

Penerapan Algoritma YOLO V11 untuk Deteksi Tingkat Kekondusifan Pembelajaran di Kelas

1st Putri Chichilia Nababan
Fakultas Rekayasa Industri
Universitas Telkom
Bandung, Indonesia

putrichichilian@telkomuniversity.ac.id

2nd Nur Ichsan Utama, S.T., M.T.,
Ph.D.

Fakultas Rekayasa Industri
Universitas Telkom
Bandung, Indonesia

nichsan@telkomuniversity.ac.id

3rd Faqih Hamami, S. Kom., M.T.
Fakultas Rekayasa Industri
Universitas Telkom
Bandung, Indonesia

faqihhamami@telkomuniversity.ac.id

Permasalahan utama dalam proses pembelajaran tatap muka adalah kurangnya pemantauan objektif terhadap perilaku mahasiswa di dalam kelas, terutama terkait tingkat fokus selama kegiatan belajar berlangsung. Perilaku tidak fokus seperti tidur dan menggunakan ponsel kerap terjadi tanpa dapat terdeteksi secara real-time oleh pengajar, yang berdampak pada penurunan efektivitas pembelajaran. Untuk mengatasi hal ini, penelitian ini mengusulkan sistem pendeteksi perilaku mahasiswa berbasis algoritma YOLOv11, yang bertugas mendeteksi tiga jenis gerakan: tidur, menggunakan ponsel, dan memperhatikan. Data citra diperoleh melalui dokumentasi video di kelas, kemudian dilabeli secara manual dan dilatih menggunakan model YOLOv11 dengan parameter utama 50 *epoch*, ukuran gambar 640 piksel, dan *batch size* 16. Model yang dikembangkan dievaluasi menggunakan metrik evaluasi seperti *accuracy*, *precision*, *recall*, *F1-score*, dan *mean Average Precision* (mAP). Hasil evaluasi menunjukkan nilai *precision accuracy* sebesar 0.9652 sebesar 0.9457, *recall* sebesar 0.9575, *F1-score* sebesar 0.9515, mAP@50 sebesar 0.9672, dan akurasi sistem mencapai 0.9662. Output dari sistem berupa file Excel (.xlsx) yang memuat *timestamp*, jenis perilaku yang terdeteksi, dan *confidence score*, serta proporsi persentase perilaku fokus dan tidak fokus dalam kelas. Berdasarkan hasil tersebut, dapat disimpulkan bahwa YOLOv11 dapat diterapkan secara efektif dalam mendeteksi dan mengklasifikasikan perilaku mahasiswa, serta memberikan gambaran yang objektif mengenai tingkat kekondusifan pembelajaran di kelas.

kata kunci: YOLOv11, deteksi perilaku, *computer vision*, fokus belajar, *deep learning*.

I. PENDAHULUAN

Kemajuan teknologi telah merevolusi metode pembelajaran, menuntut pendekatan yang adaptif dan berpusat pada siswa. Fokus belajar merupakan indikator penting dalam keberhasilan pendidikan. Namun, gangguan internal seperti rasa kantuk dan penggunaan ponsel, serta keterbatasan pengajar dalam memantau secara manual, menjadi hambatan tersendiri. Disisi lain tantangan yang dihadapi pengajar dalam memantau keterlibatan siswa secara menyeluruh, terutama di kelas dengan jumlah siswa yang besar[1]. Dengan berkembangnya teknologi computer vision, hadir potensi untuk membantu pengajar dalam memantau keterlibatan siswa secara objektif dan *real-time*. YOLO (*You*

Only Look Once) sebagai algoritma deteksi objek *real-time* menjadi solusi potensial. Penelitian ini bertujuan untuk mengimplementasikan YOLOv11 guna mendeteksi perilaku tidak fokus di kelas dan mengidentifikasi tingkat kekondusifan kelas melalui deteksi visual otomatis.

II. KAJIAN TEORI

Menyajikan dan menjelaskan teori-teori yang berkaitan dengan variabel-variabel penelitian. Poin subjudul ditulis dalam abjad.

