# FACETS: Efficient Constrained Iterative NAS for Object Detection in Robotic Perception


**Tony Tran[1] (thtran37@cougarnet.uh.edu), Bin Hu[2,3] (bhu12@uh.edu)**

Dept. of Research Computing[1], Dept. of Engineering Technology[2], Dept. of Electrical and Computer Engineering[3]
University of Houston, Houston, USA


IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION

Scan here to download the paper and poster

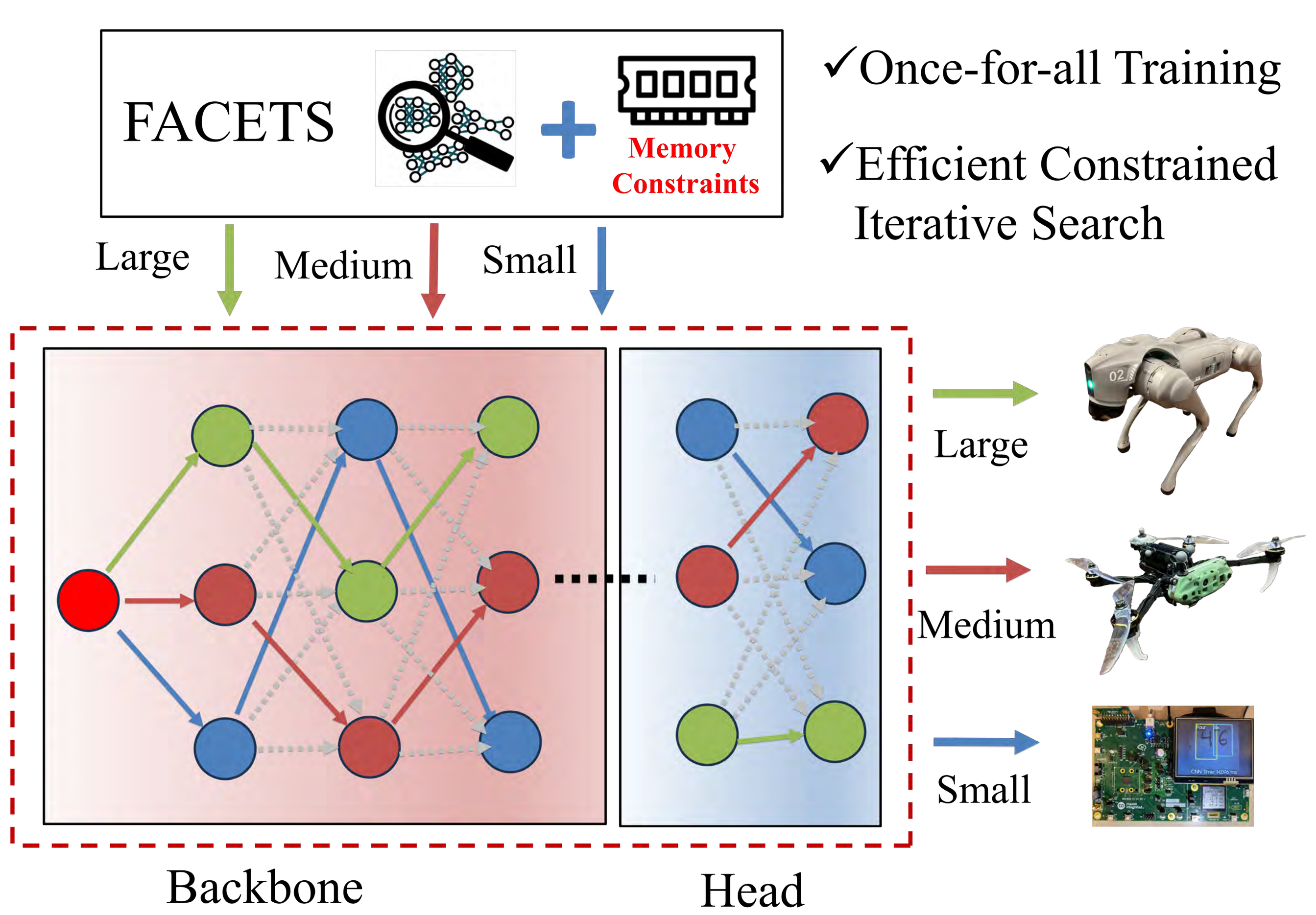
## ABSTRACT


Neural Architecture Search (NAS) for object detection is computationally costly due to multi-modular optimization. FACETS introduces an efficient, iterative NAS method that refines architectures cyclically, alternating between modules (e.g., backbone, head) with *Population Passthrough*. It achieves **4.75% higher accuracy twice as fast** as progressive methods and refines search spaces with **27% higher mean accuracy** than global search. On the MAX78000, FACETS **reduces energy by 45.4% and latency by 29.3%**.


