

## PERBANDINGAN *INSTANCE SEGMENTATION IMAGE* PADA YOLO8

Resty Wulanningrum<sup>\*1</sup>, Anik Nur Handayani<sup>2</sup>, Aji Prasetya Wibawa<sup>3</sup>

<sup>1</sup>Universitas Nusantara PGRI Kediri, Kediri, <sup>1,2,3</sup>Universitas Negeri Malang, Malang

Email: <sup>1</sup>restyw@unpkdr.ac.id, <sup>2</sup>aniknur.ft@um.ac.id, <sup>3</sup>aji.prasetya.ft@um.ac.id

<sup>\*</sup>Penulis Korespondensi

(Naskah masuk: 14 Desember 2023, diterima untuk diterbitkan: 08 Agustus 2024)

### Abstrak

Seorang pejalan kaki sangat rawan terhadap kecelakaan di jalan. Deteksi pejalan kaki merupakan salah satu cara untuk mengidentifikasi atau mengklasifikasikan antara orang, jalan atau yang lainnya. *Instance segmentation* adalah salah satu proses untuk melakukan segmentasi antara orang dan jalan. *Instance segmentation* dan penggunaan yolov8 merupakan salah satu implementasi dalam deteksi pejalan kaki. Perbandingan segmentasi pada dataset *Penn-Fudan Database* menggunakan yolov8 dengan model *yolov8n-seg*, *yolov8s-seg*, *yolov8m-seg*, *yolov8l-seg*, *yolov8x-seg*. Penelitian ini menggunakan dataset publik *pedestrian* atau pejalan kaki dengan objek multi person yang diambil dari dataset *Penn-Fudan Database*. Dataset mempunyai 2 kelas, yaitu orang dan jalan. Hasil perbandingan penggunaan model yolov8 model segmentasi yang terbaik adalah menggunakan model yolov8l-seg. Hasil penelitian didapatkan *Instance segmentation valid box* pada data orang, mAP50 tertinggi pada yolov8l-seg dengan nilai 0,828 dan mAP50-95 adalah 0,723. *Instance segmentation valid mask* pada orang nilai mAP50 tertinggi pada yolov8l-seg dengan nilai 0,825 dan mAP50-95 adalah 0,645. Pada penelitian ini, yolov8l-seg menjadi nilai terbaik dibandingkan versi yang lain, karena berdasarkan nilai mAP tertinggi pada valid mask sebesar 0,825.

**Kata kunci:** *Instance Segmentation, yolo8, pejalan kaki, perbandingan*

## COMPARISON OF *IMAGE SEGMENTATION INSTANCE* ON YOLO8

### Abstract

A pedestrian is very vulnerable to road accidents. Pedestrian detection is one way to identify or classify between people, roads or others. *Instance segmentation* is one of the processes to segment people and roads. *Instance segmentation* and the use of yolov8 is one of the implementations in pedestrian detection. Comparison of segmentation on *Penn-Fudan Database* dataset using yolov8 with *yolov8n-seg*, *yolov8s-seg*, *yolov8m-seg*, *yolov8l-seg*, *yolov8x-seg* models. This research uses a public pedestrian dataset with multi-person objects taken from the *Penn-Fudan Database* dataset. The dataset has 2 classes, namely people and roads. The results of the comparison using the yolov8 model, the best segmentation model is using the yolov8l-seg model. The results obtained *Instance segmentation valid box* on people data, the highest mAP50 on yolov8l-seg with a value of 0.828 and mAP50-95 is 0.723. *Instance segmentation valid mask* on people the highest mAP50 value on yolov8l-seg with a value of 0.825 and mAP50-95 is 0.645. In his study, yolov8l-seg is the best value compared to other versions, because based on the highest mAP value on the valid mask of 0.825.

**Keywords:** *Instance Segmentation, yolo8, pedestrian, comparison*

### 1. PENDAHULUAN

Deteksi pejalan kaki adalah untuk mengidentifikasi dan mengukur keberadaan dan posisi dari pejalan kaki dalam gambar. Pejalan kaki memiliki resiko lebih tinggi untuk terlibat kecelakaan, sehingga merupakan tugas penting dalam banyak aplikasi praktis untuk keselamatan pejalan kaki, mengemudi secara otonom dan pengawasan video (Rajendar, et al., 2022). Salah satu model yang bisa

melakukan deteksi pejalan kaki adalah menggunakan YOLOv8.

