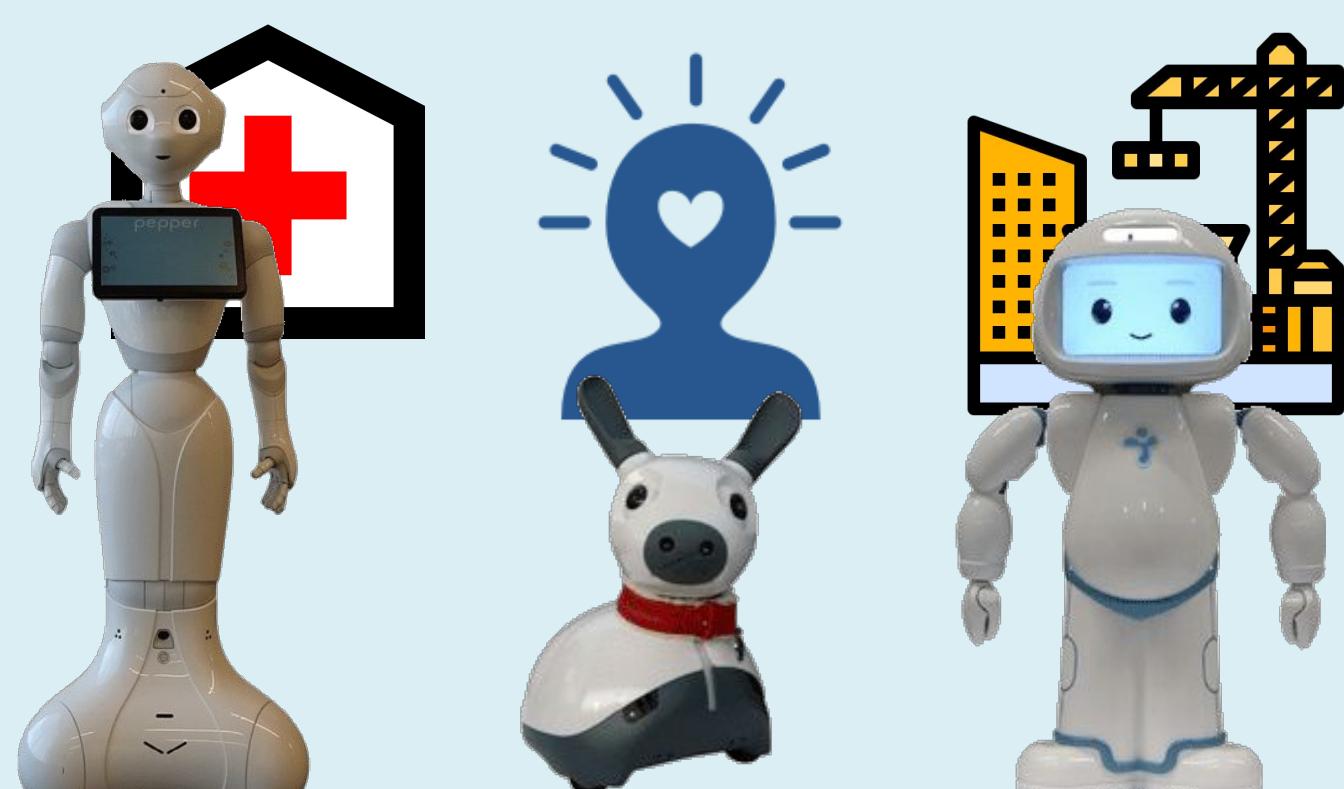




¹University of Richmond, ²University of Nebraska-Lincoln

Introduction



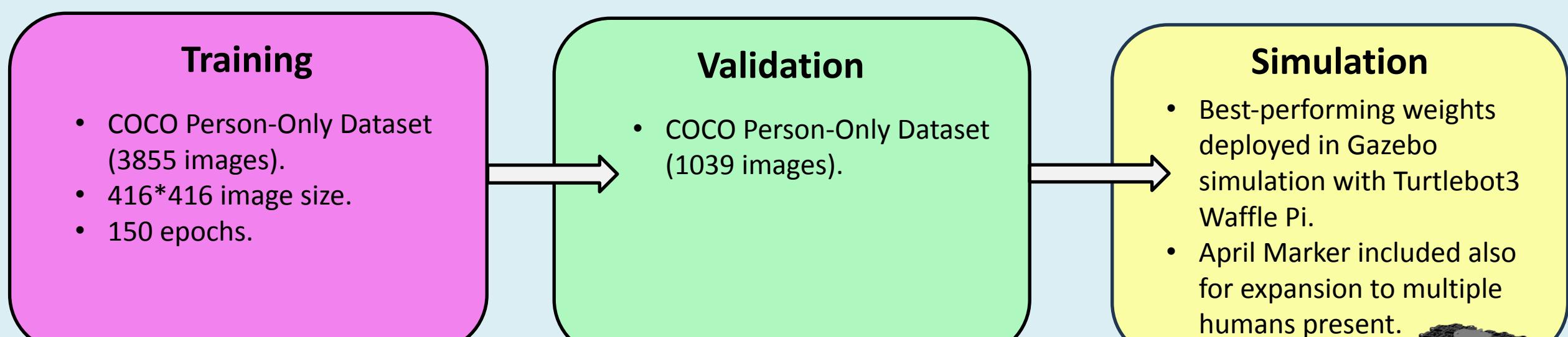
- **Human-robot interaction:** study of relationships between humans and robots.
- Applications range from facilitating socialization and rehabilitation to providing caregiving support.
- Applications frequently involve **person-following scenarios**, where