

DETEKSI MANUSIA DENGAN ALGORITMA YOLO UNTUK PEMUTARAN AUDIO OTOMATIS DI AREA TERTENTU

M Alfin Mansyur¹⁾, Nunik Pratiwi^{*2)}

1. Universitas Muhammadiyah Prof. Dr. Hamka, Indonesia

2. Universitas Muhammadiyah Prof. Dr. Hamka, Indonesia

Article Info

Kata Kunci: *You Only Look Once*; Deteksi Objek; Pelacakan Objek; Keamanan Ritel.

Keywords: *You Only Look Once*; *Object Detection*; *Object Tracking*; *Retail Security*.

Article history:

Received 1 October 2024

Revised 6 November 2024

Accepted 12 December 2024

Available online 1 March 2025

DOI :

<https://doi.org/10.29100/jipi.v10i1.5967>

* Corresponding author. Nunik

Pratiwi

E-mail address:

npratiwi@uhamka.ac.id

ABSTRAK

Indonesia memiliki sektor ritel yang luas, mulai dari pasar tradisional, supermarket, minimarket, hingga toko online. Pencurian masih menjadi tantangan besar, terutama bagi toko-toko kecil yang hanya mengandalkan CCTV sebagai sistem keamanan. Untuk membantu para pelaku usaha ritel di Indonesia dalam mencegah terjadinya pencurian, sistem keamanan ritel perlu diinovasi dengan teknologi yang dapat diintegrasikan dengan CCTV. Penelitian ini bertujuan untuk mengembangkan program yang dapat mendeteksi setiap kedatangan pengunjung dan memutar suara sebagai bentuk peringatan. Model yang digunakan adalah YOLOv5 pre-trained untuk mendeteksi pose kepala. Program diuji dengan video rekaman CCTV yang didapat dari situs web YouTube. Video dibagi menjadi beberapa segmen untuk mensimulasikan berbagai skenario saat pengunjung memasuki toko. Hasil pengujian menunjukkan bahwa program berhasil memberikan ID pada setiap objek yang terdeteksi dan memutar audio dengan waktu respons rata-rata antara 0 hingga 0,7 milidetik. Performa terbaik program dalam memberikan ID dan memutar audio terjadi pada pengujian dengan kondisi jumlah objek yang sedikit dan tanpa adanya halangan

ABSTRACT

Indonesia has a vast retail sector, ranging from traditional markets, super- markets, minimarkets, to online stores. Shoplifting remains a major chal- lenge, especially for small shops that rely solely on CCTV as their secu- rity system. To assist Indonesian retailers in preventing shoplifting, retail security systems need to be innovated with technology that can be inte- grated with CCTV. This study aims to develop a program that can detect the arrival of each visitor and play a sound as a warning. The model used is YOLOv5 pre-trained for head pose detection. The program was tested with CCTV footage obtained from the YouTube website. The video was divided into several segments to simulate various scenarios when visitors enter the store. The test results show that the program successfully pro- vides an ID for each detected object and plays audio with an average re- sponse time between 0 and 0.7 milliseconds. The program's best perfor- mance in providing IDs and playing audio occurred in tests with fewer objects and no obstructions.

I. PENDAHULUAN

Industri ritel di Indonesia saat ini sedang mengalami ekspansi yang pesat, dengan implikasi ekonomi yang signifikan. Hal ini didorong oleh pertumbuhan besar dalam jumlah pendirian ritel dan penjualan tahunan yang melebihi triliunan rupiah [1]. Namun, di balik pertumbuhan yang pesat ini, pencurian selalu menjadi perhatian utama bagi para pelaku industri ritel. Untuk menanggulangi masalah tersebut, penjaga toko menjadi kunci upaya pencegahan pencurian [2].

Kemampuan penjaga toko untuk selalu siaga memantau kedatangan pengunjung tidaklah praktis, terutama mengingat banyak toko di Indonesia yang hanya memiliki satu orang penjaga [3]. Pemasangan Closed Circuit Television (CCTV) di toko kelontong merupakan salah satu upaya untuk meningkatkan keamanan dan pengawasan terhadap aktivitas pengunjung. Namun, keterbatasan jumlah penjaga toko yang memantau CCTV sering kali membuat mereka mengalami rasa jenuh dan bosan. Kesibukan penjaga toko dalam melayani pelanggan juga dapat menurunkan kewaspadaan dan keefektifan pengawasan terhadap pengunjung yang datang. Oleh karena itu meningkatkan kualitas sistem keamanan pada masyarakat dengan memanfaatkan teknologi mutakhir adalah suatu

keharusan [4].

Terdapat penelitian yang telah menunjukkan potensi penggunaan sistem monitoring berbasis IoT dan YOLO untuk meningkatkan keamanan ruang server di Sekretariat BPSDM ESDM. Sistem ini menggabungkan ESP32Cam, sensor PIR, buzzer, dan algoritma YOLO untuk mendeteksi gerakan manusia dan mengidentifikasi objek. Ketika sensor PIR mendeteksi gerakan, buzzer akan berbunyi dan sistem akan menangkap gambar objek untuk dikirimkan ke web server. Algoritma YOLO kemudian memproses gambar untuk mengidentifikasi nama objek yang terdeteksi. Kinerja sistem dalam mendeteksi objek manusia mengalami penundaan dan akurasi penamaan objek pada aplikasi tidak selalu optimal [5].

