Is user feedback always informative?

Retrieval Latent Defending for Semi-Supervised Domain Adaptation without Source Data

Junha Song Taesoo Kim Junha Kim Gunhee Nam Thijs Kooi Jaegul Choo

Korea Advanced Institute of Science and Technology (KAIST) Lunit Inc.

Research Motive

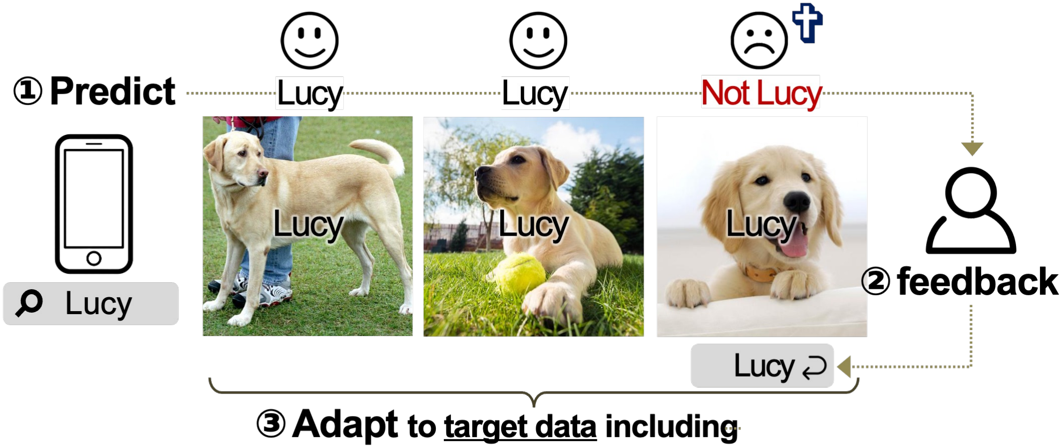
- 💁‍♂️ Semi-supervised domain adaptation (SemiSDA) adapt the source model to the target domain using both labeled and unlabeled target data.
- 🙋‍♀️ How do we collect labeled data in target devices like smartphones or medical applications?
- 🧏‍♂️ Users can provide small amounts of feedback!

Research Motive

- Semi-supervised domain adaptation (SemiSDA) adapt the source model to the target domain using both labeled and unlabeled target data.
- How do we collect labeled data in target devices like smartphones or medical applications?
- Users can provide small amounts of feedback!
- Actually, users might be more likely to provide feedback when the model makes incorrect predictions.† Did you consider this angle?

Research Motive

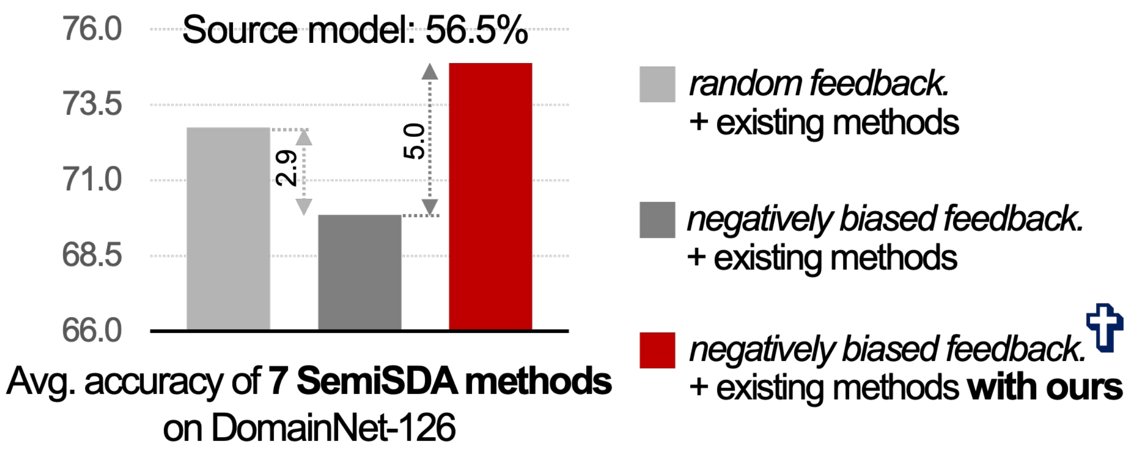
- Semi-supervised domain adaptation (SemiSDA) adapt the source model to the target domain using both labeled and unlabeled target data.
- How do we collect labeled data in target devices like smartphones or medical applications?
- Users can provide small amounts of feedback!
- Actually, users might be more likely to provide feedback when the model makes incorrect predictions.† Did you consider this angle?



