

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 State of The Art**

Terdapat beberapa penelitian yang dijadikan *State of The Art* dan dibandingkan dengan penelitian yang peneliti lakukan sekarang, diantaranya yaitu:

“Rancang Bangun Sistem Pengukuran Posisi Target Dengan Kamera Stereo Untuk Pengarah Senjata Otomatis”. Penelitian ini ditulis oleh A. M. Utama. Penelitian ini bertujuan untuk merancang sistem yang mampu mengidentifikasi target yang dipilih pada satu kamera referensi dan dikenali oleh kamera lain, mengukur posisi target terhadap senjata, mengarahkan senjata secara otomatis ke arah target, dan menentukan sudut elevasi yang tepat agar penembakan oleh senjata dapat mencapai target dengan akurat. Metode yang digunakan dalam penelitian ini merupakan stereo visual untuk menghasilkan citra yang akan digunakan dalam proses triangulasi untuk mendapatkan informasi jarak terget terhadap kamera. Terdapat kelebihan dari penelitian ini yaitu akurasi yang tinggi dalam penentuan posisi target. Dengan adanya kelebihan terdapat juga kekurangannya, yaitu sistem memiliki batasan jarak minimal dan maksimal target, yaitu 100 cm hingga 800 cm [5].

“Sistem Orientasi Objek Dengan Metode Stereo Vision Berbasis Raspberry Pi”. Penelitian ini ditulis oleh R. Ginting, R. Patmasari, dan S. Aulia. Penelitian ini bertujuan untuk mengembangkan sistem pengukuran orientasi objek dengan menggunakan metode stereo vision berbasis Raspberry Pi. Sistem ini dirancang untuk mengukur jarak dan dimensi objek secara visual dengan akurasi di atas 95% dan presisi 100%, yang dapat digunakan dalam sistem otonom untuk navigasi dan identifikasi objek. Metode yang diterapkan mencakup deteksi tepi Canny, transformasi garis Hough, dan deteksi sudut Harris, yang memanfaatkan dua kamera untuk menangkap citra dari sudut pandang berbeda. Kelebihan dari sistem ini meliputi akurasi dan presisi pengukuran yang tinggi serta waktu komputasi yang

cepat saat menggunakan metode tertentu. Untuk kekurangannya adalah stabilitas pengukuran dapat terganggu oleh variasi dalam hasil deteksi tepi dan sudut, serta peningkatan waktu komputasi ketika menggunakan metode pengolahan citra yang lebih kompleks [6].

“Rancang Bangun Vision Sensor Menggunakan Stereo Kamera Pada Sistem Manipulasi Objek”. Penelitian ini ditulis oleh Taryudi. Penelitian ini bertujuan untuk mengembangkan sensor vision menggunakan sistem stereo kamera guna mendeteksi objek tiga dimensi (3D) dalam ruang kerja yang digunakan oleh robot manipulator. Metode penelitian mencakup pengkalibrasian kamera untuk memperoleh parameter intrinsik dan ekstrinsik menggunakan papan catur dengan pola 9x7 dan ukuran kotak 40x40 mm, serta pengujian sistem untuk mendeteksi lokasi objek berwarna dalam ruang kerja menggunakan metode *triangulation*. Kelebihan dari penelitian ini adalah sifatnya yang aman dan non-invasif, akurasi dan jangkauan deteksi yang luas, serta implementasi yang sederhana. Untuk kekurangannya yaitu resolusi kamera yang rendah, ketergantungan pada proses kalibrasi yang tepat, dan kompleksitas pengolahan data dari dua kamera untuk menghasilkan informasi 3D [7].

“Metode Stereo Vision Dengan Kamera CMOS Untuk Pengukuran Jarak” Penelitian ini ditulis oleh A. Finali. Penelitian ini bertujuan untuk mengembangkan metode pengukuran jarak menggunakan teknik stereo vision dengan kamera CMOS. Sistem yang dikembangkan diharapkan dapat mendeteksi dan mengukur jarak objek secara akurat, baik dalam kondisi pencahayaan dalam ruangan maupun luar ruangan. Metode yang digunakan dalam penelitian ini melibatkan penggunaan stereo kamera dengan sensor CMOS untuk mendeteksi dan mengukur jarak objek. Proses pengolahan citra dilakukan dengan teknik *thresholding* dan *edge detection*, menggunakan perangkat lunak OpenCV. Kelebihan dari penelitian ini adalah sistem ini mampu mengukur jarak dengan tingkat kesalahan yang relatif rendah pada objek diam, terutama dalam pencahayaan dalam ruangan. Namun kekurangannya sistem sangat sensitif terhadap perubahan pencahayaan, terutama

dalam kondisi luar ruangan dengan sinar matahari yang terang, yang dapat mempengaruhi hasil pengukuran [8].

