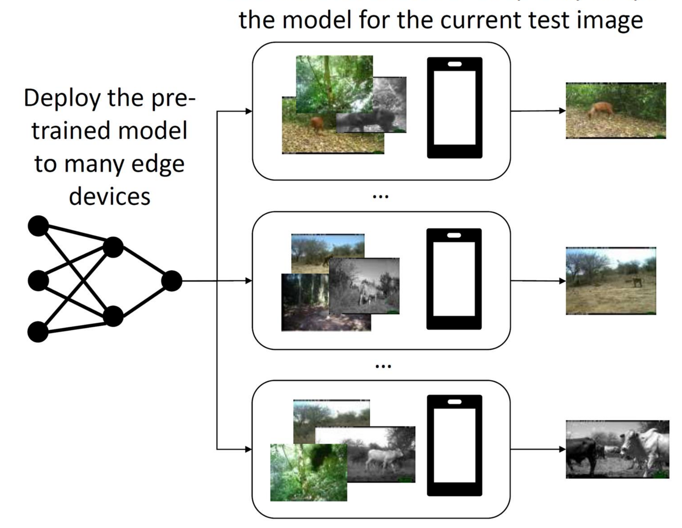
Feed-Forward Latent Domain Adaptation

Ondrej Bohdal, Da Li, Shell Xu Hu, Timothy Hospedales

Samsung Research   THE UNIVERSITY of EDINBURGH informatics

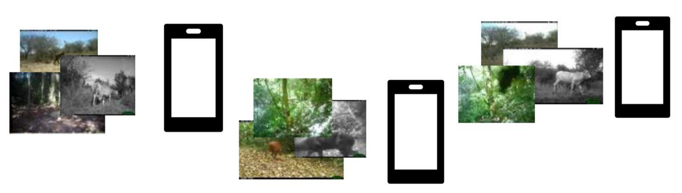
Problem Setting

- Motivation: adapt model on user's device to optimize performance on local data distribution



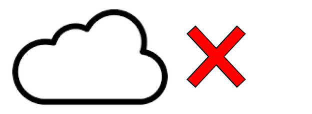
Modelling Users as Tasks

- Each user has their own set of unlabelled images: support set
  - Randomly sampled combination of domains
  - Used for adaptation of the model
- For benchmarking we have labelled test images for each user: query set
  - Coming from one of the domains available in the support set