Research Motive

- Semi-supervised domain adaptation (SemiSDA) adapt the source model to the target domain using both labeled and unlabeled target data.
- How do we collect labeled data in target devices like smartphones or medical applications?
- Users can provide small amounts of feedback!
- Actually, users might be more likely to provide feedback when the model makes incorrect predictions.† Did you consider this angle?
- Previous works hadn't considered that before. That’s an interesting observation! [doing some experiments...]
- I've been looking into existing SemiSDA methods, and I've noticed they often yield suboptimal results in the scenario.

Research Motive



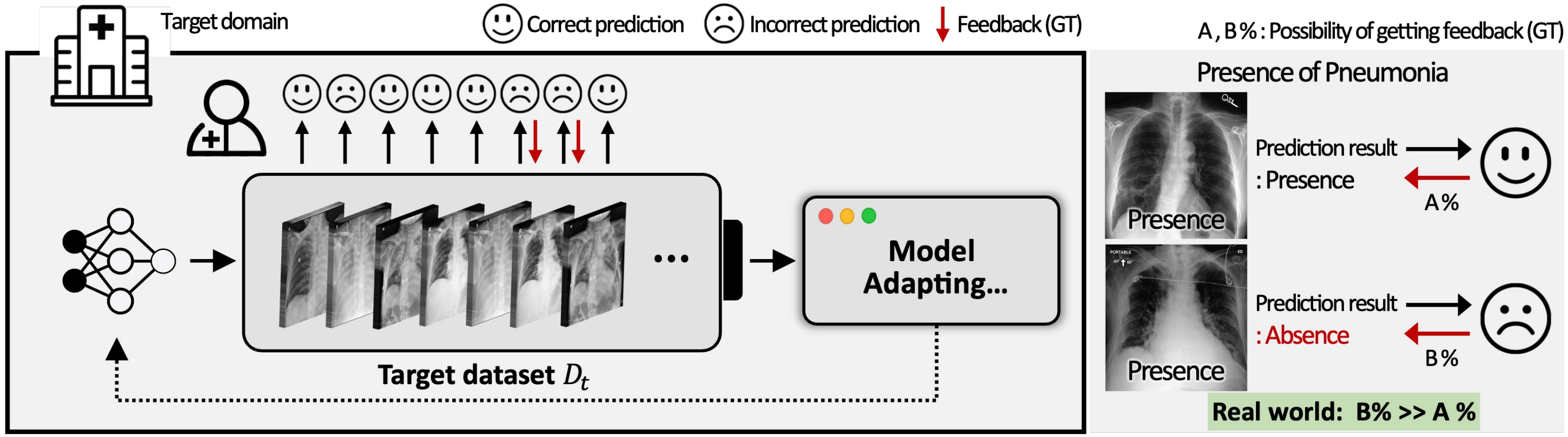
- 🤦 Previous works hadn't considered that before. That's an interesting observation! [doing some experiments...]
- 🤷 I've been looking into existing SemiSDA methods, and I've noticed they often yield suboptimal results in the scenario.

Rethinking user-provided feedback

- Users generally expect their feedback to be used as a basis of model improvement, motivating feedback on misclassified samples.
- For example, a radiologist might log a misdiagnosed chest X-ray by the model, as its accuracy directly impacts the patient's survival.

