

BAB I PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi transportasi dan meningkatnya jumlah kendaraan bermotor menyebabkan kebutuhan terhadap sistem keselamatan berbasis teknologi menjadi semakin penting. Salah satu teknologi yang berkembang pesat adalah sistem deteksi kendaraan yang mampu mengenali objek kendaraan secara otomatis melalui kamera. Teknologi ini banyak diterapkan pada sistem bantuan pengemudi maupun sistem pemantauan lalu lintas untuk meningkatkan keamanan dan kewaspadaan pengguna jalan. Seiring perkembangan kecerdasan buatan, sistem deteksi objek *real-time* menjadi solusi yang efektif untuk membantu mengenali kendaraan di sekitar pengguna jalan secara cepat dan otomatis (Ren dkk., 2018).

Real-time Object Detection (RTOD) merupakan satu teknologi yang dapat membantu mengenali objek disekitar secara *real-time* adalah sistem *RTOD* yang telah banyak diterapkan dalam kendaraan untuk meningkatkan keselamatan, seperti sistem pengereman otomatis dan peringatan tabrakan (Yurtsever dkk., 2020). Sistem ini memanfaatkan sensor dan algoritma yang cepat untuk mengidentifikasi objek di sekitar kendaraan dan mengambil tindakan preventif (Chen dkk., 2016). Penerapan teknologi serupa dalam alat bantu bagi pejalan kaki dapat memberikan manfaat yang signifikan, membantu mereka mengenali kendaraan di sekitarnya dan memberikan peringatan akan kendaraan yang akan mendekat.

Pada sistem *RTOD* terdapat proses untuk menentukan di mana lokasi sebuah objek berada dan mengenali objek yang terdeteksi atau mengkategorikan objek ke dalam kategori tertentu (Redmon dkk., 2016). Pada penelitian Xiao dkk. (2020) terdapat perkembangan *model* deteksi objek dimulai dari *Two stage detector* yang menggunakan dua tahap proses untuk mendeteksi yaitu menghasilkan *bounding box* dan setiap *bounding box* akan diklasifikasikan sesuai kategori kelasnya, contohnya seperti *Region-based Convolutional Neural Network (R-CNN)*, *Fast Region-based Convolutional Neural Network (Fast R-CNN)*, *Faster Region-based Convolutional Neural Network (Faster R-CNN)* , dan *Mask Region-based Convolutional Neural Network (Mask R-CNN)*. Sedangkan *one stage detector* proses untuk mendapatkan *bounding box* dan mengklasifikasikan objeknya dilakukan di proses yang sama,

contohnya seperti, *You Only Look Once (YOLO)*, *Faster R-CNN*, dan *Single Shot Detector (SSD)*. Dari *model* deteksi objek yang dipaparkan sebelumnya *One stage detector* lebih cepat dari pendahulunya yaitu *Two stage detector*.

YOLO merupakan salah satu *model One stage detector*, perkembangannya mulai dari *YOLOv1*, *YOLOv2*, *YOLOv3-Tiny*, *YOLOv4*, *YOLOv4-Tiny*, *YOLOv5*. Dari penelitian Gündüz & Işık (2023) pada uji coba dengan computer *GPU model* *YOLOv5* dan *YOLOv3* memberikan hasil yang lebih bagus diantara *model* lainnya. Sedangkan untuk *YOLOv3-Tiny* dan *YOLOv4-Tiny* memiliki performa yang tinggi dibandingkan *model* *YOLOv5* dan *YOLOv3*. Penelitian Bacea & Oniga (2023) mengungkapkan untuk deployment pada perangkat *mobile* atau dengan sumber daya kecil dapat menurunkan akurasi dan performa dari *model*. Sedangkan untuk pejalan kaki performa yang lebih cepat dibutuhkan untuk perangkat *mobile* dan untuk aplikasi *real-time* untuk mendeteksi objek kendaraan roda empat.