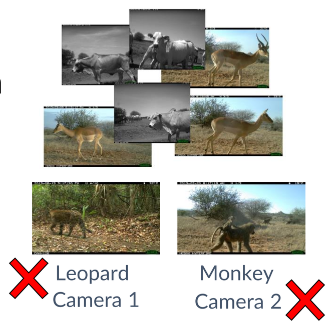


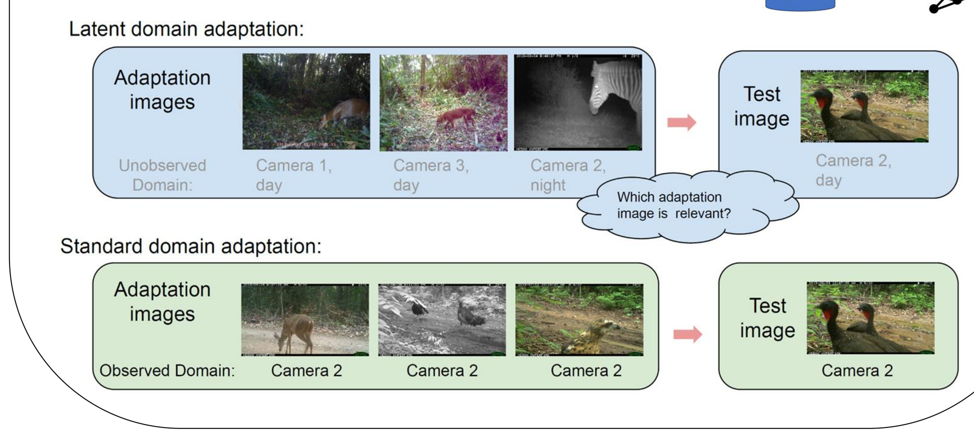
Challenges and Desiderata

- Keep data local for privacy
  - Local processing - no cloud



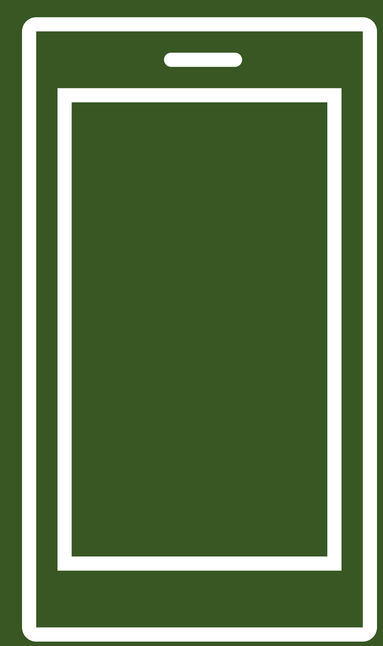
- Feed-forward
  - Backpropagation is slow and may not be supported on edge devices $\frac{\partial \ell}{\partial w}$
- Latent domains
  - User's data have a mixture of domain relevant and irrelevant examples
- No labels for user's examples
  - No class or domain labels
  - Same label space
- Source-free
  - Access only to the pre-trained model, not the source data





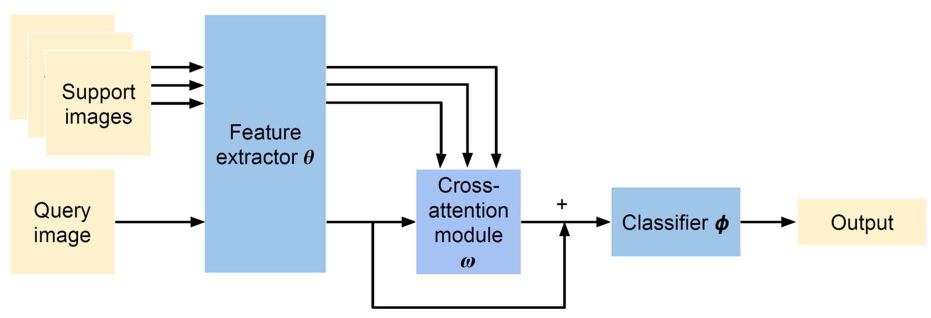
Adapt pre-trained model on-device.

Keep data private.



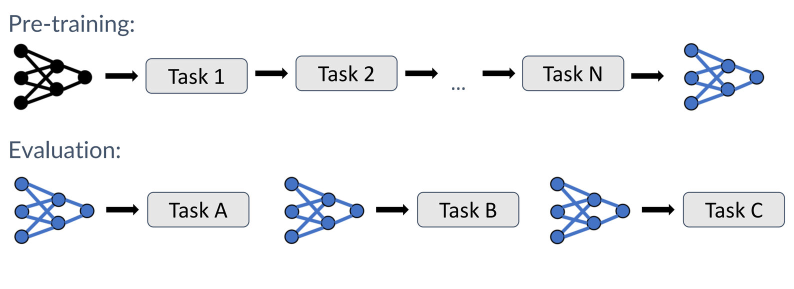
Our Solution: CXDA

- Key idea: use a cross-attention mechanism to identify and exploit relevant support instances for adapting to the query example
- Image-to-image cross attention
  - Flatten all features of an image into a vector



Overall Workflow

- Pre-training across many randomly sampled tasks
  - Inner loop: feed-forward adaptation using the support set
  - Outer loop: gradient-based training of the model parameters
- Evaluation on tasks randomly sampled from new unseen domains
  - Feed-forward adaptation on new tasks



Experiments

- Synthetic and real-world benchmarks: domains are various persons, image corruptions or camera locations
  - FEMNIST, CIFAR-C, TinyImageNet-C, iWildCam
- Domain supervision helpful but can be outperformed

