

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada beberapa tahun terakhir, komunitas *computer vision* sudah banyak berkembang dalam menjalankan tugas yang sulit. Teknik *deep learning* sekarang banyak mendemonstrasikan performanya yang sangat baik dalam deteksi obyek pada citra. Tetapi pelacakan selalu memberikan tantangan apalagi jika membahas tentang pelacakan banyak objek [1]. *Multiple Object Tracking* merupakan suatu metode untuk melacak objek yang bergerak dan jumlahnya lebih dari satu [2]. *Multiple Object Tracking* (MOT), juga disebut *Multi-Target Tracking* (MTT), adalah tugas *computer vision* yang bertujuan menganalisis video untuk mengidentifikasi dan melacak objek yang termasuk dalam satu atau beberapa kategori, seperti pejalan kaki, mobil, hewan dan benda mati, tanpa pengetahuan sebelumnya tentang penampilan dan jumlah target. Berbeda dari algoritma deteksi objek, yang keluarannya adalah kumpulan kotak pembatas persegi yang diidentifikasi olehnya koordinat, tinggi dan lebar, algoritma MOT juga mengaitkan ID target ke setiap kotak (dikenal sebagai deteksi), untuk membedakan antara objek intra-kelas [3]. MOT bertujuan untuk memprediksi lintasan beberapa target dalam video yang berurutan. Penerapan MOT sangat beragam mulai dari *autonomous driving* hingga *smart video analysis* [4]. Pengaplikasian MOT dalam dunia nyata banyak digunakan seperti untuk video keamanan, pengenalan gerakan, penglihatan robot, dan interaksi robot manusia [5]. Representasi objek merupakan faktor penting dalam membangun algoritma pelacakan objek visual yang kuat [6].

Dibalik banyaknya perkembangan MOT yang sangat baik pada bidang *computer vision*, MOT juga memiliki permasalahan - permasalahan umum seperti *drift* dari *tracking points* yang dikarenakan banyaknya variasi penampilan yang disebabkan oleh *noise*, *illumination*, pose, latar belakang yang berantakan, interaksi, oklusi, dan gerakan kamera. Sulit untuk menerapkan algoritma pelacakan visual yang sangat baik dalam hal akurasi dan ketahanan [7]. Sebagian besar metode MOT mengalami masalah terlalu banyak mengidentifikasi objek dalam jumlah

yang besar, dan menyebabkan penurunan kinerja dan gangguan akurasi pelacakan di latar belakang yang berantakan [5]. Selain itu kinerja metode MOT sangat bergantung pada detektor objek. Misalnya, pelacakan dengan deteksi yang banyak digunakan Paradigma dibangun di atas detektor objek, yang menyediakan deteksi hipotesis untuk mendorong prosedur pelacakan. Jika diberikan set yang berbeda dari hipotesis deteksi sambil memperbaiki komponen yang lain, pendekatan yang identik akan menghasilkan pelacakan hasil dengan perbedaan kinerja yang signifikan. Di komunitas, terkadang tidak ada penjelasan tentang modul deteksi yang diberikan dalam pendekatan. Ini membuat perbandingan dengan pendekatan lain tidak layak. Tolak ukur yang mapan seperti KITTI dan *MOTChallenge* berusaha untuk meringankan masalah ini dan juga bergerak ke arah yang lebih berprinsip dan evaluasi terpadu dari deteksi dan pelacakan [8]. Beberapa metode sudah dilakukan seperti menggunakan *Fast Region-Based Convolutional Network* (*Fast R-CNN*) tetapi terbatas karena kecepatan komputasinya, saat komputasi berlangsung program hanya berjalan kurang dari 10 *frame per second*(FPS) yang menyebabkan tidak optimalnya untuk dijalankan di kondisi *real-time* [4]. Pada contoh lain *Intra-category occlusion* dan kemiripan objek dapat menyebabkan ambiguitas dalam asosiasi data. Beberapa isyarat, termasuk gerakan, bentuk dan penampakan objek, digabungkan untuk mengurangi masalah ini. Di sisi lain, deteksi hasil tidak selalu dapat diandalkan. Variasi pose dan oklusi dalam adegan yang ramai sering menyebabkan kegagalan deteksi seperti positif palsu, deteksi hilang, dan pembatas tidak akurat. Metode *Region-based Fully Convolutional Network* (R-FCN) sudah hampir menyelesaikan permasalahan diatas dengan menambahkan *fitur deeply learned person re-identification* (ReID). Meskipun metode R-FCN ini sudah membuat proses komputasi menjadi lebih cepat 2-3 kali lipat tetapi untuk membuatnya optimal pada kondisi *real-time* itu masih belum cukup baik dan penggunaan metode tersebut hanya dapat mendeteksi manusia saja [9]. *You only look once* (YOLO) adalah algoritma canggih yang menggunakan *Convolutional Neural Network* (CNN) untuk melakukan deteksi objek. YOLOv4 memiliki tingkat AP dan FPS yang cukup tinggi dibandingkan dengan *object detector* lainnya, *object detector* lain seperti *EfficientDet* memiliki tingkat AP yang tinggi tetapi memiliki nilai FPS yang rendah. Penggunaan YOLOv4 sangat tepat

