



# Deteksi dan Klasifikasi Kendaraan Berbasis Algoritma *You Only Look Once (Yolov7)*

Yusuf Kautsar Rohiman, Bulkis Kanata, L Ahmad S Irfan Akbar\*

Fakultas Teknik, Program Studi Teknik Elektro, Universitas Mataram, Mataram, Indonesia

Email: <sup>1</sup>yusufkautsarr@gmail.com, <sup>2</sup>uqjkanata@unram.ac.id, <sup>3</sup>irfan@unram.ac.id

Email Penulis Korespondensi: yusufkautsarr@gmail.com

**Abstrak**—Kepadatan lalu lintas di Indonesia menimbulkan kebutuhan akan sistem pendeteksian kendaraan yang akurat untuk mendukung perencanaan infrastruktur. Penelitian ini bertujuan mengimplementasikan algoritma YOLOv7 untuk mendeteksi dan mengklasifikasikan berbagai jenis kendaraan pada citra lalu lintas. Metode yang digunakan meliputi pelatihan model menggunakan Google Colab dengan dataset Kaggle berjumlah 6.633 gambar, batch size 1, sebanyak 19 epoch, dan optimisasi dengan algoritma Stochastic Gradient Descent (SGD). Hasil pelatihan menunjukkan model mampu mencapai precision sebesar 93,22%, recall 90,64%, mAP@0.5 sebesar 94,27%, dan mAP@0.5:0.95 sebesar 69,19%, dengan waktu pelatihan selama 1 jam. Kesimpulannya, algoritma YOLOv7 efektif dalam mendeteksi dan mengklasifikasikan kendaraan, namun untuk meningkatkan performa lebih lanjut, disarankan menggunakan jumlah epoch yang lebih banyak dalam pelatihan.

**Kata Kunci:** YOLOv7; Deteksi Kendaraan; Klasifikasi Lalu Lintas; Deep Learning; Google Colab; Kaggle Dataset; Deteksi Objek

**Abstract**—The increasing traffic density in Indonesia highlights the need for an accurate vehicle detection system to support infrastructure planning. This study aims to implement the YOLOv7 algorithm for detecting and classifying various types of vehicles in traffic images. The method involves training the model using Google Colab on a Kaggle dataset consisting of 6,633 images, with a batch size of 1, 19 training epochs, and optimization using the Stochastic Gradient Descent (SGD) algorithm. The training results show that the model achieved a precision of 93.22%, recall of 90.64%, mAP@0.5 of 94.27%, and mAP@0.5:0.95 of 69.19%, with a total training time of 1 hours. In conclusion, the YOLOv7 algorithm is effective for vehicle detection and classification, although increasing the number of training epochs is recommended to further enhance model performance.

**Keywords:** YOLOv7; Vehicle Detection; Traffic Classification; Deep Learning; Google Colab; Kaggle Dataset; Object Detection

## 1. PENDAHULUAN

Negara Indonesia mempunyai kepadatan penduduk yang sangat padat, terutama dikota-kota besar yang dimana jalan selalu dipadati oleh berbagai jenis kendaraan. Pada jam sibuk banyaknya kendaraan yang membuat kemacetan di jalan. Oleh karena itu dibutuhkan pembangunan pelebaran jalan untuk menampung kendaraan yang dipadati oleh berbagai jenis kendaraan yang melintas. Agar pembangunan pelebaran jalan yang tepat pada lokasi yang sering terjadinya kepadatan, maka dibutuhkan sistem pendeteksian jenis-jenis kendaraan yang melintas di jalan raya [1]

Masalah mengenai lalu lintas menjadi salah satu permasalahan yang paling menantang dan sulit dalam melakukan manajemen kota terutama pada negara berkembang. Secara global, kemacetan lalu lintas menjadi permasalahan yang serius. Permasalahan ini menjadi menarik perhatian dari beberapa kalangan seperti ahli tata kota, pemerintahan, teknisi, dan para peneliti untuk mencari solusi mengenai kemacetan [2]. Sebagian kota di Indonesia sudah menggunakan CCTV (Closed Control Television) untuk melakukan pemantauan arus lalu lintas. Dalam melakukan proses pemantauan arus lalu lintas, CCTV akan melakukan ekstraksi informasi dari gambar seperti kecepatan laju kendaraan, kemacetan, bentuk dan jenis kendaraan, nomor kendaraan, dan pelanggaran atau kecelakaan lalu lintas [3]

Beberapa tahun belakangan, penelitian mengenai kecerdasan buatan seperti deteksi objek dapat memudahkan para peneliti untuk mengenali objek yang terdapat pada sebuah gambar. Deteksi objek adalah salah satu bidang pada computer vision. Computer vision merupakan ilmu yang mempelajari bagaimana komputer dapat melihat serta melakukan analisa pada suatu objek yang terdapat pada sebuah gambar.

Penelitian tentang pemantauan lalu lintas sudah banyak diusulkan oleh para peneliti dengan menggunakan beberapa pendekatan, seperti yang diusulkan oleh [4] penelitian menggunakan model Histogram of Oriented Gradient (HOG) dengan Support Machine Vector (SVM). Penelitian dengan menggunakan metode deep learning juga diusulkan oleh beberapa peneliti untuk melakukan pemantauan lalu lintas dengan menggunakan pendekatan seperti, Faster R-CNN [5], Convolutional Neural Network (CNN) [6], serta You Only Look Once (YOLO) [7].

Citra digital merupakan suatu matriks dimana indeks baris dan kolomnya menyatakan suatu titik pada citra tersebut dan elemen matriksnya (yang disebut sebagai elemen gambar atau piksel) menyatakan tingkat keabuan pada titik tersebut. Untuk sebuah citra digital, setiap pixel memiliki nilai integer yakni gray level yang menunjukkan amplitudo atau intensitas dari piksel tersebut. Citra merupakan fungsi dua dimensi yang kedua variabelnya yaitu nilai amplitudo dan koordinatnya merupakan nilai integer [8].

Deteksi objek adalah teknik dari computer vision dalam pembacaan pengenalan objek pada gambar ataupun video. Deteksi objek bertujuan untuk menemukan dan mengklasifikasi, karena deteksi objek dapat mendeteksi lebih dari satu objek sekaligus serta memberikan informasi tentang lokasi objek dalam gambar [9].

