

Implementasi Sistem Deteksi *Bullying* Berbasis YOLOv11 pada CCTV Sekolah dengan *Dashboard Monitoring* dan Notifikasi Telegram

Yasmin Maulidia Andini Putri¹, Rania Ayu Praesmawardhani², Hamim Thohari Mahfudhillah³, Annisa Ahmada Atusta⁴

^{1,2,3,4}MTsN 6 Malang, Jl. Raya Sukoraharjo No.37, RW.01, Kepanjen, Kabupaten Malang, Jawa Timur 65163, Indonesia

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 10-08-2026

Revisi Akhir:

Diterbitkan Online: 25-09-2026

KATA KUNCI

bullying, CCTV, YOLOv11, Flask, Telegram, computer vision

KORESPONDENSI

No HP: 085646738422

E-mail: hamimtm@gmail.com

ABSTRACT

Physical bullying in school environments often occurs spontaneously, making it difficult to monitor continuously using conventional Closed-Circuit Television (CCTV) systems. This study aims to implement SAFEYE-AI, a YOLOv11-based bullying detection system integrated with school CCTV, a Flask-based monitoring dashboard, and Telegram notifications. The research employs the AI Project Cycle framework, encompassing problem scoping, data acquisition, data exploration, modeling, evaluation, and deployment. The YOLOv11n model was trained using 3,771 images from school CCTV footage across 10 classes, comprising nine types of physical bullying actions and a student class. Training was conducted over 100 epochs using an NVIDIA RTX 4060 GPU. Evaluation results yielded a precision of 0.903, recall of 0.953, mAP@0.5 of 0.954, and mAP@0.5:0.95 of 0.686. Confidence threshold analysis identified an optimal value of 0.25, resulting in an F1-score of 0.929. Implementation on the CCTV infrastructure demonstrated the system's capability to process up to nine camera streams simultaneously. The system successfully detected kicking and fighting incidents within the school environment and transmitted incident information via Telegram. These results indicate that the integration of YOLOv11, a Flask dashboard, CCTV, and Telegram creates a bullying monitoring system that supports real-time detection and alerting in school environments.

1. PENDAHULUAN

Perundungan atau *bullying* merupakan salah satu permasalahan serius dalam lingkungan pendidikan karena dapat mengganggu rasa aman, kesejahteraan, dan perkembangan peserta didik. Perundungan dapat berbentuk verbal, psikologis, maupun fisik, dengan karakteristik berupa tindakan agresif yang dilakukan secara berulang dan melibatkan ketidakseimbangan kekuatan antara pelaku dan korban [1]. UNICEF melaporkan bahwa lebih dari sepertiga siswa usia 13–15 tahun di dunia mengalami perundungan, sementara sekitar 150 juta remaja pada kelompok usia tersebut pernah mengalami kekerasan antarteman di dalam dan sekitar lingkungan sekolah [2]. Di Indonesia, hasil Asesmen Nasional tahun 2021 menunjukkan bahwa 24,4% peserta didik berpotensi mengalami insiden perundungan di satuan pendidikan [3]. Kondisi tersebut menunjukkan bahwa diperlukan mekanisme pemantauan yang mampu membantu sekolah mengidentifikasi indikasi perundungan secara lebih cepat.

Salah satu infrastruktur yang banyak digunakan untuk mendukung keamanan sekolah adalah *Closed-Circuit Television*

(CCTV). CCTV memungkinkan aktivitas pada berbagai area sekolah direkam dan dipantau secara terus-menerus. Namun, sistem CCTV konvensional pada dasarnya masih bersifat pasif karena rekaman hanya memberikan informasi ketika terdapat petugas yang melakukan pemantauan secara langsung atau meninjau kembali video yang tersimpan [4]. Peningkatan jumlah kamera semakin memperbesar beban pemantauan karena operator harus mengamati beberapa aliran video secara bersamaan. Kondisi tersebut menyebabkan kejadian yang berlangsung secara spontan dan dalam waktu singkat berpotensi terlewatkan. Oleh karena itu, CCTV perlu didukung teknologi yang mampu menganalisis aliran video dan memberikan informasi mengenai kejadian yang memiliki indikasi risiko secara otomatis.

Perkembangan *Artificial Intelligence* (AI), khususnya *computer vision* dan *deep learning*, memberikan peluang untuk mengembangkan sistem pengawasan yang lebih aktif. Model *object detection* dapat digunakan untuk mengenali pola visual tertentu pada citra maupun video sehingga komputer tidak hanya merekam kejadian, tetapi juga mengidentifikasi objek atau aktivitas yang terdeteksi. Salah satu pendekatan yang banyak

digunakan untuk kebutuhan deteksi secara *real-time* adalah *You Only Look Once* (YOLO) karena mampu melakukan deteksi dalam satu tahap sehingga mendukung pemrosesan video secara cepat [5]. YOLOv11 merupakan pengembangan terbaru dalam keluarga YOLO yang dirancang dengan peningkatan pada ekstraksi fitur, efisiensi arsitektur, serta keseimbangan antara kecepatan dan akurasi deteksi [6]. Karakteristik tersebut menjadikan YOLOv11 berpotensi diterapkan pada sistem pengawasan berbasis CCTV yang membutuhkan proses deteksi secara *real-time*.

Meskipun demikian, penerapan *object detection* untuk mendeteksi tindakan *bullying* fisik memiliki tantangan tersendiri. Tindakan seperti memukul, menendang, menampar, mendorong, atau berkelahi tidak hanya berkaitan dengan keberadaan objek manusia, tetapi juga melibatkan interaksi dan gerakan antarmanusia. Rekaman CCTV sekolah juga memiliki variasi sudut kamera, jarak objek, pencahayaan, kepadatan siswa, occlusion, dan kompleksitas latar belakang. Oleh karena itu, model perlu dilatih menggunakan data yang memiliki karakteristik serupa dengan lingkungan tempat sistem akan diterapkan. Penelitian terdahulu mengenai deteksi kekerasan berbasis *deep learning* menunjukkan bahwa kemampuan sistem sangat dipengaruhi oleh representasi data dan kondisi lingkungan pengamatan [7], [8]. Dengan demikian, pengembangan sistem deteksi *bullying* tidak cukup hanya mengevaluasi performa model pada *dataset*, tetapi juga perlu menguji kemampuan model ketika diintegrasikan dengan infrastruktur CCTV pada lingkungan nyata.

