

# *Design of Object Detection System Using Deep Learning in Laboratory Room*

Perancangan Sistem Deteksi Objek Menggunakan Yolov5 Pada Ruangan Laboratorium

Dion Omong, M. Dwisnanto Putro, Jane I. Litouw

Dept. of Electrical Engineering, Sam Ratulangi University Manado, Kampus Bahu St., 95115, Indonesia

e-mails: [dionomong023@student.unsrat.ac.id](mailto:dionomong023@student.unsrat.ac.id)

Received: 27 March 2024; revised: 10 November 2024; accepted: 21 February 2025

*Abstract — In today's era, the challenge of object detection within laboratory spaces is paramount. The need for automated monitoring systems supporting object detection within laboratory settings, particularly for service robots, is crucial for scientific research and safety. Utilizing convolutional artificial neural networks has shown promising capabilities in object recognition. This paper proposes a novel method for detecting and localizing specific objects within laboratory environments using deep learning techniques. Employing a lightweight primary feature extractor allows for the identification of crucial object features without imposing excessive computational demands. Additionally, a depthwise separable convolution module is introduced, designed to capture essential components within multi-level convolutional patches specific to laboratory conditions. To address the challenges encountered within laboratory spaces, a new dataset is proposed, encompassing issues such as lighting variations, blurriness, occlusions, and complex backgrounds. Evaluation results indicate that the proposed model surpasses other lightweight object detection models in laboratory settings, offering high accuracy and efficiency. This model can be efficiently implemented on low-tier devices, ensuring real-time object detection and processing within laboratory environments.*

**Keywords —** object detection, service robot, convolutional neural networks, You Only Look Once version 5.

**Abstrak —** Di era sekarang ini, tantangan deteksi objek di dalam ruang laboratorium menjadi sangat penting. Kebutuhan akan sistem pemantauan otomatis yang mendukung deteksi objek di dalam laboratorium, khususnya untuk robot servis, sangat penting untuk penelitian ilmiah dan keselamatan. Memanfaatkan jaringan saraf tiruan convolutional telah menunjukkan kemampuan yang menjanjikan dalam pengenalan objek. Makalah ini mengusulkan metode baru untuk mendeteksi dan melokalisasi objek tertentu di dalam lingkungan laboratorium menggunakan teknik pembelajaran mendalam. Menggunakan ekstraktor fitur utama yang ringan memungkinkan identifikasi fitur objek yang penting tanpa membebani komputasi yang berlebihan. Selain itu, modul konvolusi yang dapat dipisahkan secara mendalam juga diperkenalkan, yang dirancang untuk menangkap komponen penting dalam patch konvolusi multi-level yang spesifik untuk kondisi laboratorium. Untuk mengatasi tantangan yang dihadapi di dalam ruang laboratorium, dataset baru diusulkan, yang mencakup masalah seperti variasi pencahayaan, kekaburan, oklusi, dan latar belakang yang kompleks. Hasil evaluasi menunjukkan bahwa model yang diusulkan melampaui model pendeteksian objek ringan lainnya di lingkungan laboratorium, menawarkan akurasi dan efisiensi yang tinggi. Model ini dapat diimplementasikan secara efisien pada perangkat tingkat rendah, memastikan pendeteksian dan pemrosesan objek secara real-time di lingkungan laboratorium.

**Kata kunci —** deteksi objek, robot layanan, jaringan saraf konvolusi, Anda Hanya Melihat Sekali versi 5.

## I. PENDAHULUAN

pengelolaan keamanan dan pengawasan objek menjadi aspek vital yang memengaruhi efisiensi operasional dan keselamatan lingkungan kerja. Ruang laboratorium sering kali menjadi tempat untuk pengembangan teknologi, eksperimen ilmiah, dan penelitian mendalam. Namun, tantangan muncul dalam pengelolaan ruangan ini, terutama terkait pemantauan aktivitas, keamanan peralatan, dan manajemen inventaris yang efektif. Tantangan utama yang muncul adalah bagaimana menciptakan sistem yang mampu secara real-time mendeteksi objek dan aktivitas yang terjadi di ruang laboratorium. Meskipun ada beberapa teknologi deteksi objek yang ada, masih terdapat kekurangan dalam hal akurasi, kecepatan, dan adaptabilitas terhadap lingkungan ruang laboratorium yang dinamis.[1]

