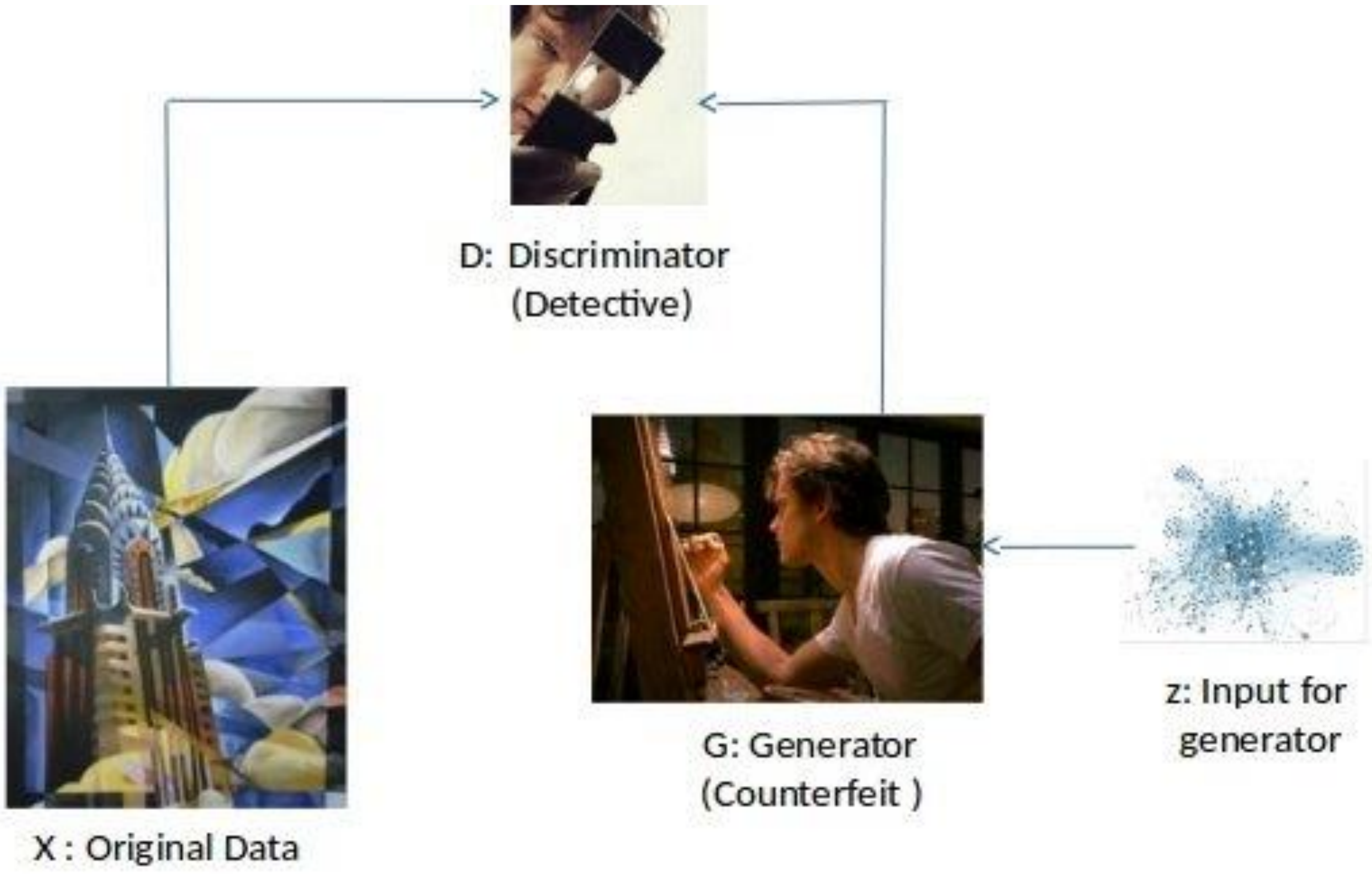
GAN Related Works

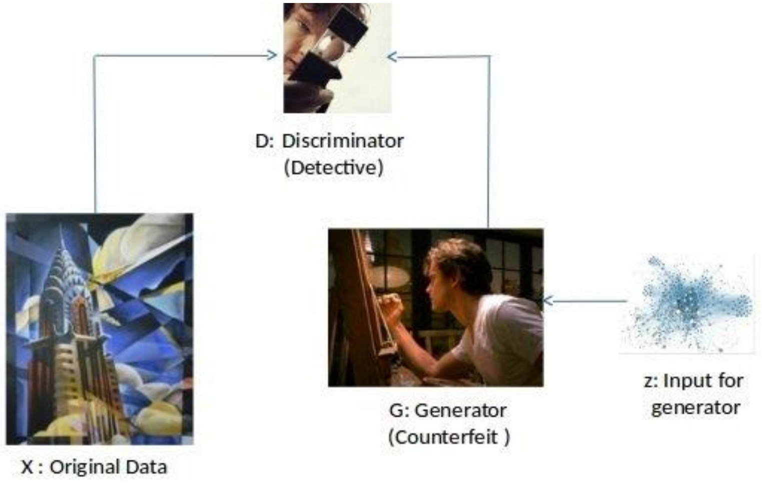
CVPR 2018 & Selective Works in ICML and NIPS

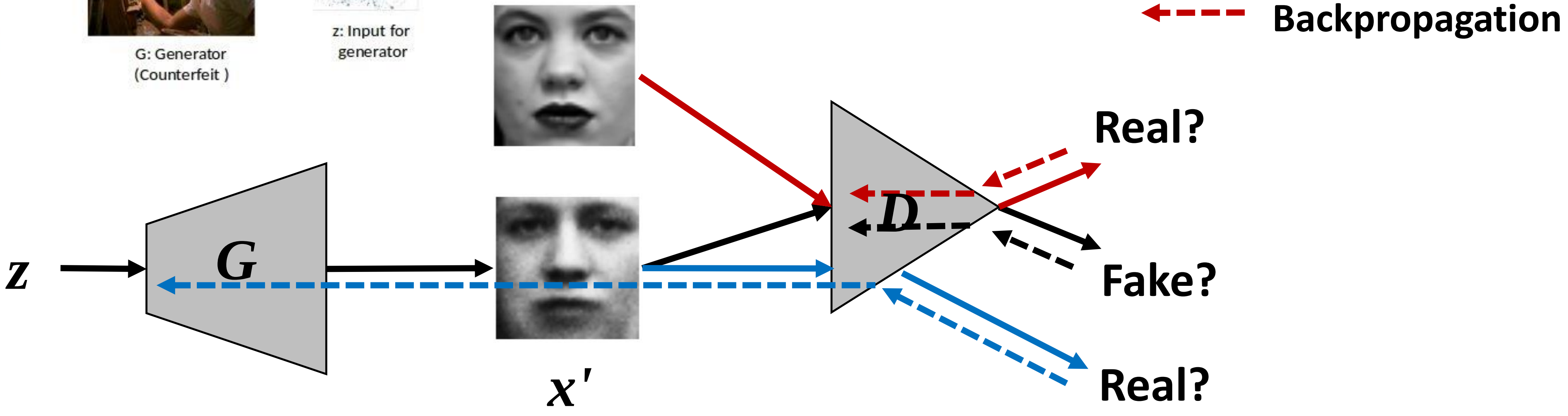
Zhifei Zhang

Generative Adversarial Networks (GANs)

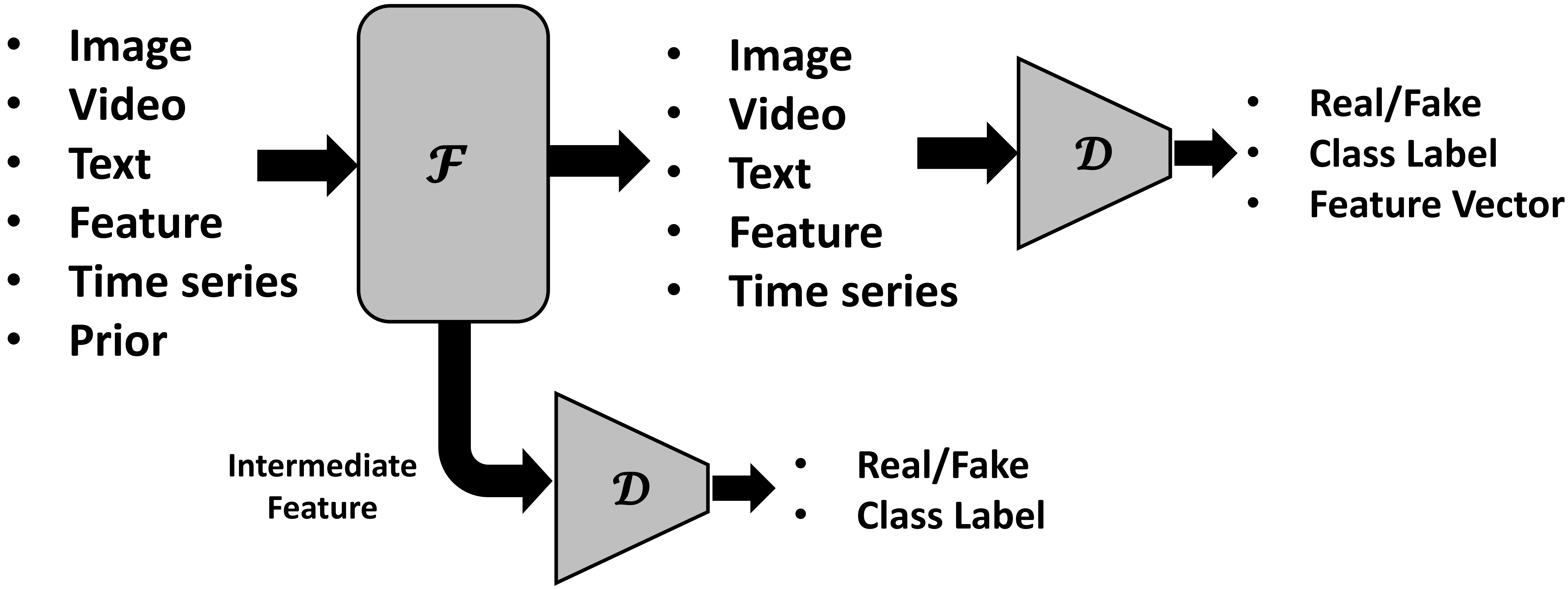


Generative Adversarial Networks (GANs)



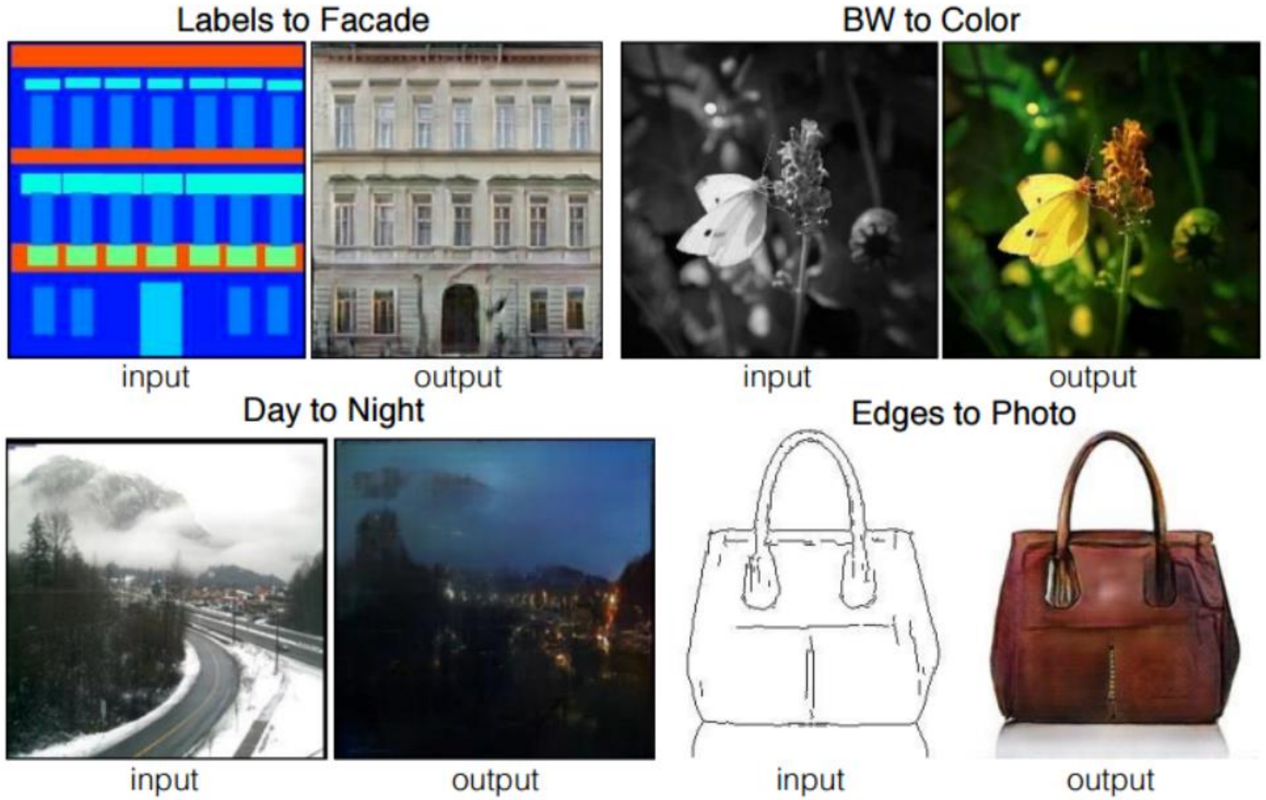


Generative Adversarial Networks (GANs)



Generative Adversarial Networks (GANs)

Image → $\mathcal{F}$ → Image



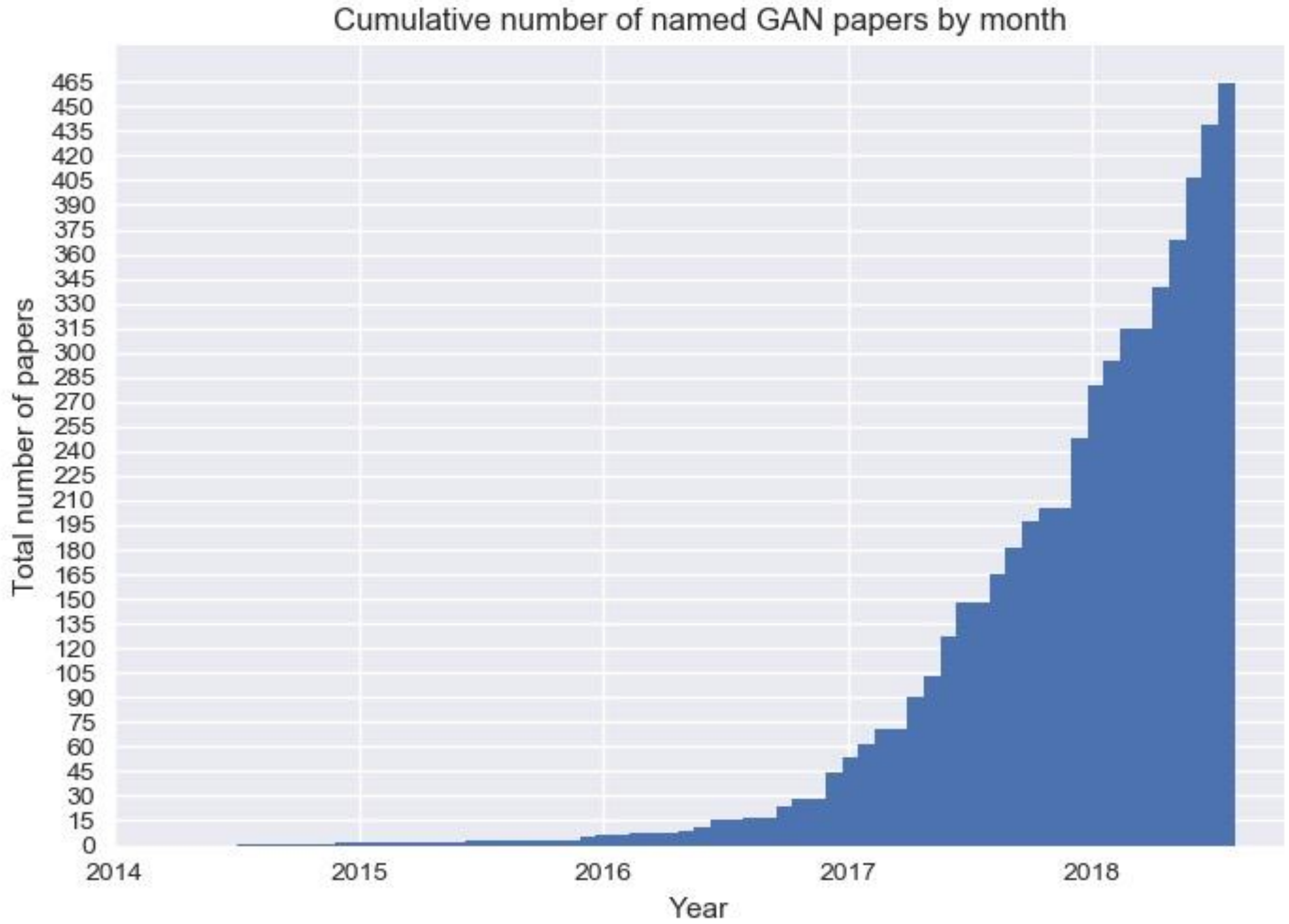
[Isola et al. CVPR2017]

