TrashDet: Iterative Neural Architecture Search for Efficient Waste Detection

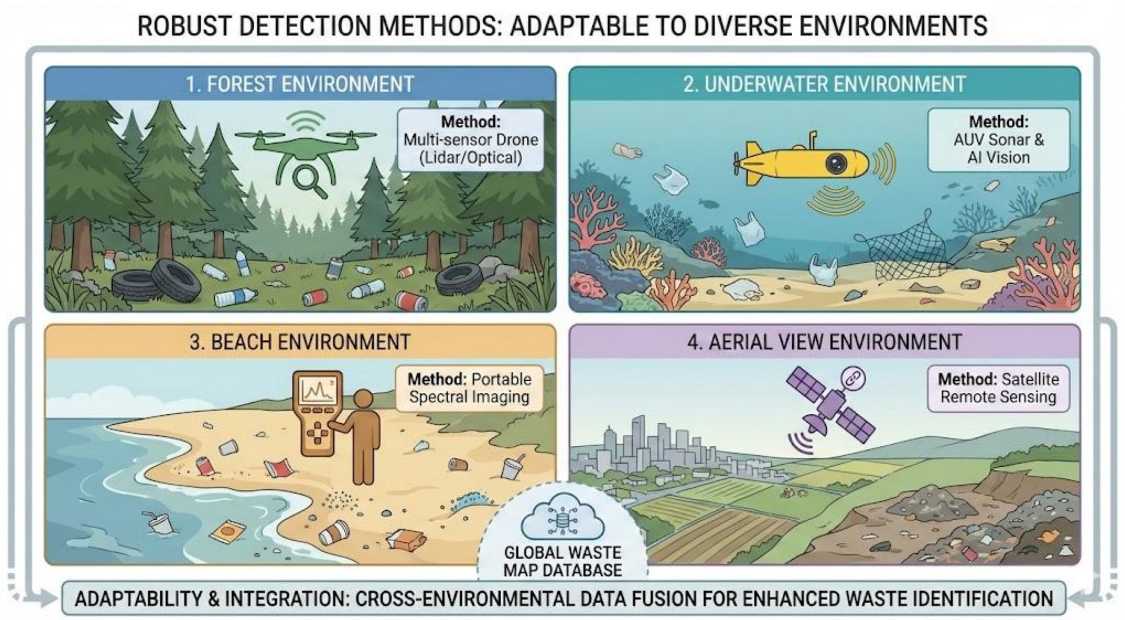
Tony Tran

UNIVERSITY OF HOUSTON

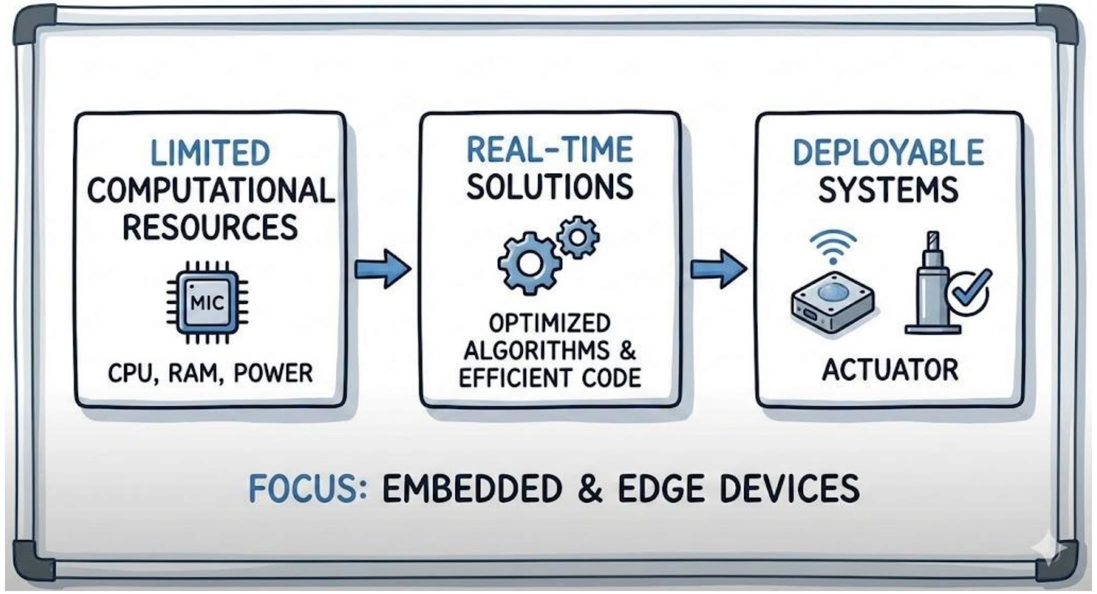
Waste Monitoring on the Edge

UNIVERSITY OF HOUSTON

The global concern regarding waste management underscores the need for new intelligent monitoring systems through Computer Vision methods