Meskipun terdapat berbagai teknologi deteksi objek yang ada, tantangan dalam mendeteksi dan mengawasi objek-objek yang berada dalam ruang laboratorium masih menjadi isu yang perlu diatasi. Kecepatan, ketepatan, dan adaptabilitas terhadap lingkungan ruang laboratorium yang dinamis masih menjadi titik fokus dalam pengembangan sistem deteksi objek.[2]

Pendekatan modern memperkenalkan Convolutional Neural Networks (CNN) sebagai teknik untuk mengekstraksi fitur yang secara efektif membedakan antara fitur objek dan latar belakang[3]. Penggunaan operasi filter memungkinkan identifikasi karakteristik unik dari objek target melalui penerapan bobot terlatih. Melalui pembelajaran mandiri, jaringan dapat secara independen menentukan bobot kernel dengan memperkirakan kesalahan prediksi. Pendekatan ini telah ditunjukkan untuk secara andal menyoroti informasi penting dari objek, meskipun dengan mengorbankan sumber daya komputasi yang signifikan ketika menggunakan lapisan konvolusi yang dalam. Beberapa penelitian[4], telah menggunakan metode ini untuk mengenali objek dalam gambar input. Khususnya, kinerja yang dihasilkan dianggap memuaskan, tanpa mengalami saturasi selama proses pelatihan.

Dalam lingkungan laboratorium, penerapan algoritma Convolutional Neural Network (CNN) telah menunjukkan keahliannya dalam mendeteksi karakteristik objek-objek khusus di dalamnya. Meskipun dihadapkan pada tantangan seperti variasi bentuk, skala, dan kondisi pencahayaan yang terbatas, metode ini tetap mempertahankan kemampuannya

dalam mengidentifikasi ciri khas dari objek laboratorium. Penelitian terdahulu telah berhasil menerapkan pendekatan ini untuk menemukan posisi objek laboratorium yang memiliki peran penting dalam pengelolaan lingkungan laboratorium[5].

Penggunaan teknik deep learning, terutama dengan dataset yang luas, terbukti efektif dalam berbagai konteks, memungkinkan identifikasi dan klasifikasi objek dengan tingkat akurasi yang tinggi di dalam lingkungan laboratorium. Penerapan algoritma cerdas, seperti model Region-Convolutional Neural Network (R-CNN) atau Faster R-CNN[5], di dalam lingkungan laboratorium bertujuan untuk meminimalkan waktu yang dibutuhkan untuk mengidentifikasi dan mengklasifikasikan objek-objek khusus. Pendekatan menggunakan algoritma R-CNN dan Fast R-CNN mengutamakan pencarian selektif dalam mengidentifikasi proposal wilayah, mempercepat proses yang pada gilirannya meningkatkan kinerja jaringan dalam mendeteksi objek yang relevan di lingkungan laboratorium.

Kehebatan metode CNN mendorong penelitian ini untuk menerapkan arsitektur berbasis konvolusi yang mempertimbangkan efisiensi. Hal ini mendorong pengembangan arsitektur yang hemat biaya yang dapat beroperasi pada perangkat dengan keterbatasan. Sebuah tulang punggung (backbone) ringan dengan cepat mengekstrak fitur-fitur objek laboratorium didukung oleh modul perhatian. Modul perhatian ini memperkuat kinerja ekstraktor untuk menangkap fitur khusus dari objek yang dituju[6]. Dalam penelitian ini, kami mengusulkan sistem visi untuk mendeteksi dan melokalisasi objek laboratorium yang dapat beroperasi dengan daya komputasi yang rendah. Sistem ini dilengkapi dengan modul peningkatan untuk meningkatkan kinerja jaringan dan menggunakan tiga lapisan deteksi. Kontribusi dari penelitian ini dapat diringkas sebagai berikut.

Pengembangan metode deteksi objek laboratorium yang baru yang dapat beroperasi dengan cepat untuk implementasi pada perangkat dengan keterbatasan.

