

DETEKSI TINGKAT KEFOKUSAN SISWA DALAM PEMBELAJARAN MENGUNAKAN ARSITEKTUR YOLOV26

Malika Harsanto, Andia Enggar Mayasari, Aulia Nabila Realisti, Janera Almasahni, Husni Thamrin

Magister Informatika, Universitas Muhammadiyah Surakarta
Jl. Ahmad Yani, Tromol Pos1, Pabelan Kartasura, Surakarta, Indonesia
L208240017@student.ums.ac.id

ABSTRAK

Tingkat kefokuskan siswa selama pembelajaran berperan penting dalam mendukung pemahaman materi, partisipasi kelas, dan keberhasilan proses belajar. Namun, pengamatan fokus siswa secara manual masih memiliki keterbatasan karena bergantung pada persepsi guru, sulit dilakukan secara konsisten, dan kurang efektif pada kelas dengan jumlah siswa yang banyak. Kondisi kelas yang dinamis, perubahan posisi siswa, pencahayaan tidak stabil, serta objek yang saling bertumpukan juga membuat proses pengamatan menjadi lebih sulit. Penelitian ini bertujuan membangun sistem deteksi tingkat kefokuskan siswa secara otomatis menggunakan arsitektur YOLOv26. Data penelitian terdiri atas dua kelas, yaitu *focus* dan *not_focus*, yang diperoleh dari Roboflow. Tahapan penelitian meliputi pengumpulan dataset, pra-pemrosesan data, pelatihan model, pengujian deteksi, dan evaluasi performa. Data dibagi menjadi data latih, validasi, dan uji. Pelatihan dilakukan selama seratus epoch menggunakan optimizer AdamW. Evaluasi dilakukan berdasarkan precision, recall, mean Average Precision atau mAP, serta confusion matrix. Hasil penelitian menunjukkan bahwa YOLOv26 memperoleh precision sebesar 89,6%, recall sebesar 88,7%, dan mAP sebesar 95,2%. Confusion matrix menunjukkan prediksi benar sebesar 97% pada kelas *focus* dan 96% pada kelas *not_focus*. Hasil ini menunjukkan bahwa YOLOv26 dapat membantu monitoring kefokuskan siswa secara cepat, objektif, dan cukup akurat di lingkungan pembelajaran. Sistem ini juga berpotensi mendukung guru dalam mengevaluasi perhatian siswa secara berkelanjutan.

Kata kunci : YOLOv26, computer vision, deteksi fokus, pembelajaran, deteksi objek

1. PENDAHULUAN

Perkembangan teknologi pembelajaran dalam beberapa tahun terakhir mendorong perubahan cara guru memantau aktivitas belajar siswa di kelas. Proses pembelajaran tidak hanya berfokus pada penyampaian materi, tetapi juga pada bagaimana perhatian dan keterlibatan siswa dapat diamati secara berkelanjutan. Tingkat kefokuskan siswa menjadi salah satu indikator penting dalam keberhasilan pembelajaran. Siswa yang kehilangan fokus cenderung mengalami penurunan pemahaman materi, keterlambatan menyelesaikan tugas, dan rendahnya partisipasi selama proses belajar berlangsung. Kondisi tersebut sering terjadi pada pembelajaran tatap muka maupun hybrid, terutama ketika guru harus mengawasi banyak siswa secara bersamaan.

Pengamatan fokus siswa secara manual masih memiliki berbagai keterbatasan. Guru umumnya menilai perhatian siswa melalui pengamatan langsung, seperti arah pandang, posisi kepala, ekspresi wajah, dan aktivitas selama pembelajaran. Cara ini dapat membantu guru memahami kondisi kelas, tetapi hasilnya sering dipengaruhi oleh subjektivitas, kelelahan, perbedaan persepsi, dan keterbatasan waktu pengawasan sehingga hasil penilaian kurang optimal[1]. Permasalahan tersebut menunjukkan bahwa pemantauan kefokuskan siswa membutuhkan cara yang lebih cepat, konsisten, dan objektif.

Pemanfaatan teknologi pengolahan citra mulai banyak diterapkan dalam bidang pendidikan untuk membantu pemantauan aktivitas belajar siswa. Teknologi ini memungkinkan sistem mengenali pola perilaku siswa melalui gambar atau video pembelajaran. Beberapa penelitian telah menggunakan metode deteksi objek berbasis YOLO untuk mengidentifikasi perilaku siswa di kelas. Penelitian Rosalina tahun 2024 menunjukkan bahwa YOLOv5 mampu mendeteksi siswa fokus dan tidak fokus siswa secara real-time dengan nilai mAP@50 mencapai 89,9%[2]. Penelitian lain menggunakan YOLOv10 untuk mendeteksi perilaku siswa di lingkungan kelas dan menghasilkan performa yang cukup baik pada berbagai kondisi pencahayaan dan posisi siswa[3]. Selain itu, pendekatan berbasis YOLOv8 dan pengenalan ekspresi wajah juga terbukti dapat membantu identifikasi tingkat perhatian siswa selama proses pembelajaran berlangsung[4].

