BAB II TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Terdapat beberapa penelitian terdahulu yang terkait dengan implementasi *Computer Vision* dan *Machine Learning* YOLOV7 untuk untuk mendeteksi sebuah objek, dapat dilihat pada Tabel 2.1 sebagai berikut.

Tabei 2.1 Penelitian terdahulu

Judul	Metode	Analisis & Hasil
A Car Target	Mendeteksi target	Dari segi ukuran model
Detection	mobil berdasarkan	YOLOv7-tiny ini
Method based	algoritma YOLOv7-	memiliki ukuran yang
on YOLOv7-	tiny.	kecil 11,7M
tiny	/11/2/3-//	dibandingkan model
		YOLOv7 yang lain. Telah
133		diverifikasi bahwa
	M2	metode ini memiliki
1 = 8	人で意識を	akurasi mAP 80,8% dan
1		77,2% untuk mendeteksi
		jenis kendaraan lain.
A Comparative	Mempertimbangkan	Hasil percobaan YOLOv7
Study of	mana yang lebih	memiliki skor presisi
YOLOv5 and	baik dari YOLOv5	52,8%, nilai <i>recall</i>
YOLOv7	dan YOLOv7 secara	56,4%, <i>mAP</i> @0,5 51,5%
Object	independen dalam	dan mAP@0.5:0,95
Detection	hal <i>precision</i> dan	31,5% sedangkan
Algorithms	recall, mAP@0.5	YOLOv5 adalah skor
	dan mAP@0.5:0.95.	presisi 62,6%, nilai <i>recall</i>
		53,4%, mAP@0.5 sebesar
		55.3% dan
	A Car Target Detection Method based on YOLOv7- tiny A Comparative Study of YOLOv5 and YOLOv7 Object Detection	A Car Target Detection Method based on YOLOv7- tiny A Comparative Study of YOLOv5 and YOLOv7 dan YOLOv7 dan YOLOv7 on YOLOv7 on young a second of the process of the proce

			mAP@0.5:0.95 sebesar
			34.2%.
Yu et al.,	Foreign	penelitian ini	YOLOv7 memiliki skor
2023	Objects	membandingkan	mAP (Mean Average
	Identification	antara algoritma	Precision) tertinggi
	of	YOLOv7 yang	92,2% dan menunjukan
	Transmission	sudah ditingkatkan	skor 19 FPS (Frames Per
	Line Based on	oleh seri YOLO	Second) kedua setelah
	Improved	lainnya (Faster-	model Centernet. Dengan
	YOLOv7	rcnn, Centernet, dan	perbandingan YOLOv7
		model detection	yang tidak ditingkatkan
1/6		lainnya).	dan akurasi rata-rata
0-1	100		meningkat 11,9%, yang
50 N		(1/11111/////	merupakan peningkatan
		10/18/16/16/16/16/16/16/16/16/16/16/16/16/16/	akurasi yang paling jelas
	3		dibandingkan dengan
	10	がな	target deteksi lainnya

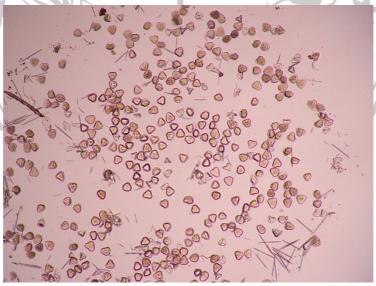
2.2 Kelapa Sawit

Kelapa Sawit dengan nama ilmiah (*Elaeis Guiniensis Jacq.*) yang berasal dari Afrika Barat dan diperkenalkan ke Indonesia pada abad ke-18 oleh kolonial Belanda. Tanaman ini awalnya ditanam di Jawa untuk keperluan hiasan, tetapi kemudian dikembangkan sebagai tanaman komersial pada awal abad ke-20. Pada saat itu, produksi minyak kelapa sawit di Indonesia masih berskala kecil. Kelapa sawit merupakan tumbuhan berkeping satu (*monokotil*) dengan bunga betina dan bunga jantan terdapat pada satu batang pohon yang sama, ketika proses perkawinan silang antara bunga betina dan jantan biasanya dibantu serangga atau agen polinasi dengan nama ilmiah (*Elaeidobius Kamerunicus*) yang nantinya akan menghasilkan buah kelapa sawit [1]. Kelapa sawit (*Elaeis*) bukanlah tanaman asli Indonesia, melainkan kelapa sawit ini tanaman asli Afrika Barat dengan iklim hutan hujan tropis seperti Kamerun dan Pantai Gading, Kelapa sawit ditanam secara komersil di Indonesia sejak tahun 1991 [2].

