

BAB 2

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 ArduCam IMX477

ArduCam IMX477 adalah modul kamera berkualitas tinggi yang dirancang untuk digunakan pada platform komputasi seperti Raspberry Pi dan NVIDIA Jetson. Kamera ini menggunakan sensor Sony IMX477, yang dikenal dengan resolusi tinggi dan kemampuan pengambilan gambar yang sangat baik.



Gambar 2.1 ArduCam IMX477

Pada Gambar 2.1 menunjukkan gambar komponen ArduCam IMX477. Spesifikasi ArduCam IMX477 dapat dilihat pada Gambar 2.1:

Tabel 2.1 spesifikasi ArduCam IMX477

No	Atribut	Detail
1.	Resolusi	4056 (H) x 3040 (V)
2.	Megapiksel	12.3 MP
3.	Tegangan Suplai	1.05 V (Digital), 1.8 V (Antarmuka), 2.8 V (Analog)
4.	Kromatisitas	Warna, RGB
5.	Jenis Rana	Rolling Shutter
6.	<i>Frame Rate</i>	15 hingga 240 fps, tergantung pada mode video
7.	Resolusi ADC	10-bit, 12-bit
8.	Ukuran Piksel	1.55 μm (H) x 1.55 μm (V)
9.	Antarmuka	MIPI 2-lane/4-lane, D-PHY V1.2
10.	Kecepatan Data	Maks. 2.1 Gbps/lane
11.	Frekuensi Clock	6-27 MHz
12.	Paket	92 pin LGA Keramik
13.	Diagonal Optik	7.857 mm
14.	Format Sensor Gambar	Tipe 1/2.3"

No	Atribut	Detail
15.	Dimensi Sensor	7.564 (H) x 5.476 (V) mm
16.	Format Video <i>Output</i>	RAW12/10/8, COMP8

2.1.2 NVIDIA Jetson AGX Orin

NVIDIA Jetson AGX Orin adalah platform komputasi yang dirancang khusus untuk aplikasi kecerdasan buatan (AI) dan robotika. NVIDIA Jetson AGX Orin digunakan dalam berbagai aplikasi, seperti robotika, kendaraan otonom, visi komputer, kesehatan, dan bioteknologi. Perangkat ini memiliki dua tipe, yaitu dengan memori 32 GB dan 64 GB. Pada penelitian ini, digunakan NVIDIA Jetson AGX Orin 32 GB.



Gambar 2.2 NVIDIA Jetson AGX Orin

Gambar 2.2 menunjukkan platform komputasi NVIDIA Jetson AGX Orin. Spesifikasi dari NVIDIA Jetson AGX Orin 32GB dapat dilihat pada Tabel 2.2:

Tabel 2.2 Spesifikasi NVIDIA Jetson AGX Orin

No	Atribut	Detail
1.	Performa AI	200 TOPS (INT8)
2.	GPU	Arsitektur NVIDIA Ampere dengan 1792 NVIDIA CUDA core dan 56 Tensor core
3.	Frekuensi Maks GPU	1.3 GHz
4.	CPU	8-core Arm Cortex-A78AE v8.2 64-bit CPU (2MB L2 + 4MB L3)
5.	Frekuensi Maks CPU	2.2 GHz
6.	DLA Accelerator	2x NVIDIA NVDLA v2
7.	Frekuensi Maks DLA	1.2 GHz
8.	Accelerator Vision	PVA v2.0
9.	Memori	32GB 256-bit LPDDR5 (204.8GB/s)
10.	Penyimpanan	64GB eMMC 5.1
11.	CSI Camera	Hingga 6 kamera (via virtual channel), 16 lane CSI-2 (16 virtual channel)
12.	Antarmuka Kamera	16 MIPI CSI-2 lanes

