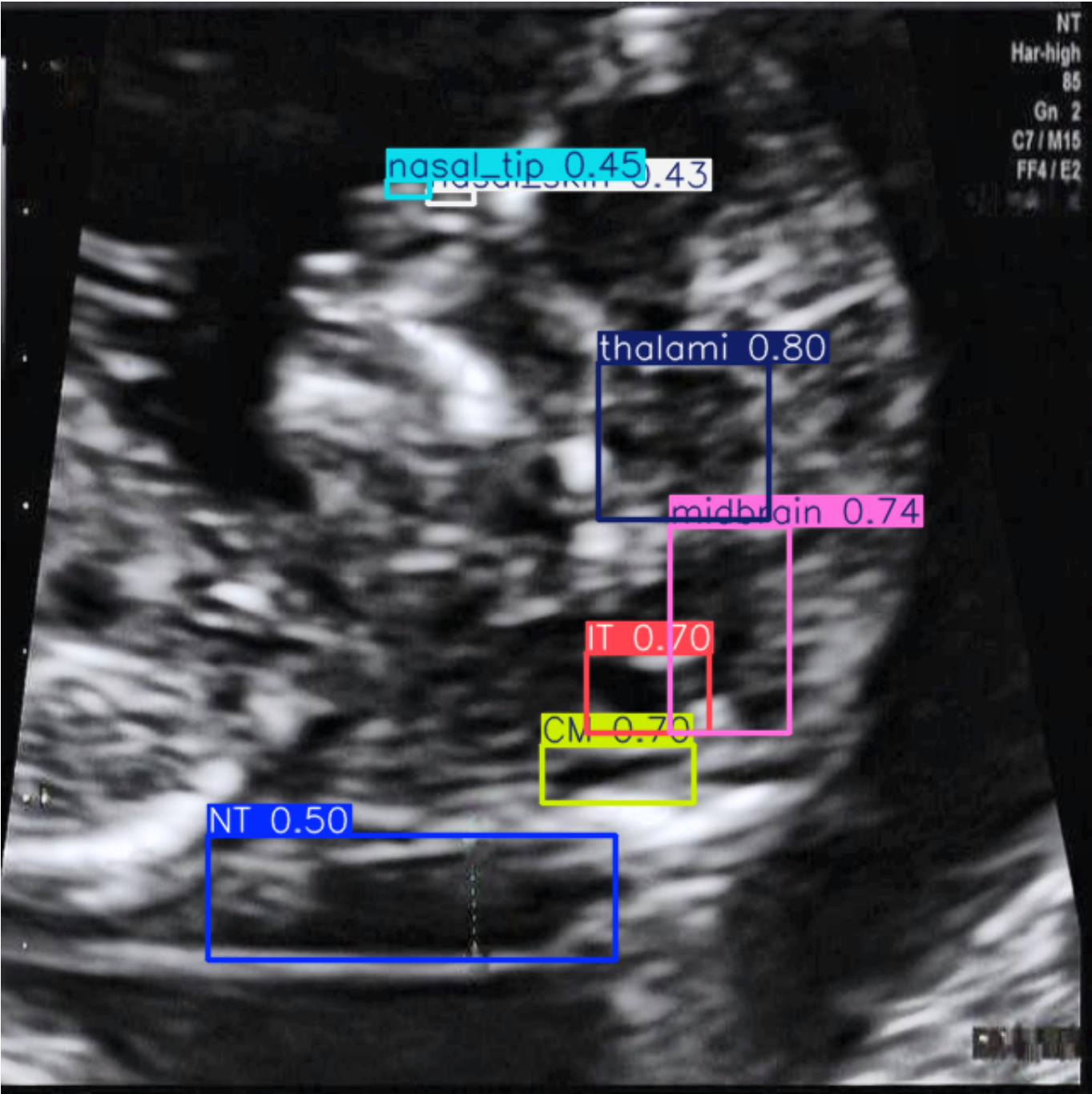


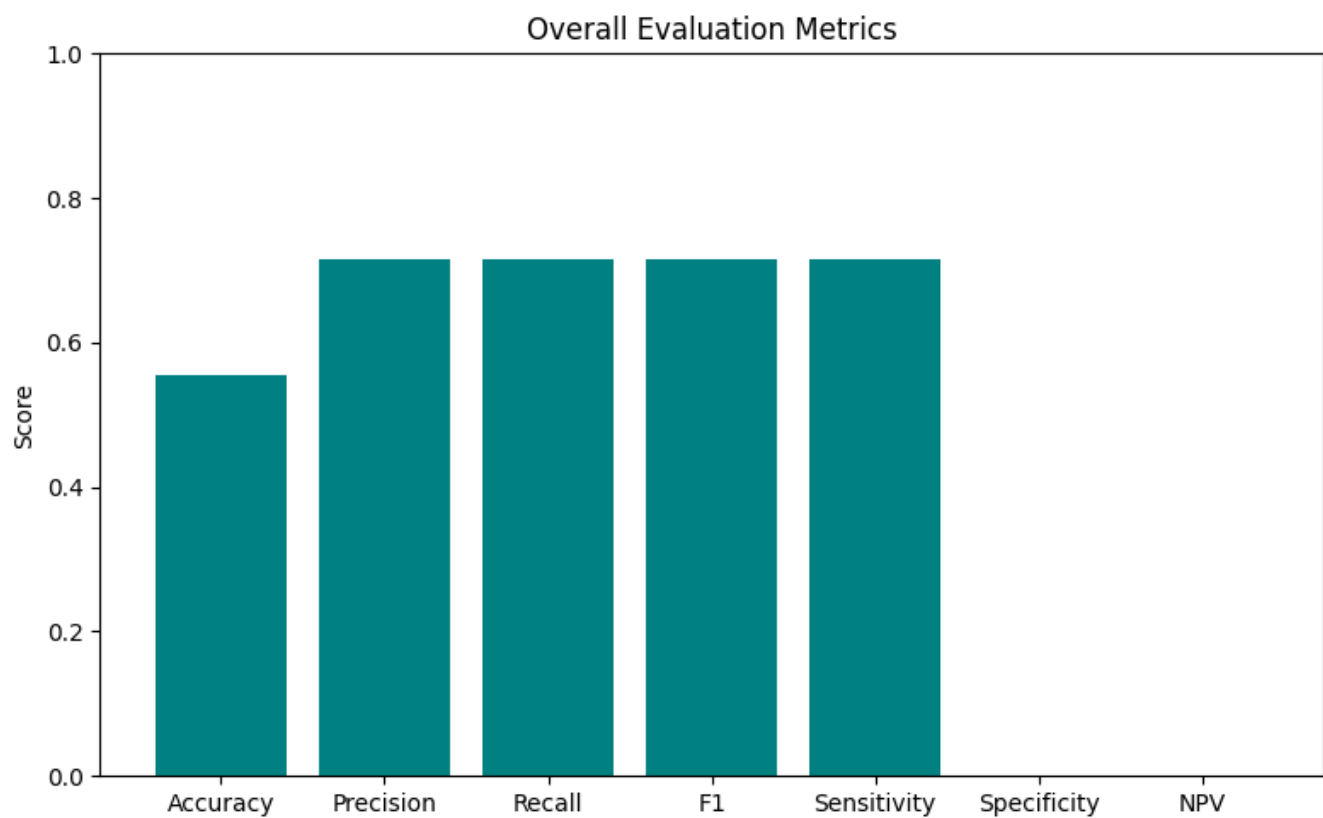
