TAAN: Task-Aware Attention Network for Few-shot Classification

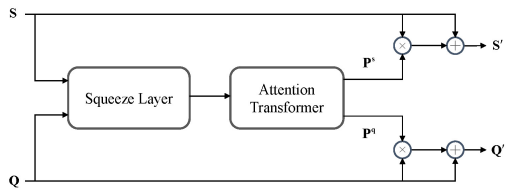
Zhe Wang, Li Liu, and FanZhang Li
School of computer science and technology, Soochow University

Introduction

- Few-shot classification aims to recognize unlabeled samples from unseen classes given only a few labeled samples.
- Metric-based methods learn to represent image data in an appropriate feature space and use a distance metric to predict image labels.
- These approaches all extract features from samples independently without looking at the entire task as a whole, and so fail to provide an enough discrimination to features.

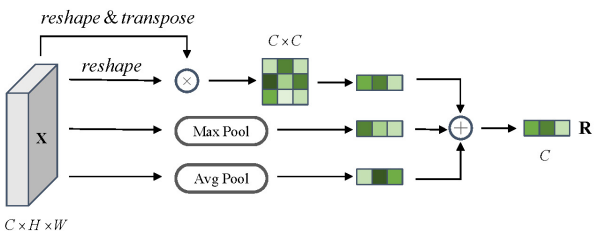
Methods

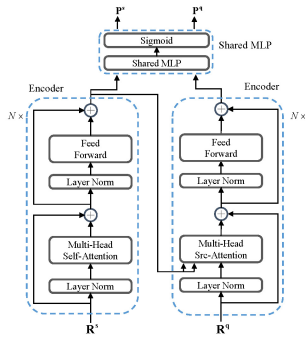
- Our method exploits the semantic relevance between instance features in a task to highlight the discriminating feature dimension.
- We introduce a ***Task-Relevant Channel Attention Module*** to model bi-level context, including instance-level context and set-level context. Then we use the set-level context to guide the generation of task-relevant channel attention for each support and query examples.



Task-Relevant Channel Attention Module

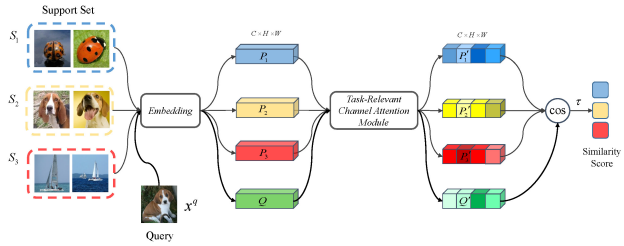
- The ***Squeeze Layer*** firstly generates the channel representation for each instance, which is treated as instance-level channel attention.





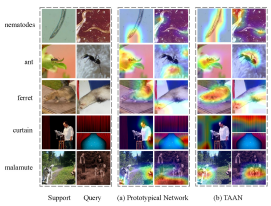
- Then, the ***Attention transformer*** helps to transform channel representation to task-level channel attention by looking at all support classes.

- The ***Task-Aware Attention Network*** consists of three modules: an embedding module, a task-relevant channel attention module and a classification module.



Experiments & Visualization Analysis

Model	Backbone	mini-ImageNet 1-shot	mini-ImageNet 5-shot	tierd-ImageNet 1-shot	tierd-ImageNet 5-shot
Model-based					
Meta-learner LSTM [2]	ConvNet-4	43.44 ± 0.77	60.60 ± 0.71	-	-
Meta-SGD [27], reported by [19]	ConvNet-4	54.24 ± 0.03	70.86 ± 0.04	62.95 ± 0.03	79.34 ± 0.06
SNAIL [15]	ResNet-12	55.71 ± 0.99	68.88 ± 0.92	-	-
Optimization-based					
MAML [16]	ConvNet-4	48.70 ± 1.75	63.15 ± 0.91	51.67 ± 1.81	70.30 ± 1.75
Cosine Classifier [4]	ResNet-12	55.43 ± 0.81	77.18 ± 0.61	61.49 ± 0.91	82.37 ± 0.67
MTL [28]	ResNet-12	61.20 ± 1.80	75.50 ± 0.80	-	-
LEO [19]	WRN-28-10	61.76 ± 0.08	77.59 ± 0.12	66.33 ± 0.05	81.44 ± 0.09
MetaOptNet [20]	ResNet-12	62.64 ± 0.61	78.63 ± 0.46	65.99 ± 0.72	81.56 ± 0.53
Metrics-based					
Matching network [8]	ConvNet-4	43.56 ± 0.84	55.31 ± 0.73	-	-
Relation network [10], reported by [20]	ConvNet-4	50.44 ± 0.82	65.32 ± 0.70	54.48 ± 0.93	71.32 ± 0.78
Prototypical network [9], reported by [20]	ConvNet-4	49.42 ± 0.78	68.20 ± 0.66	53.31 ± 0.89	72.69 ± 0.74
TADAM [21]	ResNet-12	58.5 ± 0.3	76.7 ± 0.3	-	-
Cross Attention network [29]	ResNet-12	63.85 ± 0.48	79.44 ± 0.34	69.89 ± 0.51	84.23 ± 0.37
CTM [30]	ResNet-18	62.05 ± 0.55	78.63 ± 0.06	64.78 ± 0.11	81.05 ± 0.52
CTM, data augment [30]	ResNet-18	64.12 ± 0.82	80.51 ± 0.13	68.41 ± 0.39	84.28 ± 1.73
Baselines					
Classifier-Baseline [31]	ResNet-12	58.91 ± 0.23	77.76 ± 0.17	68.07 ± 0.26	83.74 ± 0.18
Meta-Baseline [31]	ResNet-12	63.17 ± 0.23	79.26 ± 0.17	68.62 ± 0.27	83.29 ± 0.18
Prototypical Network (re-implement)	ResNet-12	59.88 ± 0.45	78.50 ± 0.33	65.65 ± 0.33	83.40 ± 0.13
Ours: TAAN	ResNet-12	**66.48 ± 0.45**	**81.40 ± 0.31**	**70.07 ± 0.51**	**84.56 ± 0.35**



	SL self-attention	SL avg-pool	SL max-pool	AT	1-shot	5-shot
(i)	-	-	-	-	62.67	78.20
(ii)	✓	✓	✓	-	63.56	79.56
(iii)	✓	-	-	✓	65.79	81.05
	-	✓	-	✓	65.83	81.02
	-	-	✓	✓	65.38	81.11
(iv)	✓	✓	✓	✓	66.59	81.54

Conclusion

In this paper, we propose the ***Task-Aware Attention Network***, a more effective and robust deep metric network with attention mechanism for few-shot learning. Our method can adaptively select the most relevant channels and generate more discriminative features for the target task.