Part III: Data-driven priors

Weakly Supervised CNN Segmentation: Models and Optimization

Ismail Ben Ayed
Christian Desrosiers
Jose Dolz
Hoel Kervadec

Data-driven priors: Advanced

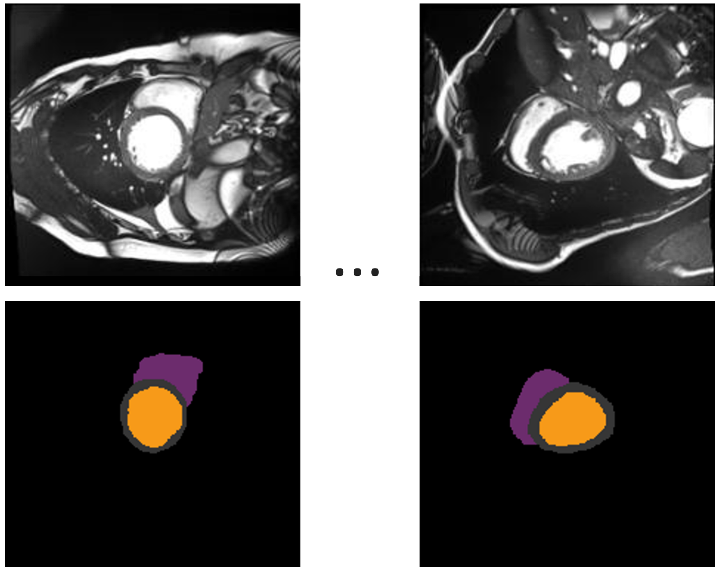
1) Adversarial learning
2) Consistency regularization
3) Unsupervised representation learning
4) Self-paced learning

Adversarial learning

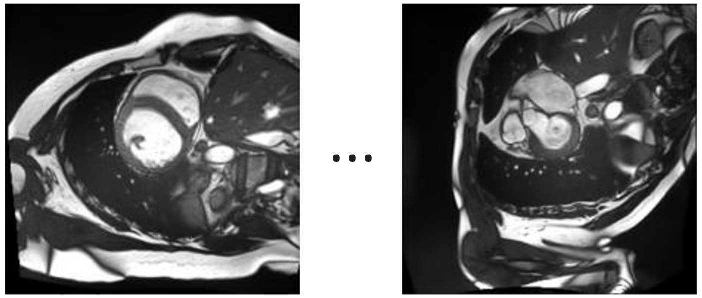
for weakly-supervised segmentation

Learning with unlabeled images

Labeled images (few)

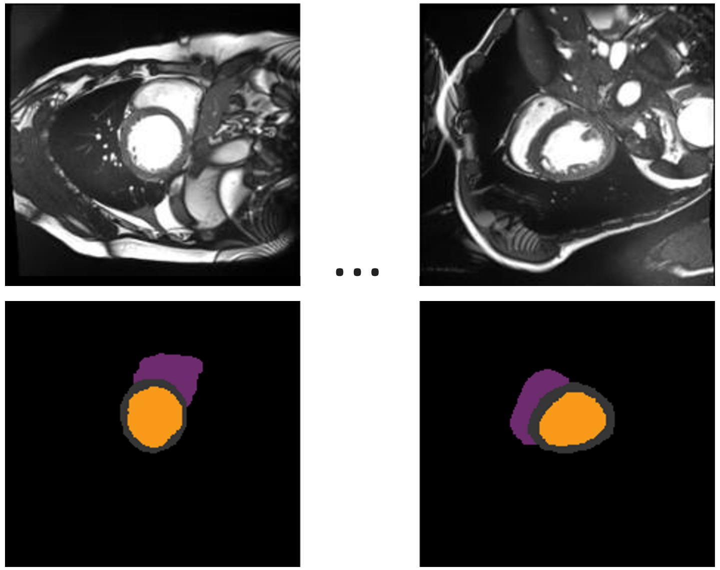


Unlabeled images (many)

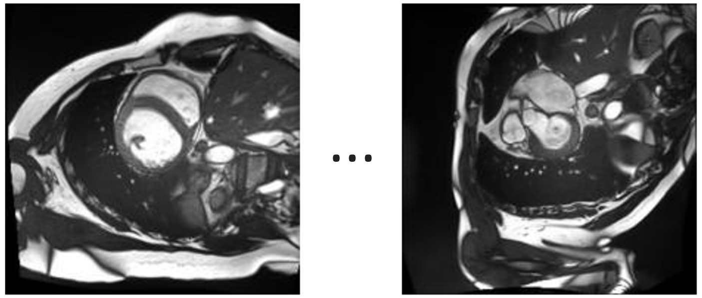


Learning with unlabeled images

Labeled images (few)



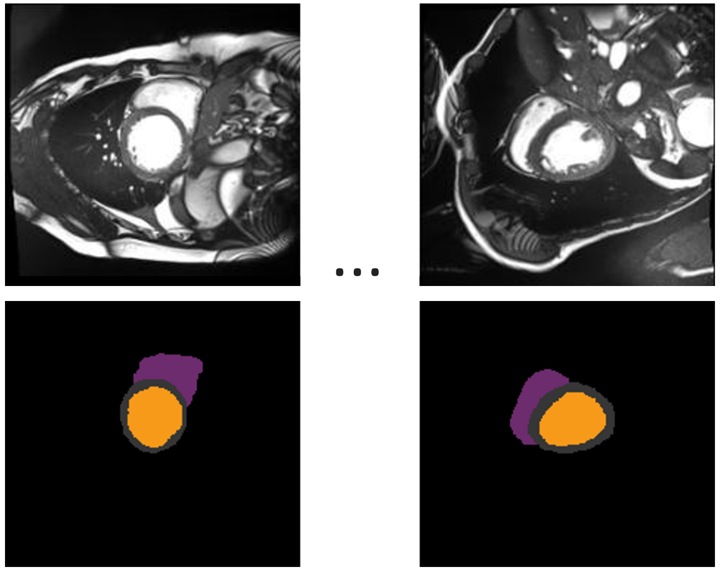
Unlabeled images (many)



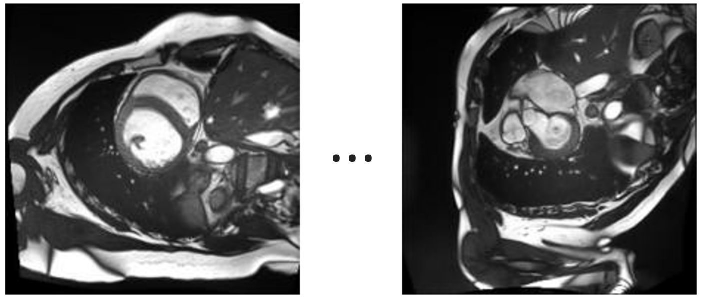
How can we use this information
to learn segmentation ?

Learning with unlabeled images

Labeled images (few)



Unlabeled images (many)



How can we use this information to learn segmentation ?

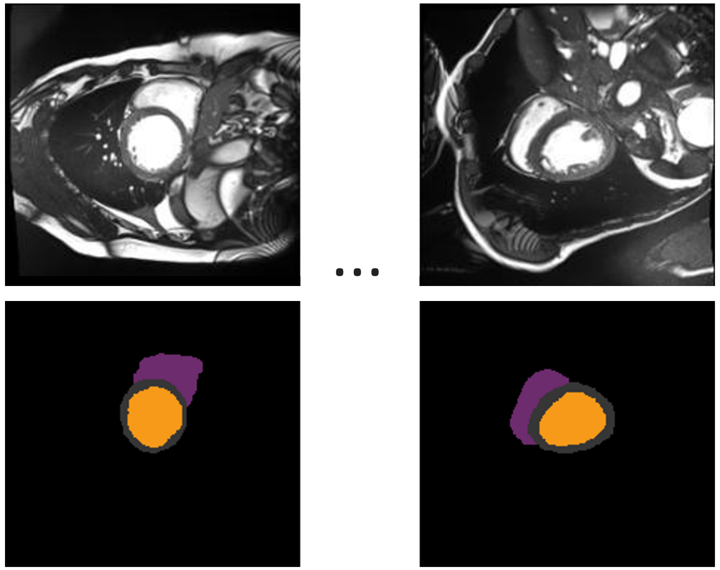
1) **Knowledge-based priors:**

- Size bounds, shape atlas, boundary smoothness, etc.
- Difficult to adapt to new domains or tasks

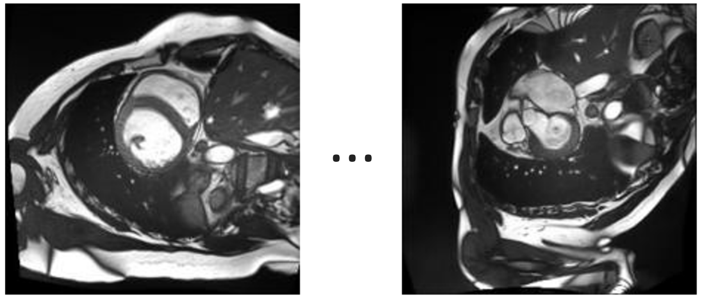
Presented in previous slides

Learning with unlabeled images

Labeled images (few)



Unlabeled images (many)



How can we use this information to learn segmentation ?

1) Knowledge-based priors:

- Size bounds, shape atlas, boundary smoothness, etc.
- Difficult to adapt to new domains or tasks

Presented in previous slides

2) Adversarial learning:

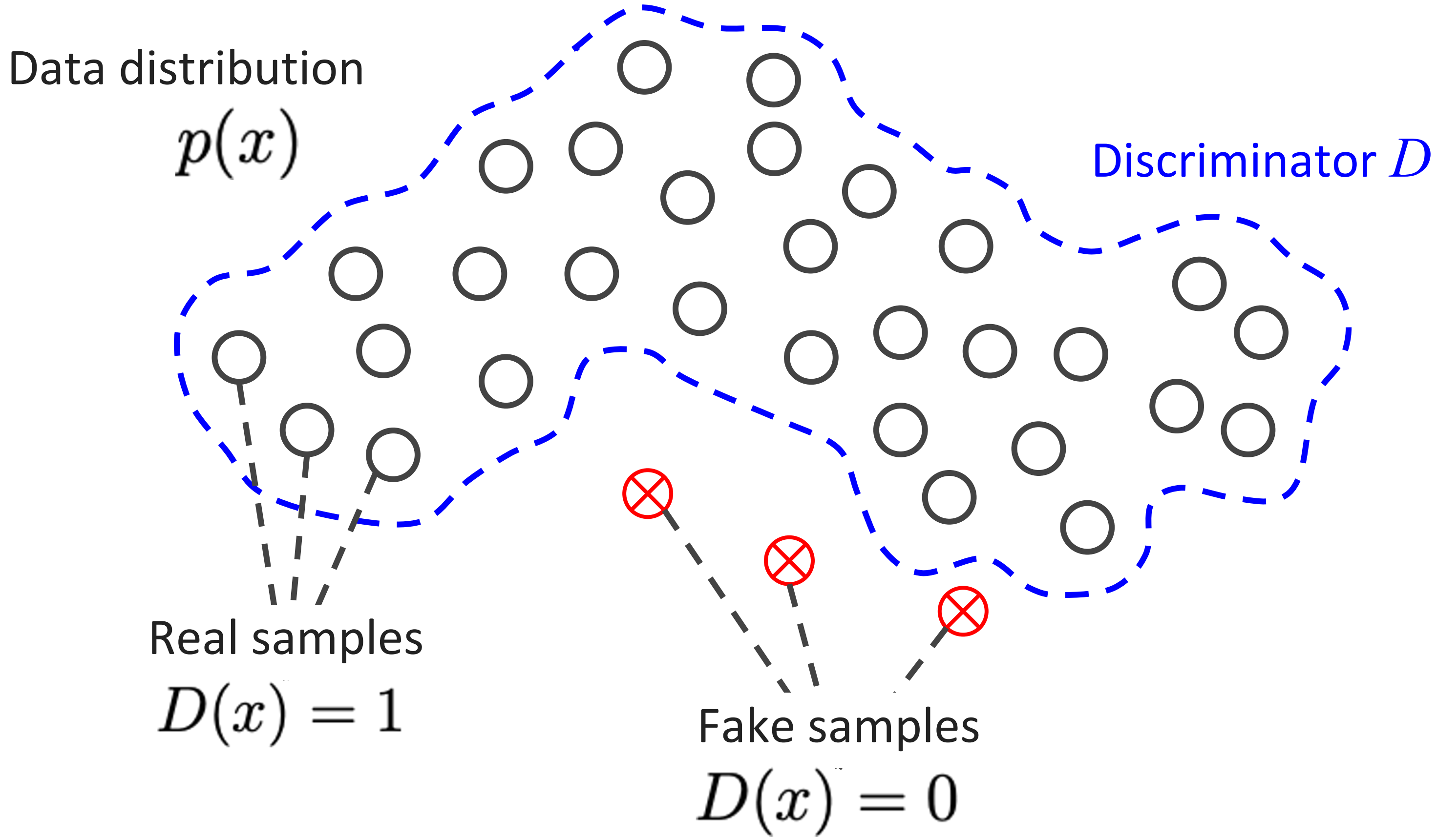
- Learn priors directly from <u>training data</u>
- Easily adapts to new domains or tasks

Discussed in next slides

Adversarial learning

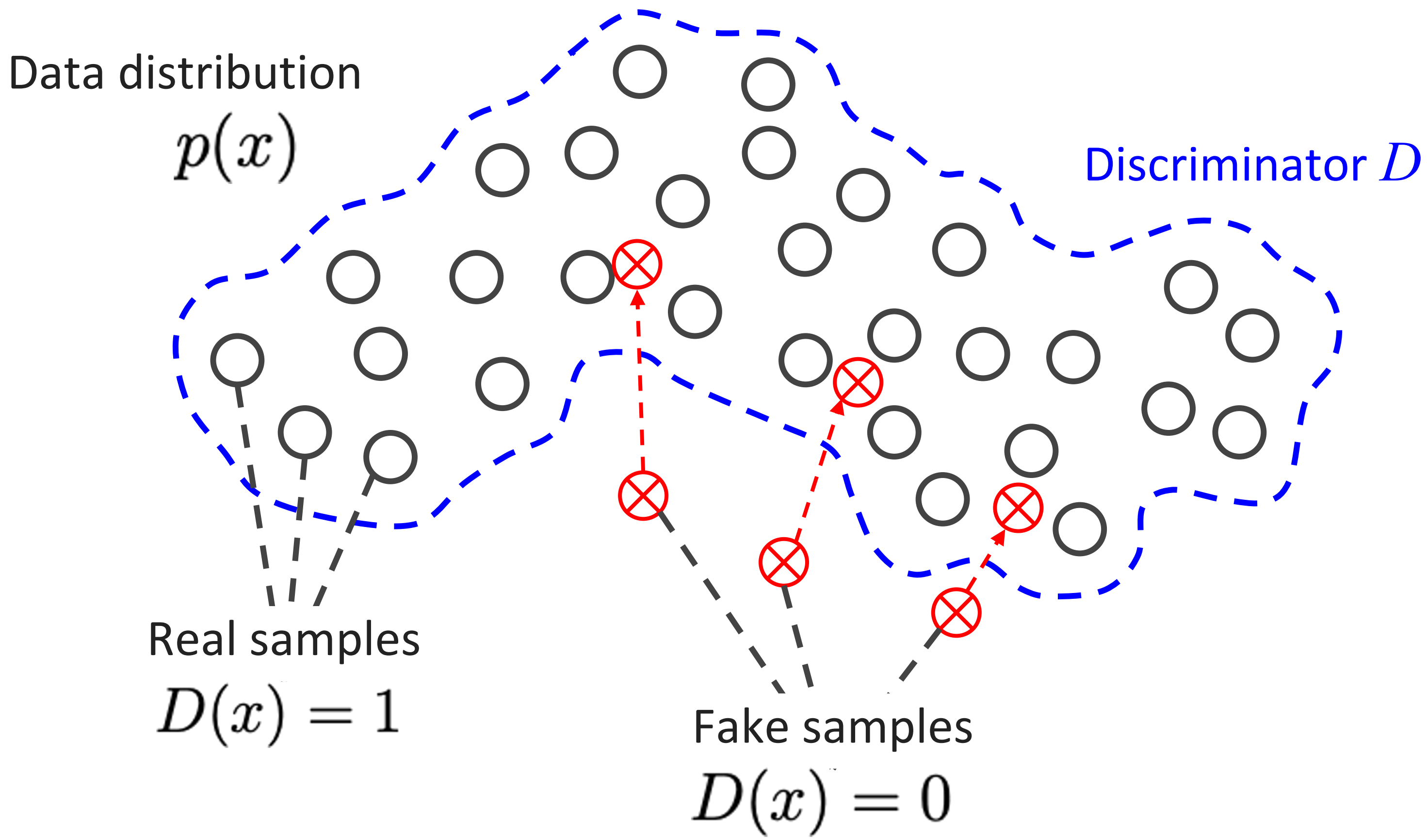
Basic idea:

Learn the data distribution using a classifier (the discriminator)



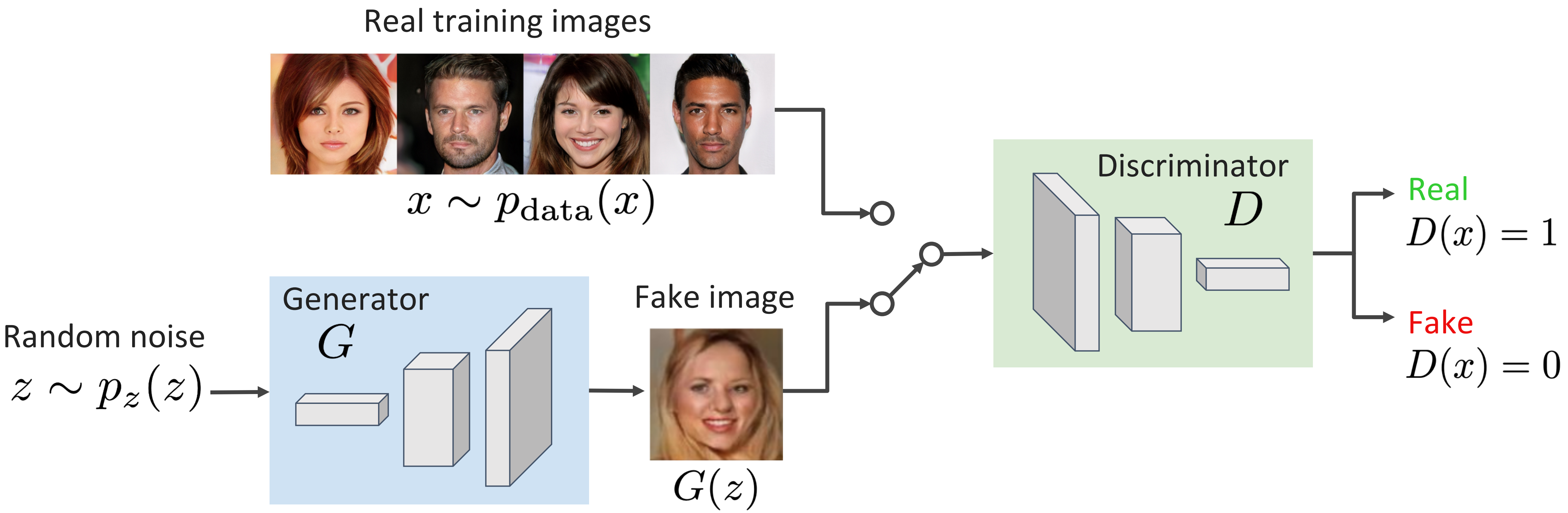
Adversarial learning

Basic idea:
Learn the data distribution using a classifier (the discriminator)

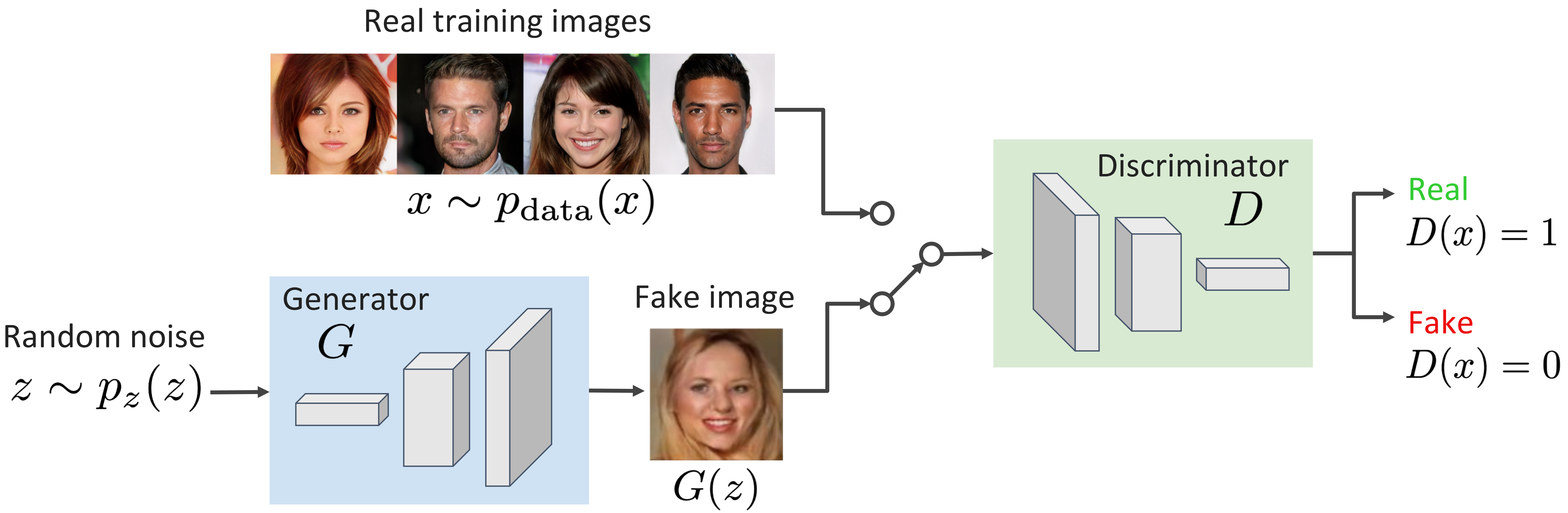


Objective: Generate samples in the distribution of real data

Generative adversarial network (GAN)

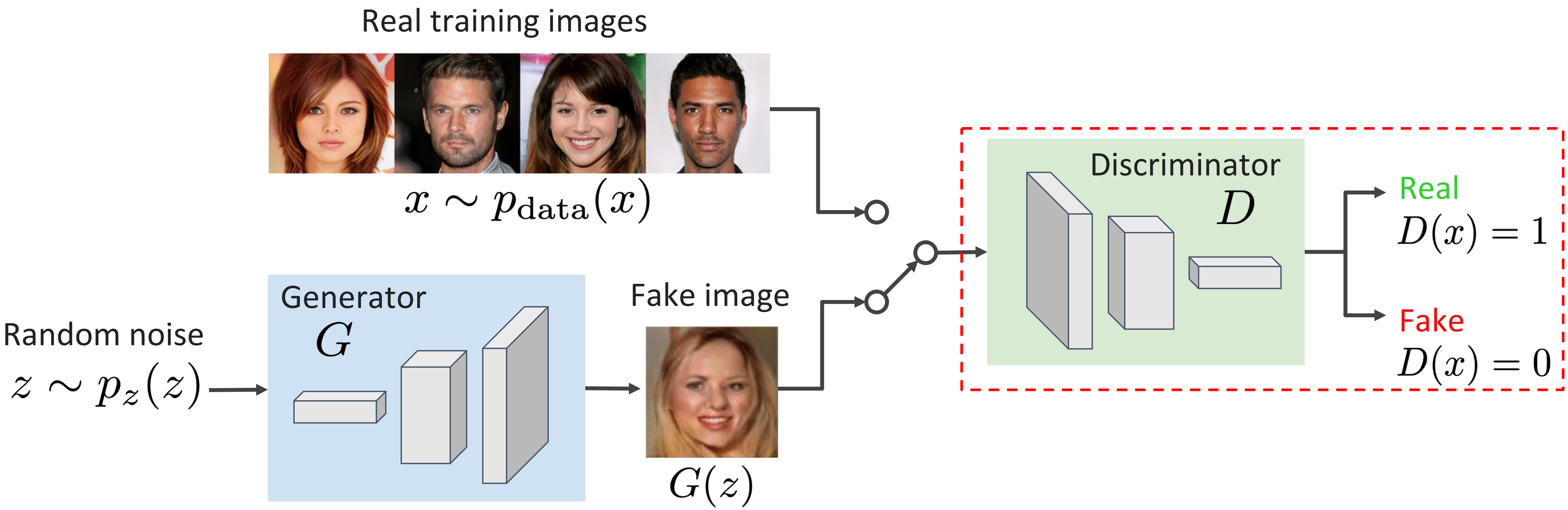


Generative adversarial network (GAN)



How to make sure that generated images look real ?

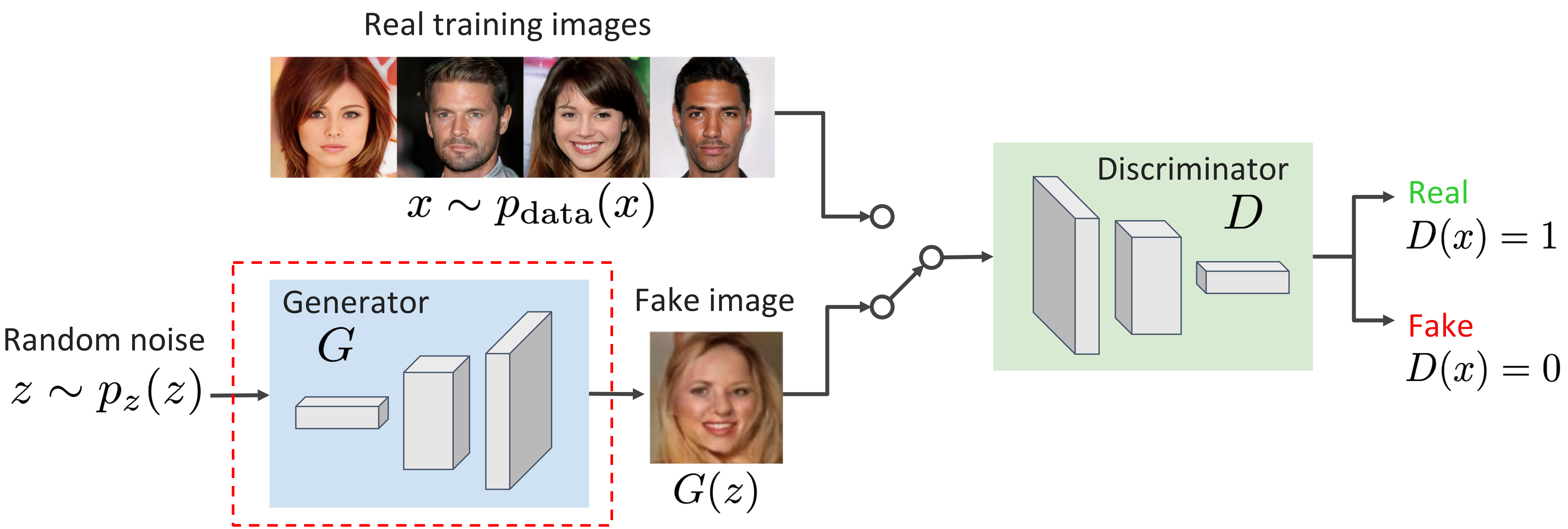
Generative adversarial network (GAN)



Training the discriminator (cross-entropy):

$$\max_D \; \underbrace{\mathbb{E}_{x \sim p_{\text{data}}(x)} \left[\log D(x) \right]}_{\text{Output '1' for real images}} + \underbrace{\mathbb{E}_{z \sim p_z(z)} \left[\log(1 - D(G(z))) \right]}_{\text{Output '0' for generated images}}$$

Generative adversarial network (GAN)

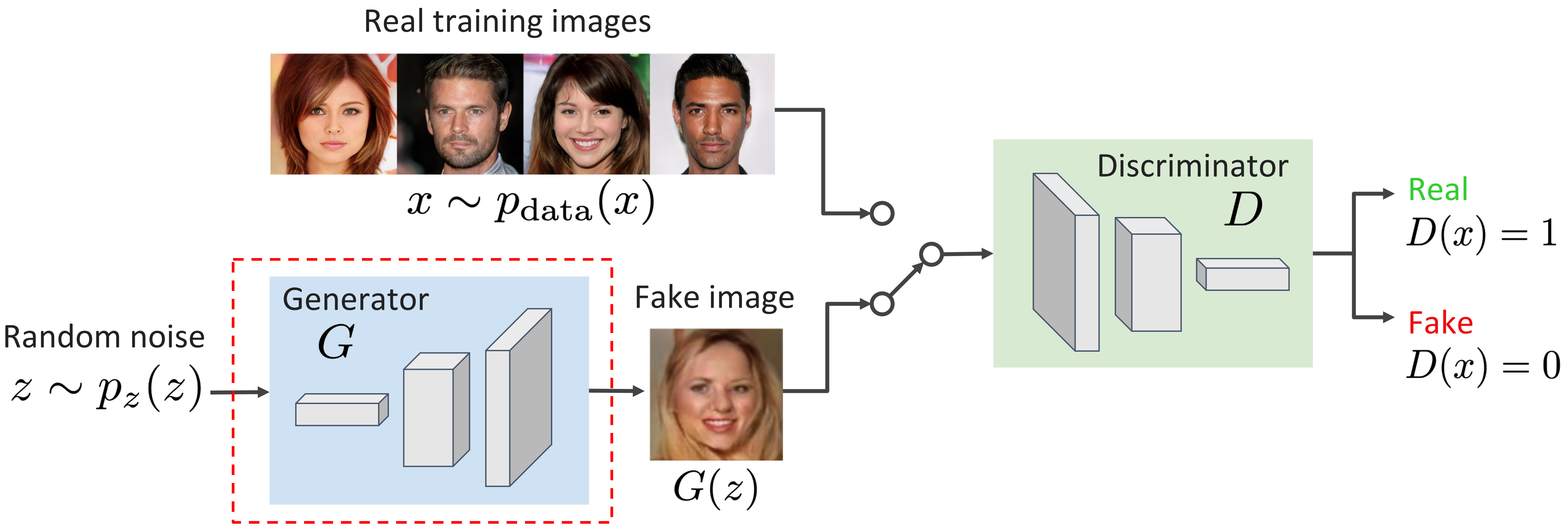


Training the generator:

$$\min_G \; \mathbb{E}_{z \sim p_z(z)} \left[\log(1 - D(G(z))) \right]$$

Fool the discriminator into predicting '1' for fake images

Generative adversarial network (GAN)