Pada penelitian lain telah mengkaji sistem keamanan pada sebuah toko kelontong menggunakan Arduino dan sensor ultrasonik untuk mendeteksi pengunjung, dengan jarak deteksi sensor yang dibatasi pada jarak 12 cm dan akan mendeteksi objek pada jarak 5 cm, jika objek terdeteksi dengan status "objek datang", DF Player akan merespon dengan memutar audio dan mengambil gambar. Ketika mendeteksi objek dengan gerakan lambat, sensor ultrasonik akan mengalami kebingungan dan menyebabkan nilai sensor menjadi acak [2].

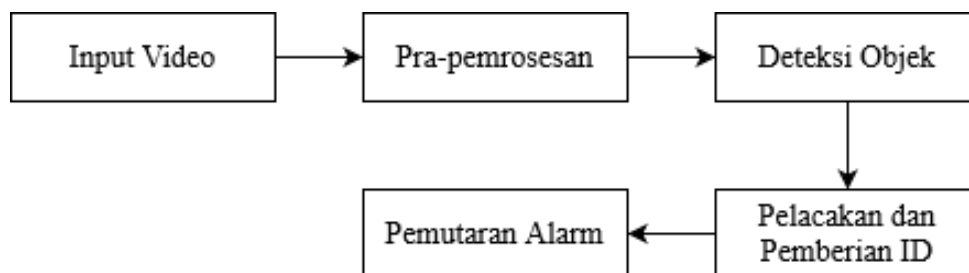
Penelitian sebelumnya juga telah mengembangkan sistem keamanan rumah menggunakan Arduino dan mikrokontroler Atmega2560 untuk mendeteksi pergerakan dengan memanfaatkan sensor PIR. Kamera IP terhubung dengan motor servo untuk memantau ruangan secara real-time melalui web. Sistem mampu mendeteksi pergerakan dengan akurat pada suhu 27 derajat Celcius dalam radius 4 hingga 7 meter dengan waktu respons rata-rata 2 detik. Mikrokontroler secara otomatis mengaktifkan alarm buzzer dan mengirimkan SMS ke pengguna [6].

Untuk membantu dalam mendeteksi kedatangan pengunjung di toko kelontong dengan memanfaatkan CCTV, diperlukan sebuah algoritma visi komputer yang dapat mengoptimalkan proses tersebut [7]. YOLO (You Only Look Once) adalah algoritma visi komputer yang terkenal dengan kemampuannya mendeteksi dan mengklasifikasikan objek secara cepat dan akurat dalam satu kali pandang [8]. Algoritma YOLO banyak diaplikasikan dalam berbagai bidang, seperti deteksi masker dan pemantauan jarak fisik [9], identifikasi kecelakaan lalu lintas [10], dan penghitungan jumlah kendaraan di jalan raya [11]. YOLOv5 adalah salah satu model deteksi objek real-time yang paling populer saat ini. Algoritma ini dipilih, karena terkenal dengan akurasinya, serta jauh lebih ringan dan lebih cepat dibandingkan EfficientDet [12]. YOLOv5 mampu memberikan performa yang cukup baik dengan akurasi, AP, AR, dan mAP sebesar 0,43, 0,492, 0,431, dan 0,555 dalam mendeteksi dan melacak orang melalui CCTV [13]. YOLOv5 memiliki arsitektur model yang sederhana dan mudah dipahami, sehingga mudah diimplementasikan dan dimodifikasi [14]. Arsitektur YOLOv5 terdiri dari empat varian, yaitu YOLOv5s, YOLOv5m, YOLOv5l, dan YOLOv5x. Setiap modelnya memiliki kompleksitas dan kinerja yang berbeda, sehingga dapat dipilih sesuai dengan kebutuhan dan keterbatasan perangkat yang digunakan [15].

Pada penelitian ini dilakukan deteksi objek berupa kepala manusia lalu melacak pergerakannya, untuk memberikan ID unik dan memutar audio ketika objek tersebut memasuki area deteksi yang ditentukan. Algoritma YOLOv5 digunakan untuk mendeteksi objek kepala, OpenCV untuk pemrosesan video, dan pygame untuk pemutaran audio. Model yang digunakan adalah "class_Yolov5_head" yang dikembangkan oleh pengguna GitHub bernama RichardoMrMu[16].

II. METODE PENELITIAN

Penelitian ini mengusulkan penggunaan YOLOv5 sebagai solusi inovatif untuk meningkatkan efektivitas pengawasan di toko kelontong dengan memanfaatkan CCTV. Bagan alur penelitian yang diusulkan digambarkan pada Gambar 1.



