

# Rancang Bangun Sistem Deteksi Manusia dengan YOLO pada video CCTV

**Diterima:**  
10 Juni 2024

**Revisi:**  
10 Juli 2024

**Terbit:**  
1 Agustus 2024

**<sup>1\*</sup>Muhamad Yusup Efendi, <sup>2</sup>Resty Wulanningrum, <sup>3</sup>Ahmad Bagus Setiawan**

<sup>1-3</sup>Universitas Nusantara PGRI Kediri

<sup>1\*</sup>[Efendi.lut@gmail.com](mailto:Efendi.lut@gmail.com), <sup>2</sup>[restyw@unpkdr.ac.id](mailto:restyw@unpkdr.ac.id), <sup>3</sup> [bagus.este@gmail.com](mailto:bagus.este@gmail.com)

**Abstrak**—Penelitian ini bertujuan untuk mengembangkan sistem deteksi objek manusia berbasis computer vision pada CCTV untuk meningkatkan efisiensi penyimpanan data. Masalah klasik sistem CCTV konvensional adalah overload penyimpanan akibat rekaman terus-menerus tanpa seleksi. Solusi yang ditawarkan adalah menerapkan algoritma deteksi objek manusia menggunakan YOLO (You Only Look Once) untuk mengoptimalkan penyimpanan dengan hanya merekam saat ada kehadiran manusia. Penelitian ini menggunakan dataset berupa 775 gambar orang dari Roboflow, di mana setiap gambar dilengkapi dengan label kotak di sekitar objek manusia. Proses pelatihan model dilakukan menggunakan algoritma YOLO versi 8 dengan arsitektur YOLOv8m-seg. Hasil pengujian menggunakan 15 sampel gambar menunjukkan bahwa model dapat mendeteksi semua objek manusia dengan akurasi 100%. Meskipun hasil ini sangat baik, perlu dilakukan pengujian lebih lanjut dengan jumlah sampel yang lebih besar untuk mengevaluasi kinerja model secara menyeluruh. Penelitian ini memberikan kontribusi dalam memecahkan masalah klasik penyimpanan CCTV dengan memanfaatkan teknologi deteksi objek manusia berbasis computer vision.

**Kata Kunci**— CCTV; Deteksi Objek; YOLO

**Abstract**— This research aims to develop a computer vision-based human object detection system on CCTV to improve data storage efficiency. The classic problem of conventional CCTV systems is storage overload due to continuous recording without selection. The solution offered is to apply a human object detection algorithm using YOLO (You Only Look Once) to optimize storage by only recording when there is human presence. This research uses a dataset of 775 images of people from Roboflow, where each image is equipped with a box label around the human object. The model training process is performed using the YOLO version 8 algorithm with YOLOv8m-seg architecture. The test results using 15 image samples show that the model can detect all human objects with 100% accuracy. Although these results are excellent, it is necessary to conduct further tests with a larger number of samples to thoroughly evaluate the performance of the model. This research contributes to solving the classic problem of CCTV storage by utilizing computer vision-based human object detection technology.

**Keywords**— CCTV; Deteksi Objek; YOLO

This is an open access article under the CC BY-SA License.



---

## Penulis Korespondensi:

Muhamad Yusup Efendi  
Departemen Penulis,  
Universitas Nusantara PGRI Kediri  
Email: Email Penulis  
ID Orcid: [<https://orcid.org/register>]  
Handphone: 08xxx

---

## I. PENDAHULUAN

Di era ketika tindak kriminal terus meningkat, keamanan menjadi isu krusial yang harus ditangani dengan serius. Salah satu solusi andalan adalah pemanfaatan teknologi CCTV yang terbukti ampuh dalam memantau keamanan 24/7[1]. Namun, sistem perekaman CCTV konvensional justru menimbulkan masalah baru, yaitu ledakan konsumsi penyimpanan akibat perekaman terus-menerus tanpa seleksi. Merekam setiap momen tanpa pandang bulu tak hanya memboroskan ruang penyimpanan, tetapi juga memperumit proses identifikasi kejadian penting dari tumpukan data tak berguna. Solusi cerdas untuk mengatasi problema ini adalah menerapkan deteksi objek manusia berbasis *computer vision*. Dengan memanfaatkan algoritma deteksi gerakan dan pengenalan bentuk tubuh, CCTV hanya akan aktif merekam ketika ada kehadiran manusia.

Kehadiran manusia yang terdeteksi CCTV merupakan salah satu penerapan konsep algoritma untuk mengenali objek manusia, menjanjikan efisiensi luar biasa. Penyimpanan dapat dioptimalkan dengan menyaring fokus hanya pada objek manusia. Identifikasi insiden juga akan lebih mudah, sementara kinerja sistem CCTV pun dapat ditingkatkan.

