



Systematic Literature Review on Object Detection Algorithms for Autonomous Systems: Evaluation of CNN Architectures, Dataset Challenges, and Implementation Strategies

Tinjauan Literatur Sistematis Tentang Algoritma Deteksi Objek untuk Sistem Otonom: Evaluasi Arsitektur CNN, Tantangan Dataset, dan Strategi Implementasi

Mohammad Yasser Arafat^{1*}, Muhammad Bagus Andra²

^{1,2}Fakultas Ilmu Komputer, Universitas Nusa Mandiri, Indonesia

E-Mail: ¹aphet.erisgie@gmail.com, ²bagus.andra.mail@gmail.com

Received Feb 13th 2026; Revised Mar 14th 2026; Accepted Mar 30th 2026; Available Online Apr 24th 2026

Corresponding Author: Mohammad Yasser Arafat

Copyright ©2026 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

CNN-based object detection systems have become a fundamental component in the development of autonomous technologies, including self-driving vehicles, mobile robots, and intelligent surveillance systems. This study conducts a systematic literature review using the PRISMA 2020 framework to evaluate CNN architectures, dataset challenges, and implementation strategies for systems with limited computational resources. From 30 identified articles, only six high-quality primary studies (2021–2025) met strict inclusion criteria across major academic databases. The analysis reveals that hybrid architectures integrating attention mechanisms and multi-scale feature pyramid networks achieve superior performance, with AttenRetina reaching a mean Average Precision (mAP) of 0.86 on the KITTI dataset. Key dataset challenges include small object detection, complex backgrounds, partial occlusion, and varying lighting conditions. These issues are addressed through dynamic loss functions and data augmentation techniques. For deployment on resource-constrained devices, lightweight models such as SSD MobileNetV2 and YOLOv8-MobileNetV3 demonstrate an optimal balance between accuracy and efficiency. Overall, this study provides a comprehensive framework for developers to select suitable architectures, prepare datasets, and design deployment strategies tailored to specific autonomous applications.

Keywords: Architectures, Autonomous Systems, Convolutional Neural Network, Dataset Challenges, Object Detection

Abstrak

Sistem deteksi objek berbasis Convolutional Neural Network (CNN) menjadi komponen penting dalam pengembangan teknologi otonom seperti kendaraan tanpa pengemudi, robot mobile, dan sistem pengawasan cerdas. Penelitian ini melakukan tinjauan literatur sistematis menggunakan kerangka PRISMA 2020 untuk mengevaluasi arsitektur CNN, tantangan dataset, serta strategi implementasi pada sistem dengan keterbatasan komputasi. Dari 30 artikel yang diidentifikasi, hanya enam artikel primer berkualitas tinggi (2021–2025) yang memenuhi kriteria seleksi ketat dari berbagai basis data akademik. Hasil analisis menunjukkan bahwa arsitektur hybrid yang menggabungkan attention mechanism dan multi-scale feature pyramid networks memberikan performa terbaik, dengan model AttenRetina mencapai mAP 0.86 pada dataset KITTI. Tantangan utama dalam dataset meliputi deteksi objek kecil, latar belakang kompleks, oklusi parsial, serta variasi pencahayaan. Masalah ini diatasi melalui penggunaan dynamic loss functions dan teknik data augmentation. Untuk implementasi pada perangkat dengan sumber daya terbatas, arsitektur ringan seperti SSD MobileNetv2 dan YOLOv8-MobileNetV3 terbukti mampu memberikan keseimbangan optimal antara akurasi dan efisiensi. Secara keseluruhan, studi ini menawarkan panduan komprehensif bagi pengembang dalam memilih arsitektur, menyiapkan dataset, dan merancang strategi deployment sesuai kebutuhan aplikasi otonom.

Kata Kunci: Arsitektur, Convolutional Neural Network, Dataset Challenges, Object Detection, Sistem Otonom



1. PENDAHULUAN

Perkembangan sistem otonom telah mengalami transformasi signifikan dalam dekade terakhir, dengan object detection menjadi komponen fundamental yang menentukan keberhasilan navigasi dan pengambilan keputusan otomatis [1], [2], [3], [4]. Sistem deteksi objek berbasis Convolutional Neural Network (CNN) telah merevolusi kemampuan kendaraan otonom, robot mobile, dan sistem pengawasan cerdas dalam mengenali serta melokalisasi objek di lingkungan kompleks secara real-time [5], [6], [7]. Teknologi ini memungkinkan sistem otonom untuk memahami konteks visual, menghindari rintangan, dan merespons situasi dinamis dengan tingkat akurasi yang terus meningkat [4], [7]. Namun, implementasi algoritma deteksi objek pada sistem otonom menghadapi tantangan kritis terkait arsitektur model, kualitas dataset pelatihan, dan strategi deployment yang efisien [8], [9], [10], [11].

Arsitektur CNN modern seperti You Only Look Only (YOLO), Single Shot Detector (SSD), EfficientDet, serta model berbasis transformer telah menunjukkan kemajuan besar dalam mengatasi trade-off antara akurasi dan kecepatan inferensi [8], [6], [7], [4], [12], [10]. Model hybrid yang mengintegrasikan konvolusi pada layer awal dengan mekanisme attention berbasis transformer pada tingkat lebih dalam telah meningkatkan efisiensi komputasi sambil mempertahankan akurasi deteksi tinggi, menjadikannya lebih layak untuk aplikasi real-time seperti autonomous driving dan edge computing [9], [13], [6], [14]. Penelitian terbaru menunjukkan bahwa arsitektur seperti SSD MobileNetv2 FPN Lite mampu mencapai tingkat akurasi klasifikasi tinggi dengan waktu inferensi cepat pada sistem kontrol embedded, yang sangat penting untuk navigasi robot mobile otonom [15], [13]. Namun, pemilihan arsitektur yang tepat harus mempertimbangkan berbagai faktor seperti kompleksitas komputasi, konsumsi memori, dan kemampuan generalization pada kondisi lingkungan yang bervariasi [16], [17].

Tantangan dataset menjadi aspek krusial yang memengaruhi performa sistem deteksi objek pada aplikasi otonom. Objek berukuran kecil, oklusi parsial, variasi pencahayaan ekstrem, dan class imbalance merupakan permasalahan yang secara konsisten menghambat akurasi deteksi [18], [19], [20]. Dataset benchmark seperti MS COCO, KITTI, Cityscapes, BDD100K, dan Waymo Open Dataset telah menyediakan standar evaluasi komprehensif, namun model yang dilatih pada dataset spesifik sering mengalami penurunan performa ketika diaplikasikan pada lingkungan berbeda dengan kondisi pencahayaan, sudut kamera, cuaca, dan kepadatan lalu lintas yang bervariasi [1], [21], [2], [3], [22]. Ketidakseimbangan representasi objek kecil dalam dataset menambah kompleksitas, mengakibatkan prediksi yang bias dan penurunan akurasi deteksi pada skenario real-world [13], [20]. Penelitian mengenai adaptasi domain, synthetic data, dan teknik data augmentation menjadi esensial untuk meningkatkan robustness model lintas berbagai kondisi operasional [23], [3], [4].