Gambar. 1. Alur penelitian

Penelitian ini terdiri dari lima tahapan. Dimulai dari input video, pra-proses, deteksi objek, pelacakan dan pemberian ID, dan pemutaran audio. Penelitian ini berfokus untuk mengetahui apakah YOLOv5 mampu mendeteksi dan memberi tahu penjaga toko tentang pengunjung yang datang.

A. Input Video

Tahap pertama dalam penelitian ini adalah pengambilan input video. Video dapat berasal dari kamera pengawas atau file video yang sudah ada. Dalam penelitian ini, video yang digunakan adalah potongan video yang diunduh dari media sosial YouTube. Video-video ini berisi aktivitas di berbagai kondisi yang digunakan untuk menjalankan skenario pengujian. Video-video ini menyediakan rangkaian frame yang akan dianalisis oleh sistem untuk mendeteksi keberadaan objek yang telah ditentukan.

B. Prapemrosesan

Tahapan berikutnya adalah prapemrosesan, di mana pengubahan ukuran frame dilakukan. Semua frame pada video diubah ukurannya menjadi 640 piksel pada sisi terpanjangnya, dengan tetap mempertahankan rasio aspek asli. Pengubahan ukuran frame menjadi lebih kecil ini bertujuan untuk mengurangi penggunaan daya dan mengoptimalkan proses deteksi objek [17]. Penggunaan resolusi video yang besar juga dapat menyebabkan tumpang tindih pada layer. Oleh karena itu, resolusi 640 piksel dipilih karena tidak terlalu kecil sehingga gambar masih dapat dilihat dengan jelas, namun juga tidak terlalu besar sehingga dapat mengurangi penggunaan daya GPU.

C. Deteksi Objek

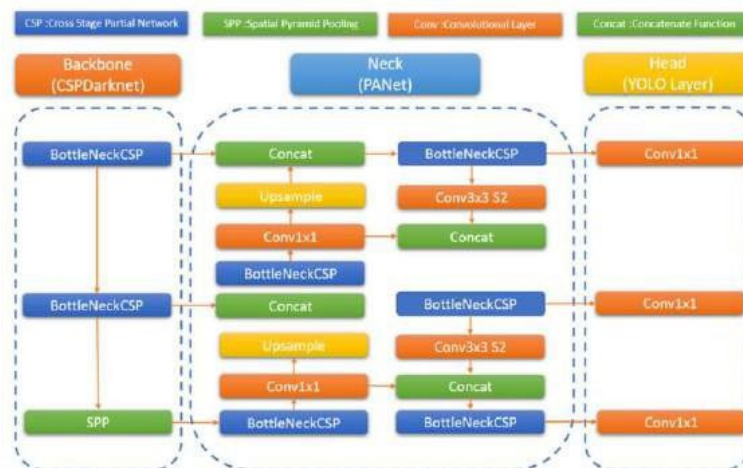
Setelah melalui tahap praproses, frame-frame video selanjutnya diproses untuk mendeteksi objek di dalamnya menggunakan algoritma YOLO. Dalam penelitian ini menggunakan model YOLOv5 pre-trained 'class_Yolov5_head' yang dikembangkan oleh RichardoMrMu untuk mendeteksi pose kepala [16]. Model ini dipilih karena memiliki akurasi dan kecepatan yang tinggi untuk deteksi kepala, seperti yang ditunjukkan dalam hasil pengujian pada gambar 2.



Gambar 2. Hasil pengujian model

YOLO (You Only Look Once) adalah algoritma pengenalan objek yang terkenal dengan kemampuannya mendeteksi dan mengklasifikasikan objek secara cepat dan akurat dalam satu kali pandang [8]. YOLO membagi gambar input menjadi grid berukuran $S \times S$. Setiap sel grid bertanggung jawab untuk mendeteksi objek jika pusat objek jatuh ke dalam sel tersebut. Model ini memprediksi beberapa bounding box dan nilai confidence untuk setiap sel grid, yang terdiri dari lima prediksi: koordinat (x, y) untuk pusat kotak, lebar (w), tinggi (h), dan nilai confidence yang mengukur seberapa yakinnya model bahwa kotak tersebut berisi objek [18].

YOLOv5 adalah salah satu model deteksi objek real-time yang paling populer saat ini. Model ini terkenal karena akurasi dan kecepatannya yang tinggi. YOLOv5 memiliki arsitektur model yang sederhana dan mudah dipahami, sehingga mudah diimplementasikan dan dimodifikasi [14]. Arsitektur YOLO V5 terdiri dari tiga komponen utama, yaitu Backbone, Neck, dan Head [19]. Seperti yang ditunjukkan pada Gambar 3.



Gambar 3. Arsitektur YOLO V5 [19]

Pada arsitektur yolo V5, backbone menggunakan arsitektur Cross Stage Partial (CSP) untuk mengekstrak fitur pada berbagai tingkatan skala. Fitur-fitur gambar kemudian diproses melalui bottleneck dan dialirkan ke Neck, yang terdiri dari PA-Net dan Special Pyramid Pooling (SPP). Neck tersusun atas serangkaian BottleNeckCSP yang dikonkatenasi dan beberapa layer konvolusi. Terakhir, Head menggabungkan fitur gambar menggunakan jaringan konvolusi bertumpuk untuk memproses prediksi kotak dan kelas objek. Lokalisasi kotak prediksi memanfaatkan algoritma pelacakan objek yang secara iteratif menyempurnakan posisi bounding box[20].