Meskipun demikian, deteksi kefokuskan siswa di lingkungan kelas masih memiliki beberapa tantangan. Lingkungan kelas bersifat dinamis karena banyak siswa berada dalam satu gambar, posisi siswa saling berdekatan, sudut kamera dapat berubah, pencahayaan tidak selalu stabil, dan gerakan siswa berlangsung secara terus-menerus. Kondisi tersebut membuat proses deteksi menjadi lebih sulit, terutama ketika objek berukuran kecil atau sebagian tertutup. Penelitian terbaru mengenai deteksi perilaku siswa di kelas menyebutkan bahwa tantangan utama sistem pengenalan perilaku terletak pada kemampuan model dalam menjaga ketepatan hasil sekaligus

mempertahankan kecepatan deteksi secara langsung[5]. Oleh karena itu, diperlukan metode yang mampu mengenali pola visual siswa dengan baik dan tetap ringan untuk diterapkan pada lingkungan pembelajaran.

Penelitian ini bertujuan untuk membangun sistem deteksi tingkat kefokuskan siswa menggunakan arsitektur YOLOv26. Sistem dirancang untuk mengenali dua kondisi utama, yaitu siswa fokus dan siswa tidak fokus, berdasarkan gambar hasil pengamatan kelas. Pemilihan YOLOv26 didasarkan pada kemampuannya dalam melakukan deteksi secara cepat dan stabil pada lingkungan kelas yang dinamis. Melalui sistem ini, proses pemantauan pembelajaran diharapkan dapat dilakukan secara lebih objektif sehingga guru memperoleh bantuan dalam mengevaluasi keterlibatan siswa selama proses belajar berlangsung.

Rencana penyelesaian masalah dalam penelitian ini dilakukan melalui beberapa tahapan. Tahap pertama adalah pengumpulan dataset citra aktivitas siswa dengan dua kelas utama, yaitu focus dan not_focus. Tahap kedua adalah pra-pemrosesan data untuk menyesuaikan ukuran gambar dan memeriksa anotasi. Tahap ketiga adalah pelatihan sistem menggunakan arsitektur YOLOv26. Tahap keempat adalah pengujian deteksi untuk melihat kemampuan sistem dalam mengenali siswa fokus dan tidak fokus. Tahap terakhir adalah evaluasi performa menggunakan precision, recall, mean Average Precision atau mAP, dan confusion matrix. Melalui tahapan tersebut, penelitian ini diharapkan mampu menghasilkan sistem pemantauan pembelajaran yang cepat, objektif, dan relevan untuk membantu guru mengamati tingkat kefokuskan siswa di kelas.

2. TINJAUAN PUSTAKA

Beberapa penelitian terdahulu telah membahas penggunaan deteksi objek untuk mengenali perilaku dan perhatian siswa di kelas. Rosalina, Bimantoro, dan Wijaya menggunakan YOLOv5 untuk mendeteksi siswa fokus dan tidak fokus selama pembelajaran. Penelitian tersebut menunjukkan bahwa metode YOLO dapat digunakan untuk mendeteksi tingkat fokus siswa secara langsung dengan nilai mAP@50 mencapai 89,9%[2]. Penelitian ini menjadi dasar bahwa deteksi objek dapat membantu proses pemantauan perhatian siswa di kelas.

Pramudita dkk. menerapkan YOLOv10 untuk mendeteksi perilaku siswa dalam analisis keterlibatan kelas. Penelitian tersebut menunjukkan bahwa deteksi perilaku siswa dapat dilakukan pada berbagai kondisi pencahayaan dan posisi siswa[3]. Hal ini relevan dengan kondisi kelas nyata yang memiliki variasi sudut pandang, gerakan siswa, serta perubahan kondisi ruangan selama pembelajaran berlangsung.

Xing mengembangkan pendekatan berbasis YOLOv8 dan pengenalan ekspresi wajah untuk mendeteksi perhatian siswa di kelas. Penelitian tersebut menekankan bahwa perhatian siswa dapat

diamati melalui ciri visual, termasuk ekspresi wajah dan posisi siswa selama pembelajaran[4]. Pendekatan ini menunjukkan bahwa pengamatan visual dapat digunakan untuk membantu guru memahami tingkat perhatian siswa secara lebih terukur.

Sheng, Li, dan Chan mengembangkan deteksi perilaku siswa secara langsung menggunakan pengembangan YOLOv8s. Penelitian tersebut menjelaskan bahwa tantangan utama dalam deteksi perilaku siswa terletak pada kemampuan sistem menjaga ketepatan hasil sekaligus mempertahankan kecepatan deteksi[5]. Tantangan tersebut penting karena lingkungan kelas memiliki banyak variasi, seperti siswa yang saling berdekatan, objek yang tertutup sebagian, dan perubahan gerak siswa.

