北京郵電大學
Beijing University of Posts and Telecommunications

Meta Self-Learning for Multi-Source Domain Adaptation: A Benchmark

Shuhao Qiu, Chuang Zhu, Wenli Zhou
Beijing University of Posts and Telecommunications
Beijing, China

Main Contributions

北京郵電大學 Beijing University of Posts and Telecommunications

- A multi-source domain adaptation dataset for text recognition
  - First dataset in the area
- A new method: Meta Self-Learning
  - Improves the quality of pseudo-label
  - Can be easily applied to any task
- A benchmark for the dataset

BUPT_AI CVSM Computer Vision and Smart Medicine

https://github.com/bupt-ai-cz/Meta-SelfLearning

Dataset Overview

北京郵電大學
Beijing University of Posts and Telecommunications

- Five different domains
- More than five million images
- A wide variety of length, appearance and corpus.

https://github.com/bupt-ai-cz/Meta-SelfLearning

BUPT_AI
CVSM
Computer Vision and Smart Medicine

Meta Self-Learning Method

北京郵電大學 Beijing University of Posts and Telecommunications

Algorithm Description

① The data from source domains with labels $\mathrm{D_s}$ are used for warm-up, which is very important for self-learning method.

② The model is evaluated on the target domain data without labels $\overline{\mathrm{D}}_\mathrm{T}$ and generates pseudo-labels.

③ The target domain data with pseudo-labels $\breve{\mathrm{D}}_\mathrm{T}$ and $\mathrm{D_s}$ are split randomly as $\overline{\mathrm{M}}$ and $\widetilde{\mathrm{M}}$.

④ Use $\overline{\mathrm{M}}$ for meta-train, the parameter is updated as $\theta' = \theta - \alpha \frac{\partial \mathrm{l}(\theta\prime;\bar{M})}{\partial \theta}$

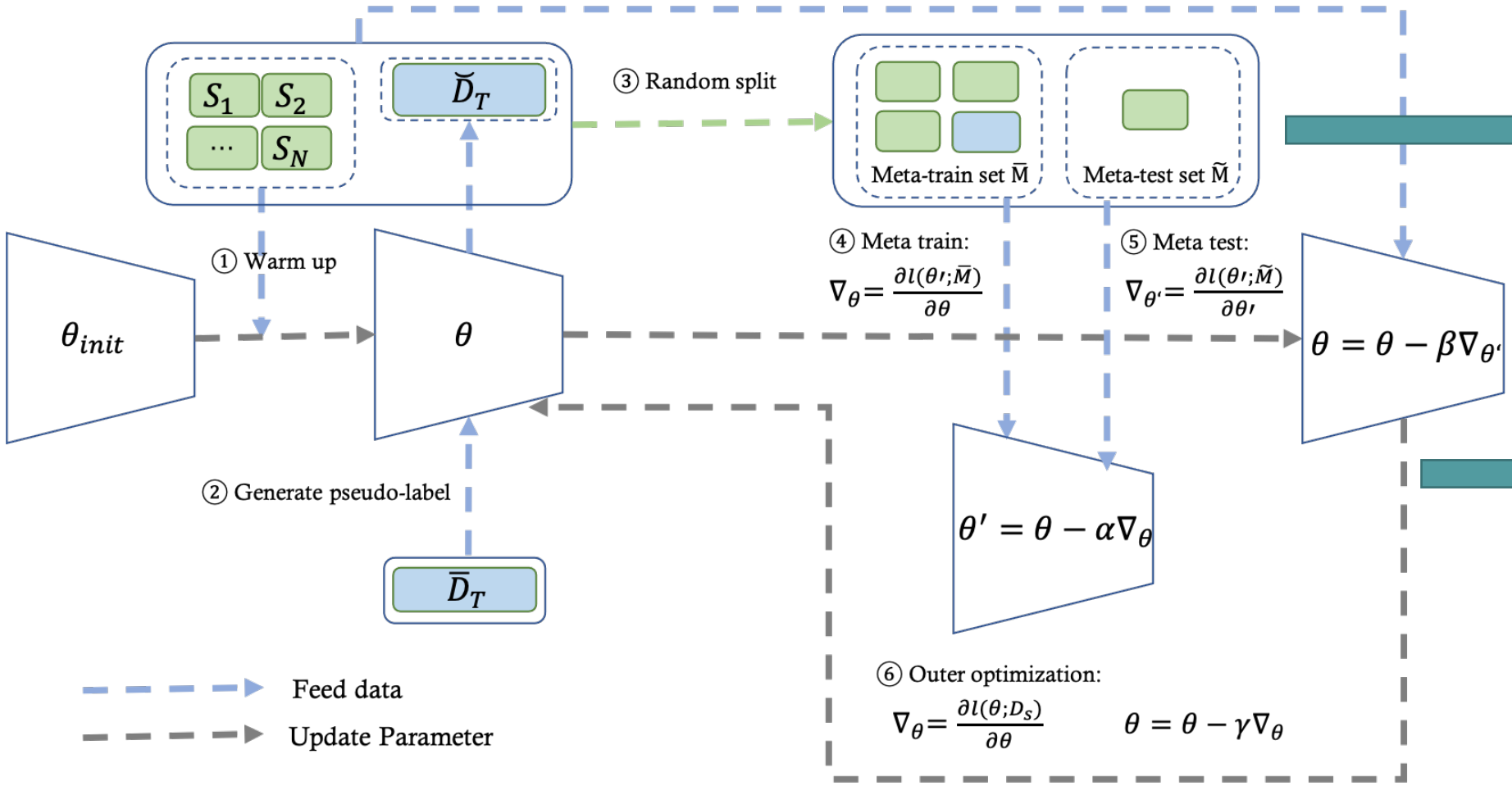
⑤ Use $\widetilde{\boldsymbol{M}}$ for meta-test, the parameter is updated as $\theta = \theta - \beta \frac{\partial l(\theta\prime;\tilde{M})}{\partial \theta\prime}$

