Evaluating the Energy Efficiency of Few-Shot Learning for Object Detection in Industrial Settings

Presenter: Prof. Panagiotis Sarigiannidis

Minds

May 13th, 2024

Funded by
the European Union

Authors

Georgios Tsoumplekas, Ilias Siniosoglou, Panagiotis Sarigiannidis
R&D Department, Metamind Innovations P.C., Kozani, Greece
{gtsoumplekas, isiniosoglou, psarigiannidis}@metamind.gr

Vladislav Li, Vasileios Argyriou
Department of Networks and Digital Media, Kingston University, Kingston upon Thames, United Kingdom
{v.li, vasileios.argyriou}@kingston.ac.uk

Ilias Siniosoglou, Panagiotis Radoglou-Grammatikis, Panagiotis Sarigiannidis
Department of Electrical and Computer Engineering, University of Western Macedonia, Kozani, Greece
{isiniosoglou, pradoglou, psarigiannidis}@uowm.gr

Sotirios K. Goudos
Physics Department, Aristotle University of Thessaloniki, Thessaloniki, Greece
sgoudo@physics.auth.gr

Ioannis D. Moscholios
Department of Informatics and Telecommunications, University of Peloponnese, Tripoli, Greece
idm@uop.gr

Panagiotis Radoglou-Grammatikis
Department of Research and Development, K3Y Ltd., Sofia, 1000, Bulgaria
pradoglou@k3y.bg

Outline

- Motivation
- Methodology
- Experimental Results
- Conclusion

Motivation

Motivation[1/2] - Introduction

AI in Industrial Applications

- Sustainability and energy efficiency are critical requirements in industrial settings
- Data can be scarce or expensive to acquire due to privacy regulations
- Hardware and bandwidth resources are limited on edge devices

Few-shot Learning

- Learning from a limited amount of data
- Reduced resource demands, alleviation of lengthy model training
- Finetuning-based approaches have shown promising results

Current Limitations

- Quantifying and enhancing the energy efficiency of finetuning-based FSL models is an underexplored topic

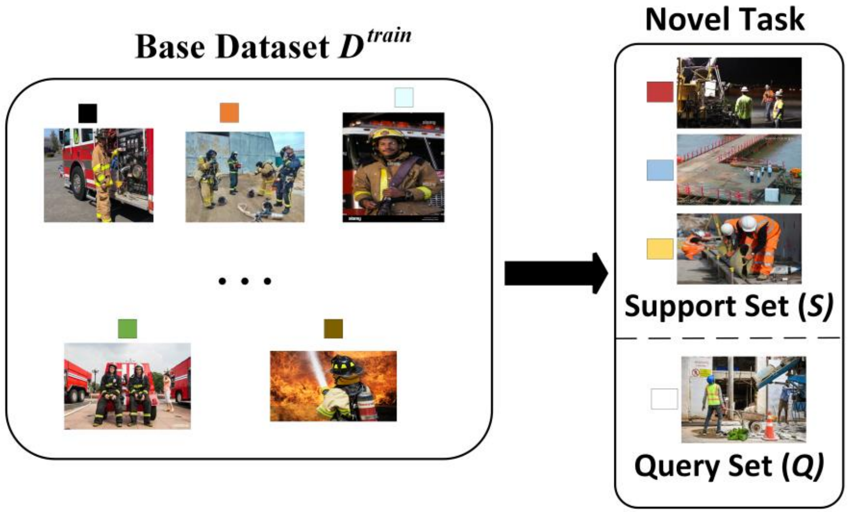
Motivation[2/2] - Contributions

- Explore the trade-off between the training performance and energy efficiency of finetuning-based methods for few-shot object detection
- Propose Efficiency Factor as a novel metric to quantify the performance vs energy consumption trade-off of FSL models
- Perform a thorough performance and energy consumption evaluation of finetuning-based object detectors on three benchmarks of volatile industrial data

Methodology

Methodology[1/3] - Problem Formulation

- **Base dataset** D^{train}
    - Contemporary object detection modules are typically pretrained in a large dataset containing abundant annotated images (e.g., MS-COCO)
    - This dataset contains images of objects belonging to N_B classes
    - Under this problem formulation this dataset constitutes the base dataset D^{train} that the Yolo model is initially trained on.