Selain kemampuan deteksi, hasil inferensi AI perlu diubah menjadi informasi yang dapat digunakan oleh pihak sekolah. Sistem yang hanya menghasilkan *bounding box* pada video belum sepenuhnya mendukung proses pemantauan dan respons terhadap kejadian. Hasil deteksi perlu disimpan sebagai informasi kejadian, dilengkapi waktu, lokasi, jenis tindakan, nilai *confidence*, dan bukti visual. Selanjutnya, informasi tersebut perlu dapat disampaikan kepada pihak yang bertanggung jawab melalui mekanisme peringatan otomatis. Integrasi antara deteksi otomatis dan mekanisme peringatan menjadi penting karena nilai praktis sistem pengawasan cerdas tidak hanya ditentukan oleh kemampuan mengenali kejadian, tetapi juga oleh kemampuan menyampaikan hasil deteksi kepada pihak yang dapat melakukan verifikasi dan tindak lanjut [9]. Pendekatan tersebut juga sejalan dengan kebutuhan pengembangan sistem *intelligent video surveillance* yang mampu menghubungkan proses deteksi dengan respons terhadap kejadian.

Berdasarkan kebutuhan tersebut, penelitian ini mengimplementasikan SAFEYE-AI, yaitu sistem deteksi *bullying* fisik berbasis YOLOv11 yang terintegrasi dengan infrastruktur CCTV sekolah, *dashboard monitoring* berbasis Flask, dan mekanisme notifikasi Telegram. Model dilatih menggunakan *dataset* yang diperoleh dari rekaman CCTV sekolah untuk mengenali sembilan kategori tindakan *bullying* fisik, yaitu memukul, menendang, menampar, mencakar, mendorong, menjegal, berkelahi, melempar, dan menyikut, serta kelas siswa sebagai representasi aktivitas normal. Sistem tidak hanya melakukan inferensi terhadap aliran video, tetapi juga menyediakan *dashboard* untuk *monitoring*, *Alert Center* untuk pencatatan kejadian, pengelolaan kamera dan model, serta

mekanisme pengiriman peringatan melalui Telegram. Integrasi tersebut membentuk alur operasional dari video CCTV, inferensi YOLOv11, identifikasi kejadian, dokumentasi visual, hingga penyampaian notifikasi.

Berbeda dengan pendekatan yang berhenti pada evaluasi model menggunakan *dataset* atau rekaman video, penelitian ini diarahkan pada implementasi model pada infrastruktur CCTV sekolah secara langsung. Penelitian juga menganalisis *confidence threshold* untuk memperoleh konfigurasi yang sesuai dengan kebutuhan sistem *monitoring*, kemudian menguji kemampuan sistem dalam memproses beberapa aliran CCTV secara simultan serta menyampaikan hasil deteksi melalui Telegram. Dengan demikian, kontribusi penelitian tidak hanya terletak pada penerapan YOLOv11 untuk mendeteksi tindakan *bullying*, tetapi juga pada integrasi model dengan sistem *monitoring* dan mekanisme peringatan pada lingkungan sekolah nyata.

Penelitian ini bertujuan untuk mengimplementasikan sistem deteksi *bullying* berbasis YOLOv11 pada CCTV sekolah, menganalisis performa model berdasarkan *precision*, *recall*, F1-score, mAP@0.5, dan mAP@0.5:0.95, menentukan *confidence threshold* yang sesuai untuk implementasi, serta mengevaluasi penerapan sistem melalui *dashboard Flask* dan notifikasi Telegram pada kondisi operasional CCTV sekolah. Hasil penelitian diharapkan dapat memberikan kontribusi terhadap pengembangan *intelligent video surveillance* berbasis AI yang tidak hanya mampu melakukan deteksi secara otomatis, tetapi juga mampu menyajikan informasi kejadian dan memberikan peringatan secara *real-time* untuk mendukung proses pemantauan keamanan di lingkungan sekolah.

2. TINJAUAN PUSTAKA

2.1 Computer vision dan Deep learning untuk Deteksi Objek

Computer vision merupakan bidang yang memungkinkan komputer memperoleh dan menganalisis informasi visual dari citra maupun video. Salah satu pendekatan yang banyak digunakan dalam *computer vision* modern adalah *deep learning*. *Deep learning* memanfaatkan jaringan saraf dengan banyak lapisan untuk mempelajari representasi dan pola kompleks dari data secara otomatis [10].

Dalam analisis citra dan video, *Convolutional Neural Network* (CNN) menjadi salah satu arsitektur yang banyak digunakan karena mampu mengekstraksi karakteristik visual dari gambar. Kemampuan tersebut memungkinkan model *deep learning* diterapkan untuk berbagai tugas, termasuk image classification, object detection, dan analisis aktivitas manusia.

Object detection memiliki perbedaan dengan klasifikasi citra karena sistem tidak hanya menentukan kelas objek, tetapi juga menentukan lokasi objek melalui *bounding box*. Pendekatan ini relevan untuk analisis video CCTV karena sistem perlu mengetahui objek atau individu yang terlibat dalam suatu kejadian sekaligus mengidentifikasi kelas tindakan yang terdeteksi. Dalam penelitian ini, *object detection* digunakan untuk mengenali tindakan *bullying* fisik pada *frame* video CCTV.

2.2 YOLOv11

You Only Look Once (YOLO) merupakan keluarga model *object detection* yang dirancang untuk melakukan deteksi objek dalam satu tahap pemrosesan. Pendekatan single-stage detection memungkinkan proses identifikasi objek dilakukan secara efisien sehingga YOLO banyak digunakan pada aplikasi yang membutuhkan pemrosesan video secara *real-time* [5].

YOLOv11 merupakan salah satu pengembangan dalam keluarga YOLO yang diperkenalkan oleh Ultralytics dan dirancang untuk memberikan keseimbangan antara akurasi deteksi dan efisiensi komputasi. Dalam penelitian ini, YOLOv11 digunakan sebagai model utama untuk mendeteksi sembilan kategori tindakan *bullying* fisik, yaitu memukul, menendang, menampar, mencakar, mendorong, menjegal, berkelahi, melempar, dan menyikut, serta satu kelas siswa sebagai representasi aktivitas normal.

Penerapan YOLO pada deteksi kekerasan juga telah dilakukan pada penelitian terdahulu. Xu et al. melalui pendekatan *Violence-YOLO* menunjukkan bahwa performa deteksi kekerasan pada lingkungan yang kompleks tidak hanya dipengaruhi oleh arsitektur model, tetapi juga oleh kemampuan *dataset* dalam merepresentasikan variasi objek, sudut kamera, dan kondisi lingkungan [8]. Hal tersebut menunjukkan pentingnya penggunaan *dataset* yang sesuai dengan karakteristik lingkungan penerapan.

Selain itu, penelitian Putri et al. menerapkan pendekatan *deep learning* untuk mendeteksi tindakan agresif seperti menendang, mendorong, dan memukul dalam konteks pemantauan *bullying* di sekolah [11]. Penelitian tersebut menunjukkan potensi *computer vision* untuk membantu proses identifikasi tindakan kekerasan, tetapi pendekatan deteksi pada penelitian ini diarahkan lebih lanjut pada integrasi dengan sistem CCTV sekolah dan mekanisme peringatan otomatis.

