir tersebut. Pada tahun 2010 hampir 97% remaja pernah mengakses materi pornografi dimana hal tersebut mempunyai dampak yang buruk dalam tumbuh kembang remaja. Untuk membantu mengatasi dampak buruk dari pornografi kita bisa menggunakan kecerdasan buatan untuk melakukan deteksi. Namun terkadang saat melakukan deteksi ada beberapa hal yang sebenarnya tidak termasuk kedalam pornografi namun ikut terlabeli sebagai pornografi. Oleh karena itu disini peneliti menganggap dibutuhkan alat yang dapat memahami sebuah gambar lebih dalam sehingga menghasilkan deskripsi secara tepat apakah gambar tersebut dapat kita beri label pornografi atau bukan.

Dalam menciptakan alat yang mampu membuat deskripsi dari suatu gambar ada beberapa algoritma maupun gabungan algoritma yang bisa digunakan. Salah satu yang peneliti gunakan disini adalah gabungan antar alogritma deep learning. Algoritma *deep learning* merupakan algoritma jaringan syaraf buatan atau *neural network* namun memiliki jaringan yang lebih besar dan lebih rumit. Pada penelitian ini, peneliti menggunakan *Convolutional Neural Network (CNN)* sebagai encoder yang akan melakukan feature extraction pada gambar kemudian digabungkan dengan *Long-short Term Memory (LSTM)* sebagai decoder yang akan melakukan prediksi rangkaian kata untuk hasil deskripsinya. Pada penelitian yang dilakukan oleh Arnav, Hankyu Jang, dan Pulkit Maloo, mereka juga menggunakan *Convolutional Neural Network (CNN)*

dan *Long-short Term Memory (LSTM)* untuk mendapatkan prediksi *caption*. Namun model yang mereka gunakan tergolong model yang sudah lama, yaitu VGG dan ResNet.[1] Selain itu data yang digunakan merupakan data yang hanya berasal dari Flickr8k saja. Berbeda dengan penelitian ini yang menggunakan model yang lebih baru yaitu InceptionV3 dan lebih optimal. Selain itu pada penelitian ini juga memiliki focus untuk mengembangkan model yang mampu memprediksi *caption* pada gambar porno. Pada penelitian ini juga diterapkan proses *transfer learning* dimana hasil encoding bisa digunakan pada network decoder lainnya sesuai dengan kebutuhan dan hasil akhir.

Penelitian ini dapat dikembangkan untuk memperkuat blokir konten pornografi yang sedang digencarkan oleh pemerintah. Saat ini blokir yang dilakukan pemerintah baru terbatas kepada keyword yang biasa digunakan pada pencarian *search engine* saja. Namun terkadang konten yang tidak mengandung pornografi juga ikut terblokir. Dengan adanya penelitian diharapkan konten pornografi bisa lebih terspesifikasi dengan detail sehingga tidak membuat konten yang bermanfaat lainnya ikut terblokir.

Peneliti mengharapkan hasil penelitian ini dapat memberikan hasil prediksi *caption* yang baik dan akurasi model yang lebih baik dari penelitian sebelumnya. Selain itu penelitian ini dapat digunakan sebagai batu loncatan untuk penelitian selanjutnya.

1.2. Rumusan Penelitian

Berdasarkan latar belakang yang telah di jabarkan, permasalahan pada penelitian ini adalah bagaimana menerapkan *Convolutional Neural Network (CNN)* dan *Long-short Term Memory (LSTM)* dalam memprediksi *caption* dari sebuah gambar. Peran dari *CNN* sendiri merupakan sebagai encoder. Encoder disini berguna untuk melakukan ekstraksi fitur. Sedangkan untuk *LSTM* merupakan sekuensial network yang akan digunakan untuk memprediksi *caption* menggunakan beam search.

Pembuktian hasil penelitian dilakukan dengan menghitung nilai akurasi prediksi dari deskripsi dengan menggunakan *BLEU score*, *METEOR*, dan *CIDEr* dimana skor persentase yang lebih tinggi menandakan hasil yang lebih baik.

1.3. Batasan Penelitian

Batasan masalah dalam penelitian ini adalah :

1. Algoritma yang digunakan adalah *Convolutional Neural Network (CNN)* dan *Long-short Term Memory (LSTM)*.
2. Input yang dipakai adalah gambar dan teks caption dari dataset Flickr8K dan gambar porno yang peneliti kumpulkan sendiri. Untuk caption dari gambar yang peneliti kumpulkan sendiri membuatnya sendiri secara manual.
3. Belum menyertakan gambar kartun maupun anime.
4. Fokus caption adalah action yang terlihat pada gambar.
5. Caption dalam Bahasa Inggris.
6. Jumlah data yang digunakan adalah sekitar 8100 gambar dan 40500 caption.
7. Bahasa pemrograman yang digunakan adalah Python dan JavaScript.
8. Untuk proses *feature extraction* peneliti menggunakan *bottleneck feature*.
9. Menggunakan *pre trained model* yang disediakan oleh keras.

1.4. Tujuan Penelitian

Tujuan penelitian ini adalah untuk mengetahui kinerja dari *CNN* dan *LSTM* dalam melakukan prediksi *caption* dari sebuah gambar. Selain itu penelitian ini juga bertujuan untuk mengukur nilai akurasi dari prediksi *caption* dari algoritma *CNN* dan *LSTM*.

1.5. Metodologi Penelitian

Metode yang digunakan dalam pengerjaan tugas akhir ini saya jabarkan seperti di bawah ini:

1.5.1 Pengumpulan Data

Pengumpulan data Flickr8k dilakukan dengan mendownload file pada website universitas Illinois. Sedangkan tambahan dataset porno diperoleh dari kaggle dengan caption yang ditambahkan secara manual sesuai dengan gambar.

1.5.2 Tahapan Penelitian

1. Analisis

Melakukan analisa terhadap syarat-syarat yang dibutuhkan dalam melakukan penelitian.

1) Analisis masalah

Melakukan pemetaan terhadap masalah yang akan diselesaikan melalui penelitian ini.

2) Analisis kebutuhan data

Memetakan kebutuhan data yang akan di gunakan pada penelitian ini, yaitu data gambar porno beserta dengan captionnya sebanyak 8100 gambar dengan jumlah gambar porno sebanyak 100 gambar.

3) Analisis kebutuhan penelitian

Memetakan kebutuhan yang diperlukan selama penelitian, yaitu kebutuhan fungsional dan kebutuhan non fungsional.

2. Perancangan

Pada tahap ini dilakukan perancangan metode kombinasi antara *CNN* sebagai feature ekstraktor dan *LSTM* sebagai caption generator yang akan diimplementasikan untuk mendapatkan sebuah model yang mampu menghasilkan prediksi *caption* dari sebuah gambar.

3. Implementasi

Pada tahap ini akan dilakukan penulisan kode program.

4. Pengujian

Pada tahap ini akan dilakukan perhitungan akurasi dari implementasi yang telah dilakukan. Metode yang digunakan adalah dengan menghitung *BLEU score*, *METEOR score*, dan *CIDEr* terhadap *caption* yang sudah di hasilkan. Hasil pengujian adalah jumlah score dimana skor yang lebih tinggi menunjukkan hasil yang lebih baik.