Jia dan He membahas pengenalan perilaku siswa di kelas menggunakan pendekatan berbasis pembelajaran mesin. Penelitian tersebut menunjukkan bahwa perilaku siswa selama pembelajaran dapat dikenali melalui data visual sehingga dapat membantu proses evaluasi keterlibatan siswa[8]. Wang juga menjelaskan bahwa pengolahan citra dapat digunakan untuk mendeteksi dan menganalisis perilaku siswa dalam kegiatan pembelajaran[12]. Sementara itu, Zhang meneliti deteksi perilaku siswa pada pembelajaran berbasis YOLOv5 dan menunjukkan bahwa metode deteksi objek dapat diterapkan untuk mengenali aktivitas siswa di lingkungan pendidikan[13].

Berdasarkan beberapa penelitian tersebut, dapat disimpulkan bahwa deteksi objek telah banyak digunakan untuk membantu pemantauan perilaku dan perhatian siswa di kelas. Namun, penelitian sebelumnya masih menghadapi tantangan pada kondisi kelas yang dinamis, seperti pencahayaan yang berubah, posisi siswa yang berdekatan, sudut kamera yang berbeda, dan objek yang tertutup sebagian. Oleh karena itu, penelitian ini menggunakan YOLOv26 untuk mendeteksi tingkat kefokuskan siswa dengan dua kelas utama, yaitu focus dan not_focus. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem pemantauan pembelajaran yang cepat, objektif, dan relevan untuk membantu guru mengevaluasi perhatian siswa di kelas.

2.1. Deteksi Objek

Deteksi objek merupakan teknologi dalam bidang pengolahan citra yang digunakan untuk mengenali dan menentukan lokasi objek pada gambar maupun video digital[6]. Teknologi ini tidak hanya mengidentifikasi jenis objek, tetapi juga menentukan posisi objek melalui kotak pembatas. Dalam perkembangannya, metode deteksi objek semakin banyak digunakan karena mampu mengenali objek secara cepat dan akurat pada berbagai kondisi visual.

Salah satu metode yang banyak digunakan dalam deteksi objek adalah YOLO. Metode ini memiliki kemampuan mendeteksi objek dalam satu tahap proses sehingga dapat digunakan pada sistem yang membutuhkan hasil cepat [7]. Dalam bidang

pendidikan, deteksi objek dapat dimanfaatkan untuk mengenali aktivitas siswa selama pembelajaran berlangsung, seperti memperhatikan materi, menulis, atau melakukan aktivitas lain yang berkaitan dengan tingkat fokus siswa[8].

2.2. Tingkat Kefokusan Siswa

Kefokusan siswa merupakan kemampuan siswa dalam memusatkan perhatian selama proses pembelajaran berlangsung[9]. Tingkat fokus memiliki peran penting karena berkaitan dengan pemahaman materi, keterlibatan dalam kelas, dan keberhasilan siswa dalam mengikuti kegiatan belajar. Siswa yang mampu mempertahankan fokus cenderung lebih mudah memahami penjelasan guru dan menyelesaikan tugas dengan baik.

Pengamatan fokus siswa secara manual masih memiliki keterbatasan. Guru harus memperhatikan banyak siswa dalam waktu yang bersamaan, sehingga proses pengamatan sering tidak dapat dilakukan secara terus-menerus. Selain itu, hasil pengamatan juga dapat dipengaruhi oleh subjektivitas, kelelahan, dan perbedaan persepsi[10]. Oleh karena itu, diperlukan sistem yang mampu membantu proses pemantauan perhatian siswa secara lebih objektif dan efisien[11].

2.3. Pengolahan Citra dalam Pembelajaran

Penerapan pengolahan citra dalam bidang pendidikan berkembang untuk membantu proses pemantauan aktivitas belajar siswa secara otomatis[12]. Teknologi ini memungkinkan sistem mengenali perilaku siswa melalui gambar maupun video pembelajaran. Dengan bantuan sistem tersebut, guru dapat memperoleh informasi tambahan mengenai kondisi perhatian siswa selama pembelajaran berlangsung.

Beberapa penelitian menunjukkan bahwa penggunaan pengolahan citra dapat membantu guru mengevaluasi keterlibatan siswa dengan lebih cepat dan konsisten[13]. Dalam hal pembelajaran di kelas, teknologi ini dapat digunakan untuk mendeteksi posisi siswa, arah perhatian, ekspresi wajah, dan aktivitas lain yang berhubungan dengan tingkat fokus. Oleh karena itu, pengolahan citra memiliki potensi untuk mendukung proses evaluasi pembelajaran secara lebih objektif.

