

# DETECTING SMOKING ACTIVITY BEHAVIOR USING YOLOV8 AND YOLOV11

## DETEKSI PERILAKU KEGIATAN MEROKOK MENGGUNAKAN YOLOV8 DAN YOLOV11

Salsabilla Azahra Putri<sup>1</sup>, Murinto<sup>\*2</sup>, Sunardi<sup>3</sup>

<sup>1,2,3</sup> Program Studi Magister Informatika, Universitas Ahmad Dahlan, Indonesia

<sup>1</sup>salsabillaazahra06@gmail.com, <sup>2</sup>murintokusno@tif.uad.ac.id, <sup>3</sup>sunardi@mti.uad.ac.id

**Abstract** - Smoking behavior in public spaces remains a major challenge in the implementation of public health policies, particularly within designated smoke-free zones. This study aims to examine whether architectural improvements and spatio-temporal modeling in object detection models can enhance the accuracy of real-time smoking behavior detection. Specifically, the performance of YOLOv8 and an experimental version, YOLOv11, is compared using a vision-based approach. A dataset of 3,000 annotated images is used, consisting of smoking and non-smoking activities such as drinking or phone use, with variations in lighting, body posture, and camera angles. The dataset was divided into 80% for training, 20% for validation, and 20% for testing, with data augmentation applied to improve generalization. YOLOv11 incorporates spatio-temporal modules and attention mechanisms not present in YOLOv8. Evaluation results show that YOLOv11 outperforms YOLOv8, achieving a Precision of 0.95, Recall of 0.91, and F1-Score of 0.93, while YOLOv8 reached 0.89, 0.87, and 0.88 respectively. These findings indicate that YOLOv11 offers a more robust and adaptive solution for automatically recognizing smoking behavior in real-world environments and supports the development of intelligent surveillance systems for enforcing smoke-free policies.

**Keywords** - Deep Learning, Object Detection, Public Surveillance, Smoking, YOLO.

**Abstrak** - Perilaku merokok di ruang publik merupakan tantangan signifikan dalam implementasi kebijakan kesehatan masyarakat, khususnya di kawasan tanpa rokok. Penelitian ini bertujuan untuk menguji apakah peningkatan arsitektur dan pemodelan spasial-temporal pada model deteksi objek dapat meningkatkan akurasi deteksi perilaku merokok secara real-time. Penelitian ini membandingkan performa YOLOv8 dan versi eksperimental YOLOv11 dengan pendekatan vision-based menggunakan dataset sebanyak 3.000 citra beranotasi. Dataset mencakup aktivitas merokok dan non-merokok seperti minum atau menggunakan ponsel dalam berbagai kondisi pencahayaan, posisi tubuh, dan sudut kamera. Data dibagi menjadi 80% untuk pelatihan, 20% validasi, dan 20% pengujian, serta dilakukan augmentasi untuk meningkatkan generalisasi model. YOLOv11 dirancang dengan modul spasial-temporal dan mekanisme perhatian (attention mechanism) yang tidak tersedia pada YOLOv8. Hasil evaluasi menunjukkan bahwa YOLOv11 unggul dengan Precision sebesar 0.95, Recall 0.91, dan F1-Score 0.93, sedangkan YOLOv8 hanya mencapai Precision 0.89, Recall 0.87, dan F1-Score 0.88. Temuan ini menunjukkan bahwa YOLOv11 lebih unggul dalam mengenali perilaku merokok secara otomatis dan adaptif di lingkungan nyata, serta berpotensi mendukung sistem pengawasan cerdas dalam menegakkan kebijakan kawasan tanpa rokok.

**Kata Kunci** - Deteksi Objek, Deep Learning, Merokok, Pengawasan Publik, YOLO.

## I. PENDAHULUAN

Perilaku merokok di ruang publik menimbulkan dampak serius terhadap kesehatan masyarakat, baik bagi perokok aktif maupun pasif. Untuk mengurangi risiko tersebut, pemerintah telah menetapkan regulasi Kawasan Tanpa Rokok (KTR) melalui berbagai kebijakan dan peraturan daerah. Namun, implementasi kebijakan ini di lapangan masih menghadapi banyak kendala, terutama dalam hal pengawasan yang konsisten dan efektif [1]. Salah satu tantangan utama dalam penegakan KTR adalah keterbatasan sumber daya manusia (SDM) dan tingginya mobilitas masyarakat di ruang publik [2]. Petugas pengawas tidak selalu mampu menjangkau seluruh area secara real-time, sehingga dibutuhkan solusi yang efisien, adaptif, dan beroperasi secara otomatis untuk mendukung pengawasan lapangan.