### Challenges

- ✓ *Optimization Complexity*: growing search space exponentially increases optimization difficulty
- ✓ *Heterogenous Search Space*: diverse subnetwork configurations challenge efficient optimization
- ✓ *Budget Constraint*: strict memory, power, and compute limits hinder detection model deployment

## FACETS in ROBOTIC PERCEPTION




**FACETS enables flexible and efficient deployment across diverse robotic platforms**

## KEY CONTRIBUTIONS

1. ***Unified Iterative NAS***: Iteratively optimizes subnetworks, **boosting accuracy by 4.75%** over progressive search.
2. ***Fast Search***: Converges **two times faster** with Population Passthrough.
3. ***Better Search Space***: Refined space delivers **27% higher mean accuracy** vs. global space.
4. ***Efficient Deployment***: Validated on MAX78000, **cuts energy by 45.4%** and **latency by 29.3%**, **improving accuracy by 4.5%**.

## ITERATIVE SEARCH PROCESS

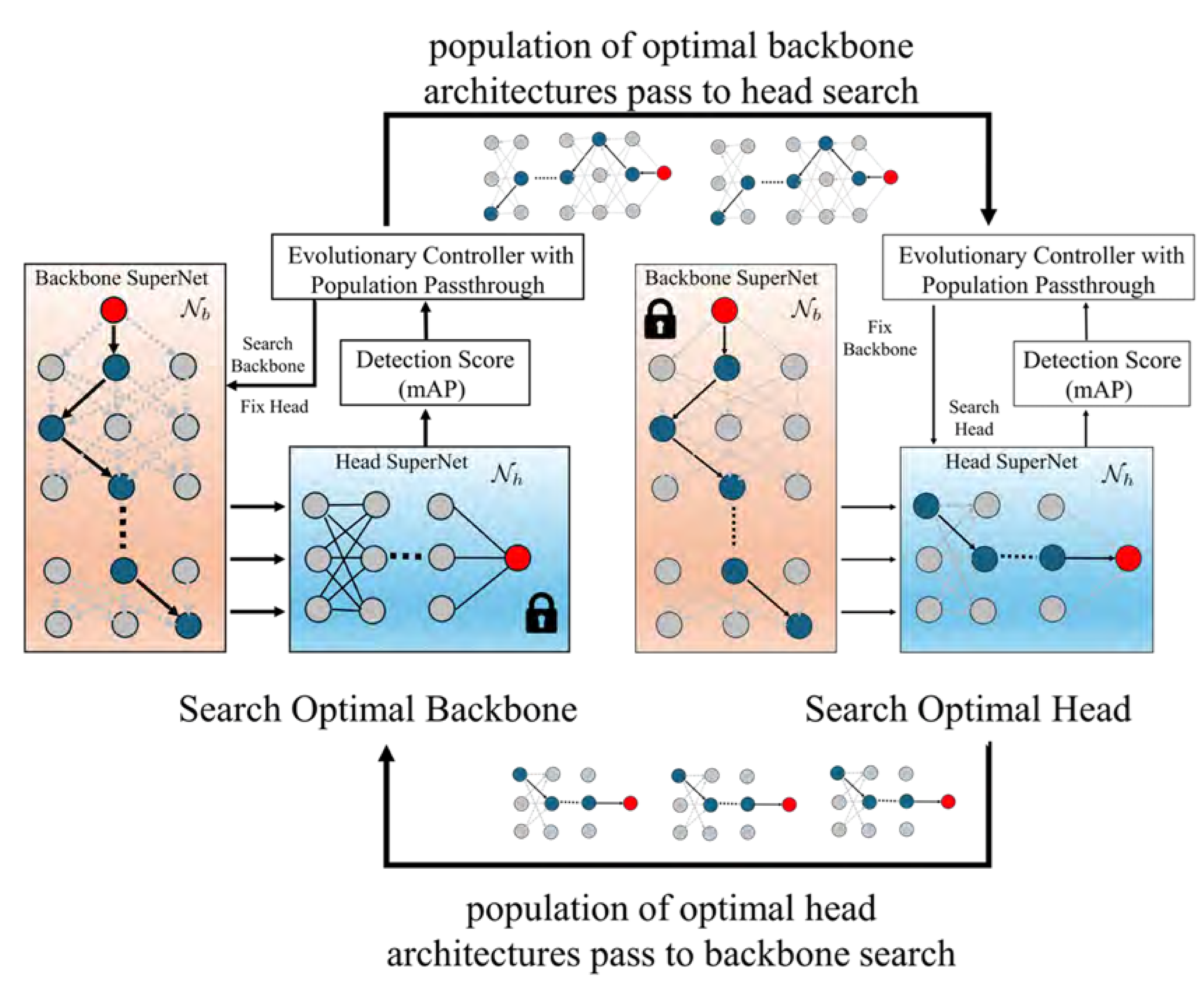
- Step 1: Search optimal backbone b with fixed head

$$\min_{b} \mathcal{L}_{val}^{det}\left[\mathcal{N}\left((b,h^0),(w_b^*,w_h^0)\right)\right]$$
$$s.t.\ w_b^* = \operatorname*{argmin}_{w_b} \mathcal{L}_{train}^{det}\left[\mathcal{N}\left((b,h^0),(w_b,w_h^0)\right)\right]$$