2.3 Integrasi Deteksi AI dengan Sistem Monitoring dan Notifikasi

Model *object detection* yang berdiri sendiri hanya menghasilkan keluaran berupa prediksi objek dan nilai *confidence*. Agar dapat digunakan dalam lingkungan operasional, keluaran tersebut perlu diintegrasikan dengan sistem yang mampu menampilkan, menyimpan, dan meneruskan informasi hasil deteksi.

Integrasi AI dengan sistem *monitoring* memungkinkan hasil inferensi ditampilkan secara *real-time* bersama informasi sumber kamera dan kejadian yang teridentifikasi. Pada penelitian ini, model YOLOv11 diintegrasikan dengan aplikasi berbasis *Flask* untuk membangun *dashboard monitoring* SAFEYE-AI. Sistem tersebut menghubungkan aliran video CCTV dengan proses inferensi, penyimpanan informasi kejadian, serta penyajian hasil deteksi kepada pengguna.

Selain *monitoring*, mekanisme notifikasi diperlukan agar pengguna tidak harus mengamati *dashboard* secara terus-menerus. Penelitian Ye et al. menunjukkan bahwa sistem deteksi kekerasan berbasis kamera dapat dilengkapi mekanisme

penyampaian alarm kepada pihak terkait melalui pesan atau media sosial [9]. Pendekatan tersebut memperlihatkan bahwa sistem deteksi akan memiliki nilai operasional yang lebih besar ketika hasil analisis video dapat segera diteruskan kepada pihak yang bertanggung jawab.

Dalam penelitian ini, Telegram digunakan sebagai media penyampaian notifikasi ketika sistem mendeteksi tindakan yang telah ditentukan. Informasi yang dikirim dapat mencakup jenis tindakan, lokasi kamera, waktu kejadian, *confidence*, dan bukti visual. Dengan demikian, integrasi YOLOv11, CCTV, *dashboard* Flask, dan Telegram membentuk alur sistem yang menghubungkan proses deteksi otomatis dengan penyampaian informasi untuk mendukung verifikasi oleh pihak sekolah.

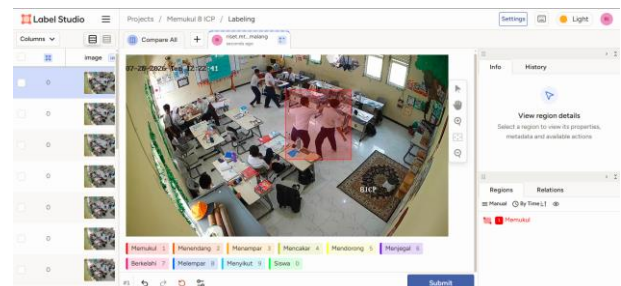
3. METODOLOGI

3.1 Jenis dan Tahapan Penelitian

Penelitian ini merupakan penelitian rekayasa teknologi (*engineering research*) yang bertujuan mengembangkan dan mengimplementasikan sistem deteksi *bullying* fisik berbasis *computer vision* menggunakan YOLOv11 pada lingkungan CCTV sekolah. Penelitian menggunakan kerangka *AI Project Cycle* yang terdiri atas *problem scoping*, *data acquisition*, *data exploration*, *modelling*, *evaluation*, dan *deployment*. Penelitian dilaksanakan pada April–Agustus 2026 di MTsN 6 Malang, meliputi pengumpulan data, pelatihan model, evaluasi, pengembangan aplikasi, dan implementasi pada infrastruktur CCTV sekolah.

3.2 Pengumpulan dan Pengolahan Dataset

Data diperoleh dari rekaman CCTV sekolah dan video simulasi tindakan kekerasan fisik. Video diekstraksi menjadi *frame* menggunakan Python dan OpenCV, kemudian diseleksi berdasarkan kualitas citra, kesesuaian aktivitas, sudut pengambilan gambar, dan keberagaman kondisi visual. Setiap *frame* dianotasi menggunakan *bounding box* melalui Label Studio dan diperiksa kembali untuk memastikan kesesuaian antara anotasi dan kelas. *Dataset* yang digunakan berjumlah 3.771 citra, terdiri atas 3.016 citra latih dan 755 citra validasi, dengan 10 kelas, yaitu memukul, menendang, menampar, mencakar, mendorong, menjegal, berkelahi, melempar, menyikut, dan siswa.



Gambar 1. Proses menganotasi objek pada setiap *frame* menggunakan *bounding box* melalui Label Studio

3.3 Pelatihan dan Evaluasi Model

Model yang digunakan adalah YOLOv11n. Pelatihan dilakukan menggunakan GPU NVIDIA RTX 4060 selama 100 epoch. Kinerja model dievaluasi menggunakan *precision*, *recall*, F1-score, mAP@0.5, dan mAP@0.5:0.95. Selain itu, dilakukan analisis terhadap beberapa nilai *confidence threshold* untuk menentukan konfigurasi yang sesuai dengan kebutuhan deteksi pada sistem CCTV.

3.4 Implementasi Sistem

Model YOLOv11 yang telah dilatih diintegrasikan ke dalam aplikasi SAFEYE-AI menggunakan Python dan Flask sebagai backend. Sistem menerima aliran video CCTV melalui protokol RTSP, melakukan inferensi secara *real-time*, kemudian menampilkan hasil deteksi pada *dashboard monitoring*. Sistem juga menyimpan informasi kejadian dan mengintegrasikan Telegram Bot untuk mengirimkan notifikasi ketika terdeteksi tindakan yang memenuhi konfigurasi sistem.



Gambar 2. Proses mengintegrasikan SAFEYE-AI dengan Infrastruktur CCTV sekolah

3.5 Pengujian Implementasi

Pengujian implementasi dilakukan dengan menerapkan model pada aliran CCTV sekolah dan mengamati kemampuan sistem dalam melakukan deteksi secara *real-time*. Evaluasi implementasi meliputi keberhasilan deteksi pada kondisi CCTV, penyajian hasil pada *dashboard*, penyimpanan bukti visual, serta keberhasilan pengiriman notifikasi melalui Telegram.

4. HASIL DAN PEMBAHASAN

4.1 Hasil Training Model YOLOv11 Menggunakan Dataset CCTV Sekolah

Model YOLOv11n dilatih menggunakan *dataset* yang diperoleh dari rekaman CCTV sekolah. *Dataset* terdiri atas 3.771 citra, dengan 3.016 citra untuk pelatihan dan 755 citra untuk validasi. *Dataset* mencakup 10 kelas, yaitu berkelahi, melempar, memukul, menampar, mencakar, mendorong, menendang, menjegal, menyikut, dan siswa. Proses *training* dilakukan selama 100 epoch menggunakan GPU NVIDIA RTX 4060 8 GB dengan waktu sekitar 41 menit.