2.4. Arsitektur YOLO

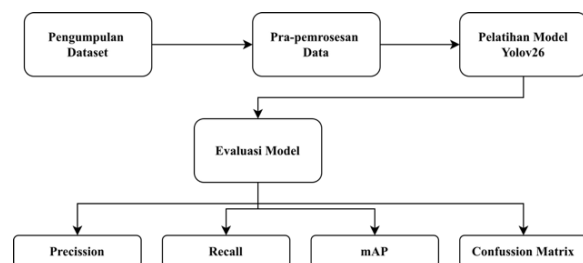
YOLO atau You Only Look Once merupakan metode deteksi objek yang bekerja dengan memprediksi lokasi dan kelas objek dalam satu tahap proses. Metode ini memiliki keunggulan pada kecepatan deteksi sehingga cocok digunakan untuk kebutuhan pemantauan secara *real-time*[14]. Keunggulan tersebut membuat YOLO banyak diterapkan pada berbagai bidang, termasuk pendidikan, keamanan, transportasi, dan pemantauan aktivitas manusia.

Pengembangan YOLO hingga versi terbaru meningkatkan kemampuan sistem dalam mengenali

objek pada kondisi yang kompleks. Lingkungan kelas termasuk kondisi yang menantang karena terdapat banyak siswa dalam satu gambar, posisi siswa dapat saling berdekatan, pencahayaan dapat berubah, dan sebagian objek dapat tertutup[15]. Pada penelitian ini digunakan arsitektur YOLOv26 untuk mendeteksi tingkat kefokuskan siswa selama proses pembelajaran berlangsung. Sistem diarahkan untuk mengenali dua kelas utama, yaitu siswa fokus dan siswa tidak fokus.

3. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan utama yang saling berkaitan untuk membangun sistem deteksi tingkat kefokuskan siswa menggunakan arsitektur YOLOv26. Tahapan penelitian dimulai dari proses pengumpulan dataset hingga evaluasi performa model deteksi. Skema tahapan penelitian yang digunakan dapat dilihat pada Gambar 1. Berikut:

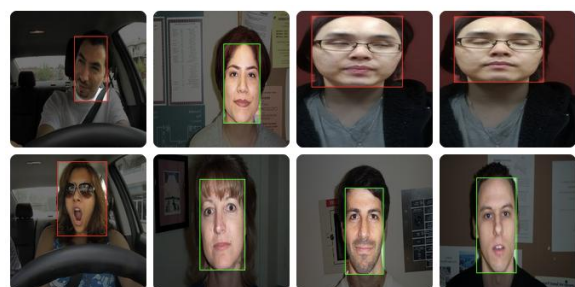


Gambar 1. Tahapan Penelitian

Tahapan penelitian pada Gambar 1 terdiri dari lima proses utama, yaitu pengumpulan dataset, pra-pemrosesan data, pelatihan model YOLOv26, pengujian deteksi, dan evaluasi model. Setiap tahapan memiliki peran penting dalam menghasilkan sistem deteksi fokus siswa yang mampu bekerja secara *real-time* dengan tingkat akurasi yang baik.

3.1. Pengumpulan Dataset

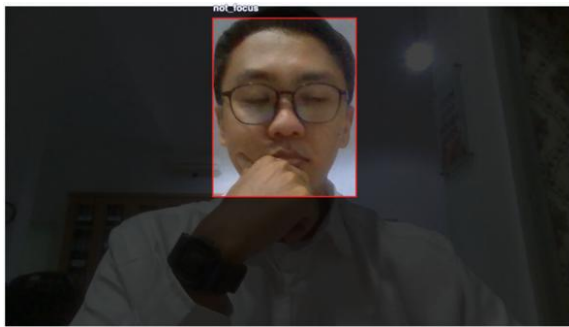
Tahap pertama penelitian dilakukan dengan mengumpulkan dataset citra aktivitas siswa selama proses pembelajaran berlangsung. Dataset diperoleh menggunakan platform Roboflow melalui project focus-detection-oztk8[16]. Data yang digunakan terdiri dari dua kelas utama, yaitu siswa fokus dan siswa tidak fokus. Pada Gambar 2. Berikut merupakan dataset yang digunakan:



Gambar 2. Dataset yang digunakan

3.2. Pra-Pemrosesan Data

Tahap pra-pemrosesan dilakukan untuk meningkatkan kualitas dataset sebelum digunakan pada proses pelatihan model. Pada tahap ini dilakukan beberapa proses seperti: Resize gambar, dan pemeriksaan anotasi. Selain itu dilakukan visualisasi distribusi dataset untuk mengetahui keseimbangan jumlah data pada setiap kelas. Tahap ini penting karena kualitas dataset sangat memengaruhi performa model deteksi objek.