	FEMNIST		CIFAR-C		TinyImageNet-C		iWildCam	
Approach	**W10%**	**Avg**	**W10%**	**Avg**	**W10%**	**Avg**	**W10%**	**Avg**
ERM	52.7 ± 1.4	77.2 ± 0.9	44.3 ± 0.5	68.6 ± 0.3	4.8 ± 0.2	26.4 ± 0.4	0.0 ± 0.0	38.7 ± 0.8
CML [39]	50.4 ± 1.3	76.0 ± 0.9	44.8 ± 0.5	69.5 ± 0.5	4.8 ± 0.5	25.7 ± 0.6	0.0 ± 0.0	38.7 ± 1.1
BN [18, 39]	52.2 ± 1.5	78.0 ± 0.7	45.4 ± 0.7	69.3 ± 0.4	5.9 ± 0.2	27.7 ± 0.3	1.9 ± 1.1	42.5 ± 0.8
Our CXDA	**53.3 ± 0.6**	**78.3 ± 0.0**	**49.4 ± 0.6**	**72.0 ± 0.3**	**6.5 ± 0.2**	**28.6 ± 0.3**	**3.6 ± 1.5**	**43.5 ± 1.5**
FT-EM (TENT) [11]	51.7 ± 1.4	77.6 ± 0.8	44.9 ± 0.6	69.2 ± 0.4	3.9 ± 0.4	25.7 ± 0.3	0.0 ± 0.0	38.6 ± 0.8
FT-IM (SHOT) [26, 31]	52.5 ± 1.2	77.5 ± 0.8	45.6 ± 0.5	69.5 ± 0.3	4.8 ± 0.4	24.6 ± 1.0	0.0 ± 0.0	38.7 ± 0.8
SF-OCDA [40]	51.5 ± 1.4	77.5 ± 0.7	46.7 ± 1.2	70.1 ± 0.5	5.5 ± 0.2	26.7 ± 0.2	0.0 ± 0.0	38.4 ± 0.6
CoTTA [36]	51.4 ± 0.4	76.8 ± 0.2	46.2 ± 0.3	69.8 ± 0.2	4.9 ± 0.5	26.0 ± 0.7	0.0 ± 0.0	38.6 ± 0.5
SLA [7]	46.0 ± 1.4	74.1 ± 0.8	40.8 ± 1.1	64.0 ± 0.7	2.5 ± 0.1	16.9 ± 0.3	0.0 ± 0.0	29.9 ± 1.4

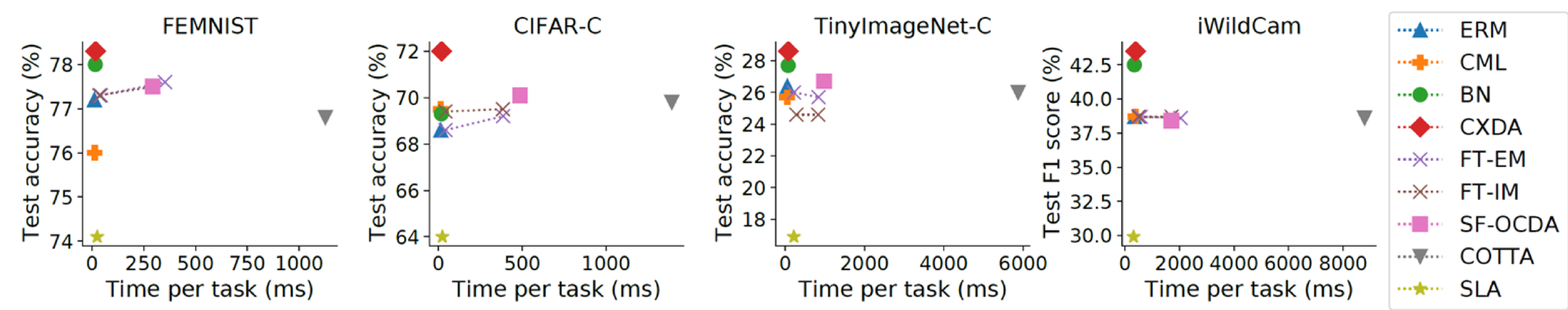
Table 1. Main benchmark results: average and worst-case (worst 10% tasks) test performance, with standard error of the mean across 3 random seeds. Accuracy is reported for all except iWildCam, where F1 score is used (%). The best results are highlighted in bold. Our CXDA approach achieves the best performance across all of the benchmarks.

Cross-attention	FEMNIST	CIFAR-C	TinyImageNet-C	iWildCam
Domain-unsupervised	78.3 ± 0.0	72.0 ± 0.3	28.6 ± 0.3	43.5 ± 1.5
Domain-supervised	79.4 ± 0.4	69.8 ± 0.4	28.6 ± 0.2	52.0 ± 1.2

Table 2. Comparison of domain-unsupervised and domain-supervised CXDA on our benchmarks. Average test accuracy for all benchmarks apart from iWildCam where F1 score is reported (%). Domain supervision is helpful in multiple cases, but can be outperformed.

