

## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Robot Pingpong Interaktif

Robot pingpong interaktif adalah robot yang memiliki kemampuan interaktif untuk bermain pingpong dengan manusia. Robot ini memiliki beberapa sensor yang memungkinkannya mendeteksi bola yang datang dan mengembalikannya ke meja pemain. Sensor-sensor ini biasanya termasuk sensor sentuh, suara, kamera, getaran, dan lainnya.

Robot pingpong interaktif dapat digunakan untuk hiburan, pelatihan, pengembangan keterampilan, dan pengujian.

Pada tahun 2022, telah dipasarkan produk robot pelontar bola pingpong FastPong [1] dengan fitur interaktif, seperti terlihat pada gambar berikut.



*Gambar 2.1: Robot Pelontar Bola Pingpong Interaktif FastPong [1]*

Robot ini menggunakan sejenis "panel" fleksibel yang dapat dibentangkan di atas permukaan meja sebagai sarana merealisasikan fitur interaktifnya. Panel ini memiliki rangkaian sensor dan elektronik didalamnya dan terhubung ke robot melalui kabel. Fitur interaktif dihasilkan ketika bola dilontarkan kembali oleh pemain dan menyentuh panel tersebut maka robot akan merespon dengan melontarkan bola lain dari tempat penampungannya ke arah pemain. Keuntungan dari solusi ini adalah panel sensor tidak terpengaruh oleh parameter lingkungan sekitar misalnya suara dan cahaya. Kelemahan dari solusi ini adalah pantulan bola

pada panel berbeda dengan pada meja, sebagian permukaan meja pingpong tertutup sehingga pemain tidak merasakan berlatih pada meja lengkap dan tidak praktis dalam hal pemasangannya.

Sebelum produk ini muncul pada tahun 2019, jenis robot pelontar bola pingpong dengan fitur interaktif ini juga sudah dikembangkan [3]. Berbeda dengan sebelumnya, teknik pendeteksian bola yang dikembalikan oleh pemain adalah menggunakan laser yang dikonfigurasi dalam bentuk matrik X dan Y, seperti pada gambar berikut.



*Gambar 2.2: Robot Pelontar Bola Pingpong Interaktif dengan Jaring laser [3]*

Keuntungan dari teknik ini adalah kecepatan respon deteksinya yang tinggi. Sedangkan kelemahannya solusi ini adalah tingkat kerumitan struktur mekanik dan elektroniknya yang tinggi, mahal dan perawatan yang sulit.

Setahun kemudian pada tahun 2021, robot pelontar bola pingpong dengan fitur interaktif telah dikembangkan [5]. Pendeteksian bola terpantulnya menggunakan solusi sensor getar [4] mirip dengan [1], perbedaannya terletak pada panel-panel tidak fiks namun bisa di atur-atur letaknya di atas meja. Kemudian pada tahun 2023, panel-panel ini dikembangkan menjadi modul-modul wireless [11].

Pendeteksian koordinat bola pingpong untuk tujuan pengembangan aplikasi pelatihan penempatan bola ke sasaran di meja pingpong tidak lepas dari sistem pendeteksian bola balasan dari pemain dan pendeteksian bola terpantul di atas meja yang dapat dideteksi dengan berbagai teknik yang telah disampaikan pada paragraf-paragraf di atas.

Produk yang memiliki tujuan yang sama dengan penelitian ini (pendeteksian koordinat bola, pendeteksian bola berada pada koordinat yang tepat dan

pengaplikasiannya untuk pelatihan ketepatan sasaran pengembalian bola oleh pemain) adalah seperti pada [1], seperti yang telah dijelaskan di atas. Perbedaan yang sangat mendasar adalah pada penelitian ini akan digunakan kamera, pengolahan citra dan kecerdasan buatan sebagai dasar desainnya.

## **2.2 Convolution Neural Network (CNN)**

Convolutional Neural Network (CNN) [14] adalah jenis jaringan saraf tiruan yang dibuat untuk mengelola pola visual, terutama di bidang komputer visi. CNN terkenal dapat menangani masalah analisis gambar dan data visual dengan menggunakan struktur spasial data.