Kemajuan teknologi kecerdasan buatan (Artificial Intelligence/AI) dan pengolahan citra (computer vision) menawarkan potensi besar untuk mengatasi masalah ini [3]. Sistem pengawasan otomatis (pengawasan cerdas) berbasis kamera telah digunakan dalam berbagai konteks, termasuk deteksi perilaku manusia, dengan memanfaatkan pembelajaran mesin (machine learning) dan pembelajaran mendalam (deep learning) untuk mengidentifikasi objek dan aktivitas secara akurat dan waktu nyata [4],[5]. Deteksi perilaku merokok secara otomatis merupakan tantangan tersendiri. Rokok berukuran kecil, sering kali tertutup tangan atau benda lain, dan sulit dibedakan dari aktivitas serupa seperti minum atau menggunakan ponsel [6]. Oleh karena itu, dibutuhkan model deteksi visual yang mampu mengenali pola-pola perilaku tersebut secara presisi dalam berbagai kondisi visual. Salah satu pendekatan populer yang digunakan adalah model deteksi objek berbasis *You Only Look Once* (YOLO), yang dikenal karena kecepatannya dan akurasinya. Perkembangan YOLO dari versi awal hingga versi terbaru seperti YOLOv8 dan YOLOv11 menunjukkan peningkatan signifikan dalam hal stabilitas dan performa [7]–[9]. YOLOv8 dikenal andal dalam berbagai kondisi pencahayaan dan sudut kamera, sementara YOLOv11 menghadirkan pembaruan arsitektur, kemampuan spasial-temporal, serta ketahanan terhadap gangguan visual yang kompleks [10]–[12].

Penelitian ini bertujuan untuk membandingkan performa YOLOv8 dan YOLOv11 dalam mendeteksi perilaku merokok di ruang publik. Sebanyak 3.000 citra beranotasi digunakan, mencakup dua kategori: merokok dan non-merokok (misalnya, minum atau menggunakan ponsel). Evaluasi dilakukan berdasarkan metrik akurasi, presisi, *recall*, dan *F1-score* untuk menilai keunggulan dan keterbatasan masing-masing model. Hasil dari studi ini diharapkan dapat menjadi referensi dalam pengembangan sistem pengawasan cerdas yang mendukung penegakan regulasi KTR secara efisien dan adaptif di lingkungan nyata.

## II. SIGNIFIKASI STUDI

### A. Tinjauan Literatur

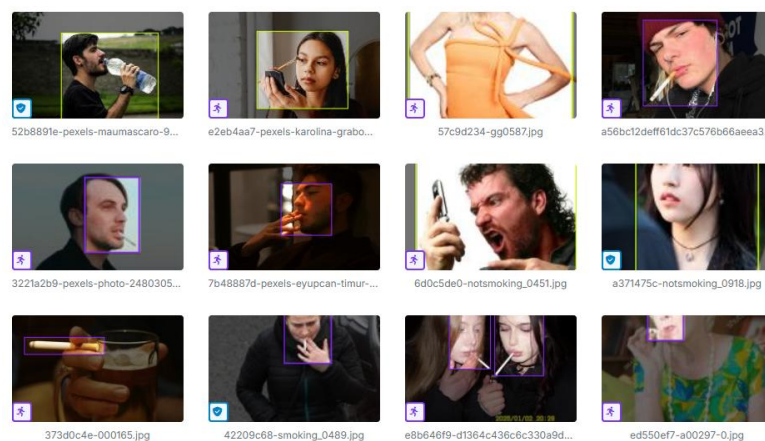
Penelitian ini memiliki kontribusi penting dalam dua dimensi utama, yaitu dimensi praktis yang berkaitan dengan penegakan kebijakan publik, serta dimensi ilmiah yang berfokus pada pengembangan metode deteksi perilaku berbasis visi komputer. Dari sisi praktis, sistem deteksi perilaku merokok secara otomatis yang dikembangkan dalam penelitian ini berpotensi menjadi solusi konkret untuk mendukung pelaksanaan kebijakan KT. Deteksi aktivitas merokok secara real-time melalui rekaman CCTV memungkinkan institusi seperti pemerintah daerah, rumah sakit, sekolah, serta pengelola fasilitas publik lainnya untuk memantau tingkat kepatuhan masyarakat terhadap regulasi tanpa harus menambah beban tenaga pengawas. Efektivitas sistem ini menjadi sangat relevan terutama di area publik dengan intensitas pergerakan manusia yang tinggi, seperti terminal, stasiun, fasilitas pendidikan, dan perkantoran [1]. Sistem ini juga dapat diintegrasikan dengan notifikasi otomatis atau sistem pelaporan berbasis Internet of Things (IoT), sehingga mampu memberikan respon cepat dan efisien terhadap pelanggaran yang terjadi [13].

