Deteksi Jenis Alas Kaki MenggunakanYOLOV8 Berbasis Citra Webcam

1st Dimash Altairs Hidayatulloh
Program Studi S1 Informatika
Universitas Telkom Surabaya
Kota Surabaya, Indonesia
dimashalta@student.telkomuniversity.ac.id

2nd Ahmad Wali Satria Bahari Johan. S.ST., M.Kom.,
Universitas Telkom Surabaya
Kota Surabaya, Indonesia
ahmadsatria@telkomuniversity.ac.id

3rd Dr. Tanzilal Mustaqim. S.Kom., M.Kom.,
Universitas Telkom Surabaya
Kota Surabaya, Indonesia
tanzilal@telkomuniversity.ac.id

***Abstrak* — Pengawasan kepatuhan penggunaan alas kaki di lingkungan kampus masih dilakukan secara manual sehingga kurang efektif pada area dengan mobilitas mahasiswa yang tinggi. Penelitian ini mengembangkan sistem deteksi alas kaki berbasis computer vision menggunakan model YOLOv8n untuk membedakan objek sepatu dan sandal secara otomatis melalui citra webcam. Dataset diperoleh dari kombinasi data primer dan sekunder yang kemudian melalui proses anotasi, augmentasi, dan pelatihan model. Sistem diimplementasikan dalam bentuk aplikasi web yang mendukung deteksi melalui unggahan gambar dan kamera secara langsung. Evaluasi dilakukan menggunakan metrik Precision, Recall, mAP50, dan mAP50-95. Hasil pengujian menunjukkan model mampu mencapai nilai mAP50 sekitar 0,85 dengan performa terbaik pada kelas sepatu. Selain itu, penelitian menemukan bahwa konsistensi anotasi dan kualitas dataset memberikan pengaruh signifikan terhadap peningkatan performa model selama proses pelatihan. Hasil penelitian menunjukkan bahwa YOLOv8n dapat diterapkan sebagai solusi monitoring otomatis penggunaan alas kaki pada lingkungan kampus dengan kebutuhan komputasi yang relatif ringan.**

***Kata kunci*— deteksi objek, alas kaki, YOLOv8, *computer vision*, citra *webcam*.**

I. PENDAHULUAN

Lingkungan kampus merupakan ruang publik akademik yang menuntut adanya tata tertib dan kedisiplinan dari seluruh mahasiswa. Salah satu bentuk aturan yang diterapkan adalah ketentuan penggunaan alas kaki yang sesuai dengan norma dan etika akademik. Di Universitas Telkom, ketentuan tersebut telah diatur secara resmi dalam peraturan akademik dan kode etik mahasiswa yang melarang penggunaan sandal serta jenis alas kaki tertentu di lingkungan kampus [1]. Meskipun aturan ini penting untuk menjaga kerapian, ketertiban, dan citra dari institusi itu sendiri, penerapannya di lapangan masih menghadapi kendala karena proses pengawasan umumnya dilakukan secara manual oleh petugas tertentu. Pada area dengan kepadatan mahasiswa yang tinggi, seperti lobi gedung dan pintu masuk fasilitas kampus, keterbatasan jumlah personel dan luasnya area pengawasan dapat menyebabkan pemantauan menjadi kurang konsisten [1].

Perkembangan *computer vision* dan *deep learning* membuka peluang untuk mendukung proses pemantauan secara otomatis melalui system deteksi objek berbasis citra ataupun video. Salah satu algoritma yang banyak digunakan dalam deteksi objek secara *real-time* adalah YOLO (*You Only Look Once*), yang dikenal efisien karena mampu melakukan deteksi dalam satu tahap pemrosesan [2]. Pada versi YOLOv8 Ultralytics menghadirkan versi yang efisiensi komputasi, termasuk pada varian nya YOLOv8n yang dirancang agar tetap ringan dan sesuai untuk perangkat dengan sumber daya yang terbatas [3]. Karakteristik tersebut menjadikan YOLOv8n relevan untuk diterapkan pada system deteksi jenis alas kaki berbasis citra *webcam* yang membutuhkan proses inferensi cepat dan responsif.

