

YOLO Nano: **A Highly Compact You Only Look Once Convolutional** **Neural Network for Object Detection**

Alexander Wong, Mahmoud Famuori, Javad Shafiee,
Francis Li, Brendan Chwyl, Jonathan Chung

DarwinAI
University of Waterloo

Object Detection

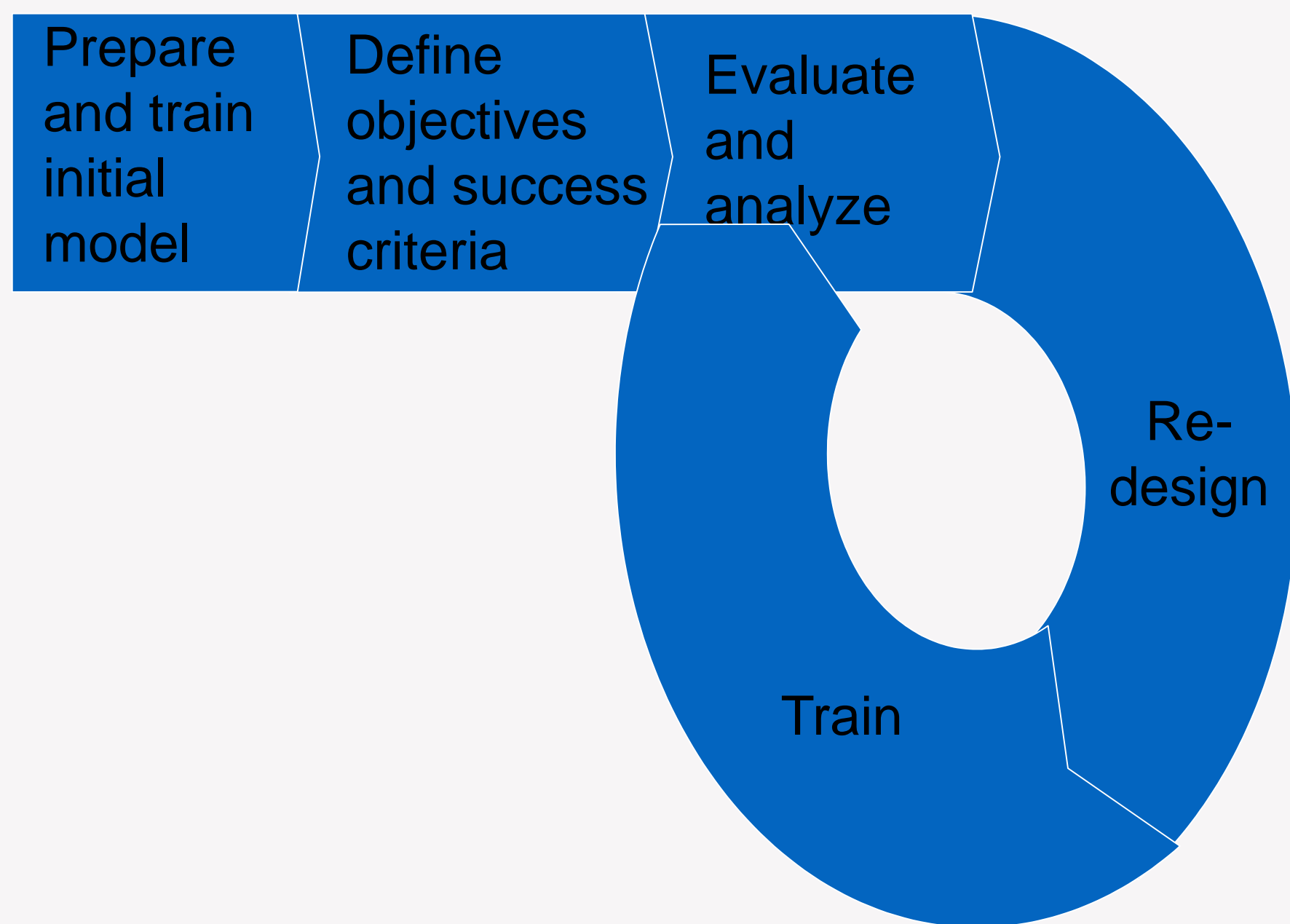
- Localize and classify objects of interest within a scene.
 - R-CNN, Mask R-CNN
 - Highly computationally complex and memory intense (Two-stage Networks).
- You Only Look Once (YOLO):
 - Single-shot architectures
 - Object detection as a regression problem
- Tiny YOLO
 - Model size Vs Performance

