

ABSTRACT

The ability of generative AI to create highly realistic images poses a serious challenge to the authenticity of visual content, leading to negative consequences such as the spread of disinformation and copyright infringement. This study aims to develop and optimize a MobileNetV2 based detection model to accurately and efficiently classify AI generated and real images. The analysis focused on two key factors: the influence of dataset size variations and the role of hyperparameter optimization using the Hyperband algorithm. The model artifact was developed using a transfer learning approach and evaluated on three data scenarios from the CIFAKE dataset: small scale (20% of the data), medium scale (50% of the data), and full data (100% of the data). The evaluation results show that the model trained on the medium-scale data (50%) achieved the most optimal performance with an accuracy of 96.76% and an F1-score 96.69%, along with an efficiency ratio of 0.244. In contrast, increasing the data to 100% did not provide a proportional improvement. The model's accuracy even slightly decreased to 95.69%, indicating a phenomenon of diminishing returns. This was also marked by a significant trade-off between precision (99.01%) and recall (92.31%). Meanwhile, the model on the small scale data (20%) proved to be the most efficient, with an overall good performance of 95.94% accuracy and a 95.92% F1-score, and the highest efficiency ratio of 0.485. It can be concluded that a larger quantity of data does not always guarantee better performance. This finding provides a practical guidance that optimizing dataset size is key to balancing a model's performance with the available computational resources.

Keywords : Generative AI, MobileNetV2, Dataset size, Hyperparameter optimization, Transfer learning

ABSTRAK

Kemampuan AI generatif dalam menciptakan gambar yang realistis menimbulkan tantangan yang serius terkait keaslian konten visual, yang memicu berbagai dampak negatif seperti penyebaran disinformasi dan pelanggaran hak cipta. Penelitian ini bertujuan untuk mengembangkan dan mengoptimalkan artefak berupa model deteksi berbasis *MobileNetV2* untuk mengklasifikasikan gambar buatan AI dan gambar asli secara akurat dan efisien. Analisis berfokus pada dua faktor kunci: pengaruh variasi ukuran dataset dan peran optimasi *hyperparameter* dengan algoritma *Hyperband*. Artefak model dikembangkan dengan pendekatan *transfer learning* dan dievaluasi pada tiga skenario data berdasarkan dataset *CIFAKE*: skala kecil (20% data), skala sedang (50% data), dan seluruh data (100% data). Hasil evaluasi menunjukkan bahwa model pada skala sedang (50% data) memiliki performa paling optimal dengan *accuracy* sebesar 96,76%, *F1-score* sebesar 96,69%, serta rasio efisiensi sebesar 0,244. Sebaliknya, penambahan data hingga 100% tidak memberikan peningkatan sepadan, dengan nilai *accuracy* yang sedikit menurun ke 95,69%, mengindikasikan fenomena *diminishing returns*. Hal ini juga ditandai dengan munculnya *trade-off* signifikan antara *precision* (99,01%) dan *recall* (92,31%). Sementara itu, model pada skala kecil (20% data) terbukti paling efisien dengan *accuracy* 95,94% dan *F1-score* 95,92%, dengan rasio efisiensi tertinggi 0,485. Dapat disimpulkan bahwa kuantitas data yang lebih besar tidak selalu menjamin performa yang lebih baik. Temuan ini memberikan panduan praktis bahwa optimalisasi ukuran data adalah kunci untuk menyeimbangkan performa model dengan ketersediaan sumber daya komputasi.

Kata Kunci : AI generatif, MobileNetV2, Ukuran dataset, Optimasi hyperparameter, Transfer learning