Kelapa sawit pertama kali masuk ke Indonesia dibawa oleh Dr. D. T. Pryce sebanyak 4 biji jenis *Dura* pada tahun 1848, yang ditanam dikebun Raya Bogor. Terdapat dua jenis kelapa sawit *Elaeis guiniensis* dan *Elaeis oleifera*, tipe kelapa sawit dibedakan berdasarkan ketebalan cangkang: *Dura, Pisifera* dan *Tenera* [2].

2.2.1 Serbuk Sari atau Pollen

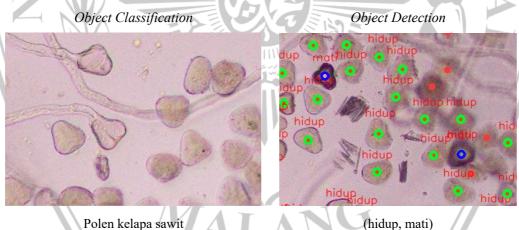
Serbuk Sari atau pollen adalah zat reproduksi yang dihasilkan dari bunga jantan kelapa sawit dengan tekstur serbuk yang kasar, serbuk sari mengandung sel kelamin jantan atau sel sperma bunga jantan dari tanaman kelapa sawit. Serbuk sari memiliki lapisan mantel keras yang dapat melindungi sel-sel sperma selama proses gerakan mereka antara benang sari ke putik tanaman berbunga atau dari kerucut jantan (cone) untuk betina pada organ reproduksi. Sel sperma ini akan bergerak menuju bunga betina kelapa sawit yang biasanya dibantu oleh angin atau serangga penyerbuk seperti Elaeidobius kamerunicus. Jika serbuk sari berhasil mencapai bunga betina, pembuahan akan terjadi, dan ini akan menghasilkan pembentukan biji kelapa sawit yang baru. Serbuk sari memiliki peran penting dalam proses reproduksi tanaman. Berikut adalah gambar 2.1 polen yang berhasil diambil dari mikroskop. Serbuk sari tanaman kelapa sawit mempunyai ukuran dengan rata-rata panjang pollen ialah 39,9 µm dan rata-rata lebarnya ialah 34,1 µm. Selama proses pembuahan terjadi, di mana inti serbuk sari menyatu dengan inti sel telur, akan membentuk embrio yang kemudian akan berkembang menjadi biji atau buah.



Gambar 2.1 Serbuk sari atau Pollen kelapa sawit

2.3 Object Detection

Object detection atau deteksi objek adalah salah satu metode computer vision yang menggunakan identifikasi dan pengenalan objek dalam gambar atau video. Deteksi objek adalah salah satu domain yang menyaksikan kesuksesan besar dalam computer vision [3]. Object detection digunakan tidak hanya untuk menentukan keberadaan objek dalam sebuah gambar dan video, melainkan untuk menemukan dan mengklasifikasikan objek ke dalam sebuah variable (kategori) yang sudah ditentukan. Ada beberapa faktor yang dapat mempengaruhi hasil deteksi, antara lain background objek, posisi objek, dan tinggi atau jarak objek pada gambar [4]. Proses ini biasanya menggunakan dua cara yaitu lokalisasi menetukan lokasi yang tepat dari objek pada gambar. Biasanya dilakukan dengan menggambar kotak pembatas (bounding box) di sekitar objek, yang akan menentukan posisi dan ukuran pada objek didalam gambar. Sedangkan klasifikasi, objek yang telah dilokalkan, selanjutnya akan di tetapkan ke kategori atau kelas tertentu. Ini melibatkan pengidentifikasian jenis atau kelas dari setiap objek yang terdeteksi, seperti orang, mobil, atau anjing, di antara serangkaian kategori yang telah ditentukan sebelumnya.