YOLOv8 (Jocher, et al., 2023) adalah versi terbaru dari rangkaian algoritma deteksi objek YOLO (*You Only Look Once*), yang dikenal dengan kemampuan dan akurasi pendeteksian *real time* (Bochkovskiy, et al., 2020), (Jocher, 2022), (Li, et al., 2023) (Li, et al., 2022), (Redmon, et al., 2022), (Wang, et al., 2017), (Wang, et al., 2022). YOLOv8

menggunakan mekanisme *anchorfree detection*, *feature pyramid network* yang disempurnakan, dan fungsi kerugian yang dimodifikasi untuk mencapai peningkatan kinerja dibandingkan pendahulunya. Ini dirancang untuk mensegmentasi objek dalam gambar dengan memprediksi *bounding box* dan *associated class probabilities* menggunakan *single convolutional neural network* (CNN) (Dumitriu, et al., 2023).

Sebelum memilih YOLOv8, peneliti mempertimbangkan model lain yang berada di peringkat teratas pengujian segmentasi instance COCO, seperti EVA (Fang, et al., 2022), FD-SwinV2-G (Wei, et al., 2022), BEiT-3 (Wei, et al., 2022), MasK Dino (Li, et al., 2022) dan SwinV2-G (Liu, et al., 2022), serta metode lama, seperti TensorMask (Chen, et al., 2019), PolarMask (Xie, et al., 2020), SipMask (Xie, et al., 2020), dan Topeng R-CNN (He, et al., 2017). Namun, peneliti memilih YOLOv8 sebagai metode dasar karena beberapa alasan. Pertama, algoritma YOLO secara konsisten menunjukkan kinerja yang kuat dalam tugas deteksi objek (Dumitriu, et al., 2023), menjadikannya kompetitif dan cocok untuk masalah *instance segmentation* pejalan kaki (*pedestrian*). Kemampuan deteksi *real-time* YOLOv8, khususnya, sangat penting untuk aplikasi yang melibatkan keselamatan publik dan tanggap darurat, seperti yang ditunjukkan oleh keberhasilan versi YOLO sebelumnya dalam berbagai tugas deteksi objek. Kedua, YOLOv8 adalah metode yang sudah mapan dengan komunitas pengguna yang luas, menyediakan sumber daya implementasi yang dapat diakses (Zhao & Song, 2023). Ketiga, mudah untuk diterapkan dan digunakan dalam berbagai ukuran versi (Lou, et al., 2023).

Penggunaan *instance segmentation* merupakan bagian dari *preprocessing data* untuk mendapatkan deteksi dan klasifikasi objek dengan baik. Sehingga bisa didapatkan hasil yang terbaik pada proses deteksi pejalan kaki.

Tujuan dan kontribusi dari penelitian ini adalah untuk mengetahui model yolov8 yang mana, untuk menghasilkan akurasi serta model yang terbaik pada data pejalan kaki.

## 2. METODE PENELITIAN

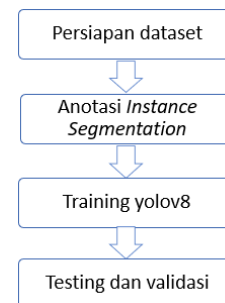
### 2.1 Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah sebagai berikut:

#### a. Persiapan dataset

Penelitian ini menggunakan dataset publik tentang gambar *pedestrian* atau pejalan kaki dengan objek multi person yang diambil dari dataset *Penn-Fudan Database for Pedestrian*. Data set terdiri dari 2 kelas, yaitu orang dan jalan. Proses *training* akan dibedakan menjadi 3 kelas, yaitu orang, jalan dan all. All disini maksudnya adalah langsung dikenali 2 kelas, yaitu orang dan jalan. Dataset yang digunakan berjumlah 210 data

dengan pembagian 80% atau 167 gambar yang digunakan sebagai data training, 14% atau 30 gambar yang digunakan sebagai *valid*, dan 6% atau 13 gambar yang digunakan sebagai *testing*.



Gambar 1. Metode penelitian



Gambar 2. Dataset anotasi orang



Gambar 3. Dataset anotasi orang dan jalan

Gambar 2 merupakan salah satu contoh dataset yang akan digunakan pada proses pelatihan dan akan dilakukan anotasi orang. Sedangkan gambar 3 merupakan salah satu contoh dataset yang akan dilakukan anotasi orang dan jalan.

