

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi informasi dan komunikasi pada era digital saat ini telah memberikan dampak yang sangat besar terhadap berbagai aspek kehidupan manusia. Salah satu bidang yang mengalami perkembangan pesat adalah *Computer Vision*. *Computer Vision* merupakan cabang dari kecerdasan buatan (*Artificial Intelligence/AI*) yang memungkinkan sistem komputer untuk mengenali, menafsirkan, dan menganalisis informasi visual dari citra maupun video. Dengan adanya teknologi ini, komputer dapat melakukan tugas-tugas yang sebelumnya hanya bisa dilakukan oleh manusia, seperti mengenali wajah, mendeteksi objek, hingga memahami pola visual dalam lingkungan sekitar [1].

Salah satu cabang utama dalam *Computer Vision* adalah *Object Detection*, yaitu proses untuk mengidentifikasi, mengklasifikasikan, dan menentukan lokasi objek - objek tertentu dalam sebuah citra digital. Teknologi ini memiliki peran penting di berbagai sektor, mulai dari keamanan berbasis CCTV, pengawasan wilayah publik, pemantauan udara dengan drone, hingga sistem transportasi cerdas (*Intelligent Transportation Systems/ITS*). Dalam bidang transportasi, *Object Detection* digunakan untuk mendukung pengembangan kendaraan otonom dan sistem bantuan pengemudi (*Advanced Driver Assistance Systems/ADAS*) [2].

Meskipun *Object Detection* telah mengalami perkembangan yang pesat, penerapannya di dunia nyata masih menghadapi berbagai tantangan teknis. Tantangan tersebut mencakup kesulitan mendeteksi objek berukuran kecil (*small object detection*), penanganan kondisi objek yang tertutup sebagian (*occlusion*), serta variasi kondisi lingkungan seperti perubahan pencahayaan, kabut, dan kualitas citra rendah. Hal ini menuntut pengembangan model *deep learning* yang tidak hanya akurat, tetapi juga *robust* dan efisien untuk dapat beroperasi secara *real-time* dan mampu menggeneralisasi pada kondisi visual yang berbeda [3].

Untuk menghadapi tantangan tersebut, berbagai algoritma *Object Detection* berbasis *deep learning* telah dikembangkan. Salah satu algoritma yang paling populer dan banyak digunakan adalah *You Only Look Once* (YOLO). YOLO merupakan algoritma deteksi objek satu tahap (*one-stage detector*) yang dikenal karena kecepatannya serta kemampuannya bekerja secara *real-time*. Sejak diperkenalkan pertama kali pada tahun 2016, YOLO telah mengalami berbagai pengembangan arsitektur mulai dari YOLOv1 hingga versi terbarunya, YOLOv12 [4]. Setiap versi membawa peningkatan signifikan pada efisiensi deteksi, mekanisme ekstraksi fitur, serta strategi pelatihan.

Berbagai penelitian sebelumnya telah memanfaatkan varian YOLO untuk meningkatkan performa deteksi rambu lalu lintas, dengan fokus pada *robustness* dan akurasi di lingkungan yang menantang. Huang et al. (2023) [1] mengusulkan modifikasi YOLOv8 dengan menambahkan *Asymptotic Feature Pyramid Network* (AFPV) untuk meningkatkan akurasi deteksi objek kecil dengan hasil mAP 59,53%. Upaya untuk meningkatkan ketahanan model terhadap variasi kondisi visual dan lingkungan yang kompleks menjadi fokus utama. Y. He et al., (2025) [5] memperkenalkan NTS-YOLO (berbasis YOLOv5) yang mengintegrasikan metode *unsupervised image enhancement* sebagai langkah *preprocessing* untuk mengatasi kondisi pencahayaan yang tidak optimal, sehingga meningkatkan performa deteksi rambu di kondisi *low-light*/malam hari. F. Li et al. (2023) [3] secara spesifik merancang STC-YOLO (*Small Traffic Sign Detection Network for Complex Environments*) berbasis YOLOv5. Penelitian ini berfokus pada penguatan arsitektur untuk mengatasi tantangan visual di lingkungan yang kompleks, seperti *occlusion* dan *blur*, dengan peningkatan *mAP* yang signifikan dibandingkan model *baseline*. Selain itu, Mahadshetti et al. (2025) [4] mengusulkan integrasi *Multiscale Feature Fusion* dan *Context-Aware Aggregation* pada YOLOv9, yang terbukti mampu mengatasi masalah deteksi objek kecil dan *occlusion*, dengan mAP mencapai 97% pada dataset TT100K dan 90,7% pada CCTSDB2021.