Gambar 2.2 Perbedaan klasifikasi dan deteksi objek

2.4 Artificial Intelligence

Artificial intelligence atau kecerdasan buatan yang didasarkan pada prinsip bahwa kecerdasan manusia dapat didefinisikan sedemikian rupa sehingga mesin dapat dengan mudah menirunya dan menjalankan tugas, dari yang paling sederhana hingga yang lebih kompleks. Tujuan dari kecerdasan buatan termasuk meniru

aktivitas kognitif manusia. Para peneliti dan pengembang di bidang ini membuat langkah yang sangat cepat dalam meniru aktivitas seperti pembelajaran, penalaran, dan persepsi, sejauh hal ini dapat didefinisikan secara konkret. Beberapa orang percaya bahwa para inovator akan segera dapat mengembangkan sistem yang melebihi kapasitas manusia untuk belajar atau menalar subjek apa pun. Namun, sebagian lainnya tetap skeptis karena semua aktivitas kognitif dicampur dengan penilaian nilai yang tunduk pada pengalaman manusia. AI akan terus berkembang dan memberikan manfaat bagi berbagai industry, lintas disiplin ilmu seperti matematika, ilmu komputer, linguistik, psikologi, dan banyak lagi.

2.4.1 Machine Learning

Machine Learning (ML) adalah cabang dari artificial intelligence di mana komputer, yang disebut mesin, belajar melakukan berbagai tugas secara otomatis [5]. Pada kehidupan sehari hari dapat kita temui ML dapat menyediakan alat pemantauan kualitas tanah bagi petani disawah, alat pemantauan ini dapat menampilkan hasil kualitas tanah dan rekomendasi yang sudah di modifikasi berdasarkan data eksperimental dan data pada lapangan. ML menggunakan fitur yang diekstrak dari input data eksperimental, data eksperimental ini digunakan untuk mengembangkan model yang mampu memprediksi hasil yang diinginkan yang diinginkan dari data baru. Machine learning menggabungkan permodelan matematika dan algoritma yang kompleks untuk melakukan tugas-tugas dengan belajar dari data yang ada. ML telah berhasil diterapkan di berbagai bidang yang membutuhkan klasifikasi, prediksi, dan rekomendasi [6]. Pembelajaran mesin dapat dibagi menjadi tiga bagian utama:

- 1. Supervised learning termasuk algoritma ML dengan data berlabel untuk melakukan tugas-tugas seperti klasifikasi atau regresi. Tujuan dari supervised learning adalah mengidentifikasi hubungan yang mendasari antara fitur input dan variabel output, sehingga dapat memprediksi respons target untuk input yang tidak terduga di masa depan fitur.
- 2. *Unsupervised learning* dilatih dengan sebuah dataset yang tidak berlabel, untuk mengidentifikasi pola dan struktur dalam data yang diberikan. Algoritma ini memungkinkan pengguna untuk melakukan tugas

- pemrosesan yang lebih kompleks dibandingkan dengan *supervised learning*. Algoritma ini terutama digunakan untuk tugas-tugas seperti pengelompokan, deteksi anomali, jaringan saraf, asosiasi fitur, dan lain-lain.
- 3. Reinforcement leraning adalah sebuah paradigma pembelajaran mesin yang melibatkan proses pengambilan keputusan berurutan untuk mencapai tujuan akhir. Reinforcement learning berbasis umpan balik, yang berarti sistem belajar berperilaku sesuai dengan hasil tindakan yang dilakukan. Setiap tindakan yang dilakukan sistem akan mendapatkan feedback, baik berupa feedback baik dan kurang baik.

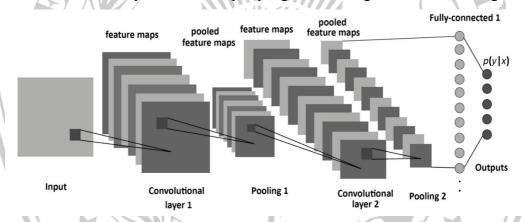
2.4.1 Deep Learning

Deep learning (DL) adalah cabang dari Machine Learning yang didasarkan pada arsitektur Artificial Neural Network. Artificial Neural Network atau ANN menggunakan lapisan-lapisan node yang saling terhubung yang biasanya disebut neuron yang bekerja bersama untuk memproses dan belajar dari data input yang ada. Dalam fully connected deep neural network, ada beberapa lapisan input dan satu atau lebih lapisan tersembunyi yang terhubung satu demi satu. Setiap neuron akan menerima input dari neuron lapisan sebelumnya. Output dari satu neuron menjadi input untuk neuron lain di lapisan jaringan berikutnya, dan proses ini akan terus berlanjut hingga lapisan terakhir yang menghasilkan output jaringan. Lapisan-lapisan jaringan saraf mengubah data input melalui serangkaian transformasi nonlinier, yang memungkinkan jaringan mempelajari representasi kompleks dari data input.