A. Fokus dalam Konteks Pembelajaran

Fokus belajar mencerminkan kemampuan siswa dalam mempertahankan perhatian selama proses pembelajaran. Gangguan terhadap fokus dapat menghambat pemahaman dan hasil belajar. Penelitian lebih lanjut menunjukkan bahwa fokus dapat dipengaruhi oleh faktor internal dan eksternal. Misalnya, siswa yang berada dalam lingkungan belajar yang tenang dan kondusif cenderung lebih mudah mempertahankan fokus, sedangkan kebisingan atau gangguan visual di sekitar kelas dapat menurunkan tingkat konsentrasi mereka [2].

B. Computer Vision

Computer Vision merupakan salah satu bidang dalam ilmu komputer yang berfokus pada pengembangan sistem yang mampu secara otomatis mengamati, mengidentifikasi, dan memahami informasi visual dari gambar maupun video. Tujuan utamanya adalah memungkinkan komputer atau mesin untuk menafsirkan serta memahami lingkungan visual di dunia nyata layaknya kemampuan manusia dalam melihat [3]

C. YOLOv11

YOLO (*You Only Look Once*) merupakan salah satu metode deteksi objek yang banyak digunakan karena kemampuannya dalam melakukan pemrosesan secara waktu nyata (*real-time*). Algoritma ini mampu memproses gambar dengan kecepatan hingga 45 *frame per second*. YOLO termasuk ke dalam kategori *one-stage detector*, yaitu pendekatan deteksi objek yang langsung melakukan prediksi kelas dan lokasi objek dalam satu proses terpadu [4].

D. Knowledge Discovery in Database (KDD)

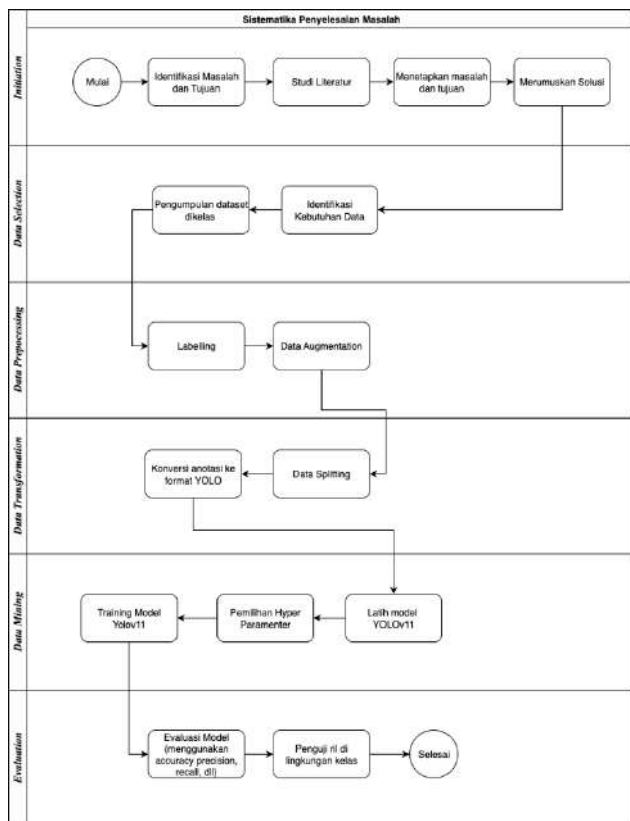
Knowledge Discovery in Database (KDD) merupakan suatu pendekatan yang digunakan untuk mengekstraksi informasi dari data historis guna mengidentifikasi pola dan hubungan yang tersembunyi dalam suatu dataset. Informasi yang diperoleh dari proses ini dapat dimanfaatkan sebagai dasar dalam pengambilan keputusan di masa mendatang [5].

E. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan bentuk lanjutan dari *Multilayer Perceptron* (MLP) yang tergolong dalam jaringan saraf tipe *feed forward*, yaitu jaringan tanpa umpan balik. CNN dirancang khusus untuk memproses data berdimensi dua, seperti gambar. Karena memiliki arsitektur yang mendalam, CNN dikategorikan sebagai bagian dari *Deep Neural Network* dan banyak digunakan dalam pengolahan data citra [6].