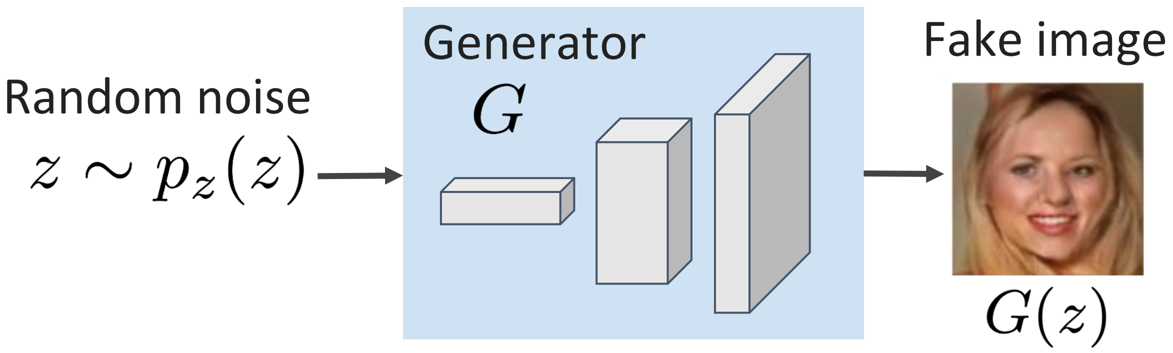
Training the whole architecture:

$$\min_G \max_D \; \mathbb{E}_{x \sim p_{\text{data}}(x)}\left[\log D(x)\right] + \mathbb{E}_{z \sim p_z(z)}\left[\log(1 - D(G(z)))\right]$$

Corresponds to a minimax problem (*more on this later...*)

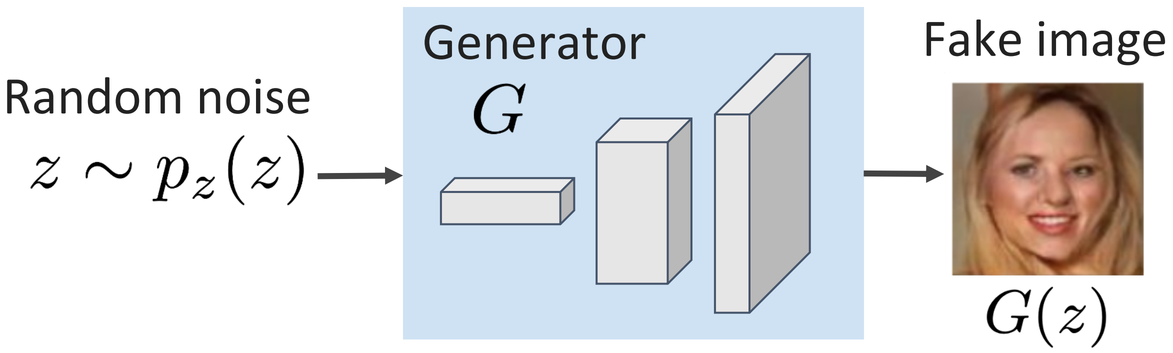
GANs for segmentation

GAN for image generation:

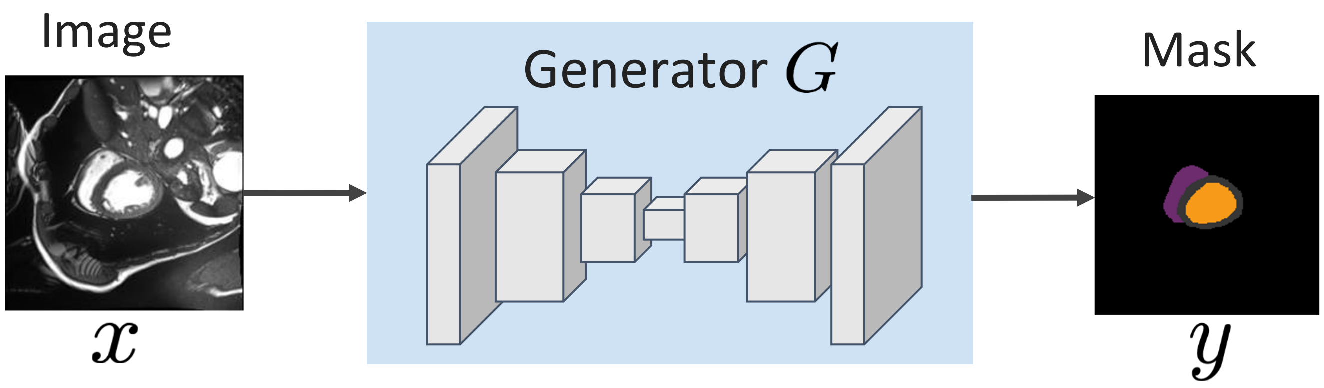


GANs for segmentation

GAN for image generation:

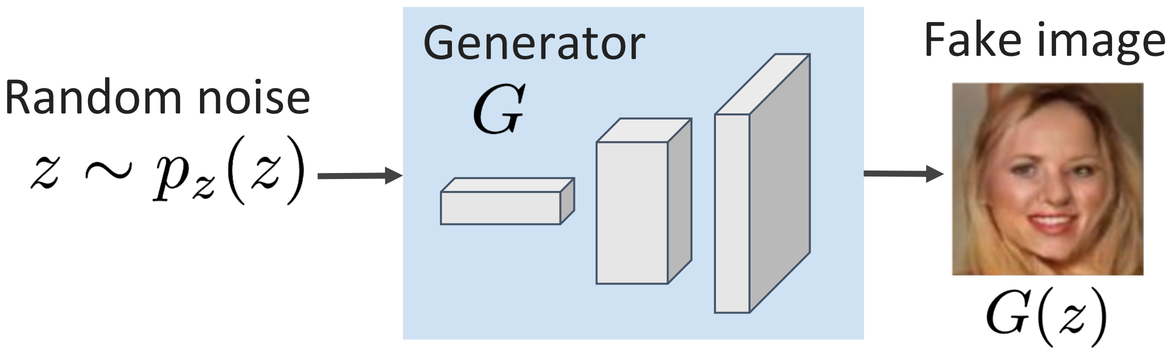


GAN for image segmentation:

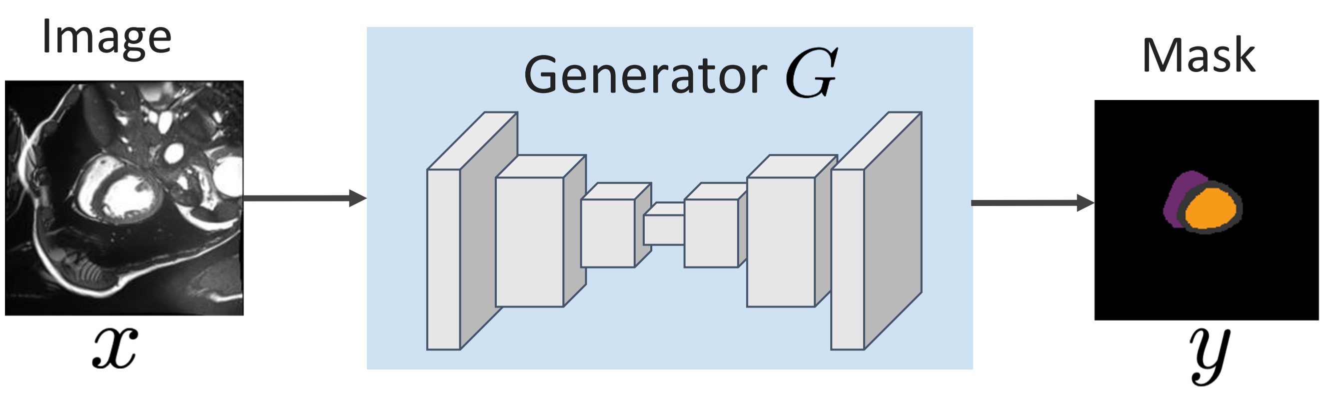


GANs for segmentation

GAN for image generation:



GAN for image segmentation:



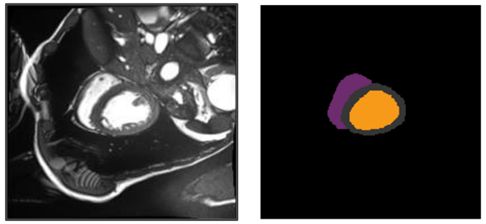
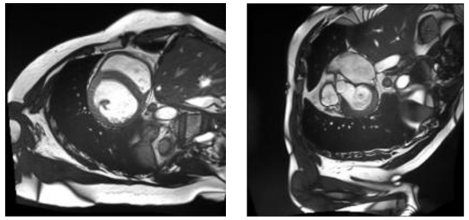
We are now modeling the distribution of *segmentation masks*

The generator is a segmentation network (encoder-decoder)

Adversarial semi-supervised segmentation

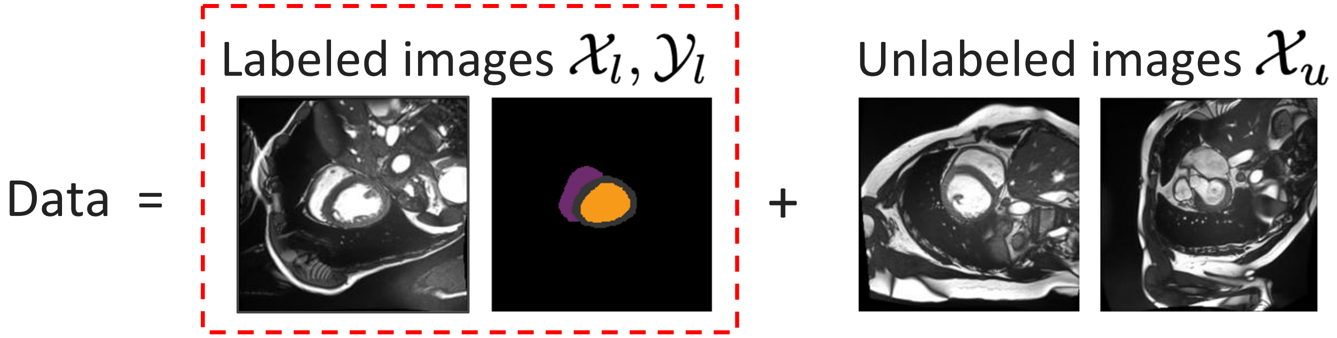
Basic idea: Learn to generate segmentation masks which can't be differentiated from ground-truth (GT)

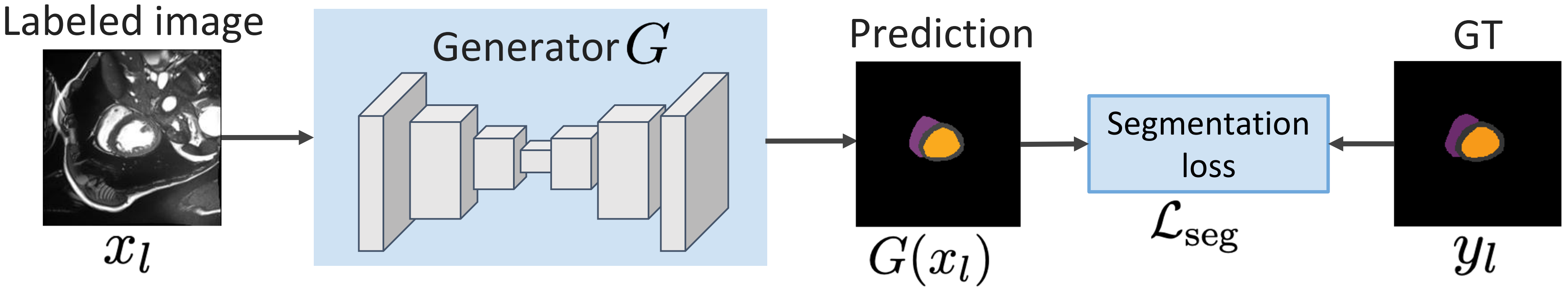
Labeled images $\mathcal{X}_l, \mathcal{Y}_l$ Unlabeled images $\mathcal{X}_u$

Data =  + 

Adversarial semi-supervised segmentation

Basic idea: Learn to generate segmentation masks which can't be differentiated from ground-truth (GT)

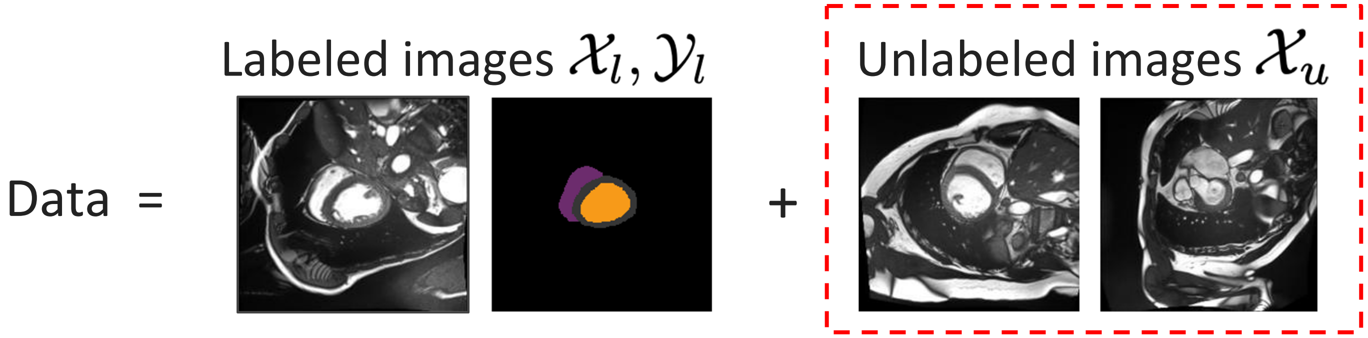


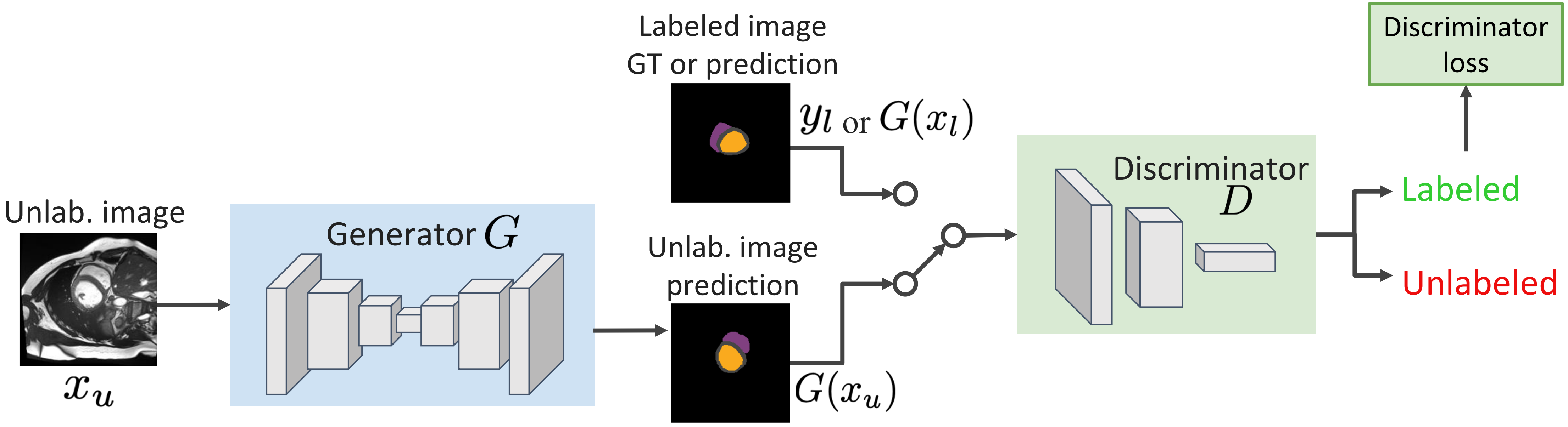


$$\mathcal{L}_{\text{sup}}(G) \;=\; \mathbb{E}_{(x_l,y_l)\sim\mathcal{X}_l,\mathcal{Y}_l}\left[\mathcal{L}_{\text{seg}}(G(x_l), y_l)\right]$$

Adversarial semi-supervised segmentation

Basic idea: Learn to generate segmentation masks which can't be differentiated from ground-truth (GT)



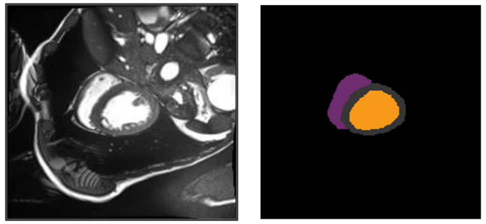
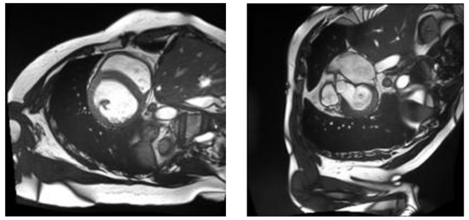


$$\mathcal{L}_{\text{adv}}(G, D) \;=\; \mathbb{E}_{x_u \sim \mathcal{X}_u}\big[\mathcal{L}_{\text{dis}}\big(D(G(x_u)), 0\big)\big] \;+\; \mathbb{E}_{x_l \sim \mathcal{X}_l}\big[\mathcal{L}_{\text{dis}}\big(D(G(x_l)), 1\big)\big]$$

Adversarial semi-supervised segmentation

Basic idea: Learn to generate segmentation masks which can't be differentiated from ground-truth (GT)

Labeled images $\mathcal{X}_l, \mathcal{Y}_l$ Unlabeled images $\mathcal{X}_u$

Data =  + 

Both labeled and unlabeled:

Controls the trade-off between the two losses

$$\min_G \max_D \ \mathcal{L}(G,D) = \underbrace{\frac{1}{|\mathcal{X}_l|}\sum_{l=1}^{|\mathcal{X}_l|}\mathcal{L}_{\text{seg}}(G(x_l),y_l)}_{\text{Supervised loss}} - \underbrace{\frac{\lambda}{|\mathcal{X}_l|+|\mathcal{X}_u|}\left(\sum_{l=1}^{|\mathcal{X}_l|}\mathcal{L}_{\text{dis}}\big(D(G(x_l)),1\big) + \sum_{u=1}^{|\mathcal{X}_u|}\mathcal{L}_{\text{dis}}\big(D(G(x_u)),0\big)\right)}_{\text{Adversarial loss}}$$

Adversarial semi-supervised segmentation

Adversarial network for semi-supervised segmentation of histological images

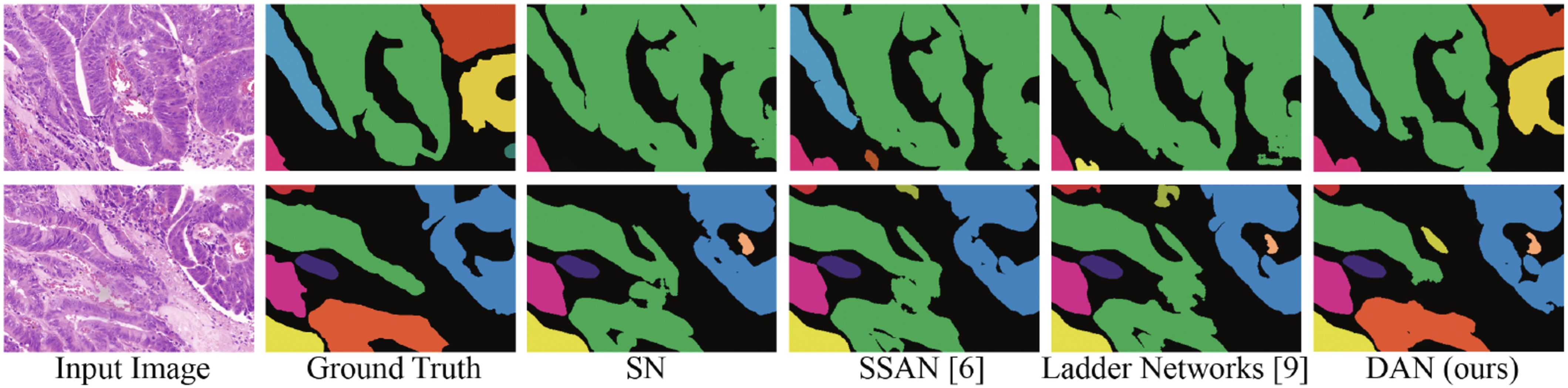
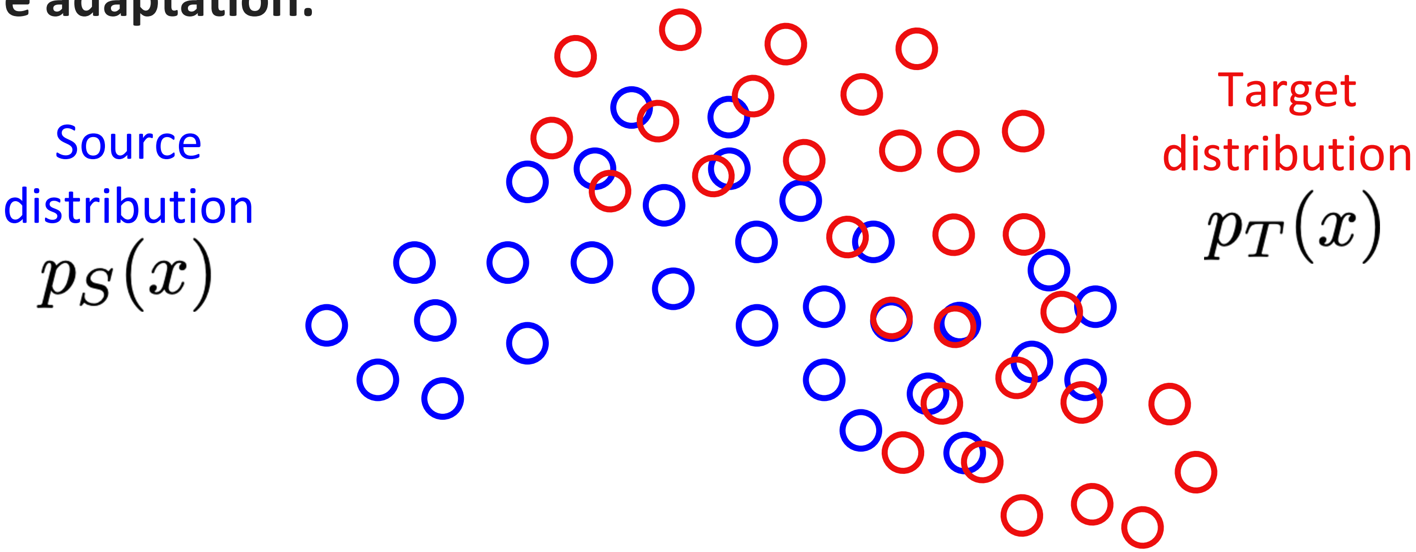


Image from Zhang, Y., et al. "Deep adversarial networks for biomedical image segmentation utilizing unannotated images." *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*. 2017.

Domain adaptation

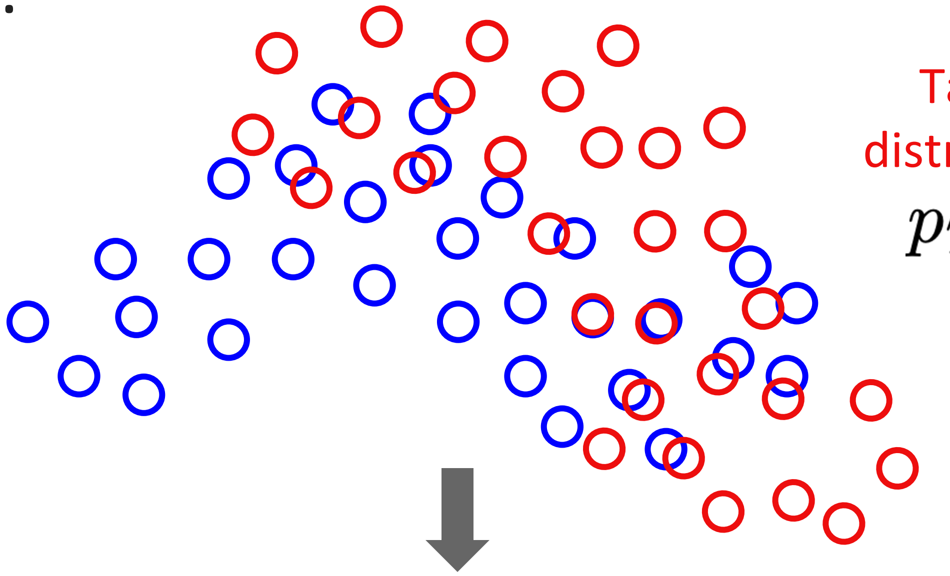
Before adaptation:



Domain adaptation

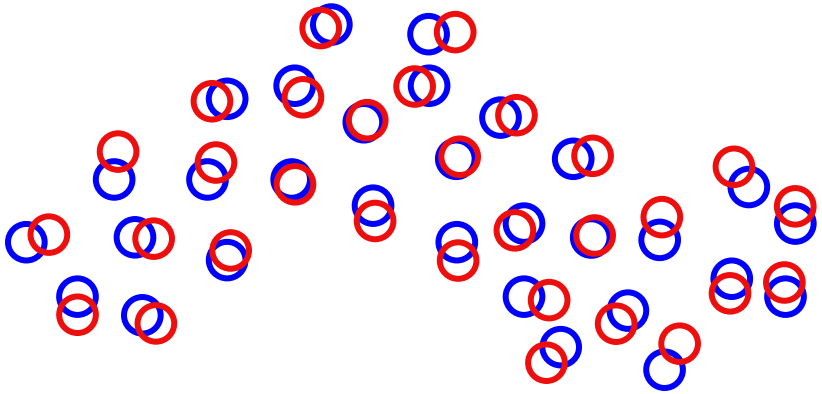
Before adaptation:

Source distribution
$p_S(x)$



Target distribution
$p_T(x)$

After adaptation:



$$p_S(x) = p_T(x)$$

Objective

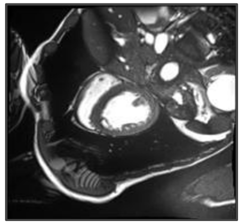
Align the distributions (input, output or representation) so that a model trained on Source data also works on Target data

Adversarial domain adaptation

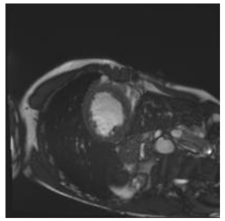
Labeled images $\mathcal{X}_S, \mathcal{Y}_S$

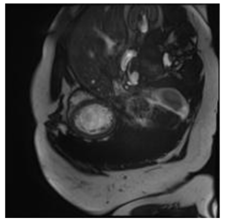
Unlabeled images $\mathcal{X}_T$

Data =



\+

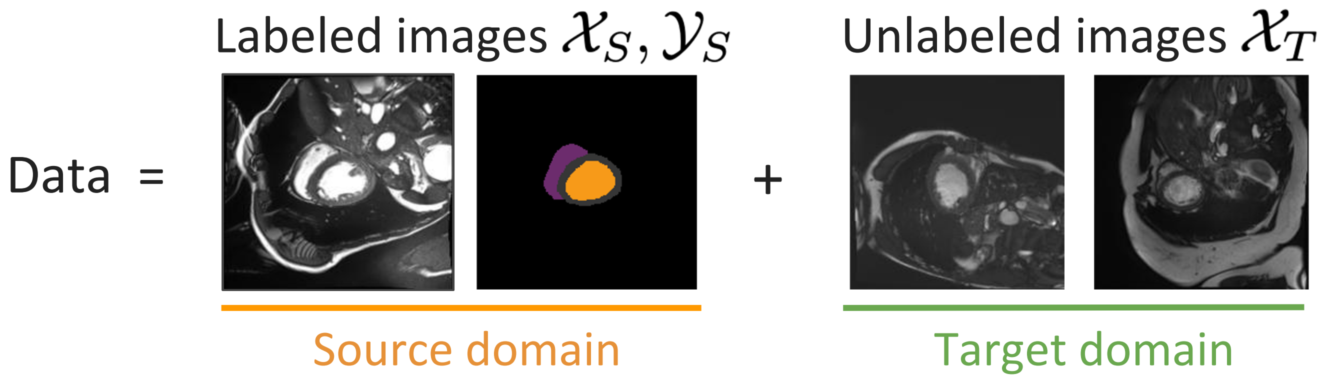


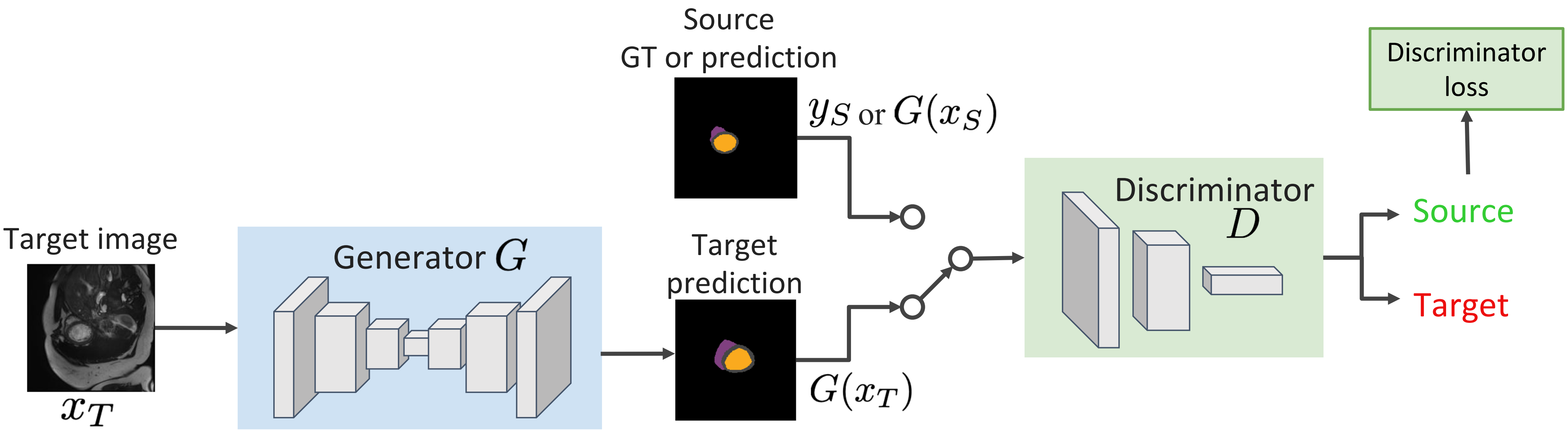


Source domain

Target domain

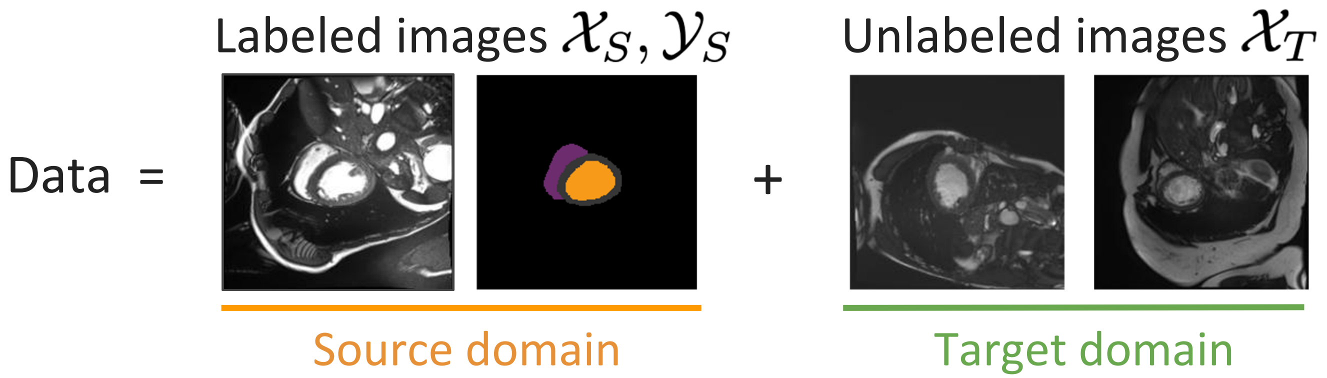
Adversarial domain adaptation

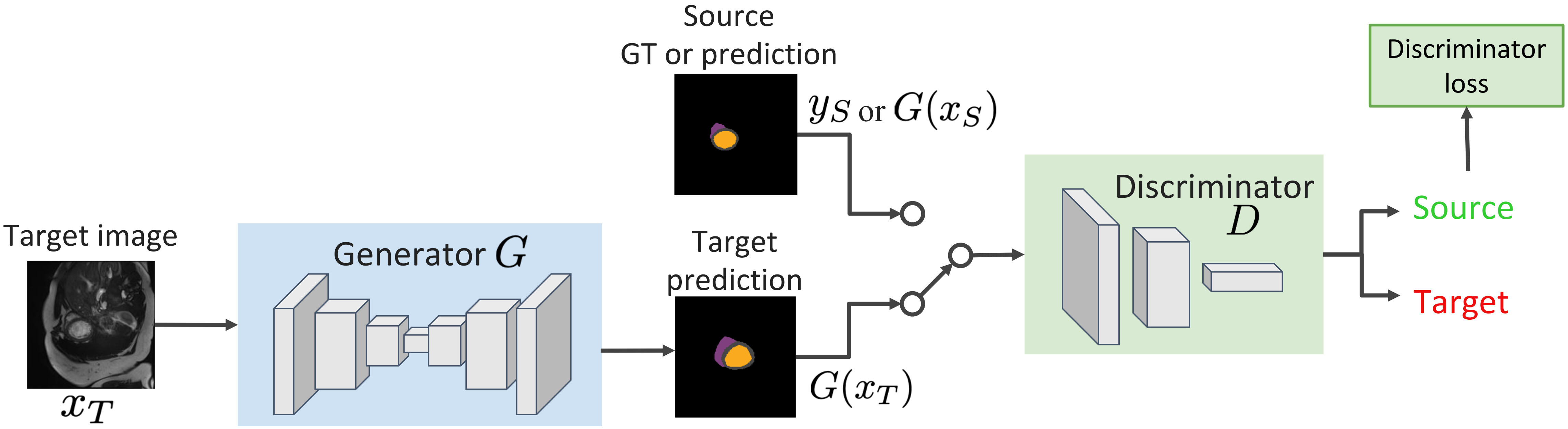




Tsai, Y.-H., et al. "Learning to adapt structured output space for semantic segmentation." IEEE Conf. on Computer Vision and Pattern Recognition. 2018.