2.5 Convolutional Neural Network

Convolution Neural Network (CNN) merupakan salah satu jenis Neural Network (NN) yang biasa digunakan dalam computer vision. Computer vision sendiri merupakan salah satu bidang kecerdasan buatan yang memungkinkan komputer memahami dan menafsirkan gambar. Convolution Neural Network sangat berguna untuk menemukan pola dalam gambar untuk mengenali objek. CNN sangat efektif untuk mengklasifikasikan data non-gambar seperti audio, deret waktu, dan sinyal. Biasanya CNN digunakan untuk mendeteksi dan mengenali

object seperti pada Gambar 2.2 perbedaan antara klasifikasi dan deteksi objek. Secara garis besar CNN tidak jauh berbeda dengan Neural Network seperti biasanya. CNN terdiri dari neuron yang memiliki bobot, bias dan actifation function. Keunggulan dari jaringan saraf convolutional ini, bahwa mereka tidak bergantung pada ekstraktor atau fitur filter yang dibuat secara manual. Sebaliknya, mereka akan melatih (train per second) dari tingkat piksel pertama sampai kategori objek terakhir [3]. Model CNN memiliki struktur standar yang terdiri dari convolution layer dan polling layer (sering kali setiap polling layer ditempatkan dengan convolution layer). Lapisan terakhir adalah terdiri dari sejumlah kecil fully connected layer, dan lapisan terakhir adalah mengklasifikasi softmax seperti yang ditunjukkan pada Gambar 2.3. CNN biasanya dilatih dengan backpropagation yang melewati Stochastic Gradient Decent (SGD) untuk menemukan bobot (weight) dan bias (biases) pada saat training, agar meminimalisir loss tertentu pada hasil training untuk memetakan input acak ke output yang sudah ditargetkan sedekat mungkin.

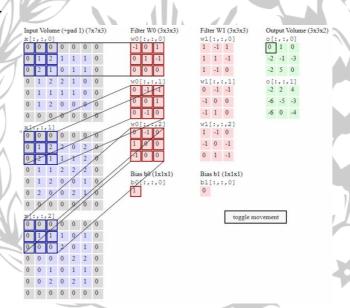


Gambar 2.3 Arsitekstur Convolution Neural Network (CNN)[7]

Arsitektur Convolution Neural Network yang lengkap juga dikenal sebagai Covnets. Covnets adalah rangkaian lapisan, dan setiap lapisan mengubah satu volume ke volume lainnya melalui fungsi terdiferensiasi. CNN biasanya mempunyai beberapa lapisan layer antara lain convolutional layer, polling layer, dan fully connected layer dll:

1. Convolution layer adalah lapisan konvolusi yang terdiri dari beberapa kernel atau filter (convolution mask), bertujuan untuk mengekstrak fitur-fitur lokal dari gambar yang akan di input. Input gambar merupakan array dua dimensi, kernel ialah parameter berupa array multidimensi dengan disesuaikan model algoritma. Lapisan ini melakukan "dot product" antara dua matriks, di mana

satu *matriks* adalah parameter yang dapat dipelajari atau dikenal sebagai *kernel. Kernel* secara spasial lebih kecil daripada citra tetapi lebih mendalam. Artinya, jika gambar terdiri dari tiga channel (RGB), tinggi dan lebar *kernel* secara spasial akan kecil, tetapi kedalamannya meluas hingga ketiga *channel* (RGB). *Convolutional layer* terdiri dari *neuron* yang tersusun sehingga membentuk sebuah filter dengan panjang dan tinggi (*pixels*). Sebagai contoh, *layer* pertama pada *feature extraction layer* biasanya adalah *conv. Layer* dengan ukuran 5x5x3. Panjang 5 *pixel*, tinggi 5 *pixel* dan tebal/ jumlah 3 buah, sesuai dengan *chanel* dari citra tersebut. Ketiga filter akan digeser keseluruh bagian permukaan pada citra. Setiap pergeseran akan dilakukan operasi (*dot*) antara *input* dan nilai dari filter, sehingga akan menghasilkan sebuah *output activation map* atau *feature map*. Proses *feature map* dapat dilihat seperti gambar 2.7.



