

## ABSTRAK

Kesenjangan antara inspirasi visual yang kaya dan tantangan artikulasi kreatif (*writer's block*) menjadi masalah utama yang menghambat proses penulisan. Penelitian ini bertujuan menjembatani kesenjangan tersebut dengan merancang sistem kecerdasan buatan dua tahap berbasis *deep learning* untuk menyediakan stimulus naratif otomatis. Metode yang diimplementasikan terdiri dari sebuah arsitektur *Convolutional Neural Network* (CNN) kustom yang dibangun untuk mendeteksi tujuh kelas objek alam dari 4.362 gambar, di mana output objeknya kemudian digunakan sebagai *prompt* untuk model *Generative Pre-trained Transformer* (GPT-2) yang telah di-*fine-tune* untuk menghasilkan narasi puitis. Hasil penelitian menunjukkan bahwa modul CNN, yang dievaluasi dengan metrik akurasi dan *Mean Squared Error* (MSE), mencapai performa puncak sebesar 61.96%. Analisis *confusion matrix* mengungkap performa ini tidak dibatasi oleh *overfitting*, melainkan oleh ambiguitas visual antar kelas. Meskipun modul GPT-2 mampu menghasilkan narasi dengan skor BERTScore F1 hingga 0.6455, temuan utama penelitian ini adalah kualitas narasinya sangat bergantung pada keakuratan input dari modul CNN, yang terbukti berfungsi sebagai *bottleneck* kritis bagi efektivitas sistem secara keseluruhan.