Strategi implementasi yang efektif menjadi faktor penentu keberhasilan deployment sistem deteksi objek pada perangkat dengan sumber daya terbatas. Teknik seperti knowledge distillation, model pruning, kuantisasi, serta pemilihan backbone ringan telah terbukti meningkatkan efisiensi tanpa mengorbankan akurasi secara signifikan, memungkinkan implementasi pada platform edge computing seperti Jetson Nano, Coral TPU, dan FPGA [16], [13], [6], [12]. Optimalisasi arsitektur melalui lightweight neural networks memungkinkan sistem otonom beroperasi dengan konsumsi daya rendah sambil mempertahankan performa deteksi yang memadai untuk aplikasi kritis seperti autonomous emergency braking dan collision avoidance [15], [13], [7]. Systematic Literature Review (SLR) diperlukan untuk mengonsolidasikan temuan terkini mengenai evaluasi arsitektur CNN, tantangan dataset, dan strategi implementasi yang telah terbukti efektif dalam konteks sistem otonom [24], [10]. Tinjauan komprehensif ini akan memberikan insight berharga bagi peneliti dan praktisi dalam merancang sistem deteksi objek yang robust, efisien, dan dapat diandalkan untuk aplikasi dunia nyata [9], [17].

Berdasarkan latar belakang yang telah dipaparkan, penelitian ini difokuskan pada tiga permasalahan utama. Pertama, bagaimana performa berbagai arsitektur CNN dalam konteks deteksi objek untuk sistem otonom, khususnya dalam hal akurasi, kecepatan inferensi, dan efisiensi komputasi. Kedua, tantangan apa saja yang muncul dari karakteristik *dataset* yang digunakan untuk melatih model deteksi objek, termasuk masalah *class imbalance*, variasi kondisi lingkungan, dan representasi objek berukuran kecil. Ketiga, strategi implementasi apa yang paling efektif untuk *deployment* algoritma deteksi objek pada sistem otonom dengan keterbatasan sumber daya komputasi.

Penelitian ini bertujuan untuk melakukan evaluasi sistematis terhadap arsitektur CNN yang digunakan dalam deteksi objek untuk sistem otonom, mengidentifikasi kelebihan dan keterbatasan masing-masing pendekatan dalam skenario aplikasi yang berbeda. Selain itu, penelitian ini menganalisis tantangan yang berkaitan dengan karakteristik *dataset* dan pengaruhnya terhadap performa model deteksi objek. Terakhir, penelitian ini menyusun rekomendasi strategi implementasi yang optimal untuk *deployment* sistem deteksi objek pada platform dengan sumber daya terbatas, mempertimbangkan aspek efisiensi dan keandalan operasional. Hasil penelitian ini diharapkan memberikan kontribusi teoretis berupa *framework* komprehensif untuk evaluasi arsitektur CNN dalam konteks deteksi objek sistem otonom, serta pemahaman mendalam tentang pengaruh karakteristik *dataset* terhadap performa model. Secara praktis, penelitian ini menyediakan panduan implementasi yang dapat digunakan oleh praktisi dan pengembang sistem otonom dalam memilih

arsitektur, menyiapkan *dataset*, dan merancang strategi *deployment* yang sesuai dengan kebutuhan aplikasi spesifik mereka, sehingga meningkatkan efisiensi pengembangan dan keandalan sistem deteksi objek dalam aplikasi dunia nyata.

2. METODE

Penelitian ini mengadopsi pendekatan *Systematic Literature Review* (SLR) yang dirancang untuk mengidentifikasi, mengevaluasi, dan menyintesis studi relevan mengenai algoritma deteksi objek berbasis CNN untuk sistem otonom. Proses tinjauan dilaksanakan mengikuti kerangka kerja PRISMA 2020 (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) yang menekankan transparansi dan reproduksibilitas dalam setiap tahapan penelitian [24]. Strategi pencarian literatur dilakukan secara komprehensif melalui empat basis data akademik utama, yaitu IEEE Xplore, ScienceDirect, Scopus, dan Google Scholar, dengan rentang waktu publikasi antara tahun 2020 hingga 2025 untuk memastikan relevansi dan kemutakhiran temuan. Pencarian awal menghasilkan total 30 artikel dari seluruh basis data sebelum dilakukan proses eliminasi duplikasi. *String* pencarian terstruktur menggunakan kombinasi kata kunci (*keywords*) berikut: "object detection", "*Convolutional Neural Network*", "*autonomous systems*", "YOLO", "*dataset challenges*", "*CNN architecture*", "*implementation strategies*", "*embedded systems*", dan "*real-time object detection*", yang dikombinasikan menggunakan operator Boolean AND/OR untuk memaksimalkan cakupan artikel yang relevan (Paul, 2025). Kriteria inklusi yang ditetapkan dalam penelitian ini mencakup: (1) artikel jurnal, prosiding konferensi, dan publikasi ilmiah *peer-reviewed*; (2) studi yang secara eksplisit membahas arsitektur CNN untuk deteksi objek; (3) penelitian yang mencakup evaluasi performa model dengan metrik kuantitatif; (4) publikasi yang membahas tantangan *dataset* atau strategi implementasi pada konteks sistem otonom; dan (5) artikel yang diterbitkan dalam rentang tahun 2020–2025. Adapun kriteria eksklusi meliputi: (1) artikel yang tidak berbahasa Inggris; (2) publikasi tanpa proses *peer-review*; (3) studi yang tidak menyediakan data empiris atau metodologi yang jelas; dan (4) penelitian yang tidak berkaitan langsung dengan aplikasi sistem otonom. Dari 30 artikel yang teridentifikasi, dilakukan eliminasi duplikasi sehingga diperoleh 20 artikel unik, yang kemudian disaring melalui *screening* judul dan abstrak menghasilkan 15 artikel, dan setelah pembacaan teks lengkap diperoleh 6 artikel primer yang memenuhi seluruh kriteria [24].

Proses seleksi artikel dilakukan melalui tahapan *screening* judul dan abstrak, diikuti dengan pembacaan teks lengkap untuk memastikan relevansi dengan rumusan masalah penelitian. Ekstraksi data dilakukan secara sistematis menggunakan matriks yang mencakup informasi mengenai jenis arsitektur CNN, *dataset* yang digunakan, metrik evaluasi performa, serta temuan utama terkait implementasi pada perangkat dengan sumber daya terbatas. Analisis dan sintesis data menggunakan pendekatan naratif tematik untuk mengidentifikasi pola, tren, dan kesenjangan penelitian yang ada, serta menyusun rekomendasi berbasis evidensi untuk pengembangan sistem deteksi objek yang robust dan efisien pada aplikasi otonom [10].