III. METODE

Sistematika penyelesaian masalah dalam penelitian ini menggunakan kerangka kerja *Knowledge Discovery in Database* (KDD).



GAMBAR 1
(Sistem Penyelesaian Masalah)

A. Initiation

Tahapan ini merupakan tahap awal dalam pelaksanaan penelitian, yang bertujuan untuk memahami konteks permasalahan dan menentukan arah solusi. Dimulai dengan identifikasi masalah dan tujuan dari penelitian, kemudian dilakukan studi literatur untuk memperkuat pemahaman terhadap topik yang diangkat, terutama tentang penerapan YOLOv11 dalam deteksi perilaku mahasiswa. Setelah itu,

masalah dan tujuan dirumuskan secara lebih terarah agar dapat diturunkan ke dalam solusi teknis yang sesuai.

B. Data Selection

Pada tahap ini dilakukan proses identifikasi kebutuhan data yang dibutuhkan untuk menyelesaikan permasalahan yang telah dirumuskan. Data yang digunakan merupakan gambar atau citra yang dikumpulkan dari lingkungan kelas secara langsung. Pengumpulan data dilakukan dengan memperhatikan variasi sudut pandang, pencahayaan, dan aktivitas mahasiswa agar model dapat belajar dari representasi visual yang beragam. Proses ini juga mencakup pemilihan jenis aktivitas yang akan diklasifikasikan (tidur, menggunakan ponsel, memperhatikan), serta penyesuaian jumlah data untuk masing-masing kelas..

C. Data Preprocessing

Setelah data terkumpul, tahap selanjutnya adalah preprocessing, yaitu proses penyiapan data agar siap digunakan dalam pelatihan model. Dalam tahap ini, gambar dilabeli secara manual menggunakan aplikasi Roboflow Annotate untuk menandai area objek dan kategori perilaku yang ditampilkan. Setelah pelabelan, dilakukan data augmentation untuk memperbanyak dan memperkaya dataset dengan memodifikasi gambar (misalnya rotasi, pencahayaan, zoom) guna meningkatkan generalisasi model. Tujuannya agar model tidak hanya belajar dari kondisi statis tetapi juga mampu mengenali pola dalam kondisi bervariasi.

D. Data Transformation

Tahapan ini merupakan proses konversi data hasil pelabelan menjadi format yang dapat digunakan oleh algoritma YOLOv11. Format YOLO memerlukan struktur khusus dalam bentuk file .txt yang memuat informasi *bounding box* dan label dalam skala normalisasi. Selain itu, pada tahap ini juga dilakukan data splitting, yaitu pembagian data ke dalam subset training, validation, dan testing (biasanya dengan rasio 80:10:10) untuk memisahkan data yang digunakan saat pelatihan dan data yang digunakan saat evaluasi agar hasilnya lebih objektif.

E. Data Mining

Pada tahap ini dilakukan proses pelatihan model YOLOv11 menggunakan dataset yang telah disiapkan. Proses training dilakukan di Google Colab dengan parameter yang telah ditentukan (misalnya jumlah *epoch*, ukuran gambar, *batch size*, dan *learning rate*). Selain itu, juga dilakukan pemilihan dan penyesuaian hyperparameter untuk mencapai performa optimal. Tujuan utama tahap ini adalah membangun model yang mampu mengenali tiga kategori perilaku (tidur, memperhatikan, menggunakan ponsel) secara otomatis melalui input gambar.

F. Evaluation

Tahap terakhir adalah evaluasi model untuk menilai performa sistem. Evaluasi dilakukan dengan menggunakan metrik evaluasi seperti *accuracy*, *precision*, *recall*, *F1-score*, *mAP@50*, dan *mAP@50-95*. Selain itu, dilakukan juga pengujian sistem di lingkungan kelas nyata untuk melihat efektivitas implementasinya secara langsung. Output akhir dari sistem berupa file .xlsx yang berisi hasil prediksi

(timestamp, jenis perilaku, confidence score), serta perhitungan persentase fokus dan tidak fokus yang digunakan untuk menilai tingkat kekondusifan kelas.