Human-Machine Collaborative Design

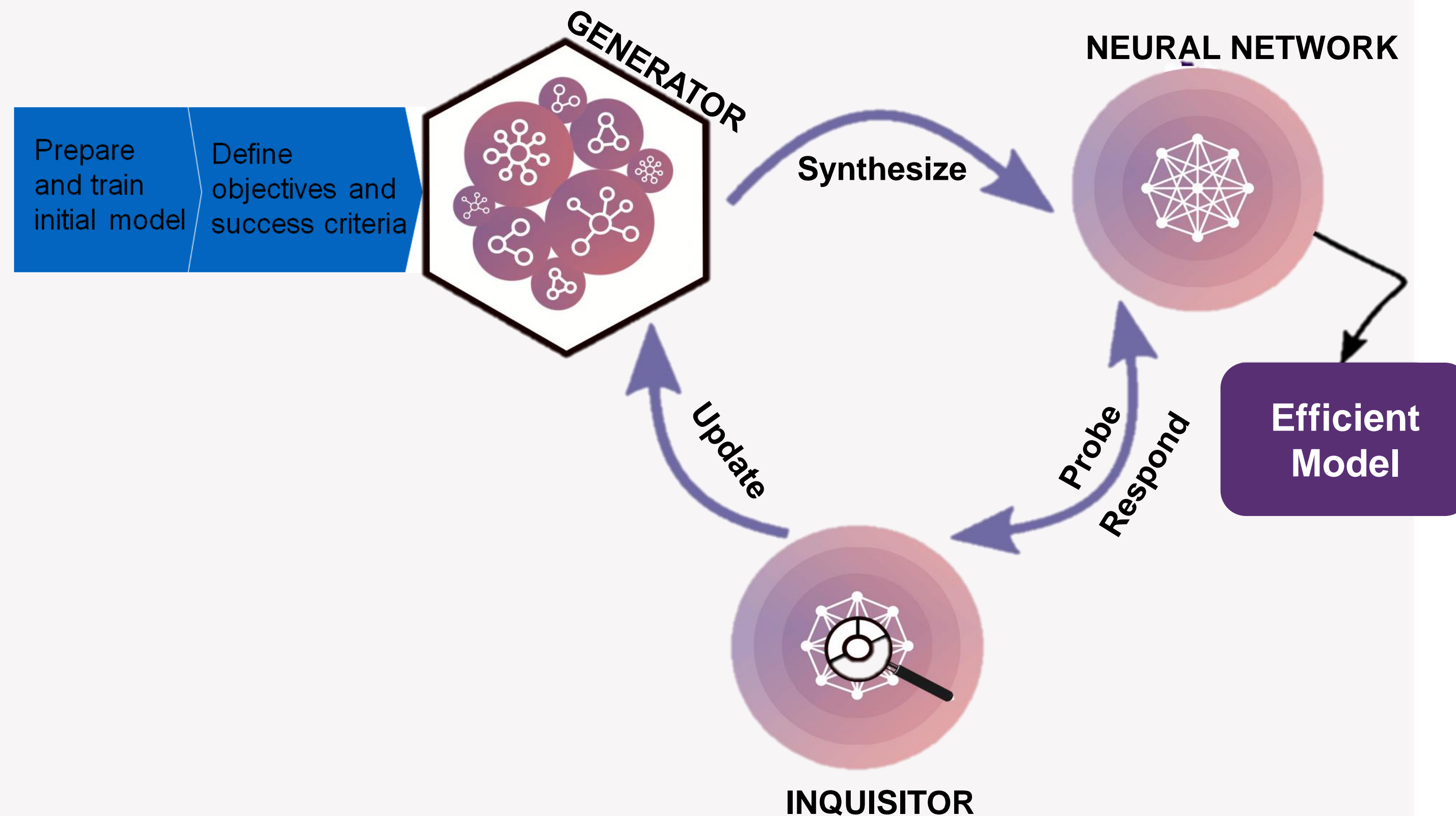
- A human-specified initial design prototype.
 - **High-level** network infrastructure for achieving high modeling accuracy.
- A set of human-specified design requirements.
 - Machines are considerably more capable at **low-level** design exploration.

