

SKRIPSI

**SISTEM DETEKSI KECEPATAN RELATIF KENDARAAN PADA
TEKNOLOGI KENDARAAN *AUTONOMOUS* MENGGUNAKAN
MEASUREMENT METHOD BERBASIS *STEREO CAMERA***

Disusun dan diajukan oleh:

MUHAMMAD FADHIL BIN BAHRUNNIDA

D121171523



DEPARTEMEN TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS HASANUDDIN

MAKASSAR

2022

LEMBAR PENGESAHAN SKRIPSI
SISTEM DETEKSI KECEPATAN RELATIF KENDARAAN PADA
TEKNOLOGI KENDARAAN AUTONOMOUS MENGGUNAKAN
MEASUREMENT METHOD BERBASIS STEREO CAMERA

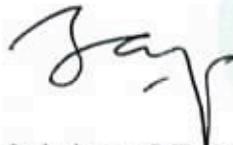
Disusun dan diajukan oleh
MUHAMMAD FADHIL BIN BAHRUNNIDA
D121171523

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Studi Program Sarjana Program Studi Teknik Informatika Fakultas Teknik Universitas Hasanuddin pada tanggal 01 November 2022 dan dinyatakan telah memenuhi syarat kelulusan.

Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,



Dr. Indrabayu, S.T., M.T., M.Bus.Sys.
Nip. 197507162002121004



A. Ais Prayogi, ST., M.Eng
Nip. 198305102014041001

Ketua Program Studi,



Dr. Indrabayu, S.T., M.T., M.Bus.Sys.
Nip. 19750716 200212 1 004

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Muh. Fadhil Bin Bahrunnida

NIM : D121171523

Departemen : Teknik Informatika

Jenjang : S1

Menyatakan dengan ini karya tulisan saya berjudul:

SISTEM DETEKSI KECEPATAN RELATIF KENDARAAN PADA
TEKNOLOGI KENDARAAN AUTONOMOUS MENGGUNAKAN
MEASUREMENT METHOD BERBASIS STEREO CAMERA

Adalah karya tulisan saya sendiri dan bukan merupakan pengambilalihan tulisan orang lain bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri.

Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini hasil karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Makassar, 31 Oktober 2022

Yang menyatakan,



Muh. Fadhil Bin Bahrunnida

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, yang telah memberikan rahmat dan karunia-Nya, sehingga dapat menyelesaikan tugas akhir dengan judul “Sistem Deteksi Kecepatan Relatif Kendaraan Pada Teknologi Kendaraan *Autonomous* Menggunakan *Measurement Method* Berbasis *Stereo Camera*” sebagai salah satu syarat dalam menyelesaikan jenjang Strata-1 di Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin.

Penulis menyadari bahwa penyusunan dan penulisan tugas akhir ini tidak dapat terselesaikan dengan baik tanpa adanya bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih banyak kepada:

1. Kedua orang tua penulis, Bapak Bahrunnida Amien dan Ibu Darmawati Umar yang selalu memberikan dukungan, doa, dan semangat yang tiada hentinya, serta selalu sabar dalam mendidik penulis sejak kecil.
2. Bapak Dr. Indrabayu, S.T., M.T., M.Bus.sys. selaku pembimbing I dan Bapak A. Ais Prayogi Alimuddin, S.T., M.Eng. selaku pembimbing II, yang senantiasa menyediakan waktu, tenaga, pikiran, dan perhatian yang luar biasa dalam mengarahkan penulis untuk menyelesaikan tugas akhir.
3. Ibu Dr. Eng Intan Sari Areni., ST., MT., serta ibu Anugrayani Bustamin, S.T., M.T. selaku dosen-dosen hebat non-pembimbing yang senantiasa mendorong penulis untuk selalu fokus dalam penelitian dan tidak pernah keberatan menerima konsultasi penulis.

4. Tim *Autonomous Car dan Data Mining Tuberkulosis Anak 2017 AIMP Unhas*; Nublan Azqalani, Alfarabi Alif, Herul Hidayat, Irzam Kasyfillah, Taslinda. Terimakasih sudah menjadi kawan yang tidak pernah bosan untuk saling menguatkan dan mendukung ditengah badai revisi yang tak kunjung usai, kalian luar biasa.
5. Sahabat penulis sejak duduk di sekolah dasar; Raihanul Abrar, Ibrahim bin Jahada, serta Usamah. Terimakasih karena tetap menjadi teman yang menyenangkan dan penuh tawa.
6. Fitriani Nasir yang sudah menjadi partner diskusi di hampir setengah bagian penelitian berjalan.
7. Muhammad Zulfahmi Sadrah dan Kanda Sofyan Tandungan yang senantiasa menjadi teman diskusi dan banyak membantu penulis dalam proses debugging kode yang ditulis.
8. Fauzan Alif Anwar kawan yang telah banyak berkontribusi dengan menyiapkan tempat dan kebutuhan peralatan dalam pembuatan sistem.
9. Saudari Ilmi Nurrahma Ismail dan Andi Priska Saskia yang telah banyak membantu penulis terkait hal-hal administratif pada lingkup tugas akhir.
10. Segenap staf dan dosen Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin yang telah membantu kelancaran penyelesaian tugas akhir.
11. Keluarga besar AIMP Research Group Unhas, yang senantiasa menjadi teman diskusi dan bertukar pikiran terkait problematika tugas akhir.

12. Teman-teman seangkatan RECOGN17ER yang telah menjadi bagian dari kehidupan kampus penulis, serta menjadi teman cerita dikala penat dalam proses penyusunan tugas akhir ini.

13. Serta pihak-pihak lain yang tidak disebutkan dan tanpa sadar telah menjadi inspirasi dan membantu penulis dalam menyelesaikan tugas akhir.

Penulis berharap semoga Tuhan Yang Maha Esa berkenan membalas segala kebaikan serta jasa dari semua pihak yang telah banyak membantu penulis dalam menyelesaikan tugas akhir ini. Penulis menyadari bahwa tugas akhir ini masih jauh dari kata sempurna, oleh karena itu penulis mengharapkan segala bentuk saran serta masukan yang membangun dari berbagai pihak. Semoga tugas akhir ini dapat memberi manfaat bagi para pembaca dan semua pihak. Aamiin.

Makassar, September 2022

Penulis,

Muh. Fadhil Bin Bahrunnida.

ABSTRAK

Autonomous Vehicle (AV) atau kendaraan tanpa awak merupakan teknologi masa depan yang sedang banyak dikembangkan perusahaan-perusahaan besar di dunia. Salah satu kemampuan dasar yang harus dimiliki oleh AV adalah kemampuan mendeteksi kecepatan kendaraan sekitar, sehingga dapat menjadi dasar keputusan kendaraan dalam bermanuver di segala jenis medan jalan. Penelitian ini mencoba menggabungkan teknologi *Computer vision* yang dapat memberikan informasi terkait objek disekitar dan digabungkan dengan model matematis yang mengadopsi prinsip-prinsip triangulasi (selanjutnya penelitian ini akan menggunakan istilah *Measurement Method* dalam model matematis yang diajukan) sehingga jarak antara kendaraan dapat diketahui secara *Real-time* dan selanjutnya menjadi dasar acuan untuk menghitung kecepatan kendaraan yang masuk dalam frame kamera. Hasil penelitian menunjukkan penggunaan metode *Stereo-YoloV5-Deppsort* yang diintegrasikan dengan model matematis *Measurement Method* mendapatkan skor MAPE 6,047% untuk deteksi dan estimasi jarak antara kendaraan dan skor MAPE 14,42% untuk deteksi dan estimasi kecepatan kendaraan.

Kata kunci: *Autonomous Vehicle, Stereo vision, YOLOv5, Deepsort, Measurement Method*

DAFTAR ISI

KATA PENGANTAR.....	ii
ABSTRAK	v
DAFTAR ISI	vi
DAFTAR GAMBAR	ix
DAFTAR TABEL	xii
BAB 1.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah.....	4
1.6 Metode Penulisan.....	5
1.7 Sistematika Penulisan.....	5
BAB 2.....	7
2.1 Autonomous Vehicle.....	7
2.2 Computer Vision	10
2.3 <i>Deep Learning</i>	12
2.4 Yolo V5	13
2.5 DeepSort.....	19
2. 6 <i>Stereo Camera</i>	24
2.7 <i>Stereo Vision System</i>	26
2.8 Kalibrasi <i>Stereo</i>	27
2.9 Rektifikasi <i>Stereo</i>	28

2.10 Pencocokan <i>Stereo</i>	29
2.11 Triangulasi.....	31
2.12 <i>Measurement Method</i>	33
BAB 3.....	35
3.1 Tahapan Penelitian	35
3.2 Waktu dan Lokasi Penelitian	36
3.3 Instrumen Penelitian.....	36
3.4 Teknik Pengambilan Data	37
3.5 Perancangan Sistem Deteksi dan <i>Tracking</i> Kendaraan.....	39
3.6 Perancangan <i>Stereo System</i>	44
3.6 Integrasi Sistem.....	50
3.7 <i>Measurement Method</i> untuk Deteksi Jarak Antar Kendaraan.....	57
3.8 <i>Measurement Method</i> untuk Deteksi Kecepatan.	61
3.9 Skenario Pengujian Sistem.....	63
3.10 Analisis Kinerja dan Evaluasi Sistem	65
BAB 4.....	69
4.1 Hasil dan Pembahasan Model Yolo V5 – DeepSORT.....	69
4.2. Hasil dan Pembahasan <i>Stereo Vision System</i>	72
4.3. Hasil dan Pembahasan Sistem Deteksi Jarak	80
4.4. Hasil dan Pembahasan Sistem Deteksi Kecepatan.....	87
4.5 Efisiensi Komputasi Program Stereo-YoloV5-Deeppsort.....	94
BAB 5.....	98
5.1 Kesimpulan	98
5.2 Saran.....	100
DAFTAR PUSTAKA.....	102