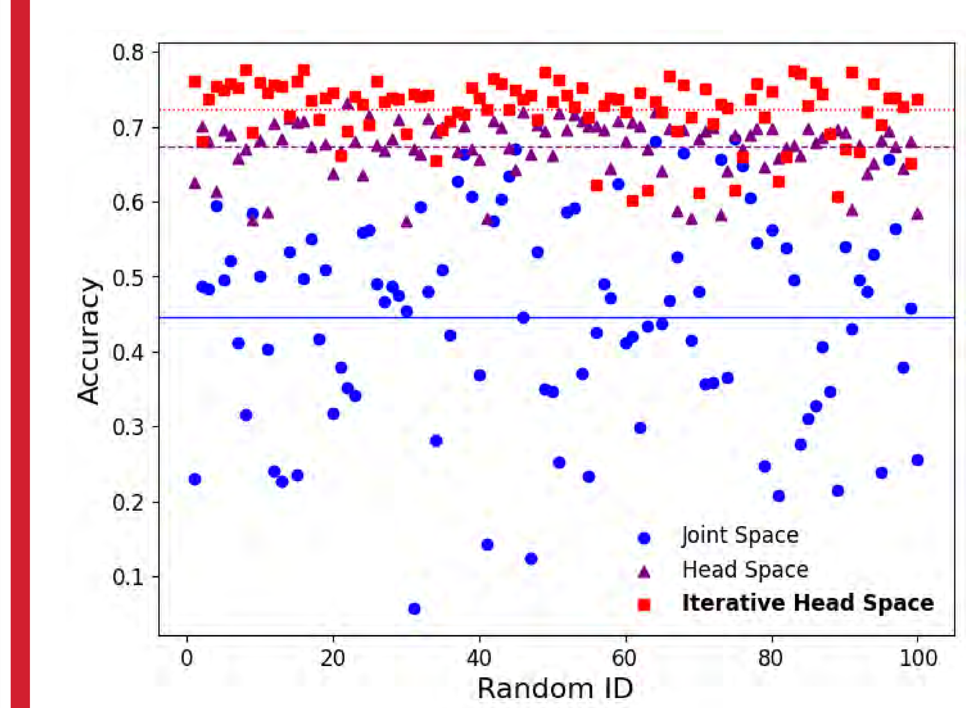
- Step 2: Search optimal head $h$ with fixed backbone

$$\min_{h} \mathcal{L}_{val}^{det}\left[\mathcal{N}\left((b^*,h),(w_b^*,w_h^*)\right)\right]$$
$$s.t.\ w_h^* = \operatorname*{argmin}_{w_h} \mathcal{L}_{train}^{det}\left[\mathcal{N}\left((b^*,h),(w_b^*,w_h)\right)\right]$$

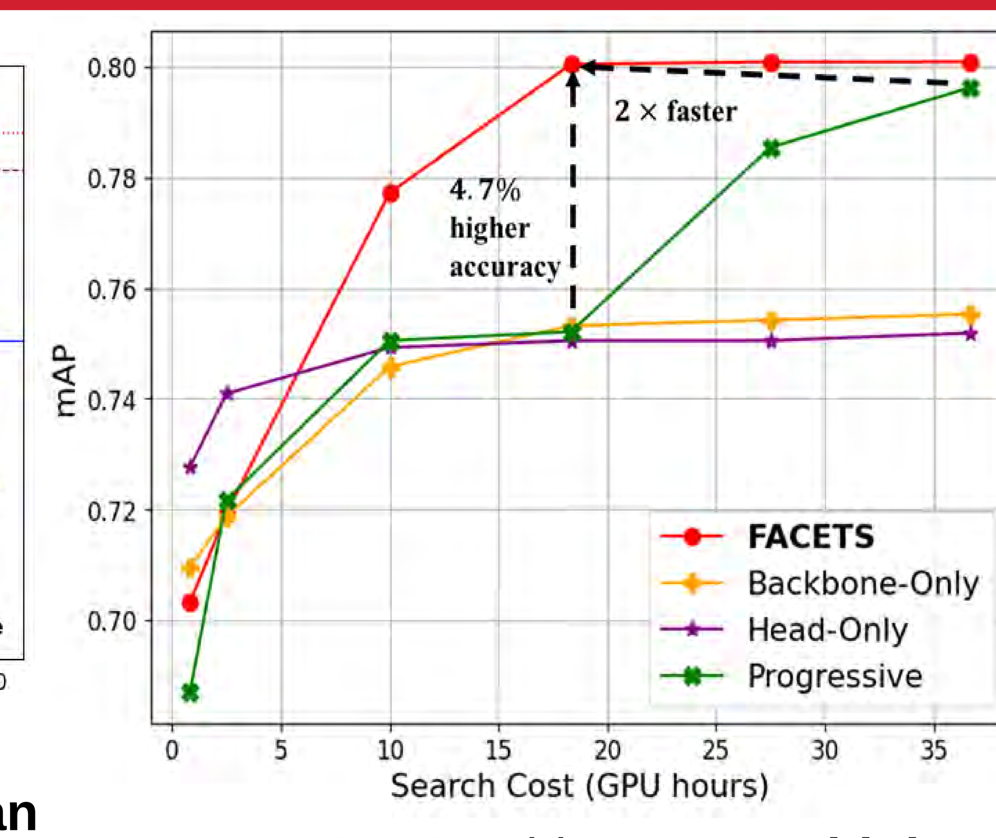
- Repeat Step 1 and Step 2 iteratively until converge