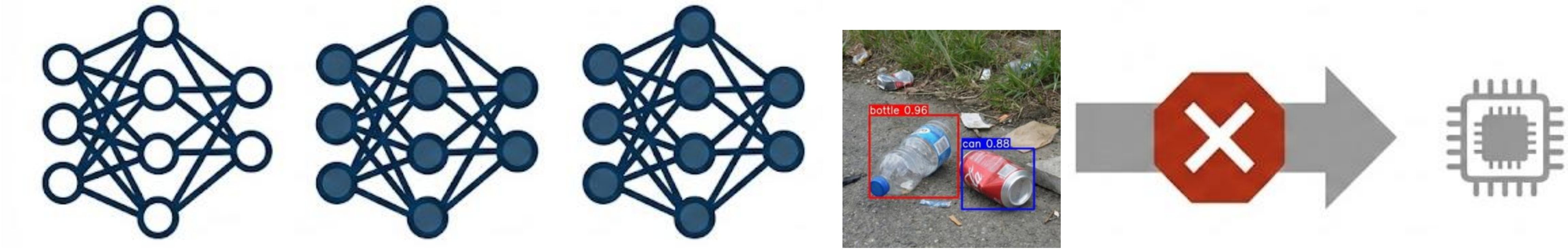
Need for Robust Detection Methods Adaptable to Diverse Environments



Development of Real-Time Solutions Deployable on Systems with Limited Computational Resources

This work aims to develop a real-time waste monitoring system for diverse environments, leveraging computer vision on resource-constrained edge devices to enable reliable general waste detection

Existing SOTA Waste Detectors



YOLOv8m
16.6 mAP50 @21M

RTMDet-l
16.9 mAP50 @52M

AltiDet-m
18.4 mAP50 @85M

Resource
Constrained Devices

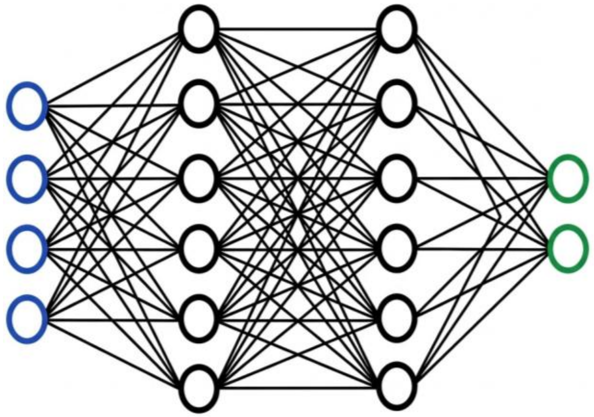
State-of-the-Art Models (e.g., YOLO, RTMDet, AltiDet) achieve good results on waste detection tasks but typically **rely on large, high-capacity architectures** with **substantial computational and parameter overhead**.

- **Infeasible Size:** They rely on backbones with tens of millions of parameters, resulting in large memory footprints
- **High Resource Demand:** The reliance on heavy convolutional backbones and dense feature hierarchies result in higher memory usage, increased latency, and elevated energy consumption
- **Conclusion:** These limitations motivate the need for more efficient detection architectures that preserve competitive accuracy while reducing computational costs

Hardware-Aware Iterative NAS Framework

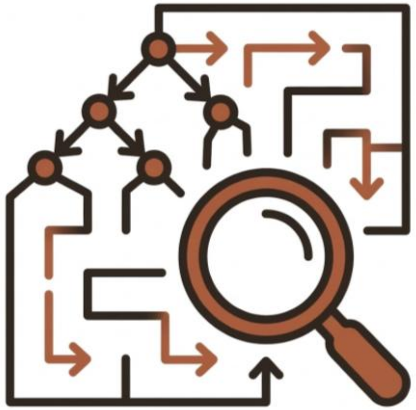
UNIVERSITY OF HOUSTON

Instead of manually designing models, we develop a framework that automatically discovers optimal detector architecture tailored to specific hardware constraints



Hardware-Constrained Supernet

A flexible, overparameterized network that contains many subnetworks, filtered by hardware constraints



Iterative Evolutionary Search

An iterative search algorithm that efficiently explores all supernet modules (i.e. backbone, neck, head) to find the best performing subnet