Hasil evaluasi menunjukkan bahwa model memperoleh *precision* sebesar 0,903, *recall* 0,953, mAP@0.5 0,954, dan mAP@0.5:0.95 sebesar 0,686. Nilai tersebut menunjukkan kemampuan deteksi yang baik, terutama pada mAP@0.5 dan *recall*. *Recall* sebesar 0,953 menunjukkan bahwa sebagian besar objek atau tindakan yang menjadi target dapat ditemukan oleh model. Nilai *recall* yang lebih tinggi daripada *precision* juga menunjukkan bahwa model cenderung mempertahankan sensitivitas deteksi, meskipun masih memungkinkan munculnya sejumlah prediksi positif yang tidak tepat.

Tabel 1. Performa model YOLOv11 secara keseluruhan

Metrik	Nilai
<i>Precision</i>	0,903
<i>Recall</i>	0,953
mAP@0.5	0,954
mAP@0.5:0.95	0,686

Jika ditinjau berdasarkan kelas, sebagian besar tindakan *bullying* memiliki performa tinggi. Kelas menyikut memperoleh mAP@0.5 tertinggi sebesar 0,995, diikuti mencakar 0,988, berkelahi dan menjegal masing-masing 0,985, menampar 0,981, melempar 0,981, dan menendang 0,973. Sebaliknya, kelas mendorong memiliki mAP@0.5 sebesar 0,867, sedangkan kelas siswa memperoleh 0,825.

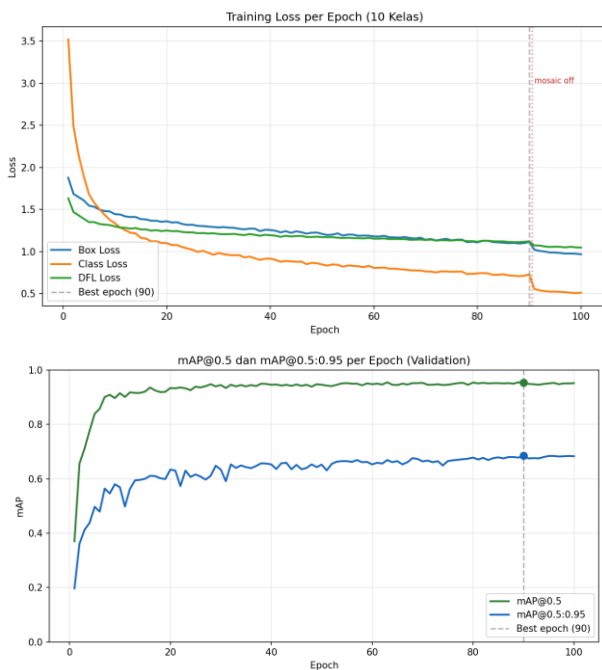
Tabel 2. Performa model berdasarkan kelas

Kelas	<i>Precision</i>	<i>Recall</i>	mAP@0.5	mAP@0.5:0.95
Berkelahi	0,974	1,000	0,985	0,683
Melempar	0,864	0,955	0,981	0,727
Memukul	0,865	0,927	0,964	0,754
Menampar	0,962	1,000	0,981	0,798
Mencakar	0,938	0,986	0,988	0,847
Mendorong	0,754	0,992	0,867	0,553
Menendang	0,914	0,936	0,973	0,755
Menjegal	0,987	0,982	0,985	0,628
Menyikut	0,989	1,000	0,995	0,627
Siswa	0,785	0,752	0,825	0,488

Perbedaan performa antarkelas menunjukkan bahwa kompleksitas visual setiap tindakan tidak sama. Kelas dengan pola gerakan yang lebih jelas seperti menyikut, mencakar, berkelahi, dan menjegal dapat dikenali dengan lebih baik. Sebaliknya, kelas mendorong dan siswa memiliki performa relatif lebih rendah. Tindakan mendorong dapat memiliki pola visual yang lebih sulit dibedakan karena perubahan posisi tubuh dan interaksi antarsiswa dapat menyerupai aktivitas normal. Sementara itu, kelas siswa merupakan kelas dengan jumlah instance paling besar, tetapi memiliki variasi pose dan aktivitas yang lebih luas sehingga karakteristik visualnya lebih beragam.

4.2 Analisis Training Loss

Kurva proses *training* memperlihatkan bahwa nilai *Box Loss*, *Class Loss*, dan *Distribution Focal Loss (DFL Loss)* terus menurun secara konsisten selama proses pelatihan. Selain itu, nilai mAP@0.5 meningkat secara cepat pada awal *training* dan kemudian stabil pada kisaran 0,94–0,95, sedangkan mAP@0.5:0.95 terus meningkat hingga mencapai nilai optimum pada *epoch* ke-90. Kondisi tersebut menunjukkan bahwa model telah mengalami proses konvergensi dengan baik tanpa indikasi *overfitting*.



Gambar 3. Grafik *Box Loss*, *Class Loss*, dan *DFL Loss* per epoch (atas), Grafik *mAP@0.5* dan *mAP@0.5:0.95* per epoch (bawah)

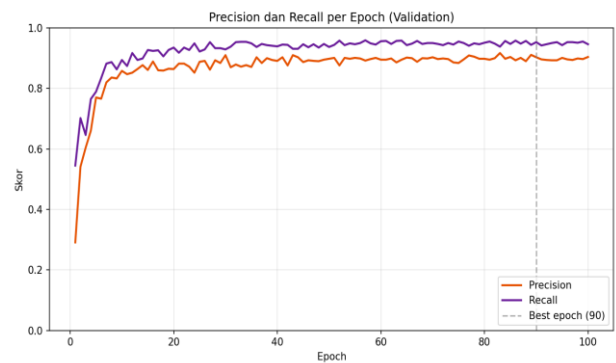
Hasil penelitian menunjukkan bahwa proses *training* menggunakan *dataset* yang diperoleh secara langsung dari CCTV sekolah mampu meningkatkan performa model YOLOv11 dalam mendeteksi tindakan *bullying* pada lingkungan nyata. Berbeda dengan proses pelatihan sebelumnya yang menggunakan *dataset* terbatas, *dataset* baru berasal dari kondisi operasional sebenarnya sehingga memuat variasi pencahayaan, sudut pengambilan gambar, kepadatan siswa, kualitas video, serta berbagai karakteristik visual yang umum ditemukan pada CCTV sekolah. Kondisi tersebut memungkinkan model mempelajari distribusi data yang lebih representatif sehingga kemampuan generalisasi model meningkat. Hal ini terlihat dari peningkatan nilai *mAP@0.5* dari 90,3% menjadi 95,4%, sedangkan *mAP@0.5:0.95* meningkat dari 61,8% menjadi 68,6%, meskipun jumlah kelas yang dideteksi bertambah dari dua kelas efektif menjadi sepuluh kelas tindakan *bullying*.

