Deep Generative Prior

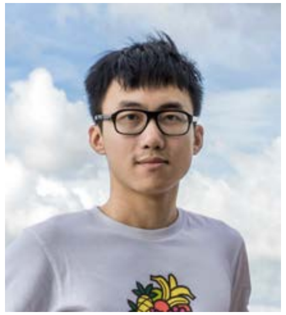
Chen Change Loy
S-Lab, Nanyang Technological University

Papers

Exploiting Deep Generative Prior for Versatile Image Restoration and Manipulation

ECCV 2020 (Oral)

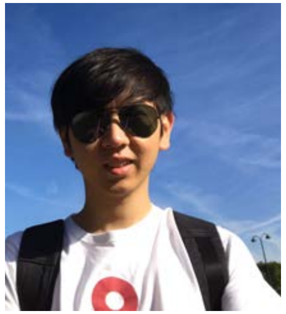
Xingang Pan, Xiaohang Zhan, Bo Dai, Dahua Lin, Chen Change Loy, Ping Luo



Do 2D GANs Know 3D Shape? Unsupervised 3D Shape Reconstruction from 2D Image GANs

ICLR, 2021 (Oral)

Xingang Pan, Bo Dai, Ziwei Liu, Chen Change Loy, Ping Luo



GLEAN: Generative Latent Bank for Large-Factor Image Super-Resolution

Arxiv preprint, 2020

Kelvin C.K Chan, Xintao Wang, Xiangyu Xu, Jinwei Gu, Chen Change Loy

香港中文大學
The Chinese University of Hong Kong

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Image Priors

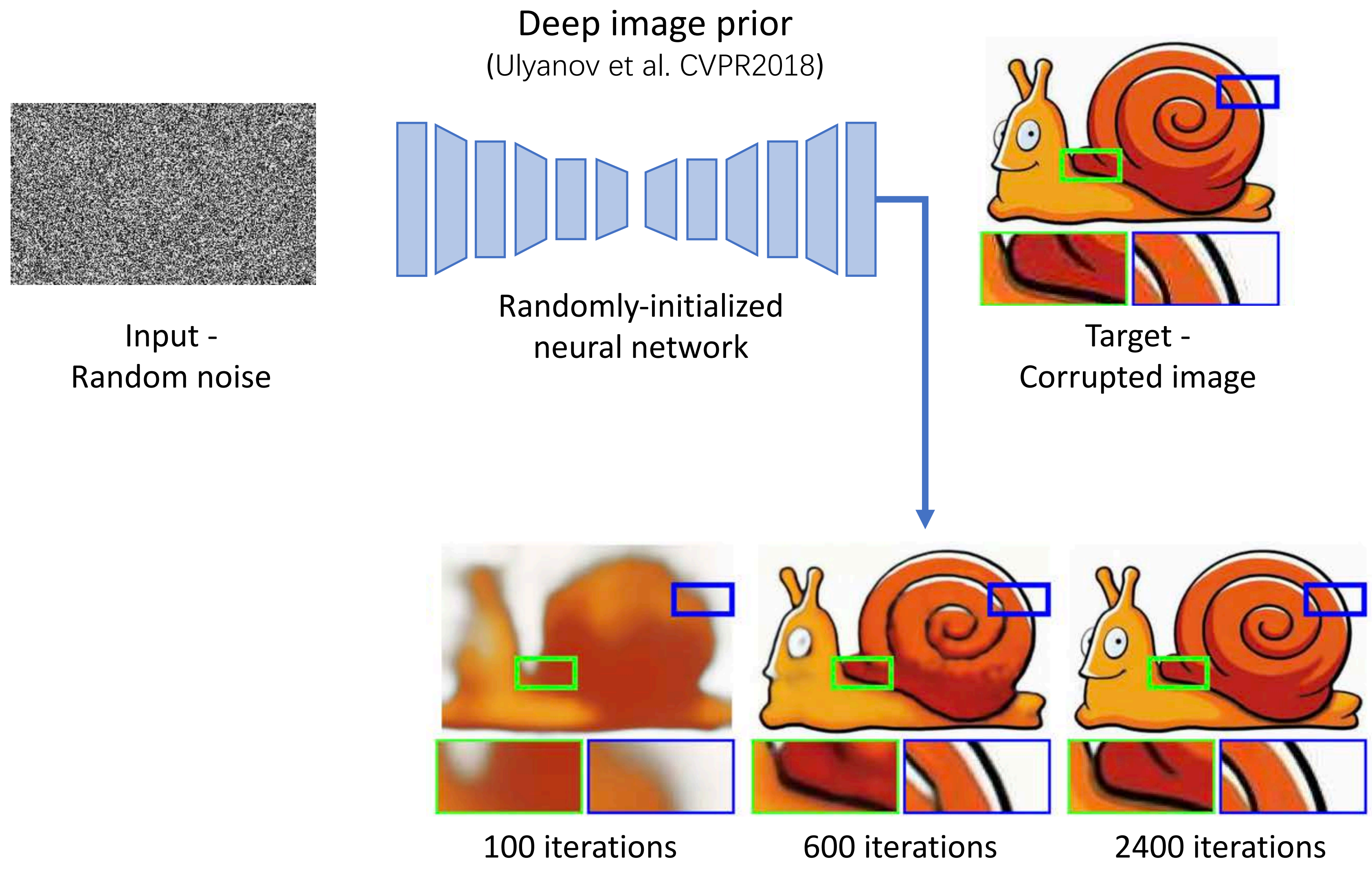
$$\min_x \underbrace{E(x; x_0)}_{\text{data term}} + \underbrace{R(x)}_{\text{image prior}}$$