Generative Synthesis

MANUAL



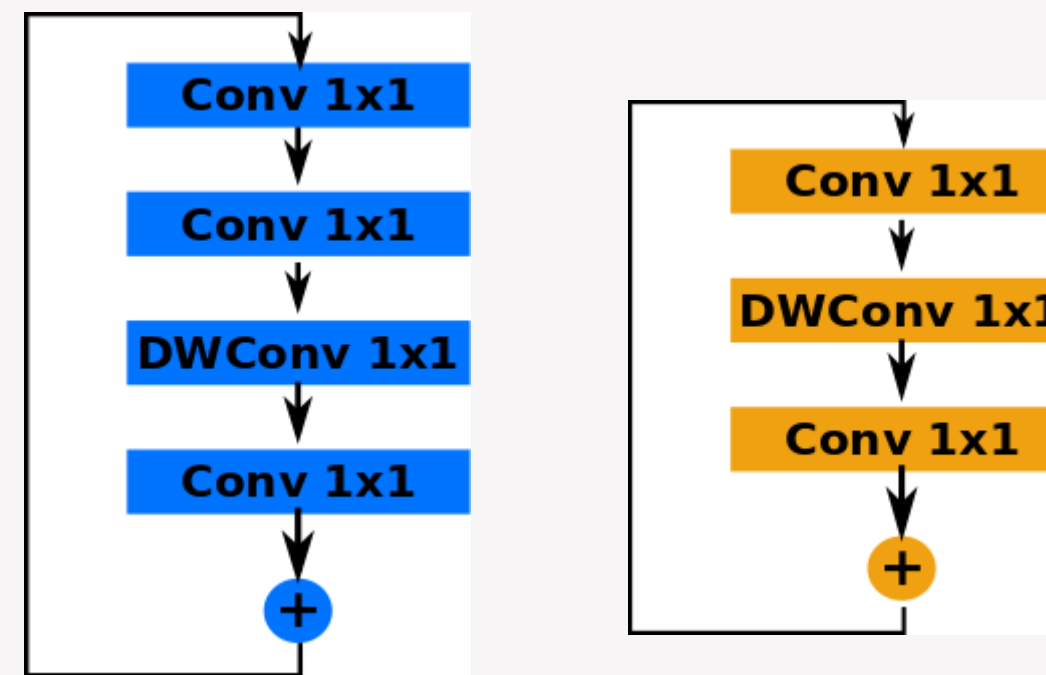
Human-Machine Collaborative Design



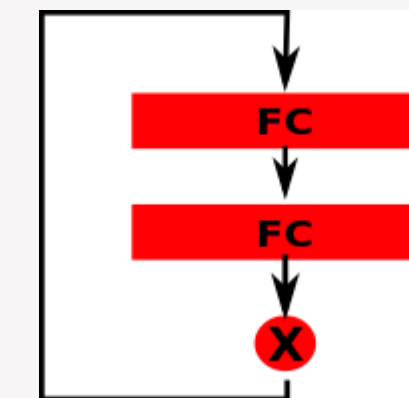
$$\mathcal{G} = \max_{\mathcal{G}} \mathcal{U}(\mathcal{G}(s)) \text{ subject to } 1_r(\mathcal{G}(s)) = 1, \forall s \in S.$$

YOLO Nano

- Residual Projection-Expansion-Projection Macroarchitecture (PEP)

