

SISTEM MONITORING DAN KENDALI AKSI BUANG SAMPAH SEMBARANGAN MENGGUNAKAN KAMERA BERBASIS 3D CNN DENGAN ARSITEKTUR R(2+1)D

Agustrian Rahmatullah^{1*}, Nyayu Latifah Husni²

^{1,2}Teknik Elektro, Politeknik negeri sriwijaya: Jl. Sriwijaya Negara, Bukit Besar, Kota Palembang, Sumatera Selatan 30139

Keywords:

Deteksi Aktivitas Membuang Sampah, Cctv, Monitoring Real-Time, YOLOv8, R(2+1)D,

Correspondent Email:

agustrianrahmatullah9@gmail.com

Pertumbuhan populasi yang cepat di kota-kota berkembang Indonesia menyebabkan volume sampah terus meningkat, dengan Kota Palembang menghasilkan sekitar 1.300 ton sampah per hari. Rendahnya kesadaran masyarakat menyebabkan perilaku membuang sampah sembarangan masih sering ditemukan di ruang publik, berdampak pada pencemaran lingkungan dan kerusakan infrastruktur. Metode pengawasan manual terbukti kurang efektif akibat keterbatasan jangkauan dan ketidakmampuan beroperasi secara terus-menerus. Penelitian ini mengusulkan sistem deteksi aktivitas membuang sampah sembarangan berbasis CCTV dengan mengintegrasikan model YOLOv8 sebagai pendeteksi manusia dan model R(2+1)D sebagai pengenalan aktivitas dalam satu pipeline monitoring terintegrasi. YOLOv8 mendeteksi keberadaan manusia pada setiap frame video dan menghasilkan bounding box sebagai masukan spasial bagi model R(2+1)D untuk mengklasifikasikan aktivitas ke dalam dua kategori, yaitu normal dan membuang sampah. Sistem juga dilengkapi fitur peringatan kedekatan objek terhadap kamera yang mengaktifkan alarm secara otomatis ketika rasio bounding box mencapai atau melebihi 60% dari luas frame. Evaluasi menggunakan confusion matrix pada 15 data uji menghasilkan precision sebesar 81,82%, recall sebesar 90%, dan accuracy sebesar 80%. Hasil ini menunjukkan bahwa integrasi YOLOv8 dan R(2+1)D mampu mendeteksi aktivitas membuang sampah secara otomatis dan real-time guna mendukung pengawasan lingkungan yang lebih efektif di kawasan perkotaan.



Copyright © 2026 The Author(s), Published by Jurnal Informatika dan Teknik Elektro Terapan. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Rapid population growth in developing cities across Indonesia has caused the volume of unmanaged waste to continue increasing, with Palembang City generating approximately 1,300 tons of waste per day. Low public awareness has resulted in persistent littering behavior in public spaces, particularly along riverbanks, leading to environmental pollution and infrastructure damage. Conventional manual surveillance methods have proven ineffective due to limited coverage and the inability to operate continuously. This study proposes a CCTV-based littering activity detection system by integrating YOLOv8 as a human detector and R(2+1)D as an activity recognizer into a single integrated monitoring pipeline. YOLOv8 detects human presence in each video frame and generates bounding boxes used as spatial input for the R(2+1)D model to classify activities into two categories: normal and littering. The system also incorporates an object-proximity warning feature that automatically activates an alarm when the bounding box ratio reaches or exceeds 60% of the total frame area. System evaluation using a confusion matrix on 15 test samples yielded a precision of 81.82%, recall of 90%, and accuracy of 80%. These results demonstrate that the integration of YOLOv8 and R(2+1)D is capable of automatically detecting littering activity in real-time, supporting more effective environmental monitoring in urban areas.

1. PENDAHULUAN

Pertumbuhan populasi yang cepat di kota-kota berkembang di Indonesia menyebabkan volume sampah yang dihasilkan terus meningkat dari waktu ke waktu. Di Kota Palembang, sejak 5 Maret 2026 total produksi sampah telah mencapai sekitar 1.300 ton per hari dan angka ini diperkirakan akan terus bertambah seiring pertumbuhan jumlah penduduk [1]. Rendahnya kesadaran masyarakat untuk membuang sampah pada tempatnya menyebabkan perilaku membuang sampah sembarangan masih sering ditemukan di ruang publik, terutama di sekitar bantaran sungai, yang berdampak pada pencemaran air, potensi banjir, kerusakan infrastruktur, serta meningkatnya penyebaran penyakit di area sekitarnya.

Pengawasan terhadap pelanggaran pembuangan sampah selama ini umumnya masih bergantung pada petugas lapangan secara manual. Metode tersebut terbukti kurang efektif mengingat luasnya wilayah yang harus diawasi serta ketidakmampuannya untuk beroperasi secara terus-menerus selama 24 jam, sehingga pelanggaran sering kali tidak terdeteksi dan perilaku tersebut terus berulang. Oleh karena itu, diperlukan sistem pengawasan berbasis kamera CCTV yang terintegrasi dengan teknologi kecerdasan buatan untuk memungkinkan deteksi perilaku secara otomatis dan real-time, sekaligus mendukung terciptanya konsep smart city.

