硕士生学术交流

汇 报 人 ： 陈润发
汇报日期 ： 2020/7/13

硕士生学术交流

内容安排

CVPR2020线上会议参会感受分享

研究问题概述Problem Definition

Paper Reading 1: CVPR20 MixNMatch

Paper Reading 2: CVPR20 SEAN

Future Works

硕士生学术交流

内容安排

CVPR2020线上会议参会感受分享

研究问题概述Problem Definition

Paper Reading 1: CVPR20 MixNMatch

Paper Reading 2: CVPR20 SEAN

Future Works

基本情况介绍

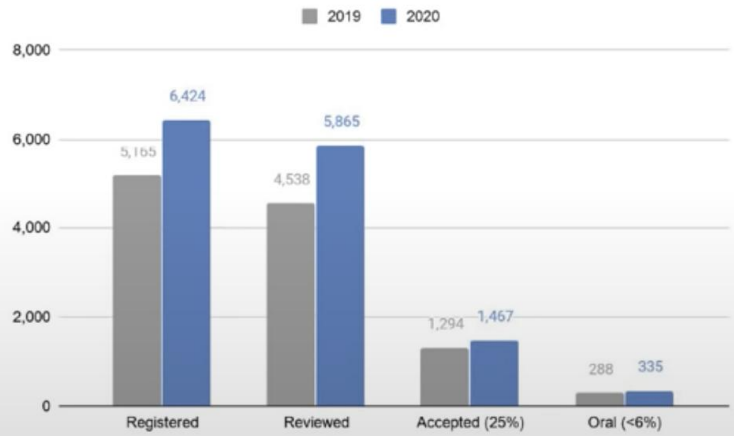
- 西雅图时间6月14日至19日在线上成功召开。
- 有7600名与会者，25个赞助商，5025个活动项目（论文/演讲等），以及1,497,800分钟的ZOOM讨论。
- 6月14日、15日、19日为Workshops和Tutorials
- 16日～18日为Main Conference（Oral和Poster）

基本情况介绍

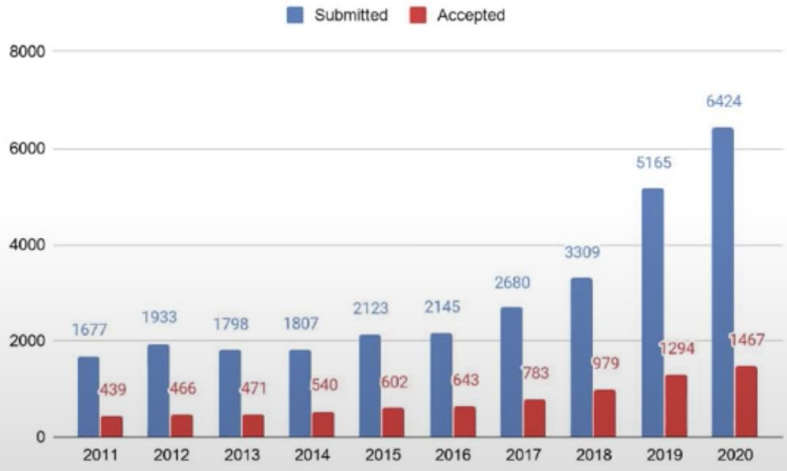
Posters/Orals

As before, papers were accepted as orals and posters purely based on the quality. There were no caps set in the paper decision process.

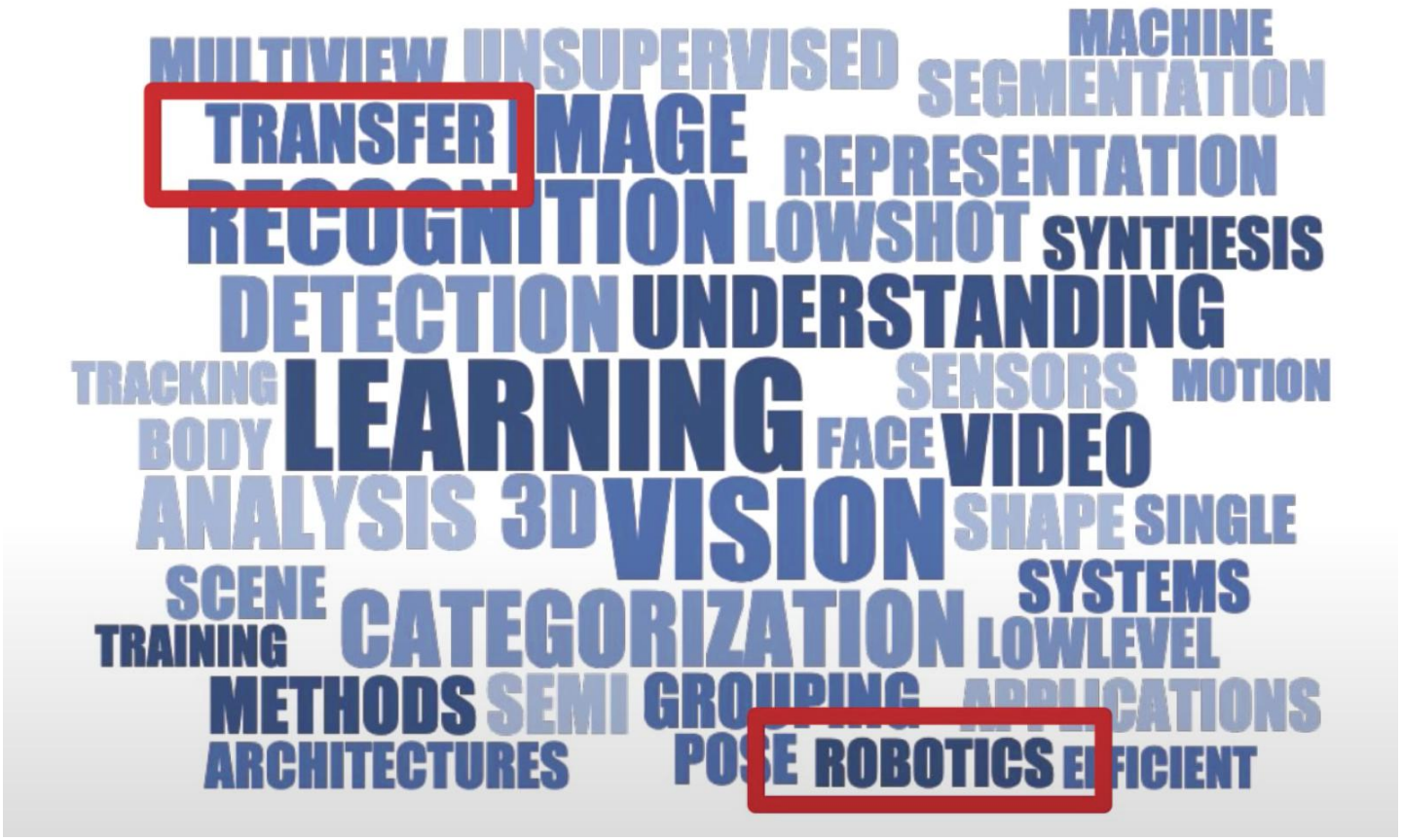
- **6,424 registered** (vs. 5,165 in 2019)
- **5,865 valid submissions** (vs. 4,538 in 2019)
- **1,467 accepted (25.0%)**
- **335 orals (5.7%)**



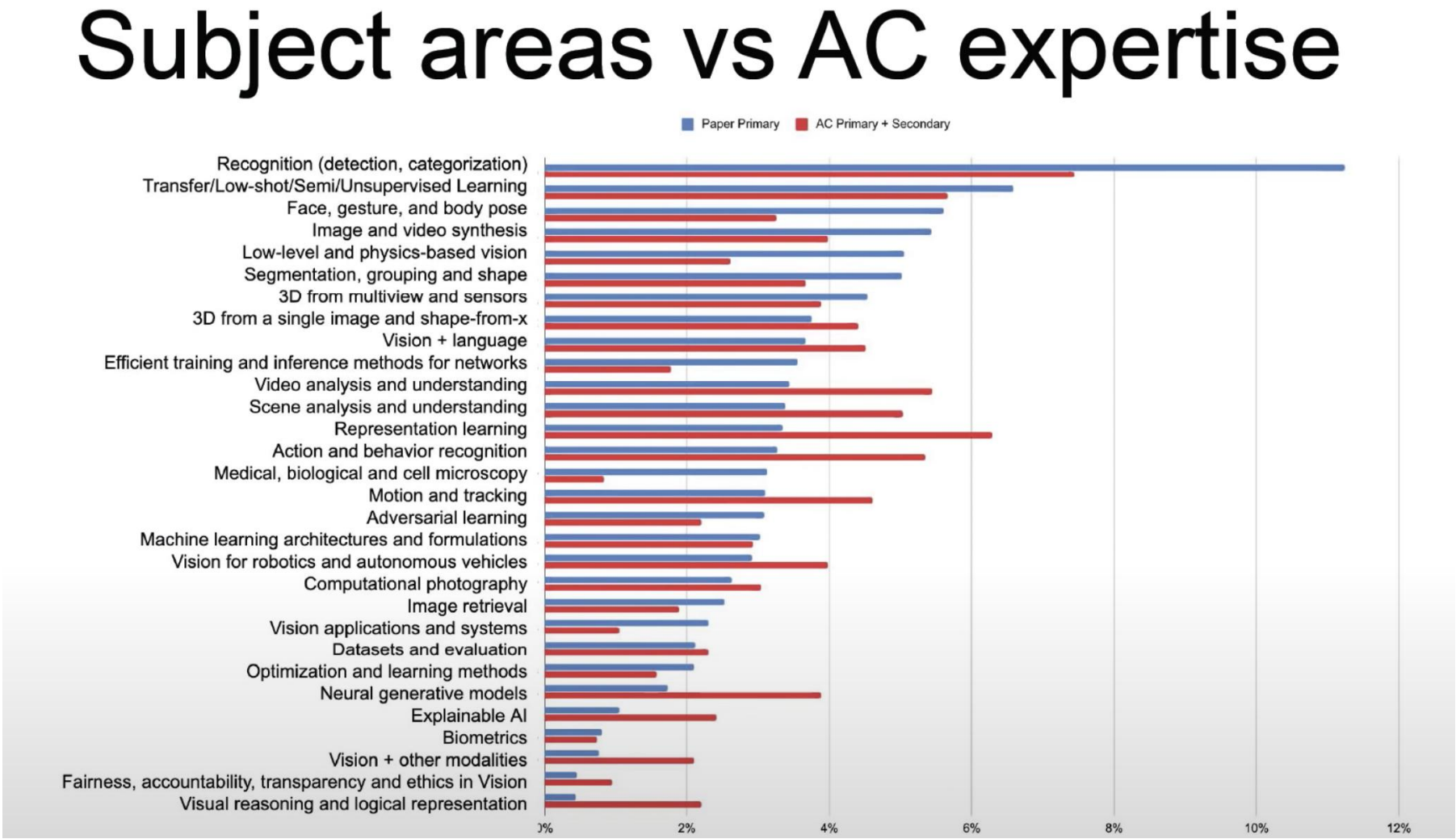
Registered vs accepted last 10 years



基本情况介绍


Popular Areas
MULTIVIEW
UNSUPERVISED
MACHINE
SEGMENTATION
TRANSFER
IMAGE
REPRESENTATION
RECOGNITION
LOWSHOT
SYNTHESIS
DETECTION
UNDERSTANDING
TRACKING
LEARNING
SENSORS
MOTION
BODY
FACE
VIDEO
ANALYSIS
3D
VISION
SHAPE
SINGLE
SCENE
SYSTEMS
TRAINING
CATEGORIZATION
LOWLEVEL
METHODS
SEMI
GROUPING
APPLICATIONS
ARCHITECTURES
POSE
ROBOTICS
EFFICIENT