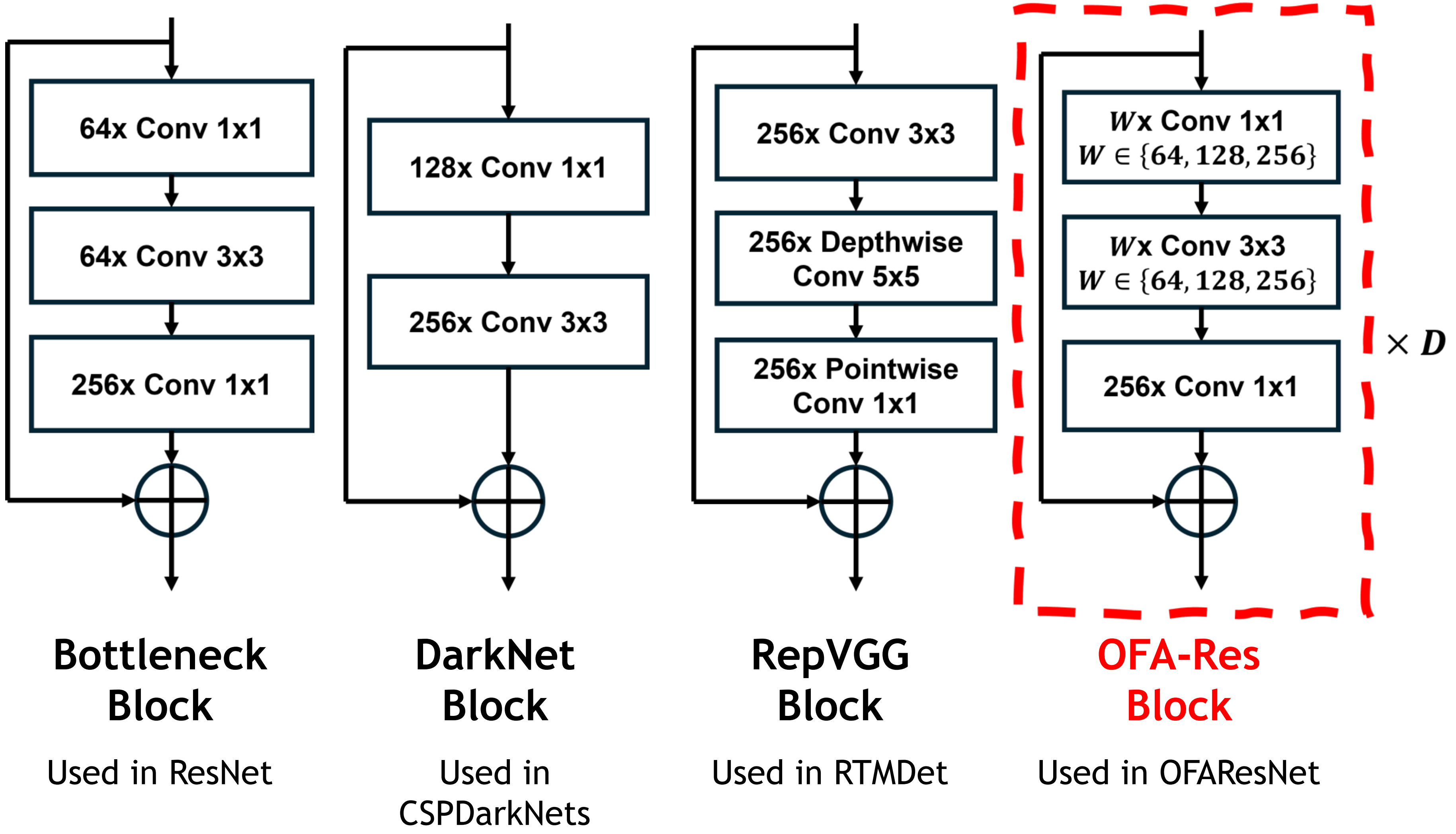
Population Passthrough

A unique search memory mechanism to make the search process stable and cost-effective

*This method yields a family of highly efficient, deployment-ready models we call **TrashDets***

Defining the Search Space: OFA Supernet

The search space explores the OFA-Res Block applied to all components (i.e. backbone, neck, and head) of object detection supernets

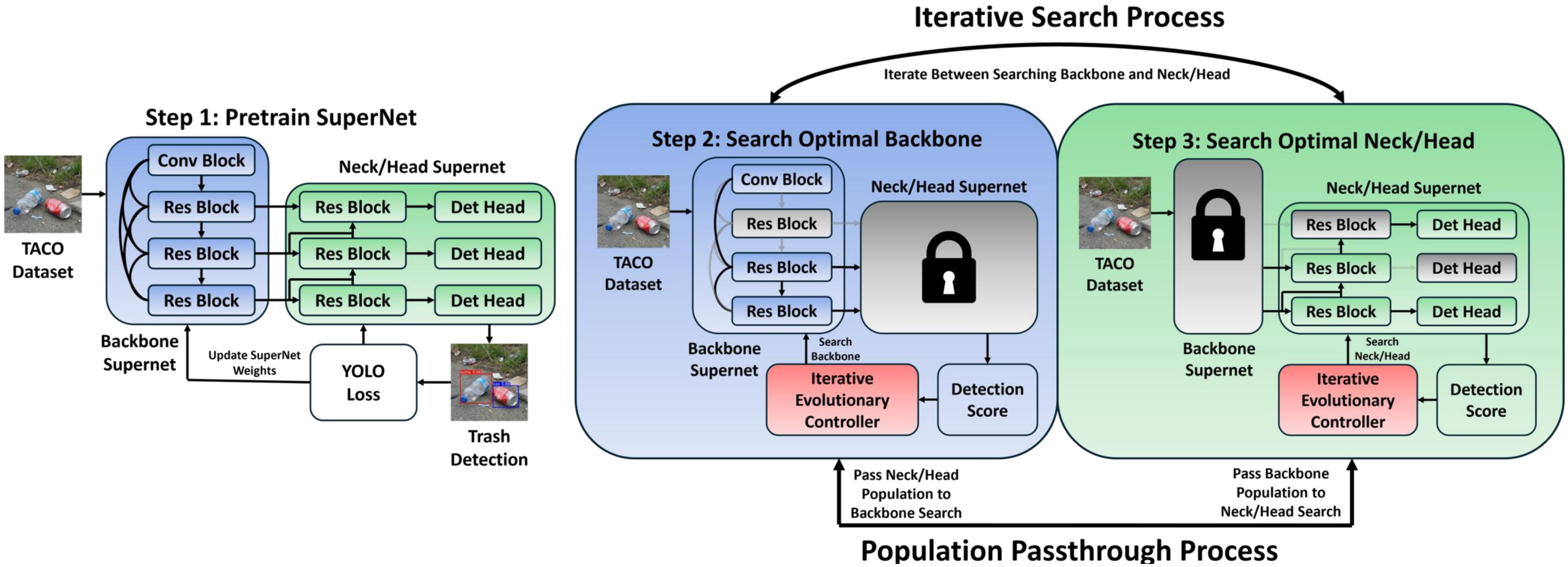


Search Dimensions:

