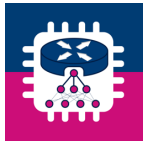
IMPERIAL

Towards Low-Energy Fine-Tuning on Constrained Devices

MSN/Coseners

Josh Millar

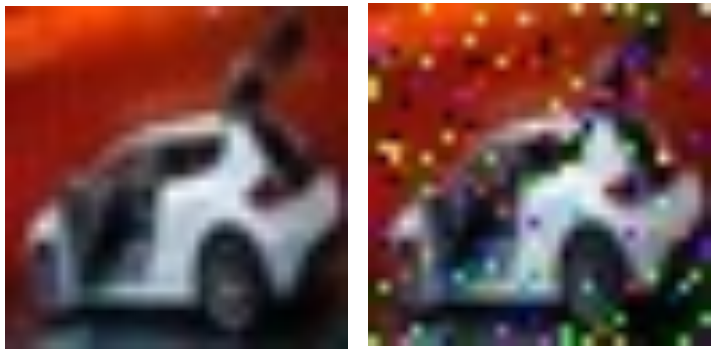
05/07/2024



Motivation

- On-device ML vulnerable to distribution shift
  - input-, feature-, and output-level
- Fine-Tuning (FT) approaches: entire model or last layer(s).

Input-level



Source Target

Feature-level

Source Target

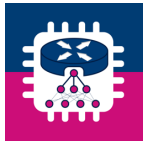
Output-level

Source **“Dog”** Target **“Yorkshire Terrier”**

Contributions

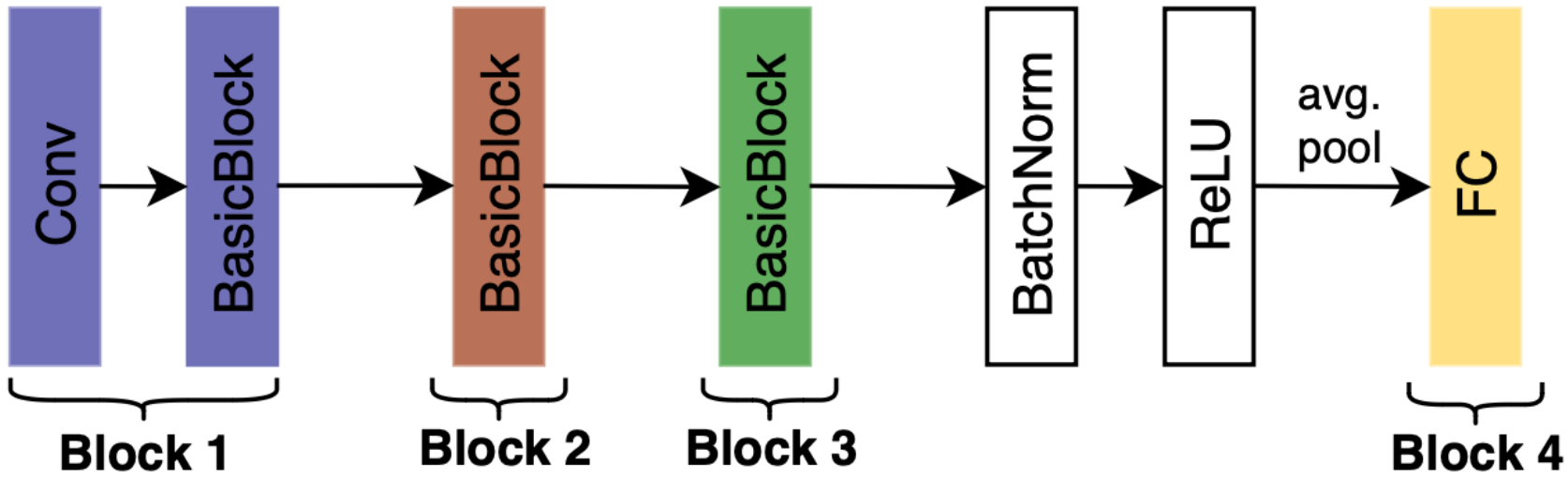
***Target Block* Fine-Tuning (TBFT)**

- FT model “blocks” corresponding to shift type present
    - i.e., output-level shift = last block(s) , input-level = first block(s), etc
- *higher accuracy* than other layer combinations
- at *reduced/equivalent energy cost*.