Total variation – denoise
(Rudin et al. Physica D)

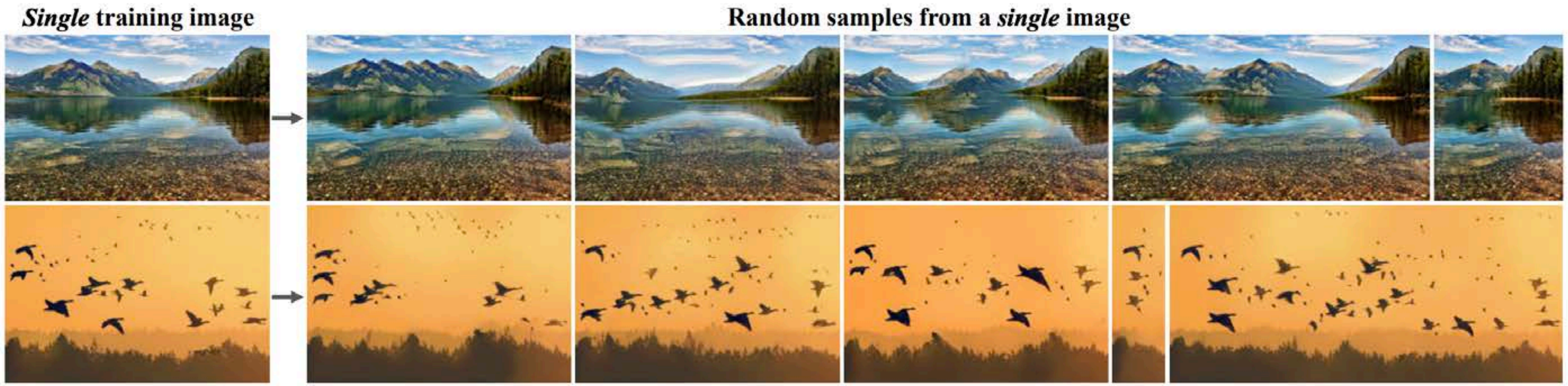
Dark channel prior – dehaze
(He et al. CVPR2009)

MRF – pixel correlation
(Roth et al. CVPR2005)

Deep Image Prior



Deep Image Prior

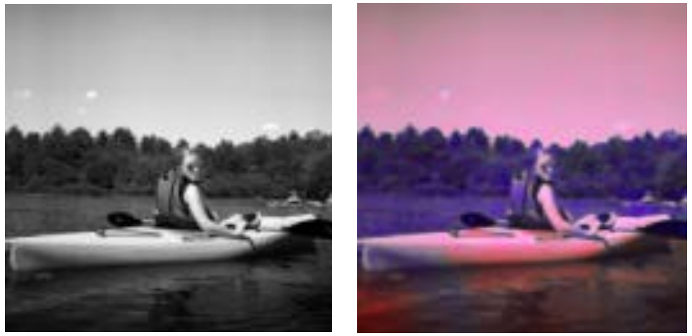


SinGAN
(T. Rott Shaham et al. ICCV2019)

Deep Image Prior

CNN and GAN are trained from a single image of interest from scratch

Limited access to image statistics beyond the input image



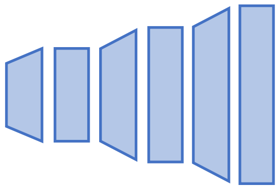
Deep image prior
(Ulyanov et al. CVPR2018)

Target　Random jittering

SinGAN
(T. Rott Shaham et al. ICCV2019)

Deep Generative Prior

$\mathbf{z} \sim \mathcal{N}(0, I)$



Latent space

Generator

Can we leverage a GAN trained on large-scale natural images for richer priors beyond a single image?

GAN is a good approximator for natural image manifold.