- **Depth**: Number of Active Blocks per Stage (2 to 8)
- **Width**: Channel Scaling Multipliers (e.g. 0.8x, 1.0x, 1.25x)
- **Expansion Ratio**: Controls Intermediate Channels within Blocks

We Apply these Blocks to Parts of the Backbone (OFA ResNet), Neck (OFA PANet), and Head (OFA Yolov3)

Iterative Evolutionary Search Algorithm



Step 1:

A large, flexible supernetwork is pretrained on the TACO dataset, containing all possible architectures

Step 2:

The search algorithm freezes the neck/head and searches for the highest performing backbone configuration, subject to predefined hardware constraints

Step 3:

With the best backbone from the previous step fixed, the search algorithm searches for the optimal neck/head configuration, subject to predefined hardware constraints

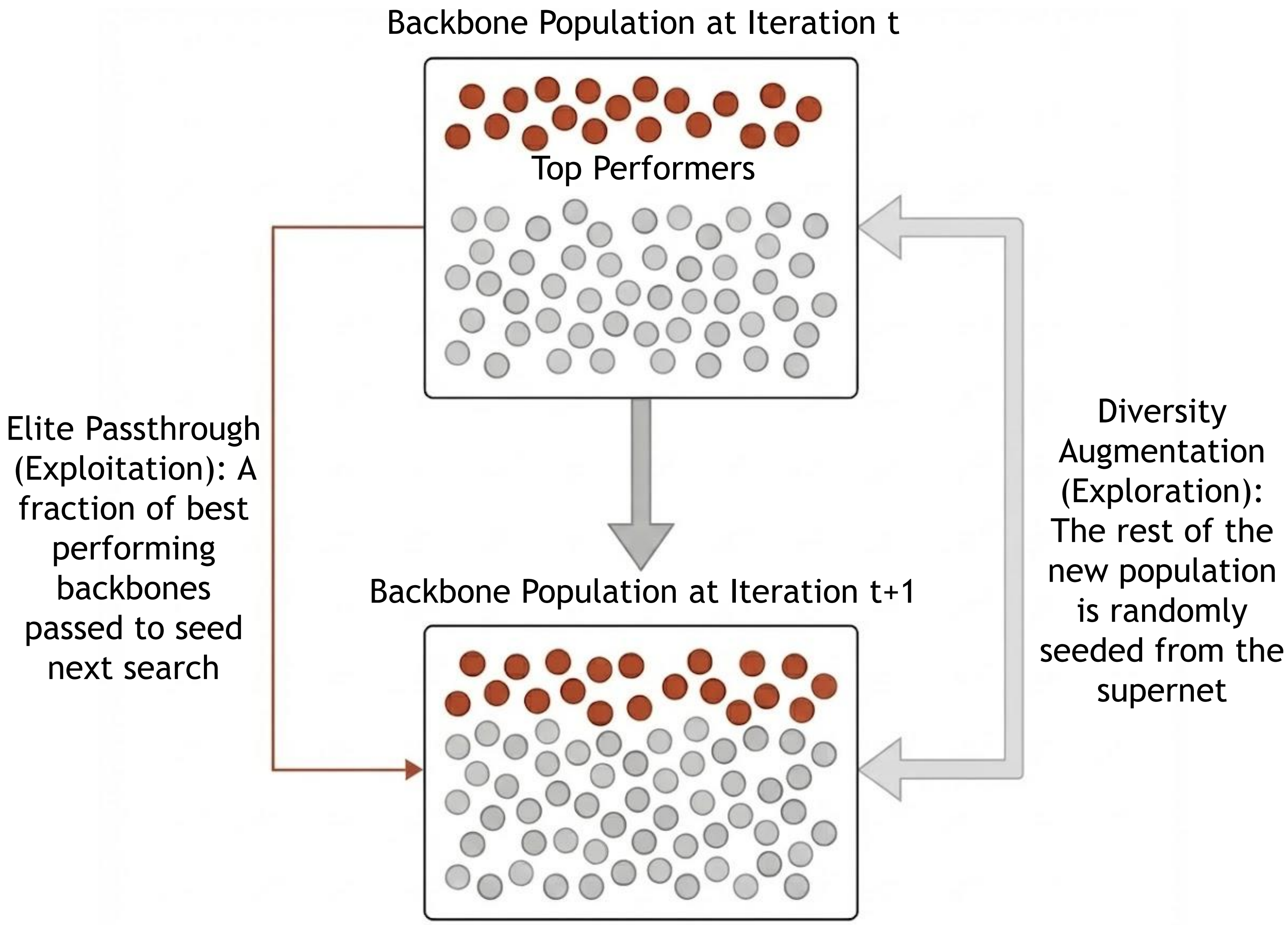
Stabilizing Search: Population Passthrough

The Challenge

When alternating between searching the backbone and the neck/head, the search process can become unstable, losing valuable progress