Pengembangan modul perhatian kelompok yang inovatif untuk memperkuat kinerja tulang punggung dengan menangkap fitur khusus dari objek laboratorium. Modul ini melakukan alokasi kelompok konvolusi dan perhatian untuk membedakan fitur objek pada berbagai skala area penangkapan. Penugasan terpisah dari setiap peta fitur dapat meningkatkan variasi informasi objek tanpa mengorbankan efisiensi energi komputasi dan parameter.

## II. METODE

### A. Pengumpulan Dataset

Dalam pengembangan model, dilakukan pengumpulan data gambar yang akan digunakan dalam proses pelatihan pada model nanti. Kumpulan data dari eksperimen kami telah dianotasi dengan cermat untuk tugas khusus objek deteksi.

Data ini mencakup gambar didalam laboratorium yang diambil dari beberapa video dengan kondisi pencahayaan pada siang dan malam hari. Kami memperoleh video tersebut secara mandiri. Keseluruhan video berisi objek yang terdiri dari meja besar, meja kecil, kersi kuliah, kursi acara, lemari, dispenser, monitor, orang, toolbox, rak, papan tulis, kardus, pintu.

Kami membuat label pada setiap objek menggunakan alat

Boobs Master[7] dan mengabaikan latar belakang yang membantu memberikan pengetahuan detektor secara tepat. Proses ini mereplikasi skenario dunia nyata dan memastikan kualitas dan keakuratan dari bounding box ground truth melalui anotasi manual. Kami secara hati-hati memberikan anotasi pada setiap gambar dengan mengidentifikasi dan menggambarkan kotak pembatas di sekitar setiap objek. Dataset yang telah disesuaikan ini memainkan peran penting dalam melatih dan mengevaluasi kinerja sistem deteksi waktu nyata kami, yang menggabungkan arsitektur jaringan yang diusulkan. Upaya ini menghasilkan total 1.619 gambar, untuk tahap pelatihan 1295, dan 324 untuk tahap pengujian.

### B. Konfigurasi Pelatihan dan Pengujian

Dalam percobaan ini, workstation menggunakan Ubuntu 18.04 dengan RAM 16 GB, ditenagai oleh CPU Intel Core i5-2320 dan dilengkapi dengan GPU NVIDIA GTX 1080 sebagai kartu grafis akselerator. Jaringan yang diusulkan disimulasikan menggunakan PyTorch dan memanfaatkan CUDA 10.1 untuk akselerasi GPU yang efisien. Proses pelatihan dilakukan lebih dari 300 epoch, memastikan model mengalami iterasi pelatihan yang cukup untuk mendeteksi pada ruangan laboratorium.

