JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin

Deteksi Kecurangan Ujian Berbasis Foto: Perbandingan Kinerja YOLOv5 dan YOLOv8

Whika Cahyo Saputro[1*], Chastine Fatichah[2], Titik Wihayanti[3]

[1,2]Teknik Informatika, Institut Teknologi Surabaya, Surabaya, Indonesia
[3]Teknik Informatika, Universitas Ma'soem, Bandung, Indonesia
Email: [1,*]6025202009@student.its.ac.id, [2]chastine@if.its.ac.id, [3,*]magister.titikw@gmail.com
Email Penulis Korespondensi: [1]6025202009@student.its.ac.id

Abstrak − Pandemi Covid-19 dan peralihan pembelajaran ke sistem daring menciptakan tantangan baru dalam pengawasan ujian. Sebuah studi sistematis menemukan bahwa tingkat kecurangan dalam ujian daring melonjak hingga 55% selama pandemi sehingga hal ini mendorong dibutuhkannya sistem pengawasan ujian yang adaptif dan cerdas melalui pemanfaatan teknologi deteksi objek menggunakan YOLO. Sejumlah penelitian terdahulu telah mengimplementasikan YOLOv5 dan YOLOv8 dalam sistem pengawasan ujian dan melaporkan performa yang tinggi pada berbagai skenario deteksi kecurangan. Namun hingga saat ini masih terbatas penelitian yang secara langsung membandingkan kinerja dua algoritma ini untuk konteks kecurangan ujian sehingga klaim keunggulan masing-masing belum sepenuhnya didukung oleh analisis komparatif yang terstandarisasi. Penelitian ini bertujuan untuk melakukan analisis perbandingan kinerja kinerja algoritma YOLOv5 dan YOLOv8 dalam mendeteksi indikasi kecurangan ujian berbasis foto. Perbandingannya didasarkan pada hasil metrik evaluasi kuantitatif yaitu *precision*, *recall*, dan *mean Average Precision* (mAP) setelah kedua model diuji dengan dataset dan jumlah *epoch* yang sama. Hasil penelitian menunjukan kedua model memliki perbedaan karakteristik kinerja, di mana YOLOv5 unggul pada nilai *precision* dan mAP, sehingga lebih akurat dan sesuai untuk sistem pengawasan ujian dengan tingkat kesalahan deteksi rendah, sementara model YOLOv8 memiliki nilai *recall* lebih tinggi, khususnya pada kelas *cheating*, yang menunjukkan sensitivitas deteksi lebih baik dan lebih sesuai untuk sistem *proctoring* yang menekankan kelengkapan deteksi kecurangan. Penelitian ini memberikan kontribusi berupa dasar rekomendasi implementatif bagi pengembangan sistem *proctoring* otomatis sesuai kebutuhan operasional, apakah menekankan minimisasi kesalahan deteksi atau kelengkapan identifikasi kecurangan.

Kata Kunci: Pengawasan Ujian, Deteksi Kecurangan Akademik, Pembelajaran Mendalam, Evaluasi Perbandingan

Abstract− The Covid-19 pandemic and the shift to online learning have created new challenges in exam supervision. A systematic study found that the rate of cheating in online exams jumped to 55% during the pandemic, prompting the need for an adaptive and intelligent exam supervision system through the use of YOLO as object detection technology. Previous studies have applied YOLOv5 and YOLOv8 to exam monitoring systems, which are claimed to have superior performance in various exam detection scenarios. However, to date, there has been limited research that directly compares the performance of these two algorithms, so the claims of their respective superiority are not yet fully supported by standardized comparative analysis. This study aims to conduct a comparative analysis of the performance of the YOLOv5 and YOLOv8 algorithms in detecting indications of photo-based exam cheating. The comparison is based on the results of quantitative evaluation metrics, namely precision, recall, and mean Average Precision (mAP) after both models were tested with the same dataset and number of epochs. The results show that the two models have different performance characteristics, with YOLOv5 excelling in precision and mAP values, making it more accurate and suitable for exam monitoring systems with low detection error rates, while the YOLOv8 model has higher recall values, especially in the cheating class, indicating better detection sensitivity and making it more suitable for systems that emphasize the completeness of cheating detection. This research contributes to the basis for implementable recommendations for the development of automated proctoring systems according to operational needs, whether emphasizing the minimization of detection errors or the completeness of cheating identification.

Keywords: Exam Proctoring, Academic Cheating Detection, Deep Learning, Comparative Evaluation

1. PENDAHULUAN

Pendidikan merupakan salah satu sektor yang terus mengalami transformasi seiring dengan teknologi yang semakin berkembang pesat. Dalam sektor pendidikan, aspek evaluasi pembelajaran menjadi bagian penting untuk mengukur kompetensi dan tingkat pemahaman peserta didik. Evaluasi pembelajaran juga merupakan alat untuk mengukur tercapainya tujuan pendidikan termasuk mengukur program yang dirancang untuk mencapai tujuan tersebut[1]. Sayangnya, dalam praktiknya kecurangan pada proses evaluasi pembelajaran atau kecurangan ujian masih menjadi permasalahan serius diberbagai institusi pendidikan [2]

Kecurangan dalam sistem ujian konvensional terjadi dalam berbagai bentuk seperti melalui catatan, interaksi dengan peserta lain hingga menggunakan perangkat elektrokik menjadi tantangan besar bagi pengawas ujian. Sistem pengawasan konvensional juga memiliki beberapa kelemahan seperti keterbatasan kapasitas pengawas ketika jumlah peserta sangat banyak dengan ruangan yang luas. Pengawasan konvensional juga kurang efektif daalam mendeteksi kecurangan seperti gerakan kepala/tangan, atau kontak tangan antar peserta secara detail dalam waktu bersamaan [3]

Pandemi COVID-19 meningkatkan masalah yang terjadi pada integritas ujian. Covid memaksa aktivitas pendidikan beralih ke sistem daring, termasuk pelaksanaan ujian. Perubahan mendadak ini menghadirkan tantangan besar dalam hal pengawasan karena banyak ujian yang beralih dilakukan secara daring. Pengawasan ujian daring

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin

mengandalkan perangkat pribadi peserta seperti laptop, komputer atau ponsel yang dilengkapi kamera. Namun, tingkat pengawasan daring juga memiliki tantangan tersendiri. Kecurangan pada ujian ini dilakukan oleh peserta dengan membuka sumber eksternal, bekerja sama dengan orang lain di luar kamera, atau bahkan menggunakan perangkat lunak tertentu untuk mencari jawaban.

Studi sistematis yang dilakukan oleh Newton [4] menemukan bahwa sebelum pandemi, sekitar 30% mahasiswa mengaku curang dalam ujian daring, namun angka ini melonjak menjadi 55% selama pandemi. Penelitian lain oleh Pleasants [5] menunjukan bahwa ujian daring tanpa pengawasan (*unproctored*), tingkat kecurangan bahkan bisa mencapai 70%.Temuan empiris tersebut memberikan gambaran bahwa kecurangan ujian tidak dapat dipandang sebagai fenomena insidental, melainkan persoalan yang perlu untuk segera ditangani. Oleh karena itu, dibutuhkan sistem pengawasan yang adaptif, cerdas, dan mampu mendeteksi aktivitas kecurangan baik di ruang tatap muka maupun dalam ujian daring.