3. HASIL DAN PEMBAHASAN

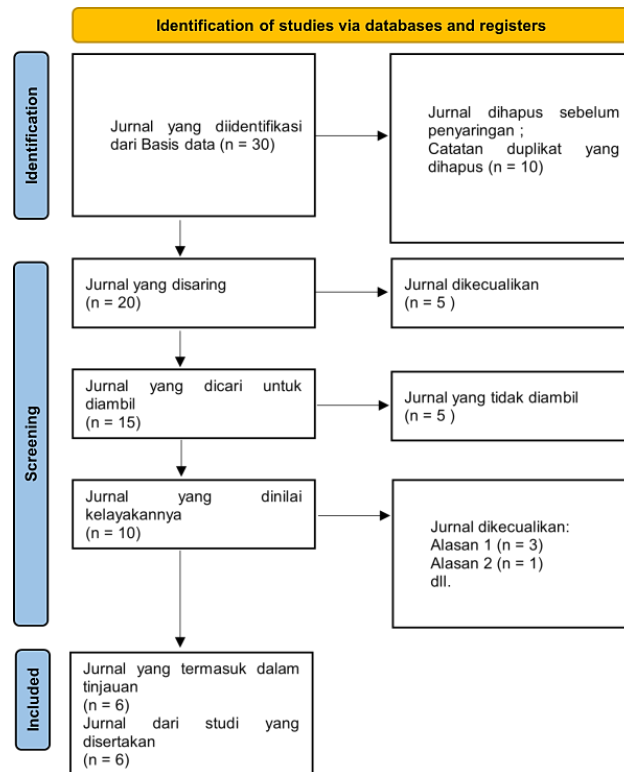
3.1. Proses Seleksi Literatur

Proses identifikasi dan seleksi literatur dalam penelitian ini mengikuti protokol PRISMA 2020 yang memastikan transparansi dan sistematika dalam setiap tahapan peninjauan. Pencarian dilakukan pada empat basis data akademik utama, yaitu IEEE Xplore, ScienceDirect, Scopus, dan Google Scholar, dengan rentang waktu publikasi 2020–2025 menggunakan kombinasi kata kunci terstruktur yang mencakup "object detection", "*Convolutional Neural Network*", "*autonomous systems*", "YOLO", "*dataset challenges*", dan "*implementation strategies*" dengan operator Boolean AND/OR. Tahap awal menghasilkan total 30 artikel dari seluruh basis data tersebut. Setelah dilakukan eliminasi duplikasi, diperoleh 20 artikel unik yang kemudian memasuki tahap *screening* berdasarkan judul dan abstrak. Lima kriteria inklusi dan empat kriteria eksklusi diterapkan secara konsisten dalam setiap tahap seleksi. Pada tahap ini, 5 artikel dieksklusikan karena tidak memenuhi kriteria relevansi dengan fokus penelitian mengenai algoritma deteksi objek untuk sistem otonom. Sebanyak 15 artikel yang lolos kemudian dievaluasi secara menyeluruh melalui pembacaan teks lengkap, dimana 9 artikel tambahan dikeluarkan dengan berbagai alasan metodologis dan substansif: tiga artikel tidak menyajikan evaluasi arsitektur CNN secara komprehensif, satu artikel menggunakan metodologi yang tidak sesuai dengan standar *peer-review*, dan beberapa studi tidak menyediakan data empiris yang memadai untuk analisis. Proses seleksi akhir menghasilkan 6 artikel primer yang memenuhi seluruh kriteria inklusi dan eksklusi yang telah ditetapkan, dengan kualitas metodologi yang terverifikasi dan relevansi tinggi terhadap pertanyaan penelitian. Keenam artikel tersebut kemudian dianalisis secara mendalam untuk mengekstraksi informasi mengenai arsitektur CNN, karakteristik *dataset*, metrik performa, dan strategi implementasi yang digunakan dalam konteks sistem otonom. Gambar 1 merupakan flowchart PRISMA untuk proses seleksi literatur.

2.1.1 Analisis Kritis dan Sintesis Temuan

Berdasarkan sintesis mendalam terhadap keenam artikel primer, terdapat beberapa temuan kritis yang melampaui sekadar rangkuman individual dan membentuk pemahaman komprehensif tentang lanskap penelitian deteksi objek untuk sistem otonom.

Pertama, terdapat kesenjangan konsistensi pelaporan metrik evaluasi. Perbandingan lintas studi menghadapi hambatan signifikan karena heterogenitas metrik yang dilaporkan. [25] melaporkan *mean Average Precision* (mAP) secara eksplisit sebesar 0.86 pada *dataset* KITTI, tidak menyajikan nilai mAP spesifik, hanya menyebutkan "peningkatan signifikan". Kondisi ini mencerminkan absennya standardisasi protokol evaluasi dalam komunitas penelitian, yang menyulitkan pengambilan keputusan berbasis evidensi bagi praktisi sistem otonom. Penelitian mendatang perlu mengadopsi protokol pelaporan terpadu seperti COCO evaluation metrics secara konsisten [21] agar perbandingan arsitektur dapat dilakukan secara objektif.



Gambar 1. Flowchart PRISMA untuk Proses Seleksi Literatur

Kedua, terdapat bias temporal dan kontekstual dalam pemilihan *dataset*. Lima dari enam studi menggunakan *dataset* benchmark yang mapan seperti KITTI, MS COCO, CCTSDB, dan SODA-D, yang meskipun memiliki nilai standardisasi tinggi, tidak selalu mencerminkan kondisi operasional sistem otonom di berbagai konteks geografis dan kultural. Secara eksplisit menggunakan *dataset real-world* lokal dari Banda Aceh, Indonesia, yang menunjukkan bahwa performa model dapat bervariasi secara signifikan ketika diaplikasikan di luar konteks *dataset* pelatihan aslinya. Temuan ini mengindikasikan adanya *domain gap* yang belum sepenuhnya diatasi oleh literatur yang ada, dan menjadi peringatan kritis bahwa model dengan performa tinggi pada benchmark standar belum tentu generalizable pada kondisi lokal yang berbeda [16].

Ketiga, adanya trade-off yang belum terselesaikan antara akurasi dan efisiensi. Arsitektur dengan performa tertinggi seperti *AttenRetina* [25] tidak dilaporkan karakteristik efisiensinya secara eksplisit, sementara arsitektur yang paling efisien seperti *SSD MobileNv2* [15] tidak melaporkan nilai mAP absolut. Hal ini menciptakan celah pengetahuan yang kritis: tidak ada satu pun studi yang secara komprehensif memetakan *Pareto frontier* antara akurasi, kecepatan inferensi, dan konsumsi memori untuk spektrum arsitektur yang luas. Kondisi ini menyulitkan pengembang sistem otonom dalam membuat keputusan arsitektur yang *informed*.