LAMPIRAN	109
----------------	-----

DAFTAR GAMBAR

Gambar 2. 1 Rangkaian proses utama pada Computer Vision.....	11
Gambar 2. 2 Perbandingan Yolov4 dengan beberapa algoritma.....	14
Gambar 2. 3 Perbandingan Yolov5 dengan EfficientDet.....	15
Gambar 2. 4 Arsitektur Yolo V5.....	17
Gambar 2. 5 Flowchart YoloV5 dalam mendeteksi objek.....	18
Gambar 2. 6 Alur kerja algoritma DeepSort (Avola et al., 2021).....	24
Gambar 2. 7 Proses utama pembentukan stereo vision.....	27
Gambar 2. 8 Model Kamera pinhole.....	28
Gambar 2. 9 Pencocokan stereo pada dua buah gambar dan hasil peta ke dalam gambar yang berisi peta disparitas	30
Gambar 2. 10 Triangulasi untuk menentukan posisi sebuah objek.....	31
Gambar 2. 11 Model geometri dasar perhitungan jarak pada kamera	32
Gambar 3. 1 Alur penelitian.....	35
Gambar 3. 2. Data primer (1), Data Youtube (2), Dataset Poribohon-BD(3).....	38
Gambar 3. 3 Proses augmentasi dengan LabelImg	38
Gambar 3. 4 Gambaran umum dataset yang digunakan	39
Gambar 3. 5 Gambar ter-anotasi dilaman roboflow.....	40
Gambar 3. 6 Tampilan hasil training yang telah dicatat <i>Weights&Biases</i>	41
Gambar 3. 7 Flowchart Proses kerja keseluruhan sistem deteksi kecepatan kendaraan	43
Gambar 3. 8 Contoh pola papan yang baik dan buruk dalam proses kalibrasi	45
Gambar 3. 9 Pola papan catur yang digunakan.....	46
Gambar 3. 10 Contoh pengambilan gambar secara stereo.....	47
Gambar 3. 11 Contoh hasil rektifikasi stereo.....	49
Gambar 3. 12 Contoh hasil pencocokan stereo.....	50
Gambar 3. 13 Potongan kode untuk mendeklarasikan port kamera.....	51
Gambar 3. 14 Potongan kode untuk menyimpan stream kamera kedalam nilai list/array.....	52

Gambar 3. 15 Potongan kode menunjukkan penampung data beberapa variabel disimpan dalam bentuk dictionary.	55
Gambar 3. 16 Potongan kode untuk menghitung kecepatan.....	57
Gambar 3. 13 Model dasar perhitungan Measurement Method secara stereo.	58
Gambar 3. 14 Model dasar perhitungan Measurement Method secara stereo.	59
Gambar 3. 19 Model simulasi penentuan adegan pengambilan kecepatan dengan Measurement Method.....	61
Gambar 3. 20 Skenario pengujian sistem deteksi jarak	64
Gambar 3. 17 Skenario pengujian sistem deteksi kecepatan	65
Gambar 4. 1 Grafik nilai mAP pada setiap step.....	70
Gambar 4. 2 Grafik nilai precision pada setiap step	70
Gambar 4. 3 Grafik nilai recall pada setiap step	71
Gambar 4. 4 Dua nilai estimasi jarak pada jarak aktual 9 meter(4.2).....	78
Gambar 4. 5 Struktur file yang berisi matriks peta transformasi undistorsi dan rektifikasi stereo kamera	79
Gambar 4. 6 Deteksi jarak pada mobil pada percobaan pertama (disparity error belum diperhitungkan)	81
Gambar 4. 7 Deteksi pada motor pada percobaan pertama (disparity error belum diperhitungkan)	81
Gambar 4. 8 Deteksi pada mobil pada percobaan terakhir (disparity error telah diperhitungkan)	82
Gambar 4. 9 Deteksi pada mobil pada percobaan terakhir (disparity error telah diperhitungkan)	82
Gambar 4. 10 Dua nilai estimasi jarak pada jarak aktual 9 meter.....	85
Gambar 4. 11 Hasil Deteksi kecepatan dengan kecepatan pengamat dan mobil di depan sama-sama 10km/jam	88
Gambar 4. 12 Hasil Deteksi kecepatan dengan kecepatan pengamat 10km/jam dan kecepatan mobil di depan 20 km/jam.....	89
Gambar 4. 13 Hasil Deteksi kecepatan dengan kecepatan pengamat 10km/jam dan kecepatan mobil di depan 60 km/jam.....	89

Gambar 4. 14 Hasil Deteksi kecepatan dengan kecepatan pengamat 20 km/jam dan kecepatan mobil di depan 10 km/jam.....	90
Gambar 4. 15 Hasil Deteksi kecepatan dengan kecepatan pengamat dan mobil di depan sama-sama 20 km/jam	90
Gambar 4. 16 Hasil Deteksi kecepatan dengan kecepatan pengamat 20 km/jam dan kecepatan mobil di depan 60 km/jam.....	91

DAFTAR TABEL

Tabel 2. 1 Taksonomi tingkat Otomasi kendaraan tanpa awak	8
Tabel 4. 1 Nilai RMSE dengan papan catur dengan ukuran satuan pola 3 cm.....	73
Tabel 4. 2 Nilai RMSE dengan papan catur dengan ukuran satuan pola 11 cm ...	73
Tabel 4. 3 Evaluasi sistem deteksi jarak pada mobil	83
Tabel 4. 4 Evaluasi sistem deteksi jarak pada motor	84
Tabel 4. 5 Evaluasi sistem deteksi kecepatan pada mobil dengan kecepatan pengamat 10 km/jam.....	92
Tabel 4. 6 Evaluasi sistem deteksi kecepatan pada mobil dengan kecepatan pengamat 20 km/jam.....	93
Tabel 4. 7 Contoh waktu komputasi pada 30 frame pertama dalam sebuah scenario pengujian.....	94

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Teknologi kendaraan *autonomous* (*Autonomous Vehicle* – AV) merupakan teknologi masa depan yang berpotensi menjadi salah satu teknologi paling disruptif di zaman kita. Tingkat manfaat prospektif AV sangat terkait dengan seberapa luas teknologi tersebut akan diterima dan diadopsi (Golbabaei et al., 2020). Dalam tulisannya, Saeed Asadi Bagloe menyatakan bahwa AV dapat menekan biaya transportasi secara signifikan. Dalam satu perkiraan, dampak social AV berdampak pada penghematan asuransi kecelakaan, pengurangan waktu perjalanan, efisiensi bahan bakar dan benefit parkir yang memungkinkan mencapai USD 2000 per tahun per kendaraan dan mungkin mencapai USD 4000 ketika biaya kecelakaan komprehensif diperhitungkan (Bagloee et al., 2016).

Menurut penelitian yang dipublikasikan oleh Deloitte, market AV akan mencapai angka 27 miliar USD pada tahun 2025, yang mana hampir 50% dari angka tersebut dihabiskan untuk pengembangan dan optimalisasi *software*, sekaligus menjadi angka terbesar jika dibandingkan biaya yang dihabiskan untuk perawatan dan *hardware*. (Deloitte 2019).

Sejauh ini, sejumlah produsen kendaraan mulai beralih untuk membiayai riset-riset yang dapat mengoptimalkan kinerja kendaraan *autonomous* (GreyB, 2021) . Kemampuan-kemampuan dasar kendaraan *autonomous* menjadi bidang riset yang gencar didalami oleh para pelaku teknologi, salah satu kemampuan dasar yang harus dimiliki oleh AV adalah kemampuan mendeteksi kecepatan kendaraan

di sekitarnya (Iqbal, 2020) yang selanjutnya menjadi salah satu informasi penting bagi AV dalam bermanuver di jalanan.

Dalam lingkup Universitas Hasanuddin sendiri, Nurul Fathanah Mustamin telah melakukan penelitian terkait deteksi jarak antara kendaraan menggunakan pendekatan metode *width-based method*. Dimana metode ini mengharuskan adanya nilai lebar spesifik setiap kendaraan, sehingga diperlukan adanya database mengenai lebar kendaraan yang akan dideteksi. Hal ini, tentu dapat menjadi masalah, mengingat banyaknya jenis kendaraan di Indonesia.

Berdasarkan uraian diatas, untuk dapat mengembangkan sistem *autonomous* yang dapat bekerja secara optimal, khususnya beradaptasi dengan kebiasaan pengguna jalan di kota Makassar, perlu didukung dengan sistem pendeteksi terkait kecepatan kendaraan secara *real-time*. Dalam hal ini, akan dirancang sebuah sistem yang akan mengimplementasikan *Measurement Method* menggunakan *Stereo-Camera* yang terpasang pada bagian depan mobil yang tidak memerlukan informasi spesifik lebar setiap kendaraan.

Penggunaan kamera tunggal untuk sistem sensor merupakan metode yang banyak digunakan pada rekonstruksi 3D berbasis citra yang ada. Namun ini membebankan batasan pada hasil yang dihasilkan. Pemandangan hanya dapat direkonstruksi ke faktor skala yang tidak diketahui jika ada kamera tunggal digunakan untuk akuisisi citra (Salman et al., 2017). Bukti nyata dalam kehidupan sehari-hari adalah penggunaan mata manusia, dimana seakan terjadi perubahan posisi sebuah objek ketika mengamati objek hanya dengan satu mata dibandingkan dengan kedua mata.

Sistem ini akan berfokus pada proses pengenalan kendaraan beserta fitur-fiturnya yang selanjutnya dikombinasikan dengan *Measurement Method* dan menjadi data acuan sistem dalam menentukan kecepatan kendaraan di depan maupun dibelakang pengamat berdasarkan perubahan *disparity* objek pada setiap framenya.

1.2 Rumusan Masalah

Adapun rumusan masalah dari penelitian ini adalah sebagai berikut:

- a. Bagaimana mengimplementasikan sistem deteksi kecepatan kendaraan sekitar dengan berbasis *stereo-camera* secara *Real-time*?
- b. Bagaimana unjuk kinerja *Measurement Method* dalam mendeteksi kecepatan kendaraan secara *Real-time*?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

- a. Untuk mengimplementasikan teknologi *stereo-camera* pada bidang *Computer vision* secara *Real-time*.
- b. Untuk mengetahui performa sistem deteksi kecepatan kendaraan menggunakan *Measurement Method* berbasis *stereo-camera*.

