

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Landasan Teori**

##### **2.1.1 Sistem Deteksi**

Sistem deteksi merujuk pada rangkaian teknologi dan algoritma yang digunakan untuk mengenali dan mengidentifikasi objek atau pola tertentu dalam data, seperti gambar atau video. Melibatkan penggunaan teknik-teknik seperti pemrosesan citra, pembelajaran mesin, dan jaringan saraf tiruan untuk melakukan tugas-tugas seperti deteksi objek, klasifikasi, segmentasi, dan pelacakan (Russakovsky dkk., 2015).

##### **2.1.2 Sepak Bola**

Sepak bola adalah olahraga tim yang dimainkan oleh dua tim, dengan tujuan mencetak gol ke gawang lawan. Pertandingan terdiri dari dua babak, masing-masing berlangsung selama 45 menit. Permainan ini, pemain menggunakan berbagai bagian tubuh untuk mengontrol bola, sementara wasit bertanggung jawab menegakkan aturan permainan. Bola, yang memiliki ukuran standar, digunakan untuk memainkan permainan. Semua elemen ini, bersama-sama, membentuk esensi dari permainan sepak bola yang menarik dan kompetitif (Nelson, 1994).

##### **2.1.3 Deteksi Objek**

Teknologi deteksi objek digunakan untuk mengidentifikasi keberadaan dan atribut suatu objek dalam gambar, termasuk parameter dan posisinya. Pengenalan objek membedakan setiap kelas objek yang ada dalam data latihan, dimulai dari identifikasi objek itu sendiri. Ada dua jenis deteksi objek, yaitu deteksi lunak (soft detection) yang hanya menentukan kehadiran objek, dan deteksi keras (hard detection) yang mampu menentukan kehadiran serta posisi objek dalam citra (Suherman dkk., 2023). Proses deteksi objek melibatkan pencarian keseluruhan gambar untuk menemukan pola geometris yang sesuai dengan objek yang dicari dalam basis data. Deteksi

terjadi ketika terdapat korelasi yang cukup tinggi antara template dan area yang diukur dalam gambar, melalui pemindaian semua posisi, skala, dan rotasi pada template untuk setiap gambar. Baru-baru ini, teknologi pengenalan objek berbasis citra telah terbukti akurat dan presisi dalam menghadapi data latihan (Jalied dan Voronkov, 2016).

#### **2.1.4 Citra Warna dan Klasifikasi Citra**

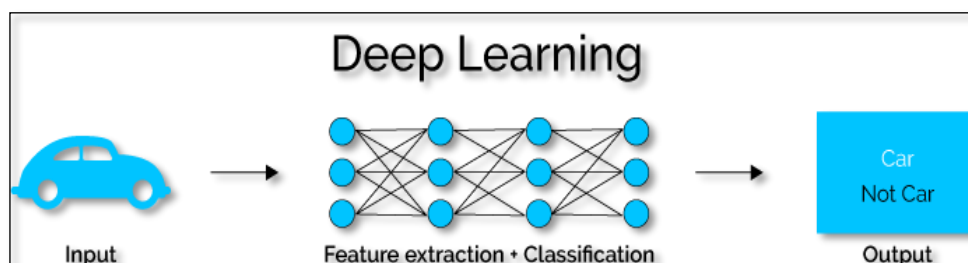
Citra warna adalah gambar visual yang menggambarkan objek, lanskap, atau fenomena dengan beragam warna. Model warna RGB merupakan cara umum untuk merepresentasikan warna dalam pengolahan citra. Model ini terdiri dari tiga komponen warna utama, yaitu merah (red), hijau (green), dan biru (blue). Tiap piksel dalam citra warna diwakili oleh campuran nilai-nilai dari ketiga komponen warna tersebut, yang menentukan kecerahan dan campuran warna pada setiap titik dalam citra (McAndrew, 2004)

Klasifikasi citra adalah proses mengelompokkan atau mengidentifikasi gambar digital ke dalam kategori atau kelas tertentu berdasarkan ciri-ciri visual atau atribut yang dimilikinya. Pola-pola dalam citra sering dikenali menggunakan teknik-teknik pengolahan citra dan pembelajaran mesin. Tujuan utama dari klasifikasi citra adalah untuk mengenali atau membedakan objek, pola, atau fitur yang ada dalam citra (Haralick dkk., 1973).

#### **2.1.5 Deep Learning**

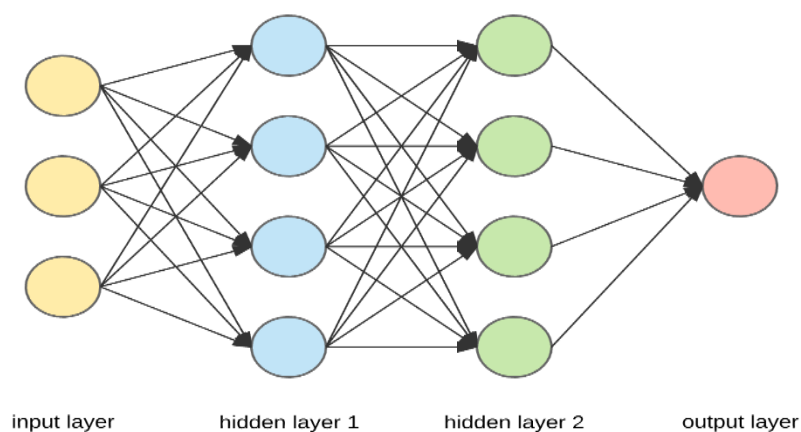
*Deep Learning* adalah algoritma jaringan saraf tiruan yang menggunakan metadata sebagai input, lalu memprosesnya melalui beberapa lapisan tersembunyi untuk menghasilkan output. Keunggulan utamanya adalah kemampuan untuk mengekstraksi fitur secara otomatis, mengurangi beban pemrograman dalam memilih fitur yang diperlukan untuk memecahkan masalah tertentu. Algoritma ini dapat digunakan untuk masalah yang memerlukan pengawasan, tanpa pengawasan, atau semi terawasi, seperti pengenalan citra, suara, atau teks. Setiap lapisan tersembunyi dalam deep learning bertanggung jawab untuk melatih fitur-fitur unik berdasarkan output dari lapisan sebelumnya.. Semakin banyak lapisan tersembunyi, algoritma ini akan semakin kompleks dan abstrak, memungkinkan penyelesaian