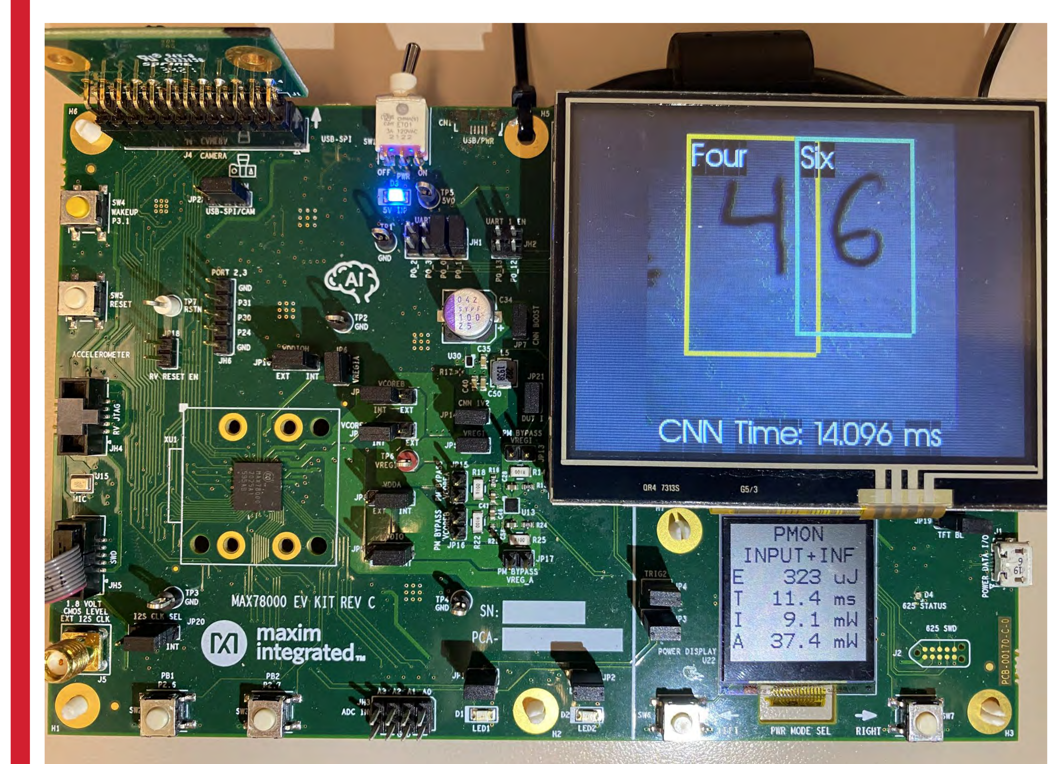

## PERFORMANCE HIGHLIGHTS




FACETS yields **27% higher mean accuracy** over global in randomly sampled architectures




FACETS achieves **4.7% higher accuracy twice as fast** as progressive methods




FACETS Deployment on MAX78000 with CNN Accelerator (**432 KB memory limit**)

- FACETS outperforms SOTA TinierSSD on MAX78000 by
  - ✓ **45.4%** less energy
  - ✓ **29.3%** less latency
  - ✓ **4.5%** improved accuracy

*Performance and Efficiency Comparison on MAX78000 Platform*

| Model | Size (kB) | Energy (uJ) | Latency (ms) | Accuracy (mAP) |
|---|---|---|---|---|
| TinierSSD | 336 | 570 | 14 | 83.6 |
| **FACETS** | **391** | **311** | **9.9** | **88.1** |

## ACKNOWLEDGEMENTS


The authors gratefully acknowledge the financial support of National Science Foundation Award (IIS -2007386)