

ABSTRAK

Daerah Istimewa Yogyakarta dikenal sebagai kota pendidikan. Namun tidak dipungkiri bahwa Yogyakarta juga merupakan daerah dengan destinasi wisatanya. Berdasarkan data Badan Pusat Statistik (BPS) pada tahun 2019 total kedatangan wisatawan baik nusantara maupun mancanegara adalah 6.549.381 wisatawan. Banyak wisatawan mengabadikan momen kunjungan wisata ke dalam sebuah foto yang diunggah melalui internet, akan tetapi tidak jarang informasi yang diberikan tidak sesuai dengan gambar tersebut. Sehingga dibutuhkan *image captioning* untuk merepresentasikan sebuah gambar menggunakan kecerdasan buatan dari *Computer Vision* dan *Natural Language Processing*.

Penelitian ini menggunakan metode *Convolutional Neural Network* (CNN) dengan ekstraksi fitur ResNet-50 dan *Gated Recurrent Unit* (GRU) untuk merancang model *image captioning* atau keterangan gambar otomatis. Model tersebut dikembangkan dengan menggunakan *attention mechanism* karena dapat menyatukan informasi yang penting dalam sebuah kalimat dan memfokuskan hasil dari GRU terhadap target kata, sehingga *output* yang dihasilkan diharapkan lebih baik dan sesuai dengan gambar. GRU dipilih karena merupakan pengembangan dari LSTM dimana GRU memiliki *gate* atau gerbang yang lebih sedikit dibandingkan LSTM sehingga memiliki komputasi pelatihan yang cenderung lebih cepat. Dataset pada penelitian ini dibatasi dengan 10 tempat wisata yang ada di Yogyakarta, dengan 50 gambar pada setiap tempat wisatanya. Data gambar diperoleh melalui internet baik *social media* ataupun dengan *search engine* pada browser dengan tiga buah *caption* pada setiap gambarnya. Sehingga total dataset adalah 500 gambar dengan 1500 *caption*.

Dari hasil penelitian ini dengan menggunakan metode CNN ekstraksi fitur ResNet-50 dan GRU didapatkan hasil evaluasi skor BLEU {1,2,3,4} sebesar {42.64, 32.56, 18.55, 16.27} dengan hasil yang cukup baik. Model mampu menghasilkan keterangan gambar berbahasa Indonesia yang sesuai dengan gambar walaupun masih ditemukan kalimat yang kurang tepat. Eksperimen dilakukan dengan menggunakan gambar yang tidak ada dalam *dataset*, model mampu menghasilkan deskripsi gambar yang benar dan dapat dipahami dari hasil keterangan tersebut.

Kata Kunci: *Image Captioning*, Pariwisata Yogyakarta, *Convolutional Neural Network*, *Gated Recurrent Unit*, ResNet-50, *Attention Mechanism*.