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut:

- a. Memberikan dorongan penggunaan teknologi *Computer vision* dalam menyelesaikan masalah transportasi, khususnya di kota Makassar.
- b. Memberikan informasi yang dapat digunakan sebagai dasar untuk melakukan riset berkelanjutan mengenai industri *autonomous vehicle*.
- c. Membangun sinergitas dunia kampus dan pihak pemerintahan demi terwujudnya program Makassar sebagai *smart city*.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

- a. Data training merupakan sinkronisasi data primer kondisi transportasi di beberapa ruas jalan kota Makassar dan data sekunder yang berkaitan dengan penelitian *autonomous vehicle* yang telah ada sebelumnya.
- b. Deteksi kecepatan pada penelitian ini akan dikerucutkan menjadi 5 kelas, yakni:
 - Mobil
 - Motor
 - Bus
 - Truck
 - Sepeda
- c. Pengambilan data primer dilakukan di siang hari di beberapa titik di kota Makassar.
- d. Posisi *stereo-camera* berada pada bagian depan kendaraan pengamat.
- e. Sistem akan diuji pada kecepatan maksimal 40 km/jam baik untuk kendaraan pengamat maupun kendaraan yang diamati.
- f. Metode kalibrasi kamera akan menggunakan *2D-Calibration*.
- g. Evaluasi keakuratan hasil kalibrasi *stereo* kamera menggunakan metode RMSE

- h. Evaluasi keakuratan sistem dalam deteksi jarak maupun kecepatan menggunakan MAPE.

1.6 Metode Penulisan

Berikut adalah metode yang digunakan pada penelitian:

- a. Metode Pengambilan Data

Data kendaraan berupa data primer dan data sekunder. Data primer merupakan video yang diambil di beberapa ruas jalan kota Makassar dan kemudian di ekstrak menjadi beberapa frame yang selanjutnya dikombinasikan dengan data sekunder.

Data sekunder bersumber dari beberapa video dashcam kendaraan di Indonesia yang ada di youtube serta dataset Poribohon-BD (Tabassum, 2020).

- b. Studi Literatur

Studi literatur dilakukan dengan mengumpulkan informasi yang dapat menjadi landasan teori melalui berbagai sumber seperti buku, jurnal, paper, internet, dan sumber lainnya.

- c. Diskusi dan Konsultasi

Diskusi dan konsultasi dilakukan dengan melakukan tanya jawab dengan dosen pembimbing, forum-forum penggiat *Computer vision* khususnya terkait *autonomous driving*, serta pihak lainnya yang dapat mendukung penyelesaian penelitian ini.

1.7 Sistematika Penulisan

Berikut adalah gambaran singkat mengenai isi tulisan secara keseluruhan:

BAB 1 PENDAHULUAN

Pada bab ini diuraikan mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah penelitian, metode penelitian, serta sistematika penulisan.

BAB 2 TINJAUAN PUSTAKA

Pada bab ini membahas landasan teori yang digunakan untuk menganalisis masalah yang akan diteliti serta hal-hal lain yang berhubungan dengan variabel-variabel data yang akan digunakan, *computer vision*, *stereo-camera*, *Deep Learning* dan *Measurement Method*.

BAB 3 METODOLOGI PENELITIAN

Pada bab ini berisi mengenai tahap penelitian, instrumen penelitian, pengumpulan data, penerapan algoritma, teknik pengolahan data, serta hasil asosiasi.

BAB 4 HASIL DAN PEMBAHASAN

Bab ini berisi tentang sistem yang telah berhasil dibangun serta pembahasannya.

BAB 5 PENUTUP

Bab ini berisi tentang kesimpulan yang didapatkan berdasarkan hasil penelitian yang telah dilakukan serta saran-saran untuk pengembangan lebih lanjut.

BAB 2

TINJAUAN PUSTAKA

2.1 Autonomous Vehicle

Autonomous vehicle (AV) atau kendaraan tanpa awak pertama kali direncanakan pada tahun 1918 (Pendleton et al., 2017), dan kemudian konsep mengenai kendaraan tanpa awak ini di pamerkan oleh General Motors pada tahun 1939 (Shladover, 2018). *Research* dan *development* mengenai teknologi kendaraan *autonomous* pertama kali diinisiasi atas kolaborasi General Motors dan Radio Corporation of America Sarnoff Laboratory pada kisaran tahun 1950 an. Sejak tahun 1964 hingga 2003 program penelitian dan pengembangan mengenai AV mulai marak beroperasi di AS, Eropa, dan Jepang yang dibawah inisiatif individu maupun kerjasama beberapa lembaga pemerintahan dan akademisi yang bertujuan untuk mengembangkan beberapa peleton bus dan truk otomatis, sistem cerdas untuk kendaraan, dan pemrosesan video terhadap deteksi pengenalan adegan saat mengemudi. (Shladover, 2018).

Penelitian mengenai AV mulai terakselerasi melalui Defense Advanced Research Projects Agency's (DARPA) program tantangan besar di AS pada tahun 2004. Tantangan tersebut mampu menghasilkan AV yang mampu melintasi medan gurun pada tahun 2005, dan pada tahun 2007 para peneliti juga berhasil menempatkan AV di jalan perkotaan melalui DARPA's Urban Challenge Program (Pendleton et al., 2017; Shladover, 2018). Sejak saat itu penelitian dan pengembangan berlanjut secara pesat baik di lingkup akademik maupun di industri.

Sejalan dengan konsep Otomasi, taksonomi mengenai kendaraan tanpa awak dibagi menjadi 4 tingkat oleh National Highway Traffic Safety Administration (NHTSA) pada tahun 2013 (Wadud et al., 2016), dan pada tahun 2014 Society of Automotive Engineers International (SAE) memperkenalkan 5 taksonomi kendaraan *autonomous* – yang kemudian diperbaharui pada tahun 2016 (Coppola and Morisio, 2016; Milakis et al., 2017; SAE International, 2016b, 2016a; Snyder, 2016). Pada tahun 2016, NHTSA mengadopsi tingkat taksonomi dan otomisasi SAE (“NHTSA., 2016,” n.d.). Taksonomi dan tingkat otomisasi SAE telah menjadi standar industri, dan juga sering dirujuk dalam literatur akademik (Rubin, 2016; Scheltes and de Almeida Correia, 2017; Shladover, 2018; Walker and Marchau, 2017). Pembagian ini kemudian dapat dilihat pada tabel 2.1 berikut:

Tabel 2. 1 Taksonomi tingkat Otomasi kendaraan tanpa awak

Level Otomasi	Kontribusi Sistem Otomasi	Kontribusi Pengendara	Contoh Fitur
Level 0 (Tidak ada sistem Otomasi)	Memberikan peringatan dan bantuan sementara	Menentukan arah, posisi, persepsi, hingga manajemen kemudi	Blind spot warning
Level 1 (Sudah ada sistem yang berperan sebagai asisten mengemudi)	Dukungan kemudi atau akselerasi rem kepada pengemudi	Manajemen kemudi atau akselerasi rem tetap masih dalam kendali	Lane centering atau kontrol kemudi otomatis
Level 2 (Sebagian kecil skema mengemudi diserahkan kepada sistem)	Dukungan kemudi dan akselerasi rem kepada pengemudi	Manajemen kemudi dan akselerasi rem tetap masih dalam kendali	Lane centering dan kontrol kemudi otomatis

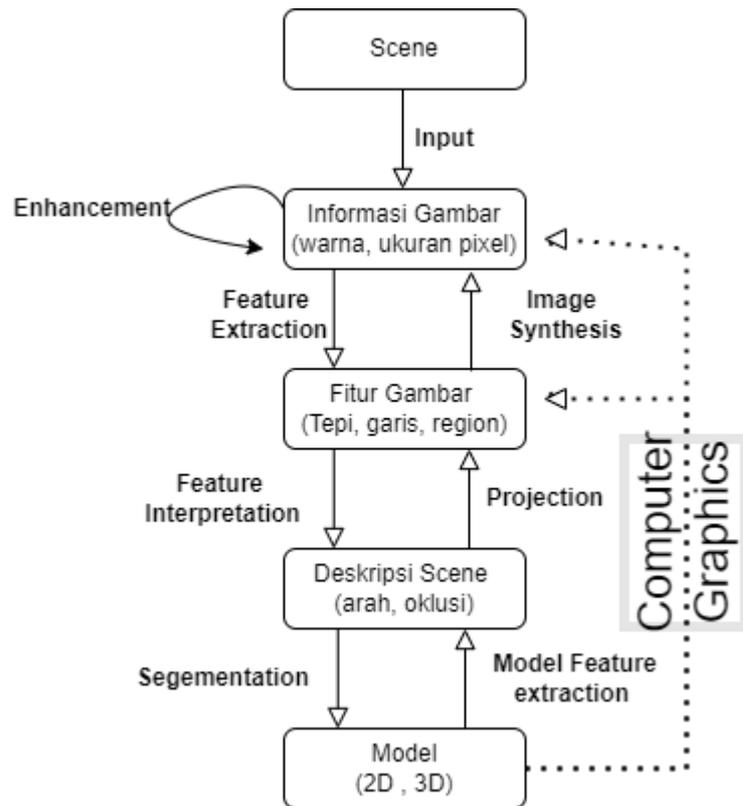
Level 3 (Fungsi asistensi sangat berperan di beberapa kondisi-kritis)	Sudah dapat menjalankan manuver sederhana	Ketika sistem meminta, pengendara harus mengambil alih	Berkendara otomatis di jalan yang lowong.
Level 4 (Full-Otomasi, kecuali beberapa skenario mengemudi)	Sudah dapat menjalankan manuver yang cukup berat, namun belum seahli pengendara manusia	Sistem sangat jarang meminta pengalihan kendali	Berkendara di daerah macet, menghindari gerak kendaraan yang dapat membahayakan.
Level 5 (Full-Otomasi, kemampuan sistem setara dengan pengendara manusia di setiap skenario pengemudian)	Sistem sudah setara kemampuan pengendara manusia	Pengendara dapat mengambil kendali jika diinginkan saja	Dapat berkendara di semua jenis medan.