masalah yang lebih rumit. Jaringan saraf dalam *deep learning* dapat terdiri dari beberapa lapisan sederhana hingga banyak lapisan kompleks, memungkinkan pemecahan masalah yang lebih kompleks (LeCun dkk., 2015).



Gambar 2. 1 Struktur model *deep learning* (LeCun dkk., 2015)

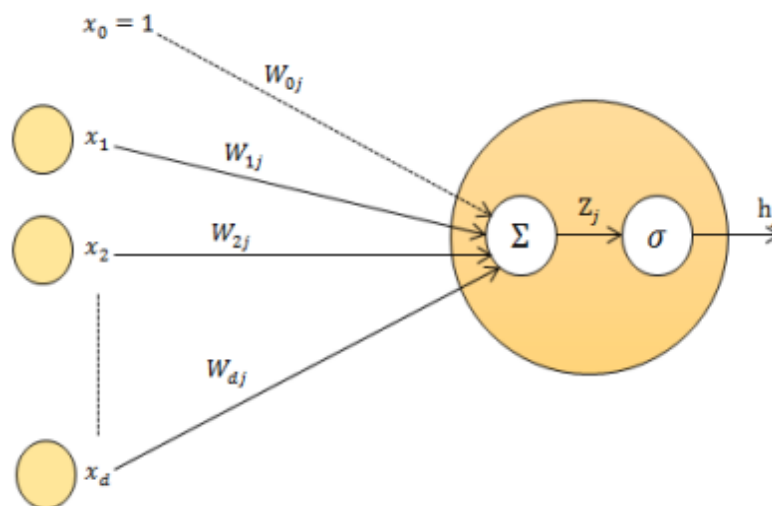
Gambar 2.1 menggambarkan struktur permodelan jaringan dalam *deep learning*. Input dari model ini adalah gambar mobil, yang kemudian melewati serangkaian lapisan neuron yang bertindak sebagai fitur ekstraksi dan klasifikasi. Setiap lapisan tersebut bertanggung jawab untuk mengenali dan mengekstraksi fitur-fitur penting dari gambar mobil. Proses ini menciptakan representasi yang semakin abstrak dan kompleks dari gambar input. Akhirnya, output dari jaringan ini adalah gambar yang telah diklasifikasikan



Gambar 2. 2 Kumpulan layer pada jaringan *deep learning*

Gambar 2.2, terdapat ilustrasi struktur jaringan pada *deep learning*. layer-layer ini melibatkan satu input layer yang ditandai dengan warna kuning, beberapa *hidden layer*, dan satu output layer yang diberi warna merah muda. Hidden layer

pertama ditampilkan dengan warna biru, sementara hidden layer kedua ditampilkan dengan warna hijau. Setiap layer memiliki *neuron-neuron* yang saling terhubung, di mana koneksi antar-neuron terjadi antara layer-layer yang berbeda. Menunjukkan bagaimana data masukan diproses melalui jaringan dan menghasilkan output yang diinginkan.



Gambar 2. 3 *Perceptron* (LeCun dkk., 2015)

Sebuah konstruksi sistem yang melibatkan sebuah neuron beserta input dan output yang terkaitnya dikenal sebagai perceptron. Sebagaimana tergambar pada ilustrasi di Gambar 3.11, sebuah *neuron* yang diwakili oleh simbol  $h_j$  menerima sejumlah input  $x_1, x_2, \dots, x_n$ , dimana nilai-nilai ini berasal dari data atau output yang dihasilkan oleh lapisan sebelumnya. Perlu dicatat bahwa  $x_0$ , yang sering disebut sebagai bias, dianggap sebagai elemen non-input yang tetap bernilai 1. Variabel-variabel  $x_1, x_2, \dots, x_n$ , yang tergantung pada himpunan bilangan ril, mengacu pada bobot atau *weights* yang menghubungkan input dengan neuron tersebut. Bobot-bobot ini menentukan fungsi dari interaksi antara input dan output dalam sistem, dan merupakan parameter yang disesuaikan oleh mesin untuk mengurangi kesalahan antara nilai output yang dihasilkan dan pola nilai yang diharapkan selama proses pembelajaran. Kompleksitas sistem dalam konteks *deep learning* menciptakan potensi untuk adanya ratusan juta bobot yang harus disesuaikan.. Untuk mencapai penyesuaian yang optimal terhadap vektor bobot,

algoritma melibatkan perhitungan vektor *gradien* yang mencerminkan perubahan kesalahan yang timbul saat bobot-bobot tersebut disesuaikan. Prosedur ini bertujuan untuk mengurangi kesalahan secara bertahap dan mencapai tingkat akurasi yang diinginkan (LeCun dkk., 2015)

