Addressing Challenging Place Recognition Tasks using Generative Adversarial Networks

分享人：徐健

1.任务

- 对于无人驾驶等任务，某些时候需要进行场景识别（place recognition）来确定车辆所处于的位置。
- 对于长时间的定位任务，由于季节，天气，光照的变化，所以进行long-term place recognition比较困难。例如，根据冬天建立的地图，在夏天拍摄的图片中进行匹配，以确定车辆位置。

解决思想

- 使用CycleGAN, 直接将冬天的图片转换为夏天的图片， 然后进行图像匹配。

Fig. 1: Learned image translations: (**Rows 1,3**): Original images (**Rows 2,4**): Winter to summer and summer to winter from the corresponding images in rows 1 and 3. Generative Adversarial Networks are able to map from one domain to another with high visual fidelity. Discriminators in each domain learn features specific to the domain that are useful for place recognition tasks. Instead of matching summer to winter (rows:1,3), which is a difficult perception task, summer images can be reliably matched to generated-summer (rows:2,3).

基础知识：

CycleGAN(Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks，ICCV2017)

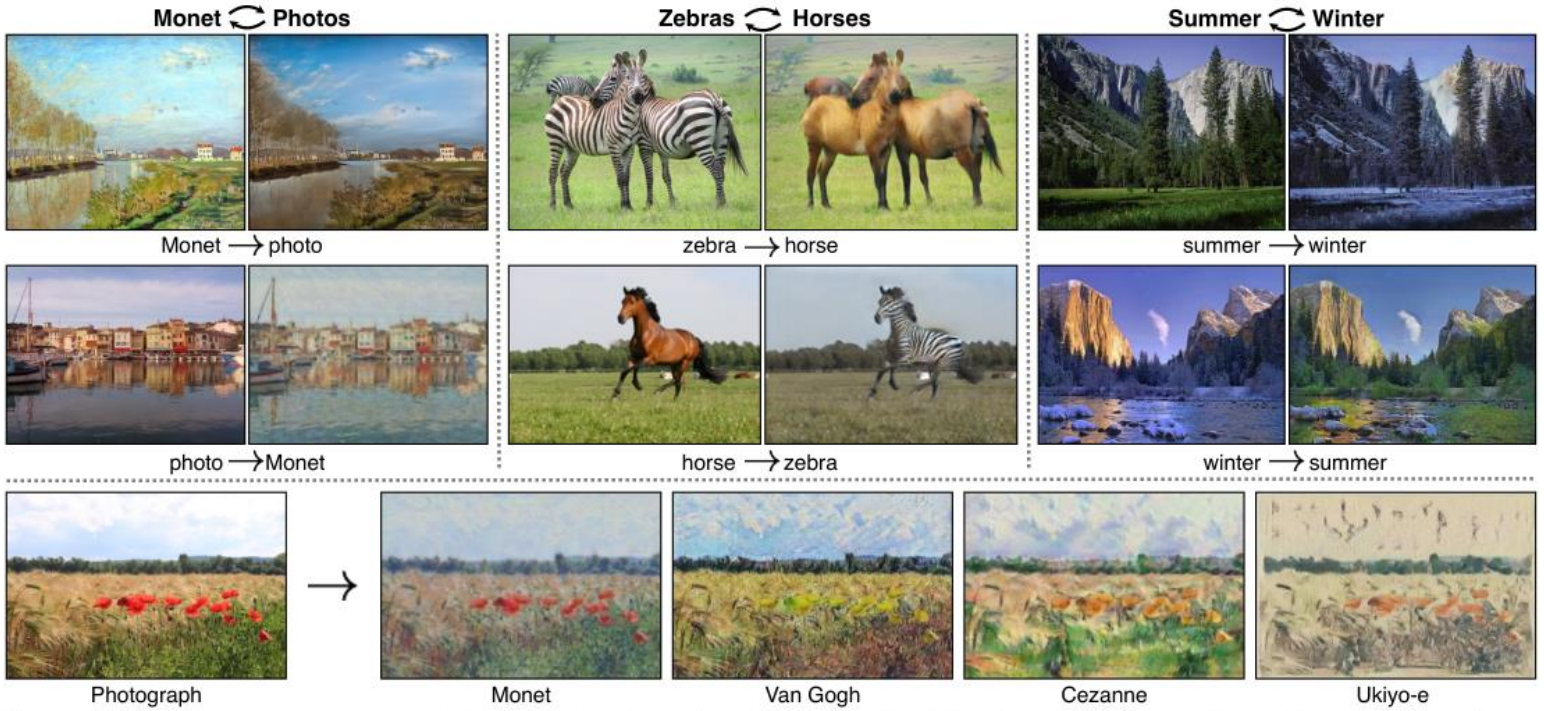
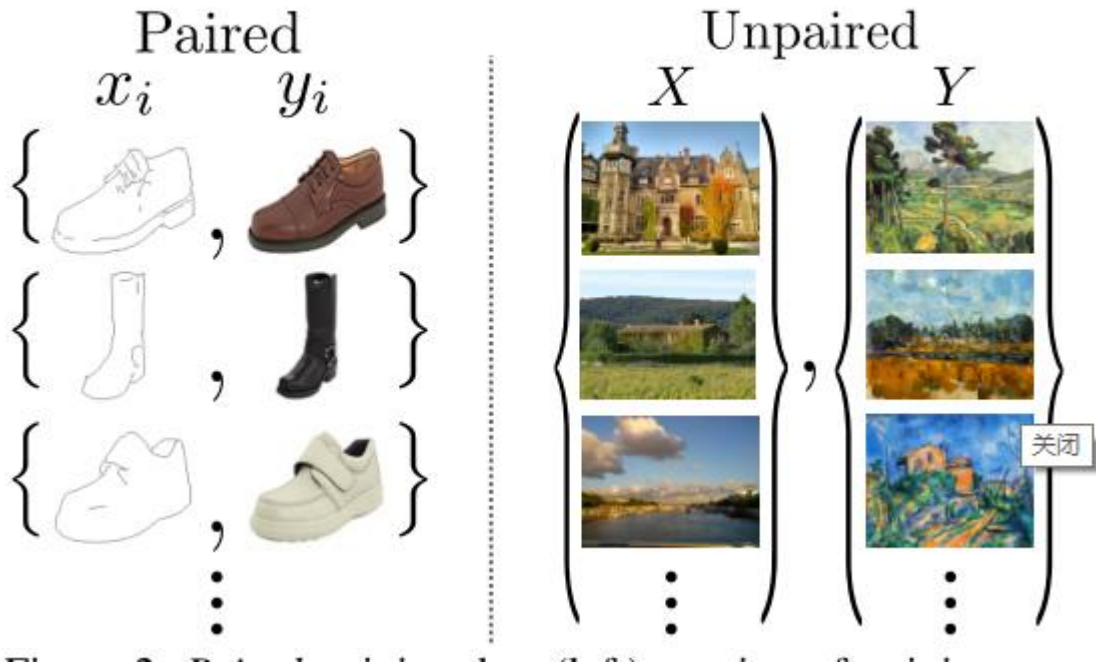