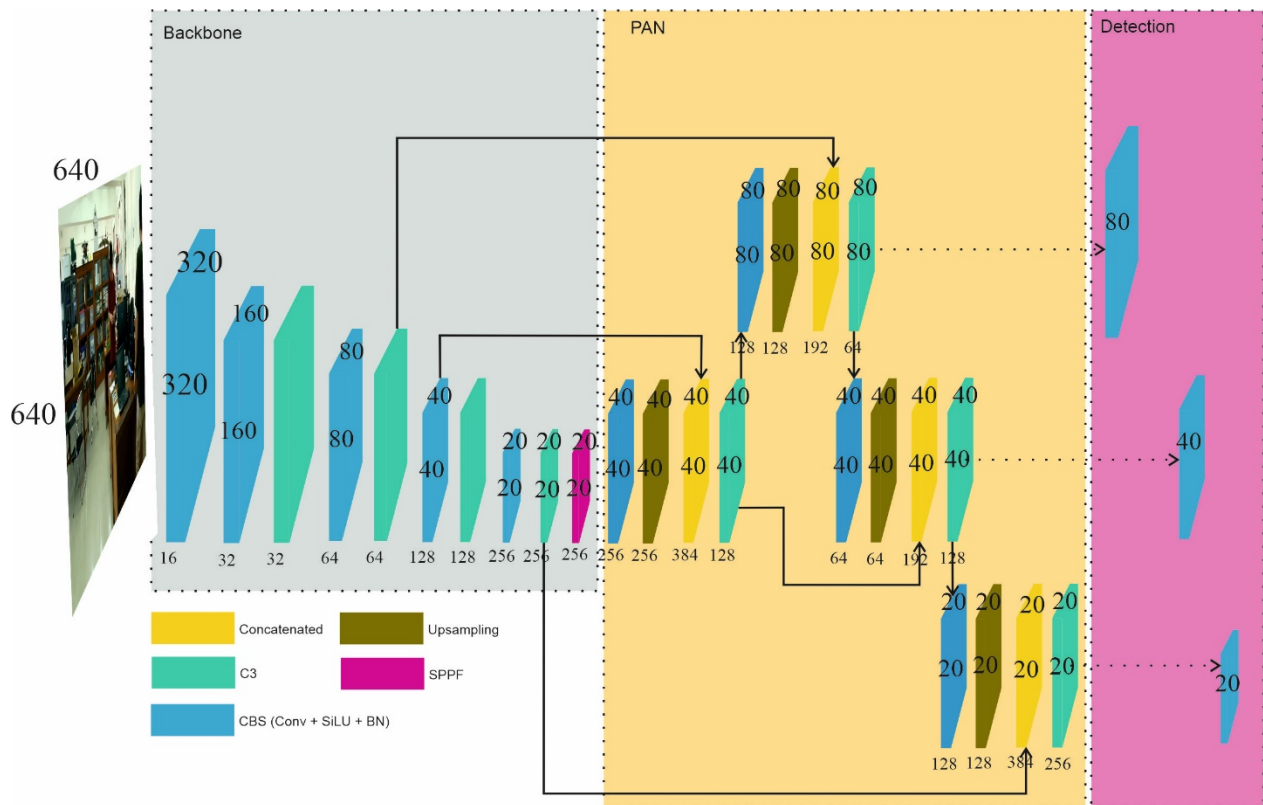
Ukuran gambar yang digunakan untuk pelatihan dan evaluasi ditetapkan pada 640x640 piksel. Laju pembelajaran sebesar 0,001 dan ukuran batch 64 digunakan selama proses pelatihan. Kemudian, ia mengadopsi fungsi loss [13] untuk menghitung perbedaan antara prediksi bounding box dengan kelas objek dan ground truth. Model ini mengukur kinerjanya pada NVIDIA Jetson Nano dengan RAM LPDDR4 4 GB 64-bit pada tahap pengujian dan penerapan. Hal ini memungkinkan kami untuk mengevaluasi kompatibilitas dan efisiensi model dalam skenario dunia nyata dengan sumber daya komputasi yang terbatas. Model performa dievaluasi berdasarkan mean average precision (mAP) yang menghitung intersection over union (IoU) pada 0.5 dan 0.5:0.95.

### B. Arsitektur Model

Pada bagian ini, arsitektur yang diusulkan dijelaskan secara rinci dan berfokus pada modul yang ditawarkan untuk meningkatkan kinerja deteksi objek dalam ruangan laboratorium. Objek di dalam ruang laboratorium memiliki karakteristik tekstur yang berbeda-beda. Dalam hal ini, arsitektur yang diusulkan bertujuan untuk mengekstraksi informasi objek dan meningkatkan efisiensi energi. Tujuannya adalah mendorong sistem visi ini agar dapat diimplementasikan pada perangkat berbiaya rendah untuk kebutuhan pengembangan teknologi di ruang laboratorium. Secara keseluruhan, arsitektur kami terinspirasi oleh YOLOv5 [13], yang dikembangkan agar lebih ringan dan terdiri dari ekstraktor utama, modul perhatian, jaringan agregat jalur, dan lapisan deteksi. Arsitektur umum detektor kami disajikan pada Gambar. 1

### C. Modul Ekstraktor Utama

Ekstraksi fitur memainkan peran penting dalam memisahkan fitur yang berharga dari fitur yang sepele. Proses ini menjadi inti utama dari jaringan pendeteksi objek, yang menggunakan lapisan konvolusi. Operasi ini secara kuat membedakan informasi target dengan menerapkan multi-kernel berbobot. Proses pembelajaran mendorong setiap filter untuk memperbarui bobot dan memberikan prediksi akhir yang akurat. Arsitektur yang diusulkan menggunakan ekstraktor utama untuk mengekstrak fitur-fitur penting sambil mengurangi dimensi peta fitur untuk mengurangi komputasi



Gambar 1 Arsitektur YOLOv5

floating-point. Selain itu, arsitektur ini meningkatkan jumlah saluran secara simultan yang dapat mengakomodasi fitur yang diekstraksi. Struktur ini adalah tipikal dari semua arsitektur CNN yang cenderung berdimensi kecil dan gemuk pada jaringan.