Sejumlah penelitian terdahulu menunjukkan bahwa pengembangan YOLOv8 pada berbagai konteks mampu memberikan performa yang baik, khususnya pada deteksi objek kecil dan kondisi visual yang kompleks [4] [5] [6] [7] [8]. Namun, penerapan pada objek alas kaki di lingkungan kampus memiliki tantangan tersendiri karena objek sepatu dan sandal memiliki bentuk yang mirip, sering berukuran relatif kecil dalam citra, serta dipengaruhi kondisi pencahayaan dan sudut pengambilan gambar. Selain itu, kualitas anotasi dan keseimbangan dataset juga berpengaruh besar terhadap hasil pelatihan model, sehingga aspek ini perlu diperhatikan sejak tahap pengumpulan data hingga evaluasi [4]- [8]. Berdasarkan hal tersebut, penelitian ini berfokus pada perancangan sistem deteksi jenis alas kaki yang mampu membedakan sepatu dan sandal secara otomatis menggunakan YOLOv8n berbasis citra *webcam* di lingkungan kampus, sekaligus mengevaluasi performanya menggunakan metrik *precision*, *recall*, mAP50, dan mAP50-95.

II. KAJIAN TEORI

Bagian ini memaparkan landasan teoritis yang melandasi penelitian deteksi jenis alas kaki menggunakan YOLOv8 berbasis citra *webcam*.

A. *Computer Vision*

Computer vision merupakan cabang kecerdasan buatan yang berfokus pada kemampuan sistem untuk memahami dan menginterpretasikan informasi visual dari gambar maupun video. Dalam penelitian ini, *computer vision* menjadi dasar untuk mengenali jenis alas kaki secara otomatis dari citra *webcam*. Pendekatan ini penting karena sistem tidak hanya membaca citra sebagai data mentah, tetapi juga mengekstraksi fitur visual yang dapat membedakan objek sepatu dan sandal [9].

B. Deteksi Objek

Deteksi objek adalah proses mengenali keberadaan objek sekaligus menentukan lokasi objek tersebut di dalam citra. Berbeda dengan klasifikasi citra, deteksi objek menghasilkan *bounding box* sebagai penanda posisi objek. Dalam penelitian ini, deteksi objek digunakan untuk mengidentifikasi dua kelas utama, yaitu sepatu dan sandal [10] [5].

C. YOLOv8

YOLO (*You Only Look Once*) merupakan pendekatan *one-stage detector* yang memproses deteksi objek dalam satu tahap sehingga efisien untuk aplikasi *real-time*. YOLOv8 dipilih karena menawarkan kombinasi akurasi dan kecepatan yang baik, serta mendukung varian ringan seperti YOLOv8n yang sesuai untuk perangkat dengan keterbatasan komputasi. Dalam penelitian ini, YOLOv8n digunakan sebagai *baseline* model untuk deteksi alas kaki berbasis citra *webcam* [3].

D. Dataset, Anotasi Data, dan *Bounding Box*

Dataset merupakan komponen utama dalam pelatihan model deteksi objek karena performa model sangat bergantung pada kualitas, variasi, dan keseimbangan data. Pada penelitian ini, dataset terdiri dari gabungan data primer dan sekunder yang dianotasi menggunakan *bounding box* sebelum diproses lebih lanjut. Anotasi yang konsisten penting karena *bounding box* yang tidak seragam dapat menurunkan kualitas pembelajaran model dan memengaruhi IoU, *precision*, serta *recall* [4] [5].

E. *Preprocessing* dan Augmentasi Data

Preprocessing dilakukan untuk menyeragamkan citra sebelum masuk ke model, misalnya *resize*, normalisasi, dan *letterboxing*. Augmentasi data digunakan untuk menambah variasi dataset tanpa menambah data baru secara langsung, misalnya melalui *flip*, rotasi, atau penyesuaian warna. Tahap ini penting karena objek alas kaki pada citra *webcam* sering muncul dalam kondisi yang bervariasi, sehingga model perlu dilatih agar lebih robust terhadap perubahan sudut, pencahayaan, dan posisi objek [7] [6] [8].

F. Metrik Evaluasi

Performa model deteksi objek dievaluasi menggunakan *precision*, *recall*, mAP50, dan mAP50-95. *Precision* menunjukkan ketepatan prediksi model, *recall* menunjukkan kemampuan model menemukan objek yang benar, sedangkan mAP digunakan untuk menilai performa deteksi secara keseluruhan pada berbagai ambang IoU. Metrik ini lebih sesuai dibanding *accuracy* karena deteksi objek melibatkan lokasi *bounding box* dan lebih dari satu objek dalam satu citra [5].