untuk dijadikan pilihan karena memiliki tingkat nilai AP dan FPS yang keduanya baik, dengan kondisi dimana dibutuhkannya FPS yang tinggi seperti pelacakan mobil hingga yang rendah seperti pelacakan manusia biasa YOLOv4 dapat melakukannya dengan mendapatkan hasil yang baik. Dengan menggunakan *output* dari umpan YOLOv4 pendeteksian objek ini ke *Simple Online dan Realtime Tracking dengan Deep Association Metric (Deep SORT)* akan didapatkan hasil pelacakan objek yang akurat dan komputasi yang ringan. *Simple Online dan Realtime Tracking (SORT)* mendapatkan nilai yang baik dalam bagian pelacakan presisi dan akurasi pada MOT Challenge, tetapi SORT menghasilkan pengalih identitas yang relatif dalam jumlah besar. Ini karena dipekerjakan metrik asosiasi hanya akurat ketika estimasi ketidakpastian rendah. Oleh karena itu metode SORT ini dikembangkan dengan mengganti metrik asosiasi dengan metrik yang lebih terinformasi yang menggabungkan informasi gerakan dan tampilan, dengan ini metode SORT berubah menjadi Deep SORT. Deep Sort memiliki nilai MOTA yang relatif kecil dibanding object tracking lainnya [10]. Penggabungan Metode YOLOv4 dan Deep SORT sebagai sebuah metode dalam MOT dapat menutupi kelemahan dari Deep SORT yang memiliki nilai MOTA yang relatif kecil.

Pada Tugas akhir ini mengusulkan untuk dilakukannya eksploitasi pada parameter dan arsitektur Deep SORT. Dari hasil yang diperoleh didapatkan bahwa metode *eksploitasi* ini berhasil meningkatkan parameter performansi pada MOTA, MOTP dan kecepatan proses. Penulis memilih Deep SORT sebagai bahan untuk di eksploitasi karena penulis yakin Deep SORT masih dapat dikembangkan lebih jauh agar metode Deep SORT kedepannya dapat mengeluarkan hasil akurasi dan presisi yang lebih baik dan juga berat komputasi yang lebih ringan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka rumusan masalah yang akan di bahas pada Tugas Akhir ini adalah:

1. Bagaimana cara melakukan eksploitasi parameter untuk dapat meringankan beban komputasi agar dapat meningkatkan kecepatan proses?
2. Bagaimana cara mengeksplotasi metode Deep SORT agar mendapatkan nilai yang lebih baik pada parameter akurasi dan presisi?

3. Bagaimana cara mengukur dan menganalisis parameter performansi akurasi, presisi dan kecepatan proses untuk kasus pelacakan *multi-object tracking*?

1.3 Tujuan Penelitian

Sementara itu tujuan dari penelitian Tugas Akhir ini adalah sebagai berikut:

- a. Melakukan eksploitasi parameter untuk dapat meringankan beban komputasi agar dapat meningkatkan kecepatan proses.
- b. Mengeksplotasi metode Deep SORT agar mendapatkan nilai yang lebih baik pada parameter akurasi dan presisi.
- c. Mengukur dan menganalisis parameter performansi performansi akurasi, presisi dan kecepatan proses untuk kasus pelacakan *multi-object tracking*.