Keempat, keterbatasan generalisasi pada kondisi *adverse weather*. Dari keenam studi, hanya [23] yang secara eksplisit mengeksplorasi performa model pada kondisi cuaca yang beragam melalui *synthetic data*. Lima studi lainnya tidak membahas robustness model terhadap kondisi seperti hujan deras, kabut, atau silau matahari langsung, yang merupakan kondisi kritis dalam operasional kendaraan otonom. Ini mengindikasikan bahwa sebagian besar arsitektur yang dievaluasi belum diuji secara menyeluruh untuk skenario *safety-critical*, sehingga klaim kemampuan *real-world deployment* perlu disikapi secara kritis [18].

Kelima, minimnya kajian tentang aspek keamanan dan *adversarial robustness*. Tidak satu pun dari keenam studi yang membahas ketahanan model terhadap serangan adversarial, seperti *adversarial patch* atau manipulasi input yang dapat mengecoh sistem deteksi. Dalam konteks sistem otonom yang *safety-critical*, kerentanan ini merupakan celah penelitian yang serius dan perlu mendapat perhatian lebih dalam penelitian berikutnya [13].

Sintesis *Systematic Literature Review* tentang algoritma deteksi objek untuk sistem dapat dilihat pada Tabel 1.

Tabel 1. Sintesis *Literature Systematic Review* tentang Algoritma Deteksi Objek untuk Sistem

No	Author	Judul	Metode	Sample	Teknik Analisa Data	Uji Hipotesis	Hasil Temuan Peneliti	Relevansi dengan Topik
1	Yan et al., 2025	FNon R-CNN: A Multi-scale Ground Object Detection and Recognition Network	Improved Faster R-CNN dengan integrasi Non-Local Blocks, Feature Pyramid Network (FPN), dan dynamic loss function	SODA-D dataset	Evaluasi detection accuracy dengan fokus pada multi-scale object detection	Pengujian performa model pada objek dengan berbagai skala, khususnya objek kecil	Model menunjukkan peningkatan signifikan dalam akurasi deteksi, terutama untuk objek kecil. FNon R-CNN berhasil meningkatkan global context modeling dan multi-level feature fusion	Sangat relevan - mengevaluasi arsitektur CNN (Faster R-CNN) untuk deteksi objek multi-skala dalam sistem transportasi cerdas, mengatasi tantangan dataset dengan objek kecil
2	Zou & Liu, 2025	Small object detection algorithm based on improved YOLOv10 for traffic sign	LTS-YOLOv10 dengan Omni-Dimensional Dynamic Convolution (ODConv), attention-guided bidirectional feature pyramid network (EMA-BiFPN), dan MPDIoU loss function	CCTSDb, TT100K, dan DFG datasets	Evaluasi mAP (mean Average Precision), precision, dan recall untuk deteksi rambu lalu lintas	Perbandingan performa dengan YOLOv10 original pada tiga dataset publik	LTS-YOLOv10 mencapai peningkatan 3.8% mAP pada dataset CCTSDb dibanding YOLOv10 original, dengan peningkatan notable pada TT100K dan DFG. Model menunjukkan akurasi, robustness, dan real-time performance yang superior	Sangat relevan - evaluasi arsitektur CNN (YOLOv10) untuk sistem autonomous driving, mengatasi tantangan deteksi objek kecil dengan <i>complex backgrounds</i> , strategi implementasi dengan parameter efisien
3	Aulia et al., 2024	A new CNN-BASED object detection system for autonomous mobile robots based on <i>real-world</i> vehicle datasets	<i>SSD MobileNetV2</i> FPN Lite 320x320 dengan retraining menggunakan <i>real-world</i> datasets	Dataset original dari Banda Aceh city dengan berbagai kondisi pencahayaan (pagi, siang, malam)	Evaluasi kuantitatif dan kualitatif menggunakan performance indicators untuk classification dan detection accuracy	Pengujian model pada kondisi pencahayaan berbeda dan objek multiple (mobil, motor, orang, becak)	Model pretrained yang dilatih ulang menunjukkan peningkatan akurasi klasifikasi dan deteksi. Model memiliki potensi untuk diimplementasikan pada AMR dengan inference time cepat dan ukuran compact	Sangat relevan - implementasi arsitektur CNN (<i>SSD MobileNet</i>) untuk autonomous mobile robots, mengatasi tantangan dataset <i>real-world</i> dengan variasi kondisi lingkungan, strategi deployment pada <i>embedded systems</i>
4	C. Li et al., 2025	A multi-objective dynamic detection	<i>YOLOv8</i> dengan <i>MobileNetV3_CA</i> network (integrasi <i>MobileNetV3</i> dan	Multiple autonomous driving datasets	Evaluasi performa multi-target detection	Perbandingan dengan <i>YOLOv8</i> traditional	Model yang mengkombinasikan attention convolution	Sangat relevan - evaluasi arsitektur

No	Author	Judul	Metode	Sample	Teknik Analisa Data	Uji Hipotesis	Hasil Temuan Peneliti	Relevansi dengan Topik
		model in autonomous driving based on an improved YOLOv8	Coordinate Attention), mosaic image enhancement, dan PANet untuk multi-scale feature fusion		dengan metrik detection accuracy dan specificity	dan variannya tanpa attention mechanism	dengan YOLOv8 menunjukkan peningkatan signifikan dalam multi-target detection. Mosaic enhancement efektif untuk small targets dan complex scenes	CNN lightweight (YOLOv8 + MobileNetV3) untuk autonomous driving, mengatasi tantangan deteksi multiple targets dalam complex traffic scenes, strategi implementasi dengan attention mechanism
5	Liu et al., 2025	Object detection algorithm for autonomous driving: Design and real-time performance analysis of <i>AttenRetina</i> model	<i>AttenRetina</i> model dengan multi-scale feature fusion module (FPN) dan attention mechanism	KITTI dan MS COCO <i>datasets</i>	Evaluasi menggunakan Precision, Recall, dan mAP sebagai key indicators	Perbandingan performa dengan mainstream models lainnya	<i>AttenRetina</i> mencapai mAP 0.86 pada KITTI <i>dataset</i> , lebih tinggi 12% dari basic model. Model menunjukkan performa excellent dalam mendeteksi objek dengan <i>complex backgrounds</i> dan occlusions	Sangat relevan - evaluasi arsitektur CNN (RetinaNet) dengan <i>attention mechanism</i> untuk autonomous driving, penggunaan <i>dataset</i> benchmark (KITTI, MS COCO), strategi optimisasi untuk meningkatkan accuracy dan robustness
6	V. Li et al., 2025	Enhancing 3D object detection in autonomous vehicles based on synthetic virtual environment analysis	3D object detection dengan augmented reality (AR) ecosystems menggunakan AI model untuk 3D bounding boxes	Synthetic <i>dataset</i> dengan simulasi berbagai kondisi environmental, lighting, dan spatiotemporal	Evaluasi detection performance pada berbagai kondisi challenging scenarios dengan fokus pada processing time dan accuracy	Pengujian pada kondisi cuaca berbeda dan variasi camera settings	Model menunjukkan peningkatan signifikan dalam processing time sambil mempertahankan accuracy. Integrasi 3D detection dengan AR <i>framework</i> dan synthetic data efektif mengatasi kompleksitas environmental	Relevan - eksplorasi 3D object detection untuk autonomous vehicles, mengatasi tantangan <i>dataset</i> dengan synthetic data untuk kondisi environmental yang beragam, strategi implementasi dengan AR ecosystem untuk real-time processing