Peningkatan performa tersebut sejalan dengan karakteristik model YOLO yang sangat dipengaruhi oleh kualitas dan keragaman *dataset*. Penelitian *Violence-YOLO* menjelaskan bahwa keberhasilan deteksi tindakan kekerasan pada lingkungan kompleks tidak hanya ditentukan oleh arsitektur model, tetapi juga oleh kemampuan *dataset* dalam merepresentasikan kondisi nyata sehingga model mampu melakukan generalisasi pada berbagai variasi objek, sudut kamera, dan lingkungan pengamatan [8]. Temuan tersebut mendukung hasil penelitian ini, dimana penggunaan *dataset* CCTV asli sekolah terbukti memberikan peningkatan performa dibandingkan *dataset* yang lebih terbatas.

4.3 Precision & Recall

Precision stabil di $\sim 0,90$ dan *recall* di $\sim 0,95$ sejak pertengahan training. *Recall* yang lebih tinggi dari *precision* berarti model cenderung jarang melewatkan kejadian nyata, dengan konsekuensi sesekali memunculkan deteksi berlebih —

karakteristik yang justru cocok untuk sistem pemantauan keamanan.



Gambar 4. *Precision* dan *Recall* per epoch

Nilai *Recall* sebesar 95,3% merupakan salah satu hasil yang paling penting pada penelitian ini. Pada sistem keamanan sekolah, false negative atau kegagalan mendeteksi kejadian *bullying* memiliki konsekuensi yang lebih serius dibandingkan *false positive*. Alarm yang salah masih dapat diverifikasi oleh guru melalui tampilan CCTV, sedangkan tindakan *bullying* yang tidak terdeteksi dapat menyebabkan keterlambatan intervensi terhadap korban. Oleh karena itu, karakteristik model SAFEYE-AI yang memiliki *recall* tinggi sangat sesuai untuk diimplementasikan sebagai sistem peringatan dini (*early warning system*).

Hasil tersebut juga lebih baik dibandingkan penelitian Putri et al. (2024) yang mengembangkan sistem pemantauan *bullying* menggunakan kombinasi CNN dan LSTM, dimana model yang dikembangkan memperoleh akurasi rata-rata sekitar 92%. Walaupun metrik evaluasi yang digunakan tidak sepenuhnya sama, hasil penelitian SAFEYE-AI menunjukkan bahwa pendekatan berbasis deteksi objek menggunakan YOLOv11 mampu menghasilkan performa yang lebih tinggi sekaligus memenuhi kebutuhan deteksi secara *real-time*. Selain itu, YOLO hanya memerlukan satu proses inferensi (single-stage detector) sehingga lebih efisien dibandingkan pendekatan dua tahap maupun kombinasi CNN-LSTM yang umumnya membutuhkan komputasi temporal lebih besar [12].

Evaluasi pada masing-masing kelas menunjukkan bahwa sebagian besar tindakan *bullying* memperoleh nilai *mAP@0.5* di atas 0,96, seperti berkelahi, menampar, mencakar, menendang, menjegal, dan menyikut. Hal ini menunjukkan bahwa model mampu mengenali karakteristik gerakan yang relatif berbeda antar kelas. Sebaliknya, kelas mendorong dan siswa masih memiliki performa lebih rendah dibandingkan kelas lainnya. Kondisi tersebut diduga disebabkan oleh tingginya kemiripan visual antara aktivitas normal siswa dengan aktivitas mendorong, serta jumlah data kelas mendorong yang relatif lebih sedikit dibandingkan kelas lainnya. Selain itu, kelas "siswa" memiliki variasi pose, orientasi tubuh, serta kondisi oklusi yang jauh lebih kompleks sehingga menjadi kelas yang paling sulit dipelajari model.

Temuan tersebut sejalan dengan penelitian Cai et al., (2024) yang melaporkan bahwa deteksi perilaku siswa di dalam kelas menjadi lebih sulit ketika jumlah siswa padat, banyak terjadi tumpang

tindih (*occlusion*), serta terdapat kemiripan postur antar aktivitas. Penelitian tersebut bahkan menambahkan mekanisme *coordinate attention* dan *dynamic convolution* untuk meningkatkan kemampuan model dalam membedakan perilaku siswa yang memiliki karakteristik visual serupa. Dengan demikian, performa yang sedikit lebih rendah pada kelas "siswa" dan "mendorong" pada penelitian ini masih merupakan fenomena yang wajar pada sistem deteksi perilaku berbasis CCTV.

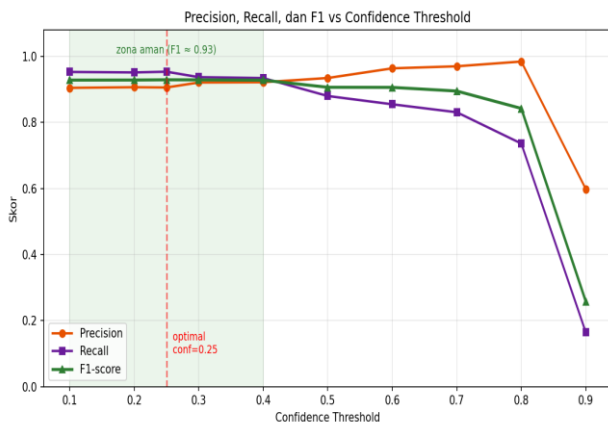
4.4 Analisis Confidence Threshold

Pengujian selanjutnya dilakukan untuk menentukan nilai *confidence threshold* yang paling sesuai pada implementasi SAFEYE-AI. Pengujian dilakukan pada rentang *threshold* 0,10 hingga 0,90.

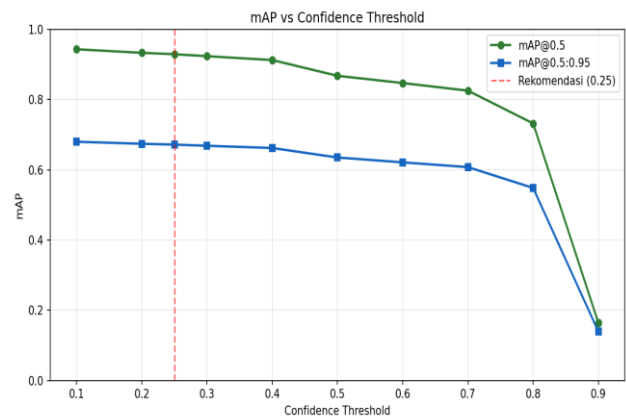
Table 1 Analisis Confidence Threshold

Confidence	Precision	Recall	F1-score	mAP@0.5	mAP@0.5:0.95
0,10	0,9046	0,9528	0,9281	0,9434	0,6799
0,20	0,9065	0,9513	0,9283	0,9333	0,6742
0,25	0,9057	0,9535	0,9290	0,9293	0,6718
0,30	0,9208	0,9370	0,9288	0,9237	0,6685
0,40	0,9217	0,9342	0,9279	0,9127	0,6621
0,50	0,9344	0,8799	0,9063	0,8678	0,6351
0,60	0,9637	0,8548	0,9060	0,8472	0,6212
0,70	0,9699	0,8307	0,8949	0,8255	0,6077
0,80	0,9843	0,7360	0,8423	0,7321	0,5483
0,90	0,5983	0,1643	0,2578	0,1637	0,1390

Hasil pengujian menunjukkan bahwa nilai *threshold* 0,25 memberikan performa terbaik dengan *Precision* sebesar 0,9057, *Recall* sebesar 0,9535, dan *F1-score* sebesar 0,9290. Nilai tersebut merupakan keseimbangan terbaik antara kemampuan mendeteksi seluruh kejadian *bullying* dan meminimalkan kesalahan deteksi.



