CNAPs: Fast and Flexible Multi-Task Classification Using Conditional Neural Adaptive Processes

James Requeima*[1,2], Jonathan Gordon*[1], John Bronskill*[1], Sebastian Nowozin[3], and Richard E. Turner[1,4]
* Equal contributors. [1]University of Cambridge, [2]Invenia Labs, [3]Google Research, and [4]Microsoft Research.

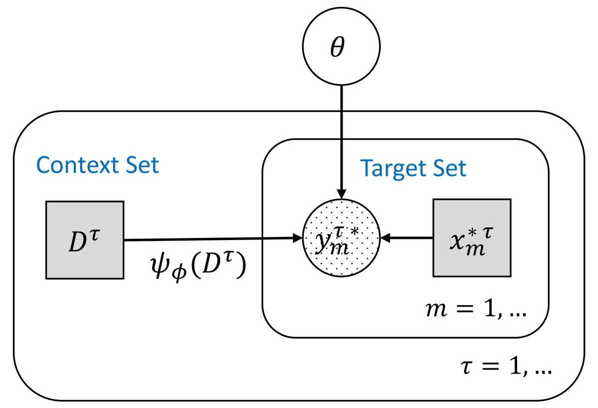
Summary

- **CNAPs**: Image classification system that *automatically adapts* to new tasks *at test time*.
- **Fast**: Adapts network parameters in a single forward pass.
- **Adaptation Networks**: Meta-learn good classifier parameters for any task.
- **Robust**: Avoids over-fitting with few and under-fitting with many training examples.
- **State-of-the-art**: Best on *Meta-Dataset*[4] multi-task, few-shot learning benchmark.
- **Versatile**: Trained model deployable to continual and active learning scenarios.

CNAPs

- Directly specify the predictive distribution in terms of the task context data[1]:
$p\left(\boldsymbol{y}^* | \boldsymbol{x}^*, \boldsymbol{\theta}, D^\tau\right) = p\left(\boldsymbol{y}^* | \boldsymbol{x}^*, \boldsymbol{\theta}, \boldsymbol{\psi}^\tau\right) = p\left(\boldsymbol{y}^* | \boldsymbol{x}^*, \boldsymbol{\theta}, \boldsymbol{\psi}_\phi\left(D^\tau\right)\right)$.