- **Novel few-shot task**
    - Most object detection datasets with industrial data are typically small
    - Under this problem formulation, these datasets constitute our *novel few-shot tasks*
    - Each *novel few-shot task (dataset with new classes)* contains N_N classes that the model has not encountered before
    - Each novel few-shot task can be broken down as follows:
    - Support set S: Used for training, contains 1-30 samples / class (also called *samples*)
    - Query set Q: Used for evaluation

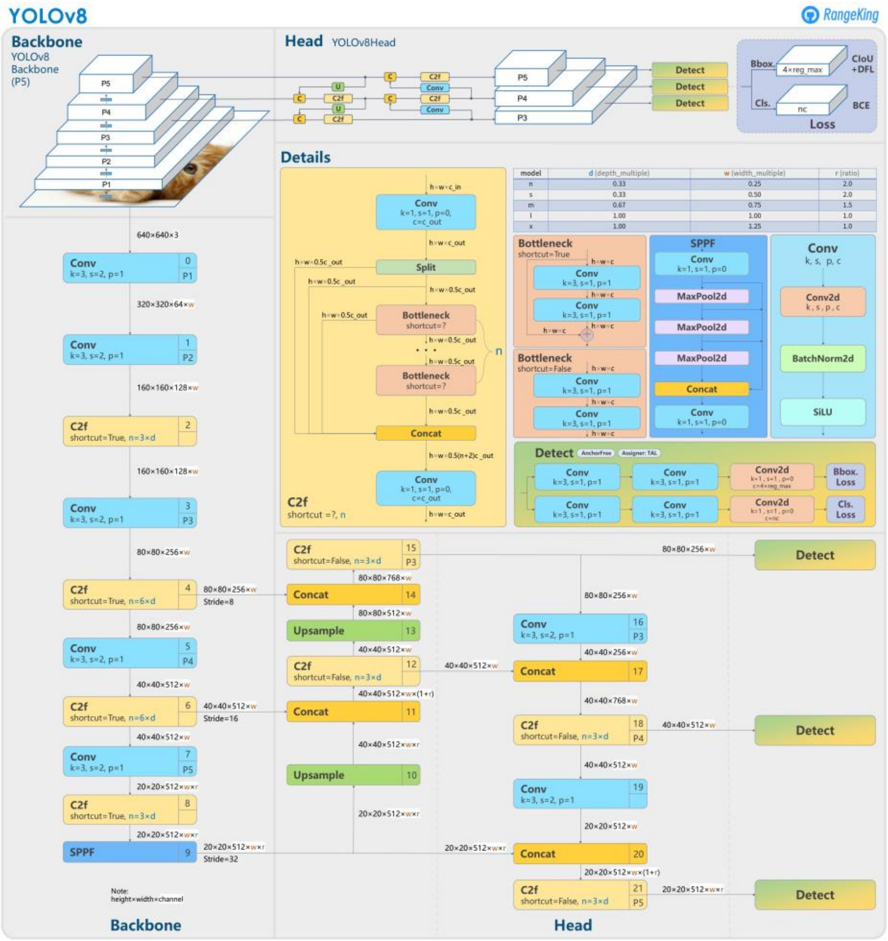
- **Objective:**
    - Develop a model that can be adapted to each novel few-shot task (e.g., learn to identify new classes) and achieve good performance in it.

Methodology[2/3] - Model Architecture

Object Detector Model: YOLOv8n (3.2M parameters)

Main Components:

- **Backbone feature extractor**
  - CSPDarknet53 based on the Feature Pyramid Network (FPN) architecture
- **Detection Head**
  - Convolutional layers
  - Three detection modules for multi-scaled object detection