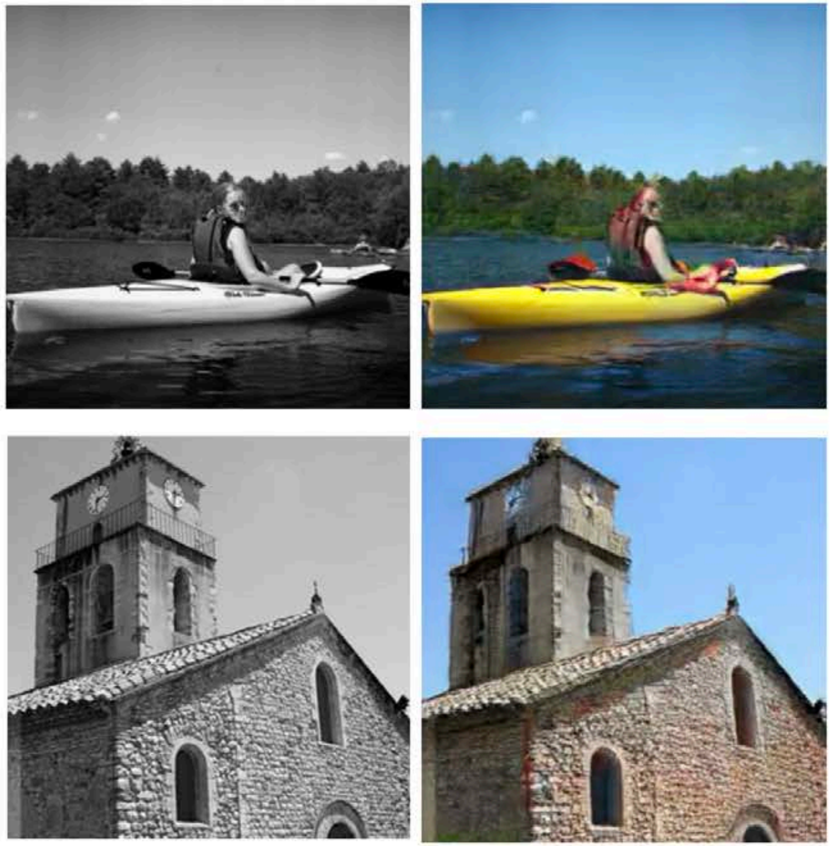
Challenges

- Cope with arbitrary images from different tasks with distinctly different natures
- Produce sharp and faithful images obeying the natural image manifold

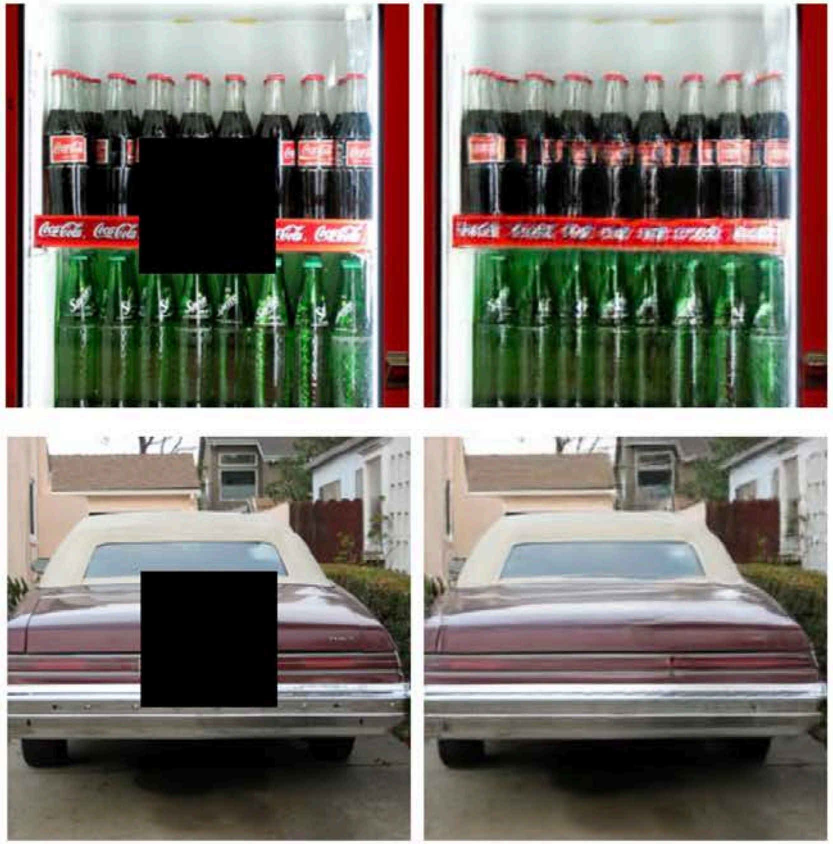
Goodfellow et al. NIPS2014

Applications

Colorization



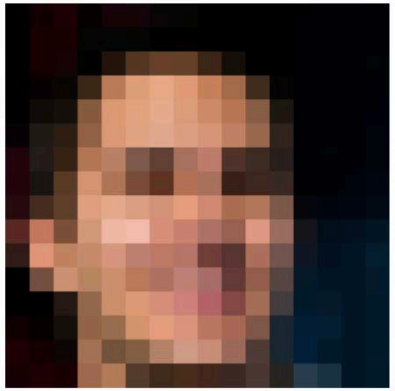
Inpainting



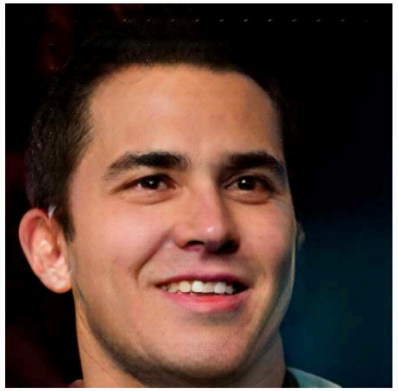
Super-resolution

Exploiting Deep Generative Prior for Versatile Image Restoration and Manipulation
ECCV 2020 (Oral)