humans and robots collaborate on tasks involving the robot's proximity to a designated person.

Problem: Many existing approaches for person-following rely on computationally expensive algorithms, making them impractical for energy-constrained unmanned ground vehicles.

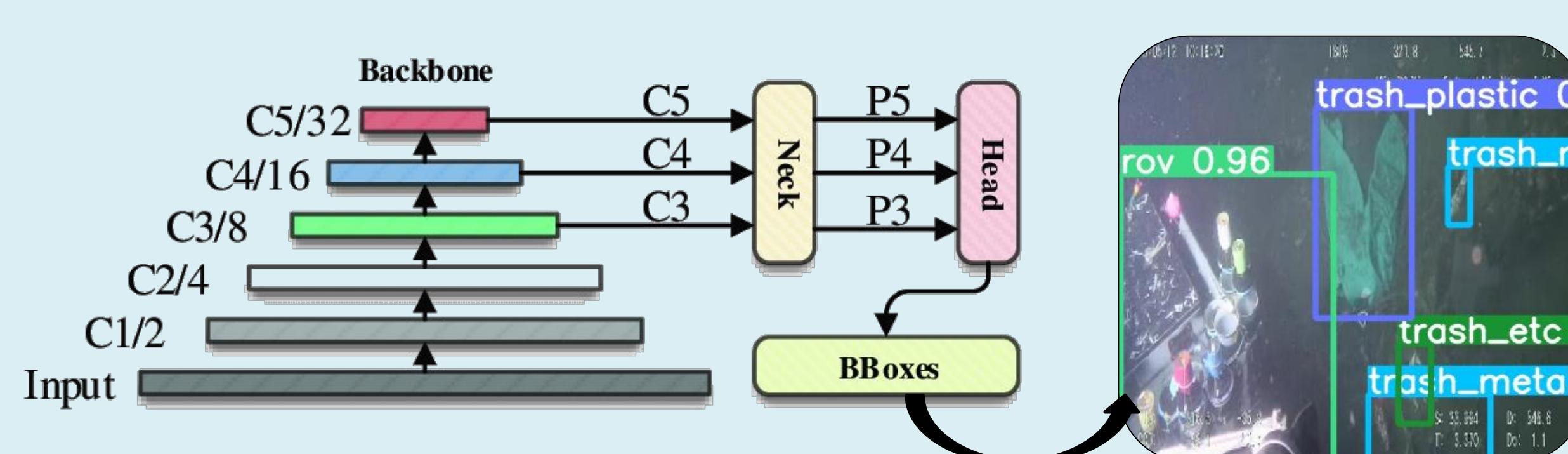


Approach

Objective: Evaluate three object detection algorithms with varying levels of energy efficiency.