Dalam beberapa tahun terakhir, perkembangan teknologi kecerdasan buatan (*Artificial Intelligence*) khususnya di bidang *computer vision*, telah membuka peluang baru dalam mengatasi tantangan ini. Teknologi *computer vision* dapat digunakan untuk meningkatkan pengawasan dan integritas dalam pelaksanaan ujian, baik secara daring maupun luring. Sistem ini mampu mendeteksi perilaku mencurigakan, mengidentifikasi plagiarisme, serta mengotomatisasi penilaian, sehingga mengurangi beban pengawas dan meningkatkan keadilan ujian [6]. Salah satu pendekatan yang paling efektif dalam mendeteksi objek secara cepat dan akurat adalah algoritma *You Only Look Once* (YOLO). YOLO merupakan metode untuk deteksi objek untuk gambar atau video yang memiliki tingkat efisiensi dan kecepatan yang tinggi [7] [8]

Penelitian terdahulu sudah menerapkan YOLO versi terbaru seperti YOLOv5 dan YOLOv8 sebagai algoritma dalam sistem pengawasan ujian. Keduanya diklaim memiliki kemampuan unggul ketika diimplementasikan dalam sistem pengawasan ujian. Seperti pada penelitian oleh Ganidisastra & Bandung [9] yang menggunakan YOLOv5 untuk sistem *face recognition* pada *proctoring m-learning*, lalu penelitian oleh Ramzan [10] yang menggunakan YOLOv5 dalam mengklasifikasi aktivitas curang selama ujian daring, dan penelitian [11] sistem pengawasan ujian otomatis berbasis kamera yang mendeteksi wajah, arah pandang, dan objek terlarang secara *real-time*. Kemudian penelitian pada YOLOv8 seperti yang dilakukan oleh Zuo [12], Xu [13] dan Shibu [14], mengklaim bahwa YOLOv8 menawarkan peningkatan yang lebih baik dari segi akurasi, kecepatan deteksi, maupun kemampuan memahami konteks spasial multi-skala yang signifikan. YOLOv8 juga diklaim kemampuan yang lebih baik dalam mendeteksi objek berukuran kecil dan perilaku kompleks seperti menoleh, berbisik, atau menggunakan ponsel dalam kondisi ujian nyata secara lebih efisien. Namun hingga saat ini belum banyak penelitian yang secara langsung membandingkan kinerja YOLOv5 dan YOLOv8 dalam sistem pengawasan ujian berbasis foto, Akibatnya, klaim keunggulan masing-masing versi YOLO dalam konteks *proctoring* ujian masih belum sepenuhnya didukung oleh analisis komparatif yang terstandarisasi.

Penelitian terbaru dalam rentang 3 tahun terakhir yaitu penelitian tahun 2023 oleh Selcuk & Serif [15] yang mengembangkan model deteksi elemen UI pada pengembangan aplikasi *mobile* menggunakan YOLOv8 yang juga dibandingkan dengan studi sebelumnya yang menggunakan YOLOv5. Hasilnya diperoleh bahwa YOLOv8 merupakan model yang paling efektif untuk tugas pengenalan elemen GUI dalam konteks penelitian ini. Penelitian tahun 2024 oleh Casas, dkk [16] yang melakukan perbandingan kinerja YOLOv5 and YOLOv8 untuk deteksi kebakaran hutan dan asap, menggunakan dataset Foggia dengan hasil model YOLOv5s memperoleh skor tertinggi di semua metrik evaluasi. Tiga penelitian lain di tahun yang sama seperti oleh Swaropp, dkk [17] melakukan kinerja komparatif model YOLOv5 dan YOLOv8 untuk deteksi kendaraan. Temuan pada penelitian ini YOLOv8 menunjukkan keefektifan dalam tugas deteksi objek kendaraan. Ma, dkk [18] melakukan perbandingan performa YOLOv5 dan YOLOv8 untuk deteksi keberadaan objek manusia sebagai tahapan dalam pengembangan prototipe sistem panduan audio untuk penyandang tunanetra yang menunjukan bahwa YOLOv8 memiliki performa keseluruhan lebih baik dibanding YOLOv5 untuk deteksi objek manusia secara real-time. Iskandar, dkk [19] menilai efektivitas model YOLOv5 dan YOLOv8 dalam mendeteksi objek dalam kondisi pencahayaan rendah. Berdasarkan perbandingan kinerja, YOLOv8 unggul dibandingkan YOLOv5 dalam nilai presisi, *recall*, F1-*Score*, dan mAP. Penelitian terbaru oleh Megaarta [20] merepresentasikan perbandingan evaluasi model YOLOv5 dan YOLOv8 untuk deteksi otomatis perilaku merokok di ruang publik. Hasil evaluasi menunjukan YOLOv8 memberikan hasil akurasi yang lebih tinggi untuk deteksi merokok secara *real-time*. Berdasarkan temuan tersebut menunjukan bahwa belum ada penelitian yang melakukan perbandingan empiris kinerja YOLOv5 dan YOLOv8 dalam konteks deteksi kecurangan ujian, sehingga penelitian ini diarahkan untuk melakukan analisis perbandingan kinerja algoritma YOLOv5 dan YOLOv8 dalam sistem deteksi kecurangan ujian berbasis foto. Perbandingan penilaian kinerja didasarkan pada hasil metrik evaluasi kuantitatif yaitu *precision*, *recall* dan mAP. Hasil penelitian ini diharapkan dapat memberikan gambaran empiris mengenai keunggulan dan keterbatasan masing-masing algoritma, serta menjadi dasar pertimbangan dalam pemilihan model deteksi objek yang paling sesuai untuk diimplementasikan pada sistem pengawasan ujian otomatis.

2. METODOLOGI PENELITIAN

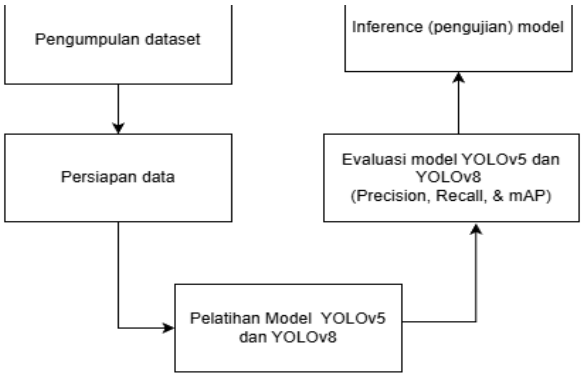
Metodologi penelitian ini diawali dengan tahap pengumpulan dataset berupa citra yang merepresentasikan perilaku ujian dalam kategori normal dan kecurangan. Dataset yang diperoleh kemudian melalui tahap persiapan data

