

Abstrak

Image Captioning adalah topik yang cukup banyak di diskusikan. Kebanyakan riset mengenai *Image Captioning* masih menggunakan bahasa Inggris, dimana bahasa lainya juga perlu dilakukan riset. Indonesia, sebagai negara terbesar di Asia-Tenggara memiliki bahasanya sendiri, yaitu Bahasa Indonesia. Bahasa Indonesia telah diajarkan di beberapa negara seperti Vietnam, Jepang, dan Australia. Pada riset ini, kami mengajukan metode berbasis attention untuk kasus *Image Captioning*, dimana ResNet101 sebagai encodernya dan LSTM dengan *Adaptive Attention* sebagai decodernya untuk kasus *Image Captioning* berbahasa Indonesia. *Adaptive Attention* bisa menentukan kapan harus melihat pada gambar dan bagian mana yang harus dilihat untuk menghasilkan kata selanjutnya. Model yang kami gunakan, dilatih menggunakan dataset MSCOCO dan Flick30K. Kedua dataset tersebut ditranslate ke Bahasa Indonesia menggunakan google translate dan translate manual. Hasil dari riset kami menghasilkan skor 0.678, 0.512, 0.375, 0.274, dan 0.990 untuk BLEU-1, BLEU-2, BLEU-3, BLEU-4, dan CIDEr secara berurutan. Model kami juga menghasilkan skor yang mirip untuk *Image Captioning* berbahasa Inggris, yang berarti model kami mampu menghasilkan performa yang sama dengan *Image Captioning* berbahasa Inggris. Kami juga mengajukan metode pengujian baru dengan melakukan survey. Hasil dari survey kami menyatakan bahwa 76.8% hasil *caption* dari model kami lebih baik dari data validasi yang di translate menggunakan google translate.

Kata Kunci : Image Captioning, Indonesia, Adaptive Attention, Bahasa Indonesia