Tulang punggung ini mengadopsi struktur YOLOv5, yang menerapkan lapisan konvolusi  $6 \times 6$  dengan langkah 2 pada tahap awal menjadi lima, yang dapat mengurangi dimensi peta. Operasi ini lebih efektif daripada pooling. Tulang punggung tetap menggunakan C3 (modul konvolusi yang lebih cepat). Modul ini menggunakan dua lapisan konvolusi di awal dan akhir modul dan menerapkan modul bottleneck pada peta fitur terpisah.

Untuk mencegah hilangnya informasi karena ekstraksi yang berlebihan, modul ini menggunakan teknik residual. Selain itu, setiap peta yang terpecah digabungkan dengan operasi penggabungan untuk memperkaya informasi. Selanjutnya, spatial pyramid pooling-fast (SPPF) diterapkan setelah modul C3 terakhir pada tulang punggung untuk memilih fitur spasial dengan cascade pooling. Modul ini dapat menangkap nilai maksimum di area cakupan yang berbeda untuk setiap level.

Untuk gambar dengan input ukuran yang bervariasi, struktur SPP[8] dalam jaringan YOLOv5 mengimplementasikan vektor fitur dengan ukuran tetap sebagai output lapisan yang terhubung penuh. Melalui penggunaan tiga kernel konvolusi dengan ukuran yang bervariasi - 3, 5, dan 9 - struktur SPP mampu meningkatkan bidang reseptif jaringan, meningkatkan kemampuan grafik fitur untuk merepresentasikan fitur, dan mengekstrak fitur melalui operasi penggabungan maksimum. Struktur SPP, yang pertama-tama melakukan operasi penggabungan maksimum  $1 \times 1$ ,  $3 \times 3$ ,  $5 \times 5$ , dan  $9 \times 9$  pada data yang ditransfer dari fungsi aktivasi normalisasi konvolusi

(Konvolusi + Normalisasi Batch + SiLU, CBS) secara paralel dan kemudian menghubungkannya dengan struktur CBS dengan penyambungan gabungan. Proses ekstraksi fitur selesai dan fusi fitur dicapai melalui struktur CBS. Namun, melalui operasi penyatuan paralel dengan berbagai ukuran kernel convolutional, struktur SPP meningkatkan komputasi program dengan mengorbankan penurunan kinerja. Hasilnya, pooling menggunakan struktur SPPF untuk meningkatkan performa pooling dengan menggunakan daya komputasi program yang lebih sedikit. Dengan struktur SPPF, tiga kernel konvolusi dengan ukuran yang sama melakukan operasi serial sebagai pengganti operasi penyatuan maksimum paralel SPP asli dengan tiga kernel dengan ukuran berbeda. Data yang ditransfer secara serial dari struktur CBS pertama-tama mengalami operasi maksimal pooling  $5 \times 5$  oleh lapisan SPPF. Selanjutnya, data dimasukkan ke dalam struktur CBS dengan penyambungan penggabungan, yang mempercepat proses penggalan informasi fitur yang lebih kaya.

#### D. Jaringan Agregat Jalur

Leher jaringan menentukan kualitas fitur dalam pendeteksian objek dengan menghubungkannya ke fitur multi-level di tulang punggung. Penggabungan fitur meningkatkan hubungan antara elemen-elemen dari setiap tahap konvolusi dengan menggunakan pendekatan upsampling dan downsampling untuk menyamakan dimensi peta. YOLOv5 mengimplementasikan jaringan agregat jalur (PAN) yang menghasilkan tiga tingkat menggunakan pendekatan jalur informasi penggabungan dari bawah ke atas. Modul ini menggunakan ekstraktor C3 untuk menyaring informasi agregat. Modul ini mengabaikan residual karena dimensi yang berbeda antara peta input dan output.