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin

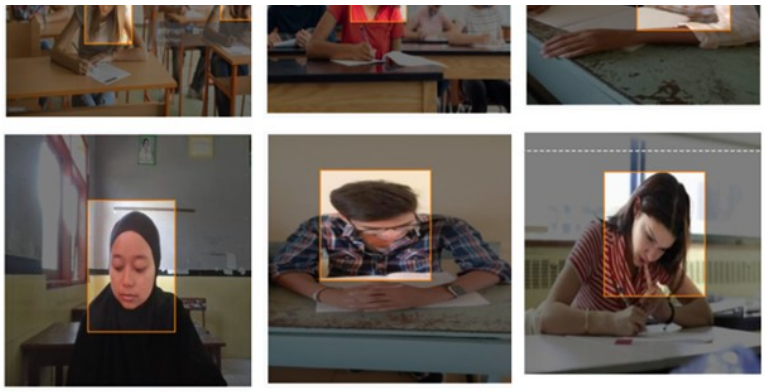
yang meliputi proses seleksi, pembersihan, anotasi, serta pembagian data untuk pelatihan dan pengujian. Selanjutnya, dilakukan pelatihan dua arsitektur deteksi objek, yaitu YOLOv5 dan YOLOv8, dengan konfigurasi parameter dan jumlah *epoch* yang sama agar hasil perbandingannya dapat berjalan dengan adil. Setelah proses pelatihan mencapai konvergensi yang stabil, kinerja kedua model dievaluasi menggunakan metrik kuantitatif yang mencakup *precision*, *recall*, dan *mean Average Precision* (mAP) untuk menilai tingkat akurasi dan sensitivitas deteksi. Tahap berikutnya adalah inferensi atau pengujian model terhadap data uji untuk mengamati kemampuan klasifikasi secara visual dan konsistensi prediksi pada kondisi nyata. Tahapan metode penelitian disajikan dalam diagram alir pada gambar 1 dibawah ini.



Gambar 1. Metodologi Penelitian

2.1 Pengumpulan dan Persiapan Data

Dataset diperoleh dari dataset anotasi publik dari *platform* Roboflow Universe. Dataset ini terdiri dari sekitar 300 data berbasis citra (*image-based)* yang merepresentasikan perilaku peserta ujian. Dataset terdiri dari dua kelas yaitu *cheating* dan no_*cheating*. Setiap citra dalam dataset kemudian dianotasi menunggunakan teknik *bounding box* untuk menandai area objek yang akan diamati. Dua kelas dataset *cheating* dan *no_cheating* dapat dilihat pada gambar 1 dan gambar 2.



Gambar 2. Dataset kelas *no cheating*

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin



Gambar 3. Dataset kelas *cheating*

Dataset kemudian dilakukan pembagian menjadi tiga data subset yaitu data latih (*train*), data validasi (*valid*) dan data uji (*test*) dengan detail pembagian datanya adalah data latih (*train*): 208 citra, data validasi (*valid*): 58 citra dan data uji (*test*): 34 citra. Pembagian dataset ini dilakukan secara otomatis pada saat dilakukan pemanggilan dataset di *platform* Roboflow. Data yang telah dikonfigurasi ini kemudian disimpan dalam file data.yml yang didalamnya memuat lokasi gambar, jumlah kelas, dan nama kelas.

2.2 Proses Pelatihan Model

Pada proses pelatihan model menggunakan dataset yang disudah dikonfigurasi dan disimpan pada file.yml, kedua model baik YOLOv5 maupun YOLOv8 kemudian dilatih menggunakan masing-masing 50 *epoch*. Ukuran citra yaitu 416x416 piksel untuk YOLOv5 dan 800 x 800 piksel untuk YOLOv8. Tabel 1 menunjukan detail konfigurasi untuk pelatihan pada masing-masing model.

Tabel 1. Konfigurasi pelatihan YOLOv5 dan YOLOv8

Aspek konfigurasi	YOLOv5	YOLOv8
Jumlah *epoch*	20 *epoch*	50 *epoch*
Konfigurasi ukuran citra	416 x 416 piksel	800 × 800 piksel
Varian YOLO	YOLOv5s	YOLOv8n

Selama proses pelatihan, kedua model melakukan pembaruan botot secara interatif menggunakan data latih dan dievaluasi berkala menggunakan data validasi.

2.3 Evaluasi Model

Usai dilakukan pelatihan pada model dan dalam prosesnya model menunjukan konvergensi yang stabil, maka selanjutnya dilakukan evaluasi terhadap kinerja model. Evaluasi kinerja pelatihan model menggunakan pendekatan berbasis metrik kuantitatif yaitu *precision*, *recall*, *mean average precision* pada ambang IoU 50% (mAP@50). Ke-tiga metrik ini merupakan metrik standar yang digunakan dalam tugas pemodelan deteksi objek. Perolehan nilai dalam ketiga metrik ini yang akan dinilai untuk mengetahui kemampuan model dalam mengenali dan mendeteksi objek dengan akurat [21].

Precision digunakan sebagai metrik yang mengukur seberapa besar proporsi deteksi positif yang benar (*true positives*) terhadap semua deteksi yang dihasilkan model (gabungan hasil *true positives* dengan *false positives*). Tujuan dari metrik ini adalah untuk menjawab berapa kelas yang terprediksi sebagai benar-benar kelas tersebut. Contohnya berapa kelas *cheating* yang terprediksi sebagai benar-benar kelas *cheating* dari semua objek yang diprediksi. Semakin tinggi nilai *precision* artinya model jarang melakukan kesalahan deteksi sehingga ketika model menyatakan bahwa sebuah objek tersebut ada maka objek tersebut memang benar-benar ada.

Recall digunakan untuk mengukur proporsi deteksi yang benar (*true positives*) diantara semua objek yang harus ditemukan dalam data (gabungan *true positives* dengan *true negatives*). Tujuannya adalah untuk menemukan berapa banyak objek sebenarnya yang mampu dideteksi oleh model. Contohnya berapa banyak objek yang berhasil dideteksi *cheating* dari keseluruhan data *cheating* yang ada. Nilai *recall* yang tinggi menunjukkan kemampuan model untuk menemukan sebagian besar objek yang relevan di dalam gambar.

mAP merupakan metrik yang mengukur rata-rata presisi yang menggabungkan antara *precision* dan *recall* di berbagai tingkat keyakinan (*confidence thresholds*) [22]. Tujuannya untuk mengukur kinerja keseluruhan model dengan

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin

mempertimbangkan hubungan antara precision dan *recall*. Nilai mAP yang tinggi menunjukkan bahwa model mampu mendeteksi objek secara akurat dan konsisten dengan tingkat kesalahan yang rendah.

2.4 Inference (Pengujian) Model

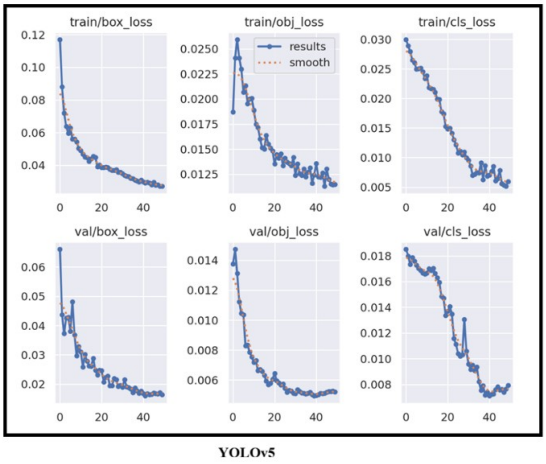
Tahap inferensi dilakukan untuk menguji kemampuan generalisasi model YOLOv5 dan YOLOv8 dalam mendeteksi perilaku kecurangan ujian pada data uji yang belum pernah dilihat oleh model selama proses pelatihan dan validasi berlangsung. Data uji yang digunakan pada tahap ini terdiri dari 34 citra statis berupa foto yang menampilkan berbagai pose kepala dan arah padangan yang merepresentasikan perilaku peserta ujian. Data uji ini bersifat independen yang tidak termasuk dalam data dataset pelatihan dan validasi sehingga dapat merepresentasikan kondisi ujian yang realistis.