#### b. Anotasi (*Instance Segmentation*)

Proses anotasi merupakan proses pemberian label pada suatu gambar. Anotasi yang digunakan dalam penelitian ini adalah menggunakan *instance segmentation* dengan menggunakan tools *roboflow*. Proses anotasi terbagi menjadi 2 kelas, yaitu orang dan jalan.

Anotasi dataset pada gambar 4 merupakan anotasi data orang dan gambar 5 mempunyai 2 kelas, yaitu kelas orang dan jalan. Pelabelan berfungsi untuk segmentasi objek orang dan jalan.



Gambar 4. Anotasi Data Orang



Gambar 5. Anotasi Data Orang dan Jalan

#### c. Training YOLOv8

Prose *training* menggunakan YOLOv8 untuk data pejalan kaki yang sudah dilakukan anotasi. Tahapan ini menggunakan 5 model yang akan digunakan sebagai perbandingan, mana yang lebih bagus untuk proses deteksi pejalan kaki. Penelitian yang akan dilakukan akan mengambil pada tugas segmentasi, yaitu *yolov8n-seg*, *yolov8s-seg*, *yolov8m-seg*, *yolov8l-seg*, *yolov8x-seg*.

#### d. Testing dan validasi

*Testing* dan validasi merupakan proses ujicoba dan pengujian dari dataset yang sudah disiapkan untuk mengetahui eavaluasinya terhadap nilai MAP50 dan MAP50-95.

### 2.2 YOLOv8

YOLOv8 adalah versi terbaru YOLO dari Ultralytics. Sebagai model mutakhir dan terancang (SOTA), YOLOv8 membangun kesuksesan versi sebelumnya, memperkenalkan fitur-fitur baru dan peningkatan untuk meningkatkan kinerja, fleksibilitas, dan efisiensi. YOLOv8 mendukung berbagai tugas vision AI, termasuk deteksi, segmentasi, estimasi pose, pelacakan, dan klasifikasi. Fleksibilitas ini memungkinkan pengguna untuk memanfaatkan kemampuan YOLOv8 di beragam aplikasi dan domain (Jocher, et al., 2023).

Model YOLOv8 diumumkan oleh Ultralytics (Los Angeles, CA, USA) pada tahun 2023. Model YOLOv8 merupakan model terbaru dari YOLO yang memadukan kecepatan tinggi dan tinggi ketepatan. Untuk tugas deteksi objek, penelitian ini membandingkan model YOLOv8 dengan model sebelumnya model, YOLOv5, yang diterbitkan oleh perusahaan yang sama (Inui, et al., 2023).

### 2.3 Instance Segmentation

*Instance Segmentasi* merupakan perkembangan dari deteksi objek dan melibatkan identifikasi objek

individual dalam gambar dan pengelompokkannya dari gambar lainnya. Output dari model *instance segmentation* adalah sekumpulan masker atau kontur yang menguraikan setiap objek dalam gambar, beserta label kelas dan skor keyakinan untuk setiap objek. *Instance segmentation* berguna tidak hanya untuk mengetahui lokasi objek dalam gambar, tetapi juga bentuk sebenarnya (Jocher, et al., 2023).

## 3. HASIL DAN PEMBAHASAN

### 3.1 Environment dan Parameter

Pelatihan dan inferensi dijalankan pada NVIDIA-SMI 525.105.17, menggunakan YOLOv8.0.20, Python 3.10.12, PyTorch 2.1.0 dan CUDA 12.0. Ukuran gambar yang dilatih  $640 \times 640$  piksel dengan 100 *epoch*, dengan *early stopping patience parameter* 50 *epoch*. Nilai parameter yang digunakan masih menggunakan *default* dari yolov8.

### 3.2 Evaluation Metric

Untuk evaluasi data validasi, melihat beberapa metrik: last epoch, untuk melihat di mana model dilatih dihentikan karena *overfitting* (dengan *patience 50 epoch*), *epoch* terbaik, yaitu epoch dengan hasil terbaik sebelum memulai *overfit*, waktu pelatihan (rata-rata per *epoch*), *frame per second* (FPS), *presisi* (2), *recall* (3), mAP50, map50:95 (5). Pada penelitian ini menggunakan *Mean Average Precision* (mAP) sebagai metrik evaluasi yang utama, khususnya mAP50. mAP ini didasarkan pada *Intersection over Union* (IoU) (1), yang diturunkan dari Jaccard indeks, dan mewakili rata-rata dari nilai presisi pada perbedaan *recall* levels untuk *IoU threshold*.