F. Penelitian Terdahulu

Penelitian terdahulu menunjukkan bahwa modifikasi YOLO pada berbagai domain dapat meningkatkan performa deteksi objek, khususnya pada objek kecil, kondisi visual kompleks, dan data yang tidak seimbang. Zhao et al. membahas masalah *class imbalance* dan *Focal Loss*; Shen et al. menunjukkan peningkatan performa melalui modifikasi *backbone* YOLOv8; Wang et al. menambahkan pendekatan multi-skala; Zhou et al. menekankan pengaruh augmentasi dan *attention*; dan Yi et al. menunjukkan pentingnya *feature fusion* untuk objek kecil. Kelima studi ini menjadi dasar bagi penelitian ini untuk menerapkan YOLOv8n pada deteksi alas kaki di lingkungan kampus [4] [5] [6] [7] [8].

III. METODE

Metodologi penelitian ini dirancang secara sistematis untuk membangun dan mengevaluasi model deteksi jenis alas kaki menggunakan YOLOv8n. Tahapan penelitian meliputi pengumpulan dataset, anotasi, *preprocessing*, augmentasi, pelatihan model, dan pengujian sistem.

A. Desain Penelitian

Penelitian ini menggunakan pendekatan eksperimen pada sistem deteksi objek berbasis *computer vision* untuk membedakan dua kelas alas kaki, yaitu sepatu dan sandal. Alur penelitian dimulai dari pengumpulan dataset, anotasi *bounding box*, *preprocessing* dan augmentasi data, pelatihan model YOLOv8n, hinga pengujian hasil deteksi pada citra *webcam* dan *input* gambar. Secara konseptual, tahapan ini terdiri dari data *layer*, model *layer*, dan *inference layer* agar proses pengembangan sistem berjalan sistematis dan mudah direproduksi.

B. Pengumpulan Dataset

Dataset yang digunakan terdiri atas dataset primer dan dataset sekunder. Dataset primer dikumpulkan secara mandiri pada area lobi kampus Universitas Telkom Surabaya menggunakan kamera *webcam*, sedangkan dataset sekunder diperoleh dari *platform* Roboflow dan beberapa video dari Youtube. Total dataset awal yang digunakan berjumlah 1.782 gambar, yang kemudian bertambah melalui augmentasi menjadi 3.026 gambar, dan setelah itu dilakukannya juga *balancing* data pada data latih (*train*) dan membuat total dataset akhir menjadi 2.314 gambar. Pembagian data dilakukan dengan rasio 70:20:10 untuk *train*, *validation*, dan *test*.

C. Anotasi dan Pembuatan Label

Proses anotasi dilakukan secara manual menggunakan *platform* Roboflow dengan metode *bounding box*. Setiap citra akan diberi label sesuai kelas objek, yaitu sepatu dan sandal, kemudian akan diperiksa ulang untuk menjaga konsistensi penempatan *bounding box* dan keseragaman dari label. Pendekatan *bounding box* dipilih agar model dapat mempelajari karakteristik visual objek secara konsisten tanpa terganggu dari latar belakang yang berlebihan.

D. *Preprocessing* dan Augmentasi Data

Preprocessing dilakukan untuk menyeragamkan citra sebelum masuk ke model. Pada penelitian ini, seluruh citra distandarisasi pada resolusi 640 x 640 piksel. Tahap *preprocessing* mencakup *resize*, normalisasi piksel ke

rentang 0-1, serta *letterboxing* untuk tetap mempertahankan aspek rasio gambar. Augmentasi dilakukan secara *offline* melalui Roboflow dan secara otomatis melalui *pipeline* YOLOv8 dengan teknik *horizontal flip*, rotasi ringan, *scaling*, *translation*, dan penyesuaian HSV untuk meningkatkan variasi data serta kemampuan generalisasi pada model.

E. Pelatihan Model

Model yang digunakan adalah YOLOv8n karena ukurannya yang ringan dan sesuai untuk pelatihan berbasis CPU. Proses pelatihan dilakukan menggunakan *framework* Ultralytics YOLOv8 dengan parameter utama berupa *epoch* 150, *batch size* 4, *learning rate* 0.003, *image size* 640 x 640 piksel, *patience* 20, *workers* 2, *confidence threshold* 0.2, dan *device* nya CPU. Konfigurasi tersebut dipilih agar proses pelatihan tetap stabil pada perangkat dengan keterbatasan komputasi. Evaluasi performa model dilakukan menggunakan metrik *precision*, *recall*, mAP50, dan mAP50-95.