基本情况介绍


Subject areas vs AC expertise
Paper Primary
AC Primary + Secondary
Recognition (detection, categorization)
Transfer/Low-shot/Semi/Unsupervised Learning
Face, gesture, and body pose
Image and video synthesis
Low-level and physics-based vision
Segmentation, grouping and shape
3D from multiview and sensors
3D from a single image and shape-from-x
Vision + language
Efficient training and inference methods for networks
Video analysis and understanding
Scene analysis and understanding
Representation learning
Action and behavior recognition
Medical, biological and cell microscopy
Motion and tracking
Adversarial learning
Machine learning architectures and formulations
Vision for robotics and autonomous vehicles
Computational photography
Image retrieval
Vision applications and systems
Datasets and evaluation
Optimization and learning methods
Neural generative models
Explainable AI
Biometrics
Vision + other modalities
Fairness, accountability, transparency and ethics in Vision
Visual reasoning and logical representation
0%
2%
4%
6%
8%
10%
12%

硕士生学术交流

内容安排

CVPR2020线上会议参会感受分享

研究问题概述Problem Definition

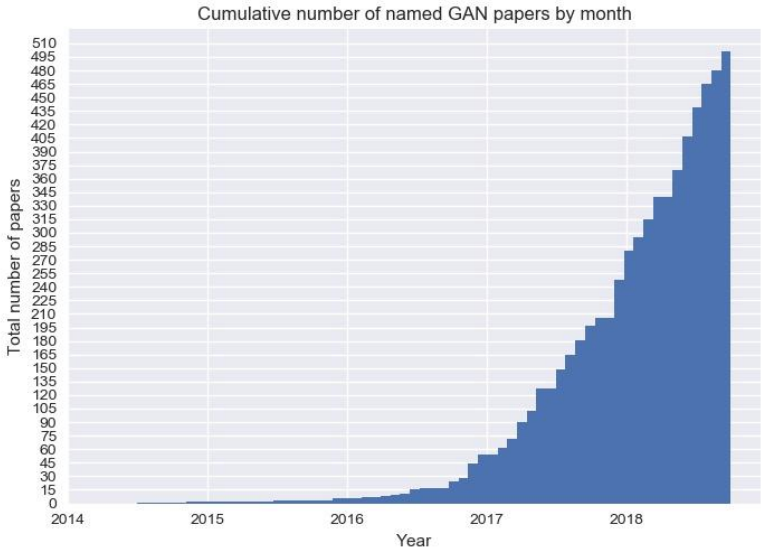
Paper Reading 1: CVPR20 MixNMatch

Paper Reading 2: CVPR20 SEAN

Future Works

生成对抗网络GAN

由于能够从训练数据中学习数据分布从而生成全新样本，GAN 为机器学习提供了一个全自动自学习的损失函数。



- 各种图像生成任务取得了成功
  - 包括人脸和自然图像生成，图像跨域转换，图像超分辨率和文本到图像合成等。
- 相关理论和技术快速发展
  - ✓ 训练不稳定及模式坍塌
  - ✓ 评价指标
  - ✘ 局部与全局收敛性及时间效率
  - ✘ 与其他生成模型的互补组合

Generator	The Evolution of GANs	Discriminator
DCGAN Generator	**ACGAN** IS:28.5	AC Discriminator
Class Batch Norm	**SNGAN** IS:36.8 FID: 27.2	Projection Discriminator Hinge Loss Spectral Normalization
Spectral Normalization Self-Attention Module	**SAGAN** IS:52.5 FID: 18.7	Self-attention Module
Large Batch Training Shared Embedding in ClassBN Truncated Trick	**BigGAN** IS:98.8 FID: 8.7	Large Batch Training Truncated Trick
Deeper Architecture	**BigGAN-deep** **IS:124.5** **FID: 5.7**	Deeper Architecture

样式（风格）编码

- 样式（风格）控制是各种图像合成和处理应用的重要组成部分，是可控的条件生成的重要手段。
- 样式通常不是由用户手动设计的，而是从参考图像中提取的。

Content Image

This image is licensed under CC-BY 3.0

+

Style Image

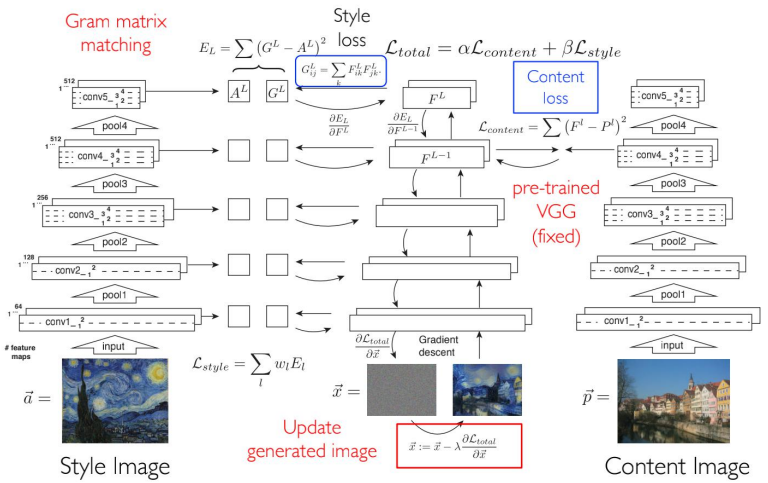
Starry Night by Van Gogh

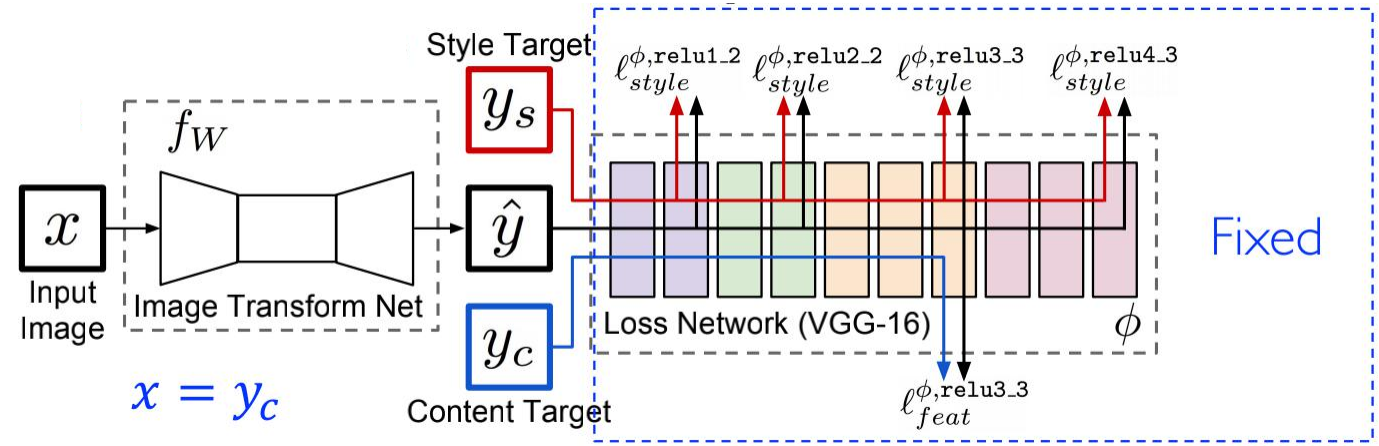
=

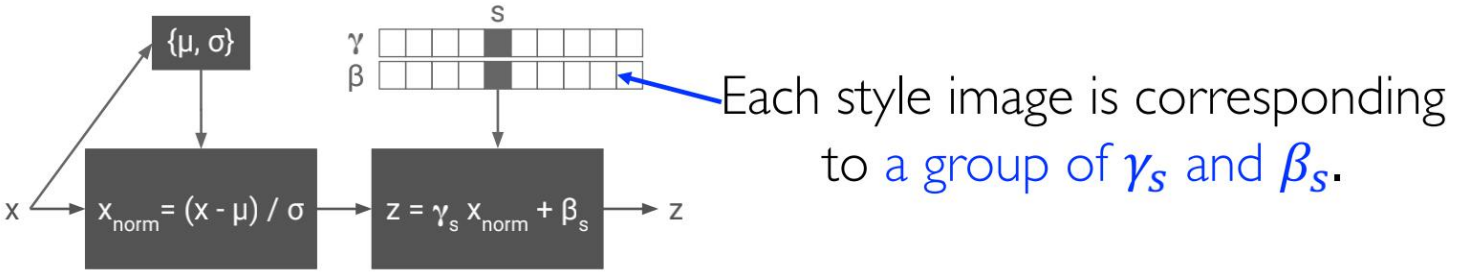
Style Transfer

样式（风格）编码

- 样式编码三大方式：
  - 编码为图像特征的统计量；
  - 编码为神经网络的权重；
  - 编码为网络规一化层的参数。



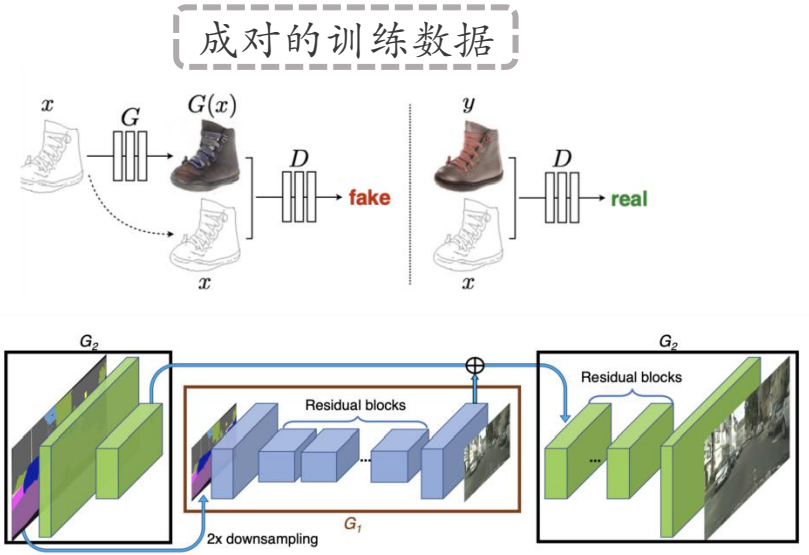


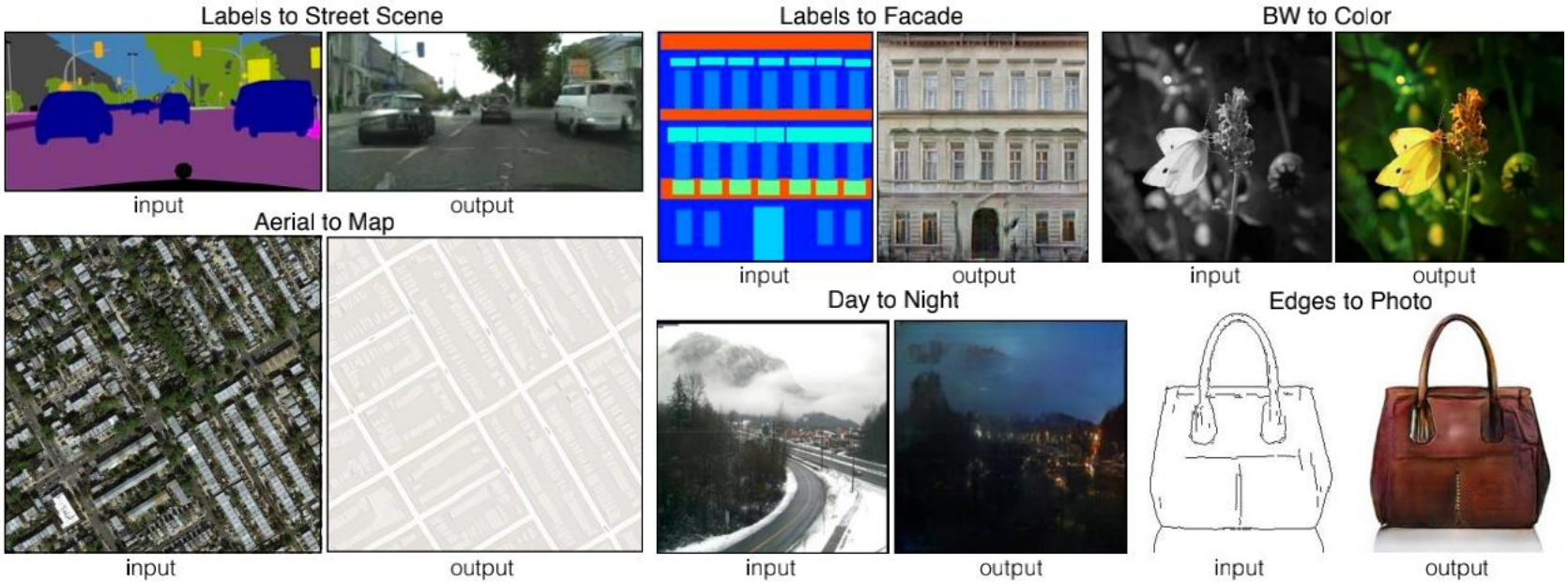


Each style image is corresponding to a group of γ_s and β_s.