Figure 1: Given any two unordered image collections X and Y, our algorithm learns to automatically "translate" an image from one into the other and vice versa: *(left)* Monet paintings and landscape photos from Flickr; *(center)* zebras and horses from ImageNet; *(right)* summer and winter Yosemite photos from Flickr. Example application *(bottom)*: using a collection of paintings of famous artists, our method learns to render natural photographs into the respective styles.

基础知识：

CycleGAN(Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks)

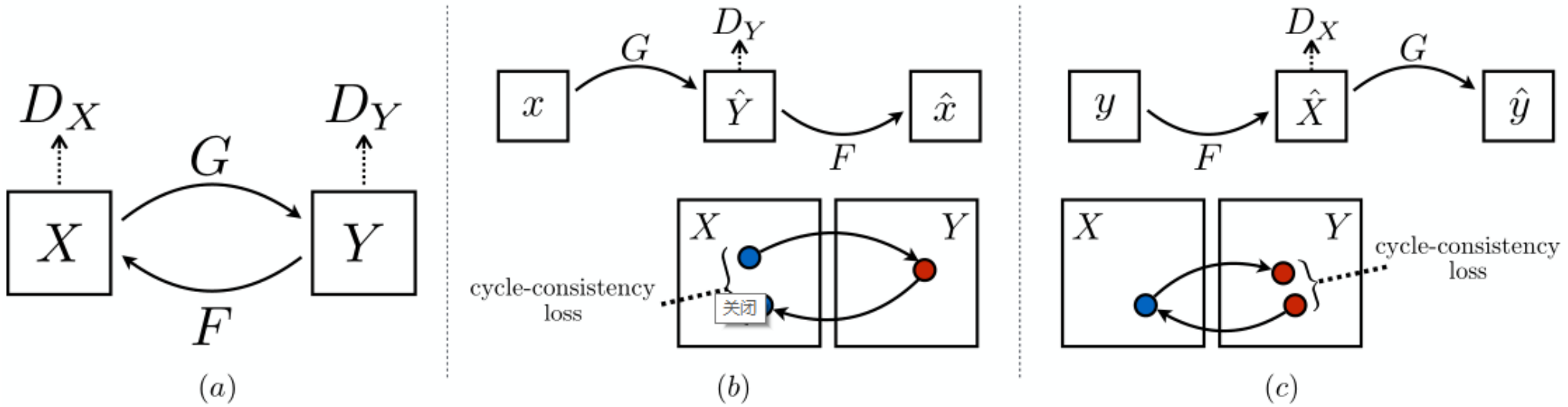
训练数据



Pix2pix： Image-to-image translation with conditional adversarial networks. In CVPR, 2017

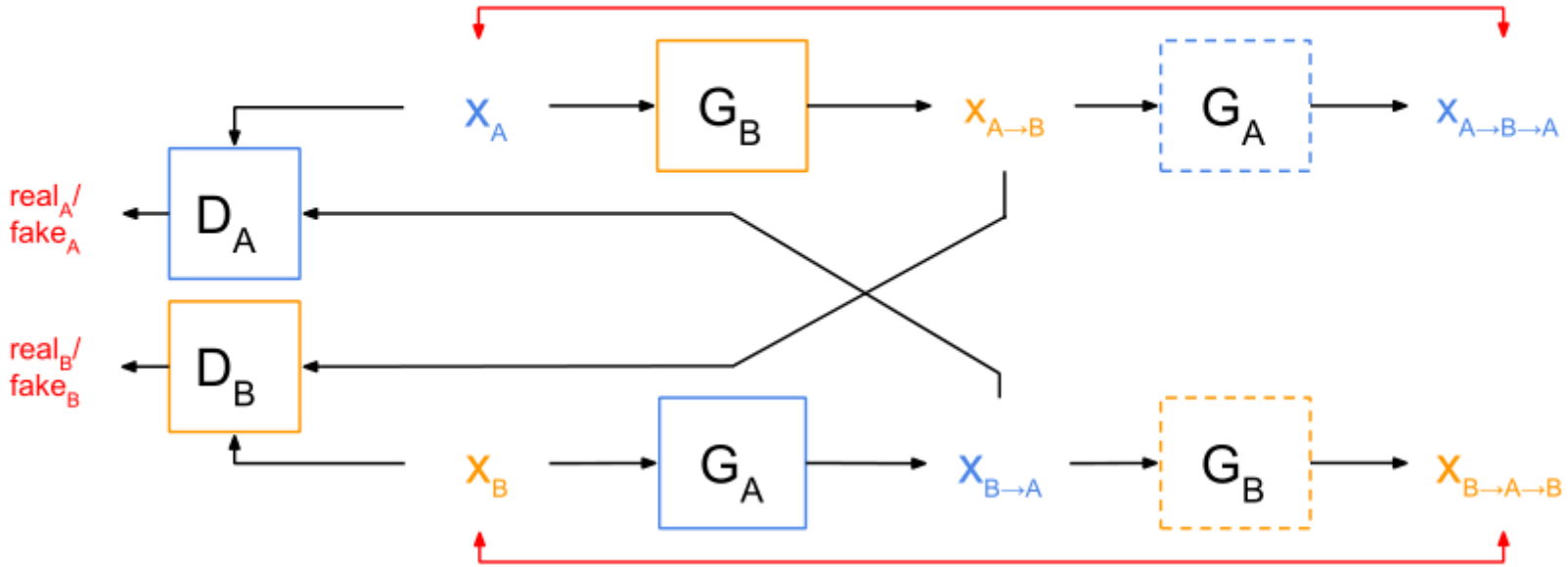
基础知识：

CycleGAN(Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks)