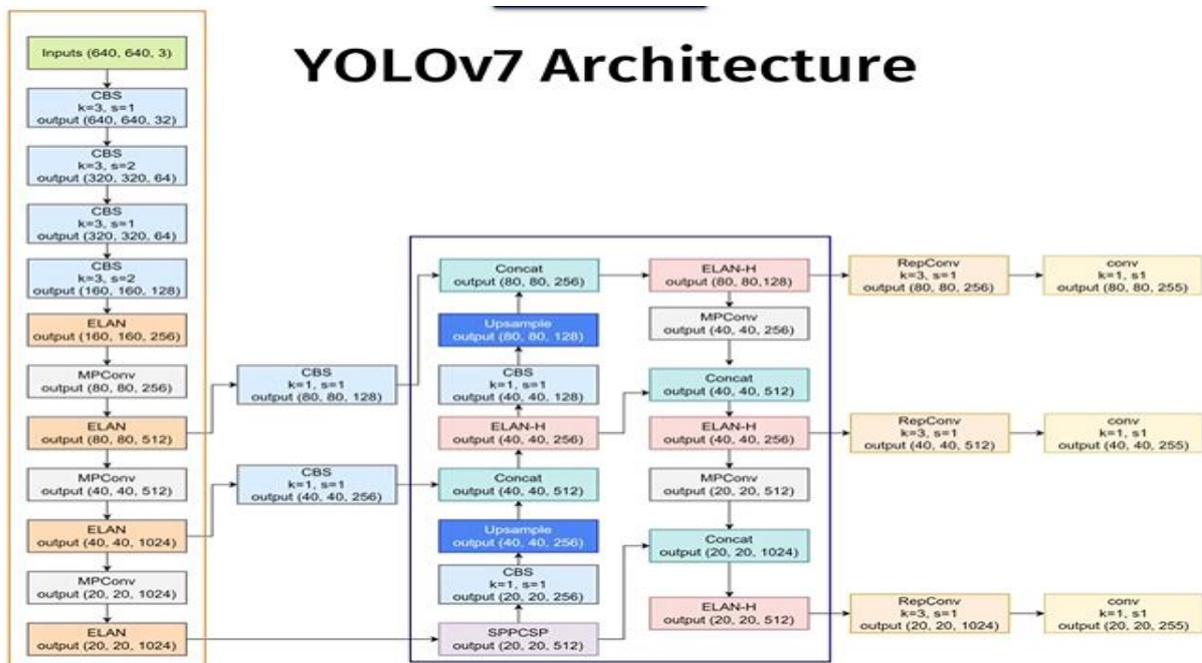
*Deep Learning* atau pembelajaran mendalam adalah salah satu bidang pada *machine learning* yang menggunakan banyak lapisan pada pengolahan informasi nonlinier dalam melakukan ekstraksi fitur, pengenalan pola, dan klasifikasi. Menurut goodfellow, deep learning atau pembelajaran mendalam adalah salah satu pendekatan yang menggunakan

konsep hierarki untuk menyelesaikan masalah pada sistem pembelajaran komputer. Konsep hierarki ini menjadikan komputer dapat mempelajari konsep yang rumit dengan melakukan penggabungan dari konsep-konsep yang lebih sederhana. Jika diilustrasikan pada sebuah graf tentang bagaimana konsep tersebut dibangun di atas konsep-konsep yang lainnya, graf ini akan memiliki banyak lapisan, hal inilah yang menjadi alasan disebut pelajaran mendalam (deep learning)[10].

*Convolutional Neural Network* adalah salah satu algoritma yang paling populer digunakan dalam *Deep learning* yang diturunkan dari *Multi-Layer Perceptron* (MLP) untuk pengolahan data dalam bentuk *grid* dua dimensi. CNN digunakan untuk mengklasifikasikan data yang sudah diberi label dengan metode belajar berbasis supervise. Model umum CNN terdiri dari empat komponen, yaitu lapisan konvolusi, lapisan *pooling*, fungsi aktivasi dan lapisan *fully connected* [11].

You Only Look Once (YOLO) adalah salah satu pendekatan untuk melakukan pendeteksian objek secara real-time berbasis Convolutional Neural Network. YOLO menggunakan pendekatan jaringan syaraf tunggal (Single neural network) untuk melakukan pendeteksian objek pada sebuah citra. Jaringan ini menggunakan fitur dari semua gambar untuk memprediksi setiap bounding box yang dapat melakukan prediksi pada kotak-kotak pembatas dan probabilitas secara langsung dalam satu evaluasi [12].

YOLOv7 menggunakan satu model CNN untuk melakukan deteksi target end-to-end. Algoritma ini mengambil seluruh gambar sebagai input dan secara langsung menghitung lokasi kotak pembatas serta kategori objek dalam lapisan keluaran. Terdiri dari empat modul utama: input, backbone, head, dan prediksi. Pada tahap awal citra akan masuk ke backbone E-Elan untuk memastikan bahwa model tidak menjadi terlalu kompleks, sehingga mengurangi risiko overfitting pada data pelatihan. Lalu citra dibadi menjadi grid dengan ukuran  $s \times s$ . Setiap grid akan diprediksi menggunakan bounding box dan menghasilkan nilai confidence. Confidence merupakan nilai yang mengidentifikasi sejauh mana bounding box tersebut berisi objek yang sesuai dengan perencanaan dan akurasi prediksi [13] seperti gambar berikut.



Gambar 1. Arsitektur YOLOv7

Metode You Only Look Once (YOLO) menjadi salah satu metode yang cepat dan akurat dalam melakukan pendeteksian objek. Metode ini mampu melakukan deteksi objek hingga 2 kali lebih cepat daripada algoritma yang lain. Pada penelitian ini, penulis menggunakan metode YOLOv7 karena memiliki beberapa peningkatan dalam mendeteksi objek dan nilai akurasi yang lebih tinggi daripada versi sebelumnya. Dengan menggunakan metode YOLOv7, diharapkan penelitian ini dapat memberikan hasil yang baik untuk mendeteksi dan mengklasifikasi kendaraan.

Pada penelitian [14] tentang “Rancang bangun penghitung dan pengidentifikasi kendaraan menggunakan *Multiple Object Tracking*”. Penelitian ini melakukan implementasi sistem penghitung dan pengidentifikasi kendaraan di jalan tol menggunakan *multiple object tracking*. Sistem menggunakan algoritma *Gaussian mixture model* dan *kalman filter* untuk mendeteksi dan melacak posisi, kecepatan arah gerak dan ukuran kendaraan dari waktu ke waktu pada tiap frame citra. Hasil dari penelitian menyatakan saat pagi hari menghasilkan hasil terbaik, dan pada malam hari hasil terburuk. Akurasi pada pagi hari sebesar 94%, siang hari 90%, sore hari 85%, dan malam hari 59%.