Secara ilmiah, penelitian ini berkontribusi terhadap pengembangan metode deteksi perilaku manusia, terutama dalam konteks situasi visual kompleks dan objek kecil yang seringkali tidak terlihat secara jelas dalam tangkapan kamera. Rokok sebagai objek target memiliki ukuran yang kecil dan sering kali tertutup oleh tangan atau wajah pengguna, sehingga menyulitkan proses deteksi. Selain itu, aktivitas merokok memiliki kemiripan visual dengan aktivitas lain seperti minum atau menggunakan ponsel, yang dapat meningkatkan potensi kesalahan klasifikasi [14]. Untuk mengatasi tantangan tersebut, penelitian ini membandingkan dua model deteksi berbasis deep learning, yaitu YOLOv8 dan YOLOv11. YOLOv8 telah dikenal karena akurasi dan kecepatannya dalam tugas deteksi objek, sedangkan YOLOv11 menghadirkan sejumlah pembaruan arsitektur yang signifikan, seperti integrasi modul spasial-temporal, penggunaan attention mechanism, serta pendekatan adaptive label assignment untuk memperkuat akurasi prediksi dalam skenario dunia nyata [15]. Temuan dalam penelitian ini menunjukkan bahwa pembaruan arsitektur pada YOLOv11 mampu meningkatkan ketahanan model terhadap gangguan visual dan variasi sudut pandang, serta memberikan akurasi deteksi yang lebih tinggi terhadap perilaku merokok yang tidak eksplisit. Oleh karena itu, penelitian ini tidak hanya menawarkan solusi teknis terhadap permasalahan pemantauan perilaku di ruang publik, tetapi juga memperluas kontribusi ilmiah dalam bidang deteksi perilaku berbasis penglihatan mesin [16].

### B. Metode Penelitian

Penelitian ini menerapkan pendekatan kuantitatif dengan metode eksperimen komputasional untuk membandingkan performa dua model deteksi objek berbasis deep learning, yaitu YOLOv8 dan YOLOv11, dalam mengidentifikasi perilaku merokok secara otomatis melalui citra visual. Dataset yang digunakan terdiri atas 3.000 citra beranotasi yang diperoleh dari berbagai sumber terbuka, seperti platform Kaggle, media sosial, kanal video daring seperti YouTube, serta dokumentasi langsung dari lingkungan publik. Citra-citra tersebut mencerminkan berbagai kondisi nyata yang kompleks, termasuk variasi pencahayaan, latar belakang, postur tubuh, serta sudut pengambilan gambar. Sebagian gambar juga menampilkan perilaku merokok yang tidak eksplisit, misalnya rokok tertutup oleh tangan atau wajah tidak menghadap kamera.

Untuk mendukung pelatihan model deteksi, setiap citra dianotasi secara manual menggunakan perangkat lunak Labellmg dan Roboflow. Proses pelabelan mencakup pemberian bounding box serta label kelas berupa *smoking* dan *not smoking*. Label ini diterapkan dengan mengikuti pedoman klasifikasi yang ketat, terutama dalam menghadapi kondisi ambigu seperti pencahayaan minim atau objek yang tertutup sebagian. Format hasil anotasi disimpan dalam ekstensi teks (.txt) sesuai standar YOLO, yang berisi informasi kelas serta koordinat objek yang telah dinormalisasi. Proses pelabelan tersebut diperlihatkan pada Gambar 2, yang menampilkan hasil anotasi pada gambar dengan dua kelas perilaku.



Gambar 2. Proses Pelabelan

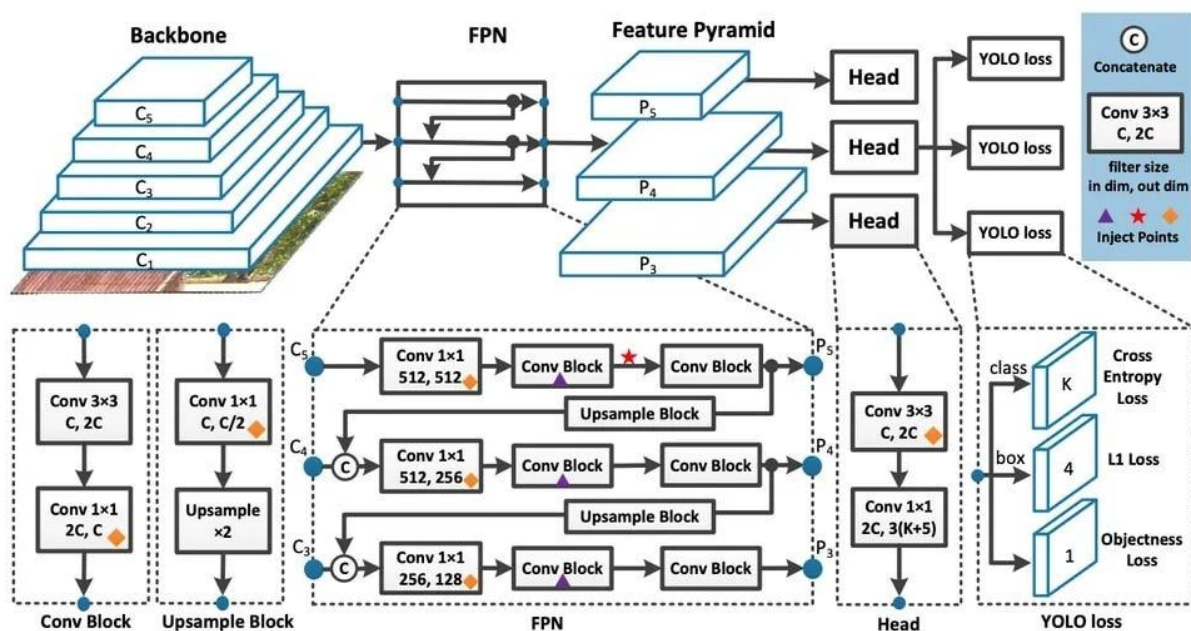
Dataset dibagi menjadi tiga subset data menggunakan metode stratifikasi, yaitu 2.400 citra (80%) untuk pelatihan, 300 citra (10%) untuk validasi, dan 300 citra (10%) untuk pengujian. Teknik augmentasi gambar diterapkan untuk memperluas keragaman data pelatihan, meliputi rotasi antara  $-15^\circ$  hingga  $+15^\circ$ , flipping horizontal dan vertikal, shear  $\pm 10^\circ$ , penyesuaian brightness dan hue  $\pm 15\%$ , serta efek blur hingga radius 2.5 piksel. Seluruh gambar kemudian di-resize ke dimensi  $640 \times 640$  piksel dengan orientasi otomatis guna memastikan konsistensi masukan ke model.