Gambar 5. Grafik Precision, Recall, dan F1 vs Confidence Threshold



Gambar 6. Grafik mAP vs Confidence Threshold

Pada *threshold* yang lebih tinggi ($\geq 0,60$), nilai *precision* memang meningkat hingga mencapai 0,96–0,98, namun *recall* mengalami penurunan yang cukup signifikan sehingga semakin banyak kejadian *bullying* yang tidak berhasil terdeteksi. Sebaliknya, *threshold* yang terlalu rendah menghasilkan *recall* tinggi tetapi meningkatkan kemungkinan munculnya false positive. Berdasarkan hasil tersebut, *threshold* 0,25 dipilih sebagai nilai optimal untuk implementasi SAFEYE-AI pada sistem *monitoring* CCTV sekolah.

Analisis *confidence threshold* menunjukkan bahwa nilai 0,25 memberikan keseimbangan terbaik antara *precision* dan *recall* dengan *F1-score* sebesar 92,9%. Menariknya, rentang *threshold* 0,10–0,40 menghasilkan nilai *F1* yang relatif stabil sehingga menunjukkan bahwa model memiliki tingkat kepercayaan prediksi yang konsisten dan tidak terlalu sensitif terhadap perubahan *threshold*. Karakteristik seperti ini menunjukkan bahwa distribusi probabilitas keluaran model cukup baik dan stabil sehingga lebih mudah diimplementasikan pada sistem nyata.

Pemilihan *threshold* 0,25 juga sesuai dengan tujuan implementasi SAFEYE-AI sebagai sistem keamanan sekolah. Pada aplikasi keselamatan (*safety-critical system*), kehilangan satu kejadian *bullying* jauh lebih berisiko dibandingkan munculnya beberapa alarm palsu. Oleh karena itu, konfigurasi *threshold* dipilih untuk mempertahankan *recall* tinggi tanpa menurunkan *precision* secara signifikan. Sebaliknya, penggunaan *threshold* yang terlalu tinggi ($\geq 0,90$) menyebabkan *recall* turun hingga sekitar 16%, sehingga sebagian besar kejadian *bullying* tidak berhasil dideteksi.

Apabila dibandingkan dengan penelitian terbaru mengenai YOLOv8 dan YOLOv11 untuk deteksi perilaku siswa di lingkungan kelas, hasil penelitian ini menunjukkan keunggulan pada aspek kompleksitas permasalahan yang diselesaikan. Penelitian tersebut hanya mengidentifikasi enam aktivitas pembelajaran seperti membaca, menulis, berdiri, atau mengangkat tangan menggunakan *dataset* yang relatif kecil, sedangkan SAFEYE-AI mendeteksi sepuluh kelas yang sebagian besar berupa interaksi fisik antar siswa dengan tingkat kesulitan yang lebih tinggi karena melibatkan kontak tubuh, gerakan cepat, dan tumpang tindih antar objek [14]. Dengan demikian, performa yang diperoleh SAFEYE-AI menunjukkan bahwa model tetap

mampu mempertahankan akurasi tinggi meskipun menghadapi skenario yang lebih kompleks.

Kontribusi utama penelitian ini dibandingkan penelitian terdahulu tidak hanya terletak pada penggunaan model YOLOv11, tetapi juga pada pengembangan *dataset bullying* yang berasal dari CCTV sekolah secara langsung, implementasi deteksi *real-time* yang terintegrasi dengan sistem *monitoring* SAFEYE-AI, serta optimasi *confidence threshold* berdasarkan kebutuhan operasional sekolah. Sebagian besar penelitian sebelumnya masih berhenti pada evaluasi model di lingkungan laboratorium atau *dataset* publik, sedangkan penelitian ini telah mengarahkan model menuju implementasi nyata pada sistem pengawasan sekolah. Dengan demikian, SAFEYE-AI memiliki potensi yang lebih besar untuk diterapkan sebagai sistem pendukung pengambilan keputusan dalam pencegahan dan penanganan *bullying* secara dini di lingkungan pendidikan.

4.5 Hasil Implementasi SAFEYE-AI Pada Infrastruktur CCTV Sekolah

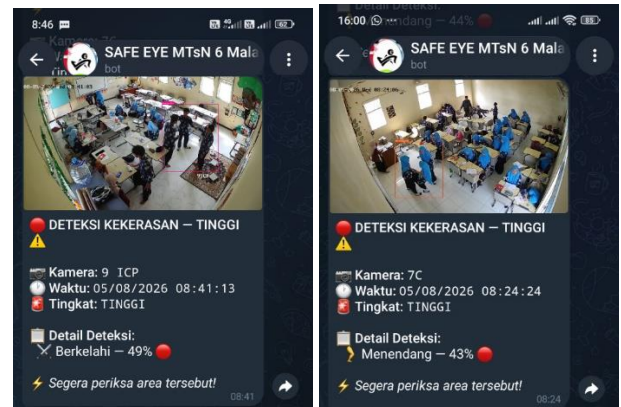
Sebagai tahap implementasi, SAFEYE-AI berhasil diintegrasikan dengan infrastruktur CCTV MTsN 6 Malang. Sekolah memiliki 40 unit kamera CCTV yang tersebar pada ruang kelas, koridor, halaman, dan area publik lainnya. Kamera terhubung melalui *Network Video Recorder* (NVR) dan dapat diakses oleh SAFEYE-AI menggunakan protokol RTSP untuk proses inferensi secara *real-time*. Implementasi ini menunjukkan bahwa model yang telah dikembangkan dapat dihubungkan dengan infrastruktur CCTV yang telah tersedia tanpa memerlukan penggantian seluruh perangkat kamera.