Pada penelitian [15] tentang “Deteksi dan Penggolongan Kendaraan dengan *Kalman Filter* dan Model *Gaussian* di Jalan Tol”. Penelitian ini menggunakan 2 jenis distribusi, yaitu distribusi *Background* dan *Foreground* untuk mendeteksi objek bergerak yang terdapat dalam file video (\*.avi) dengan resolusi 640x480. Data rekam yang digunakan berdurasi selama lima menit. Data diuji satu pertama dan empat menit terakhir, dengan jumlah durasi lima menit. Hasil



ekstraksi ciri dari kendaraan tersebut digunakan untuk melakukan penggolongan kendaraan berdasarkan dimensi pixel. Hasil segmentasi digunakan *kalman filter* untuk menghitung pelacakan posisi objek bergerak. Jika segmentasi *Bit Large Object* tidak terdapat objek yang bergerak, maka akan diteruskan pada *frame* selanjutnya. Hasil akhir deteksi sistem didapatkan dari validasi True Positif, True Negatif, False Positif, dan False Negatif dengan mencari sensitifitas dan spesifisitas pada waktu pagi, siang, dan malam.

Pada penelitian [6] tentang “Klasifikasi Jenis Kendaraan Menggunakan Metode *Convolutional Neural Network* (CNN)”. Penelitian ini bertujuan untuk klasifikasi jenis-jenis kendaraan yang tidak sesuai pada jalurnya. Menggunakan sebanyak 120 dataset citra yang terdiri dari citra mobil, motor, dan sepeda. Hasil menggunakan package keras menunjukkan akurasi sebesar 94,4% pada tahap pelatihan dan 73,3% pada tahap pengujian.

Van et al. [16] dalam penelitiannya yang berjudul “Things in the air: tagging wearable IoT information on drone videos” melaporkan bahwa evaluasi eksperimental menunjukkan tingkat pengenalan mencapai 99,5% pada berbagai jalur pejalan kaki, dan 98,6% ketika sudut kamera drone berada pada 37°. Penelitian ini merupakan yang pertama menggabungkan video dari kamera drone dengan data IoT dari sensor inersia, sehingga membuka peluang baru dalam integrasi data wearable dan penginderaan udara.

Pada penelitian [17] yang berjudul “YOLOv7-b: An Enhanced Object Detection Model for Multi-Scale and Dense Target Recognition in Remote Sensing Images”, modifikasi YOLOv7 dengan mengganti konvolusi standar dengan *deformable convolution* (DCNv2) dan menambahkan modul *Bi-level Routing Attention* (BRA) meningkatkan akurasi deteksi objek berskala kecil dan padat. Hasil eksperimen menunjukkan peningkatan mAP@0.5 sebesar 85,72% dan mAP@0.5:0.95 sebesar 66,55% pada dataset DIOR. Penulis menyarankan penggunaan augmentasi data dan integrasi mekanisme *attention* dinamis untuk meningkatkan generalisasi model pada dataset terbatas.

Dalam studi *Night Target Detection Algorithm Based on Improved YOLOv7*, peneliti mengintegrasikan modul GSConv dan ShuffleNetv2<sub>x1.5</sub> untuk mengurangi parameter model sebesar 14% sekaligus meningkatkan akurasi deteksi malam hari hingga 88,1% AP50. Teknik *Scylla-IoU* juga digunakan untuk mengoptimalkan fungsi loss dan meningkatkan ketahanan model terhadap noise [18]

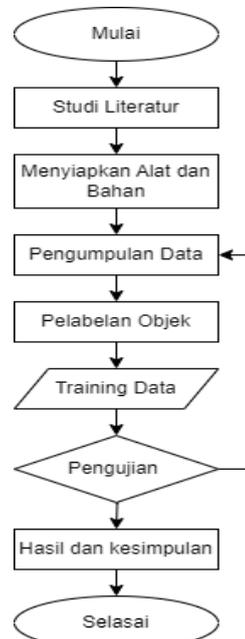
Penelitian oleh Zhang et al. (2023) dalam *Enhanced YOLOv7 Integrated with Small Target Enhancement* memperkenalkan modul *Lite Visual Center* (LVC) dan *Lite-ELAN* untuk meningkatkan deteksi objek kecil di permukaan air. Model STE-YOLO mencapai peningkatan AP50 sebesar 2,1% dan APs sebesar 1,6% dibandingkan YOLOv7 standar, dengan pengurangan parameter sebesar 14% [19]

Studi *Improved YOLOv7 for Small Object Detection Algorithm Based on Attention and Dynamic Convolution* (2023) memodifikasi YOLOv7 dengan menambahkan mekanisme *Coordinate Attention* (CA) dan *dynamic convolution*. Hasilnya, model mencapai mAP 81,1% pada dataset FloW-Img, meningkat 5,2% dari YOLOv7 dasar. Penulis menekankan pentingnya ekstraksi fitur hierarkis untuk mengurangi false positive pada objek kecil [20]

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Pada bagian ini dilakukan langkah-langkah penelitian digambarkan pada Gambar 2 yang terdiri dari studi literatur, perencanaan, pengumpulan data, implementasi, dan pengujian.



Gambar 2. Alur penelitian

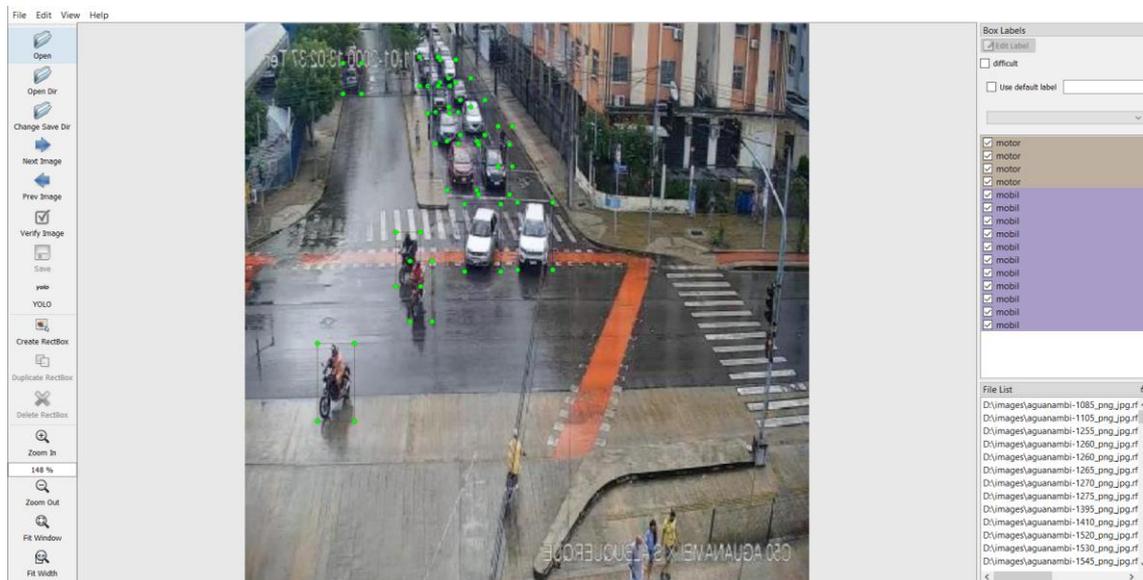
Alur penelitian secara umum dapat dilihat pada Gambar 2 dimana proses awal dimulai dengan studi literatur dan menyiapkan alat dan bahan selanjutnya ada pengumpulan data. Setelah data terkumpul maka dilakukan pelabelan objek, pelabelan atau anotasi bertujuan untuk memberikan informasi berupa nama kelas dan posisi yang berupa bounding box dari objek yang ingin dideteksi. Training data bertujuan untuk melatih komputer dengan cara mengolah gambar dan anotasi yang sudah dibuat sehingga terbentuk pola atau karakteristik dari setiap kelas yang akan menjadi bahan pertimbangan komputer dalam mencapai sebuah keputusan atau prediksi. Pengujian proses pengujian dilakukan untuk mengetahui hasil dari deteksi dan klasifikasi kendaraan. Hasil dari pengujian ini untuk mengetahui seberapa akurat dari algoritma YOLO.