Adversarial domain adaptation



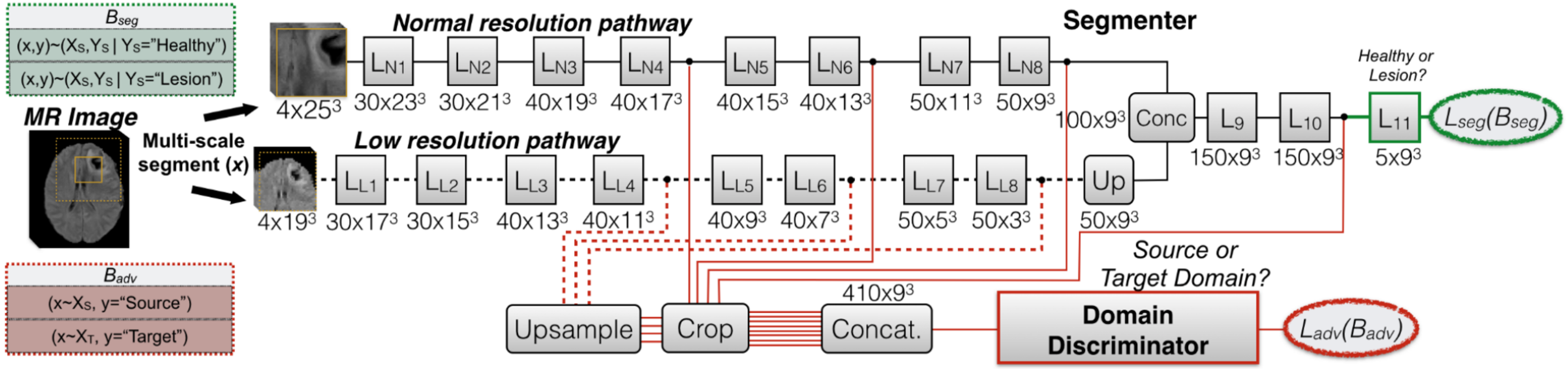


Like semi-supervised segmentation except target images are from a different domain

Tsai, Y.-H., et al. "Learning to adapt structured output space for semantic segmentation." IEEE Conf. on Computer Vision and Pattern Recognition. 2018.

Adversarial domain adaptation

Adversarial domain adaptation for brain lesion segmentation



Source domain (Database 1):

- GE, FLAIR, T2, MPRAGE, PD

Target domain (Database 2):

- SWI, FLAIR, T2, MPRAGE, PD

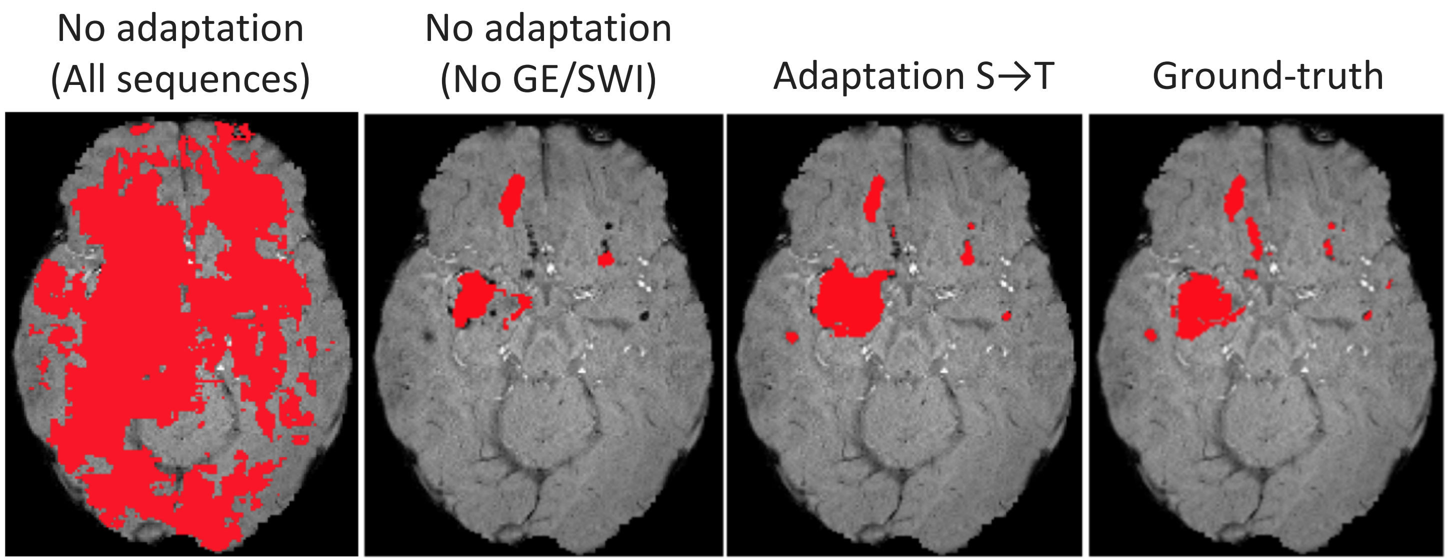
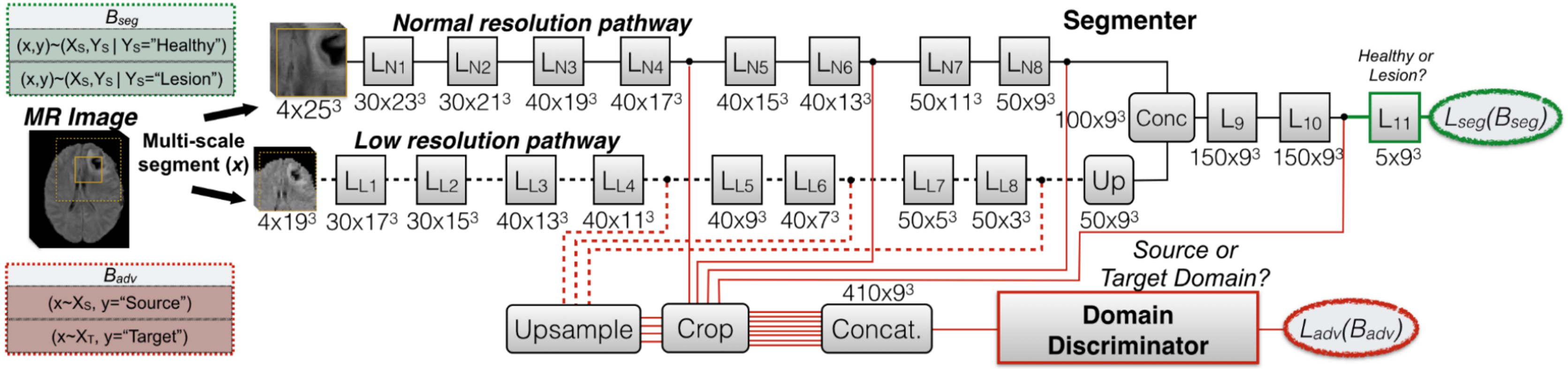


Image from Kamnitsas, K., et al. "Unsupervised domain adaptation in brain lesion segmentation with adversarial networks." *Int. Conf. on Information Processing in Medical Imaging*, 2017.

Adversarial domain adaptation

Adversarial domain adaptation for brain lesion segmentation



Adaptation done on *multi-scale feature representation*

Source domain (Database 1):

- GE, FLAIR, T2, MPRAGE, PD

Target domain (Database 2):

- SWI, FLAIR, T2, MPRAGE, PD

No adaptation (All sequences)	No adaptation (No GE/SWI)	Adaptation S→T	Ground-truth

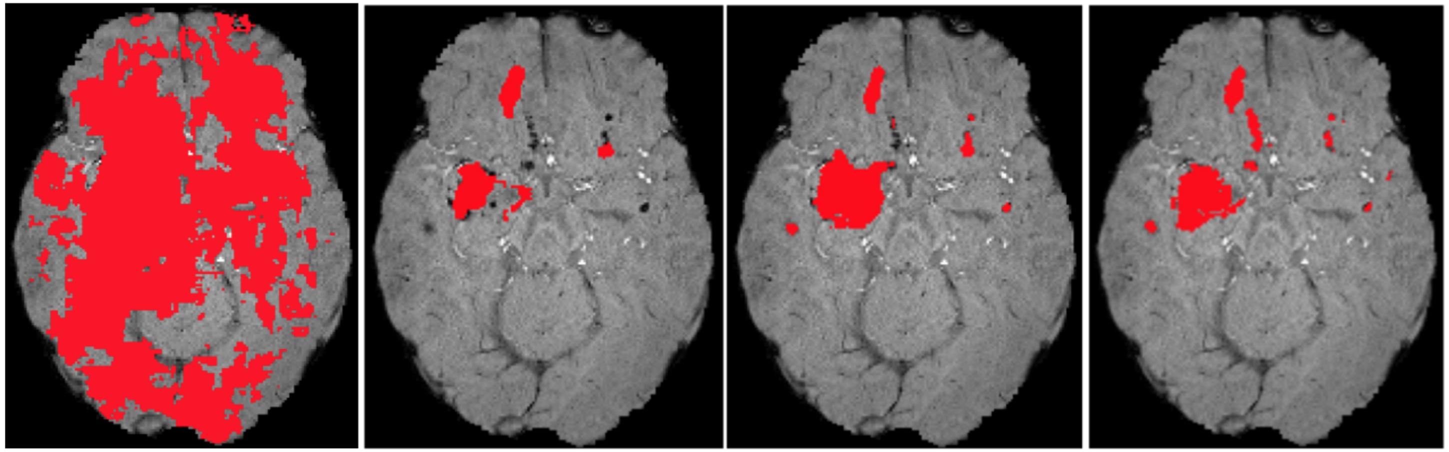


Image from Kamnitsas, K., et al. "Unsupervised domain adaptation in brain lesion segmentation with adversarial networks." *Int. Conf. on Information Processing in Medical Imaging*, 2017.

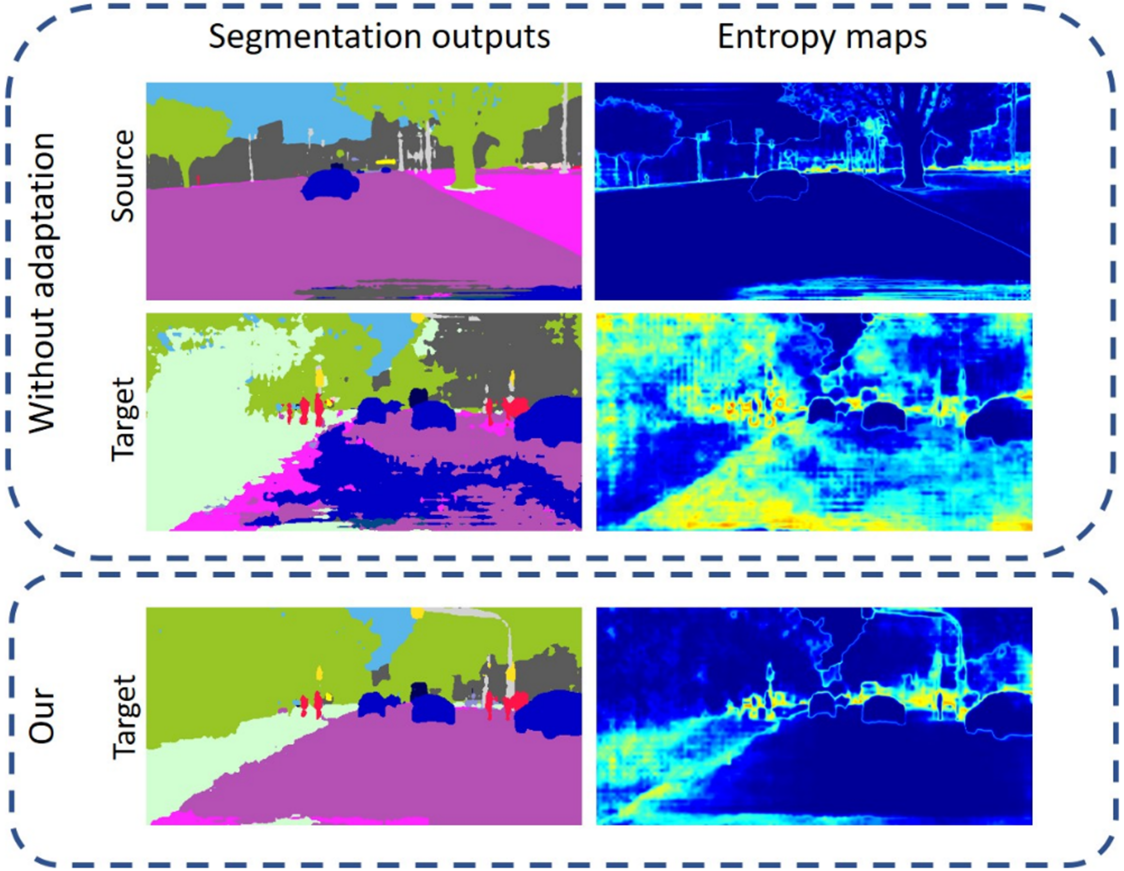
Adversarial domain adaptation

Adaptation on feature *representation* or *softmax* output. <u>What else</u> ?

Vu, T.-H., et al. "Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation." IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2019.

Adversarial domain adaptation

Adaptation on feature *representation* or *softmax* output. What else ?



Adversarial entropy minimization

- The discriminator must differentiate between source and target examples using the entropy spatial maps
- Forces the segmentation model to be consistent in its confidence across different semantic regions

Vu, T.-H., et al. "Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation." IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2019.

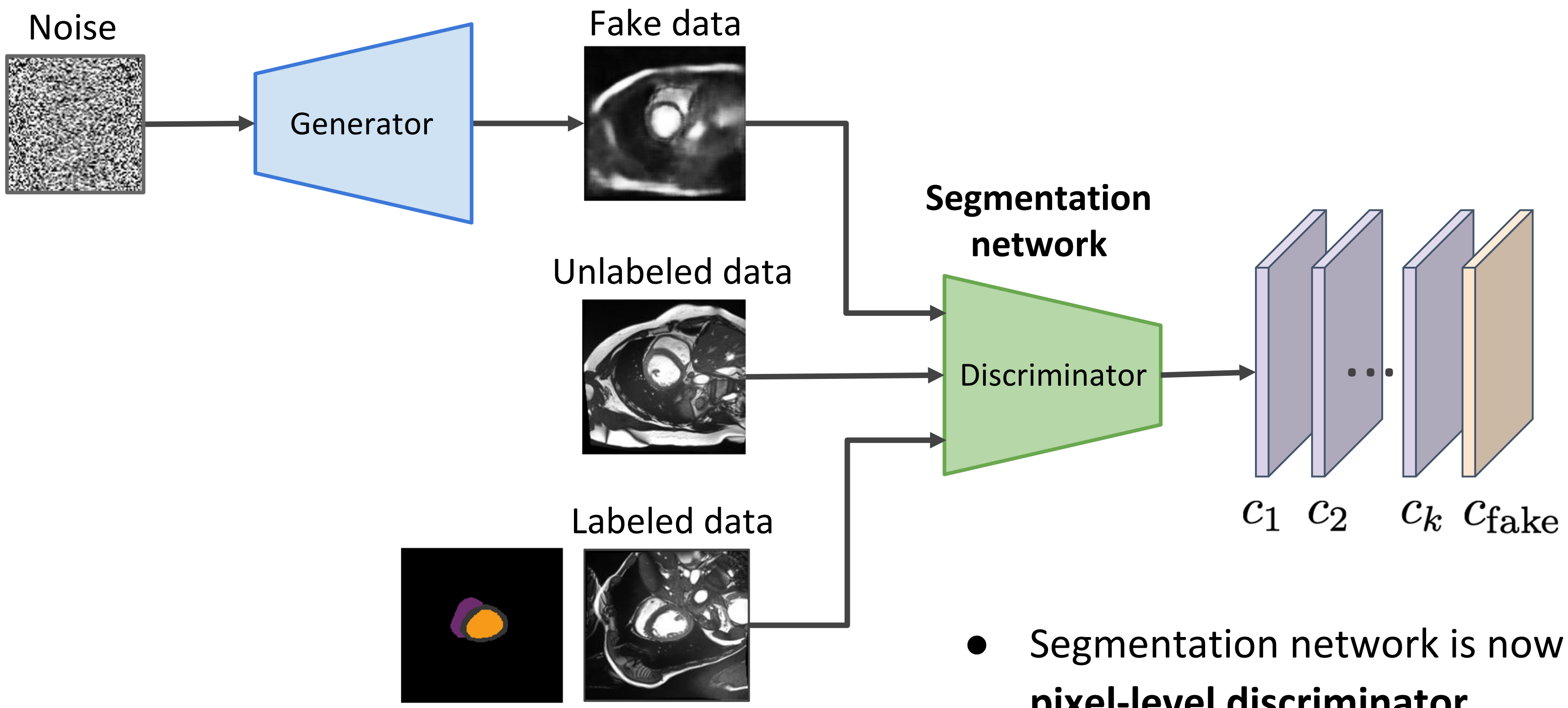
Semi-supervised segmentation with GANs

Can we use GAN-generated images to boost learning in a semi-supervised setting ?

Adapted from: Souly, N. et al.. "Semi supervised semantic segmentation using generative adversarial network." *IEEE Int. Conf. on Computer Vision*. 2017.

Semi-supervised segmentation with GANs

Can we use GAN-generated images to boost learning in a semi-supervised setting ?

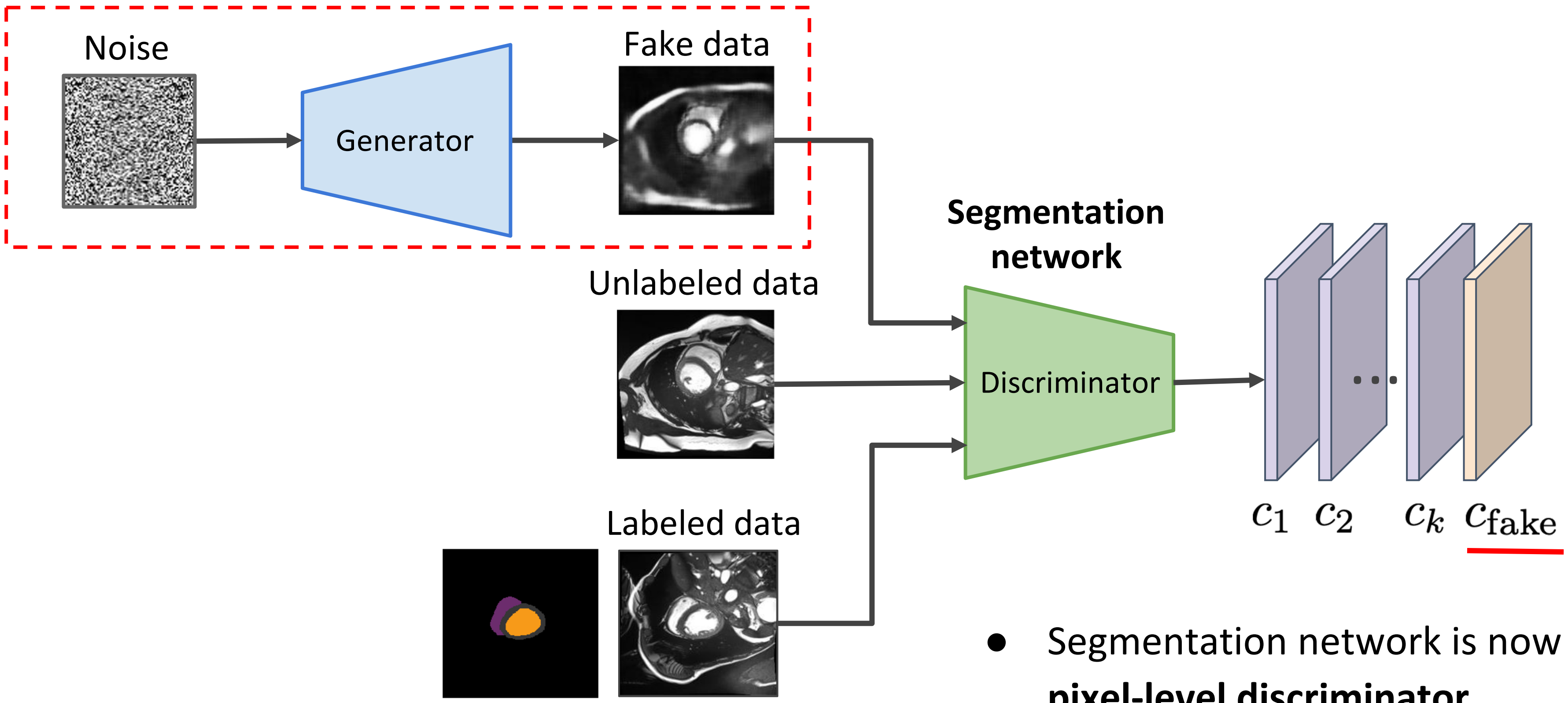


- Segmentation network is now considered as a **pixel-level discriminator**

Adapted from: Souly, N. et al.. "Semi supervised semantic segmentation using generative adversarial network." *IEEE Int. Conf. on Computer Vision*. 2017.

Semi-supervised segmentation with GANs

Can we use GAN-generated images to boost learning in a semi-supervised setting ?

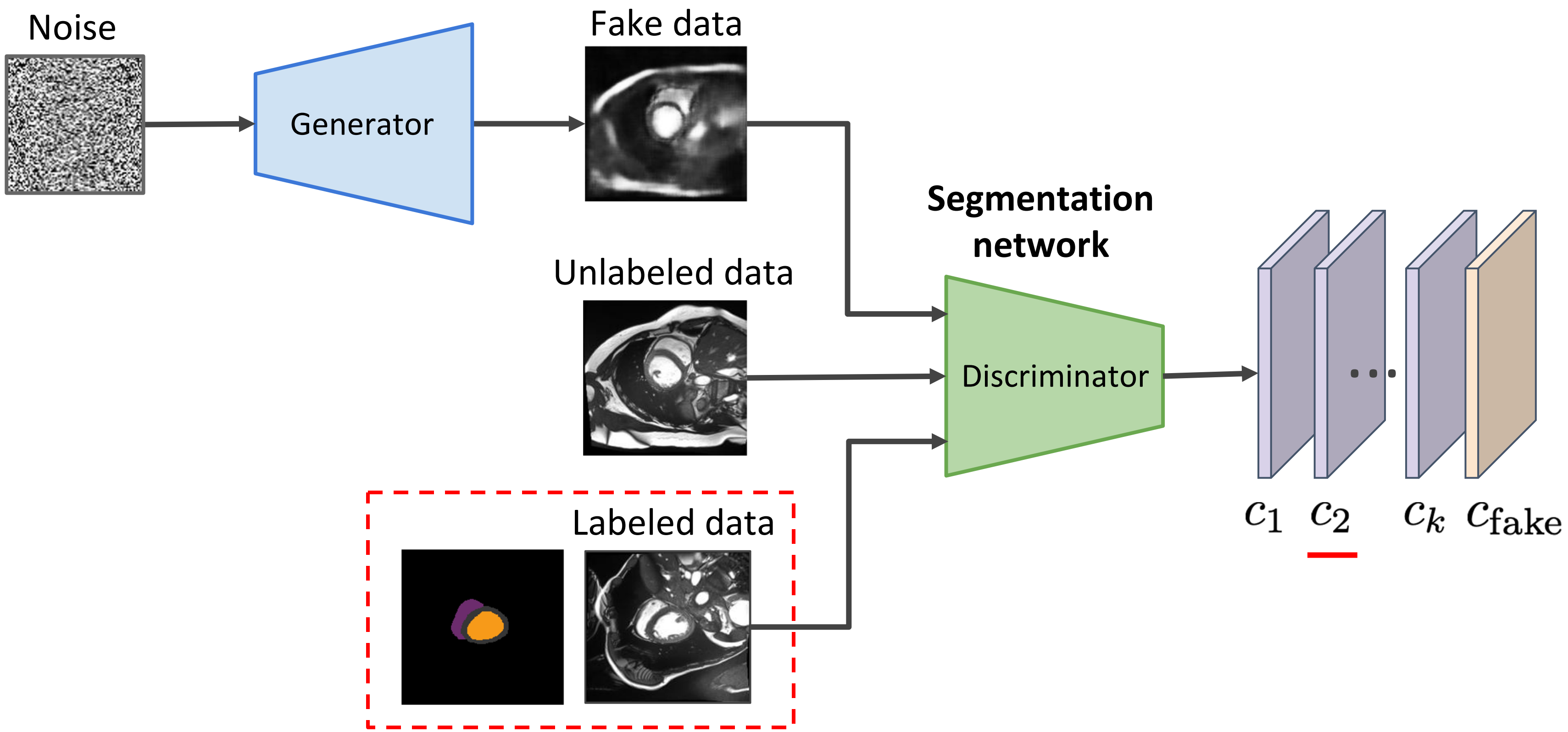


- Segmentation network is now considered as a **pixel-level discriminator**
- For each pixel, predicts the class label or an extra *fake* label

Adapted from: Souly, N. et al.. "Semi supervised semantic segmentation using generative adversarial network." *IEEE Int. Conf. on Computer Vision*. 2017.

Semi-supervised segmentation with GANs

Can we use GAN-generated images to boost learning in a semi-supervised setting ?