4.1.1 Analisis Bibliometrik

Untuk memperkuat analisis struktural dan memetakan lanskap penelitian secara komprehensif, dilakukan analisis bibliometrik menggunakan pendekatan yang merujuk pada prinsip-prinsip *bibliometric analysis* yang direkomendasikan dalam SLR kontemporer. Analisis ini mencakup identifikasi tren topik, pola kolaborasi, dan kluster penelitian berdasarkan keenam artikel primer beserta referensi yang dikutip di dalamnya.

Berdasarkan pemetaan kata kunci (*co-occurrence analysis*) dari seluruh artikel yang diidentifikasi dalam proses pencarian, terdapat tiga kluster tema utama yang mendominasi penelitian deteksi objek untuk sistem otonom dalam periode 2020–2025. Kluster pertama berpusat pada arsitektur CNN dan optimisasi model, dengan kata kunci yang sering muncul bersama meliputi "YOLO", "feature pyramid network", "attention mechanism", dan "lightweight architecture". Kluster kedua berkaitan dengan tantangan dataset dan augmentasi, mencakup "small object detection", "data augmentation", "synthetic data", dan "domain adaptation". Kluster ketiga menyangkut implementasi dan deployment, dengan kata kunci "embedded systems", "edge computing", "real-time inference", dan "knowledge distillation".

Analisis distribusi temporal menunjukkan peningkatan signifikan dalam publikasi yang mengintegrasikan *attention mechanism* dengan arsitektur *lightweight* sejak tahun 2022, yang mengindikasikan pergeseran paradigma dari optimisasi akurasi semata menuju keseimbangan akurasi-efisiensi. Tren ini konsisten dengan meningkatnya kebutuhan deployment pada perangkat edge dengan sumber daya terbatas [3]. Pola kolaborasi (*co-authorship*) dalam literatur yang dikaji menunjukkan dominasi institusi dari Asia Timur, khususnya Tiongkok, dalam publikasi terkait YOLOv8 dan YOLOv10, sementara penelitian tentang implementasi *lightweight* pada konteks lokal masih relatif terbatas, yang membuka peluang *novelty* bagi peneliti dari kawasan Asia Tenggara. Penggunaan alat bibliometrik seperti VOSviewer atau *bibliometrix* dalam R pada penelitian lanjutan sangat direkomendasikan untuk menghasilkan visualisasi jaringan pengetahuan yang lebih komprehensif, termasuk analisis *citation*, *co-citation*, dan *bibliographic coupling* yang dapat mempercepat identifikasi celah penelitian.

3.2. Evaluasi Arsitektur CNN untuk Deteksi Objek

Analisis terhadap enam studi yang diseleksi mengungkapkan diversitas arsitektur CNN yang diimplementasikan untuk sistem deteksi objek otonom, dengan variasi pendekatan yang disesuaikan dengan kebutuhan aplikasi spesifik. Penelitian Yan et al. (2025) mengembangkan FNon R-CNN yang merupakan pengembangan dari arsitektur Faster R-CNN dengan integrasi *Non-Local Blocks* dan *Feature Pyramid Network* untuk meningkatkan kemampuan deteksi multi-skala. Model ini menunjukkan peningkatan signifikan dalam *global context modeling* dan *multi-level feature fusion*, yang sangat krusial untuk mendeteksi objek dengan ukuran bervariasi dalam sistem transportasi cerdas. Implementasi *dynamic loss function* pada FNon R-CNN terbukti efektif dalam mengatasi ketidakseimbangan skala objek pada dataset SODA-D, menghasilkan akurasi deteksi yang superior khususnya untuk objek berukuran kecil yang seringkali menjadi tantangan dalam skenario *real-world*.

Pendekatan berbeda diterapkan oleh Zou & Liu (2025) yang mengembangkan LTS-YOLOv10 dengan fokus pada deteksi rambu lalu lintas berukuran kecil dengan *complex backgrounds*. Arsitektur ini mengintegrasikan *Omni-Dimensional Dynamic Convolution* (ODConv) dan *attention-guided bidirectional feature pyramid network* (EMA-BiFPN) yang memungkinkan model untuk menangkap fitur multi-skala dengan lebih efisien. Hasil eksperimen menunjukkan peningkatan 3.8% dalam *mean Average Precision* (mAP) pada dataset CCTSDB dibandingkan dengan YOLOv10 original, dengan performa notable pada dataset TT100K dan DFG. Keunggulan LTS-YOLOv10 terletak pada kemampuannya mempertahankan *real-time performance* sambil meningkatkan akurasi deteksi, menjadikannya sangat applicable untuk implementasi pada sistem *autonomous driving* yang memerlukan respons cepat dan presisi tinggi.

Dalam konteks *autonomous mobile robots*, Aulia et al. (2024) mendemonstrasikan efektivitas arsitektur SSD MobileNetV2 FPN Lite untuk aplikasi dengan keterbatasan sumber daya komputasi. Model ini dilatih menggunakan dataset *real-world* dari Banda Aceh dengan variasi kondisi pencahayaan yang mencakup pagi, siang, dan malam hari. Hasil penelitian menunjukkan bahwa model *pretrained* yang dilatih ulang dengan dataset lokal menghasilkan peningkatan signifikan dalam akurasi klasifikasi dan deteksi untuk berbagai objek termasuk mobil, motor, orang, dan becak. Arsitektur *lightweight* ini memiliki keunggulan dalam *inference time* yang cepat dan ukuran model yang kompak, membuatnya ideal untuk deployment pada *embedded systems* dengan kapasitas memori dan daya komputasi terbatas.