IoU, atau indeks Jaccard, mengukur kesamaan antara dua himpunan dan didefinisikan sebagai rasio perpotongannya terhadap gabungannya:

$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

Precision dan Recall dihitung menggunakan rumus:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (2)$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \dots\dots \quad (3)$$

Secara umum, *Average Precision* (AP) untuk satu kelas dihitung dengan memberi peringkat pada prediksi model berdasarkan prediksinya *confidence scores*, dan kemudian menghitung area di bawah kurva *presisi-recall*:

$$AP = \sum_n (Recall_n - Recall_{n-1}) \cdot Precision_n \quad (4)$$

Model yang dibuat untuk mendeteksi dua kelas (pejalan kaki dan jalan), maka mAP sama dengan AP kelas itu. Untuk mAP50, dengan menghitung AP pada ambang batas IoU 0,5. Untuk mAP50-95, menghitung nilai AP untuk setiap ambang batas IoU dari 0,5 hingga 0,95 inci kenaikan 0,05, lalu dirata-rata nilai tersebut.

$$mAP = \frac{AP_{IoU=0.5} + AP_{IoU=0.55} + \dots + AP_{IoU=0.90}}{k} \quad (5)$$

Dimana k adalah jumlah ambang batas IoU yang dipertimbangkan.

Dengan menggunakan mAP 50 dan mAP50-95, mempunyai kemampuan mengevaluasi model secara akurat untuk mengelompokkan pejalan kaki dan jalan pada batas ambang IoU yang bervariasi. Selain itu juga untuk menilai efisiensi model dengan mempertimbangkan kecepatan pelatihan dan waktu inferensi.

### 3.3 Training metode

Proses pelatihan pada semua ukuran model (n, s, m, l, dan x) dengan, membandingkan semua hasil, pelatihan, dan waktu inferensi, tetapi berfokus pada mAP50 dan FPS. Metrik yang dibandingkan untuk setiap model mewakili rata-rata, kecuali standar deviasi. Selain itu proses training metode menggunakan perbandingan layers, parameter, gradients dan GFLOPs, secara detail terlihat pada table 1.

Tabel 1. Training Metode

	yolov8n	yolov8s	yolov8m	yolov8l	yolov8x
Layers	261	261	331	401	401
Parameters	3264006	11790870	27240806	45937590	71752774
Gradients	3263990	11790854	27240790	45937574	71752758
GFLOPs	12,1	42,7	110,4	220,8	344,5

Tabel 1 menjelaskan tentang penggunaan training metode yang digunakan pada setiap *instance segmentation* pada YOLO 8, yaitu pada yolov8n menggunakan 261 layers, 3264006 parameters, 3263990 gradiens dan 12,1 GFLOPs. Sedangkan pada yolov8s menggunakan 261 layers, 11790870 parameters, 11790854 gradients dan 42,7 GFLOPs. Yolov8m menggunakan 331 layers, 27240806 parameters, 27240790 gradients dan 110,4 GFLOPs, yolov8l menggunakan 401 layers, 45937590 parameters, 45937574 gradients dan 220,8 GFLOPs. Pada yolov8x menggunakan 401 layers, 71752774 parameters, 71752758 gradients dan 344,5 GFLOPs.

### 3.4 Hasil

Ujicoba yang dilakukan pada penelitian ini meliputi proses *train box* dan *train mask*. *Train box* adalah proses pelatihan dengan menggunakan sebuah area kotak, sedangkan *train mask* adalah proses pelatihan pada area yang dilakukan *annotation* atau pelabelan terhadap objek. Objek akan dilabeli menjadi orang dan jalan. Pada saat proses pelatihan bisa membagi menjadi All, Orang dan jalan. All merupakan proses langsung *instance segmentation* pada orang dan jalan, sedangkan orang hanya berfokus pada *instance segmentation* objek orang dan jalan berfokus pada *instance segmentation* jalan.

Perbandingan *instance segmentation* pada train box dengan semua model ditunjukkan oleh Tabel 2.