⑥ Use a subet of $\boldsymbol{D_s}$ and $\breve{\boldsymbol{D}}_{\boldsymbol{T}}$ for outer optimization, $\theta = \theta - \gamma \frac{\partial l(\theta;D_S)}{\partial \theta}$

https://github.com/bupt-ai-cz/Meta-SelfLearning

BUPT AI CVSM Computer Vision and Smart Medicine

Meta Self-Learning Method



The pseudo-label image is added to the meta-update phase, which can greatly improves the accuracy of pseudo-label

Easy to apply

https://github.com/bupt-ai-cz/Meta-SelfLearning

Experiment Results

	St,Sy,D,H→C	St,Sy,D,C→H	St,Sy,C,H→D	C,St,D,H→Sy	C,Sy,D,H→St	Average
Source Only	22.43%	3.50%	29.39%	24.75%	9.24%	17.86%
MLDG [16]	23.85%	3.39%	30.31%	25.11%	12.46%	19.02%
Pseudo-Label [14]	44.97%	3.77%	51.60%	54.11%	15.00%	33.89%
Meta Self-Learning (Ours)	**58.64%**	**5.41%**	**64.09%**	**65.33%**	**16.52%**	**42.00%**

Fig 1. Experiment results of different methods on our dataset

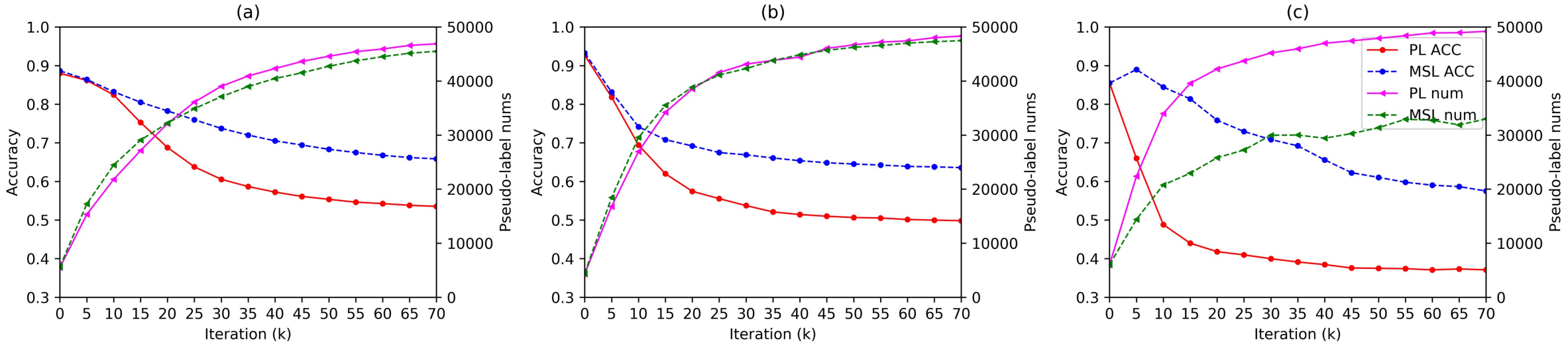


Fig 2. Accuracy of pseudo-label during training for different domain

Future Work

- Challenging in all domains.
- Better domain adaptation method.

https://github.com/bupt-ai-cz/Meta-SelfLearning

Thanks For Watching