No	Atribut	Detail
13.	Video Encode	2x 4K60, 4x 4K30, 7x 1080p60, 14x 1080p30
14.	Video Decode	1x 8K30, 2x 4K60, 4x 4K30, 6x 1080p60, 12x 1080p30
15.	Antarmuka	1x 16 lanes PCIe Gen4 (Root Port and Endpoint)
16.	Penyimpanan Tambahan	Single-lane UFS
17.	Jaringan	10/100/1000 BASE-T Ethernet
18.	Display	1x 4K60 multi-mode DP 1.4a / eDP 1.4 / HDMI 2.1
19.	I/O	4x UART, 3x SPI, 2x I2S, 8x I2C, 2x CAN, DMIC, 8x DSPK, GPIOs
20.	Daya	15W - 40W
21.	Dimensi	100mm x 87mm
22.	Konektor	699-pin Molex Mirror Mezz Connector

Jetson AGX Orin memiliki performa AI yang tinggi sehingga memungkinkan penggunaan model AI yang besar dan kompleks. Modul ini mendukung integrasi dengan berbagai sensor, sehingga memungkinkan pengembangan sistem dengan kemampuan deteksi objek dan pelacakan yang lebih baik. Versi industri dari Jetson AGX Orin juga dirancang untuk beroperasi dalam kondisi ekstrem dengan rentang suhu -40 °C hingga 85 °C.

2.1.3 Robot otonom

Robot otonom adalah mesin cerdas yang dapat melakukan tugas secara mandiri tanpa memerlukan kendali langsung dari manusia. Pada penelitian ini, robot otonom memiliki kemampuan navigasi. Kemampuan ini diperoleh dengan menggunakan teknologi seperti sensor, pengolahan citra, dan kecerdasan buatan untuk memahami dan menavigasi lingkungannya. Ini akan memungkinkan untuk menghindari rintangan dan memilih jalur terbaik menuju tujuan.



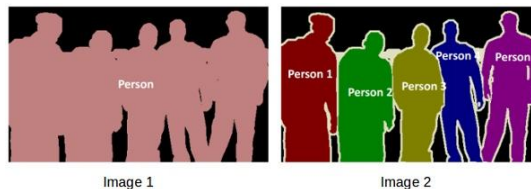
Gambar 2.3 Robot otonom

Gambar 2.3 yang ditampilkan menunjukkan platform robot otonom yang

digunakan dalam penelitian ini. Robot tersebut dirancang dengan konfigurasi empat roda berukuran besar yang memungkinkan mobilitas tinggi di berbagai jenis permukaan, termasuk lingkungan di industri yang tidak rata.

2.1.4 Image processing

Image processing atau pengolahan citra adalah bidang yang mencakup manipulasi dan analisis gambar digital menggunakan algoritma komputer. Tujuan dari proses ini adalah untuk meningkatkan kualitas gambar, mengekstrak informasi penting, dan menyiapkan gambar untuk analisis lebih lanjut. *Image processing* melibatkan konversi gambar dari format analog ke digital agar dapat diproses oleh komputer. Gambar digital terdiri dari piksel, dan setiap piksel memiliki nilai tertentu. Proses ini biasanya dilakukan dalam beberapa langkah, dimulai dari pengambilan gambar, analisis, hingga pemrosesan lebih lanjut.



Gambar 2.4 *Image processing*

Berdasarkan Gambar 2.4, ditunjukkan contoh hasil pengolahan citra pada proses segmentasi. Pada Image 1, objek manusia masih direpresentasikan dalam bentuk siluet tanpa pemisahan antarindividu. Sedangkan pada Image 2, setiap individu telah dipisahkan dan diberi label yang berbeda, sehingga memudahkan proses identifikasi dan analisis objek secara lebih spesifik.

2.1.5 Robot Operating System (ROS)



Gambar 2.5 Logo ROS

Robot Operating System (ROS) memiliki sejarah yang dimulai dari proyek pribadi Keenan Wyrobek dan Eric Berger di Stanford untuk mengatasi kendala robotika, yaitu penemuan kembali roda., dan kemudian dikembangkan oleh Willow

Garage di sebuah laboratorium penelitian robotika, mulai tahun 2007. Kini ROS dikelola oleh Open Robotics.