Mirwan Abdurrahman Hakim menawarkan terobosan dengan *You Only Look Once V3 (YOLOv3)* untuk deteksi manusia pada CCTV. Berkat akurasi 90%, sistem ini hanya merekam saat objek manusia terdeteksi, menghemat penyimpanan secara signifikan. Tiap deteksi memicu rekaman maksimal 1,6 MB dengan durasi 18 detik. Penelitian ini menjadi kontribusi berharga dalam menyelesaikan isu klasik penyimpanan pada CCTV [2].

Penelitian ini bertujuan menerapkan metode *YOLOv8* untuk deteksi objek manusia menggunakan dataset dari *Google Image/OpenImage*. Meskipun model *YOLOv8* telah dilatih dengan 80 kelas objek, pelatihan dalam penelitian ini hanya akan berfokus pada satu kelas, yaitu "person". Pendekatan ini memungkinkan kamera untuk diatur agar hanya mengenali dan merespons terhadap keberadaan objek manusia saat menerima perintah untuk merekam.

## II. METODE

### A. Analisa dan Pengumpulan data

Dataset yang digunakan dalam penelitian ini adalah *Roboflow*, di mana hanya satu kelas objek diambil, yaitu '*person*' (orang). Dataset *Roboflow* ini sudah memiliki anotasi atau label yang disertakan, sehingga tidak diperlukan aktivitas pelabelan gambar secara manual menggunakan perangkat lunak pelabelan. Jumlah gambar yang digunakan dalam penelitian ini sebanyak 917 gambar yang diambil dari dataset *Roboflow* kelas '*person*'.

### B. YOLO

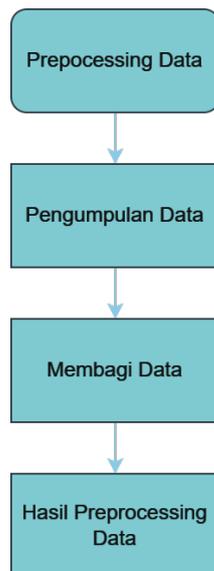
YOLO merupakan algoritma deteksi objek berbasis deep learning yang menggunakan pendekatan "one-stage". Dikembangkan pada 2015, YOLO merupakan turunan dari metode Convolutional Neural Network (CNN). Berbeda dengan algoritma sebelumnya yang melakukan deteksi proposal kandidat bounding boxes lalu verifikasi, YOLO hanya menggunakan satu lapisan Neural Network pada citra input[3].

Proses kerja YOLO diawali dengan membagi citra menjadi grid  $S \times S$ . Setiap sel dalam grid akan memprediksi  $N$  kemungkinan bounding boxes beserta probabilitasnya. Bounding box dengan probabilitas di bawah batas tertentu akan dihapus. Selanjutnya, kotak yang lolos penyaringan akan melalui Non-Maximum Suppression (NMS) untuk mendapatkan posisi objek paling akurat.

Kelebihan YOLO adalah lebih cepat, namun cenderung membuat lebih banyak kesalahan lokalisasi objek dan menghadapi kesulitan dalam mendeteksi objek berukuran kecil dan berdekatan secara rapat.

### C. Preprocessing Data

1. Mencari dan mengumpulkan data dengan objek manusia atau person
2. Membagi data menjadi data training, data valid, data testing
3. Data yang telah melalui seluruh tahapan preprocessing dapat disimpan dalam format yang sesuai dengan kebutuhan untuk kemudian digunakan dalam pelatihan atau evaluasi model.



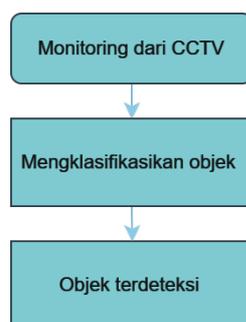
Gambar 1. Alur Preprocessing data

### D. Pelatihan *dataset*

1. Melakukan training data
2. Evaluasi model

### E. Pengujian Model

1. Inputan rekaman video dari CCTV
2. Mengklasifikasikan objek dari data yang telah dilatih
3. Objek terdeteksi sebagai Person



Gambar 2. Alur Pengujian Model

Pengujian dilakukan dengan inputan video dari rekaman CCTV. Adapun cara mendapatkan nilai akurasi dengan menggunakan rumus sebagai berikut[4].

$$Akurasi = \frac{Jumlah\ data\ benar}{Jumlah\ semua\ data} \times 100\% \quad (1)$$

### III. HASIL DAN PEMBAHASAN

#### A. Pengumpulan Data

Dalam Penelitian ini, data yang digunakan adalah 775 gambar *person* atau orang yang diambil dari *Roboflow*. Setiap gambar dilengkapi dengan label berupa kotak yang mengelilingi tubuh manusia, sehingga dapat diketahui posisi manusia dalam setiap gambar.