跨域图像转换I2I

计算机图形学和计算机视觉中的许多问题都可被视为将输入图像“翻译”为相应的输出图像的图像处理问题，例如着色，超分辨率，以及逼真的语义图像合成（标签→图像），它一种特别有用的类型，可以通过修改输入的语义布局图像轻松地进行用户控制。



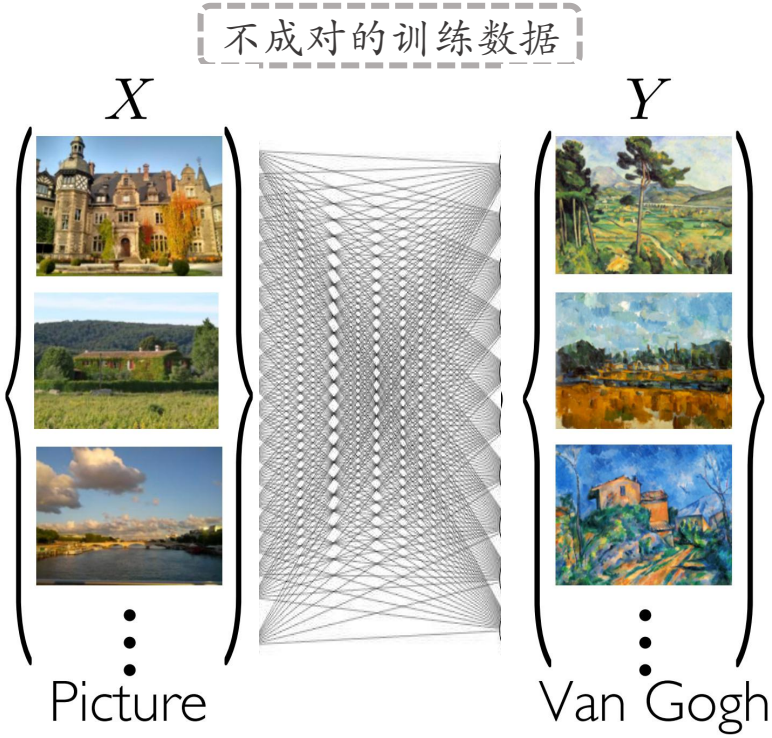


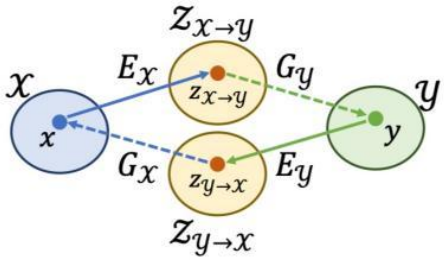
跨域图像转换I2I

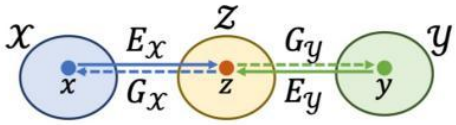
✘对齐的训练图像对难以获得，限制其应用

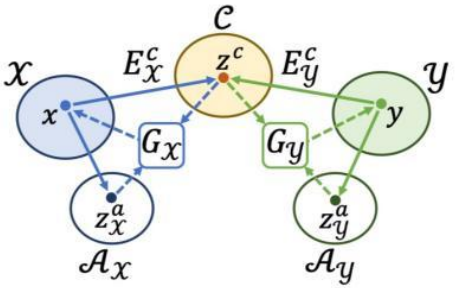
■无监督图像转换，无需成对数据

- ✓ 循环一致性损失
- ✓ 多对多映射，多模式生成
- ✓ 域共享隐层空间
- ✓ 解耦特征空间
- ✓ 注意力模块特征选择





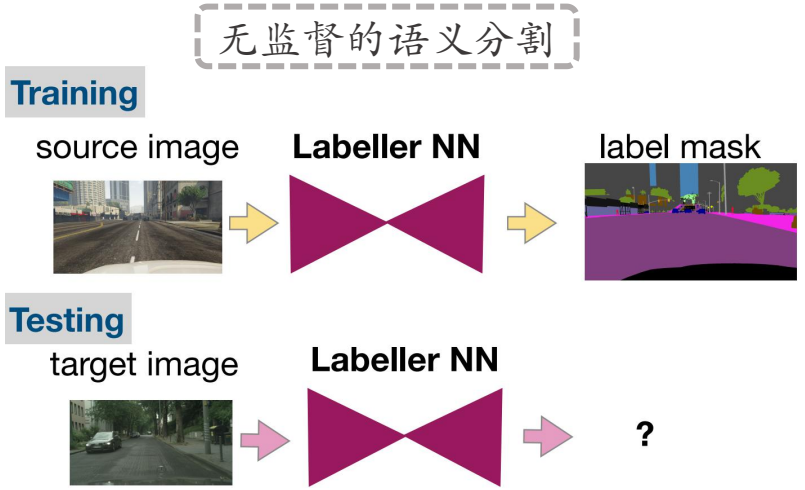


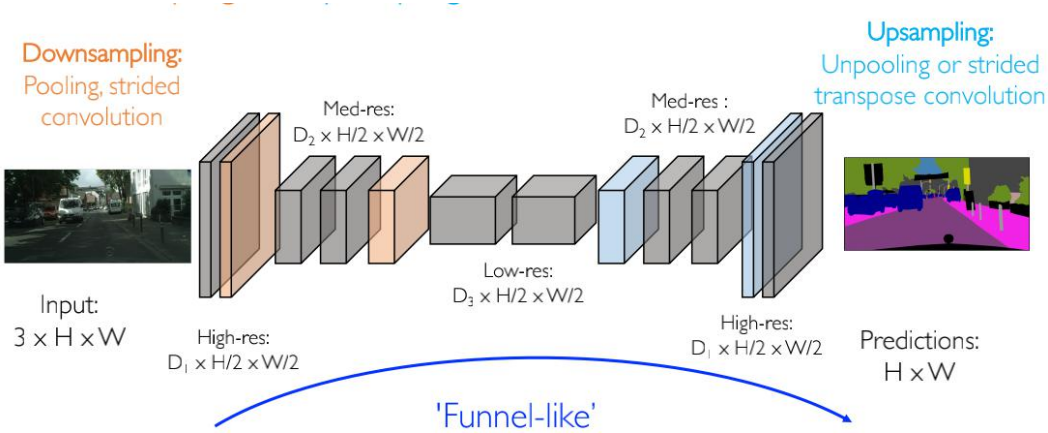


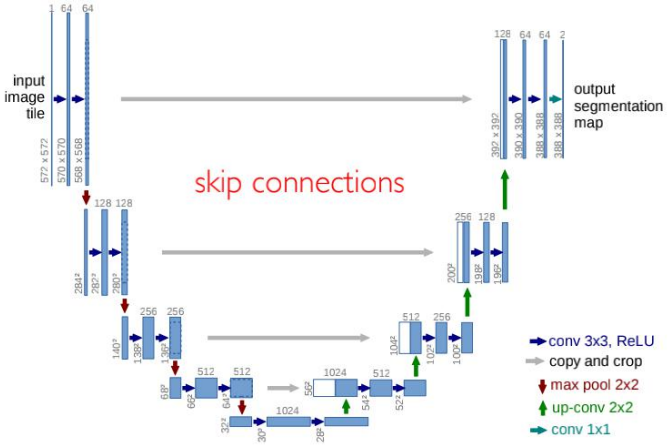
语义分割域自适应

- 语义分割是计算机视觉中最具挑战性的任务之一，它试图预测给定图像或视频帧的像素级语义标签，以方便下游应用，例如自动驾驶，视频监控和图像编辑。

- 基于CNN方法的关键是标注大量涵盖可能的场景变化的图像。

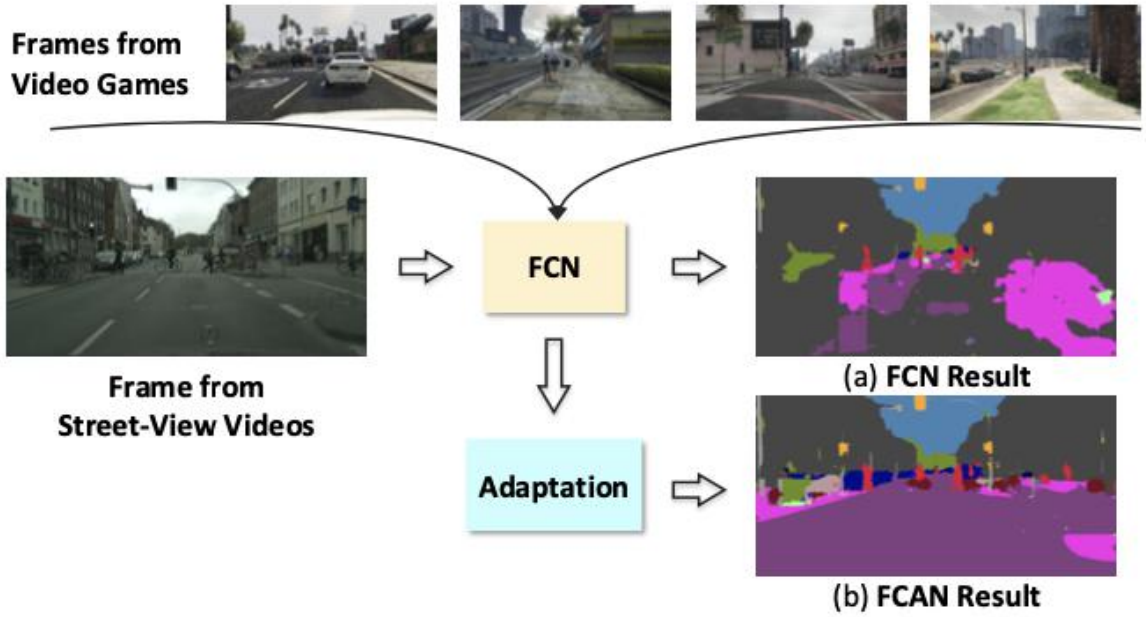






语义分割域自适应

- **利用计算机图形技术来模拟虚拟环境并自动生成图像和标签是一种经济的选择。**
  - 例如GTA5和SYNTHIA 是两个热门的城市街道合成数据集，与真实数据集（例如Cityscapes）相似且具有重叠的类别。尽管相似，但在纹理，布局，颜色和照明条件上仍然存在细微差异，这导致数据分布不同或域差异。因此，在合成数据集上训练的某个模型的性能在应用于现实场景时会急剧下降。



