

BAB V KESIMPULAN

5.1 Kesimpulan

Berdasarkan hasil penelitian mengenai optimasi hyperparameter model YOLOv8 menggunakan metode Tree-Structured Parzen Estimator (TPE) pada sistem klasifikasi tingkat kematangan apel berbasis deep learning, dapat diperoleh beberapa kesimpulan sebagai berikut.

1. Perancangan Pada dataset lama yang berjumlah 1.342 citra, OFAT menggunakan 100 trial dengan 14 epoch per trial dan membutuhkan waktu tuning lebih dari 8 jam, sedangkan TPE menggunakan 30 trial dengan 30 epoch per trial dan hanya membutuhkan waktu sekitar 2,6 jam. Dengan demikian, TPE mampu melakukan proses tuning setidaknya sekitar 5 kali lebih cepat, dengan penghematan waktu minimal 80% dibandingkan OFAT. Efisiensi tersebut diperoleh karena TPE memanfaatkan hasil trial sebelumnya melalui pemodelan probabilistik untuk mengarahkan pencarian menuju kombinasi hyperparameter yang lebih menjanjikan, sedangkan OFAT mengubah hyperparameter secara bertahap tanpa memanfaatkan hubungan antarparameter. Meskipun waktu tuning jauh lebih singkat, TPE tidak mengalami penurunan performa yang berarti; pada pelatihan ulang menggunakan hyperparameter terbaik dari dataset lama, TPE menghasilkan precision 0,990, recall 0,990, F1-score 0,995, mAP50 0,995, dan mAP50-95 0,86523, dibandingkan OFAT yang menghasilkan precision 0,982, recall 0,998, F1-score 0,990, mAP50 0,991, dan mAP50-95 0,850. TPE juga tidak menambah beban komputasi pada tahap inferensi, dengan waktu pemrosesan 10,19 ms per citra (98,16 FPS) dibandingkan OFAT sebesar 10,27 ms per citra (97,37 FPS), atau selisih kurang dari 1%.

Pada dataset baru yang berjumlah 3.344 citra, TPE diperluas menjadi 50 trial dengan 40 epoch per trial dan membutuhkan waktu optimasi sekitar 9,9 jam atau 599 menit untuk menyelesaikan seluruh proses pencarian

hyperparameter. Untuk memvalidasi hasil optimasi tersebut, konfigurasi hyperparameter terbaik dari masing-masing metode kemudian digunakan untuk melatih ulang model secara penuh pada dataset baru. Hasil pelatihan ulang menunjukkan bahwa model TPE menghasilkan F1 sebesar 0,845, mAP50 sebesar 0,875, dan mAP50-95 sebesar 0,747, sedangkan model OFAT menghasilkan recall sebesar 0,805, mAP50 sebesar 0,845, dan mAP50-95 sebesar 0,710. Hasil menunjukkan bahwa TPE menghasilkan performa deteksi yang lebih baik dibandingkan OFAT pada metrik evaluasi.

2. Berdasarkan Dataset dibangun dengan merepresentasikan variasi visual dari setiap tingkat kematangan apel, bukan dengan menganggap satu varietas sebagai satu kelas kematangan. Dataset baru menggunakan 3.344 citra hasil pengembangan dari 1.296 citra melalui augmentasi, dengan tiga kelas yaitu apple-20 (belum matang), apple-75 (setengah matang), dan apple-100 (matang). Strategi anotasi diperbaiki dengan memberikan label berdasarkan karakteristik warna yang menunjukkan tingkat kematangan, sehingga satu varietas seperti Fuji dapat masuk ke kelas yang berbeda sesuai kondisi kematangannya. Dataset juga mencakup variasi varietas, warna, pencahayaan, sudut, orientasi, ukuran, dan latar belakang. Perubahan tersebut terbukti meningkatkan kemampuan generalisasi model, termasuk dalam mengenali apel matang berwarna hijau kekuningan atau hijau kemerahan yang sebelumnya sering salah diklasifikasikan.
3. Model TPE mampu menghasilkan banyak true positive dan sedikit kesalahan. Pada pengujian sistem secara langsung dengan feeder, dari 103 apel yang diproses dalam tiga kali pengujian, diperoleh 99 true positive, 3 false positive, dan hanya 1 false negative, dengan precision 0,971, recall 0,990, dan F1-score 0,980. Sistem juga mampu mempertahankan proses deteksi dan penyortiran secara real-time pada sekitar 30 FPS, dengan throughput pengujian sebesar 33–36 apel per menit dan performa klasifikasi yang tetap tinggi. Hasil ini menunjukkan peningkatan yang cukup signifikan dibandingkan penelitian sebelumnya yang menggunakan pendekatan OFAT

untuk optimasi hyperparameter, di mana pada pengujian throughput dengan durasi 60 detik, sistem hanya mampu memproses 22 dari 24 apel pada kecepatan 0,034 m/s (22 apel/menit dengan error 8,33%). Dengan throughput 33–36 apel per menit pada sistem berbasis TPE, capaian tersebut berada pada rentang yang sebanding dengan throughput tertinggi sistem berbasis OFAT, namun dengan tingkat error yang jauh lebih rendah, yaitu hanya sekitar 2,91% (3 false positive dari 103 apel) dibandingkan error 9,52% pada sistem sebelumnya. Hal ini mengindikasikan bahwa peningkatan akurasi model hasil optimasi TPE turut berkontribusi terhadap perbaikan keandalan mekanisme penyortiran secara keseluruhan, tidak hanya pada aspek akurasi deteksi dan klasifikasi semata, melainkan juga pada konsistensi eksekusi fisik feeder, conveyor, dan diverter dalam kondisi throughput tinggi.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dijadikan acuan untuk pengembangan penelitian selanjutnya.

1. Mengembangkan dataset dengan jumlah gambar yang lebih banyak serta variasi kondisi pencahayaan, ukuran, orientasi, dan tingkat kematangan apel agar model memiliki kemampuan generalisasi yang lebih baik pada kondisi operasional nyata.
2. Mengeksplorasi metode optimasi hyperparameter lain seperti Genetic Algorithm, Particle Swarm Optimization (PSO), atau Multi-Objective Tree-Structured Parzen Estimator (MOTPE) sehingga dapat dibandingkan efektivitasnya terhadap TPE dalam meningkatkan performa model YOLOv8.
3. Mengembangkan sistem menjadi multi-grade sorting dengan menambahkan kategori kualitas berdasarkan ukuran, bentuk, cacat permukaan, maupun tingkat kerusakan sehingga sistem tidak hanya melakukan klasifikasi kematangan tetapi juga grading kualitas buah secara menyeluruh.

4. Mengoptimalkan desain mekanik feeder dan diverter agar mampu menangani kapasitas yang lebih tinggi dengan tetap mempertahankan kestabilan aliran apel serta meminimalkan potensi kemacetan selama proses sortir berlangsung.
5. Mengimplementasikan sistem monitoring berbasis antarmuka web atau Internet of Things (IoT) sehingga data hasil klasifikasi, jumlah apel pada setiap kategori, serta performa sistem dapat dipantau dan direkam secara real-time untuk keperluan analisis maupun evaluasi operasional.