Graphical Model

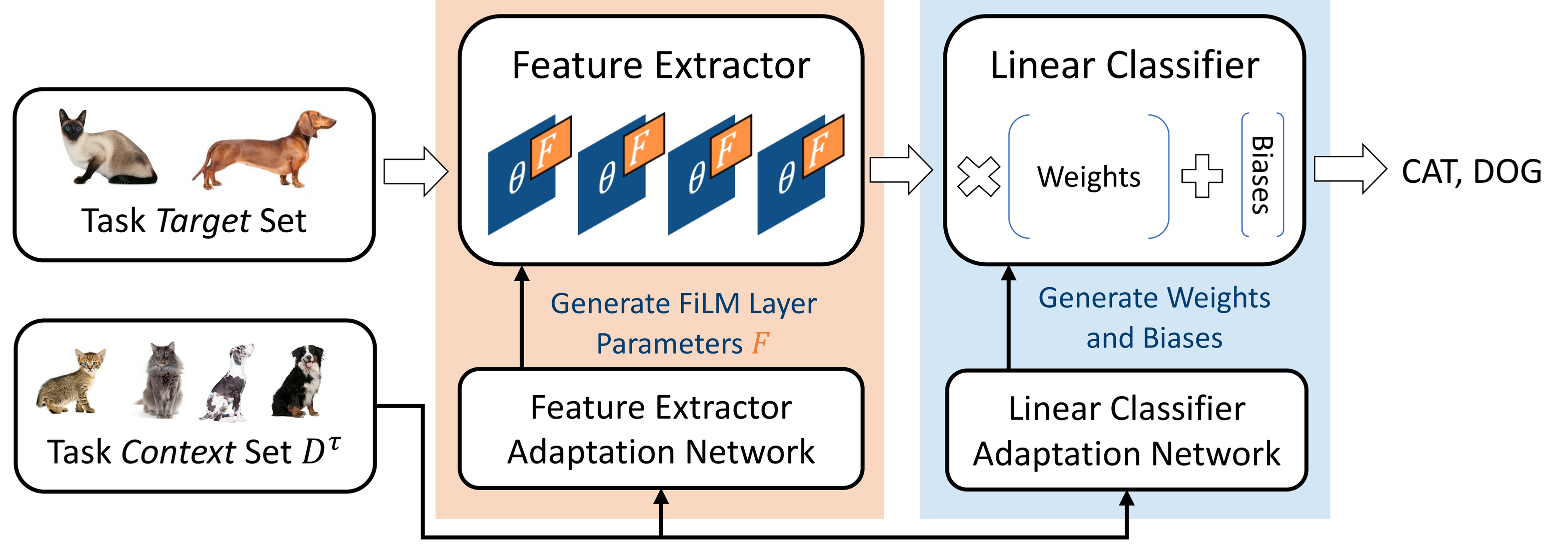


Training

We use the following two-step training procedure to simulate test time adaptation.

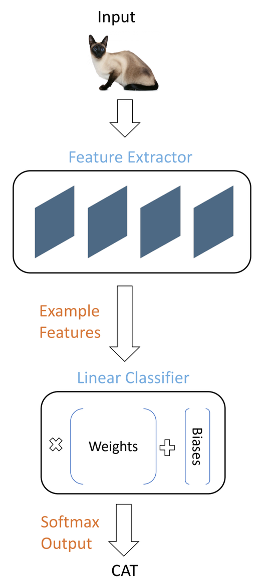
1) Pretrain and fix feature extractor parameters $\boldsymbol{\theta}$ on large dataset.
2) Train adaptation parameters $\boldsymbol{\phi}$ via maximum likelihood:
   (i) Sample a task $\tau \sim p(\tau)$.
   (ii) Randomly split data into context D^τ and target $\{(x_m^\tau, y_m^\tau)\}_{m=1}^{M_\tau}$.
   (iii) Evaluate $L^\tau = \frac{1}{M_\tau}\sum_{m=1}^{M_\tau} \log p(x_m^\tau | y_m^\tau, \boldsymbol{\theta}, \boldsymbol{\psi}_\phi(D^\tau))$.
   (iv) Update $\boldsymbol{\phi} \leftarrow \boldsymbol{\phi} + \eta \nabla_{\boldsymbol{\phi}} L^\tau$