C. Li et al. (2025) mengeksplorasi integrasi YOLOv8 dengan MobileNetV3_CA network yang mengombinasikan *attention convolution* dengan arsitektur *lightweight* untuk deteksi multi-objek dalam skenario *autonomous driving*. Penggunaan *Coordinate Attention mechanism* memungkinkan model untuk fokus pada region yang relevan sambil mempertahankan efisiensi komputasi. Teknik *mosaic image enhancement* yang diimplementasikan terbukti sangat efektif untuk meningkatkan deteksi *small targets* dan *complex scenes*, yang merupakan karakteristik umum dalam lingkungan lalu lintas urban. Model ini menunjukkan peningkatan signifikan dalam *multi-target detection accuracy* dan *specificity* dibandingkan dengan YOLOv8 tradisional tanpa *attention mechanism*, memvalidasi pentingnya mekanisme perhatian dalam meningkatkan performa deteksi pada kondisi challenging.

Liu et al. (2025) mengembangkan *AttenRetina* model yang mengintegrasikan *multi-scale feature fusion module* dengan *attention mechanism* pada arsitektur RetinaNet. Evaluasi pada dataset benchmark KITTI dan MS COCO menunjukkan bahwa *AttenRetina* mencapai mAP 0.86 pada KITTI dataset, yang merupakan peningkatan 12% dari *basic model*. Model ini menunjukkan performa excellent dalam mendeteksi

objek dengan *complex backgrounds* dan oklusi, yang merupakan tantangan utama dalam aplikasi *autonomous driving*. Kemampuan model untuk mempertahankan akurasi tinggi pada kondisi *challenging* mengindikasikan *robustness* yang superior, menjadikannya kandidat kuat untuk implementasi pada sistem produksi yang memerlukan reliabilitas tinggi.

Eksplorasi inovatif dilakukan oleh V. Li et al. (2025) yang mengintegrasikan deteksi objek 3D dengan *augmented reality (AR) ecosystems* menggunakan *dataset* sintetis. Pendekatan ini mengatasi keterbatasan *dataset real-world* dengan mensimulasikan berbagai kondisi lingkungan, pencahayaan, dan skenario spatiotemporal. Model menunjukkan peningkatan signifikan dalam *processing time* sambil mempertahankan akurasi deteksi, membuktikan bahwa integrasi 3D detection dengan *AR framework* dan *synthetic data* efektif untuk mengatasi kompleksitas lingkungan yang beragam. Strategi ini membuka peluang baru dalam pengembangan sistem deteksi objek yang *robust* tanpa ketergantungan penuh pada *dataset real-world* yang mahal dan sulit diperoleh.

Tabel 2. Perbandingan Performa Kuantitatif Arsitektur CNN untuk Deteksi Objek pada Sistem Otonom

Peneliti	Arsitektur	Dataset	mAP/ Akurasi	Peningkatan vs Baseline	Karakteristik Utama
Yan et al. (2025)	FNon R-CNN	SODA-D	Tidak disebutkan spesifik	Signifikan pada objek kecil	Multi-scale feature fusion, Non-Local Blocks
Zou & Liu (2025)	LTS-YOLOv10	CCTSDB, TT100K, DFG	+3.8% mAP (CCTSDB)	3.8% dibanding YOLOv10	ODConv, EMA-BiFPN, real-time performance
Aulia et al. (2024)	SSD MobileNetv2 FPN Lite	Real-world Banda Aceh	Akurasi klasifikasi tinggi	Peningkatan signifikan dengan retraining	Lightweight, inference time cepat, compact size
C. Li et al. (2025)	YOLOv8-MobileNetV3_CA	Multiple autonomous driving	Tidak disebutkan spesifik	Signifikan dalam multi-target	Coordinate Attention, mosaic enhancement
Liu et al. (2025)	AttenRetina	KITTI, MS COCO	mAP 0.86 (KITTI)	+12% dari basic model	Multi-scale fusion, attention mechanism
V. Li et al. (2025)	3D Detection + AR	Synthetic dataset	Tidak disebutkan spesifik	Peningkatan processing time	AR ecosystem, synthetic data simulation

Tabel 2 menyajikan rangkuman kuantitatif performa berbagai arsitektur CNN yang dievaluasi dalam studi yang dikaji. *AttenRetina* model menunjukkan performa tertinggi dengan mAP 0.86 pada *dataset* KITTI, mencerminkan efektivitas integrasi *attention mechanism* dengan *multi-scale feature fusion*. LTS-YOLOv10 mendemonstrasikan konsistensi performa dengan peningkatan terukur 3.8% pada *dataset* CCTSDB, yang mengindikasikan *robustness* model untuk deteksi objek kecil pada *complex backgrounds*. *SSD MobileNetv2 FPN Lite* menawarkan keseimbangan optimal antara akurasi dan efisiensi untuk implementasi pada *embedded systems*, meskipun nilai mAP spesifik tidak dilaporkan dalam publikasi. Variasi dalam metrik pelaporan antarsudi mencerminkan heterogenitas dalam metodologi evaluasi, yang menjadi tantangan dalam melakukan perbandingan langsung antararsitektur pada konteks aplikasi yang berbeda.

3.3. Tantangan Dataset dalam Deteksi Objek untuk Sistem Otonom

Analisis komprehensif terhadap literatur yang dikaji mengidentifikasi beberapa tantangan kritis yang berkaitan dengan karakteristik *dataset* yang digunakan dalam pelatihan model deteksi objek. Objek berukuran kecil muncul sebagai tantangan konsisten yang dihadapi oleh mayoritas penelitian, di mana deteksi yang akurat terhadap objek dengan dimensi pixel terbatas memerlukan pendekatan arsitektural khusus. Yan et al. (2025) dan Zou & Liu (2025) secara eksplisit mengatasi masalah ini melalui implementasi *multi-scale feature pyramid networks* yang memungkinkan model untuk menangkap informasi dari berbagai resolusi. Keberhasilan FNon R-CNN dalam meningkatkan deteksi objek kecil pada *dataset* SODA-D dan performa superior LTS-YOLOv10 pada rambu lalu lintas kecil mendemonstrasikan efektivitas pendekatan *multi-scale feature extraction* dalam mengatasi keterbatasan ini.

Complex backgrounds dan oklusi merupakan tantangan signifikan lainnya yang memengaruhi performa model deteksi objek. Liu et al. (2025) menunjukkan bahwa *AttenRetina* model dengan *attention mechanism* mampu mengatasi kompleksitas latar belakang dan oklusi parsial dengan lebih efektif dibandingkan model konvensional. Kemampuan model untuk memfokuskan komputasi pada region yang relevan sambil mengabaikan noise visual dari background yang kompleks menjadi kunci dalam meningkatkan akurasi deteksi pada kondisi *real-world*. Temuan ini menggarisbawahi pentingnya mekanisme perhatian dalam arsitektur modern untuk aplikasi sistem otonom yang beroperasi dalam lingkungan dinamis dan tidak terprediksi.