Applications



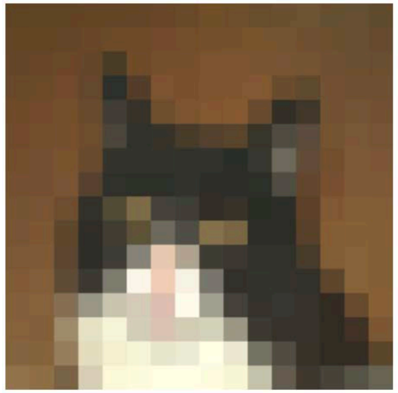
Low-Resolution



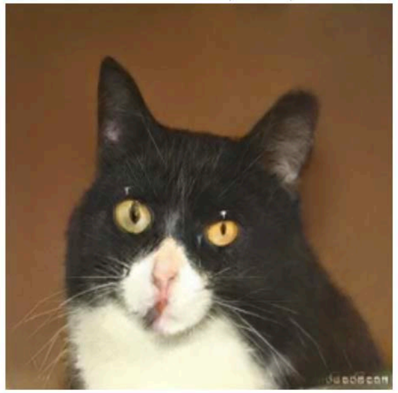
Super-Resolution

Low-Resolution

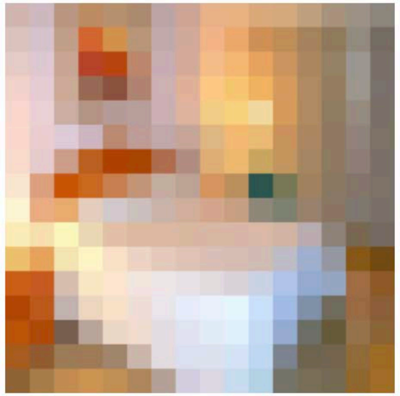
Super-Resolution



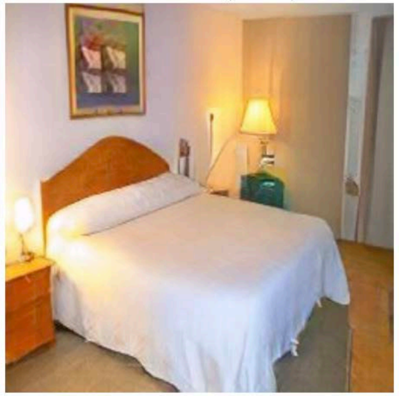
Low-Resolution



Super-Resolution



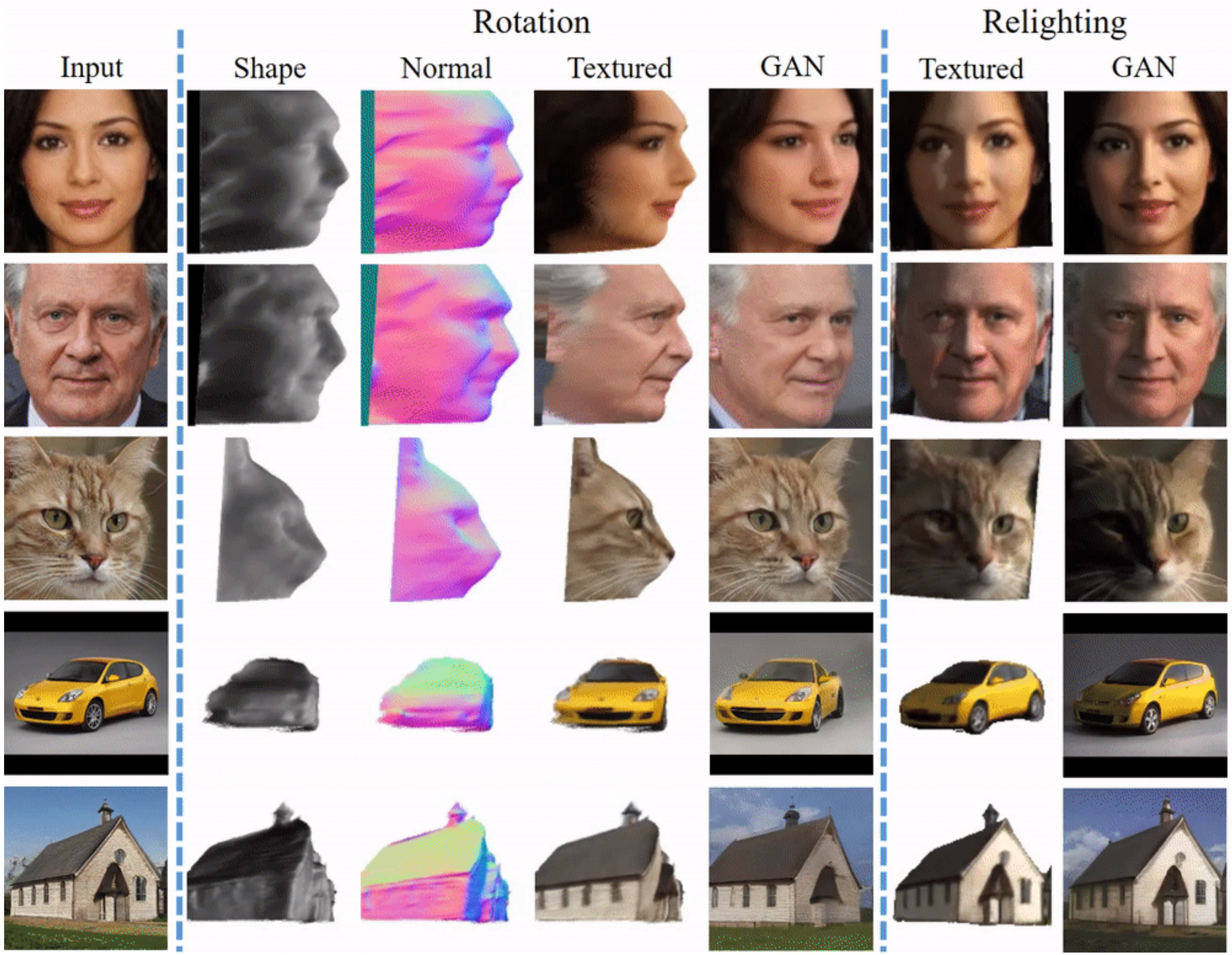
Low-Resolution



Super-Resolution

GLEAN: Generative Latent Bank for Large-Factor Image Super-Resolution
Arxiv preprint, 2020