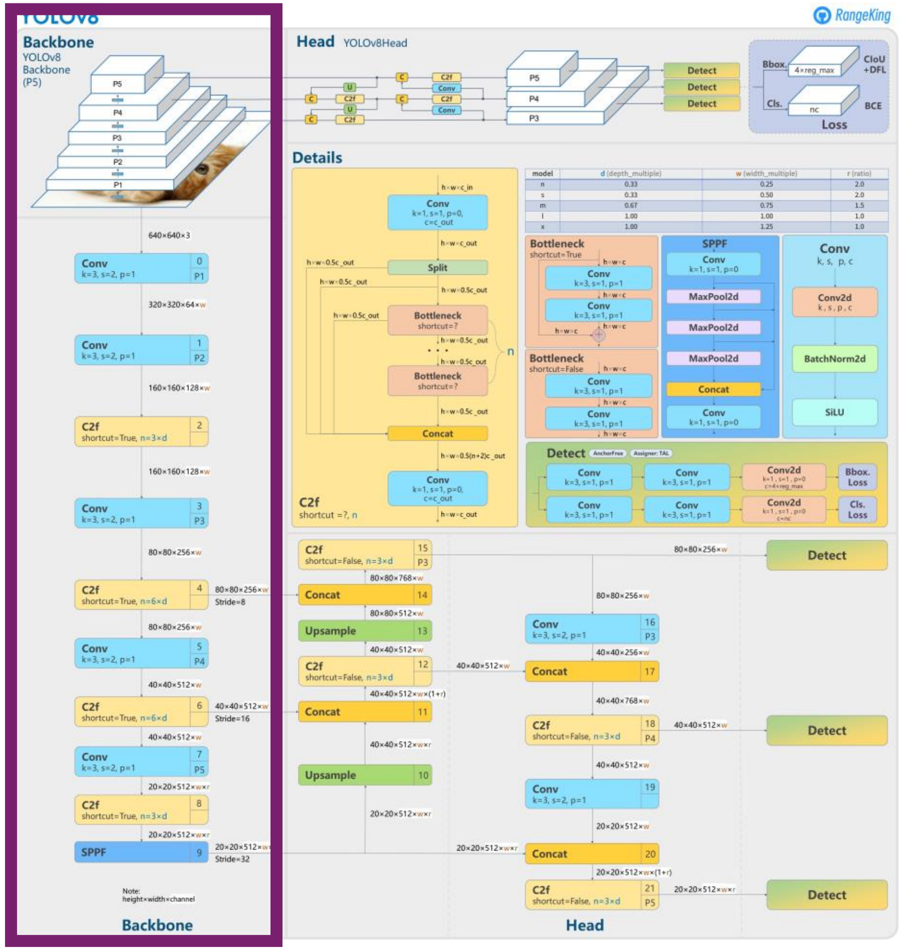


Source: https://github.com/ultralytics/ultralytics/issues/189

Methodology[2/3] - Model Architecture

<u>Backbone</u>

- **Primary Function:**
  - Acts as the feature extractor in the neural network.
- **Process Details:**
  - Processes input images through convolutional layers.
  - Generates feature maps capturing essential visual details.
- **Impact on Performance:**
  - Influences the speed of object detection.

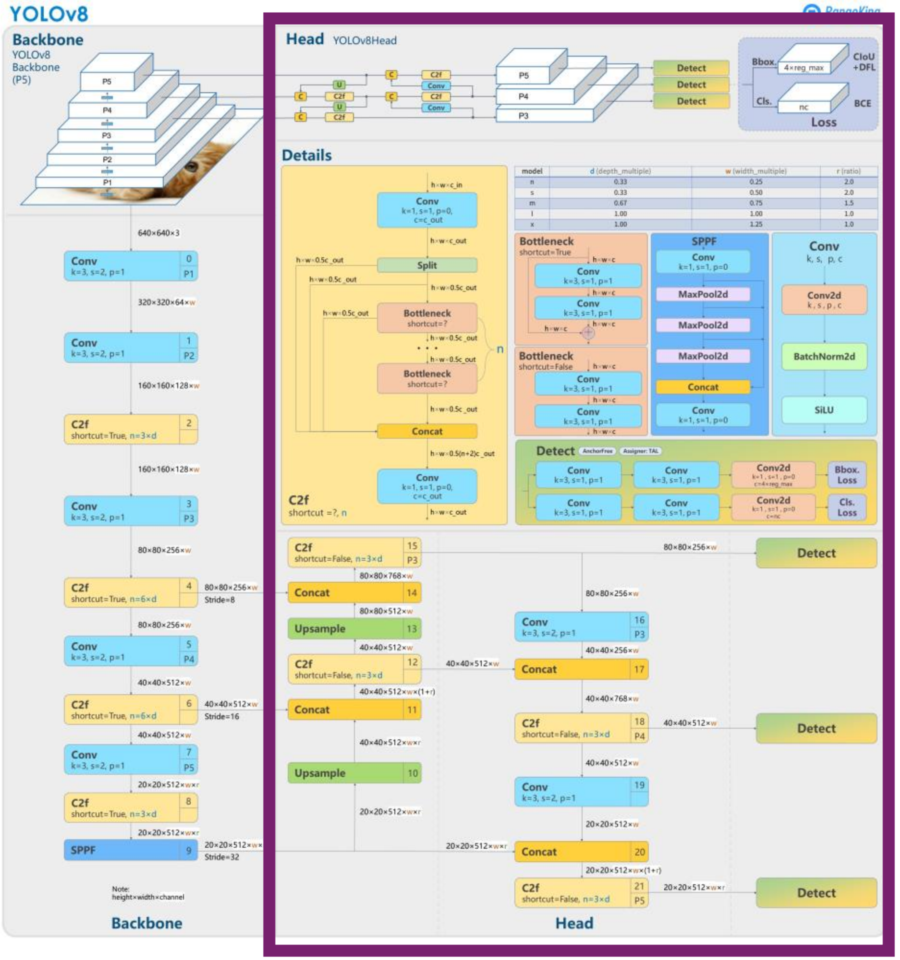


Source: https://github.com/ultralytics/ultralytics/issues/189

Methodology[2/3] - Model Architecture

Detection Head

- **Primary Function:**
  - Responsible for making final predictions in the network.
- **Process Details:**
  - Receives processed feature maps from the backbone.
  - Applies convolutional layers specifically designed to predict class probabilities, object locations, and size.
- **Outputs:**
  - Generates bounding boxes around detected objects.
  - Assigns class labels and confidence scores to each detection.