Dalam sistem pengawasan berbasis computer vision, algoritma YOLOv8 banyak digunakan untuk tahap deteksi manusia karena terbukti unggul dalam mendeteksi objek dalam satu kali inferensi dengan kecepatan tinggi [2],[3],[4]. Setelah individu dikenali, sistem perlu mengidentifikasi aktivitas yang dilakukan melalui pendekatan Three-Dimensional Convolutional Neural Network (3D CNN), yang mampu mempelajari informasi spasial dan temporal secara simultan dari rangkaian frame video [5]. Namun, penelitian Husni et al. (2025) menunjukkan bahwa arsitektur 3D CNN memiliki keterbatasan signifikan dalam hal komputasi, di mana tingginya kompleksitas parameter menyebabkan proses pelatihan menjadi tidak stabil serta kurang efisien apabila diterapkan secara real-time pada sistem CCTV [6], [7]. Kesenjangan inilah yang menjadi permasalahan utama dalam pengembangan

sistem pengenalan aktivitas berbasis video yang efisien.

Sebagai alternatif, arsitektur R(2+1)D yang diperkenalkan oleh Tran et al. (2018) terbukti lebih efisien dengan memecah konvolusi tiga dimensi menjadi konvolusi spasial dua dimensi dan konvolusi temporal satu dimensi, sehingga mampu meningkatkan kinerja pengenalan aktivitas sekaligus mengurangi beban komputasi secara signifikan dibandingkan 3D CNN konvensional [8]. Selain itu, informasi bounding box hasil deteksi YOLOv8 juga dapat dimanfaatkan untuk mendeteksi objek yang berada terlalu dekat dengan kamera, yang berpotensi mengurangi kualitas pemantauan dan mengganggu keamanan perangkat monitoring apabila terjadi kontak fisik—suatu aspek yang belum banyak diperhatikan pada penelitian sejenis sebelumnya.

Berdasarkan permasalahan dan kesenjangan tersebut, penelitian ini mengusulkan sistem deteksi aktivitas membuang sampah sembarangan berbasis CCTV dengan mengintegrasikan model YOLOv8 sebagai pendeteksi manusia dan model R(2+1)D sebagai pengenalan aktivitas, dilengkapi fitur peringatan kedekatan objek terhadap kamera, ke dalam satu sistem monitoring yang terintegrasi. Tujuan penelitian ini adalah membangun sistem yang mampu mendeteksi keberadaan manusia, mengidentifikasi aktivitas membuang sampah sembarangan secara otomatis dan real-time, serta mengevaluasi kinerja sistem menggunakan metode Confusion Matrix dengan parameter precision, recall, dan accuracy, sehingga dapat mendukung pengawasan lingkungan yang lebih efektif dan terwujudnya lingkungan perkotaan yang lebih bersih, aman, dan berkelanjutan.

2. TINJAUAN PUSTAKA

2.1. YOLOv8

YOLOv8 merupakan generasi terbaru yang dirancang untuk meningkatkan akurasi deteksi, kecepatan inferensi, serta efisiensi komputasi melalui perbaikan pada arsitektur backbone, detection head, dan fungsi loss yang digunakan [9]. Selain mendukung deteksi objek, YOLOv8 juga dapat digunakan untuk tugas klasifikasi citra, segmentasi objek, pelacakan objek, dan estimasi pose manusia sehingga menjadi salah satu model yang banyak

digunakan dalam penelitian maupun implementasi sistem berbasis kecerdasan buatan [9]. Pada penelitian ini, YOLOv8 dimanfaatkan untuk mendeteksi keberadaan manusia pada rekaman CCTV.

YOLOv8 bekerja berdasarkan prinsip object detection, yaitu mendeteksi keberadaan objek sekaligus menentukan lokasinya melalui dua proses utama berupa lokalisasi objek menggunakan bounding box dan klasifikasi objek berdasarkan kategori yang telah dipelajari selama pelatihan [3][4]. Pendekatan ini memungkinkan sistem untuk memproses seluruh frame video dalam satu kali inferensi sehingga sangat cocok diterapkan pada sistem pengawasan berbasis CCTV yang membutuhkan pemrosesan secara real-time [10] [4]. Pada penelitian ini, YOLOv8 dimanfaatkan untuk mendeteksi keberadaan manusia pada rekaman CCTV sebagai tahap awal sebelum dilakukan analisis aktivitas membuang sampah sembarangan.

2.2. R(2+1)D

R(2+1)D merupakan arsitektur jaringan saraf konvolusi berbasis residual yang dirancang khusus untuk analisis data video dalam bidang Video Action Recognition, yaitu teknologi yang memungkinkan komputer mengenali dan mengklasifikasikan aktivitas manusia berdasarkan analisis informasi spasial dan temporal dari rangkaian frame video [11] [8]. Teknologi ini banyak diterapkan pada sistem pengawasan berbasis CCTV untuk mendeteksi aktivitas tertentu secara otomatis tanpa memerlukan pemantauan langsung oleh manusia [11].

Arsitektur R(2+1)D sendiri bekerja dengan memisahkan konvolusi tiga dimensi menjadi dua tahap, yaitu konvolusi spasial dua dimensi dan konvolusi temporal satu dimensi pada setiap lapisannya, pendekatan yang diperkenalkan oleh Tran et al. untuk meningkatkan kemampuan model dalam mempelajari fitur spasial dan temporal secara lebih efektif sekaligus mengurangi kompleksitas parameter dibandingkan konvolusi 3D konvensional [8]. R(2+1)D menerima masukan berupa klip video dalam format $B \times C \times T \times H \times W$ dengan resolusi 112×112 piksel yang diproses melalui lapisan konvolusi awal, batch normalization, fungsi aktivasi

ReLU, dan max pooling sebelum memasuki beberapa kelompok residual block. Setiap residual block terdiri atas konvolusi spasial dan temporal yang bekerja secara berurutan untuk mengekstraksi karakteristik objek serta dinamika gerakan antar frame, dengan jumlah kanal yang meningkat secara bertahap seiring bertambahnya kedalaman jaringan. Pada tahap akhir, fitur yang diperoleh diringkas menggunakan global average pooling dan diteruskan ke lapisan fully connected untuk menghasilkan probabilitas kelas melalui fungsi softmax [8]. Dibandingkan dengan arsitektur 3D CNN konvensional, R(2+1)D memiliki keunggulan berupa jumlah parameter yang lebih efisien, kemampuan representasi yang lebih baik melalui penambahan fungsi nonlinier di antara konvolusi spasial dan temporal, serta proses optimasi yang lebih mudah sehingga mampu menghasilkan performa pengenalan aktivitas yang lebih baik pada data video [8].

