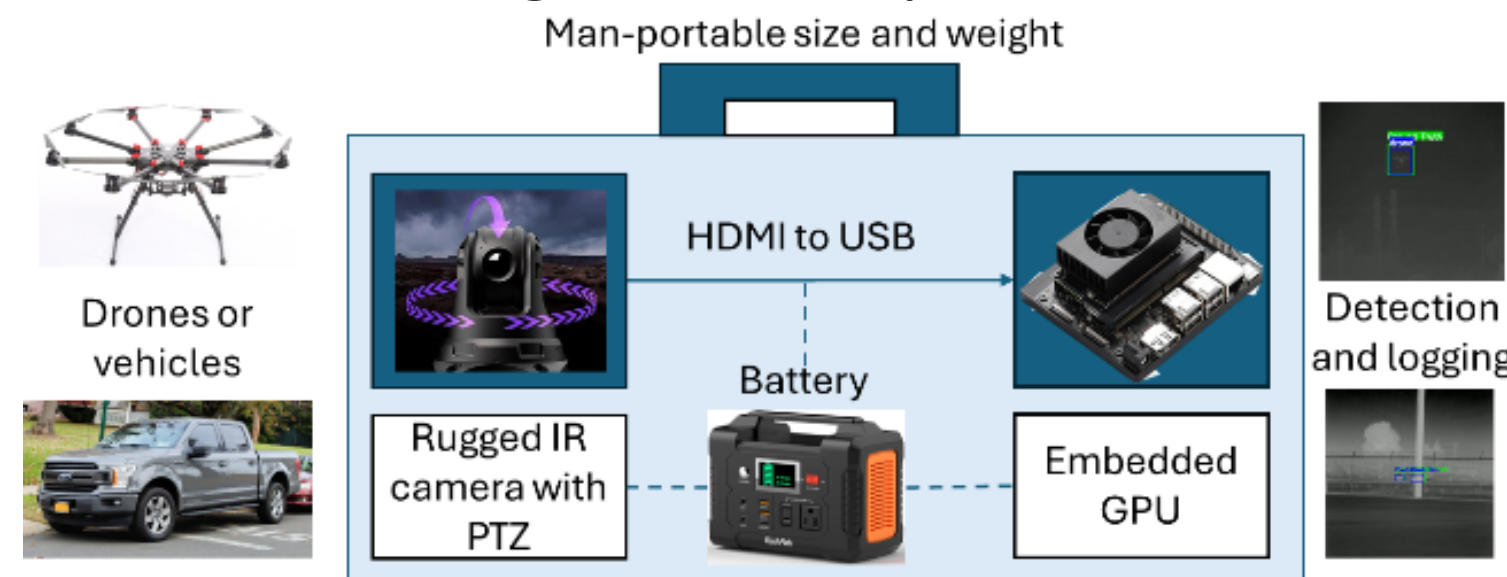


Executive Summary

Prior work at the UT Arlington Research Institute developed a man-portable computer vision system for detecting drones and vehicles in infrared (IR) imagery [1]. Though several experiments were conducted, hyperparameter tuning (HPT) was not performed, limiting detection performance.



To address this, HPT was conducted via a genetic algorithm to find good training-time parameters, followed by a grid search to find testing-time parameters. mAP50-95 accuracy went from 0.54 to 0.59, representing nearly a 10% performance gain over the work of Joshi et al.

Future work includes further HPT and exploring additional NVIDIA Jetson model conversion techniques such as post-training quantization (PTQ).

Background

The rapid increase in the use of small Unmanned Aerial Systems (sUAS) has introduced significant security vulnerabilities in public and restricted spaces. These platforms, while beneficial for several applications, can also be used for unauthorized surveillance or kinetic strikes. Standard electro-optical (EO) cameras often fail in low-light, fog, or high-glare environments. To address this, a Dark 30 Long-Wave Infrared (LWIR) camera, which provides critical day-night visibility was used. An embedded NVIDIA Jetson Orin and a field battery was also used.