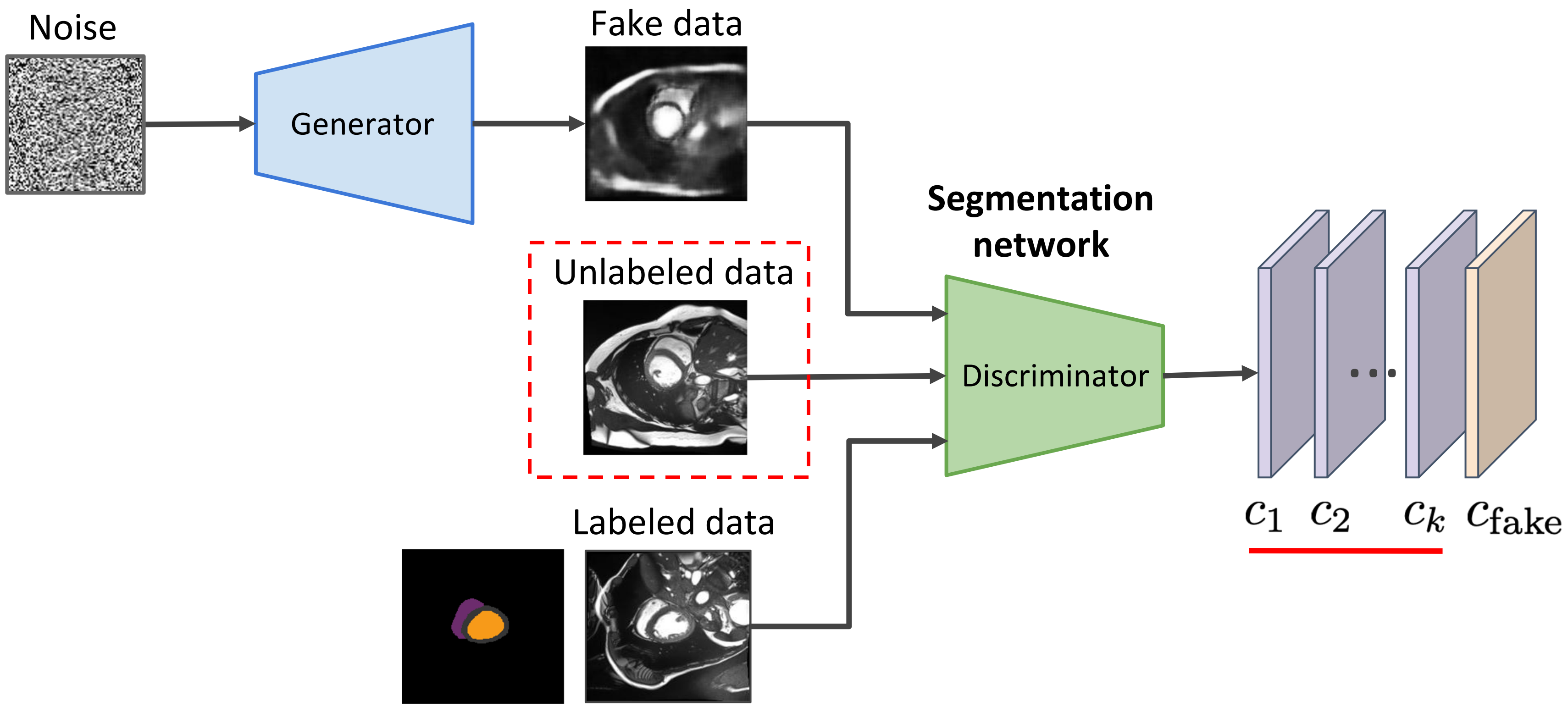
Labeled data: Predict the ground-truth label as in standard supervised segmentation

$$\mathcal{L}_{\text{sup}}(D) = \mathbb{E}_{(x,y)\sim p_{\text{data}}(x,y)}\left[-\sum_i \log p(Y_i = y_i \,|\, x)\right]$$

Adapted from: Souly, N. et al.. "Semi supervised semantic segmentation using generative adversarial network." *IEEE Int. Conf. on Computer Vision*. 2017.

Semi-supervised segmentation with GANs

Can we use GAN-generated images to boost learning in a semi-supervised setting ?



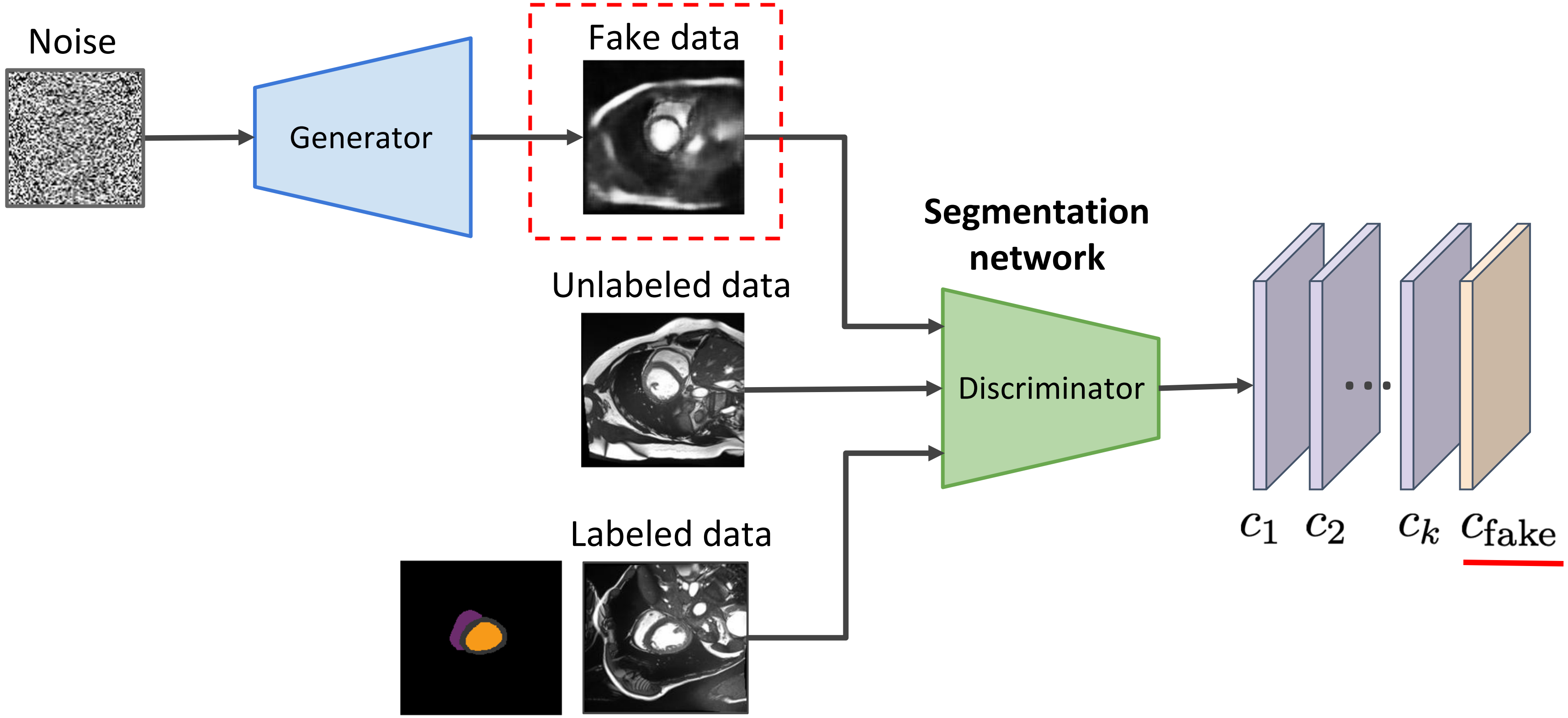
Unlabeled data: Predict the any label <u>except</u> fake

$$\mathcal{L}_{\text{unsup}}(D) = \mathbb{E}_{x\sim p_{\text{data}}(x)}\Big[-\sum_i \log p(Y_i \neq \text{fake}\,|\,x)\Big] = \mathbb{E}_{x\sim p_{\text{data}}(x)}\Big[-\sum_i \log\big(1 - p(Y_i = \text{fake}\,|\,x)\big)\Big]$$

Adapted from: Souly, N. et al.. "Semi supervised semantic segmentation using generative adversarial network." *IEEE Int. Conf. on Computer Vision*. 2017.

Semi-supervised segmentation with GANs

Can we use GAN-generated images to boost learning in a semi-supervised setting ?



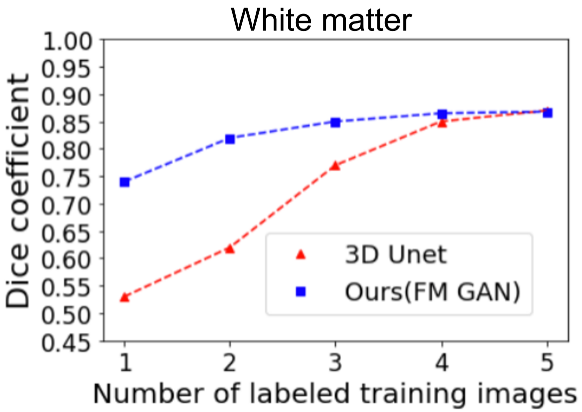
Fake data: Predict the label fake at <u>every pixel</u>

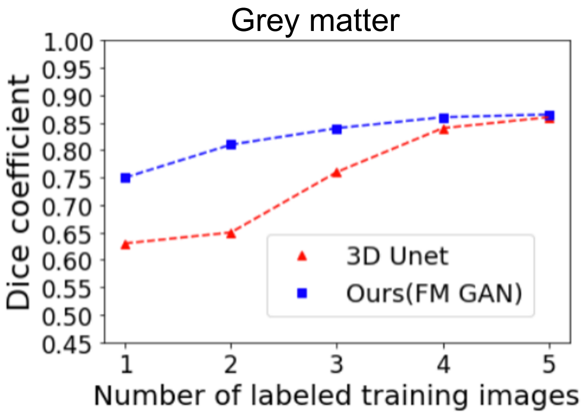
$$\mathcal{L}_{\text{fake}}(G, D) \;=\; \mathbb{E}_{z \sim p_z(z)}\Big[-\sum_i \log p(Y_i = \text{fake} \,|\, G(z))\Big]$$

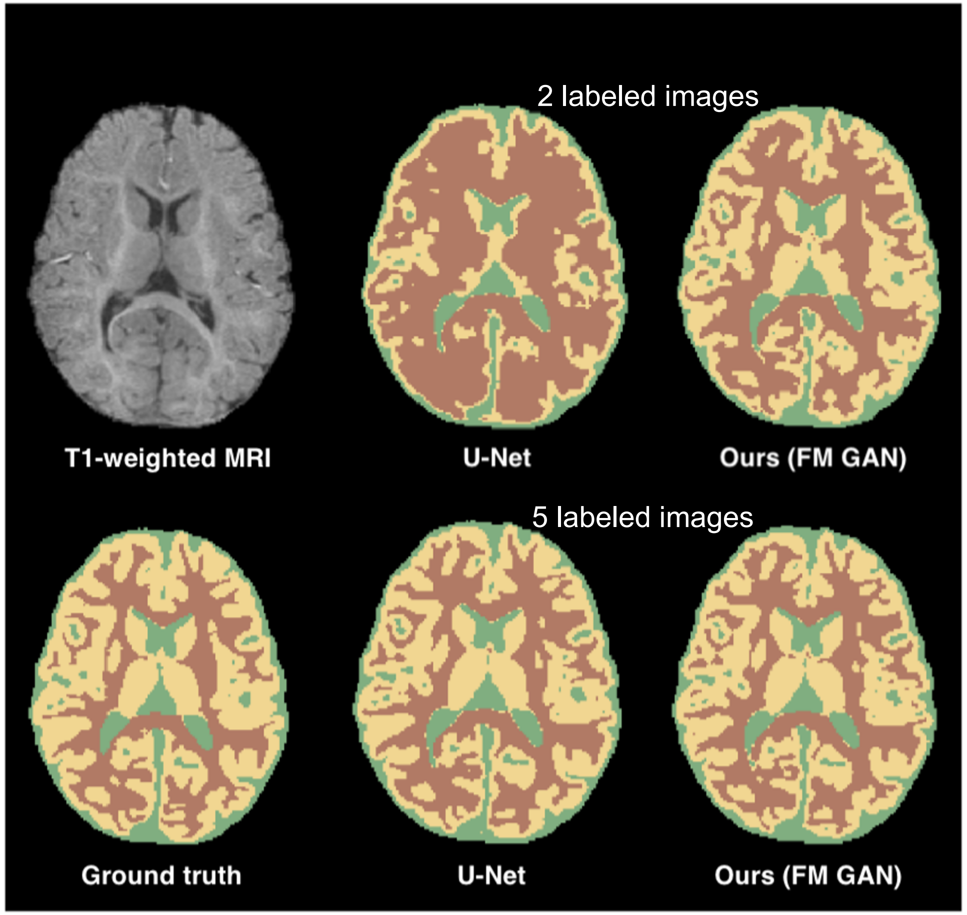
Adapted from: Souly, N. et al.. "Semi supervised semantic segmentation using generative adversarial network." *IEEE Int. Conf. on Computer Vision*. 2017.

Semi-supervised segmentation with GANs

Application to brain segmentation with very few training images







Mondal, A.K., Dolz, J. and Desrosiers, C. "Few-shot 3D multi-modal medical image segmentation using generative adversarial learning." *arXiv* (2018).

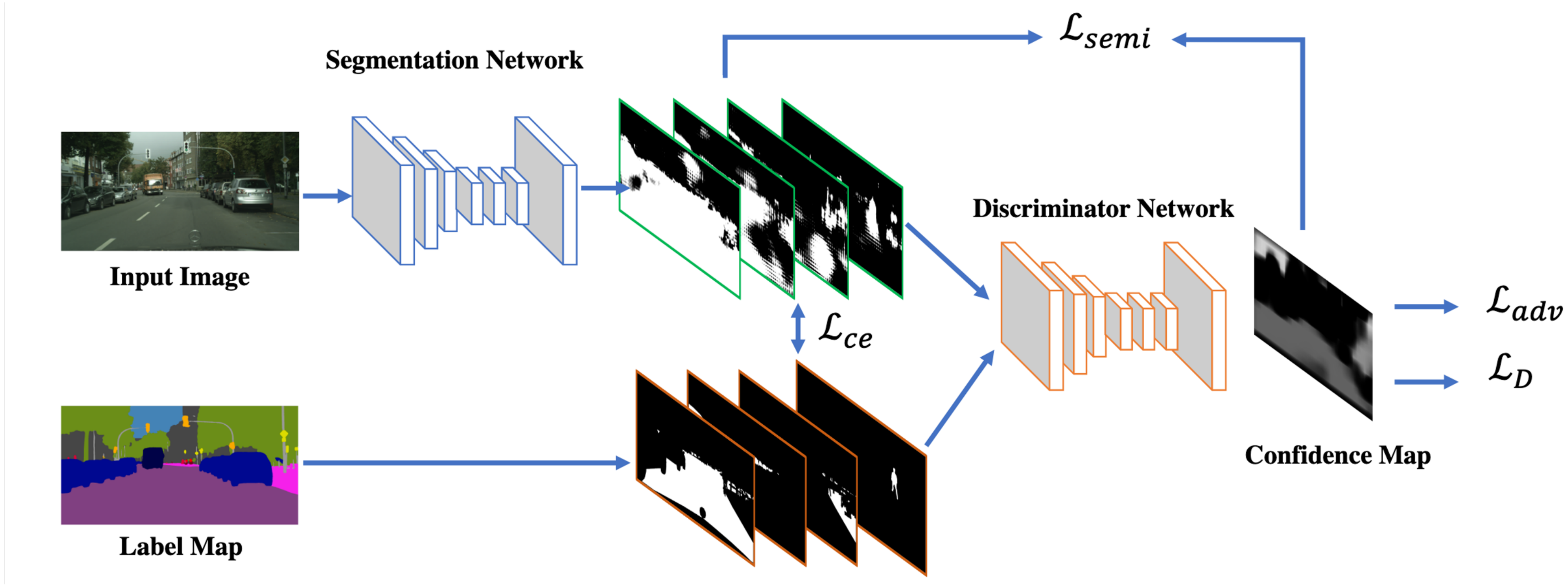
Adversarial model for self-training

How else can we leverage discriminator predictions at the pixel-level ?

Hung, W.-C. et al. "Adversarial learning for semi-supervised semantic segmentation." *British Machine Vision Conference* (*BMVC*), 2018.

Adversarial model for self-training

How else can we leverage discriminator predictions at the pixel-level ?

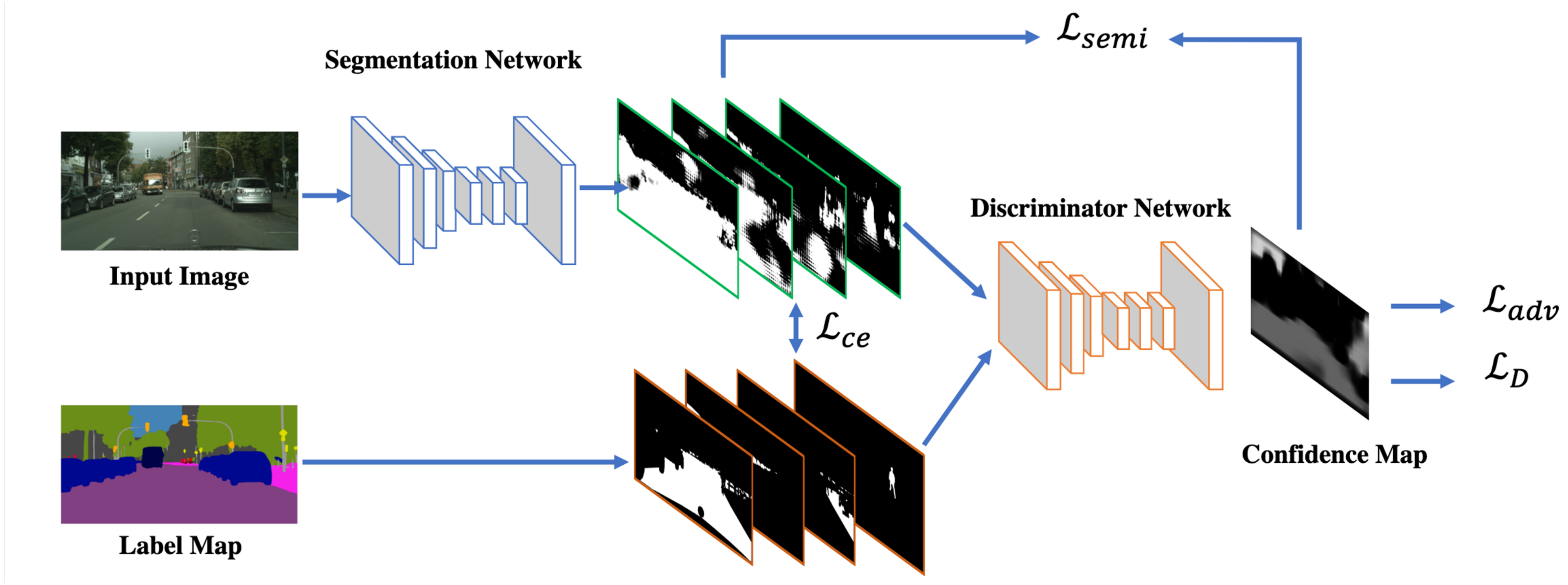


- The discriminator must discriminate between prediction and ground-truth (GT) at each pixel
- Consider the discriminator GT-class probabilities as confidence scores
- Use high-confidence predictions on unlabeled images as pseudo-labels for self-training

Hung, W.-C. et al. "Adversarial learning for semi-supervised semantic segmentation." *British Machine Vision Conference* (*BMVC*), 2018.

Adversarial model for self-training

How else can we leverage discriminator predictions at the pixel-level ?



Select pixels with confidence above a given threshold

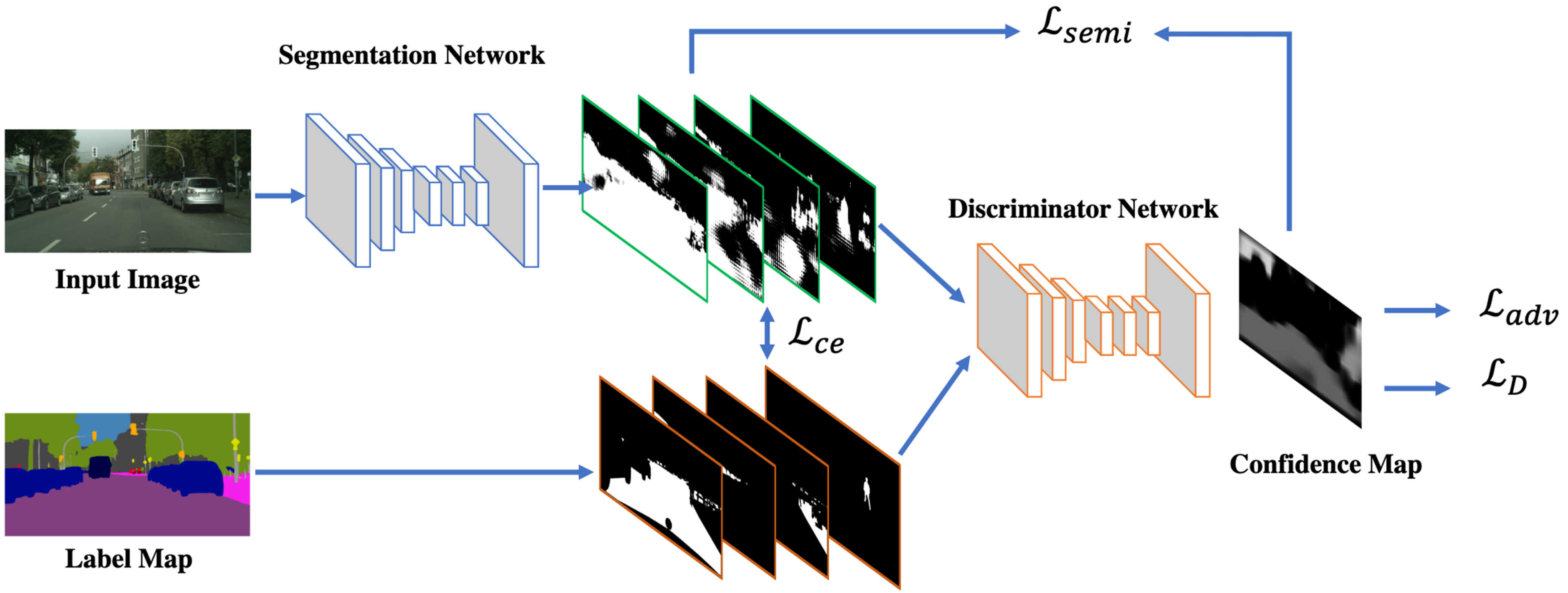
$$\mathcal{L}_{semi} = -\sum_{h,w}\sum_{c\in C} I(D(S(\mathbf{X}_n))^{(h,w)} > T_{semi}) \cdot \hat{\mathbf{Y}}_n^{(h,w,c)} \log(S(\mathbf{X}_n)^{(h,w,c)})$$

$$\hat{\mathbf{Y}}_n^{(h,w,c^*)} = 1 \text{ if } c^* = \arg\max_c S(\mathbf{X}_n)^{(h,w,c)}$$

Hung, W.-C. et al. "Adversarial learning for semi-supervised semantic segmentation." *British Machine Vision Conference* (*BMVC*), 2018.

Adversarial model for self-training

How else can we leverage discriminator predictions at the pixel-level ?



Use class with highest probability as pseudo-label

$$\mathcal{L}_{semi} = -\sum_{h,w}\sum_{c\in C} I(D(S(\mathbf{X}_n))^{(h,w)} > T_{semi}) \cdot \boxed{\hat{\mathbf{Y}}_n^{(h,w,c)}} \log(S(\mathbf{X}_n)^{(h,w,c)})$$

$$\boxed{\hat{\mathbf{Y}}_n^{(h,w,c^*)} = 1 \text{ if } c^* = \arg\max_c S(\mathbf{X}_n)^{(h,w,c)}}$$

Hung, W.-C. et al. "Adversarial learning for semi-supervised semantic segmentation." *British Machine Vision Conference* (*BMVC*), 2018.

Cycle GANs for domain adaptation

How can we learn a model to segment target images without paired images or GT ?

Source domain

Image

Ground-truth

Target domain

Unlabeled Image

Hoffman, J., et al. "CyCADA: Cycle-Consistent Adversarial Domain Adaptation." *Int. Conf. on Machine Learning* (ICML). 2018.

Cycle GANs for domain adaptation

How can we learn a model to segment target images without paired images or GT ?

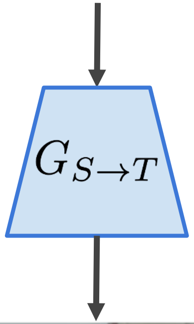
Source domain

Image

Ground-truth

Target domain

Unlabeled Image



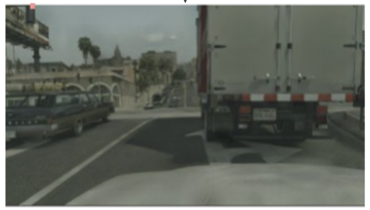


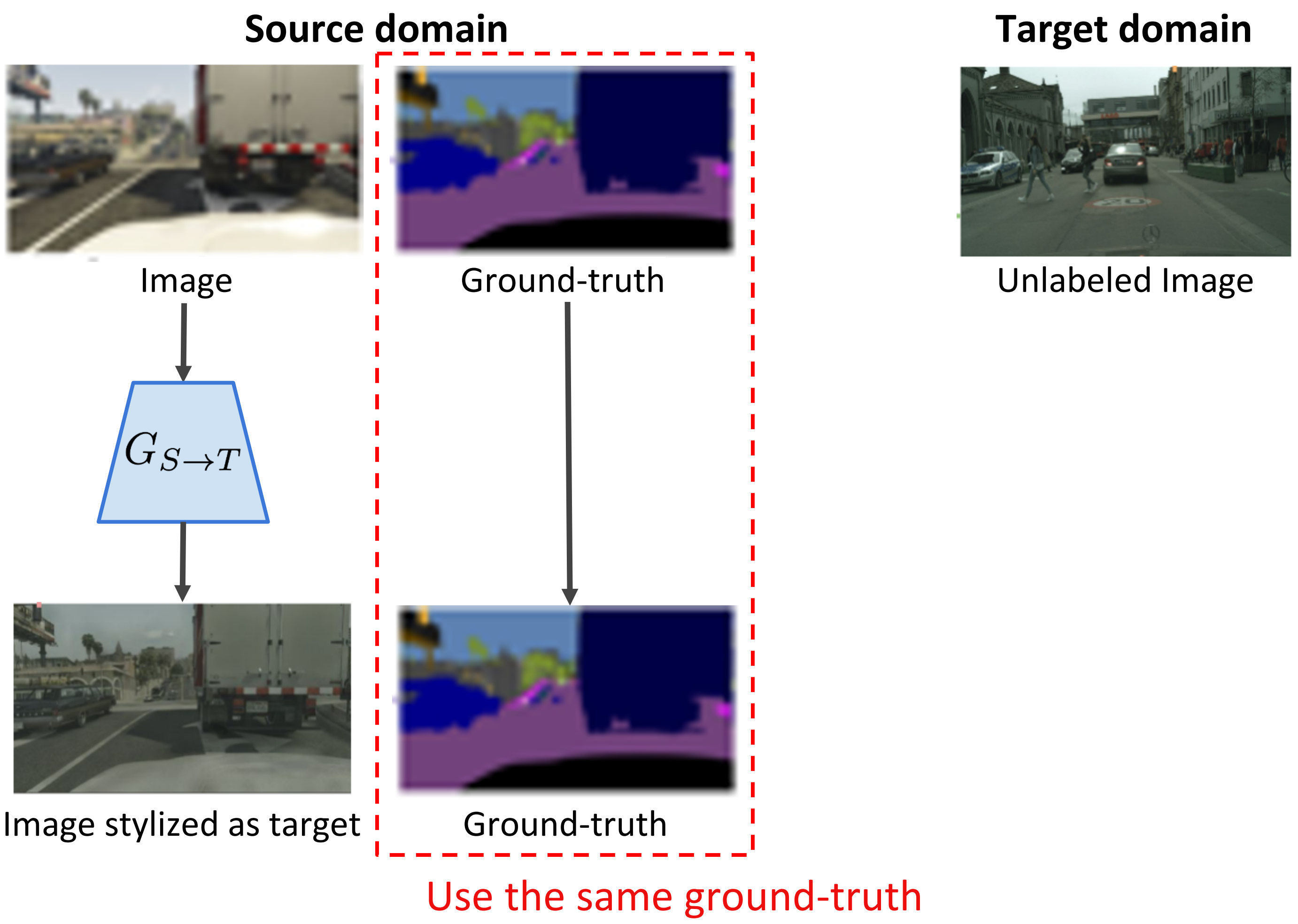
Image stylized as target

Learn an unpaired style transfer model from source to target domain

Hoffman, J., et al. "CyCADA: Cycle-Consistent Adversarial Domain Adaptation." *Int. Conf. on Machine Learning* (ICML). 2018.

Cycle GANs for domain adaptation

How can we learn a model to segment target images without paired images or GT ?

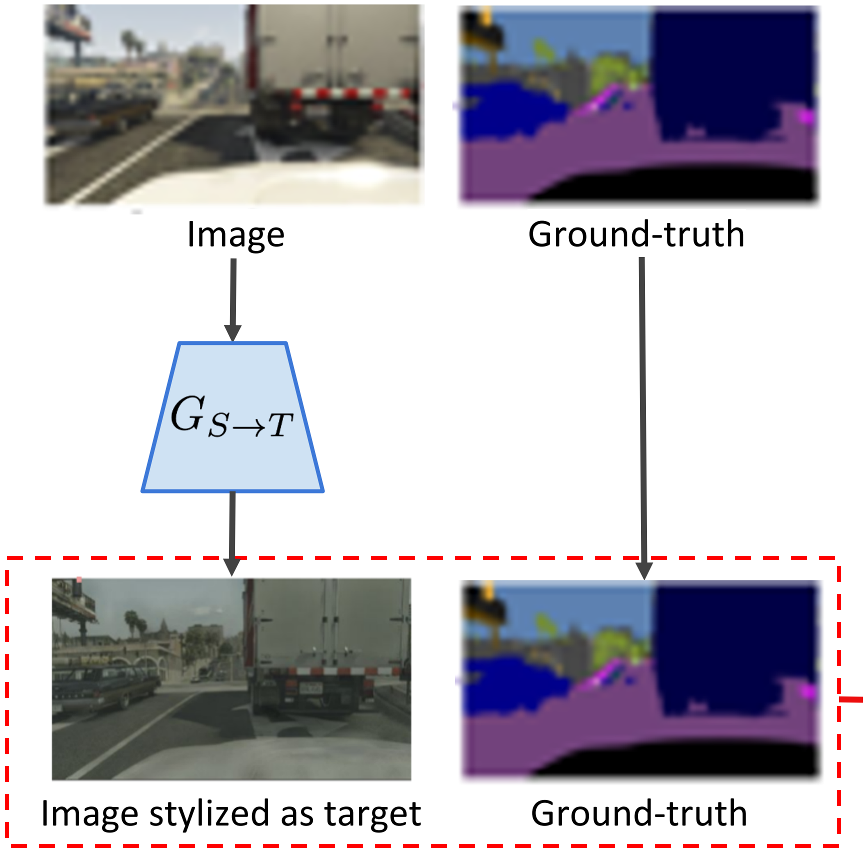


Hoffman, J., et al. "CyCADA: Cycle-Consistent Adversarial Domain Adaptation." *Int. Conf. on Machine Learning* (ICML). 2018.