Operasi yang dilakukan pada sebuah perceptron terdiri dari dua bagian yang terpisah, yaitu operasi linier yang melibatkan nilai-nilai bobot yang diberikan, dan hasil dari operasi linier tersebut kemudian diolah menggunakan operasi non-linier atau aktivasi. Setelah operasi linier selesai, hasilnya akan melewati proses transformasi menggunakan fungsi aktivasi.

$$Z_j = \sum_{i=0}^d W_{ij} X_i \quad (2.1)$$

$$h_j = \sigma(Z_j) \quad (2.2)$$

Operasi aktivasi/non-linear  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ , terdapat beragam fungsi aktivasi yang dapat diterapkan pada neuron tersembunyi. Salah satu fungsi aktivasi yang digunakan dalam penelitian ini adalah softmax layer. Softmax layer umumnya digunakan dalam kasus klasifikasi multikelas, di mana lapisan output biasanya terdiri dari lebih dari satu neuron. Misalnya, jika  $y$  merupakan sebuah vektor dengan  $m$  elemen, softmax didefinisikan sebagai berikut:

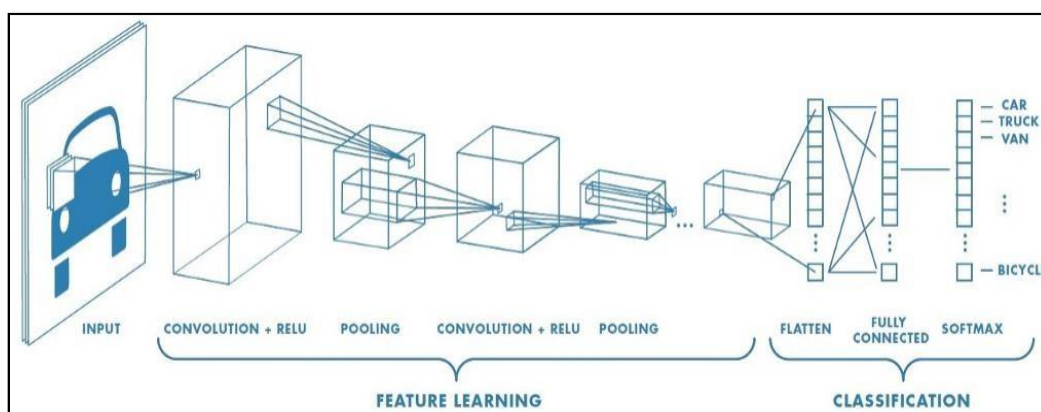
$$\sigma(a_j) = \frac{\exp(a_j)}{\sum_{k=1}^m \exp(a_k)} \quad (2.3)$$

Hasilnya  $\sum_{j=1}^m \sigma(a_j) = 1$  *softmax*.

### 2.1.6 Convolutional Neural Networks

*Convolutional Neural Networks* (CNN) adalah jenis arsitektur jaringan saraf tiruan yang dirancang khusus untuk pengolahan citra. CNN terdiri dari beberapa lapisan yang berbeda, termasuk lapisan konvolusi yang bertugas untuk mengekstraksi fitur dari citra input. Setiap lapisan konvolusi menggunakan filter atau kernel untuk memindai citra input dan mengekstraksi fitur-fitur penting. Kemudian, hasil konvolusi tersebut diolah oleh lapisan lain seperti lapisan aktivasi

(misalnya ReLU), pooling, dan lapisan penghubung (fully connected) untuk melakukan klasifikasi atau regresi. CNN telah menjadi sangat populer dalam berbagai aplikasi pengolahan citra, termasuk pengenalan objek, deteksi wajah, segmentasi gambar, dan banyak lagi, karena kemampuannya dalam menangkap struktur hierarkis dari data citra((Yamashita dkk., 2018).



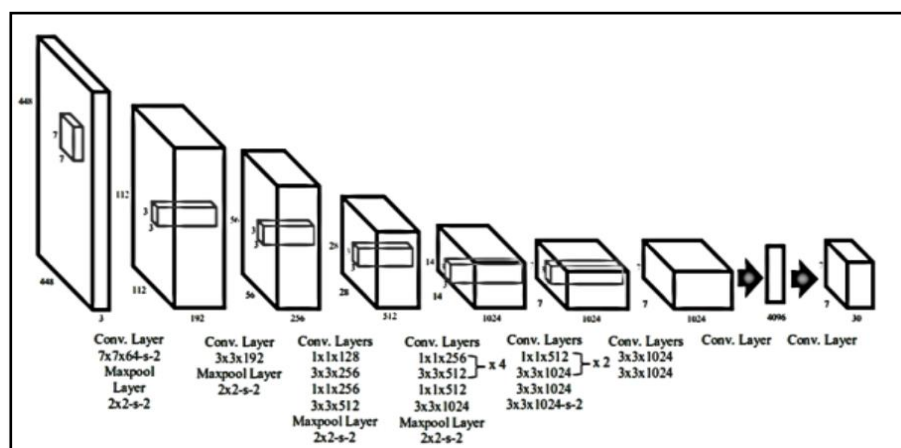
Gambar 2. 4 Arsitektur CNN(S. Patel dan Pingel, 2019)