Optimization via Hyperparameter Tuning (HPT)

While previous work established a baseline, this study improves robustness by transitioning from manual configuration of hyperparameters to automated Hyperparameter Tuning (HPT). This approach maximizes mAP accuracy scores, ensuring better detection for real-time threat mitigation.

Genetic algorithms have been used to tune neural networks in other contexts as well (e.g., [2]).

Experimental Setup

Dataset and Computer Vision Model: This research utilizes a custom infrared drone dataset developed at UTARI. To ensure model robustness and generalization, standard data augmentations and a train/test split were used.

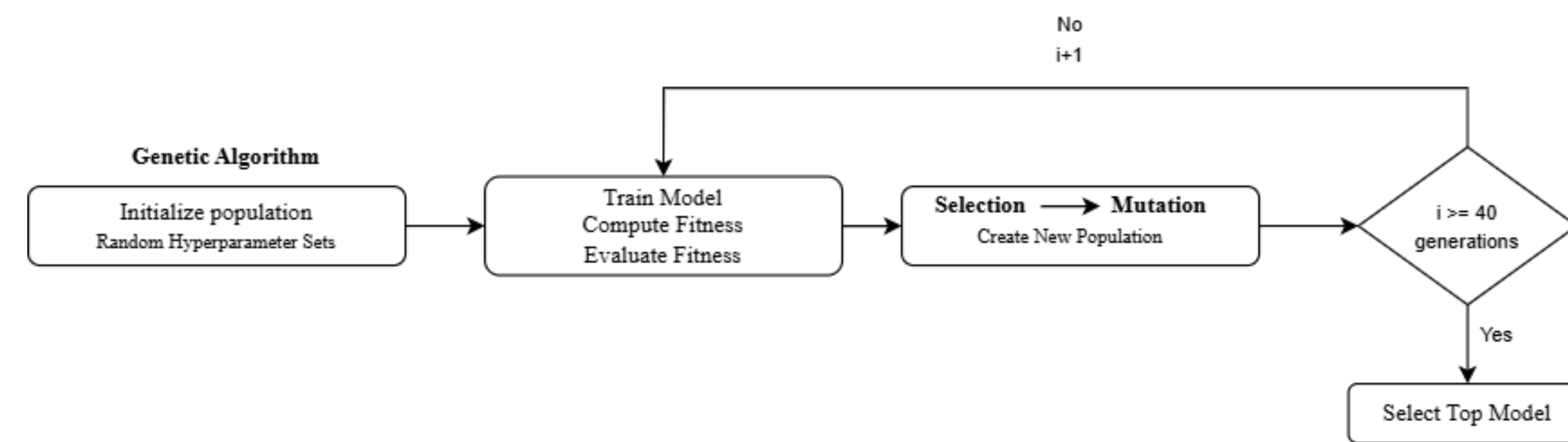
Experiments tested several computer vision models and concluded by using a YOLOv12 model on a desktop and embedded GPUs.

Hyperparameter Tuning (HPT):

A genetic algorithm (GA) was first used to efficiently explore the space for training-time parameters [3].

After this, a grid search for testing-time parameters was performed to maximize detection performance.

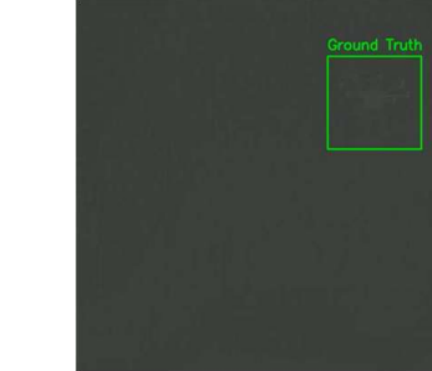
This two-stage approach combines GA's efficient global exploration with grid search's precise threshold optimization.



Training image



Testing image



The dataset is challenging due to low contrast, changes in scale, etc.

