# Model Pembuat Keterangan Gambar Berita Generatif: Studi Kasus pada Proyek AINGS = Generative News Images Captioning Model: A Case Study on the AINGS Project

Shadqi Marjan Sadiya, author



---

Abstrak

Jurnalisme di era digital mengalami transformasi signifikan, dengan kemunculan online platform dan media sosial. Hal ini membawa tantangan baru dalam penyampaian informasi yang akurat dan menarik kepada khalayak umum. Penelitian-penelitian sebelumnya sudah mencoba untuk menyelesaikan masalah mengenai pembuatan berita secara otomatis menggunakan prompt singkat, maka dari itu penelitian ini ditujukan untuk melengkapi hal yang esensial untuk pembuatan artikel berita online, yaitu pembuatan keterangan gambar secara otomatis. Penelitian ini berfokus untuk menentukan Vision-Language Model (VLMs) yang paling optimal untuk membuat keterangan gambar dalam konteks artikel berita dalam Bahasa Indonesia. Penelitian dilakukan dengan 3 (tiga) pendekatan VLMs, yakni GoodNews, Transform and Tell, dan VisualNews. Pengembangan VLMs dilakukan dengan melatih masing-masing model secara terpisah. Selanjutnya VLMs dievaluasi dengan metrik penilaian BLEU, ROUGE, dan CIDEr. Hasil yang didapatkan oleh penulis menunjukkan bahwa performa pendekatan model VisualNews jauh lebih baik dibandingkan pendekatan model GoodNews dan Transform and Tell. Model ini mencapai nilai persentase BLEU-4 sebesar 6.93%, ROUGE-L sebesar 23.54%, dan CIDEr sebesar 42.66%.

.....Journalism in the digital era has undergone significant transformations with the rise of online platforms and other social medias. This change carries along new challenges in presenting information which are both accurate and interest-grabbing for the general society. Previous researches have tried to tackle the challenges in regards to automatic creation of articles using short prompts. Therefore, this research is intended to complete that which is essential in the creation of online news articles; automatic image captioning. Our research's focus is determining which is the most optimal Vision-Language Model (VLMs) to create captions in the Indonesian language. The research is undergone using 3 (three) VLMs approaches, being GoodNews, Transform and Tell, and VisualNews. The VLMs development will be evaluated using the metrics BLEU, ROUGE, and CIDEr. Results gained from this research shows that the performance of the VisualNews model is far superior when compared to the GoodNews or Transform and Tell VLMs. The VisualNews model reached a score of 6.93% on the BLEU metric, 23.54% on the ROUGE-L, and a score of 42.66% using the CIDEr metric.