Gambar 2. 4 merupakan arsitektur dari *Convolutional Neural Network* dengan input berupa gambar mobil, gambar mobil pertama kali dimasukkan sebagai input ke dalam jaringan. Proses dimulai dengan lapisan konvolusi yang bertugas untuk mengekstraksi fitur-fitur penting dari gambar tersebut, seperti tepi, sudut, atau tekstur yang khas dari mobil. Hasil konvolusi kemudian melalui fungsi aktivasi ReLU, yang membantu dalam memperkenalkan non-linearitas dan meningkatkan kemampuan jaringan untuk mempelajari pola-pola yang kompleks. Hasilnya kemudian melewati lapisan *pooling*, di mana dimensi spasialnya direduksi untuk mengurangi *overfitting* dan meningkatkan invariansi terhadap pergeseran kecil dalam citra.. Tahap konvolusi dan pooling ini dilakukan beberapa kali, berturut-turut, untuk mengekstraksi fitur-fitur yang semakin abstrak dan kompleks dari gambar mobil. Melalui serangkaian lapisan ini, hasilnya kemudian diubah menjadi vektor satu dimensi melalui lapisan *flatten*, sehingga dapat diproses oleh lapisan *fully connected* yang terdiri dari *neuron-neuron* yang terhubung sepenuhnya. Lapisan-lapisan ini bertugas untuk melakukan klasifikasi berdasarkan fitur-fitur yang terdapat dalam vektor tersebut. Akhirnya, output dari lapisan *fully connected*

diumpankan ke *softmax* layer, yang mengonversi output menjadi distribusi probabilitas atas kelas-kelas yang mungkin, dan akhirnya menentukan kelas prediksi dari gambar mobil tersebut. Dengan demikian, arsitektur CNN ini secara bertahap mengekstraksi fitur-fitur penting dari gambar mobil dan memprosesnya untuk klasifikasi.

### 2.1.7 You Look Only Once (YOLO)

YOLO (You Only Look Once) memiliki kemampuan untuk melakukan deteksi objek secara *real-time* karena menggunakan arsitektur CNN (Convolutional Neural Network). Algoritma ini dikembangkan oleh Joseph Redmon dan Ali Farhadi pada tahun 2015, terinspirasi oleh model *GoogLeNet* dalam klasifikasi citra (Redmon dkk., 2016).



Gambar 2. 5 Arsitektur YOLO (Redmon dkk., 2016)

Gambar 2.5 menampilkan arsitektur dari YOLO, yang secara keseluruhan sangat sederhana. Pada tahap awal, sistem menerima citra input berukuran 448 x 448 piksel dengan 3 saluran warna. Citra ini kemudian melewati satu lapisan jaringan konvolusi, menghasilkan output berukuran 7 x 7 x 30. Di sini, ukuran 7 x 7 mewakili grid sel, sementara 30 adalah jumlah kotak pembatas yang didasarkan pada jumlah kelas dan komponen dalam setiap kotak (Dwiyanto dkk., 2022). Setiap operasi konvolusi memiliki parameter seperti ukuran kernel *filter*, jumlah *filter*, *padding*, dan *stride*. Penggunaan *padding* membantu mengurangi kehilangan informasi di tepi citra, sedangkan *stride* menentukan pergeseran *kernel filter*.

Penerapan teknik *zero-padding* memungkinkan *kernel* filter beroperasi secara efektif di sepanjang tepi citra, menjaga konsistensi informasi di seluruh gambar (Dwiyanto dkk., 2022).

### **2.1.8 Computer Vision**

*Computer Vision* adalah studi tentang cara komputer memahami dan menganalisis gambar dan video untuk mendapatkan pemahaman tentang dunia di sekitarnya, mirip dengan kemampuan sistem penglihatan manusia (Szeliski, 2022). *Computer Vision*, merupakan disiplin ilmu yang memberikan kemampuan pada komputer atau mesin lain untuk melihat dan merespons terhadap objek-objek yang terdeteksi, memungkinkannya untuk menampilkan objek digital dan mengumpulkan data. Banyak bidang dalam *Computer Vision* berfokus pada pengembangan aplikasi praktis yang dapat diterapkan dalam berbagai konteks dunia nyata, seperti kontrol kualitas dalam manufaktur, pengenalan karakter optik, sistem bantuan pengemudi, pengawasan, fotografi, dan hiburan (Kae dkk., 2013).

### **2.1.9 Roboflow**

*Platform* yang digunakan untuk mengelola dan menganotasi dataset gambar untuk keperluan pembelajaran mesin dan pengembangan aplikasi berbasis citra. Pengguna dapat dengan mudah mengimpor, menganotasi, menyaring, dan mengelola dataset gambar menggunakan *Roboflow*. Fitur-fitur seperti alat penandaan otomatis, pengelompokan dataset, dan integrasi dengan berbagai platform pembelajaran mesin populer seperti *TensorFlow* dan *PyTorch* tersedia dalam platform ini. *Roboflow* mempermudah proses persiapan dataset, yang merupakan langkah penting dalam pengembangan model pembelajaran mesin berbasis citra.

### **2.1.10 Google Colab**

Google Colab adalah layanan *cloud* yang disediakan oleh *Google* untuk menjalankan dan mengelola *notebook Python* yang berbasis di *cloud*. Pengguna dapat menulis dan mengeksekusi kode *Python*, menyimpan dan membagikan *notebook*, serta melakukan analisis data secara kolaboratif menggunakan *Google Colab*. Layanan ini memberikan akses ke sumber daya komputasi yang kuat dan



terkelola sepenuhnya oleh *Google*, sehingga memungkinkan pengguna untuk melakukan tugas-tugas komputasi yang kompleks tanpa perlu khawatir tentang infrastruktur teknis.