Gambar 7. Implementasi SAFEYE-AI pada infrastruktur CCTV sekolah SAFEYE-AI mampu mendeteksi tindakan *bullying* berupa kasus penendangan pada CCTV

Meskipun tersedia 40 kamera, kemampuan pemrosesan secara simultan dipengaruhi oleh kapasitas perangkat komputasi. Berdasarkan pengujian, komputer dengan Intel Core i7 Generasi ke-8, RAM 32 GB, dan GPU NVIDIA GeForce RTX 4060 mampu memproses maksimal sembilan aliran CCTV secara bersamaan dengan kondisi operasional yang stabil. Temuan ini menunjukkan bahwa penerapan sistem *intelligent surveillance* tidak hanya bergantung pada performa model, tetapi juga pada kemampuan perangkat dalam menangani beberapa aliran video secara paralel. Hal tersebut sejalan dengan penelitian terdahulu yang menunjukkan bahwa peningkatan jumlah kamera menyebabkan beban pemrosesan dan pemantauan meningkat sehingga otomatisasi dan efisiensi komputasi menjadi aspek penting dalam sistem pengawasan *real-time* [7].

Hasil implementasi menunjukkan bahwa SAFEYE-AI mampu mendeteksi indikasi *bullying* pada kondisi lingkungan sekolah yang sebenarnya. Sistem mendeteksi tindakan menendang di kelas 7C dengan *confidence* 43% dan tindakan berkelahi di kelas 9 ICP dengan *confidence* 49%. Keberhasilan deteksi pada lingkungan nyata menunjukkan bahwa model tidak hanya dapat bekerja pada data pengujian, tetapi juga dapat menghasilkan inferensi ketika berhadapan dengan variasi posisi siswa, sudut kamera, jarak objek, dan kondisi ruangan yang terdapat pada rekaman CCTV sekolah.



Gambar 8. Notifikasi *real time* deteksi tindakan *bullying* yang dikirim ke telegram guru

Setiap kejadian yang terdeteksi juga menghasilkan notifikasi melalui Telegram kepada administrator dan Guru BK. Informasi yang diteruskan meliputi bukti visual, lokasi kamera, waktu kejadian, jenis tindakan, tingkat kejadian, dan nilai *confidence*. Mekanisme ini memungkinkan pengguna memperoleh informasi tanpa harus mengamati seluruh aliran CCTV secara terus-menerus. Dengan demikian, implementasi SAFEYE-AI mengubah hasil deteksi AI menjadi informasi peringatan yang dapat segera diverifikasi oleh pihak sekolah.

Temuan tersebut sejalan dengan penelitian *A Video-Based DT-SVM School Violence Detecting Algorithm* yang memanfaatkan kamera pengawasan sekolah untuk mendeteksi kekerasan dan mengirimkan alarm kepada guru atau orang tua melalui pesan maupun media sosial [15]. Kesamaan tersebut menunjukkan bahwa integrasi antara deteksi otomatis dan mekanisme peringatan merupakan komponen penting agar hasil analisis video memiliki manfaat operasional. Namun, SAFEYE-AI mengembangkan pendekatan tersebut menggunakan YOLOv11 dan integrasi langsung dengan infrastruktur CCTV sekolah serta Telegram, sehingga proses deteksi dan penyampaian informasi berada dalam satu sistem.

Temuan tersebut juga memiliki kesamaan dengan penelitian Januriansyah & Samopa, (2025) yang mengembangkan sistem *early warning* berbasis YOLOv11 dan Telegram untuk mendeteksi kekerasan dan mengirimkan notifikasi secara instan. Penelitian tersebut melaporkan kemampuan pemrosesan *real-time* dan telah melakukan validasi menggunakan rekaman CCTV, tetapi peneliti juga menekankan perlunya pengujian lebih lanjut pada kondisi dunia nyata yang beragam. Dalam konteks penelitian ini, implementasi langsung pada CCTV MTsN 6 Malang memberikan bukti tambahan bahwa model YOLOv11 dapat dioperasikan pada lingkungan sekolah nyata, meskipun

cakupan pengujian masih terbatas pada satu lingkungan sekolah dan maksimal sembilan kamera secara simultan.

Nilai *confidence* sebesar 43% dan 49% pada kedua kejadian menunjukkan bahwa tingkat keyakinan model pada kondisi nyata lebih moderat dibandingkan performa agregat model pada *dataset* pengujian. Kondisi tersebut dapat dipengaruhi oleh perbedaan sudut kamera, pencahayaan, resolusi video, jarak objek, kepadatan siswa, dan occlusion. Oleh karena itu, hasil implementasi ini tidak hanya menunjukkan keberhasilan sistem, tetapi juga mengidentifikasi kebutuhan pengembangan *dataset* yang lebih beragam dan pengujian pada kondisi CCTV yang lebih luas.

Dengan demikian, implementasi SAFEYE-AI menunjukkan bahwa sistem telah mampu menjalankan alur CCTV → inferensi YOLOv11 → deteksi kejadian → dokumentasi visual → notifikasi Telegram dalam lingkungan sekolah nyata. Hasil ini memperkuat posisi SAFEYE-AI sebagai sistem pendukung pemantauan dan peringatan awal terhadap indikasi *bullying*, dengan Guru BK tetap berperan dalam melakukan verifikasi dan menentukan tindak lanjut terhadap kejadian yang terdeteksi.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini berhasil mengimplementasikan sistem deteksi *bullying* fisik berbasis YOLOv11 pada infrastruktur CCTV sekolah yang terintegrasi dengan *dashboard monitoring* berbasis *Flask* dan notifikasi Telegram. Model YOLOv11n yang dilatih menggunakan 3.771 citra dari *dataset* CCTV sekolah menghasilkan *precision* sebesar 0,903, *recall* 0,953, mAP@0.5 sebesar 0,954, dan mAP@0.5:0.95 sebesar 0,686. *Confidence threshold* 0,25 memberikan keseimbangan terbaik dengan F1-score sebesar 0,929. Implementasi pada lingkungan sekolah menunjukkan bahwa sistem mampu memproses hingga sembilan aliran CCTV secara simultan, mendeteksi tindakan *bullying* pada kondisi nyata, menampilkan hasil deteksi melalui *dashboard*, serta mengirimkan notifikasi dan bukti visual melalui Telegram. Dengan demikian, integrasi YOLOv11, CCTV, *Flask*, dan Telegram dapat mendukung sistem *monitoring* dan peringatan dini terhadap indikasi *bullying* di lingkungan sekolah.

5.2 Saran

Pengembangan selanjutnya disarankan untuk memperluas *dataset* dengan variasi kondisi pencahayaan, sudut kamera, jarak objek, jumlah siswa, dan kondisi lingkungan yang lebih beragam. Selain itu, pengujian dapat dilakukan pada lebih banyak kamera dan sekolah untuk mengetahui kemampuan generalisasi sistem. Optimasi model dan perangkat komputasi juga diperlukan agar jumlah aliran CCTV yang dapat diproses secara simultan dapat ditingkatkan tanpa mengurangi kemampuan deteksi.