Text → $\mathcal{F}$ → Image

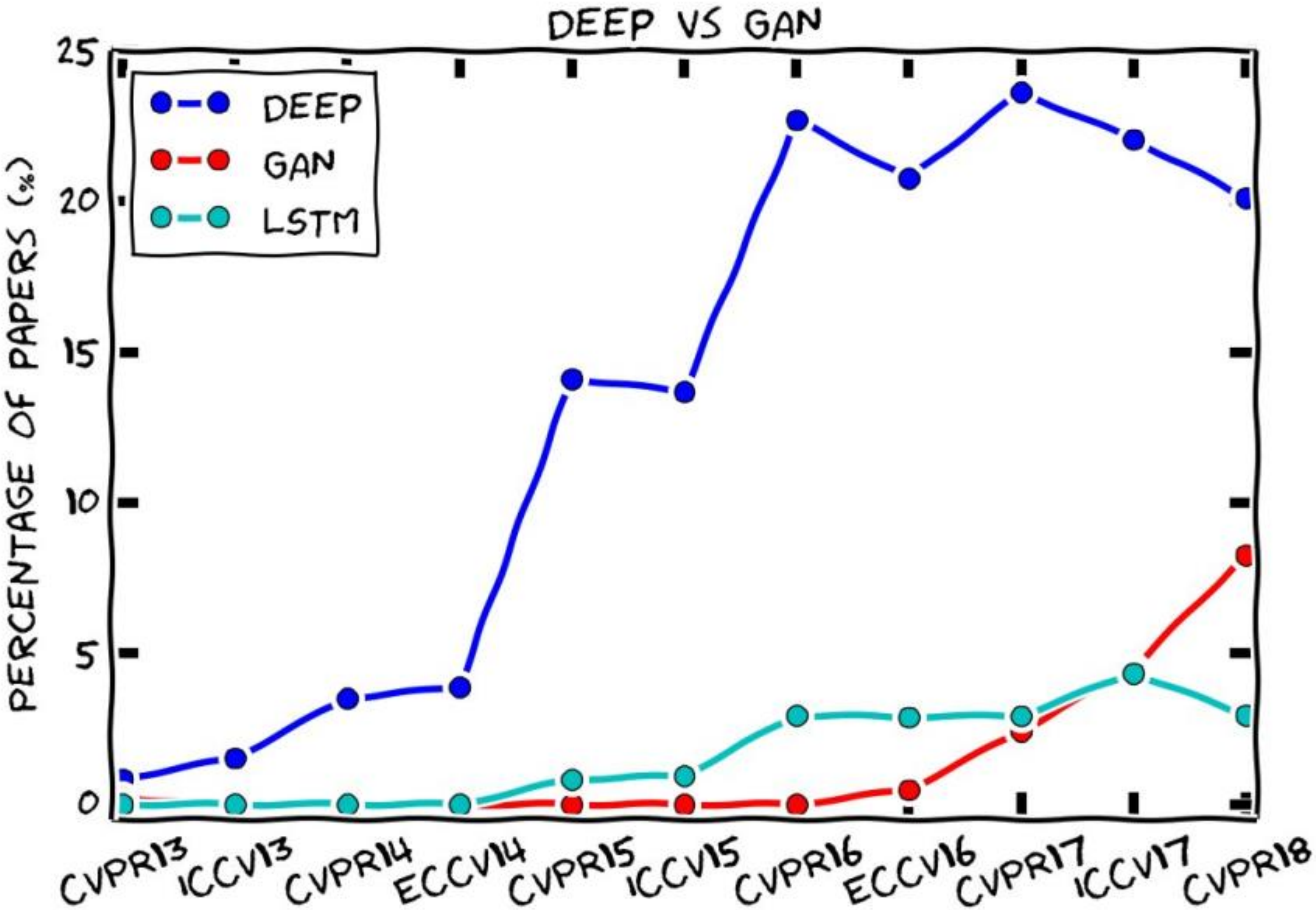
this bird is red with white and has a very short beak

[Xu et al. CVPR2018]

Generative Adversarial Networks (GANs)



https://github.com/hindupuravinash/the-gan-zoo

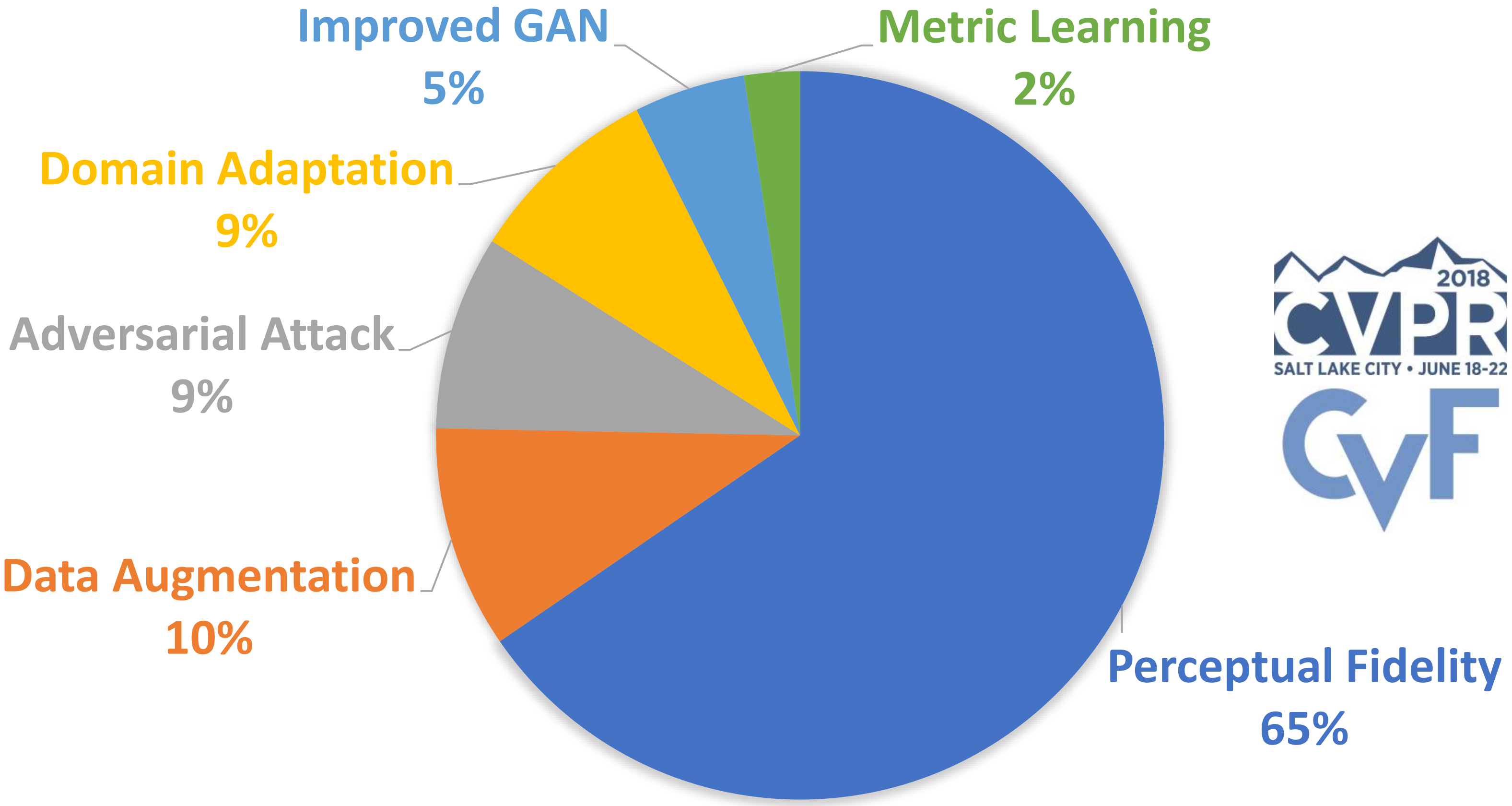


https://medium.com/syncedreview/cvpr-2018-kicks-off-best-papers-announced-d3361bcc6984

Applications of GANs

1. Perceptual Fidelity
2. Data Augmentation
3. Adversarial Attack
4. Domain Adaptation
5. Improved GAN
6. Metric Learning

Keywords : GAN, Adversarial → 81 papers



(1/6) Perceptual Fidelity

Learning Face Age Progression: A Pyramid Architecture of GANs

PairedCycleGAN: Asymmetric Style Transfer for Applying and Removing Makeup

Super-FAN: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with GANs

AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks

MoCoGAN: Decomposing Motion and Content for Video Generation

Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks

Deformable GANs for Pose-based Human Image Generation

Cross-View Image Synthesis using Conditional GANs

DA-GAN: Instance-level Image Translation by Deep Attention Generative Adversarial Networks

SeGAN: Segmenting and Generating the Invisible

Deep Photo Enhancer: Unpaired Learning for Image Enhancement from Photographs with GANs

UV-GAN: Adversarial Facial UV Map Completion for Pose-invariant Face Recognition

Multi-Content GAN for Few-Shot Font Style Transfer

From source to target and back: Symmetric Bi-Directional Adaptive GAN

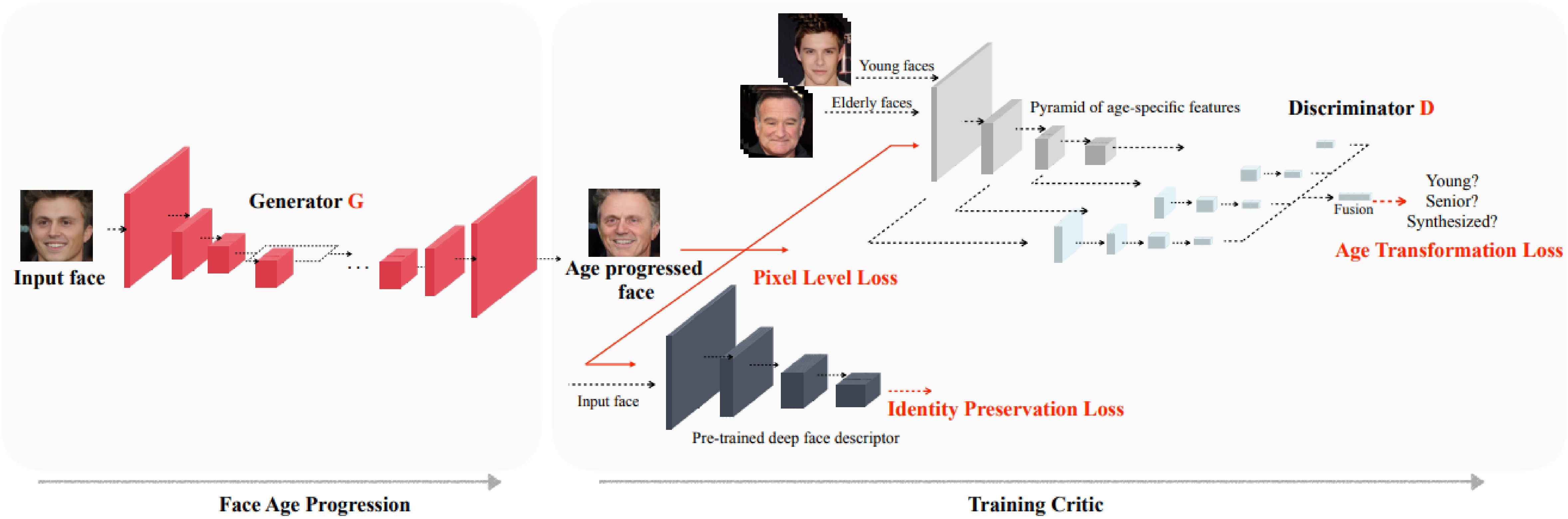
DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks

Self-Supervised Adversarial Hashing Networks for Cross-Modal Retrieval
Unsupervised Deep Generative Adversarial Hashing Network
Visual Feature Attribution using Wasserstein GANs
TextureGAN: Controlling Deep Image Synthesis with Texture Patches
StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation
High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs
SketchyGAN: Towards Diverse and Realistic Sketch to Image Synthesis
ST-GAN: Spatial Transformer Generative Adversarial Networks for Image Compositing
CartoonGAN: Generative Adversarial Networks for Photo Cartoonization
Finding Tiny Faces in the Wild with Generative Adversarial Network
Multistage Adversarial Losses for Pose-Based Human Image Synthesis
Hallucinated-IQA: No-Reference Image Quality Assessment via Adversarial Learning
A Generative Adversarial Approach for Zero-Shot Learning from Noisy Texts
Zero-Shot Visual Recognition using Semantics-Preserving Adversarial Embedding Networks
Adversarial Complementary Learning for Weakly Supervised Object Localization
Conditional Generative Adversarial Network for Structured Domain Adaptation
Duplex Generative Adversarial Network for Unsupervised Domain Adaptation
Deep Adversarial Subspace Clustering
Stacked Conditional Generative Adversarial Networks for Jointly Learning Shadow Detection and Shadow Removal

Weakly Supervised Facial Action Unit Recognition through Adversarial Training
Learning to Generate Time-Lapse Videos Using Multi-Stage Dynamic Generative Adversarial Networks
Attentive Generative Adversarial Network for Raindrop Removal from A Single Image
FaceID-GAN: Learning a Symmetry Three-Player GAN for Identity-Preserving Face Synthesis
GAGAN: Geometry-Aware Generative Adversarial Networks
Adversarially Learned One-Class Classifier for Novelty Detection
3D Human Pose Estimation in the Wild by Adversarial Learning
Crowd Counting via Adversarial Cross-Scale Consistency Pursuit
Generative Adversarial Learning Towards Fast Weakly Supervised Detection
Logo Synthesis and Manipulation with Clustered Generative Adversarial Networks
Are You Talking to Me? Reasoned Visual Dialog Generation through Adversarial Learning
Photographic Text-to-Image Synthesis with a Hierarchically-nested Adversarial Network
Label Denoising Adversarial Network (LDAN) for Inverse Lighting of Faces
Generative Adversarial Image Synthesis with Decision Tree Latent Controller
Eye In-Painting with Exemplar Generative Adversarial Networks
Face Aging with Identity-Preserved Conditional Generative Adversarial Networks
Single Image Dehazing via Conditional Generative Adversarial Network
VITAL: VIsual Tracking via Adversarial Learning
Translating and Segmenting Multimodal Medical Volumes with Cycle- and Shape-Consistency Generative Adversarial Network

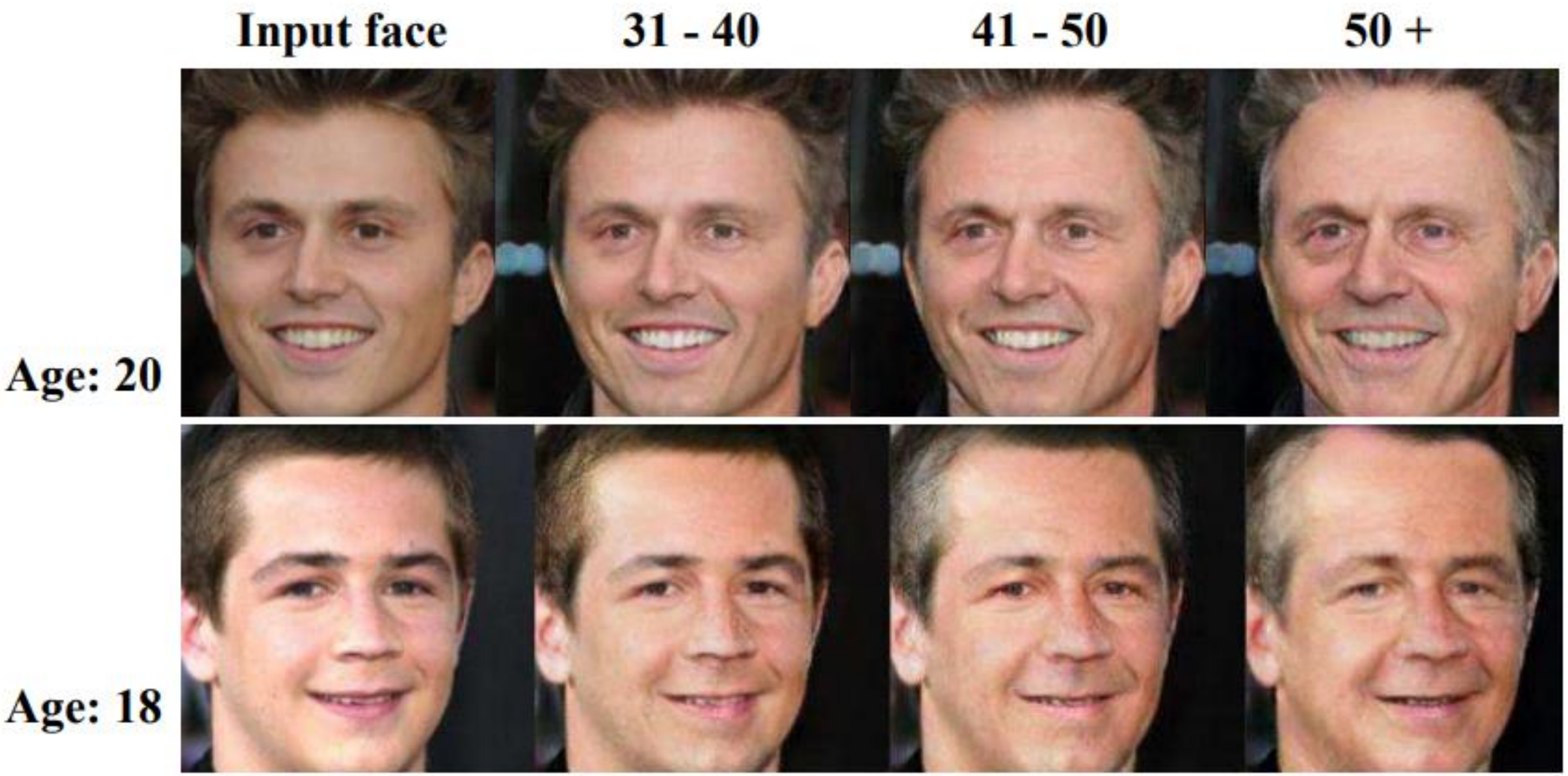
(1/6) Perceptual Fidelity

The discriminator has become a common regularization for image fidelity, and it tends to be multi-task.