Variasi kondisi pencahayaan dan environmental merupakan faktor krusial yang mempengaruhi generalisasi model. Aulia et al. (2024) secara eksplisit mengatasi tantangan ini dengan melatih model menggunakan *dataset* yang mencakup berbagai kondisi pencahayaan dari pagi hingga malam hari. Pendekatan ini menghasilkan model yang lebih robust terhadap perubahan kondisi lighting, mengurangi degradasi performa yang sering terjadi ketika model diaplikasikan pada kondisi berbeda dari data pelatihan. V. Li et al. (2025) mengambil pendekatan komplementer dengan menggunakan *synthetic data* yang mensimulasikan berbagai kondisi lingkungan, menawarkan solusi scalable untuk meningkatkan diversitas *dataset* tanpa memerlukan pengumpulan data *real-world* yang ekstensif.

Class imbalance dalam *dataset* menjadi concern yang implisit namun signifikan, khususnya dalam konteks deteksi objek kecil yang sering underrepresented dalam *dataset* standar. Implementasi *dynamic loss function* pada FNon R-CNN dan *MPDIoU loss function* pada LTS-YOLOv10 merupakan strategi untuk mengatasi ketidakseimbangan ini dengan memberikan bobot yang lebih tinggi pada kelas yang sulit atau jarang muncul. Pendekatan berbasis loss function ini terbukti efektif dalam meningkatkan sensitivitas model terhadap objek yang challenging, meningkatkan performa overall pada distribusi kelas yang tidak seimbang.

3.4. Strategi Implementasi untuk Sistem dengan Sumber Daya Terbatas

Implementasi algoritma deteksi objek pada platform dengan keterbatasan komputasi memerlukan strategi optimisasi yang sofisticated untuk mencapai keseimbangan antara akurasi dan efisiensi. Aulia et al. (2024) mendemonstrasikan keberhasilan implementasi *SSD MobileNetV2 FPN Lite* pada *autonomous mobile robots*, dimana arsitektur *lightweight* ini menghasilkan *inference time* yang cepat dengan ukuran model yang kompak. Strategi ini memungkinkan *deployment* pada *embedded systems* seperti Raspberry Pi atau Jetson Nano tanpa memerlukan hardware akselerator eksternal yang mahal. Penggunaan *MobileNet* sebagai *backbone* mengurangi jumlah parameter model secara signifikan sambil mempertahankan kemampuan ekstraksi fitur yang memadai, menjadikannya solusi praktis untuk aplikasi *edge computing*.

Integrasi *attention mechanism* dengan arsitektur *lightweight* yang diimplementasikan oleh C. Li et al. (2025) menawarkan pendekatan inovatif untuk meningkatkan efisiensi komputasi. *Coordinate Attention* dalam *YOLOv8-MobileNetV3_CA* memungkinkan model untuk mengalokasikan sumber daya komputasi secara selektif pada region yang informatif, mengurangi computational overhead tanpa mengorbankan akurasi deteksi. Strategi ini particularly beneficial untuk sistem yang beroperasi dengan batasan daya dan thermal constraints, seperti *autonomous drones* atau *wearable devices* yang memerlukan operasi berkelanjutan dengan konsumsi energi minimal.

Pendekatan *multi-scale feature fusion* yang diimplementasikan oleh Liu et al. (2025) dan Yan et al. (2025) menunjukkan bahwa kompleksitas arsitektural tidak selalu berkorelasi negatif dengan efisiensi implementasi. Dengan desain yang cermat, *Feature Pyramid Networks* dapat dioptimasi untuk memproses fitur pada berbagai skala secara paralel, memanfaatkan capabilities modern GPU atau TPU yang dirancang untuk operasi parallel processing. Implementasi yang efisien dari FPN memungkinkan sistem untuk mencapai *real-time performance* pada hardware mid-range, memperluas aksesibilitas teknologi deteksi objek tingkat lanjut untuk aplikasi dengan budget terbatas.

3.5. Rumusan Hipotesis

Berdasarkan analisis komprehensif terhadap literatur yang dikaji, beberapa hipotesis dapat dirumuskan untuk penelitian lanjutan dalam domain deteksi objek untuk sistem otonom. Pertama, integrasi *attention mechanism* dengan arsitektur CNN *lightweight* secara signifikan meningkatkan efisiensi komputasi sambil mempertahankan atau meningkatkan akurasi deteksi dibandingkan dengan arsitektur tradisional tanpa mekanisme perhatian. Kedua, penggunaan *multi-scale feature pyramid networks* menghasilkan performa superior dalam mendeteksi objek dengan variasi ukuran yang signifikan, khususnya untuk objek kecil yang challenging. Ketiga, model yang dilatih menggunakan kombinasi *dataset real-world* dengan variasi kondisi pencahayaan dan *synthetic data* menunjukkan generalisasi yang lebih baik dibandingkan model yang hanya dilatih pada *dataset* homogen. Keempat, implementasi *dynamic loss functions* yang mengadaptasi bobot berdasarkan kesulitan deteksi objek efektif mengatasi *class imbalance* dan meningkatkan sensitivitas model terhadap kelas yang underrepresented. Kelima, arsitektur berbasis *MobileNet* dengan optimisasi *FPN Lite* memungkinkan *deployment* pada *embedded systems* dengan *inference time* di bawah 100 ms sambil mempertahankan akurasi deteksi di atas 85% mAP pada *dataset* standar.

4.1 KESIMPULAN

Tinjauan sistematis terhadap enam studi primer mengungkapkan bahwa perkembangan arsitektur CNN untuk deteksi objek pada sistem otonom telah mencapai kemajuan signifikan dalam mengatasi tantangan akurasi dan efisiensi komputasi. Arsitektur *hybrid* yang mengintegrasikan mekanisme *attention* dengan *feature pyramid networks* menunjukkan performa superior, khususnya *AttenRetina* yang mencapai mAP 0.86 pada *dataset* KITTI dengan peningkatan 12% dari model dasar. Implementasi *lightweight architectures* seperti *SSD MobileNetV2*, *FPN Lite* dan *YOLOv8-MobileNetV3* membuktikan bahwa

keseimbangan antara akurasi tinggi dan kecepatan inferensi dapat tercapai untuk *deployment* pada *embedded systems*. Tantangan *dataset* yang teridentifikasi meliputi deteksi objek kecil, *complex backgrounds* dengan oklusi parsial, serta variasi kondisi pencahayaan yang ekstrem, dimana solusi melalui *multi-scale feature extraction* dan *dynamic loss functions* terbukti efektif. Strategi implementasi dengan *attention mechanism* selektif dan arsitektur modular memungkinkan sistem beroperasi pada perangkat dengan keterbatasan sumber daya sambil mempertahankan performa *real-time*. Penggunaan *synthetic data* dan *augmented reality ecosystems* membuka peluang baru dalam mengatasi keterbatasan *dataset real-world* tanpa mengorbankan kualitas pelatihan model. Temuan ini mengonfirmasi bahwa pemilihan arsitektur harus disesuaikan dengan konteks aplikasi spesifik, dengan mempertimbangkan *trade-off* antara kompleksitas model, akurasi deteksi, dan kapabilitas *hardware* yang tersedia untuk mencapai implementasi optimal pada sistem otonom.

