

ABSTRAK

KHDTK Wanadipa Universitas Diponegoro memiliki keanekaragaman hayati yang tinggi (114 spesies), namun proses identifikasi tanaman di lapangan masih bergantung pada metode manual yang lambat dan ketersediaan pakar, sehingga membatasi akses edukasi bagi masyarakat. Meskipun deep learning (CNN) menawarkan solusi, sebagian besar penelitian hanya berfokus pada akurasi dan mengabaikan efisiensi komputasi yang krusial untuk aplikasi real-time. Penelitian ini bertujuan untuk menemukan arsitektur backbone optimal untuk model YOLOv5 yang memberikan keseimbangan (trade-off) terbaik antara akurasi dan efisiensi. Empat backbone (EfficientNetB1, ResNet152, VGG16-BN, MobileNet-Large) dibandingkan dengan baseline YOLOv5N untuk diimplementasikan dalam sistem informasi berbasis web.

Metodologi penelitian dimulai dengan pengumpulan dataset 13 spesies tanaman dari KHDTK, menghasilkan 631 gambar yang kemudian diaugmentasi menjadi 1.553 gambar. Setiap model (empat varian backbone kustom dan satu baseline) dilatih selama 300 epoch menggunakan hardware (NVIDIA A100) dan hyperparameter yang seragam untuk menjamin perbandingan yang adil. Evaluasi performa dilakukan dengan mengukur metrik akurasi (Precision, Recall, mAP@0.5, mAP@0.5:0.95) dan efisiensi komputasi (GFLOPs dan FPS), dengan fokus utama menemukan trade-off terbaik melalui perhitungan Efficiency Score (mAP/GFLOPs).

Hasil penelitian menunjukkan bahwa baseline, yaitu model YOLOv5N standar tanpa modifikasi backbone yang digunakan sebagai acuan pembandingan utama, memiliki akurasi tertinggi (mAP@0.5: 0,980) namun dengan biaya komputasi yang relatif tinggi (24,1 GFLOPs). Di antara backbone kustom, EfficientNetB1 (mAP@0.5: 0,911) menjadi yang paling akurat dan secara bersamaan paling efisien secara komputasi (6,1 GFLOPs), mencapai Nilai Efisiensi tertinggi (14,93), 3,7 kali lebih baik daripada baseline (4,07). Pengukuran Frames Per Second (FPS) turut dilakukan dan menunjukkan EfficientNetB1 (47,33 FPS) sedikit lebih lambat dibandingkan baseline (80,18 FPS), namun korelasi negatif sedang hingga kuat antara GFLOPs dan FPS ($p = -0,600$) mengindikasikan selisih ini lebih disebabkan oleh keterbatasan optimasi implementasi dibandingkan kelemahan arsitektur. Disimpulkan bahwa EfficientNetB1 tetap merupakan backbone paling optimal berdasarkan keseimbangan akurasi dan efisiensi komputasi arsitektural, sehingga dipilih untuk implementasi akhir pada aplikasi web.

Kata kunci: Object Detection; Identifikasi Jenis Tanaman; YOLOv5; Rekonstruksi Model