Unlabeled Data Help: Minimax Analysis and Adversarial Robustness

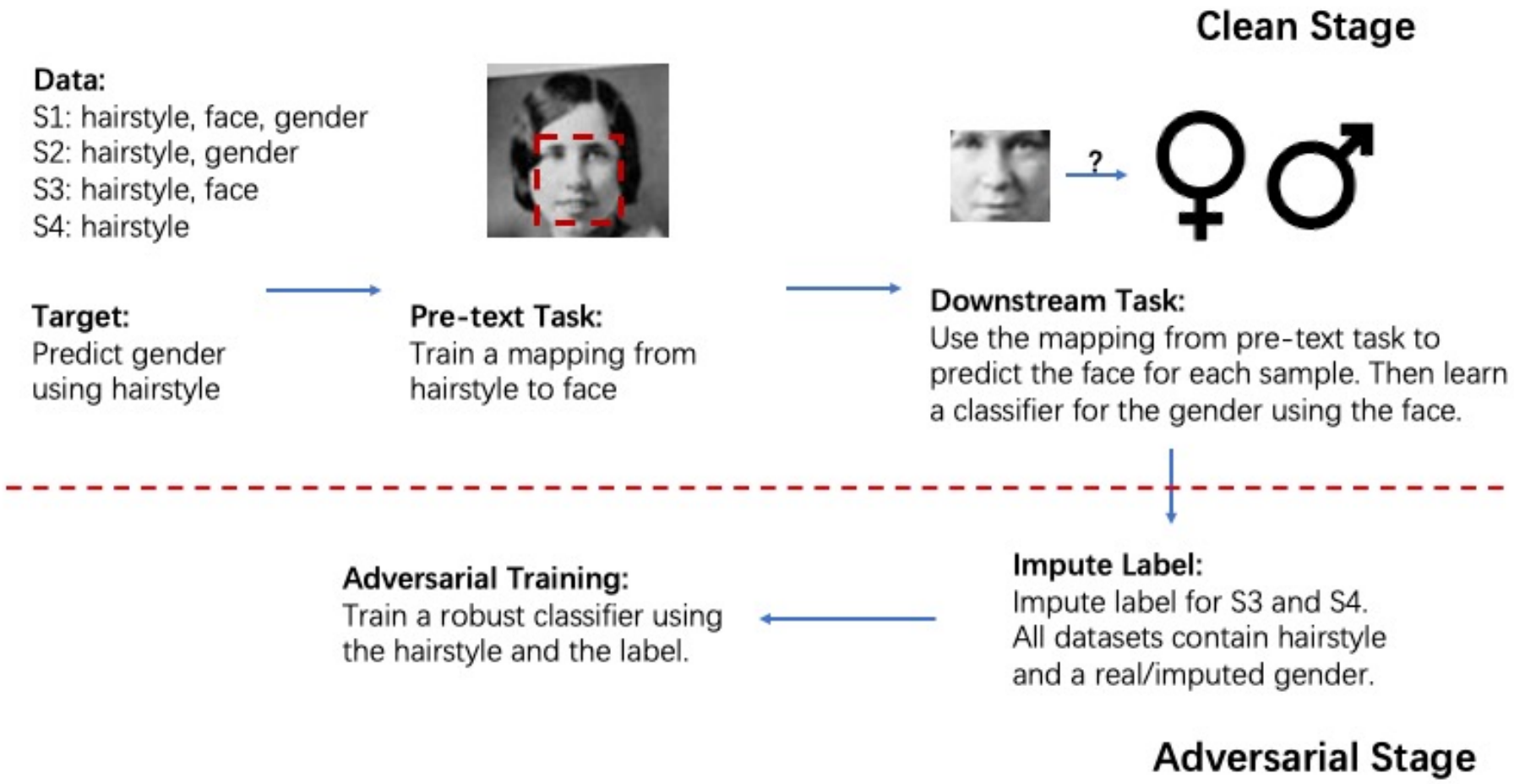
Yue Xing, Qifan Song, Guang Cheng

Purdue University, UCLA

Introduction

- We study
  - The lower bound and convergence upper bound of reconstruction-based self-supervised learning.
  - Its application into adversarial training.
- **Why SSL:** SSL improves generalization with the help of additional unlabeled data with extra attributes (Lee et. al., 2021).
  - No literature studying about its optimality.
- Whether the improvement of SSL can be inherited by adversarial training?
- **Minimax:** the best possible convergence rate that can be achieved by **any** estimator in the worst case given finite samples.
  - The existence of extra unlabeled data will change the minimax bound.

Training Procedure



Theoretical Result

- If X_1 (hairstyle) and X_2 (face) are conditionally independent (CI) with each other given the label (gender), then

$$\inf_{\widehat{f}} \sup_{\mathcal{P}} \mathbb{E}R(\widehat{f}, \epsilon) - R^*(\epsilon) = \Omega\left(\frac{d_1}{n_1 + n_2} \wedge \frac{d_1}{n_1 + n_3}\right)$$

- Otherwise,

$$\inf_{\widehat{f}} \sup_{\mathcal{P}} \mathbb{E}R(\widehat{f}, \epsilon) - R^*(\epsilon) = \Omega\left(\frac{d_1}{n_1 + n_2} \wedge \left(\frac{d_2}{n_1 + n_2} + \frac{d_1 + d_2}{n_1 + n_3}\right)\right)$$

- In addition, the upper bounds for clean and adversarial training attain these rates.
- **Why CI:** based on Lee et. al. (2021), CI is a key factor affecting SSL.
- Based on our results, SSL is optimal. CI changes the information limit, so both lower and upper bounds get changed.

References

- Lee, Jason D., et al. "Predicting what you already know helps: Provable self-supervised learning." *Advances in Neural Information Processing Systems* 34 (2021).

Thank you!