Multi-Task Knowledge Distillation for Eye Disease Prediction

Our implementation of knowledge distillation on multi task learning

Malwina Wojewoda, Mateusz Sperkowski, Szymon Rećko

Warsaw University of Technology

April 9, 2022

Table of Contents

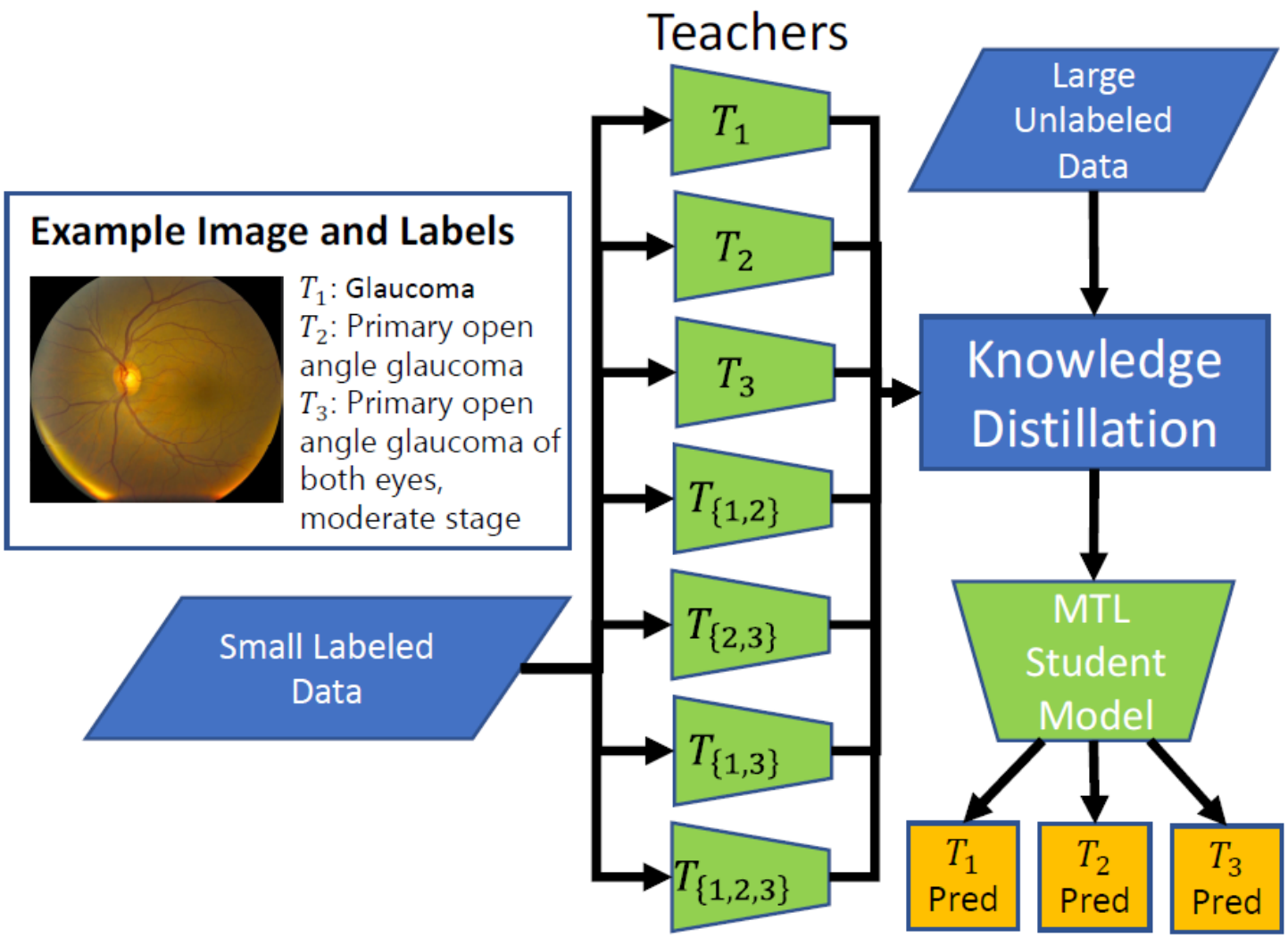
Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA

Paper

Given a fundus image, authors of [Chelaramani et al., 2021] aim to evaluate various solutions for learning deep neural classifiers using small labeled data for three tasks related to eye disease prediction.The problem is challenging because of small data size, need for predictions across multiple tasks, handling image variations, and large number of hyper-parameter choices. Their solution is to create MTL-based teacher ensemble method for knowledge distillation.

Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA

Architecture



Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA

Literature Review

- Overviews of Computer Vision application in the medical field [Shen et al., 2017] and [Razzak et al., 2018]
- Multi Task Learning: first article on the topic [Caruana, 1997]. Current day overview and a glimpse into the future [Ruder, 2017].
- Survey covering knowledge distilation [Gou et al., 2021] from many different perspectives and approaches.
- First papers on combinations of these transfer learning methods appear around 2016. [Li and Bilen, 2020] propose another approach to this problem and [Liu et al., 2019] propose usage of those methods in NLP task.

Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA

Division of work

- Selecting datasets
- Data processing
- Creating MTL teachers models
- Distill knowledge to student model
- Evaluate results

Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA

Software and data

Software:

- PyTorch and Jupyter Notebooks - neural network development and data processing
- GitHub - code versioning and collaboration
- Google Colab - neural network training

Data:

- Fundus images from paper
- Chest x-ray
- CT scans

Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA

Planned Experiments

Testing datasets from different background, as in medical and from another field.
Comparing results from this architecture to baseline models and further testing, experiments depending on the implementation results.

Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA

References I


Caruana, R. (1997).
Multitask learning.
Machine learning, 28(1):41–75.


Chelaramani, S., Gupta, M., Agarwal, V., Gupta, P., and Habash, R. (2021).
Multi-task knowledge distillation for eye disease prediction.
In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 3983–3993.


Gou, J., Yu, B., Maybank, S. J., and Tao, D. (2021).
Knowledge distillation: A survey.
International Journal of Computer Vision, 129(6):1789–1819.

Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA

References II


Li, W.-H. and Bilen, H. (2020).
Knowledge distillation for multi-task learning.
In *European Conference on Computer Vision*, pages 163–176. Springer.


Liu, X., He, P., Chen, W., and Gao, J. (2019).
Improving multi-task deep neural networks via knowledge distillation for natural language understanding.
arXiv preprint arXiv:1904.09482.


Razzak, M. I., Naz, S., and Zaib, A. (2018).
Deep learning for medical image processing: Overview, challenges and the future.
Classification in BioApps, pages 323–350.


Ruder, S. (2017).
An overview of multi-task learning in deep neural networks.
arXiv preprint arXiv:1706.05098.

Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA

References III

Shen, D., Wu, G., and Suk, H.-I. (2017).
Deep learning in medical image analysis.
Annual review of biomedical engineering, 19:221–248.

Wydział Matematyki i Nauk Informacyjnych
POLITECHNIKA WARSZAWSKA