

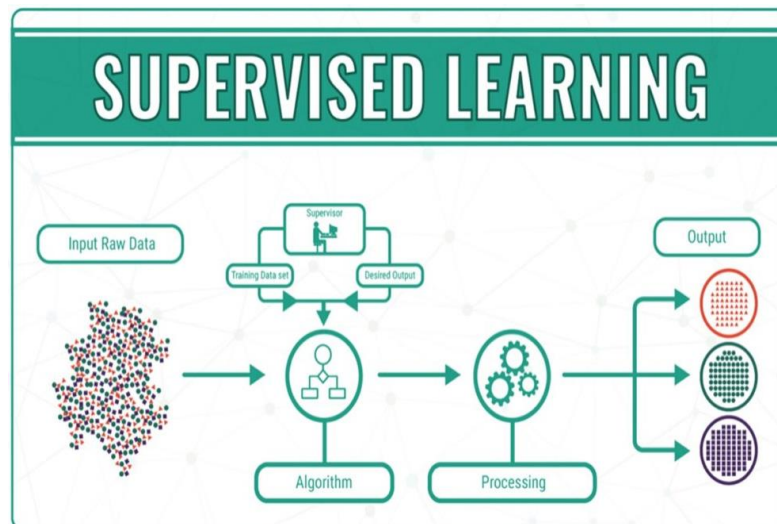
BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 *Machine Learning*

Machine learning atau yang kita kenal dalam bahasa Indonesia dengan “Pembelajaran Mesin” merupakan salah satu bidang keilmuan yang memungkinkan sistem komputer untuk belajar dari data dan pengalaman-pengalamannya untuk meningkatkan kinerja seiring waktu tanpa adanya pemrograman eksplisit. Dalam kehidupan sehari-hari kita sering menjumpai bidang keilmuan ini, contohnya prediksi kata pada *keyboard smartphone* dan juga pengenalan wajah ketika membuka *smartphone* kita (Dinata & Hasdyna, 2020). Secara umum, *machine learning* dibagi menjadi 3 kategori yakni *Supervised Learning*, *Unsupervised Learning*, dan *Reinforcement Learning* (Somvanshi & Chavan, 2016)

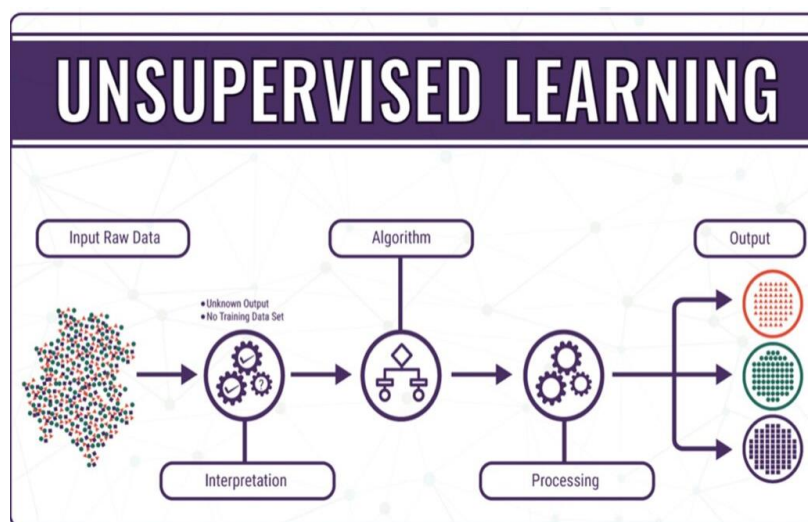


Gambar 2.1 *Supervised Learning*

Sumber : LearnEveryone

Supervised learning merupakan metode pembelajaran dimana sistem diberikan contoh input yang telah diberi label dan diharuskan

memetakan input tersebut ke *output* yang sesuai. Seperti yang dapat dilihat pada gambar 2.1, terdapat *supervisor* yang memberikan *training* pada *dataset* dan menginisiasikan *output* yang diinginkan. *Supervised learning* menganalisis data pelatihan dan menghasilkan fungsi disimpulkan yang dapat digunakan untuk memetakan contoh baru. Tujuan optimalnya adalah agar algoritma dapat dengan benar menentukan label kelas untuk contoh yang belum pernah dilihat sebelumnya. *Supervised learning* juga terbagi menjadi 2 kategori yakni regresi dan klasifikasi. Regresi merujuk pada variabel *output* yang berupa rentang nilai nyata atau kontinu, sedangkan klasifikasi merujuk pada variabel *output* berupa kategori atau *labelling*. (Brownlee, 2016)

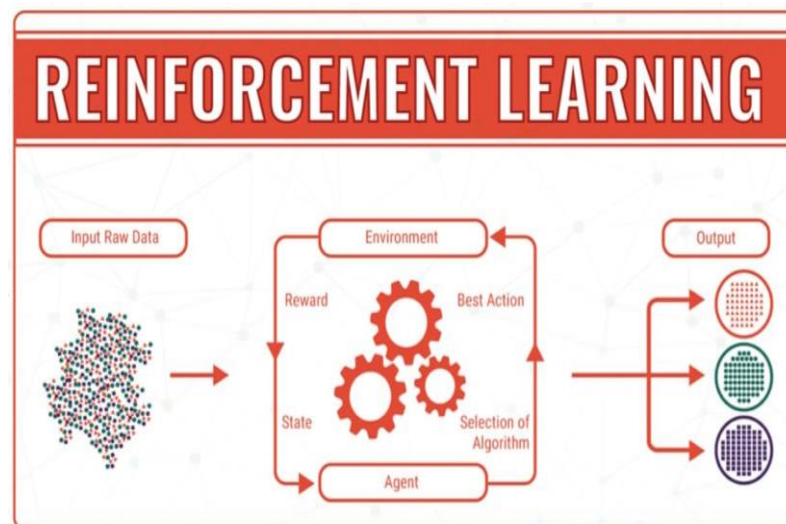


Gambar 2.2 Unsupervised Learning

Sumber :LearnEveryone

Unsupervised learning sering disebut sebagai pendekatan “cluster” karena tidak ada kebutuhan untuk memberi label pada *dataset*, dan hasilnya tidak mengidentifikasi contoh dalam kelas-kelas yang telah ditentukan (Thupae et al., 2018). Sehingga model atau algoritma mencoba untuk memahami struktur atau pola dalam data tanpa memiliki label *output* yang sesuai. Dengan kata lain, dalam *unsupervised learning*, tidak ada “guru” yang memberikan petunjuk atau label pada data pelatihan (Brownlee, 2016). Dari gambar 2.2, input raw data atau data mentah yang diinputkan kemudian ditafsirkan

menggunakan algoritma untuk yang akan menghasilkan *output* baru menurut interpretasi sistem. Tujuannya adalah untuk mengeksplorasi dan mengidentifikasi pola tersembunyi atau struktur yang mungkin tidak diketahui sebelumnya. Seperti *supervised learning*, metode pembelajaran ini juga dikelompokkan menjadi 2 kategori yakni *clustering* dan *association*. *Clustering* (pengelompokan) merupakan metode yang memisahkan data ke dalam kelompok atau kluster berdasarkan kesamaan karakteristik, misalnya mengelompokkan pelanggan berdasarkan perilaku pembelian. Sedangkan *association* atau asosiasi merupakan metode yang mencari hubungan atau korelasi antara variabel dalam data, misalnya orang yang membeli produk “A” juga cenderung membeli produk “B” (Brownlee, 2016).