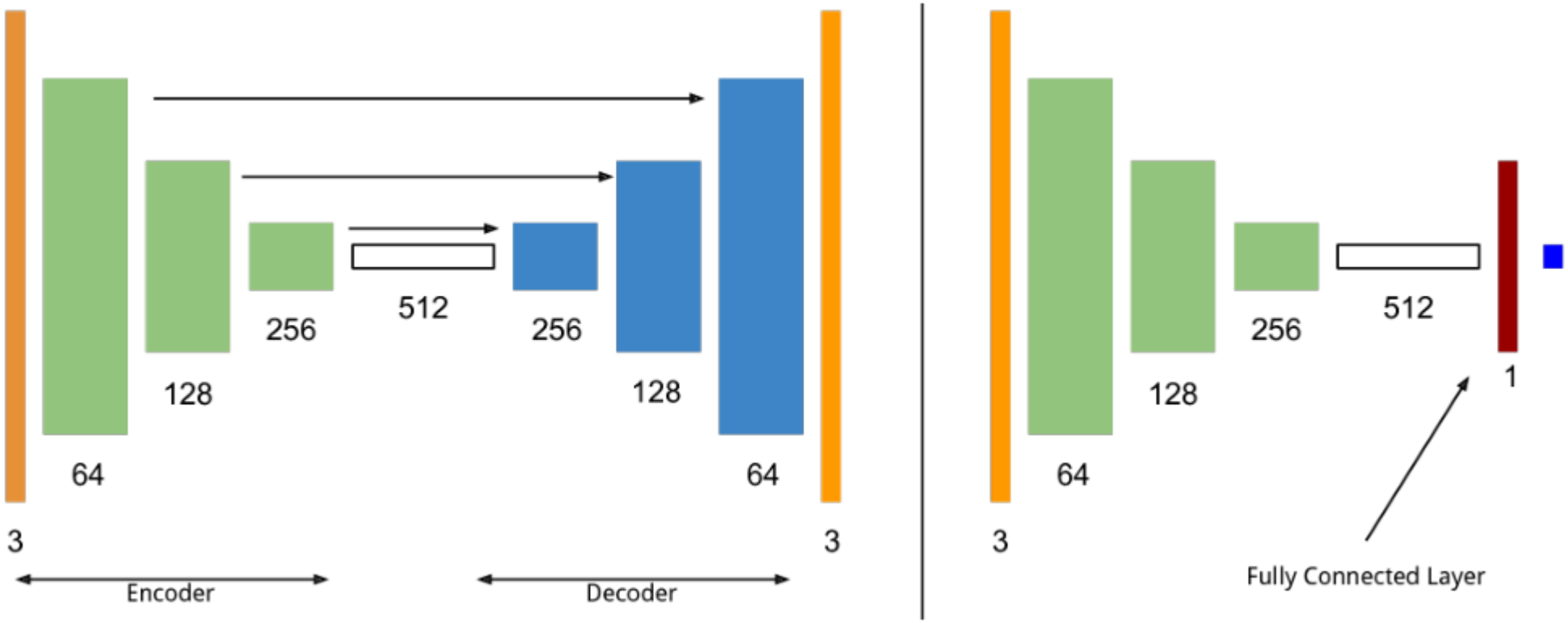


网络结构和损失设计

本文网络结构

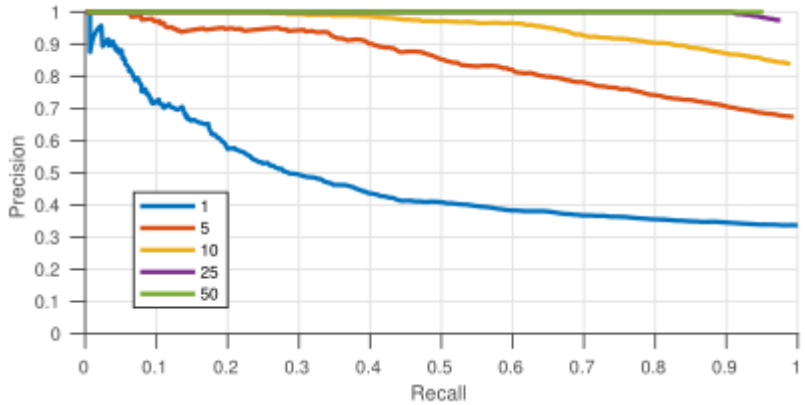

x_A
G_B
$x_{A\to B}$
G_A
$x_{A\to B\to A}$
$real_A$/
$fake_A$
D_A
$real_B$/
$fake_B$
D_B
x_B
G_A
$x_{B\to A}$
G_B
$x_{B\to A\to B}$