“Implementasi Metode Stereo Vision Pada Robot Tempur Cia Versi N2Mr3 Dengan Menggunakan Dua Kamera” Penelitian ini ditulis oleh Nadhif Misbachul Hidayat, Nur Rachman, dan M. Huda. Penelitian ini bertujuan untuk mengembangkan sistem penginderaan visual menggunakan metode stereo vision pada robot tempur. Dengan menggunakan dua kamera yang dipasang sejajar, sistem ini diharapkan dapat mengurangi korban personil akibat kontak senjata langsung dengan musuh di medan pertempuran. Metode penelitian ini mencakup pengukuran jarak *triangulation* dengan bantuan dua stereo kamera vision beresolusi 1080HD. Kelebihan penelitian ini meliputi penggunaan kamera dengan resolusi tinggi membantu mengurangi tingkat kesalahan pengukuran, terutama pada jarak minimal 30 cm dengan kesalahan kurang dari 2,5% dan kekurangannya adalah sistem ini dapat mengalami penurunan kinerja dalam kondisi pencahayaan rendah atau malam hari [9].

Dari hasil peninjauan terhadap penelitian-penelitian sebelumnya, tidak ada yang secara spesifik menggunakan stereo kamera untuk aplikasi pembasmi hama tanaman atau penyemprotan pestisida. Oleh karena itu, penelitian ini dilakukan untuk menjawab kebutuhan tersebut, dengan fokus pada pengembangan sistem yang dapat membantu petani mengatasi masalah hama secara lebih efektif melalui penggunaan teknologi stereo kamera pada lengan robot.

## **2.2 Stereo Kamera (Webcam Logitech C270)**

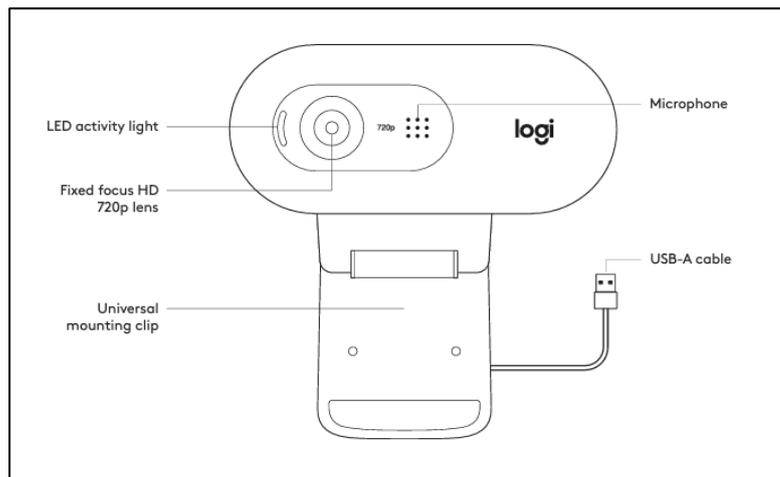
Stereo kamera merupakan jenis kamera yang memiliki dua buah atau lebih lensa dengan sensor gambar yang terpisah atau frame film untuk setiap lensanya. Stereo kamera juga dapat dibuat dengan menggunakan dua buah kamera yang terpisah dan diposisikan secara sejajar. Visualisasi dan ekstraksi informasi spasial dari gambar digital, diambil dari dua sudut pandang yang disebut sebagai *stereo vision*. Dalam kasus umum *stereo vision*, dua kamera ditempatkan secara horizontal dengan jarak dasar tertentu, menghadap ke arah yang sama. Proses kerja dimulai

dengan kalibrasi kamera, diikuti oleh penyempurnaan gambar dan rektifikasi gambar stereoskopik. Jendela stereoskopik harus disesuaikan agar nyaman saat dilihat [10].



