

ISSN (ONLINE) 2598-9936



INDONESIAN JOURNAL OF INNOVATION STUDIES
PUBLISHED BY
UNIVERSITAS MUHAMMADIYAH SIDOARJO

Table Of Contents

Journal Cover	1
Author[s] Statement	3
Editorial Team	4
Article information	5
Check this article update (crossmark)	5
Check this article impact	5
Cite this article.....	5
Title page	6
Article Title	6
Author information	6
Abstract	6
Article content	7

Originality Statement

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

Conflict of Interest Statement

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright Statement

Copyright © Author(s). This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

Indonesian Journal of Innovation Studies

Vol. 26 No. 4 (2025): October

DOI: 10.21070/ijins.v26i4.1851

EDITORIAL TEAM

Editor in Chief

Dr. Hindarto, Universitas Muhammadiyah Sidoarjo, Indonesia

Managing Editor

Mochammad Tanzil Multazam, Universitas Muhammadiyah Sidoarjo, Indonesia

Editors

Fika Megawati, Universitas Muhammadiyah Sidoarjo, Indonesia

Mahardika Darmawan Kusuma Wardana, Universitas Muhammadiyah Sidoarjo, Indonesia

Wiwit Wahyu Wijayanti, Universitas Muhammadiyah Sidoarjo, Indonesia

Farkhod Abdurakhmonov, Silk Road International Tourism University, Uzbekistan

Bobur Sobirov, Samarkand Institute of Economics and Service, Uzbekistan

Evi Rinata, Universitas Muhammadiyah Sidoarjo, Indonesia

M Faisal Amir, Universitas Muhammadiyah Sidoarjo, Indonesia

Dr. Hana Catur Wahyuni, Universitas Muhammadiyah Sidoarjo, Indonesia

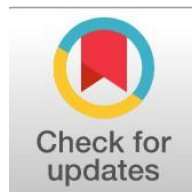
Complete list of editorial team ([link](#))

Complete list of indexing services for this journal ([link](#))

How to submit to this journal ([link](#))

Article information

Check this article update (crossmark)



Check this article impact (*)



Save this article to Mendeley



(*) Time for indexing process is various, depends on indexing database platform

Risk Detection System for Children Putting Objects into Mouth Based on Computer Vision using YOLOv11n

Sistem Deteksi Risiko Anak Memasukkan Benda ke dalam Mulut Berbasis Computer Vision menggunakan YOLOv11n

Nofri Ramadan, nofri.ramadan.2205356@students.um.ac.id, (1)

Program Studi Teknik Informatika, Universitas Negeri Malang, Indonesia

Muis Muhtadi, muis.muhtadi.ft@um.ac.id, ()

Program Studi Teknik Informatika, Universitas Negeri Malang, Indonesia

Muhammad Zulfiqar Rafi, muhammad.zulfiqar.2205356@students.um.ac.id, ()

Program Studi Teknik Informatika, Universitas Negeri Malang, Indonesia

Rinakit Estu Waluyo, rinakitestu@gmail.com, ()

Program Studi Teknik Informatika, Universitas Negeri Malang, Indonesia

⁽¹⁾ Corresponding author

Abstract

Background: Young children commonly place objects near or into their mouths, creating safety concerns that require constant supervision. **Specific Background:** Advances in computer vision enable real-time recognition of hands, faces, and objects, allowing automated identification of behaviors that may lead to hazardous mouth-related interactions. **Knowledge Gap:** Few systems combine object detection and human pose estimation specifically to assess risks related to toddler hand–mouth–object interactions in real-time environments. **Aims:** This study develops a risk detection system using YOLOv11n to recognize hands, faces, and objects while classifying conditions into safe or risky based on Euclidean distance between hand and mouth keypoints. **Results:** The system produces 92% accuracy in scenarios without objects and 74% in scenarios with objects, demonstrating its capability to differentiate between safe and risky conditions. **Novelty:** This research introduces an integrated spatial analysis approach that evaluates real-time proximity among hands, mouth, and objects, rather than detecting these elements independently. **Implication:** The system provides practical potential for real-time child safety monitoring, offering earlier awareness of mouth-related object risks in various activity settings.