Robot Operating System (ROS) merupakan kerangka kerja perangkat lunak berbasis middleware yang digunakan untuk mengembangkan aplikasi robotika. ROS menyediakan berbagai tools dan library yang memungkinkan pengembang menghubungkan dan mengendalikan komponen robot secara efisien tanpa harus menulis kode dari awal. Arsitektur ROS didasarkan pada prinsip modularitas, yaitu aplikasi robotik dibangun dari beberapa node yang masing-masing menjalankan tugas tertentu, seperti mengendalikan motor atau membaca sensor. Komunikasi antar node dilakukan melalui topik atau layanan, sehingga memungkinkan pengembangan yang lebih terstruktur dan fleksibel. ROS menyediakan berbagai pustaka dan alat untuk menangani berbagai tugas robotik, termasuk navigasi, manipulasi, dan visi komputer. Alat visualisasi seperti RViz dan simulasi dengan Gazebo dapat sangat membantu pengembang dalam memodelkan dan menguji aplikasi robotik tanpa memerlukan perangkat keras.

ROS memiliki 2 versi utama selama pengembangannya, yaitu ROS 1 dan ROS 2. Pertama kali dirilis pada tahun 2010, ROS 1 telah menjadi standar *de facto* dalam dunia robotika. Namun, sejak ROS 2 dirilis pada tahun 2017, versi ini telah meningkatkan keandalan secara signifikan, dukungan untuk sistem terdistribusi, serta kemampuan untuk bekerja di berbagai platform seperti Linux, Windows, dan macOS. ROS 2 juga mengatasi banyak keterbatasan ROS 1, termasuk kemampuan mengelola sistem *real-time* dan memungkinkan komunikasi yang lebih andal.

2.1.6 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) adalah jenis *neural network* yang dirancang khusus untuk memproses data visual, seperti gambar dan video. *Neural network* adalah program atau model *machine learning* yang membuat keputusan dengan cara yang mirip dengan otak manusia, dengan menggunakan proses yang meniru cara neuron biologis bekerja sama untuk mengidentifikasi fenomena, mempertimbangkan pilihan, dan sampai pada kesimpulan. CNN merupakan bagian dari *deep learning* dan memiliki arsitektur yang memungkinkan pengenalan pola visual secara efisien. *Convolutional Neural Network* (CNN) terdiri dari beberapa lapisan, seperti *Convolutional Layer*, *Pooling Layer*, dan *Fully-Connected Layer*.

1) *Convolutional Layer*

Berfungsi untuk mendeteksi fitur penting dalam gambar, seperti tepi, tekstur, dan pola. Proses ini menggunakan filter (kernel) untuk mengekstrak fitur pada berbagai skala.

2) *Pooling Layer*

Setelah lapisan konvolusi, *Pooling Layer* berfungsi untuk mengurangi dimensi feature maps, sehingga mengurangi jumlah parameter dan komputasi dalam jaringan. Ini membantu menghindari *overfitting* dan mempercepat proses pelatihan.

3) *Fully-Connected Layer*

Pada lapisan ini, neuron-neuron dihubungkan sepenuhnya dengan neuron pada lapisan sebelumnya. Lapisan ini bertugas mengklasifikasikan fitur yang telah diekstraksi ke dalam kategori tertentu.

CNN bekerja dengan memproses gambar sebagai input dalam bentuk *array* piksel yang kemudian dianalisis melalui beberapa tahapan utama. Pada tahap pertama, jaringan melakukan ekstraksi fitur menggunakan *Convolutional Layer*, di mana pola seperti tepi, tekstur, dan bentuk objek diidentifikasi melalui filter (kernel) yang bergerak melintasi gambar. Setelah fitur utama dikenali, sistem melakukan pengurangan dimensi menggunakan *Pooling Layer* yang bertujuan menyederhanakan data tanpa menghilangkan informasi penting, sehingga meningkatkan efisiensi pemrosesan. Selanjutnya, fitur yang telah diekstraksi akan diproses pada *fully-connected layer*, di mana CNN melakukan klasifikasi berdasarkan pola yang telah dipelajari selama pelatihan. Proses ini menghasilkan *Output* berupa label objek yang terdeteksi, yang kemudian dapat digunakan untuk berbagai aplikasi, termasuk deteksi objek pada sistem robot otonom

