

# Augmented Cyclic Consistency Regularization for Unpaired Image-to-Image Translation

Takehiko Ohkawa<sup>1</sup>, Naoto Inoue<sup>1</sup>, Hirokatsu Kataoka<sup>2</sup>, Nakamasa Inoue<sup>3</sup>

<sup>1</sup>The University of Tokyo, <sup>2</sup>AIST, <sup>3</sup>Tokyo Institute of Technology



Tokyo Tech

## Purpose

- Stabilizing the training of image-to-image translation



## Approach

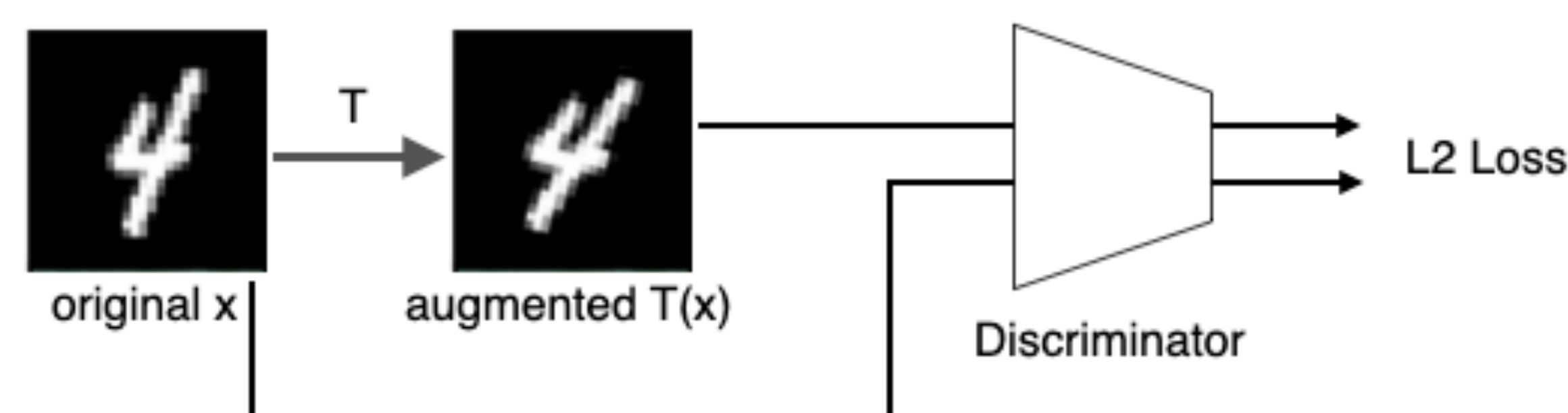
- Consistency regularization on real, fake, and recon samples

## Related work

- CycleGAN [J.Y Zhu+, ICCV'17]
- Consistency regularization [M.Sajjadi+, NeurIPS'16]