Highlights

- The system identifies hand–mouth proximity to classify safe and risky situations.
- Integrated pose estimation and object detection enable spatial risk assessment in real time.
- The model supports early awareness of mouth-related object risks during toddler activities.

Keywords

Object Detection, Pose Estimation, Computer Vision, YOLO, Child Safety Monitoring

Published date: 2025-12-12

I. Pendahuluan

Perlindungan terhadap keselamatan anak menjadi perhatian utama, khususnya ketika mereka berada di lingkungan yang mengandung berbagai risiko dari benda atau objek berbahaya [1]. Anak-anak pada umumnya senang mengeksplorasi lingkungannya dengan cara memasukkan berbagai benda ke dalam mulut, sehingga insiden tertelannya benda asing maupun zat berbahaya kerap menjadi masalah serius bagi keluarga [2]. Masalah kegawatan seperti aspirasi benda asing atau tersedak pada usia toddler merupakan kondisi yang serius dan dapat mengancam nyawa bila tidak segera ditangani, sehingga pengawasan intensif dari orang tua maupun pengasuh menjadi aspek yang sangat krusial pada tahap perkembangan ini [3].

Namun, pengawasan yang tidak memadai, seperti meninggalkan anak tanpa pendampingan, sering dikaitkan dengan berbagai dampak negatif terhadap kesehatan dan keselamatan anak [4]. Menurut data SUSY SAFE, yaitu registry internasional terbesar yang mengumpulkan kasus kecelakaan benda asing pada anak usia 0–14 tahun di negara-negara Uni Eropa dan beberapa negara non-EU, makanan menjadi penyebab hingga 26% cedera akibat aspirasi maupun tertelannya benda asing. Faktor penyebab yang paling sering dilaporkan adalah tulang ayam dan ikan (32%), kacang tanah (22%), serta biji-bijian (16%) [5]. Fakta ini menunjukkan bahwa pengawasan orang tua maupun pengasuh yang tidak cukup dapat menyebabkan risiko cedera, sehingga diperlukan strategi tambahan yang lebih efektif untuk melindungi anak dari potensi bahaya di lingkungannya.

Dalam beberapa tahun terakhir, kemajuan di bidang computer vision dan machine learning telah membuka peluang baru untuk meningkatkan keselamatan anak melalui sistem deteksi otomatis [6]. Teknologi ini mampu memberikan prediksi yang mudah dipahami, baik berupa lokasi objek maupun segmentasi visual, sehingga sistem pemantauan dapat memberikan peringatan dini sekaligus menjelaskan bagaimana suatu situasi dapat membahayakan anak [7]. Teknik yang banyak digunakan adalah object detection dan human pose estimation (HPE), di mana object detection berfungsi menemukan serta mengklasifikasikan objek, sementara HPE dapat memperkirakan posisi tubuh untuk mendeteksi aktivitas berisiko seperti memasukkan benda asing ke mulut [8], [9].

Salah satu arsitektur yang banyak dimanfaatkan dalam deteksi objek adalah YOLO (You Only Look Once) [10]. Model ini telah berkembang pesat, bahkan ini dapat digunakan untuk pose estimation melalui varian seperti YOLO Pose yang mampu menghubungkan keypoint tubuh dengan anchor [11]. Versi terbarunya, YOLOv11, menghadirkan peningkatan signifikan dalam hal akurasi, efisiensi, serta kemampuan adaptasi untuk berbasis tugas, mulai dari ibjek detection hingga instance segmentation dan pose estimation [12].