Gambar 2.3 Reinforcement Learning

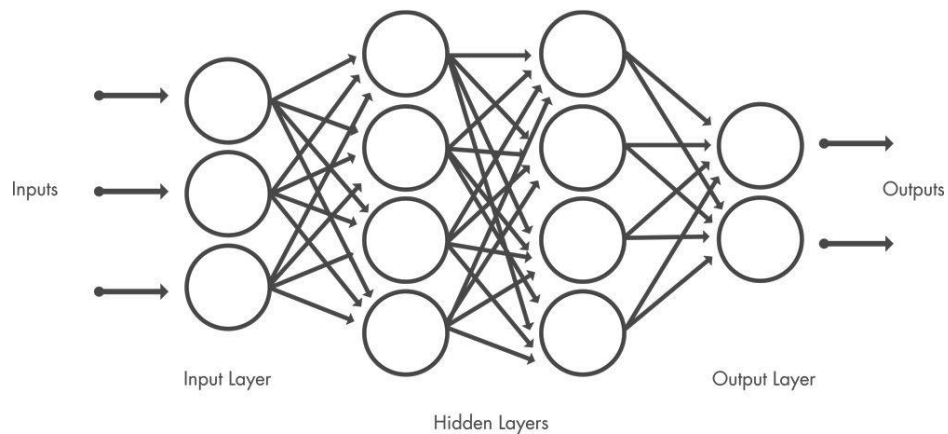
Sumber : LearnEveryone

Reinforcement learning (RL) merupakan jenis teknologi pembelajaran online yang berbeda dari *supervised learning* dan *unsupervised learning*. Dalam *reinforcement learning*, agen cerdas (robot, komputer cerdas, ataupun agen virtual lainnya) belajar dengan berinteraksi dengan lingkungan dan menerima sinyal penguatan sebagai umpan balik. Dapat dilihat dari gambar 2.3, agent akan mengamati kondisi lingkungan saat ini (*state*) untuk menentukan

algoritma yang terbaik pada kondisi saat ini, kemudian lingkungan akan memberikan umpan balik (*reward*) dari tindakan yang diambil sebelumnya, proses ini akan diulang berkali-kali untuk memaksimalkan total *reward* yang diperoleh dari lingkungan. Pada dasarnya pembelajaran ini menggunakan teori konsep cara belajar hewan. Dalam konteks ini, RL mencoba mereplikasi cara hewan dan manusia belajar dari pengalaman melalui percobaan dan kesalahan. Sehingga RL tidak memerlukan pengetahuan awal atau panduan sebelumnya, pengetahuan yang diperoleh mesin berasal dari pengalaman yang dilakukan secara coba-coba (*trial and error*) dengan terus berinteraksi pada lingkungan yang dinamis (Qiang Wang & Zhongli Zhan, 2011).

2.1.2 Convolutional Neural Network

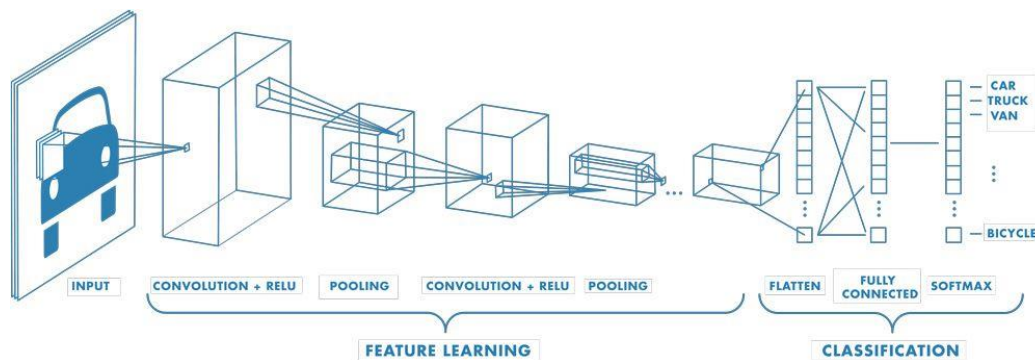
Convolutional neural network (CNN) merupakan jenis jaringan saraf tiruan yang umumnya digunakan untuk memahami, mengolah, memproses data bidang dua dimensi seperti gambar (Browne & Ghidary, 2003). CNN terbukti sangat baik dalam mengeksekusi masalah yang memiliki input data berupa susunan grid yang akan diprediksi, seperti deret waktu (grid 1-D) atau gambar (grid 2-D) (Ketkar & Moolayil, 2021). Pada dasarnya, *Convolutional Neural Network* (CNN) adalah pengembangan dari *Multilayer Perceptron* (MLP) yang dirancang khusus untuk mengolah data dua dimensi. Sehingga cara kerja CNN juga memiliki kesamaan dengan MLP, tetapi perbedaannya terletak pada representasi setiap neuron. Dalam CNN, setiap neuron direpresentasikan dalam bentuk dua dimensi, berbeda dengan MLP di mana setiap neuron hanya berukuran satu dimensi (Putra I Wayan Suartika E. et al., 2016).



Gambar 2.4 Lapisan CNN

Sumber : mathworks.com

Pada dasarnya CNN terdiri dari lapisan *input*, lapisan *output*, dan banyak lapisan yang tersembunyi (*hidden layer*), yang ditunjukkan pada gambar 2.4. Umumnya pada lapisan tersembunyi tersebut terdiri dari lapisan konvolusi, *Rectified Linear Unit* (ReLU), dan *pooling* (Ketkar & Moolayil, 2021) .



Gambar 2.5 Hidden Layer

Sumber : mathworks.com

Dari gambar 2.5 dapat kita lihat bahwa terdapat 3 lapisan inti pada bagian *feature learning*, bagian ini merupakan bagian dari hidden layer atau lapisan tersembunyi, *feature learning* merupakan proses di mana jaringan secara otomatis mengekstrak fitur-fitur penting dari data input yang kompleks seperti gambar. Berikut penjelasan mengenai lapisan tersembunyi:

- Lapisan konvolusi berfungsi untuk mengekstraksi fitur dari input menggunakan operasi konvolusi dengan filter atau kernel. Filter ini akan bergeser ke seluruh gambar untuk mendeteksi pola dan fitur lokal seperti tepi, sudut, atau tekstur.
- *Rectified Linear Unit* merupakan lapisan yang memungkinkan pelatihan dapat dilakukan dengan lebih cepat dan efektif, yakni dengan memetakan nilai negatif menjadi nol dan mempertahankan nilai positif.
- Lapisan Pooling digunakan untuk mengurangi dimensi spatial dari representasi yang dihasilkan oleh lapisan konvolusi.