The Solution

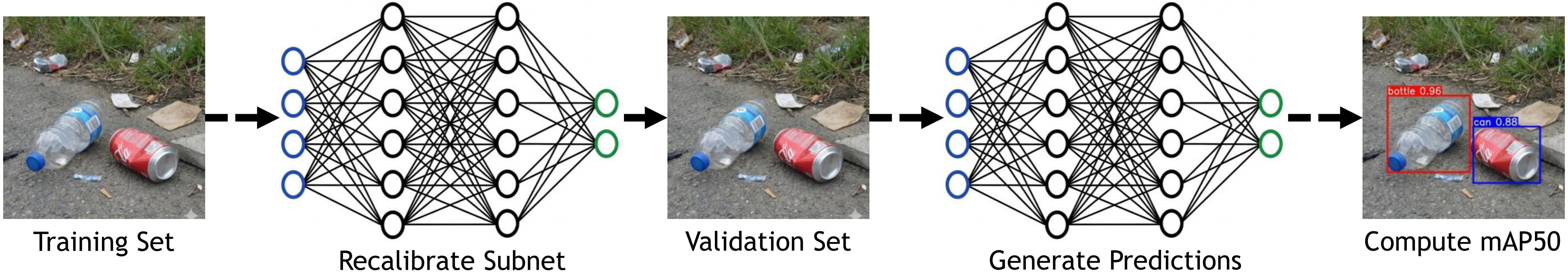
We adopt a novel *Population Passthrough* mechanism to preserve high-performing candidates across search cycles, leading to faster and more stable convergence