Berbagai penelitian telah menunjukkan efektivitas YOLO dalam konteks berbeda. Misalnya, YOLOv8 digunakan untuk memantau kondisi bayi melalui deteksi emosi dan gerakan secara real-time [13]. YOLOv11 juga dioptimalkan khusus untuk deteksi anak di lingkungan medis dengan akurasi tinggi melalui Child Detection Dataset [14]. Penelitian lain mengembangkan sistem alarm jatuh berbasis YOLOv8 yang berhasil mencapai akurasi 96,33% [15]. Bahkan, YOLOv5 dikombinasikan dengan metode hybrid untuk klasifikasi emosi lintas budaya anak-anak dan mencapai akurasi hingga 96% [16]. Dalam konteks berbeda, YOLO juga telah digunakan untuk deteksi benda berbahaya berskala kecil di area kargo bandara [17], deteksi facial landmark [18], serta hand gesture recognition yang lebih ringan untuk interaksi manusia-komputer [19]. Sementara itu, sistem berbasis CCTV juga telah memanfaatkan YOLO untuk mendeteksi anak dan objek berbahaya seperti pisau dan gunting dengan akurasi memadai, meskipun performanya menurun pada kondisi multi-objek [20].

Dari berbagai studi sebelumnya, YOLO terbukti fleksibel dan efektif di beragam scenario. Namun, sebagian besar penelitian masih berfokus pada konteks umum seperti deteksi emosi, gesture, atau keselamatan kerja, sementara penerapannya untuk pemantauan keselamatan balita di rumah masih jarang dilakukan. Tantangan seperti ukuran tangan dan mulut balita yang kecil serta interaksi dengan objek dalam kondisi beragam juga belum banyak diteliti. Penelitian ini bertujuan mengembangkan sistem pemantauan balita berbasis YOLO yang mampu mendeteksi tangan, mulut, dan objek secara real time. System ini ditujukan untuk mengenali potensi bahaya, khususnya ketika tangan atau objek mendekati area mulut sehingga risiko kecelakaan seperti memasukkan benda berbahaya dapat diminimalkan. Dengan demikian, penelitian ini memberikan kontribusi terhadap pengembangan teknologi pemantauan anak yang lebih adaptif, praktis, dan sesuai kebutuhan keluarga modern.

II. Metode

Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan desain pengembangan sistem computer vision. Penelitian bertujuan untuk mengembangkan dan mengevaluasi sistem deteksi risiko balita memasukkan benda ke mulut berbasis object detection dan human pose estimation (HPE). Variabel utama penelitian meliputi variabel independen berupa model deep learning (YOLOv11n) dan variabel dependen berupa performa sistem deteksi risiko yang diukur melalui metrik akurasi, presisi, dan recall.

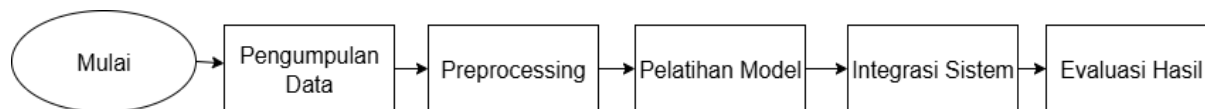
1. Waktu dan Tempat Penelitian

Seluruh proses penelitian, yang mencakup pengumpulan data, preprocessing, pelatihan model, integrasi, dan evaluasi sistem, dilaksanakan pada periode Penelitian, September hingga November 2025. Penelitian dilakukan secara simulatif pada lingkungan komputasi dengan spesifikasi: prosesor Intel Core i7 generasi ke-11, RAM 16 GB, penyimpanan SSD 512 GB, serta GPU NVIDIA RTX 3050Ti dengan VRAM 4 GB. Sistem operasi yang digunakan adalah Windows 11, dan pemrograman dilakukan menggunakan Python versi 3.12.2 melalui Jupyter Notebook dengan pustaka OpenCV, Ultralytics YOLO, dan [ISSN 2598-9936 \(online\)](https://www.scribd.com/document/99361892/ISSN-2598-9936-online), <https://ijins.umsida.ac.id>, published by [Universitas Muhammadiyah Sidoarjo](https://www.scribd.com/document/99361892/ISSN-2598-9936-online)

PyTorch.

2. Prosedur Penelitian

Prosedur penelitian ini dijabarkan melalui beberapa tahapan utama yang dapat diringkas dalam alur kerja penelitian seperti terlihat pada Gambar 1.