Experimental Test Plan

The training time search space focused on three critical parameters designed to maximize model convergence and generalization:

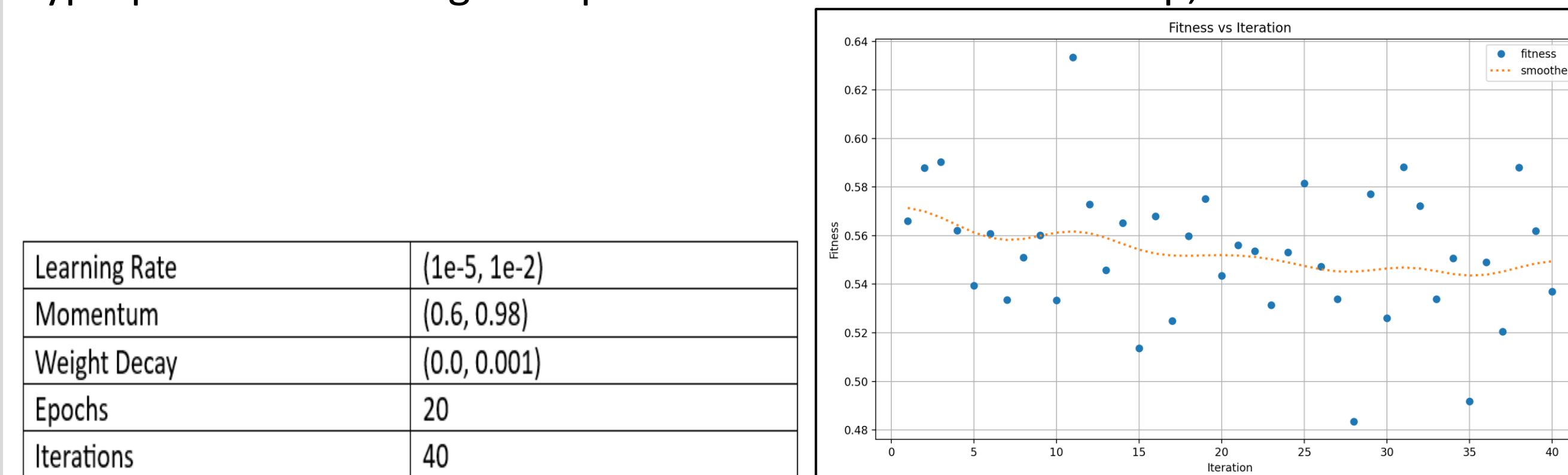
- Learning Rate:** Dictates the step size at each iteration while traversing the loss landscape toward a minimum.
- Momentum:** Accumulates past gradients to smooth updates and reduce oscillations, especially in ravines.
- Weight Decay:** Regularizes training by penalizing large weights, reducing potential overfitting.

Evaluation Metric: Fitness

Fitness is a single scalar value used by the GA to rank the best hyperparameter combinations. **Fitness** = $0.1 * mAP50 + 0.9 * mAP50-95$

The testing time search space optimized **confidence** and **intersection over union** (IoU) thresholds. A detection counts if the network's confidence is above the first threshold and its IoU with ground truth is also above the second threshold.