Proses pengujian dilakukan dengan ambang *confidence threshold* sebesar 0.25 untuk YOLOv8 dan 0.1 untuk YOLOv5. Nilai ambang ini dipilih karena memberikan keseimbangan antara sesitivitas deteksi (*recall*) dan akurasi prediksi (*precision*) yang umumnya digunakan dalam YOLO terutama YOLOv8 [23]. Hasil pengujian disajikan dalam bentuk visualisasi deteksi yang memuat citra objek yang terdeteksi dengan ditandai *bounding box* dan label kelas hasil prediksinya sebagai *cheating* atau *no cheating.*

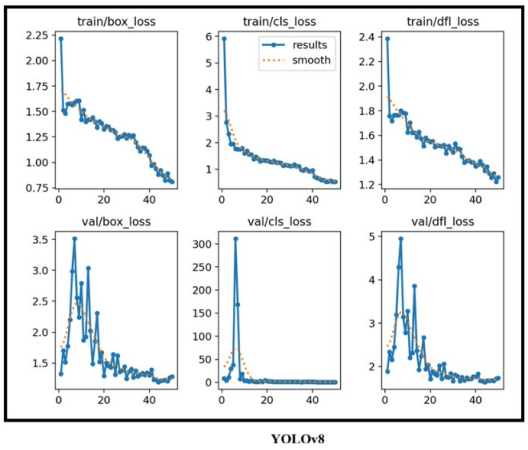
3. HASIL DAN PEMBAHASAN

3.1 Hasil

Hasil pelatihan pada kedua model menunjukan proses konvergensi yang stabil. Hasil proses pelatihan untuk model YOLOv5 dapat dilihat pada gambar 4.



Gambar 4. Hasil proses pelatihan model YOLOv5



Gambar 5. Hasil proses pelatihan model YOLOv8

Pada gambar 4 untuk YOLOv5 terlihat bahwa pada awal proses pelatihan baik pada data *train* dan validasi mulai menunjukan kondisi *loss* yang turun stabil pada *epoch* ke 40 dan mencapai kondisi maksimal pada *epoch* ke 50. Nilai *box loss* pada data *train* turun dari 0.12 ke 0.03 yang artinya model cukup baik dalam menentukan nilai kesalahan

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin

dalam menentukan posisi *bounding box*. Nilai *object loss* pada data train juga mengalami penurunan dari 0.025 ke 0.012 yanga artinya model cukup baik dalam mengenali ada atau tidaknya objek di suatu area. Nilai *classification loss* pada data train ikut mengalami penurunan dari 0.03 ke 0.005, artinya model makin tepat mengenali kelas objek. Pada grafik data validasi juga menunjukan tren penurunan nilai *loss* yang serupa.

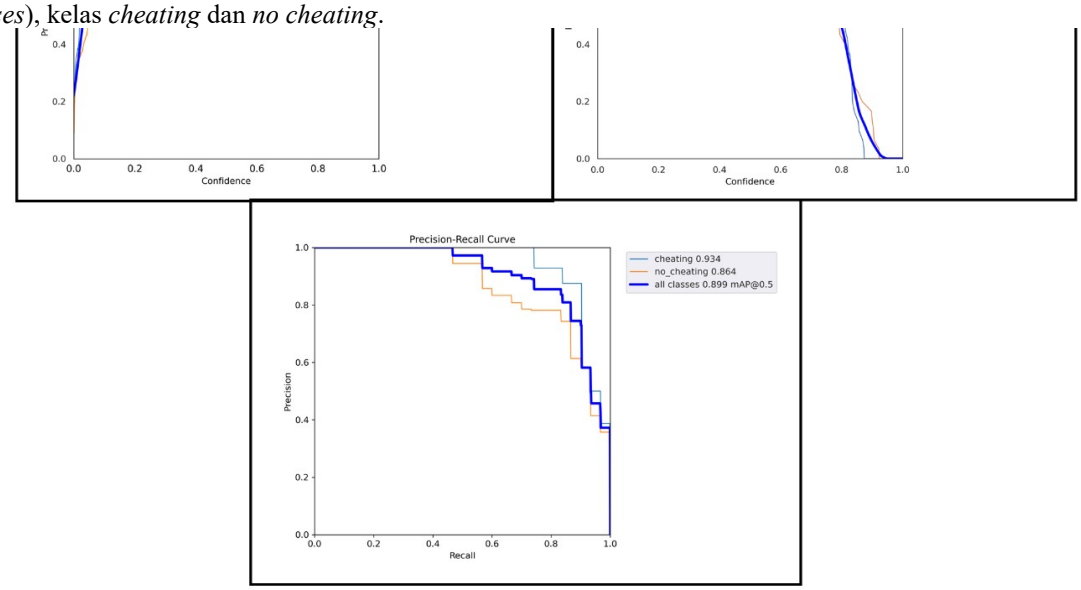
Pada gambar 5 untuk YOLOv8, proses pelatihan untuk model tersebut juga menunjukan konvergensi yang stabil pada *epoch* ke 40 hingga 50. Pada data *train*, nilai *box loss* menurun dari 2.2 ke 0.8. Nilai *classification loss* turun dari 6 ke 1, lalu nilai *distribution focal loss* atau dfl *loss* ikut mengalami penurunan dari 2.4 ke 1.2. Sedangkan untuk grafik pada data validasi meski prosesnya sempat fluktuatif namun juga menunjukan tren penurunan yang serupa dengan data train.

Hasil metrik evaluasi untuk model YOLOv5 dapat dilihat tabel 2 berikut ini.

Tabel 2. Hasil metrik evaluasi model YOLOv5

Kelas Yolo V5	Foto	Contoh	*Precision* (P)	*Recall* (R)	mAP50
All classes	58	61	0.831	0.851	0.899
Cheating	58	31	0.928	0.835	0.934
No Cheating	58	30	0.734	0.867	0.864

Metrik evaluasi pada model YOLOv5 untuk semua kelas atau *all classes* menunjukan hasil yang tinggi yaitu *precision*: 0.831, *recall*: 851, mAP50: 0.899. Kemudian untuk kelas *cheating* hasilnya yaitu *precision*: 0.928, *recall*: 0.835, mAP50: 0.934. Lalu untuk kelas *no cheating* hasilnya yaitu *precision*: 0.734, *recall*: 0.867, mAP50: 0.864. Berdasarkan hasil evaluasi model YOLOv5 memiliki tingkat akurasi yang tinggi untuk ketiga kelas baik semua kelas (*all classes*), kelas *cheating* dan *no cheating*.