2.3. Confusion Matrix

Confusion matrix adalah alat evaluasi yang digunakan untuk mengukur kinerja model klasifikasi dengan membandingkan hasil prediksi terhadap kondisi sebenarnya.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Gambar 1 Tabel confusion Matrix

Terdapat empat komponen utama dalam confusion matrix, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN), yang selanjutnya digunakan sebagai dasar perhitungan metrik evaluasi [12] [13].

Berdasarkan keempat komponen tersebut, tiga metrik utama yang digunakan dalam penelitian ini adalah accuracy, precision, dan recall. Accuracy mengukur tingkat ketepatan prediksi model secara keseluruhan terhadap seluruh data uji, dihitung menggunakan Persamaan (2.1). Precision mengukur seberapa tepat model dalam memprediksi kelas positif dari seluruh prediksi positif yang dihasilkan,

dihitung menggunakan Persamaan (2.2). Sementara itu, recall mengukur kemampuan model dalam mendeteksi seluruh data yang benar-benar termasuk ke dalam kelas positif, dihitung menggunakan Persamaan (2.3).

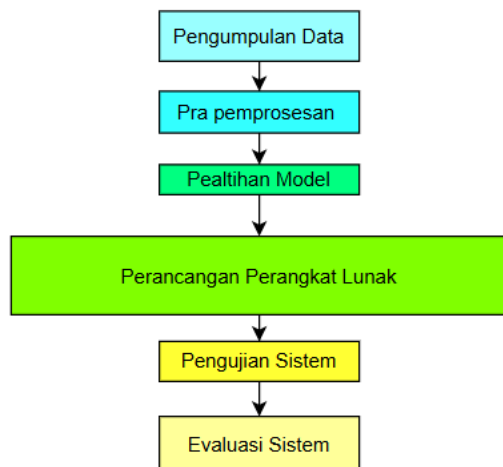
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (2.1)$$

$$precision = \frac{TP}{TP+FP} \times 100 \quad (2.2)$$

$$Recall = \frac{TP}{TP+FN} \times 100 \quad (2.3)$$

3. METODE PENELITIAN

Penelitian ini mengembangkan sistem deteksi aktivitas membuang sampah secara otomatis pada rekaman CCTV dengan mengintegrasikan model YOLOv8 dan Residual (2+1)D Convolutional Neural Network (R(2+1)D). Penggunaan kombinasi kedua model dilakukan untuk mengoptimalkan proses deteksi objek manusia dan klasifikasi aktivitas berdasarkan informasi spasial serta temporal pada video [8] [14],



Gambar 2 Kerangka Penelitian

Tahapan penelitian meliputi pengumpulan data, prapemrosesan, pelatihan model, perancangan perangkat lunak, pengujian sistem. Serta evaluasi, Alur penelitian secara keseluruhan ditunjukkan pada **Gambar 2**.

3.1. Pengumpulan Dataset

Data yang digunakan dalam penelitian ini terdiri atas dataset citra manusia untuk pelatihan model deteksi objek dan dataset video aktivitas manusia untuk pelatihan model klasifikasi aksi.

Dataset citra manusia digunakan untuk melatih model YOLOv8 agar mampu mendeteksi objek manusia secara akurat pada berbagai kondisi lingkungan. Sementara itu, dataset video terdiri atas dua kelas aktivitas, yaitu aktivitas normal dan aktivitas membuang sampah ditampilkan pada **Tabel 3.1**.

Tabel 3. 1 Dataset

Jenis Data	Tipe Data	Jumlah Data
Video aksi membuang sampah	Video	270
Video Selain aksi buang sampah	Video	270
Citra Objek Manusia	image	1.688

Pengumpulan data dilakukan melalui pengambilan video secara langsung menggunakan kamera serta pemanfaatan dataset dari sumber terbuka yang relevan. Data dikumpulkan dengan variasi sudut pandang kamera, jarak objek, latar belakang, dan kondisi pencahayaan untuk meningkatkan kemampuan generalisasi model terhadap kondisi nyata di lapangan. Selanjutnya, data dibagi menjadi data pelatihan, validasi, dan pengujian.

3.2. Prapemrosesan dan Pelatihan Model YOLOv8

Sebelum proses pelatihan, seluruh dataset citra melalui tahap anotasi menggunakan platform Roboflow. Setiap objek manusia diberikan label berupa bounding box untuk menunjukkan posisi objek yang akan dideteksi oleh model. Proses anotasi ini menghasilkan dataset berlabel yang digunakan sebagai masukan pada tahap pelatihan YOLOv8.

Tahap prapemrosesan meliputi penyesuaian ukuran citra, normalisasi data, serta augmentasi berupa rotasi dan flipping. Augmentasi dilakukan untuk meningkatkan variasi data sehingga model mampu mengenali objek manusia pada berbagai kondisi dan sudut pengambilan gambar.