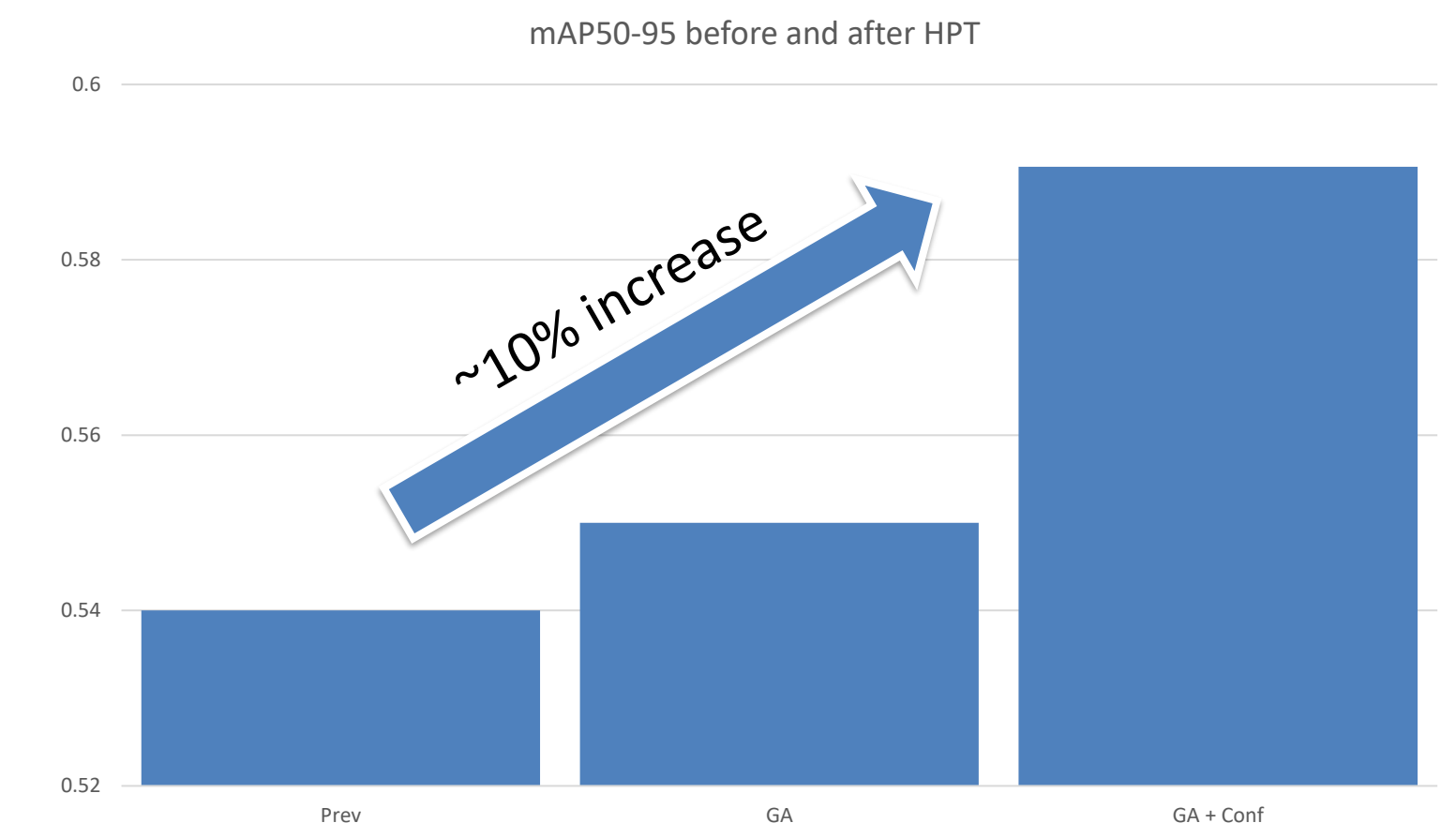
Hyperparameter tuning took place on a dual **RTX 5090** setup, and took **~16 hours**.