[Yang et al. CVPR2018] Learning Face Age Progression: A Pyramid Architecture of GANs

(1/6) Perceptual Fidelity



[Yang et al. CVPR2018] Learning Face Age Progression: A Pyramid Architecture of GANs

(2/6) Data Augmentation

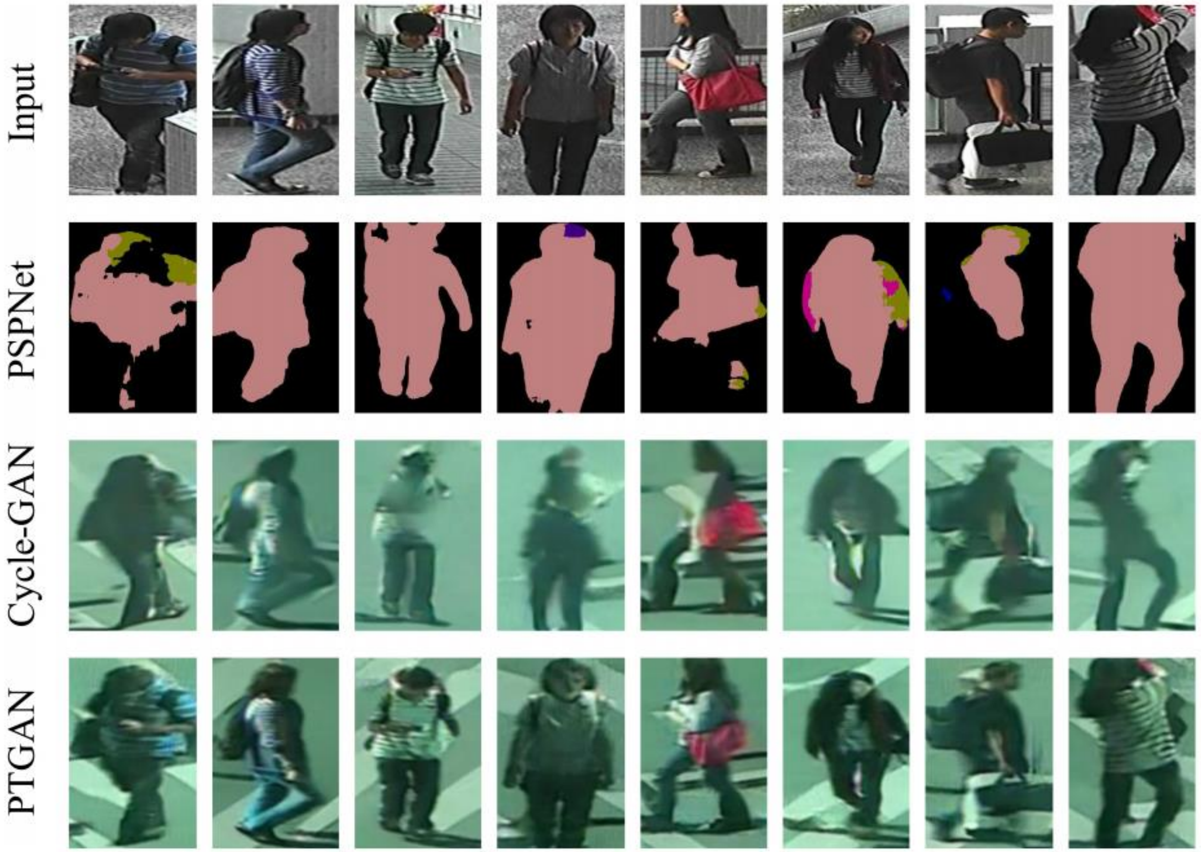
GANerated Hands for Real-Time 3D Hand Tracking from Monocular RGB
Person Transfer GAN to Bridge Domain Gap for Person Re-Identification (transfer data from other dataset)
HashGAN: Deep Learning to Hash with Pair Conditional Wasserstein GAN
Adversarial Data Programming: Using GANs to Relax the Bottleneck of Curated Labeled Data
Jointly Optimize Data Augmentation and Network Training: Adversarial Data Augmentation in Human Pose Estimation
Image Blind Denoising With Generative Adversarial Network Based Noise Modeling
SINT++: Robust Visual Tracking via Adversarial Positive Instance Generation
Adversarially Occluded Samples for Person Re-identification

Roughly, there are two ways of data augmentation:

- **Generate training data**
- **Transfer data from other dataset (1 paper)**

(2/6) Data Augmentation

Figure 1: Illustration of the domain gap between *CUHK03* and *PRID*. It is obvious that, *CUHK03* and *PRID* present different styles, *e.g.*, distinct lightings, resolutions, human race, seasons, backgrounds, *etc.*, resulting in low accuracy when training on *CUHK03* and testing on *PRID*.



[Wei et al. CVPR2018] Person Transfer GAN to Bridge Domain Gap for Person Re-Identification

(3/6) Adversarial Attack

On the Robustness of Semantic Segmentation Models to Adversarial Attacks

Defense against Adversarial Attacks Using High-Level Representation Guided Denoiser

Defense against Universal Adversarial Perturbations

Generative Adversarial Perturbations

Art of singular vectors and universal adversarial perturbations

Deflecting Adversarial Attacks with Pixel Deflection

Boosting Adversarial Attacks with Momentum

Attack vs. Protection

(3/6) Adversarial Attack

Universal perturbation

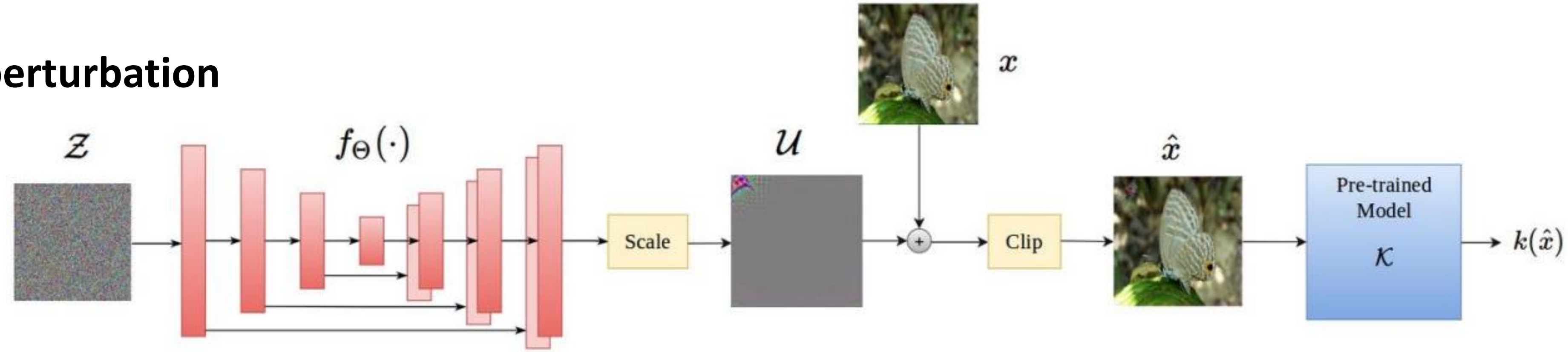
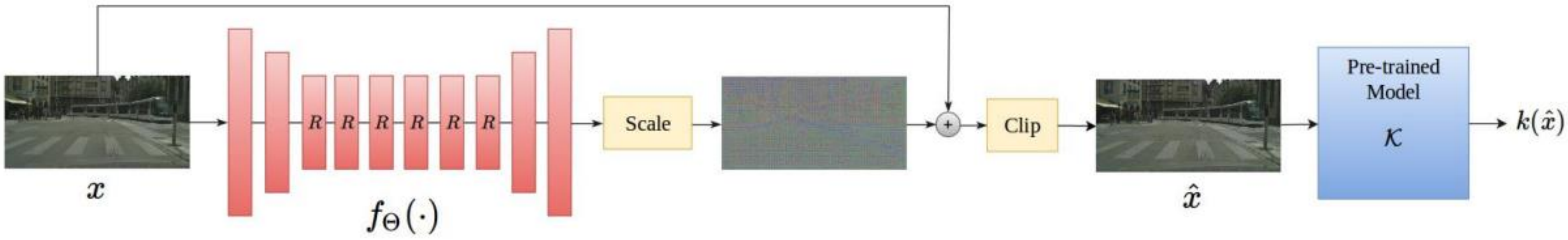
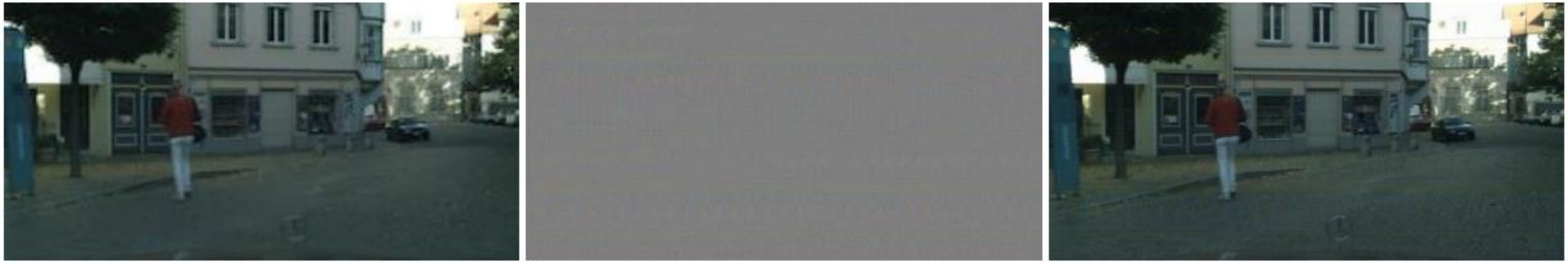


Image-dependent perturbation



[Poursaeed et al. CVPR2018] Generative Adversarial Perturbations

(3/6) Adversarial Attack

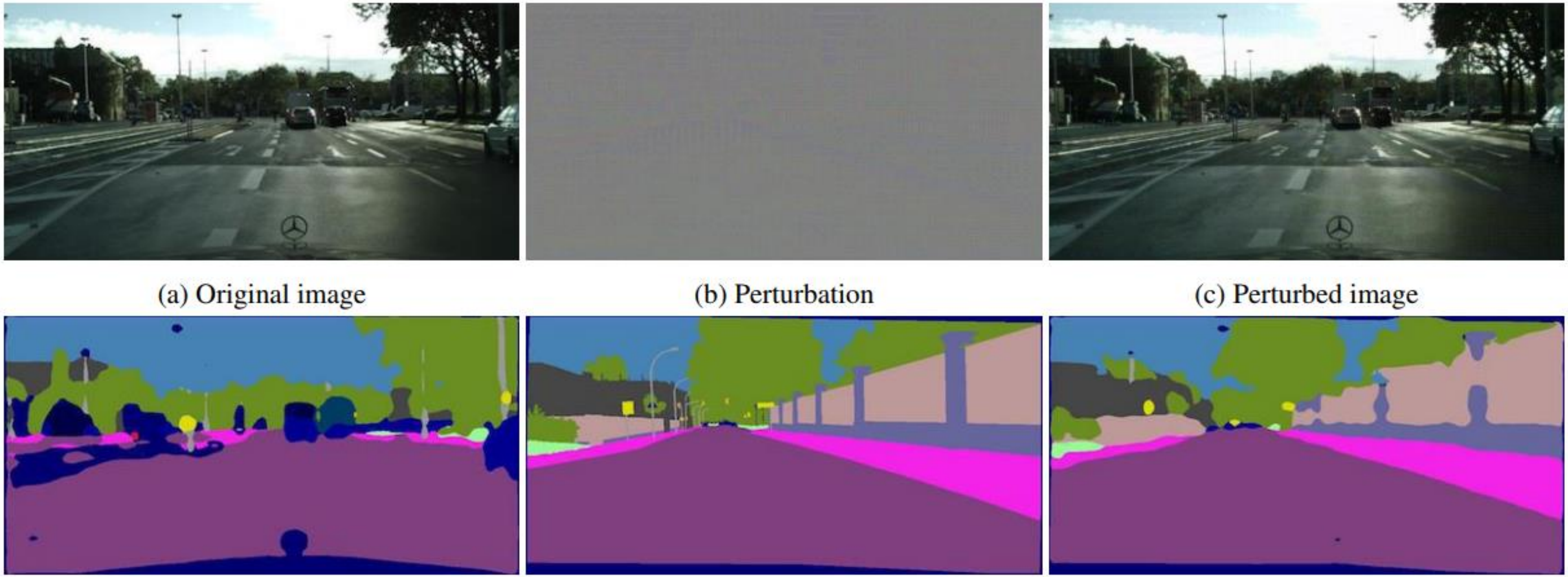


(a) Original image (b) Perturbation (c) Perturbed image

(d) Prediction for original image (e) Groundtruth (f) Prediction for perturbed image