#### B. Training

Perintah akan menjalankan proses pelatihan model segmentasi YOLO (*You Only Look Once*) versi 8 dengan model YOLOv8m-seg. Data pelatihan yang digunakan diambil dari lokasi dataset yang ditentukan dalam file *data.yaml*. Proses pelatihan akan dilakukan selama 100 epoch dengan ukuran gambar input sebesar 640x640 piksel. Tujuan dari pelatihan ini adalah untuk menghasilkan model yang dapat melakukan segmentasi objek, dalam hal ini memisahkan objek manusia dari latar belakangnya, secara akurat berdasarkan dataset yang disediakan.

Setelah proses pelatihan selesai, model terbaik (*best model*) akan disimpan dalam file *best.pt*. File ini berisi parameter-parameter model yang telah dilatih dan dapat digunakan untuk melakukan inferensi atau pengujian lebih lanjut.

```
!yolo task=segment mode=train model=yolov8m-seg.pt data={dataset.location}/data.yaml epochs=100 imgsz=640
```

Gambar 3. Custom Training data Google Colab

#### C. Pengujian Model

Proses pengujian model dilakukan untuk mengetahui tingkat akurasi dari model yang telah dilatih. Hal ini dilakukan dengan menentukan jumlah sampel data yang akan diuji. Kemudian, dihitung jumlah data yang terdeteksi dengan benar dari keseluruhan sampel uji. Nilai akurasi model dapat diperoleh dengan menghitung persentase antara jumlah data yang terdeteksi dengan benar terhadap total jumlah sampel uji.

Parameter yang digunakan dalam pengujian mencakup berbagai posisi dan aktivitas yang mungkin dilakukan oleh manusia. Dalam contoh ini, jumlah sampel yang diuji adalah 15 gambar. Setiap gambar akan dievaluasi apakah objek manusia terdeteksi atau tidak, yang ditandai dengan adanya kotak pembatasan pada objek manusia. Selain itu, setiap objek yang terdeteksi juga akan memiliki nilai *confidence* atau tingkat kepercayaan terhadap deteksi tersebut.



Gambar 4. Hasil percobaan Siang hari



Gambar 5. Hasil percobaan malam hari

Hasil pengujian dapat disajikan dalam bentuk Tabel 1, yang menampilkan informasi terkait sampel uji, status deteksi, dan nilai confidence untuk setiap objek manusia yang terdeteksi.

Tabel 1. Hasil Percobaan

No	Jarak	Status	True or False
1.	1 meter	Dekat, Menghadap ke depan	True
2.	1 meter	Dekat, Menghadap ke depan	True
3.	1,5 meter	Dekat, Menghadap ke depan	True
15	1 meter	Dekat, Menghadap ke depan	True

Keterangan:

Jarak dekat = Jarak manusia dari kamera CCTV adalah 0 – 2 meter

Hasil pengujian dengan 15 sampel tersebut sehingga menghasilkan akurasi sebagai berikut:

$$Akurasi = \frac{15}{15} \times 100\% = 100\%$$

Jadi, dengan 15 sampel uji, model berhasil mendeteksi semua objek manusia dengan benar, sehingga menghasilkan akurasi sebesar 100%. Hal ini menunjukkan bahwa model yang dilatih memiliki performa yang sangat baik dalam mendeteksi keberadaan objek manusia pada gambar.

Perlu diperhatikan bahwa akurasi 100% pada 15 sampel ini belum tentu mewakili kinerja model secara keseluruhan. Untuk mendapatkan gambaran akurasi yang lebih menyeluruh, perlu dilakukan pengujian dengan jumlah sampel yang lebih besar.

#### IV. KESIMPULAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa, Proses pengujian dilakukan untuk mengetahui tingkat akurasi model. Pengujian menggunakan 15 sampel gambar yang mencakup berbagai posisi dan aktivitas manusia. Hasil pengujian menunjukkan bahwa model berhasil mendeteksi semua objek manusia dengan benar, sehingga menghasilkan akurasi 100%.

Meskipun akurasi 100% pada 15 sampel ini menunjukkan performa yang sangat baik, perlu dilakukan pengujian dengan jumlah sampel yang lebih besar untuk mendapatkan gambaran akurasi yang lebih menyeluruh.

Secara keseluruhan, penelitian ini telah menghasilkan model segmentasi objek manusia yang memiliki akurasi tinggi berdasarkan pengujian awal. Namun, evaluasi lebih lanjut dengan sampel yang lebih besar perlu dilakukan untuk memastikan konsistensi dan kinerja model secara menyeluruh.