D. Pelacakan dan Pemberian ID

Setelah objek terdeteksi, langkah selanjutnya adalah melacak objek menggunakan perhitungan nilai Intersection over Union (IOU). IOU adalah metrik yang digunakan untuk mengevaluasi akurasi bounding box yang diprediksi oleh model. IOU dihitung dengan membagi luas tumpang tindih antara kotak prediksi dan kotak kebenaran tanah (ground truth) dengan luas gabungan dari kedua kotak tersebut [19]. Nilai IoU (Intersection over Union) dalam deteksi objek bervariasi antara 0 dan 1, dengan nilai 1 menunjukkan tumpang tindih sempurna antara kotak prediksi dan anotasi ground truth. Semakin tinggi nilai IoU, semakin dekat kotak prediksi dengan objek sebenarnya, yang menunjukkan akurasi model deteksi objek yang tinggi[21]. Sebaliknya, nilai IoU yang rendah menunjukkan bahwa kotak prediksi tidak tumpang tindih atau tumpang tindihnya sangat kecil dengan anotasi ground truth [21]. Hal ini menandakan bahwa model deteksi objek kurang akurat dalam mengidentifikasi dan melokalisasi objek dalam gambar atau video.

Pelacakan dilakukan untuk menentukan apakah objek yang terdeteksi adalah objek yang sudah ada dalam daftar pelacak. Ketika objek yang terdeteksi tidak ada dalam daftar pelacak dan berada di dalam area pemberian ID, maka objek tersebut akan diberikan ID dan ditambahkan ke dalam daftar pelacak. Sedangkan untuk objek yang sudah ada dalam daftar pelacak, maka posisinya akan terus diperbarui untuk mempertahankan ID yang telah diberikan pada objek tersebut.

E. Pemutaran Audio

Langkah penting berikutnya adalah proses pemutaran audio ketika objek yang dilacak atau yang memiliki ID unik memasuki area deteksi yang telah ditentukan. Program memeriksa apakah pusat objek (titik tengah bawah dari bounding box) berada di dalam area deteksi. Jika objek yang dilacak memasuki area deteksi dan belum memutar audio sebelumnya, maka program akan memutar audio. Untuk menghindari pemutaran berulang, setiap objek hanya dapat sekali memicu pemutaran audio, sehingga tidak dapat memutar audio untuk kedua kalinya ketika objek yang sama kembali memasuki area deteksi. Modul pygame digunakan dalam program ini mengatur pemutaran audio [22].

Program akan dievaluasi dengan melakukan pengujian menggunakan input berupa video rekaman CCTV yang didapatkan dari situs web youtube. Video yang berisi rekaman orang berjalan akan dibagi menjadi beberapa segmen untuk mensimulasikan berbagai skenario saat pengunjung memasuki toko. Program akan diuji untuk mendeteksi dan memberikan bounding box pada setiap objek kepala, lalu memberikan ID pada objek yang masuk atau melewati area berwarna biru, dan memutar audio sebanyak satu kali pada setiap objek yang memiliki ID ketika masuk ke area berwarna hijau. Objek yang berjalan melalui arah sebaliknya akan dianggap sebagai pengunjung yang hendak keluar dari toko. Oleh karena itu, audio tidak akan diputar karena objek tersebut tidak memiliki ID. Pengujian ini tidak dipengaruhi oleh kondisi pencahayaan, noise, maupun cuaca.

III. HASIL DAN PEMBAHASAN

Pada bagian ini akan membahas hasil dari pengujian yang telah dilakukan. Pembahasan dimulai dari input video, mencakup karakteristik data uji dalam berbagai kondisi. Selanjutnya, Prapemrosesan yang terdiri dari perubahan ukuran pada video yang dimasukkan. Tahap deteksi objek dievaluasi untuk melihat performa model dalam berbagai skenario. Kemudian, pelacakan dan pemberian ID dibahas untuk memastikan akurasi identifikasi objek. Terakhir, mekanisme pemutaran audio sebagai peringatan saat objek terdeteksi.

A. Input Video

Pada tahap awal, video yang telah disiapkan dimasukkan ke dalam program. Video yang digunakan berasal dari saluran YouTube The Real Samui Webcam yang menayangkan video jalanan secara langsung selama 24 jam <https://t.ly/-rGjh>. Video ini menampilkan berbagai orang berjalan yang akan disimulasikan sebagai kondisi saat pengunjung hendak memasuki toko.

Video dibagi menjadi lima segmen, dengan potongan video yang memiliki karakteristik jumlah orang berbeda diambil untuk mensimulasikan berbagai skenario saat pengunjung memasuki toko. Karakteristik dari setiap video dapat dilihat pada Tabel 1.