Applications



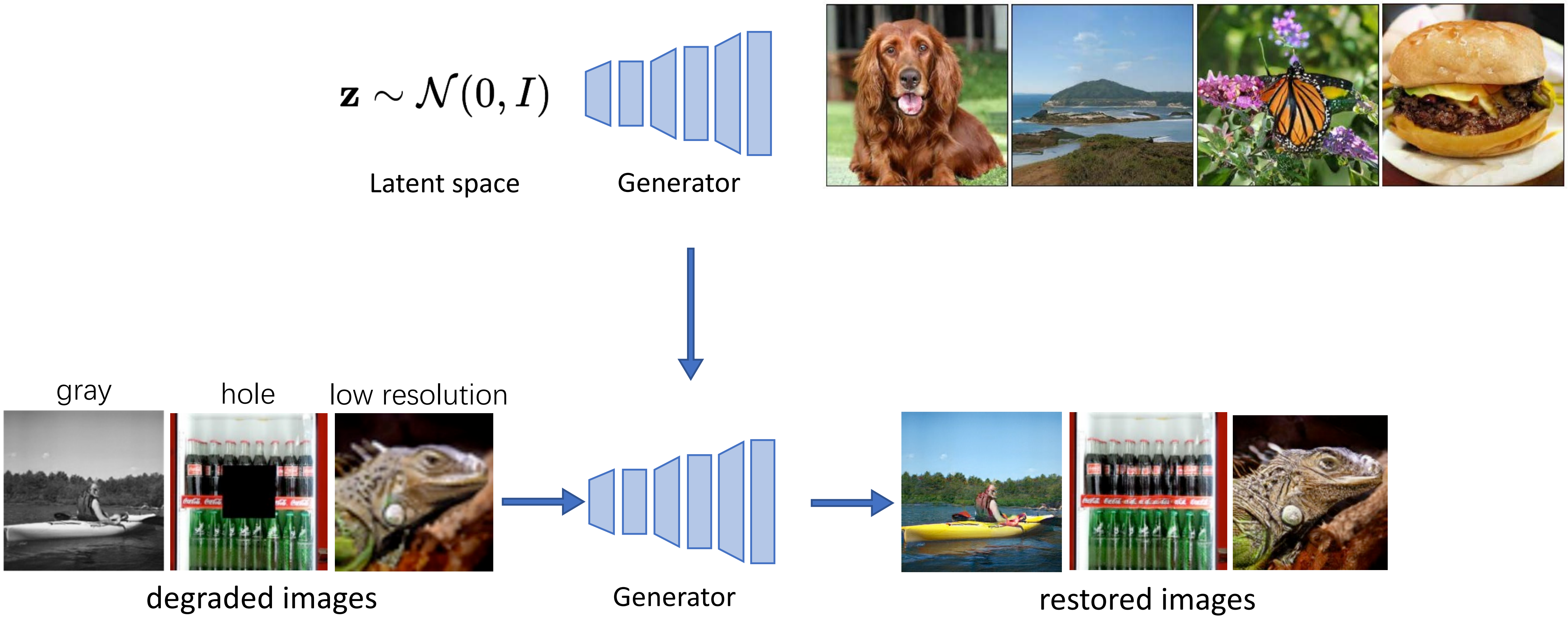
Do 2D GANs Know 3D Shape? Unsupervised 3D Shape Reconstruction from 2D Image GANs
Arxiv preprint, 2020

