ICLR

Unsupervised Meta-Learning via In-Context Learning

Anna Vettoruzzo, Lorenzo Braccaioli, Joaquin Vanschoren, Marlena Nowaczyk

Code

Contacts

HALMSTAD UNIVERSITY

UNIVERSITY OF TRENTO

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Motivation

Unsupervised data

Unlabeled data is everywhere.

Generalization

Train once and apply to different datasets.

Fast inference

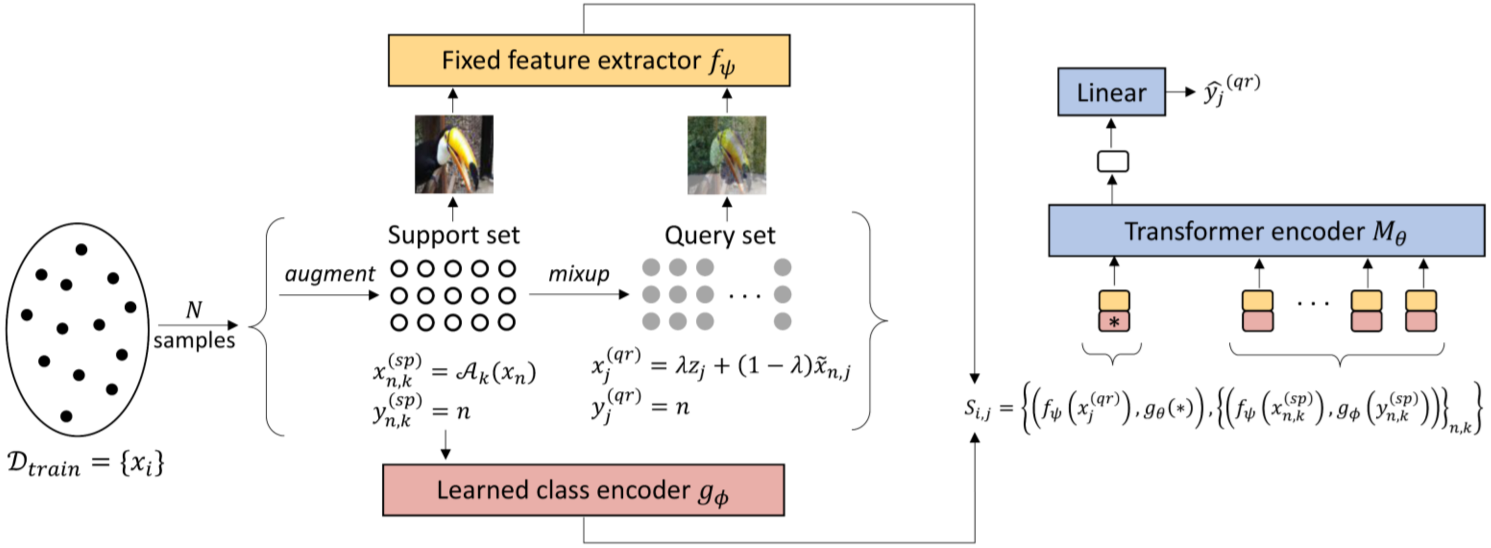
In-context learning for quick predictions, without fine-tuning.

Method

CAMeLU: Context **A**ware **Me**ta-**L**earning in **U**nsupervised scenarios.

Key features:

- Propose a **novel task creation mechanism** for unsupervised meta-learning.
- Reframe meta-learning as a **non-causal sequence modeling problem**.
- Meta-learn a transformer model **for efficient cross-domain generalization**.