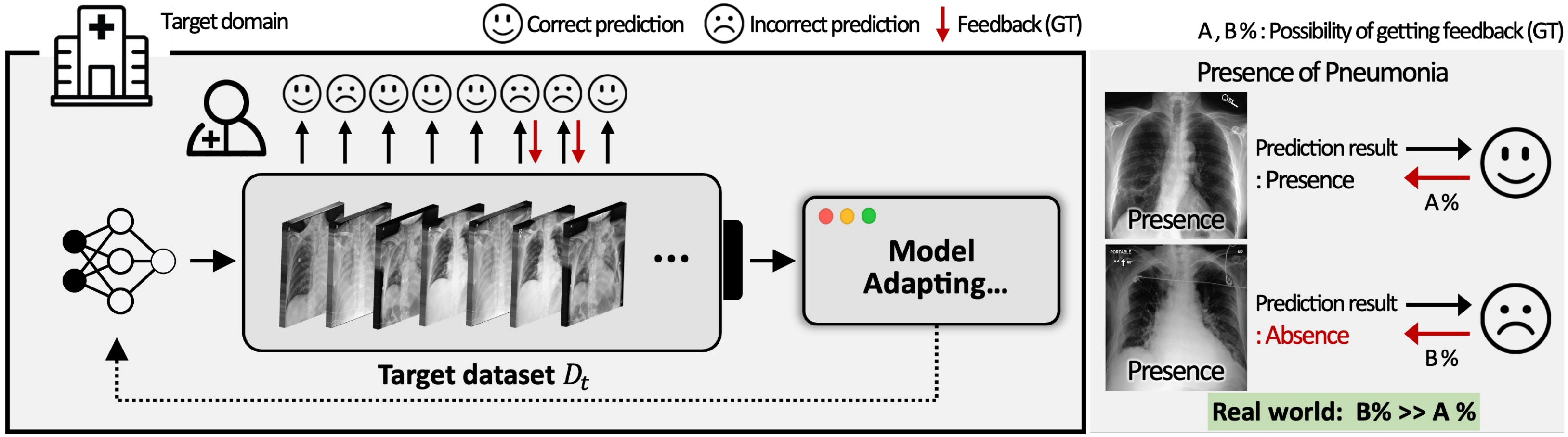
Rethinking user-provided feedback

- Users generally expect their feedback to be used as a basis of model improvement, motivating feedback on misclassified samples.
- For example, a radiologist might log a misdiagnosed chest X-ray by the model, as its accuracy directly impacts the patient's survival.



Rethinking user-provided feedback

- Users generally expect their feedback to be used as a basis of model improvement, motivating feedback on misclassified samples.
- For example, a radiologist might log a misdiagnosed chest X-ray by the model, as its accuracy directly impacts the patient's survival.



- We introduce this novel view called Negatively Biased Feedback (NBF).

Influence of NBF on SemiSDA

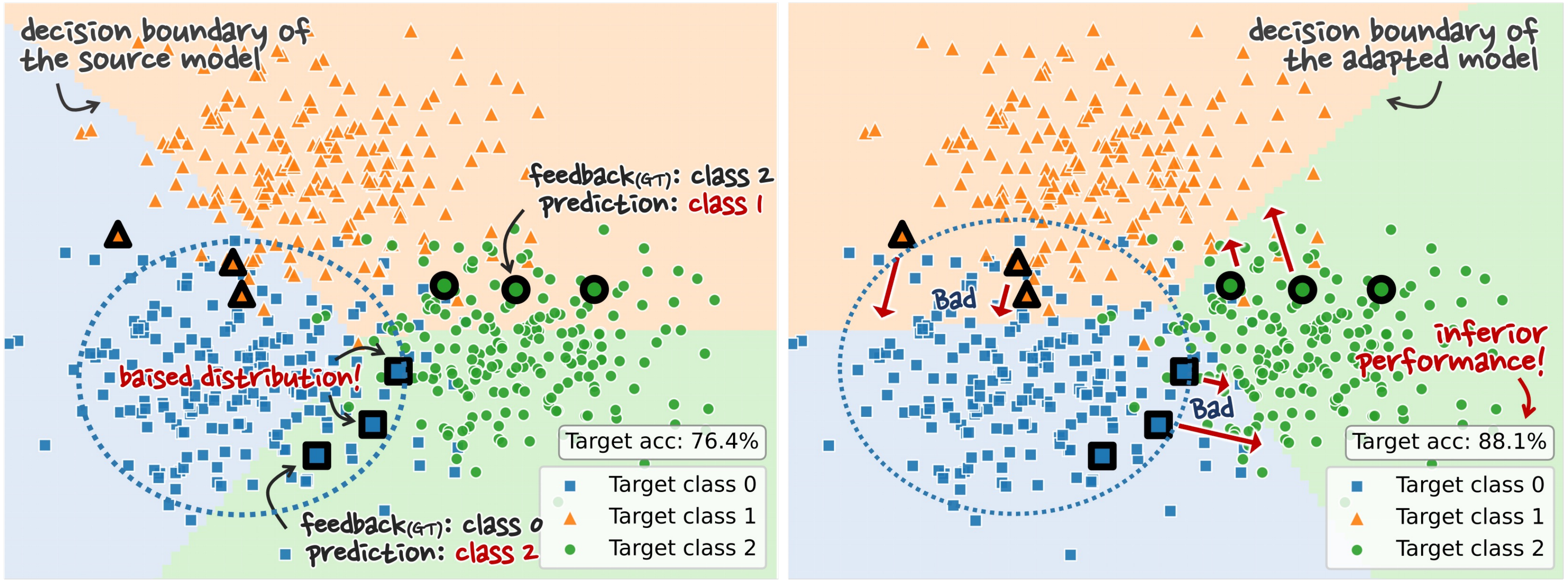
- We conduct extensive studies, revealing interesting insights:

Influence of NBF on SemiSDA

- We conduct extensive studies, revealing interesting insights:
    a. User-provided feedback in practice (NBF) has a biased distribution.
    b. NBF causes the source model to be adapted with suboptimal results.