$$L_{cr} = \|D(x) - D(T(x))\|^2$$

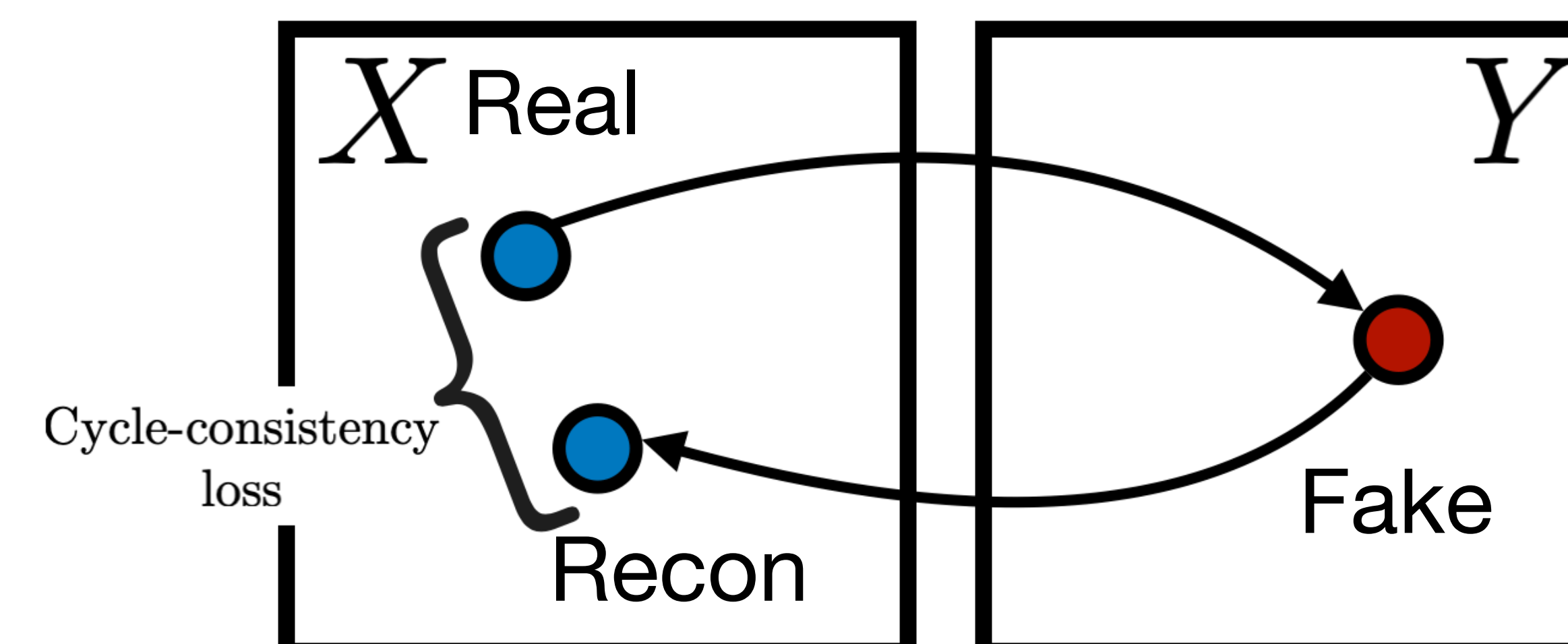
T: data augmentation



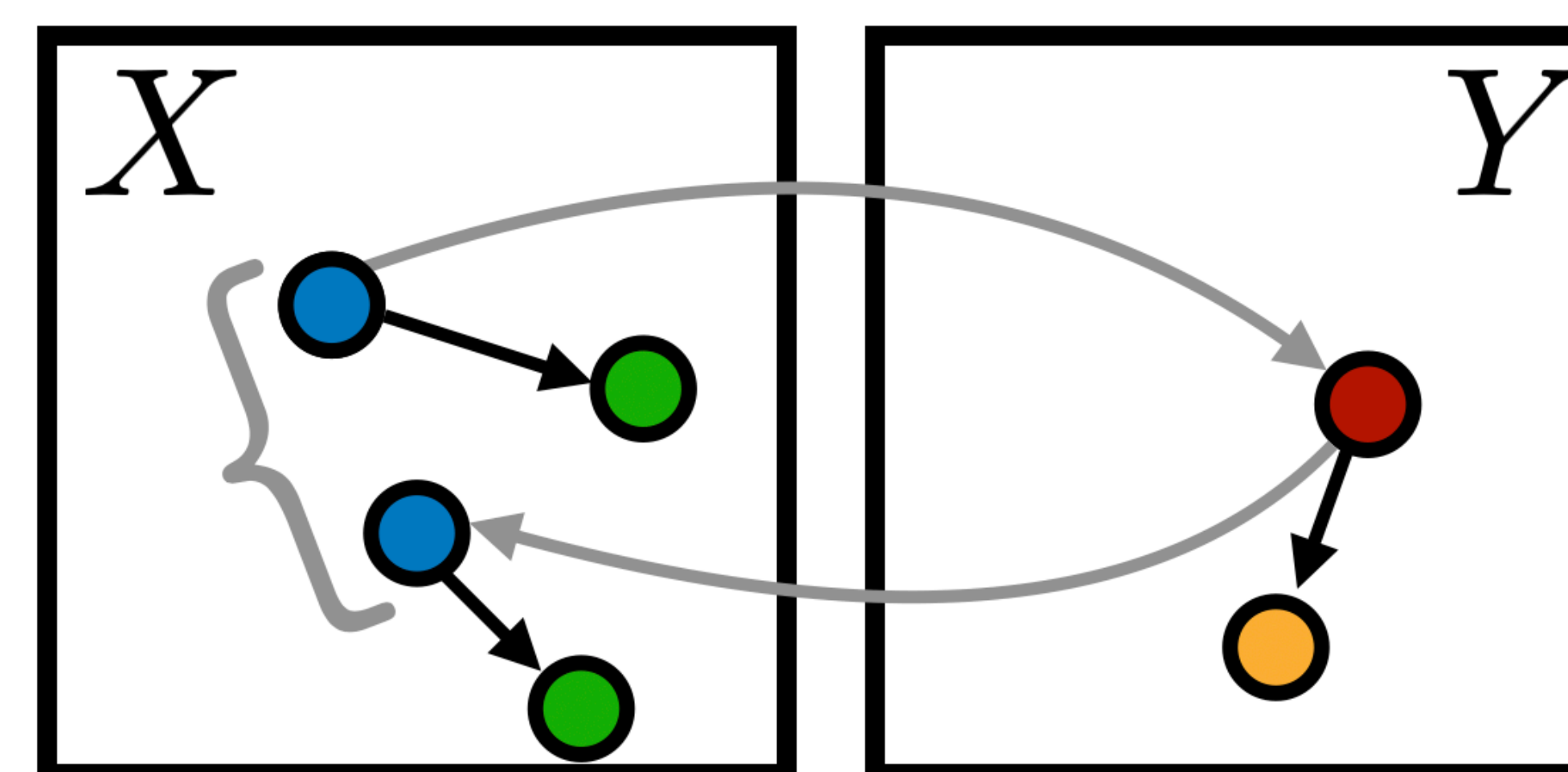
- CR-GAN [H Zhang+, ICLR'20]