YOLOv5 Basic Architecture:



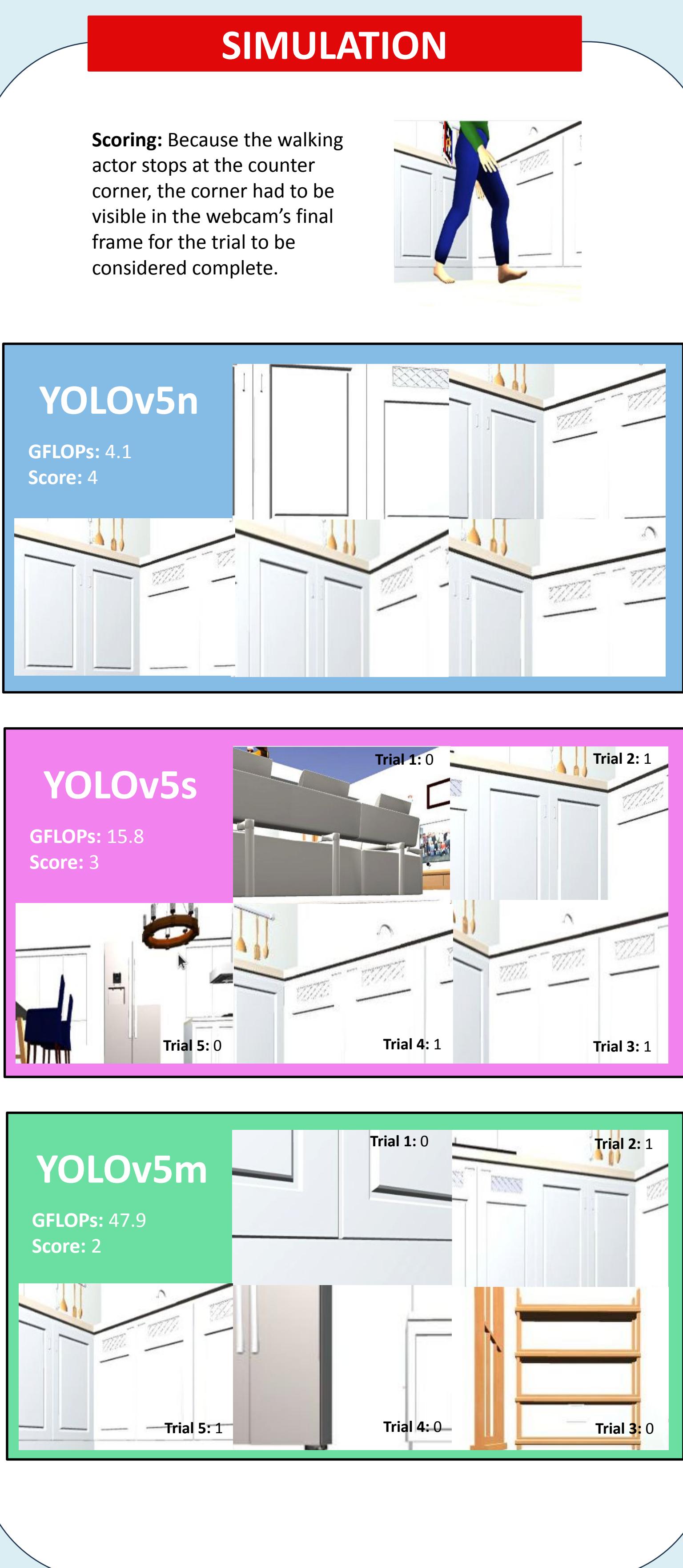
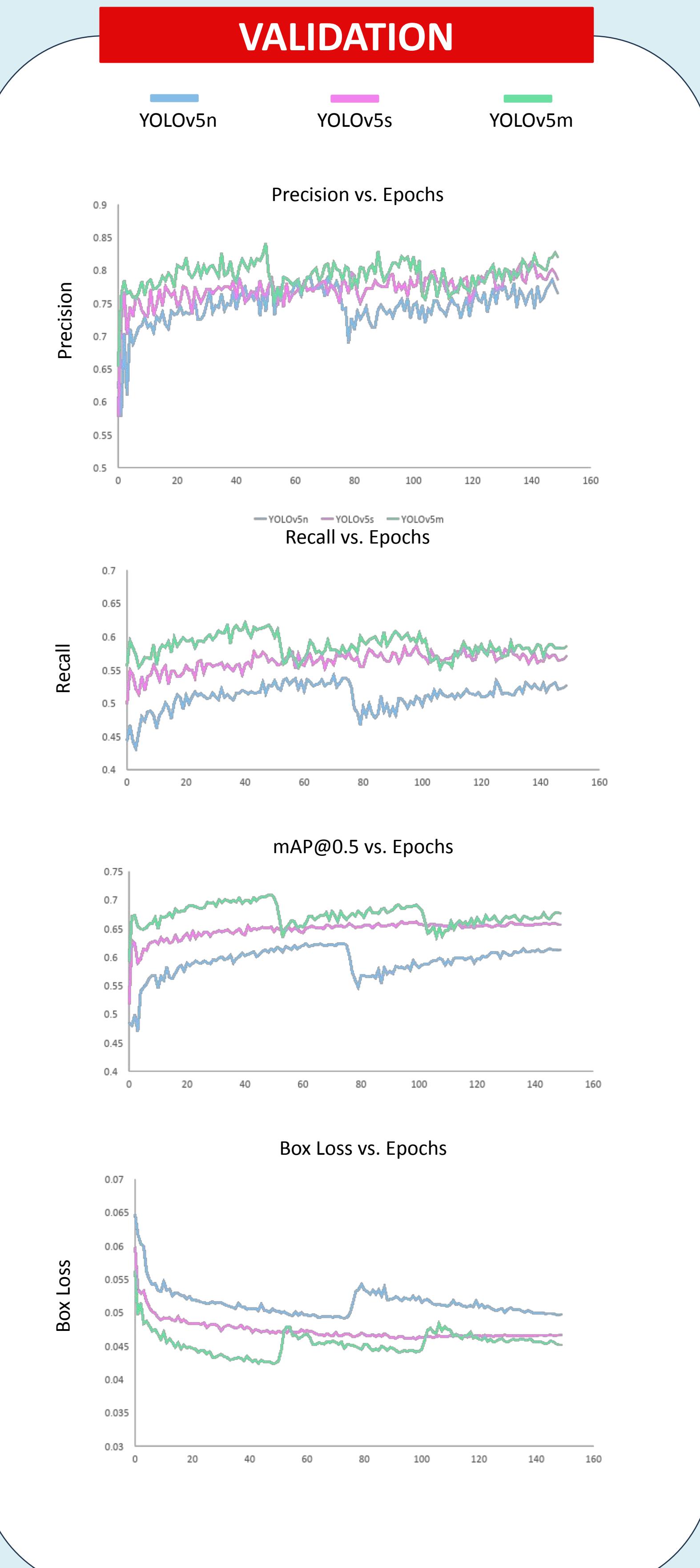