YOLOv12 merupakan versi terbaru dari algoritma YOLO yang diperkenalkan pada tahun 2025 dengan pendekatan *attention-centric*

*architecture*. Model ini menghadirkan berbagai inovasi seperti *Area Attention (A<sup>2</sup>)*, *Residual Efficient Layer Aggregation Network (R-ELAN)*, serta *FlashAttention*, yang meningkatkan efisiensi komputasi dan generalisasi tanpa mengorbankan kecepatan inferensi [2],[6]. YOLOv12 menawarkan keseimbangan ideal antara kecepatan dan akurasi, menjadikannya salah satu model paling efisien untuk tugas deteksi objek *real-time*.

Meskipun YOLOv12 menunjukkan potensi tinggi, performa model deteksi objek tidak hanya ditentukan oleh arsitektur model, tetapi juga sangat bergantung pada kualitas dan keragaman data pelatihan. Pada kasus deteksi rambu lalu lintas, model sering menghadapi kondisi visual yang kompleks seperti pencahayaan rendah, *blur* akibat pergerakan kendaraan, gangguan cuaca, *occlusion*, serta ukuran objek yang relatif kecil dalam citra. Kondisi tersebut menyebabkan model sulit melakukan generalisasi apabila hanya dilatih menggunakan dataset dengan variasi terbatas. Oleh karena itu, diperlukan pendekatan yang mampu meningkatkan keberagaman data tanpa harus melakukan pengumpulan data baru secara besar-besaran [3].

Salah satu pendekatan yang banyak digunakan untuk mengatasi permasalahan tersebut adalah *data augmentation*. Teknik ini bertujuan memperkaya data pelatihan melalui berbagai transformasi citra seperti rotasi, *flipping*, perubahan *brightness* dan *contrast*, penambahan *noise*, *blur*, *mosaic augmentation*, hingga *copy-paste augmentation*. Dengan adanya variasi buatan tersebut, model dapat belajar mengenali objek pada berbagai kondisi visual yang lebih beragam sehingga meningkatkan *robustness*, mengurangi risiko *overfitting*, serta memperbaiki performa deteksi pada data baru yang belum pernah dilihat sebelumnya [1], [7], [8].

Berbagai penelitian terdahulu menunjukkan bahwa penerapan teknik *Data augmentation* memberikan kontribusi signifikan terhadap peningkatan performa model deteksi rambu lalu lintas terutama pada deteksi objek berukuran kecil (*small object detection*), kondisi pencahayaan rendah, serta situasi dengan *occlusion*. Misalnya, penelitian oleh Shi et al. (2024) [9] membandingkan performa model sebelum dan sesudah penerapan *data augmentation* pada deteksi rambu lalu lintas. Pada dataset CTSD, sebelum

augmentasi (*baseline*), model memperoleh nilai  $mAP@0.5$  sebesar 85,4% dan  $mAP@0.5:0.95$  sebesar 62,0%. Setelah augmentasi diterapkan, performa meningkat menjadi 88,9% untuk  $mAP@0.5$  dan 64,5% untuk  $mAP@0.5:0.95$ . Sementara itu, pada dataset GTSDDB, nilai  $mAP@0.5$  meningkat dari 85,2% menjadi 89,2%, dan  $mAP@0.5:0.95$  meningkat dari 52,1% menjadi 56,0%. Hasil tersebut menunjukkan bahwa penerapan *data augmentation* terbukti mampu meningkatkan akurasi deteksi secara signifikan khususnya pada objek kecil dan kondisi lingkungan yang kompleks.

Penelitian lain oleh Jingzhan Ge (2023) [10] pada dataset TT100K juga menunjukkan bahwa penggunaan beberapa skenario *data augmentation* menghasilkan performa yang lebih baik dibandingkan dataset *baseline*. Dalam penelitian tersebut, Version 1 digunakan sebagai dataset dasar (*baseline*) dan dibandingkan dengan beberapa versi dataset hasil augmentasi. Hasil eksperimen menunjukkan bahwa metode augmentasi yang diterapkan efektif dalam meningkatkan performa model, terutama pada kelas rambu dengan jumlah data yang terbatas serta distribusi data yang tidak seimbang.