Keberhasilan sistem deteksi objek tidak hanya ditentukan oleh pemilihan *model*, tetapi juga dipengaruhi oleh konfigurasi parameter yang digunakan selama proses pelatihan dan inferensi. Dua parameter yang memiliki peran penting adalah *threshold confidence* dan *anchor box*. Nilai *threshold confidence* digunakan untuk menentukan tingkat *confidence* minimum agar suatu objek dianggap valid sebagai hasil deteksi, sedangkan *anchor box* berfungsi sebagai referensi awal dalam proses prediksi lokasi objek (Redmon & Farhadi, 2016a). Penentuan nilai yang kurang tepat dapat menyebabkan meningkatnya kesalahan deteksi, baik berupa objek yang tidak terdeteksi maupun munculnya deteksi palsu. Oleh karena itu, diperlukan pengujian untuk memperoleh kombinasi *threshold confidence* dan konfigurasi *anchor box* yang mampu menghasilkan keseimbangan antara tingkat akurasi deteksi yang diukur menggunakan *Mean Average Precision (mAP)* dan kecepatan pemrosesan yang diukur menggunakan *Frames per Second (FPS)* (R. Huang dkk., 2018).

Keterbatasan kemampuan perangkat *mobile* dalam melakukan pemrosesan data secara *real-time* menjadi tantangan utama dalam implementasi sistem deteksi objek. Perangkat *mobile* memiliki keterbatasan pada kapasitas memori, daya komputasi, serta konsumsi energi dibandingkan perangkat komputasi dengan spesifikasi tinggi. Akibatnya, proses deteksi objek secara terus-menerus dapat

menurunkan performa aplikasi dan menghasilkan latensi yang tinggi. Yang mana aplikasi berbasis keselamatan membutuhkan waktu respons yang cepat agar mampu memberikan informasi secara *real-time* kepada pengguna (J. Huang dkk., 2019). Banyak penelitian sistem deteksi kendaraan lebih berfokus pada peningkatan tingkat akurasi deteksi tanpa mempertimbangkan efisiensi komputasi pada perangkat *mobile*. *Model* dengan akurasi tinggi umumnya memiliki ukuran parameter yang besar dan memerlukan proses komputasi yang kompleks sehingga kurang sesuai pada perangkat dengan spesifikasi rendah. Oleh karena itu, diperlukan *model* deteksi objek yang mampu menjaga keseimbangan antara kecepatan deteksi dan tingkat akurasi agar dapat digunakan secara optimal pada sistem berbasis *mobile* (J. Huang dkk., 2019).

Meskipun perangkat *mobile* modern telah dilengkapi dengan *Graphical Processing Unit (GPU)* untuk mendukung kebutuhan grafis dan multimedia, tidak seluruh *GPU mobile* dapat dimanfaatkan secara optimal untuk proses komputasi deep learning. Hal ini terjadi karena pada implementasi *on-device inference*, proses deteksi objek umumnya memanfaatkan *framework* seperti *TensorFlow Lite* yang menggunakan *GPU Delegate* atau *Neural Networks API (NNAPI)* untuk melakukan akselerasi inferensi. Namun, kompatibilitas akselerasi tersebut sangat bergantung pada dukungan arsitektur *hardware*, *driver*, serta kemampuan komputasi *GPU* yang dimiliki perangkat. Beberapa *GPU* pada perangkat *mobile* kelas rendah hanya dirancang untuk kebutuhan rendering grafis sehingga tidak memiliki dukungan yang optimal terhadap *pipeline* komputasi *neural network*. Akibatnya, proses inferensi tidak dapat dijalankan melalui *GPU* dan sistem secara otomatis melakukan *fallback* ke *Central Processing Unit (CPU)* sebagai perangkat komputasi utama. Kondisi ini menyebabkan performa deteksi objek *real-time* menjadi terbatas karena seluruh proses inferensi bergantung pada kemampuan *CPU* perangkat yang relatif lebih rendah dibandingkan *GPU* dengan dukungan komputasi yang memadai (Lane dkk., 2016; Google, 2024).