Gambar 6. Kurva nilai *precision* dan *recall* untuk model YOLOv5

Kurva pada gambar 6 menunjukan hubungan antara nilai *precision* dan *recall* dari hasil evaluasi model YOLOv5 yang berubah ketika menaikkan atau menurunkan ambang *confidence threshold*. Pada kurva bagian *Precision Confidence Curve* menampilkan nilai *precision* yang mengalami kenaikan pada *all clasess* (semua kelas). Pada *confidence* rendah (<0.2), nilai *precision* pada *all classes* masih sekitar 0.4–0.6. Ketika *confidence* meningkat pada angka 0.4–0.8, nilai *precision* naik signifikan hingga 0.9 dan mencapai puncak pada *confidence* >0.8 yaitu diangka 1.0 yang artinya model dapat memprediksi tanpa kesalahan pada ambang tersebut. *Tren* yang sama berlaku untuk kelas *cheating* yang bahkan sudah mampu mencapai nilai *precision* 0.9 dengan *confidence* rendah (0.2), namun untuk kelas *no cheating* terjadi tren berbeda yang lebih lambat dan baru mencapai kenaikan maksimal ketika berada di *confidence* tinggi (0.8).

Pada kurva *Recall Confidence Curve* di gambar 4 menunjukkan hubungan antara nilai *recall* dan *confidence threshold* pada model YOLOv5. Pada *confidence* rendah (<0.2), nilai *recall* pada *all classes* berada di kisaran 0.95–1.0, yang berarti model mampu mendeteksi hampir seluruh objek. Ketika *confidence* meningkat (0.4–0.8), nilai *recall* mulai

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin

menurun secara bertahap dari 0.9 menjadi sekitar 0.6, menandakan model semakin selektif terhadap prediksi. Pada *confidence* tinggi (>0.8), nilai *recall* turun tajam mendekati 0.0, karena hanya deteksi dengan *confidence* tinggi yang diterima. Pola serupa juga terlihat pada kelas *cheating* dan *no cheating* dengan performa yang relatif seimbang.

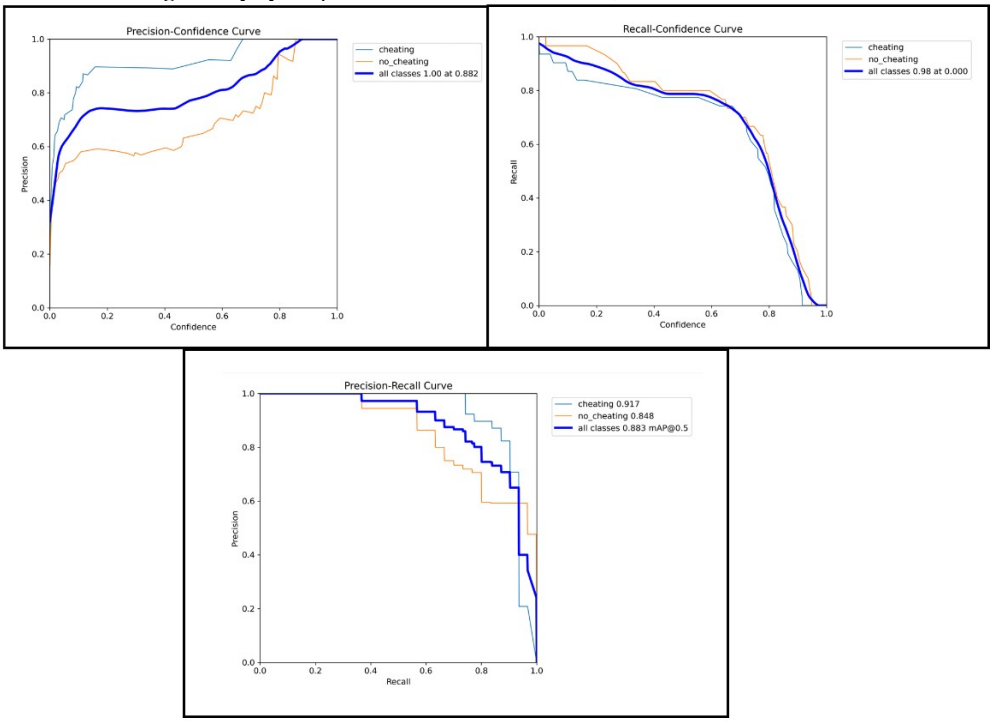
Kurva pada *Precision Recall Curve* di gambar 6 menunjukan hubungan antara nilai *precision* dan *recall* secara bersamaan untuk model YOLOv5. Kurva tersebut memperlihatkan keseimbangan antara ketepatan prediksi (*precision*) dan kemampuan model dalam menemukan seluruh objek yang benar (*recall*) pada berbagai ambang *confidence threshold*. Hasil pada kurva kelas *cheating* memiliki nilai AP (*avarage precision*) tertinggi yaitu 0.934, diikuti kelas *no_cheating* yaitu 0.864. Nilai mAP@0.5 sebesar 0.899 menandakan performa model yang sangat baik dalam mendeteksi perilaku kecurangan.

Selanjutnya hasil metrik evaluasi model YOLOv8 ditunjukan pada tabel

Tabel 3. Hasil metrik evaluasi model YOLOv8

Kelas Yolo V5	Foto	Contoh	*Precision* (P)	*Recall* (R)	mAP50
All classes	58	61	0.841	0.886	0.889
Cheating	31	31	0.908	0.839	0.943
No Cheating	27	30	0.773	0.933	0.836

Pada model YOLOv8 hasil metrik evaluasi juga memiliki tingkat akurasi yang tinggi untuk semua kelas. Kelas *all classes* menunjukan hasil *precision*: 0.841, *recall*: 0.886, dan mAP50:0.889. Kelas *cheating* hasilnya yaitu *precision*: 0.908, *recall*: 0.839, mAP50: 0.943. Kelas *no cheating* hasilnya yaitu *precision*: 0.773, *recall*: 0.933, dan mAP50: 0.836.



Gambar 7. Kurva nilai *precision* dan *recall* untuk model YOLOv8

Gambar 7 menunjukan kurva untuk metrik evaluasi pada model YOLOv8. Pada kurva *Precision–Confidence* menampilkan nilai *precision* yang mengalami peningkatan seiring dengan kenaikan *confidence threshold*. Kelas *cheating* menunjukkan peningkatan yang stabil dengan *precision* maksimal 1.0 pada *confidence* >6. Kelas *no cheating* peningkatannya melambat dan baru mencapai maksimal 1.0 pada *confidence* >8. Sementara *all class* peningkatannya stabil namun mencapai nilai maksimum 1.0 pada *confidence* >8 seperti kelas *no cheating*. Sebaliknya pada kurva *Recall–Confidence* menunjukkan pola penurunan *recall* ketika *confidence threshold* meningkat. Hal ini terjadi pada ketiga kelas baik *all class*, *cheating* dan juga *no cheating*. Masing-masing mencapai nilai *recall* maksimal 1 pada *confidence* <2. Sementara itu, kurva *Precision–Recall* memperlihatkan keseimbangan antara *precision* dan *recall* pada berbagai *confidence threshold*. Kelas *cheating* memperoleh nilai *Average Precision* (AP) sebesar 0,917, sedangkan kelas *no_cheating* memperoleh AP sebesar 0,848. Nilai mAP@0.5 sebesar 0,883 untuk seluruh kelas menunjukkan bahwa model memiliki performa yang baik dan konsisten dalam mendeteksi perilaku kecurangan maupun kondisi normal.