Cycle GANs for domain adaptation

How can we learn a model to segment target images without paired images or GT ?

Source domain



Target domain

Unlabeled Image

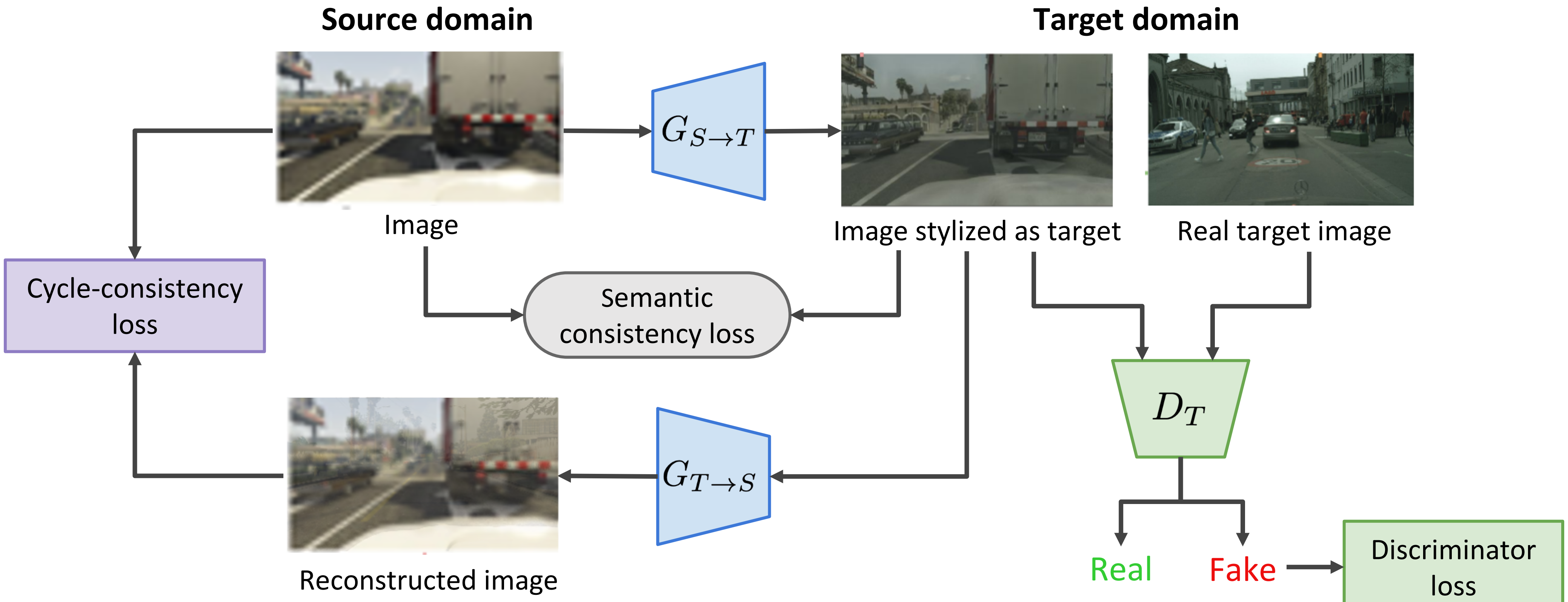
Learn the network for segmenting target images

How to learn the unpaired style transfer network ?

Hoffman, J., et al. "CyCADA: Cycle-Consistent Adversarial Domain Adaptation." *Int. Conf. on Machine Learning* (ICML). 2018.

Cycle GANs for domain adaptation

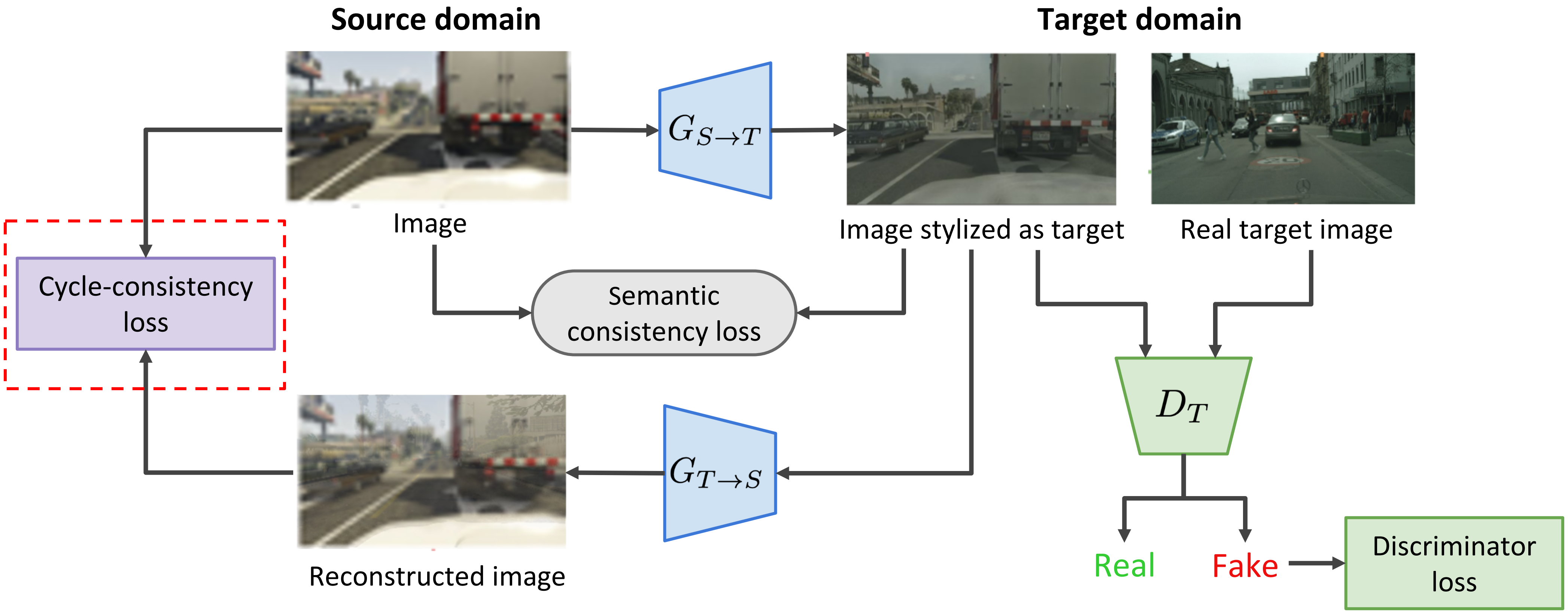
How can we learn a model to segment target images without paired images or GT ?



Hung, W.-C. et al. "Adversarial learning for semi-supervised semantic segmentation." *British Machine Vision Conference* (*BMVC*), 2018.

Cycle GANs for domain adaptation

How can we learn a model to segment target images without paired images or GT ?

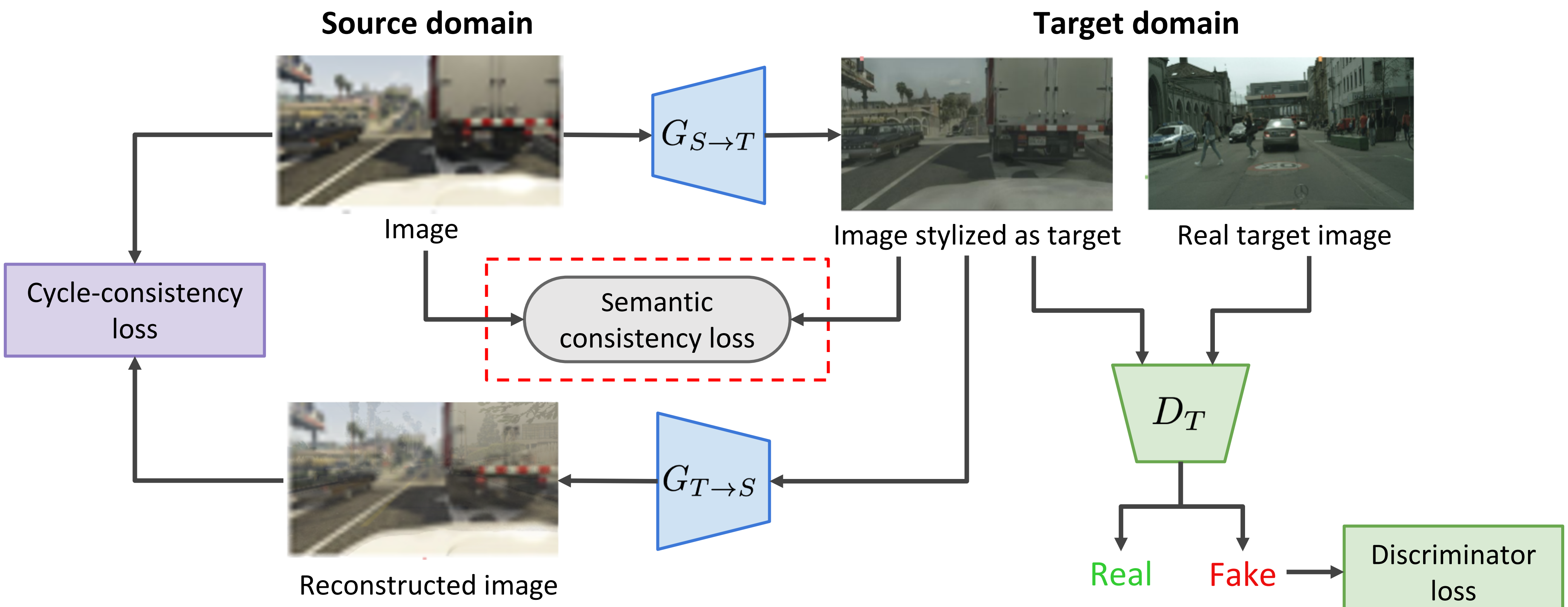


Cycle consistency loss: $L_{\text{cycle}}(G_{S\to T}, G_{T\to S}) = \mathbb{E}_{x\sim p_S(x)}\left[\|x - G_{T\to S}(G_{S\to T}(x))\|_1\right]$

Hoffman, J., et al. "CyCADA: Cycle-Consistent Adversarial Domain Adaptation." *Int. Conf. on Machine Learning* (ICML). 2018.

Cycle GANs for domain adaptation

How can we learn a model to segment target images without paired images or GT ?

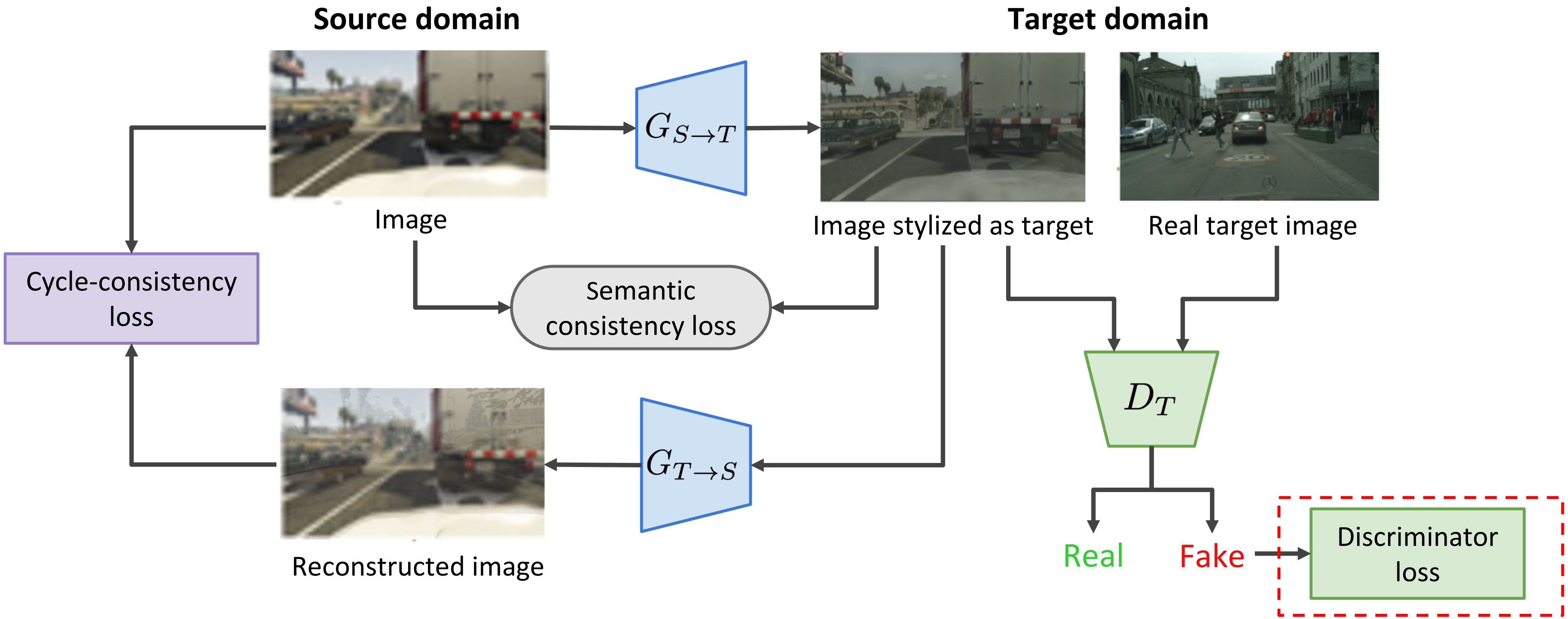


Semantic consistency loss: Segmentation for the source image and its stylized target version should be consistent

Hoffman, J., et al. "CyCADA: Cycle-Consistent Adversarial Domain Adaptation." *Int. Conf. on Machine Learning* (ICML). 2018.

Cycle GANs for domain adaptation

How can we learn a model to segment target images without paired images or GT ?

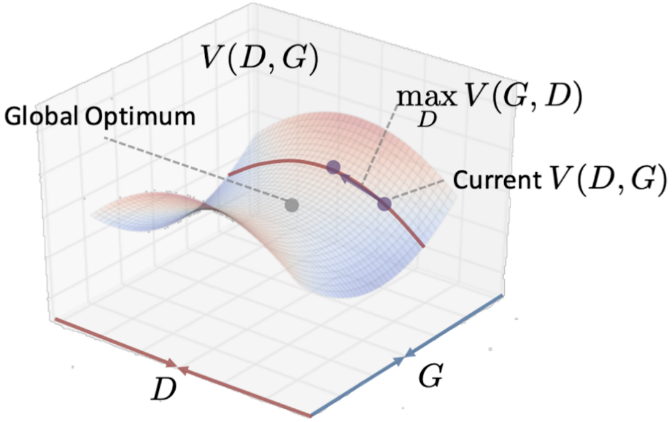


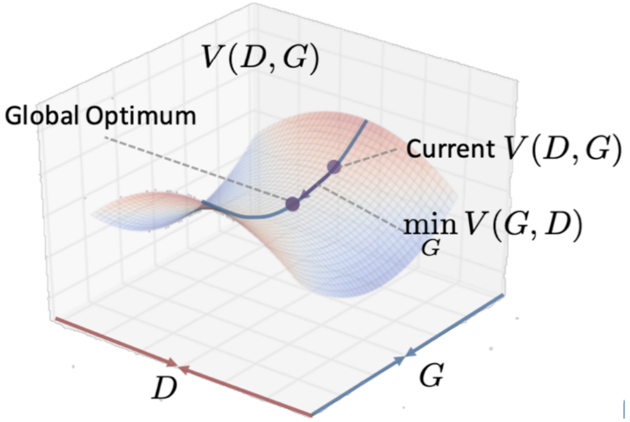
Discriminator loss: Target images generated from source should look like real target ones

Hoffman, J., et al. "CyCADA: Cycle-Consistent Adversarial Domain Adaptation." *Int. Conf. on Machine Learning* (ICML). 2018.

Challenges of adversarial learning

1) Unstable optimization of minimax problem

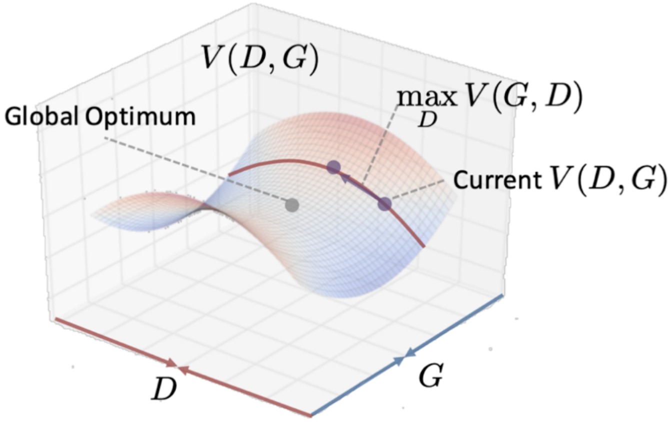


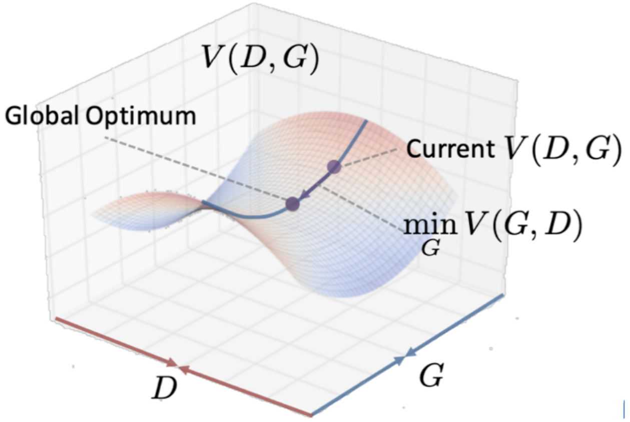


Ghosh, A., et al. "Multi-agent diverse generative adversarial networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
Chang, Mark. “Generative Adversarial Networks”, published online, 2016

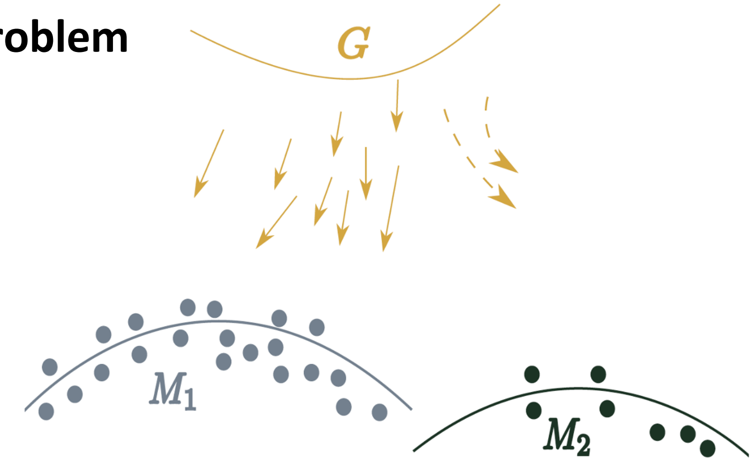
Challenges of adversarial learning

1) Unstable optimization of minimax problem





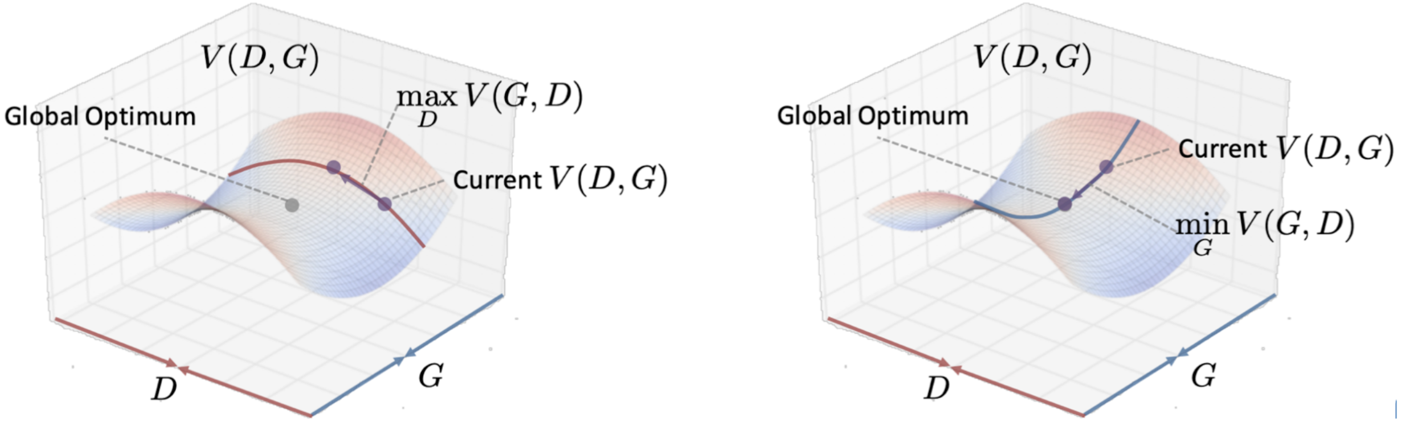
2) Mode collapse problem



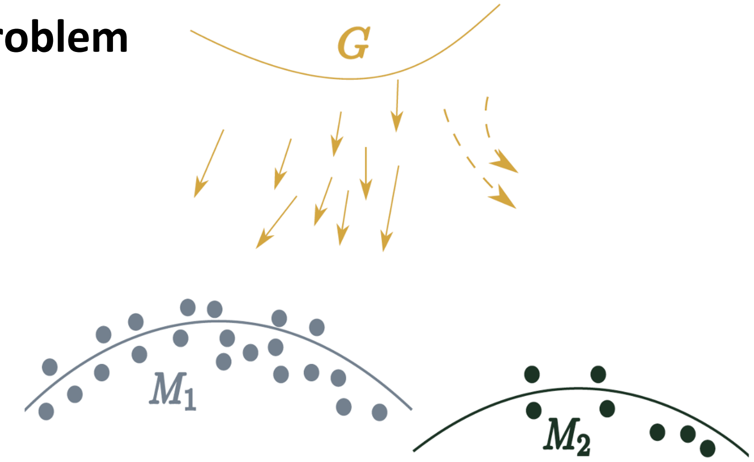
Ghosh, A., et al. "Multi-agent diverse generative adversarial networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
Chang, Mark. "Generative Adversarial Networks", published online, 2016

Challenges of adversarial learning

1) Unstable optimization of minimax problem



2) Mode collapse problem



Various solutions:

- Spectral normalization (Miyato *et al.*, 2018)
- Wasserstein GANs (Arjovsky *et al.*, 2017)
- LSGANs (Mao *et al.*, 2017)
- etc.

Ghosh, A., et al. "Multi-agent diverse generative adversarial networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
Chang, Mark. "Generative Adversarial Networks", published online, 2016

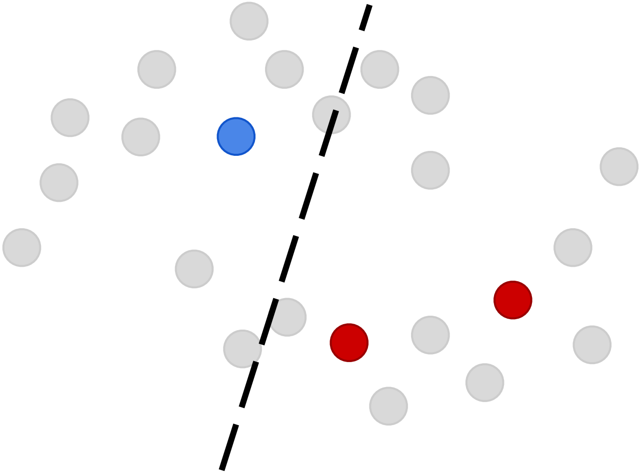
Consistency regularization

for weakly-supervised segmentation

Consistency regularization for SSL

How to better use unlabeled data ?

Vanilla supervised learning

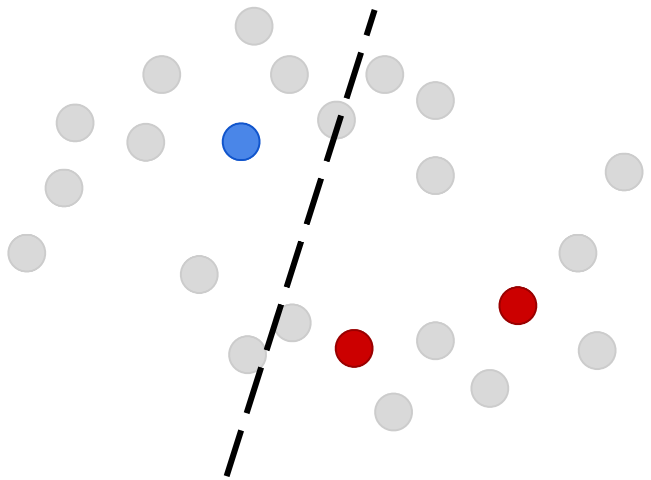


- Consider only *labeled* samples
- Overfits when few training samples

Consistency regularization for SSL

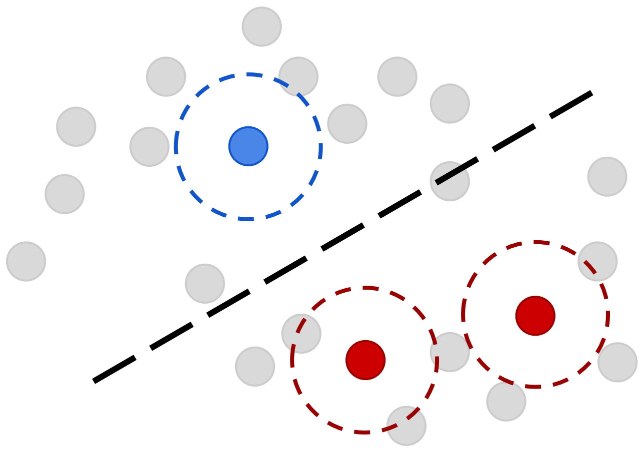
How to better use unlabeled data ?

Vanilla supervised learning



- Consider only *labeled* samples
- Overfits when few training samples

Data augmentation

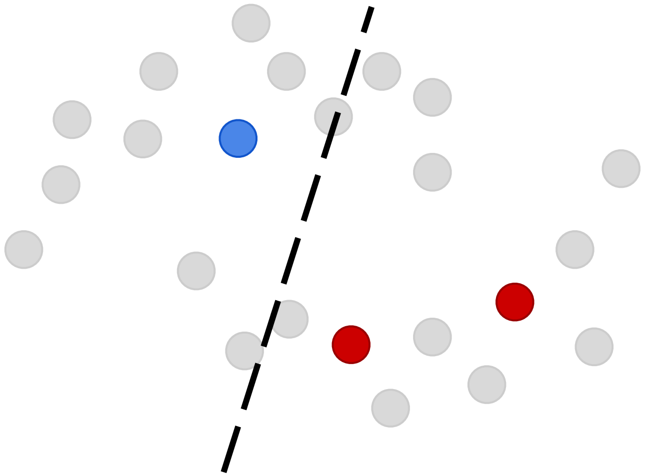


- Transform *labeled* examples to augment the training set
- Better generalization, but not enough for semi-supervised learning

Consistency regularization for SSL

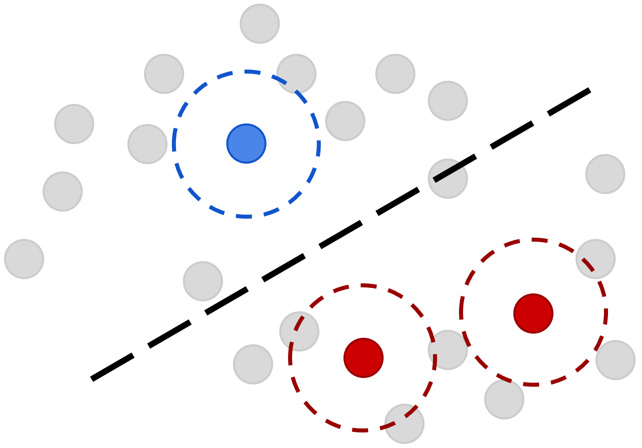
How to better use <u>unlabeled data</u> ?

Vanilla supervised learning



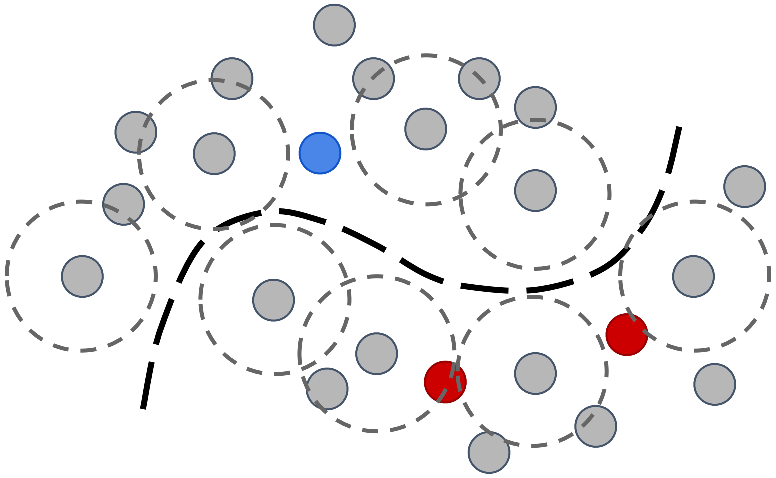
- Consider only *labeled* samples
- Overfits when few training samples

Data augmentation



- Transform *labeled* samples to augment the training set
- Better generalization, but not enough for semi-supervised learning

Consistency regularization



- Perturb *unlabeled* samples with noise or guided transformations
- Impose the network to have consistent outputs for perturbed samples

SSL methods using consistency regularization

Basic transformation consistency (Γ-model)

$$\mathcal{L}(\theta;\, \mathcal{D}_l, \mathcal{D}_u) \;=\; \frac{1}{|\mathcal{D}_l|} \sum_{(x,y)\in\mathcal{D}_l} \ell_{\text{sup}}\big(f(x), y\big) \;+\; \frac{\lambda}{|\mathcal{D}_u|} \sum_{x\in\mathcal{D}_u} \mathbb{E}_{T\sim p_T}\Big[\ell_{\text{reg}}\big(T(f(x)),\; f(T(x))\big)\Big]$$

SSL methods using consistency regularization

Basic transformation consistency (Γ-model)

Standard supervised loss

$$\mathcal{L}(\theta; \mathcal{D}_l, \mathcal{D}_u) = \boxed{\frac{1}{|\mathcal{D}_l|} \sum_{(x,y)\in\mathcal{D}_l} \ell_{\text{sup}}\big(f(x), y\big)} + \frac{\lambda}{|\mathcal{D}_u|} \sum_{x\in\mathcal{D}_u} \mathbb{E}_{T\sim p_T}\Big[\ell_{\text{reg}}\big(T(f(x)),\, f(T(x))\big)\Big]$$

Cross-entropy, Dice, etc.