Selanjutnya, pertanyaan tentang seberapa cepat kendaraan Otonom akan membangun dominasi pasar menjadi pernyataan yang sulit, Seberapa cepat pasar akan mengadopsi kendaraan ini tidak diketahui, dan sebagian besar penelitian dilakukan dengan melihat tingkat pertumbuhan teknologi sebelumnya. Pada tahun 2012, panel anggota IEEE memperkirakan bahwa 75% dari armada akan otonom pada tahun 2040 (Richard Reald, n.d.). The Victoria Transport Policy Institute (Litman, 2015) memprediksi penyerapan yang lebih lambat - dengan 75% pasar dicapai pada tahun 2060. Ini didasarkan pada perbandingan dengan teknologi lain pada kendaraan, seperti transmisi otomatis, navigasi on-board dan kendaraan hybrid, yang semuanya membutuhkan waktu beberapa dekade untuk mencapai penguasaan pasar yang signifikan.

2.2 Computer Vision

Computer vision (CV) atau visi komputer dikembangkan dengan memiliki dua tujuan ganda. Dari sudut ilmu sains biologi, visi komputer dimaksudkan untuk memahami model komputasi dari sistem visual manusia. Dari sudut ilmu rekayasa titik pandang, CV bertujuan membangun sistem otonom yang dapat melakukan beberapa tugas yang dapat dilakukan oleh sistem visual manusia (bahkan diharapkan dapat melampauinya dalam beberapa kasus). Banyaknya tantangan di bidang ini terkait dengan ekstraksi informasi 3D dan temporal dari 2D dengan waktu pengambilan yang variatif sebagaimana satu atau lebih kamera memperoleh informasi yang terangkum dalam adegan dinamis (Huang, 1996). Tentu saja, kedua tujuan tersebut berkaitan erat. Keberagaman mata manusia dan sifat sistem visual manusia sering memberikan inspirasi kepada para insinyur yang merancang sistem visi komputer. Sebaliknya, algoritma visi komputer dapat menawarkan wawasan tentang bagaimana sistem visual manusia bekerja dalam sekian baris kode.

Pemrosesan CV pada lingkup objek 3D terdiri dari banyak komponen yang melakukan fungsi tertentu dalam koordinasi dengan komponen lain. Gambar 2.1 merupakan blok rangkaian utama dan tipe informasi atau data yang terlibat dari proses tersebut. (Shirai, 2012)



Gambar 2. 1 Rangkaian proses utama pada *Computer Vision*

Sebagaimana yang tersaji pada gambar 2.1, Ketika sebuah adegan/scene masuk ke dalam program, informasi dasar dari gambar tersebut akan diekstraksi baik komponen warna pada gambar, ukuran tiap objek dalam pixel serta informasi pixel keseluruhan adegan itu sendiri. Informasi yang diekstraksi bisa berupa raw gambar yang original, maupun scene yang telah mengalami beberapa enhancement baik itu kontras, saturasi, dll. Setelah itu scene tersebut akan diinterpelasi untuk mendapatkan data tepi, garis, region, dan data apapun yang berkaitan dengan pembentukan dasar tiap objek pada scene tersebut. Langkah terakhir merupakan segmentasi untuk mendapatkan deskripsi mengenai scene tersebut yang selanjutnya akan terbentuk matriks model 2D maupun 3D dari satu scene tersebut.

2.3 Deep Learning

Deep Learning merupakan sub bidang *Machine learning* yang berkaitan dengan algoritma yang terinspirasi oleh struktur dan fungsi otak manusia. Kemampuan inti *Deep Learning* adalah kemampuan untuk membentuk tingkat abstraksi yang lebih tinggi dari representasi dalam data dan pola mentah. Data latihan yang digunakan untuk melatih model adalah data lapangan yang Reall, dan akan diolah menggunakan metode one hot encoding, sehingga model yang didapatkan berisi data kategorikal yang dapat dicocokkan sewaktu-waktu dengan data masukan baru. (Voican, 2021).

Deep Learning memungkinkan model komputasi yang terdiri dari beberapa lapisan pemrosesan untuk mempelajari representasi dari data dengan beberapa tingkat abstraksi. Metode-metode ini telah secara dramatis meningkatkan state-of-the-art dalam pengenalan suara, pengenalan objek visual, deteksi objek dan banyak domain lainnya seperti penemuan obat dan genomik. *Deep Learning* menemukan struktur rumit dalam kumpulan data besar dengan menggunakan algoritma backpropagation untuk menunjukkan bagaimana mesin harus mengubah parameter internalnya yang digunakan untuk menghitung representasi di setiap lapisan dari representasi di lapisan sebelumnya. Jaring konvolusi yang dalam telah menghasilkan terobosan dalam pemrosesan gambar, video, ucapan, dan audio, sedangkan jaring berulang telah menyoroti data berurutan seperti teks dan ucapan (LeCun et al., 2015).

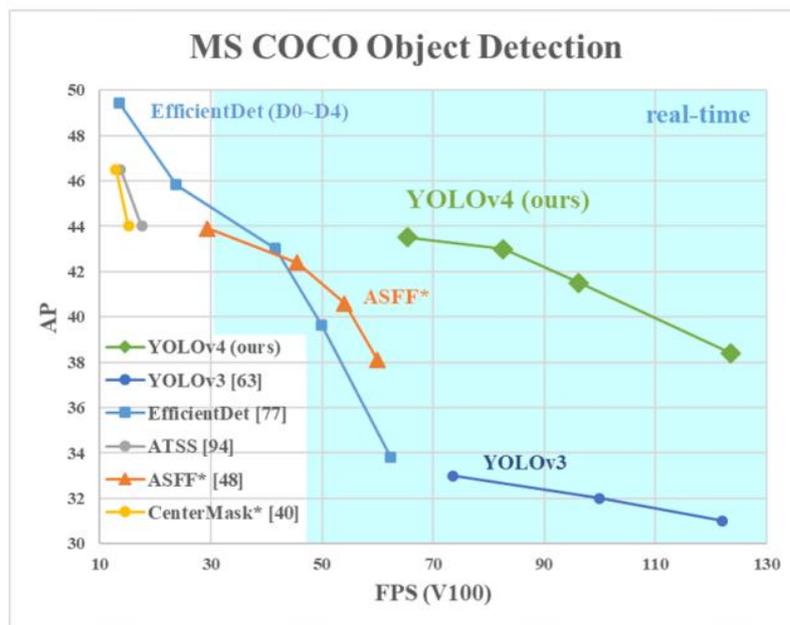
Dalam penerapannya terhadap kasus *computer vision*, *Deep Learning* seringkali dipercaya dapat menyelesaikan kasus *Computer vision* dengan akurasi

yang baik dibandingkan dengan metode *Machine learning* konvensional (O'Mahony et al., 2020). Dalam penerapannya, penggunaan *Deep Learning* tidak saja baik secara akurasi namun juga dapat sekaligus menyelesaikan tiga ranah pada CV, baik deteksi objek pada gambar, deteksi objek pada video, hingga deteksi objek secara *real-time* serta lebih jauh *Deep Learning* juga diharapkan dapat menyelesaikan tantangan yang lebih advance pada *Computer vision* seperti menghitung jumlah objek, *mentracking* objek, dan lain-lain.

2.4 Yolo V5

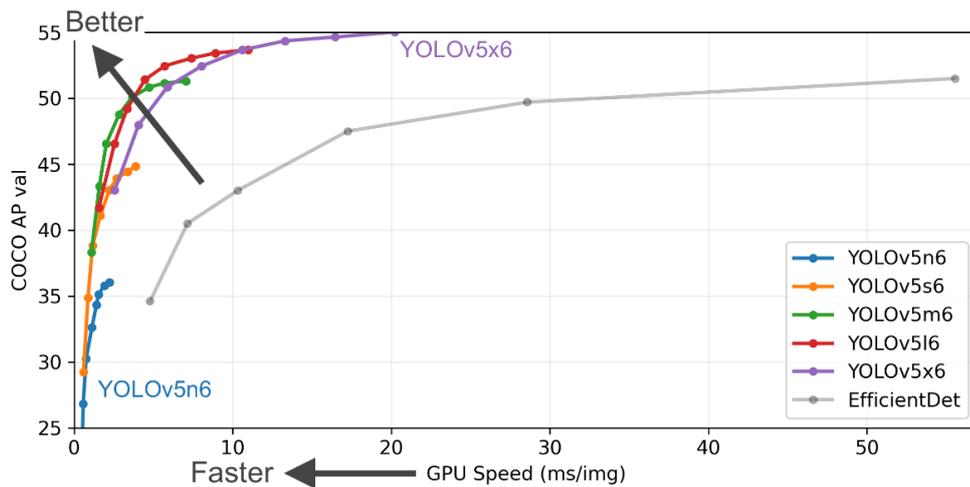
Yolo "*You only look once*" merupakan algoritma populer dan banyak digunakan dalam bidang *computer vision*. Menggunakan jaringan saraf untuk menyediakan deteksi objek secara *Real-time*. Algoritma ini populer karena kecepatan dan akurasinya. Ini telah digunakan dalam berbagai aplikasi untuk mendeteksi sinyal lalu lintas, orang, meteran parkir, dan hewan. Pertama kali diperkenalkan oleh Joseph Redmon pada tulisannya "*You Only Look Once: Unified, Real-time Object Detection*" yang kemudian menjadi salah satu inovasi pada ranah deteksi objek secara *Real-time* pada tahun 2016 (Redmon et al., 2016). Kemudian pada tahun 2017 Joseph juga memperkenalkan versi terbaru yang dikenal dengan "*YOLO9000: Better, Faster, Stronger*" (Redmon and Farhadi, 2016) dan berlanjut pada April 2018 YOLO juga mendapat versi terbarunya "*Yolov3: An Incremental Improvement*" (Redmon and Farhadi, 2018, p. 3). Kemudian pada tahun 2020, Alexey Bochkovskiy memperkenalkan Yolov4 dalam tulisan "*YOLOv4: Optimal Speed and Accuracy of Object Detection*" (Bochkovskiy et al., 2020), YOLOv4

diperkenalkan dengan beberapa hal baru yang mencengangkan, Ini mengungguli YOLOv3 dengan margin tinggi dan juga memiliki jumlah presisi rata-rata yang signifikan jika dibandingkan dengan algoritma lain seperti EfficientDet, hal ini dapat dilihat pada gambar 2.2 berikut:



Gambar 2. 2 Perbandingan Yolov4 dengan beberapa algoritma.