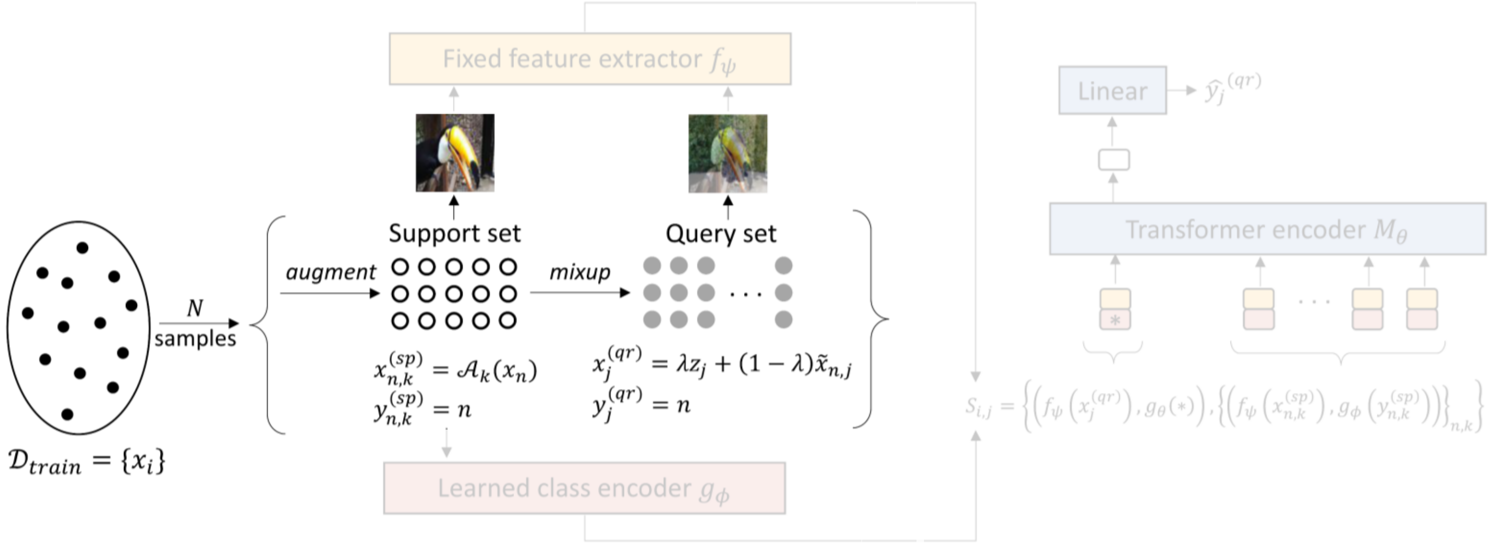


Method

CAMeLU: Context **A**ware **Me**ta-**L**earning in **U**nsupervised scenarios.

Key features:

- Propose a **novel task creation mechanism** for unsupervised meta-learning.
- Reframe meta-learning as a **non-causal sequence modeling problem**.
- Meta-learn a transformer model **for efficient cross-domain generalization**.