TABEL I
 TABEL KARAKTERISTIK VIDEO

No	Nama Video	Durasi	Jumlah Objek	Keterangan
1	Single Entry	10 detik	1	Menampilkan satu orang yang berjalan lurus
2	Multiple Entry	10 detik	3	Menampilkan beberapa orang yang berjalan lurus
3	Reentry	60 detik	3	Menampilkan satu orang yang berjalan lurus keluar dari sudut pandang kamera dan kembali masuk
4	Dengan Hambatan	15 detik	11	Menampilkan sekelompok orang yang berjalan lurus sambil berdekatan sehingga menutupi anak-anak yang ada di belakang
5	Durasi	30 detik	10	Satu Orang Berdiam Diri disaat Orang lain Berjalan lurus

B. Prapemrosesan

Pada langkah ini, setiap frame video diubah ukurannya dari resolusi aslinya menjadi resolusi yang lebih kecil. Contohnya, video asli dengan resolusi 1920x1080 piksel diubah menjadi 640x(rasio asli). Resolusi 640 piksel dipilih karena cukup besar untuk menjaga kejernihan gambar, namun juga cukup kecil untuk menghemat penggunaan daya GPU. Hasil proses resize dapat dilihat pada Gambar 4.



Gambar. 4. Hasil Proses Resize

C. Deteksi Objek

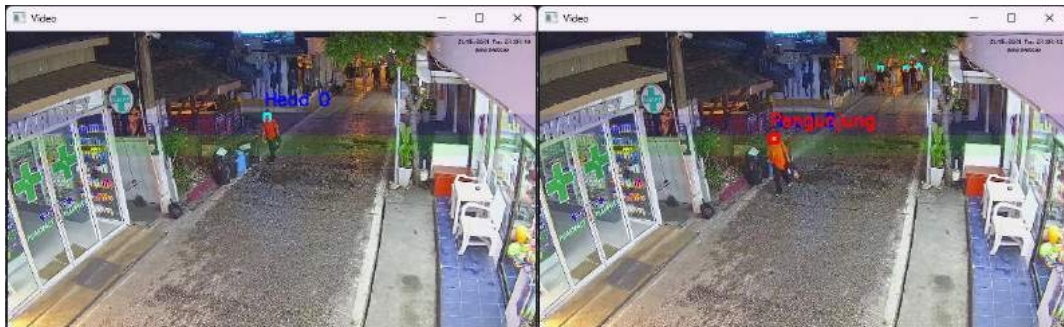
Dua area poligon ditetapkan dalam program ini: pts_id untuk pemberian ID, ditandai dengan poligon berwarna biru, dan pts_audio untuk pemutaran audio, ditandai dengan poligon berwarna hijau yang terletak setelah pts_id. Pengujian dilakukan untuk memverifikasi bahwa objek yang melewati pts_id akan mendapatkan ID dan kesempatan untuk memutar audio. Saat objek dengan ID terus bergerak dan memasuki pts_audio, audio akan diputar. Pengujian dibatasi pada pergerakan objek yang bergerak lurus melewati pts_id dan kemudian pts_audio sebagai simulasi pengunjung masuk. Sebaliknya, objek yang melewati pts_audio terlebih dahulu dan kemudian pts_id dianggap sebagai pengunjung keluar, sehingga audio tidak diputar.

Proses dimulai dengan mendeteksi objek kepala manusia di setiap frame video menggunakan model YOLOv5, yang kemudian diberikan bounding box berwarna hijau dan titik merah sebagai pusat objeknya. Setelah itu, pusat dari bounding box objek diuji terhadap (pts_id). Jika pusat objek berada dalam area pemberian ID, maka objek yang tidak memiliki ID akan diberikan ID baru. Setelah ID diberikan, proses pelacakan objek dilakukan untuk mempertahankan ID objek yang telah diberikan. Pelacakan menggunakan Intersection over Union (IoU) untuk memperbarui posisi objek dalam daftar pelacak (object_tracker). Setelah itu, pusat bounding box diuji terhadap (pts_audio). Jika objek yang memiliki ID berada di dalam area pemutaran audio dan belum pernah memutar audio sebelumnya, maka audio akan diputar sekali untuk satu ID. Hasil dari pengujian dapat dilihat pada tabel 2.

