

SISTEM DETEKSI MANUSIA PADA AREA TERBATAS MENGUNAKAN YOLOV7 BERBASIS CCTV

Ridho Bagus Abdillah¹, Yohanni Syahra²

ridhoabdillah890@gmail.com¹, yohannisyastra@umsu.ac.id²

Universitas Muhammadiyah Sumatera Utara

ABSTRAK

Pengawasan keamanan gedung kampus berbasis Closed-Circuit Television (CCTV) pada umumnya masih bergantung pada pemantauan manual sehingga rentan terhadap kelelahan operator dan keterlambatan respons. Penelitian ini bertujuan merancang dan mengimplementasikan sistem deteksi manusia otomatis pada area terbatas gedung kampus menggunakan You Only Look Once versi 7 (YOLOv7), serta mengevaluasi kinerjanya pada variasi pencahayaan dan jumlah objek. Zona pengawasan dibatasi dengan Region of Interest (ROI) berbentuk poligon yang dipadukan dengan parameter waktu operasional; pelanggaran ditetapkan apabila titik tengah bounding box manusia berada di dalam ROI di luar jam operasional. Sistem dibangun dengan Python, PyTorch, dan OpenCV menggunakan bobot YOLOv7 prelatih pada dataset COCO tanpa pelatihan ulang. Pengujian memakai 18.000 frame sampel dari 50 jam rekaman CCTV pada kondisi siang dan malam (inframerah). Hasil menunjukkan precision 99,47%, recall 98,61%, F1-score 99,04%, mAP@0,5 96,45%, dan kecepatan rata-rata 47 FPS pada GPU NVIDIA GeForce RTX 3050 Ti. Rerata confidence score menurun dari 87,4% (siang) menjadi 64,2% (malam), dan tingkat deteksi menurun seiring bertambahnya jumlah objek. Dari 150 kejadian pelanggaran, 142 (94,7%) teridentifikasi dengan waktu respons rerata 0,8 detik. Akurasi penentuan posisi terhadap ROI paling rendah di area perbatasan (94,40%). Hasil ini menunjukkan bahwa integrasi YOLOv7, ROI, dan parameter waktu layak digunakan sebagai alat bantu pemantauan otomatis pada lokasi uji.

Kata Kunci: Deteksi Manusia, YOLOv7, CCTV, Region Of Interest, Area Terbatas, Computer Vision.

ABSTRACT

Conventional Closed-Circuit Television (CCTV) surveillance in campus buildings still relies largely on manual monitoring, which is prone to operator fatigue and delayed response. This study designs and implements an automated human detection system for restricted campus areas using You Only Look Once version 7 (YOLOv7) and evaluates its performance under varying lighting conditions and numbers of people per frame. Surveillance zones are constrained by a polygonal Region of Interest (ROI) combined with an operating-hours parameter; a violation is registered when the center point of a person's bounding box lies inside the ROI outside operating hours. The system was built with Python, PyTorch, and OpenCV using pretrained YOLOv7 weights from the COCO dataset without retraining. Testing used 18,000 sampled frames from 50 hours of CCTV footage recorded in daytime and at night (infrared mode). The system achieved a precision of 99.47%, recall of 98.61%, F1-score of 99.04%, mAP@0.5 of 96.45%, and an average speed of 47 FPS on an NVIDIA GeForce RTX 3050 Ti GPU. The mean confidence score dropped from 87.4% (day) to 64.2% (night), and the detection rate declined as the number of people per frame increased. Of 150 violation events, 142 (94.7%) were identified with a mean response time of 0.8 s. Spatial classification accuracy was lowest near the ROI boundary (94.40%). These results indicate that combining YOLOv7, ROI, and time parameters is feasible as an automated monitoring aid at the single test site.

Keywords: Human Detection, YOLOv7, CCTV, Region Of Interest; Restricted Area, Computer Vision.

PENDAHULUAN

Pengawasan keamanan gedung kampus umumnya mengandalkan kamera Closed-

Circuit Television (CCTV) yang dipantau petugas secara manual. Cara ini memiliki keterbatasan dalam hal konsistensi, kelelahan operator, dan kecepatan respons, terutama pada area yang luas dan di luar jam operasional ketika jumlah petugas terbatas. Kemajuan computer vision dan deep learning memungkinkan sistem pengawasan berkembang dari sekadar perekam visual menjadi sistem cerdas yang menganalisis objek dan aktivitas secara otomatis.