Model YOLOv8 kemudian dilatih menggunakan dataset yang telah dianotasi. YOLOv8 dipilih karena memiliki kemampuan deteksi objek secara real-time dengan tingkat akurasi yang tinggi. Hasil keluaran model

berupa koordinat bounding box yang menunjukkan lokasi objek manusia pada setiap frame video.

3.3. Prapemrosesan dan Pelatihan Model R(2+1)D

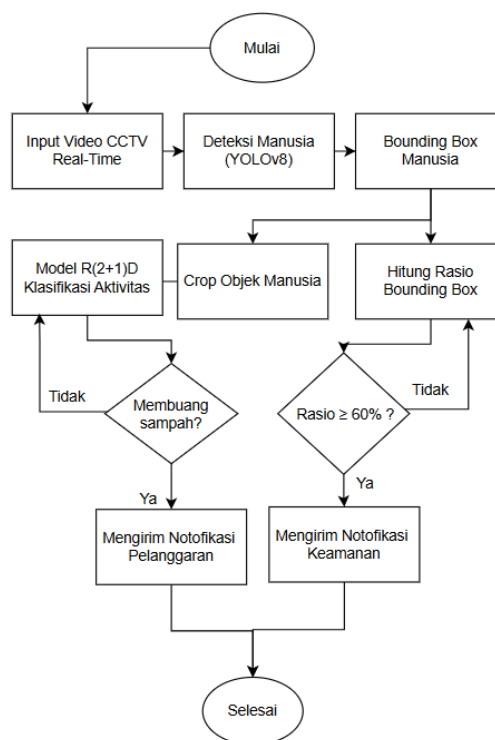
Dataset video yang digunakan untuk klasifikasi aktivitas terlebih dahulu dikonversi menjadi urutan frame. Setiap video diekstraksi menjadi sejumlah frame yang merepresentasikan rangkaian gerakan aktivitas manusia. Tahapan ini bertujuan agar model dapat mempelajari perubahan gerakan yang terjadi dari waktu ke waktu.

Selanjutnya dilakukan proses prapemrosesan berupa normalisasi nilai piksel, konversi grayscale, perubahan ukuran citra menjadi 112×112 piksel, random horizontal flip, dan color jittering. Tahapan ini bertujuan menghasilkan data yang seragam serta meningkatkan ketahanan model terhadap variasi kondisi lingkungan.

Model R(2+1)D digunakan untuk mempelajari karakteristik gerakan berdasarkan informasi spasial dan temporal dari urutan frame video. Arsitektur ini memisahkan proses ekstraksi fitur spasial dan temporal sehingga mampu meningkatkan performa pengenalan aktivitas pada data video. Hasil keluaran model berupa klasifikasi aktivitas ke dalam dua kelas, yaitu aktivitas normal dan aktivitas membuang sampah.

3.4. Perancangan Sistem

Sistem yang dirancang pada penelitian ini mengintegrasikan model deteksi objek YOLOv8 dan model pengenalan aktivitas R(2+1)D ke dalam satu pipeline monitoring berbasis video. Sistem dirancang untuk melakukan pengawasan secara otomatis terhadap aktivitas yang terjadi pada area pengamatan tanpa memerlukan pemantauan manual secara terus-menerus oleh pengguna. Sumber data utama yang digunakan berupa video dari kamera CCTV yang diproses secara real-time sehingga sistem dapat memberikan respons langsung terhadap setiap aktivitas yang terdeteksi.



Gambar 3 Alur Sistem

Alur kerja sistem secara keseluruhan dimulai dari pengambilan video oleh kamera, kemudian setiap frame dianalisis oleh YOLOv8 untuk mendeteksi keberadaan manusia. Hasil deteksi selanjutnya digunakan sebagai dasar untuk pengenalan aktivitas menggunakan model R(2+1)D sekaligus proses deteksi kedekatan objek terhadap kamera. Seluruh informasi yang dihasilkan dari kedua proses tersebut ditampilkan pada antarmuka monitoring, dan notifikasi dikirimkan secara otomatis apabila terdeteksi aktivitas membuang sampah atau kondisi objek yang terlalu dekat dengan kamera.

Integrasi seluruh proses mulai dari pengambilan video, deteksi manusia, pengenalan aktivitas, hingga pengiriman notifikasi berjalan dalam satu rangkaian kerja yang saling terhubung. Perancangan ini memungkinkan sistem tidak hanya berfungsi sebagai alat pemantauan pasif, tetapi juga sebagai sistem pendukung aktif yang membantu pengguna dalam mengawasi area pengamatan secara lebih efektif dan responsif.

3.4.1. Sistem Deteksi Aksi Buang Sampah

Proses monitoring diawali dengan deteksi manusia menggunakan model YOLOv8, yang

menganalisis setiap frame dari kamera untuk mengetahui keberadaan manusia pada area pengamatan. YOLOv8 dipilih karena memiliki kecepatan deteksi dan akurasi yang tinggi sehingga sesuai untuk diterapkan pada sistem real-time. Hasil deteksi ditampilkan dalam bentuk bounding box yang mengelilingi objek manusia, sekaligus berfungsi sebagai referensi spasial untuk tahap pemrosesan selanjutnya.

Setelah manusia berhasil dideteksi, sistem melakukan ekstraksi citra berdasarkan koordinat bounding box untuk memisahkan area objek dari latar belakang. Area yang telah diekstraksi kemudian disusun menjadi rangkaian frame yang digunakan sebagai masukan bagi model R(2+1)D. Dengan hanya memfokuskan analisis pada area yang mengandung manusia, model dapat mengenali pola gerakan secara lebih akurat tanpa terganggu oleh informasi latar belakang yang tidak relevan.