**Gambar 2.1** Webcam Logitech C270

Kamera yang digunakan untuk penelitian ini adalah kamera webcam logitech c270. Berikut spesifikasi lengkapnya [11]:

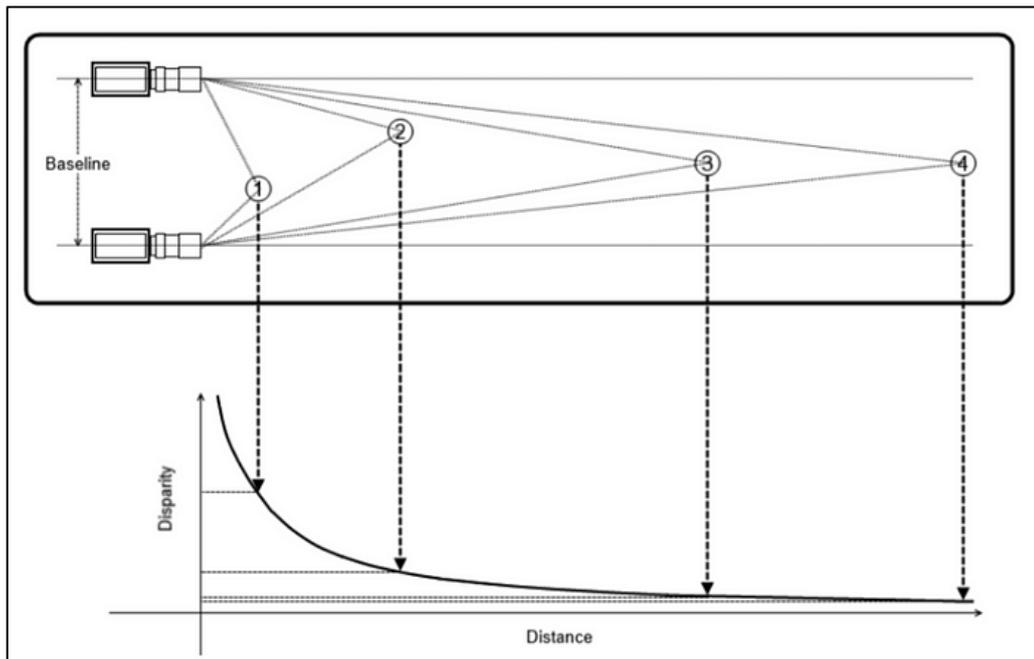


**Gambar 2.2** Spesifikasi webcam Logitech C270

**Tabel 2.1** Spesifikasi webcam Logitech C270

Spesifikasi	Detail
Resolusi	1280 x 720p
Frame rate	30 fps
Sudut pandang	55°

Microfon	Mono
Tipe kabel	USB-A
Panjang kabel	1.5 meter



**Gambar 2.3** Visualisasi Stereo Kamera

Pada gambar 2.3 dapat dilihat terdapat prinsip dasar pengukuran jarak menggunakan stereo kamera dan *triangulation*. Terdapat *baseline* yang merupakan jarak horizontal antara dua kamera yang digunakan untuk pengambilan gambar stereo. Semakin besar *baseline*, semakin besar disparitas yang dapat diukur, dan ini meningkatkan akurasi pengukuran jarak. Kemudian ada disparitas (*Disparity*) yang merupakan perbedaan posisi objek pada dua gambar yang diambil oleh dua kamera. Disparitas diukur sebagai perbedaan koordinat piksel dari objek yang sama pada dua gambar. Disparitas ini berbanding terbalik dengan jarak, semakin jauh objek, semakin kecil disparitasnya. Hasil pengukuran jarak ini bisa didapatkan menggunakan teknik *triangulation* dengan rumus :

$$Z = \frac{fxb}{d} \quad (1)$$

- Z = Jarak dalam meter
- f = *Focal length* dalam piksel
- b = *Baseline* dalam meter
- d = Disparitas dalam piksel

*Focal length* adalah jarak antara lensa kamera dan sensor gambar ketika subjek berada dalam fokus.

Disparitas (*Disparity*) adalah perbedaan posisi suatu titik tertentu dalam gambar yang ditangkap oleh kamera kiri dan kanan pada sistem *stereo vision* dengan rumus :

$$d = x_L - x_R \quad (2)$$

$x_L$  = Koordinat horizontal objek pada gambar yang diambil oleh kamera kiri

$x_R$  = Koordinat horizontal objek pada gambar yang diambil oleh kamera kanan

*Baseline* adalah jarak antara dua kamera dalam sistem *stereo vision* [12].



**Gambar 2.4** Gambar perbandingan tomat untuk pencarian disparitas

Terlihat pada gambar 2.4 pencarian disparitas bisa dicari dengan melihat perbedaan posisi objek yang sama ketika terlihat dari sudut pandang yang berbeda. Disparitas (d) adalah selisih antara posisi dari objek dalam gambar kiri dan kanan.