Gambar 3. Anotasi Dataset

3.3. Pelatihan Model YOLOv26

Tahap berikutnya adalah pelatihan model menggunakan arsitektur YOLOv26. Model dilatih menggunakan dataset hasil preprocessing untuk mengenali objek siswa fokus dan tidak fokus secara otomatis. Pada penelitian ini proses pelatihan dilakukan menggunakan library Ultralytics dengan parameter yang disajikan pada Tabel 1. berikut:

Tabel 1. Parameter pelatihan

Parameter	Nilai
Epoch	100
Optimizer	AdamW
Learning rate	0.01
momentum	0.937
Weight decay	0.0005

YOLOv26 bekerja menggunakan metode single stage detector sehingga proses deteksi objek dilakukan dalam satu tahap inferensi. Model akan mempelajari pola visual dari dataset kemudian memprediksi lokasi objek dan kelas objek secara simultan.

3.4. Evaluasi Model

Evaluasi performa model untuk mengetahui tingkat keberhasilan sistem deteksi fokus siswa. Evaluasi dilakukan menggunakan beberapa parameter pengujian sesuai dengan persamaan (1)-(3), yaitu:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (1)$$

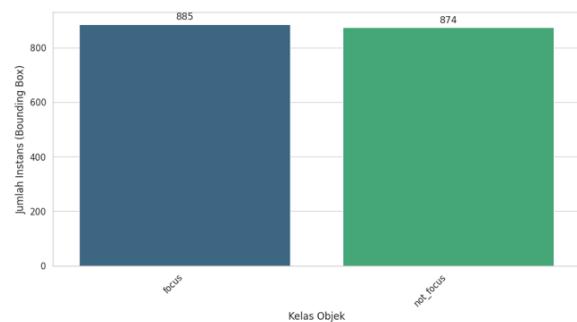
$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i \quad (3)$$

4. HASIL DAN PEMBAHASAN

Pada penelitian ini dilakukan implementasi arsitektur YOLOv26 untuk mendeteksi tingkat kefokus siswa selama proses pembelajaran berlangsung. Sistem dikembangkan menggunakan dataset citra aktivitas siswa yang terdiri dari dua kelas utama, yaitu fokus dan tidak fokus. Proses pelatihan dilakukan menggunakan framework Ultralytics YOLO dengan memanfaatkan GPU NVIDIA A100-SXM4-40GB MIG 3g.20gb untuk mempercepat proses komputasi model.

Dataset yang digunakan telah melalui tahap preprocessing dan augmentasi sebelum digunakan pada proses pelatihan. Data kemudian dibagi menjadi data training sebanyak 1759 gambar, validation sebanyak 200 gambar, dan testing sebanyak 100 gambar agar model mampu melakukan generalisasi dengan baik terhadap data baru. Pada Gambar 4. berikut visualisasi dari distribusi kelas pada dataset yang digunakan untuk pelatihan.



Gambar 4. Distribusi Dataset

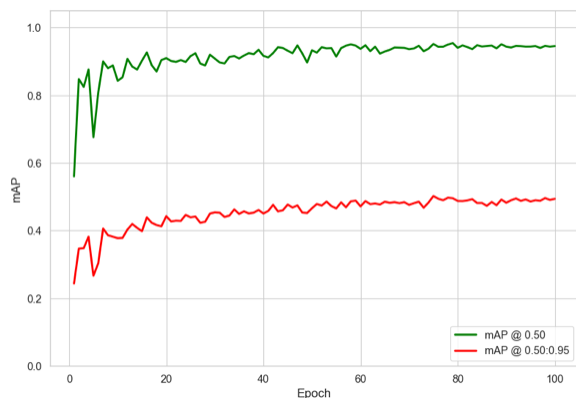
Berdasarkan Gambar 4, dataset yang digunakan terdiri dari dua kelas utama, yaitu kelas *fokus* dan *not_focus*. Jumlah data pada kelas *fokus* sebanyak 885 objek, sedangkan kelas *not_focus* sebanyak 874 objek. Hasil visualisasi menunjukkan bahwa distribusi dataset pada kedua kelas relatif seimbang. Selisih jumlah data antar kelas tidak terlalu signifikan sehingga dapat membantu model dalam mempelajari pola objek secara lebih stabil selama proses pelatihan. Dataset yang seimbang sangat penting pada proses klasifikasi karena dapat mengurangi risiko *bias* model terhadap salah satu kelas tertentu.

Proses pelatihan model YOLOv26 dilakukan menggunakan dataset fokus siswa yang telah melalui tahap preprocessing dan augmentasi data. Pelatihan dilakukan untuk mengetahui kemampuan model dalam mengenali objek siswa fokus dan tidak fokus selama proses pembelajaran berlangsung. Hasil pelatihan model diperoleh berdasarkan parameter precision, recall, dan mean Average Precision (mAP). Nilai evaluasi model ditunjukkan pada Tabel 2.