Metode terdekat diterapkan untuk menggandakan ukuran dimensi peta pada tahap pertama dan kedua. Metode ini menghasilkan peta fitur terbesar pada lapisan deteksi. Di sisi lain, dua blok konvolusi 3 x 3 dengan langkah 2 diterapkan secara berurutan setelah tahap kedua untuk menghasilkan peta menengah dan kecil. Interkoneksi informasi dengan fitur tulang punggung menerapkan operasi penggabungan, yang diperlukan untuk memperkuat fitur kombinasi sebelum diteruskan ke modul deteksi.

### E. Lapisan Deteksi

Lapisan Deteksi dalam YOLOv5 memainkan peran penting dalam menghasilkan prediksi untuk deteksi objek. Ini menghasilkan tiga peta fitur skala yang sesuai dengan grid berukuran 80x80, 40x40, dan 20x20. Ukuran grid yang berbeda ini memungkinkan deteksi objek kecil, sedang, dan besar, mengakomodasi berbagai skala dalam gambar input[9].

Dalam setiap sel grid, dihasilkan tiga bingkai prediksi, mencakup skor kepercayaan untuk keberadaan objek dan informasi posisional. Bingkai prediksi ini berfungsi untuk melokalisasi dan mengklasifikasikan objek yang ada di sel grid masing-masing. Skor kepercayaan mewakili keyakinan model terhadap keberadaan objek, sementara informasi posisional mencakup koordinat kotak pembatas seperti lebar, tinggi, dan pusat objek yang terdeteksi.

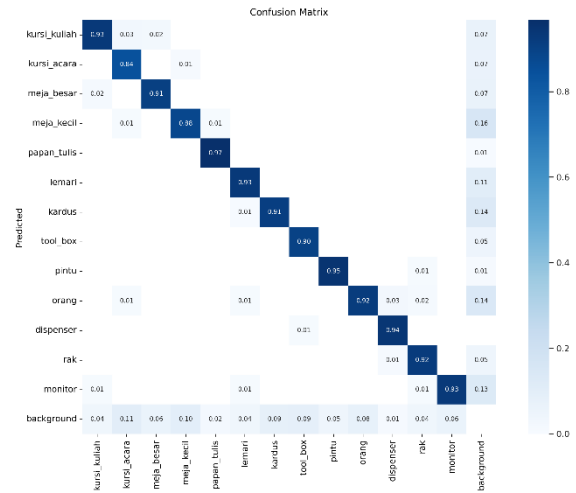
Setelah pembuatan bingkai prediksi, algoritma Non-Maximum Suppression (NMS) diterapkan untuk menghilangkan prediksi yang redundan. NMS memastikan hanya prediksi yang paling percaya diri dan tidak tumpang tindih yang dipertahankan, mencegah deteksi ganda dari objek yang sama dan menyempurnakan set prediksi akhir.

## III. HASIL DAN PEMBAHASAN

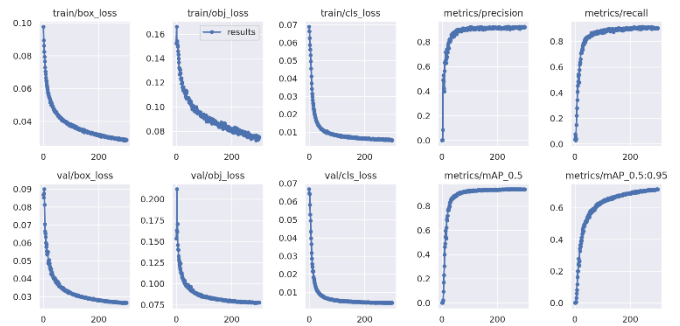
Pada bagian ini, kinerja model yang diusulkan dievaluasi pada dataset penyusutan Bunaken dan dibandingkan dengan pendeteksian objek berbasis YOLOv5 yang lebih ringan. Model ini memanfaatkan kinerja YOLO yang dapat melokalisasi objek secara real-time[10]. Penelitian ini juga menganalisis biaya efisiensi dan mengukur kecepatan model pada Jetson Nano.