Penggunaan *GPU Delegate* pada *TensorFlow Lite* juga dipengaruhi oleh kompatibilitas operasi pada *model* deep learning yang digunakan. Tidak seluruh layer atau operasi dalam *model* deteksi objek dapat dieksekusi melalui *pipeline GPU mobile*, terutama pada perangkat dengan spesifikasi rendah. Ketika terdapat

operasi yang tidak didukung, proses inferensi akan dialihkan kembali ke *CPU* sehingga menyebabkan peningkatan latency dan penurunan performa aplikasi *real-time*. Oleh karena itu, pengembangan sistem deteksi objek berbasis *mobile* perlu mempertimbangkan kompatibilitas serta optimasi *model* agar dapat berjalan secara efisien pada perangkat dengan sumber daya terbatas (Howard dkk., 2017; R. Huang dkk., 2018).

Dari keterbatasan kompatibilitas *GPU* perangkat *mobile* dalam melakukan pemrosesan data secara *real-time* jadi *model YOLO-LITE* dipilih karena metode *Real-Time Object Detection* atau bisa disingkat *RTOD One stage detector* ini yang dikembangkan untuk perangkat *non-GPU* seperti perangkat *mobile*. Metode yang menjadi basis *YOLO-LITE* yaitu *YOLOv2* yang juga digunakan untuk mendeteksi objek namun dengan masih memerlukan sumber daya komputasi yang relatif besar dan tidak optimal untuk perangkat dengan keterbatasan sumber daya seperti perangkat *mobile* (R. Huang dkk., 2018).

Kemudian, informasi mengenai keberadaan kendaraan saja belum cukup untuk mendukung sistem keselamatan pejalan kaki secara optimal. Sistem juga perlu mengetahui seberapa dekat kendaraan tersebut terhadap posisi pengguna agar tingkat risiko dapat dinilai dengan lebih baik. Informasi jarak memungkinkan sistem memberikan peringatan yang lebih relevan berdasarkan kedekatan kendaraan yang terdeteksi. Dengan demikian, integrasi antara deteksi objek dan estimasi jarak dapat menghasilkan sistem yang tidak hanya mampu mengenali tipe kendaraan, tetapi juga memberikan informasi yang lebih bermakna untuk mendukung pengambilan keputusan dan meningkatkan kewaspadaan pengguna terhadap potensi bahaya di sekitarnya. Penggunaan estimasi jarak menggunakan kamera *monocular* juga masih memiliki berbagai tantangan, terutama dalam menjaga tingkat akurasi terhadap berbagai jenis kendaraan. Perbedaan dimensi kendaraan seperti Hatchback, SUV, MPV, Pickup, dan Van dapat memengaruhi hasil estimasi jarak, karena tinggi objek yang terdeteksi pada citra berbeda-beda. Oleh karena itu, diperlukan pendekatan *multicalibration monocular distance estimation* untuk meningkatkan akurasi estimasi jarak berdasarkan tipe kendaraan yang terdeteksi (Zhong dkk., 2024).

Pada proses estimasi jarak berbasis kamera *monocular*, pendekatan *multicalibration monocular distance estimation* diperlukan untuk meningkatkan akurasi estimasi terhadap berbagai jenis kendaraan. Metode *monocular distance estimation* umumnya menggunakan satu parameter kalibrasi untuk seluruh objek sehingga menghasilkan tingkat kesalahan yang cukup tinggi ketika diterapkan pada kendaraan dengan dimensi berbeda. Perbedaan tinggi kendaraan seperti Hatchback, SUV, MPV, Pickup, dan Van menyebabkan perubahan proporsi objek pada citra sehingga memengaruhi hasil estimasi jarak. Dengan pendekatan *multicalibration*, setiap kategori kendaraan dapat menggunakan parameter kalibrasi yang berbeda sesuai karakteristik dimensinya sehingga estimasi jarak menjadi lebih adaptif dan akurat pada berbagai kondisi lingkungan (Zhong dkk., 2024).