Model R(2+1)D mengklasifikasikan aktivitas yang diamati ke dalam dua kategori, yaitu aktivitas normal dan aktivitas membuang sampah. Model membandingkan pola gerakan yang terdeteksi dengan pola yang telah dipelajari pada tahap pelatihan, kemudian menghasilkan prediksi dengan nilai confidence tertinggi. Apabila aktivitas membuang sampah terdeteksi, sistem secara otomatis mengirimkan notifikasi kepada pengguna sehingga tindakan yang diperlukan dapat segera dilakukan tanpa harus terus-menerus memantau layar monitoring.

3.4.2. Sistem Peringatan Kedekatan Objek terhadap Kamera

Selain digunakan untuk pengenalan aktivitas, informasi bounding box yang dihasilkan oleh YOLOv8 juga dimanfaatkan untuk memperkirakan tingkat kedekatan objek terhadap kamera. Fitur ini bekerja berdasarkan prinsip bahwa semakin dekat posisi objek terhadap kamera, semakin besar ukuran bounding box yang dihasilkan relatif terhadap luas keseluruhan frame [15]. Rasio kedekatan dihitung menggunakan Persamaan (3.1), yang mengukur persentase luas frame yang ditempati oleh bounding box objek yang terdeteksi.

Nilai ambang batas sebesar 60% ditetapkan berdasarkan hasil pengamatan selama pengujian sistem. Apabila nilai rasio bounding box mencapai atau melebihi ambang

batas tersebut, sistem menganggap objek berada terlalu dekat dengan kamera dan secara otomatis mengaktifkan alarm serta mengirimkan notifikasi kepada pengguna. Untuk mencegah pengiriman alarm yang berlebihan, sistem menerapkan mekanisme cooldown setelah setiap alarm dikirimkan, sehingga notifikasi tidak dikirim secara berulang ketika objek tetap berada pada posisi yang sama dalam jangka waktu tertentu.

Hasil deteksi kedekatan objek ditampilkan pada antarmuka monitoring bersama dengan hasil deteksi manusia dan klasifikasi aktivitas, sehingga pengguna memperoleh informasi status pemantauan secara komprehensif dan real-time. Fitur tambahan ini meningkatkan fungsionalitas sistem monitoring secara keseluruhan, tidak hanya dalam mendeteksi perilaku membuang sampah, tetapi juga dalam menjaga kualitas pengawasan dan keamanan fisik perangkat monitoring. berdasarkan Persamaan (3.1):

$$R = \frac{LB_{Box}}{L_{Frame}} \times 100 \quad (3.1)$$

dengan:

RR R = rasio bounding box terhadap frame (%)

$LB_{Box} = L_{\{BBox\}}$ LB_{Box} = luas bounding box objek manusia (piksel²)

$L_{Frame} = L_{\{Frame\}}$ L_{Frame} = luas keseluruhan frame video (piksel²)

3.5. Pengujian sistem

3.5.1. Pengujian Sistem terintegrasi

Pengujian sistem dilakukan untuk mengevaluasi kinerja model kecerdasan buatan yang telah dikembangkan dalam mendeteksi keberadaan manusia serta mengklasifikasikan aktivitas pembuangan sampah berdasarkan data visual dari kamera CCTV. Pengujian menggunakan data video secara real-time karena pendekatan ini mampu memberikan gambaran yang lebih representatif terhadap kondisi operasional sistem yang sebenarnya, di mana model dapat dievaluasi pada kondisi lingkungan yang dinamis seperti perubahan pencahayaan, pergerakan objek, dan variasi cuaca.

Proses pengujian dilakukan menggunakan video yang diperoleh dari kamera yang dipasang pada area bantaran sungai sebagai lokasi target penerapan sistem. Setiap frame video diproses secara real-time oleh model YOLOv8 untuk mendeteksi keberadaan manusia, kemudian hasil deteksi berupa bounding box digunakan sebagai masukan bagi model R(2+1)D untuk mengidentifikasi aktivitas yang dilakukan.

3.5.2. Pengujian Sistem Peringatan Kedekatan Objek

Pengujian sistem peringatan kedekatan objek dilakukan untuk memastikan mekanisme alarm dapat bekerja sesuai rancangan, yaitu mengirimkan peringatan ketika rasio bounding box terhadap luas frame mencapai ambang batas tertentu. Pengujian dilakukan dalam dua skenario, yaitu kondisi objek dekat kamera dan kondisi objek jauh dari kamera, dengan nilai ambang batas yang ditetapkan sebesar 60% berdasarkan Persamaan (3.1)

Tabel 3. 2 Pengujian Sistem Peringatan Kedekatan Objek

No	Kondisi Pengujian	Rasio Bounding Box	Status Alarm
1.	Objek dekat kamera	$\geq 60\%$	Terkirim
2.	Objek jauh dari kamera	$< 60\%$	Tidak terkirim

Berdasarkan hasil pengujian pada Tabel 3.2, sistem berhasil mengirimkan alarm ketika rasio bounding box mencapai atau melebihi 60% dari luas frame, dan tidak mengirimkan alarm ketika rasio berada di bawah ambang batas tersebut. Hasil ini menunjukkan bahwa mekanisme deteksi kedekatan objek berbasis bounding box telah berfungsi sesuai rancangan, sehingga fitur ini dapat digunakan sebagai mekanisme tambahan untuk menjaga keamanan perangkat monitoring dari objek yang berada terlalu dekat dengan kamera.