## 2.2 Pengumpulan dan Pengolahan Data

Pada langkah ini dilakukan pengumpulan data yang diperlukan dalam penelitian. Dataset diperoleh dari website Kaggle dengan jumlah gambar sebanyak 6633, data tersebut kemudian diunduh dan diekstrak disalah satu localdisk. Setelah melakukan pengekstrakan dataset tersebut akan dilakukan proses labeling. Data tersebut berupa foto yang berfungsi untuk pemberian label pada jenis kendaraan. Data masukan dari penelitian ini berupa foto dengan format .jpg.

Pelabelan atau anotasi bertujuan untuk memberikan informasi berupa nama kelas dan posisi yang berupa bounding box dari objek yang ingin dideteksi. Tahap pelabelan objek dilakukan ketika data akuisisi yang berupa dataset yang diperoleh dari website kaggle telah diambil. Proses pelabelan dibagi menjadi 5 kelas, yaitu bicycle, bus, car, motorbike, dan person.

Proses anotasi citra menggunakan software LabelImg. Pelabelan ini menggunakan anotasi dalam format anotasi YOLO. Hasil dari anotasi tersebut adalah data yang terdapat informasi letak kotak pembatas dan labelnya dalam bentuk .txt. Pada .txt file terdapat baris file yang memiliki format `<object-class> <x_center> <y_center> <width> <height>`, dimana pada `<object-class>` merupakan bilangan bulat yang menyatakan kelas objek, `<x_center>` dan `<y_center>` adalah koordinat pusat persegi kotak pembatas, `<width>` dan `<height>` adalah nilai float relatif terhadap dimensi gambar. Hasil dari .txt file dapat dilihat pada Gambar 3.



**Gambar 3.** Proses anotasi dataset menggunakan LabelImg

Berikut merupakan hasil Gambar 4 yang sudah diberikan label:

	aguanambi-1085_p...	9/5/2023 11:45 AM	JPG File	75 KB
	aguanambi-1085_p...	11/27/2023 11:22 AM	Text Document	1 KB
	aguanambi-1105_p...	9/5/2023 11:45 AM	JPG File	75 KB
	aguanambi-1105_p...	11/27/2023 2:20 PM	Text Document	1 KB
	aguanambi-1255_p...	9/5/2023 11:45 AM	JPG File	82 KB
	aguanambi-1255_p...	11/27/2023 2:23 PM	Text Document	1 KB
	aguanambi-1260_p...	9/5/2023 11:45 AM	JPG File	82 KB
	aguanambi-1260_p...	11/27/2023 2:25 PM	Text Document	1 KB
	aguanambi-1260_p...	9/5/2023 11:45 AM	JPG File	82 KB
	aguanambi-1260_p...	11/27/2023 2:28 PM	Text Document	1 KB
	aguanambi-1265_p...	9/5/2023 11:45 AM	JPG File	81 KB
	aguanambi-1265_p...	11/27/2023 2:30 PM	Text Document	1 KB

**Gambar 4.** Gambar sudah diberikan label



```

0 0.160938 0.639844 0.065625 0.139063
0 0.289062 0.420312 0.043750 0.096875
0 0.311719 0.477344 0.039062 0.107813
0 0.460938 0.218750 0.025000 0.071875
1 0.421094 0.382031 0.064062 0.120313
1 0.514062 0.378906 0.062500 0.120313
1 0.384375 0.253125 0.053125 0.087500
1 0.441406 0.260156 0.054688 0.089063
1 0.357812 0.185938 0.040625 0.056250
1 0.403125 0.167187 0.043750 0.062500
1 0.342969 0.135156 0.039062 0.048438
1 0.380469 0.122656 0.039062 0.051562
1 0.189844 0.097656 0.032813 0.054688
1 0.359375 0.084375 0.028125 0.040625
1 0.325781 0.085156 0.032813 0.051562
    
```

Gambar 5. Hasil dari .txt file

Pemisahan dataset dilakukan melalui command prompt (CMD) dengan proporsi 80% train, 10% validasi, dan 10% test. Setelah proses pemisahan selesai, data tersebut diunggah ke Google Drive untuk memudahkan akses dan ekstraksi melalui platform Google Colab dalam rangka integrasi ke proses training data.

### 2.3 Training Data

Pelatihan data dilakukan menggunakan Google Colab. Pada tahap awal pelatihan YOLOv7, diperlukan format khusus untuk dataset kustom yang ditulis dalam file .txt. Proses pelatihan memanfaatkan parameter seperti jumlah epoch, batch size, dan optimizer. Selain itu, diterapkan mekanisme callback untuk menghentikan pelatihan secara otomatis jika nilai kerugian meningkat selama proses berlangsung.