Selain parameter tinggi objek referensi, *focal length* kamera merupakan komponen penting dalam proses *monocular distance estimation*. Dalam *model* kamera *pinhole*, hubungan antara ukuran objek pada dunia nyata dan ukuran objek pada citra dipengaruhi secara langsung oleh nilai *focal length*. Semakin besar nilai *focal length*, semakin besar pula ukuran objek yang diproyeksikan pada sensor untuk jarak yang sama. Oleh karena itu, ketepatan nilai *focal length* sangat berpengaruh terhadap akurasi hasil estimasi jarak yang diperoleh dari citra (Zhong dkk., 2024).

Berdasarkan permasalahan tersebut, penelitian ini berfokus pada pengembangan *model RTOD* untuk objek 5 tipe kendaraan roda empat dengan parameter *threshold* dan *anchor box* optimal yang digunakan untuk sistem deteksi di perangkat *mobile* dengan hasil *Mean Average Precision* dan *Frames Per Second* terbesar. Serta kalibrasi *focal length* terbaik berdasarkan nilai *Mean Average Precentage Error* pada aplikasi yang digunakan untuk mengukur jarak antara kamera dari prespektif pejalan kaki dengan kendaraan roda empat.

1.2 Perumusan Masalah

Berdasarkan uraian penjelasan latar belakang tersebut, maka penulis mendapatkan rumusan masalah sebagai berikut:

1.2.1 Permasalahan

Berdasarkan latar belakang, permasalahan dalam penelitian ini berfokus pada bagaimana menerapkan metode *YOLO-LITE* untuk mendeteksi kendaraan roda empat secara *real-time* pada perangkat *mobile* dengan sumber daya komputasi terbatas, bagaimana mengimplementasikan metode *Monocular Camera Distance Estimation* untuk memperkirakan jarak kendaraan roda empat menggunakan satu kamera pada perangkat *mobile*, serta bagaimana menerapkan *multicalibration* untuk perhitungan *focal length* pada lima tipe kendaraan roda empat guna meningkatkan akurasi estimasi jarak pada sistem *Monocular Camera Distance Estimation*.

1.2.2 Metode Usulan

Metode yang diusulkan pada penelitian ini terdiri atas dua tahapan utama, yaitu deteksi objek dan estimasi jarak. Tahap pertama menggunakan metode *YOLO-LITE* untuk mendeteksi serta mengklasifikasikan lima tipe kendaraan roda empat, yaitu Hatchback, MPV, SUV, Pickup, dan Van. Metode *YOLO-LITE* dipilih karena merupakan *model* deteksi objek ringan yang dirancang untuk perangkat dengan keterbatasan sumber daya komputasi.

Tahap kedua menggunakan metode *Monocular Camera Distance Estimation* untuk memperkirakan jarak antara kamera perangkat *mobile* dan kendaraan yang terdeteksi. Proses estimasi jarak dilakukan berdasarkan informasi hasil deteksi objek serta parameter kalibrasi yang diperoleh melalui pendekatan *multicalibration* pada parameter *focal length*.

1.2.3 Pertanyaan Penelitian

Dari rumusan masalah diatas didapatkan pertanyaan penelitian sebagai berikut:

1. Berapa nilai *threshold confidence* serta *anchor box* yang menghasilkan *Mean Average Precision* dan *Frames Per Second* optimal dari *model RTOD YOLO-LITE* untuk mendeteksi kendaraan roda empat pada perangkat *mobile* dengan 5 variasi tipe kendaraan roda empat?
2. Berapa nilai *focal length* optimal berdasarkan nilai *Mean Average Percentage Error (MAPE)* pada kalibrasi jarak asli 250cm, 300cm, 350cm, 400cm, 450cm, 500cm, untuk pengukuran estimasi jarak dari hasil deteksi *model RTOD YOLO-LITE* pada perangkat *mobile* dengan setiap tipe objek kendaraan roda empat?