3.6. Evaluasi Sistem Terintegrasi

Evaluasi sistem terintegrasi dilakukan untuk menilai kinerja keseluruhan sistem dalam

mendeteksi dan mengidentifikasi aktivitas membuang sampah secara otomatis, dengan fokus pada kemampuan model YOLOv8 dan R(2+1)D dalam bekerja secara terintegrasi. Sistem menjalankan dua fungsi utama secara berurutan, yaitu deteksi keberadaan manusia menggunakan YOLOv8 yang hasilnya digunakan sebagai masukan bagi model R(2+1)D untuk mengidentifikasi aktivitas, khususnya membedakan aktivitas membuang sampah dari aktivitas normal pada area pengamatan.

Evaluasi dilakukan menggunakan data video real-time dari lokasi pengujian dengan mengamati berbagai faktor seperti perubahan pencahayaan, variasi latar belakang, posisi objek, dan gangguan visual yang dapat memengaruhi hasil deteksi dan klasifikasi. Untuk memperoleh hasil evaluasi yang objektif dan terukur, digunakan metrik accuracy, precision, dan recall yang dihitung berdasarkan confusion matrix dengan empat kategori hasil prediksi, yaitu true positive (TP), true negative (TN), false positive (FP), dan false negative (FN).

Hasil evaluasi sistem terintegrasi digunakan untuk menilai tingkat keandalan sistem secara menyeluruh sekaligus mengidentifikasi kendala yang muncul selama pengoperasian, seperti kegagalan deteksi objek, kesalahan klasifikasi, maupun penurunan performa akibat perubahan kondisi lingkungan. Melalui evaluasi ini diharapkan diperoleh gambaran komprehensif mengenai kemampuan sistem dalam mendukung proses pemantauan lingkungan secara real-time, serta menjadi dasar pengembangan dan penyempurnaan sistem pada penelitian selanjutnya.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pelatihan YOLOv8

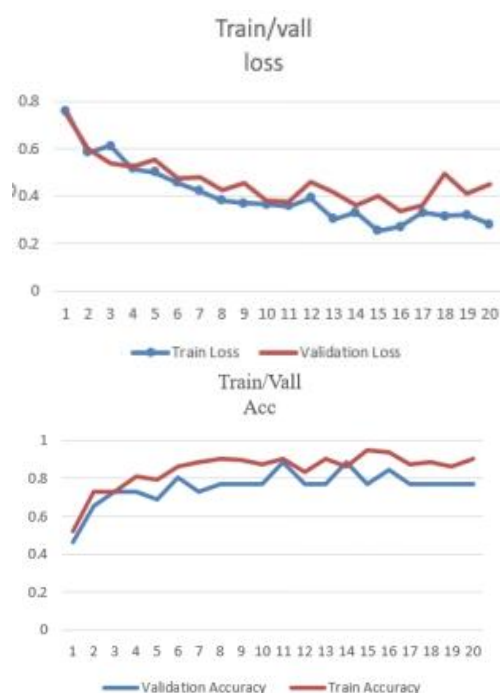
Tabel 4. 1 Hasil Pelatihan Yolov8

Class	Precision	Recall	mAP50	mAP95
Person	1.00	0.67	0.486	0.261

Berdasarkan Tabel 4.1, model YOLOv8 menghasilkan nilai precision sebesar 1,00, recall sebesar 0,67, mAP50 sebesar 0,486, dan mAP50-95 sebesar 0,261. Nilai precision 100% menunjukkan tidak adanya false positive, yang berarti seluruh objek yang terdeteksi

merupakan objek manusia yang benar. Sementara itu, nilai recall sebesar 0,67 mengindikasikan masih terdapat beberapa objek manusia yang tidak terdeteksi, kemungkinan akibat variasi sudut pengambilan gambar, ukuran objek yang kecil, atau objek yang tertutup sebagian. Meskipun demikian, performa model dinilai cukup memadai untuk mendukung proses deteksi manusia sebagai tahap awal dalam sistem.

4.2. Hasil Pelatihan R(2+1)D



Gambar 4 Grafik Pelatihan model R(2+1)D

Berdasarkan grafik pelatihan pada Gambar 4, training loss menurun dari 0,76 menjadi 0,28 pada epoch ke-16, dengan training accuracy mencapai 94,79%. Validation loss terbaik diperoleh sebesar 0,3376 dan validation accuracy tertinggi sebesar 88,46%, sedangkan model yang dipilih melalui early stopping menghasilkan validation accuracy sebesar 84,62%. Selisih antara training dan validation accuracy mengindikasikan adanya overfitting ringan, namun penggunaan augmentasi data, dropout, dan early stopping berhasil menjaga kestabilan performa model.

4.3. Hasil Pengujian Sistem Terintegrasi

4.3.1. Hasil Pengujian Sistem Deteksi Aktivitas Membuang Sampah

Pengujian sistem dilakukan untuk mengevaluasi kinerja integrasi model YOLOv8 dan R(2+1)D dalam mendeteksi aktivitas membuang sampah secara real-time. YOLOv8 bertugas mendeteksi keberadaan manusia pada setiap frame video dan menghasilkan bounding box sebagai referensi lokasi objek, yang selanjutnya digunakan sebagai masukan bagi model R(2+1)D untuk mengklasifikasikan aktivitas ke dalam dua kategori, yaitu normal dan buang sampah. Apabila aktivitas membuang sampah terdeteksi, sistem secara otomatis memberikan notifikasi peringatan kepada pengguna.

