

Abstract

Deploying AI models in a new environment requires significant costs related to the collection of labeled data for model re-training. Source-free domain adaptation (SFDA) addresses this issue by adapting the model to a new environment **using only unlabeled data from the target environment**. On the basis of the theory of self-training, we show **a unified theoretical understanding of the existing SFDA methods** that were individually developed. We also **build an improved SFDA method** based on two useful insights derived from this theoretical understanding and **achieve consistent performance improvements over conventional methods** under various conditions. SFDA is attracting significant attention due to its practical utility in that it enables flexible adaptation of the model to new environments while preserving the privacy and copyright of the source training data. Our theoretical understanding of SFDA will **improve the reliability of SFDA techniques and serve as a foundation for the creation of new SFDA methods**.

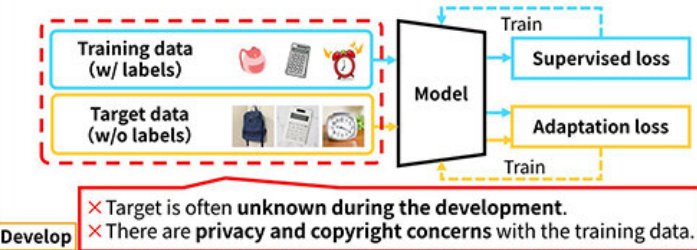
What's Domain Adaptation (DA)?

- Adapting a model to a target data environment that is different from the training data one.

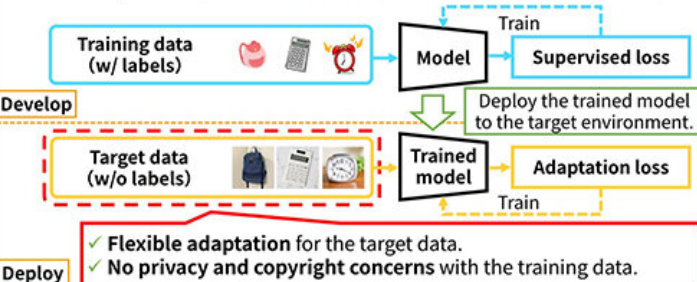


Source-free DA (SFDA)

- Conventional DA requires **both training and target data** for adaptation.



- SFDA requires **only unlabeled target data** for adaptation (**no labeled data** is needed for adaptation).



Q. What makes SFDA successful without labels?

Contribution 1: Theoretical Understanding

- Two key elements** derived from the theory of self-training

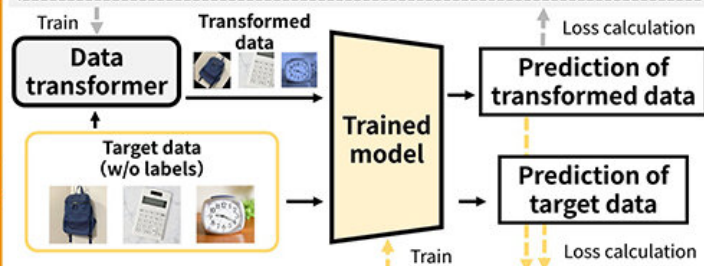


- Two insights** derived from the theoretical understanding

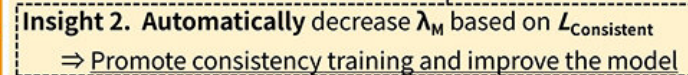
- Balance of Data Transformation**
Strength of data transformation used in consistency training **critically affects the model performance**.
Moderate (green checkmark), Too weak (red X), Too strong (red X).
- Balance Consistency and Diversity**
As consistency training progress, diversity training **becomes unnecessary**.
Continuing diversity training **may hinder the overall training effect**.

Contribution 2: Improved Method

Transformation training loss: $L_{\text{Boost}} + \lambda_A L_{\text{Reduce}}$



Model training loss: $L_{\text{Consistent}} + \lambda_M L_{\text{Diverse}}$



References

- [1] Y. Mitsuzumi, A. Kimura, H. Kashima, "Understanding and improving source-free domain adaptation from a theoretical perspective," in *Proc. The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 28515-28524, 2024.
- [2] https://www.youtube.com/watch?v=SnWqZ_lb93Y

Contact

Yu Mitsuzumi, Recognition Research Group, Media Information Laboratory