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin

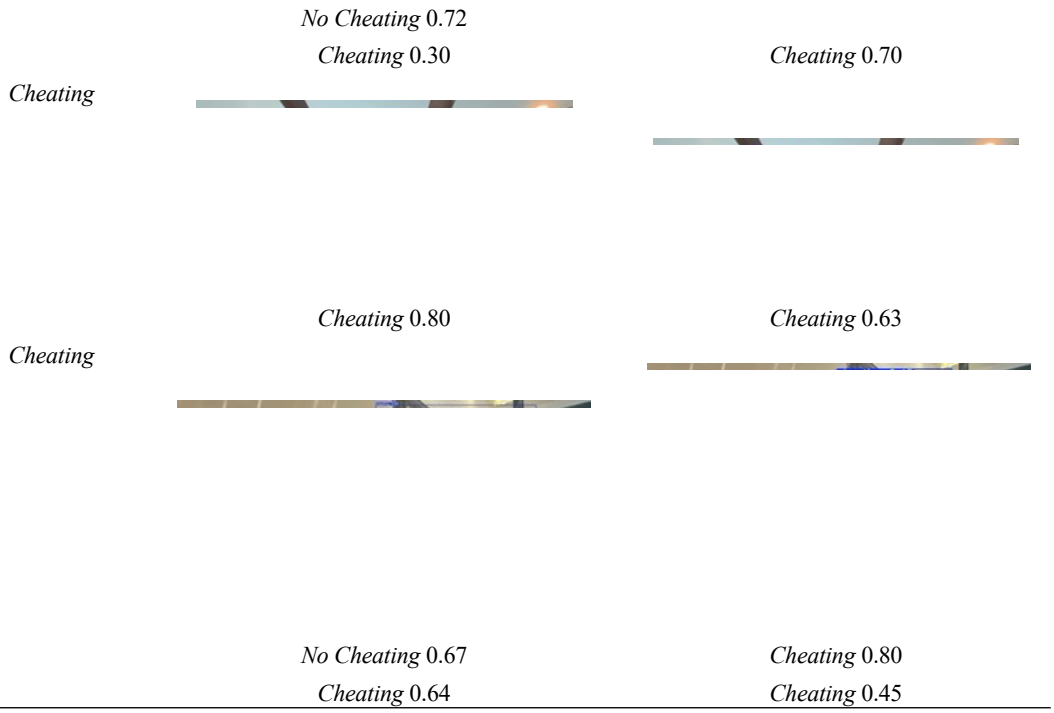
Tabel 4. Hasil Inferensi (Pengujian) pada model YOLOv5 dan YOLOv8

Label Kelas	Hasil Prediksi	
	YoloV5	YoloV8
No Cheating		
Cheating	*No Cheating* 0.89	*No Cheating* 0.80
Cheating	*Cheating* 0.86	*Cheating* 0.77
No Cheating	*Cheating* 0.76	*Cheating* 0.74
Cheating	*No Cheating* 0.92	*No Cheating* 0.82
Cheating	*Cheating* 0.80	*Cheating* 0.74
No Cheating	*Cheating* 0.88	*Cheating* 0.78

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin



Tabel 4 menyajikan hasil inferensi yang divisualisaikan melalui citra perilaku ujian beserta label hasil prediksi dan nilai confidence threshold-nya. Gambar dalam tabel 4 hanya menampilkan beberapa data yang mewakili hasil pengujianan dan tidak menampilkan keseluruhan 34 citra yang diuji. Berdasarkan hasil tersebut kedua model mampu memprediksi perilaku ujian sesuai kelasnya. Hasil pengujian menunjukan bahwa YOLOv5 dan YOLov8 dapat mengenali kondisi no cheating dengan nilai confidence threshold yang cukup tinggi. Pada kelas no cheating, YOLOv5 menunjukan hasil confidence threshold 0.72 sampai 0.92, jauh lebih tinggi dibanding YOLOv8 yang memperoleh hasil 0.70-0.82. Pada kelas cheating juga menunjukan pola serupa dimana YOLOv5 memperoleh hasil 0.76-0.88 sementara YOLOv8 memperoleh hasil lebih rendah dalam rentang 0.63-0.78. Selain itu kedua model juga masih melakukan kesalahan prediksi pada baris ke-7 dimana YOLOv8 memprediksi label no cheating sebagai cheating dengan nilai confidence threshold 0.70 dan pada baris ke 9 YOLOv5 juga melakukan kesalahan prediksi data cheating sebagai no cheating dengan confidence threshold 0.67.

3.2 Pembahasan

Berdasarkan hasil pelatihan kedua model menunjukan proses konvergensi yang stabil pada epoch terakhir yang menandakan bahwa model mampu mempelajari pola data dengan baik. Penurunan nilai loss yang serupa pada data training maupun validasi juga menunjukan tidak ada overfitting siginifikan dari kedua model.

Tabel 4. Perbandingan Metrik Evaluasi YOLOv5 dan YOLOv8

Kelas	Model	Precision	Recall	mAP@0.5
All classes	YOLOv5	0.831	0.851	0.899
	YOLOv8	0.841	0.886	0.889
Cheating	YOLOv5	0.928	0.835	0.934
	YOLOv8	0.841	0.886	0.889
No Cheating	YOLOv5	0.734	0.867	0.864
	YOLOv8	0.773	0.933	0.836

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

JURNAL MEDIA INFORMATIKA [JUMIN]
Volume 7 Nomor 2 April 2026, Page 257-268
ISSN 2808-005X (media online)
Available Online at http://ejournal.sisfokomtek.org/index.php/jumin

Dari sisi metrik hasil evaluasi pada kedua model yang dilatih seperti ditunjukan pada tabel 4. Terlihat bahwa kedua model sama-sama memiliki kinerja yang baik dalam mendeteksi kecurangan ujian berbasis foto. Kinerja baik terlihat dari perolehan nilai pada metrik *precision*, *recall* dan MAP yang relatif tinggi pada kedua model. Namun keduanya memiliki karakteristik kinerja yang berbeda yang akan dipaparkan pada sub bab berikut ini.

3.2.1 Perbandingan Kinerja Model Kelas *All Classes*

Berdasarkan hasil pada kelas *all classes* YOLOv8 lebih unggul pada metrik *precision* dan juga *recall*. YOLOv8 memperoleh nilai *precision* 0.841 dan dan *recall* 0.886, sementara YOLOv5 memperoleh *precision* 0.831 dan dan *recall* 0.851. Hal ini menunjukan bahwa YOLOv8 lebih baik dalam mendeteksi keseluruhan objek yang ada di dalam citra yang tercermin dari nilai *recall* yang lebih *tinggi*. YOLOv8 juga memiliki tingkat ketepatan yang lebih tinggi dalam menemukan objek yang sebenarnya yang tercermin dari nilai *precision* yang lebih tinggi. Namun, nilai mAP untuk *all classes*, YOLOv5 lebih unggul dibanding YOLOv8, dimana YOLOv5 memperoleh skor 0.899, sedangkan YOLOv8 memperoleh 0.889. Hal ini mengindikasikan bahwa YOLOv5 memiliki kemampuan deteksi yang jauh lebih baik dalam menilai keseimbangan antara ketepatan (*precision*) dan daya tangkap (*recall*). Perbedaan ini menunjukkan adanya kompromi kinerja (*trade-off*) antara *precision* dan *recall* pada kedua model yang diuji.