Gambar 1. Alur proses penelitian

Tahapan penelitian secara garis besar mencakup Pengumpulan dan Preprocessing Dataset, Pelatihan Model, Integrasi Sistem Deteksi, dan Evaluasi Performa Sistem. Tahapan ini dilakukan secara sekuensial, dimulai dari penyiapan data masukan yang optimal hingga pengujian kinerja model yang terintegrasi.

3. Data Penelitian

Target atau sasaran penelitian ini adalah aktivitas balita yang berinteraksi dengan objek berbahaya. Subjek penelitian berupa citra yang memuat tiga entitas utama: wajah, tangan, dan objek. Dataset diperoleh dari sumber public bernotasi sebagai sampel respresentif. Data deteksi tangan berasal dari Ultralytics Hand Keypoints Dataset, sedangkan data wajah diambil dari Roboflow yang menyediakan anotasi keypoints. Untuk deteksi objek, digunakan model YOLOv11n pretrained dengan COCO Dataset yang berisi berbagai kategori objek umum.

4. Instrumen dan Teknik Pengumpulan Data

Instrumen utama dalam penelitian ini adalah Model YOLOv11n (arsitektur deep learning) yang diimplementasikan melalui framework Ultralytics. Teknik pengumpulan data dilakukan melalui kurasi dan pengunduhan dataset publik yang relevan, diikuti oleh tahapan preprocessing data citra. Tahap preprocessing ini sangat krusial untuk menyiapkan input yang optimal bagi model. Prosesnya meliputi konversi citra ke grayscale, dilanjutkan dengan resizing citra menjadi ukuran 640 x 640 piksel sesuai format input YOLOv11n. Selain itu, dilakukan augmentasi berupa rotasi, penyesuaian kecerahan (brightness adjustment), dan horizontal flip guna meningkatkan keragaman data latih. Seluruh anotasi juga disesuaikan dalam format YOLO agar koordinat objek dan keypoints dapat terbaca dengan benar oleh model.

5. Pelatihan Model

Pelatihan model dilakukan menggunakan dua pendekatan yang berbeda. Model YOLOv11n-Pose digunakan untuk mendeteksi tangan dan wajah, karena model ini memiliki kemampuan dalam mengenali titik-titik kunci (keypoints) posisi bagian tubuh manusia. Model ini dilatih ulang (fine-tuning) menggunakan dataset tangan dan wajah yang telah disiapkan sebelumnya, dengan bobot awal (pretrained weights) dari YOLOv11n yang sudah melalui pelatihan pada dataset umum. Sementara itu, model YOLOv11n pretrained dimanfaatkan untuk mendeteksi berbagai objek berbahaya yang berpotensi menimbulkan risiko pada anak. Model ini menggunakan bobot bawaan dari dataset COCO yang telah dilatih sebelumnya, sehingga dapat langsung mengenali objek umum tanpa memerlukan pelatihan tambahan.

6. Integrasi Sistem dan Makna Data

Sistem pada tahap integrasi menggabungkan hasil deteksi dari model pose dan object detection untuk menganalisis hubungan spasial di antara ketiganya. Ketika tangan dan objek terdeteksi dalam satu frame, jarak Euclidean antara keypoints tangan dan area mulut dihitung. Data dimaknai sebagai kondisi “risky” jika jarak tersebut berada di bawah ambang batas (threshold) tertentu, sedangkan kondisi lainnya diklasifikasikan sebagai “safe”. Proses ini memungkinkan sistem untuk secara otomatis mengenali potensi bahaya, seperti upaya anak memasukkan benda asing ke dalam mulutnya, sehingga mampu mendukung upaya pencegahan insiden sejak dini, dan ini merupakan penjabaran data yang diperoleh kaitannya dengan permasalahan penelitian.