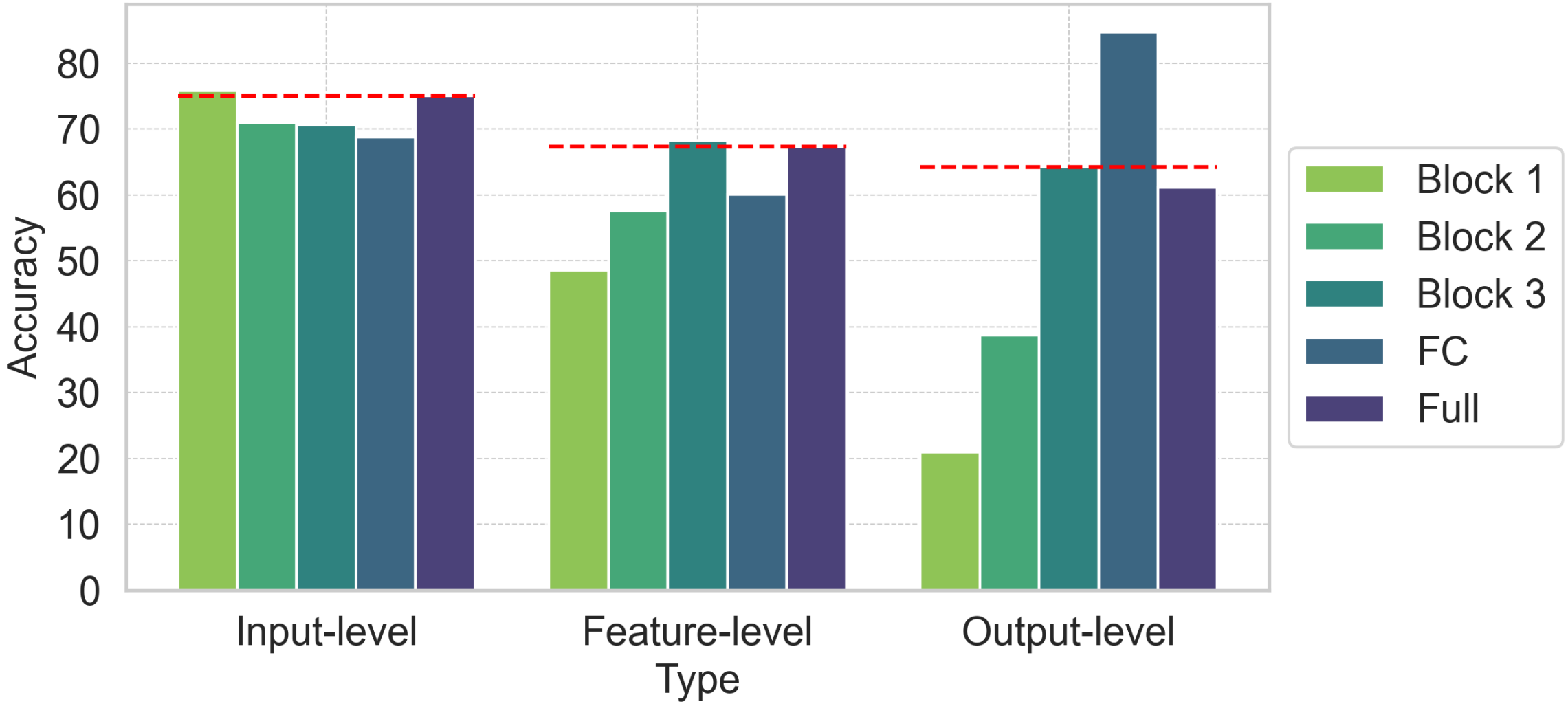


Experiments

- **ResNet-26**: Inherent “block” structure
- *Input-level*: CIFAR-C, *Feature-level*: Living17, *Output-level*: CIFAR-Flip