Hasil prediksi sistem kemudian dibandingkan dengan kondisi sebenarnya dan dievaluasi menggunakan confusion matrix berdasarkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN),

Tabel 4. 2 Hasil pengujian sistem

No	Hasil Identifikasi	Aksi Sebenarnya	Prediksi Aksi	Deskripsi
1.		Buang Sampah	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Postif (Tp)
2.		Buang Sampah	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Postif (Tp)
3.		Buang Sampah	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Postif (Tp)
4.		Buang Sampah	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Postif (Tp)
5.		Buang Sampah	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Postif (Tp)

6.		Buang Sampah	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Postif (Tp)
7.		Buang Sampah	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Postif (Tp)
8.		Buang Sampah	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Postif (Tp)
9.		Buang Sampah	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Postif (Tp)
10.		Buang Sampah	Normal	Mendeteksi Aksi Buang Sampah Dengan Data Negatif (Fn)
11.		Normal	Normal	Tidak Mendeteksi Aksi Buang Sampah Dengan Data Negatif (Tn)
12.		Normal	Normal	Tidak Mendeteksi Aksi Buang Sampah Dengan Data Negatif (Tn)
13.		Normal	Normal	Tidak Mendeteksi Aksi Buang Sampah Dengan Data Negatif (Tn)
14.		Normal	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Negatif (Fp)
15.		Normal	Buang Sampah	Mendeteksi Aksi Buang Sampah Dengan Data Negatif (Fp)

Sebagaimana ditunjukkan pada Tabel 4.2. Dari 15 data pengujian, sistem berhasil mengklasifikasikan 9 aktivitas membuang sampah secara benar (TP), 3 aktivitas normal

secara benar (TN), 1 aktivitas membuang sampah yang tidak terdeteksi (FN), dan 2 aktivitas normal yang salah diklasifikasikan sebagai membuang sampah (FP).

Berdasarkan hasil Confusion Matrix tersebut, dilakukan perhitungan nilai precision, recall, dan accuracy untuk mengetahui performa sistem secara lebih terukur. Ketiga parameter tersebut digunakan karena dapat memberikan gambaran mengenai kemampuan sistem dalam mendeteksi aktivitas membuang sampah dari berbagai aspek.

$$precision = \frac{9}{9+2} \times 100\% = 81,82\%$$

$$Recall = \frac{9}{9+1} \times 100\% = 90\%$$

$$Accuracy = \frac{9 + 3}{9 + 3 + 2 + 1} \times 100\% = 80\%$$

Tabel 4.3 Hasil Perhitungan Confusion Matrix Pengujian

Jumlah Data	Precision	Recall	Accuracy
15	81,82%	90%	80%

Berdasarkan hasil tersebut, sistem menghasilkan nilai precision sebesar 81,8%, recall sebesar 90%, dan accuracy sebesar 80%, yang menunjukkan bahwa sistem mampu mendeteksi aktivitas membuang sampah dengan cukup baik meskipun masih terdapat ruang untuk peningkatan dalam mengurangi kesalahan klasifikasi.

4.3.2. Hasil Pengujian Sistem Peringatan Kedekatan Objek terhadap Kamera

Pengujian sistem peringatan kedekatan objek dilakukan untuk memverifikasi kemampuan sistem dalam mendeteksi objek yang berada terlalu dekat dengan kamera berdasarkan rasio bounding box hasil deteksi YOLOv8 terhadap luas keseluruhan frame. Sistem dirancang untuk mengaktifkan alarm secara otomatis apabila rasio tersebut mencapai atau melebihi ambang batas 60%, dan tidak mengirimkan alarm ketika rasio masih berada di bawah nilai tersebut. Pengujian dilakukan pada dua kondisi utama, yaitu objek dekat dan objek jauh dari kamera, dengan hasil yang ditunjukkan pada Tabel 4.3.

Tabel 4. 4 Hasil Pengujian Sistem Peringatan Kedekatan Objek

No	Hasil Bounding Box	Rasio Bounding Box	Status Alarm	Hasil
1.		$\geq 60\%$	Alarm Terkirim	Terkirim
2.		$< 60\%$	Alarm Tidak terkirim	Tidak Terkirim

Berdasarkan Tabel 4.3, sistem berhasil mengaktifkan alarm pada kondisi objek dekat (rasio $\geq 60\%$) dan tidak mengaktifkan alarm pada kondisi objek jauh (rasio $< 60\%$), sesuai dengan rancangan yang telah ditetapkan. Selain itu, mekanisme cooldown yang diterapkan juga berjalan dengan baik sehingga alarm tidak dikirim secara berulang ketika objek tetap berada pada posisi yang sama, sehingga dapat mengurangi kemungkinan false alarm selama proses monitoring berlangsung. Pendekatan berbasis rasio bounding box ini memiliki keunggulan karena tidak memerlukan sensor tambahan maupun perhitungan jarak yang kompleks, sehingga proses komputasi tetap efisien dan dapat diterapkan secara real-time.

5. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem monitoring deteksi aktivitas membuang sampah sembarangan berbasis CCTV dengan mengintegrasikan model YOLOv8 dan R(2+1)D dalam satu pipeline pemantauan yang terintegrasi. Sistem yang dikembangkan mampu mendeteksi keberadaan manusia secara real-time menggunakan YOLOv8 dengan nilai precision 100% dan recall 0,67, yang menunjukkan bahwa seluruh objek yang terdeteksi merupakan objek manusia yang benar meskipun masih terdapat beberapa objek yang tidak terdeteksi akibat variasi kondisi lingkungan. Model R(2+1)D berhasil dilatih dengan training accuracy sebesar 94,79% dan validation accuracy sebesar 84,62%, dengan loss validasi terbaik sebesar 0,3376 pada epoch ke-16.