IV. HASIL DAN PEMBAHASAN

Bagian ini berisi paparan objektif peneliti terhadap hasil-hasil penelitian berupa penjelasan dan analisis terhadap penemuan-penemuan penelitian, penjelasan serta penafsiran dari data dan hubungan yang diperoleh, serta pembuatan generalisasi dari penemuan. Apabila terdapat hipotesis, maka pada bagian ini juga menjelaskan proses pengujian hipotesis beserta hasilnya.

Hasil penelitian harus disajikan secara jelas dan sistematis supaya mudah dibaca dan dipahami. Penyajian hasil penelitian dapat dilakukan dengan cara deskriptif (naratif), menggunakan tabulasi, tabel atau grafik, atau dengan menggunakan gabungan dua atau ketiganya secara sekaligus. Penggunaan ketiga cara tersebut disesuaikan dengan jenis data dan sejauh mana deskripsi data akan dijelaskan. Misalnya, pada awal peneliti memaparkan narasi temuannya, kemudian didukung dengan sajian data dalam bentuk tabulasi, tabel atau grafik. Peneliti juga menyajikan data-data hasil penelitian, kemudian didukung grafik dilanjutkan deskripsi naratif [10 pts]. Berikan kemungkinan pengembangan atau penelitian ke depan terkait penelitian ini

A. Pengujian di kelas nyata



GAMBAR 2
(Pengujian di kelas nyata)

B. Output hasil

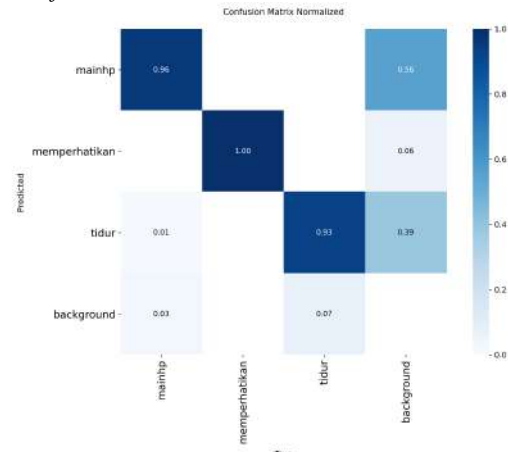
563	2025-06-25 09:45:27	memperhatikan	0,863360639
564	2025-06-25 09:45:28	main hp	0,965924209
565	2025-06-25 09:45:28	tidur	0,858285481
566	2025-06-25 09:45:28	tidur	0,941256412
567	2025-06-25 09:45:28	tidur	0,640668603
568	2025-06-25 09:45:28	memperhatikan	0,559492171
569	2025-06-25 09:45:28	memperhatikan	0,938974667
570	2025-06-25 09:45:28	main hp	0,825299314
571	2025-06-25 09:45:28	main hp	0,891963904
572	2025-06-25 09:45:28	memperhatikan	0,835838021
573	2025-06-25 09:45:28	memperhatikan	0,836210168
574		TOTAL TIDUR	125
575		TOTAL MAIN_HP	135
576		TOTAL NORMAL	312
577		TOTAL KONDUSIF	312
578		TOTAL TIDAK KONDUSIF	260
579		PERSEN KONDUSIF (%)	0,545454545
580		PERSEN TIDAK KONDUSIF (%)	0,454545455
581		STATUS DOMINAN	KONDUSIF

GAMBAR 3

(Output hasil)

Berdasarkan tabel hasil pengujian tersebut dapat diketahui kondisi kelas dapat dikatakan kondusif terhitung dari perhitungan perilaku kondusif yaitu 54% dihitung dari jumlah gerakan memperhatikan dan perilaku tidak kondusif yaitu 45% dihitung dari jumlah gerakan tidur dan menggunakan ponsel sehingga dikalkulasikan menjadi tidak kondusif.