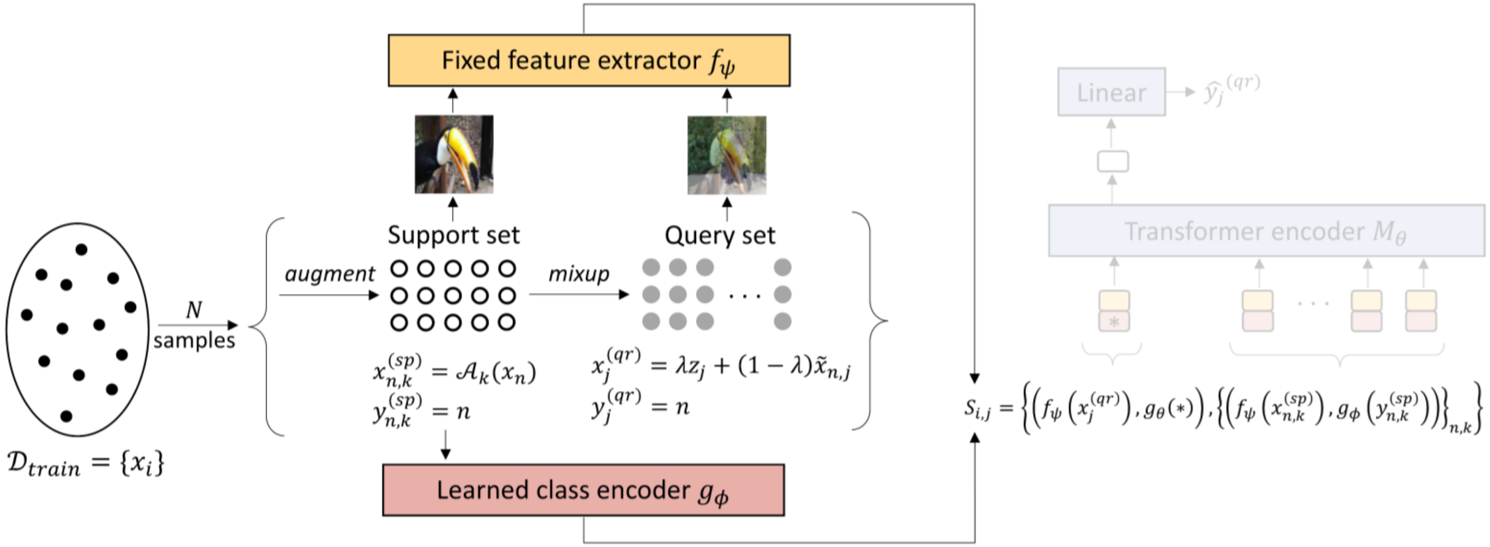


Method

CAMeLU: Context **A**ware **Me**ta-**L**earning in **U**nsupervised scenarios.

Key features:

- Propose a **novel task creation mechanism** for unsupervised meta-learning.
- Reframe meta-learning as a **non-causal sequence modeling problem**.
- Meta-learn a transformer model **for efficient cross-domain generalization**.

