



Comparative Analysis of Transformer Models in Object Detection and Relationship Determination on COCO Dataset

Raihan Atsal Hafizh, Kemas Rahmat Saleh Wiharja*, Muhammad Arya Fikriansyah

School of Computing, Informatics, Telkom University, Bandung, Indonesia

Email: ¹raihanatsal@student.telkomuniversity.ac.id, ^{2,*}bagindokemas@telkomuniversity.ac.id,

³muhammadaryaf@student.telkomuniversity.ac.id

Correspondence Author Email: bagindokemas@telkomuniversity.ac.id

Abstrak—Penelitian ini bertujuan untuk menggali integrasi antara model deteksi objek dan prediksi hubungan guna meningkatkan interpretasi gambar, dengan fokus pada pertanyaan inti: Apa tantangan yang memerlukan Analisis Komparatif antara Model Deteksi Objek dan Model Transformer dalam Penentuan Hubungan? Model deteksi objek yang tangguh menunjukkan kinerja unggul, terutama pada ambang batas Intersection over Union (IoU) yang lebih rendah dan untuk objek yang lebih besar, menciptakan dasar yang kokoh untuk analisis lanjutan. Evaluasi dilakukan terhadap model transformer, seperti GIT, GPT-2, dan PromptCap, dengan menitikberatkan pada kemampuan generasi bahasa mereka dan menampilkan metrik kinerja yang mencolok, termasuk metrik berbasis kata kunci yang inovatif. Penelitian ini transparan dalam menghadapi batasan terkait kendala dataset dan potensi tantangan dalam generalisasi model, memberikan dasar yang jelas untuk landasan penelitian. Melalui evaluasi bersama model deteksi objek dan transformer, penelitian ini memberikan wawasan berharga tentang dinamika interaksi antara pemahaman visual dan linguistik dalam memahami gambar. Dengan mengakui keterbatasan, termasuk kendala data dan generalisasi model, penelitian ini membuka peluang untuk penyempurnaan di masa depan, mengatasi keterbatasan yang teridentifikasi dan menjelajahi berbagai domain aplikasi. Pendekatan komprehensif terhadap pemahaman interaksi antara elemen visual dan tekstual turut berkontribusi pada perkembangan yang terus-menerus dalam penelitian visi komputer dan pemrosesan bahasa alami.

Kata Kunci: Deteksi Objek; Prediksi Relasi; Model Transformer; Pemahaman Gambar; Memahami Gambar