硕士生学术交流

内容安排

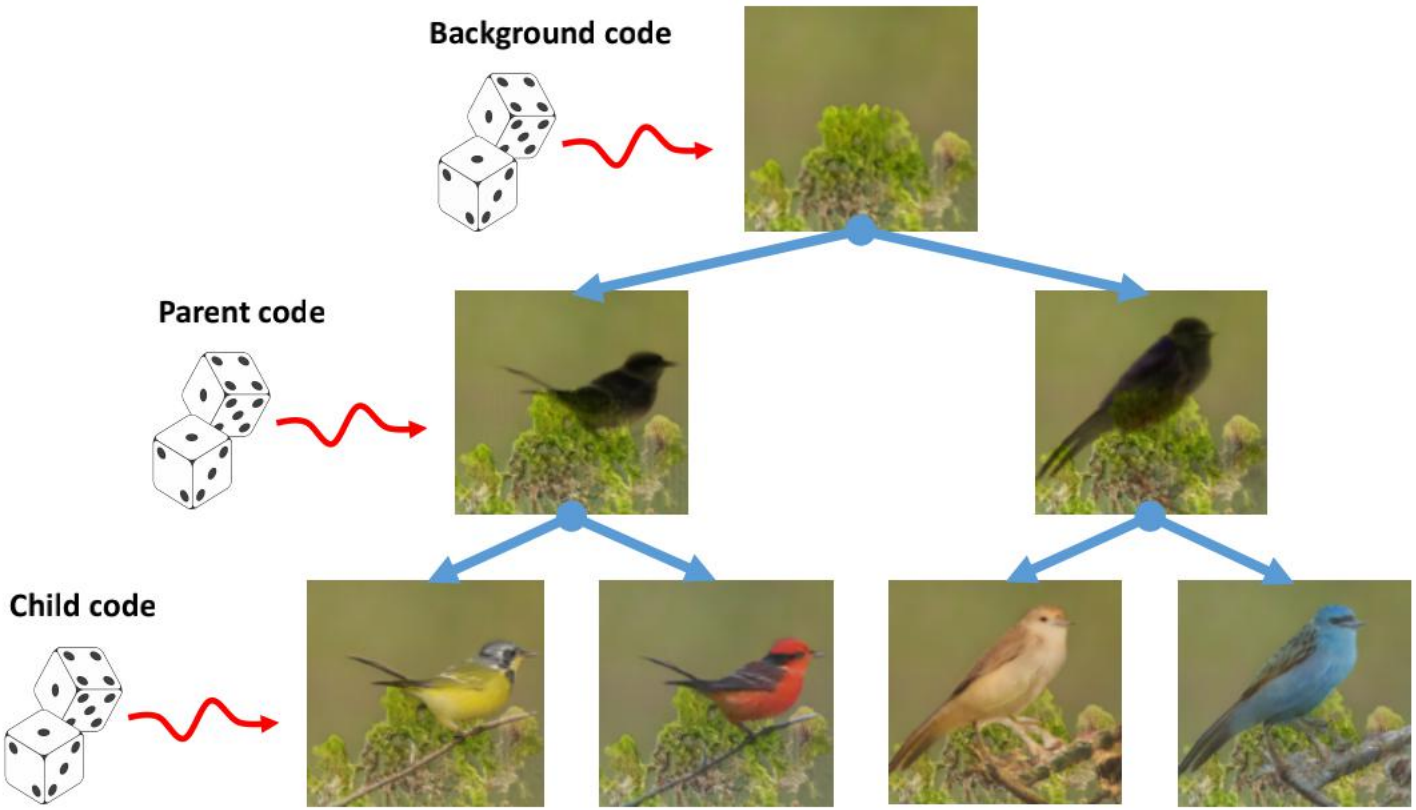
CVPR2020线上会议参会感受分享

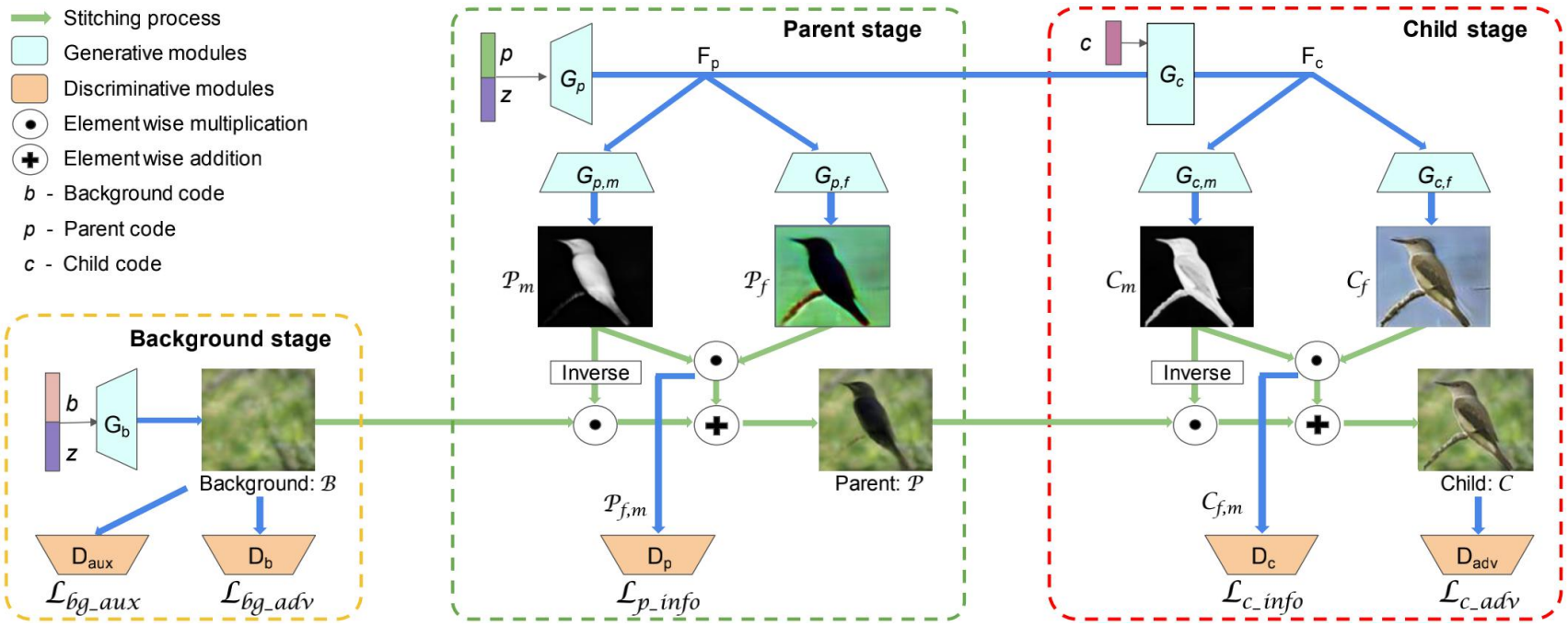
研究问题概述Problem Definition

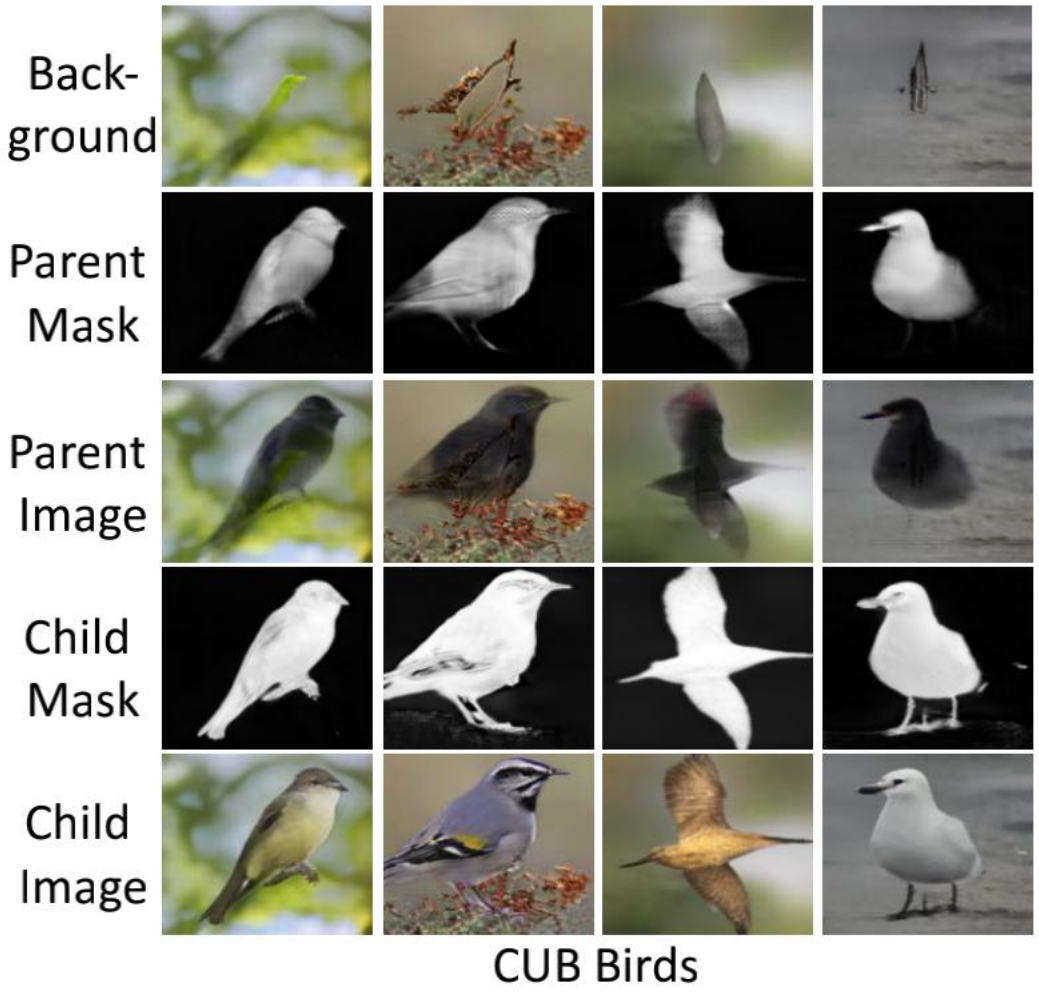
Paper Reading 1: CVPR20 MixNMatch

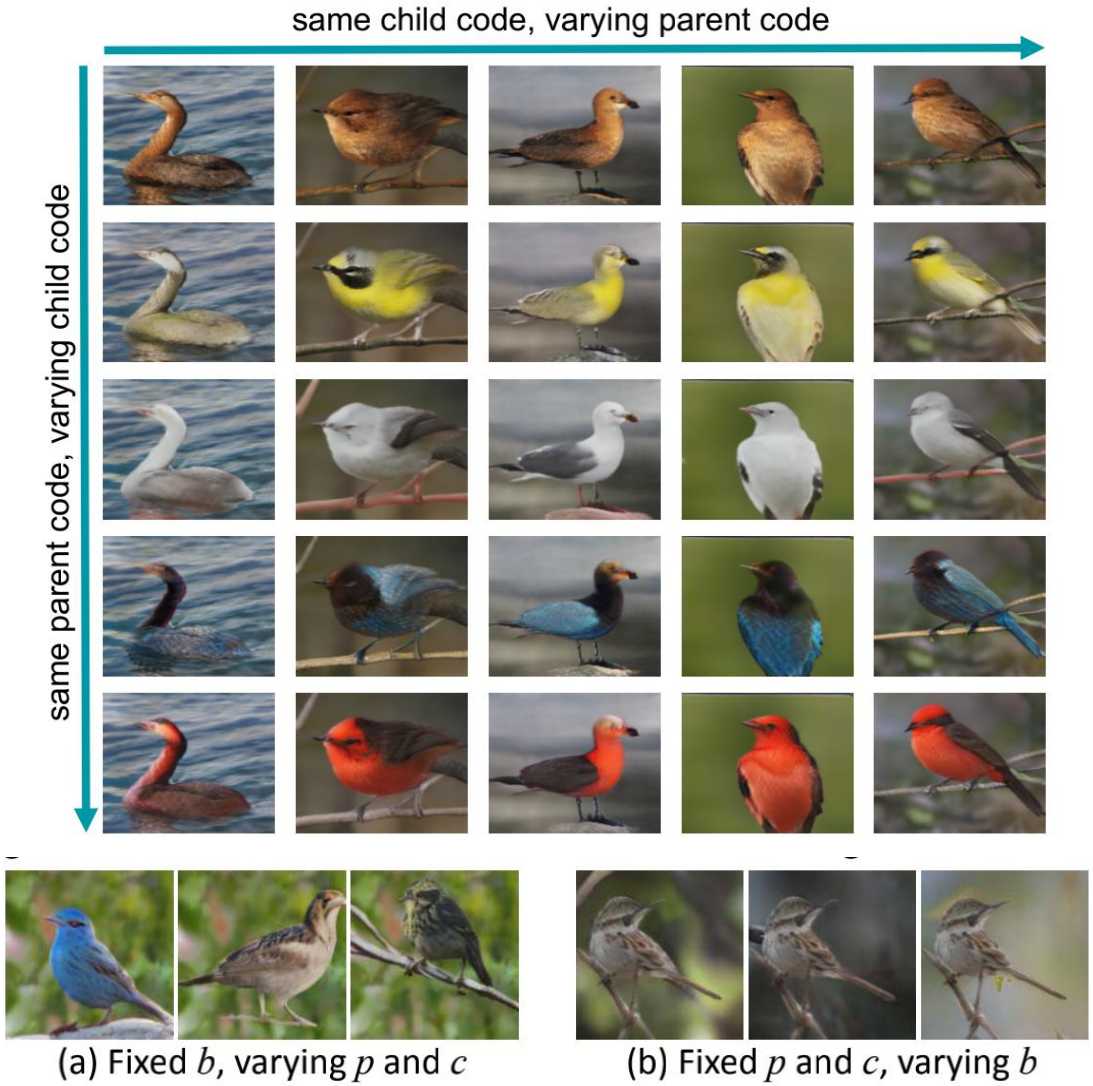
Paper Reading 2: CVPR20 SEAN

Future Works


Background code
Parent code
Child code

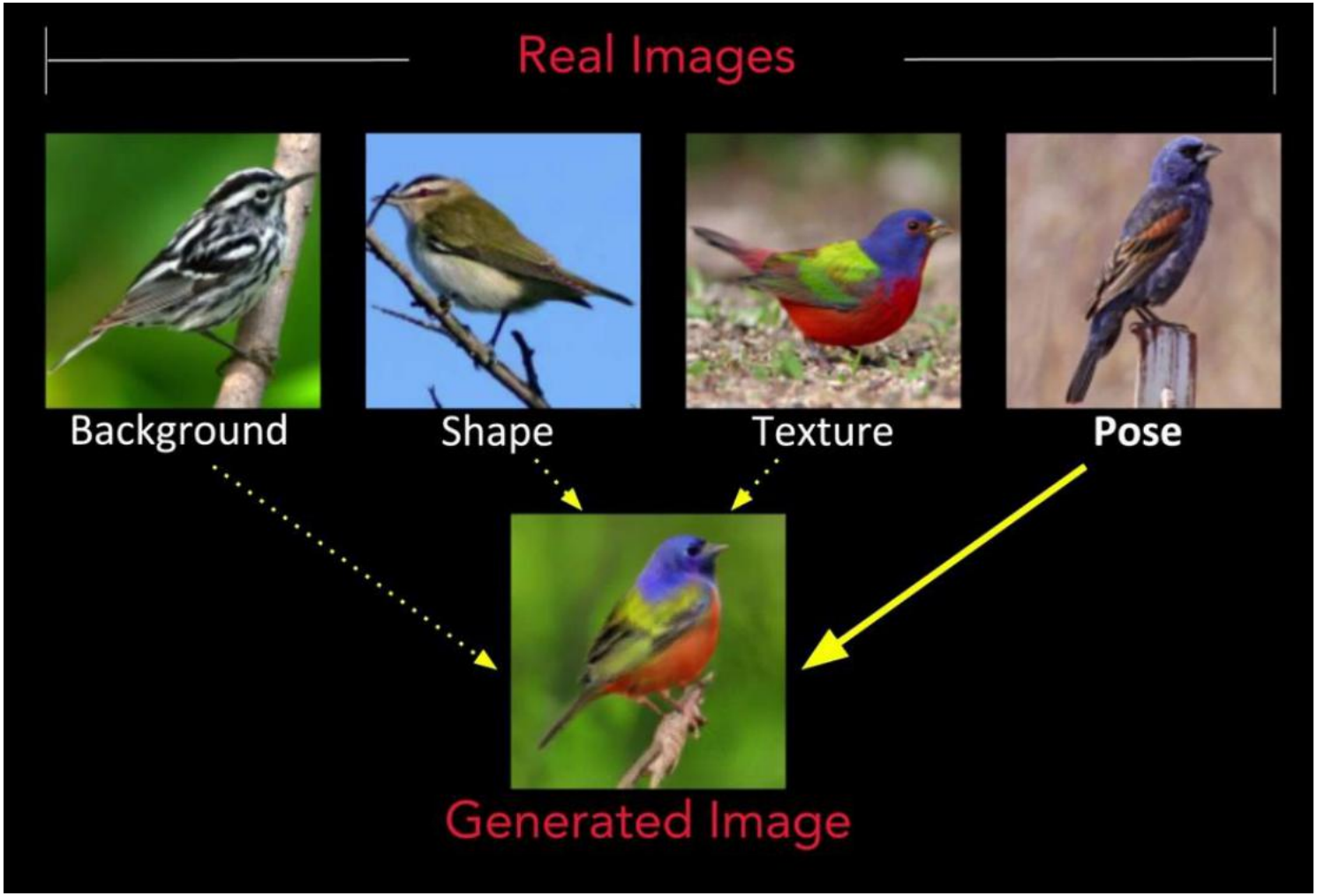

Stitching process
Generative modules
Discriminative modules
Element wise multiplication
Element wise addition
b - Background code
p - Parent code
c - Child code
Background stage
b
z
G_b
Background: $\mathcal{B}$
D_{aux}
D_b
$\mathcal{L}_{bg\_aux}$