本文网络结构

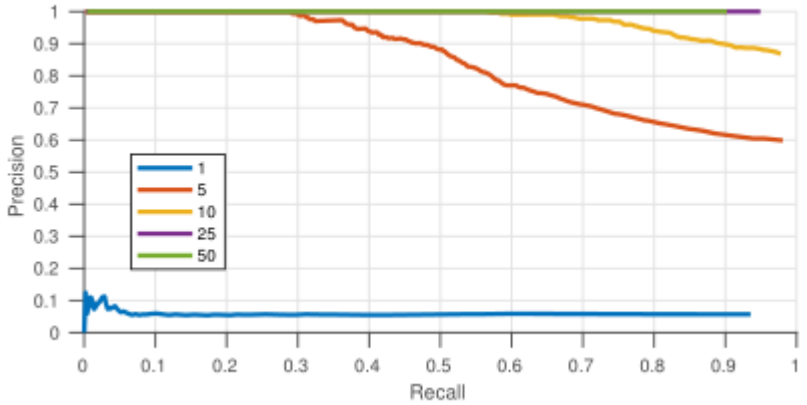


1. 生成的图片直接接SeqSLAM
2. 全连接层输出作为representation

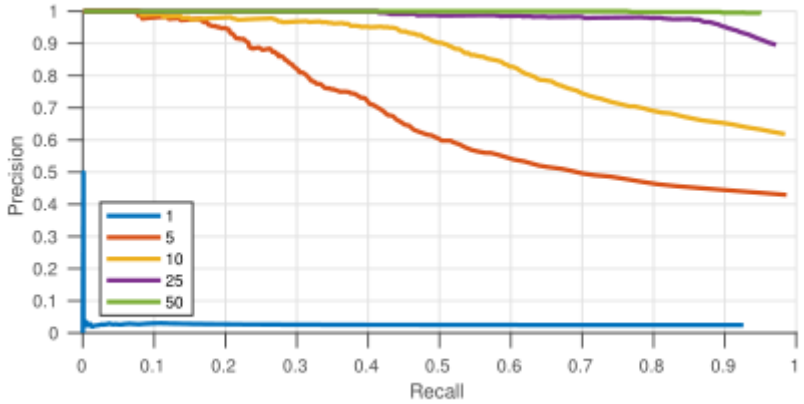
实验结果



全连接层输出作为representation
（多张图片时直接串联）



生成的图片直接接SeqSLAM