[Poursaeed et al. CVPR2018] Generative Adversarial Perturbations

(3/6) Adversarial Attack

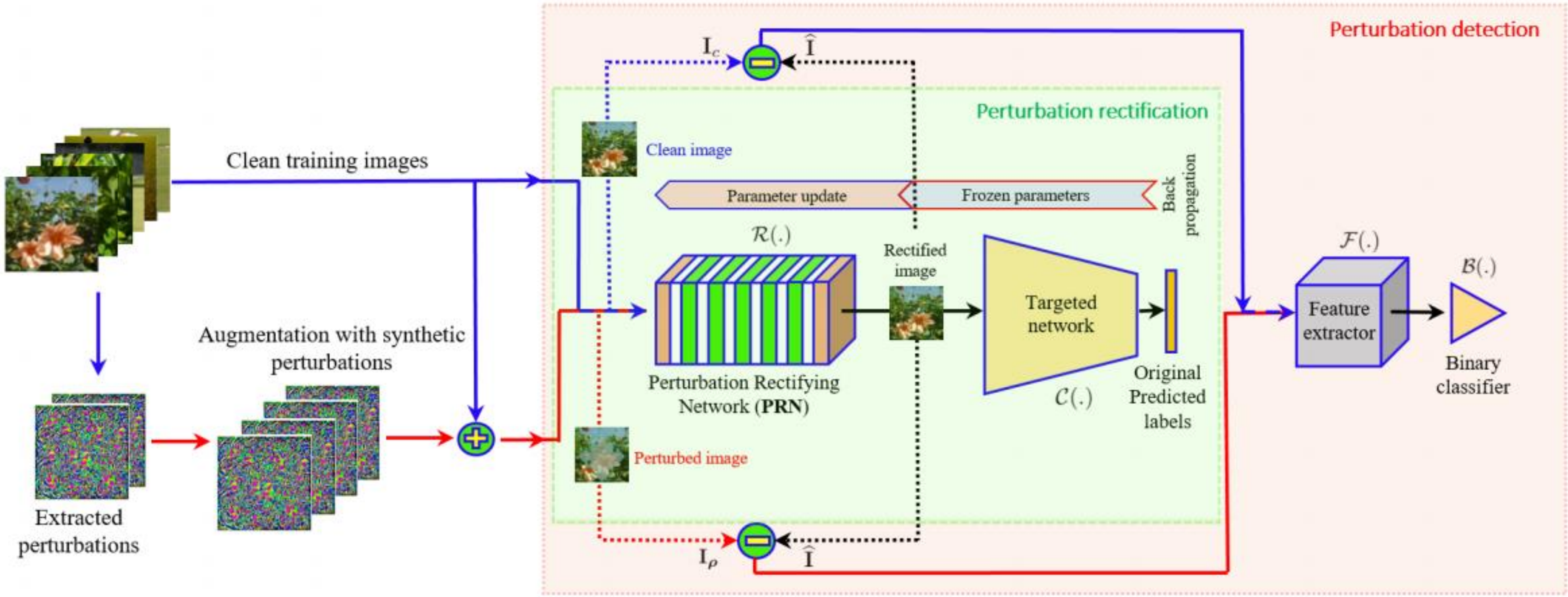


(a) Original image (b) Perturbation (c) Perturbed image

(d) Prediction for original image (e) Target (f) Prediction for perturbed image

[Poursaeed et al. CVPR2018] Generative Adversarial Perturbations

(3/6) Adversarial Attack



[Akhtar et al. CVPR2018] Defense against Universal Adversarial Perturbations

(4/6) Domain Adaptation

Partial Transfer Learning with Selective Adversarial Networks

Collaborative and Adversarial Network for Unsupervised domain adaptation

Domain Generalization with Adversarial Feature Learning

Adversarial Feature Augmentation for Unsupervised Domain Adaptation

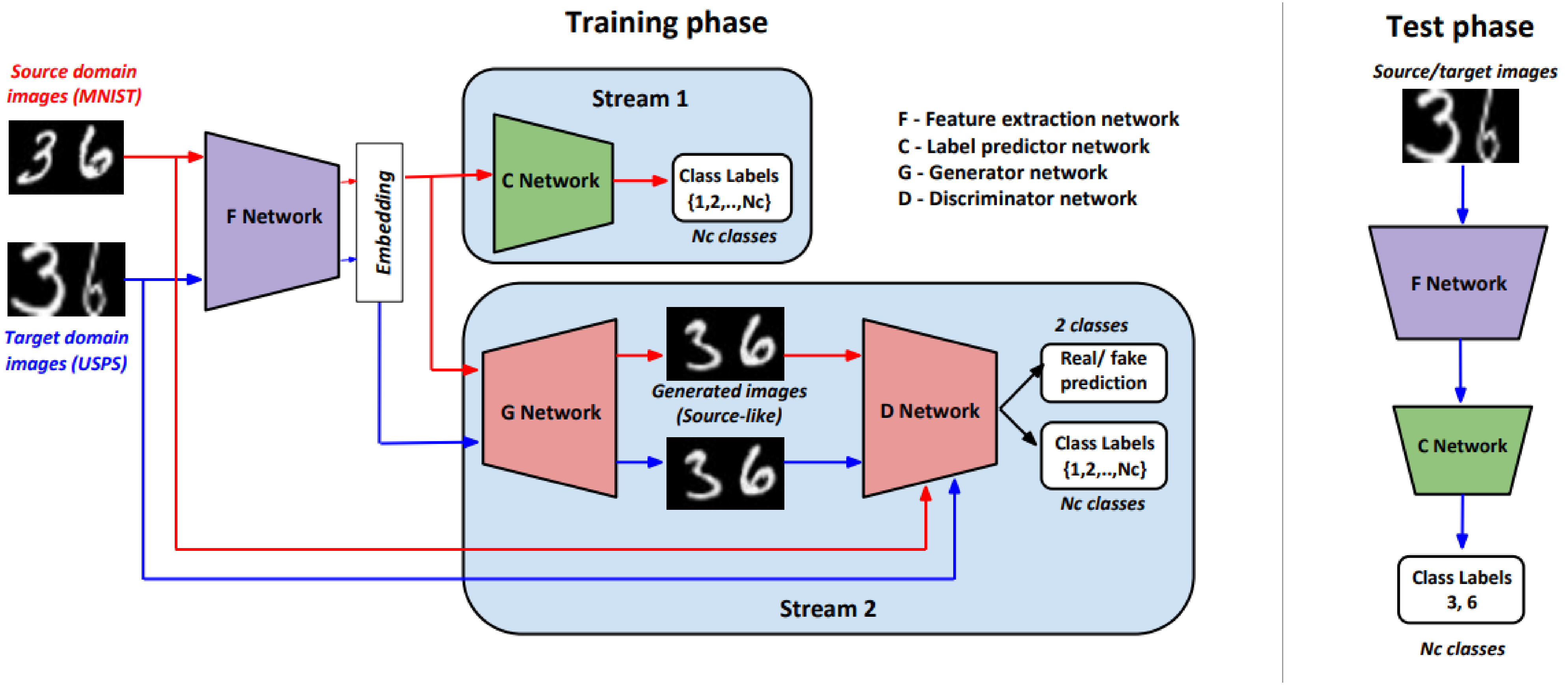
Re-weighted Adversarial Adaptation Network for Unsupervised Domain

Importance Weighted Adversarial Nets for Partial Domain Adaptation

Generate To Adapt: Aligning Domains using Generative Adversarial Networks

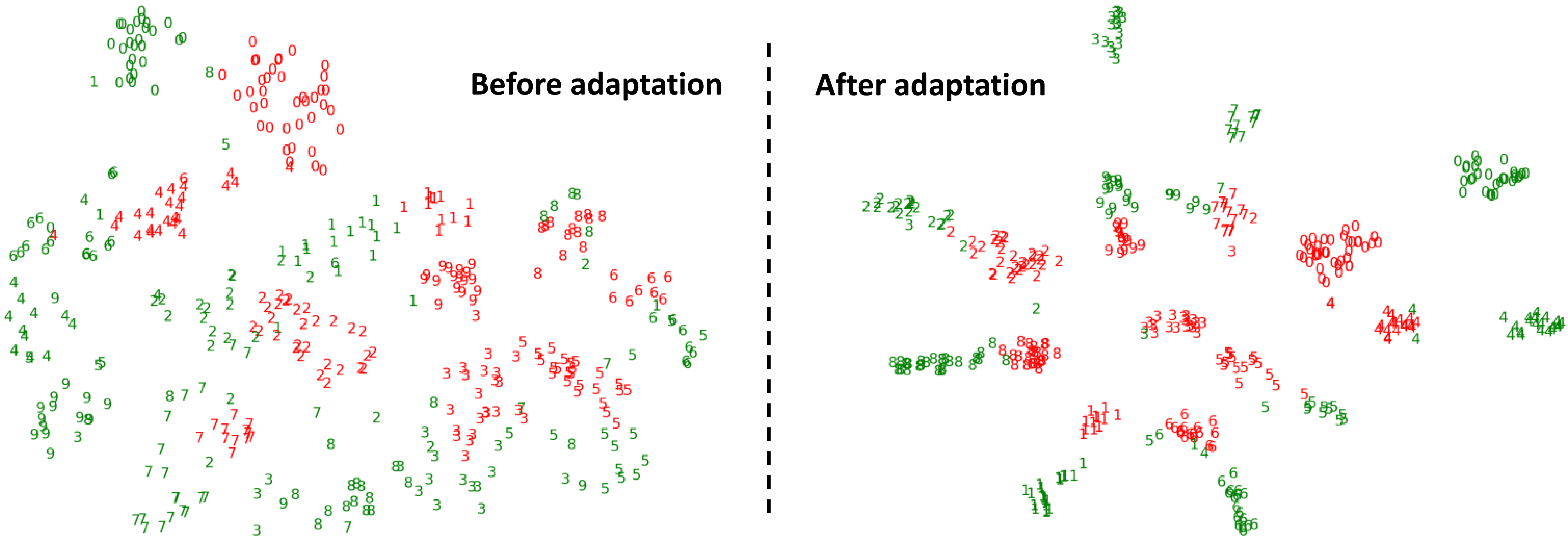
Align two domains in the feature space

(4/6) Transfer Learning (Domain Adaptation)



[Sankaranarayanan et al. CVPR2018] Generate To Adapt: Aligning Domains using Generative Adversarial Networks

(4/6) Transfer Learning (Domain Adaptation)

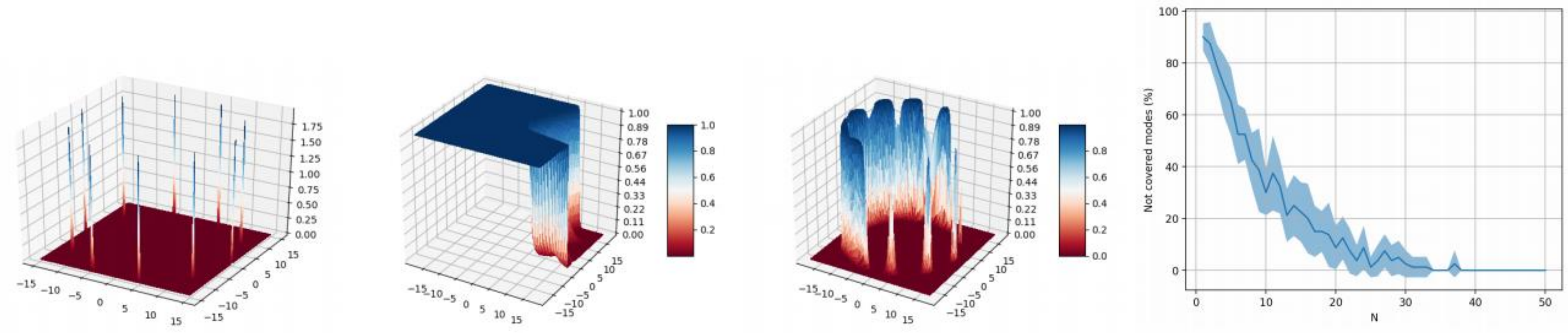


SVHN → MNIST

[Sankaranarayanan et al. CVPR2018] Generate To Adapt: Aligning Domains using Generative Adversarial Networks

(5/6) Improved GAN

SGAN: An Alternative Training of Generative Adversarial Networks
Global versus Localized Generative Adversarial Nets
Matching Adversarial Networks
Multi-Agent Diverse Generative Adversarial Networks

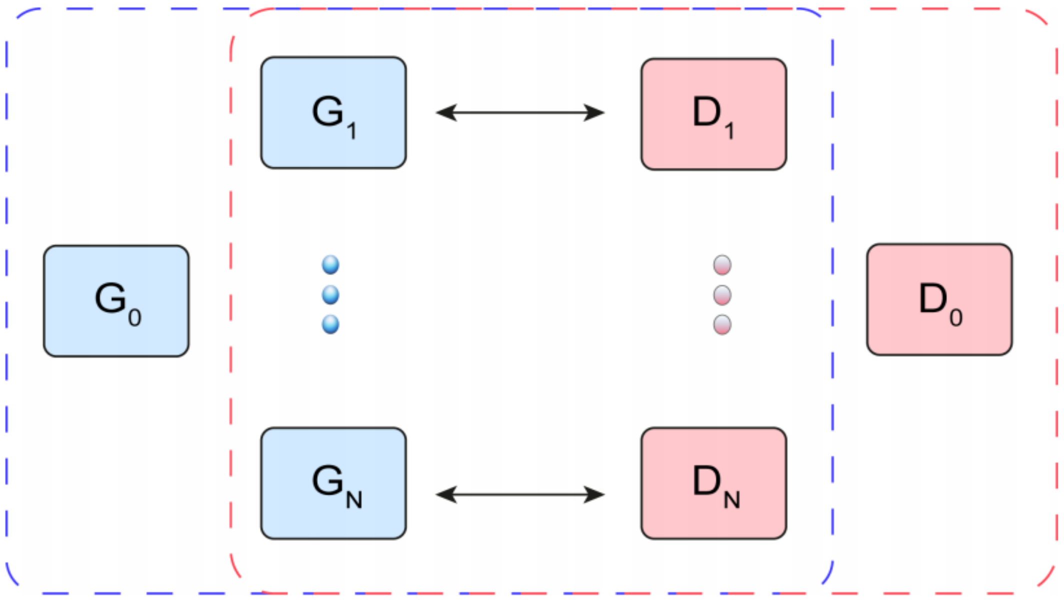


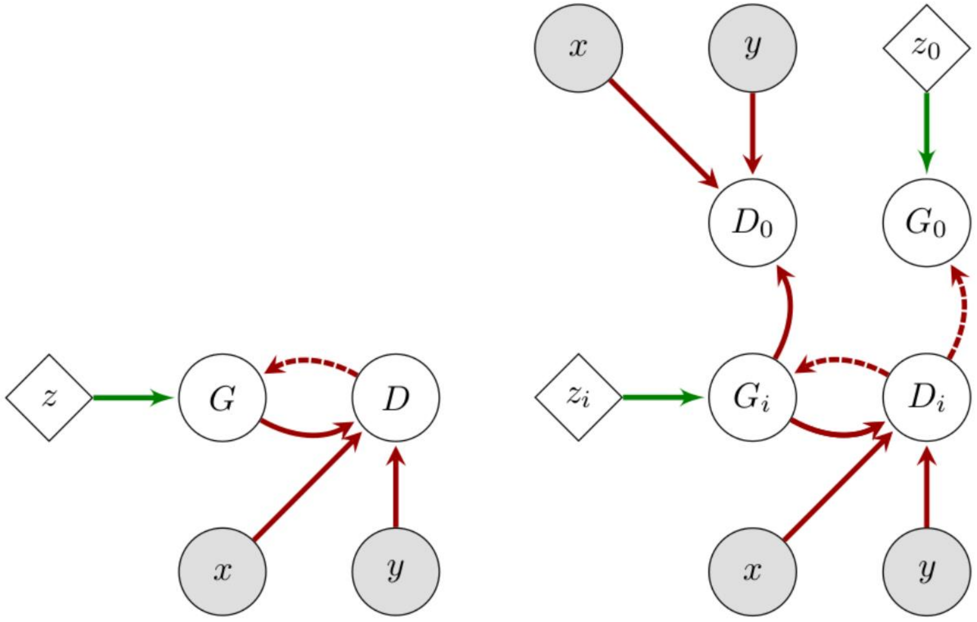
(a) Real data (10-GMM) (b) Discriminator output (c) S-Discriminator output (d) Not-covered modes (%)

[Chavdarova and Fleuret, CVPR2018] SGAN: An Alternative Training of Generative Adversarial Networks

The probability that a mode of the real data will not be covered goes down exponentially as the number of independent training pairs increases.