Influence of NBF on SemiSDA

- We conduct extensive studies, revealing interesting insights:
  a. User-provided feedback in practice (NBF) has a biased distribution.
  b. NBF causes the source model to be adapted with suboptimal results.



Counterintuitive Effect of NBF

- Our intuitive reasoning probably suggests that NBF provides more information than RF by correcting more source model deficiencies, and thus leads to better adaptation performance.
- Our work highlights the importance of careful design when using user feedback in real-world scenarios.

Prerequisite: Previous SemiSDA methods

- Their model adaptation combines cross-entropy loss for labeled data with consistency regularization on multi-view unlabeled data.

$$\mathcal{L}_{sup} = \frac{1}{B}\sum_{b=1}^{B}\mathcal{H}(y_{lb}^{b}, f_{\theta}(x_{lb}^{b})), \quad \mathcal{L}_{unsup} = \frac{1}{\mu \cdot B}\sum_{b=1}^{\mu \cdot B}\mathcal{H}(\hat{y}_{ulb}^{b}, f_{\theta}(\Omega(x_{ulb}^{b})))$$

Prerequisite: Previous SemiSDA methods

- Their model adaptation combines cross-entropy loss for labeled data with consistency regularization on multi-view unlabeled data.

$$\mathcal{L}_{sup} = \frac{1}{B}\sum_{b=1}^{B}\mathcal{H}(y_{lb}^{b}, f_\theta(x_{lb}^{b})), \quad \mathcal{L}_{unsup} = \frac{1}{\mu \cdot B}\sum_{b=1}^{\mu \cdot B}\mathcal{H}(\hat{y}_{ulb}^{b}, f_\theta(\Omega(x_{ulb}^{b})))$$

- Existing methods, overlooking the realistic setup of NBF, suffer from inadequate adaptation performance.