TABEL II
 TABEL HASIL PENGUJIAN

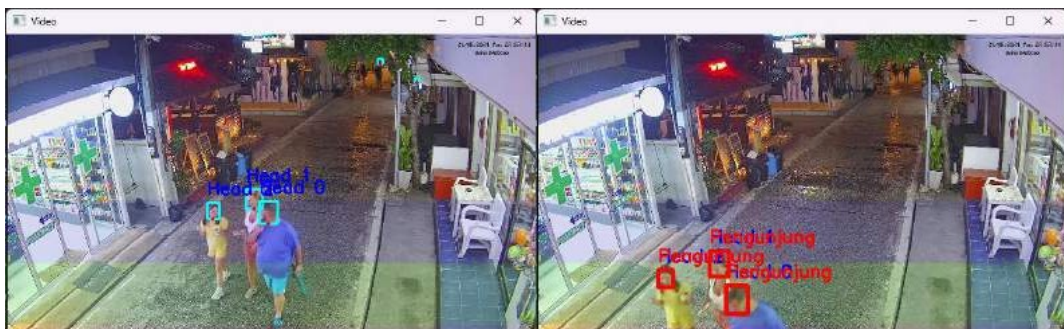
Video	Nama Video	Jumlah Objek	Objek Terdeteksi	Pemberian ID	Pemutaran Audio	Average Response Time (ms)
1	Single Entry	1	1	1	1	0
2	Multiple Entry	3	3	3	3	0,3
3	Reentry	3	3	4	4	0,5
4	Dengan Hambatan	11	11	10	10	0,8
5	Durasi	10	10	10	10	0,7

1. Pada pengujian video Single Entry, program berhasil mendeteksi 1 objek kepala dan memberikan ID, lalu berhasil memutar audio sebanyak 1 kali. Untuk hasil pengujian video Single Entry dapat dilihat pada gambar 5.



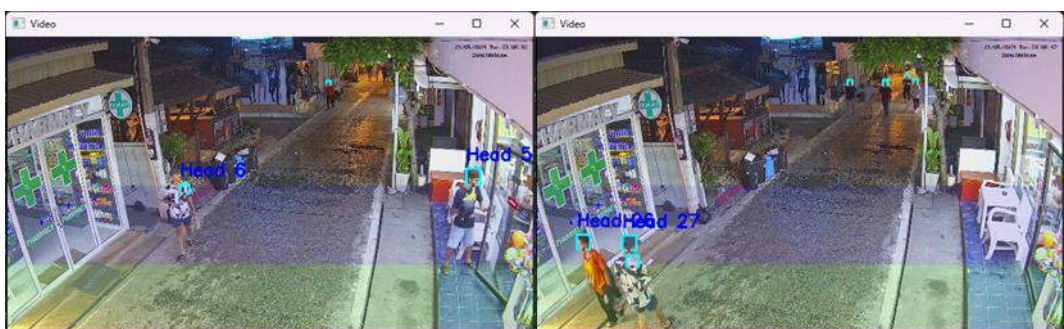
Gambar. 5. Hasil Pengujian Video Single Entry

2. Pada pengujian video Multiple Entry, program berhasil mendeteksi 3 objek kepala dan memberikan ID pada masing masing objek, lalu berhasil memutar audio sebanyak 3 kali. Untuk hasil pengujian video Multiple Entry dapat dilihat pada gambar 6.



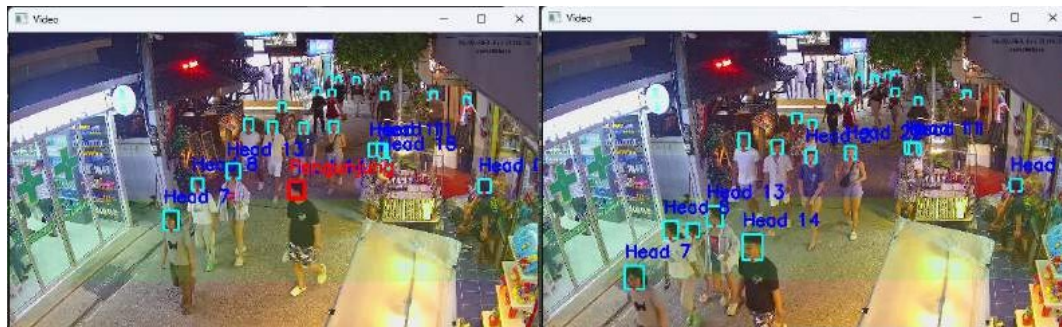
Gambar. 6. Hasil Pengujian Video Multiple Entry

3. Pada pengujian video Reentry, program berhasil mendeteksi 2 objek kepala dan memberikan ID untuk masing masing objek. Kemudian disaat objek dengan label head6 keluar dari sudut pandang kamera dan kembali masuk ke area pemberian ID, ID baru berupa label Head27 diberikan. Sehingga objek tersebut dapat memutar audio untuk kedua kalinya. Untuk hasil pengujian video Reentry dapat dilihat pada gambar 7.



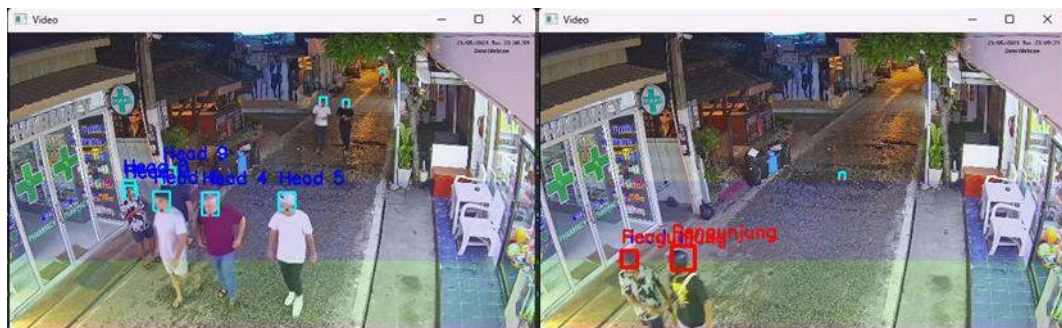
Gambar. 7. Hasil Pengujian Video Reentry