SSL methods using consistency regularization

Basic transformation consistency (Γ-model)

Transformation consistency loss

$$\mathcal{L}(\theta; \mathcal{D}_l, \mathcal{D}_u) = \frac{1}{|\mathcal{D}_l|} \sum_{(x,y)\in\mathcal{D}_l} \ell_{\text{sup}}\big(f(x), y\big) + \boxed{\frac{\lambda}{|\mathcal{D}_u|} \sum_{x\in\mathcal{D}_u} \mathbb{E}_{T\sim p_T} \Big[\ell_{\text{reg}}\big(T(f(x)),\, f(T(x))\big)\Big]}$$

Random transformation: rotation, flip, crop, etc.

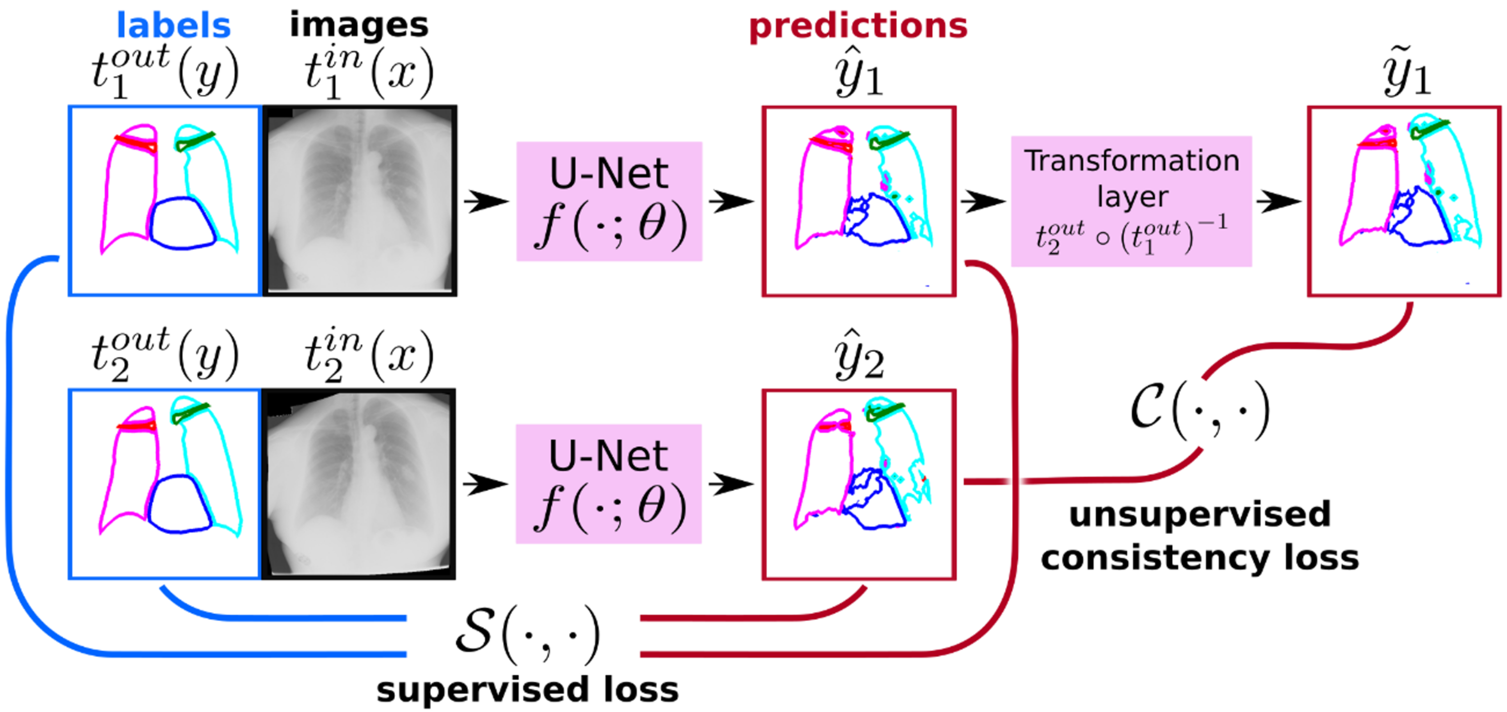
Regularization loss imposing ***transformation equivariance***

L2 regularization loss:

$$\ell_{\text{reg}}\big(T(f(x)),\, f(T(x))\big) = \big\|T(f(x)) - f(T(x))\big\|_2^2$$

SSL methods using consistency regularization

Application to chest X-ray segmentation:

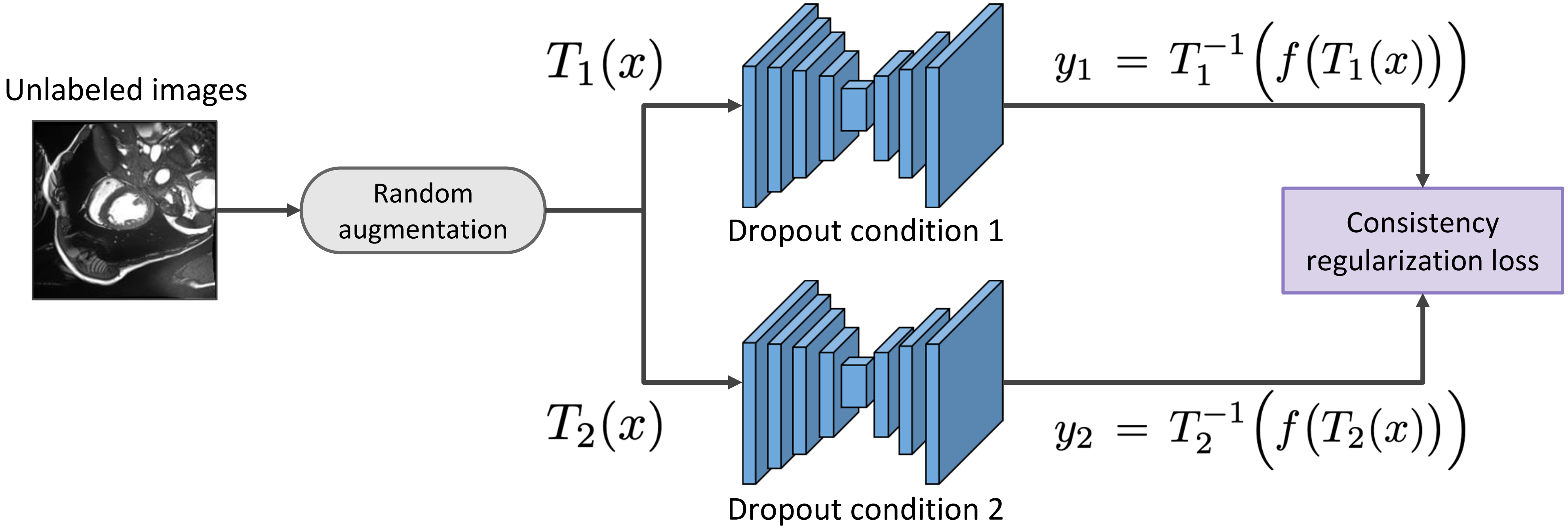


Transformations are random elastic deformations

Bortsova, G., et al.. "Semi-supervised medical image segmentation via learning consistency under transformations." MICCAI, 2019.

SSL methods using consistency regularization

Self-ensembling (Π-model):

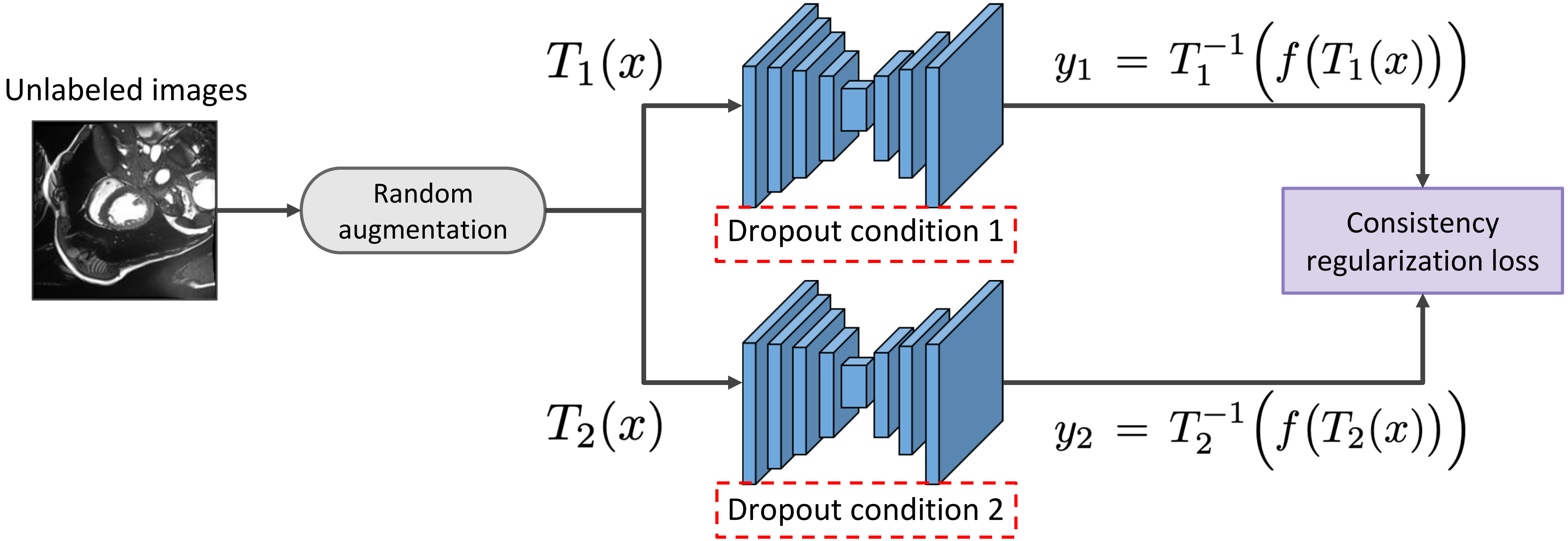


Key idea:

- Applying <u>different dropouts</u> on the <u>same network</u> gives an ensemble of models
- Also leverages random image transformations

SSL methods using consistency regularization

Self-ensembling (Π-model):

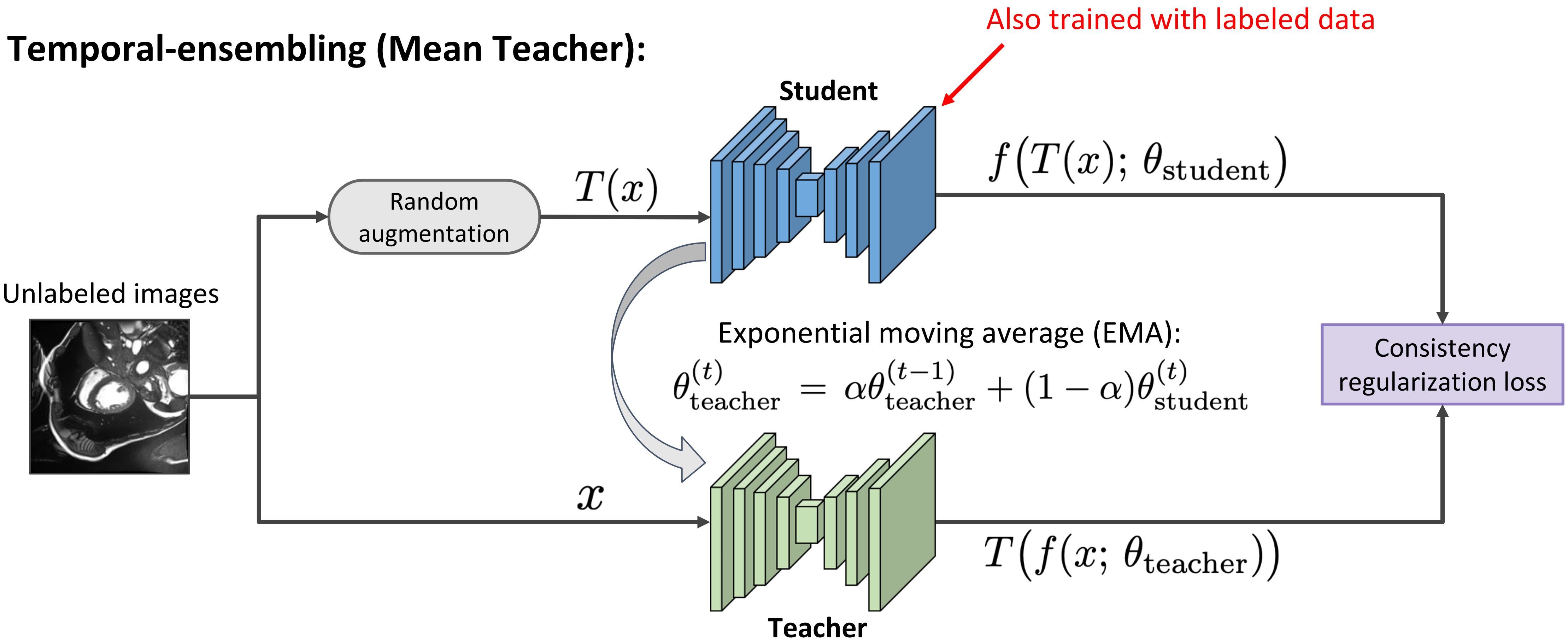


Key idea:

- Applying <u>different dropouts</u> on the <u>same network</u> gives an ensemble of models
- Also leverages random image transformations

SSL methods using consistency regularization

Temporal-ensembling (Mean Teacher):

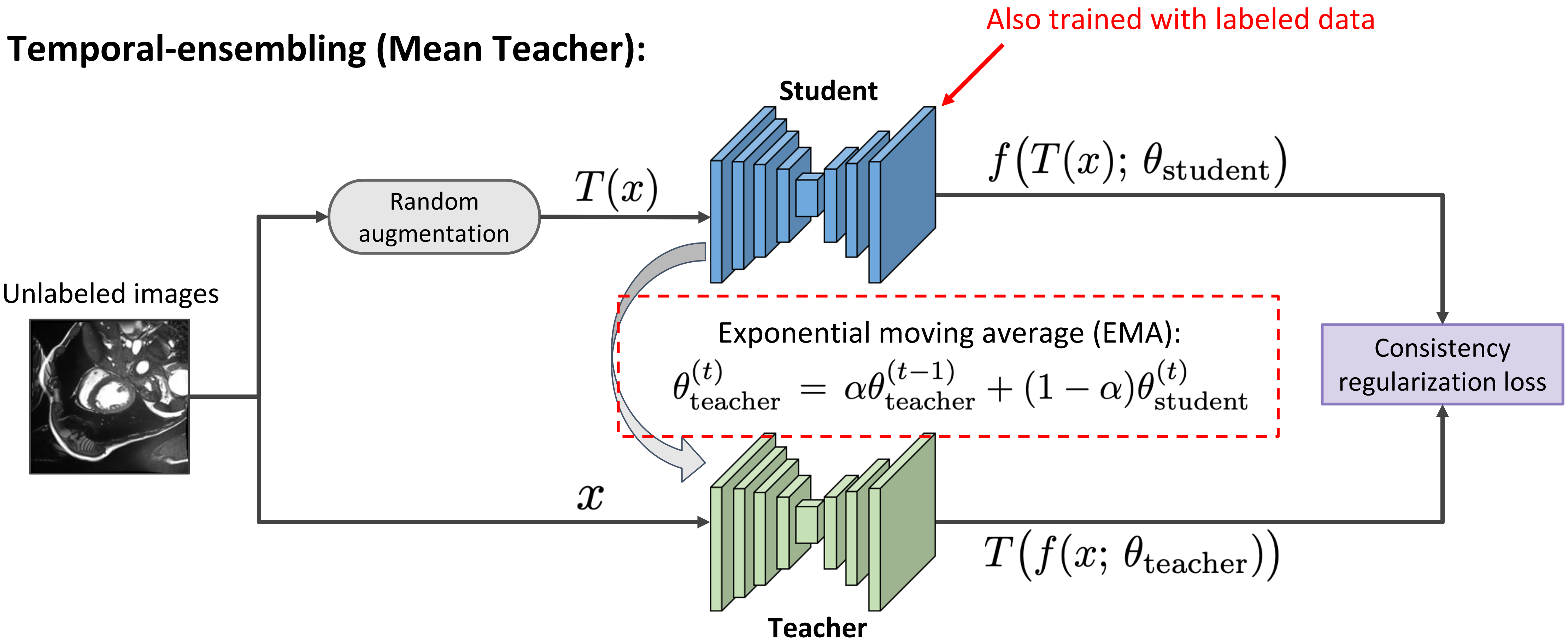


Key idea:

- Consistency between the predictions of a Teacher and a Student network
- The Teacher's weights are an EMA of the Student's at previous training iterations ($\alpha \approx 1$)
- *__Note__*: original Temporal Ensembling computes the EMA on outputs for each sample

SSL methods using consistency regularization

Temporal-ensembling (Mean Teacher):

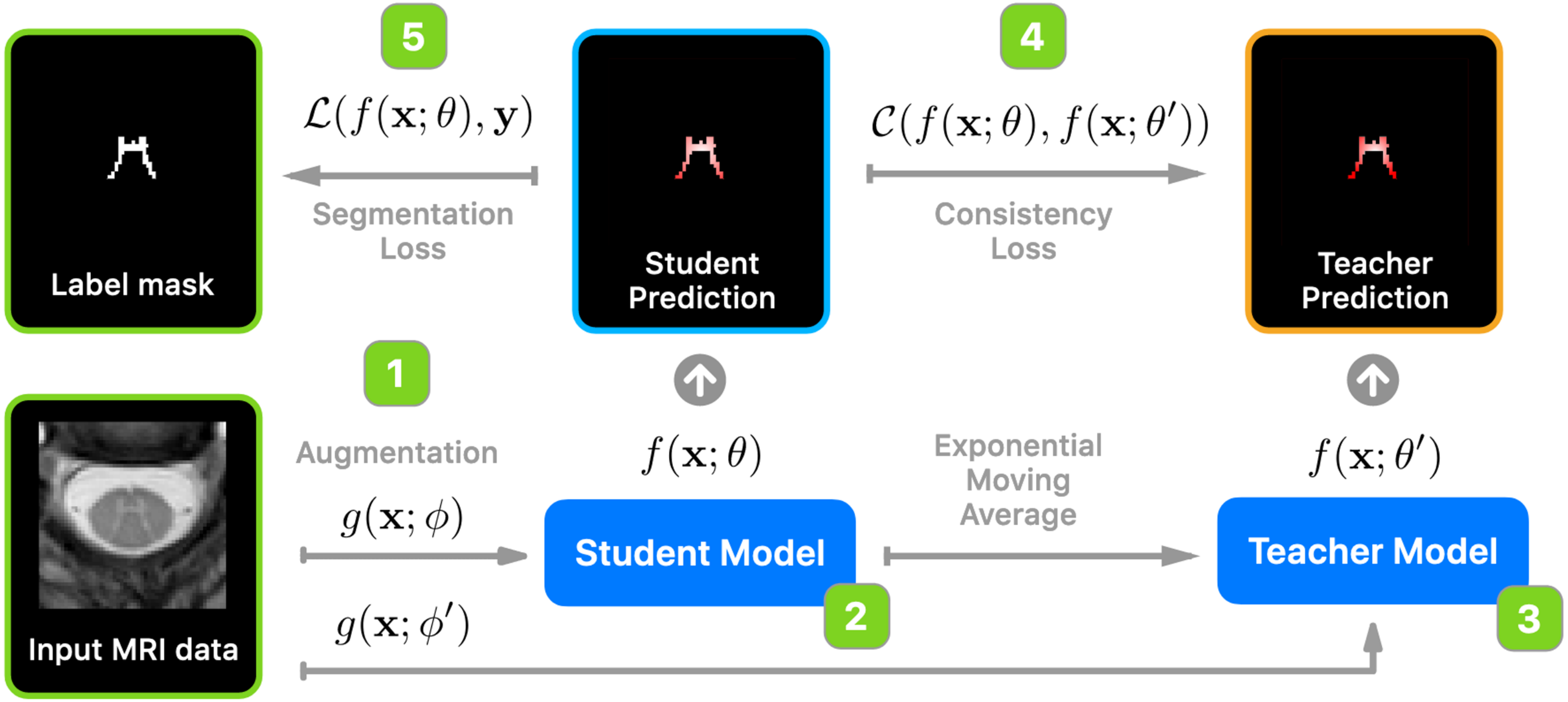


Key idea:

- Consistency between the predictions of a Teacher and a Student network
- The Teacher's weights are an EMA of the Student's at previous training iterations
- *<u>Note</u>*: original Temporal Ensembling computes the EMA on outputs for each sample

SSL methods using consistency regularization

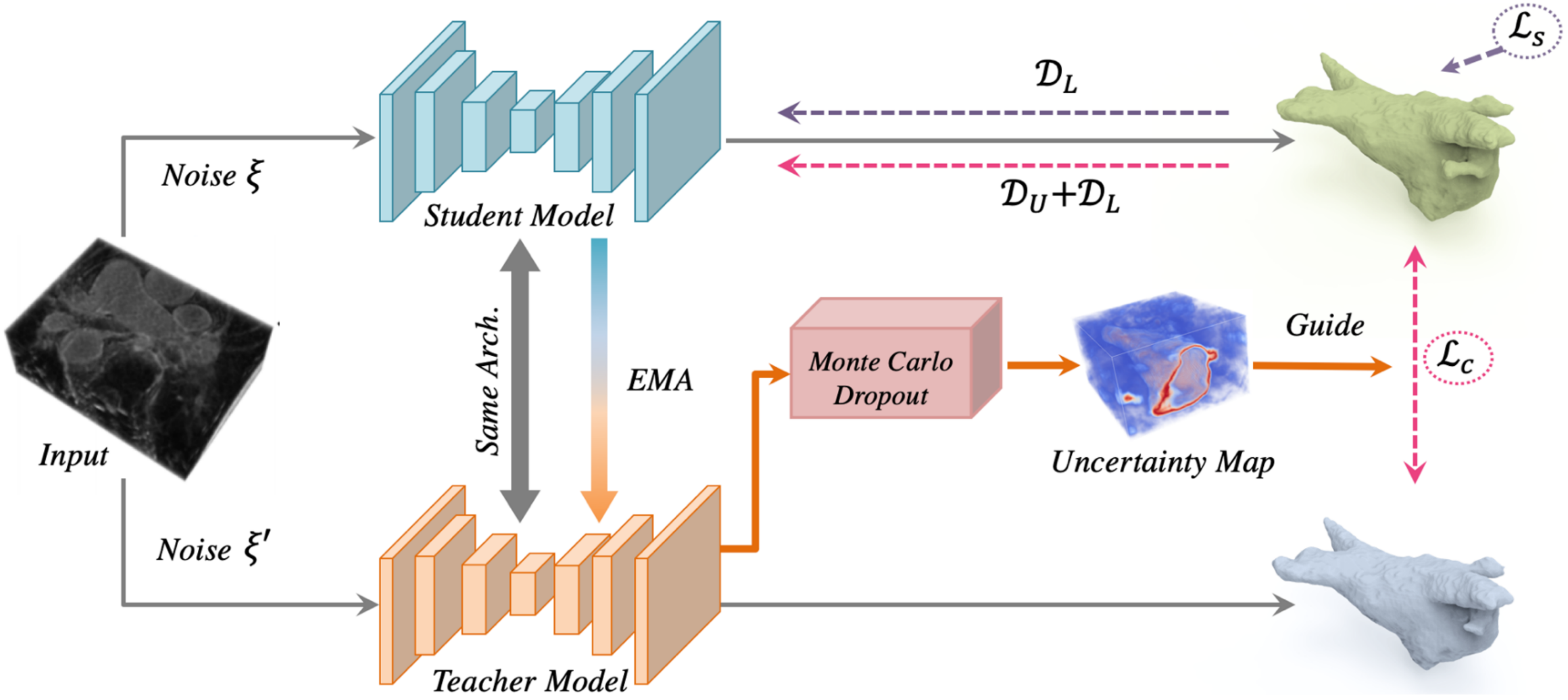
Application of Mean Teacher to segmenting MRI spinal cord gray matter



Perone, C.S. and Cohen-Adad, J. "Deep semi-supervised segmentation with weight-averaged consistency targets." Deep learning in medical image analysis and multimodal learning for clinical decision support, 2019 (*extended version in Neuroimage*)

SSL methods using consistency regularization

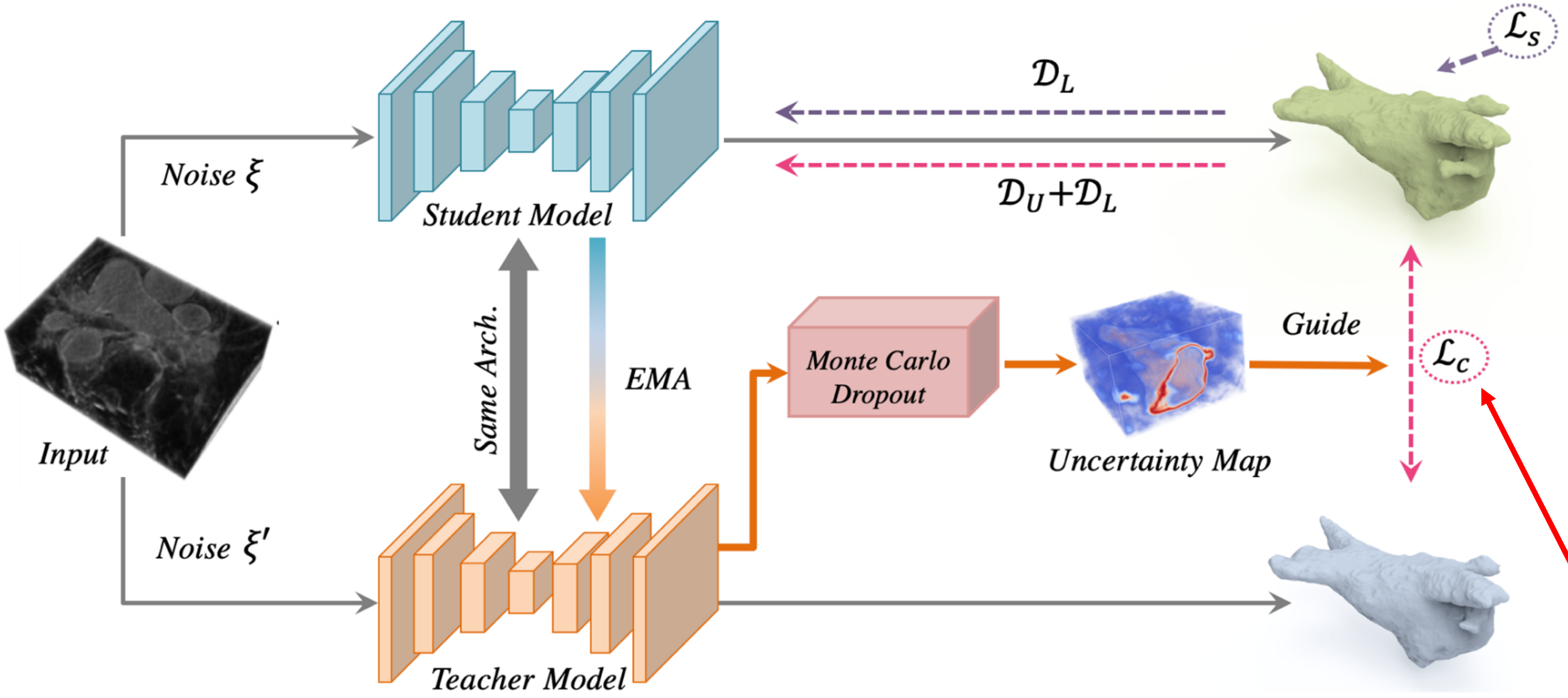
Uncertainty-aware self-ensembling



Yu, L et al.. "Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium segmentation." MICCAI, 2019.