Salah satu pendekatan yang banyak digunakan untuk deteksi objek real-time adalah You Only Look Once (YOLO), yaitu detektor satu tahap (one-stage) yang memprediksi lokasi dan kelas objek dalam satu kali proses inferensi (Redmon dkk., 2016). Pendekatan ini berbeda dari detektor dua tahap yang memisahkan pencarian wilayah objek dan klasifikasi (Zhao dkk., 2019). Keluarga YOLO terus berkembang, antara lain YOLOv4 (Bochkovskiy dkk., 2020) dan YOLOv7 yang dilaporkan melampaui detektor objek yang ada pada rentang kecepatan 5–160 FPS (Wang dkk., 2023), sehingga relevan untuk pengawasan berbasis CCTV yang menuntut pemrosesan secara langsung.

Deteksi manusia berbasis YOLO untuk pengawasan cerdas telah dikaji, misalnya oleh Nguyen dkk. (2021) yang merancang model deteksi manusia berbasis YOLO dengan keseimbangan antara akurasi dan waktu pemrosesan untuk perangkat tepi (edge). Namun, kajian semacam itu pada umumnya berfokus pada kinerja deteksi pada skenario umum. Sepengetahuan penulis, pembatasan zona pengawasan berdasarkan ruang dan waktu operasional, yang menentukan apakah kehadiran seseorang merupakan potensi pelanggaran, belum banyak dikaji secara terintegrasi dengan evaluasi pada berbagai kondisi pencahayaan.

Pada lingkungan kampus terdapat area yang menjadi zona terbatas di luar jam operasional, sehingga keberadaan manusia pada area tersebut pada waktu tertentu dapat dikategorikan sebagai potensi pelanggaran keamanan, misalnya akses tidak sah atau aktivitas mencurigakan. Metode Region of Interest (ROI) memungkinkan pengawasan dibatasi pada zona tertentu sehingga sistem tidak hanya mendeteksi manusia tetapi juga menentukan apakah manusia tersebut berada di dalam zona terbatas. Di sisi lain, variasi pencahayaan (siang dan malam), sudut kamera, dan kepadatan objek menjadi tantangan bagi sistem deteksi berbasis CCTV sehingga kinerjanya perlu dianalisis pada kondisi operasional yang berbeda.

Penelitian ini bertujuan (1) mengimplementasikan YOLOv7 pada sistem deteksi manusia berbasis CCTV untuk area terbatas gedung kampus di luar jam operasional yang dipadukan dengan ROI dan parameter waktu; (2) mengevaluasi kinerja sistem menggunakan precision, recall, F1-score, mean Average Precision (mAP), dan frame per second (FPS) pada variasi pencahayaan dan jumlah objek; serta (3) mengevaluasi kemampuan sistem mengidentifikasi pelanggaran pada area terbatas. Ruang lingkup penelitian dibatasi pada kelas person menggunakan model YOLOv7 pralatih, satu kamera CCTV dengan sudut dan posisi tetap pada satu gedung kampus, tanpa pengenalan wajah dan tanpa membedakan orang berizin dan tidak berizin.

METODE PENELITIAN

1. Rancangan Penelitian

Penelitian ini merupakan penelitian eksperimental kuantitatif dengan pendekatan implementatif, yaitu merancang, membangun, dan menguji sistem dengan hasil yang diukur secara numerik. Tahapan penelitian meliputi (1) studi literatur, (2) perancangan sistem, (3) implementasi, (4) pengambilan data dan pengujian, serta (5) analisis dan evaluasi hasil.

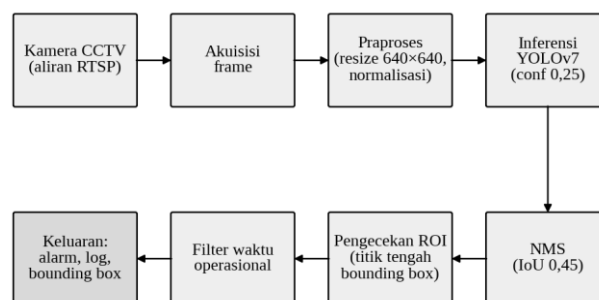
2. Lokasi dan Akuisisi Data

Penelitian dilaksanakan di area depan dan koridor Gedung Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI), Universitas Muhammadiyah Sumatera Utara, Medan. Area ini dipilih karena aksesnya terbatas setelah jam operasional sehingga berpotensi menjadi titik rawan aktivitas tidak berkepentingan. Jam operasional ditetapkan pukul 07.00–17.00 WIB; waktu di luar rentang tersebut dianggap sebagai waktu terbatas.