4. Pada pengujian video Dengan Hambatan, program berhasil mendeteksi 10 objek kepala dan kesulitan mendeteksi 1 objek berupa anak-anak yang tertutup oleh objek lain. Sehingga hanya memberikan sebanyak 10 ID dan memutar audio sebanyak 10 kali. Untuk hasil pengujian video Dengan Hambatan dapat dilihat pada gambar 8.



Gambar. 8. Hasil Pengujian Video Dengan Hambatan

5. Pada pengujian video Durasi, program berhasil mendeteksi 10 objek kepala dan memberikan ID pada masing-masing objek. Audio diputar sebanyak 10 kali, disaat objek dengan label head1 yang hanya berdiam diri akhirnya bergerak masuk ke area Audio. Untuk hasil pengujian video Durasi dapat dilihat pada gambar 9.



Gambar. 9. Hasil Pengujian Video Durasi

Hasil pengujian menunjukkan bahwa program berhasil memberikan ID dan memutar audio dengan cepat dan tanpa delay, pada setiap objek yang terdeteksi melewati area yang telah ditentukan. Pada skenario di mana orang berjalan dari arah berlawanan atau keluar area, audio berhasil tidak diputar. Hal ini menunjukkan bahwa program mampu menjalankan fungsi utamanya dengan baik, terutama pada kondisi dengan jumlah orang yang sedikit dan saling berjauhan. Namun, program terkadang mengalami kendala dalam mempertahankan ID yang telah diberikan kepada objek. Hal ini terjadi ketika model kesulitan mendeteksi objek di tengah keramaian, di mana terdapat banyak orang dan benda asing yang menghalangi. Akibatnya, ID objek sering menghilang dan muncul kembali, sehingga menyebabkan audio gagal diputar. Hal ini menunjukkan bahwa kemampuan model dalam mendeteksi objek perlu ditingkatkan dengan melatihnya menggunakan data set yang lebih sesuai. Jika dibandingkan dengan penelitian sebelumnya yang menggunakan sensor ultrasonik dan PIR, penelitian ini menggunakan algoritma YOLOv5 untuk mendeteksi objek kepala manusia dan mampu memutar audio dengan respon rata-rata 0 hingga 0.7 milidetik.

IV. KESIMPULAN

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa penggunaan algoritma You Only Look Once (YOLO) menggunakan model YOLOv5 pre-trained 'class_Yolov5_head' yang dikembangkan oleh RichardoMrMu sangat mempengaruhi kinerja program dalam memberikan ID dan memutar audio. Performa program yang terbaik dalam memberikan ID dan memutar audio terjadi pada pengujian dengan jumlah objek yang sedikit dan saling berjauhan. Seiring dengan meningkatnya jumlah objek dan kondisi keramaian, program mengalami kendala dalam mempertahankan ID objek, yang menunjukkan bahwa model perlu ditingkatkan kemampuannya dalam mendeteksi dan melacak objek. Untuk penelitian lebih lanjut, disarankan untuk melatih model menggunakan dataset yang lebih beragam dan sesuai dengan kondisi nyata di lapangan.