(5/6) Improved GAN





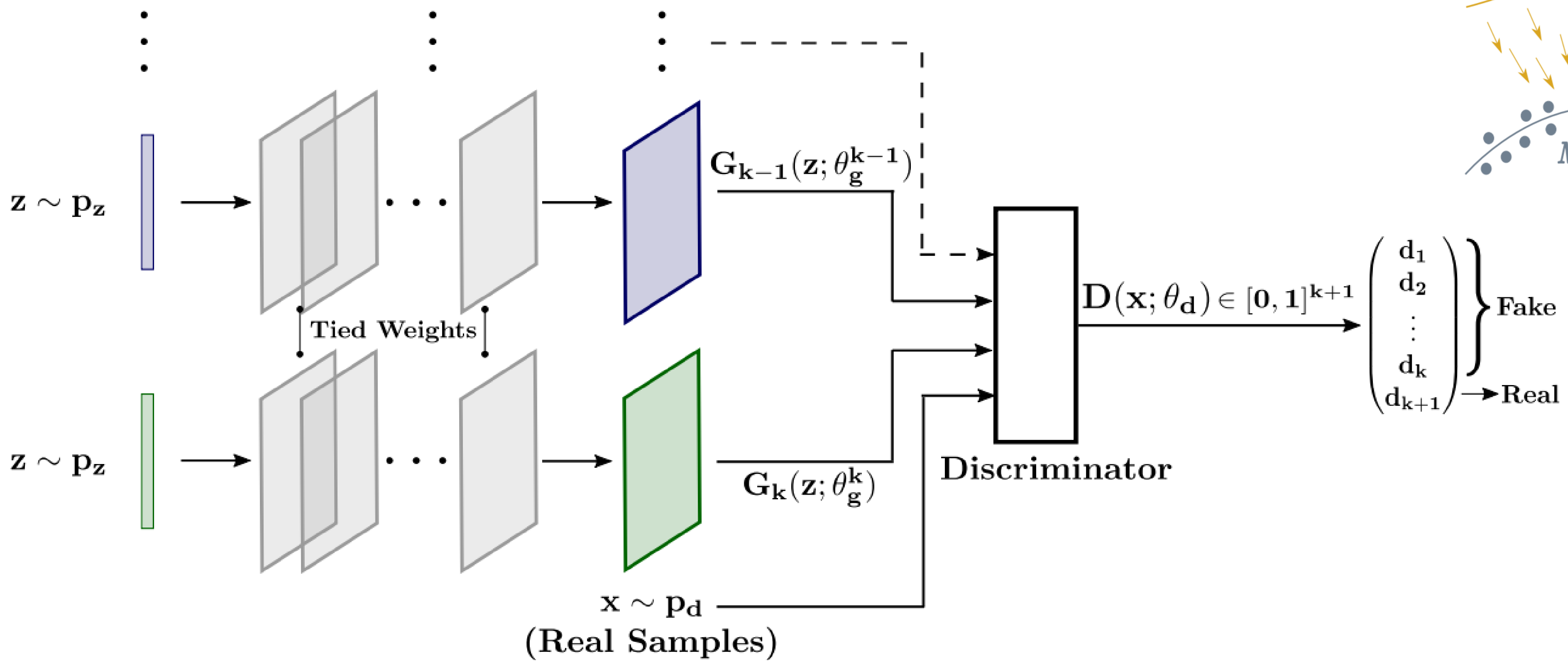
(a) GAN training (b) SGAN training

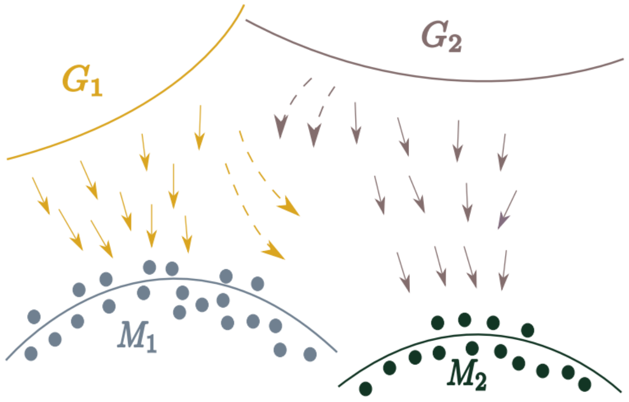
Questions:

- **What is the computational complexity?**
- **Why is D_0 needed as it is not related the training of G_0?**

[Chavdarova and Fleuret, CVPR2018] SGAN: An Alternative Training of Generative Adversarial Networks

(5/6) Improved GAN





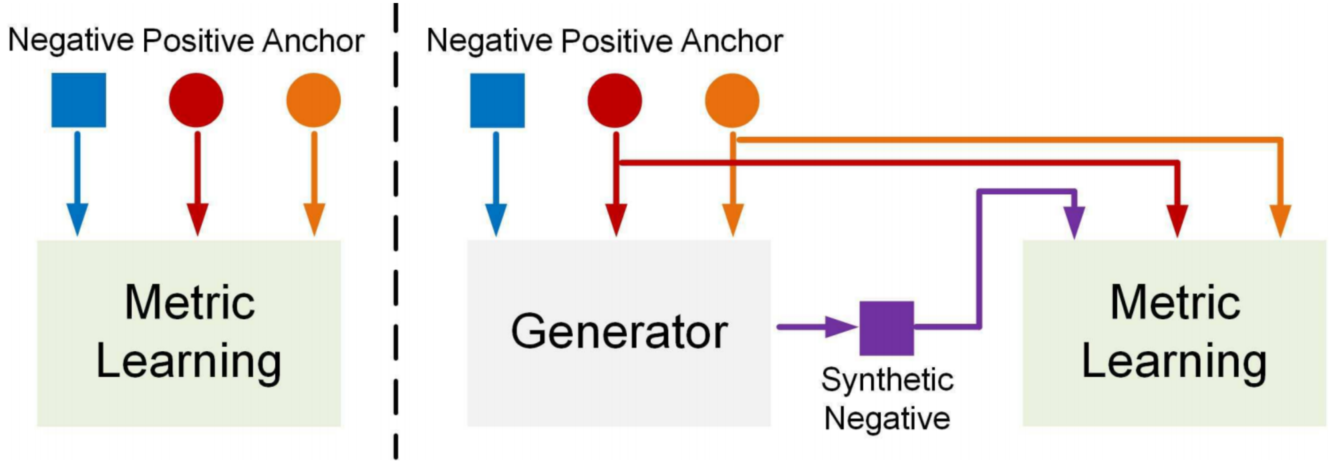
[Ghosh et al. CVPR2018] Multi-Agent Diverse Generative Adversarial Networks

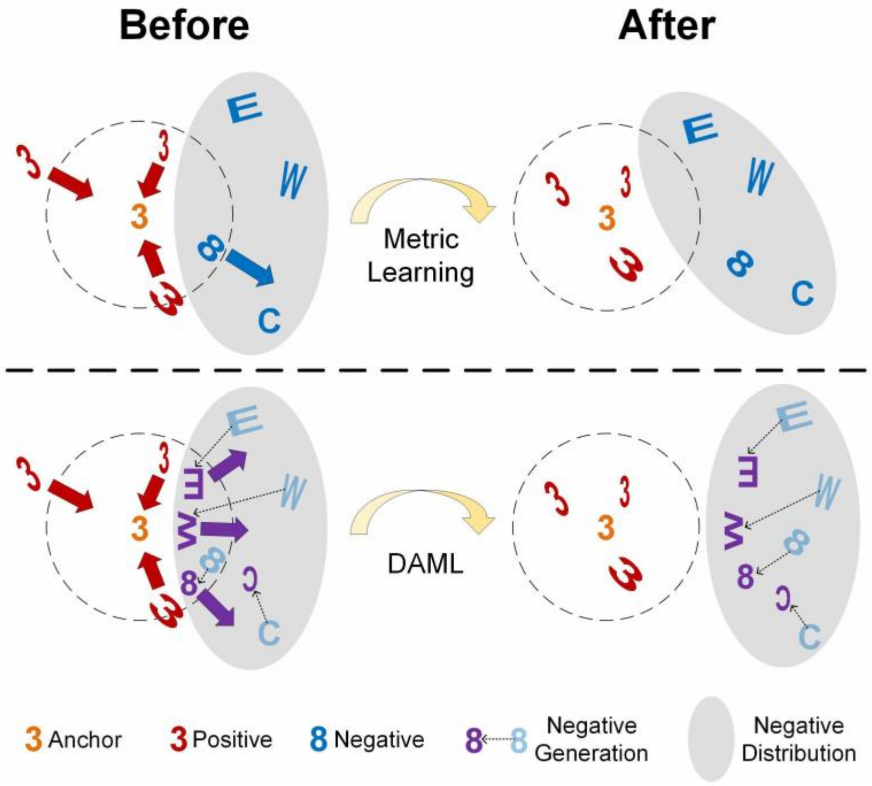
(6/6) Metric Learning

Deep Adversarial Metric Learning

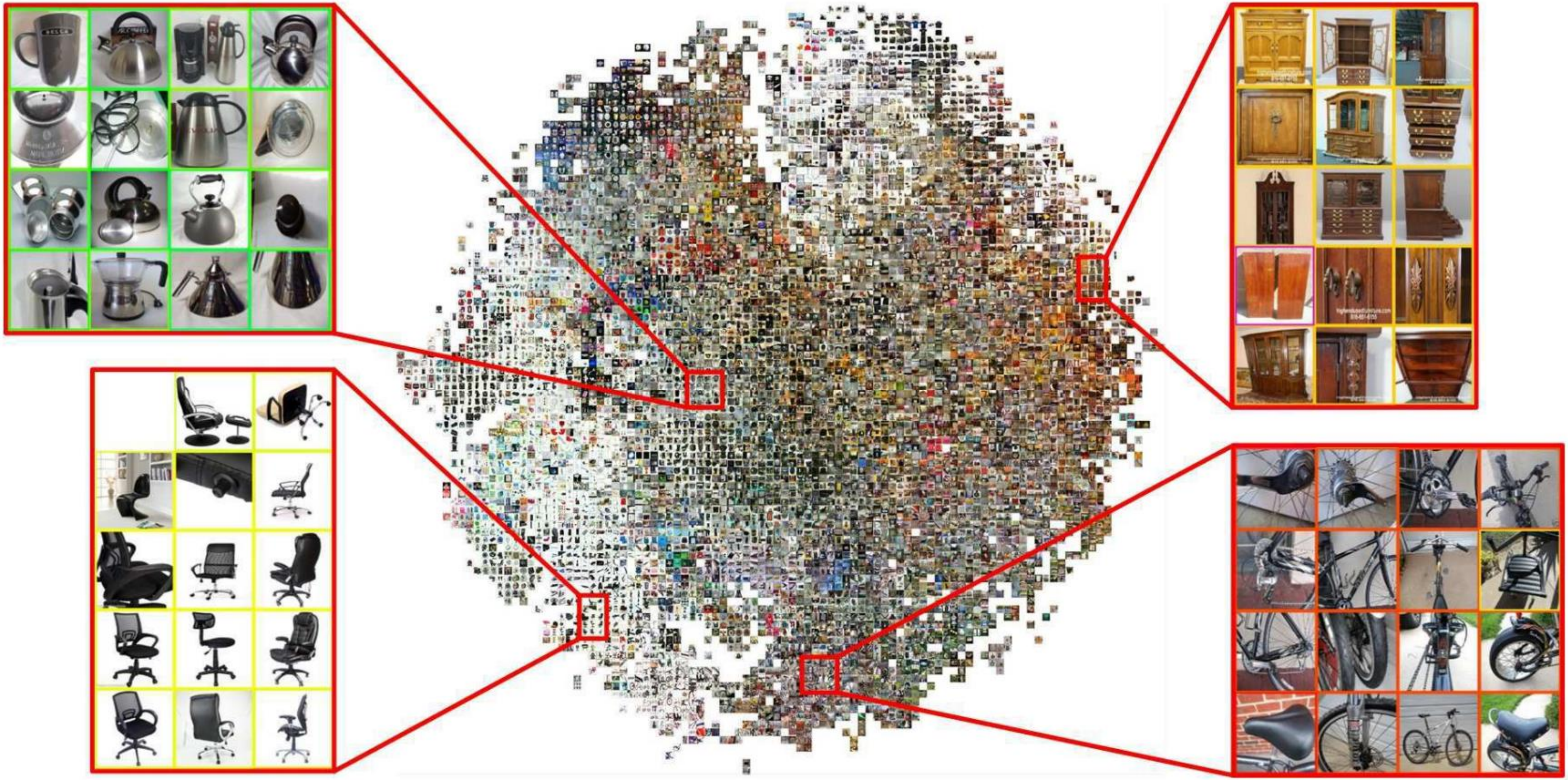
Multi-Task Adversarial Network for Disentangled Feature Learning

The training procedure largely relies on hard negative samples





[Duan et al. CVPR2018] Deep Adversarial Metric Learning



[Duan et al. CVPR2018] Deep Adversarial Metric Learning

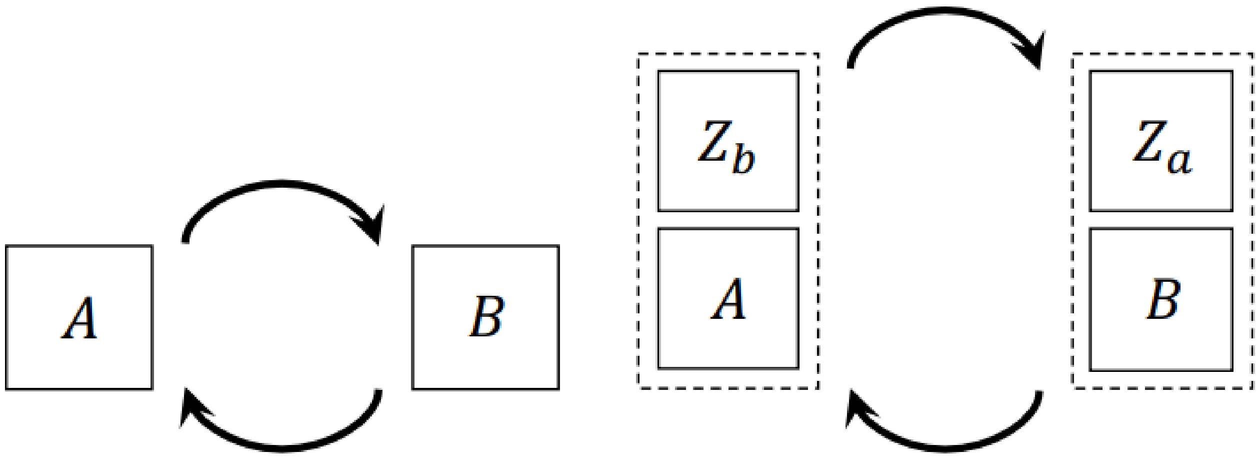
Some Interesting Works

- **Many-to-Many Domain Transfer**
- **Evaluation Metric for GANs**
- **Self-Attention GAN**
- **Video-to-Video Synthesis**

[1/4] Many-to-Many Domain Transfer

Motivation: Relax the limitation of one-to-one mapping

Idea: CycleGAN (unpaired training) + Domain condition



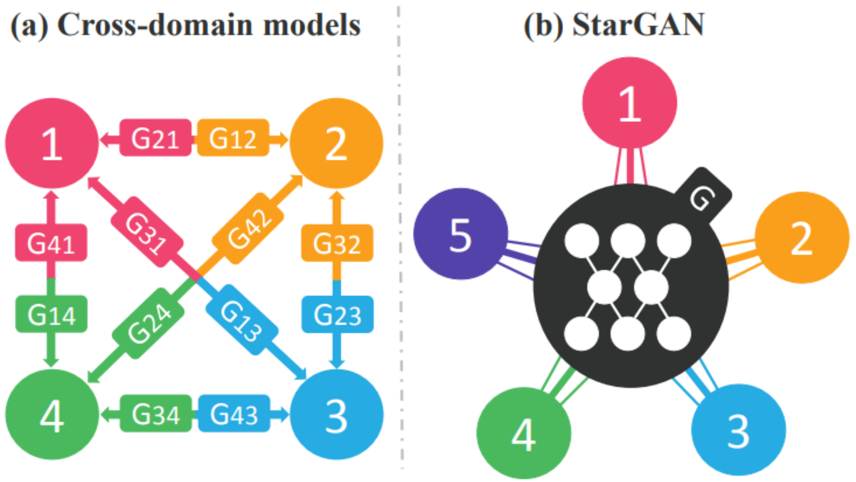
(a) CycleGAN (b) Augmented CycleGAN

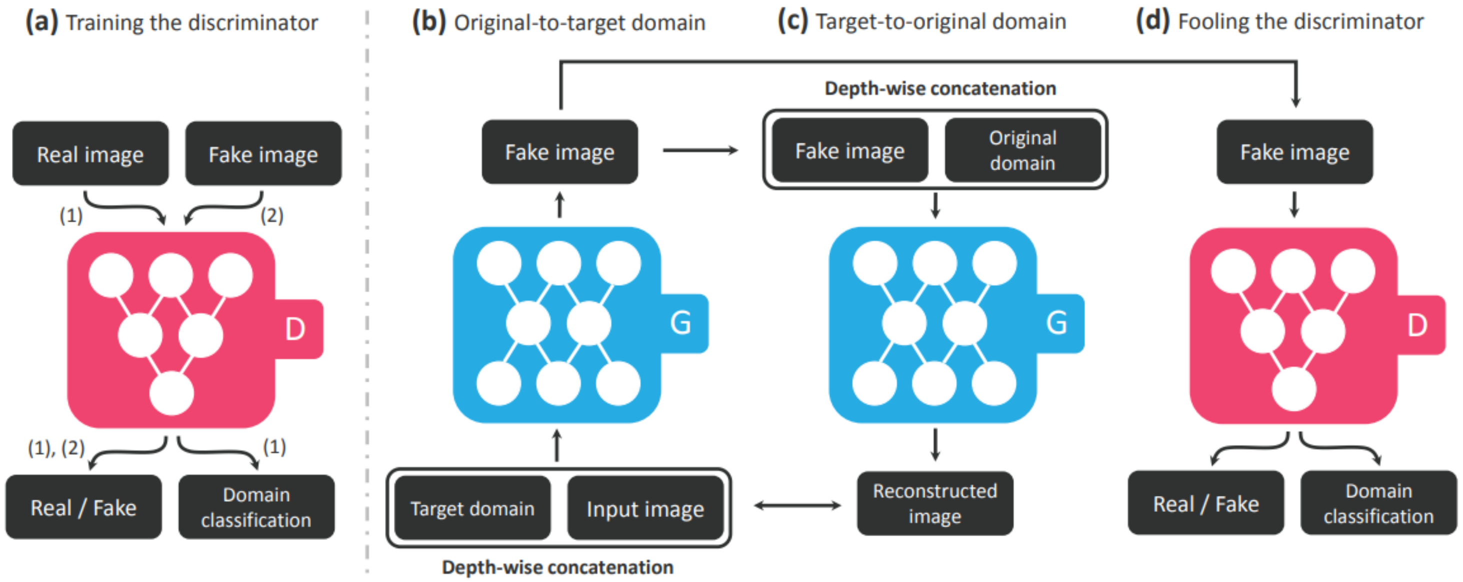
[Almahairi et al. ICML 2018] Augmented CycleGAN: Learning Many-to-Many Mappings from Unpaired Data