# Fetal Ultrasound AI Evaluation Report

Generated on: 2025-06-17 20:57:44

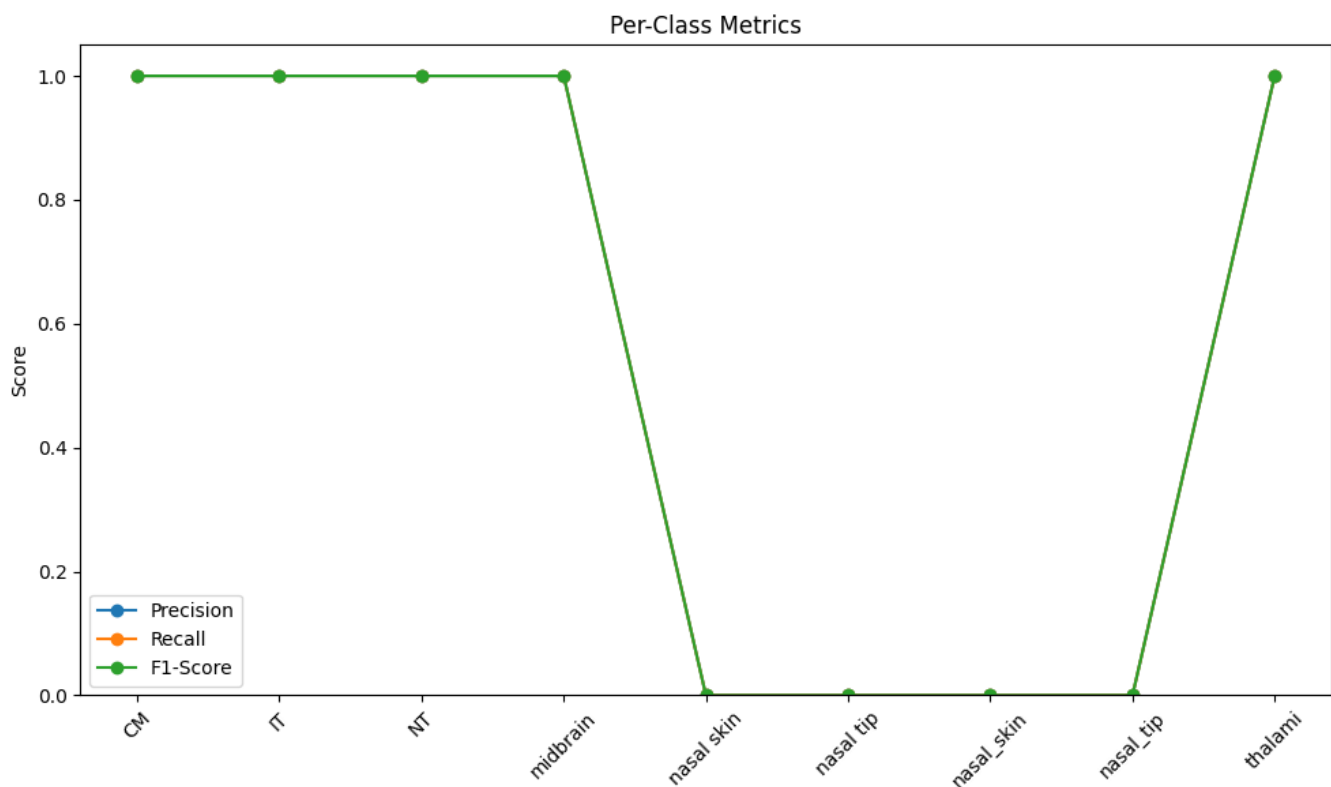
## Prediction Output (Detected Biomarkers)