Table 2 menunjukkan hasil pelatihan box dengan *instance segmentation* pada All (orang dan jalan) nilai mAP50 yang tertinggi pada yolov8m-seg yaitu 0,66 dan nilai mAP50-95 adalah 0,538. *Instance segmentation* pada orang nilai mAP50 tertinggi pada yolov8l-seg dengan nilai 0,828 dan mAP50-95 adalah 0,721. Sedangkan *instance segmentation* pada jalan nilai tertinggi pada yolov8x-seg dengan nilai mAP50 0,512 dan mAP50-95 sebesar 0,366.

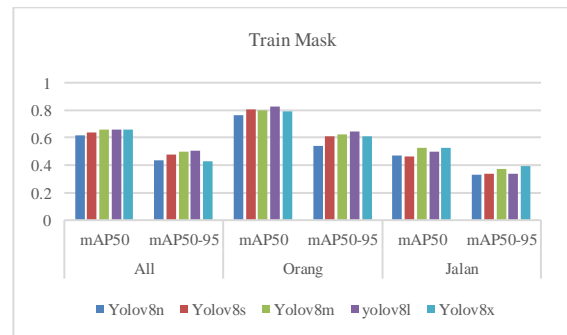


Gambar 6 Grafik Train Box

Gambar 6 menunjukkan perbandingan *instance segmentation* yolov8n, yolov8s, yolov8m, yolov8l, yolov8x. Pada objek orang memiliki nilai mAP50 dan mAP50-95 yang paling tinggi dibandingkan objek all dan jalan, dengan nilai tertinggi mAP50 dan mAP50-95 pada yolov8l-seg.

Table 3 menunjukkan hasil pelatihan masker atau pelabelan dengan *instance segmentation* pada All (orang dan jalan) nilai mAP50 yang tertinggi pada yolov8l-seg yaitu 0,662 dan nilai mAP50-95 adalah 0,503. *Instance segmentation* pada orang nilai mAP50 tertinggi pada yolov8l-seg dengan nilai 0,825 dan mAP50-95 adalah 0,645. Sedangkan *instance segmentation* pada jalan nilai tertinggi pada yolov8x-seg dengan nilai mAP50 0,524 dan mAP50-95 sebesar 0,339.

Gambar 7 menunjukkan perbandingan *instance segmentation* yolov8n, yolov8s, yolov8m, yolov8l, yolov8x. Pada objek orang memiliki nilai mAP50 dan mAP50-95 yang paling tinggi dibandingkan objek all dan jalan, dengan nilai tertinggi mAP50 dan mAP50-95 pada yolov8l-seg.



Gambar 7 Train Mask



Tabel 2 *Train Box*

		Yolov8n	Yolov8s	Yolov8m	yolov8l	Yolov8x
All	mAP50	0,587	0,615	0,66	0,635	0,659
	mAP50-95	0,424	0,506	0,538	0,52	0,53
Orang	mAP50	0,777	0,816	0,828	0,828	0,807
	mAP50-95	0,618	0,698	0,717	0,721	0,695
Jalan	mAP50	0,396	0,415	0,492	0,442	0,512
	mAP50-95	0,231	0,314	0,358	0,319	0,366

Tabel 3 *Train Mask*

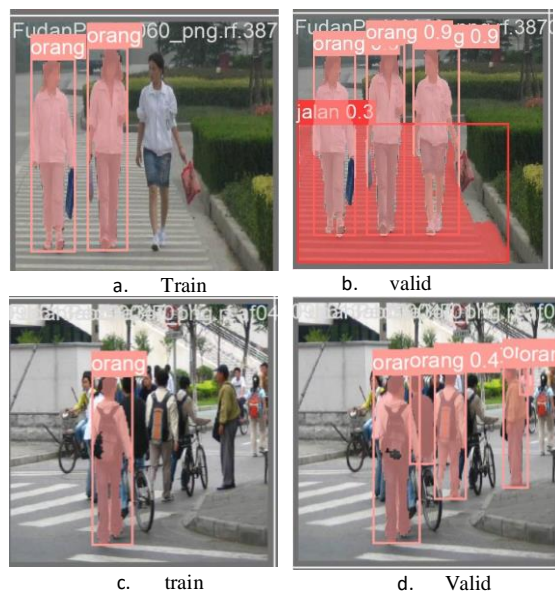
		Yolov8n	Yolov8s	Yolov8m	Yolov8l	Yolov8x
All	mAP50	0,617	0,636	0,662	0,662	0,657
	mAP50-95	0,433	0,475	0,497	0,503	0,429
Orang	mAP50	0,763	0,808	0,802	0,825	0,789
	mAP50-95	0,537	0,612	0,622	0,645	0,612
Jalan	mAP50	0,471	0,465	0,523	0,5	0,524
	mAP50-95	0,329	0,337	0,372	0,339	0,394