Pelatihan dilakukan secara independen pada dua model, yaitu YOLOv8 dan YOLOv11, menggunakan framework PyTorch dan pustaka Ultralytics di lingkungan GPU berbasis CUDA. Parameter pelatihan yang digunakan untuk kedua model dirinci dalam Tabel 1, yang mencakup jumlah epoch, ukuran citra, dan batch size. Augmentasi lanjutan seperti mixup, cutout, dan random perspective juga diterapkan selama proses pelatihan untuk meningkatkan kemampuan generalisasi. Evaluasi dilakukan setiap epoch menggunakan data validasi, dan model terbaik disimpan berdasarkan nilai mean Average Precision (mAP) tertinggi.

TABEL I  
PARAMETER PELATIHAN MODEL

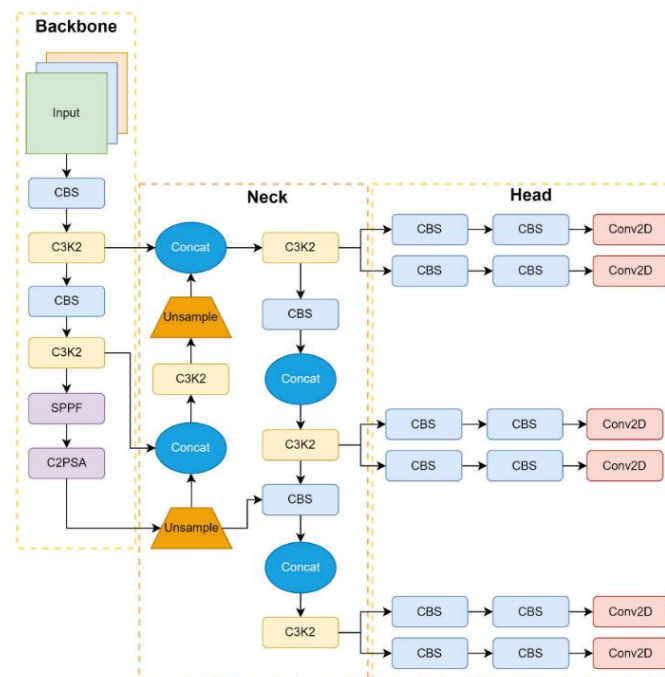
| Parameter  | Nilai |
|------------|-------|
| Epoch      | 100   |
| Image Size | 640   |
| Batch Size | 64    |

Dua model yang digunakan dalam penelitian ini, YOLOv8 dan YOLOv11, merupakan arsitektur deteksi objek satu tahap yang efisien dan cocok diterapkan pada sistem surveillance berbasis edge maupun cloud. YOLOv8, sebagaimana divisualisasikan pada Gambar 3, mengadopsi desain modular dengan tiga komponen utama: Backbone C2f, Neck PANet modifikasi, dan Head anchor-free. Desain ini memungkinkan pemrosesan yang cepat tanpa mengorbankan akurasi. YOLOv8 juga mendukung auto-anchor learning, adaptive loss, serta kemampuan pelacakan dan segmentasi objek secara real-time, menjadikannya pilihan yang andal untuk deteksi objek kecil seperti rokok dalam berbagai kondisi pencahayaan dan latar [17].



Gambar 3. Arsitektur YOLOv8 [18]

Sementara itu, YOLOv11 merupakan iterasi lanjutan yang dirancang untuk menghadirkan peningkatan akurasi dan stabilitas dalam skenario kompleks. Seperti ditunjukkan pada Gambar 4, arsitekturnya mencakup Backbone Transformer-enhanced CSPNet, yang meningkatkan pemahaman spasial dan semantik citra, Neck CSPFusion dan BiFPN, serta Head prediktif dengan adaptive label assignment. Inovasi utama pada YOLOv11 mencakup integrasi attention mechanism untuk memfokuskan perhatian model pada area penting dalam citra, kemampuan spasial-temporal untuk memahami hubungan antar objek dalam urutan visual, serta penerapan augmentasi pencahayaan dinamis dan blur adaptif untuk meningkatkan ketahanan terhadap noise visual. Arsitektur ini membuat YOLOv11 lebih unggul dalam mengenali perilaku merokok yang tidak eksplisit, seperti saat rokok tersembunyi sebagian atau pengguna tidak berorientasi langsung ke kamera [18], [19].