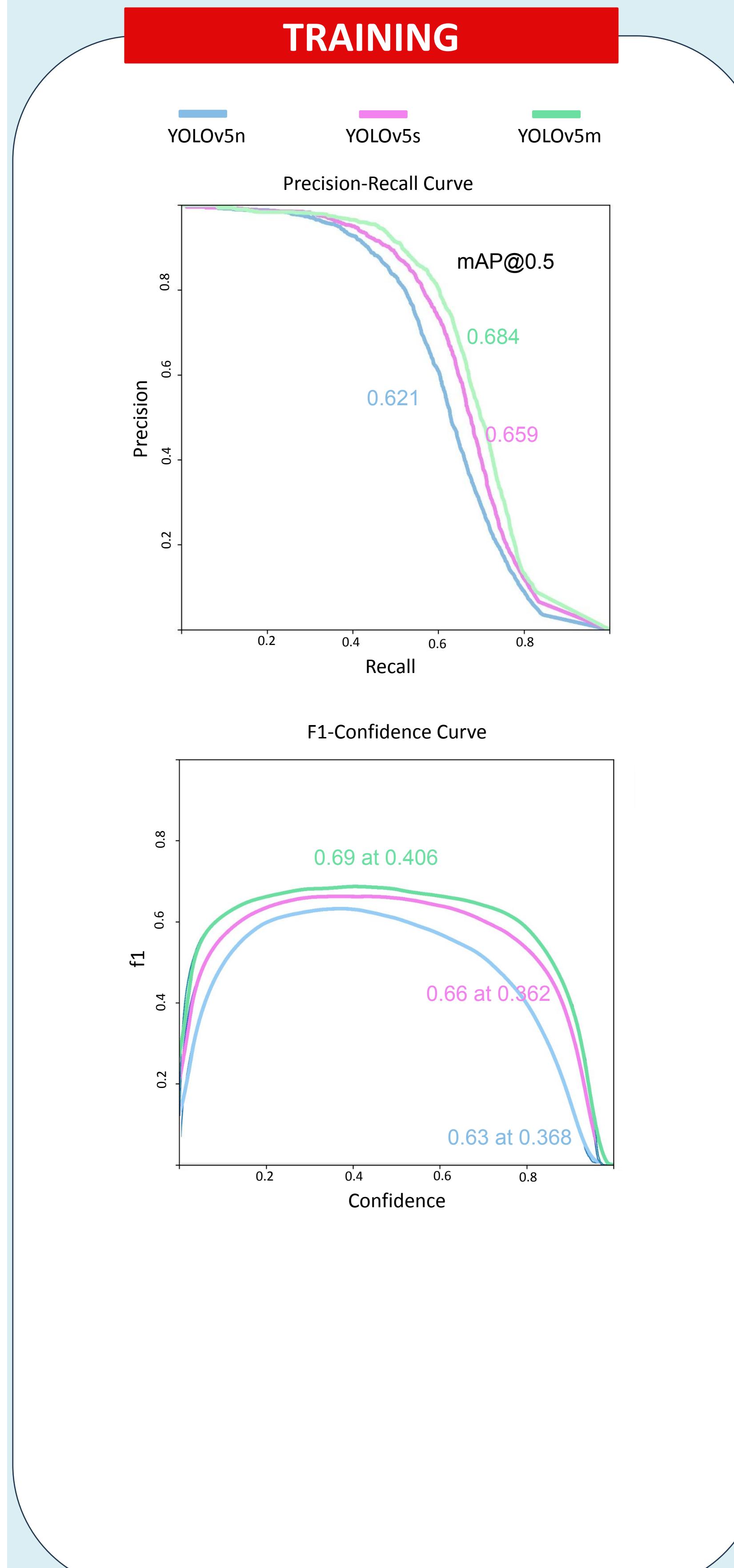
YOLOv5 Architecture Variations:

