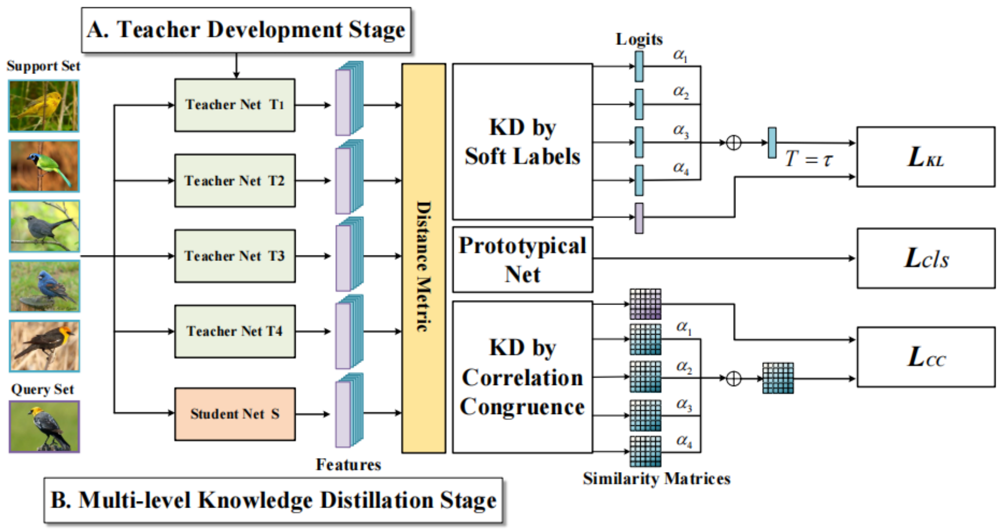
Teachers Cooperation: Team-Knowledge Distillation for Multiple Cross-Domain Few-Shot Learning

Zhong JI, Jingwei NI, Xiyao LIU, Yanwei PANG

Frontiers of Computer Science, DOI: 10.1007/s11704-022-1250-2

Problems & Ideas

- Problems of Multiple Cross-Domain Few-Shot Learning:
  - Domain shift between the auxiliary set and the test set.
  - Multiple source domain data may cause the problem of knowledge interference, resulting in negative knowledge transfering.
- Ideas: A multi-teacher knowledge distillation based method that incorporates task-oriented knowledge distillation and multi-level knowledge distillation.



The architecture illustration of the proposed TKD-Net on 5-way 1-shot classification. It first respectively pre-trains teacher models on multiple seen domains by supervised learning in the teacher development stage. Then multi-level knowledge is distilled from the cooperation of teachers to the student model by soft labels and correlation congruence. The second stage is under the paradigm of meta-learning. The final student model is employed to few-shot classification tasks on the unseen domain.

Main Contributions

- Contributions:
  - We propose a knowledge distillation based method for multiple cross-domain few-shot learning, named Team-Knowledge Distillation Networks (TKD-Net). It incorporates task-oriented knowledge distillation and multiple cooperation among teachers to train a good student with better generalization ability on the unseen domain.
  - To further exploit multi-level knowledge from teacher models, we distill both response-based knowledge and relation-based knowledge into the student model and introduce a novel objective function to guide the training of the student model.

5-way 1-shot				
Method	CUB	Cars	Places	Plantae
MatchingNet [1]	39.24 ± 0.55%	30.33 ± 0.45%	41.46 ± 0.59%	31.71 ± 0.51%
ProtoNet [3]	36.54 ± 0.52%	29.38 ± 0.42%	40.12 ± 0.59%	31.42 ± 0.49%
RelationNet [4]	39.29 ± 0.54%	30.46 ± 0.44%	40.86 ± 0.61%	32.81 ± 0.52%
MAML(w/o pre-train) [5]	35.06 ± 0.54%	31.12 ± 0.54%	36.14 ± 0.56%	30.95 ± 0.49%
MAML [5]	35.50 ± 0.53%	26.76 ± 0.42%	39.12 ± 0.60%	31.35 ± 0.49%
Proto-MAML [15]	36.05 ± 0.53%	29.46 ± 0.44%	38.71 ± 0.57%	31.20 ± 0.49%
LFT [13] on MatchingNet	34.20 ± 0.53%	30.15 ± 0.46%	39.43 ± 0.60%	29.50 ± 0.39%
LFT [13] on RelationNet	39.70 ± 0.58%	**32.59 ± 0.54%**	39.92 ± 0.59%	33.11 ± 0.56%
TKD-Net (Ours)	**40.69 ± 0.51%**	31.36 ± 0.46%	**45.71 ± 0.60%**	**34.53 ± 0.52%**

5-way 5-shot				
Method	CUB	Cars	Places	Plantae
MatchingNet [1]	53.71 ± 0.53%	38.39 ± 0.48 %	55.96 ± 0.58%	45.00 ± 0.54%
ProtoNet [3]	54.79 ± 0.56%	41.76 ± 0.55%	59.91 ± 0.56%	42.99 ± 0.51%
RelationNet [4]	55.29 ± 0.52%	43.58 ± 0.56%	58.62 ± 0.58%	46.01 ± 0.55%
MAML(w/o pre-train) [5]	53.20 ± 0.54%	43.71 ± 0.56%	53.91 ± 0.57%	44.70 ± 0.53%
MAML [5]	52.66 ± 0.52%	43.43 ± 0.53%	56.61 ± 0.58%	42.72 ± 0.55%
Proto-MAML [15]	57.21 ± 0.54%	45.06 ± 0.56%	58.38 ± 0.57%	47.45 ± 0.55%
LFT [13] on MatchingNet	49.09 ± 0.53%	42.42 ± 0.53%	54.15 ± 0.54%	43.32 ± 0.53%
LFT [13] on RelationNet	55.53 ± 0.59%	**46.05 ± 0.55%**	53.17 ± 0.55%	45.66 ± 0.56%
TKD-Net (Ours)	**57.85 ± 0.53%**	43.93 ± 0.52%	**64.96 ± 0.56%**	**48.61 ± 0.53%**

Multiple cross-domain few-shot classification accuracy of TKD-Net with ±95% confidence intervals. Left: 5-way 1-shot classification result; Right: 5-way 5-shot classification result.