Architecture

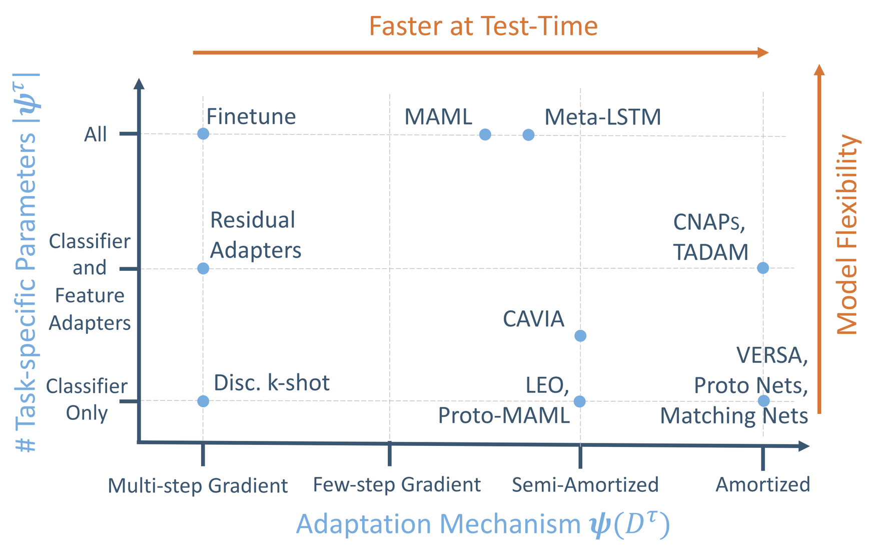


Adapting a Classification Model

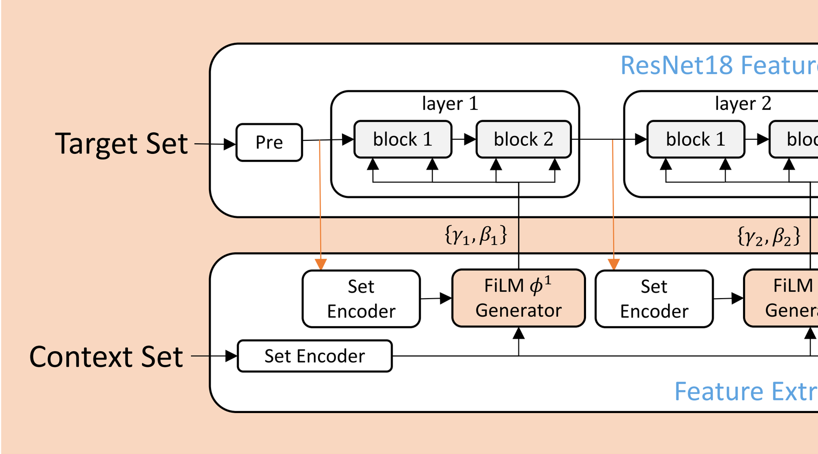
Classification Model

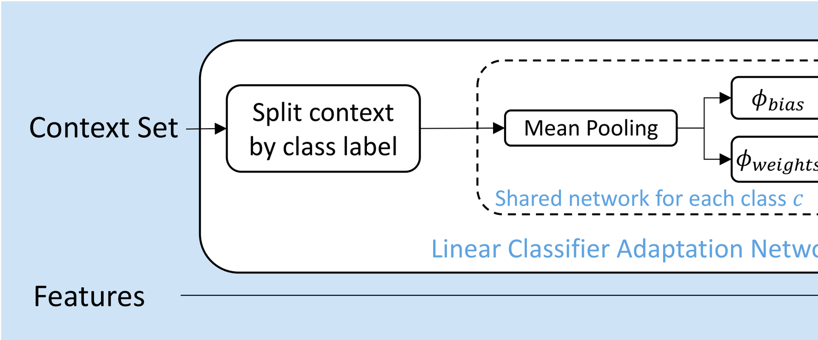


Parameter Adaptation



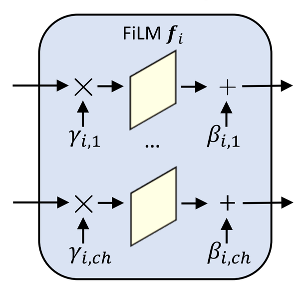
Adaptation Networks ψ



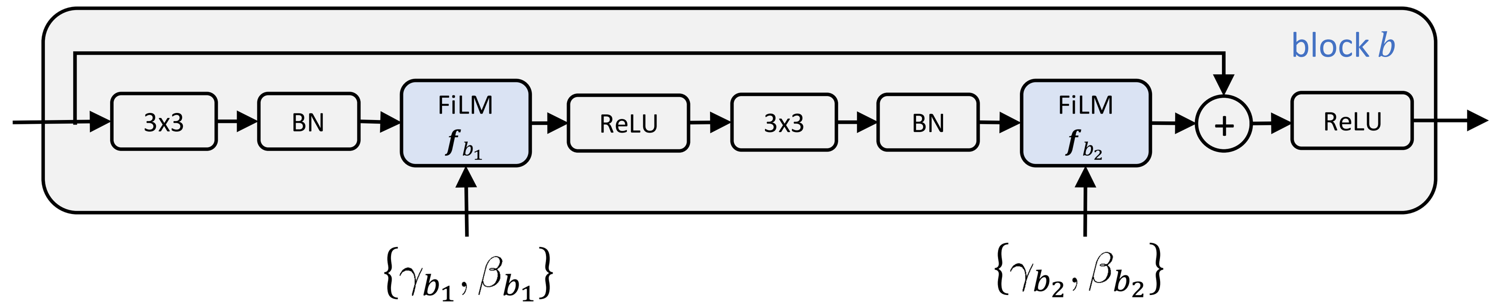


FiLM Layers

FiLM[2] Layer