Experimental Results

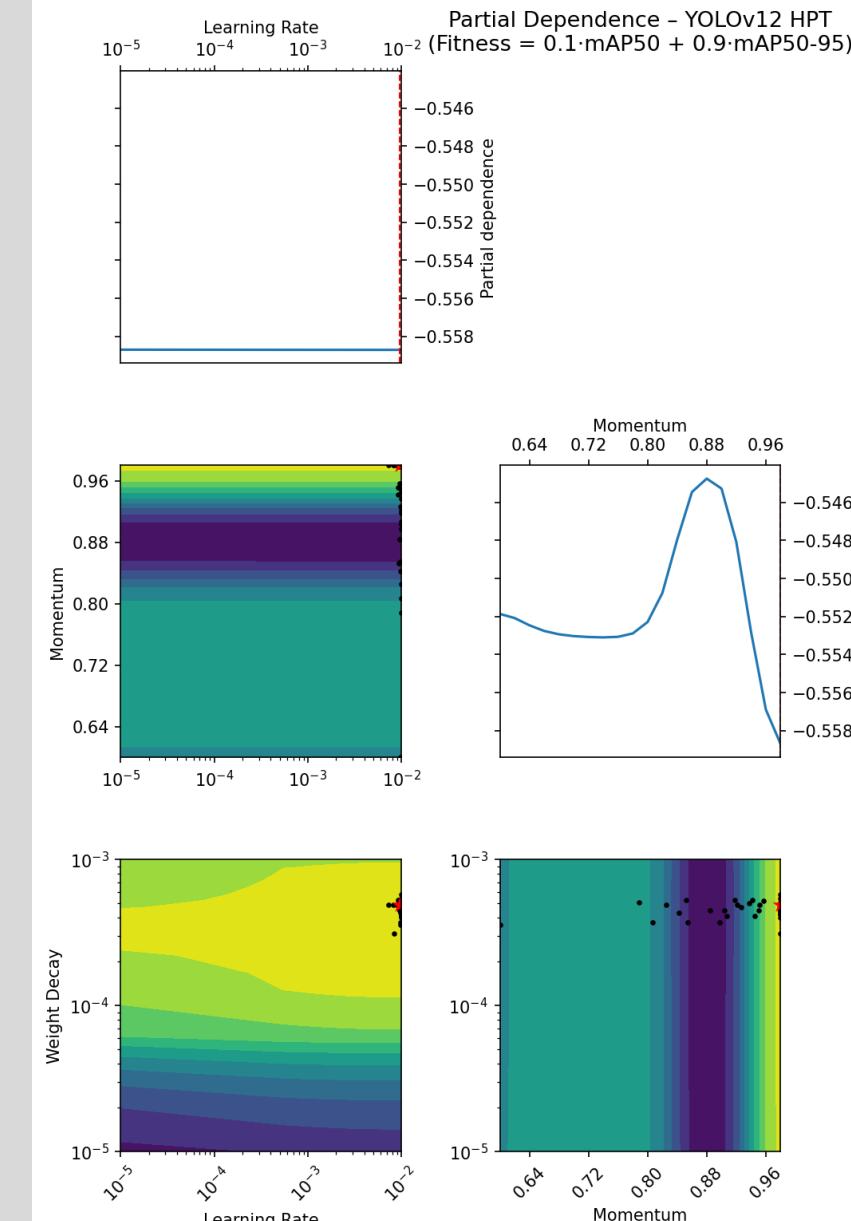
Following the extraction of the optimal hyperparameters from the genetic algorithm, the model underwent secondary tuning -- a grid search for the confidence and intersection over union (IoU) thresholds.

This combined strategy of a genetic algorithm search and post-training threshold calibrations generated a substantial performance lift, resulting in a final mAP50-95 of 0.59.