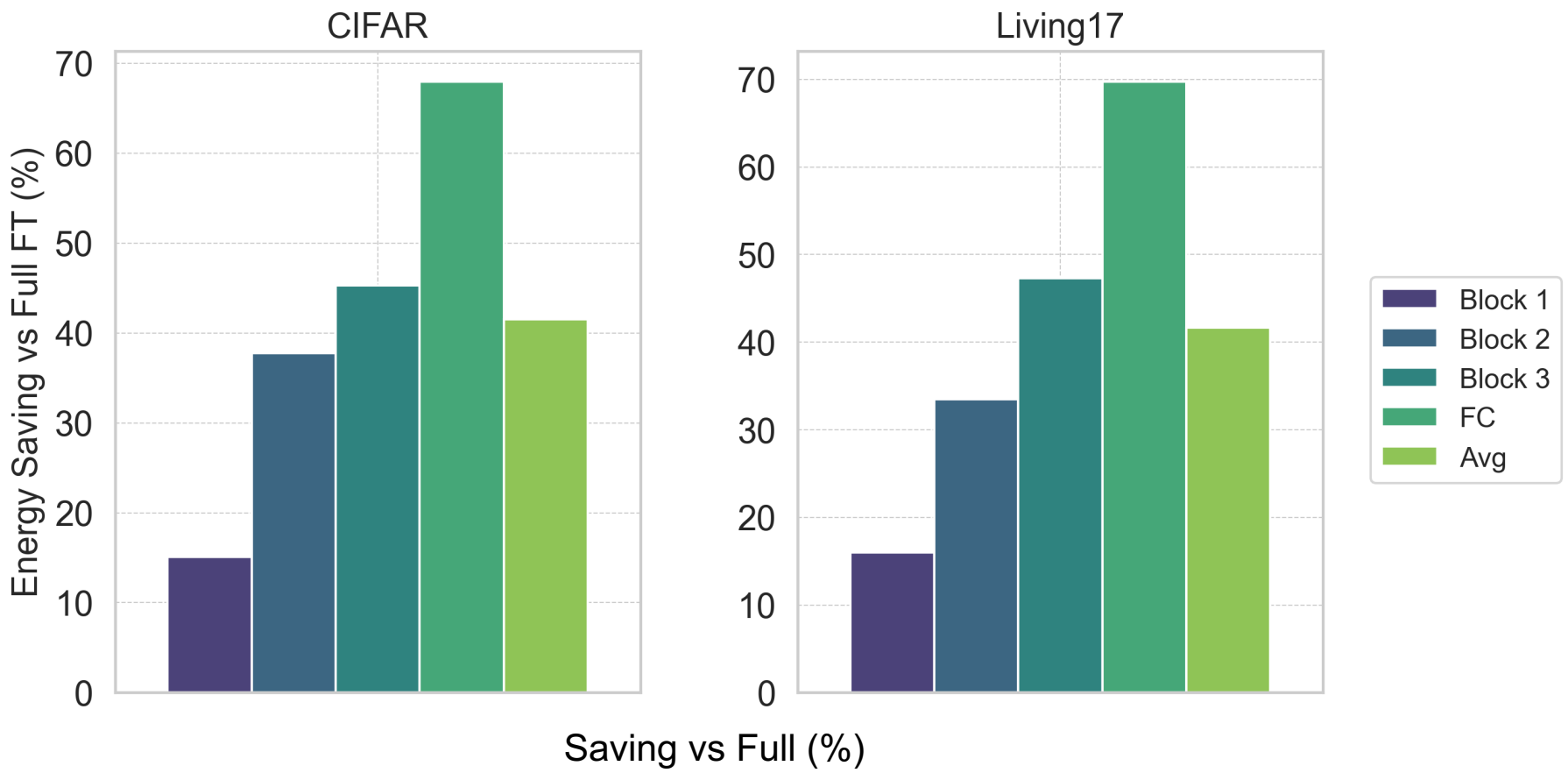


Results



- average **15.3% improvement** in **accuracy** over full FT

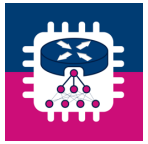
Results (cont.)



- energy saving (%) vs full FT (FFT), calculated as $ES = (E_{FFT} - E_{TBFT})/E_{FFT}$
- **average energy saving** of **41.57%** over FFT

Limitations + Future Work

- **Detecting shift type**
    - TBFT relies on knowledge of shift type …unrealistic
    - Block/parameter selection approaches are gradient-based + costly.
        - instead, select layers based on incoming data.
- **Unsupervised FT**
- **Handling complex and composite shifts**



Goal: long-term and inexpensive sensing devices

Terracorder

- Multi-sensor environment/*biodiversity* sensing device
- using off-the-shelf components
- and ML for classification + sensor duty cycling

initial results: ~70uA deep-sleep, 1.5+ years on 1 charge with 6700mAh *

* excluding comms and networking

Thank you