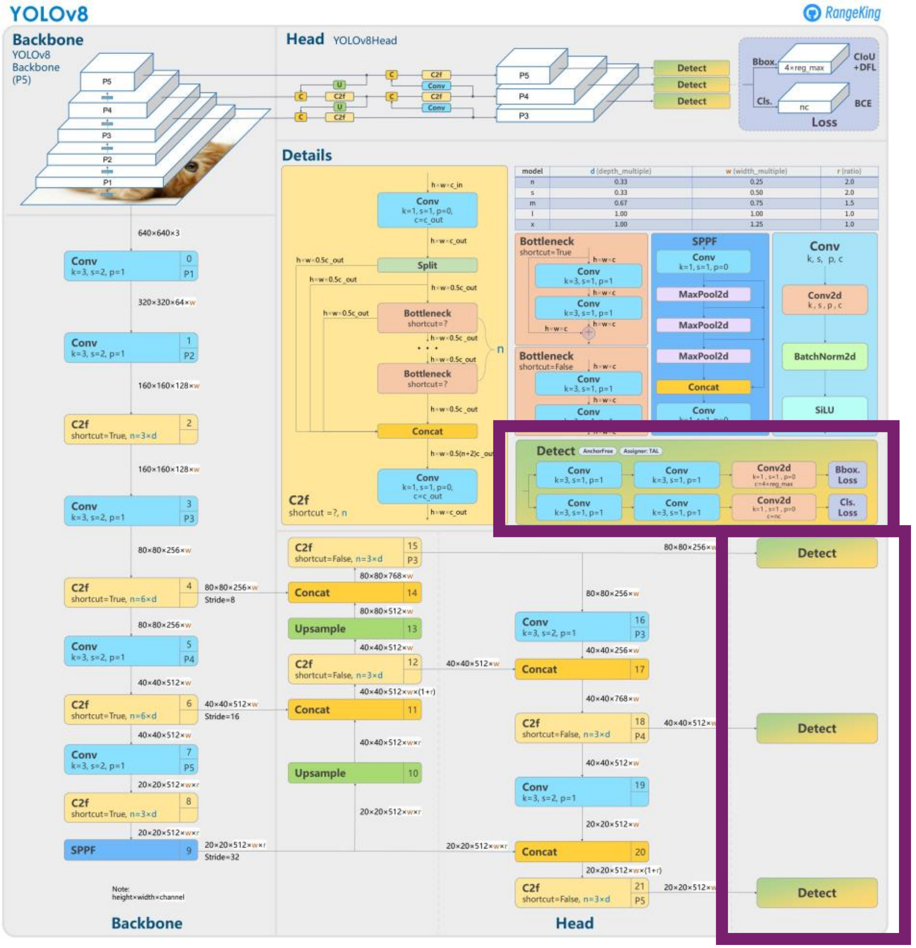


Source: https://github.com/ultralytics/ultralytics/issues/189

Methodology[2/3] - Model Architecture

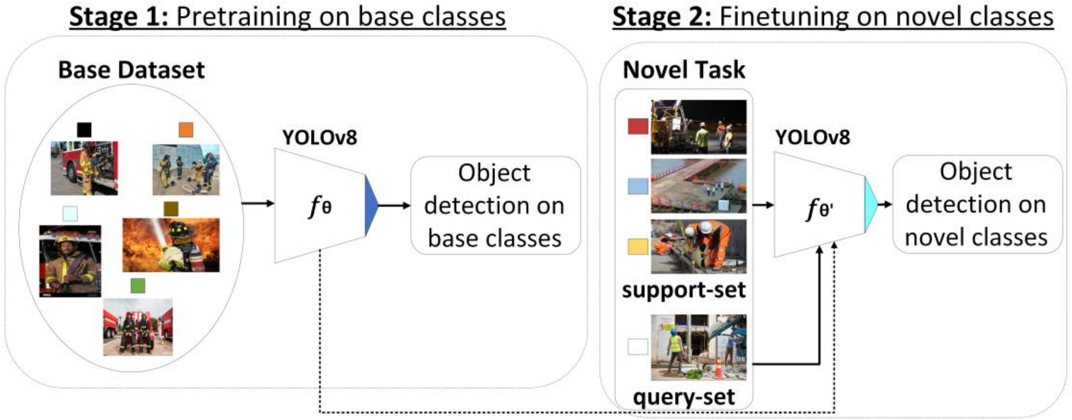
Detection Module

- **Primary Function:**
  - Localizes and classifies objects within the image.
- **Process Details:**
  - Integrates feature maps from both the backbone and neck.
  - Utilizes multiple scales for detection to enhance accuracy at various object sizes.
- **Components:**
  - Consists of multiple detection layers tailored for different scales.
  - Each layer predicts bounding boxes, class probabilities, and objectness scores.



