

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil temuan dari penelitian ini dapat disimpulkan beberapa hal sebagai berikut:

1. *Trade-off* antara ukuran model, kecepatan inferensi, dan akurasi deteksi menunjukkan hasil yang bervariasi tergantung pada format target dan tingkat *quantization* yang diterapkan. Representasi FP32 dan FP16 mempertahankan akurasi yang konsisten dengan nilai $mAP@0.5$ stabil pada 99.5% di seluruh format, menunjukkan bahwa *quantization* FP16 tidak menyebabkan penurunan akurasi yang berarti. Sementara itu, *quantization* ke INT8 memberikan hasil yang lebih beragam. TensorRT-INT8 mencatatkan waktu inferensi tercepat 3.5 ms dengan pengurangan ukuran model hingga 48.2%, namun mengalami penurunan $mAP@0.5:0.95$ menjadi 75.5%. OpenVINO-INT8 justru menunjukkan peningkatan performa dengan $mAP@0.5:0.95$ mencapai 76.7% dan *F1-score* sebesar 99.9% karena efek regularisasi *quantization*. Sementara TFLite-INT8 mengalami penurunan $mAP@0.5:0.95$ menjadi 67.3% dan *precision* turun ke 95.5%, namun mempertahankan *recall* sebesar 100% yang menguntungkan untuk aplikasi moderasi konten.
2. Evaluasi pengaruh *quantization* terhadap konsumsi sumber daya sistem pada perangkat seluler menggunakan TFLite menunjukkan efisiensi yang signifikan. Model TFLite-INT8 berhasil mengurangi beban CPU hingga 6.5%, menurunkan penggunaan RAM sebesar 25.2%, dan menghemat konsumsi daya baterai hingga 26.2% dibandingkan dengan versi FP32. Hasil ini membuktikan bahwa teknik *quantization* tidak hanya efektif dalam mengurangi ukuran model dan mempercepat inferensi, tetapi juga memberikan dampak positif yang substansial terhadap efisiensi energi dan penggunaan sumber daya pada perangkat seluler, menjadikannya solusi yang ideal untuk implementasi deteksi iklan judi *online* secara *real-time*.

5.2 Saran

Berikut adalah beberapa rekomendasi yang dapat diperbaiki dalam melakukan pengembangan model dan dapat dilakukan pada penelitian selanjutnya:

1. Penelitian ini menggunakan *dataset* yang berfokus pada konten Instagram dengan jumlah dan variasi yang terbatas. Disarankan untuk memperluas *dataset* dengan menambahkan sampel dari platform media sosial lainnya seperti Facebook, TikTok, dan YouTube, serta meningkatkan diversitas konten iklan judi *online* untuk meningkatkan generalisasi model.
2. Penelitian ini melakukan evaluasi pada satu jenis perangkat Android kelas menengah. Disarankan untuk menguji performa model pada berbagai perangkat seluler dengan spesifikasi perangkat keras yang berbeda, termasuk perangkat iOS, untuk memvalidasi konsistensi performa dan mengidentifikasi optimalisasi spesifik platform. Evaluasi pada perangkat dengan sumber daya yang lebih terbatas juga penting untuk memastikan bahwa model dapat diterapkan secara luas di berbagai platform.
3. Penelitian ini menggunakan pendekatan deteksi objek berbasis YOLO untuk mengenali logo judi *online* secara visual. Namun, beberapa konten iklan judi juga memuat elemen teks yang dapat lolos dari deteksi visual murni. Oleh karena itu, disarankan untuk menggabungkan metode deteksi objek (seperti YOLO) dengan *Optical Character Recognition* (OCR) guna meningkatkan cakupan deteksi, khususnya terhadap teks promosi atau alamat situs yang tersembunyi di dalam gambar. Kombinasi ini dapat memperkuat sistem dalam mendeteksi konten iklan judi yang bersifat tekstual maupun visual secara bersamaan. Penelitian selanjutnya dapat mengeksplorasi integrasi *pipeline* antara model deteksi objek dan OCR, serta mengukur performanya terhadap konten campuran yang lebih kompleks.