Singkatnya, CNN mampu memetakan fungsi dari input gambar ke *output* kelas dengan cara mengekstrak fitur secara hierarkis dari fitur dasar hingga fitur yang lebih kompleks.

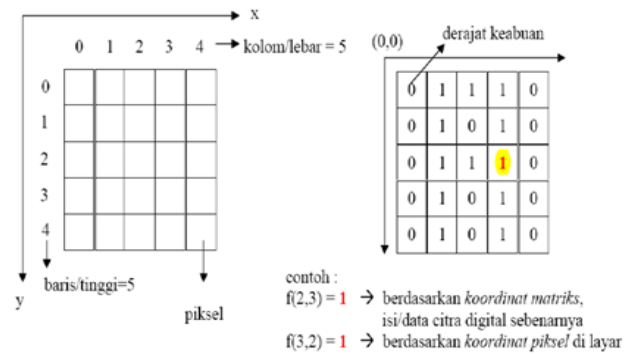
2.1.3 Model (YOLO)

Sejak pertama kali diperkenalkan pada tahun 2015 oleh Joseph Redmon dkk. berjudul '*You Only Look Once: Unified, Real-Time Object Detection*'. Makalah tersebut mengubah paradigma deteksi objek, dengan menggambarkan pada dasarnya sebagai masalah regresi satu langkah, dimulai dari piksel gambar dan bergerak menuju kotak pembatas (*Bounding Box*) serta probabilitas kelas. Pendekatan yang diusulkan, berdasarkan konsep '*unified*', memungkinkan prediksi simultan beberapa kotak pembatas dan probabilitas kelas, sehingga meningkatkan kecepatan dan akurasi. Sejak diperkenalkan pertama kali, detektor objek YOLO (*You Only Look Once*) telah mengalami perkembangan yang pesat, dengan versi terbaru yakni model YOLO-v8 yang dirilis pada bulan Januari 2023 (Hussain, 2023).

YOLOv8 merupakan model yang dikembangkan oleh perusahaan Ultralytics yang juga pengembang model YOLOv5, model ini rilis pada Januari 2023. Beberapa fitur yang diunggulkan pada YOLOv8 yakni arsitektur backbone dan neck yang canggih, *anchor-free split* Ultralytics *head*, mengoptimalkan keseimbangan antara kecepatan dan akurasi, dan beragam model pra-latih. Dengan adanya fitur-fitur tersebut memberikan peningkatan dalam hal akurasi dan kecepatan pada YOLOv8 jika dibandingkan dengan algoritma YOLOv7, v6-2.0, dan v5-7.0 seperti yang terlihat pada gambar 2.7.

2.1.4 Citra

Citra merupakan representasi dua dimensi dari objek fisik nyata dalam tiga dimensi (Putri Riani Asti, 2016). Secara matematis citra dapat dianggap sebagai fungsi kontinu dari intensitas cahaya pada bidang dua dimensi. Fungsi dua dimensi $f(x,y)$ dapat dianggap sebagai representasi matematis dari citra. Fungsi ini mewakili nilai ciri atau intensitas cahaya pada setiap titik elemen (x,y) dalam citra, yang didapatkan dari peralatan sensor yang merekam nilai-nilai tersebut. Suatu citra digital $a[m,n]$ dapat dianggap sebagai ruang diskrit berdimensi-dua yang berasal dari citra analog $a(x,y)$ melalui proses sampling, yang sering disebut sebagai digitalisasi. Proses ini mengonversi citra kontinu dua dimensi $a(x,y)$ menjadi representasi citra diskrit $a[m,n]$ dengan menggunakan grid diskrit pada titik-titik (m,n) . (Arnita et al., 2022). Untuk memahami lebih mudah dapat dilihat pada gambar 2.8.



Gambar 2.8 Fungsi Dua Dimensi Citra

Sumber : putuadisusanta.wordpress.com

Secara umum, terdapat 3 jenis citra yang digunakan dalam pemrosesan citra, yaitu citra berwarna, citra berskala keabuan, dan citra biner.

- Citra berwarna atau citra RGB mendeskripsikan warna sebagai kombinasi positif dari komponen R (*red/merah*), G (*green/hijau*), dan B (*blue/biru*) sehingga membentuk suatu susunan warna yang sangat luas.
- Citra berskala keabuan atau citra *grayscale* mendeskripsikan warna dengan gradasi hitam dan putih sehingga akan menghasilkan citra berwarna keabuan.
- Citra biner atau *monochrome* mendeskripsikan warna dengan 2 kemungkinan nilai, yakni 0 dan 1. Dengan nilai 0 menyatakan warna hitam dan nilai 1 menyatakan warna putih. Sehingga citra yang dihasilkan berupa gambar hitam dan putih.

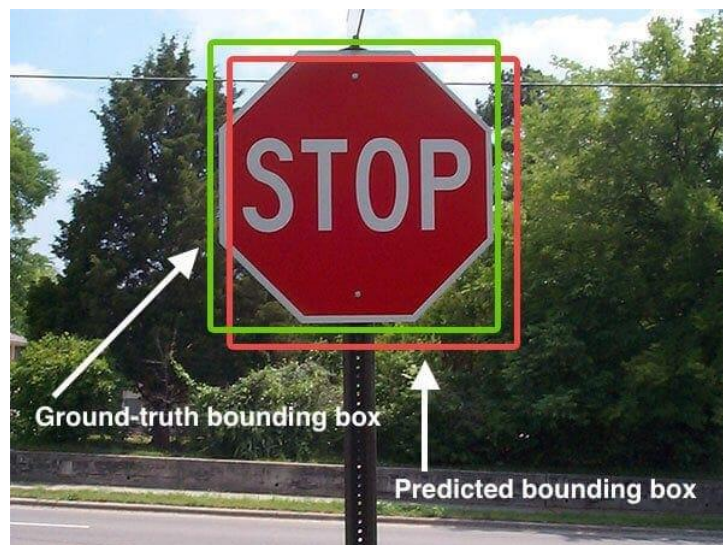
2.1.5 Mean Average Precision (mAP)

Mean Average Precision (mAP) merupakan metrik untuk mengukur akurasi prediksi yang menghitung rerata dari nilai presisi rata-rata berbagai kelas dalam. Metrik ini umum digunakan dalam menguji performa deteksi objek dan segmentasi data. Sebelum mengenal lebih dalam mengenai perhitungan mAP, terlebih dahulu

akan dijelaskan mengenai *Intersection of Union* (IoU), *Confusion Matrix*, *Precision*, dan *Recall*.

- *Intersection over Union* (IoU)

Intersection over Union (IoU) merupakan metode evaluasi yang digunakan untuk mengukur akurasi deteksi objek terhadap suatu *dataset*. IoU memerlukan dua area yang akan dihitung, yaitu area hasil interseksi dan area hasil gabungan dari dua *bounding box*, yang merupakan *bounding box* dari *ground-truth* (*bounding box* aktual) dan *bounding box* yang dihasilkan oleh model deteksi (Shianto et al., 2019). Dengan ilustrasi yang ditunjukkan pada gambar 2.9.