7. Evaluasi Kinerja Sistem

Evaluasi dilakukan dengan membandingkan hasil prediksi model terhadap ground truth menggunakan sejumlah frame yang mewakili dua kondisi spesifik: anak di dekat objek berbahaya dan anak tanpa objek di sekitar. Penilaian performa sistem menggunakan tiga metrik utama: Akurasi (Accuracy), Presisi (Precision), dan Recall. Ketiga metrik ini dipilih karena mampu menggambarkan tingkat ketepatan dan sensitivitas sistem dalam mengklasifikasikan kondisi anak. Perhitungan setiap metrik dilakukan menggunakan persamaan (1), (2), dan (3).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

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

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

TP (True Positive) merupakan jumlah frame yang benar terdeteksi sebagai kondisi berisiko (risky), TN (True Negative) adalah jumlah frame yang benar terdeteksi sebagai kondisi aman (safe), FP (False Positive) menunjukkan jumlah frame yang salah diklasifikasikan sebagai risky, dan FN (False Negative) merupakan jumlah frame yang salah diklasifikasikan sebagai safe.

III. Hasil dan Pembahasan

Sebelum memaparkan hasil deteksi, penting dicatat bahwa system memiliki beberapa keterbatasan. Kinerja deteksi sangat dipengaruhi kualitas video uji, terutama pada kondisi pencahayaan buruk, gerakan cepat, atau resolusi rendah. Selain itu, system hanya dapat mengenali objek yang terdapat dalam model COCO, sehingga objek di luar kategori tersebut tidak dapat diklasifikasikan dengan tepat. Deteksi wajah juga bergantung pada keterlihatan wajah, jika tidak menghadap kamera atau tertutup, bounding box dan keypoints mulut sulit terdeteksi secara akurat. Hal serupa juga terjadi pada deteksi tangan apabila tangan tidak terlihat secara penuh, tertutup oleh benda lain, atau berada di luar bidang pandang kamera, yang dapat memengaruhi hasil analisis jarak antara tangan dan mulut.

Berdasarkan hasil implemantasi system, deteksi dilakukan pada beberapa frame yang diambil dari video simulasi aktivitas balita. Gambar 2 menunjukkan hasil deteksi pada kondisi ketika tangan atau objek tidak berada di dekat mulut balita. Dalam situasi ini, sistem berhasil memberikan informasi secara tepat dengan mengklasifikasikan status situasi sebagai safe.



Gambar 2. Visualisasi status safe

Sementara itu, Gambar 3 memperlihatkan kondisi ketika tangan balita mendekati atau menyentuh objek di area mulut. Pada kondisi ini, sistem secara otomatis mengubah status menjadi risky berwarna merah setelah jarak Euclidean antara keypoints tangan atau mulut berada di bawah ambang batas (threshold) yang telah ditentukan.



Gambar 3. Visualisasi status risky

Hasil pengujian model menghasilkan nilai Confusion Matrix yang menggambarkan distribusi prediksi sistem terhadap kondisi sebenarnya pada setiap kategori. Matriks ini digunakan untuk menilai sejauh mana sistem mampu membedakan antara situasi anak yang berada di dekat objek dan kondisi tanpa objek di sekitarnya. Berdasarkan hasil pengujian terhadap seluruh frame uji (200 frame total), diperoleh nilai Confusion Matrix seperti yang ditampilkan pada Tabel 1.

Kondisi	TP	TN	FP	FN
Dengan Objek	73	1	5	21
Tanpa Objek	91	1	2	6

Table 1. Confussion Matrix pada Kondisi Anak dengan dan tanpa Objek

Evaluasi kinerja sistem dilakukan terhadap 200 frame yang mewakili kondisi dengan objek (100 frame) dan kondisi tanpa objek (100 frame). Kinerja sistem dievaluasi menggunakan metrik Akurasi, Presisi (Precision), dan Recall (Sensitivitas) untuk memahami kualitas klasifikasi secara lebih detail. Hasil pengujian menunjukkan bahwa sistem mampu mencapai kinerja yang baik dalam mengenali kedua kondisi. Tingkat akurasi sistem secara keseluruhan, serta rincian metrik untuk masing-masing kondisi dirangkum dalam Tabel 2.