C. Confusion Matrix



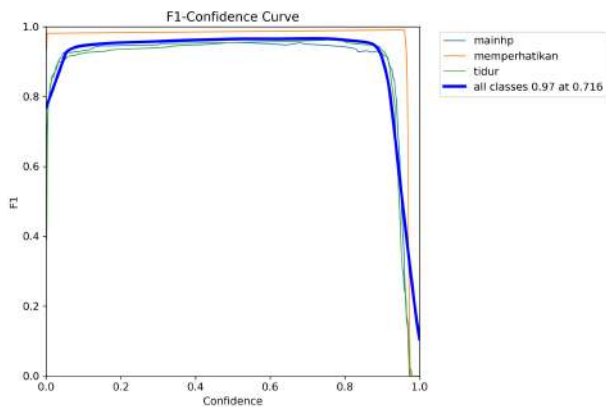
GAMBAR 4
(Confusion Matrix)

Gambar di atas menunjukkan *Confusion Matrix* hasil pengujian model YOLOv11 terhadap empat kelas: mainhp, memperhatikan, tidur, dan background. Model menunjukkan akurasi tinggi pada kelas memperhatikan (1.00), mainhp (0.96), dan tidur (0.93). Namun, kelas background banyak salah diklasifikasikan sebagai mainhp (0.56) dan tidur (0.39), yang menunjukkan bahwa model masih kesulitan membedakan latar. Secara umum, model efektif mengenali gestur utama, meskipun diperlukan perbaikan pada deteksi background untuk meningkatkan akurasi keseluruhan.

D. Accuracy

Dalam penelitian ini, hasil pengujian menunjukkan bahwa model YOLOv11 yang digunakan mampu mencapai tingkat akurasi sebesar 96.62%. Nilai ini diperoleh dari perbandingan antara jumlah prediksi yang benar dan keseluruhan data yang diuji, dengan rincian TP = 922, TN = 910, FP = 38, dan FN = 28. Perolehan *accuracy* sebesar 96.62% mengindikasikan bahwa model mampu mengklasifikasikan perilaku mahasiswa baik fokus maupun tidak fokus dengan tingkat ketepatan yang sangat tinggi. Hasil ini memperkuat validitas model dalam konteks deteksi visual perilaku di lingkungan pembelajaran.

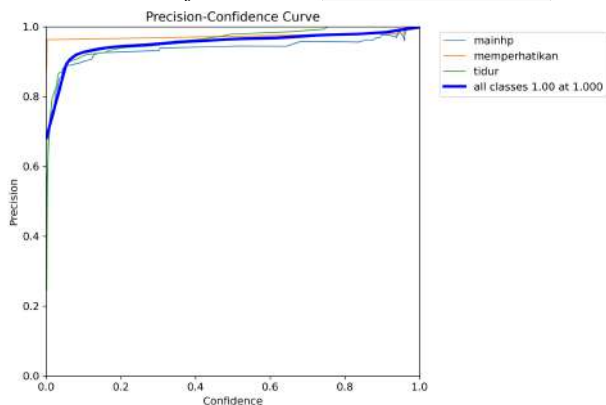
E. F1 Confidence Curve



GAMBAR 5
(F1 Confidence Curve)

Gambar di atas menunjukkan *F1-Confidence Curve* yang mengilustrasikan hubungan antara nilai confidence threshold dan *F1-score* untuk masing-masing kelas: mainhp, memperhatikan, dan tidur, serta total gabungan (*all classes*). *F1-score* adalah metrik harmonik dari *precision* dan *recall*, ideal digunakan saat distribusi kelas tidak seimbang. Kurva menunjukkan bahwa performa terbaik dicapai pada *threshold confidence* sekitar 0.716 dengan *F1-score* gabungan sebesar 0.97. Kelas memperhatikan tampil paling stabil dan tinggi, menunjukkan deteksi yang konsisten. Sementara itu, kelas mainhp dan tidur mengalami penurunan *F1* pada *threshold* di atas 0.9 akibat berkurangnya jumlah prediksi valid. Penurunan tajam di *confidence* mendekati 1.0 terjadi karena *recall* menurun drastis saat prediksi sangat dibatasi, sehingga memengaruhi skor *F1* secara keseluruhan.