Keseluruhan versi yolo yang telah disebutkan sebelumnya dibangun diatas framework Darknet, dan pada tahun 2020 tepatnya pada tanggal 9 juni, secara unofficial Glenn Jocher merilis Yolov5 pada repositorynya. Implementasi Yolov5 sempat menjadi kontroversi disebabkan Glenn Jocher, sampai tulisan ini dibuat belum merilis makalah resmi terkait Yolov5, selain itu Yolov5 diimplementasi secara natif di framework PyTorch, berbeda dengan versi Yolo sebelumnya yang dibangun diatas framework Darknet (Kanjee, 2022). Gambar 2.3 menunjukkan perfomansi beberapa versi Yolov5 terhadap algoritma EfficientDet pada dataset Coco.



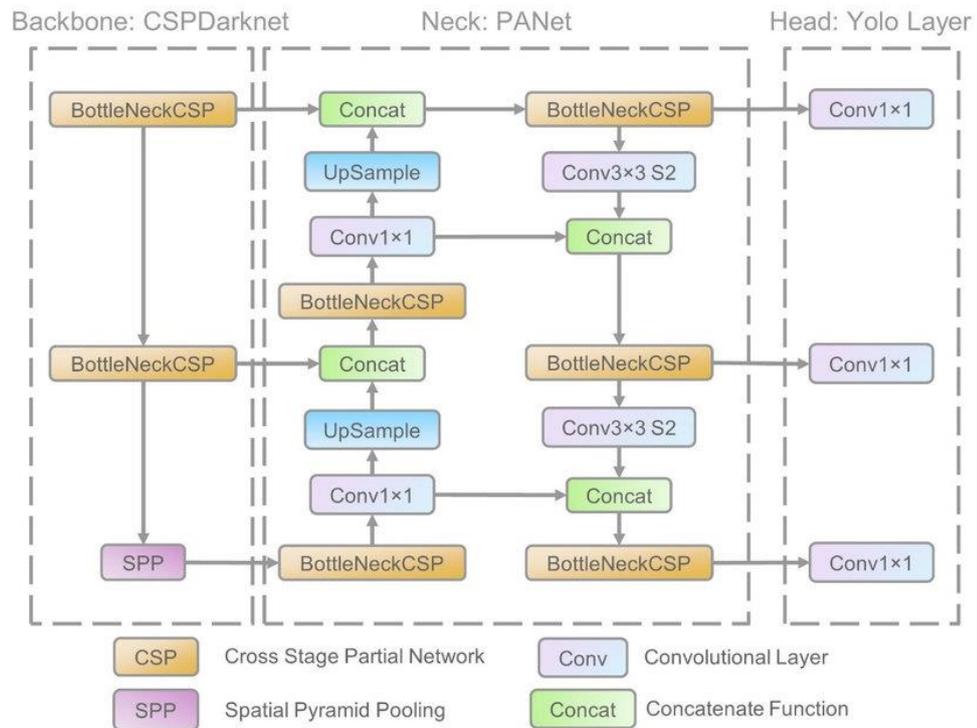
Gambar 2. 3 Perbandingan Yolov5 dengan EfficientDet

Berdasarkan gambar 2.3 peneliti mengajukan Yolov5 sebagai metode *Deep Learning* dalam mendeteksi kendaraan sekitar, hal ini disebabkan kebutuhan komputasi yang cepat dengan konsumsi waktu yang lebih efisien. Sebagaimana model *Deep Learning* pada umumnya, arsitektur Yolo V5 juga terbentuk atas 3 komponen utama; *Model Backbone*, *Model Neck*, Dan *Model Head* (Gutta, 2021).

Model *backbone* yang digunakan Yolo V5 adalah CSPDarkNet (*Cross Stage Partian Dark Network*). Fungsi utama dari model *backbone* adalah mengekstraksi fitur dari citra yang diinput. Pemilihan CSPDarkNet sebagai model *backbone* dikarenakan algoritma ini didesain untuk meningkatkan kapabilitas pembelajaran algoritma *Deep Learning* yang berbasis CNN(Wang et al., 2020). Selain itu, metode ini juga membantu memecahkan masalah informasi gradien berulang dalam tulang punggung skala besar, dan mengintegrasikan perubahan gradien ke dalam peta fitur, sehingga mengurangi parameter dan FLOPS (operasi titik-mengambang per detik) model, yang tidak hanya memastikan kecepatan dan akurasi inferensi, tetapi juga mengurangi ukuran model(Xu et al., 2021).

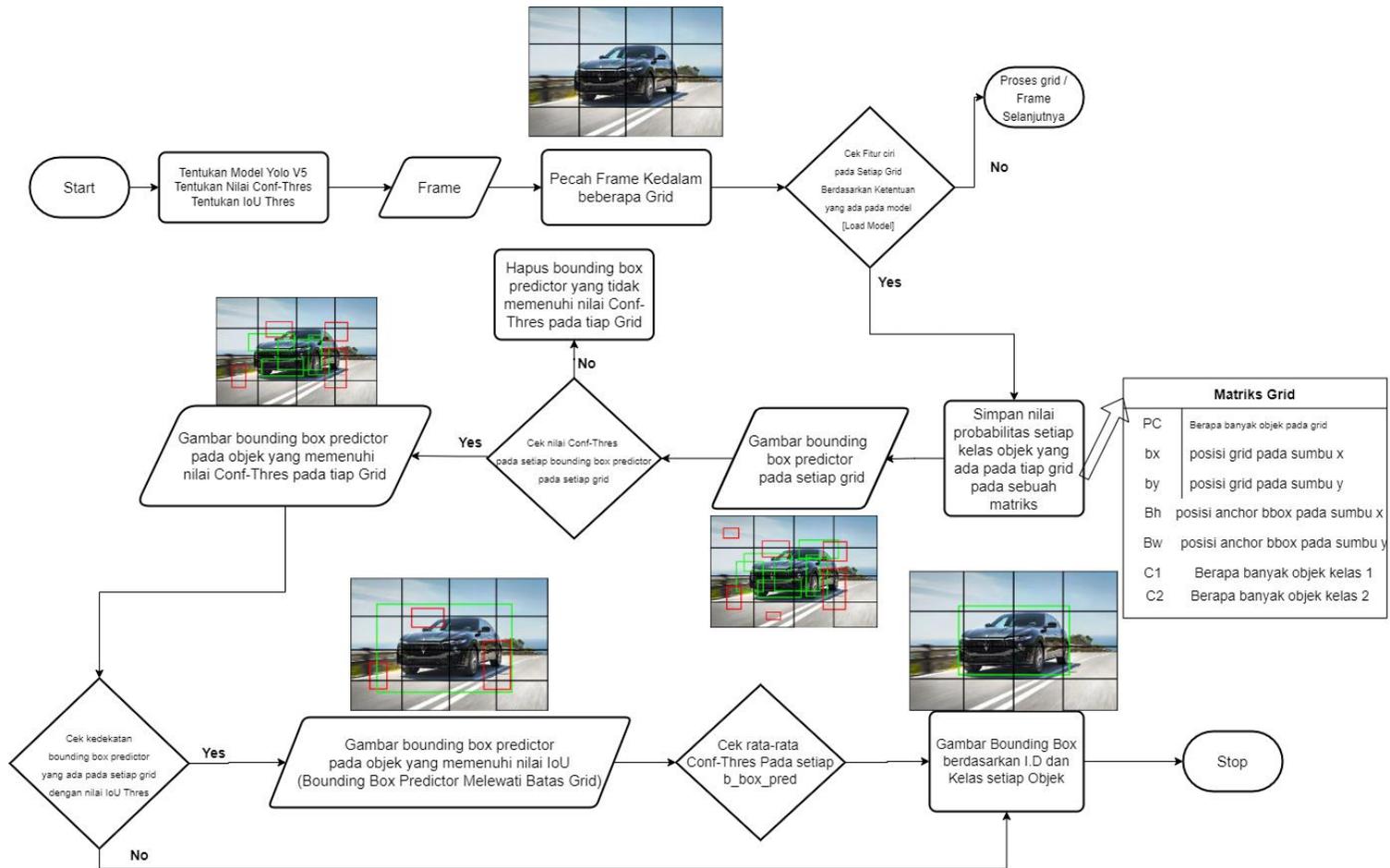
Untuk bagian kedua, Yolo V5 menggunakan PANet (*Path Aggregation Network*) sebagai *Model Neck*. *Model Neck* berfungsi untuk menghasilkan piramida fitur. Piramida fitur membantu model untuk menggeneralisasi dengan baik pada penskalaan objek. Ini membantu untuk mengidentifikasi objek yang sama dengan ukuran dan skala yang berbeda. Piramida fitur sangat berguna dan membantu model untuk bekerja dengan baik pada data yang tidak terlihat (Team, n.d.). Yolov5 PANet sebagai *Neck*-nya untuk meningkatkan arus informasi. PANet mengadopsi struktur jaringan piramida fitur (FPN) baru dengan jalur bottom-up yang ditingkatkan, yang meningkatkan penyebaran fitur tingkat rendah. Pada saat yang sama, penggabungan fitur adaptif, yang menghubungkan kisi fitur dan semua level fitur, digunakan untuk membuat informasi yang berguna di setiap level fitur menyebarkan langsung ke subnetwork berikut. PANet meningkatkan pemanfaatan sinyal lokalisasi yang akurat di lapisan bawah, yang jelas dapat meningkatkan akurasi lokasi objek (Xu et al., 2021).