Gambar 2.4 Feature Map

Pada proses *feture map* pada gambar 2.4 proses pergeseran *stride* setiap *pixel*, *feature map* tersebut menggunakan 2 *stride* yang akan melompat 2 kotak setiap bergeaser, semakin kecil nilai *stride* maka akan menghasilkan nilai yang bagus. kernel bergeser melintasi tinggi dan lebar gambar-menghasilkan representasi gambar dari wilayah reseptif. Dari *feature map* diatas untuk mengitung dimensi dapat dirumuskan sebagai berikut (1):

$$V = \frac{W - F + 2P}{S} + 1 \tag{2.1}$$

V= Volume

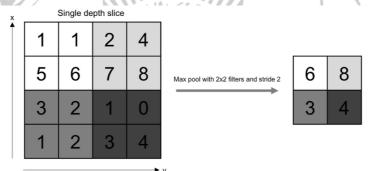
P= Zero Padding

W= Panjang (tinggi input)

S = Stride

F= Pnajang (tinggi filter)

2. Polling layer ini bertanggung jawab untuk mengurangi ukuran spasial dari Convolved Feature. Hal ini untuk mengurangi beban komputasi untuk memproses data dengan pengurangan dimensi. Selain itu, lapisan ini berguna untuk mengekstrak fitur dominan yang bersifat rotasi dan posisi invarian, sehingga menjaga proses training model secara efektif. Ada dua jenis pooling: Max Pooling dan Average Pooling. Max Pooling berguna untuk mengembalikan nilai maksimum dari bagian gambar yang dicakup oleh Kernel. Di sisi lain, Average Pooling mengembalikan rata-rata semua nilai dari bagian gambar yang dicakup oleh kernel.



Gambar 2.4 Polling layer

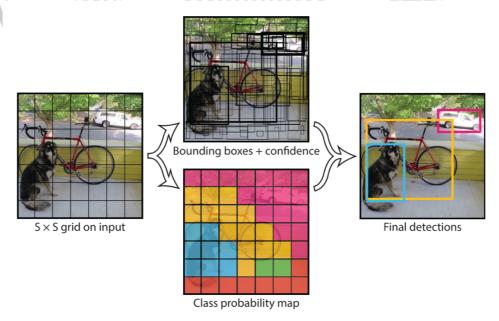
3. Fully Connected terdiri dari weight dan biases bersama dengan neuron dan digunakan untuk menghubungkan neuron di antara dua lapisan yang berbeda. Lapisan ini biasanya ditempatkan sebelum lapisan output dan membentuk beberapa lapisan terakhir dari Arsitektur CNN. Neuron pada lapisan ini memiliki konektivitas penuh dengan semua neuron pada lapisan sebelumnya dan lapisan berikutnya seperti yang terlihat pada FCNN biasa. Inilah sebabnya mengapa hal ini dapat dihitung seperti biasa dengan perkalian matriks yang diikuti dengan efek bias. Lapisan fully connected sangat membantu memetakan representasi antara input dan output.

CNN akan kurang akurat ketika diberikan dataset yang kecil tetapi CNN dapat menunjukkan akurasi yang signifikan/pemecahan rekor ketika diberi kumpulan

dataset gambar yang besar seperti Gambar 2.3 Arsitekstur *Convolution Neural Network* (CNN) [3].