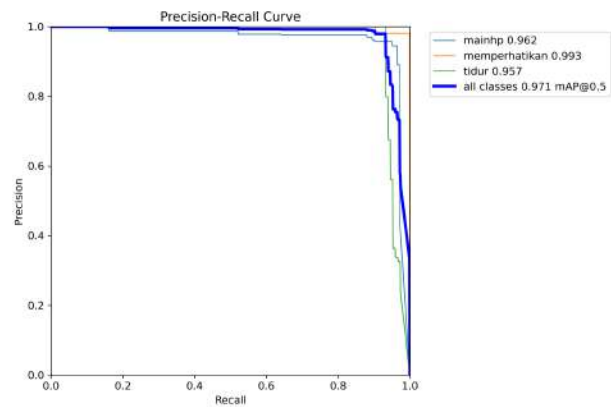
F. Precision Confidence Curve



GAMBAR 6
(Precision Confidence Curve)

Grafik di atas menunjukkan *Precision Confidence Curve* yang menggambarkan hubungan antara ambang kepercayaan (*confidence threshold*) dan tingkat presisi untuk tiga kelas: mainhp, memperhatikan, dan tidur. Terlihat bahwa presisi meningkat seiring naiknya *confidence*, dengan nilai maksimum mendekati 1.0 pada *threshold* tinggi. Kelas memperhatikan menunjukkan presisi yang stabil dan tinggi, sementara mainhp dan tidur sedikit berfluktuasi di *confidence* rendah. Kurva gabungan (*all classes*) mencapai presisi sempurna pada *confidence* 1.0, meskipun dengan jumlah prediksi yang lebih sedikit.

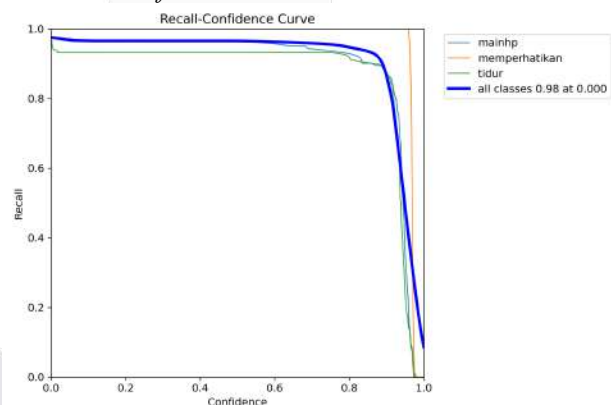
G. Precision Recall Curve



GAMBAR 7
(Precision Recall Curve)

Gambar di atas menunjukkan *Precision-Recall Curve* untuk tiga kelas: mainhp, memperhatikan, dan tidur, dengan garis biru tebal mewakili gabungan seluruh kelas. Grafik ini menunjukkan bahwa model memiliki performa deteksi yang sangat baik, ditandai dengan nilai *mAP@0.5* sebesar 0.971. Kelas memperhatikan memiliki performa tertinggi dengan *mAP* 0.993, diikuti oleh mainhp (0.962) dan tidur (0.957). Kurva yang mendekati sudut kanan atas menandakan presisi dan *recall* tinggi secara bersamaan. Sedikit penurunan presisi terjadi saat *recall* mendekati maksimum, yang umum terjadi pada model klasifikasi.