Model *Head* utamanya digunakan untuk melakukan bagian deteksi akhir. Ini menerapkan kotak jangkar pada fitur dan menghasilkan vektor keluaran akhir dengan probabilitas kelas, skor objektivitas, dan kotak pembatas (Team, n.d.). Dalam *Head* model YOLO v5 sama dengan versi YOLO V3 dan V4 sebelumnya. Model *Head* ini sering disebut dengan Yolo Layer. Menghasilkan 3 ukuran yang berbeda (18×18 , 36×36 , 72×72) dari peta fitur untuk mencapai prediksi multi-skala, memungkinkan model untuk menangani objek kecil, sedang, dan besar (Xu et al., 2021). Gambar 2.4 menggambarkan secara umum arsitektur Yolo V5 (Xu et al., 2021) :



Gambar 2. 4 Arsitektur Yolo V5

Setelah sebuah model terbentuk, maka model siap untuk digunakan dalam mendeteksi objek berdasarkan kelas yang telah dilatih. Flowchart proses kerja YoloV5 dalam mendeteksi objek pada penelitian ini dapat dilihat pada gambar 2.5 berikut:



Gambar 2. 5 Flowchart YoloV5 dalam mendeteksi objek

2.5 DeepSort

DeepSort (*Deep association metric of Simple Online and Real time Tracking*) adalah pendekatan berbasis *Deep Learning* untuk melacak kustom objek dalam video (Wojke et al., 2017). Dalam penelitian awal mengenai DeepSort, DeepSort digunakan untuk melacak individu yang ada dalam rekaman video pengawasan. Metode ini bekerja dengan memanfaatkan pola yang dipelajari melalui objek yang terdeteksi di gambar yang kemudian digabungkan dengan informasi temporal untuk memprediksi lintasan terkait objek yang menarik. Ini akan melacak setiap objek yang dipertimbangkan dengan pemetaan pengenalan unik untuk analisis statistik lebih lanjut. DeepSort juga berguna untuk menangani tantangan terkait seperti oklusi, berbagai sudut pandang, kamera non-stasioner, dan anotasi data pelatihan.

Untuk pelacakan yang efektif, Kalman-Filter (R. E. KALMAN, n.d.) dan algoritma Hungaria digunakan. Kalman-Filter secara rekursif digunakan untuk asosiasi yang lebih baik, dan dapat memprediksi posisi masa depan berdasarkan posisi saat ini. Sedangkan Algoritma Hungaria digunakan untuk asosiasi dan atribusi id yang mengidentifikasi apakah suatu objek di frame saat ini sama dengan frame sebelumnya. Dalam mendeteksi, DeepSort menggambarkan setiap target dengan ruang delapan dimensi dalam model kecepatan konstan linier, sebagai berikut: dimana (u, v) mewakili lokasi piksel horizontal dan vertikal dari pusat target, λ dan h masing-masing mewakili skala (Real) dan rasio aspek kotak pembatas target. Perlu diingat bahwa rasio aspek dianggap konstan. Variabel lainnya adalah kecepatan masing-masing variabel. Ketika deteksi dikaitkan dengan

target, *bounding box* yang terdeteksi digunakan untuk memperbarui status target dimana komponen kecepatan diselesaikan secara optimal melalui kerangka Kalman-Filter. Nantinya, standar Kalman-Filter digunakan dengan kecepatan gerak konstan dan linier model observasi, dimana koordinat pembatas (u, v, λ, h) diambil sebagai pengamatan langsung dari keadaan objek. Jika tidak ada deteksi yang dikaitkan dengan target, keadaannya diprediksi tanpa koreksi menggunakan model kecepatan linier (persamaan 2.1).

$$x = [u, v, \lambda, h, x', y', \lambda', h']^T \quad (2.1)$$

u, v = posisi *centroid* setiap *bounding box* objek
 λ = nilai aspek rasio *bounding box* objek
 h = tinggi *bounding box*
 x', y', λ', h' = *velocity factors* setiap variable-variabel sebelumnya dalam koordinat gambar (berkaitan dengan model *simple linear velocity* pada *Kalman filter*)

Skenario tracking didefinisikan dalam ruang delapan dimensi ($u, v, \lambda, h, x', y', \lambda', h'$). Dimana nilai u, v menyatakan posisi *centroid* dari setiap *bounding box*, nilai λ merepresentasikan nilai rasio tinggi-lebar *bounding box*, dan h merepresentasikan tinggi *bounding box*. Untuk setiap track T , mulai dari asosiasi pengukuran terakhir yang berhasil λ_T , jumlah total frame dihitung. Dengan prediksi positif dari Kalman-Filter, penghitung bertambah dan kemudian ketika trek dikaitkan dengan pengukuran, Nilainya diatur ulang ke 0. Selanjutnya, jika trek yang diidentifikasi melebihi waktu maksimum yang telah ditentukan, maka objek tersebut dianggap telah meninggalkan adegan dan trek yang sesuai akan dihapus dari set trek. Dan jika tidak ada trek yang tersedia untuk beberapa objek yang

terdeteksi maka hipotesis trek baru dimulai untuk setiap trek tak dikenal dari objek baru yang terdeteksi yang tidak dapat dipetakan ke trek yang ada. Untuk tiga frame pertama, trek baru diklasifikasikan sebagai tidak terbatas sampai pemetaan pengukuran yang berhasil dihitung. Jika trek tidak berhasil dipetakan dengan pengukuran maka itu akan dihapus dari kumpulan trek. Algoritma Hungaria (Kuhn, 1955) kemudian digunakan untuk memecahkan masalah pemetaan antara pengukuran yang baru tiba dan keadaan Kalman yang diprediksi dengan mempertimbangkan informasi gerakan dan tampilan dengan bantuan *Mahalanobis-Distance* yang dihitung di antara objek yang dapat di formulakan menjadi (persamaan 2.2):

$$d^{(1)}(i, j) = (d_j - y_i)^T S_i^{-1} (d_j - y_i) \quad (2.2)$$

$d^{(1)}(i, j)$	= nilai sampel <i>Mahalanobis-Distance</i>
d_j	= nilai bounding box terdistribusi
y_i	= nilai rata-rata sample pada tiga frame pertama
S_i	= kovarians sampel (dengan bias terkoreksi)
	matriks dari n pengamatan dalam sampel yang diamati

dimana proyeksi distribusi lintasan ke-i ke dalam ruang pengukuran diwakili oleh (y_i, S_i) dan deteksi kotak pembatas ke-j oleh d_j . Jarak Mahalanobis mempertimbangkan ketidakpastian ini dengan memperkirakan jumlah standar deviasi, deteksi jauh dari lokasi trek rata-rata. Selanjutnya, dengan menggunakan matrik ini, dimungkinkan untuk mengecualikan asosiasi yang tidak mungkin dengan membatasi jarak Mahalanobis. Keputusan ini dilambangkan dengan indikator yang bernilai 1 jika hubungan antara jalur ke-i dan deteksi ke-j dapat diterima dengan persamaan 2.3:

$$b_{i,j}^{(1)} = 1[d^{(1)}(i, j) < t^{(1)}] \quad (2.3)$$

$b^{(1)}(i,j)$ = nilai normalisasi sampel *Mahalanobis-Distance*
 $d^{(1)}(i,j)$ = nilai sampel *Mahalanobis-Distance*
 $t^{(1)}$ = Mahalanobis threshold (9.4877)

meskipun jarak Mahalanobis bekerja dengan efisien tetapi gagal di lingkungan dimana gerakan kamera terjadi, oleh karena itu matrik lain diperkenalkan untuk masalah penugasan. Ini matrik kedua mengukur jarak cosinus terkecil antara trek ke-i dan deteksi ke-j di ruang tampilan sebagai berikut:

$$d^{(2)}(i, j) = \min\{1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in \mathbb{R}^2\} \quad (2.4)$$

r_j = konstanta descriptor penampil pada deteksi ke-j
 r_k = konstanta descriptor penampil pada tracking ke-j
 \mathbb{R}^2 = Keseluruhan nilai matrik jarak kosinus terkecil

dan ambang batas yang sesuai diukur untuk indikator ini pada kumpulan data pelatihan terpisah. Untuk membangun masalah asosiasi, kedua matrik digabungkan menggunakan jumlah berbobot:

$$c_{i,j} = \lambda d^{(1)}(i, j) + (1 - \lambda) d^{(2)}(i, j) \quad (2.5)$$

$c_{i,j}$ = jumlah bobot gabungan matriks jarak kosinus terkecil dan matriks jarak Mahalanobis

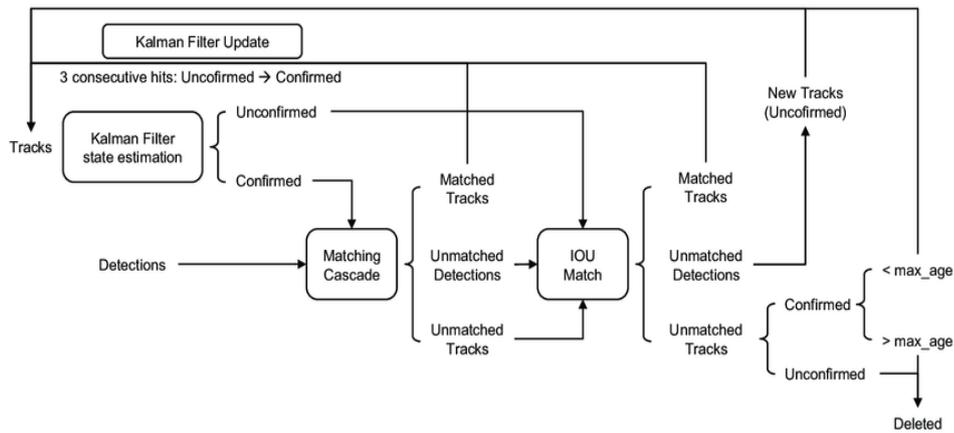
dimana asosiasi dapat diterima jika berada di dalam gerbang wilayah kedua matrik:

$$b_{i,j} = \prod_{m=1} 2b_{i,j}^{(m)}. \quad (2.6)$$

$b_{i,j}$ = matriks gerbang wilayah jarak kosinus terkecil dan matriks jarak Mahalanobis
 \prod = produk cartesian

Pengaruh setiap matrik pada biaya asosiasi gabungan dapat dikontrol melalui hyperparameter λ diberi kode warna berdasarkan asosiasinya dengan index setiap objek. Dalam menetapkan deteksi ke target yang ada, masing-masing target geometri kotak pembatas diperkirakan dengan memprediksi lokasi dalam bingkai saat ini. Matriks biaya penugasan adalah kemudian dihitung sebagai jarak persimpangan-over-union atau *Intersection-of-Union* (IOU) antara setiap deteksi dan semua kotak pembatas yang diprediksi dari target yang ada. Tugas diselesaikan secara optimal menggunakan algoritma Hungaria. Selain itu, minimal IOU dikenakan untuk menolak tugas dimana deteksi untuk target tumpang tindih kurang dari IOUmin.