$\mathcal{L}_{bg\_adv}$
Parent stage
p
z
G_p
F_p
$G_{p,m}$
$G_{p,f}$
$\mathcal{P}_m$
$\mathcal{P}_f$
Inverse
$\mathcal{P}_{f,m}$
Parent: $\mathcal{P}$
D_p
$\mathcal{L}_{p\_info}$
Child stage
c
G_c
F_c
$G_{c,m}$
$G_{c,f}$
C_m
C_f
Inverse
$C_{f,m}$
Child: C
D_c
D_{adv}
$\mathcal{L}_{c\_info}$
$\mathcal{L}_{c\_adv}$


Back-
ground
Parent
Mask
Parent
Image
Child
Mask
Child
Image
CUB Birds



(a) Fixed b, varying p and c

(b) Fixed p and c, varying b


Real Images
Background
Shape
Texture
Pose
Generated Image

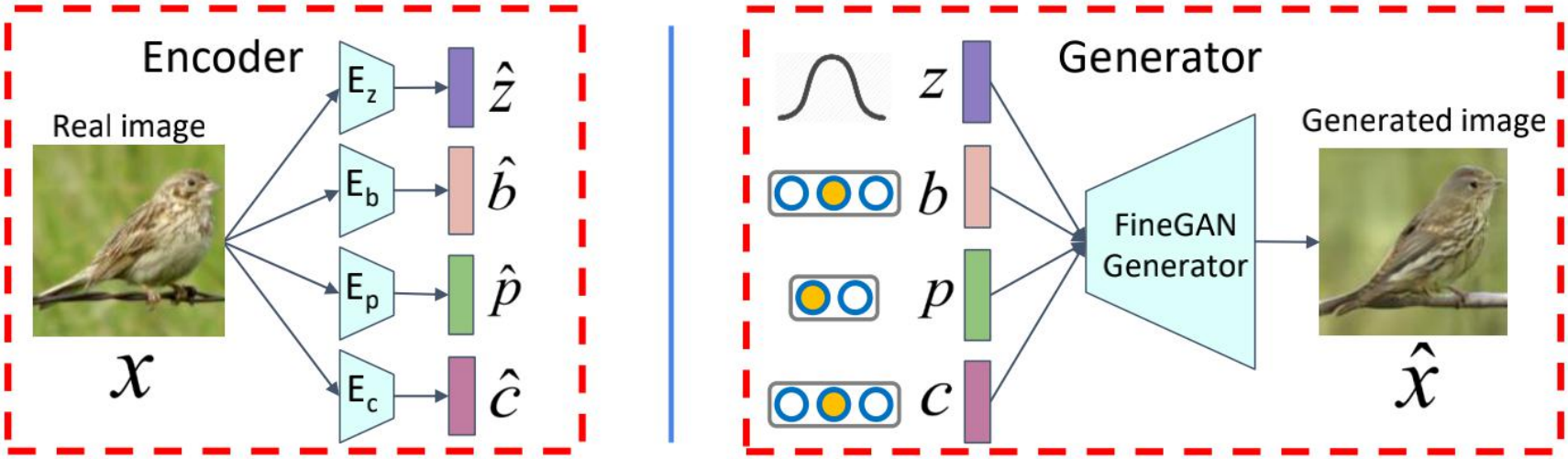
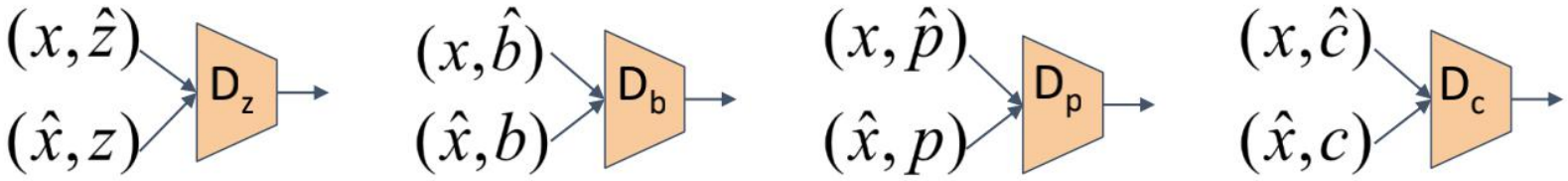
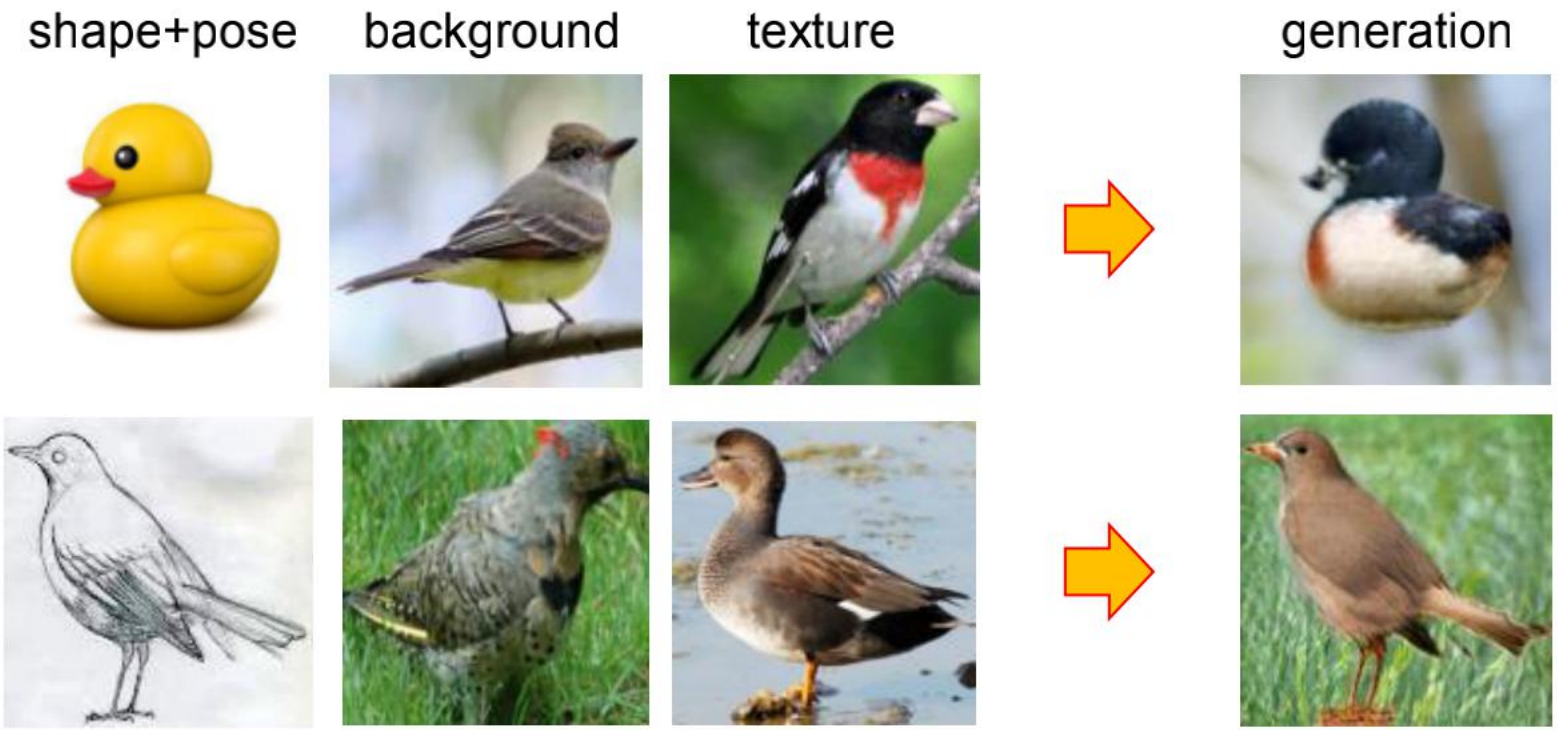