### 2.1.11 *Python*

Bahasa *Python* adalah bahasa pemrograman tingkat tinggi yang serbaguna, mudah dipelajari, dan populer dalam pengembangan perangkat lunak. *Python* diciptakan oleh Guido van Rossum pada tahun 1991 dan dirilis sebagai perangkat lunak *open-source*. Beberapa fitur khas *Python* termasuk sintaksis yang bersih dan mudah dibaca, dukungan untuk paradigma pemrograman berorientasi objek dan fungsional, serta beragam modul dan pustaka yang luas untuk berbagai keperluan pengembangan.

### 2.1.12 *Confusion Matrix*

*Confusion matrix* merupakan alat evaluasi penting dalam analisis klasifikasi yang memberikan gambaran rinci tentang bagaimana model mengelompokkan data ke dalam berbagai kategori. Terdiri dari empat kategori, yaitu *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN). Beberapa metrik evaluasi umum yang terkait dengan *confusion matrix* meliputi akurasi, presisi, recall, dan skor F1:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (2.4)$$

$$Precision = \frac{TP}{TP + FP} \quad (2.5)$$

$$Recall = \frac{TP}{TP + FN} \quad (2.6)$$

$$F1 - Score = 2 \times \frac{precision \times recall}{precision + recall} \quad (2.7)$$

## 2.2 *State-of-the-Art*

Analisis olahraga, khususnya dalam konteks pertandingan sepak bola, telah mengalami perkembangan berkat kemajuan dalam *computer vision* dan *deep learning*. Penelitian dalam identifikasi elemen-elemen dalam pertandingan sepak bola telah dominan mengarah pada pengembangan sistem deteksi yang berbasis *Convolutional Neural Network* (CNN). Sistem-sistem ini bertujuan untuk secara

otomatis mengenali berbagai elemen seperti pemain, wasit, dan bola dalam rekaman pertandingan. Memanfaatkan kekuatan arsitektur CNN, sistem deteksi ini menunjukkan kemampuan yang menjanjikan dalam mengidentifikasi dan melacak elemen-elemen tersebut dengan akurat sepanjang pertandingan. Sistem deteksi ini memberikan kontribusi penting pada analisis olahraga dan evaluasi kinerja.

Penelitian ke-1(Komorowski dkk., 2019) dengan judul “*FootAndBall: Integrated player and ball detector*” dalam jurnal “*Electronic Letters on Computer Vision and Image Analysis 20(1):99-116, 2021*”, membahas inovasi dalam deteksi bola dan pemain dalam rekaman video sepak bola, dengan fokus pada peningkatan akurasi dan efisiensi dalam analisis otomatis permainan sepak bola. Penelitian ini mengidentifikasi tantangan dalam menemukan bola dalam rekaman video jarak jauh dan mengusulkan metode *FootAndBall*, yang didasarkan pada jaringan saraf efisien dengan desain piramida fitur untuk meningkatkan akurasi deteksi. Evaluasi menunjukkan bahwa *FootAndBall* mencapai tingkat akurasi yang tinggi dalam mendeteksi bola dan pemain, serta memiliki potensi untuk digunakan dalam otomatisasi analisis video sepak bola di klub dan akademi sepak bola. Parameter yang digunakan untuk menilai kinerja algoritma adalah objek, yaitu pemain dan bola. Nilai *ball AP*, diperoleh 0.999, kemudian *player AP* sebesar 0.921, dengan *mAP* mencapai 0.915.

Penelitian ke-2(Wang dan Li, 2022) dengan judul “*Deep Learning-Based Football Player Detection in Videos*” dalam jurnal “*Computational Intelligence and Neuroscience Volume 2022, Article ID 3540642, 8 pages*”, penelitian ini memperkenalkan algoritma analisis video sepak bola berbasis jaringan saraf konvolusi untuk mendeteksi pemain secara *real-time*. Metode ini menggunakan lima blok konvolusi untuk mengekstrak fitur pemain dengan resolusi berbeda dan menggabungkannya untuk meningkatkan akurasi deteksi. Hasil eksperimen menunjukkan efektivitas algoritma tersebut. Kesimpulannya, disebutkan bahwa algoritma ini dapat mendeteksi pemain secara langsung dari video pertandingan, dengan kecepatan proses tinggi dan kemampuan untuk beradaptasi dengan berbagai ukuran gambar. Parameter yang digunakan untuk menilai kinerja algoritma adalah objek, yaitu pemain dan bola. Algoritma yang dikenalkan mendapatkan nilai *mAP*

0.915 untuk dataset *ISSIA-CNR* dan untuk dataset *Soccer player detection* mendapatkan 0.932.

Penelitian ke-3(Naik dan Hashmi, 2021) dengan judul “*Ball and player detection & tracking in soccer videos using improved yolov3 model*” dalam jurnal “*Research Square (Research Square)*”, membahas penggunaan model YOLOv3 berbasis *deep learning* untuk deteksi bola dan pemain dalam video sepak bola siaran langsung. Pendekatan ini penting karena dapat membantu pelatih dalam mengevaluasi kinerja tim dan membuat keputusan yang dioptimalkan, serta membantu wasit dalam membuat keputusan pada momen-momen penting dalam pertandingan. Metode ini menggunakan algoritma SORT untuk pelacakan yang mencapai kinerja tinggi dan mengatasi tantangan seperti oklusi sebagian besar waktu. Kesimpulannya, penelitian ini menyediakan landasan untuk analisis pemain dan bola dalam video sepak bola dengan potensi pengembangan di masa depan untuk analisis pemain yang lebih mendalam dan pemetaan pola permainan.