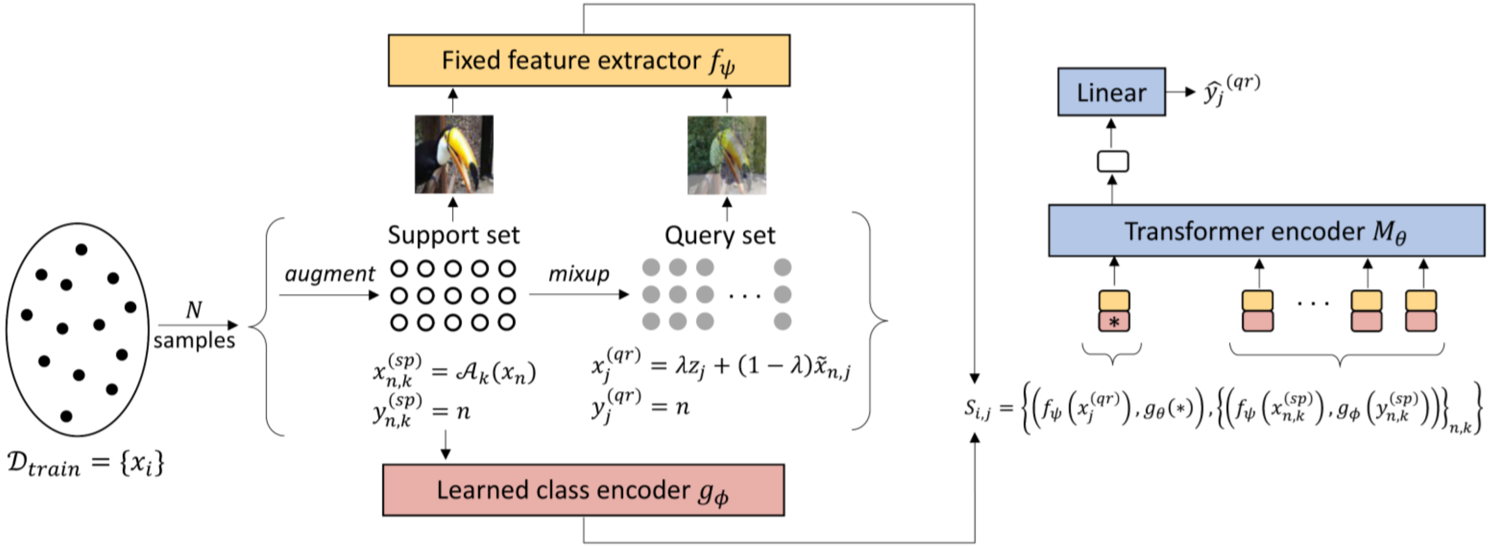


Method

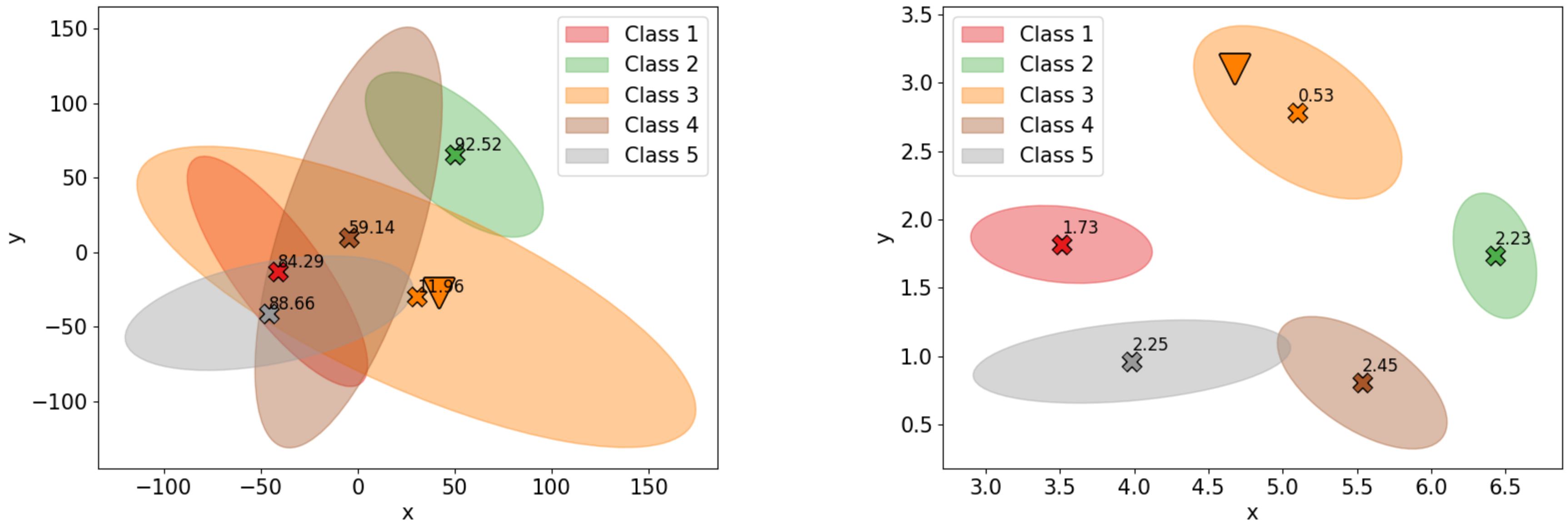
CAMeLU: Context **A**ware **Me**ta-**L**earning in **U**nsupervised scenarios.

Key features:

- Propose a **novel task creation mechanism** for unsupervised meta-learning.
- Reframe meta-learning as a **non-causal sequence modeling problem**.
- Meta-learn a transformer model **for efficient cross-domain generalization**.



Feature space representation



(a) Feature extractor

(b) Transformer encoder

Figure 1: Visualization of clustered embeddings obtained with CAMeLU after the feature extractor and the transformer encoder on a 5-way 5-shot task sampled from the CUB dataset.