3.2.2 Perbandingan Kinerja Model Kelas *Cheating*

Pada kelas *Cheating*, YOLOv5 menunjukan performa yang sangat baik dengan memperoleh *precision* 0.928, *recall* 0.835, dan mAP@0.5 sebesar 0.934. Nilai *precision* yang sangat tinggi menunjukan bahwa YOLOv5 dapat memprediksi objek kecurangan dengan sangat tepat sesuai dengan kondisi nyata. Pada YOLOv8 untuk kelas ini juga menunjukan performa yang baik, dengan memperoleh *precision* 0.841, *recall* 0.886, dan mAP@0.5 sebesar 0.889. Berbeda dengan YOLOv5, pada YOLOv8 nilai *recall* jauh lebih tinggi menunjukkan bahwa YOLOv8 lebih mampu mendeteksi 80% lebih untuk perilaku kecurangan yang ada pada data uji. Dengan demikian, YOLOv8 lebih agresif dalam mendeteksi kecurangan, namun berpotensi menghasilkan lebih banyak prediksi yang kurang tepat dibandingkan YOLOv5.

3.2.3 Perbandingan Kinerja Model Kelas *No Cheating*

Pada kelas *no cheating*, YOLOv5 memperoleh *precision* 0.734, *recall* 0.867, dan mAP@0.5 sebesar 0.864. Nilai *recall* yang lebih tinggi dibanding *precision*, menujukan bahwa YOLOv5 lebih baik dalam mendeteksi seluruh aktivitas normal namun lebih mudah melakukan kesalahan prediksi dimana aktivitas kecurangan (*cheating*) mungkin bisa terdeteksi sebagai aktivitas normal (*no cheating).* Pada model YOLOv8 untuk kelas *no cheating* memperoleh *precision* 0.773 dan *recall* 0.933. Nilai recall yang sangat tinggi menunjukan bahwa YOLOv8 hampir mampu mengenali seluruh kondisi normal yang ada di dalam data uji.

3.2.4 Perbandingan Kinerja mAP untuk Semua Kelas

Selanjutnya untuk nilai mAP@0.5 pada ketiga kelas, YOLOv5 memperoleh nilai jauh lebih tinggi pada semua kelas yang artinya bahwa YOLOv5 memiliki keseimbangan yang lebih baik antara *precision* dan *recall* dibandingkan YOLOv8 pada seluruh kelas. Hal ini menunjukan bahwa, model YOLOv5 tidak hanya banyak menemukan objek, tetapi juga melakukannya dengan tingkat kesalahan yang lebih rendah secara konsisten.

3.2.5 Perbandingan Hasil Inferensi

Berdasarkan hasil visualisasi inferensi, baik YOLOv5 dan YOLOv8 memprediksi perilaku ujian ke dalam kelas yang sesuai. Namun kedua model menunjukan perbedaan karakteristik prediksi dengan hasil tingkat keyakinan (*confidence threshold*) dan pola kesalahan klasifikasi.

Pada kelas *no cheating*, YOLOv5 menghasilkan nilai *confidence* yang lebih tinggi dan stabil yaitu pada rentang 0.72–0.92, hal ini menunjukan bahwa model mampu mengenali kondisi ujian normal dengan tingkat kepercayaan yang tinggi. YOLOv8 juga memiliki kemampuan yang baik dalam memprediksi kelas ini, namun dengan nilai *confidence* yang lebih moderat yaitu 0.70–0.82. Tren serupa juga terlihat pada kelas *cheating*, YOLOv5 memberikan prediksi dengan *confidence* lebih tinggi (0.76–0.88) dibandingkan YOLOv8 (0.63–0.78). Hal ini menunjukan bahwa YOLOv5 memiliki kemampuan yang lebih selektif dalam melakukan prediksi sehingga hasil prediksinya memiliki tingkat *confidence* yang tinggi.

Meskipun demikian, kedua model masih mengalami kesalahan pada beberapa kasus tertentu. YOLOv8 keliru mengklasifikasikan kondisi *no cheating* sebagai *cheating* dengan *confidence* 0.70, sedangkan YOLOv5 salah memprediksi kelas *cheating* sebagai *no cheating* dengan *confidence* 0.67. Kesalahan ini menunjukkan bahwa perilaku ujian yang ambigu masih menjadi tantangan bagi kedua model dalam proses inferensi.

4. KESIMPULAN

Penelitian ini membandingkan YOLOv5 dan YOLOv8 untuk deteksi kecurangan ujian berbasis foto. Berdasarkan perolehan nilai pada tiga metrik evaluasi yang tinggi, kedua model terbukti layak diterapkan pada sistem

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

pengawasan ujian, namun dengan karakteristik yang berbeda. Berdasarkan hasil perbandingan terhadap hasil metrik evaluasi dan hasil pengujian maka dapat disimpulkan bahwa YOLOv5 menghasilkan nilai *precision* dan nilai mAP yang lebih tinggi sehingga model ini lebih unggul dalam ketepatan dan kestabilan deteksi, sementara YOLOv8 menghasilkan nilai *recall* yang lebih tinggi pada semua kelas sehingga model ini lebih unggul dalam kemampuan menemukan sebanyak mungkin indikasi kecurangan. Dengan demikian pemilihan model dapat disesuaikan dengan kebutuhan sistem pengawasan ujian. Jika membutuhkan sistem yang menekankan pada tingkat akurasi tinggi dan minim kesalahan maka model YOLOv5 lebih direkomendasikan, namun jika lebih menekankan pada kemampuan dalam deteksi kecurangan yang menyeluruh maka dapat menggunakan model YOLOv8.

Meskipun penelitian ini memberikan analisis komparatif menggunakan metrik evaluasi yang terstandarisasi namun masih memiliki keterbatasan yang perlu diperhatikan. Pertama dataset yang digunakan masih terbatas pada dataset berbasis foto dengan jumlah terbatas, sehingga belum sepenuhnya merepresentasikan variasi situasi ujian dengan kondisi lebih kompleks seperti perbedaan pencahayaan, kualitas kamera dan juga kepadatan peserta. Kedua, konfigurasi pelatihan menggunakan jumlah *epoch* dan parameter dasar yang sama, namun belum melakukan eksplorasi *hyperparameter tuning* lanjutan yang berpotensi meningkatkan performa masing-masing model secara lebih optimal. Ketiga, evaluasi masih berfokus pada metrik kuantitatif utama (*precision*, *recall*, dan mAP) namun belum melakukan analisis mendalam terhadap aspek efisiensi komputasi seperti waktu inferensi dan penggunaan sumber daya perangkat, yang penting untuk implementasi sistem *proctoring real-time*.

Untuk penelitian selanjutnya disarankan untuk melakukan pengujian pada dataset yang lebih besar dan beragam, termasuk skenario ujian nyata. Melakukan eksplorasi teknik augmentasi data yang lebih mendalam dengan optimasi *hyperparameter*, serta penerapan model pada arsitektur *real-time* berbasis *video streaming* sehingga dapat memberikan gambaran performa yang lebih komprehensif. Terakhir, penelitian lanjutan juga dapat difokuskan pada evaluasi terhadap efisiensi komputasi dan performa menggunakan perangkat dengan spesifikasi berbeda untuk memastikan kesiapan implementasi sistem pada lingkungan operasional yang sesungguhnya.