Outline

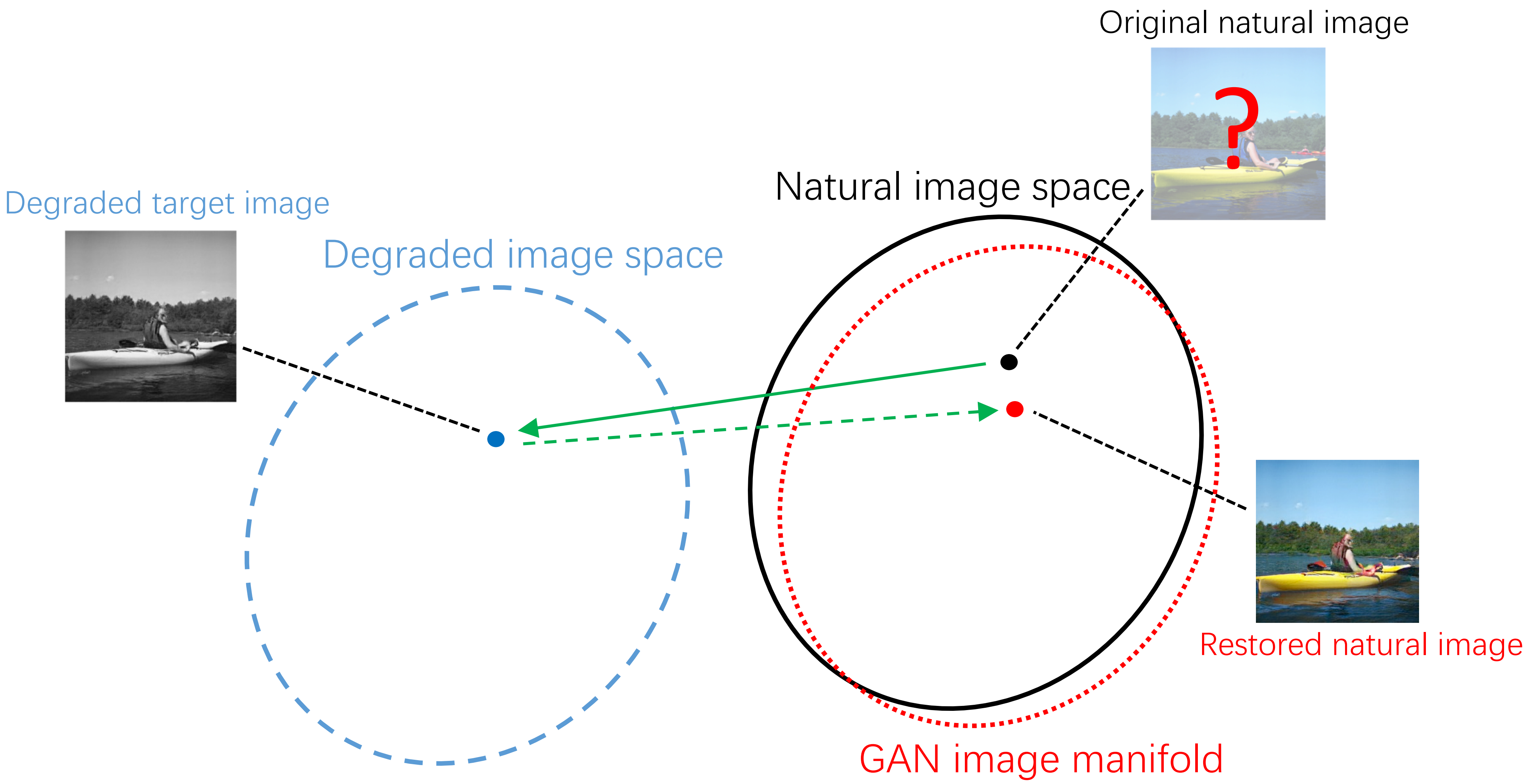
1. How to exploit generic image prior in pretrained GANs?
2. How to exploit pretrained GAN as a latent bank for large-factor image super-resolution?
3. Do 2D GANs model 3D geometry? Shape-from-X where X=GAN?

Motivation

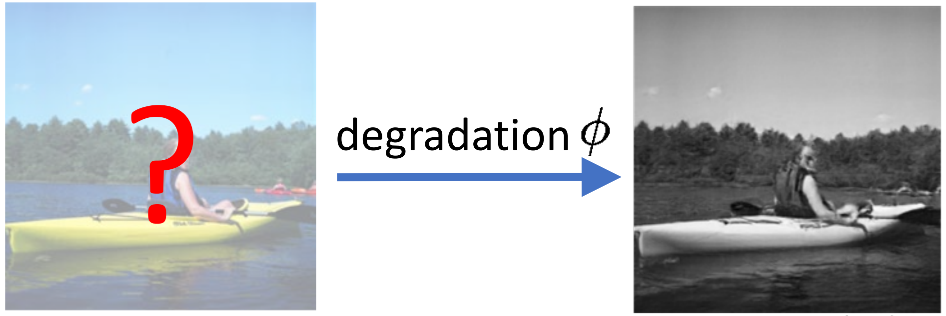
Our goal: exploit generic image prior of pretrained GAN