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. Hasil Pelatihan Model

Hasil pelatihan model YOLOv8n menunjukkan bahwa model mampu mencapai performa yang cukup baik pada subset validasi. Nilai *precision* berada pada kisaran 0,82–0,85, *recall* 0,75–0,80, mAP50 sekitar 0,85, dan mAP50–95 pada kisaran 0,62–0,64. Grafik pelatihan juga menunjukkan tren penurunan *training loss* dan *validation loss* seiring bertambahnya *epoch*, disertai peningkatan metrik evaluasi hingga model mencapai kondisi konvergen.

TABEL 1
(HASIL EVALUASI MODEL YOLOV8N)

Metrik	Nilai
Precision	0,82 – 0,85
Recall	0,75 – 0,80
mAP50	~0,85
mAP50–95	0,62 – 0,64

Berdasarkan hasil berikut, model menunjukkan kemampuan deteksi yang baik pada ambang IoU standar, namun ketepatan lokalisasi *bounding box* masih menurun ketika dievaluasi dengan ambang yang lebih ketat. Hal ini terlihat dari selisih antara mAP50 dan mAP50-95.

B. Analisis *Confusion Matrix*

Hasil *confusion matrix* menunjukkan bahwa model memiliki performa yang lebih baik pada kelas sepatu dibandingkan kelas sandal. Pada kelas sandal, terdapat 68 prediksi benar, 4 salah diklasifikasikan sebagai sepatu, dan 32 tidak terdeteksi sebagai objek. Pada kelas sepatu, terdapat 674 prediksi benar, 14 salah diklasifikasikan sebagai sandal, dan 177 tidak terdeteksi sebagai objek. Selain itu, *background* masih menghasilkan *false positive*, yaitu 14 prediksi sandal dan 55 prediksi sepatu. Kondisi ini menunjukkan bahwa model masih lebih kuat mengenali sepatu dibanding sandal.

TABEL 2
(CONFUSION MATRIX MODEL DETEKSI ALAS KAKI)

Actual* / *Predicted	**Sandal**	**Sepatu**	***Background***
Sandal	68	4	32
Sepatu	14	674	177
Background	14	55	-

TABEL 3
(CONFUSION MATRIX TERNORMALISASI (%))

***Actual* / *Predicted* (%)**	**Sandal**	**Sepatu**
Sandal	65	4
Sepatu	2	78

C. Hasil Pengujian Sistem

Pengujian sistem dilakukan dalam dua skenario, yaitu deteksi *real-time* menggunakan *webcam* dan deteksi menggunakan *upload* gambar. Hasil pengujian pada kedua skenario menunjukkan status berhasil, dengan objek sepatu dan sandal dapat ditampilkan melalui *bounding box* dan label kelas pada antarmuka sistem.

TABEL 4
(HASIL PENGUJIAN SISTEM DETEKSI ALAS KAKI)

Skenario Pengujian	Mode	Hasil Deteksi	Keterangan
Pengujian 1	*Real-time webcam*	Berhasil	Objek sepatu dan sandal terdeteksi dengan *bounding box* dan label kelas
Pengujian 2	*Upload* gambar	Berhasil	Sistem menampilkan hasil deteksi

			pada citra statis yang diunggah

D. Pembahasan

Nilai *precision* yang lebih tinggi dibandingkan *recall* menunjukkan bahwa model cenderung lebih selektif dalam memberi prediksi, sehingga sebagian besar prediksi yang muncul memang benar, tetapi masih ada objek yang terlewat. Selisih mAP50 dan mAP50–95 juga menandakan bahwa model masih lebih stabil saat penilaian lokalisasi dilakukan pada ambang IoU yang lebih longgar daripada yang ketat.

Performa kelas sepatu yang lebih baik dibanding sandal sangat masuk akal karena distribusi dataset tidak seimbang. Pada data yang digunakan, jumlah objek sepatu jauh lebih banyak daripada sandal, sehingga model lebih sering melihat pola sepatu selama pelatihan. Ketidakseimbangan ini ikut menjelaskan mengapa kelas sandal masih lebih sulit dipelajari secara konsisten.