DAFTAR PUSTAKA

- [1] A. E. Pribadi, E. Darwiyanto, and J. H. Husen, "Perencanaan Strategis Sistem Informasi Pada Ritel Toko Swalayan Karunia," *e-Proceeding of Engineering*, vol. 7, no. 1, 2020.
- [2] A. H. P. Subakti, A. M. H. Pardede, and M. A. S. Syari, "RANCANGAN SISTEM NOTIFIKASI KEDATANGAN PEMBELI DENGAN SUARA MENGGUNAKAN ARDUINO," *JTIK (Jurnal Teknik Informatika Kaputama)*, vol. 7, no. 1, 2023, doi: 10.59697/jtik.v7i1.24.
- [3] A. A. Syukur, "Implementasi Webcam sebagai Pendeteksi Wajah pada Sistem Keamanan Perumahan menggunakan Image Processing," *ELECTRICES*, vol. 2, no. 1, 2020, doi: 10.32722/ees.v2i1.2791.
- [4] G. N. Rizkatama, A. Nugroho, and A. F. Suni, "Sistem Cerdas Penghitung Jumlah Mobil untuk Mengetahui Keterse- diaan Lahan Parkir berbasis Python dan YOLO v4," *Edu Komputika Journal*, vol. 8, no. 2, 2021, doi: 10.15294/edu- komputika.v8i2.47865.
- [5] F. A. Saputra and J. C. Chandra, "Prototipe Sistem Keamanan Ruang Server Otomatis Menggunakan ESP32CAM dan Algoritma You Only Look Once (YOLO)," *Jurnal TICOM: Technology of Information and Communication*, vol. 11, no. 1, pp. 62–67, 2022.
- [6] E. A. Z. Hamidi, M. R. Effendi, and M. R. Ramdani, "Prototipe Sistem Keamanan Rumah Berbasis Web dan SMS Gateway," *TELKA - Telekomunikasi, Elektronika, Komputasi dan Kontrol*, vol. 6, no. 1, 2020, doi: 10.15575/telka.v6n1.56-65.
- [7] J. Nishchal, S. Reddy, and P. N. Navya, "Automated Cheating Detection in Exams using Posture and Emotion Anal- ysis," in *Proceedings of CONECCCT 2020 - 6th IEEE International Conference on Electronics, Computing and Com- munication Technologies*, 2020. doi: 10.1109/CONECCCT50063.2020.9198691.
- [8] A. Riansyah and A. H. Mirza, "Pendeteksi Mobil Berdasarkan Merek dan Tipe Menggunakan Algoritma YOLO," *SMATIKA JURNAL*, vol. 13, no. 01, 2023, doi: 10.32664/smatika.v13i01.719.
- [9] B. Hardiansyah and A. Primasetya, "Sistem Deteksi Penggunaan masker (Face Mask Detection) Menggunakan AI- goritma Deep Learning YOLOv4," *STAINS (SEMINAR NASIONAL TEKNOLOGI & SAINS)*, vol. 2, no. 1, 2023.
- [10] A. Srinivasan, A. Srikanth, H. Indrajit, and V. Narasimhan, "A Novel Approach for Road Accident Detection using DETR Algorithm," in *2020 International Conference on Intelligent Data Science Technologies and Applications, IDSTA 2020*, 2020. doi: 10.1109/IDSTA50958.2020.9263703.
- [11] C. J. Lin, S. Y. Jeng, and H. W. Lioa, "A Real-Time Vehicle Counting, Speed Estimation, and Classification System Based on Virtual Detection Zone and YOLO," *Math Probl Eng*, vol. 2021, 2021, doi: 10.1155/2021/1577614.
- [12] N. H. Tasnim, S. Afrin, B. Biswas, A. A. Anye, and R. Khan, "Automatic classification of textile visual pollutants using deep learning networks," *Alexandria Engineering Journal*, vol. 62, 2023, doi: 10.1016/j.aej.2022.07.039.
- [13] F. Bimantoro, I. G. P. S. Wijaya, and M. R. Aohana, "Pendeteksian Kecurangan Ujian Melalui CCTV Menggunakan Algoritma YOLOv5," *Seminar Nasional Teknologi & Sains*, vol. 3, no. 1, 2024, doi: 10.29407/stains.v3i1.4360.
- [14] B. Yan, P. Fan, X. Lei, Z. Liu, and F. Yang, "A real-time apple targets detection method for picking robot based on improved YOLOv5," *Remote Sens (Basel)*, vol. 13, no. 9, 2021, doi: 10.3390/rs13091619.
- [15] G. Jocher, "ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation (v7.0)," <https://github.com/ultralytics/yolov5/tree/v7.0>, 2022.
- [16] RichardoMrMu, "RichardoMrMu/class_yolov5_head," https://github.com/RichardoMrMu/class_yolov5_head.
- [17] H. Feng, G. Mu, S. Zhong, P. Zhang, and T. Yuan, "Benchmark Analysis of YOLO Performance on Edge Intelli- gence Devices," *Cryptography*, vol. 6, no. 2, 2022, doi: 10.3390/cryptography6020016.
- [18] M. A. Maulana, "SISTEM PENDETEKSI PHYSICAL DISTANCE PADA ANTRIAN MENGGUNAKAN METODE YOLO (YOU ONLY LOOK ONCE) V3," *Jurnal IPTEK*, vol. 7, no. 1, 2023, doi: 10.31543/jii.v7i1.225.
- [19] Y. Egi, M. Hajyzadeh, and E. Eyceyurt, "Drone-Computer Communication Based Tomato Generative Organ Count- ing Model Using YOLO V5 and Deep-Sort," *Agriculture (Switzerland)*, vol. 12, no. 9, 2022, doi: 10.3390/agricul- ture12091290.
- [20] Z. Chen *et al.*, "Plant Disease Recognition Model Based on Improved YOLOv5," *Agronomy*, vol. 12, no. 2, 2022, doi: 10.3390/agronomy12020365.
- [21] H. Husnan, C. Fatichah, and R. Dikairono, "Deteksi Objek Menggunakan Metode YOLO dan Implementasinya pada Robot Bawah Air," *Jurnal Teknik ITS*, vol. 12, no. 3, 2023, doi: 10.12962/j23373539.v12i3.122326.
- [22] M. Mahendru and S. K. Dubey, "Real time object detection with audio feedback using Yolo vs. Yolo_V3," in *Pro- ceedings of the Confluence 2021: 11th International Conference on Cloud Computing, Data Science and Engineer- ing*, 2021. doi: 10.1109/Confluence51648.2021.9377064.