Deep Generative Prior

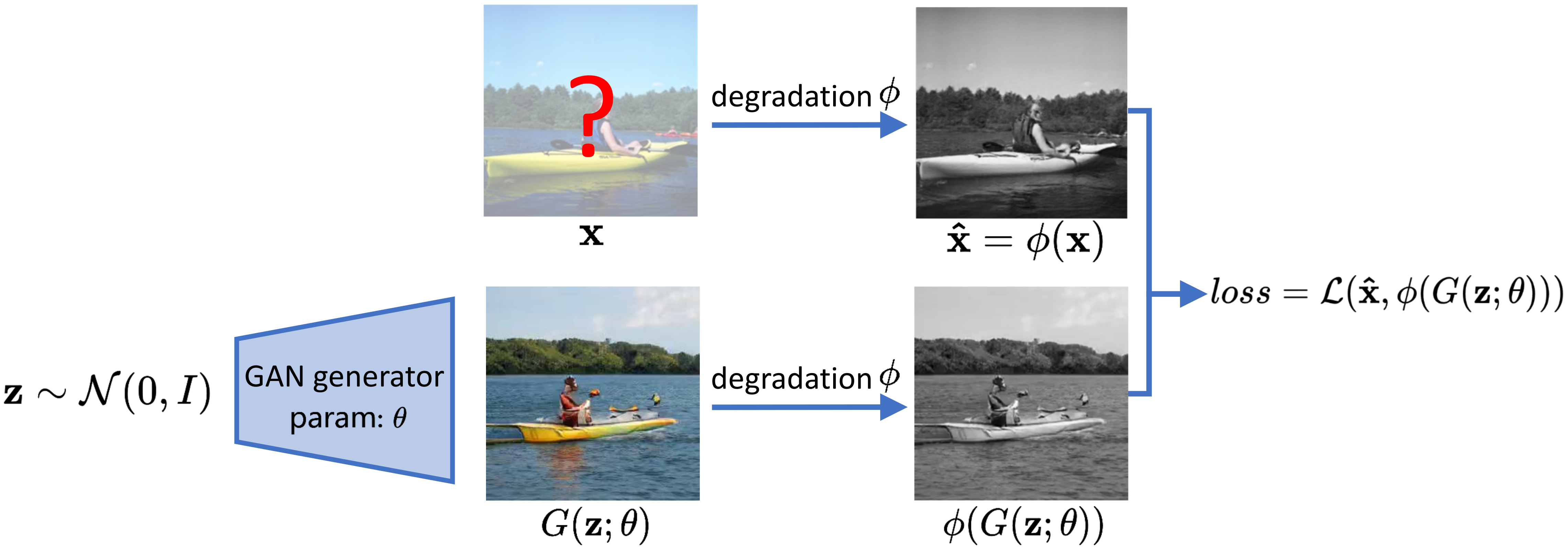


Deep Generative Prior



$\mathbf{x}$ $\quad\quad$ $\hat{\mathbf{x}} = \phi(\mathbf{x})$

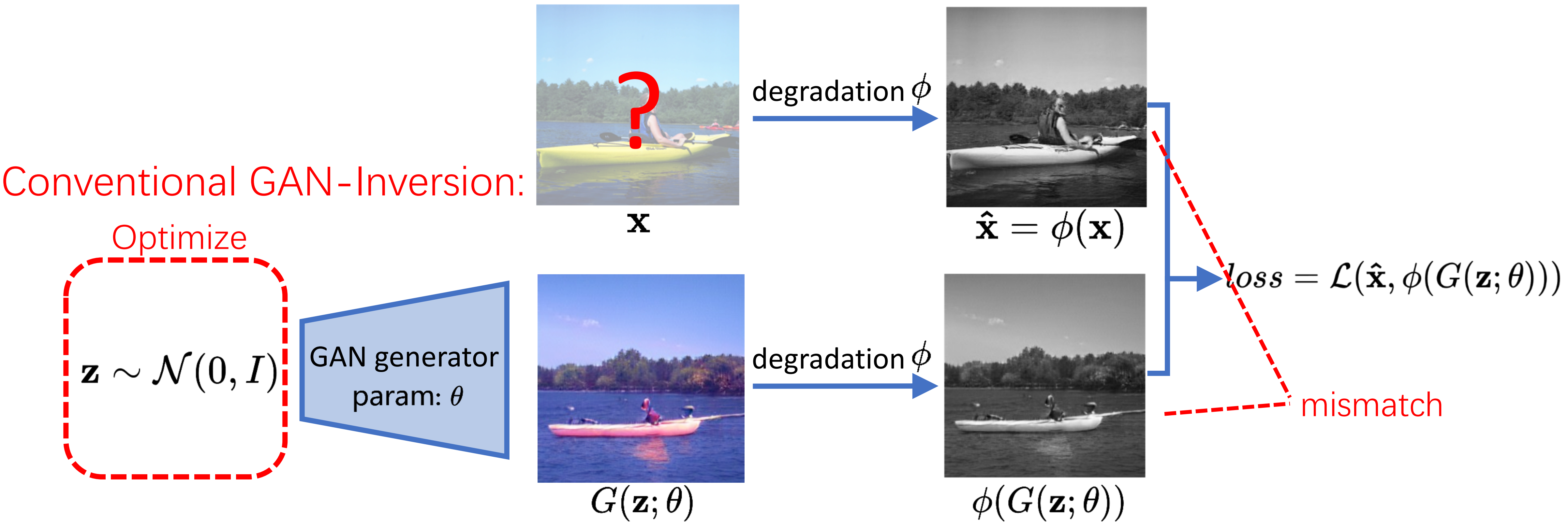
Deep Generative Prior



color restored

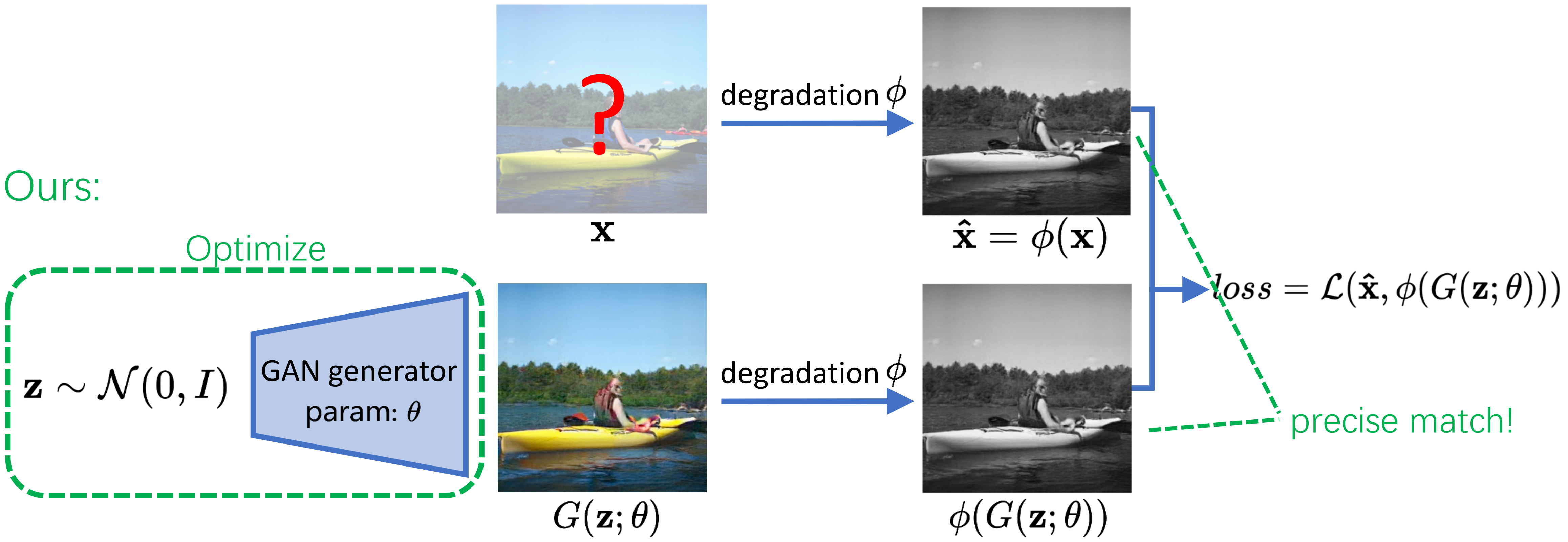
reconstruct in the gray space

Challenges of GAN-Inversion



$$\mathbf{z}^* = argmin_{\mathbf{z} \in R^d} \mathcal{L}(\hat{\mathbf{x}}, \phi(G(\mathbf{z}; \theta)))$$

Solution 1: Relax the Generator for GAN-Inversion



$$\theta^*, \mathbf{z}^* = argmin_{\theta,\mathbf{z}} \mathcal{L}(\hat{\mathbf{x}}, \phi(G(\mathbf{z};\theta)))$$ (Relaxed GAN-inversion)