1.4 Batasan Masalah

Pada penelitian di Tugas Akhir memiliki batasan masalah yang membatasi penelitiannya, antara lain:

1. Dataset yang digunakan merupakan seluruh dataset *official pre-trained YOLOv4* yang *output* nya dibatasi menjadi 4 kelas yaitu: *person, car, cat and dog*.
2. Simulasi metode dilakukan dengan menggunakan program *python*.
3. Spesifikasi *tools* yang digunakan dalam penelitian ini sebagai berikut:
 - *Python 3.7*
 - *Tensorflow 2.3.0*
 - *GPU GTX 1660 TI*
4. Pengukuran dan analisis yang dipakai adalah *Multi-Object Tracking Accuracy (MOTA)*, *Multi-Object Tracking Precision (MOTP)* dan *Processing speed (FPS)*.

1.5 Metode Penelitian

Metode yang digunakan dalam menyelesaikan penelitian di Tugas Akhir dengan pendekatan seperti:

1. Studi literatur

Tahap ini berupa pengumpulan referensi berupa *paper*, jurnal, *conference*, buku, artikel maupun referensi lain yang berhubungan dengan *object detection*, *object tracking*, *multiple object tracking*, *deep SORT*, *YOLO*, dan *evaluating MOTA and MOTP*.

2. Perancangan sistem

Sistem ini dirancang sesuai dengan flowchart yang telah dibuat. Untuk merealisasikan system tersebut menggunakan *python 3.7*, *tensorflow 2.3.0* dan aplikasi *anaconda*.

3. Proses eksplotasi metode

Tahap ini dilakukan dengan mengeksplotasi metode pada Parameter performansi yang telah ditetapkan yaitu antara lain seperti *batch size*, *matching threshold*, *n init*, *max age* dan *max cosine distance*. Parameter tersebut diubah nilainya secara berulang kali hingga dapat mendapatkan performansi akurasi dan kecepatan komputasi yang lebih baik dari sebelumnya.

4. Pengujian metode eksploitasi

Tahap ini dilakukan dengan menggunakan MOTA dan MOTP metic, ini adalah program untuk menghitung MOTA dan MOTP sesuai dengan standar *MOTChallenge* yang berlaku untuk mendapatkan nilai dari MOTA dan MOTP dari hasil test yang telah dieksploitasi. Untuk pengujian kecepatan proses dilakukan pada saat proses *tracking* berlangsung dengan menambahkan fitur pada program agar dapat menampilkan hasil perhitungan rata - rata FPS.

5. Analisis performansi hasil pengujian

Tahap ini menganalisis kinerja system berdasarkan parameter yang telah ditentukan sebelumnya yaitu MOTA, MOTP, dan kecepatan proses

6. Kesimpulan

Tahap ini dilakukan penyusunan laporan dari berbagai percobaan dan analisis, exsploitasi terbaik telah terpilih untuk metode Deep SORT ini. Dimana pada tahap ini akan dilakukan penarikan kesimpulan dari berbagai skenario.

1.6 Sistematika Penulisan

Sistematika penulisan dalam Tugas Akhir ini adalah sebagai berikut:

1. BAB 2 KONSEP DASAR

Bab ini membahas terkait konsep dasar dan tinjauan pustaka yang digunakan dalam penelitian seperti *multiple object tracking*, YOLO, Deep SORT, CNN, Bahasa Python, dan Tensorflow

2. BAB 3 PERANCANGAN SISTEM

Bab ini menjelaskan desain sistem yang akan dirancang, perancangan sistem, parameter performansi, dan spesifikasi perangkat yang digunakan.

3. BAB 4 HASIL DAN ANALISIS

Bab ini berisi data hasil pengujian sistem yang dilakukan dan analisis hasil pengujian yang didapat yaitu MOTA, MOTP dan kecepatan proses

4. BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari hasil analisis terhadap pengujian sistem dan saran untuk penelitian selanjutnya untuk meningkatkan performansi sistem *multiple object tracking*.