Nano YOLOv5	Small YOLOv5s	Medium YOLOv5m
4 MB _{FP16} 6.3 ms _{V100} 28.4 mAP _{COCO}	14 MB _{FP16} 6.4 ms _{V100} 37.2 mAP _{COCO}	41 MB _{FP16} 8.2 ms _{V100} 45.2 mAP _{COCO}

Webcam on Turtlebot in Home Environment:



Results



Conclusions

- Based on lower success rate of YOLOv5m in completing the human-following task compared to YOLOv5n and YOLOv5s, **speed likely should be prioritized over accuracy in this task and tasks like it.**
- April marker code not being tracked was leading cause of task incompleteness due to walking actor passing out of scene, suggesting contribution of lower processing rate.
- Only person or not person is being decided by YOLO, and false positives would not likely have April marker in real-world simulation.

Future Directions

- **Additional Gazebo simulation:** Design and test other applications of human-following program (e.g. hospitals, construction sites) to increase robustness against other distractors and similarly dressed workers.
- **Deployment of code of Turtle Bot.** Adapt scenario to real-world household environment.
- **Experiment with training hyperparameters:** Validate YOLOv5 architectures with different optimization algorithms, learning rates, number of epochs on different scenes.
- **Train other architectures for human identification:** Single-Shot Multibox Detection, other YOLO algorithms (e.g. spiking YOLO, YOLOv8).