From memorization to generalization

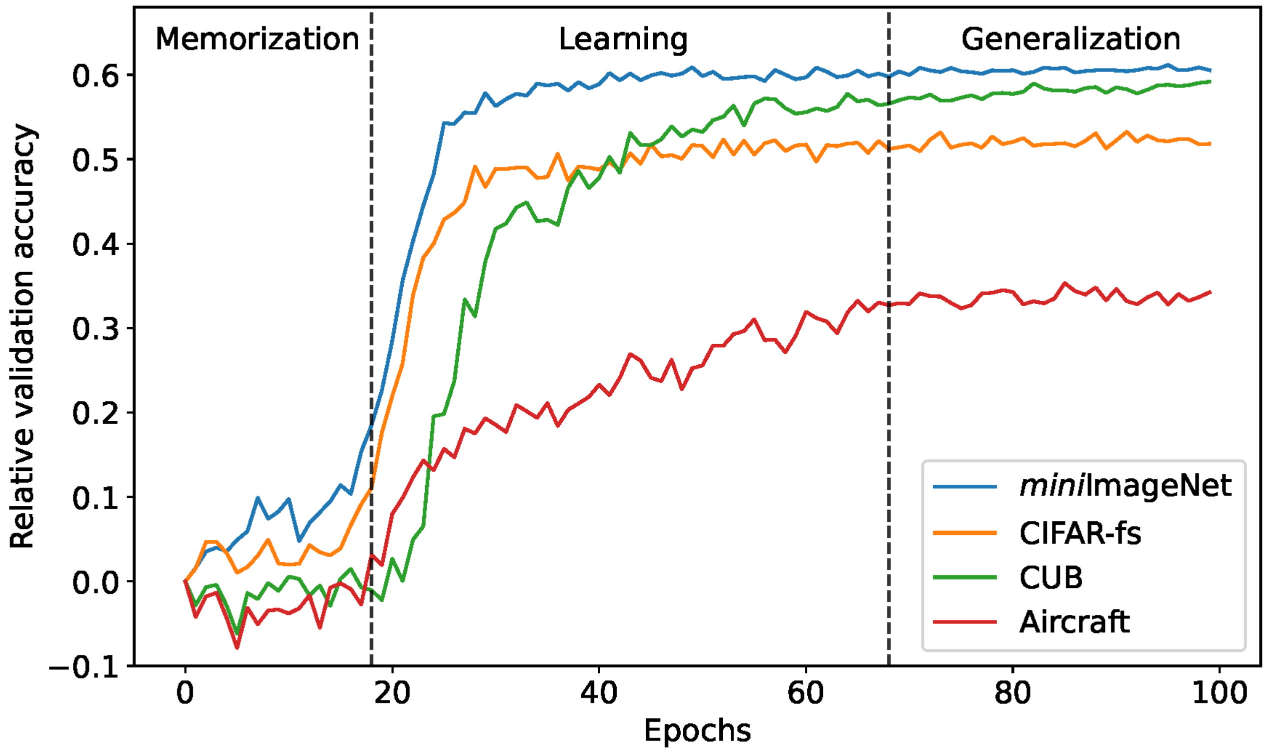


Figure 2: Learning trend when transferring knowledge from a different prior dataset (ImageNet-964).

Comparative results

Table 1: Accuracy (%) on 5-way 5-shot tasks. The symbol † indicates results affected by data leakage.

	*mini*ImageNet		CIFAR-fs		CUB		Aircraft		Meta-iNat	
Method	5w1s	5w5s	5w1s	5w5s	5w1s	5w5s	5w1s	5w5s	5w1s	5w5s
In-Domain										
CACTUs-MAML	43.30	54.21	42.00	56.64	31.19	36.81	24.06	27.26	20.13	21.84
CACTUs-ProtoNet	48.85	62.52	50.90	64.52	33.93	44.41	26.27	30.88	27.30	29.08
UMTRA	39.93	50.73	32.93	46.13	27.06	36.6	22.40	31.73	28.96	37.12
Meta-GMVAE	55.38†	65.10†	52.02	64.18	33.59	39.09	24.83	27.60	34.22	40.23
PsCo	47.29	64.85	42.21	62.92	33.09	51.02	26.19	38.80	36.97	55.88
Cross-Domain										
PsCo	67.89	90.17	53.34	76.22	43.35	70.19	29.87	38.20	46.21	70.05
CAMeLU	**76.51**	**92.14**	**61.79**	**80.43**	**65.52**	**80.35**	**33.17**	**39.11**	**57.27**	**75.45**
CAML (supervised)	81.75	92.31	59.44	75.27	54.63	66.81	28.92	32.06	50.86	67.07

Conclusion

- We present a novel approach to unsupervised meta-learning through in-context learning.
- We propose a novel task creation mechanism.
- Our approach
  - ✓ generalizes cross-domain,
  - ✓ is scalable,
  - ✓ makes fast, in-context predictions without fine-tuning.

Code

Contacts

TU/e

ICLR

Unsupervised Meta-Learning via In-Context Learning

Anna Vettoruzzo, Lorenzo Braccaioli, Joaquin Vanschoren, Marlena Nowaczyk

Code

Contacts

HALMSTAD UNIVERSITY

UNIVERSITY OF TRENTO

EINDHOVEN UNIVERSITY OF TECHNOLOGY