Lapisan konvolusi dan pooling yang digunakan oleh CNN membedakannya dari jaringan saraf tiruan konvensional. Dengan cara yang mirip dengan cara manusia mengenali objek dalam gambar, lapisan-lapisan ini dimaksudkan untuk mengekstraksi fitur visual dari data gambar.

Selain segmentasi semantik, pengenalan wajah, dan deteksi objek, CNN telah membuat kemajuan besar dalam bidang pengenalan gambar. Ini telah digunakan secara luas dalam berbagai aplikasi, termasuk analisis medis dan pengenalan tulisan tangan, serta mobil otonom.

## **2.3 Framework Yolo (You Only Look Once)**

YOLO, yang berarti "You Only Look Once", adalah framework atau pendekatan kecerdasan buatan yang memungkinkan objek diidentifikasi dalam gambar atau video secara real-time [15]. Karena kemampuan mendeteksi objek dengan cepat dan akurat, metode ini sangat populer di komunitas komputer visi dan pengolahan citra.

Menurut framework YOLO, masalah deteksi objek adalah masalah regresi. Artinya, YOLO melihat gambar secara keseluruhan dan mencoba menemukan kotak pembatas (bounding box) dan menemukan jenis objek yang ada di dalamnya dalam satu langkah daripada menggunakan metode tradisional yang membagi gambar menjadi beberapa wilayah kecil dan mengidentifikasi objek di masing-masing wilayah.

Keunggulan utama YOLO adalah kecepatan; dia dapat mendeteksi objek secara real-time bahkan dalam video berkecepatan tinggi berkat pendekatan langsung dan efektifnya. YOLO juga dapat mendeteksi berbagai objek secara bersamaan dalam satu gambar dengan mudah, seperti terdapat pada penelitian berikut [12][13].

YOLO, bagaimanapun, memiliki kekurangan. Deteksi objek yang sangat kecil atau dengan perbedaan skala yang signifikan mungkin tidak seakurat metode deteksi lainnya. Selain itu, YOLO cenderung kurang memperhatikan detail objek dalam deteksi tumpang tindih atau sangat dekat.

YOLO adalah teknik deteksi objek yang menggunakan jaringan saraf tiruan, terutama jaringan konvolusi, untuk mengidentifikasi dan mengklasifikasikan objek dalam gambar atau video. Kombinasi antara konsep deteksi objek, klasifikasi multi-label, dan penggunaan jaringan saraf adalah inti dari bagaimana YOLO berfungsi.

## **2.4 Pengenalan Objek**

Di bidang kecerdasan buatan, pengenalan objek didefinisikan sebagai kemampuan sistem komputer atau perangkat lunak untuk mengidentifikasi dan mengklasifikasikan objek yang terlihat dalam gambar atau video. Pengenalan objek adalah salah satu aspek penting dalam bidang komputer visi, yang merupakan cabang dari kecerdasan buatan yang berfokus pada pemahaman dan interpretasi data visual. Tujuan utamanya adalah untuk memungkinkan mesin "melihat" dan "memahami" konten visual seperti manusia.

Beberapa tahapan digunakan dalam proses pengenalan objek, seperti berikut:

1. Deteksi Objek: Dengan menggambar kotak pembatas, juga dikenal sebagai kotak pembatas, Anda dapat menemukan lokasi dan ukuran objek dalam gambar. Ini menunjukkan di mana objek berada.
2. Ekstraksi Fitur: Proses mengubah gambar menjadi representasi numerik yang dapat diolah oleh algoritma komputer dikenal sebagai ekstraksi fitur. Ekstraksi pola warna, bentuk, tekstur, dan fitur lainnya yang membantu dalam mengenali objek dapat termasuk dalam kategori ini.

3. **Klasifikasi Objek:** Setelah fitur diekstraksi, algoritma klasifikasi digunakan untuk mengenali objek berdasarkan fitur-fiturnya. Ini melibatkan proses membandingkan fitur-fitur yang diekstraksi dengan pola yang telah dipelajari dalam pelatihan sebelumnya.
4. **Pelabelan dan Identifikasi:** Setelah objek terdeteksi dan diklasifikasikan, sistem dapat memberi mereka label atau nama yang sesuai. Misalnya, sistem dapat mengenali objek sebagai "manusia" ketika mereka digambarkan sebagai orang.