Data diperoleh dari kamera CCTV Hikvision DS-2CE16D0T-IR beresolusi 1920×1080 piksel pada 30 FPS dengan fitur inframerah (infrared, IR). Rekaman siang hari diambil pada pukul 09.00–15.00 WIB (citra RGB tanpa IR) dan rekaman malam hari pada pukul 20.00–23.00 WIB (mode IR, citra grayscale). Total rekaman berdurasi 50 jam, terdiri atas 25 jam siang dan 25 jam malam, dengan variasi aktivitas berupa orang berjalan sendiri, berjalan berkelompok, berhenti di area tertentu, dan melintas di sekitar batas ROI. Karena pengujian seluruh frame (secara teoretis 5.400.000 frame) tidak praktis, digunakan systematic sampling dengan mengambil satu frame setiap 10 detik sehingga diperoleh 18.000 frame sampel (9.000 siang dan 9.000 malam). Setiap frame divalidasi agar tidak rusak dan memuat informasi visual yang jelas.

3. Arsitektur Sistem

Alur sistem ditunjukkan pada Gambar 1. Video CCTV dibaca sebagai aliran RTSP menggunakan OpenCV, diubah ukurannya menjadi 640×640 piksel, dinormalisasi ke rentang $[0, 1]$, lalu diproses YOLOv7 pada GPU. Keluaran model berupa bounding box, confidence score, dan ID kelas disaring dengan ambang confidence 0,25 dan hanya kelas person (ID 0) yang diteruskan; selanjutnya NMS dengan ambang IoU 0,45 menghapus deteksi ganda. Deteksi kemudian diperiksa terhadap ROI dan filter waktu operasional, dengan keluaran berupa tampilan bounding box, peringatan teks dan alarm suara, serta log kejadian.



Gambar 1. Arsitektur dan alur pemrosesan sistem.

Sistem dibangun dengan Python 3.10.11, PyTorch 2.0.1 (CUDA 11.7), dan OpenCV 4.8.0 pada Windows 11 Pro. Model yang digunakan adalah bobot pralatih yolov7.pt dari repositori resmi YOLOv7 tanpa pelatihan ulang. Spesifikasi perangkat dan parameter model disajikan pada Tabel 1 dan Tabel 2.

Tabel 1. Spesifikasi perangkat keras dan lunak

Komponen	Spesifikasi
Prosesor	Intel Core i7-11800H, 2,30 GHz
RAM	8 GB DDR4
GPU	NVIDIA GeForce RTX 3050 Ti (4 GB VRAM), CUDA 11.7, cuDNN 8.5.0
Penyimpanan	SSD NVMe 512 GB
Kamera	Hikvision DS-2CE16D0T-IR, 1920×1080, IR
Perangkat lunak	Windows 11 Pro 22H2; Python 3.10.11; PyTorch 2.0.1; OpenCV 4.8.0

Tabel 2. Parameter model YOLOv7

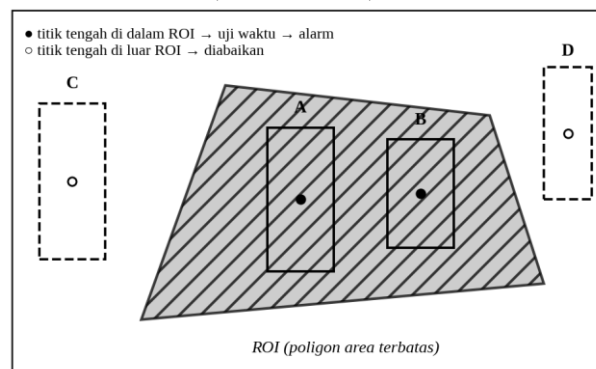
Parameter	Nilai
Model	YOLOv7 prlatih (yolov7.pt)
Dataset pelatihan	COCO (80 kelas)
Ukuran masukan	640 × 640 piksel
Ambang <i>confidence</i>	0,25
Ambang IoU (NMS)	0,45
Kelas target	0 (<i>person</i>)