[1/4] Many-to-Many Domain Transfer

Motivation: Relax the limitation of one-to-one mapping

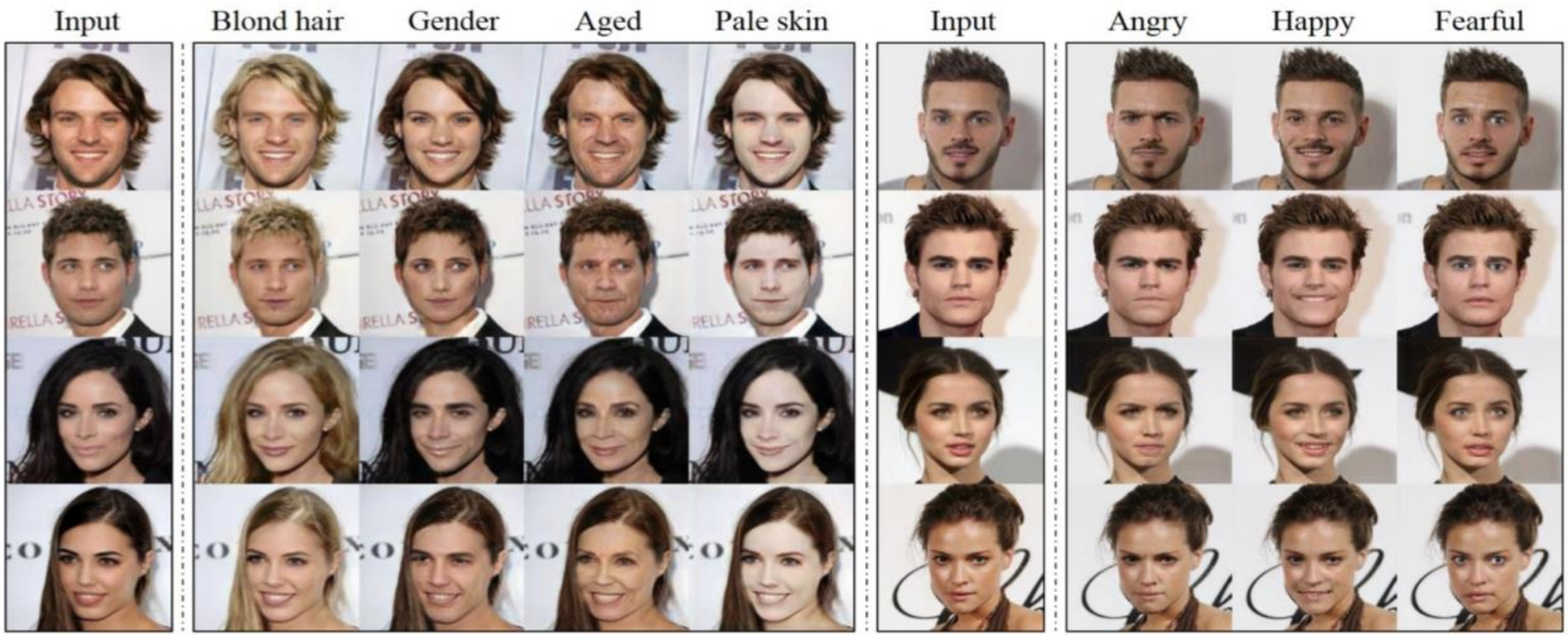
Idea: CycleGAN (unpaired training) + Domain condition





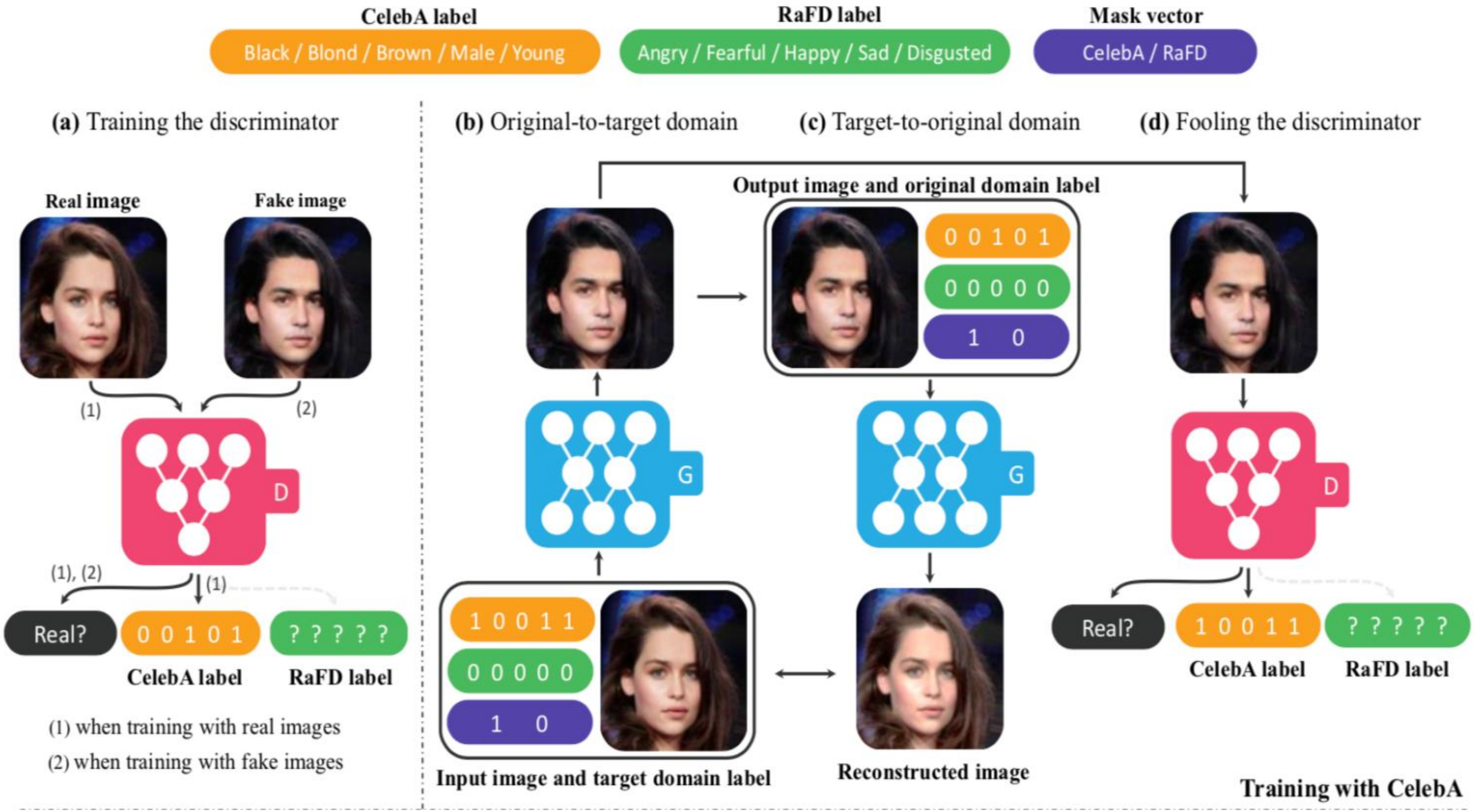
[Choi et al. CVPR 2018] StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation