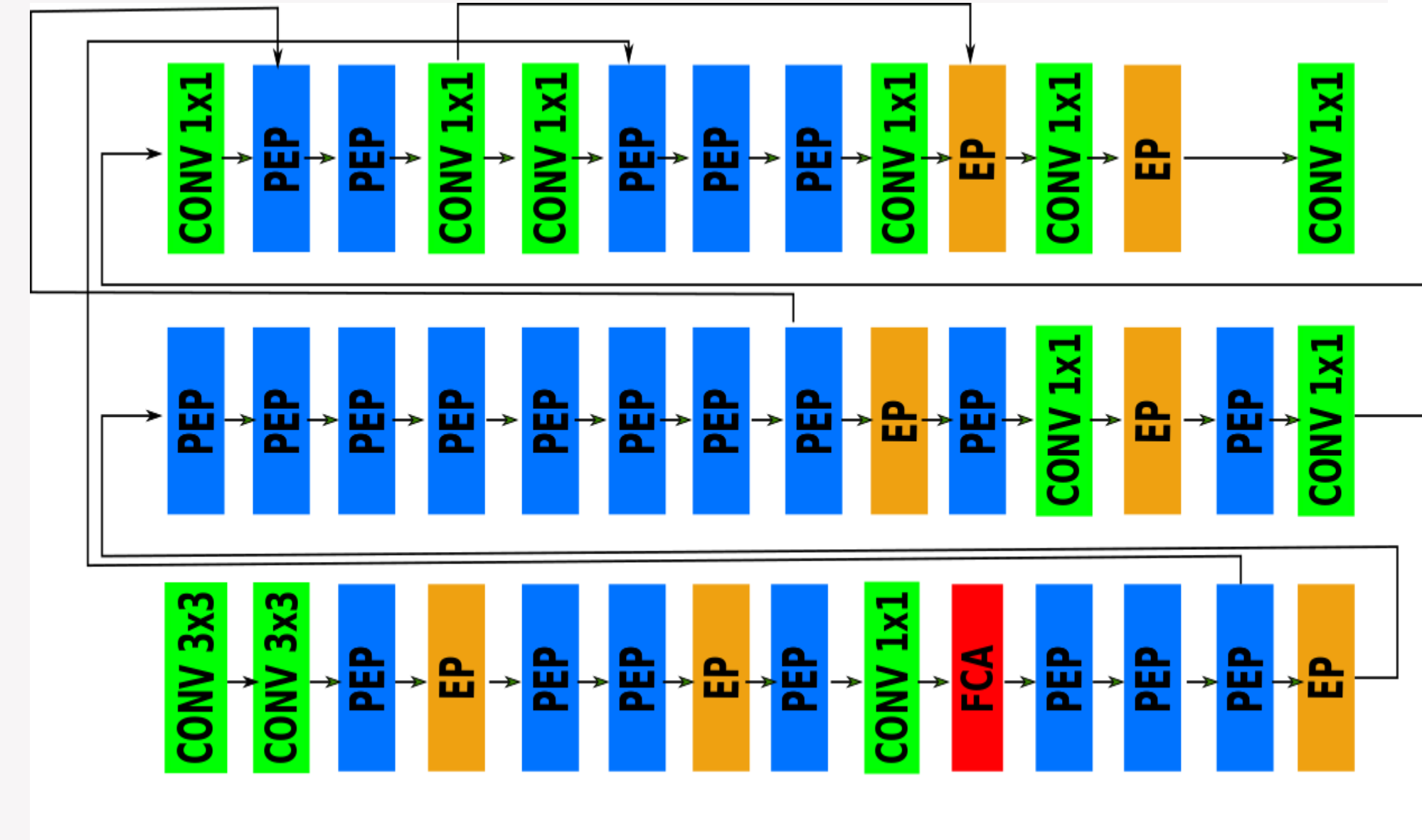


- Fully-connected Attention Macroarchitecture



- Macroarchitecture and Microarchitecture Heterogeneity

- Each component is uniquely tailored
- A very strong balance between computational complexity and model expressiveness.



Experimental Results

- Training on VOC2007/2012 dataset: 20 different objects
- Test on VOC2007
- Inference speeds of ~ 26.9 FPS, at 15W power budgets.
- Inference speeds of ~ 48.2 FPS, at 30W power budgets.

Model Name	Model size	mAP (VOC 2007)	computational cost (ops)
Tiny YOLOv2 [13]	60.5MB	57.1%	6.97B
Tiny YOLOv3 [14]	33.4MB	58.4%	5.52B
YOLO Nano	4.0MB	69.1%	4.57B

THANK YOU!

Javad Shafiee

javad@darwinai.ca

mjshafiee@uwaterloo.ca

www.DarwinAI.ca