2.6 You Only Look Once (YOLO) & YOLOv7

YOLO (You Only Look Once) merupakan algoritma citra digital (object detection) yang berjalan menggunakan teknik deep learning dan algoritma ini bekerja berdasarkan prinsip single shot, maksudnya jaringan arsitektur ini diatur sedemikian rupa agar dalam satu lintasan bingkai dapat mendeteksi beberapa objek sekaligus dan secara serentak. Maka dari itu YOLO akan menggunakan gambar atau video sebagai input yang nantinya akan menghasilkan prediksi deteksi gambar secara langsung, tanpa perlu melewati langkah proses deteksi secara terpisah. Menurut [8] YOLO adalah jaringan saraf convolutional tunggal yang membagi gambar input menjadi satu set sel grid, jadi tidak seperti klasifikasi gambar atau deteksi wajah, setiap sel grid dalam algoritma YOLO akan memiliki vektor terkait dalam output yang memberi tahu kita jika ada objek di dalamnya. Model YOLO ini sangat presisi yang memungkinkan model dapat mendeteksi objek di dalam bingkai gambar. YOLO menerapkan jaringan saraf ke seluruh gambar untuk memprediksi kotak pembatas dan probabilitasnya seperti gambar 2.6 Architecture object detection. Model ini sangat canggih sehingga YOLO akan belajar lebih banyak dari waktu ke waktu dan meningkatkan akurasi prediksinya dari waktu ke waktu.



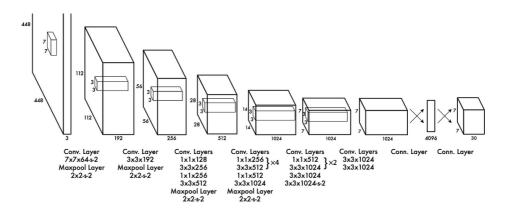
Gambar 2.5 Ilustrasi YOLO (You Only Look Once) [9]

Algoritma YOLO melakukan prediksi objek berupa *bounding box* B dengan *confidence score* (nilai konfidensi) C dan membagi input gambar atau video menjadi S x S *grid. Confidence score* mencerminkan seberapa yakin model tersebut bahwa kotak berisi objek dan dapat melihat seberapa akurat model memperkirakan kotak yang diprediksi. Nilai konfidensi dapat didefinisikan sebagai persamaan (2.2) berikut:

$$Confidance = Pr(Object) \times IoU \tag{2.2}$$

Pr() adalah nilai dari objektifitas dalam cell, Object berarti ketika terdapat objek dalam cell maka nilai 0 dan bernilai 1 terdapat sebuah objek dalam cell, IoU merupakan nilai rasio antara kotak ground truth dan kotak prediksi. Jika tidak ada objek di dalam cell, nilai konfidensi akan menjadi nol. Jika tidak, nilai konfidensi sama dengan intersection over union (IoU) antara kotak yang diprediksi dan kebenaran dasar. Grid cell yang kecil bertanggung jawab melakukan prediksi jika titik tengah suatu objek jatuh pada grid. Bounding box mengandung 5 elemen prediksi antaralain x, y, w, h dan confidence score. Koordinat x dan y mempresentasikan titik tengah dari box relatif pada grid cell. Elemen w (lebar) dan h (tinggi) dianggap relatif terhadap keseluruhan gambar [9].

Berdasarkan Gambar 2.5 System akan membagi gambar dengan dimensi grid $S \times S$ dan setiap sel grid akan memprediksi B bounding boxes, nilai confidence pada setiap kotak, dan class probability map untuk kelas C. Prediksi ini di rumuskan sebagai tensor menjadi $S \times S \times (B * 5 + C)$. Secara umum, arsitektur YOLO akan digambarkan secara berlapis lapis atau biasa disebut *convolution layer*.



Gambar 2.6 Architecture YOLO

Arsitektur jaringan YOLO terinspirasi oleh model GoogLeNet untuk klasifikasi gambar. YOLO memiliki 24 lapisan konvolusi yang diikuti dengan 2 *fully connected layers*. Pada gambar 2.6 Jaringan deteksi YOLO memiliki 24 lapisan konvolusi diikuti 2 *fully connected layer*. Lapisan konvolusional 1 × 1 akan mengurangi ruang pada fitur dari lapisan sebelumnya. Lapisan *convolution* di *training* dengan kalsifikasi ImageNet dengan setengah resolusi (gambar input 224 × 224) kemudian menggandakan resolusi untuk deteksi 448 × 448.[9].