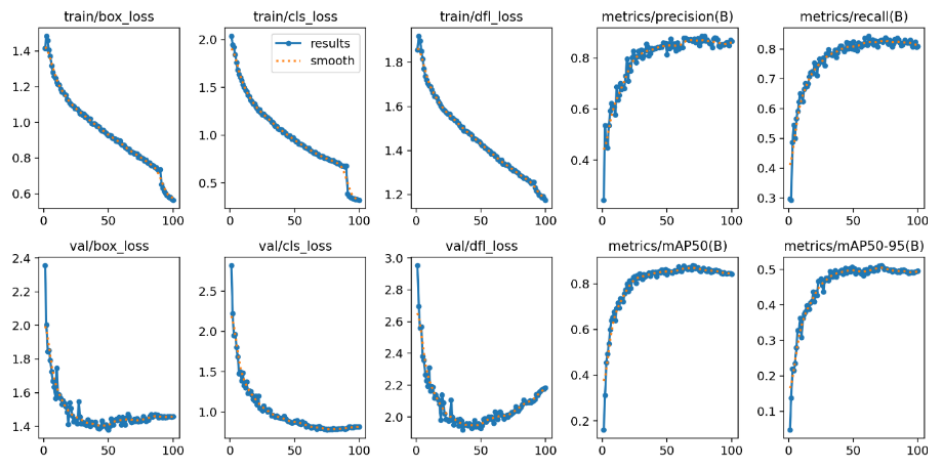
Gambar 4. Arsitektur YOLOv11

Dengan pendekatan metodologis ini, penelitian bertujuan memberikan gambaran komprehensif mengenai efektivitas masing-masing model dalam tugas deteksi perilaku merokok, serta mengidentifikasi kontribusi arsitektural pada performa deteksi secara keseluruhan.

### III. HASIL DAN PEMBAHASAN

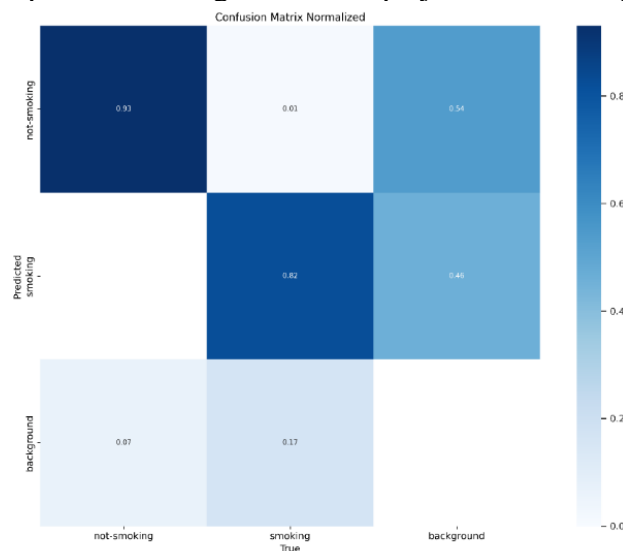
#### A. Hasil YOLOv8

Hasil pelatihan model YOLOv8 menunjukkan konvergensi stabil selama 100 epoch, dengan penurunan progresif pada metrik *train/box\_loss*, *train/cls\_loss*, dan *train/dfl\_loss*. Metrik validasi juga menunjukkan pola penurunan yang konsisten, menandakan proses pelatihan berhasil membentuk representasi fitur yang generalis terhadap data baru. Seperti ditunjukkan pada Gambar 5, precision mencapai 0.89 dan recall sebesar 0.87, sedangkan nilai *mean Average Precision* (mAP) pada *IoU threshold* 0.5 mencapai 0.91, dan mAP@[0.5:0.95] sebesar 0.73. Ini menandakan bahwa model mampu mendeteksi perilaku merokok dengan akurasi tinggi di berbagai kondisi spasial, meskipun dalam kasus tertentu masih terjadi kesalahan klasifikasi.



Gambar 1. Hasil Evaluasi Model YOLOv8

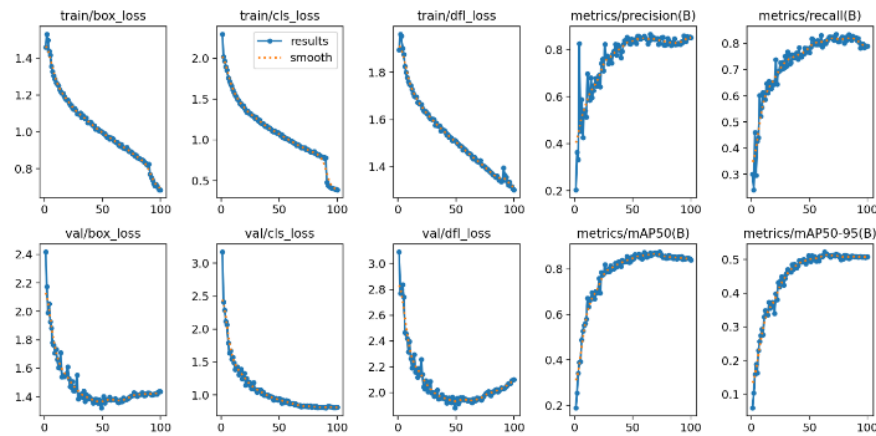
Analisis confusion matrix pada Gambar 6 menunjukkan bahwa YOLOv8 memiliki tingkat akurasi prediksi 93% pada kelas *not smoking* dan 82% pada kelas *smoking*. Namun, masih terdapat beberapa false positive, terutama saat individu memegang objek kecil seperti sedotan atau pena, yang secara visual menyerupai rokok. Selain itu, pencahayaan ekstrem atau posisi rokok yang tertutup tangan juga menyebabkan kesalahan deteksi. Untuk memperkuat performa, beberapa strategi lanjutan dapat diterapkan, seperti optimasi *confidence threshold*, penguatan augmentasi terhadap citra low-light atau kompleks, serta pelabelan ulang untuk memperjelas citra ambigu.