Gambar 2.9 Ilustrasi IoU

Sumber : pyimagesearch.com

IoU dihitung dengan rumus:

$$IoU = \frac{\text{Area of Overlap (area interseksi)}}{\text{Area of Union (area gabungan)}} \quad (2.1)$$

Nilai IoU berkisar antara 0 dan 1. Nilai 1 menunjukkan bahwa *bounding box* hasil deteksi sepenuhnya bersamaan dengan *bounding box ground-truth*, sementara nilai 0 menunjukkan tidak ada interseksi antara kedua *bounding box* tersebut. IoU digunakan untuk mengevaluasi sejauh mana *bounding box* yang dihasilkan oleh model

mendekati atau cocok dengan *ground truth bounding box*, sehingga memberikan gambaran yang lebih baik tentang kualitas deteksi objek.

- *Confusion Matrix*

Confusion matrix merupakan sebuah tabel yang berfungsi untuk menganalisis sejauh mana klasifier mampu mengenali tuple dari kelas yang berbeda (W. A. Kurniawan et al., 2019). Dalam *confusion matrix* terdapat 4 kondisi yakni :

Tabel 2.1 Confusion Matrix

<i>Confusion Matrix</i>		Keadaan Sebenarnya	
		Positif (1)	Negatif (0)
Prediksi Model	Positif (1)	TP	FP
	Negatif (0)	FN	TN

Dari tabel 2.1 terdapat berbagai kondisi yang terjadi antara keadaan aktual dan prediksi model, berikut penjelasannya :

1. *True Positive* (TP), merupakan kondisi ketika model mengkasifikasikan data (prediksi) dengan nilai positif dan dengan keadaan sebenarnya yang bernilai positif.
2. *False Positif* (FP), merupakan kondisi ketika model mengkasifikasikan data (prediksi) dengan nilai positif sedangkan keadaan sebenarnya yang bernilai negatif.
3. *False Negatif* (FN), merupakan kondisi ketika model mengkasifikasikan data (prediksi) dengan nilai negatif sedangkan keadaan sebenarnya bernilai positif.
4. *True Negatif* (TN), merupakan kondisi ketika model mengkasifikasikan data (prediksi) dengan nilai negatif dan keadaan sebenarnya bernilai negatif.

Jumlah setiap kondisi tersebut akan digunakan sebagai data dalam persamaan *precision* dan *recall*, untuk mendapatkan nilai akhir yakni hasil persamaan nilai mAP.

- *Precision*

Precision merupakan perbandingan antara jumlah data yang termasuk dalam kategori positif yang diklasifikasikan secara benar oleh sistem dengan keseluruhan data yang sistem klasifikasikan sebagai positif (Shianto et al., 2019).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2.2)$$

Dari persamaan 2.2 dapat dilihat bahwa presisi dapat dihitung dengan membagi jumlah *True Positive* (TP) oleh jumlah keseluruhan data yang diklasifikasikan sebagai positif. Presisi memberikan informasi tentang seberapa akurat model dalam mengidentifikasi *instance* yang sebenarnya positif, dan nilai presisi yang tinggi menunjukkan bahwa sebagian besar *instance* yang diklasifikasikan sebagai positif oleh model memang benar-benar positif.

- *Recall*

Recall, juga dikenal sebagai *sensitivity* atau *true positive rate*, merupakan pengukuran untuk data dengan klasifikasi positif yang benar oleh sistem (Shianto et al., 2019). *Recall* dapat dihitung menggunakan persamaan :

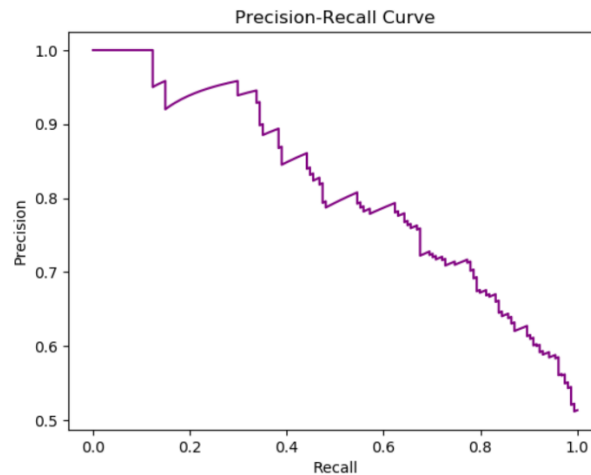
$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2.3)$$

Pada persamaan 2.3, *recall* menggambarkan kemampuan model untuk mengenali dan mengklasifikasikan secara benar *instance* yang sebenarnya positif. Jika *recall* tinggi, itu berarti model cenderung tidak melewatkan *instance* positif yang sebenarnya. *Recall* dan *Precision* adalah dua metrik evaluasi klasifikasi yang sering digunakan bersama-sama untuk memberikan gambaran yang lebih lengkap tentang kinerja model.

- *Precision-Recall Curve*

Precision-recall curve merupakan suatu plot yang digunakan dalam pengklasifikasi dan pemulihan informasi, terutama dalam konteks pengklasifikasi biner. Kurva ini memvisualisasikan hubungan

antara presisi dan recall untuk berbagai nilai ambang batas (*threshold*) yang digunakan dalam pengambilan keputusan klasifikasi.



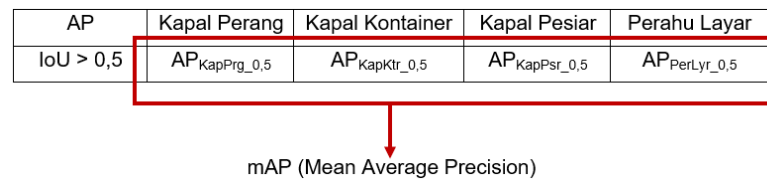
Gambar 2.10 Contoh Kurva *Precision-Recall*

Sumber = statology.org

Kurva *Precision-Recall* memvisualisasikan perubahan presisi dan *recall* ketika ambang batas (*threshold*) IoU digeser. Jika model memiliki nilai recall yang tinggi namun precision yang rendah, model cenderung mengidentifikasi sebagian besar sampel yang sebenarnya positif, tetapi juga memberikan banyak hasil positif palsu, yaitu kasus di mana model mengklasifikasikan sampel yang sebenarnya negatif sebagai positif. Sedangkan jika nilai *precision* tinggi namun *recall* yang rendah, model cenderung memberikan hasil yang benar saat mengidentifikasi sampel sebagai positif, tetapi kemungkinan tidak mengidentifikasi semua sampel yang sebenarnya positif, atau dapat melewatkan beberapa sampel positif. Untuk memperoleh model yang baik maka nilai *precision* dan *recall* harus tetap tinggi pada berbagai *threshold* IoU. Sehingga kurva *precision-recall* yang baik seperti pada gambar 2.10.