Jarak IOU dari kotak pembatas secara implisit menangani oklusi jangka pendek yang disebabkan oleh target yang lewat. Khususnya, ketika target ditutupi oleh penghalang objek, hanya occluder yang terdeteksi, karena jarak IOU tepat mendukung deteksi dengan skala yang sama. Hal ini memungkinkan kedua target occluder dikoreksi dengan deteksi sementara target yang dicakup tidak terpengaruh karena tidak ada penetapan terbuat. Gambar 2.6 berikut menggambar secara sederhana bagaimana deepSort bekerja (Avola et al., 2021).



Gambar 2. 6 Alur kerja algoritma DeepSort (Avola et al., 2021)

2. 6 Stereo Camera

Kamera *stereo* adalah jenis kamera dengan dua atau lebih lensa dengan sensor gambar atau bingkai film terpisah untuk setiap lensa (Liu and Chen, 2009). Hal ini memungkinkan kamera untuk mensimulasikan penglihatan binokular manusia, dan oleh karena itu memberikannya kemampuan untuk menangkap gambar tiga dimensi, sebuah proses yang dikenal sebagai fotografi *stereo*. Kamera *stereo* dapat digunakan untuk membuat tampilan *stereo* dan gambar 3D untuk film, atau untuk pencitraan jarak jauh. Jarak antara lensa dalam kamera *stereo* tipikal (jarak intra-aksial) adalah sekitar jarak antara mata seseorang (dikenal sebagai jarak intraokular) dan sekitar 6,35 cm, meskipun garis dasar yang lebih panjang (jarak antar-kamera yang lebih besar) menghasilkan 3 dimensi yang lebih ekstrim. Kamera *stereo* dapat dibuat dengan memasang dua kamera memiliki konfigurasi yang sama pada basis yang sama. Batasan paling penting dalam mengambil sepasang stereoskopik gambar adalah sebagai berikut:

- a. Kamera harus disejajarkan secara horizontal,
- b. Gambar harus diambil pada saat yang sama

Gambar 3D yang mengikuti teori di balik kamera *stereo* juga dapat dibuat lebih murah dengan mengambil dua gambar dengan kamera yang sama, tetapi menggerakkan kamera beberapa inci ke kiri atau ke kanan. Jika gambar diedit sehingga setiap mata melihat gambar yang berbeda, maka gambar tersebut akan tampak 3D. Metode ini memiliki masalah dengan objek yang bergerak dalam pandangan yang berbeda, meskipun bekerja dengan baik dengan benda mati.

Dalam visi komputer, kadang-kadang mungkin untuk memperoleh informasi dari kamera yang tidak dikalibrasi, tetapi kamera kalibrasi mutlak diperlukan ketika berhadapan dengan matrik informasi untuk mendapatkan informasi 3D. Dengan kata lain, parameter kamera internal dan eksternal yang diperoleh dari proses kalibrasi diperlukan saat membangun 3D struktur adegan dari koordinat piksel gambar.

Formulasi matematis yang memperkirakan perilaku kamera diperlukan. Parameter ekstrinsik secara unik untuk mengidentifikasi transformasi antara referensi dunia nyata dan bingkai referensi kamera yang tidak diketahui seperti nilai rotasi dan translasi. Sedangkan, parameter intrinsik adalah karakteristik internal kamera seperti *focal length*, *principal point* dan koefisien distorsi. Persamaan 2.7 diturunkan oleh (Hartley and Zisserman, 2004) mewakili model kamera lubang jarum yang menunjukkan pemetaan proyektif dari koordinat dunia (x, y, z) ke koordinat piksel (u, v). Tujuan utama dari ini matriks transformasi adalah untuk menemukan hubungan antara Titik 3D pemandangan dengan titik proyeksi 2D di bidang gambar.

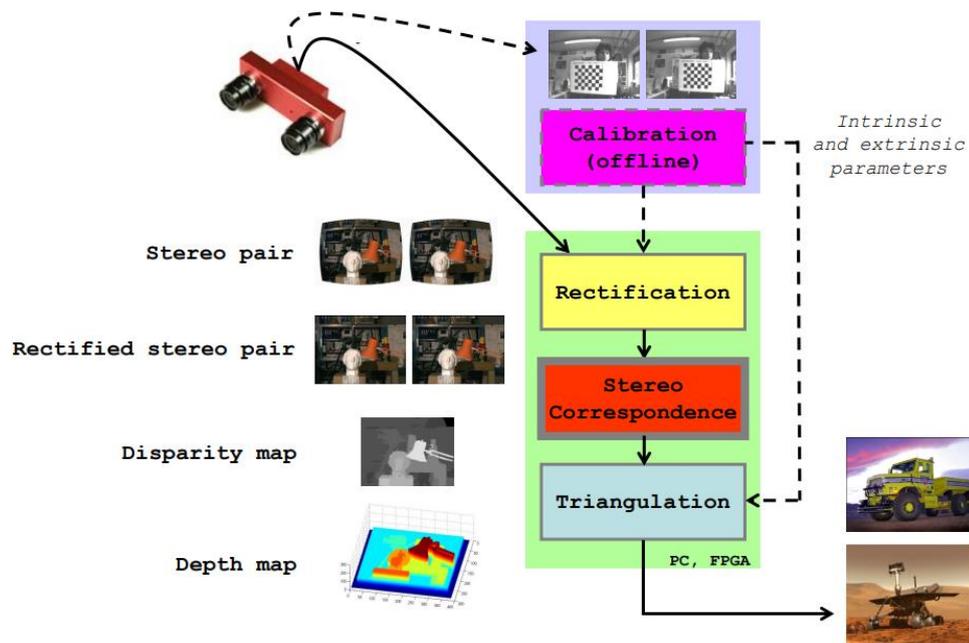
$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix}}_{[\mathbf{R} \quad \mathbf{t}]} \underbrace{\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}}_{\tilde{\mathbf{M}}} \quad (2.7)$$

dimana; α and β : Focal length
 u_0 and v_0 : Principal point
 γ : Koefisien kemiringan
 r : Rotational vector
 t : Translation vector

2.7 Stereo Vision System

Stereo-vision-system adalah teknik yang bertujuan untuk menyimpulkan ke dalam dari dua atau lebih kamera. Ini memainkan peran penting dalam visi komputer. Gambar tunggal tidak memiliki informasi ke dalam atau 3D. *Stereo vision* mengambil dua gambar pemandangan dari sudut pandang yang berbeda biasanya disebut sebagai gambar kiri dan kanan menggunakan dua kamera. *Stereo vision* mirip dengan penglihatan manusia binokular (bermata dua) menangkap dua pandangan yang berbeda dari sebuah adegan dan otak memproses dan mencocokkan kesamaan pada kedua gambar dan perbedaan memungkinkan otak untuk membangun informasi yang mendalam. (George and George, 2014).

Proses *Stereo vision* melibatkan empat langkah utama yaitu *Calibration*, *Rectification*, *Stereo-correspondence* atau *Stereo Matching*, dan Triangulasi hal ini dapat dilihat pada gambar 2.7 (Bradski and Kaehler, 2011).



Gambar 2. 7 Proses utama pembentukan *stereo vision*

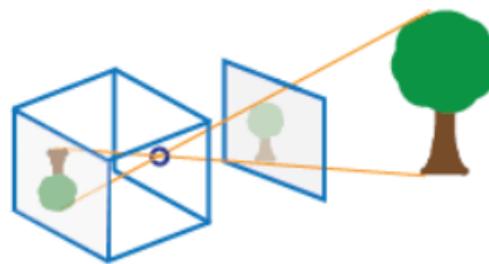
Penjelasan detail mengenai proses pada gambar 2.7 kemudian dapat dilihat pada penjabaran poin 2.8 – 2.12.

2.8 Kalibrasi *Stereo*

Fungsi kalibrasi kamera dalam *Stereo vision* adalah menentukan parameter-parameter geometrik lensa dan mengevaluasi baik kinerja maupun stabilitas antara lensa dan kamera. Untuk mengkalibrasikan kamera, kita harus mengekstrak semua titik sudut dari tiap citra dan kemudian menghitung parameter intrinsik dan ekstrinsik kamera. Parameter intrinsik menunjukkan ciri-ciri internal kamera, seperti: panjang fokus, titik tengah citra, koefisien distorsi, dan parameter penyimpangan lensa, sedangkan parameter ekstrinsik, seperti: orientasi rotasi dan translasi yang menghubungkan dua kamera (Raajan et al., 2012).

Untuk kesederhanaan analisis, penelitian ini menggunakan model kamera *pinhole*. Model kamera *pinhole* tidak memperhitungkan distorsi lensa karena kamera *pinhole* yang ideal tidak memiliki lensa. Untuk merepresentasikan kamera nyata secara akurat, model kamera lengkap yang digunakan oleh penelitian ini mencakup perhitungan distorsi lensa radial dan tangensial. Gambar 2.8 menunjukkan bagaimana model kamera *pinhole* memproyeksikan sebuah objek.

Ada dua jenis distorsi yaitu distorsi tangensial dan radial. Radial distorsi muncul sebagai akibat dari bentuk lensa, distorsi ini adalah nol pada (optik) pusat imager atau bidang gambar dan meningkat saat kita bergerak menuju pinggiran. Distorsi tangensial muncul dari proses perakitan kamera secara keseluruhan. Distorsi ini disebabkan oleh cacat produksi akibat lensa tidak berfungsi tepat sejajar dengan bidang pencitraan. Kedua distorsi ini secara matematis dihapus selama proses kalibrasi menghasilkan gambar yang tidak terdistorsi.



Gambar 2. 8 Model Kamera *pinhole*

2.9 Rektifikasi *Stereo*

Rektifikasi citra adalah proses untuk memproyeksikan dua atau lebih citra ke dalam satu bidang citra. Proses ini memperbaiki penyimpangan lensa dengan mentransformasikan citra ke dalam sistem koordinat standar (Trucco and Verri,

1998). Karena dua kamera menangkap citra dari sudut yang berbeda, citra yang dihasilkan pun berada pada bidang yang berbeda. Rektifikasi dapat dibedakan menjadi dua jenis, yaitu: rektifikasi pada kamera terkalibrasi dan rektifikasi pada hasil citra dari kamera yang tidak terkalibrasi.