### 2.3 Tanaman Tomat



**Gambar 2.5** Tanaman Tomat

Tanaman tomat merupakan salah satu jenis buah yang dikategorikan sebagai sayuran penting di Indonesia. Buahnya memiliki sumber vitamin dan mineral yang sangat berguna bagi tubuh manusia. Pertumbuhan tomat akan didukung dengan adanya unsur hara yang cukup. Tomat berada di urutan kelima produksi tanaman sayuran di Indonesia yang produksinya naik-turun setiap tahun [13]. Selain itu buah tomat juga cukup penting karena memiliki cita rasa yang lezat, baik berupa buah segar atau saus. Tomat merupakan tanaman yang sering dijumpai di pasaran, warnanya yang cerah sungguh menarik dan memikat para pembeli [14].

### 2.4 Pengolahan Citra

Pengolahan citra adalah teknik yang digunakan untuk memanipulasi atau mengubah citra dengan tujuan tertentu. Ini bisa digunakan untuk memperbaiki kualitas citra, mengubah citra, mengubah citra analog menjadi digital, atau bahkan untuk mengevaluasi kualitas suatu produk tanpa merusak produk itu sendiri. Citra dalam perwujudan dapat bermacam-macam, mulai dari gambar putih pada sebuah

foto yang tidak bergerak sampai pada gambar berwarna yang bergerak pada televisi. Beberapa aspek pada pengolahan citra :

1. Pemrosesan Citra. Melibatkan transformasi citra menjadi citra lain. Masukan citra dan keluaran pun citra, keluaran citra cenderung memiliki kualitas lebih baik.
2. Citra Digital. Citra digital adalah representasi citra melalui *sampling* secara ruang dan waktu. *Sampling* secara ruang berdasarkan koordinat sinyal (x,y), sedangkan secara waktu melibatkan sederetan citra yang bergerak, yang dikenal sebagai video digital.
3. Peningkatan Kualitas Citra. Teknik-teknik pengolahan citra dapat digunakan untuk memperbaiki kualitas citra, seperti mengurangi *noise*, meningkatkan kontras, atau mempertajam citra.
4. Penggunaan. Pengolahan citra digital dapat digunakan dalam berbagai bidang, seperti pengawasan keamanan, pengenalan pola, pengolahan citra medis, dan masih banyak lagi [15].

## 2.5 Deteksi Objek dan Pengenalan Objek

Deteksi objek merupakan proses menemukan *instance* objek dari kelas tertentu, seperti wajah, mobil, dan pohon, dalam gambar atau video. Algoritma deteksi objek memanfaatkan pembelajaran mesin (*Machine Learning*) atau pembelajaran mendalam (*Deep Learning*) untuk menghasilkan hasil yang signifikan. Ketika manusia melihat gambar atau video secara langsung mereka akan dengan cepat mengenali objek tersebut, berbeda dengan komputer yang memerlukan komputasi yang kompleks untuk mengenali objek. Tujuan dari deteksi objek adalah untuk menemukan dan menentukan lokasi dari suatu objek dalam gambar atau video [16].

Pengenalan objek adalah kemampuan suatu sistem komputer untuk mengidentifikasi dan memahami objek-objek dalam gambar atau video. Pengenalan objek merupakan salah satu permasalahan dari penglihatan komputer (*Computer Vision*) yang berhubungan dengan pengenalan suatu objek dalam citra. Terdapat

dua pendekatan dalam pengenalan objek yang menggunakan pembelajaran mendalam (*Deep Learning*) yaitu pendekatan berdasarkan *region proposal* atau metode dua tahap [17].

## 2.6 *Machine Learning*

*Machine Learning* (ML) adalah mesin yang dikembangkan untuk bisa belajar dengan sendirinya tanpa arahan dari penggunanya. Teknologi ini memungkinkan mesin untuk mempelajari data yang ada dan melakukan tugas-tugas tertentu sesuai dengan apa yang ia pelajari. Istilah *machine learning* pertama kali dikemukakan oleh beberapa ilmuwan matematika seperti Adrien Marie Legendre, Thomas Bayes dan Andrey Markov pada tahun 1920-an dengan mengemukakan dasar-dasar *machine learning* dan konsepnya.

*Machine Learning* dapat dibagi menjadi 3, diantaranya :

1. *Supervised Learning*, adalah jenis *machine learning* yang memanfaatkan data yang telah diberi label untuk melatih algoritma dan model sehingga dapat mengklasifikasikan atau memprediksi data dengan akurasi yang tinggi dan dibawah pengawasan. Yang termasuk dalam *supervised learning* diantaranya adalah *classification* dan *regression*.
2. *Unsupervised Learning* adalah teknik *machine learning* yang memungkinkan sistem komputer untuk menganalisis dan mengelompokkan data yang tidak diberi label dan pembelajarannya tanpa pengawasan. Yang termasuk dalam *unsupervised learning* diantaranya adalah *clustering* dan *dimensionality reduction*.
3. *Reinforcement Learning* adalah algoritma *machine learning* yang bisa membuat mesin bekerja secara otomatis untuk menentukan perilaku ideal sehingga dapat memaksimalkan kinerja algoritma. *Reinforcement learning* cocok untuk analisis data eksplorasi, strategi *cross-selling*, segmentasi pelanggan, dan pengenalan gambar [18].