## Proposed method

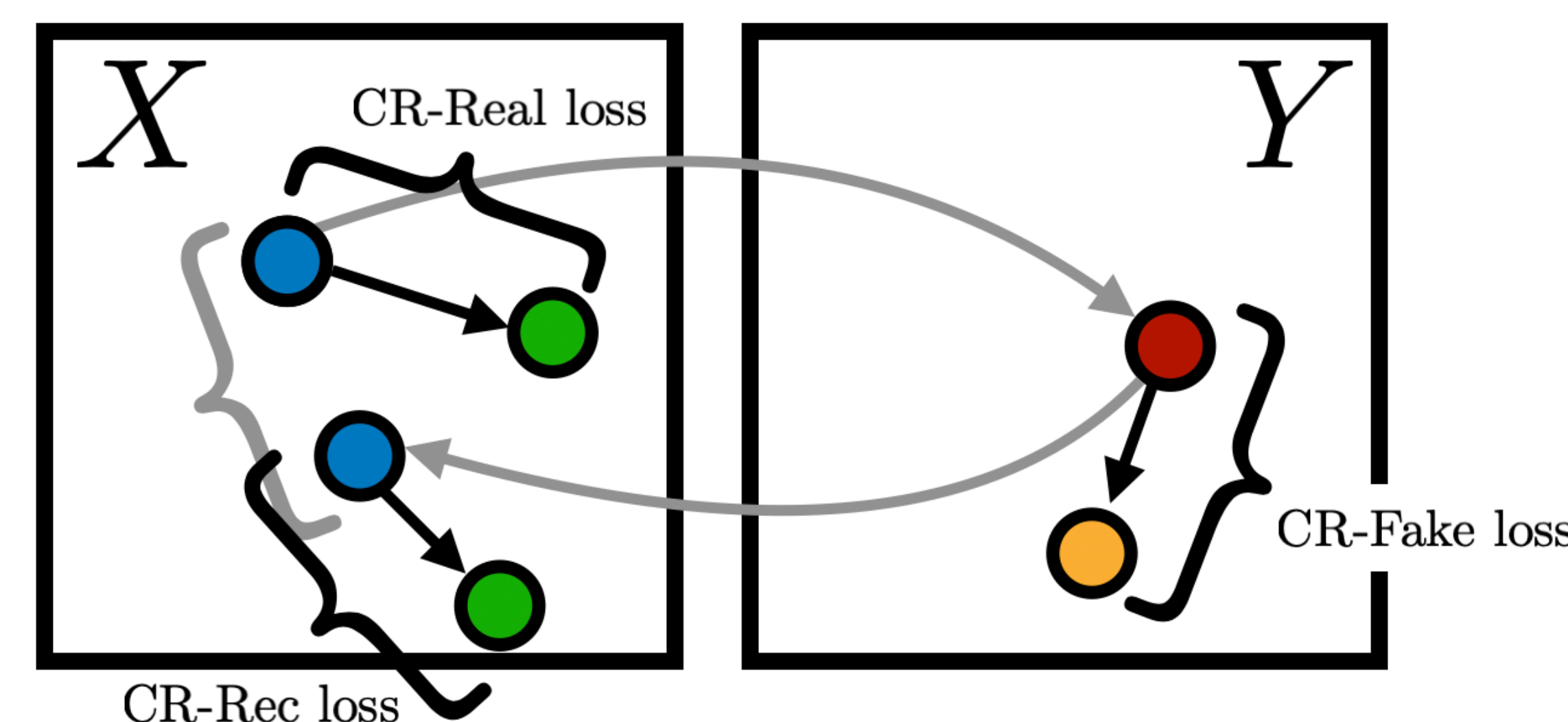
### Step1: CycleGAN



### Step2: Data augmentation for each sample



### Step3: Consistency regularization for each pair of x and T(x) (x=real, fake, recon)



## Results

### Digits



Model	MNIST -> MNIST-M	MNIST-M -> MNIST
CycleGAN	97.7	92.2
CR-CycleGAN	97.7	94.3
Ours	<b>98.0</b>	<b>94.5</b>

Model	MNIST -> SVHN	SVHN -> MNIST
CycleGAN	47.1	28.2
CR-CycleGAN	43.7	29.6
Ours	<b>51.0</b>	<b>31.9</b>

Accuracy ▲ (%)

### Maps ↔ Aerial photo



Model	Map -> Photo	Photo -> Map
CycleGAN	15.9	<b>2.3</b>
CR-CycleGAN	12.0	<b>2.3</b>
Ours	<b>11.7</b>	<b>2.3</b>

MSE (with GT) \* 100 ▼

### Cityscapes labels ↔ Photo



Model	Label -> Photo	Photo -> Label
CycleGAN	72.8	5.7
CR-CycleGAN	62.5	6.5
Ours	<b>56.1</b>	<b>4.0</b>

MSE (with GT) \* 100 ▼