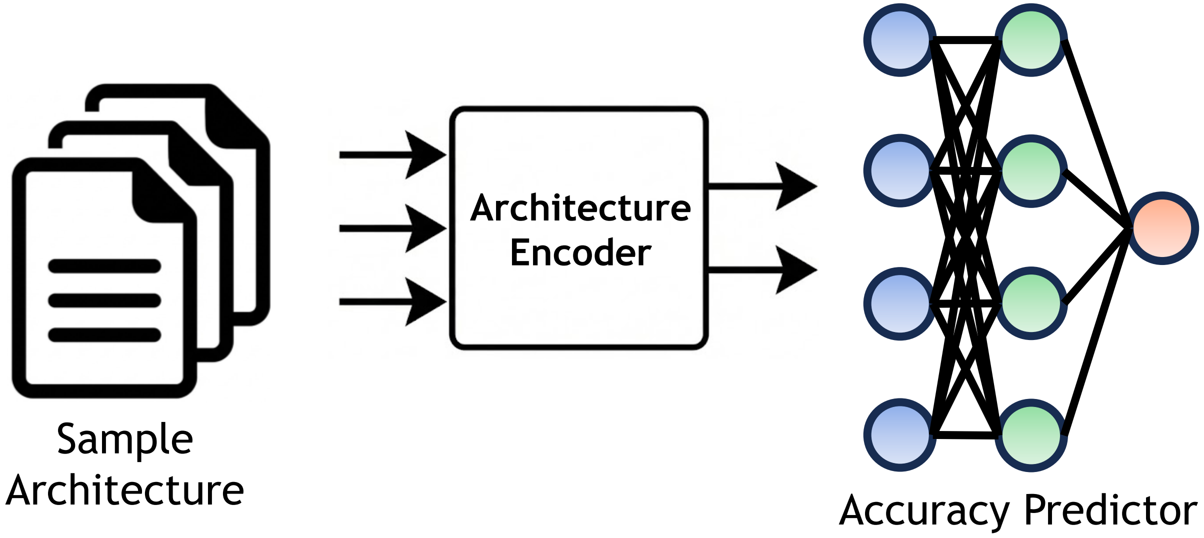
Evaluation Proxy: Accuracy Predictor

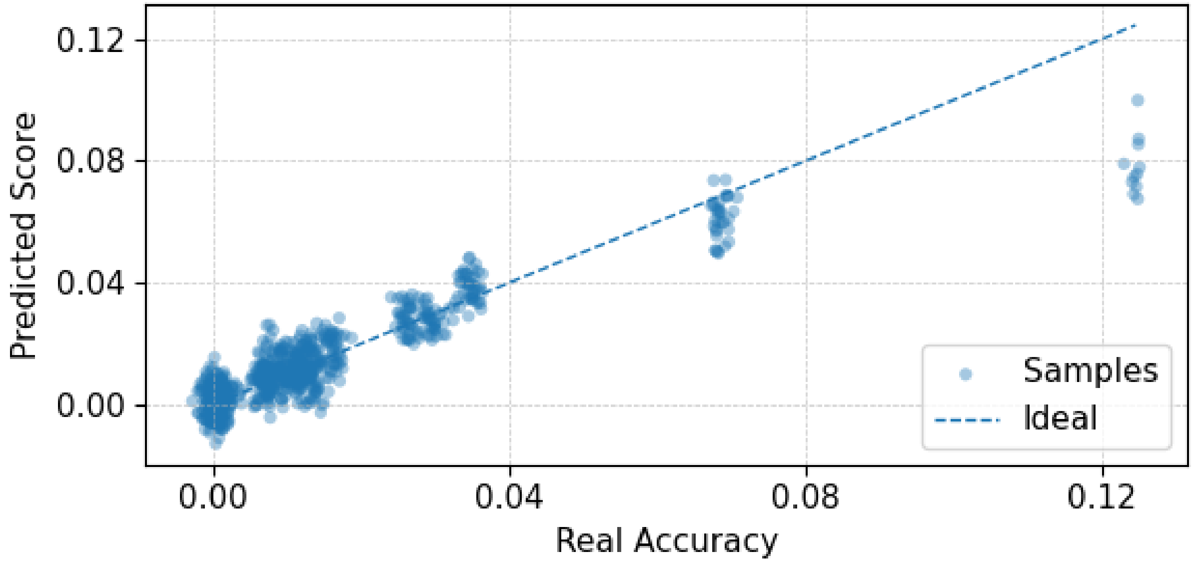
UNIVERSITY OF HOUSTON

Standard Evaluation Metric: mAP50



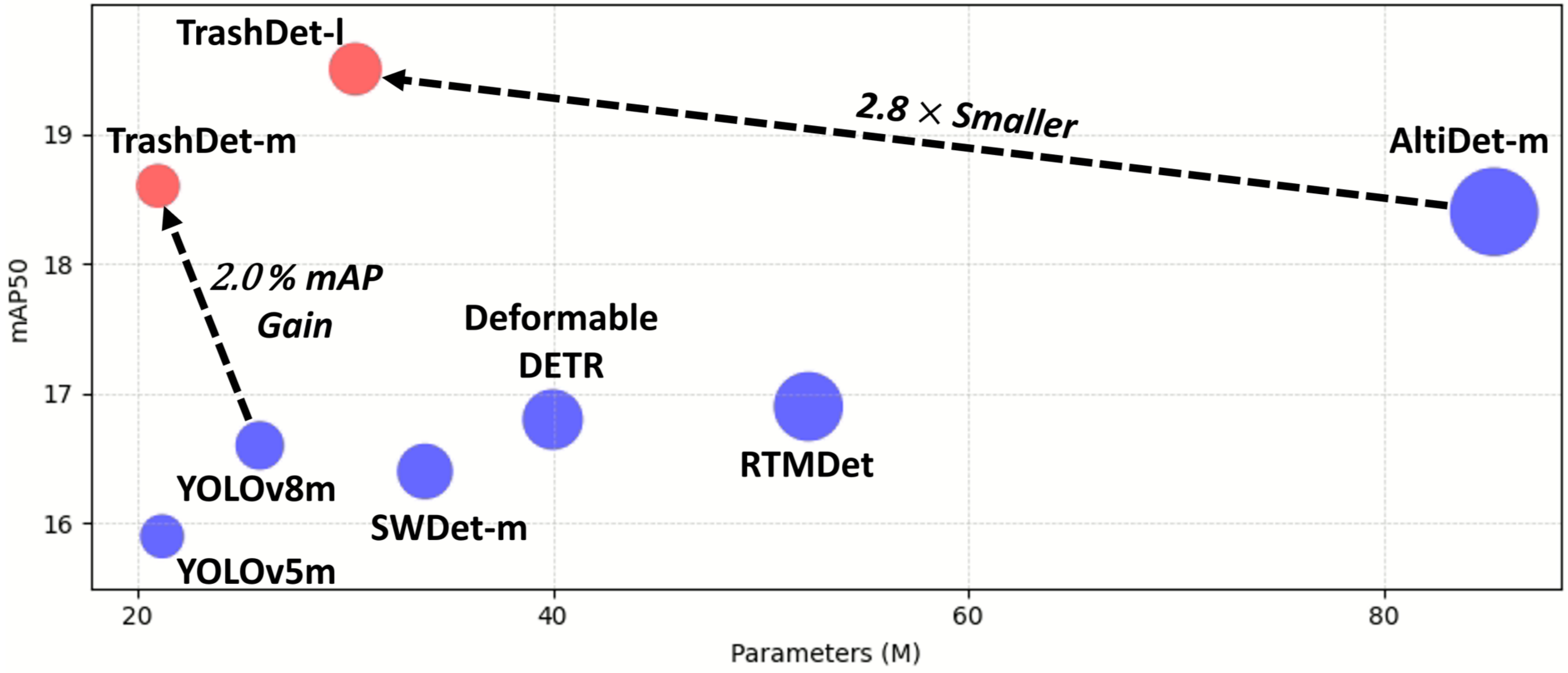
Score Proxy: Accuracy Predictor





TrashDet for Efficient Waste Detection

UNIVERSITY OF HOUSTON



Our framework discovers models that are significantly more compact and accurate than the existing detectors on the TACO dataset