4. ROI dan Logika Identifikasi Pelanggaran

ROI ditetapkan secara manual pada satu frame referensi. Peneliti mengklik titik-titik batas area terbatas menggunakan fungsi mouse event OpenCV sehingga terbentuk poligon tertutup yang koordinat pikselnya disimpan dalam berkas konfigurasi JSON agar dapat digunakan ulang. ROI mencakup area depan gedung dan koridor gedung. Untuk setiap manusia yang terdeteksi, titik tengah bounding box dihitung dengan Persamaan (1).

$$c = ((x_{min} + x_{max})/2, (y_{min} + y_{max})/2) \quad (1)$$

Titik tersebut diuji terhadap poligon ROI menggunakan algoritma ray casting atau fungsi `cv2.pointPolygonTest()` pada OpenCV. Objek dengan titik tengah di luar ROI diabaikan, sedangkan objek dengan titik tengah di dalam ROI diteruskan ke pemeriksaan waktu (Gambar 2). Suatu kejadian dinyatakan sebagai pelanggaran apabila tiga syarat terpenuhi bersamaan: (i) terdeteksi manusia, (ii) titik tengah bounding box berada di dalam ROI, dan (iii) waktu sistem berada di luar jam operasional. Pada kondisi ini sistem mengaktifkan alarm dan menyimpan log yang memuat timestamp, jumlah objek, koordinat bounding box, rerata confidence score, status ROI, dan status waktu operasional.



Gambar 2. Skema pengecekan titik tengah bounding box terhadap ROI. Objek A dan B berada di dalam ROI dan diteruskan ke uji waktu; objek C dan D berada di luar ROI dan diabaikan.

5. Skenario Pengujian dan Metrik Evaluasi

Pengujian dirancang secara faktorial dengan empat faktor: pencahayaan (siang, malam), jumlah objek per frame (1, 2, dan 3 orang atau lebih), posisi objek terhadap ROI (di dalam, di luar, dan di perbatasan), serta waktu operasional (dalam dan luar jam operasional). Kombinasinya menghasilkan 36 skenario yang masing-masing diulang tiga kali sehingga terdapat 108 percobaan. Selain itu dilakukan pengujian identifikasi pelanggaran pada 150 kejadian ketika objek berada di dalam ROI di luar jam operasional.

Status true positive (TP), false positive (FP), dan false negative (FN) ditentukan melalui verifikasi manual terhadap frame sampel dengan ambang IoU 0,5. Kinerja deteksi dihitung dengan Persamaan (2)–(4).

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (2)$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (3)$$

$$\text{F1} = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall}) \quad (4)$$

Metrik lain yang digunakan adalah mAP@0,5 (mAP pada IoU 0,5), FPS, rerata confidence score, tingkat deteksi (persentase frame berisi n orang yang berhasil terdeteksi), akurasi penentuan zona spasial terhadap ROI, dan waktu respons (selang waktu dari deteksi hingga alarm aktif).

HASIL DAN PEMBAHASAN

1. Implementasi dan Data Uji

Sistem berhasil direalisasikan dan mampu membaca aliran CCTV, mendeteksi manusia, menentukan posisinya terhadap ROI, mengevaluasi waktu operasional, mengaktifkan alarm, dan mencatat log. Waktu inferensi YOLOv7 pada GPU RTX 3050 Ti berkisar 18–22 ms per frame, dan kecepatan rata-rata seluruh alur pemrosesan adalah 47 FPS (48 FPS siang dan 46 FPS malam), lebih tinggi daripada laju video kamera (30 FPS) sehingga setiap frame dapat diproses tanpa antrean.

Tabel 3 merangkum data uji. Dari 18.000 frame sampel, 4.320 frame (24%) memuat manusia dengan total 5.760 objek. Proporsi frame berisi manusia pada siang hari (32%) lebih tinggi daripada malam hari (16%); angka ini terutama mencerminkan tingkat aktivitas di lokasi dan bukan ukuran kinerja deteksi.