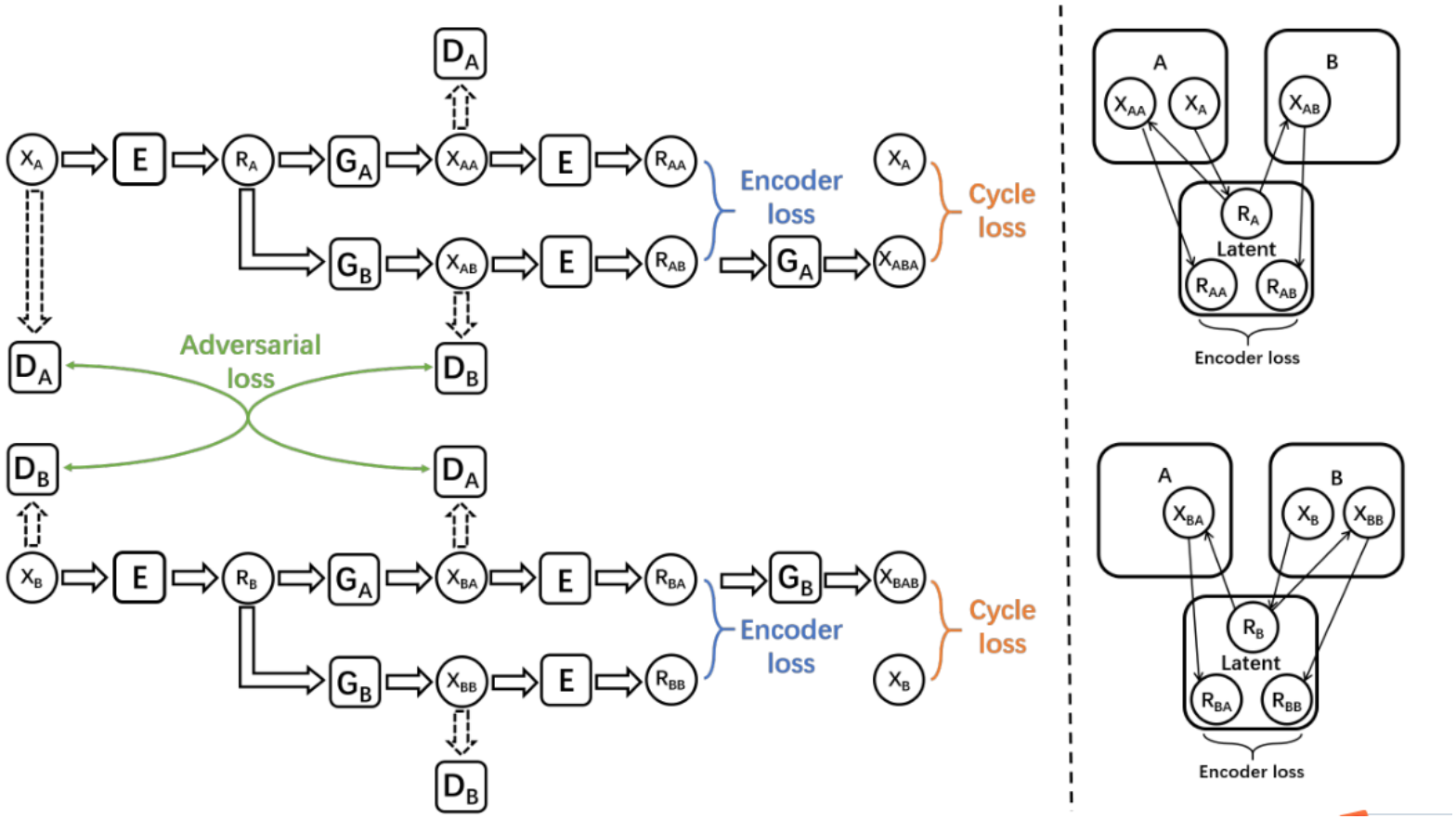


SeqSLAM

我的想法

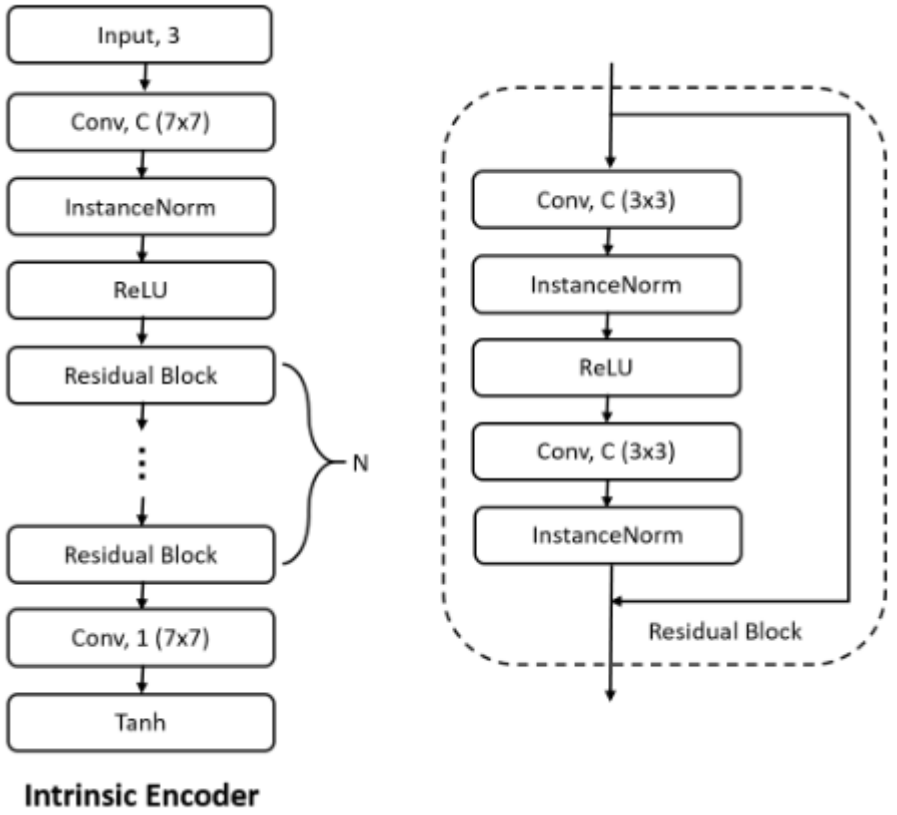
- 本文从一个domain（冬天）转到另一个domain（夏天）。假设提前知道了季节，并且仅仅是两两互相转换。
- 能不能设计一种不是转换，而是直接挖掘出多个domain中的潜在稳定结构/内容。