Tabel 4 *Valid Box*

		Yolov8n	Yolov8s	Yolov8m	yolov8l	Yolov8x
All	mAP50	0,586	0,615	0,66	0,634	0,66
	mAP50-95	0,424	0,499	0,539	0,52	0,53
Orang	mAP50	0,777	0,816	0,828	0,828	0,807
	mAP50-95	0,621	0,695	0,719	0,723	0,695
Jalan	mAP50	0,396	0,415	0,492	0,441	0,512
	mAP50-95	0,228	0,304	0,358	0,318	0,366

### 3.5 Pengujian

Penelitian ini membandingkan semua model segmentasi yolov8n-seg, yolov8s-seg, yolov8m-seg, yolov8l-seg, yolov8x-seg menggunakan *cross validation* pada gambar.

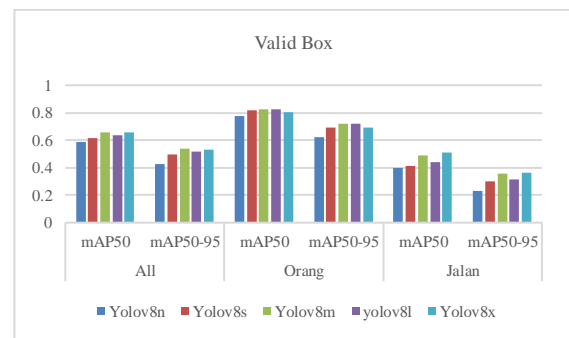
Gambar 8. *Train dan Valid*

Gambar 8, merupakan contoh sampel penggunaan Batch 1 pada yolov8-seg. Pada table tersebut ada 2 model yaitu *train* (ditunjukkan oleh a dan c) dan *Valid* (ditunjukkan oleh b dan d). Train merupakan proses pelatihan dari dataset sedangkan valid adalah proses validasi dari instance segmentation. Kolom valid akan muncul nilai dari

proses pengenalan dengan memunculkan nilai *Frame Per Second* (FPS) dari tiap objek yang terdeteksi.

Hasil validasi *train box* untuk objek pejalan kaki terdiri dari 3 pengenalan, yaitu all (orang dan jalan), orang dan jalan yang ditunjukkan pada table .

Table 4 menunjukkan hasil validasi box dengan instance segmentation pada All (orang dan jalan) nilai mAP50 yang tertinggi pada yolov8m-seg yaitu 0,66 dan nilai mAP50-95 adalah 0,539. Instance segmentation pada orang nilai mAP50 tertinggi pada yolov8l-seg dengan nilai 0,828 dan mAP50-95 adalah 0,723. Sedangkan instance segmentation pada jalan nilai tertinggi pada yolov8x-seg dengan nilai mAP50 0,512 dan mAP50-95 sebesar 0,366.

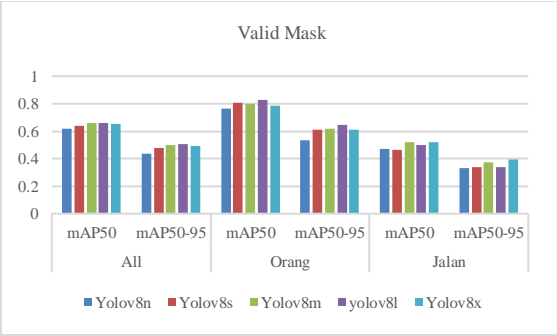
Gambar 9 *Valid Box*

Gambar 9 menunjukkan perbandingan instance segmentation yolov8n, yolov8s, yolov8m, yolov8l, yolov8x. Pada objek orang memiliki nilai mAP50 dan mAP50-95 yang paling tinggi dibandingkan objek all dan jalan, dengan nilai tertinggi mAP50 dan mAP50-95 pada yolov8l-seg.