Tabel 3. Distribusi data uji

Kondisi	Durasi (jam)	Frame sampel	Frame berisi manusia	Objek terdeteksi
Siang	25	9.000	2.880 (32%)	4.032
Malam	25	9.000	1.440 (16%)	1.728
Total	50	18.000	4.320 (24%)	5.760

2. Pengaruh Pencahayaan

Tabel 4 membandingkan hasil deteksi siang dan malam hari. Rerata confidence score turun dari 87,4% menjadi 64,2% (selisih 23,2 poin persentase), dan rentang nilainya bergeser dari 72–96% menjadi 45–81%. False positive meningkat dari 12 menjadi 18 kejadian (0,42% menjadi 1,25% dari frame berisi manusia) dan false negative dari 28 menjadi 52 kejadian (0,97% menjadi 3,61%).

Tabel 4. Perbandingan hasil deteksi siang dan malam

Parameter	Siang	Malam	Selisih
Rerata <i>confidence</i> (%)	87,4	64,2	-23,2
Rentang <i>confidence</i> (%)	72–96	45–81	–
Objek per frame (rerata)	1,4	1,2	-0,2
<i>False positive</i> (kejadian)	12 (0,42%)	18 (1,25%)	+6
<i>False negative</i> (kejadian)	28 (0,97%)	52 (3,61%)	+24

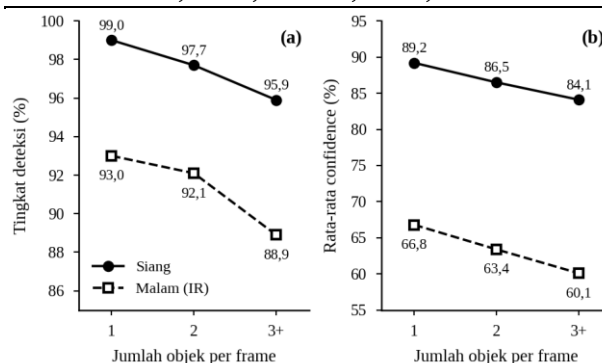
Persentase FP dan FN pada Tabel 4 dihitung terhadap frame berisi manusia (2.880 frame siang dan 1.440 frame malam). Penurunan kinerja pada malam hari diduga terkait dengan karakteristik citra IR yang bersifat grayscale, berkontras lebih rendah, dan tidak memuat informasi warna, sedangkan model dilatih pada COCO yang didominasi citra berwarna sehingga terdapat perbedaan distribusi piksel. False positive pada siang hari umumnya berasal dari bayangan pohon atau tiang yang menyerupai bentuk manusia, dan pada malam hari dari objek non-manusia berbentuk serupa pada citra IR. False negative berkaitan dengan objek yang jauh, tertutup sebagian, berada di luar sudut pandang kamera, atau bergerak cepat pada citra IR yang kurang jelas.

3. Pengaruh Jumlah Objek

Jumlah frame uji pada skenario 1, 2, dan 3+ objek berturut-turut adalah 1.200, 960, dan 720 (siang) serta 600, 480, dan 360 (malam). Hasilnya disajikan pada Tabel 5 dan Gambar 3. Tingkat deteksi turun dari 99,0% menjadi 95,9% pada siang hari, dan dari 93,0% menjadi 88,9% pada malam hari, ketika jumlah objek bertambah dari satu menjadi tiga atau lebih. Rerata *confidence* turun 5,1 poin pada siang dan 6,7 poin pada malam hari.

Tabel 5. Kinerja deteksi berdasarkan jumlah objek (S = siang, M = malam)

Jumlah objek	Tingkat deteksi S / M (%)	<i>Confidence</i> S / M (%)	FPS
1	99,0 / 93,0	89,2 / 66,8	52
2	97,7 / 92,1	86,5 / 63,4	48
3+	95,9 / 88,9	84,1 / 60,1	41



Gambar 3. Pengaruh jumlah objek per frame terhadap (a) tingkat deteksi dan (b) rerata *confidence* score.