Intrinsic Encoder



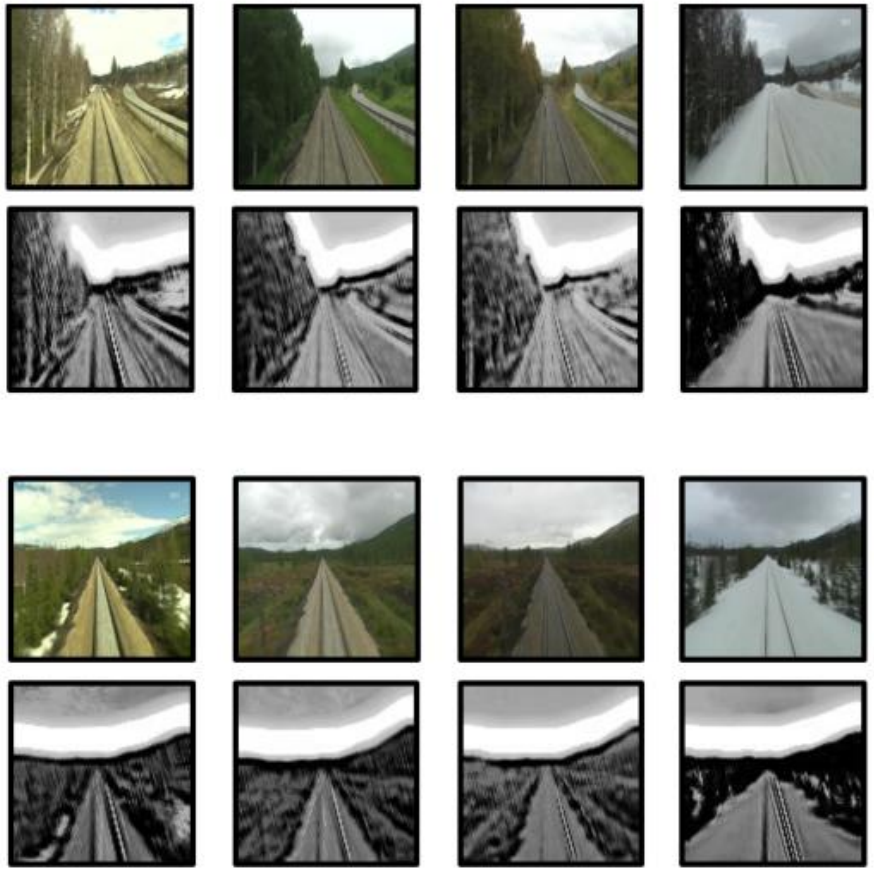
损失设计

Intrinsic Encoder