Sebelum membahas mengenai mAP perlu diketahui mengenai AP (*Average precision*) dalam tiap-tiap kelas klasifikasi. AP didefinisikan sebagai luas area dibawah kurva *precision-recall*, sedangkan nilai *precision* dan *recall* dapat diperoleh dari rumus yang telah dijelaskan

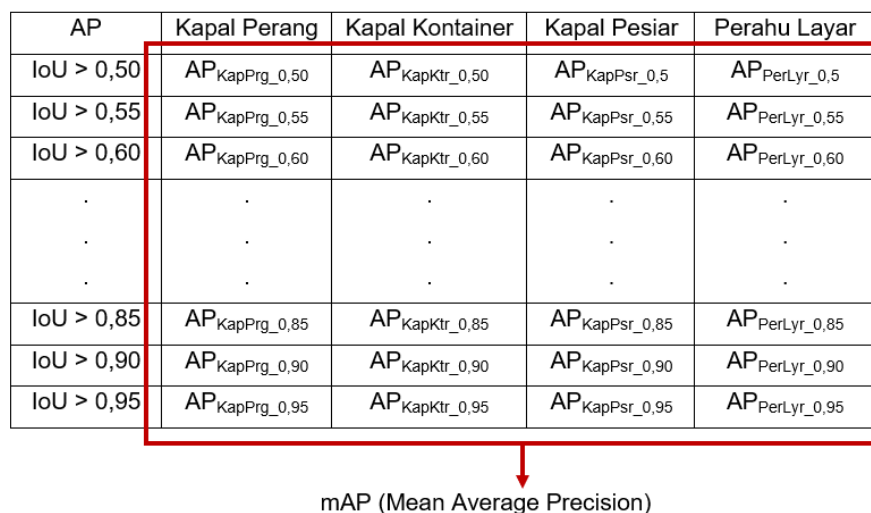
sebelumnya. Jika *confusion matrix* nilai-nilainya diperoleh dengan mengacu pada dasar batas (*threshold*) *IoU dataset Pascal VOC (Visual Object Classes)*, yang mana jika nilai $IoU > 0,5$ maka dinilai kedalam *True Positive* sedangkan jika $IoU \leq 0,5$ maka dianggap sebagai *False Positive*. Juga jika ada dua atau lebih *prediction bounding box* terhadap satu *ground truth bounding box* maka *prediction bounding box* dengan *confidence* yang tertinggi yang dianggap sebagai *True Positive* sedangkan sisanya dianggap sebagai *False Positive* sedangkan jika tidak ada *prediction bounding box* dalam satu *ground truth bounding box* maka dianggap sebagai *False Negative*.



Gambar 2.11 MAP PASCAL VOC

“Sumber : Diolah oleh peneliti”

Setelah memperoleh grafik *precision* terhadap *recall* kemudian diperoleh AP dari nilai integral *precision* terhadap *recall* disetiap titik kontinu dari 0 hingga 1 (PASCAL VOC setelah 2008). Sehingga ketika telah mendapatkan nilai AP untuk seluruh kelas maka nilai tersebut dirata-ratakan kemudian menjadi nilai mAP. Ilustrasi perhitungan mAP dengan menggunakan Pascal VOC dapat dilihat pada gambar 2.11.



Gambar 2.12 MAP COCO

“Sumber : Diolah oleh peneliti”

Jika *confusion matrix* nilai-nilainya diperoleh dengan mengacu pada dasar batas (*threshold*) IoU dataset COCO (*Common Objects in Context*), yang mana dikenal dengan istilah IoU AP[.50: .05: .95] yakni dengan menghitung nilai AP dengan IoU 0,5 hingga IoU 0,95 dengan kenaikan batas (*threshold*) 0,05 pada setiap kelas. Setelah memperoleh grafik *precision* terhadap *recall* kemudian diperoleh AP dari membagi grafik tersebut menjadi 101 titik dimulai dari *precision* saat *recall* = 0, saat = 0.01, saat = 0.02, ..., saat = 1. Sehingga ketika telah mendapatkan nilai AP untuk seluruh kelas maka nilai tersebut dirata-ratakan kemudian menjadi nilai mAP, seperti yang ditunjukkan pada gambar 2.12.

2.1.6 Inertial Measurement Unit (IMU)

Sebelum memahami IMU, perlu diketahui mengenai INS yakni sistem navigasi yang menggunakan prinsip dasar fisika untuk melacak posisi, kecepatan, dan orientasi objek bergerak tanpa memerlukan sumber eksternal seperti GPS. INS memanfaatkan sensor inersial yang terintegrasi dalam satu platform yang disebut *Inertial Measurement Unit* (IMU) (Magnusson & Odenman, 2012). *Inertial Measurement Unit* (IMU) merupakan sebuah sensor yang digunakan untuk mengukur percepatan linear, percepatan sudut, dan orientasi dari sebuah objek (A. H. Kurniawan & Rivai, 2018).

IMU menggunakan sensor inersial, seperti *accelerometer* dan *gyroscope*, untuk mengukur gerakan objek dalam ruang tiga dimensi. *Accelerometer* digunakan untuk mengukur percepatan linier objek dalam tiga sumbu yakni x, y, dan z. Percepatan linier adalah perubahan kecepatan dalam satu arah tertentu. Sedangkan *gyroscope* digunakan untuk mengukur laju perubahan sudut objek dalam tiga sumbu: roll, pitch, dan yaw. Laju perubahan sudut ini mengindikasikan seberapa

cepat objek berputar dalam ruang tiga dimensi. Terkadang imu juga menggunakan sensor magnetometer yang mengukur medan magnetik di sekitar objek. Ini dapat membantu dalam menentukan orientasi objek terhadap medan magnet bumi.

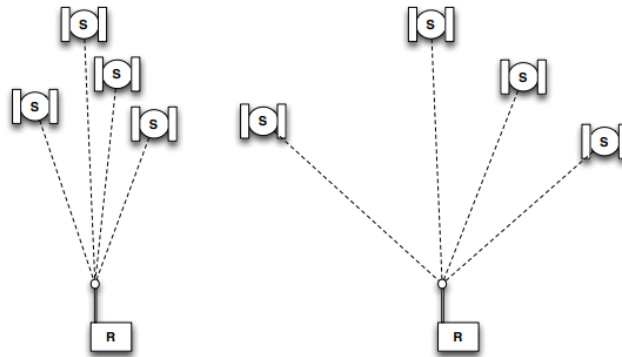
Dalam IMU terdapat istilah *Degree of Freedom* (DOF), DOF mengacu pada jumlah derajat kebebasan yang dapat diukur oleh sensor dalam unit tersebut. Jumlah DOF menentukan seberapa lengkapnya informasi gerakan yang dapat diperoleh dari IMU. Secara umum IMU yang menggunakan sensor *accelerometer* dan *gyroscope* memiliki 6 DOF yakni tiga DOF untuk gerakan translasi (Gerakan sepanjang sumbu x, y, dan z) dan tiga DOF untuk gerakan rotasi (Rotasi sepanjang sumbu x (*pitch*), y (*roll*), dan z (*yaw*)). Sedangkan pada IMU yang menggunakan sensor *accelerometer*, *gyroscope*, dan *magnetometer* memiliki 9 DOF dengan 3DOF yang merujuk pada kemampuan sensor untuk mendeteksi medan magnet dan memberikan informasi tentang orientasi relatif objek terhadap medan magnet Bumi.