Selisih siang–malam relatif konsisten pada seluruh skenario (22,4–24,0 poin untuk *confidence* dan 5,6–7,0 poin untuk tingkat deteksi), sehingga pengaruh pencahayaan tampak lebih dominan daripada jumlah objek, sedangkan kombinasi malam hari dan 3+ objek menghasilkan kinerja terendah (88,9% dan 60,1%). Penurunan akibat bertambahnya objek diduga disebabkan oleh oklusi parsial, bounding box yang berdekatan atau tumpang

tindih yang mempersulit NMS, dan kompleksitas visual yang meningkat. Kecepatan pemrosesan turun dari 52 FPS (satu objek) menjadi 41 FPS (tiga objek atau lebih), tetapi masih di atas 30 FPS.

4. Akurasi Penentuan Posisi terhadap ROI

Tabel 6. Akurasi penentuan zona spasial objek terhadap ROI

Posisi objek	Frame uji	Benar	Akurasi (%)
Di dalam ROI	2.160	2.148	99,44
Di luar ROI	2.160	2.155	99,77
Perbatasan ROI	500	472	94,40

Penentuan posisi sangat akurat untuk objek yang jelas berada di dalam (99,44%) maupun di luar ROI (99,77%) (Tabel 6). Akurasi menurun di area perbatasan (94,40%; 28 dari 500 frame salah) karena titik tengah bounding box berfluktuasi melintasi garis batas poligon akibat gerak langkah dan perubahan postur. Kondisi ini menunjukkan perlunya buffer zone atau temporal smoothing, yaitu konfirmasi pelanggaran hanya setelah objek berada di dalam ROI selama n frame berurutan, untuk menekan alarm palsu di sekitar batas zona.

5. Identifikasi Pelanggaran Area

Tabel 7. Hasil identifikasi pelanggaran (di dalam ROI, luar jam operasional)

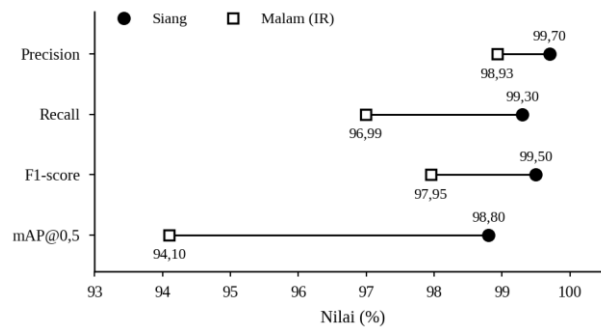
Parameter	Nilai
Total kejadian	150
Teridentifikasi	142 (94,7%)
Gagal teridentifikasi	8 (5,3%)
Rerata <i>confidence score</i>	72,3%
Rerata waktu respons	0,8 detik
Alarm aktif	142 (100% dari yang teridentifikasi)

Dari 150 kejadian objek di dalam ROI di luar jam operasional, 142 kejadian (94,7%) berhasil diidentifikasi sebagai pelanggaran dan seluruhnya memicu alarm dengan waktu respons rerata 0,8 detik (Tabel 7), yang memadai untuk keperluan peringatan hampir seketika. Rerata confidence sebesar 72,3% berada di antara rerata siang (87,4%) dan malam hari (64,2%). Delapan kejadian yang gagal (5,3%) berpotensi menjadi celah keamanan; penelaahan penyebabnya menjadi bahan penelitian lanjutan.

6. Evaluasi Kinerja Keseluruhan

Tabel 8. Rekapitulasi metrik evaluasi kinerja sistem

Metrik	Siang	Malam	Keseluruhan
TP	4.004	1.676	5.680
FP	12	18	30
FN	28	52	80
<i>Precision (%)</i>	99,70	98,93	99,47
<i>Recall (%)</i>	99,30	96,99	98,61
<i>F1-score (%)</i>	99,50	97,95	99,04
<i>mAP@0,5 (%)</i>	98,80	94,10	96,45
FPS	48	46	47



Gambar 4. Perbandingan precision, recall, F1-score, dan mAP@0,5 pada kondisi siang dan malam.