Source: https://github.com/ultralytics/ultralytics/issues/189

Methodology[3/3] - Proposed Approach



1st stage: Pretrain model in the base dataset D^{train} (MS-COCO as the base dataset)

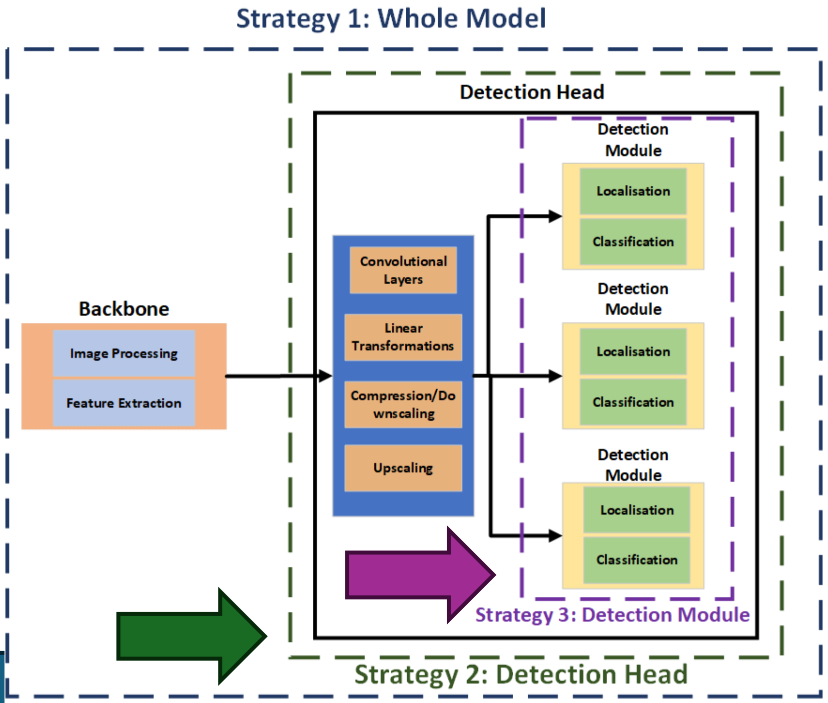
2nd stage: Finetune different modules of the model in the novel few-shot dataset (PPE, CS, Fire) ***to enable the model recognize the new classes***

Finetuning strategies:

We employ three strategies for finetuning:

- Whole model (full finetuning), finetuned parameters: 3.2M
- Detection Heads (partial finetuning), finetuned parameters: 1.7M
- Detection Modules (partial finetuning), finetuned parameters: 750K

These strategies train different parts of the Yolov8 model to investigate the best optimization technique under Few-shot learning.



Experimental Results

Results[1/5] - Datasets

Utilization of 3 open-source datasets of objects found in industrial settings

Dataset	No. Classes	Classes	Purpose
PPE Dataset	4	Helmet, Gloves, Mask, Cloth	Localize and identify PPE for first responders
Construction Safety (CS) Dataset	3	Helmets, Vests, other PPE	Identify presence/absence of PPE in industrial settings
Fire Dataset	1	Fire scenes	Locate and recognize fires

PPE Dataset

CS Dataset

Fire Dataset

Results[2/5] - Evaluation Metrics

- **Object Detection Performance:** Mean Average Precision (mAP)

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$$

- **Energy Efficiency:** Energy consumption (Wh) during finetuning stage

- **Efficiency Factor (EF):** Novel metric that takes into consideration both mAP and energy consumption (EC)