Image-code *paired* distributions need to be matched



Adversarial loss used to match *paired* distributions

Cartoon/Sketch → Image



硕士生学术交流

内容安排

CVPR2020线上会议参会感受分享

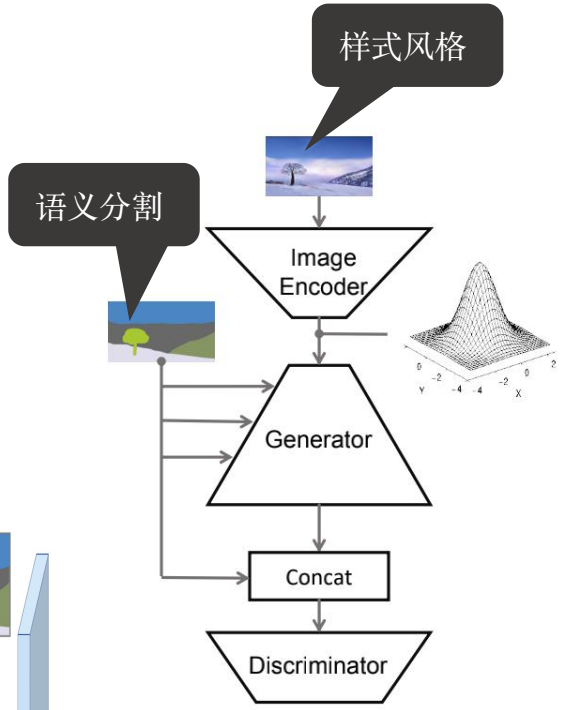
研究问题概述Problem Definition

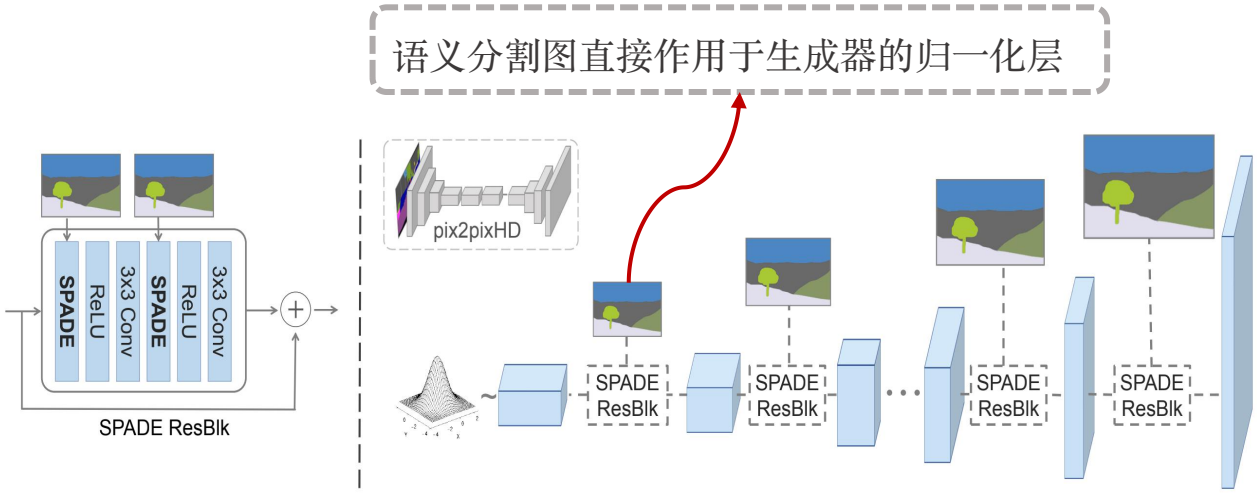
Paper Reading 1: CVPR20 MixNMatch

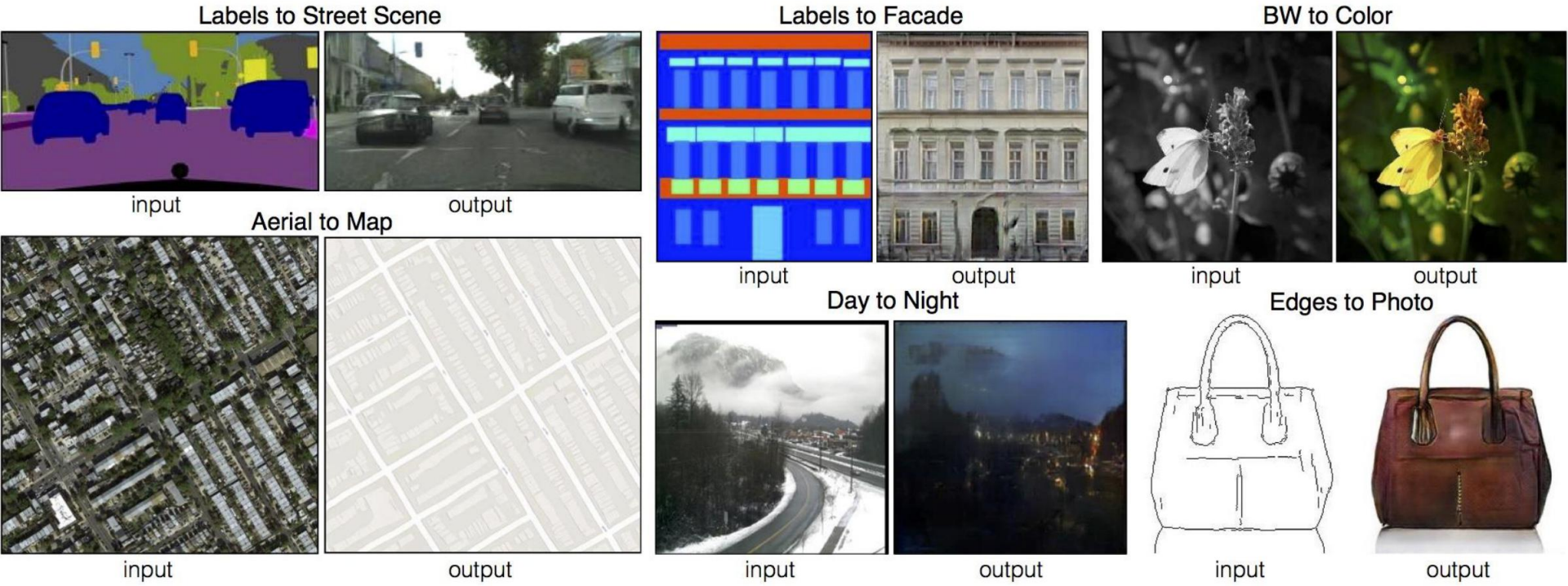
Paper Reading 2: CVPR20 SEAN

Future Works

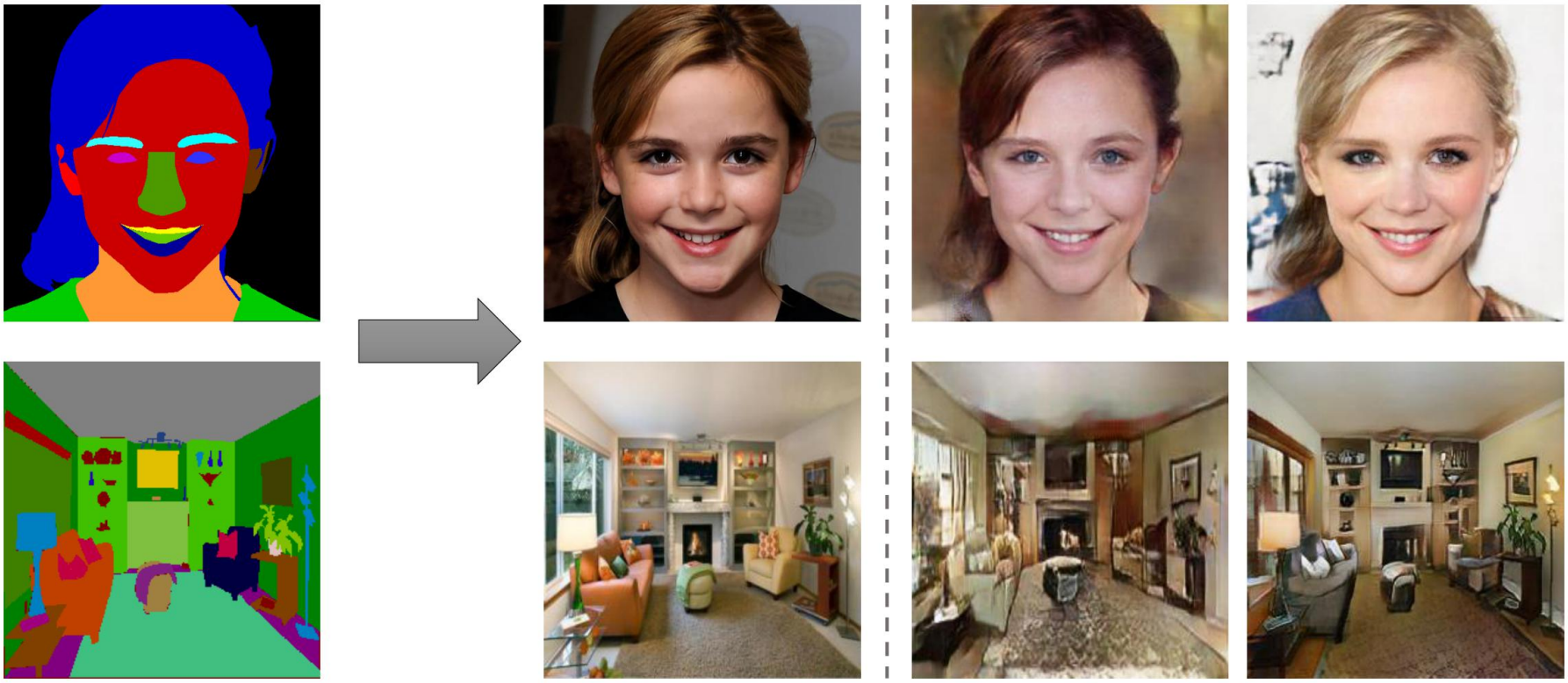
- 成对监督图像转换工作
  SPADE(CVPR‘19Oral)
  - 支持多样化输出
  - 输出满足分割语义和场景风格的合成图