2.1.7 Global Positioning System (GPS)

GPS adalah sistem navigasi berbasis satelit yang memungkinkan pengguna di seluruh dunia untuk menentukan lokasi, waktu, dan kecepatan mereka secara akurat. GPS terdiri dari sejumlah satelit yang mengorbit Bumi dan stasiun kontrol di darat yang mengontrol dan memantau sistem secara keseluruhan. Dengan 24 satelit yang beroperasi pada 6 orbit yang berbeda, GPS memiliki keakuratan dan cakupan yang sangat luas.

Untuk mengetahui lokasi perangkat pengguna/GPS, GPS mengirimkan sinyal yang kemudian diterima oleh satelit kemudian satelit mengirimkan kembali sinyal-sinyal tersebut sehingga diterima oleh pengguna. Selisih waktu yang diperoleh dalam pengiriman dan penerimaan sinyal akan dihitung sehingga dapat diketahui jarak

pengguna dengan satelit. Dengan memperoleh informasi jarak relatif ke tiga atau lebih satelit, penerima GPS dapat melakukan trilaterasi untuk menentukan posisi pengguna dengan akurasi yang tinggi.



Gambar 2.13 Ilustrasi DOP

Sumber : (Magnusson & Odenman, 2012)

Evaluasi akurasi pengukuran posisi GPS dapat dilakukan menggunakan berbagai informasi, termasuk standar deviasi, *signal to noise ratio*, dan *dilution of precision* (DOP) yang dapat dilihat pada gambar 2.17. *Dilution of Precision* (DOP) adalah sebuah konsep dalam navigasi GPS yang mengukur seberapa baik kondisi geometri antara satelit GPS yang terlihat dan penerima GPS. DOP digunakan untuk mengevaluasi seberapa akurat atau presisi estimasi posisi yang diberikan oleh GPS. Semakin rendah nilai DOP, semakin baik geometri antara satelit dan semakin tinggi akurasi posisi yang diharapkan, dan sebaliknya.

2.2 Hasil Penelitian Terdahulu

Penelitian terdahulu berfungsi sebagai acuan dalam membuat penelitian ini sehingga penelitian ini dapat berlandaskan baik secara teoritis maupun empiris. Berikut hasil penelitian terdahulu yang berkaitan dengan penelitian yang akan dilakukan :

- a) Cen et al., (2023) dengan judul "*An Improved Ship Classification Method Based on YOLOv7 Model with Attention Mechanism*". Penelitian ini membahas mengenai metode klasifikasi kapal berbasis model YOLOv7 yang ditingkatkan dengan menambahkan *Residual Convolutional Block Attention Module* (RCBAM), penambahan ini bertujuan untuk meningkatkan akurasi dan kecepatan pelatihan dalam mendeteksi kapal kecil dari citra pengawasan video. Dengan memanfaatkan *transfer learning* dan metode penggabungan fitur, model YOLOv7-RCBAM menunjukkan peningkatan signifikan dalam *recall*, *precision*, *F1 score*, dan *mean Average Precision* (mAP), yang mencapai mAP tertinggi 97.59%, mengungguli versi YOLO lainnya seperti YOLOv4 dan YOLOv5. Model ini menunjukkan bahwa integrasi *attention mechanism* CBAM dan pendekatan residual dapat secara efektif memperkuat ekstraksi fitur dan meningkatkan kemampuan anti-gangguan dalam skenario pengawasan video.
- b) Wu et al., (2023) dengan judul "*Ship Detection and Recognition Based on Improved YOLOv7*". Penelitian ini membahas mengenai peningkatan model YOLOv7 untuk deteksi dan pengenalan kapal dengan mengganti *fixed anchor boxes* dengan *anchor boxes* yang lebih sesuai berdasarkan ukuran kapal dalam *dataset*, serta memperkenalkan modul *multi-scale feature fusion* untuk penangkapan fitur yang lebih efisien. Teknik augmentasi data juga diterapkan untuk meningkatkan keberagaman dan ketahanan data. Hasil eksperimen menunjukkan bahwa model yang diusulkan secara signifikan mengungguli pendekatan terkini dalam hal akurasi dan

ketahanan deteksi, dengan mAP akhir sebesar 90,15%, yang berpotensi untuk digunakan dalam aplikasi praktis di domain maritim.

- c) Chang et al., (2022) dengan judul “*Modified Yolov3 for Ship Detection with Visible and Infrared Images*”. Penelitian ini membahas mengenai pengembangan model Yolov3 yang dimodifikasi untuk deteksi kapal menggunakan gambar terlihat dan inframerah, yang bertujuan untuk meningkatkan efisiensi operasi dan manajemen pelabuhan. Modifikasi yang dilakukan meliputi pengurangan ukuran gambar input, jumlah filter konvolusi, dan penambahan modul *spatial pyramid pooling* (SPP) untuk memperbaiki ekstraksi fitur. Menggunakan data set yang terdiri dari 5557 gambar inframerah dan cahaya terlihat dari enam jenis kapal, hasil eksperimen menunjukkan bahwa model yang diusulkan meningkatkan *mean average precision* (mAP) sebesar 5.8%, kecepatan pemrosesan atau *Frame Per Second* (FPS) sebesar 8%, dan mengurangi jumlah operasi *floating point* (BFLOPs) sekitar 47.5%. Kesimpulannya, model Yolov3-spp yang dimodifikasi mengungguli Yolov3 asli dalam hal akurasi deteksi dan efisiensi komputasi, dengan masa depan penelitian diarahkan pada mekanisme atensi dan data set yang lebih lengkap.
- d) Han et al., (2021) dengan judul “*ShipYOLO: An Enhanced Model for Ship Detection*”. Penelitian ini membahas mengenai model deteksi kapal yang ditingkatkan dari YOLO-V4 (*Ship YOLO*), dirancang untuk memenuhi kebutuhan deteksi kapal dengan kecepatan dan akurasi yang lebih tinggi. Model ini mengadopsi RCSPDarknet sebagai jaringan *backbone* dengan reparameterisasi struktur untuk performa real-time yang lebih baik, modul DSPP (*Dilated Convolution and Max-Pooling*) untuk memperkuat fitur deteksi kapal kecil, dan struktur piramida fitur AtFPN (*Attention-based Feature Pyramid Network*) untuk meningkatkan akurasi. Hasil eksperimen menunjukkan bahwa ShipYOLO unggul dalam hal kecepatan dan akurasi dibandingkan YOLO-V3 dan YOLO-V4, terutama dalam mendeteksi kapal

berukuran kecil, menjadikannya solusi yang efektif untuk aplikasi deteksi kapal dalam kondisi *real-time*.