Table 5 menunjukkan hasil validasi box dengan instance segmentation pada All (orang dan jalan) nilai mAP50 yang tertinggi pada yolov8l-seg yaitu 0,662 dan nilai mAP50-95 adalah 0,503. Instance segmentation pada orang nilai mAP50 tertinggi pada yolov8l-seg dengan nilai 0,825 dan mAP50-95 adalah 0,645. Sedangkan instance segmentation pada jalan nilai tertinggi pada yolov8x-seg dengan nilai mAP50 0,522 dan mAP50-95 sebesar 0,394.

Tabel 5 Valid Mask

		Yolov8n	Yolov8s	Yolov8m	yolov8l	Yolov8x
All	mAP50	0,617	0,636	0,662	0,662	0,656
	mAP50-95	0,433	0,475	0,496	0,503	0,493
Orang	mAP50	0,763	0,808	0,802	0,825	0,789
	mAP50-95	0,537	0,613	0,621	0,645	0,612
Jalan	mAP50	0,471	0,465	0,522	0,498	0,522
	mAP50-95	0,329	0,338	0,372	0,341	0,394



Gambar 10 Valid Mask

Gambar 10 menunjukkan perbandingan instance segmentation yolov8n, yolov8s, yolov8m, yolov8l, yolov8x. Pada objek orang memiliki nilai mAP50 dan mAP50-95 yang paling tinggi dibandingkan objek all dan jalan, dengan nilai tertinggi mAP50 dan mAP50-95 pada yolov8l-seg.

Pada ujicoba yang dilakukan di atas, nilai yang paling tinggi menggunakan mAP50 pada model yolov8l. untuk kasus saya ini yg terbaik menggunakan model yolov8l dg nilai mAP sebesar 82,8% untuk bounding box dan 82,5% untuk valid mask.

4. KESIMPULAN

Berdasarkan pengujian yang telah dilakukan dengan membandingkan model yolov8n-seg, yolov8s-seg, yolov8m-seg, yolov8l-seg dan yolov8x-seg pada data gambar banyak orang pejalan kaki dengan menggunakan anotasi data orang dan pejalan kaki, maka didapatkan hasil penelitian yaitu: *instance segmentation valid box* tertinggi orang nilai mAP50 tertinggi pada yolov8l-seg dengan nilai 0,828 dan mAP50-95 adalah 0,723 dan *Instance segmentation valid mask* tertinggi pada orang nilai mAP50 tertinggi pada yolov8l-seg dengan nilai 0,825 dan mAP50-95 adalah 0,645

DAFTAR PUSTAKA

ALZUBAIDI, ET AL., 2021. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal Big Data*, Volume 8, pp. 1-74. <https://doi.org/10.1186/s40537-021-00444-8>  
ANAGNOSTIS, A., TAGARAKIS, A. C., KATERIS, D. & MOYSIADIS, V., 2021. Orchard mapping with deep learning semantic segmentation. *Sensor*, Volume 21, pp. 1-20. <https://doi.org/10.3390/s21113813>  
BOCHKOVSKIY, A., WANG, C. Y. & LIAO, H. M., 2020. Yolov4: Optimal speed and accuracy

of object detection. *arXiv preprint*, p. arXiv:2004.10934. <https://doi.org/10.48550/arXiv.2004.10934>  
CHEN, X., GIRSHICK, R., HE, K. & DOLLAR, P., 2019. *Tensormask: A foundation for dense object segmentation*. s.l., In Proceedings of the IEEE/CVF international conference on computer vision, pages 2061–2069. <https://doi.org/10.1109/ICCV.2019.00215>  
DUMITRIU, A. et al., 2023. *Rip Current Segmentation: A Novel Benchmark and YOLOv8 Baseline Results*. s.l., In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1261–1271). <https://doi.org/10.1109/CVPRW59228.2023.00133>  
FANG, Y. ET AL., 2022. Eva: Exploring the limits of masked visual representation learning at scale. *Computer Vision and Pattern Recognition*, Volume arXiv preprint arXiv:2211.07636. <https://doi.org/10.1109/CVPR52729.2023.01855>  
HE, K., GKIOXARI, G., DOLLAR, P. & GIRSHICK, R., 2017. *Mask R-CNN*. s.l., In Proceedings of the IEEE international conference on computer vision, pages 2961–2969. <https://doi.org/10.48550/arXiv.1703.06870>  
INUI, A. et al ., 2023. Detection of elbow OCD in the ultrasound image by artificial intelligence using YOLOv8. *Applied Sciences*, 13(13), p. 7623. <https://doi.org/10.3390/app13137623>  
JOCHER, G., 2022. *YOLOv5 by Ultralytics*. [Online] Available at: <https://github.com/ultralytics/yolov5/> [Diakses 4 November 2023].  
JOCHER, G., CHAURASIA, A. & QIU, J., 2023. *YOLO by Ultralytics*. [Online] Available at: <https://github.com/ultralytics/ultralytics> [Diakses 2 November 2023].  
LI, C. et al., 2023. YOLOv6 v3.0: A Full-Scale Reloading. *Computer Vision and Pattern Recognition*, Volume arXiv:2301.05586. <https://doi.org/10.48550/arXiv.2301.05586>  
LI, C. et al ., 2022. Yolov6: A single-stage object detection framework for industrial applications. *Computer Vision and Pattern Recognition*, Volume arXiv:2209.02976. <https://doi.org/10.48550/ARXIV.2209.02976>  
LI, F. et al., 2022. Mask DINO: Towards a unified transformer-based framework for object