Hasil pengujian sistem terintegrasi pada 15 data uji menunjukkan nilai precision sebesar

81,82%, recall sebesar 90%, dan accuracy sebesar 80%. Nilai recall 90% mengindikasikan bahwa sistem mampu mendeteksi sebagian besar aktivitas membuang sampah yang terjadi, sementara nilai precision 81,82% menunjukkan masih terdapat beberapa kesalahan klasifikasi pada aktivitas normal. Selain itu, fitur peringatan kedekatan objek terhadap kamera berbasis rasio bounding box berhasil bekerja sesuai rancangan, dengan alarm yang hanya aktif ketika rasio bounding box mencapai atau melebihi ambang batas 60% dari luas frame, tanpa memerlukan sensor tambahan maupun perhitungan jarak yang kompleks.

Secara keseluruhan, integrasi YOLOv8 dan R(2+1)D terbukti mampu membangun sistem pengawasan lingkungan yang lebih efisien dan responsif dibandingkan metode pemantauan manual konvensional, serta berpotensi mendukung terwujudnya lingkungan perkotaan yang lebih bersih dan berkelanjutan. Untuk penelitian selanjutnya, disarankan untuk memperluas dataset pelatihan, mengoptimalkan nilai threshold, serta menguji sistem pada berbagai kondisi lingkungan yang lebih beragam guna meningkatkan akurasi dan keandalan sistem secara keseluruhan.

DAFTAR PUSTAKA

- [1] Nadiya, "Palembang Hasilkan 1.300 Ton Sampah/Hari, PLTSA Ditarget Oktober Beroperasi.," detikSumbagsel. [Online]. Available: <https://www.detik.com/sumbagsel/berita/d-8385826/palembang-hasilkan-1-300-ton-sampah-hari-pltsa-ditarget-oktober-beroperasi>
- [2] M. Deshpande, "ENERGY-EFFICIENT CCTV CAMERA SYSTEM FOR REAL- TIME HUMAN DETECTION USING YOLOV8 MODEL," no. M1, pp. 47–59, 2024, doi: 10.15587/1729-4061.2024.310401.
- [3] D. S. Mahardhika and M. A. I. Pakereng, "Deteksi Objek Secara Real-Time Berbasis YOLOv8 dan Algoritma DeepSORT," vol. 5, no. 1, pp. 671–686, 2026.
- [4] I. Pratama and A. Muliantara, "Deteksi Objek pada Citra Menggunakan Model YOLO," vol. 2, pp. 469–474, 2024.
- [5] N. Combining, C. Block, A. Module, and U. M. Frames, "Anomaly Detection Based on a 3D Convolutional Neural Network Combining Convolutional Block Attention Module Using Merged Frames," 2023.

- [6] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning Spatiotemporal Features with 3D Convolutional Networks".
- [7] S. Ji and K. Yu, "3D Convolutional Neural Networks for Human Action Recognition," 2012, doi: 10.1109/TPAMI.2012.59.
- [8] D. Tran, H. Wang, L. Torresani, J. Ray, Y. Lecun, and M. Paluri, "A Closer Look at Spatiotemporal Convolutions for Action Recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 6450–6459, 2018, doi: 10.1109/CVPR.2018.00675.
- [9] F. O. R. R. Eal, T. V Ision, C. V Jul, C. Science, and H. Hd, "YOLOV5, YOLOV8 AND YOLOV10: THE GO-TO DETECTORS FOR REAL-TIME VISION," pp. 1–12, 2024.
- [10] A. Peryanto *et al.*, "Rancang Bangun Klasifikasi Citra Dengan Teknologi Deep Learning Berbasis Metode Convolutional Neural Network," vol. 8, pp. 138–147, 2019.
- [11] A. Ferrari, D. Micucci, M. Mobilio, and P. Napolitano, "Trends in human activity recognition using smartphones," *J. Reliab. Intell. Environ.*, vol. 7, no. 3, pp. 189–213, 2021, doi: 10.1007/s40860-021-00147-0.
- [12] Rina, "Memahami Confusion Matrix: Accuracy, Precision, Recall, Specificity, dan F1-Score untuk Evaluasi Model Klasifikasi," medium. [Online]. Available: <https://esairina.medium.com/memahami-confusion-matrix-accuracy-precision-recall-specificity-dan-f1-score-610d4f0db7cf>
- [13] C. Rozikin, U. S. Karawang, and J. Barat, "ANALISIS SENTIMEN ULASAN PENGGUNA PADA GAME THEOTOWN MENGGUNAKAN SUPPORT VECTOR MACHINE," vol. 14, no. 2.
- [14] J. Solawetz and Francesco, "Apa itu YOLOv8? Panduan Lengkap," Roboflow. Accessed: Dec. 05, 2025. [Online]. Available: <https://blog.roboflow.com/what-is-yolov8/#:~:text=YOLOv8 adalah model YOLO tercanggih terbaru yang,peningkatan arsitektur serta pengalaman pengembang dibandingkan YOLOv5.>
- [15] E. E. Idohen, A. M. Olatunde, and U. A. Umar, "Object distance estimation using a monovision camera," *LAES Int. J. Robot. Autom.*, vol. 12, no. 4, pp. 325–331, 2023, doi: 10.11591/ijra.v12i4.pp325-331.