- e) Pratama et al., (2020) dengan judul “*Deep learning-based object detection and geographic coordinate estimation system for GeoTiff imagery*”. Penelitian ini mengembangkan sistem berbasis *deep learning* untuk menganalisis citra udara GeoTiff guna mendeteksi objek seperti pesawat, kapal, dan mobil, serta menghitung koordinat geografisnya. Dengan menggunakan *Convolutional Neural Network* (CNN) yang dimodifikasi dari model YOLO untuk mendeteksi objek dan menghitung koordinatnya. Pengujian menunjukkan akurasi deteksi objek sebesar 81,05% dan presisi 93,29%, dengan *Mean Absolute Error* (MAE) estimasi koordinat bervariasi dari $0,000012^\circ$ hingga $0,000034^\circ$. Meskipun akurasinya lebih rendah dibandingkan metode sebelumnya karena klasifikasi objek menjadi tiga kelas, sistem ini unggul dalam kemampuan mengestimasi koordinat geografis yang tidak ditemukan pada penelitian sebelumnya.

Berikut rangkuman dan perbedaan pada penelitian terdahulu dengan penelitian ini, dalam bentuk tabel:

Tabel 2.2 Penelitian Terdahulu

No.	Nama Peneliti	Judul	Metode	Hasil	Perbedaan
1.	Jian Cen , Hao Feng , Xi Liu, Yongjian Hu, Haoliang Li, Haisheng Li, and Weisheng Huang	<i>An Improved Ship Classification Method Based on YOLOv7 Model with Attention Mechanism(Cen et al., 2023)</i>	Dalam penelitian ini, penulis memodifikasi arsitektur model YOLOv7 menjadi menjadi YOLOv7- <i>residual convolutional block attention module</i> (YOLOv7-RCBAM). Metode ini menggabungkan mekanisme konvolusi atensi (<i>Convolutional Attention Mecanism</i>) dan koneksi residual (<i>Residual Connection</i>) ke dalam YOLOv7.	Model YOLOv7-RCBAM mencapai kinerja lebih baik dengan mAP (mean Average Precision) sebesar 97.59%, menunjukkan kemampuannya dalam ekstraksi fitur objek pada lapisan-lapisan dalam. YOLOv7-RCBAM juga dapat dengan akurat mengidentifikasi kapal dalam lingkungan yang kompleks dengan kondisi kegelapan dan noise, mencapai mAP sebesar 96.13%.	Penelitian ini berfokus pada pengembangan pada model YOLOv7 serta tidak memperkirakan koordinat kapal.
2.	Wei Wu , Xiulai Li, Zhuhua Hu and Xiaozhang Liu	<i>Ship Detection and Recognition Based on Improved YOLOv7</i> (Wu et al., 2023)	Dalam penelitian ini, penulis mengembangkan model YOLOv7 dengan mengganti <i>fixed anchor boxes</i> pada model konvensional menjadi model yang lebih cocok, yang dirancang khusus berdasarkan distribusi ukuran kapal dalam <i>dataset</i> . Peneliti juga memperkenalkan modul penggabungan fitur	Model YOLOv7 yang dikembangkan oleh peneliti mengungguli model terkini dalam hal akurasi deteksi dan ketahanan, dengan memperoleh nilai mAP yang mencapai 90,15%. Model menunjukkan akurasi dan nilai mAP tinggi, terutama dalam mengidentifikasi perahu nelayan kecil, sehingga menjadi opsi yang cocok	Penelitian ini berfokus pada pengembangan pada model YOLOv7 serta tidak memperkirakan koordinat kapal.

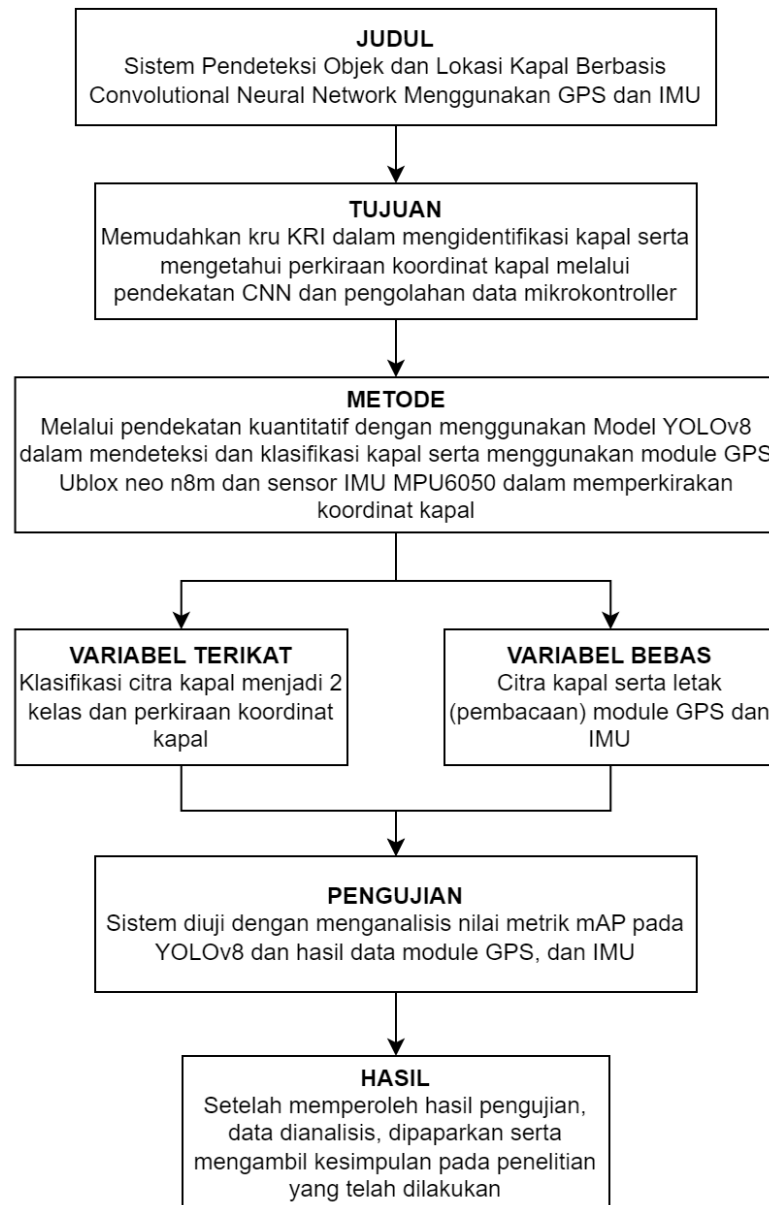
			<p>multiscale baru yang memungkinkan sistem untuk mengklasifikasi kapal secara efisien. <i>Preprocessing</i> data juga ditingkatkan melalui penerapan teknik augmentasi data, untuk memperkuat keragaman dan ketahanan data serta meningkatkan kemampuan model untuk mengenali kapal dengan variasi yang lebih besar.</p>	<p>untuk deteksi identifikasi kapal di jalur air pedalaman (<i>Inland Waterways</i>).</p>	
3.	<p>Lena Chang, Yi-Ting Chen, Jung-Hua Wang, and Yang-Lang Chang</p>	<p><i>Modified Yolov3 for Ship Detection with Visible and Infrared Images</i>(Chang et al., 2022)</p>	<p>dalam penelitian ini, peneliti mengembangkan model YOLOv3 untuk mendeteksi citra kapal dengan citra yang berupa citra terlihat (<i>Visible Image</i>) dan citra inframerah (<i>Infrared Image</i>). Penelitian ini melakukan penyesuaian pada arsitektur YOLOv3 dengan memilih ukuran gambar input yang tepat, filter konvolusi yang lebih sedikit, dan skala deteksi yang lebih sedikit juga. Selanjutnya model dimodifikasi dengan menambahkan <i>Spatial Pyramid Pooling</i> (SPP)</p>	<p>Hasil eksperimen pada <i>dataset</i> menunjukkan bahwa arsitektur jaringan yang dimodifikasi mencapai kinerja yang dapat diterima dalam deteksi kapal, dengan <i>mean average precision</i> (mAP) sebesar 93,2%, pemrosesan 104 frame per detik (FPS), dan 29,2 miliar operasi <i>floating point</i> (BFLOPs).</p>	<p>Penelitian ini menggunakan model YOLOv3 yang dikembangkan, menggunakan citra infra merah serta tidak memperkirakan koordinat kapal.</p>