Tabel 2. Hasil pelatihan model

Kelas	Precision (%)	Recall (%)	mAP(%)
All	89.6	88.7	95.2
Focus	89.5	90.0	95.2
Not_Focus	89.7	87.5	95.2

Berdasarkan Tabel 2, secara keseluruhan, hasil pelatihan menunjukkan bahwa arsitektur YOLOv26 mampu digunakan untuk mendeteksi tingkat kefokus siswa secara real-time dengan performa yang cukup stabil. Model juga menunjukkan kemampuan generalisasi yang baik terhadap variasi data pembelajaran yang digunakan pada penelitian ini. Selain hasil evaluasi pada Tabel 2, performa pelatihan model juga dianalisis menggunakan grafik *mean Average Precision* (mAP) dan grafik *loss* selama proses training berlangsung. Visualisasi tersebut digunakan untuk mengetahui perkembangan kemampuan model dalam mendeteksi objek siswa fokus dan tidak fokus pada setiap epoch pelatihan:

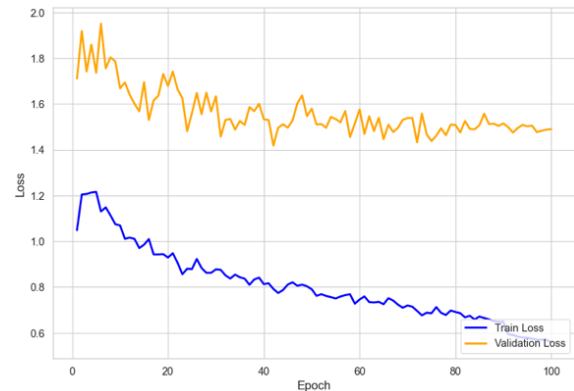


Gambar 5. Hasil mAP pelatihan

Berdasarkan Gambar 5, nilai $mAP@0.50$ mengalami peningkatan secara bertahap selama proses pelatihan berlangsung. Pada awal epoch, nilai mAP masih berada pada kisaran rendah karena model belum mampu mengenali pola objek secara optimal. Namun setelah beberapa epoch, nilai mAP meningkat secara signifikan dan mulai stabil mendekati nilai 0.95 pada akhir pelatihan. Peningkatan nilai mAP menunjukkan bahwa model YOLOv26 mampu mempelajari karakteristik objek siswa fokus dan tidak fokus dengan baik. Stabilitas grafik pada akhir pelatihan menunjukkan bahwa model telah mencapai kondisi konvergen dan mampu melakukan deteksi objek secara lebih akurat. Selain itu, nilai $mAP@0.50:0.95$ juga mengalami peningkatan secara konsisten hingga mencapai kisaran 0.50 pada akhir epoch. Hal tersebut menunjukkan bahwa kemampuan lokalisasi bounding box model cukup baik pada berbagai tingkat Intersection over Union (IoU).

Berdasarkan Gambar 6, nilai *train loss* mengalami penurunan secara bertahap selama proses pelatihan berlangsung. Pada awal epoch, nilai loss masih cukup tinggi karena model belum mampu mengenali pola data dengan baik. Seiring bertambahnya epoch, nilai *train loss* terus menurun hingga mencapai kondisi stabil pada akhir pelatihan. Sedangkan nilai *validation loss* juga menunjukkan tren penurunan meskipun mengalami fluktuasi pada beberapa epoch. Fluktuasi tersebut terjadi karena model diuji menggunakan data validation yang

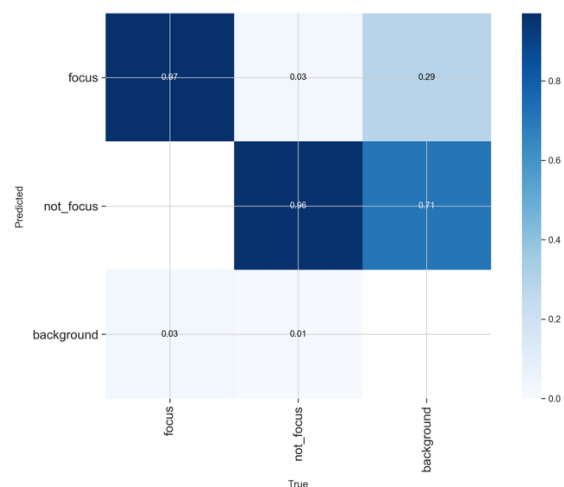
memiliki variasi kondisi berbeda dibanding data training. Namun secara keseluruhan nilai validation loss tetap berada pada kondisi stabil sehingga menunjukkan bahwa model tidak mengalami overfitting yang signifikan.