Kondisi	Accuracy	Precision	Recall
Dengan Objek	74%	47,1%	49,07%
Tanpa Objek	92%	63,5%	56,07%

Table 2. Perbandingan Performa Sistem pada Kondisi Anak dengan dan tanpa Objek

Hasil pengujian terhadap 200 frame menunjukkan adanya perbedaan kinerja yang cukup signifikan pada sistem. Pada situasi ketika anak berada di lingkungan yang terdapat berbagai objek, sistem hanya mampu mencapai akurasi sekitar 74%, dengan precision sebesar 47,16% dan recall 49,07%. Sebaliknya, ketika pengujian dilakukan tanpa keberadaan objek tambahan, akurasi sistem meningkat signifikan hingga mencapai 92%, disertai precision 63,57% dan recall 56,07%. Perbedaan hasil ini menunjukkan bahwa tingkat kerumitan visual akibat adanya objek di sekitar anak memiliki pengaruh besar terhadap kinerja sistem, terutama dalam proses pendeteksian dan pelacakan keypoints pada tangan dan wajah.

Fenomena *occlusion* ini menjadi penyebab utama meningkatnya kesalahan prediksi (*False Positive* dan *False Negative*), yang pada akhirnya menurunkan nilai precision dan recall pada kondisi dengan objek (lihat Tabel 1 dan Tabel 2). Ketika tangan anak sebagian tertutup oleh objek atau beririsan dengan latar belakang yang kompleks, model mengalami kesulitan dalam mempertahankan *spatial consistency* pada setiap *frame*. Selain itu, hasil tersebut juga menunjukkan bahwa sistem cenderung lebih sensitif terhadap konteks visual sederhana, di mana area tangan dan wajah terdeteksi dengan jelas tanpa gangguan dari benda lain. Temuan ini menegaskan bahwa pose estimation masih memiliki keterbatasan ketika digunakan pada lingkungan yang kurang terkontrol, terutama saat terdapat banyak objek gerherak dan kondisi pencahayaan yang tidak stabil. Walaupun model mampu membaca pola gerak dan ekspresi anak, kinerjanya tetap sangat dipengaruhi oleh kualitas visual serta jarak kamera dengan subjek.

Meski demikian, system secara keseluruhan tetap menunjukkan kemampuan adaptasi yang baik dalam membedakan kondisi anak saat ada maupun tidak ada objke. Kolaborasi antara pose estimation dan objek detection memungkinkan identifikasi posisi tangan dan wajah dengan akurasi spasial yang tinggi, sekaligus mendeteksi keberadaan dan jarak objek yang berpotensi membahayakan. Integrasi kedua model ini membuat system mampu menilai resiko interaksi anak dengan objek secara cepat dan realtime.

Kelebihan pendekatan multi-model ini dibanding model tunggal antara lain: (1) pendeteksian posisi tubuh tetap stabil meskipun objek bergerak, (2) sistem mampu menilai konteks interaksi bukan hanya keberadaan objek, dan (3) deteksi risiko tidak hanya berbasis klasifikasi tetapi juga ukuran jarak spasial. Hal ini tidak dapat dicapai oleh model tunggal seperti YOLO saja ataupun MediaPipe saja.

Beberapa faktor eksternal masih menjadi tantangan dalam menjaga konsistensi performa sistem, seperti variasi pencahayaan, sudut pengambilan kamera, visibilitas objek, serta perbedaan warna kulit anak yang dapat memengaruhi akurasi deteksi *landmark*. Selain itu, gerakan anak yang terlalu cepat sering menimbulkan *motion blur* pada *frame* video, sehingga posisi tangan atau wajah tidak terdeteksi secara utuh oleh model. Kualitas citra yang rendah juga dapat menurunkan ketajaman fitur visual yang diperlukan model untuk mengenali *keypoints* secara konsisten.

Jika dibandingkan dengan penelitian sebelumnya, sistem dalam penelitian ini memiliki arah pengembangan yang sejalan dengan beberapa studi terdahulu. Penelitian ini menawarkan novelty berupa integrasi beberapa model deteksi wajah, tangan dan objek yang dipadukan dengan perhitungan Euclidean distance sebagai indikator risiko perilaku secara langsung. Pendekatan ini memperluas fungsi YOLO dari sekedar pendeteksi objek statis menjadi alat analisis perilaku berbasis hubungan antar objek (objek human interaction analysis).