## DAFTAR PUSTAKA

- [1] O. Rizan and J. Jend Sudirman Selindung Kel Selindung Kec Gabek Kota Pangkalpinang, "RANCANGAN APLIKASI MONITORING KAMERA CCTV UNTUK PERANGKAT MOBILE BERBASIS ANDROID."
- [2] M. Abdurrahman Hakim, T. Rohana, D. Sulistya Kusumaningrum, and P. Karawang, "Seminar Nasional Hasil Riset Prefix-RTR PEREKAMAN OTOMATIS BERDASARKAN DETEKSI OBJEK MANUSIA PADA CCTV MENGGUNAKAN METODE YOU ONLY LOOK ONCE V3 (YOLOV3)," 2020.
- [3] M. Yusqi Alfian Thoriq, K. Eka Permana, I. Agustien Siradjuddin, T. Informatika, U. Trunojoyo Madura, and J. Raya Telang Kamal, "DETEKSI WAJAH MANUSIA BERBASIS ONE STAGE DETECTOR MENGGUNAKAN METODE YOU ONLY LOOK ONCE (YOLO)," 2023. [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index>
- [4] K. C. Leowis, J. Raharjo, and N. Ibrahim, "RANCANG BANGUN SISTEM PENGENALAN WAJAH DI AREA PUBLIK BERBASIS VIDEO MENGGUNAKAN METODE PRINCIPAL COMPONENT ANALYSIS (PCA) DAN VIOLA JONES DESIGN OF THE PUBLIC FACE RECOGNITION SYSTEM BASED ON VIDEO USING PRINCIPAL COMPONENT ANALYSIS (PCA) AND VIOLA JONES METHODS."
- [5] F. Sthevanie, A. Kurniawan, K. Nur, and R. #3, "Deteksi Helm pada Video Pengendara Sepeda Motor menggunakan Ekstraksi Ciri Histogram of Oriented Gradients", doi: 10.21108/indojc.2020.5.1.377.
- [6] M. Zulfikri, K. Abd Latif, R. Hammad, M. Syahrir, and P. Studi, "Deteksi dan Estimasi Kecepatan Kendaraan dalam Sistem Pengawasan Lalu Lintas Menggunakan Pengolahan Citra Detection and Estimation of Vehicle Speed in Traffic Control Systems Using Image Processing."
- [7] Y. Adil Hidayat, Y. Diah Rosita, and L. Ardiantoro, "DETEKSI UKURAN BOTOL KEMASAN DENGAN METODE HAAR CASCADE DAN SHAPE AREA FEATURE," *APPLIED SCIENCE, ENGINEERING, AND TECHNOLOGY*, vol. 2, no. 1, 2023.
- [8] M. F. Mustaqim, A. Nugroho, D. Alfa, and F. Suni, "Sistem Deteksi Kecepatan Kendaraan Menggunakan Metode Haar Cascade untuk Keamanan Berkendara," *Edu ElektriKa Journal*, vol. 10, no. 2.
- [9] A. Habib Fitriansyah and E. Rachmawati, "PENGENALAN JALAN BERLUBANG BERBASIS VISION MENGGUNAKAN PYRAMID HISTOGRAM OF ORIENTED GRADIENTS," vol. 10, no. 3, 2023, doi: 10.25126/jtiik.2023106820.
- [10] O. V. Putra and I. N. Gustri, "Sistem Deteksi Marka Jalan Berbasis Convolutional Neural Network," *Journal of Computer Engineering, Network, and Intelligent Multimedia*, vol. 1, no. 1, pp. 1–13, Feb. 2023, doi: 10.59378/jcenim.v1i1.2.
- [11] J. Homepage, B. Purnama, E. Insanudin, F. Labib, and N. Sayyid Furqoon, "BATIK: Jurnal Pengembangan dan Pengabdian Masyarakat Multikultural Implementation of Computer Vision for Truck Detection Implementasi Computer Vision untuk Deteksi Truk," vol. 2, pp. 17–23, 2024, doi: 10.57152/batik.v2i1.1223.
- [12] J. H. Sri Wisna *et al.*, "Jurnal Sustainable: Jurnal Hasil Penelitian dan Industri Terapan," vol. 09, no. 01, pp. 8–14, 2020.
- [13] B. Putra, G. Pamungkas, B. Nugroho, and F. Anggraeny, "DETEKSI DAN MENGHITUNG MANUSIA MENGGUNAKAN YOLO-CNN."
- [14] L. Rahma, H. Syaputra, A. H. Mirza, and S. D. Purnamasari, "Objek Deteksi Makanan Khas Palembang Menggunakan Algoritma YOLO (You Only Look Once)," 2021.
- [15] D. Nafis Alfarizi, R. Agung Pangestu, D. Aditya, M. Adi Setiawan, and P. Rosyani, "Penggunaan Metode YOLO Pada Deteksi Objek: Sebuah Tinjauan Literatur Sistematis," 2023. [Online]. Available: <https://jurnalmahasiswa.com/index.php/aidanspk>