Sementara itu, He et al. (2025) [5] melalui metode *NTS-YOLO* mengombinasikan *image enhancement* dan augmentasi berbasis *brightness shift* guna meningkatkan deteksi rambu di kondisi malam hari. Pendekatan lain dilakukan oleh Soufi & Valdenegro-Toro (2019) [11] dan Sielemann et al. (2024) [12] yang menggunakan model *Generative Adversarial Networks (GAN)* untuk menghasilkan citra sintetis rambu dalam berbagai kondisi seperti buram, kotor, atau tertutup sebagian. Kajian literatur terkini oleh Flores-Calero et al. (2024) [13] juga menegaskan bahwa kombinasi teknik augmentasi geometris (rotasi, flipping), fotometrik (brightness dan contrast), serta *mosaic* dan *CutMix* merupakan strategi paling efektif untuk memperkuat *robustness* model di lingkungan kompleks.

Temuan tersebut menunjukkan bahwa evaluasi perbandingan antara model tanpa augmentasi (*baseline*) dan model dengan augmentasi menjadi penting untuk mengukur kontribusi nyata dari teknik *data augmentation* terhadap performa deteksi objek. Namun, sebagian besar penelitian terdahulu masih berfokus pada varian YOLO sebelumnya seperti YOLOv5, YOLOv8, dan

YOLOv9, sedangkan kajian mengenai pengaruh *data augmentation* terhadap performa native YOLOv12 sebagai versi terbaru dengan pendekatan *attention-centric architecture* masih sangat terbatas. Padahal, YOLOv12 menawarkan efisiensi komputasi yang lebih baik, kemampuan generalisasi yang lebih kuat, serta kecepatan inferensi yang tetap optimal untuk deteksi objek secara *real-time*.

Berdasarkan kondisi tersebut, penelitian ini difokuskan untuk menganalisis pengaruh *data augmentation* terhadap performa YOLOv12 dalam mendeteksi rambu lalu lintas dengan membandingkan dua skenario pelatihan, yaitu model tanpa augmentasi (*baseline*) dan model dengan penerapan augmentasi data. Evaluasi dilakukan menggunakan metrik Precision, Recall, *mean Average Precision* (mAP), serta *latency* untuk mengukur akurasi dan efisiensi inferensi model. Hasil penelitian ini diharapkan dapat memberikan gambaran yang lebih jelas mengenai kontribusi *data augmentation* terhadap peningkatan akurasi, efisiensi inferensi, serta *robustness* model YOLOv12 dalam sistem deteksi rambu lalu lintas berbasis *Computer Vision*.

## 1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, maka permasalahan yang akan diteliti dalam skripsi ini dapat dirumuskan sebagai berikut:

1. Bagaimana menerapkan model YOLOv12 dan teknik *Data augmentation* untuk mendeteksi rambu lalu lintas guna mengetahui pengaruh augmentasi terhadap hasil deteksi rambu lalulintas?
2. Bagaimana kinerja model YOLOv12 dan teknik *Data augmentation* dalam mendeteksi rambu lalu lintas ?

## 1.3 Batasan Masalah

Agar penelitian ini lebih terfokus dan berjalan sesuai dengan tujuan yang telah ditetapkan, maka ruang lingkup penelitian ini dibatasi pada hal-hal berikut:

1. Objek penelitian dibatasi pada rambu lalu lintas di Indonesia yang terdapat dalam dataset Traffic\_Indonesia dari Roboflow Universe.

2. Algoritma deteksi objek yang digunakan difokuskan pada YOLOv12m tanpa melakukan modifikasi terhadap arsitektur internal (native model).
3. Tahap *preprocessing* dibatasi pada penyesuaian ukuran citra menjadi  $640 \times 640$  piksel, sesuai dengan standar input YOLO.
4. Penelitian ini menerapkan Teknik kombinasi *data augmentation* pada tahap pelatihan untuk meningkatkan keragaman data dan ketahanan (*robustness*) model terhadap variasi kondisi citra. Teknik augmentasi yang diterapkan meliputi: Transformasi Geometris (seperti *random rotation* dengan batasan derajat kecil, dan *random scaling*), Perubahan Warna (*HSB adjustment*), Mosaic Augmentation, dan Efek Visual Lingkungan (*Motion Blur*, *Gaussian Noise*).

#### 1.4 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Menerapkan model YOLOv12 dan teknik *Data augmentation* pada dataset rambu lalu lintas untuk mengidentifikasi pengaruh augmentasi terhadap hasil deteksi.
2. Mengevaluasi kinerja model YOLOv12 dalam mendeteksi rambu lalu lintas melalui perbandingan performa antara model tanpa augmentasi dan model dengan augmentasi.