			untuk meningkatkan kinerja jaringan dalam ekstraksi fitur.		
4.	Xu Han, Lining Zhao, Yue Ning, and Jingfeng Hu	<i>ShipYOLO: An Enhanced Model for Ship Detection</i> (Han et al., 2021)	<p>Penelitian ini mengembangkan model YOLOV4 menjadi model SHIPYOLO yakni model yang hanya mendeteksi kapal saja. Secara garis besar peneliti membagi peningkatannya kedalam tahap; yang pertama Optimasi CSPDarknet Backbone Network dengan melakukan pembaruan (<i>upgrade</i>) Feature Extraction Component dan pembentukan RCSPDarknet Backbone Network. kedua peneliti mendesain DSPP untuk Informasi Spasial Kapal Skala Kecil, dengan menambahkan modul baru yakni <i>Amplified Receptive Field Module</i> serta meningkatkan akuisisi informasi spasial. ketiga, peningkatan Struktur Feature Pyramid (PAFPN) dan penyesuaian untuk</p>	<p>ShipYOLO menunjukkan keunggulan dalam kecepatan dan akurasi, bahkan dengan ukuran input yang berbeda. Dengan ukuran input 320 x 320, ShipYOLO meningkatkan FPS dan mAP@5:5:95 (mAP90) sebesar 23.7% dan 13.6% (10.6%) dibandingkan dengan YOLO-V4 pada PC dengan GPU NVIDIA 1080Ti</p>	<p>Penelitian berfokus dalam membuat model baru berdasarkan model YOLOv4 serta tidak memperkirakan letak koordinat kapal</p>

			skala yang berbeda		
5.	B M Pratama, D Gunawan, and R A G Gultom	<i>Deep learning-based object detection and geographic coordinate estimation system for GeoTiff imagery</i> (Pratama et al., 2020)	Penelitian ini berfokus pada pengaplikasian Algoritma <i>Convolutional Neural Network</i> untuk mendeteksi objek dan memperkirakan koordinat geografis pesawat terbang, kapal, dan mobil melalui citra udara <i>GeoTiff</i> .	Prototipe sistem menunjukkan akurasi dan presisi deteksi objek sebesar 81,05% dan 93,29%, serta memiliki nilai MAE yang bervariasi antara 0,000012 hingga 0,000034 untuk estimasi koordinat objek. Rentang nilai yang kecil menunjukkan tingkat presisi yang baik dalam menentukan koordinat geografis objek.	Penelitian ini menggunakan citra objek yang diambil melalui citra udara Geotiff

Berdasarkan kelima penelitian terdahulu yang bersangkutan dan menjadi dasar dari penelitian ini, penulis menemukan bahwa penelitian yang memiliki kaitan dan paling mendekati penelitian ini yakni penelitian yang dilakukan oleh Pratama et al., 2020. Pada penelitian tersebut peneliti mengkaji mengenai deteksi objek dengan menampilkan titik lokasi perkiraan objek yang dideteksi menggunakan citra geotiff (yakni citra yang diperoleh dari objek udara), sehingga penelitian ini memiliki kesamaan yang menginspirasi penulis dalam membuat sistem deteksi dan perkiraan lokasi kapal musuh. Namun dalam penerapannya penulis menggunakan citra yang diperoleh langsung dari kapal sehingga berbeda dengan penelitian terdahulu yang menggunakan citra geotiff, serta perkiraan lokasi yang penulis angkat dalam penelitian ini berupa perkiraan lokasi kapal pendeteksi sehingga dalam keadaan nyata perkiraan lokasi bukan berupa lokasi nyata dari kapal musuh melainkan lokasi perkiraan dimana citra tersebut dideteksi.

2.3 Kerangka Berpikir



Gambar 2.14 Kerangka Berpikir

“Sumber : Diolah oleh peneliti”

Dari gambar 2.14 sesuai dengan judul penelitian, dalam penelitian ini lingkup pendeteksian objek dan perkiraan lokasi hanya pada kapal laut dengan berbasis metode klasifikasi convolutional neural network serta penggunaan sensor GPS dan IMU. Penelitian ini juga bertujuan untuk memudahkan kru KRI dalam mengidentifikasi kapal dan mengetahui perkiraan koordinat kapal melalui hasil penelitian ini. Metode yang

digunakan dalam penelitian diperoleh melalui pendekatan kuantitatif dengan menggunakan YOLOv8 sebagai model dalam mendeteksi dan mengklasifikasi kapal serta memanfaatkan sensor GPS dan IMU dalam memperoleh perkiraan koordinat dan orientasi kapal. Tentunya untuk memperoleh hasil dalam penelitian, ditentukan variabel terikat berupa klasifikasi kapal menjadi 2 kelas dan perkiraan koordinat kapal dengan variabel bebas berupa citra kapal serta letak pembacaan sensor GPS dan IMU, kedua variabel ini diidentifikasi dengan memanipulasi variabel bebas dan mengamati variabel terikat. Dari eksperimen tersebut sistem akan diuji dengan menganalisis metrik mAP dan selisih *error* pada pembacaan sensor GPS dan IMU. Sehingga dari penelitian tersebut peneliti memperoleh hasil dari penelitian yang telah dilakukan.

2.4 Hipotesis

Dalam penelitian ini, penulis merumuskan hipotesis sebagai berikut:

“Sistem deteksi objek dan lokasi kapal berbasis *Convolutional Neural Network* dan sensor GPS dan IMU akan mampu mendeteksi berbagai jenis kapal kedalam kelas yang ditentukan dengan tingkat akurasi yang tinggi secara real-time yang akan ditinjau melalui tingginya nilai *mean Average Precision* (mAP) serta memberikan titik perkiraan koordinat kapal melalui telegram. Sehingga penelitian diharapkan dapat membantu TNI AL dalam mengatasi permasalahan di perairan Indonesia.”