TrashDet Quantitative Comparison

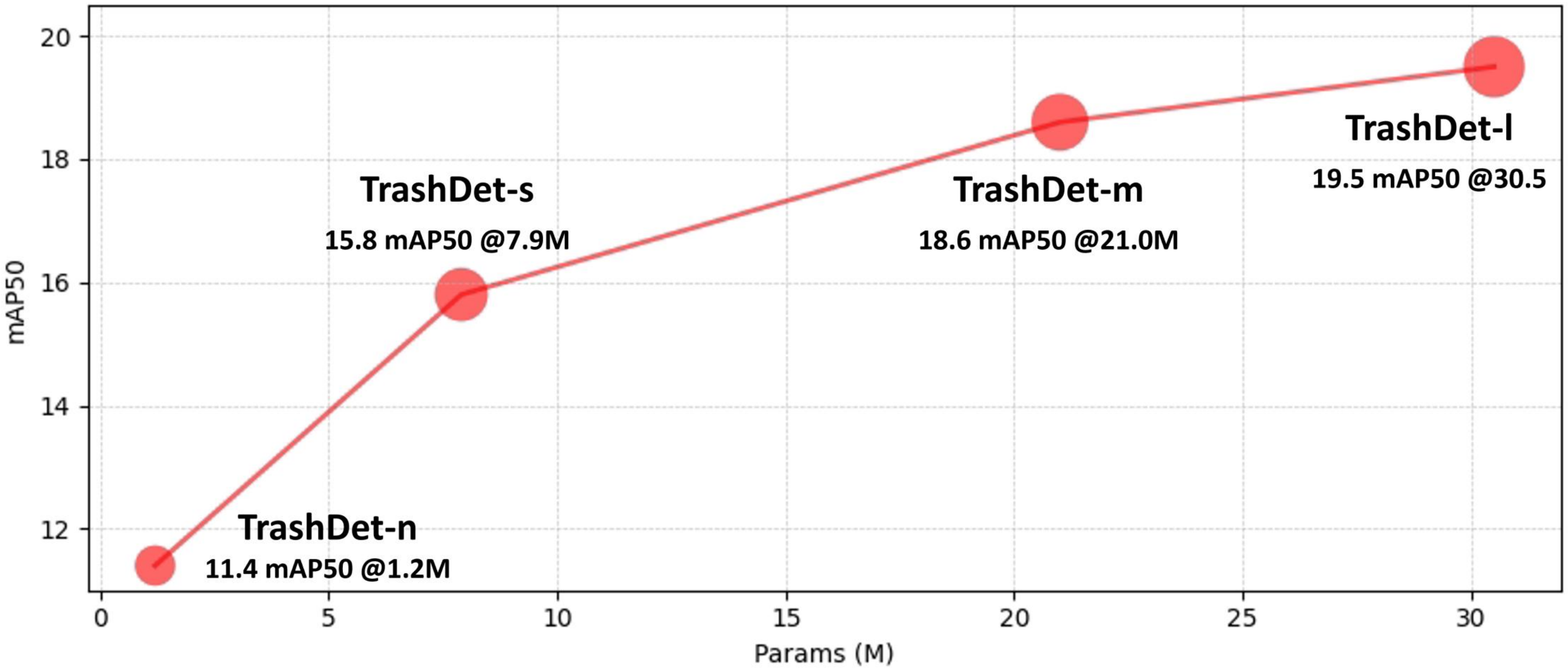
UNIVERSITY OF HOUSTON

Method	Backbone	Neck	Head	Params	AR	mAP50
YOLOv5m [9]	CSPDarknet [25]	SPPF [6] + PANet [12]	Yolov3 [20]	21.2M	**22.3**	15.9
YOLOv8m [8]	CSPDarknet [25]	SPPF [6] + PANet [12]	Yolov8 [8]	25.9M	16.6	16.6
ELASTIC-m (Ours)	**OFA ResNet**	**OFA PANet**	**OFA Yolov3**	**21.0M**	19.1	**18.6**
SWDet-m [28]	ADA [27]	EAFPN	Yolov3 [20]	33.85M	21.0	16.4
Deformable DETR [29]	ResNet-101 [7]	DETR Encoder	DETR Decoder [2]	40M	**30.3**	16.8
RTMDet [15]	RTMDet-l	PANet [12]	RTMDet	52.3M	19.4	16.9
AltiDet-m [10]	ADA + HRFE [26]	A-IFPN	Yolov3 [20]	85.3M	22.4	18.4
ELASTIC-l (Ours)	**OFA ResNet**	**OFA PANet**	**OFA Yolov3**	**30.5M**	18.6	**19.5**

TrashDet-l achievies the highest mAP50 of 19.5, outperforming the strongest baseline (AltiDet-m) while using nearly one-third of the parameters

TrashDets Family: Scalable Models

Our NAS framework can be targeted for different parameter budgets, creating a range of models