SSL methods using consistency regularization

Uncertainty-aware self-ensembling



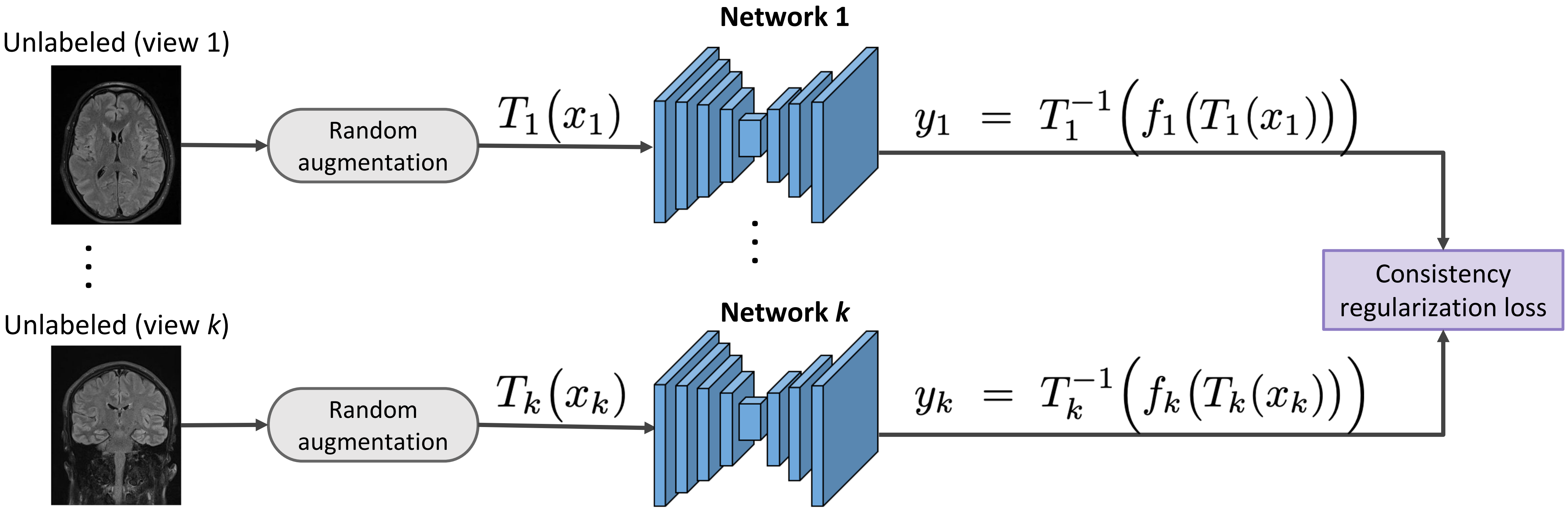
Enforces consistency only in low-uncertainty regions of the image

$$\mathcal{L}_c(f', f) = \frac{\sum_v \mathbb{I}(u_v < H) \left\| f'_v - f_v \right\|^2}{\sum_v \mathbb{I}(u_v < H)},$$

Yu, L et al.. "Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium segmentation." MICCAI, 2019.

SSL methods using consistency regularization

Muti-view co-training

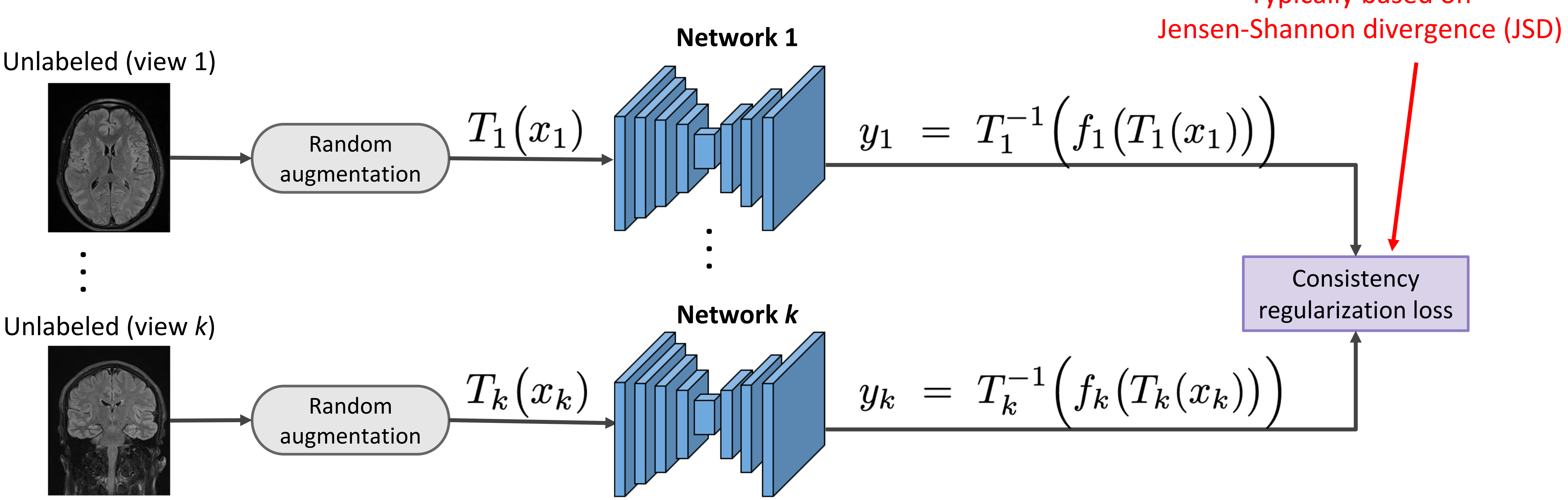


Key idea:

- Supposes the existence of separate, complementary views of the data
- Use high-confidence predictions for a given view as pseudo-labels in other views

SSL methods using consistency regularization

Muti-view co-training

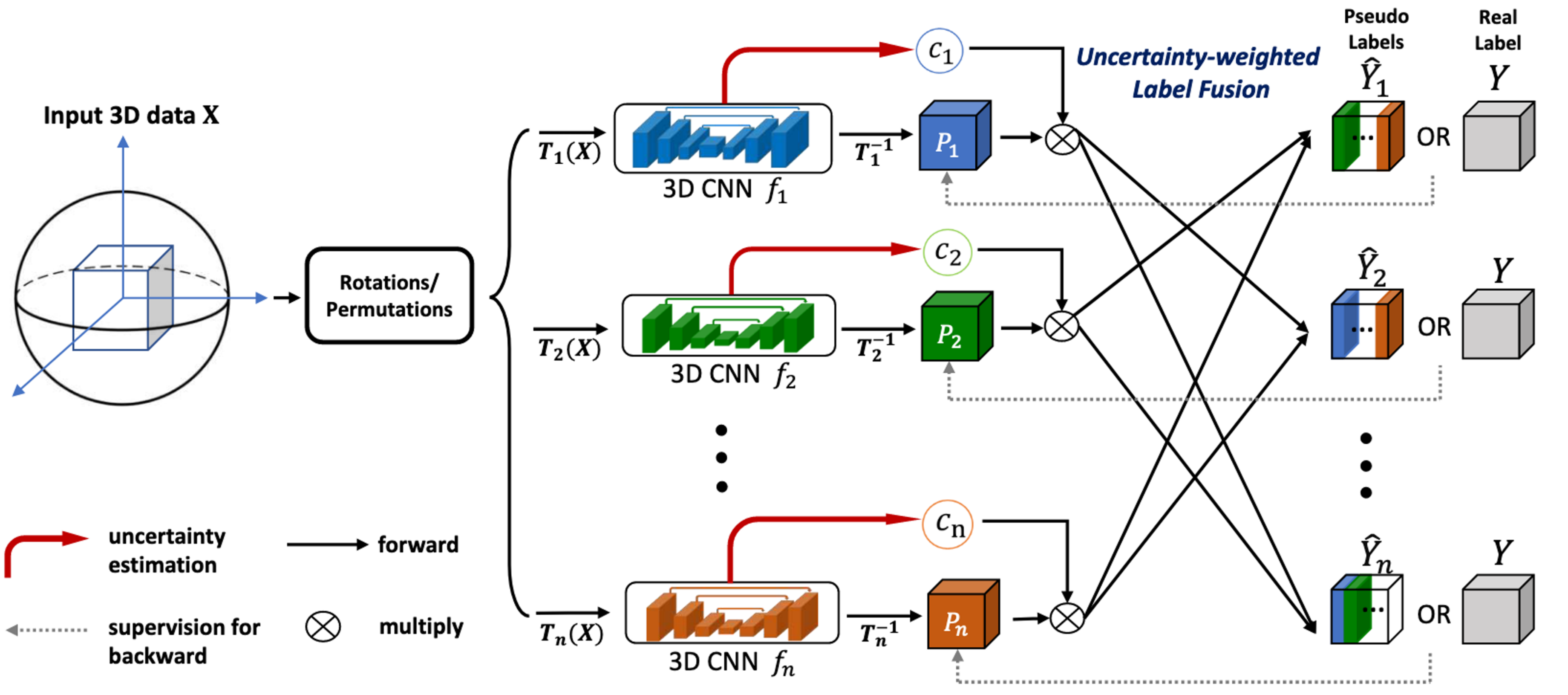


Key idea:

- Supposes the existence of separate, complementary views of the data
- Use high-confidence predictions for a given view as pseudo-labels in other views

SSL methods using consistency regularization

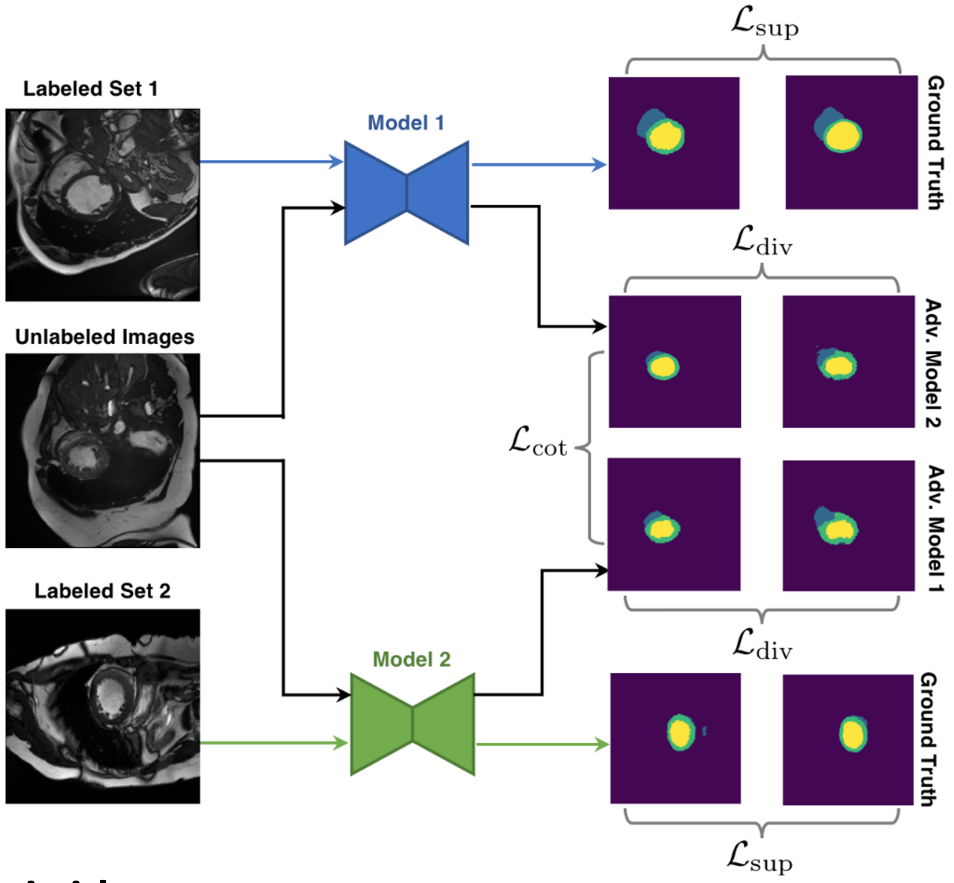
Application of multi-view co-training for pancreas and liver tumor segmentation



Xia, Y., et al. "3D semi-supervised learning with uncertainty-aware multi-view co-training." WACV 2020.

SSL methods using consistency regularization

How to apply co-training without different views ?



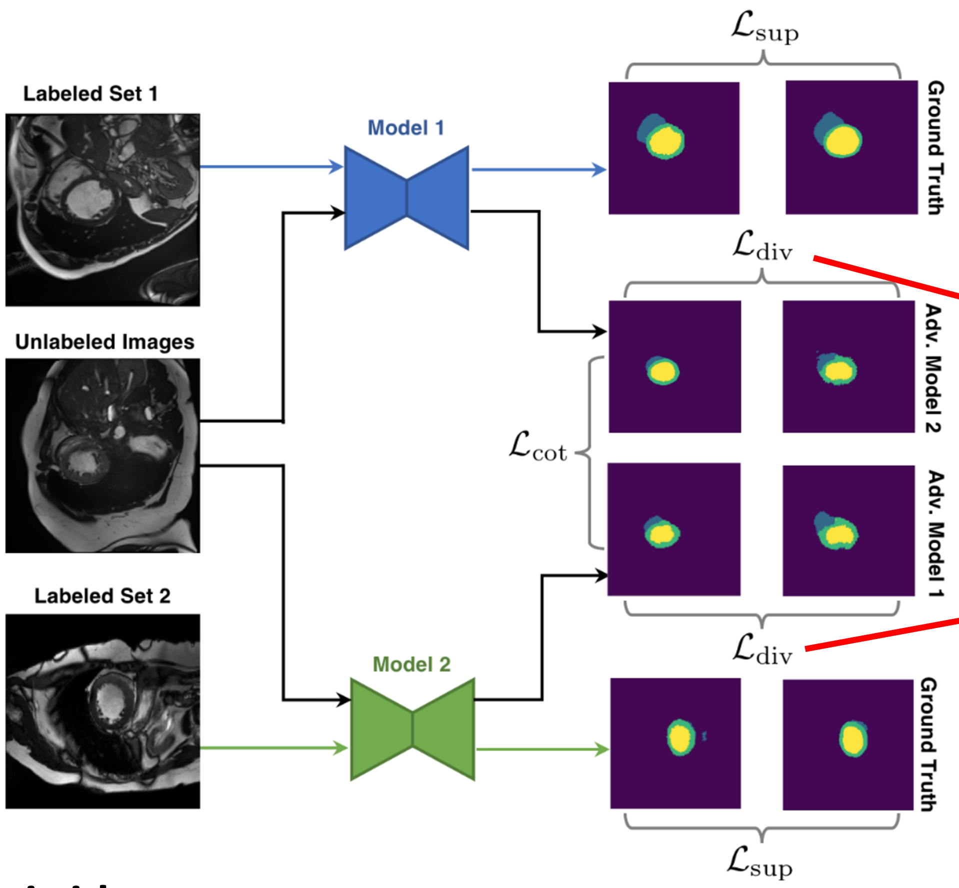
Basic idea:

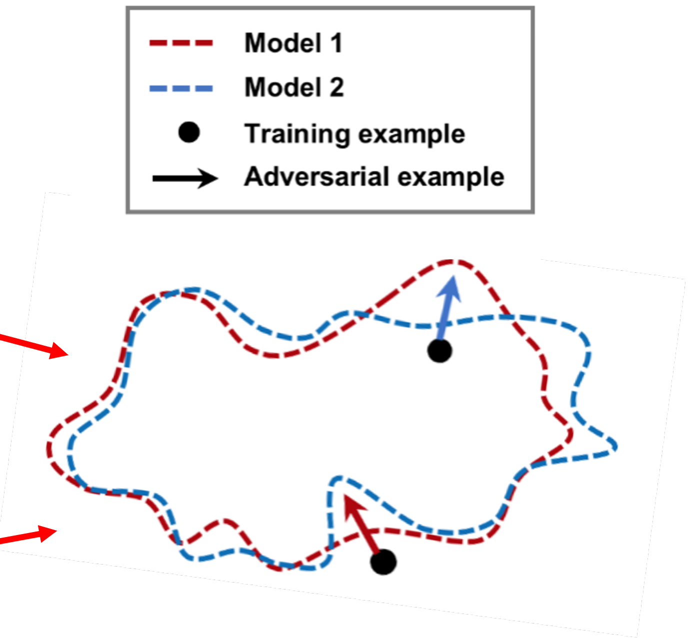
- Use adversarial examples to generate diversity in the representation learned by models

Peng, J. et al.. "Deep co-training for semi-supervised image segmentation." Pattern Recognition, 2020

SSL methods using consistency regularization

How to apply co-training without different views ?





Force a model to follow perturbations along the boundary of other models

Basic idea:

- Use adversarial examples to generate diversity in the representation learned by models

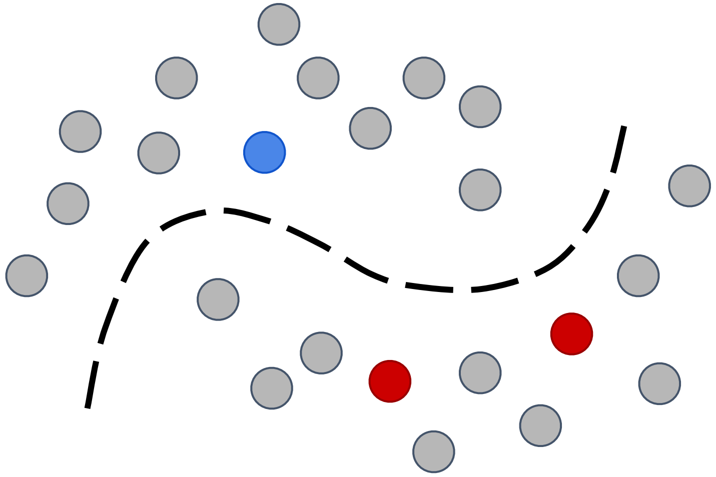
Peng, J. et al.. "Deep co-training for semi-supervised image segmentation." Pattern Recognition, 2020

Unsupervised representation learning

for weakly-supervised segmentation

Unsupervised representation learning (URL)

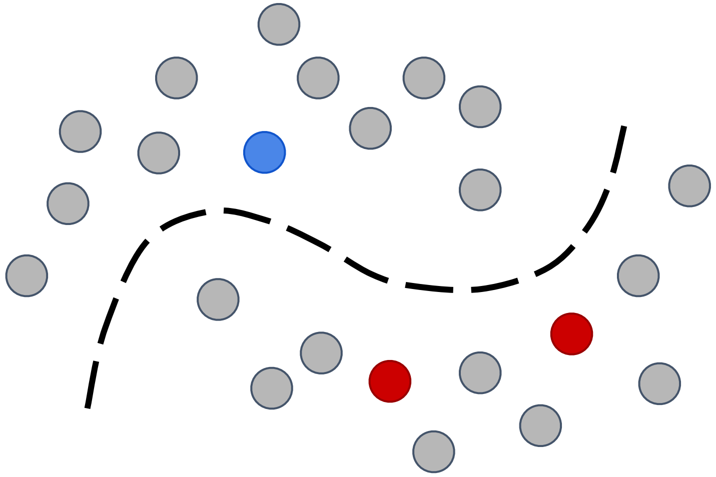
Traditional SSL



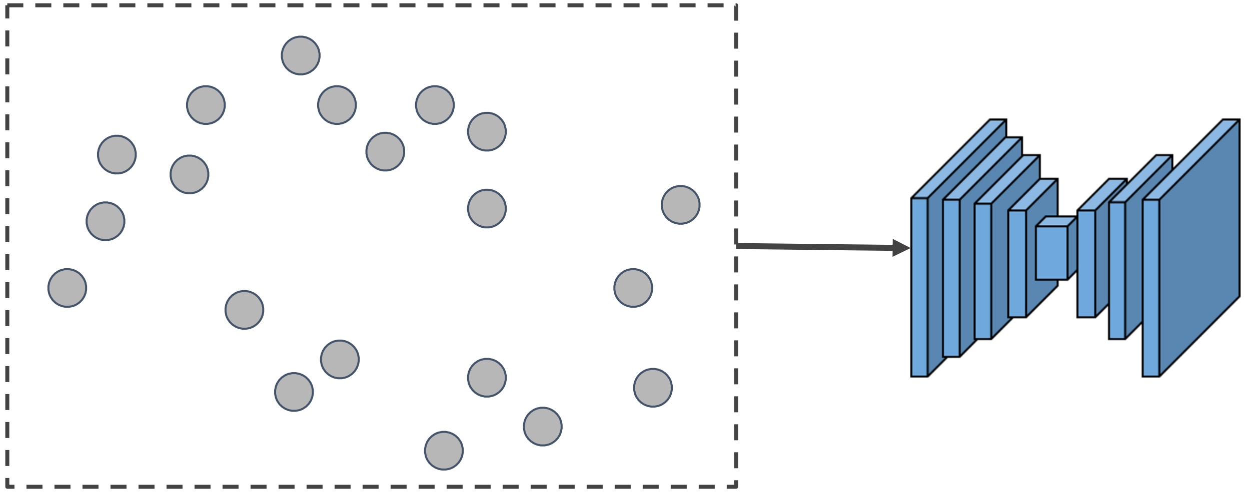
- Train a model simultaneously with both labeled and unlabeled data

Unsupervised representation learning (URL)

Traditional SSL



Unsupervised representation learning

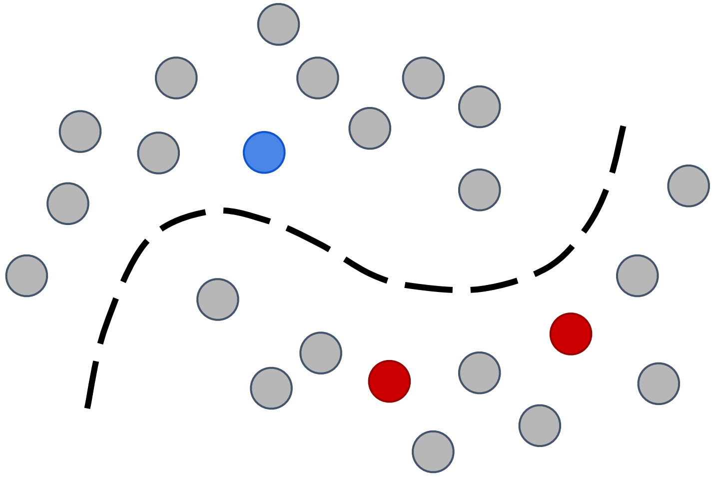


- Train a model simultaneously with both labeled and unlabeled data

- In an upstream step, use <u>only unlabeled data</u> to learn a representation useful to downstream tasks
- **Examples**:
  - Self-supervised learning
  - Contrastive learning

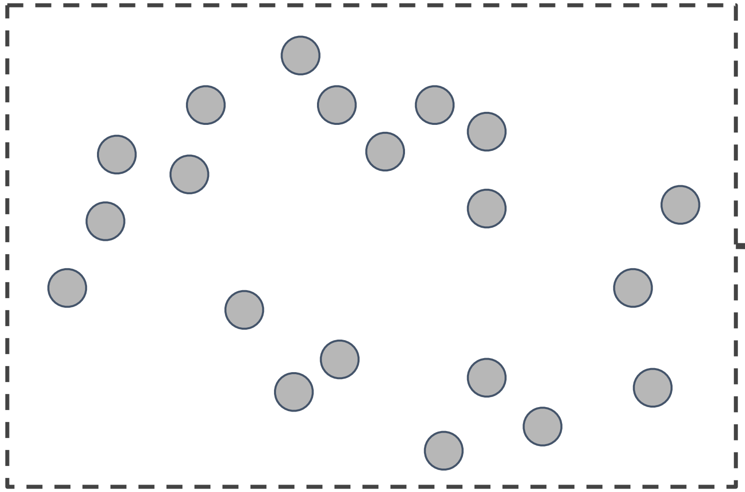
Unsupervised representation learning (URL)

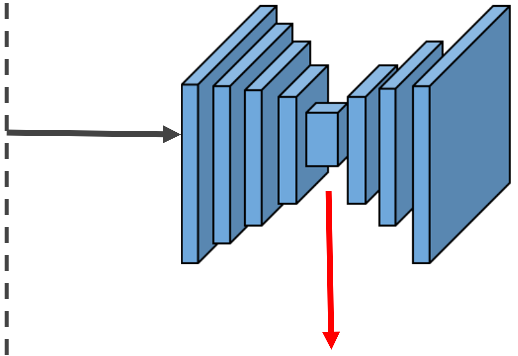
Traditional SSL



- Train a model simultaneously with both labeled and unlabeled data

Unsupervised representation learning

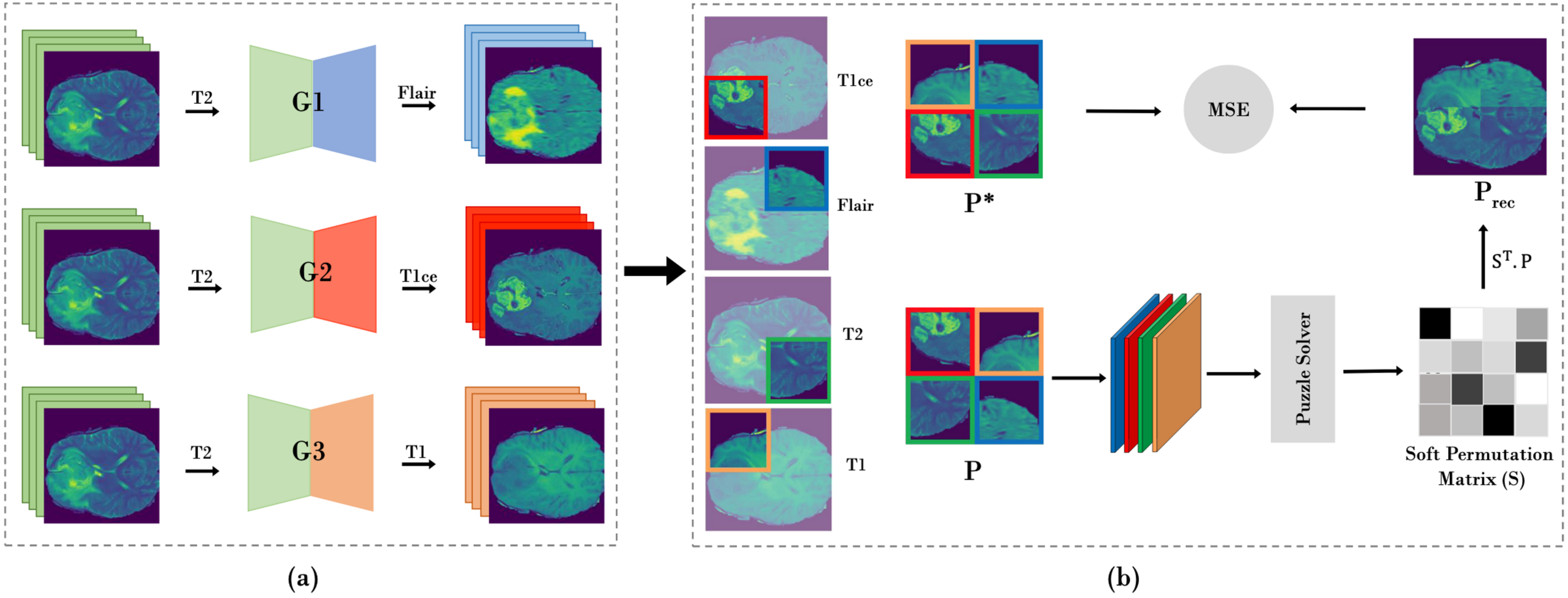




- In an upstream step, use <u>only unlabeled data</u> to learn a representation useful to downstream tasks
- **Examples**:
  - Self-supervised learning
  - Contrastive learning

Approaches for URL

Self-supervised learning:



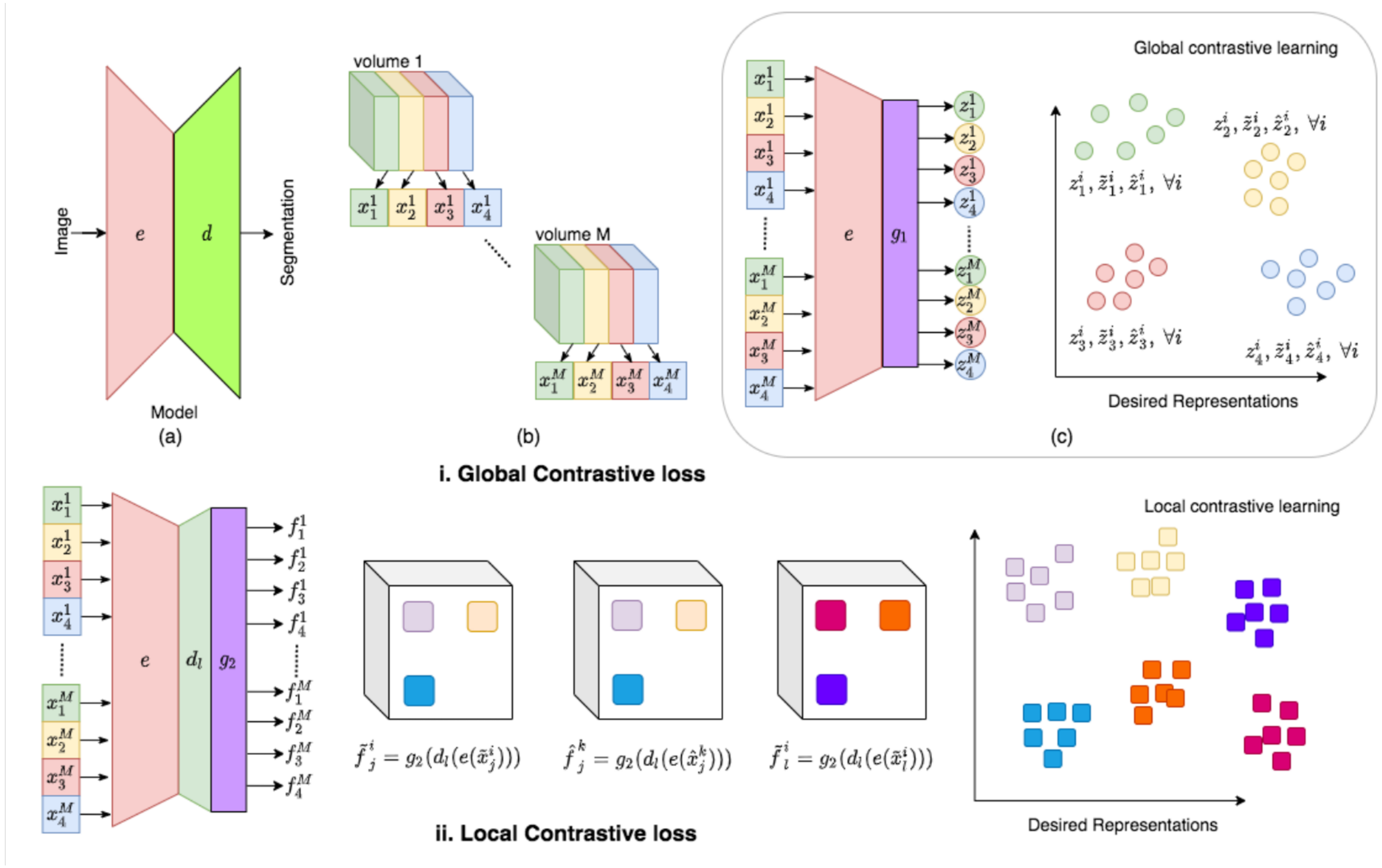
Basic idea:

- Learn to solve a pretext task which does not require annotations
- Example: find the correct order of permuted patches (*see above*)

Taleb, A., et al. "Multimodal self-supervised learning for medical image analysis." arXiv preprint (2019).