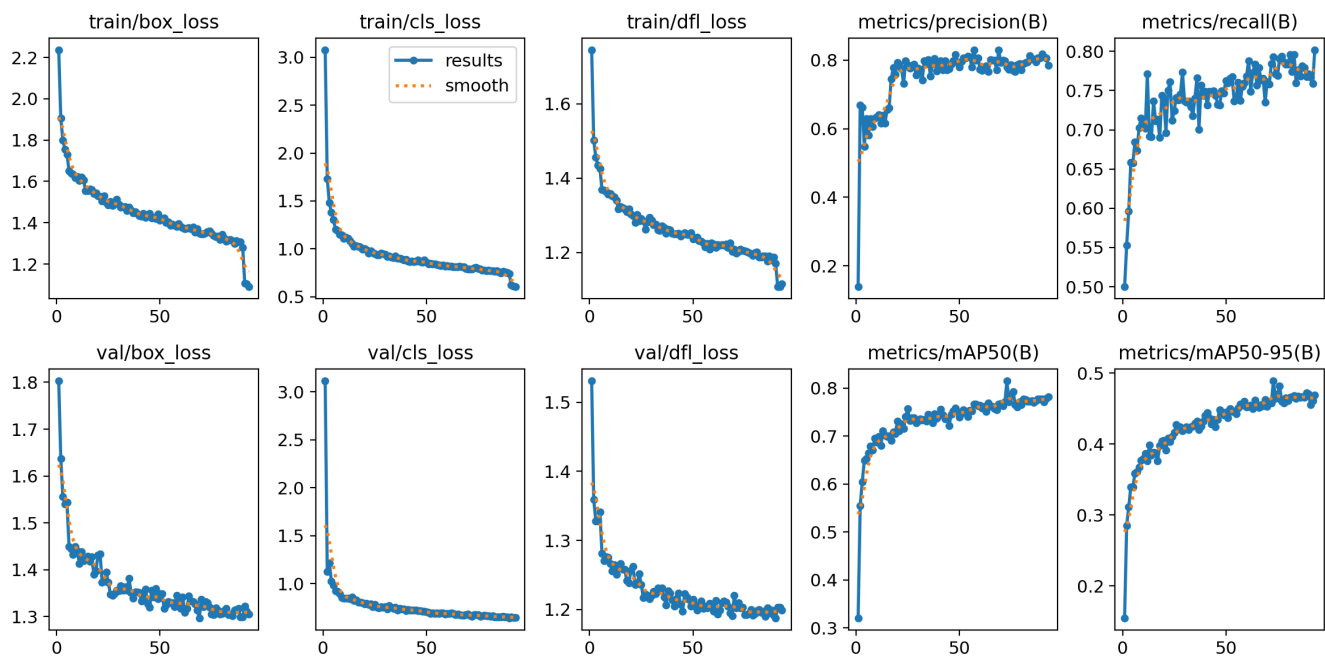
## Overall Evaluation Metrics



## Per-Class Performance (Precision, Recall, F1)



# YOLOv8 Training Performance Graphs



## Understanding the Training Graphs:

- train/box\_loss: Measures error in bounding box placement. Lower is better.
- train/cls\_loss: Classification error during training. Lower means better class accuracy.
- train/dfl\_loss: Distribution Focal Loss for precise bounding box edge alignment.
- val/box\_loss: Bounding box error on unseen validation images.
- val/cls\_loss: Classification loss on validation set. Rising values may indicate overfitting.
- val/dfl\_loss: Validation version of DFL loss.
- metrics/precision(B): Measures how accurate the positive predictions are. High = fewer false positives.
- metrics/recall(B): How many true objects were found. High = fewer missed detections.
- metrics/mAP50: Average precision at 50% IoU. High values indicate good detection.
- metrics/mAP50-95: Mean Average Precision across 10 IoU thresholds. Higher = better generalization.

## Summary Notes:

- This prediction was made using a YOLOv8 object detection model trained on fetal ultrasound images.
- Detected structures were matched with ground truth annotations for evaluation.
- Metrics reported include: Accuracy, Precision, Recall, F1-Score, Sensitivity, Specificity, NPV.
- The bar and line charts visualize model performance.

- Confusion matrix is omitted because multi-label classification makes it less meaningful.