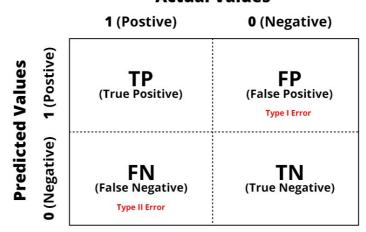
YOLOv7 merupakan *detektor objek real-time* mutakhir yang melampaui semua detektor objek yang dikenal dalam hal kecepatan dan akurasi dalam kisaran 5 FPS hingga 160 FPS. YOLOv7 memiliki akurasi tertinggi (56,8% AP) di antara semua pendeteksi *objek real-time* yang dikenal dengan 30 FPS atau lebih tinggi pada GPU V100. Selain itu, YOLOv7 mengungguli detektor objek lain seperti YOLOR, YOLOX, Scaled-YOLOv4, YOLOv5, dan banyak lagi dalam hal kecepatan dan akurasi [10]. Arsitektur model YOLOv7 masih berdasarkan pada model YOLOv4, YOLO-R, dan Scaled YOLOv4. YOLOv7 merupakan hasil percobaan untuk meningkatkan hasil deteksi yang bagus, dan membuat model lebih akurat dalam mendeteksi gambar. Model YOLOv7 tidak hanya mencoba untuk mengoptimalkan arsitektur, akan tetapi juga bertujuan untuk mengoptimalkan *training model*, agar saat menggunakan modul untuk meningkatkan akurasi saat mendeteksi objek.

Untuk mengetahui baik atau tidaknya model hasil *train* dan *validasi* dari YOLO v7 dapat dilihat dari parameter scoring model pembelajaran YOLO. Berikut adaalah parameter yang pada umumnya digunakan untuk penilaian hasil model pembelajaran YOLO.

1. Confusion Matrix

Confusion matrix adalah matriks yang digunaka untuk meringkas kinerja model pembelajaran mesin pada sekumpulan data uji. Pada dasarnya confusion matrix memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh sistem (model) dengan hasil klasifikasi sebenarnya. Confusion matrix berbentuk tabel matriks yang menggambarkan kinerja model klasifikasi pada serangkaian data uji yang nilai sebenarnya diketahui.

Actual Values



Gambar 2.7 Confusion matrix

- *True Positive* (TP): Model memprediksi label dan mencocokkan dengan benar sesuai dengan kebenaran dasar.
- *True Negative* (TN): Model tidak memprediksi label dan bukan bagian dari kebenaran dasar.
- False Positives (FP): Model memprediksi sebuah label, tetapi bukan bagian dari kebenaran dasar (Type I Error).
- False Negatives (FN): Model tidak memprediksi sebuah label, tetapi label tersebut merupakan bagian dari ground truth. (Type II Error).

2. Accuracy

Akurasi merupakan salah satu metrics yang paling sering diketahui pada pemodelan klasifikasi. Jadi akurasi adalah persentase model memprediksi benar dibanding dengan seluruh prediksinya. Akurasi dapat dihitung menggunakan persamaan (2.3) berikut:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{2.3}$$

3. Precision

Precision merupakan perhitungan ketepatan klasifikasi pada jumlah data berlabel positif atau data kelas minoritas yang memang benar secara actual merupakan kelas positif. Presisi dapat dihitung menggunakan persamaan (2.4) berikut.

$$Precision = \frac{TP}{TP + FP} \tag{2.4}$$

4. Recall

Recall (disebut juga sensitivity) adalah perhitungan ketepatan klasifikasi pada jumlah data positif yang teridentifikasi benar sebagai kelas positif. F-measure sendiri adalah relative importance antar precision dan recall. Recall dapat dihitung menggunakan persamaan (2.5) berikut.

$$Recall = \frac{TP}{TP + FN} \tag{2.5}$$

5. F1 Score

Skor F1 adalah *Harmonic Mean* antara presisi dan *recall*. Kisaran Skor F1 adalah [0, 1]. *F1-Score* digunakan untuk menghitung kombinasi presisi dan recall. F1-Score akan menggunakan harmonic mean dari presisi dan *recall*. *F1-score* dapat dihitung menggunakan persamaan (2.6) berikut:

$$F1 \, Score = 2 \times \frac{Precision. Recall}{Precision + Recall} \tag{2.6}$$

6. mAP

Mean Average Precision (mAP) adalah metrik yang digunakan untuk mengevaluasi model deteksi objek, seperti Fast R-CNN, YOLO, Mask R-CNN, dll. Nilai Average Precision (AP) dihitung berdasarkan nilai recall dari 0 hingga 1. mAP dihitung dengan mencari AP untuk setiap kelas dan kemudian dirata-ratakan pada sejumlah kelas. AP dapat di formulakan seperti pada persamaan (2.7) berikut.

$$mAP = \frac{1}{N} \sum_{i=1}^{n} AP_i \tag{2.7}$$

Di mana AP_i adalah AP untuk kelas ke-i dan N adalah jumlah total kelas yang dievaluasi. mAP menggabungkan trade-off antara nilai presisi dan recall serta mempertimbangkan hasil prediksi false positive (FP) dan false negative (FN).