Penelitian ke-4(Naik dkk., 2022) dengan judul “*DeepPlayer-Track: player and referee tracking with jersey color recognition in soccer*” dalam jurnal “*IEEE Access, 10, 32494–32509*”, mengusulkan *DeepPlayer-Track* untuk melacak pemain dan wasit dalam video sepak bola dengan akurasi tinggi. Menggunakan YOLOv4 untuk deteksi objek dan JCD-SORT untuk pelacakan pemain ganda, pendekatan ini berhasil mencapai akurasi pelacakan 96% dan 60% pada metrik MOTA dan GMOTA dengan kecepatan deteksi 23 *frame* per detik. Kontribusi utamanya adalah penggunaan YOLOv4 untuk deteksi objek dan pengenalan warna jersey untuk identifikasi pemain dan wasit, serta penerapan model pelacakan JCD-SORT untuk mencegah pergantian identitas pemain. Dengan demikian, metode ini berhasil mengatasi masalah pergantian identitas dalam pelacakan objek dalam video sepak bola secara efektif.

Penelitian ke-5(Husein dkk., 2023) dengan judul “*Detecting and Tracking Player in Football Videos Using Two-Stage Mask R-CNN Approach*” dalam “*IAIC International Conference Series (2023) 4(1) 132-138*”, membahas penggunaan model *Mask R-CNN* untuk segmentasi dan deteksi pemain sepak bola berdasarkan

dataset video pertandingan sepak bola dari Liga Inggris yang diperoleh dari *YouTube*. Dua pendekatan yang berbeda dalam model Mask R-CNN diterapkan dan diuji berdasarkan metrik evaluasi *Precision* (P), *Recall* (R), *Average Precision* (AP), dan *Mean Average Precision* (mAP). Meskipun kedua pendekatan menghasilkan hasil yang komparabel, pendekatan dengan menggunakan *DenseNet* menunjukkan kinerja yang sedikit lebih baik dibandingkan dengan menggunakan *ResNet-101*. Namun, pengetahuan yang belum memadai tentang karakteristik jenis gambar dan distribusi yang tidak merata dari berbagai jenis data yang diperoleh dari video acak menunjukkan masih ada ruang untuk peningkatan dalam model yang dilatih.

Penelitian ke-6(Hiemann dkk., 2021) dengan judul “*Enhancement of speed and accuracy trade-off for sports ball detection in videos—finding fast moving, small objects in real time*” dalam jurnal “*Sensors (2021) 21(9) 3214*”, penelitian ini mengembangkan pendekatan deteksi bola waktu nyata berbasis model YOLOv3 dengan penyesuaian arsitektur jaringan dan pelatihan untuk meningkatkan akurasi dan kecepatan deteksi. Integrasi informasi gerak dan pengetahuan domain khusus digunakan untuk meningkatkan kemampuan generalisasi jaringan. Hasilnya menunjukkan peningkatan keseimbangan antara kecepatan deteksi dan akurasi, menunjukkan potensi untuk penerapan dalam waktu nyata pada berbagai olahraga.

Penelitian ke-7(Buric dkk., 2019) dengan judul “*Player tracking in sports videos*” dalam “*2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom) (2019) 334-340*”, membahas pendekatan pelacakan pemain bola tangan menggunakan tiga metode pelacakan berbasis deteksi: algoritma *Hungarian* standar, algoritma pelacakan daring dan *real-time* (SORT), dan ekstensinya, *Deep SORT*. Metode-metode ini diuji pada dataset kustom video bola tangan. Berbagai tantangan dalam pelacakan pemain termasuk perubahan cahaya, gerakan cepat, dan penutupan pemain satu sama lain. Temuan menunjukkan bahwa penggunaan detektor objek berbasis CNN dan metode pelacakan berbasis deteksi dapat menjadi pilihan menarik untuk pelacakan pemain dalam video olahraga, meskipun masih ada masalah dengan pelacakan pemain yang

sementara keluar dari pandangan. Penelitian mendemonstrasikan potensi peningkatan pelacakan dengan menggunakan detektor objek yang lebih sukses dan menyoroti perlunya penelitian lebih lanjut untuk memperbaiki pelacakan bola dalam olahraga tim.

Penelitian ke-8(Kausalya, 2022) dengan judul “*Object Recognition in Soccer Sports Videos*” dalam jurnal “*2022 Smart Technologies, Communication and Robotics (STCR) (2022) 1-5*”, membahas penggunaan metode deteksi objek CNN dalam pengenalan objek pada video sepak bola. Metode tersebut dievaluasi dalam menangani tantangan seperti gerakan cepat, objek tersembunyi, perubahan iluminasi, dan perubahan latar belakang dalam situasi yang berbeda. Pendekatan yang diusulkan bertujuan untuk memberikan ringkasan pendekatan deteksi objek yang ada dan menguji kinerjanya pada video sepak bola, dengan fokus pada pengenalan objek kunci seperti pemain dan bola. Metode deteksi yang diuji termasuk R-CNN dan YOLO, dengan analisis kinerja yang disajikan dalam berbagai situasi, baik dalam pengenalan objek maupun pengenalan aksi.