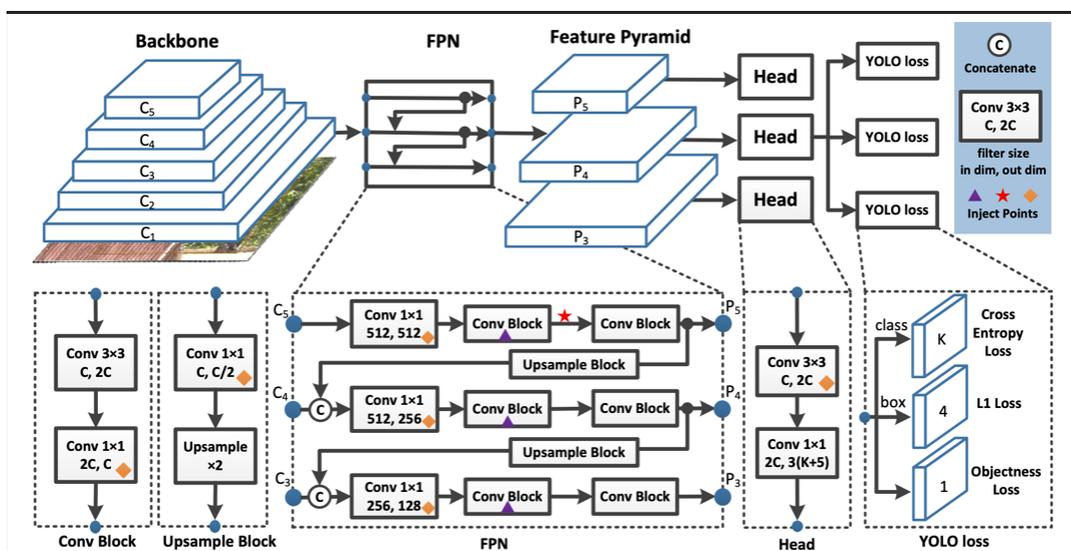
### 2.4 Pengujian

Pada bagian ini, dilakukan pengujian untuk mengevaluasi kinerja model YOLOv7 dalam mendeteksi dan mengklasifikasikan kendaraan. Dataset yang digunakan diperoleh dari website Kaggle, yang terdiri dari lima kelas: *bicycle*, *bus*, *car*, *motorbike*, dan *person*. Data tersebut melalui proses pelabelan menggunakan format anotasi YOLO

## 3. HASIL DAN PEMBAHASAN

### 3.1. Arsitektur YOLOv7

YOLOv7 (You Only Look Once versi 7) adalah generasi terbaru dari rangkaian model YOLO. Model ini termasuk dalam kategori deteksi objek satu tahap. Dalam prosesnya, citra masukan diproses melalui backbone untuk mengekstraksi fitur. Fitur-fitur tersebut kemudian dikombinasikan dan diolah lebih lanjut di bagian neck sebelum diteruskan ke head untuk melakukan prediksi. YOLO bertugas menentukan posisi dan kategori objek dengan menggambar kotak pembatas, dan menyempurnakan hasil prediksinya menggunakan teknik non-maximum suppression (NMS).



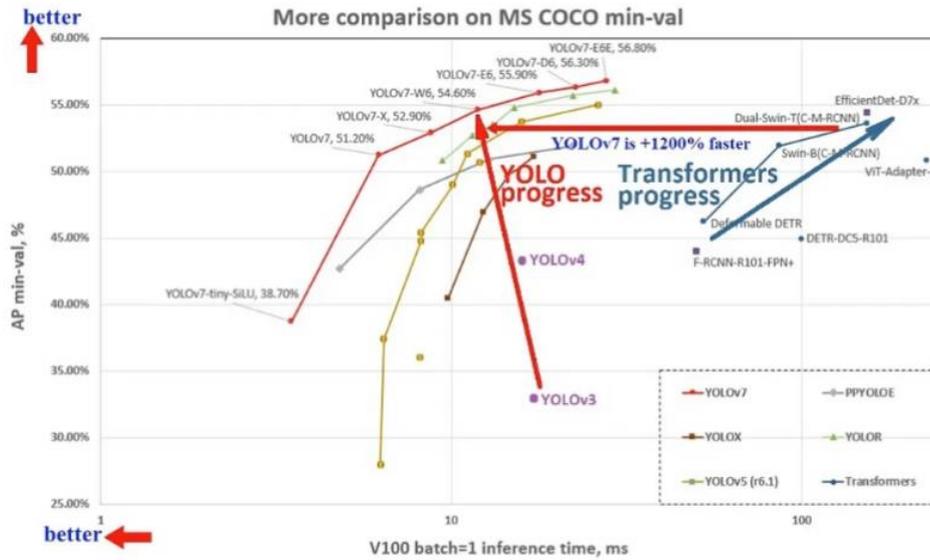
Gambar 6. Arsitektur jaringan YOLO seperti yang digambarkan dalam PP-YOLO

Model YOLO v7 ditulis oleh WongKinYiu dan Alexey Bochkovskiy (AlexeyAB). AlexeyAB mengambil obor YOLO dari penulis aslinya, Joseph Redmon, Ketika Redmon keluar dari industri Computer Vision (CV) karena masalah etika. AlexeyAB mempertahankan fork YOLOv3 untuk sementara waktu sebelum merilis YOLOv4, sebuah peningkatan dari model sebelumnya.



WongKinYiu memasuki tahap penelitian CV dengan jaringan Cross Stage Partial, yang memungkinkan YOLOv4 dan YOLOv5 membangun tulang punggung yang lebih efisien. Dari sana, WonKinYiu melaju kedepan memberikan kontribusi besar bagi keluarga penelitian YOLO dengan Scaled-YOLOv4. Model ini memperkenalkan penskalaan CSP YOLOv4 yang efisien untuk mencapai peta cangkih COCO.

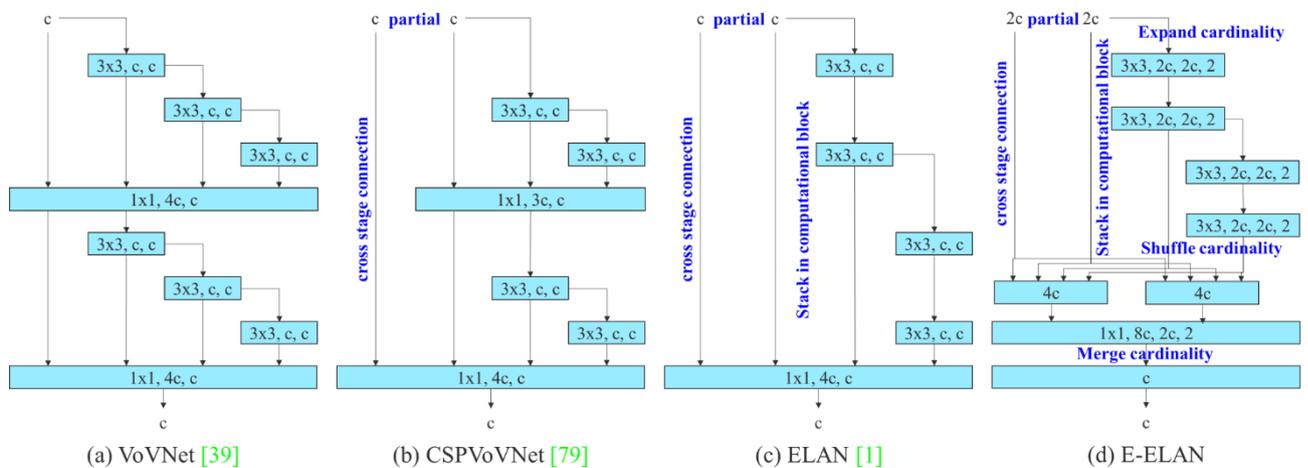
Penulis YOLOv7 berupaya menghadirkan inovasi dalam teknologi deteksi objek dengan merancang arsitektur jaringan yang mampu memprediksi bounding box secara lebih akurat dibandingkan metode lain, meskipun berada pada tingkat kecepatan inferensi yang setara.