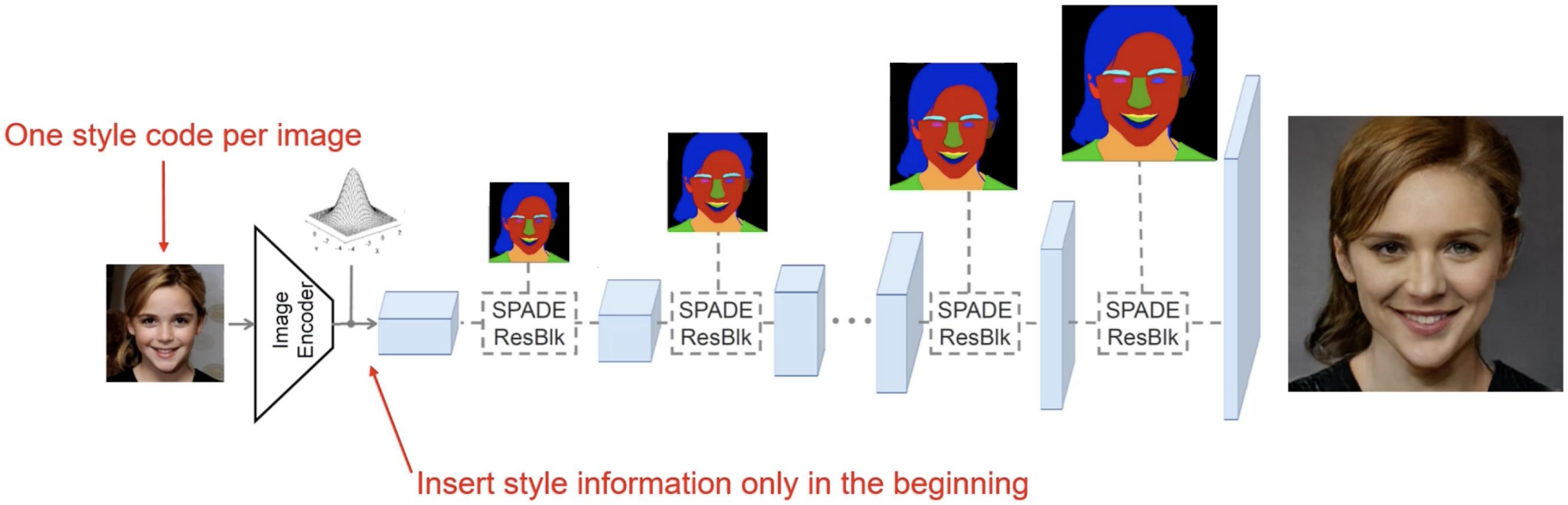
pix2pix (Isola et al.)



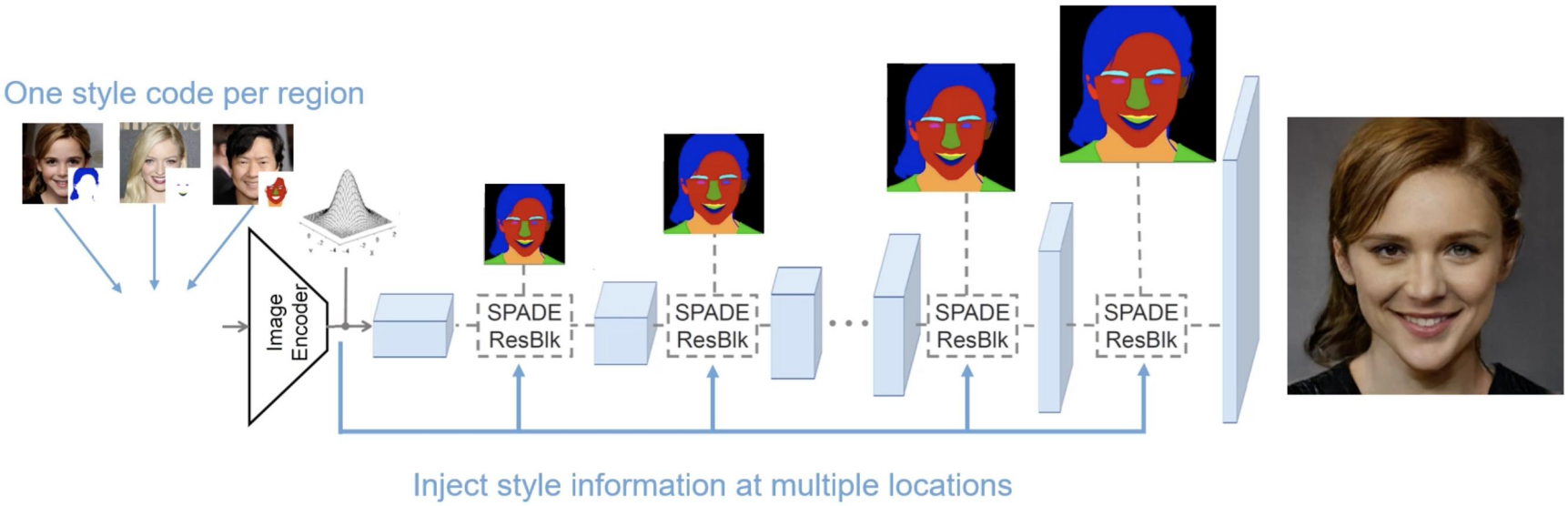
pix2pixHD (Wang et al.) & SPADE (Park et al.)

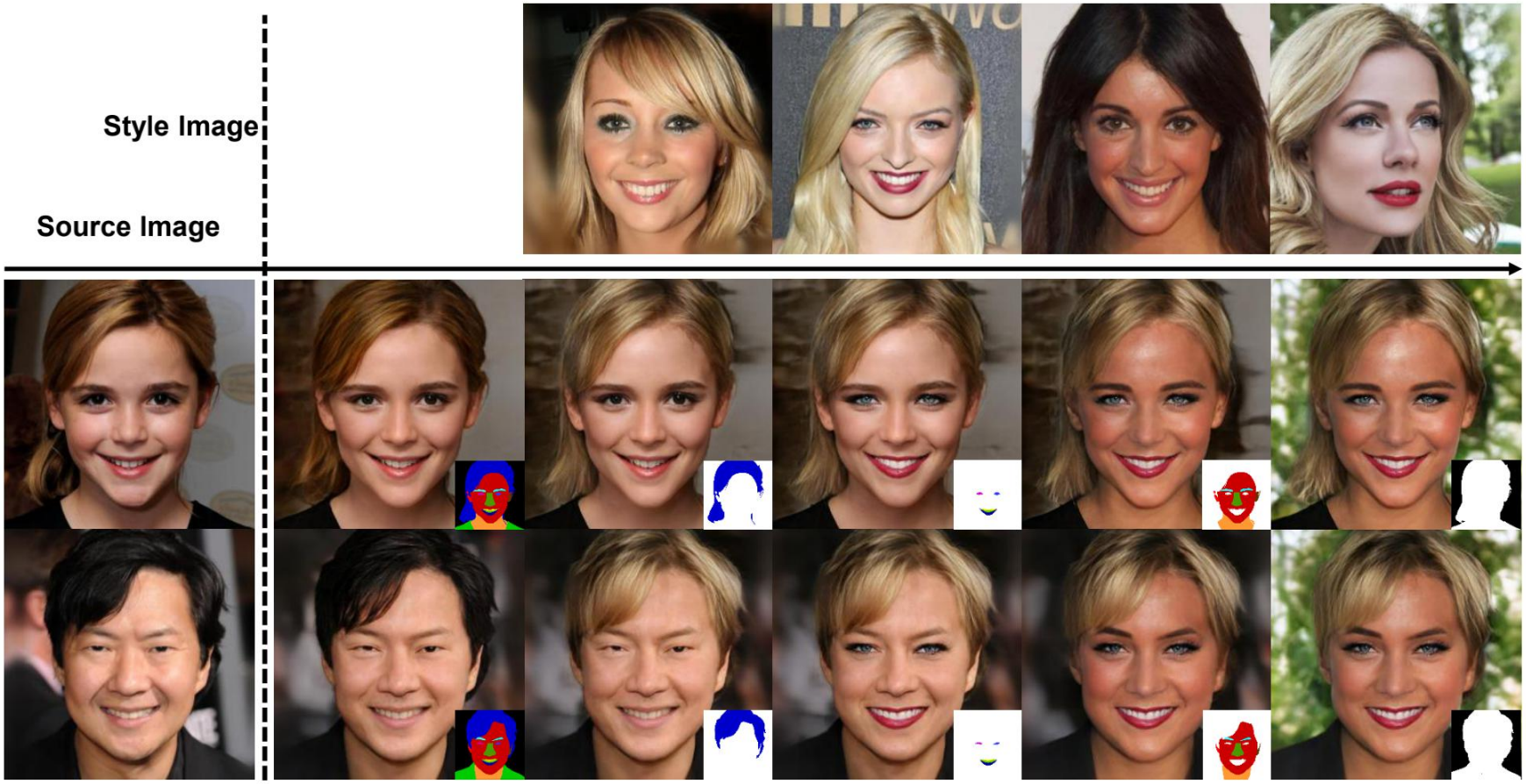
SPADE (Park et al.)