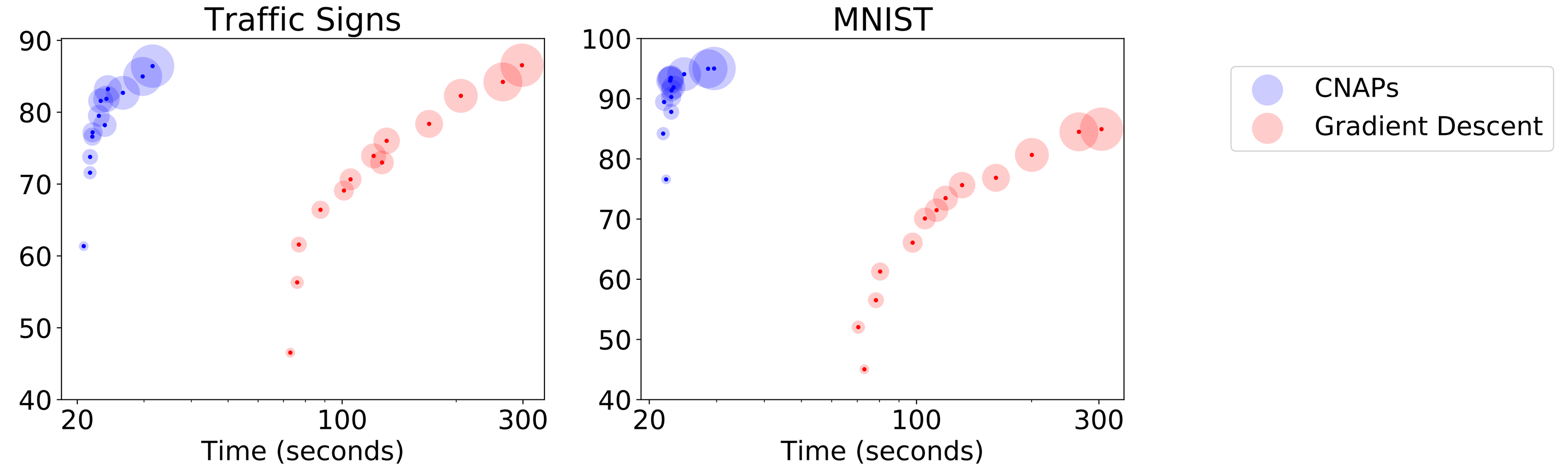


FiLM Layers Added to a ResNet[3] Block



CNAPs vs. Gradient Adaptation

- Accuracy on 5-way classification tasks as a function of processing time. CNAPs $> 5\times$ faster in adapting to an unseen task.
- Dot size reflects shot number (1 to 25 shots). CNAPs is more accurate at low shot and resists over-fitting.