REFERENCES

[1] I. L, "Evaluasi dalam Proses Pembelajaran," *Adaara J. Manaj. Pendidik. Islam*, vol. 9, no. 2, hal. 920–935, Agu 2019, doi: 10.35673/ajmpi.v9i2.427.

[2] F. Noorbehbahani, A. Mohammadi, dan M. Aminazadeh, "A systematic review of research on cheating in online exams from 2010 to 2021," *Educ. Inf. Technol.*, vol. 27, no. 6, hal. 8413–8460, Jul 2022, doi: 10.1007/s10639-022-10927-7.

[3] F. Mahmood *et al.*, "Implementation of an Intelligent Exam Supervision System Using Deep Learning Algorithms," *Sensors*, vol. 22, no. 17, hal. 6389, Agu 2022, doi: 10.3390/s22176389.

[4] P. Newton dan K. Essex, "How common is cheating in online exams and did it increase during the COVID-19 pandemic? A Systematic Review," 22 November 2022. doi: 10.21203/rs.3.rs-2187710/v1.

[5] J. Pleasants, J. M. Pleasants, dan B. Pleasants, "Cheating on Unproctored Online Exams: Prevalence, Mitigation Measures, and Effects on Exam Performance," *Online Learn.*, vol. 26, no. 1, Mar 2022, doi: 10.24059/olj.v26i1.2620.

[6] M. G. Méndez-Ortega, E. P. Herrera-Granda, A. E. P. Malte, dan R. B. H. Enríquez, "Supervision and Control of Students during Online Assessments Applying Computer Vision Techniques: A Systematic Literature Review," *Univers. J. Educ. Res.*, vol. 9, no. 5, hal. 1000–1013, Mei 2021, doi: 10.13189/ujer.2021.090513.

[7] H. Z. Amrulloh dan F. A. N. Muhammad, "Deteksi Rambutan Matang dan Busuk Menggunakan Algoritma YOLOv9," *JIKO (Jurnal Inform. dan Komputer)*, vol. 9, no. 1, hal. 53, 2025, doi: 10.26798/jiko.v9i1.1382.

[8] L. Afriyanti, "Kombinasi Teknik Pemrosesan Citra untuk Peningkatan Pendeteksian Objek Pada Carla Simulator," *JIKO (Jurnal Inform. dan Komputer)*, vol. 9, no. 3, hal. 524, 2025, doi: 10.26798/jiko.v9i3.1958.

[9] A. H. S. Ganidisastra dan Y. Bandung, "An Incremental Training on Deep Learning Face Recognition for M-Learning Online Exam Proctoring," in *2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, IEEE, Apr 2021, hal. 213–219. doi: 10.1109/APWiMob51111.2021.9435232.

[10] M. Ramzan, A. Abid, M. Bilal, K. M. Aamir, S. A. Memon, dan T.-S. Chung, "Effectiveness of Pre-Trained CNN Networks for Detecting Abnormal Activities in Online Exams," *IEEE Access*, vol. 12, hal. 21503–21519, 2024, doi: 10.1109/ACCESS.2024.3359689.

[11] A. P. Hendrawan, E. Wijayanti, dan A. A. Chamid, "Design of an Exam Cheating Detection System Application Based on Machine Learning with the Computer Vision Method," *J. Teknol. Inform. dan Komput.*, vol. 11, no. 2, hal. 509–521, Jul 2025, doi: 10.37012/jtik.v11i2.2704.

[12] Y. Zuo, S. S. Chai, dan K. L. Goh, "Cheating Detection in Examinations Using Improved YOLOv8 with Attention Mechanism," *J. Comput. Sci.*, vol. 20, no. 12, hal. 1668–1680, 2024, doi: 10.3844/jcssp.2024.1668.1680.

[13] J. Lu, N. Song, W. Zhang, J. Wang, Z. Luo, dan Y. Wang, "Cheating Recognition in Examination Halls Based on Improved YOLOv8," *Proc. - 2024 Int. Conf. Artif. Intell. Things Syst. AIoTSys 2024*, 2024, doi: 10.1109/AIoTSys63104.2024.10780486.

[14] Shruti Maria Shibu, "Object Detection for Real-Time Malpractice Detection in Classrooms Using Computer

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**

Vision," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 33s, hal. 131–140, Apr 2025, doi: 10.52783/jisem.v10i33s.5464.

[15] B. Selcuk dan T. Serif, "A Comparison of YOLOv5 and YOLOv8 in the Context of Mobile UI Detection," 2023, hal. 161–174. doi: 10.1007/978-3-031-39764-6_11.

[16] E. Casas, L. Ramos, E. Bendek, dan F. Rivas-Echeverria, "YOLOv5 vs. YOLOv8: Performance Benchmarking in Wildfire and Smoke Detection Scenarios," *J. Image Graph.*, vol. 12, no. 2, hal. 127–136, 2024, doi: 10.18178/joig.12.2.127-136.

[17] A. Swaroop, A. Satsangi, M. Sameer, dan G. Ahmad, "Performance Evaluation of YOLOv5 and YOLOv8 for Vehicle Detection: A Comparative Study," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE, Jun 2024, hal. 1–6. doi: 10.1109/ICCCNT61001.2024.10723901.

[18] N. Ma Muriyah, J. H. Sim, dan A. Yulianto, "Evaluating YOLOv5 and YOLOv8: Advancements in Human Detection," *J. Inf. Syst. Informatics*, vol. 6, no. 4, hal. 2999–3015, Des 2024, doi: 10.51519/journalisi.v6i4.944.

[19] R. J. Iskandar, C. Fatichah, dan A. Yuniarti, "Object Detection in Low-Light Conditions: A Comparison using YOLOv5 and YOLOv8," in *2024 4th International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*, IEEE, Jul 2024, hal. 558–563. doi: 10.1109/ICSINTESA62455.2024.10748090.

[20] M. Megaarta, "Comparative Evaluation of YOLOv5 and YOLOv8 Models in Detecting Smoking Behavior," *J. Artif. Intell. Eng. Appl.*, vol. 4, no. 3, hal. 2048–2056, Jun 2025, doi: 10.59934/jaiea.v4i3.1089.

[21] F. Nurdiyansyah, I. Akbar, dan L. Ursaputra, "Segmentasi Berbasis Warna Untuk Pengelompokan Kualitas Cacing Anc Menggunakan Yolov8," *JIKO (Jurnal Inform. dan Komputer)*, vol. 9, no. 1, hal. 239, 2025, doi: 10.26798/jiko.v9i1.1779.

[22] B. Cheng, R. Girshick, P. Dollár, A. C. Berg, dan A. Kirillov, "Boundary IoU: Improving object-centric image segmentation evaluation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, hal. 15329–15337, 2021, doi: 10.1109/CVPR46437.2021.01508.

[23] C. Bhalerao, "Understanding Hyper-parameter-tuning of YOLO's," towardsai.net. Diakses: 7 Januari 2026. [Daring]. Tersedia pada: https://towardsai.net/p/l/understanding-hyper-parameter-tuning-of-yolos?utm_source=chatgpt.com

This is an open access article under the CC–BY-SA license

Terakreditasi SINTA 5 SK :72/E/KPT/2024
Submitted: **23/12/2025**; Accepted: **30/01/2025**; Published: **04/04/2026**