Gambar 7. Yolov7 dievaluasi pada posisi kiri atas, menunjukkan performa yang lebih cepat dan akurat dibandingkan dengan jaringan sejenis lainnya.

### 3.2. Ektensi Efisiensi Lapisan Agregasi

Efisiensi lapisan konvolusional pada bagian backbone jaringan YOLO memainkan peran penting dalam memastikan kecepatan inferensi yang optimal. WongKinYiu memulai pendekatan efisiensi ini melalui penggunaan *Cross Stage Partial Networks*. Dalam pengembangan YOLOv7, para penulis melanjutkan eksplorasi pada topik tersebut dengan mempertimbangkan kebutuhan memori untuk menyimpan lapisan serta jarak propagasi balik gradien. Semakin pendek jarak yang ditempuh oleh gradien, semakin efektif jaringan dalam proses pembelajaran. Untuk agregasi lapisan akhir, mereka mengadopsi *E-ELAN*, yang merupakan versi pengembangan dari blok komputasi *ELAN*.



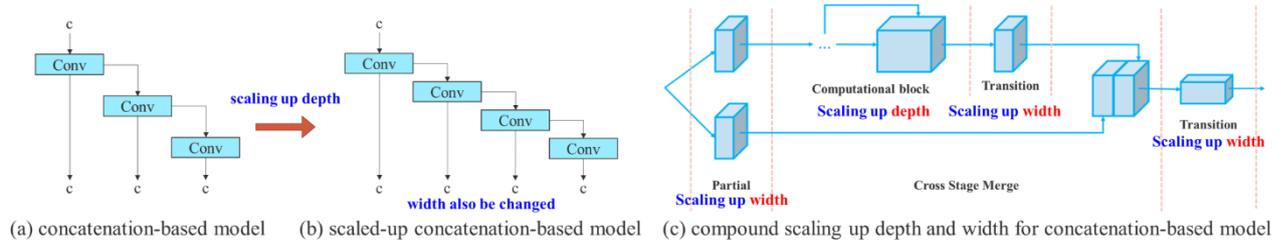
Gambar 8. Jaringan Agregasi Lapisan Efisien (E-ELAN) yang Telah Diperluas

### 3.3. Teknik Penskalaan Model

Model deteksi objek umumnya dirancang dalam beberapa varian ukuran untuk menyesuaikan kebutuhan berbagai aplikasi yang memerlukan tingkat akurasi dan kecepatan inferensi yang berbeda. Dalam proses pengembangan, aspek seperti kedalaman jaringan, lebar jaringan, serta resolusi pelatihan menjadi pertimbangan utama. Pada YOLOv7, pendekatan yang digunakan melibatkan penskalaan simultan terhadap kedalaman dan lebar jaringan, disertai dengan integrasi lapisan-

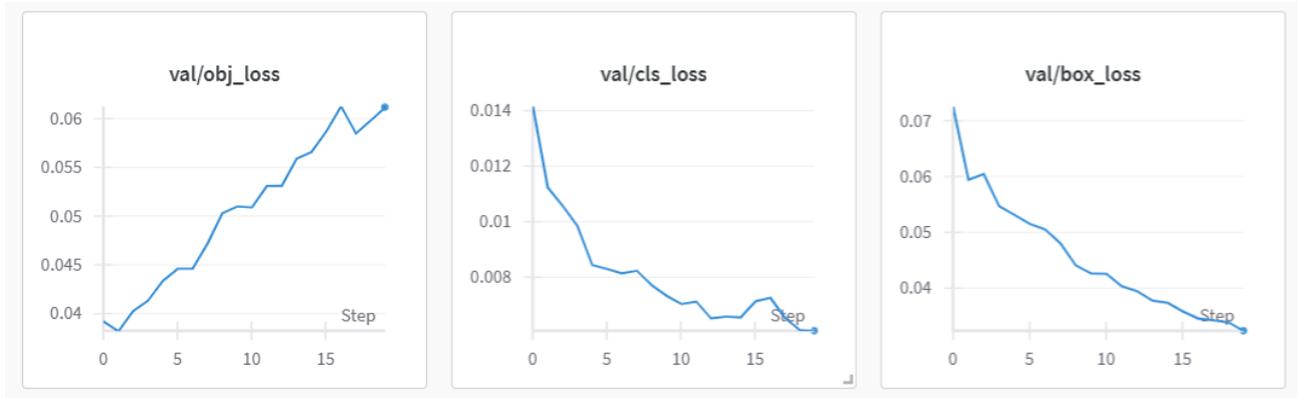


lapisan tertentu. Hasil dari studi ablation menunjukkan bahwa pendekatan ini mampu mempertahankan efisiensi arsitektur model, sekaligus memungkinkan penyesuaian terhadap berbagai ukuran model tanpa mengorbankan performa.



**Gambar 9.** Penskalaan model berbasis penggabungan (concatenation)

### 3.4. Pembahasan



**Gambar 10.** Visualisasi hasil *validation*

Gambar 10 pada grafik menunjukkan bahwa nilai *objectness loss* pada data validasi mengalami kenaikan seiring bertambahnya step (epoch), peningkatan ini menunjukkan model mengalami overfitting sehingga performa pada data validasi menurun. Selanjutnya terdapat gambar grafik *classification loss* dimana nilai loss ini menurun secara konsisten, menandakan bahwa model semakin baik dalam mengklasifikasikan objek yang terdeteksi pada data validasi. Gambar grafik *bounding box regression loss* memiliki nilai loss yang juga menurun, menandakan bahwa prediksi posisi dan ukuran bounding box semakin akurat terhadap objek pada data validasi.