Keunikan penelitian ini terlihat dari kemampuan system mengidentifikasi perilaku berisiko melalui analisis hubungan spasial antara tangan, mulut, dan objek secara real-time, bukan hanya mendeteksi objek secara terpisah. Hal ini menunjukkan kontribusi baru yang tidak ditemui pada penelitian sebelumnya. Temuan tersebut juga memperkuat potensi penerapan system sebagai child safety monitoring system di rumah, sekolah, maupun area public. Ke depan, pengembangan dapat difokuskan pada peningkatan ketahanan terhadap perubahan pencahayaan, variasi sudut kamera, serta optimasi performa pada video dengan kualitas rendah. Selain itu, perluasan kemampuan sistem dalam mendeteksi berbagai bentuk perilaku berisiko lainnya seperti melempar, menggigit, atau menjilat benda akan memperluas cakupan penerapan sistem ini pada berbagai skenario keselamatan anak di lingkungan nyata.

IV. Kesimpulan

Penelitian ini telah berhasil mengembangkan dan mengevaluasi sistem deteksi risiko balita memasukkan benda ke mulut secara real-time berbasis integrasi object detection dan human pose estimation menggunakan arsitektur YOLOv11n, sehingga tujuan penelitian berhasil tercapai. Sistem ini secara otomatis mampu mengidentifikasi potensi bahaya perilaku anak, seperti interaksi tangan balita dengan objek yang memiliki risiko memasukkan benda ke mulut, melalui analisis jarak Euclidean. Temuan ini menunjukkan bahwa sistem memiliki kemampuan dalam membedakan kondisi aman dan berisiko, menegaskan bahwa pendekatan kombinatorik computer vision memiliki potensi besar untuk diterapkan dalam sistem pemantauan keselamatan anak (child safety monitoring system). Meskipun efektif, performa sistem menurun pada kondisi dengan kompleksitas visual tinggi, seperti occlusion, pencahayaan rendah, atau pergerakan cepat. Oleh karena itu, penelitian selanjutnya perlu diarahkan pada peningkatan robustness model terhadap kondisi lingkungan yang bervariasi, serta optimalisasi performa deteksi pada video real-time dengan objek bergerak cepat. Fokus prioritas ini dipilih karena kedua aspek tersebut memiliki pengaruh terbesar terhadap akurasi sistem dan relevansi langsung pada implementasi dunia nyata. Selain itu, pengembangan kemampuan sistem untuk mendeteksi perilaku berisiko lainnya dapat menjadi arah penelitian lanjutan setelah aspek prioritas tersebut terpenuhi.

Ucapan Terima Kasih

Penulis menyampaikan terima kasih kepada seluruh pihak yang telah mendukung penyelesaian penelitian ini. Penulis juga berterima kasih kepada pihak-pihak yang telah berperan dalam pengumpulan data, pengolahan informasi, serta pengujian system. Seluruh dukungan dan kerja sama tersebut sangat berarti bagi kelancaran penelitian ini. Semoga hasil penelitian ini dapat berkontribusi pada pengembangan teknologi dan menjadi dasar bagi penelitian selanjutnya.