#### 1.5 Manfaat Penelitian

##### 1. Manfaat Teoritis

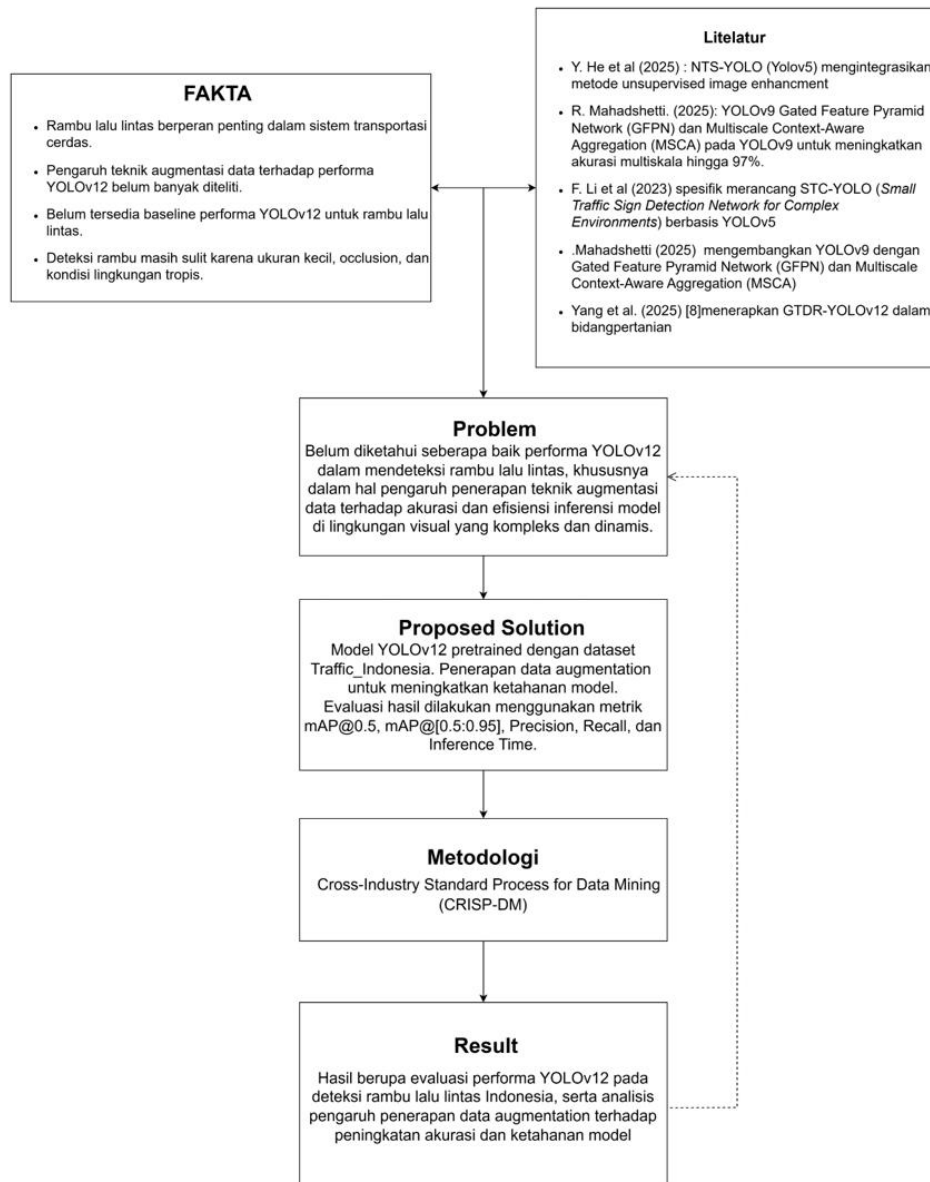
Penelitian ini memberikan kontribusi dalam pengembangan ilmu *Computer Vision*, khususnya terkait pengaruh *data augmentation* terhadap performa dan generalisasi model YOLOv12. Selain itu, penelitian ini dapat menjadi referensi dan *baseline* bagi penelitian selanjutnya dalam deteksi objek berbasis *deep learning*.

##### 2. Manfaat Praktis

Penelitian ini dapat menjadi acuan dalam pengembangan sistem deteksi rambu lalu lintas yang lebih akurat dan *robust*. Selain itu, memberikan

gambaran implementasi *data augmentation* untuk meningkatkan performa model tanpa perlu modifikasi arsitektur.