Gambar 6. Hasil Loss pelatihan

Penurunan nilai loss menunjukkan bahwa model semakin mampu meminimalkan kesalahan prediksi selama proses training berlangsung. Stabilitas grafik loss pada akhir pelatihan juga menunjukkan bahwa proses pembelajaran model berjalan dengan baik dan parameter model telah mencapai kondisi optimal. Berdasarkan hasil visualisasi grafik mAP dan loss, dapat disimpulkan bahwa arsitektur YOLOv26 memiliki performa pelatihan yang stabil dan mampu melakukan generalisasi dengan baik terhadap dataset fokus siswa yang digunakan pada penelitian ini.

Selain itu, evaluasi performa model dilakukan menggunakan *confusion matrix* untuk mengetahui tingkat keberhasilan klasifikasi pada masing-masing kelas. Confusion matrix digunakan untuk membandingkan hasil prediksi model dengan label sebenarnya pada data pengujian. Hasil confusion matrix ditunjukkan pada Gambar 7.



Gambar 7. Confusion Matrix

Berdasarkan Gambar 7, model YOLOv26 mampu melakukan klasifikasi objek dengan cukup baik pada kedua kelas, yaitu *focus* dan *not_focus*. Nilai diagonal pada confusion matrix menunjukkan tingkat prediksi benar yang dihasilkan model terhadap masing-masing kelas. Pada kelas *focus*, model memperoleh nilai prediksi benar sebesar 0.97 atau 97%. Hasil tersebut menunjukkan bahwa sebagian besar siswa yang termasuk kategori fokus berhasil dikenali dengan baik oleh sistem. Tingginya nilai klasifikasi pada kelas ini menunjukkan bahwa model mampu mengenali pola visual siswa yang memperhatikan proses pembelajaran secara konsisten. Sedangkan pada kelas *not_focus*, model memperoleh nilai prediksi benar sebesar 0.96 atau 96%. Hal tersebut menunjukkan bahwa model juga mampu mengenali siswa yang tidak fokus selama pembelajaran berlangsung dengan tingkat akurasi yang cukup tinggi.

Walaupun demikian, masih terdapat beberapa kesalahan klasifikasi pada kedua kelas. Pada kelas *focus* terdapat sekitar 0.03 atau 3% data yang salah diklasifikasikan sebagai *not_focus*. Kesalahan tersebut umumnya terjadi ketika siswa memiliki posisi kepala yang menunduk atau sebagian wajah tertutup sehingga pola visual objek menjadi sulit dikenali oleh model. Selain itu, pada kelas *not_focus* juga terdapat kesalahan prediksi sebesar 0.01 atau 1% terhadap kategori lain. Kesalahan ini dipengaruhi oleh kondisi pencahayaan ruangan, sudut pengambilan gambar, dan objek siswa yang saling bertumpukan (*occlusion*). Confusion matrix juga menunjukkan bahwa tingkat kesalahan deteksi terhadap *background* relatif kecil. Hal tersebut menunjukkan bahwa model mampu membedakan objek siswa dengan latar belakang gambar secara cukup baik sehingga proses deteksi menjadi lebih stabil.

Secara keseluruhan, hasil confusion matrix menunjukkan bahwa model YOLOv26 memiliki kemampuan klasifikasi yang baik dalam membedakan siswa fokus dan tidak fokus selama proses pembelajaran berlangsung. Tingginya nilai prediksi benar pada masing-masing kelas menunjukkan bahwa model mampu bekerja secara efektif pada lingkungan pembelajaran nyata.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, arsitektur YOLOv26 mampu digunakan untuk mendeteksi tingkat kefokuskan siswa selama proses pembelajaran secara real-time dengan performa yang cukup baik. Model berhasil mengenali dua kelas utama, yaitu *focus* dan *not_focus*, dengan nilai precision sebesar 89.6%, recall sebesar 88.7%, dan mAP sebesar 95.2%. Hasil pengujian menggunakan confusion matrix juga menunjukkan bahwa model mampu melakukan klasifikasi dengan tingkat prediksi benar sebesar 97% pada kelas *focus* dan 96% pada kelas *not_focus*. Selain itu, grafik mAP dan loss menunjukkan bahwa proses pelatihan berjalan stabil serta model mampu melakukan generalisasi dengan

baik terhadap variasi dataset yang digunakan. Walaupun demikian, masih terdapat beberapa kesalahan deteksi pada kondisi pencahayaan rendah, objek yang saling bertumpukan (*occlusion*), dan posisi wajah siswa yang tertutup sebagian. Oleh karena itu, penelitian selanjutnya dapat dikembangkan dengan menambahkan jumlah dataset yang lebih beragam, menggunakan pendekatan multi-camera, serta mengintegrasikan metode analisis perilaku lain seperti *head pose estimation* atau *eye tracking* untuk meningkatkan akurasi deteksi fokus siswa pada lingkungan pembelajaran yang lebih kompleks.