**Gambar 11.** Visualisasi hasil train

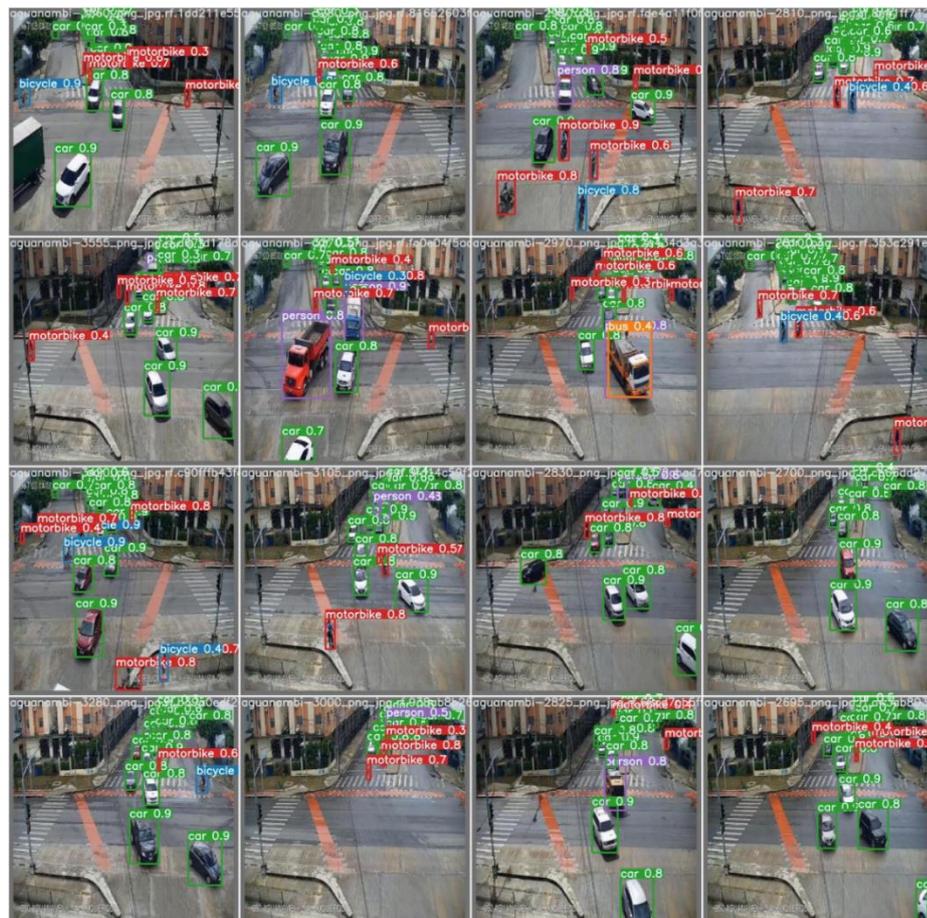
Grafik pada Gambar 11 menunjukkan bahwa nilai *objectness loss* pada data training mengalami penurunan secara konsisten dari awal hingga akhir training, menandakan model semakin baik dalam membedakan antara objek dan background pada data training. Selanjutnya terdapat gambar grafik *classification loss* dimana nilainya juga terus menurun yang menandakan bahwa model semakin akurat dalam mengklasifikasikan objek yang terdeteksi. Gambar *bounding box regression loss* penurunan loss ini berarti prediksi posisi dan ukuran bounding box oleh model semakin mendekati ground truth pada data training.



Gambar 12. Visualisi hasil training

Pada Gambar 12, Grafik recall menunjukkan seberapa baik model dalam menemukan semua objek yang ada di gambar. Nilai recall meningkat pesat di awal training dan terus naik hingga mendekati 0.9, menandakan model semakin mampu mendeteksi objek yang sebenarnya ada. Grafik precision menunjukkan seberapa akurat prediksi model, yaitu proporsi prediksi objek yang benar dari seluruh prediksi yang dibuat. Precision juga meningkat stabil hingga mendekati 0.95, artinya sebagian besar prediksi objek oleh model memang benar. Grafik mean Average Precision (mAP) dengan threshold IoU 0.5 hingga 0.95 adalah metrik komprehensif untuk mengukur performa deteksi objek pada berbagai tingkat ketelitian. Nilai mAP ini meningkat secara konsisten hingga sekitar 0.7, menunjukkan model semakin baik dalam mendeteksi dan mengklasifikasikan objek pada berbagai ukuran dan posisi. Grafik ini menunjukkan mAP pada threshold IoU 0.5 saja. Nilai mAP\_0.5 meningkat tajam di awal dan terus naik hingga mendekati 0.95, menandakan model sangat baik dalam mendeteksi objek dengan toleransi posisi yang lebih longgar.

Sebuah system dikembangkan untuk secara otomatis mendeteksi kecelakaan yang terkait dengan deteksi kendaraan yang terkait dengan kendaraan dalam gambar dan video. Metode yang diusulkan berhasil diimplementasikan. Kami melatih dataset dan hasil akhir adalah kotak batas disekitar objek yang dideteksi berikut adalah hasilnya:



Gambar 13. Deteksi menggunakan YOLOv7

Tabel 1. Hasil evaluasi

Precision	Recall	F1	mAP@0.5	mAP@0.5:0.95
93.22%	90.64%	0.91	94.27%	69.19%



Hasil dari deteksi objek menunjukkan bahwa model memiliki fungsi tingkat presisi sebesar 93.22%, yang berarti Sebagian besar objek yang diidentifikasi oleh model adalah objek yang benar-benar ada. Nilai recall sebesar 90.64% menunjukkan bahwa model mampu mendeteksi Sebagian besar objek yang sebenarnya ada dalam dataset. Skor F1 sebesar 0.91 mengidentifikasi bahwa model memiliki keseimbangan yang baik antara precision dan recall. Selain itu, mAP@0.5 mencapai nilai 94.27% yang menunjukkan bahwa model memiliki dapat mengidentifikasi objek dengan baik dengan batas IoU sebesar 0.5. Namun mAP@0.5:0.9 yang mencapai 69.19% menunjukkan bahwa performa model dalam mengidentifikasi objek dengan Tingkat tumpang tindih yang lebih ketat (IoU 0.95) mungkin sedikit lebih rendah.

#### 4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan algoritma YOLOv7 untuk deteksi dan klasifikasi kendaraan dari citra lalu lintas, dengan hasil precision sebesar 93,22%, recall 90,64%, mAP@0.5 sebesar 94,27%, dan mAP@0.5:0.95 sebesar 69,19%. Capaian ini menunjukkan bahwa model mampu mendeteksi dan mengklasifikasikan kendaraan secara akurat dan efisien, sehingga berkontribusi secara signifikan terhadap tujuan penelitian, yaitu mendukung manajemen lalu lintas dan perencanaan pembangunan jalan melalui penyediaan data kepadatan kendaraan yang lebih akurat. Dengan deteksi kendaraan yang cepat dan presisi tinggi, sistem ini dapat membantu mengidentifikasi area dengan lalu lintas padat secara lebih efektif. Untuk meningkatkan kinerja lebih lanjut, disarankan menambah jumlah epoch dalam proses pelatihan serta menerapkan teknik augmentasi data guna memperkuat ketahanan model terhadap variasi kondisi di lapangan.