Secara keseluruhan sistem mencapai precision 99,47%, recall 98,61%, F1-score 99,04%, dan mAP@0,5 96,45% pada 47 FPS (Tabel 8). Precision yang tinggi menunjukkan bahwa alarm palsu akibat objek non-manusia jarang terjadi, sedangkan recall yang tinggi menunjukkan sebagian besar manusia di area pengawasan terdeteksi. Pada malam hari, recall turun 2,31 poin dan mAP@0,5 turun 4,70 poin dibandingkan siang hari, sementara precision hanya turun 0,77 poin (Gambar 4). Dampak kondisi IR dengan demikian terutama berupa objek yang terlewat dan bukan deteksi yang keliru. Nilai-nilai ini diperoleh pada satu lokasi dan satu kamera dengan verifikasi terhadap frame sampel (satu frame per 10 detik) sehingga belum menggambarkan kinerja pada lokasi lain.

7. Keterbatasan dan Arah Pengembangan

Penelitian ini memiliki beberapa keterbatasan. Pertama, pengujian hanya menggunakan satu kamera dengan sudut tetap pada satu lokasi. Kedua, YOLOv7 dipakai tanpa pelatihan ulang sehingga belum disesuaikan dengan citra IR dan sudut pandang CCTV setempat. Ketiga, aturan berbasis titik tengah bounding box sensitif di sekitar batas ROI, dan jam operasional ditetapkan secara manual. Keempat, sistem hanya mendeteksi keberadaan manusia tanpa membedakan pihak berizin dan tidak berizin. Kelima, evaluasi memakai frame sampel sehingga aspek temporal belum dievaluasi menyeluruh.

Pengembangan lanjutan dapat mencakup pelatihan ulang (fine-tuning) dengan data IR dan beragam sudut kamera; integrasi pelacak objek (DeepSORT atau ByteTrack) beserta temporal smoothing dan buffer zone; antarmuka grafis untuk mengatur poligon ROI secara dinamis; perbandingan dengan YOLO versi lebih baru (YOLOv8–YOLOv10) serta optimasi TensorRT atau OpenVINO untuk perangkat tepi seperti NVIDIA Jetson; dan integrasi notifikasi multichannel (Telegram, WhatsApp, sirene IoT, serta dashboard pemantauan terpusat).

KESIMPULAN

Sistem deteksi manusia pada area terbatas berbasis YOLOv7 dan CCTV yang dipadukan dengan ROI poligon serta parameter waktu operasional berhasil diimplementasikan dan mampu mendeteksi manusia, menentukan posisinya terhadap ROI, mengidentifikasi pelanggaran, mengaktifkan alarm, dan mencatat log. Pada 18.000 frame sampel, sistem mencapai *precision* 99,47%, *recall* 98,61%, *F1-score* 99,04%, mAP@0,5 96,45%, dan kecepatan rata-rata 47 FPS. Kinerja menurun pada malam hari (rerata *confidence* 64,2% dibandingkan 87,4% pada siang hari) dan ketika jumlah objek bertambah (tingkat deteksi siang hari turun dari 99,0% menjadi 95,9%). Sebanyak 142 dari 150 kejadian pelanggaran (94,7%) teridentifikasi dengan waktu respons rerata 0,8 detik, sedangkan akurasi penentuan posisi paling rendah terjadi di area perbatasan ROI (94,40%). Hasil ini menunjukkan kelayakan sistem sebagai alat bantu pemantauan otomatis pada lokasi uji. Penerapan pada lingkungan lain memerlukan penyesuaian pencahayaan, posisi kamera, dan cakupan area, serta penelitian lanjutan berupa *fine-tuning* pada citra IR dan penambahan pelacakan objek.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Program Studi Teknologi Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara atas dukungan selama pelaksanaan penelitian ini.

DAFTAR PUSTAKA

- Bochkovski, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. Dalam *Computer Vision – ECCV 2014* (hlm. 740–755). Springer.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. Dalam *Computer Vision – ECCV 2016* (hlm. 21–37). Springer.
- Nguyen, H. H., Ta, T. N., Nguyen, N. C., Bui, V. T., Pham, H. M., & Nguyen, D. M. (2021). YOLO based real-time human detection for smart video surveillance at the edge. Dalam *2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)* (hlm. 439–444). IEEE.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Dalam *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (hlm. 779–788). <https://doi.org/10.1109/CVPR.2016.91>
- Wang, C.-Y., Bochkovski, A., & Liao, H.-Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Dalam *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (hlm. 7464–7475).
- Zhao, Z.-Q., Zheng, P., Xu, S., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212–3232. <https://doi.org/10.1109/TNNLS.2018.2876865>