## 1.6 Kerangka Pemikiran



Gambar 1. 1 Kerangka Pemikiran

Pada gambar 1.1 Kerangka berpikir dalam penelitian ini diawali dari fakta bahwa rambu lalu lintas memiliki peranan penting dalam mendukung sistem transportasi cerdas (*Intelligent Transportation System*). Rambu lalu lintas berfungsi sebagai media komunikasi visual yang menyampaikan perintah, larangan, atau peringatan kepada pengguna jalan. Namun, proses deteksi rambu di lingkungan nyata masih menghadapi tantangan yang signifikan seperti ukuran rambu yang kecil

dalam citra, occlusion oleh kendaraan atau pepohonan, serta pengaruh kondisi lingkungan tropis seperti pencahayaan tidak stabil, kabut, dan permukaan rambu yang pudar atau berkarat.

Selain itu, hingga saat ini belum tersedia *baseline* performa YOLOv12 yang secara khusus dievaluasi pada rambu lalu lintas. Padahal, YOLOv12 merupakan varian terbaru dari keluarga YOLO yang memiliki keunggulan dalam hal efisiensi arsitektur, kecepatan inferensi tinggi, dan kemampuan adaptasi terhadap objek berukuran kecil. Di sisi lain, pengaruh teknik augmentasi data terhadap peningkatan performa dan ketahanan (*robustness*) YOLOv12 terhadap variasi kondisi lingkungan juga belum banyak diteliti.

Berdasarkan tinjauan literatur, beberapa penelitian terdahulu telah berkontribusi terhadap peningkatan performa algoritma YOLO pada berbagai domain. Y. He et al. (2025) [5] memperkenalkan NTS-YOLO (berbasis YOLOv5) yang mengintegrasikan metode unsupervised image enhancement sebagai langkah *preprocessing* untuk mengatasi masalah pencahayaan yang tidak memadai, F. Li et al. (2023) [3] secara spesifik merancang STC-YOLO (*Small Traffic Sign Detection Network for Complex Environments*) berbasis YOLOv5. Mahadshetti (2025) [4] mengembangkan YOLOv9 dengan Gated Feature Pyramid Network (GFPN) dan Multiscale Context-Aware Aggregation (MSCA) untuk meningkatkan akurasi multi-skala hingga 97%, dan Yang et al. (2025) [14] menerapkan GTDR-YOLOv12 dalam bidang pertanian dengan modifikasi augmentasi data.

Berdasarkan fakta dan literatur tersebut, permasalahan utama yang diangkat dalam penelitian ini adalah belum diketahuinya seberapa baik performa YOLOv12 dalam mendeteksi rambu lalu lintas, terutama dalam konteks penerapan teknik augmentasi data terhadap akurasi deteksi, ketepatan klasifikasi, dan efisiensi inferensi model pada kondisi visual yang kompleks dan dinamis.

Penelitian ini menggunakan model YOLOv12 pretrained pada dataset Traffic\_Indonesia dari Roboflow Universe dengan teknik augmentasi data seperti rotasi, kontras, mosaic untuk meningkatkan ketahanan model. Evaluasi dilakukan menggunakan mAP, Precision, Recall, dan Inference Time berdasarkan metodologi CRISP-DM, dengan hasil akhir berupa analisis performa YOLOv12 dalam

mendeteksi rambu lalu lintas serta pengaruh augmentasi terhadap akurasi dan ketahanan model.

### **1.7 Sistematika Penulisan**

Sistematika penulisan dalam penelitian ini disusun untuk memberikan gambaran yang terstruktur mengenai isi dari setiap bab dalam skripsi, sebagai berikut:

#### **BAB I PENDAHULUAN**

Bab ini berisi latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan. Bab ini menjadi dasar dan gambaran umum mengenai penelitian yang dilakukan.

#### **BAB II KAJIAN LITERATUR**

Bab ini membahas landasan teori yang relevan dengan penelitian, meliputi konsep *Computer Vision*, *Object Detection*, algoritma YOLO (khususnya YOLOv12), serta teknik *data augmentation*. Selain itu, bab ini juga memuat penelitian terdahulu yang menjadi referensi dan dasar pengembangan penelitian.

#### **BAB III METODOLOGI PENELITIAN**

Bab ini menjelaskan tahapan penelitian yang dilakukan, meliputi jenis penelitian, dataset yang digunakan, tahapan preprocessing, penerapan model YOLOv12, implementasi teknik *data augmentation*, serta metode evaluasi kinerja model.

#### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini menyajikan hasil eksperimen yang telah dilakukan, termasuk hasil pelatihan dan pengujian model, analisis performa model tanpa augmentasi dan dengan augmentasi, serta pembahasan terhadap hasil yang diperoleh berdasarkan metrik evaluasi seperti precision, recall, dan mAP.

#### **BAB V SIMPULAN DAN SARAN**

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan serta saran untuk pengembangan penelitian selanjutnya.