#### REFERENCES

- [1] D. Iskandar Mulyana and M. A. Rofik, "Implementasi Deteksi Real Time Klasifikasi Jenis Kendaraan Di Indonesia Menggunakan Metode YOLOV5," *J. Pendidik. Tambusai*, vol. 6, no. 3, pp. 13971–13982, 2022, doi: 10.31004/jptam.v6i3.4825.
- [2] Nurhawanti, "SISTEM PENDETEKSI SEPEDA MOTOR PELANGGAR MARKA JALAN MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORKS (CNNs)," *Galang Tanjung*, vol. 2, no. 2504, pp. 1–9, 2019.
- [3] M. Harahap, J. Elfrida, P. Agusman, M. Rafael, R. Abram, and K. Andrianto, "Sistem Cerdas Pemantauan Arus Lalu Lintas Dengan YOLO (You Only Look Once v3)," *Semin. Nas. APTIKOM*, pp. 367–376, 2019.
- [4] D. Alamsyah, S. Gi, and M. Palembang, "Pengenalan Mobil pada Citra Digital Menggunakan HOG-SVM" *J. Informatika dan Sistem Informasi*, Vol 3, No 2, 2017, [10.35957/jatisi.v3i2.79](https://doi.org/10.35957/jatisi.v3i2.79)
- [5] A. Arinaldi, J. A. Pradana, and A. A. Gurusinga, "Detection and classification of vehicles for traffic video analytics," *Procedia Comput. Sci.*, vol. 144, pp. 259–268, 2018, doi: 10.1016/j.procs.2018.10.527.
- [6] N. Fadlia and R. Kosasih, "Klasifikasi Jenis Kendaraan Menggunakan Metode Convolutional Neural Network (Cnn)," *J. Ilm. Teknol. dan Rekayasa*, vol. 24, no. 3, pp. 207–215, 2019, doi: 10.35760/tr.2019.v24i3.2397.
- [7] M. Fachrie, "A Simple Vehicle Counting System Using Deep Learning with YOLOv3 Model," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 3, pp. 462–468, 2020, doi: 10.29207/resti.v4i3.1871.
- [8] Z. A. Fikriya, M. I. Irawan, and S. Soetrisno., "Implementasi Extreme Learning Machine untuk Pengenalan Objek Citra Digital," *J. Sains dan Seni ITS*, vol. 6, no. 1, 2017, doi: 10.12962/j23373520.v6i1.21754.
- [9] N. K. Negoro, E. Utami, and A. Yaqin, "Klasifikasi Deteksi Penggunaan Masker Menggunakan Metode Convolutional Neural Network," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 8, no. 2, pp. 664–674, 2023, doi: 10.29100/jupi.v8i2.3748.
- [10] S. R. Dewi, "Deep Learning Object Detection Pada Video," *Deep Learn. Object Detect. Pada Video Menggunakan Tensorflow Dan Convolutional Neural Netw.*, pp. 1–60, 2018, [Online]. Available: [https://dspace.uui.ac.id/bitstream/handle/123456789/7762/14611242\\_Syarifah Rosita Dewi\\_Statistika.pdf?sequence=1](https://dspace.uui.ac.id/bitstream/handle/123456789/7762/14611242_Syarifah%20Rosita%20Dewi_Statistika.pdf?sequence=1)
- [11] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," *Procedia Comput. Sci.*, vol. 132, pp. 679–688, 2018, doi: 10.1016/j.procs.2018.05.069.
- [12] G. Plastiras, C. Kyrkou, and T. Theodorides, "Efficient convnet-based object detection for unmanned aerial vehicles by selective tile processing," *ACM Int. Conf. Proceeding Ser.*, 2018, doi: 10.1145/3243394.3243692.
- [13] M. G. A. Binuri, T. Haryanti, and M. A. Haq, "Penerapan Algoritma YOLO v7 Sebagai Deteksi Kecelakaan Kendaraan Pada Lalu Lintas," *Comput. Insight J. Comput. Sci.*, vol. 4, no. 2, pp. 7–14, 2022.
- [14] L. Rahmawati and K. Adi, "Rancang bangun penghitung dan pengidentifikasi kendaraan menggunakan Multiple Object Tracking," *Youngster Phys. J.*, vol. 6, no. 1, pp. 70–75, 2017.
- [15] R. F. Waliulu, "Deteksi dan Penggolongan Kendaraan dengan Kalman Filter dan Model Gaussian di Jalan Tol," *J. Sist. Inf. Bisnis*, vol. 8, no. 1, p. 1, 2018, doi: 10.21456/vol8iss1pp1-8.
- [16] L. Da Van, L. Y. Zhang, C. H. Chang, K. L. Tong, K. R. Wu, and Y. C. Tseng, "Things in the air: tagging wearable IoT information on drone videos," *Discov. Internet Things*, vol. 1, no. 1, 2021, doi: 10.1007/s43926-021-00005-8.
- [17] Y. Song, H. Yang, L. Huang, and S. Huang, "YOLOv7-b: An Enhanced Object Detection Model for Multi-Scale and Dense Target Recognition in Remote Sensing Images," *Int. J. Adv. Comput. Sci. Appl.*, vol. 16, no. 2, pp. 233–248, 2025, doi: 10.14569/IJACSA.2025.0160225.
- [18] Z. Bowen, L. Huacai, Z. Shengbo, C. Xinqiang, and X. Hongwei, "Night target detection algorithm based on improved YOLOv7," *Sci. Rep.*, vol. 14, no. 1, pp. 1–10, 2024, doi: 10.1038/s41598-024-66842-z.
- [19] J. Yu, H. Zheng, L. Xie, L. Zhang, M. Yu, and J. Han, "Enhanced YOLOv7 integrated with small target enhancement for rapid detection of objects on water surfaces," *Front. Neurorobot.*, vol. 17, 2023, doi: 10.3389/fnbot.2023.1315251.
- [20] K. Li, Y. Wang, and Z. Hu, "Improved YOLOv7 for Small Object Detection Algorithm Based on Attention and Dynamic Convolution," *Appl. Sci.*, vol. 13, no. 16, 2023, doi: 10.3390/app13169316.