UCAPAN TERIMA KASIH

Bila ada, silahkan dituliskan pada bagian ini. Ucapan terima kasih hanya berkaitan dengan sumber dana penelitian dengan menuliskan nomor kontrak.

DAFTAR PUSTAKA

- [1] S. Metha, "Pertanggung jawaban pidana atas tindakan perundungan fisik oleh pelaku anak di bawah umur," *Soll. J. Kaji. Kontemporer Huk. Dan Masy.*, vol. 1, no. 02, 2023.
- [2] UNICEF, "Half of world's teens experience peer violence in and around school – UNICEF." Accessed: Aug. 21, 2025. [Online]. Available: https://www.unicef.org/eca/press-releases/half-worlds-teens-experience-peer-violence-and-around-school-unicef?utm_source=chatgpt.com
- [3] F. N. Fintari and M. Fauziah, "Analisis Dampak Bullying Verbal Terhadap Kemampuan Berbicara Peserta Didik Di Tingkat Sekolah Dasar.," *J. Pendidik. Indones.*, vol. 5, no. 10, 2024.
- [4] A. P. Sidiq, A. Darmawan, O. Supriadi, and O. A. Rozak, "Implementasi Sistem Monitoring CCTV Berbasis IP untuk Peningkatan Keamanan di SMK Khazanah Kebajikan," *J. Keilmuan Dan Keislaman.*, pp. 690–698, 2025.
- [5] Y. Yanto, F. Aziz, and I. Irmawati, "YOLO-V8 peningkatan algoritma untuk deteksi pemakaian masker wajah," *JATI J. Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1437–1444, 2023.
- [6] D. A. Farizki and A. L. Pramana, "ANALISIS KOMPARATIF KINERJA YOLOV5, YOLOV8, DAN YOLO11 PADA DETEKSI OBJEK LALU LINTAS MALAM HARI MENGGUNAKAN CLAHE," *JATI J. Mhs. Tek. Inform.*, vol. 10, no. 4, pp. 6757–6766, 2026.
- [7] K. B. Sahay, B. Balachander, B. Jagadeesh, G. Anand Kumar, R. Kumar, and L. Rama Parvathy, "A real time crime scene intelligent video surveillance systems in violence detection framework using deep learning techniques," *Comput. Electr. Eng.*, vol. 103, p. 108319, Oct. 2022, doi: 10.1016/j.compeleceng.2022.108319.
- [8] W. Xu, D. Zhu, R. Deng, K. Yung, and A. W. H. Ip, "Violence-YOLO: Enhanced GELAN Algorithm for Violence Detection," *Appl. Sci.*, vol. 14, no. 15, p. 6712, Jan. 2024, doi: 10.3390/app14156712.
- [9] L. Ye, L. Wang, H. Ferdinando, T. Seppänen, and E. Alasaarela, "A Video-Based DT-SVM School Violence Detecting Algorithm," *Sensors*, vol. 20, no. 7, p. 2018, Apr. 2020, doi: 10.3390/s20072018.
- [10] H. A. Pratiwi, M. Cahyanti, and M. Lamsani, "Implementasi deep learning flower scanner menggunakan metode convolutional neural network," *J. Inform.*, vol. 2, no. 1, pp. 41–59, 2021.
- [11] S. A. Putri, A. Rifai, and I. Nawawi, "Aplikasi Cerdas Sistem Deteksi Tindak Kekerasan Fisik Untuk Pengawasan Perundungan Dengan Convolutional Neural Network," *JSAI J. Sci. Appl. Inform.*, vol. 7, no. 2, 2024.
- [12] S. A. Putri, A. Rifai, and I. Nawawi, "Sistem Cerdas Deteksi Tindak Kekerasan Untuk Pengawasan Perundungan Dengan Model Deep Learning," *JSAI J. Sci. Appl. Inform.*, vol. 7, no. 2, pp. 332–340, Jun. 2024, doi: 10.36085/jsai.v7i2.6451.
- [13] F. Cai, Y. Du, and X. Qin, "An Improved YOLOv7-CO Classroom Behavior Detection Algorithm Combining Coordinate Attention Mechanism and Omni-Dimensional Dynamic Convolution," in *Artificial Intelligence and Human-Computer Interaction*, Y. Ye and P. Siarry, Eds., SAGE Publications, 2024. doi: 10.3233/FAIA240167.

- [14] M. Sofhia, "Comparative Evaluation of YOLOv8 and YOLOv11 for Student Behavior Detection in Classroom CCTV Environments." Accessed: Aug. 06, 2026. [Online]. Available: <https://www.jurnal.polgan.ac.id/index.php/sinkron/article/view/15648>
- [15] L. Ye, L. Wang, H. Ferdinando, T. Seppänen, and E. Alasaarela, "A Video-Based DT-SVM School Violence Detecting Algorithm," *Sensors*, vol. 20, no. 7, p. 2018, Apr. 2020, doi: 10.3390/s20072018.
- [16] M. S. Januriansyah and F. Samopa, "YOLOv11-Based Early Warning System for Violence and Weapon Detection," in *2025 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, Sakhir, Bahrain: IEEE, Nov. 2025, pp. 1–6. doi: 10.1109/3ict68299.2025.11442141.

BIODATA PENULIS



Yasmin Maulidia Andini Putri

Lahir di Malang pada 4 Februari 2012. Saat ini, ia sedang menempuh pendidikan kelas IX di MTsN 6 Malang pada program Cambridge International Class. Penulis memiliki ketertarikan yang tinggi dalam bidang riset, khususnya Artificial Intelligence (AI) dan pengembangan pakan alternatif ternak.



Rania Ayu Praesawardhani

Lahir di Malang pada 25 Mei 2012. Saat ini, ia sedang menempuh pendidikan di kelas IX MTsN 6 Malang pada program Kelas Olimpiade dan Riset. Penulis memiliki ketertarikan yang tinggi dalam bidang riset, khususnya non-Newtonian fluid, pakan alternatif untuk ternak, serta Artificial Intelligence (AI).



Hamim Thohari Mahfudhillah, S.Pd., M.Pd. lahir di Malang pada 7 Oktober 1991. Penulis merupakan guru mata pelajaran IPA dan Riset di MTsN 6 Malang. Beliau menyelesaikan pendidikan S1 dan S2 Pendidikan Biologi di Universitas Negeri Malang. Selain aktif di bidang riset, penulis juga telah menerbitkan berbagai buku ilmiah. Email: hamimtm@gmail.com



Annisa Ahmada Atusta, S.Pd., M.Pd. Merupakan guru MTsN 6 Malang pada mata pelajaran IPA dan Riset. Penulis lahir di Tulungagung, 4 Februari 1997. Lulusan S1 dan S2 pendidikan Biologi Universitas Negeri Malang. Email: annisaahmadaa@gmail.com