

V. PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang sudah dilakukan, dapat disimpulkan sebagai berikut.

1. Penelitian ini berhasil membuat model *image captioning* pada gambar objek wisata di Purbalingga menggunakan arsitektur *transformer*. Model yang dikembangkan mampu menghasilkan deskripsi otomatis dari gambar-gambar objek wisata. Data yang digunakan mencakup gambar dari lima objek wisata di Purbalingga, antara lain D'Las, Goa Lawa, Owabong, Purbasari, dan Sanggaluri. Seluruh proses pengembangan dan implementasi dilakukan menggunakan bahasa pemrograman *Python* dengan pendekatan *deep learning* serta *library tensorflow* dan *keras*. Model arsitektur *transformer* terdiri dari lapisan ekstraksi fitur gambar menggunakan ResNet50 serta beberapa *layer encoder* dan *decoder* dengan *attention mechanism* dan *feed-forward neural network*.
2. Evaluasi model *image captioning* dilakukan menggunakan metrik BLEU. Hasil evaluasi menunjukkan bahwa skor BLEU- $\{1, 2, 3, 4\}$ terbaik yang diperoleh adalah $\{0.672, 0.559, 0.489, 0.437\}$. Skor ini dicapai dengan menggunakan kombinasi *hyperparameter* dimensi *embedding* 128, jumlah *layer 2*, dan jumlah *head 2*.
3. Perubahan dimensi *embedding*, jumlah *layer*, dan jumlah *head* memiliki pengaruh terhadap waktu pelatihan dan skor BLEU. Peningkatan dimensi *embedding* meningkatkan waktu pelatihan secara drastis dan menurunkan skor BLEU setelah titik optimal pada dimensi 128. Penambahan jumlah *layer* meningkatkan waktu pelatihan, dengan skor optimal tercapai pada 2 *layer*, sementara lebih dari 4 *layer* menurunkan skor BLEU. Perubahan jumlah *head* memiliki dampak minimal pada waktu pelatihan dan skor BLEU, dengan sedikit penurunan skor BLEU pada jumlah *head* yang lebih besar.

4. Model *image captioning* pada gambar objek wisata di Purbalingga diimplementasikan dalam aplikasi berbasis web menggunakan metode SDLC *waterfall*. Sistem ini memiliki dua jenis pengguna, yaitu admin dan *guest*, dengan hak akses berbeda. Desain sistem menggunakan *Unified Modeling Language* (UML), dan implementasi dilakukan menggunakan *framework Flask* serta basis data *MySQL*. Aplikasi ini memungkinkan pengguna memotret atau mengunggah gambar objek wisata, mendapatkan deskripsi gambar otomatis dalam bahasa Indonesia, dan mendengarkan *caption*-nya yang dibacakan secara audio menggunakan fitur *text-to-speech* berbasis *Web Speech API*. Hasil pengujian dengan metode *blackbox* menunjukkan hasil valid di setiap kasus uji, menandakan bahwa sistem berjalan sesuai kebutuhan dan layak digunakan. Aplikasi ini diharapkan dapat memberikan manfaat dalam menguraikan deskripsi/*caption* gambar objek wisata di Purbalingga secara otomatis serta meningkatkan aksesibilitas bagi penyandang tunanetra.

5.2 Saran

Beberapa saran yang dapat diperhatikan untuk pengembangan penelitian dan sistem di masa mendatang adalah sebagai berikut:

1. Memperluas dataset dengan menambahkan lebih banyak gambar dari berbagai objek wisata dan mengambil gambar dari berbagai sudut dan kondisi yang berbeda. Hal ini akan membantu model belajar lebih banyak variasi dan meningkatkan akurasi *caption* yang dihasilkan.
2. Menggunakan arsitektur model yang lebih kompleks atau menggabungkan beberapa teknik seperti *attention mechanisms* yang lebih canggih atau menggunakan model yang lebih besar seperti GPT atau BERT untuk meningkatkan kualitas *caption* yang dihasilkan.
3. Menggunakan metrik evaluasi tambahan seperti METEOR atau ROUGE untuk memberikan penilaian yang lebih komprehensif terhadap kualitas deskripsi yang dihasilkan oleh model.
4. Menambahkan fitur-fitur tambahan pada aplikasi, seperti fitur untuk

mendeteksi dan mengenali objek dalam gambar secara otomatis dan *real time*, serta menyediakan pilihan bahasa lain untuk deskripsi dan fitur *text-to-speech*.