2.1.7 Deteksi objek dengan YOLOv8

YOLOv8 (*You Only Look Once v8*) adalah algoritma deteksi objek *real-time* berakurasi tinggi yang menggunakan pendekatan *single-stage detector*, yang artinya mendeteksi semua objek dalam satu kali proses inferensi gambar tanpa memerlukan tahap terpisah seperti region proposal. Cara kerja algoritma ini adalah dengan memasukkan gambar ke dalam model, lalu model langsung mendeteksi semua objek sekaligus dalam satu proses tanpa perlu langkah tambahan. YOLOv8

mengekstrak fitur dari gambar, menggabungkan informasi dari berbagai ukuran, lalu memprediksi lokasi dalam kotak pembatas (*bounding box*) serta jenis objek yang terdeteksi.



Gambar 2.6 Contoh hasil deteksi objek dengan YOLOv8

Berdasarkan Gambar 2.6, ditunjukkan hasil deteksi objek dengan YOLOv8, di mana sistem mampu mengenali beberapa objek dalam satu gambar dan menandainya menggunakan *bounding box* beserta label serta nilai kepercayaan (*confidence score*). Proses ini melibatkan analisis citra masukan, ekstraksi fitur menggunakan *Convolutional Neural Network* (CNN), serta penentuan lokasi objek dalam bentuk koordinat *bounding box*.

2.2 Penelitian Terdahulu

Dalam upaya pengembangan sensor multi-modal pada robot otonom untuk keperluan inspeksi, penulis merujuk pada beberapa penelitian sebelumnya sebagai referensi untuk penelitian lebih lanjut mengenai sistem sensor multi-modal pada robot otonom. Beberapa penelitian yang menjadi landasan pengembangan ini akan diuraikan lebih lanjut.

Penelitian (Naufalimam dkk., 2021) berjudul “Penerapan Deteksi dan Pelacakan Objek pada Robot Otonom Pengumpul Bola Tennis Meja Menggunakan Pengolahan Citra” mengembangkan robot otonom yang mampu mendeteksi dan melacak bola tenis meja melalui pengolahan citra. Proses pendeteksian bola dilakukan dalam dua tahap: *pre-processing* dan metode deteksi kontur. *Pre-processing* dilakukan dengan mengubah citra bola berformat RGB menjadi HSV, lalu citra HSV di-*threshold*. Citra kemudian diperhalus melalui operasi morfologi.

Setelah *pre-processing*, dilakukan pendeteksian bentuk bola menggunakan *contour detection* untuk menentukan koordinat bola. Koordinat bola diperlukan untuk pergerakan motor servo yang mengikuti perpindahan bola tersebut. Sistem yang dikembangkan mampu mendeteksi jarak bola dengan akurasi tertinggi 99,607% ketika kamera diam dan 88,304% ketika kamera bergerak. Akurasi pendeteksian ini menunjukkan bahwa sistem pengolahan citra yang digunakan cukup efektif untuk aplikasi robotika.

Pada penelitian (Mahandi, 2021) berjudul “Deteksi Objek untuk Robot Bergerak Menggunakan Kamera Omnidirectional Berbasis Fitur Warna”, robot dapat mendeteksi objek yang bergerak dan mendekatinya. Robot ini menerapkan pemrosesan citra berupa filter warna untuk mendeteksi objek. Kamera pada robot juga dibekali sebuah cermin pada bagian atasnya yang ditujukan untuk memantulkan kondisi sekeliling robot, yang memungkinkan kamera robot dapat menangkap sebesar 360 derajat, atau disebut *omnidirectional camera*. Fitur warna dapat digunakan sebagai salah satu solusi untuk mendeteksi objek menggunakan kamera *omnidirectional*. Kamera ini juga membantu robot mendeteksi objek di sekelilingnya dan bernavigasi mendekati objek dengan performa yang baik, yaitu sebesar 90%.

Kemudian, penelitian (Zophie dkk., 2022) mengimplementasikan algoritma YOLO (*You Only Look Once*) untuk mendeteksi objek menggunakan webcam. Algoritma YOLO dikenal karena kecepatan dan akurasi dalam mendeteksi berbagai jenis objek dalam satu *frame*. Penelitian ini mencakup tahapan mulai dari studi literatur, perancangan dan pemrograman menggunakan Python, hingga pengujian algoritma pada perangkat keras. Fokus penelitian ini adalah pendeteksian kendaraan bermotor, namun teknik yang digunakan dapat diaplikasikan pada berbagai jenis objek lain dalam konteks robotika.