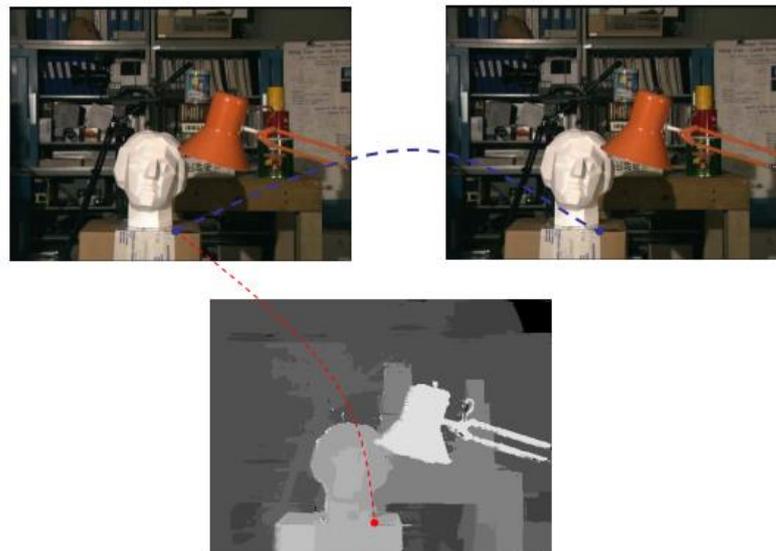
Pada *stereo* kamera yang terkalibrasi, rektifikasi citra memproyeksikan dua citra ke dalam satu bidang citra sehingga ciri-ciri yang sama pada masing-masing citra akan terlihat pada baris yang sama. Proyeksi citra ini membuat citra terlihat seperti dua kamera pada posisi sejajar. Pada rektifikasi dengan kamera tak terkalibrasi mengakibatkan parameter intrinsik dan ekstrinsik kamera tidak diketahui, sehingga proses ini membutuhkan beberapa titik yang berkorespondensi antara dua citra. Untuk menghasilkan titik berkorespondensi ini, kita harus mengumpulkan titik-titik piksel yang serupa dari kedua citra dan kemudian memilih penggabungan yang mungkin antara keduanya.

2.10 Pencocokan *Stereo*

Pencocokan *Stereo* atau Korespondensi *Stereo* bertujuan untuk menemukan titik-titik homolog dalam dua gambar dari pemandangan yang sama diambil menggunakan dua kamera yang berbeda. Diberi poin A di gambar kiri akan ada titik B yang sesuai di gambar kanan. Perbedaan posisi antara titik-titik yang bersesuaian (A dan B) di keduanya gambar disebut sebagai disparitas. Keluaran dari langkah ini adalah peta disparitas. Sepasang Gambar kiri dan kanan dari pemandangan yang sama yang diambil menggunakan kamera *stereo* disebut *stereo-pair*. Ada dua kelas

utama dari algoritma korespondensi yaitu algoritma berbasis korelasi dan berbasis fitur (Xu and Zhang, 1996).

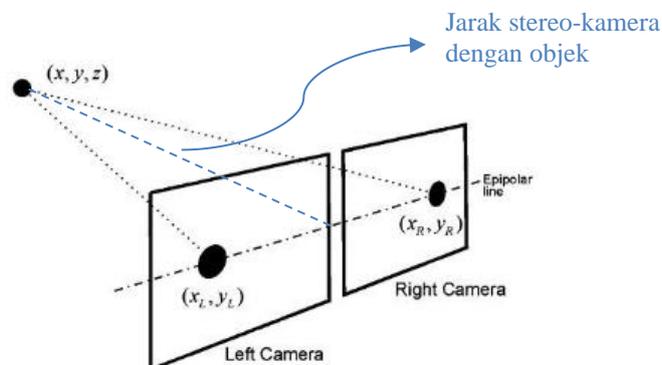
Dalam korespondensi metode berbasis korelasi adalah dicapai dengan mencocokkan intensitas gambar. Hal ini menghasilkan peta disparitas yang padat. Dalam korespondensi metode berbasis fitur dicapai dengan mencocokkan fitur gambar seperti tepi. Teknik ini lebih cepat dan tidak sensitif terhadap perubahan iluminasi dan menghasilkan peta disparitas yang jarang. Beberapa tantangan dalam pencocokan *stereo* adalah inkonsistensi warna, daerah tidak bertekstur, dan masalah oklusi. Gambar 2.9 merupakan bagaimana proses pencocokan stereo (*stereo matching*) menghasilkan peta disparitas



Gambar 2. 9 Pencocokan *stereo* pada dua buah gambar dan hasil peta ke dalam gambar yang berisi peta disparitas

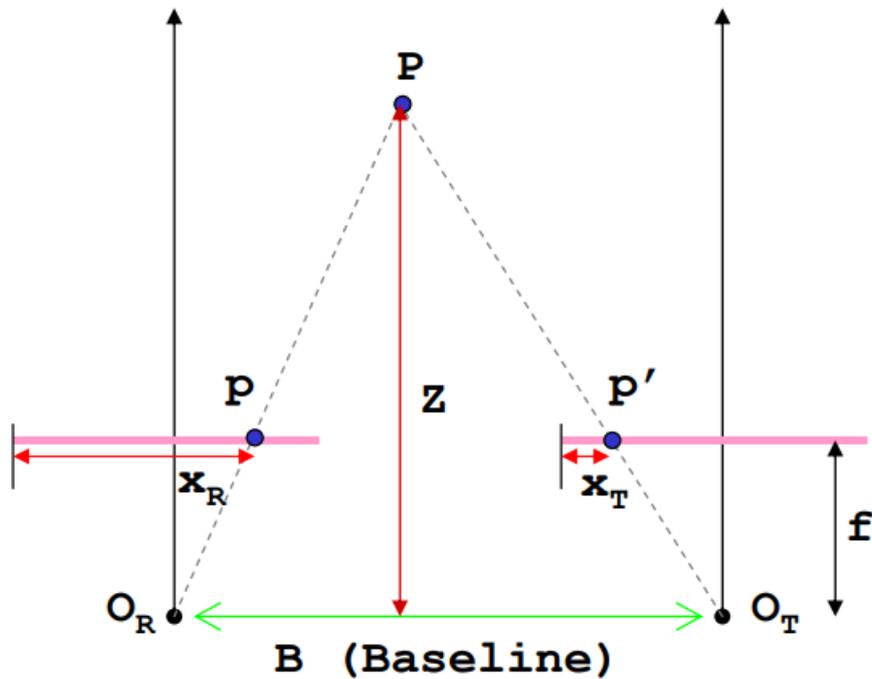
2.11 Triangulasi

Triangulasi merupakan sub-bidang dari keilmuan trigonometri dan geometri dasar. Triangulasi merupakan proses mencari koordinat sebuah objek pada peta 3D serta menentukan jarak titik tersebut dengan mengukur sudut antara titik tersebut dan dua titik referensi lainnya yang sudah diketahui posisi dan jarak antara keduanya. Koordinat dan jarak ditentukan dengan menggunakan hukum sinus. Pada penelitian ini, teknik triangulasi bertujuan untuk menghitung ke dalam atau posisi 3D titik-titik dalam gambar dari peta disparitas dan geometri *stereo* kamera yang telah terkalibrasi. Gambar 2.10 menunjukkan bagaimana prinsip triangulasi dapat menentukan posisi sebuah objek.



Gambar 2. 10 Triangulasi untuk menentukan posisi sebuah objek

Pada langkah ini kalkulasi untuk menentukan jarak sebuah objek terhadap kamera *stereo* yang telah melalui tiga tahap diatas dan dapat dijelaskan dengan perhitungan geometri pada gambar 2.11 berikut (perlu diingat bahwa kondisi berikut merupakan kondisi optimal dimana benda terletak di antara wilayah pandang kedua kamera);



Gambar 2. 11 Model geometri dasar perhitungan jarak pada kamera

Seperti dapat dilihat, dua segitiga besar \$(P, O_R, O_L)\$ dan \$(P, p, p')\$ terbentuk dari geometri *stereo* ini. Dengan demikian, hubungan sisi-sisi yang bersesuaian dapat diturunkan menjadi formula 2.8:

$$\frac{b}{Z} = \frac{(b + x_T) - x_R}{Z - f} \rightarrow Z = \frac{b \cdot f}{x_R - x_T} = \frac{b \cdot f}{d} \quad (2.8)$$

dimana; b : jarak antara lensa (cm)

Z : Jarak benda terhadap kamera (m)

F : focal length (mm)

x_t : posisi objek pada bidang kamera kiri (pixel)

x_r : posisi objek pada bidang kamera kanan (pixel)

D : disparitas (pixel)

2.12 Measurement Method

Measurement method (Metode pengukuran) merupakan metode untuk mendefinisikan suatu barang fisik menjadi satuan nilai tertentu, baik dengan perhitungan langsung maupun tidak langsung melalui perbandingan maupun kalkulasi besaran yang sebelumnya tidak diketahui dengan besaran standar yang diketahui (Khurana, n.d.).

Dalam upaya mendefinisikan nilai, *Measurement Method* dapat dilakukan dengan beberapa cara (Vaidya, 2021):

1. Pengukuran langsung (*Direct Method*) berarti menemukan nilai dengan menggunakan alat terkalibrasi secara langsung tanpa melibatkan proses matematis. Contohnya adalah mengukur voltase dengan voltmeter
2. Pengukuran secara tidak langsung (*Indirect Method*) adalah mengukur besaran lain dan nilai yang diperlukan ditentukan oleh hubungan matematis. Hal ini biasanya disebabkan metode pengukuran lain sudah tidak memungkinkan. Contohnya adalah mengukur volume berdasarkan nilai massa dan massa jenis.
3. Pengukuran komparatif (*Comparative Method*), dimana besaran yang akan diukur dibandingkan dengan nilai yang diketahui dari besaran yang sama. dalam metode ini hanya penyimpangan dari pengukur utama yang ditentukan, misalnya, bagaimana seorang seniman membandingkan ukuran kertas kanvas dengan ukuran objek yang hendak di gambar menggunakan pensilnya.