

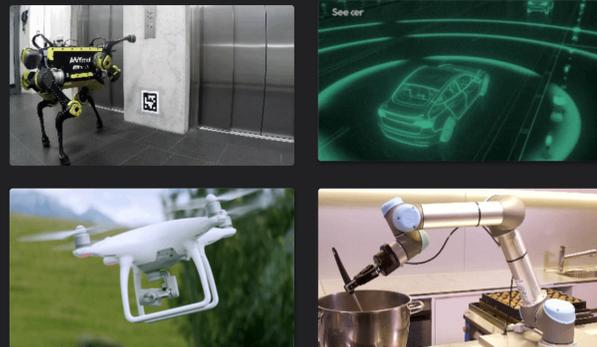
Last-Mile Delivery of Computer Vision with Test-Time Adaptation

Xin Wang

Ph.D. student at UC Berkeley

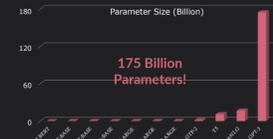


Recent Achievements in AI



AI technologies are revolutionizing our lives.

Driving Factors



Immense labeled datasets Large scale models

Key Challenges



No access to immense labeled data



Limited computation resources

Last Mile Delivery of Computer Vision

Generic models on benchmark datasets



Models suitable for any test scenario

Goal:

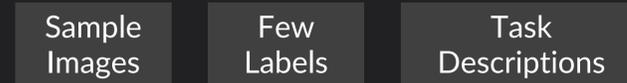
Adapt generic models to any test scenario and prediction task with **limited data, labels and computation.**

Solution: Test-Time Adaptation

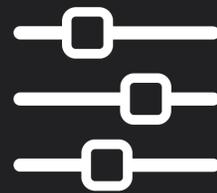


Use test-time information to adjust model **parameters** and **structures.**

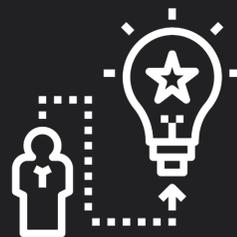
Test-time information:



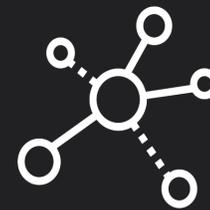
How to Use Test-Time Information?



Test-Time Fine-Tuning

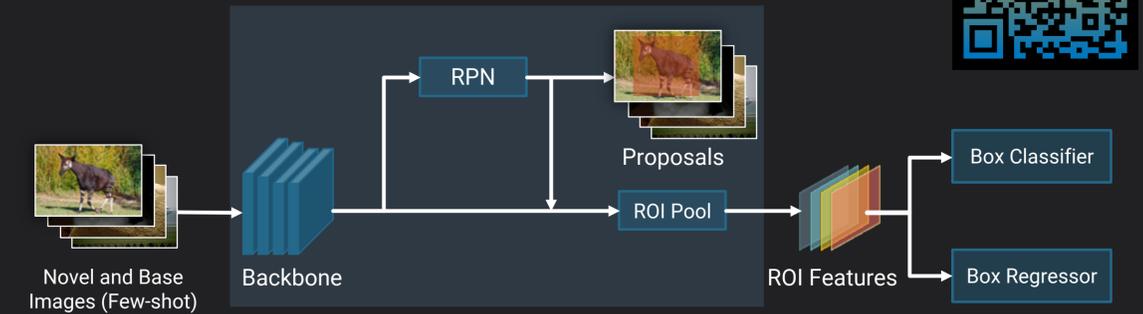


Task-Aware Weight Generation



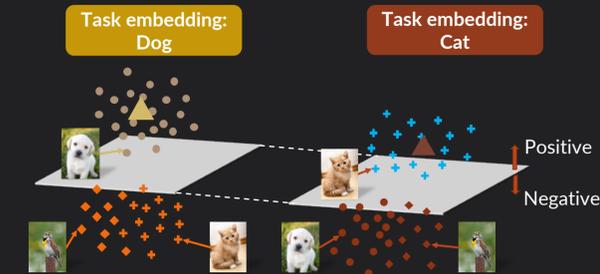
Dynamic Routing

Test-Time Fine-Tuning



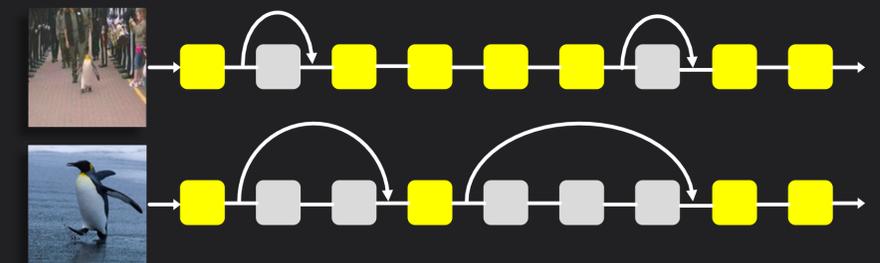
- For few-shot object detection, we first train an object detector on abundant examples from the base classes. At inference, we fine-tune the last layer of the detector on the few labeled examples.
- Model **parameters** are adjusted through fine-tuning **labeled frames** at test time.

Task-Aware Weight Generation



- **Zero-Shot Learning:** generalize to new task without access to labeled data.
- Task descriptions as input to the meta learner to generate task-specific weights.
- Model **parameters** are adjusted to unseen tasks based on **task descriptions.**

Dynamic Routing



- Depths of the convolutional network is decided by the input image
- Easy images skip more layers while challenging examples skill more layers.
- Model **structures** are adjusted on a per **input** basis to achieve reduced computation.