Gambar 2. Confusion Matrix YOLOv8

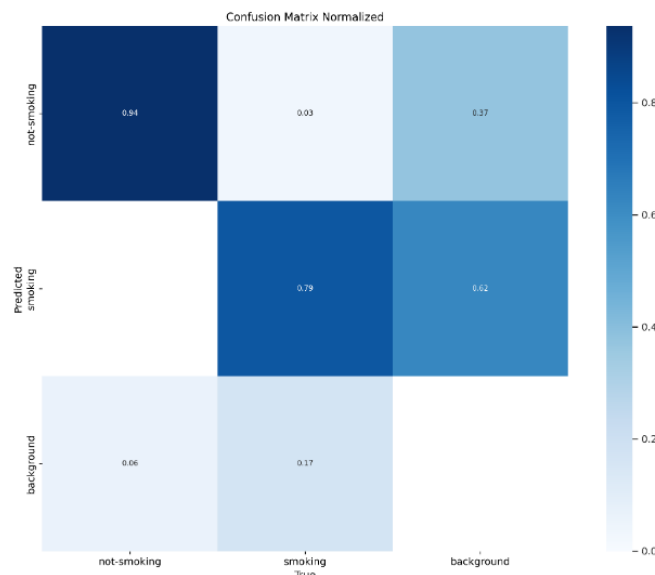
### B. Hasil YOLOv11

Model YOLOv11 menunjukkan peningkatan performa yang signifikan dan stabil selama pelatihan. Penurunan metrik *train/box\_loss* dan *val/box\_loss* hingga titik konvergen mengindikasikan bahwa model tidak mengalami overfitting dan memiliki kemampuan generalisasi yang baik. Sebagaimana ditampilkan pada Gambar 7, nilai precision mencapai 0.95 dan recall 0.91. mAP@0.5 meningkat menjadi 0.93 dan mAP@0.5:0.95 menjadi 0.76. Peningkatan ini menunjukkan bahwa YOLOv11 lebih unggul dalam mendeteksi objek target secara tepat di berbagai kondisi tumpang tindih (*overlap*), bahkan dalam situasi visual yang kompleks.



Gambar 3. Hasil Evaluasi Model YOLOv11

Performa superior YOLOv11 dapat dijelaskan dari aspek arsitekturalnya. Modul *Transformer-enhanced CSPNet* dan *BiFPN* memungkinkan model menangkap hubungan spasial dan semantik yang lebih dalam, sehingga mampu mengenali pola perilaku merokok meskipun sebagian objek tertutup atau tampak dari sudut ekstrem. *Attention mechanism* yang terintegrasi juga membantu model untuk memfokuskan perhatian pada area penting seperti tangan atau wajah, meningkatkan akurasi klasifikasi objek kecil seperti rokok.



Gambar 4. Confusion Matrix YOLOv11

Confusion matrix pada Gambar 8 menunjukkan bahwa YOLOv11 mencatat akurasi 94% pada kelas not smoking dan 79% pada kelas smoking. Meski akurasi kelas smoking masih dapat ditingkatkan, model ini secara umum berhasil mengurangi false positive yang sebelumnya lebih sering muncul pada YOLOv8. Dalam satu contoh, YOLOv8 salah mengklasifikasikan individu yang memegang sedotan sebagai smoking, sementara YOLOv11 berhasil mengidentifikasi aktivitas tersebut sebagai not smoking, berkat pemrosesan spasial-temporal yang lebih kuat.

### C. Perbandingan Kinerja YOLOv8 dan YOLOv11

Evaluasi kinerja YOLOv8 dan YOLOv11 dilakukan untuk menilai kemampuan masing-masing model dalam mendeteksi aktivitas merokok secara otomatis dan akurat. Analisis ini mencakup perbandingan metrik-metrik utama seperti precision, recall, F1-score, serta mAP@0.5 dan mAP@0.5:0.95. Dataset uji yang digunakan sama untuk kedua model guna menjamin konsistensi dan objektivitas hasil. Perbandingan kinerja YOLOv8 dan YOLOv11 dapat dilihat pada Tabel 2.