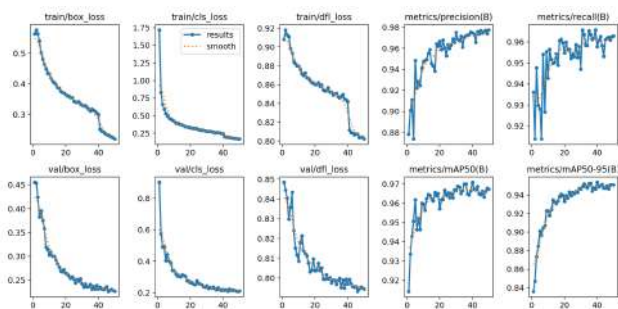
H. Recall Confidence Curve



GAMBAR 8
(Recall Confidence Curve)

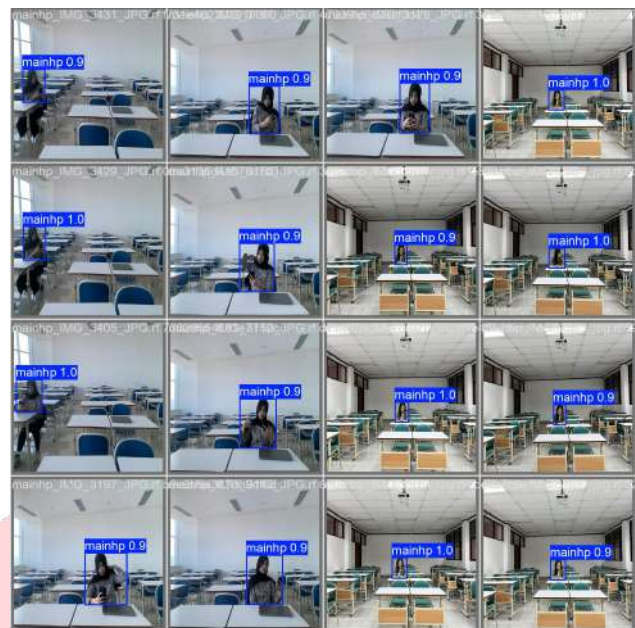
Gambar di atas memperlihatkan *Recall-Confidence Curve*, yang menunjukkan hubungan antara tingkat keyakinan (*confidence*) model dan sensitivitasnya (*recall*) dalam mendeteksi objek. Pada *confidence* rendah (0.00), *recall* mencapai hingga 0.98, menandakan kemampuan tinggi dalam menangkap objek yang relevan. Namun, seiring meningkatnya *confidence*, nilai *recall* menurun karena model menjadi lebih selektif. Kelas memperhatikan memiliki performa paling stabil, sedangkan mainhp dan tidur menunjukkan sedikit fluktuasi. Grafik ini membantu menentukan ambang *confidence* yang optimal untuk menjaga keseimbangan antara sensitivitas dan ketepatan deteksi.

I. Training and Validation Loss



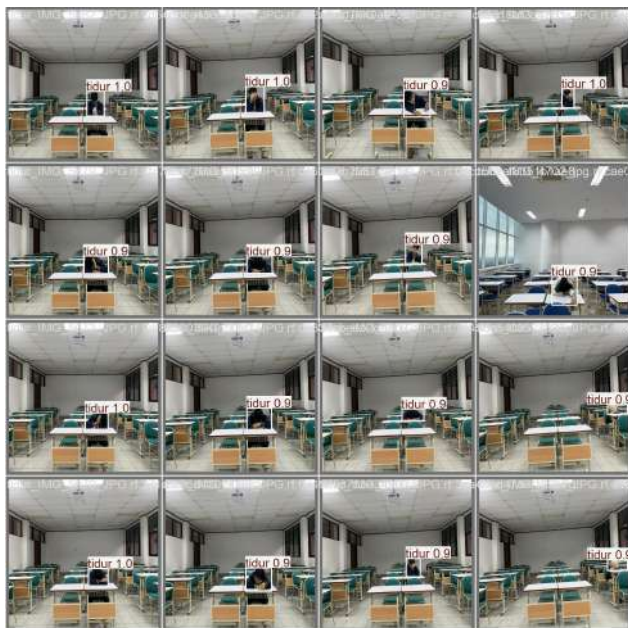
GAMBAR 9
(Training and Validation Loss)

Gambar di atas menampilkan grafik hasil pelatihan dan validasi model YOLOv11 selama 50 *epoch*. Nilai *box loss*, *classification loss*, dan *distribution focal loss* pada data pelatihan dan validasi menunjukkan tren penurunan yang stabil, menandakan proses pembelajaran yang efektif tanpa indikasi *overfitting*. Selain itu, metrik evaluasi seperti *precision*, *recall*, serta *mAP@50* dan *mAP@50-95* menunjukkan peningkatan konsisten, baik pada data pelatihan maupun validasi. Hal ini mencerminkan bahwa model mampu mendeteksi dan mengklasifikasikan gerakan siswa dengan akurasi yang baik dan generalisasi yang kuat.