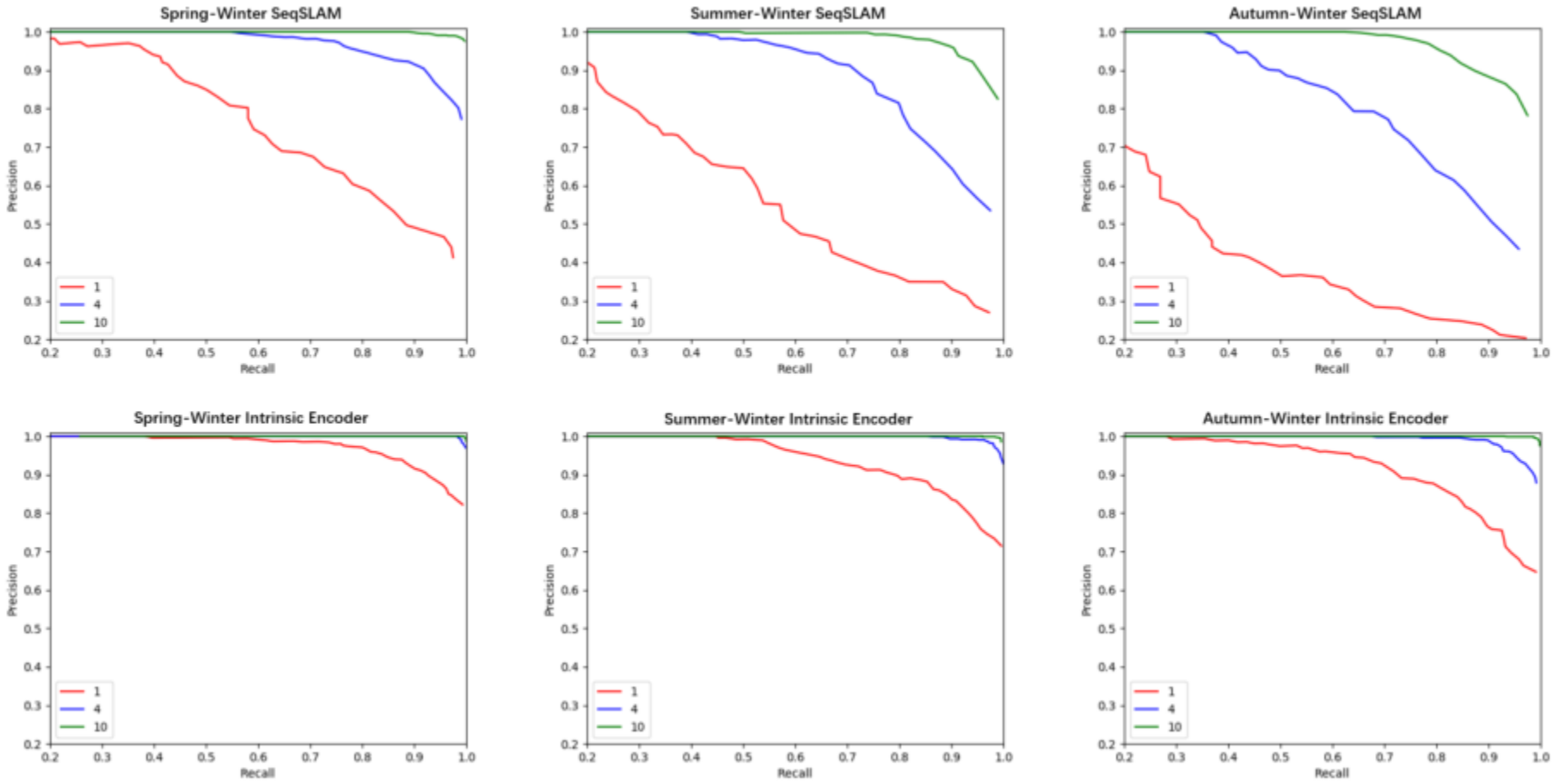


Encoder 结构

实验结果



实验结果



多季节 PR曲线