Penelitian berjudul “Deteksi Objek pada Prototipe *Autonomous Car* Menggunakan Sensor Fusion Kamera Mono dan Sensor Ultrasonik” (Hija dkk., 2021). Penelitian ini mengembangkan sistem deteksi objek pada prototipe mobil otonom dengan menggunakan fusi sensor yang menggabungkan sensor ultrasonik HC-SR04 dan kamera mono IMX219-160. Sensor ini dipasang pada purwarupa *Autonomous Car* Jetson AI Car dengan mini-PC Jetson Nano sebagai pemroses

data. Deteksi objek pada kamera menggunakan *Deep Neural network* (DNN) YOLO untuk memperoleh hasil yang lebih akurat. Hasil penelitian menunjukkan bahwa teknologi sensor fusion dapat meningkatkan kemampuan deteksi objek secara signifikan dalam berbagai kondisi operasional. Penelitian tersebut menunjukkan bahwa penggunaan kamera dan teknik pengolahan citra sangat efektif untuk mendeteksi objek pada robot otonom. Oleh karena itu, penelitian ini akan mengembangkan sistem pendeteksian objek berbasis kamera pada robot otonom dengan fokus pada peningkatan akurasi dan kemampuan deteksi dalam berbagai kondisi operasional.

Penelitian (Helnawan dkk., 2023) berjudul “Sistem Segmentasi Jalan dan Objek untuk Kendaraan Otonom Menggunakan Kamera RGB-D”. Penelitian ini membahas teknik segmentasi gambar yang diterapkan pada sistem mobil otonom untuk meningkatkan kemampuan deteksi objek, seperti sinyal lalu lintas, pejalan kaki, dan kendaraan lain. Penelitian ini dilakukan oleh tim di Universitas Teknologi Sepuluh November (ITS) yang memanfaatkan kamera Intel RealSense D435i untuk mengambil gambar beresolusi tinggi serta mengimplementasikan algoritma pengolahan citra untuk segmentasi dan klasifikasi objek. Hasil penelitian menunjukkan bahwa teknik segmentasi yang dikombinasikan dengan *machine learning* dapat mencapai akurasi deteksi lebih dari 90% pada kondisi pencahayaan yang bervariasi. Penelitian ini memberikan kontribusi penting bagi pengembangan sistem sensor multi-modal dengan menekankan peran segmentasi gambar dalam meningkatkan keandalan deteksi objek pada robot otonom.

Tujuan utama penelitian “Pengujian Deteksi Objek Manusia Menggunakan Jetson Nano Dengan Model Ssd Mobilenetv2” (Abdul Aziz dkk., 2024) adalah mengevaluasi dan membandingkan kinerja model SSD MobileNetV2 pada Jetson Nano dalam situasi tertentu untuk menjalankan tugas deteksi objek secara *real-time*. NVIDIA Jetson Nano menunjukkan performa yang sangat baik dalam mendeteksi objek manusia menggunakan model SSD MobileNetV2, dengan *frame rate* mencapai 24 FPS pada mode daya 10 watt dan 17 FPS pada mode daya 5 watt. Meskipun daya dipangkas 50%, performanya tetap tinggi dibandingkan mikroprosesor setingkatnya. Akurasi deteksi pada kondisi lingkungan terang mencapai 86%, remang-remang 53%, objek ramai 85%, dan objek tidak ramai 98%.

Faktor yang memengaruhi *frame rate* adalah daya konsumsi, di mana daya yang lebih besar menghasilkan *frame rate* lebih baik, sedangkan akurasi dipengaruhi oleh kondisi objek seperti tertutup, terpotong, atau buram, serta kualitas kamera dalam kondisi pencahayaan rendah yang dapat menimbulkan motion blur.