TABEL II  
PERBANDINGAN KINERJA YOLOv8 DAN YOLOv11

| Metrik       | YOLOv8 | YOLOv11 |
|--------------|--------|---------|
| Precision    | 0.89   | 0.95    |
| Recall       | 0.87   | 0.91    |
| F1-Score     | 0.88   | 0.93    |
| mAP@50       | 0.91   | 0.93    |
| mAP@0.5:0.95 | 0.73   | 0.76    |

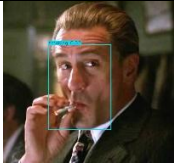
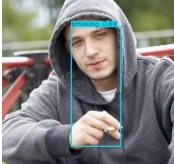
YOLOv11 unggul secara konsisten dalam seluruh metrik. Precision yang lebih tinggi menunjukkan kemampuan untuk mengurangi deteksi palsu, sedangkan recall dan F1-score yang lebih besar memperlihatkan sensitivitas model yang lebih baik terhadap objek target. Ini diperkuat oleh arsitektur YOLOv11 yang dilengkapi dengan adaptive label assignment dan modul pemrosesan visual tingkat lanjut.

Sementara YOLOv8 masih kompetitif dan lebih ringan untuk implementasi edge computing, YOLOv11 menawarkan akurasi dan ketahanan visual yang lebih tinggi, sehingga lebih layak untuk diterapkan di lingkungan pengawasan dengan tantangan visual tinggi seperti area publik terbuka, ruang semi-gelap, atau latar ramai.

#### D. Visualisasi Hasil Deteksi

Visualisasi berikut menyajikan perbandingan hasil deteksi antara model YOLOv8 dan YOLOv11 terhadap objek dengan kelas *smoking* dan *not smoking*. Masing-masing kolom menyajikan gambar asli, hasil deteksi YOLOv8, dan hasil deteksi YOLOv11. Observasi dilakukan untuk mengidentifikasi kualitas bounding box, akurasi klasifikasi, serta ketahanan model terhadap kompleksitas latar belakang.

TABEL III  
HASIL DETEKSI YOLOv8 & YOLOv11

| Deskripsi Kasus                                                                | Gambar Asli                                                                         | YOLOv8                                                                              | YOLOv11                                                                              | Kesimpulan                                                         |
|--------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------|
| Seseorang memegang rokok secara jelas di area terbuka                          |  |  |  | YOLOv8 & YOLOv11 berhasil deteksi kegiatan merokok                 |
| Dua orang dalam satu frame, hanya satu yang merokok.                           |  |  |  | YOLOv8 salah mendeteksi orang yang tidak merokok                   |
| Individu memegang rokok sebagian tertutup tangan                               |  |  |  | YOLOv8 & YOLOv11 berhasil deteksi kegiatan merokok tertutup tangan |
| Objek tidak merokok tetapi membawa benda serupa (misalnya sedotan atau pensil) |  |  |  | YOLOv8 & YOLOv11 berhasil deteksi kegiatan tidak merokok           |

Visualisasi deteksi mengungkapkan bahwa YOLOv11 memiliki ketangguhan lebih baik dalam mengenali konteks perilaku, bahkan ketika objek terlihat sebagian atau dalam kondisi pencahayaan rendah. Deteksi yang salah pada YOLOv8, seperti mengklasifikasikan jari tangan sebagai rokok, berhasil dihindari oleh YOLOv11 karena adanya attention mechanism dan semantic feature fusion yang lebih efektif.

#### IV. KESIMPULAN

Penelitian ini mengevaluasi dan membandingkan dua model deteksi objek berbasis deep learning, yakni YOLOv8 dan YOLOv11, dalam tugas mendeteksi perilaku merokok melalui citra dan video. Berdasarkan hasil pelatihan dan pengujian terhadap dataset beranotasi, YOLOv11 menunjukkan performa yang lebih unggul dibandingkan YOLOv8 dalam hampir semua metrik utama, dengan precision sebesar 0.95, recall 0.91, dan mAP@0.5:0.95 sebesar 0.76. Keunggulan ini didukung oleh arsitektur yang lebih kompleks dan adaptif, termasuk penggunaan attention mechanism dan modul spasial-temporal yang memungkinkan deteksi objek kecil seperti rokok secara lebih akurat meskipun dalam kondisi visual sulit. Meskipun demikian, penelitian ini masih memiliki keterbatasan, terutama pada cakupan variasi aktivitas non-merokok dalam dataset yang dapat menyebabkan false positive dalam konteks dunia nyata. Tantangan generalisasi terhadap latar dan sudut pandang ekstrem juga menunjukkan bahwa sistem ini perlu diuji lebih lanjut di lingkungan real-time. Oleh karena itu, pengayaan data pelatihan dan eksplorasi model ensemble atau fine-tuning domain-spesifik menjadi arah potensial untuk pengembangan selanjutnya. Dengan demikian, YOLOv11 dapat direkomendasikan sebagai model deteksi utama untuk sistem pemantauan cerdas berbasis CCTV dalam mendukung penegakan kebijakan Kawasan Tanpa Rokok (KTR) secara otomatis dan efisien di ruang publik.