Peningkatan performa model juga sangat dipengaruhi oleh konsistensi anotasi. Pada percobaan awal dengan *tight bounding box*, performa masih rendah. Setelah anotasi diperbaiki menjadi *bounding box* persegi panjang yang konsisten di seluruh dataset, mAP50 meningkat hingga 0,85 dan mAP50–95 menjadi 0,62. Ini menunjukkan bahwa kualitas label punya pengaruh langsung terhadap hasil pelatihan YOLOv8n.

Secara fungsional, sistem sudah berhasil menjalankan deteksi pada dua mode penggunaan, yaitu *webcam real-time* dan gambar statis. Artinya, integrasi model ke dalam antarmuka web sudah berjalan sesuai tujuan awal penelitian sebagai *proof-of-concept* untuk pemantauan alas kaki di lingkungan kampus.

V. KESIMPULAN

Berdasarkan hasil penelitian, sistem deteksi jenis alas kaki berbasis *computer vision* menggunakan YOLOv8n berhasil dirancang dan diimplementasikan untuk membedakan sepatu dan sandal pada citra *webcam* di lingkungan kampus. Model menunjukkan performa yang cukup baik dengan *precision* pada kisaran 0,82–0,85, *recall* 0,75–0,80, mAP50 sekitar 0,85, dan mAP50–95 pada kisaran 0,62–0,64. Hasil pengujian juga menunjukkan bahwa sistem mampu berjalan pada dua skenario, yaitu deteksi *real-time* melalui *webcam* dan deteksi menggunakan gambar statis yang diunggah. Di sisi lain, performa model masih lebih kuat pada kelas sepatu dibandingkan sandal, yang dipengaruhi oleh ketidakseimbangan distribusi dataset serta konsistensi anotasi *bounding box* selama proses pelatihan. Secara keseluruhan, penelitian ini menunjukkan bahwa YOLOv8n mampu digunakan sebagai *baseline* model untuk deteksi alas kaki berbasis citra webcam serta memberikan kontribusi berupa dataset alas kaki kampus dan evaluasi performa model pada lingkungan kampus dengan keterbatasan sumber daya komputasi.

REFERENSI

[1] TELKOM UNIVERSITY, "LARANGAN BAGI MAHASISWA TELKOM UNIVERSITY." Accessed: Jan. 23, 2026. [Online]. Available: https://telkomuniversity.ac.id/aturan-dan-sanksi/

[2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Dec. 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91.

[3] Ultralytics, "Ultralytics. (2023). YOLOv8 documentation." Accessed: Jan. 28, 2026. [Online]. Available: https://github.com/ultralytics/ultralytics

[4] Y. Shen, X. Jiang, Y. Zhao, and W. Xie, "Barefoot Footprint Detection Algorithm Based on YOLOv8-StarNet," *Sensors*, vol. 25, no. 15, Aug. 2025, doi: 10.3390/s25154578.

[5] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object Detection With Deep Learning: A Review," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.

[6] J. Zou, T. Song, S. Cao, B. Zhou, and Q. Jiang, "Dress Code Monitoring Method in Industrial Scene Based on Improved YOLOv8n and DeepSORT," *Sensors*, vol. 24, no. 18, Sep. 2024, doi: 10.3390/s24186063.

[7] X. Wang, H. Gao, Z. Jia, and Z. Li, "BL-YOLOv8: An Improved Road Defect Detection Model Based on YOLOv8," *Sensors (Basel)*, vol. 23, no. 20, Oct. 2023, doi: 10.3390/s23208361.

[8] H. Yi, B. Liu, B. Zhao, and E. Liu, "Small Object Detection Algorithm Based on Improved YOLOv8 for Remote Sensing," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 17, pp. 1734–1747, 2024, doi: 10.1109/JSTARS.2023.3339235.

[9] R. Szeliski, "Computer Vision: Algorithms and Applications 2nd Edition," 2022. doi: https://doi.org/10.1007/978-3-030-34372-9.

[10] J. Wang, T. Zhang, Y. Cheng, and N. Al-Nabhan, "Deep learning for object detection: A survey," Apr. 23, 2021, *Tech Science Press*. doi: 10.32604/CSSE.2021.017016.