Tabel 1 Perbandingan kinerja model yang diusulkan dan jaringan lain.

Model	mAP@50	mAP@50:95	Gflops
Yolov3-tiny	0.926	0.667	12.9
<b>Yolov5n</b>	<b>0.936</b>	<b>0.708</b>	<b>4.2</b>
Yolov5s	0.938	0.747	15.9
Yolov7-tiny	0.945	0.73	13.3



Gambar 2 Confusion matrices



Gambar 3 Reslut

### A. Evaluasi pada Dataset

Model yang diusulkan dalam penelitian ini telah berhasil mencapai kinerja yang signifikan dalam evaluasi presisi objek pada rentang IoU 0,5 (mAP@50) hingga 0,93 (mAP@50:95). Dibandingkan dengan model-model yang tercantum pada Tabel 1, YOLOv5n, meskipun memperoleh mAP@50 sebesar 0.708 yang sedikit lebih rendah daripada YOLOv5s, berhasil mencapai kecepatan komputasi yang lebih tinggi dengan Gflops sebesar 4.2.

Pentingnya kinerja tinggi dengan konsumsi daya yang efisien telah membuat YOLOv5n menjadi pilihan yang menarik dalam pengembangan sistem deteksi objek. Meskipun akurasi sedikit lebih rendah dibandingkan dengan beberapa model lain yang tercantum dalam tabel, pengurangan kompleksitas komputasi dapat menjadi faktor penentu dalam situasi di mana kecepatan eksekusi sangat kritis[11].

Selain itu, penting untuk dicatat bahwa evaluasi kinerja sebuah model tidak hanya bergantung pada nilai mAP, tetapi juga mempertimbangkan faktor-faktor seperti kecepatan inferensi dan kebutuhan sumber daya. Oleh karena itu, YOLOv5n, dengan keseimbangan antara akurasi dan kecepatan komputasi yang dihasilkan, dapat menjadi pilihan yang optimal dalam berbagai skenario aplikasi deteksi objek.





#### IV. KESIMPULAN DAN SARAN

Dalam perancangan sistem deteksi objek pada ruangan laboratorium, YOLOv5n telah berhasil mencapai kinerja optimal. Dengan tingkat frame per detik (FPS) mencapai 14,31, YOLOv5n menunjukkan peningkatan yang signifikan dibandingkan dengan arsitektur deteksi objek lainnya.

Kecepatan frame yang tinggi menjadi indikator utama efisiensi sistem, dan hasil ini menegaskan bahwa YOLOv5n dapat diandalkan untuk mendeteksi objek secara cepat, terutama pada perangkat dengan keterbatasan daya komputasi.

Penelitian ini menghasilkan temuan yang menarik, dengan YOLOv5n mencapai nilai mAP sebesar 61,6%. Keunggulan ini menjadikannya lebih baik dibandingkan dengan arsitektur sejenis seperti YOLOv5s, YOLOv3-tiny, dan YOLOv7-tiny. Keberhasilan tidak hanya terletak pada tingkat akurasi yang tinggi, tetapi juga pada efisiensi model. YOLOv5n dapat beroperasi pada kecepatan waktu nyata 14,31 FPS bahkan pada perangkat kelas bawah, mengukuhkannya sebagai pilihan optimal untuk aplikasi deteksi objek dalam ruangan laboratorium yang memiliki keterbatasan sumber daya komputasi.