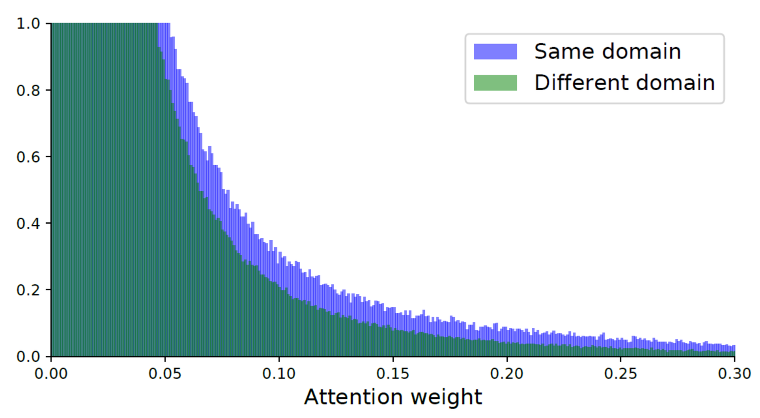
Speed Evaluation

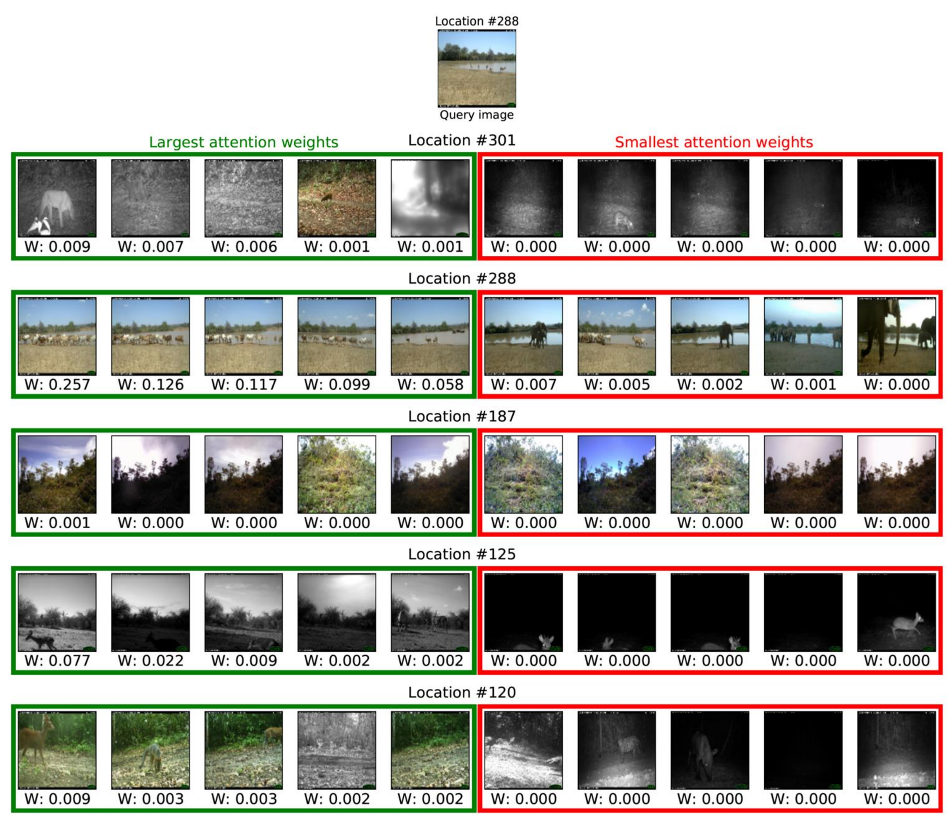
- Our CXDA:
  - Best performance
  - Capable of real-time adaptation with similar speed as the other feed-forward baselines
  - Significantly faster than the back-propagation based approaches



Analysis of Attention Weights

- Significant weight spent on attending to examples in different domains
  - Exploiting knowledge transfer beyond the boundaries of the standard domain annotation
- Overall more attention to the in-domain instances
  - Learned to match query instances with corresponding domain instances in the support set





Summary

- New highly practical problem setting for resource-constrained devices
  - Unlabelled data
  - Mixture of domains
  - Feed-forward adaptation
- Novel solution based on cross attention that selects relevant examples and uses them for real-time adaptation
- Project page: https://ondrejbohdal.github.io/cxda