2.7 Hyperparameter

Hyperparameter digunakan untuk mengintrol selama proses training dan menetukan nilai parameter model yang diinginkan. Salah satu yang mempengaruhi kinerja pada model untuk mendapatkan model yang terbaik adalah hyperparater, Hyperparamter tidak bisa digunakan selama proses training berlangsung, akan

tetapi digunakan sebelum proses taraing secara manual pada program. Berikut adalah beberapa *hyperparameter* yang umum di YOLOV7:

1. Optimizer

Optimizer merupakan metode yang digunakan dalam training jaringan syaraf. Fungsi dari optimizer untuk memperoleh nilai weight (bobot) yang memberikan nilai terbaik. Penelitian yang dilakukan dalam proses pembuatan sistem arsitektur YOLOV7 menggunakan Stochastic Gradient Descent (SGD) yang merupakan teknik pembelajaran dalam suatu sistem yang melakukan update untuk setiap satu data. Stochastic Gradient Descent mempunyai konsep mirip dengan Batch, dengan cara membagi dataset ke dalam beberapa Batch. Sehingga SGD akan terus update nilai weight tanpa menunggu menyelesaikan 1 epoch saat proses training.

2. Learning Rate

Learning rate merupakan salah satu hyperparamter penting yang digunakan untuk mengontrol seberapa cepat model dapat disesuaikan. Ketika proses training model jaringan syaraf, diperlukan beberapa teknik optimasi berdasarkan algoritma gradient descent. Setelah menghitung gradien fungsi kerugian sehubungan dengan bobot, gradien tersebut memiliki arah ke titik optimum lokal. digunakan learning rate untuk menyetel bobot ke arah tersebut dan mengoptimalkan model.

3. Epoch

Epoch merupakan jumlah iterasi selama proses training, yang digunakan untuk mengolah semua batch data untuk masuk kedalam optimizer dan keluar dari hyperparameter optimizer

4. Batch Size

Batch size adalah salah satu hyperparameter yang paling penting dalam pelatihan, batch size juga mewakili jumlah sampel yang digunakan dalam satu lintasan maju dan mundur melalui jaringan dan memiliki dampak langsung pada akurasi dan efisiensi komputasi proses pelatihan. Ukuran batch dapat dipahami sebagai pertukaran antara akurasi dan kecepatan.

5. Weight

Weight (bobot) merupakan koefisien dari fungsi yang akan diselesaikan. Weight merupakan nilai nyata yang dikelompokan dengan fitur yang menunjukan pentingnya fitur dalam memprediksi nila akhir. Weight ini akan terus mengalami optimasi salama proses training dan akan menghasilkan nilai paling optimum.

6. Iteration

Iterasi adalah jumlah *batch* yang diperlukan untuk menyelesaikan satu *epoch* yang digunakan untuk mengukur selama proses pelatihan. Jumlah iterasi sama dengan jumlah *batch* dalam satu *epoch*, dan dihitung dengan membagi jumlah total sampel dalam kumpulan data pelatihan dengan ukuran *batch*.

2.8 Streamlit

Streamlit adalah sebuah framework open source dari library Python untuk memudahkan saat membuat aplikasi web machine learning dan data science dengan cepat dan membuat visualisasi yang menawan, selain bersifat open source, streamlit juga open sharing. Library Python ini dibangun untuk memudahkan pengguna untuk mengubah data script menjadi aplikasi berbasis web yang interaktif. Streamlit menawarkan visualisasi melalui antarmuka yang interaktif dan dapat mengkomunikasikan informasi yang dapat dengan mudah dipahami oleh user. Proses deploy pada streamlit terbilang cukup mudah, hanya dengan melalui platform cloud-sharing.

MALA