DAFTAR PUSTAKA

- [1] A. Alim Murtopo, M. Aditdy, P. Septiana Ananda, and G. Gunawan, "PENERAPAN COMPUTER VISION UNTUK MENDETEKSI KELENGKAPAN ATRIBUT SISWA MENGGUNAKAN METODE CNN," *PROSISKO J. Pengemb. Ris. Dan Obs. Sist. Komput.*, vol. 11, no. 2, pp. 247–258, Sep. 2024, doi: 10.30656/prosisko.v11i2.8752.
- [2] R. Rosalina, F. Bimantoro, and I. G. P. Suta Wijaya, "STUDENT FOCUS DETECTION USING YOU ONLY LOOK ONCE V5 (YOLOV5) ALGORITHM," *J. Tek. Inform. Jutif*, vol. 5, no. 5, pp. 1203–1211, May 2024, doi: 10.52436/1.jutif.2024.5.5.1977.
- [3] R. Pramudita, M. R. Fauzan, I. N. Faza, J. Kustija, I. Hartopo, and M. A. Rizqulloh, "Student Behavior Detection Using YOLOv10 for Classroom Engagement Analysis," vol. 15, no. 2, 2026.
- [4] C. Xing, "Classroom Attention Detection Using YOLOv8 and ResNet-based Facial Expression Recognition Model," *Appl. Comput. Eng.*, vol. 177, no. 1, pp. 144–149, Aug. 2025, doi: 10.54254/2755-2721/2025.BJ25725.
- [5] X. Sheng, S. Li, and S. Chan, "Real-time classroom student behavior detection based on improved YOLOv8s," *Sci. Rep.*, vol. 15, no. 1, p. 14470, Apr. 2025, doi: 10.1038/s41598-025-99243-x.
- [6] Y. Pan, G. Wang, and J. Yu, "Overview of deep learning YOLO algorithm," in *Fourth International Conference on Computer Vision, Application, and Algorithm (CVAA 2024)*, H. Yuan and L. Leng, Eds., Chengdu, China: SPIE, Jan. 2025, p. 3. doi: 10.1117/12.3055712.
- [7] C. Li and M. Long, "Hubei University of Technology," in *Student Engagement Across Pacific Asia*, L. K. Fryer, R. B. King, and L. M. Zeng, Eds., Singapore: Springer Nature Singapore, 2024, pp. 75–92. doi: 10.1007/978-981-97-0558-0_5.
- [8] Q. Jia and J. He, "Student Behavior Recognition in Classroom Based on Deep Learning," *Appl. Sci.*, vol. 14, no. 17, p. 7981, Sep. 2024, doi: 10.3390/app14177981.

- [9] D. Bao and W. Su, "Research on the Detection and Analysis of Students' Classroom Behavioral Features Based on Deep CNNs," *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, p. 3615865, Aug. 2023, doi: 10.1145/3615865.
- [10] L. Susanti, N. K. Daulay, and B. Intan, "Sistem Absensi Mahasiswa Berbasis Pengenalan Wajah Menggunakan Algoritma YOLOv5," *JURIKOM J. Ris. Komput.*, vol. 10, no. 2, p. 640, Apr. 2023, doi: 10.30865/jurikom.v10i2.6032.
- [11] Q. Jiao and X. Qi, "Deep Learning-Based Visual Attention and Learning Behavior Analysis in Interactive Educational Technology," *Deep Learn.*, 2024.
- [12] W. Wang, "Application of deep learning algorithm in detecting and analyzing classroom behavior of art teaching," *Syst. Soft Comput.*, vol. 6, p. 200082, Dec. 2024, doi: 10.1016/j.sasc.2024.200082.
- [13] K. Zhang, "Detection and analysis of student behavior in college labor education courses based on YOLOv5 network," *J. Comput. Methods Sci. Eng.*, vol. 24, no. 2, pp. 1057–1069, Mar. 2024, doi: 10.3233/JCM-247308.
- [14] A. Saryoko and F. Aziz, "YOLO MODEL DETECTION OF STUDENT NEATNESS BASED ON DEEP LEARNING: A SYSTEMIC LITERATURE REVIEW," *JITK J. Ilmu Pengetah. Dan Teknol. Komput.*, vol. 11, no. 2, pp. 454–460, Nov. 2025, doi: 10.33480/jitk.v11i2.6986.
- [15] X. Wang, Y. Gao, Y. Zhang, and L. Du, "Intelligent Classroom Student Behavior Recognition Based on Enhanced YOLOv8," in *Proceedings of the 2025 4th International Conference on Artificial Intelligence and Education*, Nanjing China: ACM, Nov. 2025, pp. 1136–1141. doi: 10.1145/3797552.3797731.
- [16] "Focus Detection." [Online]. Available: <https://universe.roboflow.com/focus-dataset/focus-detection-oztk8>