GAMBAR 11
(Hasil prediksi menggunakan ponsel)

J. Hasil Prediksi



GAMBAR 10
(Hasil prediksi tidur)

Gambar dinomori secara berurutan. Letak penulisannya di bawah gambar disertai dengan penjelasan. Contoh: Gambar 1(A)

V. KESIMPULAN

Berdasarkan hasil penelitian, algoritma YOLOv11 terbukti efektif dalam mendeteksi dan mengklasifikasikan perilaku mahasiswa di kelas, khususnya tidur, menggunakan ponsel, dan memperhatikan. Model menunjukkan performa tinggi dengan akurasi 0.9486, *precision* 0.9457, *recall* 0.9575, *F1-score* 0.9516, *mAP@50* sebesar 0.971, dan *mAP@50-95* sebesar 0.767. Sistem menghasilkan output dalam format Excel (.xlsx) berisi timestamp, jenis perilaku, dan *confidence score*, serta menghitung proporsi fokus dan tidak fokus sebagai indikator kekondusifan kelas. Penerapan YOLOv11 ini berpotensi membantu dosen memantau dan mengevaluasi kelas secara lebih objektif dan berbasis data.

REFERENSI

- [1] E. Bozkir, C. Kosel, T. Seidel, dan E. Kasneci, "Automated Visual Attention Detection using Mobile Eye Tracking in Behavioral Classroom Studies," Mei 2025, [Daring]. Tersedia pada: <http://arxiv.org/abs/2505.07552>
- [2] A. Amrulloh, N. darajaatul Aliyah, dan D. Darmawan, "Pengaruh Kebiasaan Belajar, Lingkungan Belajar Dan Motivasi Belajar Terhadap Prestasi Belajar Siswa MTS Darul Hikmah Langkap Burneh Bangkalan," *AL-MIKRAJ Jurnal Studi Islam dan Humaniora (E-ISSN 2745-4584)*, vol. 5, no. 01, hlm. 188–200, Jul 2024, doi: 10.37680/almikraj.v5i01.5656.
- [3] T. Sutisna, A. Rachmat Raharja, E. Hariyadi, dan V. Hafizh Cahaya Putra, "Penggunaan Computer Vision

untuk Menghitung Jumlah Kendaraan dengan Menggunakan Metode SSD (Single Shoot Detector),” *INNOVATIVE: Journal Of Social Science Research*, vol. 4, hlm. 6060–6067, 2024.

- [4] D. I. Mulyana dan R. Ferdiansyah Putra, “Tinggi Ilmu Komputer Cipta Karya Informatika,” *Jurnal Teknologi Informasi dan Komunikasi*, vol. 8, no. 1, 2024, doi: 10.35870/jti.
- [5] F. Meila Azzahra Sofyan, A. Putri Riyandoro, D. Fitriani Maulana, J. Haerul Jaman, S. Informasi, dan U. Singaperbangsa Karawang, “Penerapan Data Mining dengan Algoritma C5.0 Untuk Prediksi

Penyakit Stroke,” *Juli*, 2023, [Daring]. Tersedia pada:

- [6] <https://ojs.trigunadharma.ac.id/index.php/jsk/index>
P. Adi Nugroho, I. Fenriana, dan R. Arijanto, “Implementasi Deep Learning Menggunakan Convolutional Neural Network (Cnn) Pada Ekspresi Manusia,” *JURNAL ALGOR*, vol. 2, no. 1, 2020, [Daring]. Tersedia pada: <https://jurnal.buddhidharma.ac.id/index.php/algor/index>