Approaches for URL

Contrastive learning:



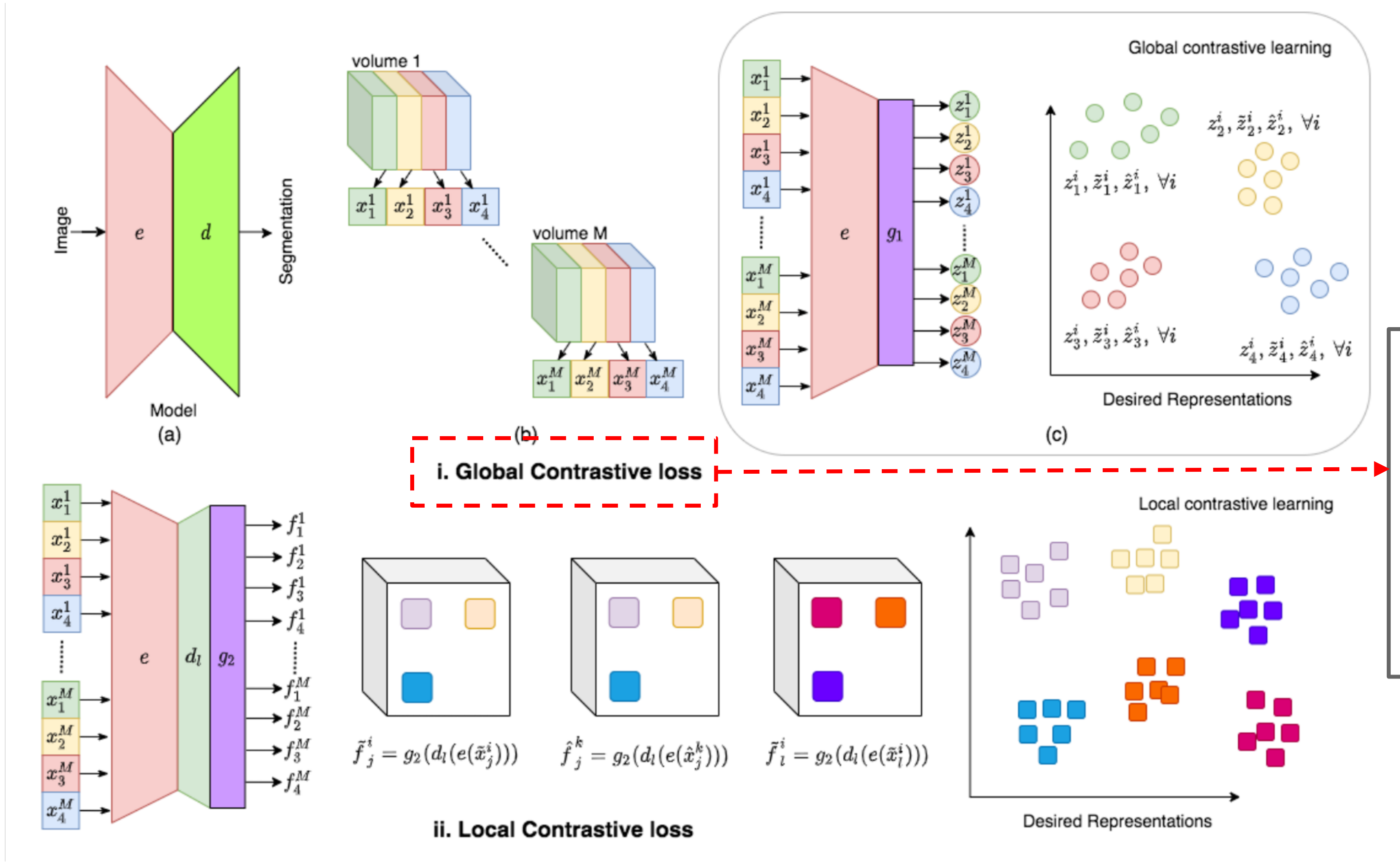
Basic idea:

- Train with pairs of images that <u>match</u> (e.g., same position in volume, same image under different transformations, etc.) or not
- Find a representation that is similar for matching pairs and different for non-matching ones

Chaitanya, K., et al. "Contrastive learning of global and local features for medical image segmentation with limited annotations." NeurIPS 2020.

Approaches for URL

Contrastive learning:



Global contrastive loss

$$-\log \frac{\boxed{e^{\text{sim}(\tilde{z},\hat{z})/\tau}}}{\boxed{e^{\text{sim}(\tilde{z},\hat{z})/\tau}} + \boxed{\sum_{\bar{x}\in\Lambda^-} e^{\text{sim}(\tilde{z},g_1(e(\bar{x})))/\tau}}}$$

$$\tilde{z} = g_1(e(\tilde{x})),\ \hat{z} = g_1(e(\hat{x})).$$

Matching pair

Non-matching pairs

Basic idea:

- Train with pairs of images that <u>match</u> (e.g., same position in volume, same image under different transformations, etc.) or not
- Find a representation that is similar for matching pairs and different for non-matching ones

Chaitanya, K., et al. "Contrastive learning of global and local features for medical image segmentation with limited annotations." NeurIPS 2020.

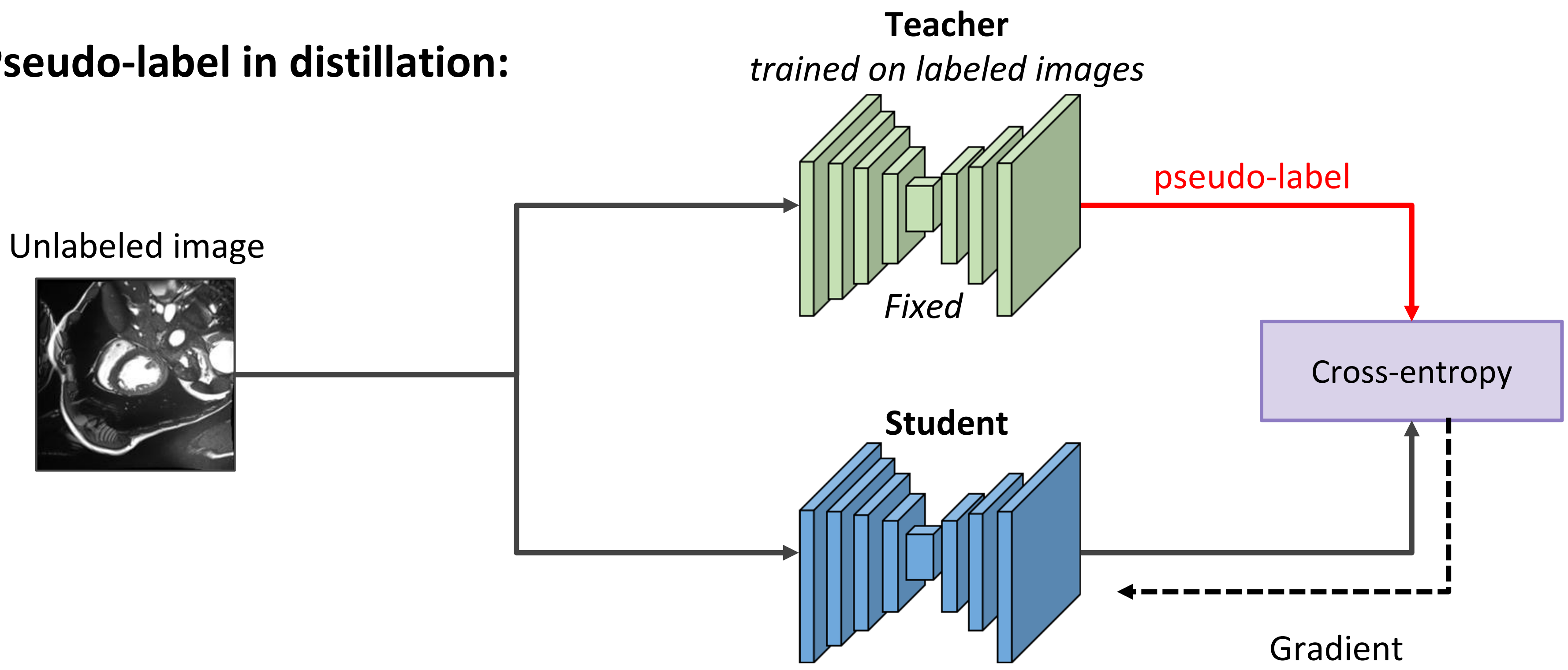
Self-paced learning

for weakly-supervised segmentation

Self-paced learning (SPL)

- Semi-supervised learning methods based on self-training, knowledge distillation or co-training use pseudo-labels on unlabeled images

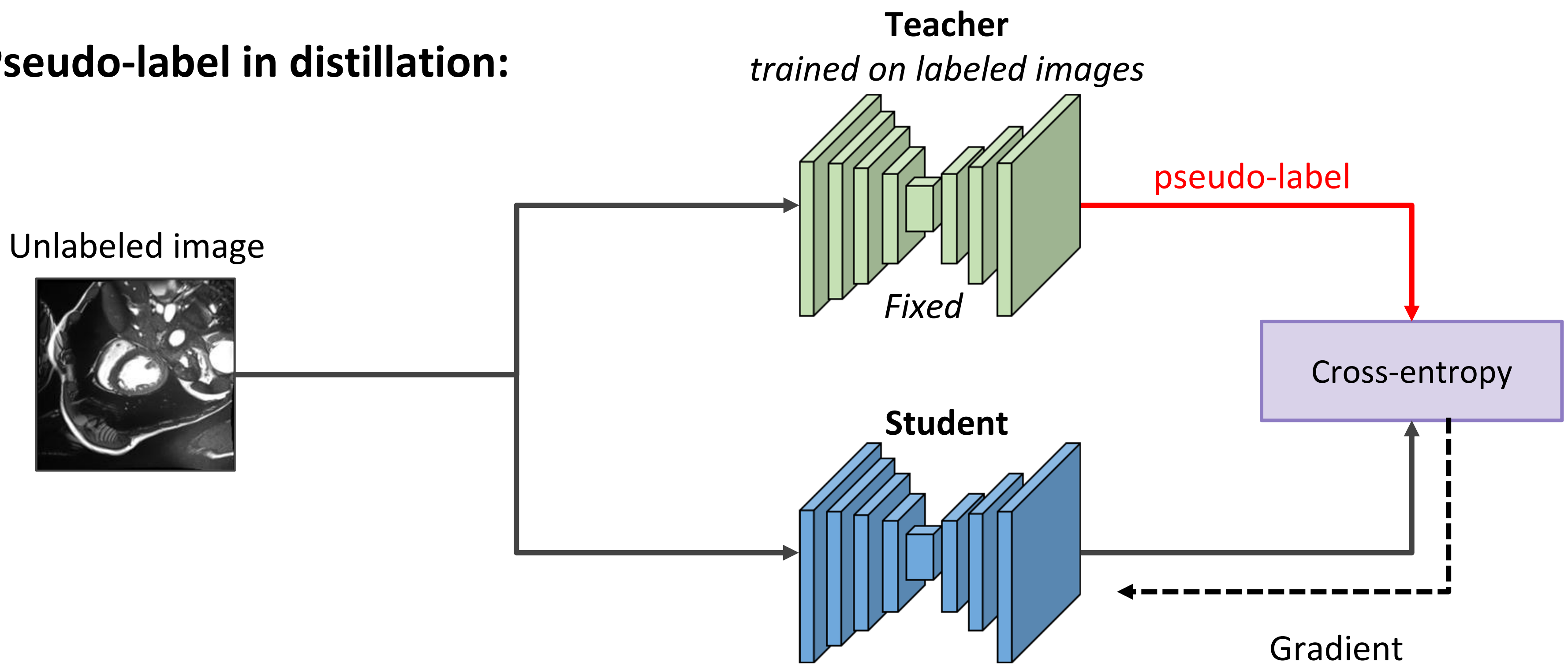
Pseudo-label in distillation:



Self-paced learning (SPL)

- Semi-supervised learning methods based on self-training, knowledge distillation or co-training use <u>pseudo-labels</u> on unlabeled images

Pseudo-label in distillation:



Problem: if pseudo-labels are <u>noisy</u>, this might <u>hurt</u> the student's training

Self-paced learning (SPL)

Standard learning: $$\mathcal{L}(\theta) \;=\; \frac{1}{|\mathcal{D}|} \sum_{(x_i, y_i) \in \mathcal{D}} \ell(f_\theta(x_i), y_i)$$

$$\Downarrow$$

Self-paced learning: $$\mathcal{L}_{\mathrm{sp}}(\theta, w) \;=\; \frac{1}{|\mathcal{D}|} \sum_{(x_i, y_i) \in \mathcal{D}} w_i \, \ell(f_\theta(x_i), y_i) \;+\; \mathcal{R}_\gamma(w)$$

Self-paced learning (SPL)

Ensures that a weight w_i

- Decreases w.r.t. loss $\ell(f_\theta(x_i), y_i)$
- Increases w.r.t. to learning pace γ

Standard learning:

$$\mathcal{L}(\theta) = \frac{1}{|\mathcal{D}|} \sum_{(x_i, y_i) \in \mathcal{D}} \ell(f_\theta(x_i), y_i)$$

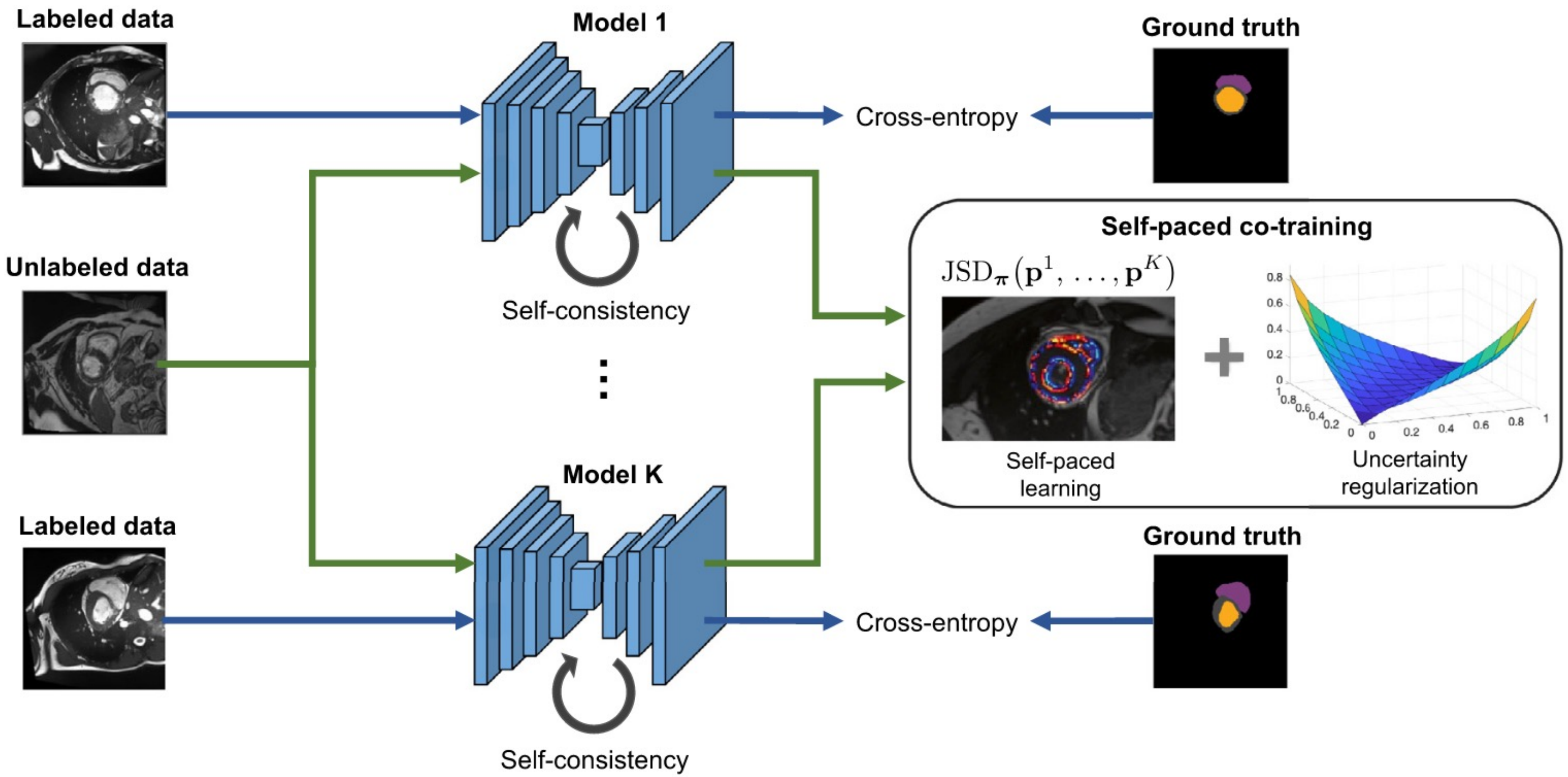
Self-paced learning:

$$\mathcal{L}_{\mathrm{sp}}(\theta, w) = \frac{1}{|\mathcal{D}|} \sum_{(x_i, y_i) \in \mathcal{D}} \boxed{w_i} \ell(f_\theta(x_i), y_i) + \boxed{\mathcal{R}_\gamma(w)}$$

- Controls the usefulness *or* easiness of (x_i, y_i)
- Optimized along with network parameters θ

Applications of SPL

SPL for co-training:



Basic idea:

- Imposes models to make predictions similar to their confidence-weighted average
- The importance of a pixel in the loss is proportional to the total confidence of models for this pixel

Wang, P., et al. " Self-paced and self-consistent co-training for semi-supervised image segmentation." Medical Image Analysis 2020.

Applications of SPL

SPL for contrastive representation learning:

$$\mathcal{L}_{\text{con}} = -\frac{1}{N}\sum_{i=1}^{N}\frac{1}{|\mathcal{P}_i^+|}\sum_{j\in\mathcal{P}_i^+}\log\frac{\exp(z_i^\intercal z_j/\tau)}{\sum_k \exp(z_i^\intercal z_k/\tau)}$$

Importance of a positive pair (i,j)

$$\mathcal{L}_{\text{SP-con}} = -\frac{1}{N}\sum_{i=1}^{N}\frac{1}{|\mathcal{P}_i^+|}\sum_{j\in\mathcal{P}_i^+}\boxed{w_{ij}}\log\frac{\exp(z_i^\intercal z_j/\tau)}{\sum_k \exp(z_i^\intercal z_k/\tau)} + \boxed{\mathcal{R}_\gamma(w)}$$

Peng, J., et al. " Self-Paced Contrastive Learning for Semi-supervised Medical Image Segmentation with Meta-labels." to appear in NeurIPS 2021

Applications of SPL

SPL for contrastive representation learning:

$$\mathcal{L}_{\text{con}} = -\frac{1}{N}\sum_{i=1}^{N}\frac{1}{|\mathcal{P}_i^+|}\sum_{j\in\mathcal{P}_i^+}\log\frac{\exp(z_i^\intercal z_j/\tau)}{\sum_k \exp(z_i^\intercal z_k/\tau)}$$

Importance of a positive pair (*i*,*j*)

$$\mathcal{L}_{\text{SP-con}} = -\frac{1}{N}\sum_{i=1}^{N}\frac{1}{|\mathcal{P}_i^+|}\sum_{j\in\mathcal{P}_i^+}\boxed{w_{ij}}\log\frac{\exp(z_i^\intercal z_j/\tau)}{\sum_k \exp(z_i^\intercal z_k/\tau)} + \boxed{\mathcal{R}_\gamma(w)}$$

Basic idea:

- Meta-labels in constrastive learning (i.e., *positive pairs*) can be noisy.
- Give more importance to high-confidence ones during training.

Peng, J., et al. " Self-Paced Contrastive Learning for Semi-supervised Medical Image Segmentation with Meta-labels." to appear in NeurIPS 2021

Concluding remarks

- Adversarial, consistency regularization and URL methods can help learn segmentation priors without having to explicitly model them
- Enhances learning in a weakly-supervised setting by restricting plausible segmentations of partially-labeled or unlabeled images
- Helps adapt segmentation models across different data domains (e.g., acquisition modality or site)
- Not a silver bullet, can be very challenging at times (e.g., adversarial instability)
- Lots of exciting opportunities for future research

Thank you

Questions ?

https://github.com/LIVIAETS/miccai_2020-weakly_supervised_tutorial

A survey on the presented content is coming very soon!

Ismail Ben Ayed, Associate Professor at ETS Montreal.

ismail.benayed@etsmtl.ca

Christian Desrosiers, Associate Professor at ETS Montreal.

christian.desrosiers@etsmtl.ca

Jose Dolz, Assistant Professor at ETS Montreal

jose.dolz@etsmtl.ca

Hoel Kervadec, PhD. ETS Montreal

hoel@kervadec.science

References

[1] Bateson M, Kervadec H, Dolz J, Lombaert H, Ben Ayed I. Constrained domain adaptation for segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2019 Oct 13 (pp. 326-334).

[2] Belkin M, Niyogi P, Sindhwani V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. Journal of machine learning research. 2006;7(Nov):2399-434.

[3] Ben Ayed I, Wang M, Miles B, Garvin GJ. TRIC: Trust region for invariant compactness and its application to abdominal aorta segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2014 Sep 14 (pp. 381-388).

[4] Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends® in Machine learning. 2011 Jul 26;3(1):1-22.

[5] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2001 Nov;23(11):1222-39.

[6] Carass A, Cuzzocreo JL, Han S, Hernandez-Castillo CR, Rasser PE, Ganz M, Beliveau V, Dolz J, Ben Ayed I, Desrosiers C, Thyreau B et al. Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images. NeuroImage. 2018 Dec 1;183:150-72.

[7] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE transactions on pattern analysis and machine intelligence. 2017 Apr 27;40(4):834-48.

[8] Chen M, Artières T, Denoyer L. Unsupervised Object Segmentation by Redrawing. In Advances in neural information processing systems 2019.

[9] Cordts M, Omran M, Ramos S, Rehfeld T, Enzweiler M, Benenson R, Franke U, Roth S, Schiele B. The cityscapes dataset for semantic urban scene understanding. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 3213-3223).

[10] Dolz J, Desrosiers C, Ben Ayed I. 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. NeuroImage. 2018 Apr 15;170:456-70.

[11] Dou Q, Ouyang C, Chen C, Chen H, Glocker B, Zhuang X, Heng PA. PnP-AdaNet: Plug-and-play adversarial domain adaptation network with a benchmark at cross-modality cardiac segmentation. IEEE Access, 7:99065–99076, 2019.

[12] Fechter T, Adebahr S, Baltas D, Ben Ayed I, Desrosiers C, Dolz J. Esophagus segmentation in CT via 3D fully convolutional neural network and random walk. Medical physics. 2017 Dec 1;44(12):6341-52.

[13] Ganaye PA, Sdika M, Benoit-Cattin H. Semi-supervised learning for segmentation under semantic constraint. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2018 Sep 16 (pp. 595-602).

[14] Ghosh A, Kulharia V, Namboodiri VP, Torr PH, Dokania PK. Multi-agent diverse generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018 (pp. 8513-8521).

[15] Grandvalet Y, Bengio Y. Semi-supervised learning by entropy minimization. In Advances in neural information processing systems 2005 (pp. 529-536).

[16] Hoffman J, Tzeng E, Park T, Zhu JY, Isola P, Saenko K, Efros AA, Darrell T. Cycada: Cycle-consistent adversarial domain adaptation. In International Conference on Machine Learning (ICML). 2018

References

[17] Hung WC, Tsai YH, Liou YT, Lin YY, Yang MH. Adversarial learning for semi-supervised semantic segmentation. In the British Machine Vision Conference (BMVC) 2018.

[18] Ji Z, Shen Y, Ma C, Gao M. Scribble-Based Hierarchical Weakly Supervised Learning for Brain Tumor Segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2019 Oct 13 (pp. 175-183).

[19] Jia Z, Huang X, Eric I, Chang C, Xu Y. Constrained deep weak supervision for histopathology image segmentation. IEEE Transactions on Medical Imaging. 2017 Jul 7;36(11):2376-88.

[20] Kervadec H, Dolz J, Tang M, Granger E, Boykov Y, Ben Ayed I. Constrained-CNN losses for weakly supervised segmentation. Medical image analysis. 2019 May 1;54:88-99.

[21] Kervadec H, Dolz J, Yuan J, Desrosiers C, Granger E, Ben Ayed I. Constrained Deep Networks: Lagrangian Optimization via Log-Barrier Extensions. arXiv preprint arXiv:1904.04205. 2019 Apr 8.

[22] Kervadec H, Dolz J, Granger E, Ben Ayed I. Curriculum semi-supervised segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2019 Oct 13 (pp. 568-576).

[23] Krähenbühl P, Koltun V. Efficient inference in fully connected crfs with gaussian edge potentials. In Advances in neural information processing systems 2011 (pp. 109-117).

[24] Krause A, Perona P, Gomes RG. Discriminative clustering by regularized information maximization. In Advances in neural information processing systems 2010 (pp. 775-783).

[25] Larrazabal AJ, Martinez C, Ferrante E. Anatomical Priors for Image Segmentation via Post-Processing with Denoising Autoencoders. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2019.

[26] Lee DH. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In Workshop on Challenges in Representation Learning, ICML 2013 Jun 21 (Vol. 3, p. 2).

[27] Lin D, Dai J, Jia J, He K, Sun J. Scribblesup: Scribble-supervised convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016 (pp. 3159-3167).

[28] Marin D, Tang M, Ayed IB, Boykov Y. Beyond Gradient Descent for Regularized Segmentation Losses. In IEEE conference on Computer Vision and Pattern Recognition (CVPR) 2019.

[29] Mondal AK, Dolz J, Desrosiers C. Few-shot 3D multi-modal medical image segmentation using generative adversarial learning. arXiv preprint arXiv:1810.12241. 2018 Oct 29.

[30] Mondal AK, Agarwal A, Dolz J, Desrosiers C. Revisiting CycleGAN for semi-supervised segmentation. arXiv preprint arXiv:1908.11569. 2019 Aug 30.

[31] Njeh I, Sallemi L, Ayed IB, Chtourou K, Lehericy S, Galanaud D, Hamida AB. 3D multimodal MRI brain glioma tumor and edema segmentation: a graph cut distribution matching approach. Computerized Medical Imaging and Graphics. 2015 Mar 1;40:108-19.

[32] Oktay O, Ferrante E, Kamnitsas K, Heinrich M, Bai W, Caballero J, Cook SA, De Marvao A, Dawes T, O'Regan DP, Kainz B. Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. IEEE transactions on medical imaging. 2017 Sep 26;37(2):384-95.

[33] Pathak D, Krahenbuhl P, Darrell T. Constrained convolutional neural networks for weakly supervised segmentation. In Proceedings of the IEEE international conference on computer vision 2015 (pp. 1796-1804).

[34] Painchaud N, Skandarani Y, Judge T, Bernard O, Lalande A, Jodoin PM. Cardiac MRI Segmentation with Strong Anatomical Guarantees. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2019 Oct 13 (pp. 632-640).

References

[35] Qu H, Wu P, Huang Q, Yi J, Riedlinger GM, De S, Metaxas DN. Weakly Supervised Deep Nuclei Segmentation using Points Annotation in Histopathology Images. In International Conference on Medical Imaging with Deep Learning 2019 May 24 (pp. 390-400).

[36] Ravishankar H, Venkataramani R, Thiruvenkadam S, Sudhakar P, Vaidya V. Learning and incorporating shape models for semantic segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2017 Sep 10 (pp. 203-211).

[37] Souly N, Spampinato C, Shah M. Semi supervised semantic segmentation using generative adversarial network. InProceedings of the IEEE International Conference on Computer Vision 2017 (pp. 5688-5696).

[38] Tang M, Perazzi F, Djelouah A, Ben Ayed I, Schroers C, Boykov Y. On regularized losses for weakly-supervised CNN segmentation. In Proceedings of the European Conference on Computer Vision (ECCV) 2018 (pp. 507-522).

[39] Tsai YH, Hung WC, Schulter S, Sohn K, Yang MH, Chandraker M. Learning to adapt structured output space for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018 (pp. 7472-7481).

[40] Vu TH, Jain H, Bucher M, Cord M, Pérez P. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2019 (pp. 2517-2526).

[41] Wang L, Nie D, Li G, Puybareau É, Dolz J, Zhang Q, Wang F, Xia J, Wu Z, Chen J, Thung KH et al. Benchmark on automatic 6-month-old infant brain segmentation algorithms: the iSeg-2017 challenge. IEEE transactions on medical imaging. 2019 Feb 27.

[42] Weston J, Ratle F, Mobahi H, Collobert R. Deep learning via semi-supervised embedding. In Neural Networks: Tricks of the Trade 2012 (pp. 639-655). Springer, Berlin, Heidelberg.

[43] Zhang Y, Yang L, Chen J, Fredericksen M, Hughes DP, Chen DZ. Deep adversarial networks for biomedical image segmentation utilizing unannotated images. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2017 Sep 10 (pp. 408-416).

[44] Zhang Y, David P, Gong B. Curriculum domain adaptation for semantic segmentation of urban scenes. In Proceedings of the IEEE International Conference on Computer Vision 2017 (pp. 2020-2030).

[45] Zhang Y, David P, Foroosh H, Gong B. A Curriculum Domain Adaptation Approach to the Semantic Segmentation of Urban Scenes. IEEE transactions on pattern analysis and machine intelligence. 2019 Mar 6.

[46] Zhou Y, Li Z, Bai S, Wang C, Chen X, Han M, Fishman E, Yuille A. Prior-aware Neural Network for Partially-Supervised Multi-Organ Segmentation. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) 2019

[47] Zhu Q, Du B, Yan P. Boundary-weighted Domain Adaptive Neural Network for Prostate MR Image Segmentation. arXiv preprint arXiv:1902.08128. 2019 Feb 21.

[48] Zhu X, Ghahramani Z, Lafferty JD. Semi-supervised learning using gaussian fields and harmonic functions. In Proceedings of the 20th International conference on Machine learning (ICML-03) 2003 (pp. 912-919).

[49] Zou Y, Yu Z, Vijaya Kumar BV, Wang J. Unsupervised domain adaptation for semantic segmentation via class-balanced self-training. In Proceedings of the European Conference on Computer Vision (ECCV) 2018 (pp. 289-305).

[50] Zou Y, Yu Z, Liu X, Kumar BV, Wang J. Confidence Regularized Self-Training. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) 2019