Penelitian ke-9(Patel dan Kamdar, 2023) dengan judul “*OBJECT DETECTION IN HOCKEY SPORT VIDEO VIA PRETRAINED YOLOV3 BASED DEEP LEARNING MODEL.*” dalam jurnal “*ICTACT Journal on Image & Video Processing (2023) 13(3)*”, membahas penggunaan model deteksi objek berbasis *deep learning* YOLOv3 untuk mendeteksi objek dalam video olahraga hoki lapangan. Hasil eksperimen menunjukkan bahwa model YOLOv3 pre-trained menghasilkan hasil yang memuaskan dalam mendeteksi objek pada dataset yang disiapkan oleh peneliti sendiri, termasuk empat entitas utama dalam hoki, yaitu Tim 1 (AUS), Tim 2 (BEL), Bola Hoki, dan Wasit, dengan memodifikasi hiperparameter model *pre-trained* tersebut. Metode ini memberikan kontribusi signifikan dalam pengembangan analisis video olahraga, khususnya dalam mendukung pemantauan kinerja pemain dan evaluasi tim dalam situasi tertentu. Mencapai akurasi deteksi objek hoki sebesar 91.3%, penelitian ini memberikan dasar yang menjanjikan untuk pengembangan lebih lanjut dalam deteksi acara dari olahraga hoki lapangan.

Penelitian ke-10(Chen dan Poullis, 2023) dengan judul “*Tracking and Identification of Ice Hockey Players*” dalam “*International Conference on Computer Vision Systems (2023) 3-16*”, membahas kerangka kerja komprehensif untuk identifikasi dan pelacakan pemain dalam hoki es, menggunakan jaringan saraf dalam kedalaman yang dilatih pada data permainan aktual. Komponen utama termasuk deteksi, identifikasi, dan pelacakan pemain. Eksperimen menunjukkan kerangka kerja mencapai presisi rata-rata (AP) 67,3 dan akurasi pelacakan objek Berganda (MOTA) 80,2 untuk deteksi dan pelacakan pemain. Akurasi identifikasi tim dan nomor pemain adalah 82,39% dan 87,19% masing-masing. Secara keseluruhan, ini merupakan kemajuan signifikan dalam pelacakan dan identifikasi pemain, berpotensi digunakan dalam berbagai aplikasi olahraga.

Penelitian ke-11(Cioppa dkk., 2020) dengan judul “*Multimodal and multiview distillation for real-time player detection on a football field*” dalam “*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (2020) 880-881*”, mengusulkan sebuah sistem untuk memantau okupansi lapangan olahraga publik dengan biaya rendah, khususnya lapangan sepak bola. Sistem ini menggunakan kamera *fisheye wide-angle* murah yang dibantu oleh kamera termal *narrow-angle* tunggal untuk mendeteksi pemain. Melalui pendekatan distilasi pengetahuan, penulis melatih sebuah jaringan dengan modalitas dan sudut pandang yang berbeda antara guru dan siswa. Metode augmentasi data khusus dikombinasikan dengan algoritma deteksi gerakan untuk menangani pelatihan di wilayah kamera *fisheye* yang tidak tercakup oleh kamera termal. Evaluasi menunjukkan bahwa solusi ini efektif dalam mendeteksi pemain di seluruh lapangan sepak bola yang direkam oleh kamera *fisheye*, baik secara kuantitatif maupun kualitatif, terutama dalam distilasi online di mana siswa dapat mendeteksi pemain secara *real-time* dan terus beradaptasi dengan kondisi video terbaru.

Penelitian ke-12(Liu dkk., 2021) dengan judul “*Detecting and matching related objects with one proposal multiple predictions*” dalam “*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2021) 4520-*

4527”, mengusulkan sebuah metode yang efisien untuk mendeteksi dan mencocokkan kotak pemain serta objek terkait (contohnya, pemain yang memegang tongkat), yang sangat dibutuhkan dalam aplikasi di mana deteksi hanya terhadap pemain saja tidak mencukupi. Metode ini mengadopsi pendekatan asosiasi implisit untuk prediksi ganda melalui kotak proposal yang sama, sehingga mampu mendeteksi dan mencocokkan objek terkait tanpa memerlukan biaya tambahan. Hasil eksperimen menunjukkan bahwa pendekatan ini sangat efektif terutama untuk aplikasi yang memerlukan pencocokan yang akurat, terutama dalam situasi di mana kotak pemain dan objek terkait memiliki tumpang tindih yang signifikan.

Penelitian ke-13(Vicente-Martínez dkk., 2023) dengan judul “*Adaptation of YOLOv7 and YOLOv7\_tiny for soccer-ball multi-detection with DeepSORT for tracking by semi-supervised system*” dalam jurnal “*Sensors (2023) 23(21) 8693*”, membahas Membahas terobosan dalam teknologi deteksi dan pelacakan objek, khususnya dalam olahraga sepak bola, studi ini menggunakan jaringan *semi-supervised* dan arsitektur *Convolutional Neural Network* (CNN) YOLOv7, serta focal loss function, mencapai akurasi 95% dalam deteksi bola sepak. Dataset terperinci berisi 6331 gambar memperkuat validasi model dalam situasi nyata, sementara pemahaman tentang teknologi dalam evaluasi kinerja pemain diperluas melalui penggunaan sistem analisis video dan kamera. Metode penelitian yang cermat, dengan kombinasi pembelajaran *semi-supervised* dan *transfer learning*, serta penggunaan *DeepSORT* untuk pelacakan lintasan bola, menambah nilai signifikan pada penelitian ini. Studi ini menawarkan solusi efektif dalam mengatasi tantangan deteksi dan pelacakan bola dalam sepak bola. Rencana penelitian mendatang akan memperdalam analisis lintasan dan mempertimbangkan faktor-faktor seperti stabilitas kamera untuk meningkatkan akurasi deteksi.