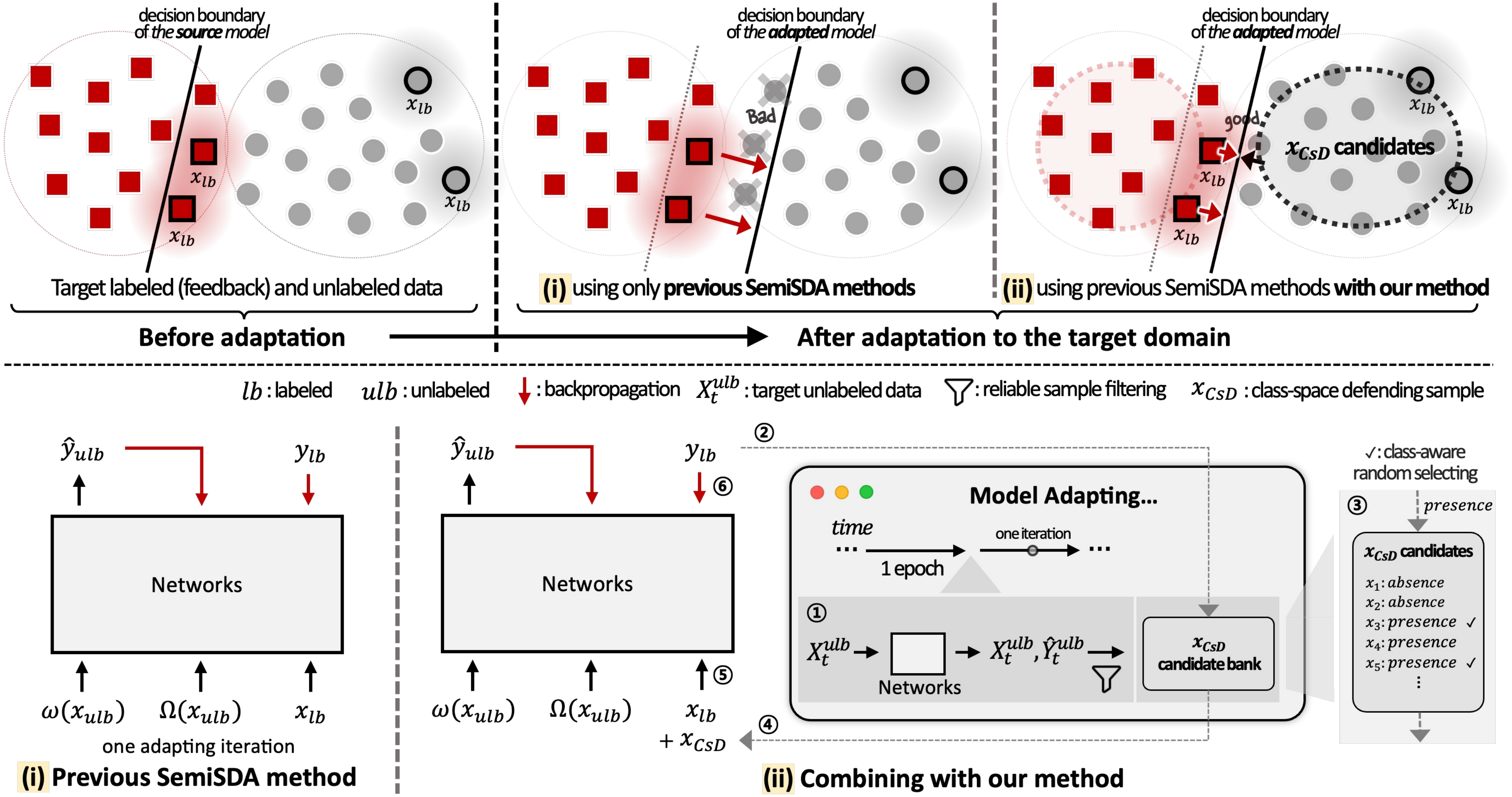
Prerequisite: Previous SemiSDA methods

- Their model adaptation combines cross-entropy loss for labeled data with consistency regularization on multi-view unlabeled data.

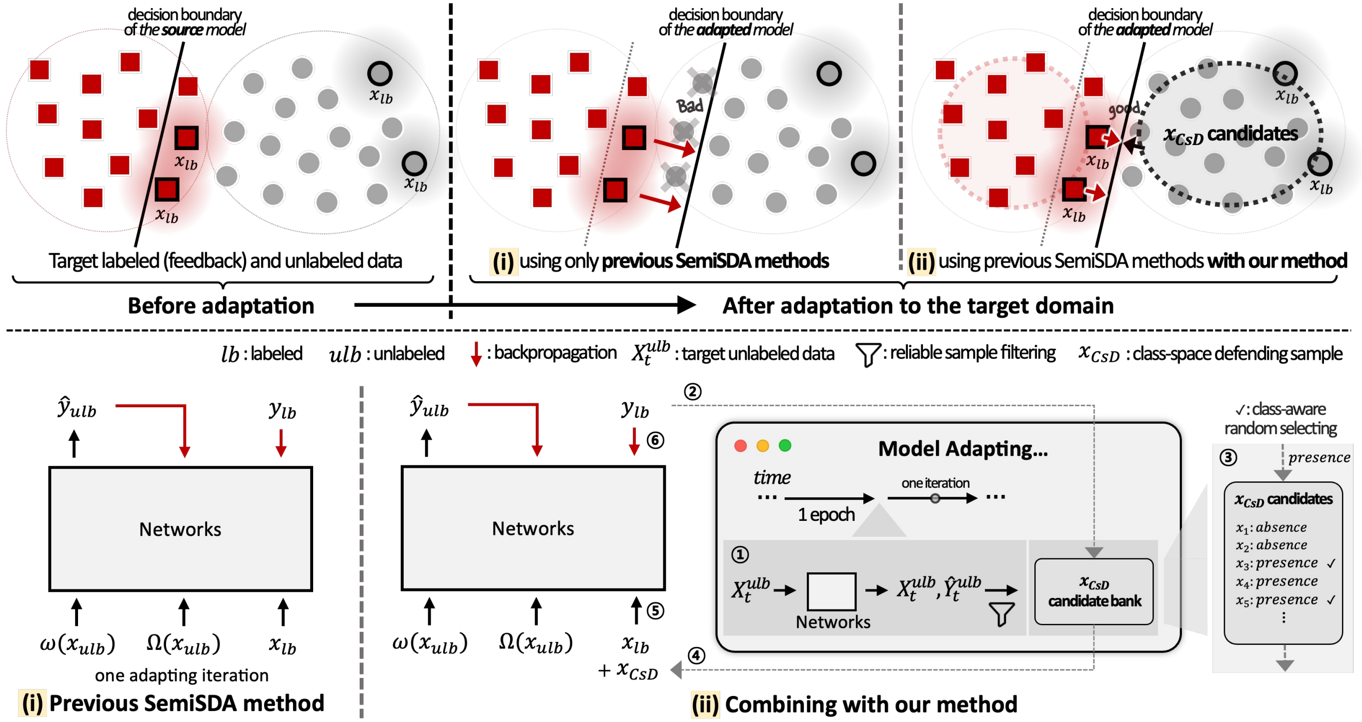
$$\mathcal{L}_{sup} = \frac{1}{B}\sum_{b=1}^{B}\mathcal{H}(y_{lb}^{b}, f_{\theta}(x_{lb}^{b})), \quad \mathcal{L}_{unsup} = \frac{1}{\mu \cdot B}\sum_{b=1}^{\mu \cdot B}\mathcal{H}(\hat{y}_{ulb}^{b}, f_{\theta}(\Omega(x_{ulb}^{b})))$$