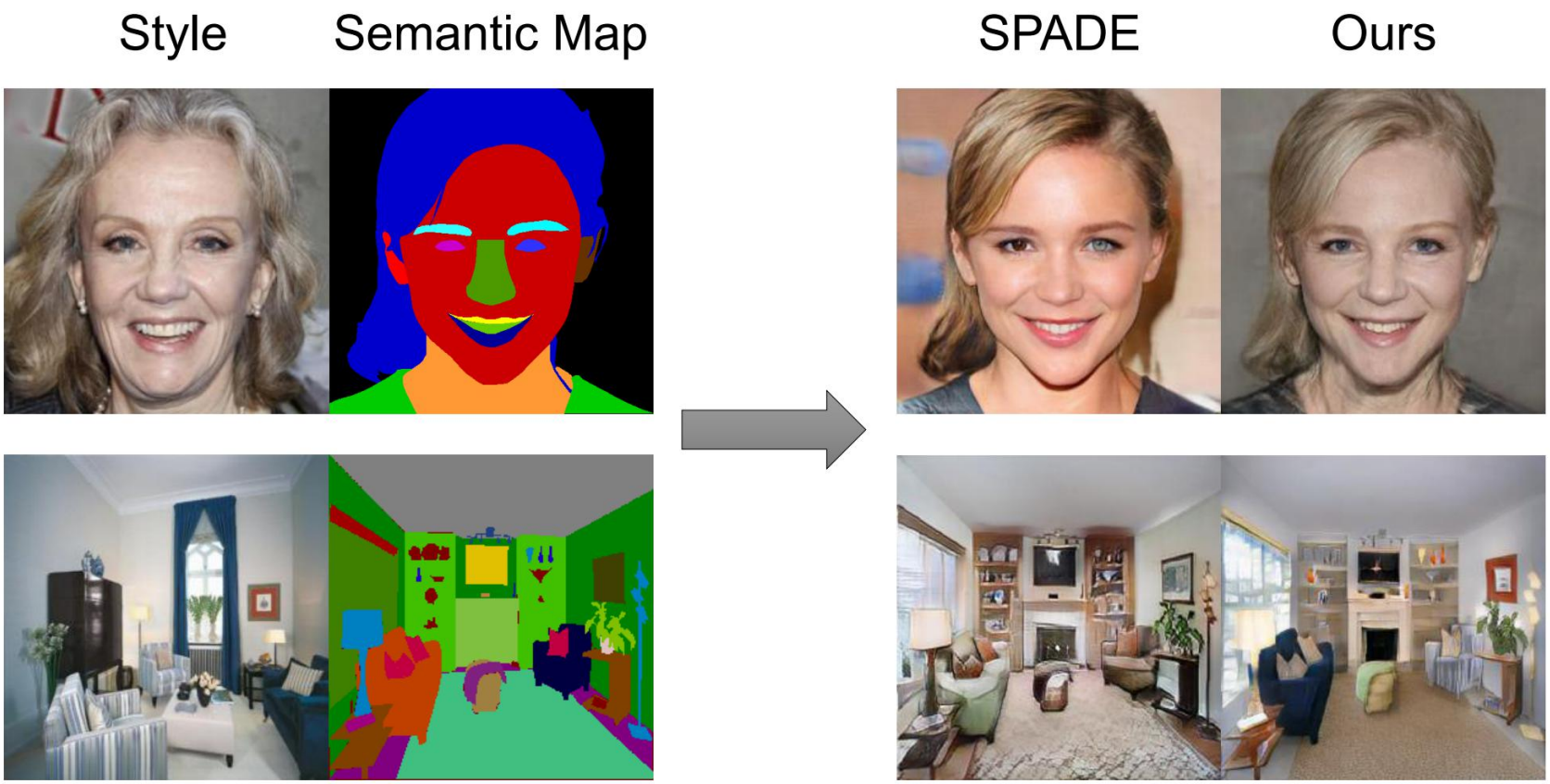
Two shortcomings

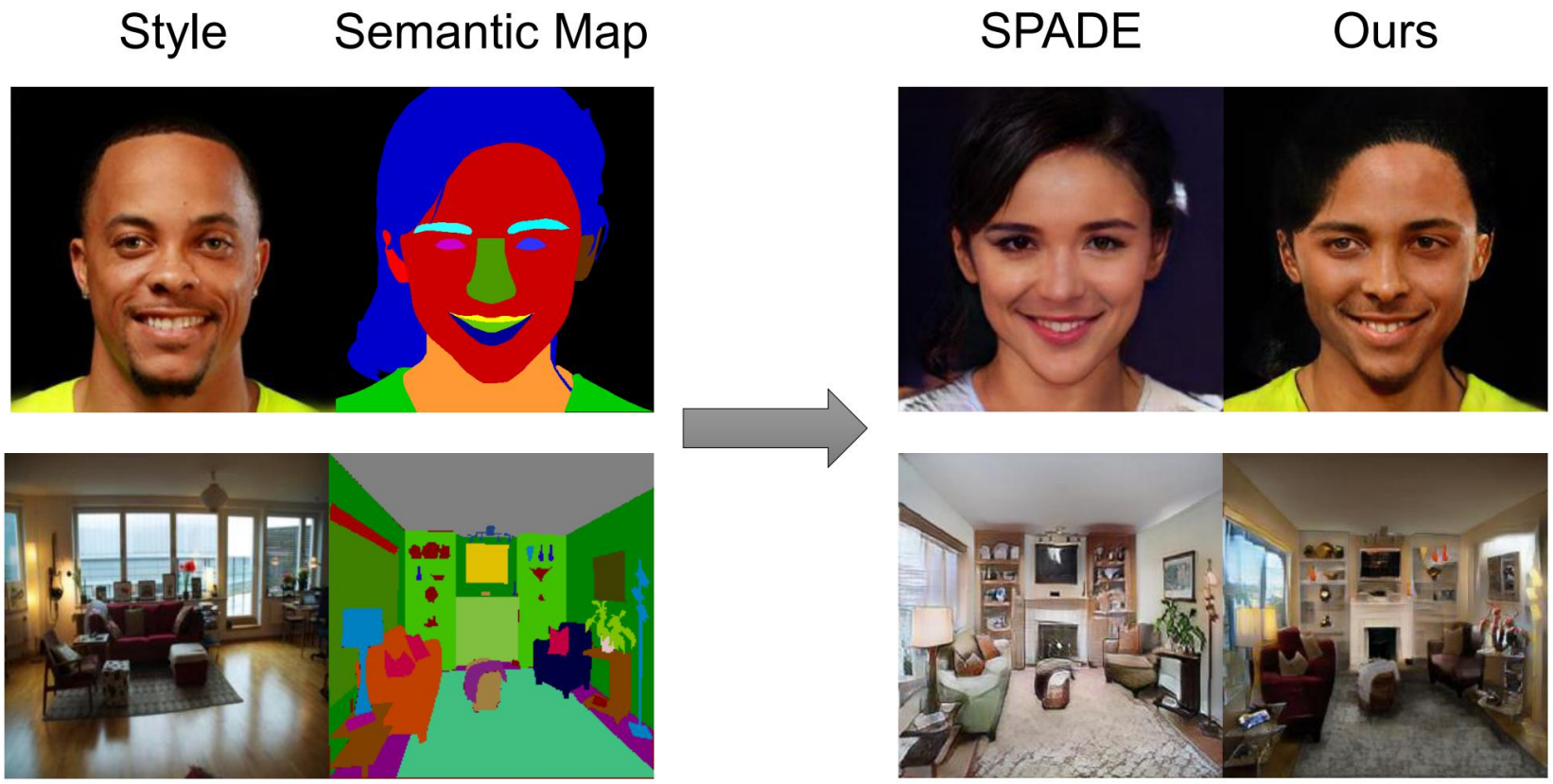


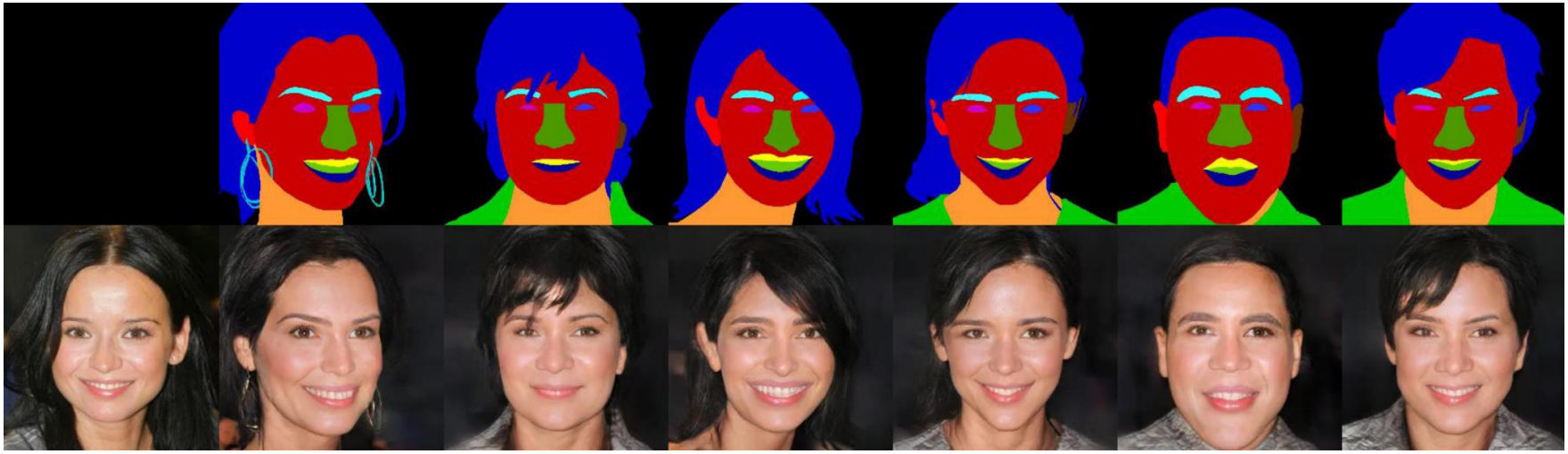
Two modifications




Style Image
Source Image


Style
Semantic Map
SPADE
Ours


Style
Semantic Map
SPADE
Ours



硕士生学术交流

内容安排

CVPR2020线上会议参会感受分享

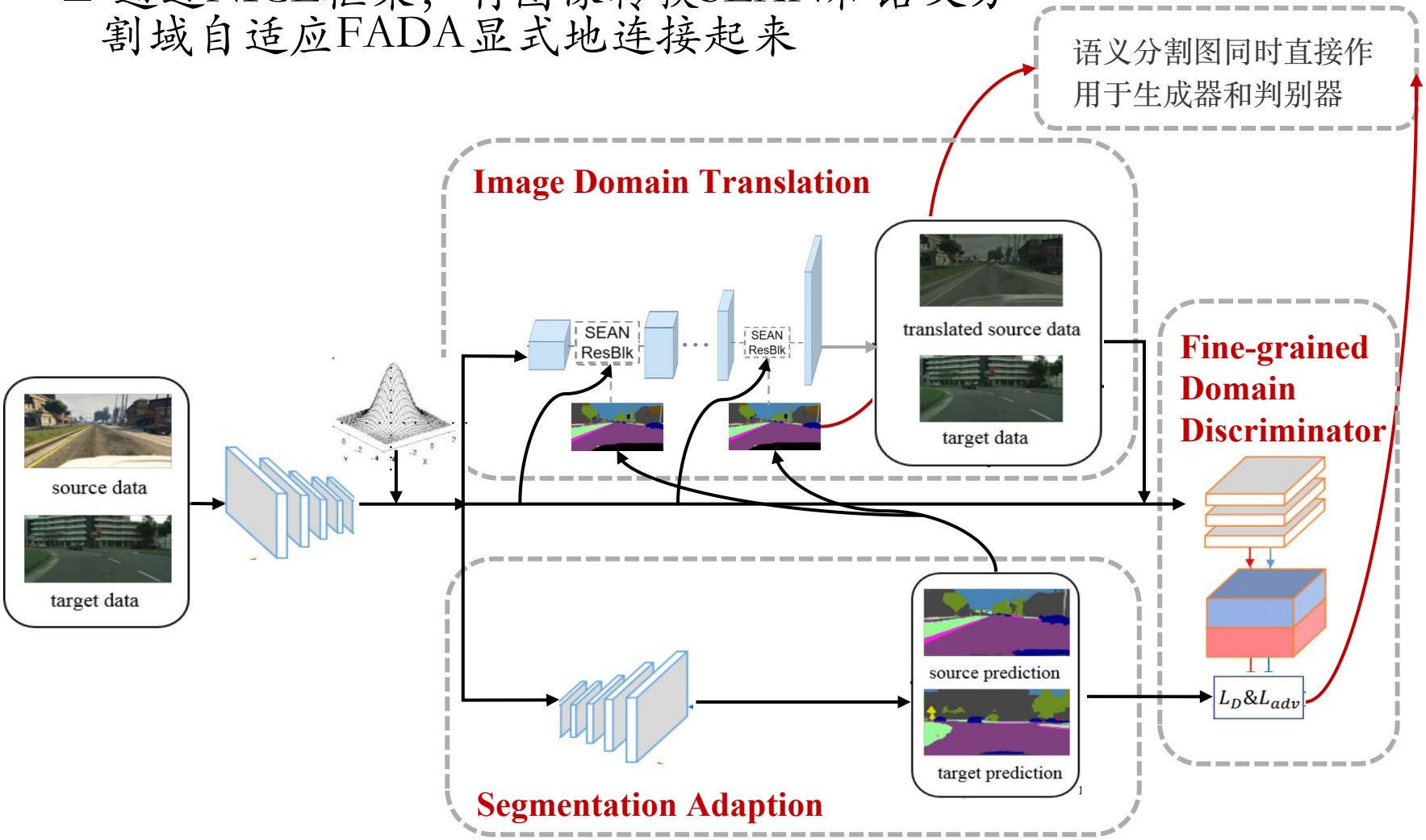
研究问题概述Problem Definition

Paper Reading 1: CVPR20 MixNMatch

Paper Reading 2: CVPR20 SEAN

Future Works

- 通过NICE框架，将图像转换SEAN和语义分割域自适应FADA显式地连接起来



谢谢！
请老师同学们批评指正！