Tabel 2.3 Penelitian terdahulu

Kategori	Peneliti 1	Peneliti 2	Peneliti 3	Peneliti 4	Peneliti 5	Peneliti 6
Nama	Muhammad Syechan Naufalimam, M. Zakiyullah Romdlony, Willy Anugrah Cahyadi (2021)	Yogi Dwi Mahandi (2021)	Jeremia Zophie, Hendri Himawan Triharminto (2022)	Muhammad Abdul Aziz1, A Sjamsjiar Rachman, I Made Budi Suksmadana (2024)	Akhmad Ibnu Hija, Katherin Indriawati, Mohammad Harwin Prayoga (2021)	Aldy Helnawan, Muhammad Attamimi, dan Astria Nur Irfansyah (2023)
Metode	Contour Detection & <i>Image processing</i>	Teknik filtering	Algoritma <i>You Only Look Once</i> (YOLO)	Eksperimen	Sensor fusion	<i>Machine learning</i>
Platform / Perangkat	Raspberry Pi 4B + Webcam Logitech B525	Robot dengan Kamera Omnidirectional	Webcam + Python + Darknet/Darkflow	NVIDIA Jetson Nano	Jetson Nano + Kamera Mono + Ultrasonik	Intel RealSense D435i + RTX 3060
Jenis Sensor / Kamera	Webcam Logitech B525	Kamera Omnidirectional 360°	Webcam	Webcam Logitech C270	Kamera Mono IMX219 + HC-SR04	Kamera RGB-D-NIR
Objek yang Dideteksi	Bola tenis meja	Bola pada robot sepak bola	Mobil dan motor	Manusia	Kendaraan & objek sekitar	Jalan, kendaraan, manusia, objek jalan
Pengujian Sistem	Akurasi tracking & sudut servo	Navigasi robot terhadap objek	Akurasi YOLO berdasarkan cahaya & resolusi	<i>Frame rate</i> , confusion matrix, pencahayaan	Akurasi sensor fusion & delay	Akurasi segmentasi & waktu proses
Kondisi Pengujian	Kamera diam dan bergerak	Lingkungan robot sepak bola	Cahaya terang & gelap	Terang, remang, ramai, tidak ramai	Objek diam dan bergerak	Outdoor & minim cahaya
Parameter Pengujian	Akurasi deteksi & tracking servo	Jarak dan sudut objek	Akurasi, confidence, resolusi	FPS, precision, recall, F1-score	FPS, delay, <i>error</i> sensor	Akurasi segmentasi & processing time
Hasil Penelitian	Akurasi deteksi mencapai 99.607%	Berhasil mendekati objek dengan performa 90%	Akurasi YOLO mencapai 69.56%	Jetson Nano 10W memiliki FPS lebih tinggi	mAP mencapai 98.68%	Akurasi segmentasi 39.60%–63.71%

Berdasarkan kajian terhadap beberapa penelitian terdahulu pada Tabel 2.3, diketahui bahwa penelitian-penelitian tersebut telah membahas topik yang relevan dengan penelitian ini, namun memiliki batasan dalam hal jangkauan pengamatan, jenis objek yang dapat terdeteksi, dan parameter pengujian yang dipakai. Kebanyakan penelitian menggunakan satu kamera atau kamera yang dapat melihat ke segala arah, dengan tujuan mendeteksi bola, kendaraan, atau objek umum lainnya. Selain itu, pengujian biasanya hanya mengukur akurasi, *frame rate*, atau kinerja algoritma, tanpa melihat lebih dalam pengaruh jarak deteksi, tingkat cahaya, dan waktu deteksi.

Penelitian ini berbeda dari penelitian sebelumnya. Penelitian ini menciptakan sistem untuk mendeteksi manusia menggunakan enam kamera IMX477 yang dipasang di sekitar robot. Hal ini memungkinkan robot memiliki pandangan 360°. Sistem ini beroperasi pada NVIDIA Jetson AGX Orin dengan menggunakan ROS 2 dan YOLOv8 untuk mendeteksi manusia secara langsung. Selain mengukur seberapa akurat, penelitian ini juga menguji jarak maksimal yang dapat dideteksi, intensitas cahaya paling rendah, serta waktu deteksi pada berbagai kecepatan robot. Dengan begitu, penilaian kinerja sistem menjadi lebih lengkap untuk penggunaan robot inspeksi.