Konsep True Positif (TP), True Negatif (TN), False Positif (FP), dan False Negatif (FN) sangat penting untuk analisis kinerja dalam banyak situasi, terutama ketika menilai model klasifikasi dan pengujian.

1. **True Positif (TP):** Ini adalah kasus ketika model klasifikasi memprediksi kelas positif dengan benar, artinya model benar-benar mengidentifikasi kasus yang sebenarnya adalah positif.

2. **True Negatif (TN):** Ini adalah kasus ketika model klasifikasi memprediksi kelas negatif dengan benar, artinya model benar-benar mengidentifikasi kasus yang sebenarnya adalah negatif.

3. **False Positif (FP):** Ini adalah kasus ketika model klasifikasi secara salah memprediksi kelas positif ketika sebenarnya adalah negatif. Dengan kata lain, model membuat kesalahan dengan mengklasifikasikan sesuatu sebagai positif ketika itu seharusnya negatif.

4. **False Negatif (FN):** Ini adalah kasus ketika model klasifikasi secara salah memprediksi kelas negatif ketika sebenarnya adalah positif. Dengan kata lain, model membuat kesalahan dengan mengklasifikasikan sesuatu sebagai negatif ketika seharusnya positif.

Akurasi, presisi, sensitivitas (recall), dan spesifisitas adalah beberapa metrik evaluasi model yang umumnya digunakan. Dengan menggunakan nilai-nilai TP, TN, FP, dan FN, Anda dapat mengukur seberapa baik model Anda dalam mengklasifikasikan data dan memahami tingkat kesalahan yang dibuat oleh model tersebut. Dengan informasi ini, Anda dapat mengoptimalkan model Anda dan membuat keputusan yang lebih baik dalam konteks aplikasi Anda, seperti

membuat model klasifikasi dalam pembelajaran mesin atau dalam bidang kedokteran untuk mendiagnosis penyakit.

## 2.5 LOSS dan mAP

"LOSS" dan "mAP" adalah metrik yang digunakan untuk mengukur kinerja model pada data pelatihan dalam pelatihan model machine learning, terutama dalam tugas-tugas seperti pengenalan objek, deteksi objek, atau klasifikasi. Mari kita bahas keduanya secara lebih rinci:

1. LOSS (Kehilangan): Loss adalah metrik yang menunjukkan seberapa jauh prediksi model dari nilai target yang sebenarnya. Tujuannya adalah untuk mengetahui seberapa baik model belajar menghubungkan masukan ke keluaran yang diharapkan. Loss biasanya digunakan dalam tugas-tugas seperti klasifikasi dan deteksi objek untuk mengukur perbedaan antara prediksi model dan label yang benar pada setiap contoh data pelatihan.

Bergantung pada tugas yang dilakukan, berbagai jenis fungsi loss digunakan. Cross-Entropy Loss (untuk klasifikasi) dan Mean Squared Error (MSE) Loss (untuk regresi) adalah beberapa contoh fungsi loss dan berbagai variasi loss lainnya yang cocok untuk tujuan tertentu.

2. mAP (mean Average Precision): Metrik mAP biasanya digunakan untuk mengevaluasi kinerja model dalam tugas deteksi objek. Ini menentukan seberapa baik model dapat mengidentifikasi dan menempatkan kotak pembatas—juga dikenal sebagai kotak pembatas—tepat di sekitar objek yang ada dalam gambar.

mAP adalah nilai rata-rata dari nilai Precision-Recall (PR) yang dihitung pada berbagai ambang batas (threshold). Precision mengukur seberapa banyak dari deteksi yang dihasilkan model adalah benar, sementara Recall mengukur seberapa banyak dari objek yang sebenarnya ditemukan oleh model.

mAP memberikan gambaran tentang seberapa baik model mendeteksi dan mengklasifikasikan objek pada berbagai ambang batas.

Untuk mengukur kinerja model selama pelatihan, LOSS dan mAP digunakan. LOSS menunjukkan seberapa baik model belajar dari data pelatihan, sementara mAP menunjukkan kemampuan model untuk menemukan objek dan

menempatkan kotak pembatas dengan benar. Tujuannya adalah untuk membantu mengoptimalkan model agar bekerja lebih baik pada data pengujian.