- detection and segmentation. *Computer Vision and Pattern Recognition*, Volume arXiv:2206.02777. <https://doi.org/10.48550/arXiv.2206.02777>
- LIU, Z. et al., 2022. *Swin transformer v2: Scaling up capacity and resolution*. arXiv preprint arXiv:2206.02777, In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. <https://doi.org/10.1109/CVPR52688.2022.01170>
- LOU, H. et al., 2023. DC-YOLOv8: Small-Size Object Detection Algorithm Based on Camera Sensor. *Electrics*, 12(10), p. 2323. <https://doi.org/10.3390/electronics12102323>
- RAJENDAR, S. et al., 2022. Prediction of stopping distance for autonomous emergency braking using stereo camera pedestrian detection. *Materials Today: Proceedings*, 51(1), pp. 1224-1228. <https://doi.org/10.1016/j.matpr.2021.07.211>
- REDMON, J., DIVVALA, S., GIRSHICK, R. & FARHADI, A., 2022. *You only look once: Unified, real-time object detection*. s.l., IEEE conference on computer vision and pattern recognition. <https://doi.org/10.1109/CVPR.2016.91>
- SETIYADI, A., UTAMI, E. & ARIATMANTO, D., 2023. Analisa Kemampuan Algoritma YOLOv8 Dalam Deteksi Objek Manusia Dengan Metode Modifikasi Arsitektur. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 7(2), pp. 891-901. <https://doi.org/10.30645/j-sakti.v7i2.694>
- WANG, C. Y., BOCHKOVSKIY, A. & LIAO, H. M., 2022. *Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors*. s.l., IEEE. <https://doi.org/10.1109/CVPR52729.2023.00721>
- WANG, C. Y. et al., 2017. Yolo9000: Better, faster, stronger. *Computer Vision and Pattern Recognition (IEEE)*, pp. 6517-6525. <https://doi.org/10.1109/CVPR.2017.690>
- WEI, Y., HU, H. & XIE, Z., 2022. Image as a foreign language: BEiT pretraining for all vision and visionlanguage tasks. *Computer Vision and Pattern Recognition*, Volume arXiv:2208.10442. <https://doi.org/10.48550/arXiv.2208.10442>
- WEI, Y., HU, H., XIE, Z. & ZHANG, Z., 2022. Contrastive learning rivals masked image modeling in fine-tuning via feature distillation. *Computer Vision and Pattern Recognition*, Volume arXiv:2205.14141. <https://doi.org/10.48550/arXiv.2205.14141>
- XIE, E. et al., 2020. *Polarmask: Single shot instance segmentation with polar representation*. s.l., In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 12193–12202. <https://doi.org/10.1109/CVPR42600.2020.01221>
- XIE, E. et al., 2020. *Sipmask: Spatial information preservation for fast image and video instance segmentation*. Glasgow, UK, In Computer Vision–ECCV 2020: 16th European Conference. [https://doi.org/10.1007/978-3-030-58568-6\\_1](https://doi.org/10.1007/978-3-030-58568-6_1)
- ZHAO, X. & SONG, Y., 2023. Improved Ship Detection with YOLOv8 Enhanced with MobileViT and GSConv. *Electronics*, 12(22), p. 4666. <https://doi.org/10.3390/electronics12224666>

*Halaman ini sengaja dikosongkan.*