Penelitian mendatang perlu mengeksplorasi integrasi *transformer-based architectures* dengan CNN untuk meningkatkan kemampuan global *context modeling* sambil mempertahankan efisiensi komputasi pada *edge devices*. Pengembangan *dataset* yang lebih komprehensif dengan representasi seimbang untuk objek kecil dan kondisi pencahayaan ekstrem menjadi prioritas untuk meningkatkan generalisasi model pada skenario *real-world* yang beragam. Implementasi teknik *knowledge distillation* dan *neural architecture search* dapat dimanfaatkan untuk mengoptimalkan arsitektur *lightweight* tanpa mengorbankan akurasi deteksi secara signifikan. Standardisasi metrik evaluasi dan protokol pengujian yang konsisten diperlukan untuk memfasilitasi perbandingan objektif antar-arsitektur pada berbagai konteks aplikasi sistem otonom. Kolaborasi antara peneliti dan industri dalam mengembangkan *benchmark datasets* yang mencerminkan kompleksitas lingkungan operasional aktual akan mempercepat transisi dari penelitian akademis ke implementasi praktis. Eksplorasi lebih lanjut terhadap *synthetic data generation* dengan domain *randomization* dapat mengurangi ketergantungan pada pengumpulan data *real-world* yang mahal dan *time-consuming*. Pengembangan *framework* yang mempertimbangkan aspek *interpretability* dan *explainability* dalam model deteksi objek menjadi penting untuk membangun kepercayaan dan akseptabilitas sistem otonom dalam aplikasi *safety-critical* seperti *autonomous driving* dan *industrial robotics*.

REFERENSI

- [1] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3354–3361. doi: 10.1109/CVPR.2012.6248074.
- [2] M. Cordts *et al.*, "The Cityscapes Dataset for Semantic Urban Scene Understanding," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 3213–3223. doi: 10.1109/CVPR.2016.350.
- [3] F. Yu *et al.*, "BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 2636–2645. doi: 10.1109/CVPR42600.2020.00271.
- [4] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO Series in 2021," *arXiv*, 2021.
- [5] E. Liang, D. Wei, F. Li, H. Lv, and S. Li, "Object detection model of vehicle-road cooperative autonomous driving based on improved YOLO11 algorithm," *Sci. Rep.*, vol. 15, no. 1, pp. 1–14, 2025, doi: 10.1038/s41598-025-18263-9.
- [6] M. Tan, R. Pang, and Q. V Le, "EfficientDet: Scalable and Efficient Object Detection," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 10781–10790. doi: 10.1109/CVPR42600.2020.01079.
- [7] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv*, 2020.
- [8] J. Cao, B. Peng, M. Gao, H. Hao, X. Li, and H. Mou, "Object Detection Based on CNN and Vision-Transformer: A Survey," *IET Comput. Vis.*, 2025.
- [9] A. Balasubramaniam and S. Pasricha, "Object Detection in Autonomous Vehicles: Status and Open Challenges," pp. 1–6, 2022.
- [10] W. M. Lim, "What Is Qualitative Research? An Overview and Guidelines," *Australas. Mark. J.*, vol. 33, no. 2, pp. 199–229, 2025, doi: 10.1177/14413582241264619.
- [11] W. Liu *et al.*, "SSD: Single Shot MultiBox Detector," in *European Conference on Computer Vision*, 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0_2.
- [12] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 7464–7475.
- [13] I. Rodriguez-Conde, C. Campos, and F. Fdez-Riverola, *Optimized convolutional neural network architectures for efficient on-device vision-based object detection*, vol. 34, no. 13. Springer London, 2022. doi: 10.1007/s00521-021-06830-w.
- [14] C. Li, Y. Zhu, and M. Zheng, "A multi-objective dynamic detection model in autonomous driving based on an improved YOLOv8," *Alexandria Eng. J.*, vol. 122, pp. 453–464, 2025, doi:

- 10.1016/j.aej.2025.03.020.
- [15] U. Aulia, I. Hasanuddin, M. Dirhamsyah, and N. Nasaruddin, "A new CNN-BASED object detection system for autonomous mobile robots based on real-world vehicle datasets," *Heliyon*, vol. 10, no. 15, p. e35247, 2024, doi: 10.1016/j.heliyon.2024.e35247.
- [16] M. Vaithianathan, "Real-Time Object Detection and Recognition in FPGA-Based Autonomous Driving Systems," *Int. J. Comput. Trends Technol.*, vol. 72, no. 4, pp. 145–152, 2024, doi: 10.14445/22312803/ijctt-v72i4p119.
- [17] A. S. Nagiub, S. Ghoniemy, M. Fayez, and H. Khaled, "3D Object Detection for Autonomous Driving: A Comprehensive Review," in *6th International Conference on Computer and Informatics*, 2024, pp. 268–278. doi: 10.1109/ICCI61671.2024.10485120.
- [18] S. Franceschini, M. M. Autorino, M. Ambrosanio, V. Pascazio, and F. Baselice, "An Ultrasound Prototype for Remote Hand Movement Sensing: The Finger Tapping Case," *Sensors*, vol. 25, no. 1, pp. 1–10, 2025, doi: 10.3390/s25010123.
- [19] Z. Yan, "FNon R-CNN: A multi-scale ground object detection and recognition network," *High-Confidence Comput.*, p. 100359, 2025, doi: 10.1016/j.hcc.2025.100359.
- [20] Y. Zou and S. Liu, "Small object detection algorithm based on improved YOLOv10 for traffic sign," *Transp. Res. Interdiscip. Perspect.*, vol. 32, p. 101501, 2025, doi: 10.1016/j.trip.2025.101501.
- [21] T.-Y. Lin *et al.*, "Microsoft COCO: Common Objects in Context," in *European Conference on Computer Vision*, 2014, pp. 740–755. doi: 10.1007/978-3-319-10602-1_48.
- [22] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in *Advances in Neural Information Processing Systems*, 2015.
- [23] V. Li *et al.*, "Enhancing 3D object detection in autonomous vehicles based on synthetic virtual environment analysis," *Image Vis. Comput.*, vol. 154, no. December 2024, p. 105385, 2025, doi: 10.1016/j.imavis.2024.105385.
- [24] M. J. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *Bmj*, vol. 372, 2021, doi: 10.1136/bmj.n71.
- [25] G. Liu, W. Jiang, C. Sun, N. Ning, R. Wang, and A. Buhari, "Object detection algorithm for autonomous driving: Design and real-time performance analysis of AttenRetina model," *Alexandria Eng. J.*, vol. 123, no. February, pp. 392–402, 2025, doi: 10.1016/j.aej.2025.02.063.