Penelitian ke-14 (Mustamo, 2018) , dengan judul “*Object detection in sports: TensorFlow Object Detection API case study*” dalam ‘*University of Oulu, Degree Programme in Mathematical Sciences. Bachelor’s Thesis, 43 p*’ membahas penggunaan *TensorFlow Object Detection API* dalam mendeteksi pemain sepak bola, dengan fokus pada model SSD dengan *Mobilenet*. Penelitian ini bertujuan

untuk menguji hipotesis seputar penggunaan model pra-pelatihan tanpa penyetelan lanjutan, keberhasilan model setelah dilatih ulang dengan data sendiri, serta tantangan yang dihadapi dalam mendeteksi pemain yang berada dalam kerumunan atau berada di dekat dinding belakang lapangan menggunakan dataset yang berisi gambar-gambar yang diekstrak dari rekaman video pertandingan sepak bola.. Hasil penelitian menunjukkan bahwa model pra-pelatihan tidak efektif dalam mendeteksi pemain dalam gambar uji, namun model yang dilatih ulang memberikan hasil yang memuaskan. Studi ini juga menyarankan pengembangan model dengan arsitektur *meta* yang berbeda, *pre-processing* gambar yang lebih baik, serta pengujian yang lebih komprehensif menggunakan dataset yang lebih besar dan lebih bervariasi.

Penelitian ke-15(Pawar, 2023) dengan judul “*FOOTBALL DETECTION WITH YOLOV3*” dalam jurnal “*International Research Journal of Modernization in Engineering Technology and Science*”, membahas penggunaan model YOLOv3 dalam mendeteksi bola dan pemain sepak bola dalam video siaran langsung. Studi ini menyoroti tantangan dalam mendeteksi bola dan pelacakan pemain dalam analisis video sepak bola. Pentingnya teknologi visi komputer dan pembelajaran mendalam dalam meningkatkan pemahaman permainan juga ditekankan. Penelitian ini mengusulkan penggunaan model pra-pelatihan untuk deteksi pemain, sementara model dilatih sendiri untuk deteksi bola melalui metode eksperimen dan pengujian. Teknik pelacakan yang diusulkan menggunakan algoritma SORT, menggabungkan *filter Kalman* dan tumpang tindih kotak pembatas untuk menangani tantangan seperti oklusi. Studi ini memberikan kontribusi penting dalam mengembangkan teknologi analisis video sepak bola yang lebih canggih dan efektif.



### 2.3 Matriks Penelitian

Berdasarkan penelitian terkait yang telah dirangkum sebelumnya, dapat disimpulkan dalam bentuk matriks penelitian seperti yang ditunjukkan dalam tabel berikut.

Tabel 2. 1 Matriks Penelitian

No	Peneliti	Ruang Lingkup							
		Keterangan Label		Analisa Objek		Elemen yang dideteksi			
		Bahasa Indonesia	Bahasa Asing	Deteksi	Identifikasi	Pemain	Penjaga Gawang	Wasit	Bola
1	(Komorowski dkk., 2019)			✓		✓			✓
2	(Wang dan Li, 2022)			✓		✓			✓
3	(Naik dan Hashmi, 2021)			✓		✓			✓
4	(Naik dkk., 2022)		✓	✓	✓	✓		✓	✓
5	(Husein dkk., 2023)			✓		✓			

6	(Hiemann dkk., 2021)		✓	✓	✓	✓			✓
7	(Buric dkk., 2019)			✓					✓
8	(Kausalya, 2022)		✓	✓	✓	✓			✓
9	(Patel dan Kamdar, 2023)		✓	✓	✓	✓		✓	✓
10	(Chen dan Poullis, 2023)			✓		✓			
11	(Cioppa dkk., 2020)			✓		✓			
12	(Liu dkk., 2021)		✓	✓	✓	✓			
13	(Vicente-Martínez dkk., 2023)		✓	✓	✓				✓
14	(Mustamo, 2018)		✓	✓	✓	✓			
15	(Pawar, 2023)		✓	✓	✓	✓			✓

Tabel 2.1 menampilkan matriks penelitian yang telah dilakukan dalam fokus deteksi objek dalam konteks pertandingan olahraga. Dari 15 penelitian yang ditampilkan, hanya satu penelitian yang mencakup deteksi objek dengan keberagaman kelas, seperti pemain, bola, dan wasit yaitu penelitian yang dilakukan oleh (Patel dan Kamdar, 2023) dengan judul “OBJECT DETECTION IN HOCKEY SPORT VIDEO VIA PRETRAINED YOLOV3 BASED DEEP LEARNING MODEL” . Penelitian yang akan dilakukan akan memperbarui label identifikasi dengan menggunakan Bahasa Indonesia dan juga akan melakukan klasifikasi objek dengan keberagaman objek, termasuk pemain, wasit, dan bola.