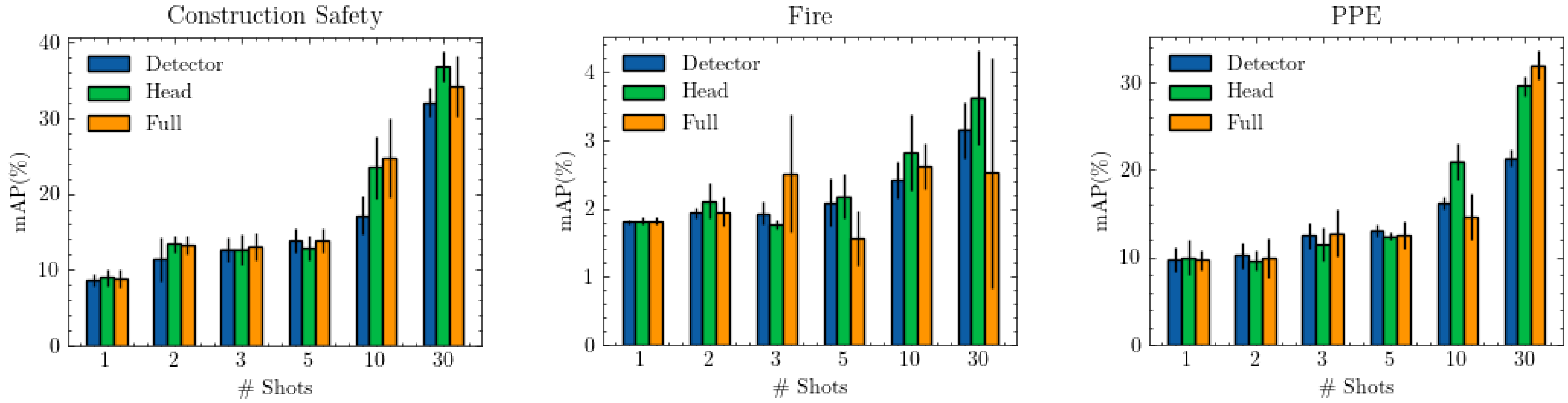
$$EF = \frac{mAP}{1 + EC}$$

- Higher EF values given to models with high mAP during evaluation and low energy consumption during finetuning

Results[3/5] - Model Performance

Main Insights:

- Increasing the number of shots *(images per class)* leads to increased model performance
- Comparable performance for all finetuning strategies in most cases for the same number of shots

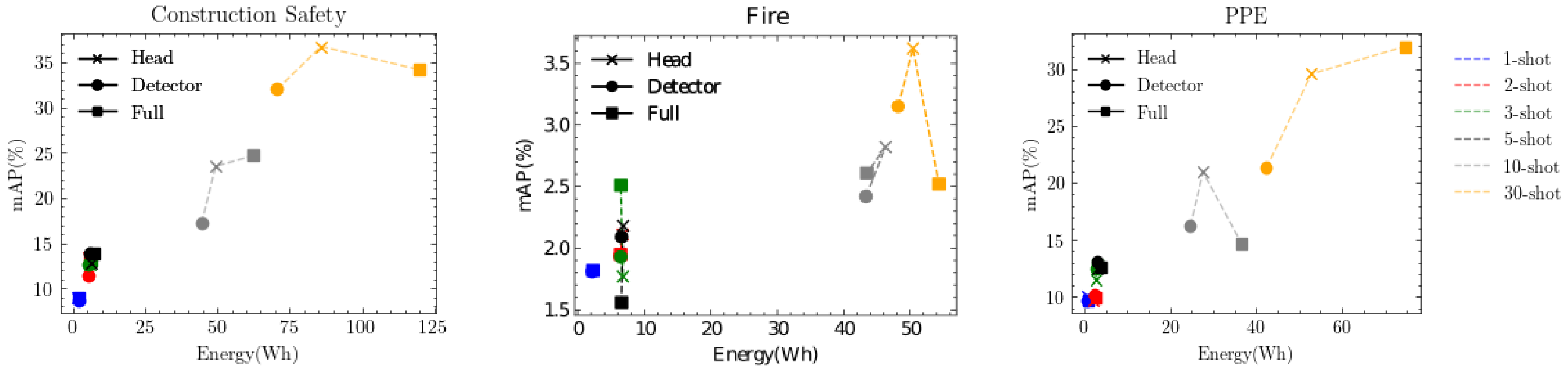


- Notably, the Full model achieves worst results than the Detector/Head because the whole model needs bigger dataset to train efficiently.

Results[4/5] - Energy Efficiency

Main Insights:

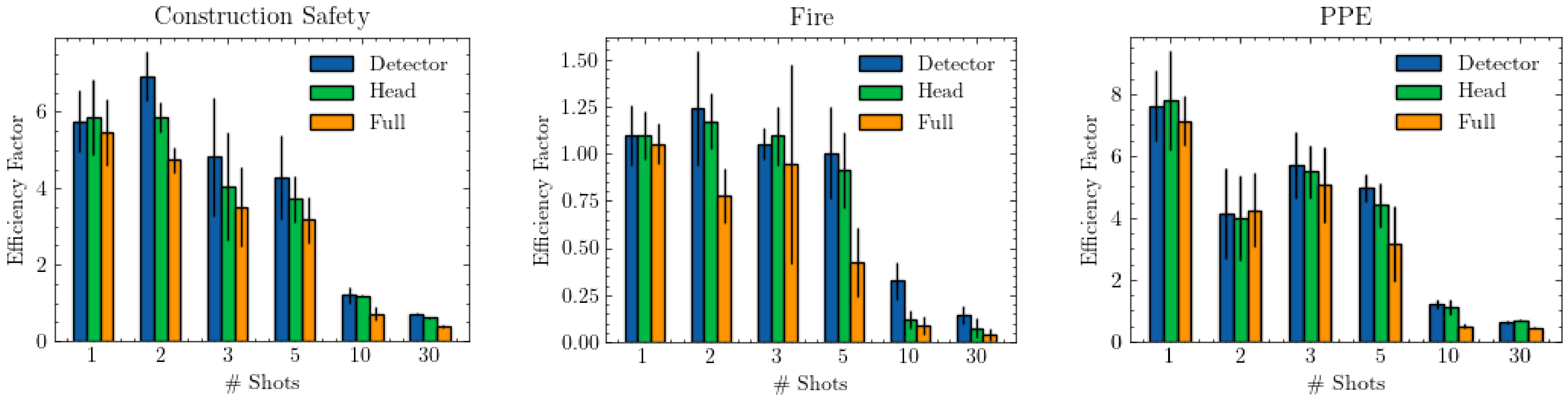
- Reducing the number of finetuned parameters leads to reduced energy consumption
- Trade-off between model performance and energy efficiency
- Finetuning only the detection heads achieves the best performance vs efficiency balance