#### KUTIPAN

- [1] L. Ali, F. Alnajjar, M. M. A. Parambil, M. I. Younes, Z. I. Abdelhalim, and H. Aljassmi, "Development of YOLOv5-Based Real-Time Smart Monitoring System for Increasing Lab Safety Awareness in Educational Institutions.," *Sensors (Basel)*, vol. 22, no. 22, Nov. 2022, doi: 10.3390/s22228820.
- [2] J. Wang *et al.*, "Research on Improved YOLOv5 for Low-Light Environment Object Detection," *Electronics*, vol. 12, no. 14. 2023. doi: 10.3390/electronics12143089.
- [3] E. M. Ditria, S. Lopez-Marcano, M. Sievers, E. L. Jinks, C. J. Brown, and R. M. Connolly, "Automating the Analysis of Fish Abundance Using Object Detection: Optimizing Animal Ecology With Deep Learning.," *Frontiers in Marine Science*, vol. 7. 2020. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fmars.2020.00429>
- [4] H. Leo, F. Arnia, and K. Munadi, "Fine Tuning CNN Pre-trained Model Based on Thermal Imaging for Obesity Early Detection.," *J. Rekamaya Elektr.*, vol. 18, Apr. 2022, doi: 10.17529/jre.v18i1.25100.
- [5] K. O. P. P. Nugraha and A. P. Rifai, "Convolutional Neural Network for Identification of Personal Protective Equipment Usage Compliance in Manufacturing Laboratory.," *J. Ilm. Tek. Ind.*, vol. 22, no. 1, pp. 11–24, 2023, doi: 10.23917/jiti.v22i1.21826.
- [6] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions.," *J. Big Data*, vol. 8, no. 1, p. 53, 2021, doi: 10.1186/s40537-021-00444-8.
- [7] Q. Hou, "Coordinate Attention for Efficient Mobile Network Design".
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition.," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8691 LNCS, no. PART 3, pp. 346–361, 2014, doi: 10.1007/978-3-319-10578-9\_23.
- [9] Z. Li, "Road Aerial Object Detection Based on Improved YOLOv5.," *J. Phys. Conf. Ser.*, vol. 2171, no. 1, 2022, doi: 10.1088/1742-6596/2171/1/012039.
- [10] S. Jalal, H. Ahmed, and M. Ahmed, *Design a Robust Real-Time Trash Detection System Using YOLOv5 Variants*. 2023. doi: 10.1109/GlobConET56651.2023.10149899.
- [11] W. Qiao, H. Guo, and E. Huang, "2023-Real-time Detection of Slug Flow in Subsea Pipelines by Embedding Yolo Object Detection Algorithm into Jetson Nano.," *J. Mar. Sci. Eng.*, vol. 11, Aug. 2023, doi: 10.3390/jmse11091658.
- [12] Y. Li, L. He, M. Zhang, Z. Cheng, W. Liu, and Z. Wu, "Improving the Performance of the Single Shot Multibox Detector for Steel Surface Defects with Context Fusion and Feature Refinement.," *Electronics*, vol. 12, p. 2440, May 2023, doi: 10.3390/electronics12112440.
- [13] T. Shi, Y. Ding, and W. Zhu, "YOLOv5s\_2E: Improved YOLOv5s for Aerial Small Target Detection.," *IEEE Access*, vol. PP, p. 1, Jan. 2023, doi: 10.1109/ACCESS.2023.3300372.



Penulis Bernama lengkap Dion Omong, anak kedua dari tiga bersaudara. Lahir di Desa Tolombukan Satu pada tanggal 02 April 2001. Penulis menempuh Pendidikan di TK Pertiwi Desa Liwutung (2006-2007), selanjutnya SD Liwutung Desa

Tolmbukan Barat (2007-2013), selanjutnya SMP 2 Ratahan (2013-2016) dan menyelesaikan sekolah tingkat atas di SMK 2 Ratahan (2016-2019). Tahun 2019 penulis melanjutkan studi di Fakultas Teknik, Jurusan Teknik Elektro, Universitas Sam Ratulangi, Manado. Pada tahun 2021 penulis memilih konsentrasi minat Teknik Kendali. Penulis melaksanakan kerja praktker di ULPTP Lahendong, Kota Tomohon selama 4 bulan dari tanggal 26 Juli 2022 sampai 24 Oktober 2022 dan melaksakan Kuliah Kerja Terpadu Angkatan 134 di Desa Tumpa Baru Kecamatan Tumpa, Kabupaten Minahasa Selatan. Selama studi di Fakultas Teknik Universitas Sam Ratulangi Jurusan Teknik Elektro, Penulis merupakan anggota aktif Himpunan Mahasiswa Elektro (HME) Unsrat, UKM Edukasi Robotika (EURO), dan ketua Control Engineering Community (CEC) 2022-2023, penulis juga pernah menjadi asisten Praktikum Sistem Kendali.