- Existing methods, overlooking the realistic setup of NBF, suffer from inadequate adaptation performance.
- we focus on developing a solution that (i) can easily combine with existing DA methods without modifying their core strategies and (ii) can be applied to a wide range of benchmarks.

Retrieval Latent Defending



Retrieval Latent Defending



Our RLD involves the following steps:

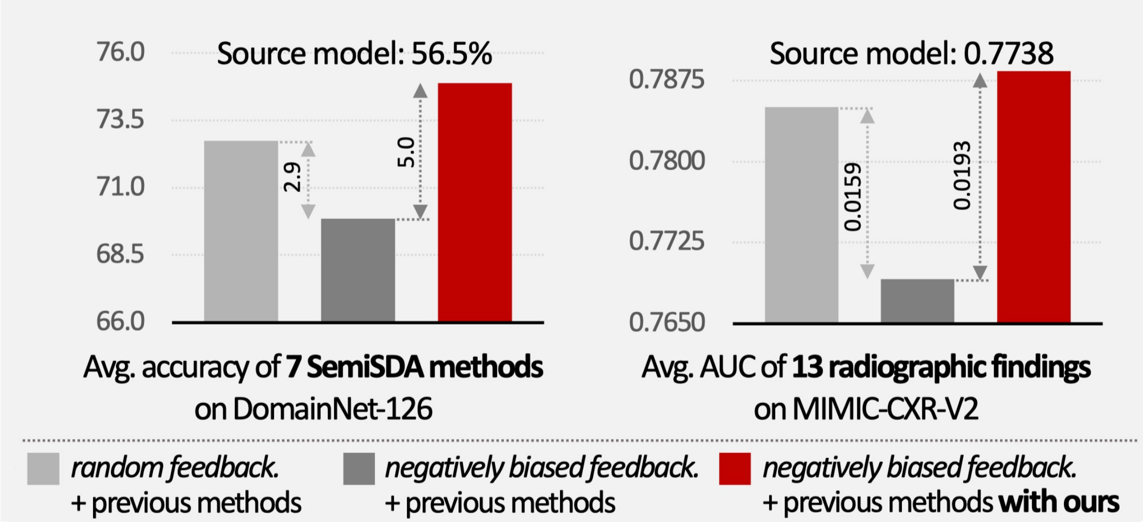
① Prior to each epoch, we generate a candidate bank of data points.

②~④ For each adapting iteration, we balance the mini-batch by retrieving latent defending samples from the bank.

⑤~⑥ The model is then adapted using the reconfigured mini-batch and following the baseline SemiSDA approach.

Experimental Results

- We evaluate NBF's unexpected influence using various benchmarks, including Image classification 🐕🐈, Medical image diagnosis 🩻🏥, and Semantic segmentation 🚗🚋.



	dataset	classical setup	NBF setup	+ours
🐕🐈	DomainNet	67.6	64.5 (-3.1)	72.0 (+7.5)
🩻🏥	MIMIC-CXR	.785	.769 (-.016)	.785 (+.019)
🚗🚋	GTA5, Cityscapes	55.3	53.0 (-2.3)	56.3 (+3.3)