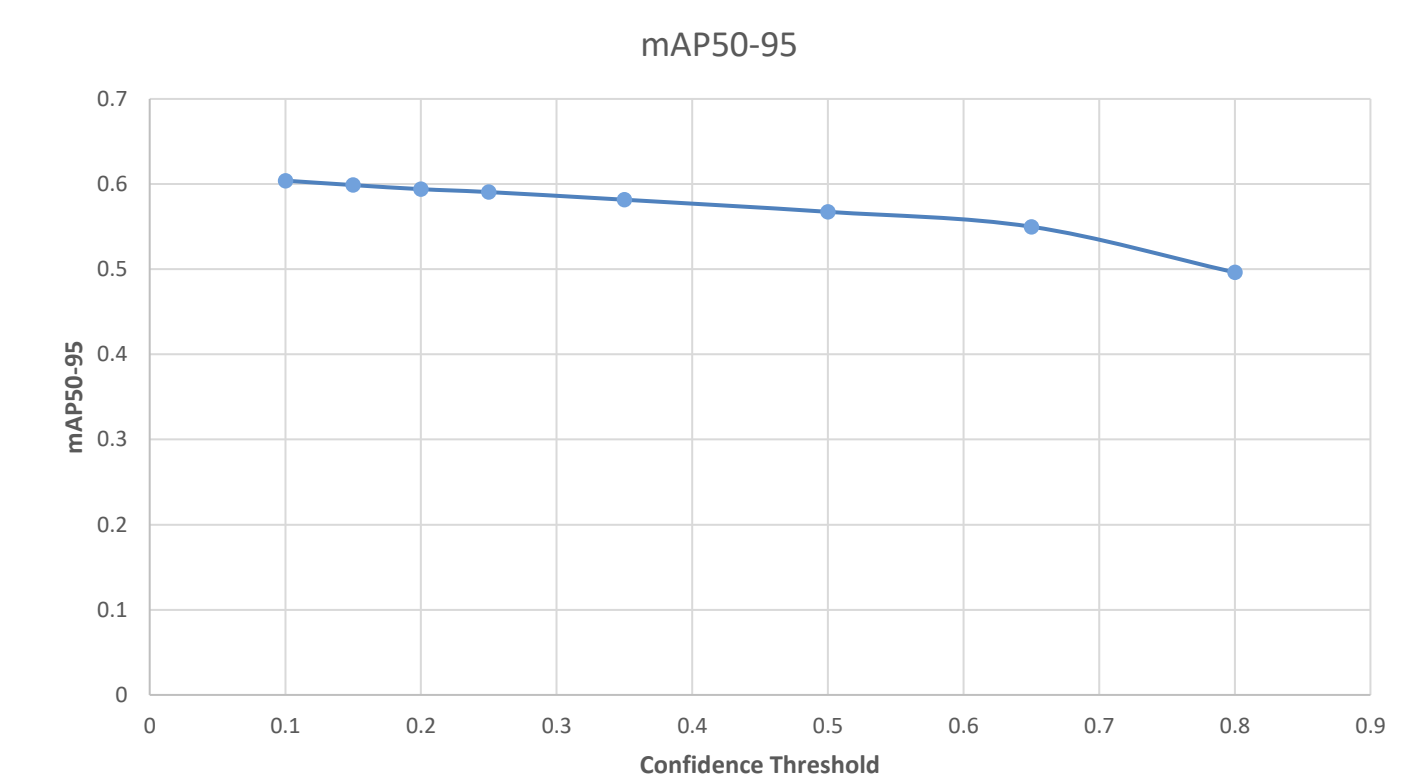
Training time parameters

- * Learning rate: low to medium->better
- * Momentum: medium to high->better
- * Weight decay: low to medium->better



Testing time parameters:

- * Lower confidence->higher mAP50-95
- * IoU did not have an effect (not shown)



Conclusions

This work shows that HPT improves small-target detection capabilities—mAP50-95 went up by ~10% from 0.54 to 0.59.

Future work can include additional HPT or exploring additional model conversion methods for the NVIDIA Jetson such as post quantization training (PQT).

References

- Joshi, R., Tambe, V., Lundberg, C., & Chelian, S. (2026) Towards field-deployable, man-portable, infrared target detection of drones and vehicles using transformer-based computer vision models and embedded GPUs. In Proc. Int'l Conf. on AI, Robotics, and Control 2026, Savannah, USA.
- Chelian, S., Goldfarb, S. E., & Bhattacharyya, R. (2018). U.S. Patent No. 10,133,983. Washington, DC: U.S. Patent and Trademark Office.
- <https://docs.ultralytics.com/guides/hyperparameter-tuning/>

Acknowledgements: This project was supported by CSE REU. Dr. Suhas Chelian advised the work (UT Arlington Research Institute, Automation and Intelligent Systems).