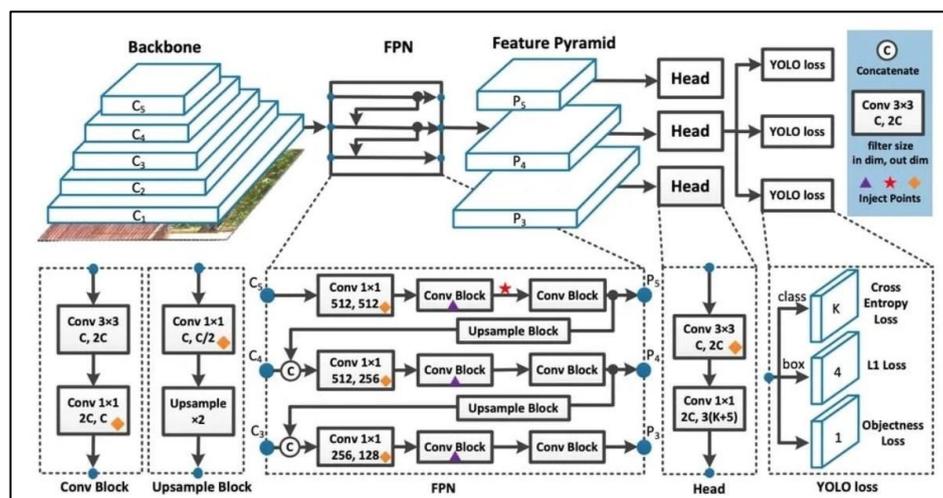
## 2.7 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan jenis jaringan saraf tiruan yang terutama digunakan untuk memproses data yang memiliki struktur grid, seperti gambar. CNN dapat digunakan untuk mendeteksi dan mengenali suatu objek pada suatu gambar. CNN terdiri dari neuron yang memiliki *weight*, *bias*, dan *activation function* [19].

### 2.7.1 You Only Look Once Version 8 (YOLOv8)

You only look once (YOLO) merupakan sebuah algoritma yang dikembangkan untuk mendeteksi objek secara langsung atau *real-time*. YOLO memanfaatkan jaringan saraf tunggal untuk melakukan deteksi dan pengenalan objek, di mana jaringan ini secara langsung memprediksi koordinat *bounding box* dan probabilitas kelas objek dalam satu langkah deteksi. YOLO menggunakan CNN untuk mendeteksi objek dalam gambar.

YOLO membagi gambar input menjadi grid berukuran  $S \times S$ , di mana setiap sel dalam grid bertanggung jawab untuk mendeteksi objek jika pusat objek berada dalam sel tersebut. Setiap sel memprediksi sejumlah *bounding box* (B) dan *confidence score* yang menunjukkan probabilitas bahwa *bounding box* tersebut berisi objek serta seberapa tepat *bounding box* tersebut [20].



Gambar 2.6 Arsitektur YOLOv8

Arsitektur YOLOv8 terdiri dari tiga bagian utama: Backbone, FPN (Feature Pyramid Network), dan Head. Backbone berfungsi untuk mengekstraksi fitur dasar dari gambar input melalui beberapa lapisan konvolusi. FPN kemudian menggabungkan fitur dari berbagai skala dalam backbone untuk menghasilkan peta fitur yang kaya, yang disebut Feature Pyramid. Selanjutnya, Head memproses peta fitur ini untuk menghasilkan prediksi bounding box dan klasifikasi objek. Proses prediksi ini dioptimalkan menggunakan tiga jenis fungsi kerugian (loss function): Cross Entropy Loss, L1 Loss, dan Objectness Loss, yang semuanya digunakan untuk meningkatkan akurasi deteksi objek selama pelatihan [21].

## **2.8 Dataset**

Dataset merupakan sekumpulan data yang disusun secara terstruktur. Dataset biasanya dipresentasikan dalam bentuk tabel. Dataset digunakan untuk memperhatikan hubungan antar variabel. Khususnya jika jumlah data dan variabel yang diteliti bervariasi. Dalam penelitian ini, dataset yang digunakan adalah data mengenai buah tomat busuk yang terkena hama. Dataset tersebut akan digunakan untuk pembuatan model *object detection* untuk buah tomat busuk yang terkena hama.