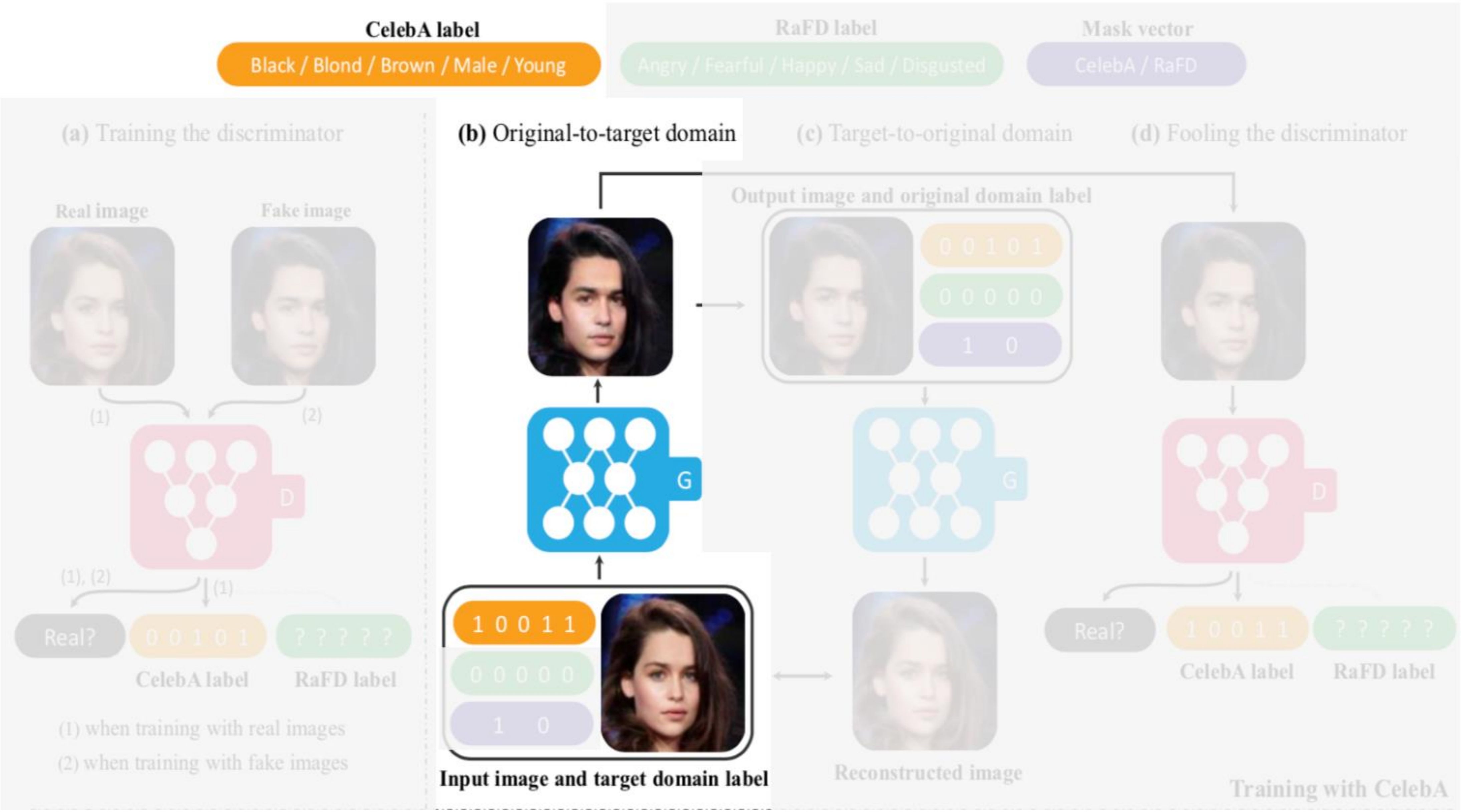
Datasets

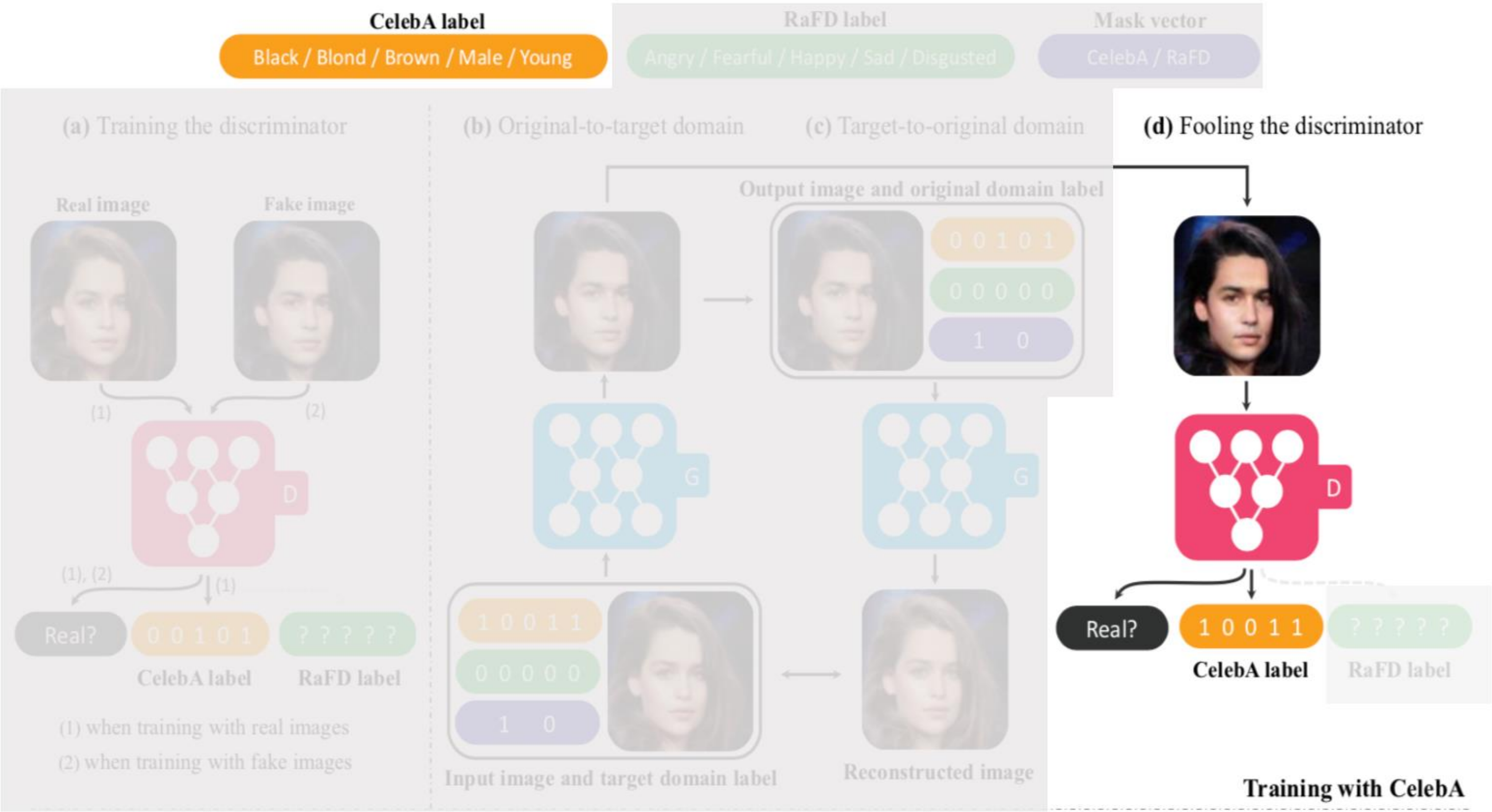


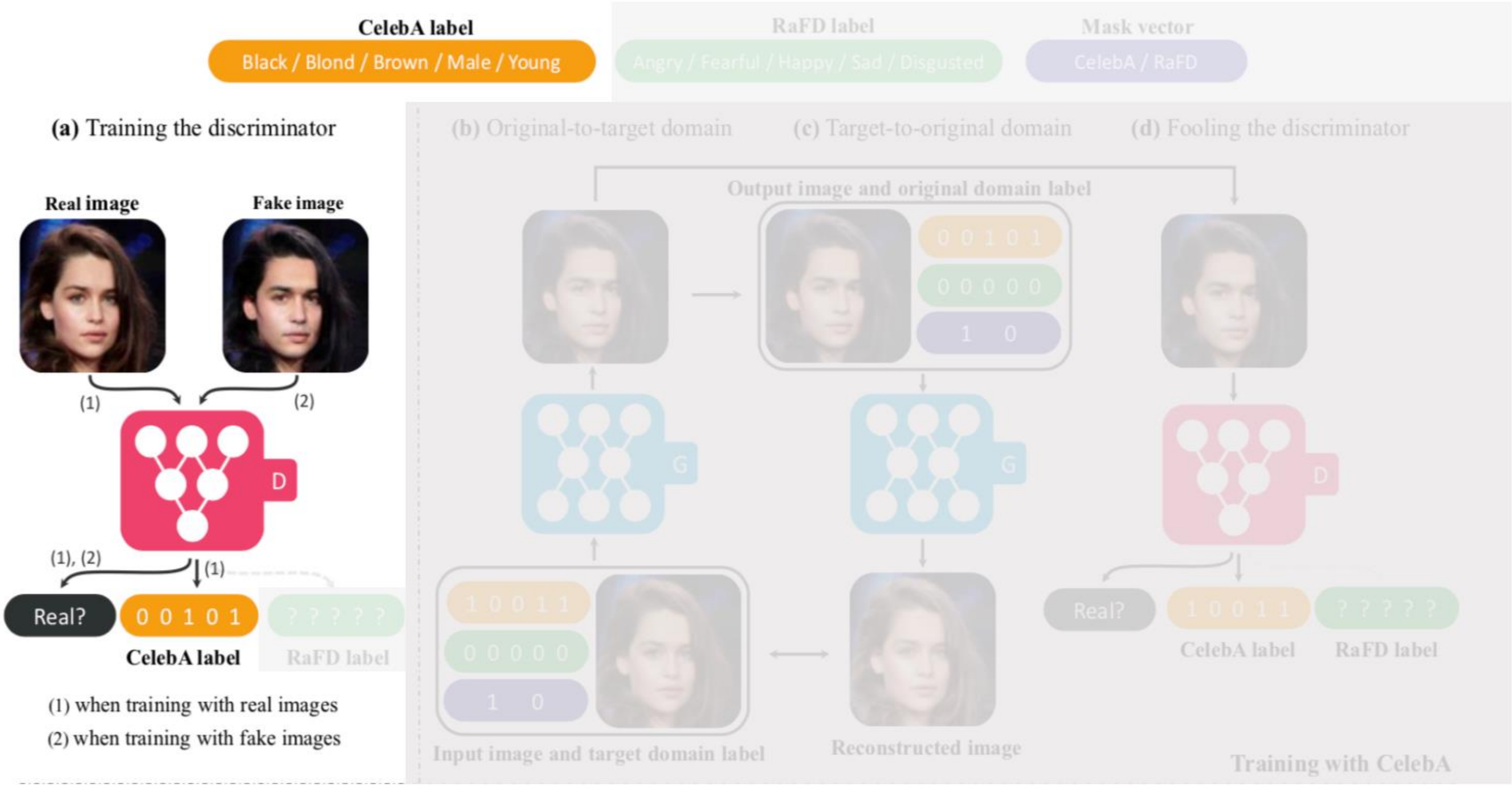
CelebA

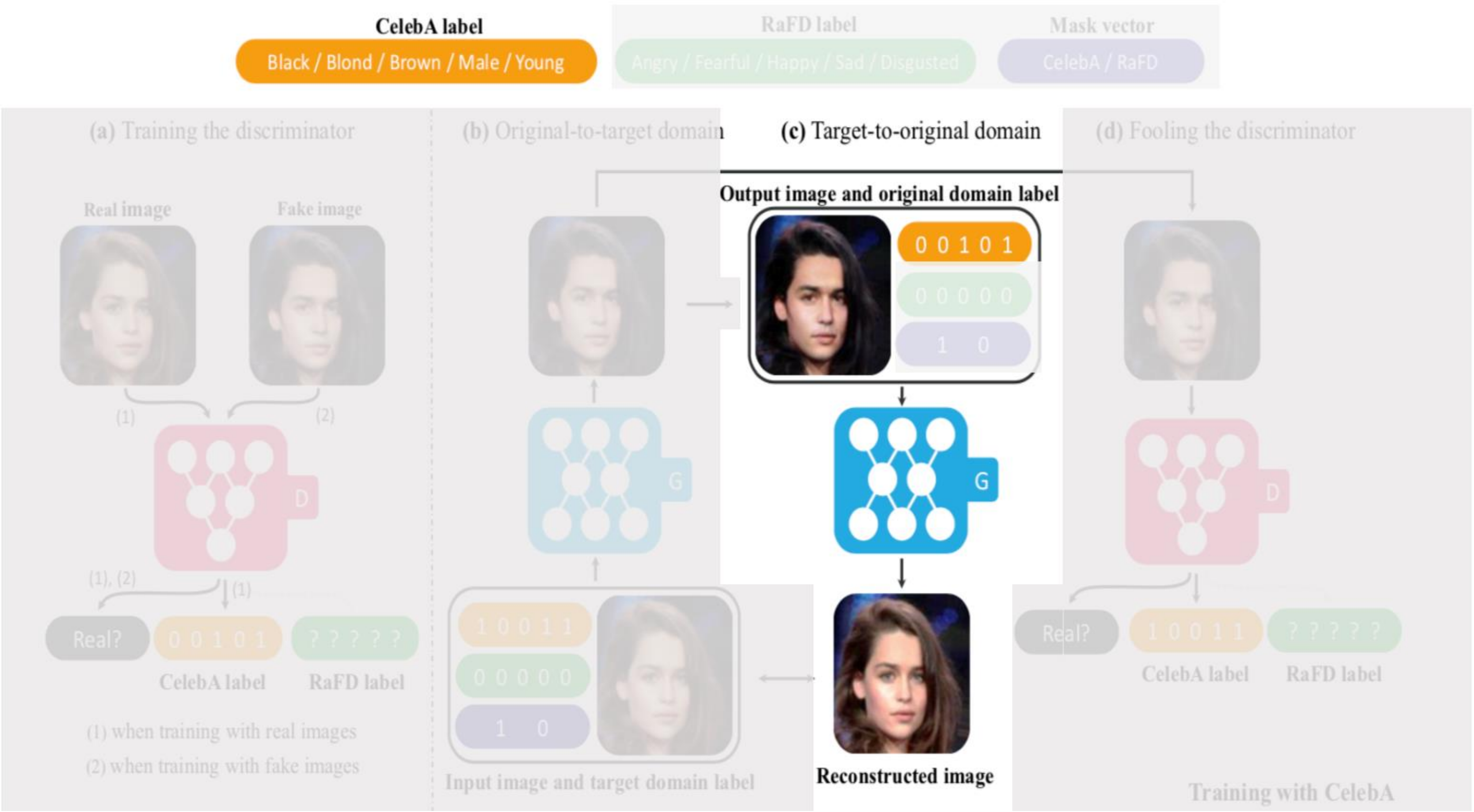
RaFD

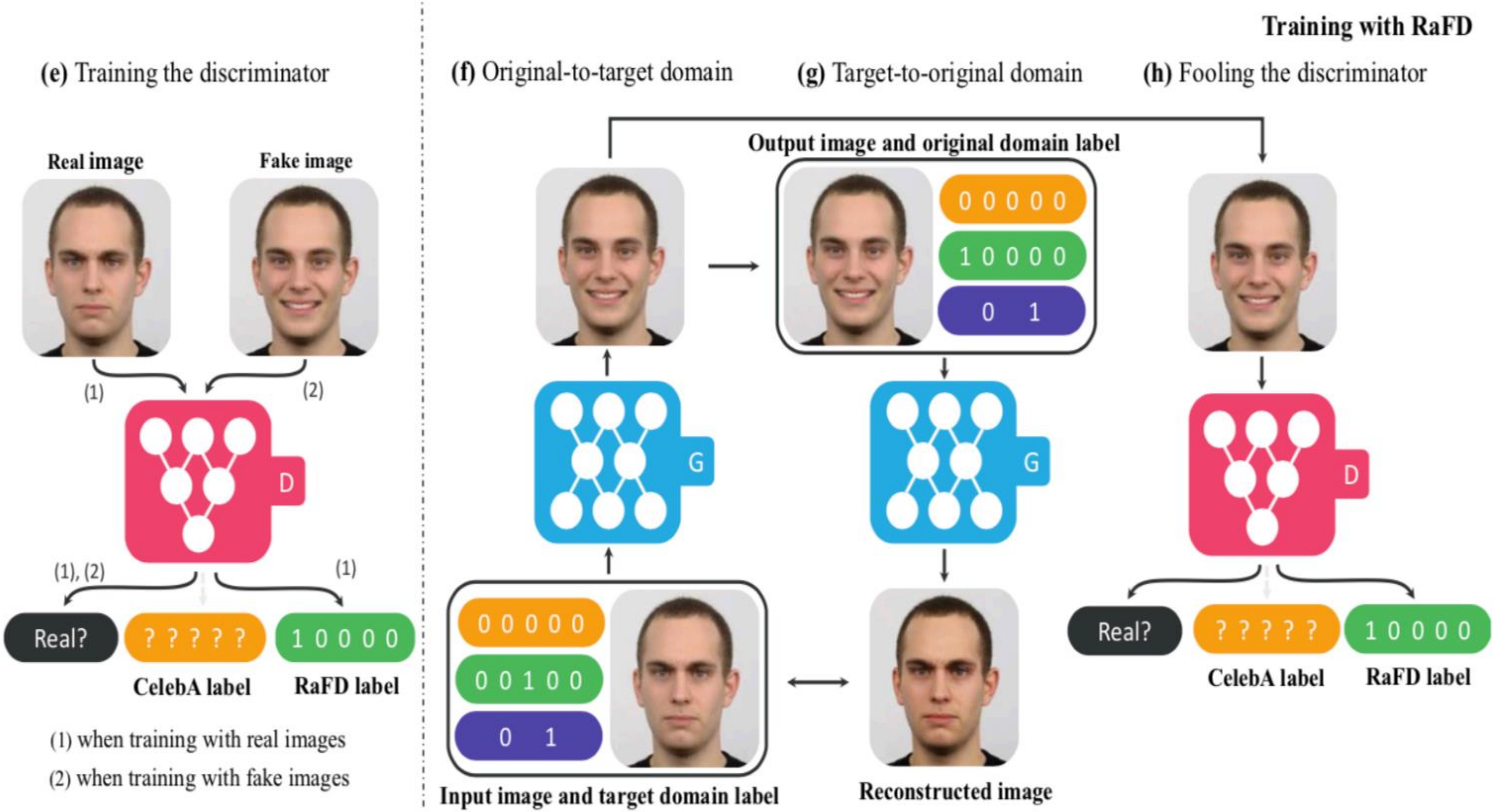

CelebA label
Black / Blond / Brown / Male / Young
RaFD label
Angry / Fearful / Happy / Sad / Disgusted
Mask vector
CelebA / RaFD
(a) Training the discriminator
(b) Original-to-target domain
(c) Target-to-original domain
(d) Fooling the discriminator
Real image
Fake image
(1)
(2)
D
(1), (2)
(1)
Real?
0 0 1 0 1
? ? ? ? ?
CelebA label
RaFD label
(1) when training with real images
(2) when training with fake images
Output image and original domain label
0 0 1 0 1
0 0 0 0 0
1 0
G
G
1 0 0 1 1
0 0 0 0 0
1 0
Input image and target domain label
Reconstructed image
D
Real?
1 0 0 1 1
? ? ? ? ?
CelebA label
RaFD label
Training with CelebA


CelebA label
Black / Blond / Brown / Male / Young
RaFD label
Angry / Fearful / Happy / Sad / Disgusted
Mask vector
CelebA / RaFD
(a) Training the discriminator
Real image
Fake image
(1)
(2)
D
(1), (2)
(1)
Real?
0 0 1 0 1
? ? ? ? ?
CelebA label
RaFD label
(1) when training with real images
(2) when training with fake images
(b) Original-to-target domain
G
1 0 0 1 1
0 0 0 0 0
1 0
Input image and target domain label
(c) Target-to-original domain
Output image and original domain label
0 0 1 0 1
0 0 0 0 0
1 0
G
Reconstructed image
(d) Fooling the discriminator
D
Real?
1 0 0 1 1
? ? ? ? ?
CelebA label
RaFD label
Training with CelebA


CelebA label
Black / Blond / Brown / Male / Young
RaFD label
Angry / Fearful / Happy / Sad / Disgusted
Mask vector
CelebA / RaFD
(a) Training the discriminator
(b) Original-to-target domain
(c) Target-to-original domain
(d) Fooling the discriminator
Real image
Fake image
(1)
(2)
D
(1), (2)
(1)
Real?
0 0 1 0 1
? ? ? ? ?
CelebA label
RaFD label
(1) when training with real images
(2) when training with fake images
Output image and original domain label
0 0 1 0 1
0 0 0 0 0
1 0
G
G
1 0 0 1 1
0 0 0 0 0
1 0
Input image and target domain label
Reconstructed image
D
Real?
1 0 0 1 1
? ? ? ? ?
CelebA label
RaFD label
Training with CelebA

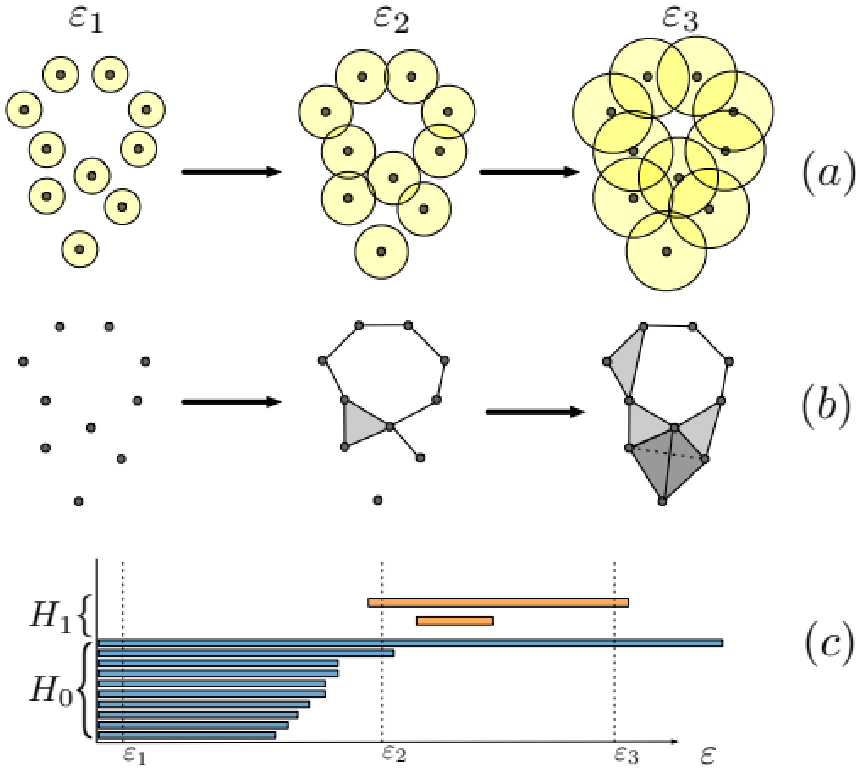

CelebA label
Black / Blond / Brown / Male / Young
RaFD label
Angry / Fearful / Happy / Sad / Disgusted
Mask vector
CelebA / RaFD
(a) Training the discriminator
(b) Original-to-target domain
(c) Target-to-original domain
(d) Fooling the discriminator
Output image and original domain label
Real image
Fake image
(1)
(2)
0 0 1 0 1
0 0 0 0 0
1 0
G
D
(1), (2)
(1)
Real?
0 0 1 0 1
? ? ? ? ?
CelebA label
RaFD label
1 0 0 1 1
0 0 0 0 0
1 0
Real?
1 0 0 1 1
? ? ? ? ?
CelebA label
RaFD label
(1) when training with real images
(2) when training with fake images
Input image and target domain label
Reconstructed image
Training with CelebA