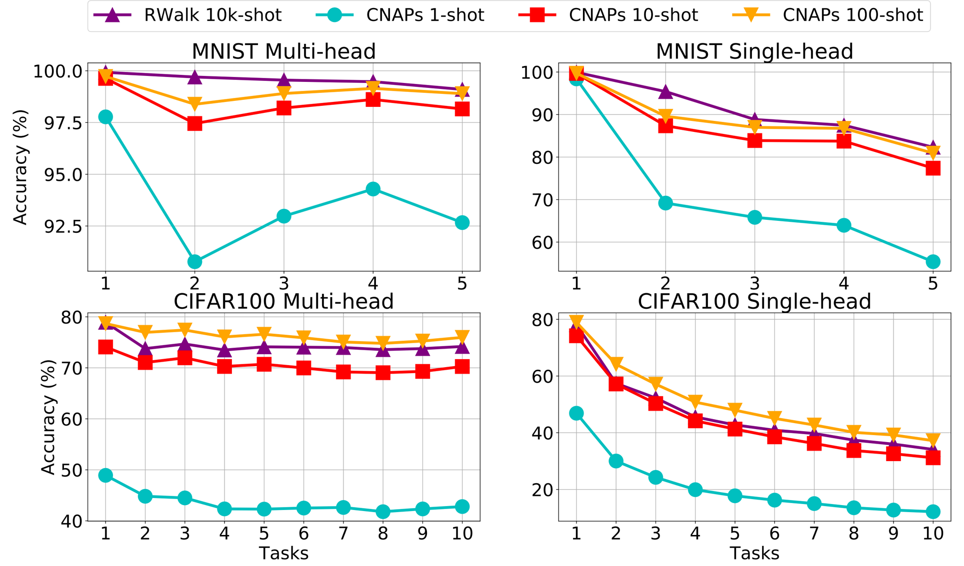


Meta-Dataset[4] Few Shot Image Classification

Dataset	Finetune	MatchingNet	ProtoNet	fo-MAML	Proto-MAML	CNAPs (no FiLM)	CNAPs (no AR)	CNAPs
ILSVRC	43.1	36.1	44.5	32.4	47.9	43.8	**51.3**	**52.3**
Omniglot	71.1	78.3	79.6	71.9	82.9	60.1	**88.0**	**88.4**
Aircraft	72.0	69.2	71.1	52.8	74.2	53.0	76.8	**80.5**
Birds	59.8	56.4	67.0	47.2	70.0	55.7	**71.4**	**72.2**
Textures	**69.1**	61.8	65.2	56.7	67.9	60.5	62.5	58.3
Quick Draw	47.0	60.8	64.9	50.5	66.6	58.1	**71.9**	**72.5**
Fungi	38.2	33.7	40.3	21.0	42.0	28.6	**46.0**	**47.4**
VGG Flower	85.3	81.9	86.9	70.9	**88.5**	75.3	**89.2**	86.0
Traffic Signs	**66.7**	55.6	46.5	34.2	52.3	55.0	60.1	60.2
MSCOCO	35.2	28.8	39.9	24.1	**41.3**	**41.2**	**42.0**	**42.6**

Continual Learning

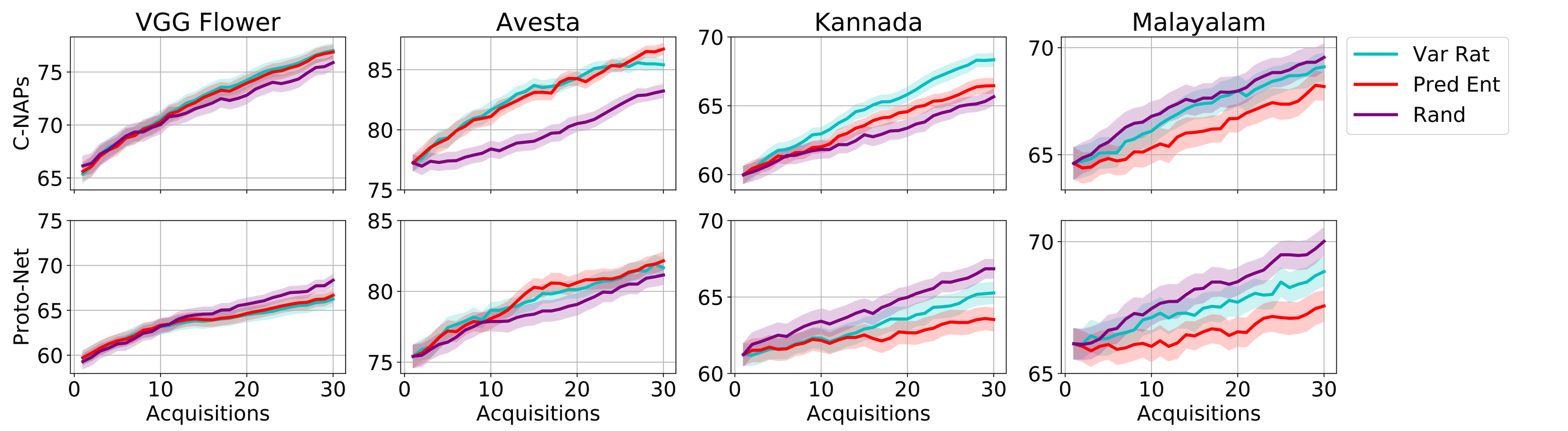
- Compute running averages for ψ^τ, otherwise not specifically trained to do continual learning.
- Model performs incremental updates using the new data and old model, and does not need to access old data.



Method	MNIST Multi	MNIST Single	CIFAR100 Multi	CIFAR100 Single
SI	**99.3**	57.6	73.2	22.8
EWC	**99.3**	55.8	72.8	23.1
VCL	98.5	-	-	-
RWalk	**99.3**	**82.5**	74.2	34.0
CNAPs	98.9	80.9	**76.0**	**37.2**

Active Learning

- Requires accurate data efficient learning that returns well-calibrated uncertainty estimates.
- CNAPs and prototypical networks using variation ratios and predictive entropy compared against random acquisition.



1. M. Garnelo et al. "Conditional neural processes." arXiv preprint arXiv:1807.01613, 2018.
2. E. Perez, et al. "Film: Visual reasoning with a general conditioning layer." AAAI, 2018.
3. K. He et al. "Deep residual learning for image recognition." CVPR, 2016.
4. E. Triantafillou et al. "Meta-dataset: A dataset of datasets for learning to learn from few examples." arXiv:1903.03096, 2019.

Paper: arxiv.org/abs/1906.07697 Code: github.com/cambridge-mlg/cnaps