- Finetuning the detector modules is the most energy-efficient choice, but the performance (mAP) is poor.
- Finetuning the whole model has the second-best performance, and it is too energy-consuming.
- **<u>Finetuning the detection heads leads to the best compromise between performance and energy preservation.</u>**

Results[5/5] - Performance / Efficiency Trade-off

Main Insights:
- Finetuning less parameters reduces energy consumption but does not necessarily lead to increased EF values
- Best EF values achieved when the number of shots is small



- As the number of shots *(i.e., images per class)* increases, the energy consumption needed to finetune the model outweighs any performance advantage and as a result the EF metric drops significantly.

Conclusion

Conclusion[1/1]

Takeaway points:

- Energy efficiency evaluation of FSL approaches has been an underexplored topic
- This study: focus on finetuning-based approaches for object detection in industrial datasets
- Examination of full and partial finetuning approaches on YOLOv8
- Trade-off between model performance and energy efficiency
- Introduction of Efficiency Factor (EF) as a metric that captures and quantifies this trade-off

Acknowledgement

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No. 101070181 (TALON).

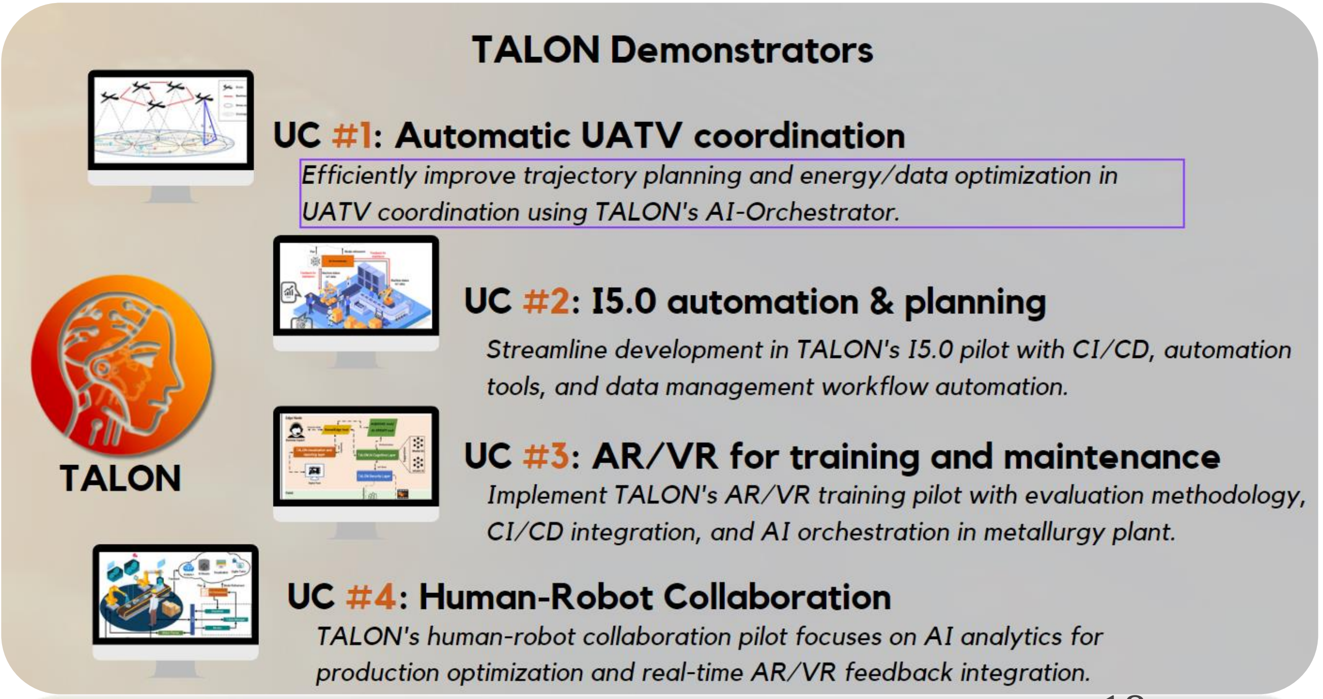
The TALON Project

Call: HORIZON-CL4-2021-HUMAN-01
Topic: HORIZON-CL4-2021-HUMAN-01-01
Type of action: RIA
Total Budget: € 3.769.382,50
Period: 36 Months