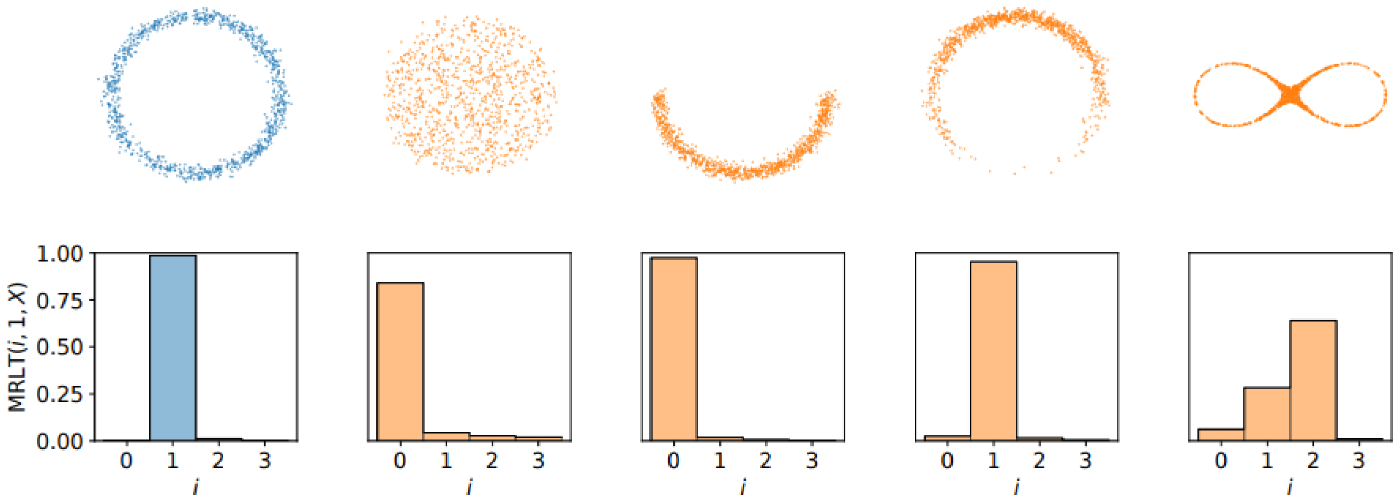

CelebA label
Black / Blond / Brown / Male / Young
RaFD label
Angry / Fearful / Happy / Sad / Disgusted
Mask vector
CelebA / RaFD
(a) Training the discriminator
Real image
Fake image
(1)
(2)
D
(1), (2)
(1)
Real?
0 0 1 0 1
? ? ? ? ?
CelebA label
RaFD label
(1) when training with real images
(2) when training with fake images
(b) Original-to-target domain
G
1 0 0 1 1
0 0 0 0 0
1 0
Input image and target domain label
(c) Target-to-original domain
Output image and original domain label
0 0 1 0 1
0 0 0 0 0
1 0
G
Reconstructed image
(d) Fooling the discriminator
D
Real?
1 0 0 1 1
? ? ? ? ?
CelebA label
RaFD label
Training with CelebA


Training with RaFD
(e) Training the discriminator
(f) Original-to-target domain
(g) Target-to-original domain
(h) Fooling the discriminator
Real image
Fake image
(1)
(2)
D
(1), (2)
(1)
Real?
? ? ? ? ?
1 0 0 0 0
CelebA label
RaFD label
(1) when training with real images
(2) when training with fake images
Output image and original domain label
0 0 0 0 0
1 0 0 0 0
0 1
G
G
D
0 0 0 0 0
0 0 1 0 0
0 1
Input image and target domain label
Reconstructed image
Real?
? ? ? ? ?
1 0 0 0 0
CelebA label
RaFD label

[2/4] Evaluation Metric for GANs

Motivation: There is still no convicting metric to evaluate GANs





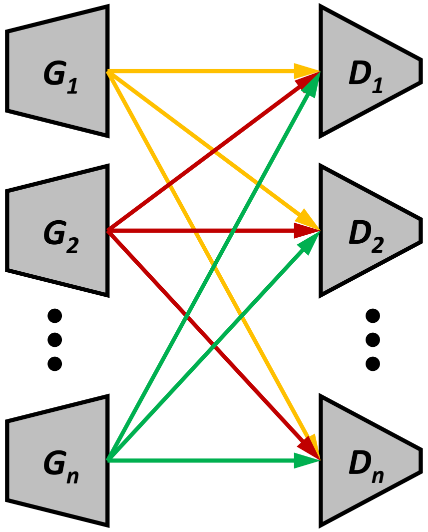
Con:

- **High computational complexity**
- **Landmark selection is tricky**
- **Lack of experiment on natural images**

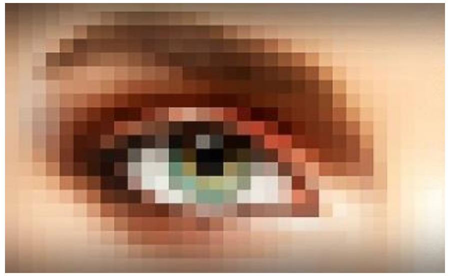
[Khrulkov and Oseledets, ICML 2018] Geometry Score: A Method For Comparing Generative Adversarial Networks

[2/4] Evaluation Metric for GANs

Motivation: There is still no convicting metric to evaluate GANs



Which is better?



Pro:

- **Relative comparison**

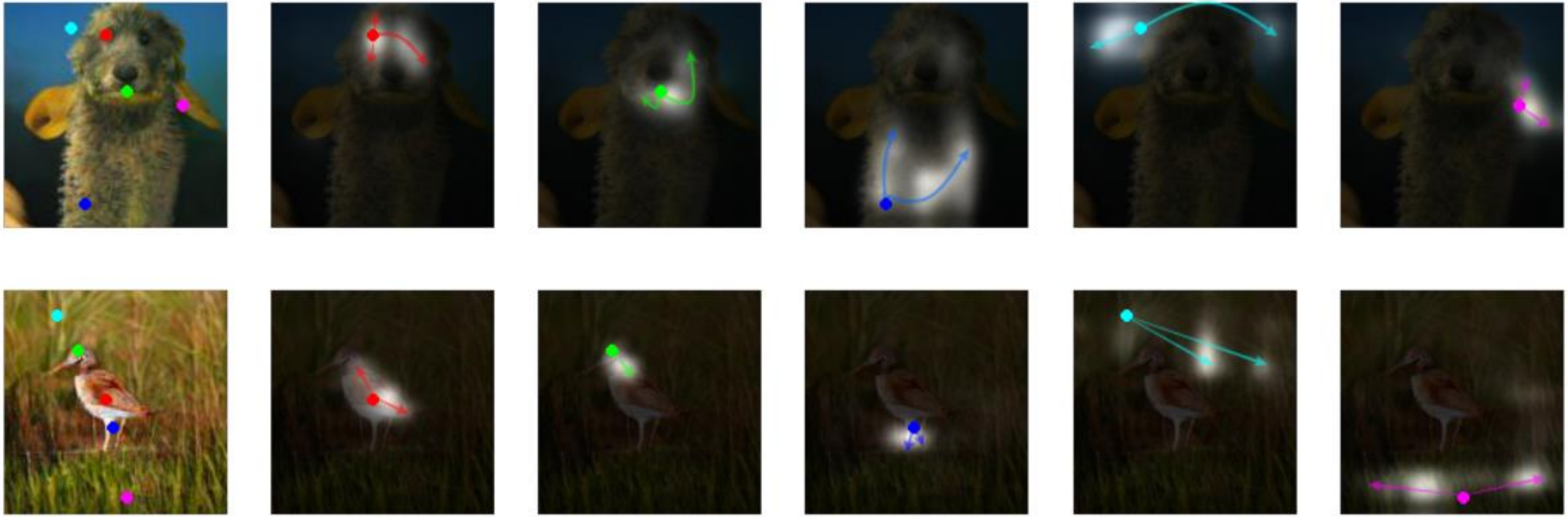
Con:

- **Lack of theoretical support**

[Olsson et al. 2018] Skill Rating for Generative Models

[3/4] Self-Attention GAN

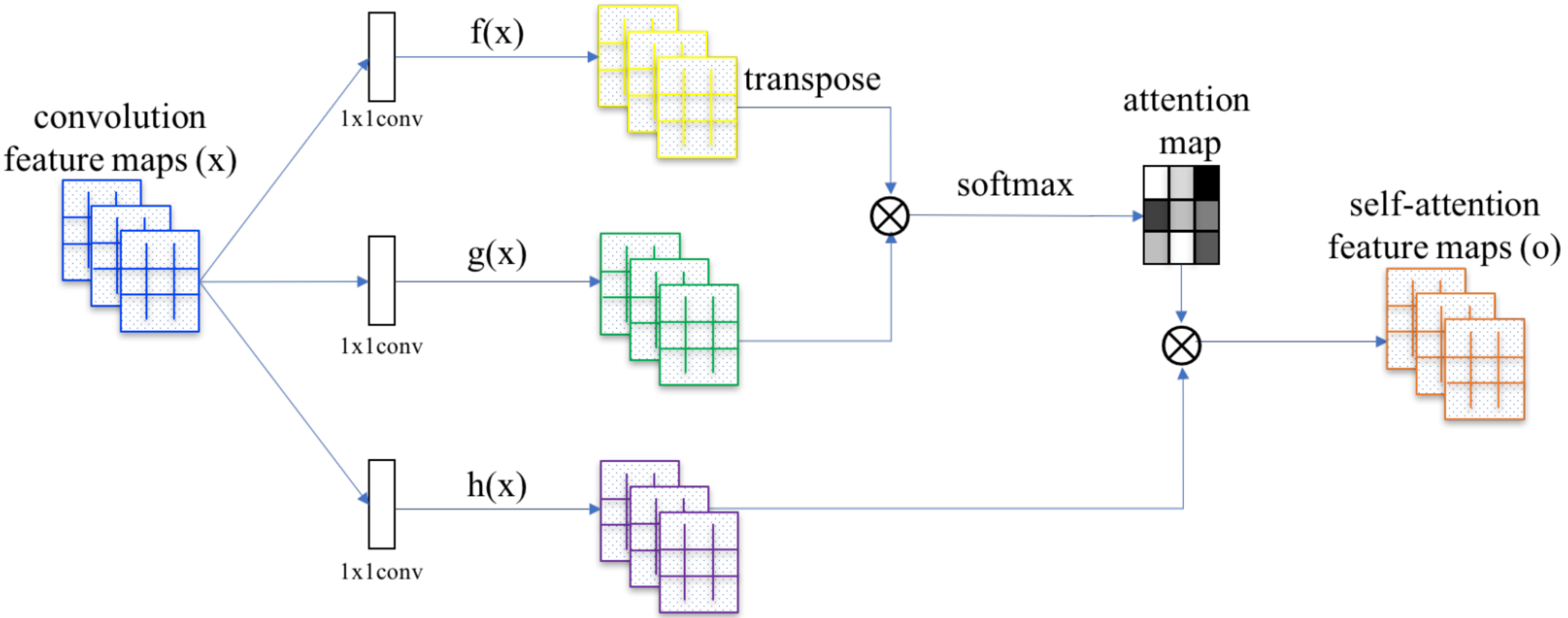
Motivation: Specially local dependency → long-range dependency (attention-driven)



[Zhang et al. 2018] Self-Attention Generative Adversarial Networks

[3/4] Self-Attention GAN

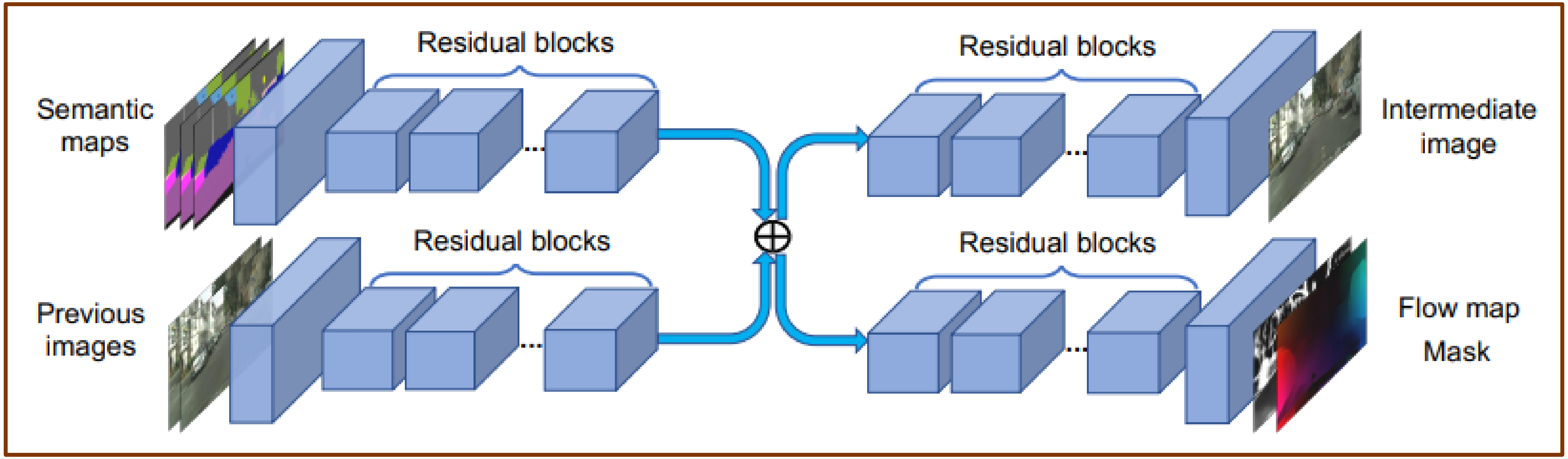
Motivation: Specially local dependency → long-range dependency (attention-driven)



[Zhang et al. 2018] Self-Attention Generative Adversarial Networks

[4/4] Video-to-Video Synthesis

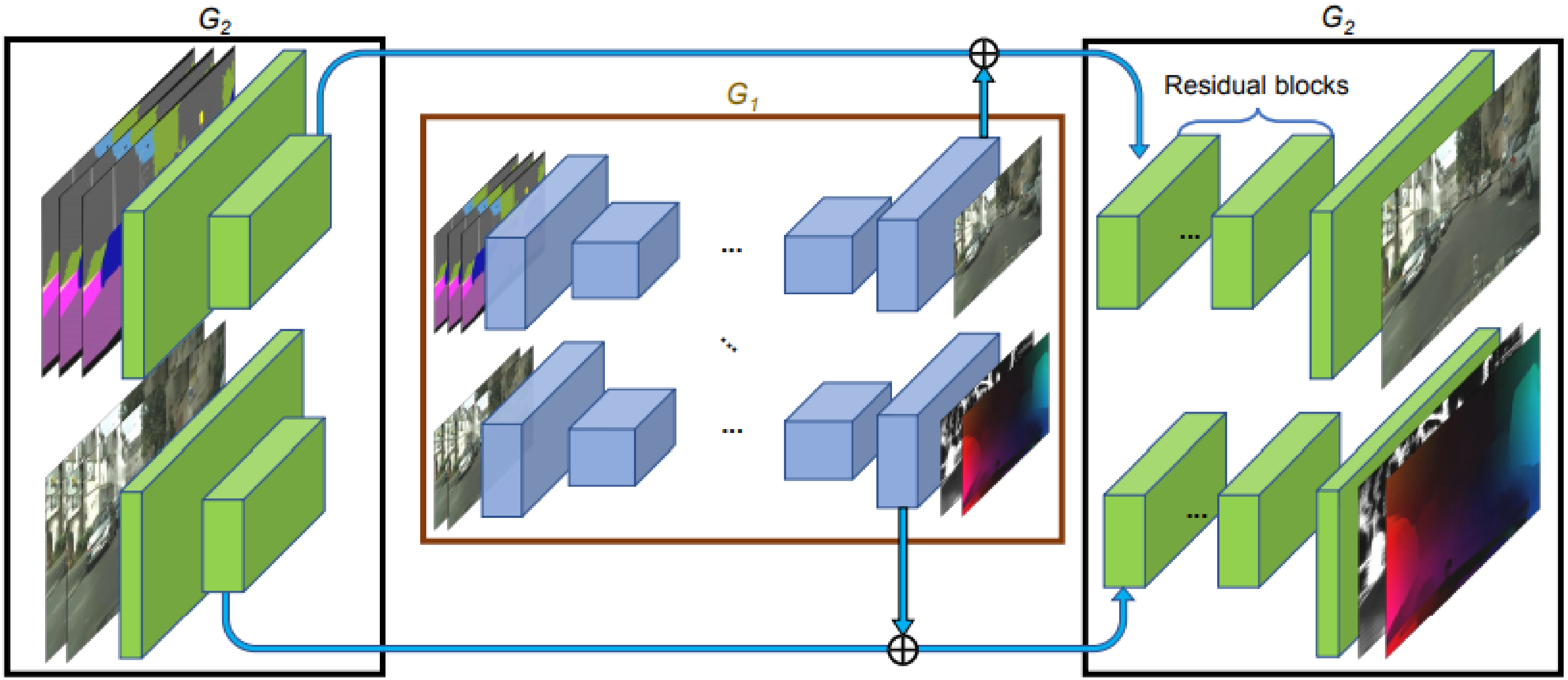
Motivation: Temporal dynamics is less explored in image-to-image translation



[Wang et al. NIPS2018] Video-to-Video Synthesis

[4/4] Video-to-Video Synthesis

Motivation: Temporal dynamics is less explored in image-to-image translation



[Wang et al. NIPS2018] Video-to-Video Synthesis

[4/4] Video-to-Video Synthesis

Motivation: Temporal dynamics is less explored in image-to-image translation

https://tcwang0509.github.io/vid2vid/

Summary

1. GANs have become a common penalty
2. Tend to be multi-task/multiple GANs
3. Less fundamental improvement of GANs
4. Image evaluation is still a challenge topic
5. Video and 3D scene synthesis are imperative

Thank you