References

1. K. Sivakrishna, J. S. Vaishnavi, A. Sreelekha, S. Sindhu, C. Varshitha, and G. Kalakoti, "Object Detection and Hazard Alert System for Child Safety on Robot using YOLO," in 2025 5th International Conference on Expert Clouds and Applications (ICOECA), IEEE, 2025, pp. 1094–1098.
2. H. M. Isa, S. A. Aldoseri, A. S. Abduljabbar, and K. A. Alsulaiti, "Accidental ingestion of foreign bodies/harmful materials in children from Bahrain: A retrospective cohort study," World J. Clin. Pediatr., vol. 12, no. 4, p. 205, 2023.
3. S. M. Damanik, E. Sitorus, and I. Mertajaya, "Penerapan Upaya Pencegahan dan Penanganan Aspirasi Benda Asing dan Kejang Demam pada Anak di Rumah," J. Comunitas Serv., vol. 3, no. 2, pp. 653–661, 2021.
4. T. Sultana, M. Ruiz-Casares, R. Iwo, M. Janus, and J. I. Nazif-Muñoz, "Maternal Education and Children Home Alone in 63 Low-and Middle-Income Countries," Glob. Pediatr. Health, vol. 11, p. 2333794X241258179, 2024.
5. M. Karišik, "Foreign body aspiration and ingestion in children," Acta Clin. Croat., vol. 62, no. Supplement 1, pp. 105–112, 2023.

6. F. A. Khan and A. Dey, "Towards Enhancing Child Safety: A Deep Learning Approach to Detect Child Safe and Unsafe Objects," in 2024 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), IEEE, 2024, pp. 123–128.
7. S. Zhu et al., "Cribnet: Enhancing infant safety in cribs through vision-based hazard detection," in 2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition (FG), IEEE, 2024, pp. 01–08.
8. C. Zheng et al., "Deep learning-based human pose estimation: A survey," *ACM Comput. Surv.*, vol. 56, no. 1, pp. 1–37, 2023.
9. A. Tripathi, M. K. Gupta, C. Srivastava, P. Dixit, and S. K. Pandey, "Object detection using yolo: A survey," in 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), IEEE, 2022, pp. 747–752.
10. P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo algorithm developments," *Procedia Comput. Sci.*, vol. 199, pp. 1066–1073, 2022.
11. D. Maji, S. Nagori, M. Mathew, and D. Poddar, "Yolo-pose: Enhancing yolo for multi person pose estimation using object keypoint similarity loss," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 2637–2646.
12. R. Khanam and M. Hussain, "YOLOv11: An Overview of the Key Architectural Enhancements." 2024. [Online]. Available: <https://arxiv.org/abs/2410.17725>
13. T. Arunkumar, S. Maheswaran, P. Dineshkumar, K. Geetha, A. Sureshkumar, and S. Praveenkumar, "Design and Implementation of an Astute Infant Monitoring System Based on YOLO v8 Algorithm," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, 2024, pp. 1–6.
14. S. Diop, F. Jouen, J. Bergounioux, and I. Trabelsi, "Fine-Tuned YOLO Model for Monitoring Children Across Medical Scenes Based on a Large-Scale Real-World Dataset for Children Detection," *IEEE Access*, 2025.
15. Z. Yang, B. Tsui, J. Ning, and Z. Wu, "Falling detection of toddlers based on improved YOLOv8 models," *Sensors*, vol. 24, no. 19, p. 6451, 2024.
16. A. M. Mekala et al., "Fused YOLO and Traditional Features for Emotion Recognition from Facial Images of Tamil and Russian Speaking Children: A Cross-Cultural Study," *IEEE Access*, 2025.
17. M. P. Salim, J. Ong, I. S. Edbert, and D. Surhatono, "Object detection for child Learning media," in 2022 8th International Conference on Science and Technology (ICST), IEEE, 2022, pp. 1–6.
18. Q. Wu, X. Wang, N. Li, S. Fong, L. Zhang, and J. Yang, "Real-Time Face and Facial Landmark Joint Detection based on End-to-End Deep Network," *IEEE Trans. Instrum. Meas.*, 2025.
19. S. Wang, C. Guo, R. Yang, Q. Zhang, and H. Ren, "A lightweight vision-based measurement for hand gesture information acquisition," *IEEE Sens. J.*, vol. 23, no. 5, pp. 4964–4973, 2022.
20. J. H. Tan and C. P. Goh, "Enhancing Child Safety: Computer Vision-Based Accident Detection for Infants and Toddlers," in 2024 3rd International Conference on Digital Transformation and Applications (ICDXA), 2024, pp. 1–5. doi: 10.1109/ICDXA61007.2024.10470712.