TALON

TALON Pillars

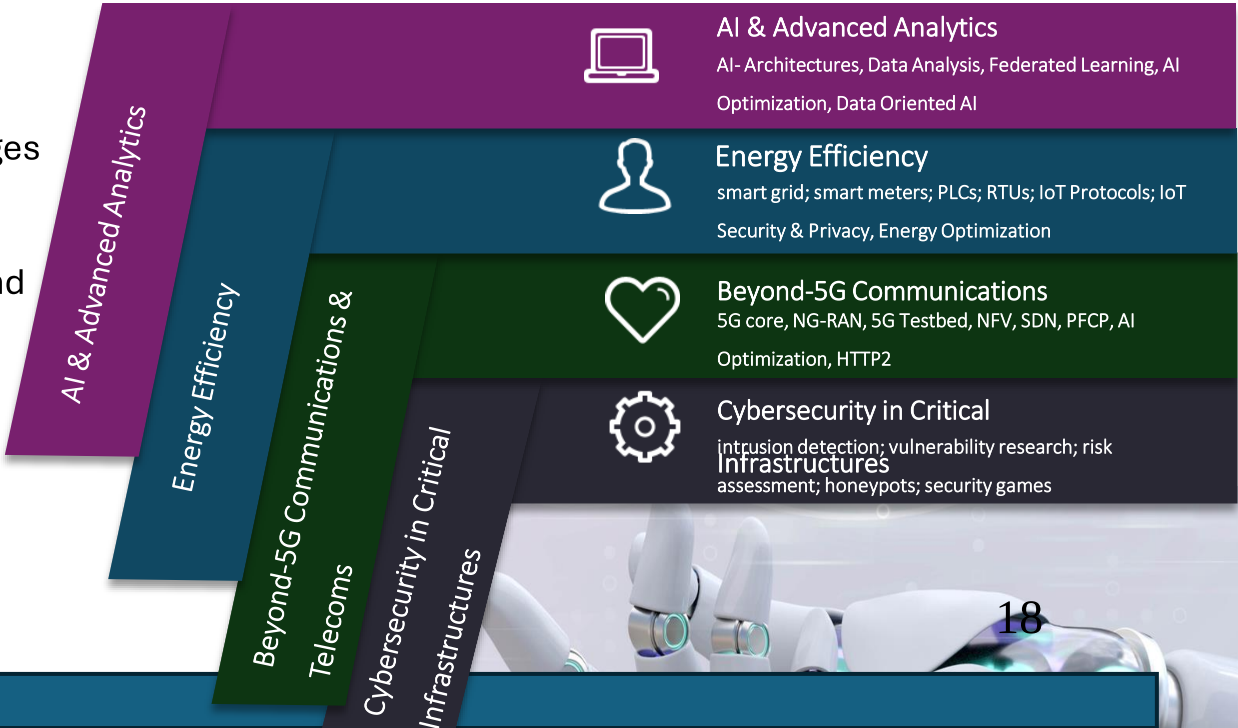
- AI orchestrator for autonomous and dynamic scalability as well as greener AI networks
- Distributed blockchain for high-security, privacy and trust in a heterogeneous application environment
- Flexible E2C deployment for “almost-zero latency” and high-computational capabilities near sensors
- DTs and HIL to boost AI explainability, trust-worthiness and transparency



MetaMind Innovations

MINDS is the first spin-off technology company of the University of Western Macedonia. The company specializes in advancing the Internet of Things (IoT) with Artificial Intelligence (AI) and Virtual Reality (VR) technologies, offering smart solutions for critical infrastructure. It was founded in 2021 by Professor Panagiotis Sarigiannidis and a team of academics and researchers, with a focus on the development of leading, reliable and customized technological solutions that enhance business efficiency, security and innovation.

MINDS addresses these challenges with leading edge solutions that harness the power of artificial intelligence, machine learning and IoT technologies.



Thank you for your attention!

psarigiannidis@metamind.gr

https://metamind.gr/

https://www.linkedin.com/company/metamind-innovations/

TALON

https://talon-project.eu/

Funded by the European Union