ABSTRAK

EVALUASI ALGORITMA YOLOV8 DAN YOLOV11 DALAM KLASIFIKASI OTOMATIS TINGKAT KEMATANGAN TANDAN BUAH SEGAR KELAPA SAWIT

Oleh

Bella Amelia

Tingkat kematangan tandan buah segar (TBS) menentukan rendemen, asam lemak bebas (ALB), dan stabilitas mutu minyak sawit. Ketidaktepatan panen menimbulkan kerugian nilai dan inefisiensi proses. Penilaian manual terhadap tandan sawit yang akan dipanen masih bersifat subjektif, dipengaruhi pencahayaan, oklusi pelepah, serta variabilitas varietas dan lokasi kebun. Karena itu, algoritma *computer vision* berbasis *one-stage detector* menawarkan otomasi yang konsisten dan cepat untuk operasional berskala besar. Penelitian ini mengkaji kelayakan dua *backbone computer vision* yaitu YOLOv8 dan YOLOv11, untuk klasifikasi otomatis tingkat kematangan TBS, sekaligus mengevaluasi ketahanan lintas domain data. Eksperimen dilakukan pada dua himpunan data yaitu data *image* dari Roboflow (8.400 citra teranotasi kualitas dan anotasi tertata) dan *dataset* lapangan (625 citra hasil augmentasi) yang merepresentasikan variasi kondisi nyata. Evaluasi model mencakup *precision*, *recall*, F1-score, *mean Average Precision* (mAP), akurasi, serta kecepatan inferensi (FPS). Uji keadilan ukuran sampel dilakukan dengan menyetarakan jumlah data (625 citra) antara subset Roboflow dan *dataset* lapangan. Hasil menunjukkan YOLOv8 lebih stabil dan seimbang di beragam skenario. Pada Roboflow, YOLOv8s menjadi pilihan utama dengan akurasi 87,8%, *precision* (90,5%), *recall* (90,8%), F1-score 90,7%, mAP 94,8%, dan kecepatan 1.123 FPS. Keunggulan Roboflow ada pada volume besar, kualitas gambar yang lebih baik, dan anotasi rapi. Saat jumlah data disetarakan menjadi 625 citra, subset Roboflow tetap unggul pada mAP (95,2%) dan *precision* (94,3%), sementara *dataset* lapangan lebih baik pada *recall* (87,6%), akurasi (78,1%), dan kecepatan (1.079 FPS), menandakan adaptasi model terhadap kondisi kebun yang beragam. Kelebihan spesifik turut teridentifikasi: YOLOv8m unggul pada *precision* dan mAP di subset Roboflow, namun nilai *recall* tinggi untuk dataset lapangan; dilain pihak, YOLOv11s menonjol pada *recall* dan mAP untuk dataset lapangan meski konsistensinya lebih rendah. Secara keseluruhan, YOLOv8s direkomendasikan untuk implementasi praktis karena konsistensi dan keseimbangan kinerja, sementara YOLOv8m atau YOLOv11s layak dipilih bila prioritasnya presisi maksimum atau cakupan deteksi dalam kondisi tertentu.

Kata kunci: Kelapa Sawit, Klasifikasi, YOLOv8, YOLOv11, Deep Learning

ABSTRACT

EVALUATION OF YOLOV8 AND YOLOV11 ALGORITHMS IN AUTOMATIC CLASSIFICATION OF RUITNESS LEVEL OF FRESH OIL PALM FRUIT BUNCHES

By

Bella Amaelia

Fresh fruit bunch (FFB) ripeness strongly determines palm-oil yield, free-fatty-acid formation, and product stability; misgrading at harvest depresses value and propagates inefficiencies downstream. Manual inspection is inherently subjective and sensitive to lighting, frond occlusion, and cultivar/site variability. Consequently, computer vision based on one-stage detectors provides consistent, high-throughput automation for estate-scale operations. This study assesses two state-of-the-art backbones i.e., YOLOv8 and YOLOv11, for automatic FFB ripeness classification and examines cross-domain robustness. This study evaluate two datasets: Roboflow (8,400 annotated images with high visual quality and curated labels) and a field dataset (625 augmented images) representing real-world variability. Metrics include precision, recall, F1-score, mean Average Precision (mAP), accuracy, and inference speed (FPS). A fair-size test is conducted by equalizing sample counts (625 images) between a Roboflow subset and the field dataset. Results indicate that YOLOv8 delivers more stable, balanced performance across scenarios. On Roboflow, YOLOv8s emerges as the primary choice, achieving 87.8% accuracy, 90.7% F1-score, 94.8% mAP, and 1,123 FPS. The advantages of Roboflow, large volume, higher image quality, and clean annotations, resulted into strong precision (90.5%), recall (90.8%), and mAP (94.8%) for YOLOv8s. When data volume is equalized to 625 images, the Roboflow subset still leads in mAP (95.2%) and precision (94.3%), whereas the field dataset yields higher recall (87.6%), accuracy (78.1%), and speed (1,079 FPS), indicating model adaptability to diverse on-farm conditions. Model-specific strengths also emerge: YOLOv8m excels in precision and mAP on the Roboflow subset and delivers high recall in the field dataset, while YOLOv11s stands out for recall and mAP in field dataset tests despite lower overall consistency. Overall, YOLOv8s is recommended for practical deployment due to its consistent, well-balanced performance, whereas YOLOv8m or YOLOv11s are viable when the priority is maximum precision or broader detection coverage under specific operating conditions.

Keywords: Oil Palm, Classification, YOLOv8, YOLOv11, Deep Learning