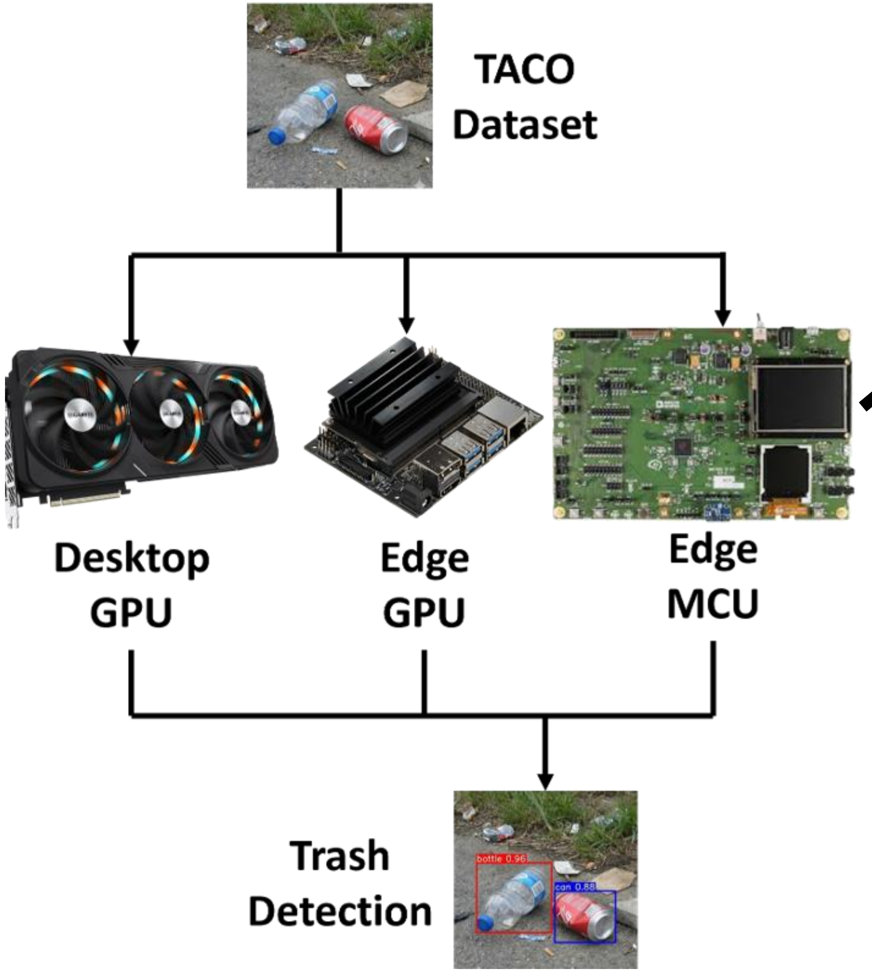


This provides practitioners with scalable options for diverse deployment targets

Deploying on Resource Constrained MCU

UNIVERSITY OF HOUSTON

Step 4: Deploy Tailored Models



Target Device: MAX78002

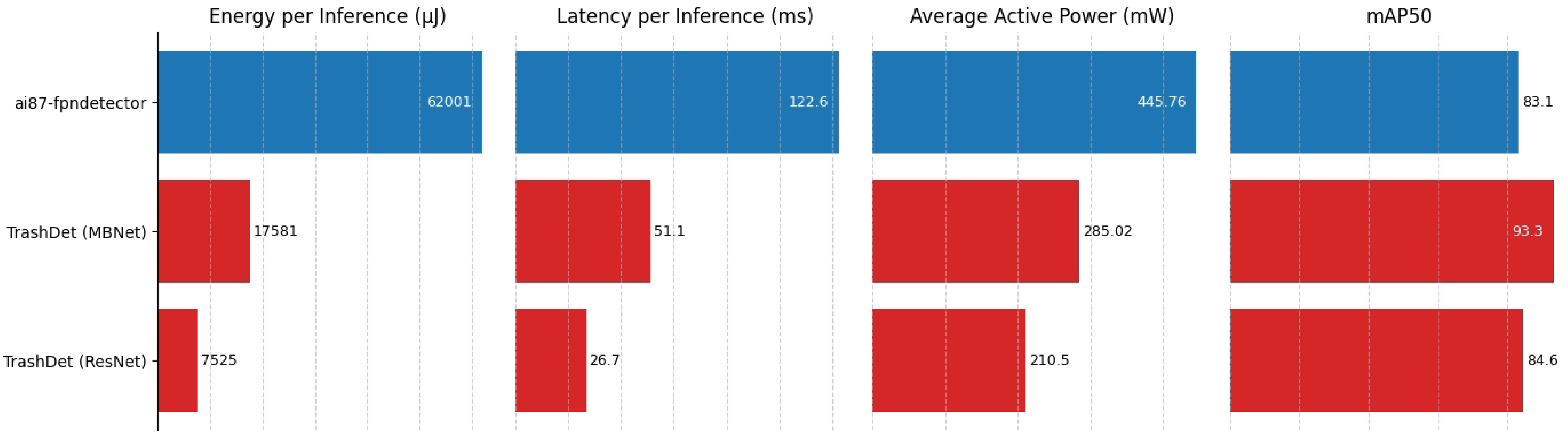
A low-power microcontroller with specialized CNN accelerator

Hardware Constraints:

- ✓ Operators: Supported kernels sizes, padding, and stride. Only specific pooling and activation functions allowed
- ✓ Memory: Kernel memory restricted to 2,340 KiB and data memory restricted to 80 KiB when not using streaming mode
- ✓ Network: Max channel dimensions (<2048) and total layers capped at 128

Quantitative Efficiency Comparison

UNIVERSITY OF HOUSTON



87.9%

Energy Reduction Compared Against TrashDet (ResNet)

78.2%

Latency Reduction Compared Against TrashDet (ResNet)

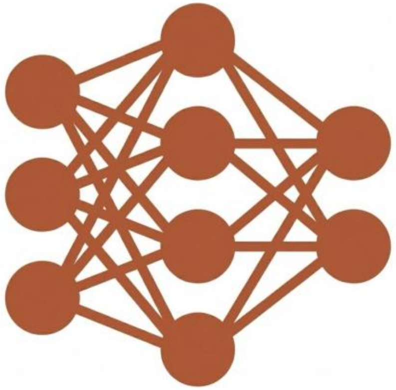
52.8%

Power Reduction Compared Against TrashDet (ResNet)

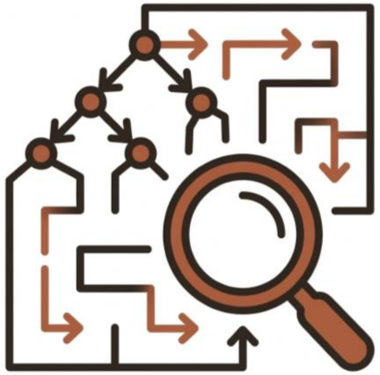
10.2%

mAP50 Increase Compared Against TrashDet (MBNet)

Conclusion



Contribution 1: A unified OFA ResDets supernet for object detection, covering all components of the network (i.e. backbone, neck, head), filtered by hardware constraints



Contribution 2: An iterative evolutionary search strategy stabilized by a population passthrough mechanism, enabling efficient and stable modular optimization

Contribution 3: The discovery of TrashDets family of models, which outperforms SOTA baselines on the TACO benchmark and demonstrates significant on-device improvements when deployed on a microcontroller