#### REFERENSI

- [1] Fariz Kahendra, Bagoes Widjanarko, and Farid Agushybana, "Faktor-Faktor yang Mempengaruhi Implementasi Kebijakan Kawasan Tanpa Rokok : Literature Review," *Media Publ. Promosi Kesehat. Indones.*, vol. 6, no. 3, pp. 430–435, Mar. 2023.
- [2] R. A. A'yuni and N. Nasrullah, "Implementasi Peraturan Daerah Tentang Kawasan Tanpa Rokok Di Dinas Kesehatan Kota Yogyakarta," *Media Law Sharia*, vol. 2, no. 2, Apr. 2021.
- [3] I. Anas and S. Zakir, "Artificial Intelligence: Solusi Pembelajaran Era Digital 5.0," *J-SAKTI (Jurnal Sains Komput. dan Inform.)*, vol. 8, no. 1, p. 35, Mar. 2024.
- [4] Salsabilla Azahra Putri, Murinto, and Sunardi, "Implementasi Algoritma YOLOv5 untuk Otomatisasi Iklan Layananan Publik tentang Larangan Merokok," *J. Inform. Polinema*, vol. 11, no. 2, pp. 195–202, Feb. 2025.
- [5] R. Sha and T. Guha, "Detection of Smoking in inside the office Using Deep Learning for CCTV Camera," in *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2022, pp. 1–9.
- [6] J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Mach. Learn. Knowl. Extr.*, vol. 5, no. 4, pp. 1680–1716, Nov. 2023.
- [7] M. Hussain, "YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection," *Machines*, vol. 11, no. 7, p. 677, Jun. 2023.

- [8] R. Varghese and S. M., “YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness,” in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, 2024, pp. 1–6.
- [9] F. M. Talaat and H. ZainEldin, “An improved fire detection approach based on YOLO-v8 for smart cities,” *Neural Comput. Appl.*, vol. 35, no. 28, pp. 20939–20954, Oct. 2023.
- [10] L. He, Y. Zhou, L. Liu, and J. Ma, “Research and Application of YOLOv11-Based Object Segmentation in Intelligent Recognition at Construction Sites,” *Buildings*, vol. 14, no. 12, p. 3777, Nov. 2024.
- [11] L. Zhang, Z. Sun, H. Tao, M. Wang, and W. Yi, “Research on Mine-Personnel Helmet Detection Based on Multi-Strategy-Improved YOLOv11,” *Sensors*, vol. 25, no. 1, p. 170, Dec. 2024.
- [12] M. Mao and M. Hong, “YOLO Object Detection for Real-Time Fabric Defect Inspection in the Textile Industry: A Review of YOLOv1 to YOLOv11,” *Sensors*, vol. 25, no. 7, p. 2270, Apr. 2025.
- [13] Zaenuar Erfandi, D. Hartanti, and J. Maulindar, “Implementasi Internet of Things (IoT) Untuk Sistem Pemantauan Kebakaran Dini Dengan Notifikasi Telegram dan Alarm,” *Infotek J. Inform. dan Teknol.*, vol. 8, no. 1, pp. 86–93, Jan. 2025.
- [14] Z. Hu and C. Lv, *Vision-Based Human Activity Recognition*. Singapore: Springer Nature Singapore, 2022.
- [15] E. S. dos Santos Júnior, T. Paixão, and A. B. Alvarez, “Comparative Performance of YOLOv8, YOLOv9, YOLOv10, and YOLOv11 for Layout Analysis of Historical Documents Images,” *Appl. Sci.*, vol. 15, no. 6, p. 3164, Mar. 2025.
- [16] Y. Liao, L. Li, H. Xiao, F. Xu, B. Shan, and H. Yin, “YOLO-MECD: Citrus Detection Algorithm Based on YOLOv11,” *Agronomy*, vol. 15, no. 3, p. 687, Mar. 2025.
- [17] Y. Wang, Y. Piao, H. Wang, H. Zhang, and B. Li, “An Improved Forest Smoke Detection Model Based on YOLOv8,” *Forests*, vol. 15, no. 3, p. 409, Feb. 2024.
- [18] C. Zhang, H. Cheng, R. Wu, B. Ren, Y. Zhu, and N. Peng, “Development of a Traffic Congestion Prediction and Emergency Lane Development Strategy Based on Object Detection Algorithms,” *Sustainability*, vol. 16, no. 23, p. 10232, Nov. 2024.
- [19] C. Zou, S. Yu, Y. Yu, H. Gu, and X. Xu, “Side-Scan Sonar Small Objects Detection Based on Improved YOLOv11,” *J. Mar. Sci. Eng.*, vol. 13, no. 1, p. 162, Jan. 2025.

### UCAPAN TERIMA KASIH

Kami menyampaikan terima kasih yang sebesar-besarnya kepada Direktorat Penelitian dan Pengabdian kepada Masyarakat (DPPM) Kemdiktisaintek atas dukungan dana penelitian melalui kontrak nomor 126/C3/DT.05.00/PL/2025; 0498.12/LL5-INT/AL.04/2025; 083/PTM/LPPM.UAD/VI/2025. Kami juga mengucapkan terima kasih kepada Universitas Ahmad Dahlan atas dukungan yang telah diberikan.