1.3 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan sebagai berikut:

1. Untuk menentukan nilai *threshold confidence* serta *anchor box* dengan nilai *Mean Average Precision* dan *Frames Per Second* optimal guna memperoleh performa terbaik *model RTOD YOLO-LITE* dalam mendeteksi kendaraan roda empat pada perangkat *mobile*.
2. Mengetahui nilai kalibrasi *focal length* optimal berdasarkan nilai *Mean Average Percentage Error* untuk jarak asli pengukuran 250cm, 300cm, 350cm, 400cm, 450cm, 500cm dari hasil deteksi *model RTOD YOLO-LITE* pada perangkat *mobile* untuk setiap tipe objek kendaraan roda empat.

1.4 Batasan-batasan

Batasan-batasan masalah untuk penelitian ini adalah sebagai berikut:

1. Kalibrasi hanya dilakukan pada kendaraan roda empat tipe *MPV*, *Hatchback*, *Van*, *Pickup* dan *SUV* saja dengan tinggi rata-rata dari setiap tipe.
2. Penelitian ini menggunakan *dataset* citra dengan jumlah 2250 yang terdiri dari kendaraan roda empat bertipe *MPV*, *Hatchback*, *Van*, *Pickup* dan *SUV*. *Dataset* juga menyertakan anotasi berupa *bounding box* untuk lokasi dari kendaraan roda empat serta kelas tipe kendaraan roda empat.
3. Pengukuran jarak hanya dilakukan menggunakan *built-in single camera* dari perangkat *mobile* serta jarak yang sudah ditentukan.
4. Resolusi *input* kamera yang digunakan untuk uji coba adalah 640x480 *pixel*.

5. Jarak yang digunakan pada eksperimen adalah 250cm, 300cm, 350cm, 400cm, 450cm, 500cm.

1.5 Sistematika Penulisan Skripsi

Berikut sistematika penulisan laporan yang disesuaikan dengan format Skripsi Program Studi Teknik Informatika Fakultas Teknik Universitas Trunodjoyo Madura:

BAB I : PENDAHULUAN

Pada bab ini diuraikan latar belakang masalah sehingga dilakukan penelitian, metode usulan yang digunakan dalam penelitian, tujuan dan manfaat, serta batasan-batasan yang terdapat dalam penelitian ini.

BAB II : KAJIAN PUSTAKA

Pada bab ini dijelaskan mengenai *Object Detection*, *YOLO-LITE*, *YOLOv2*, *Grid*, *Intersection over Union (IoU)*, *Anchor box*, *K-Means* untuk menentukan ukuran *Anchor box*, *Elbow Method* untuk menentukan jumlah dari *Anchor box* yang Optimal, Arsitektur *YOLO-LITE*, Perhitungan Akurasi, *Non Maximum Supression*, *Monocular Camera Distance Estimation*, Integrasi Model untuk Aplikasi *Mobile*, Merek, Tipe, & Model kendaraan roda empat, dan penelitian terkait yang menjadi acuan dalam penulisan proposal ini.

BAB III : METODE USULAN

Pada bab ini diuraikan metode usulan yang akan digunakan pada penelitian ini mencakup detail tahapan-tahapan yang dilakukan untuk menyelesaikan permasalahan deteksi dengan metode *YOLO-LITE* dan pengukuran jarak dengan metode *Monocular Camera Distance Estimation*.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini berisi hasil dan pembahasan penelitian yang telah dilakukan ini mencakup detail tahapan-tahapan yang dilakukan untuk menyelesaikan permasalahan deteksi dengan metode *YOLO-LITE* dan pengukuran jarak dengan metode *Monocular Camera Distance Estimation*.

BAB V : KESIMPULAN DAN SARAN

Bab ini berisi mengenai kesimpulan yang telah didapatkan selama penelitian berlangsung serta saran-saran untuk penelitian selanjutnya.

REFERENSI

Pada halaman referensi berisi tentang daftar referensi atau pustaka yang dijadikan sebagai acuan dalam penelitian.