Evaluation Metrics — YOLOv8 Instance Segmentation Model (Teeth Detection & Segmentation) on Validation dataset

📦 Bounding Box Metrics

Metric	Value
Precision (B)	0.9861
Recall (B)	0.9893
mAP50 (B)	0.9937
mAP50-95 (B)	0.7271

🎭 Segmentation Mask Metrics

Metric	Value
Precision (M)	0.9873
Recall (M)	0.9887
mAP50 (M)	0.9940
mAP50-95 (M)	0.6817

“The model achieved very high performance.
For bounding boxes, mAP50 is 0.9937, and for masks 0.9940.
Under stricter IoUs (mAP50–95), it achieved 0.72 for boxes and 0.68 for masks, which is expected in segmentation tasks.

Precision and recall are both around 98–99%, meaning the model detects almost all teeth with very few false alarms.

Additionally, the fitness score—a combined metric YOLO uses to decide the ‘best’ model checkpoint—was 1.4087, which is *significantly high*.
A higher fitness score means the model has an excellent balance of precision, recall, box accuracy, mask accuracy, and confidence calibration, confirming that this checkpoint is highly optimized and reliable.”