Conventional Loss

Target $\hat{\mathbf{x}}$ —— $G(\mathbf{z};\theta)$ during finetuning →

MSE

Perceptual Loss

Solution 2: Discriminator Feature-Matching Loss

L1 distance in the discriminator feature space:

$$\mathcal{L}(\mathbf{x}_1, \mathbf{x}_2) = \sum_{i \in \mathcal{I}} \|D(\mathbf{x}_1, i), D(\mathbf{x}_2, i)\|_1$$

$D\,(\mathbf{x}, \mathrm{i})$ returns the feature of $\mathbf{x}$ at the i'th block of the discriminator.

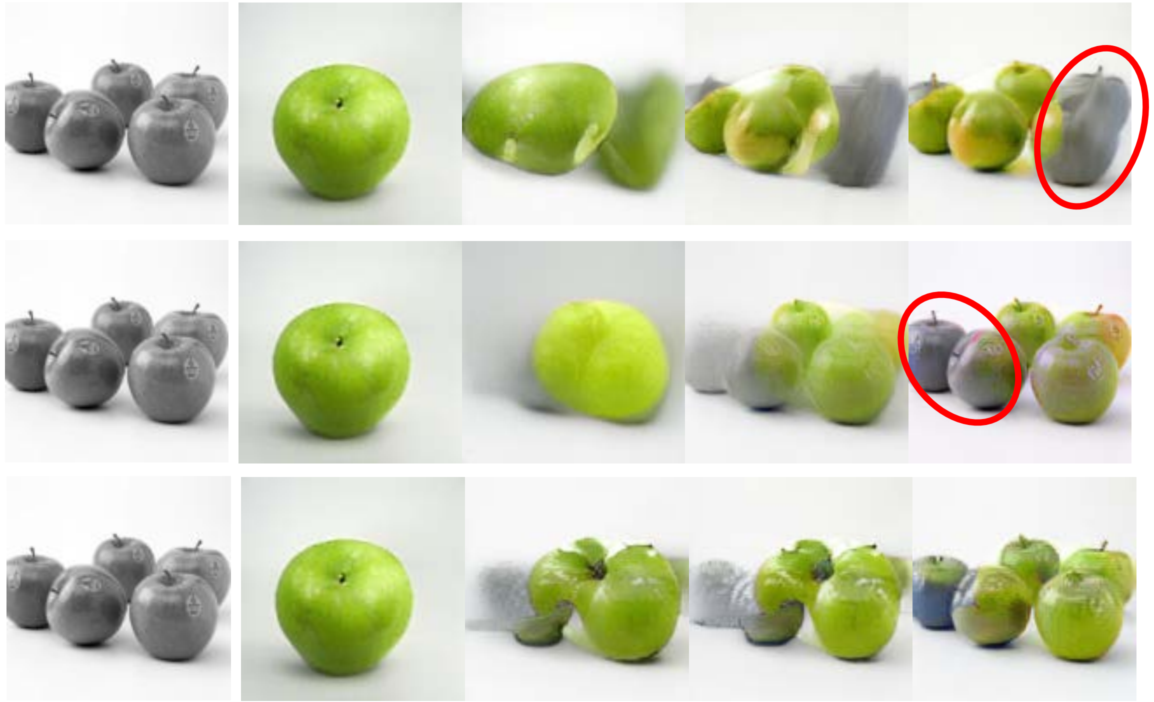
Conventional Loss

Target $\hat{\mathbf{x}}$ —— $G(\mathbf{z};\theta)$ during finetuning ⟶

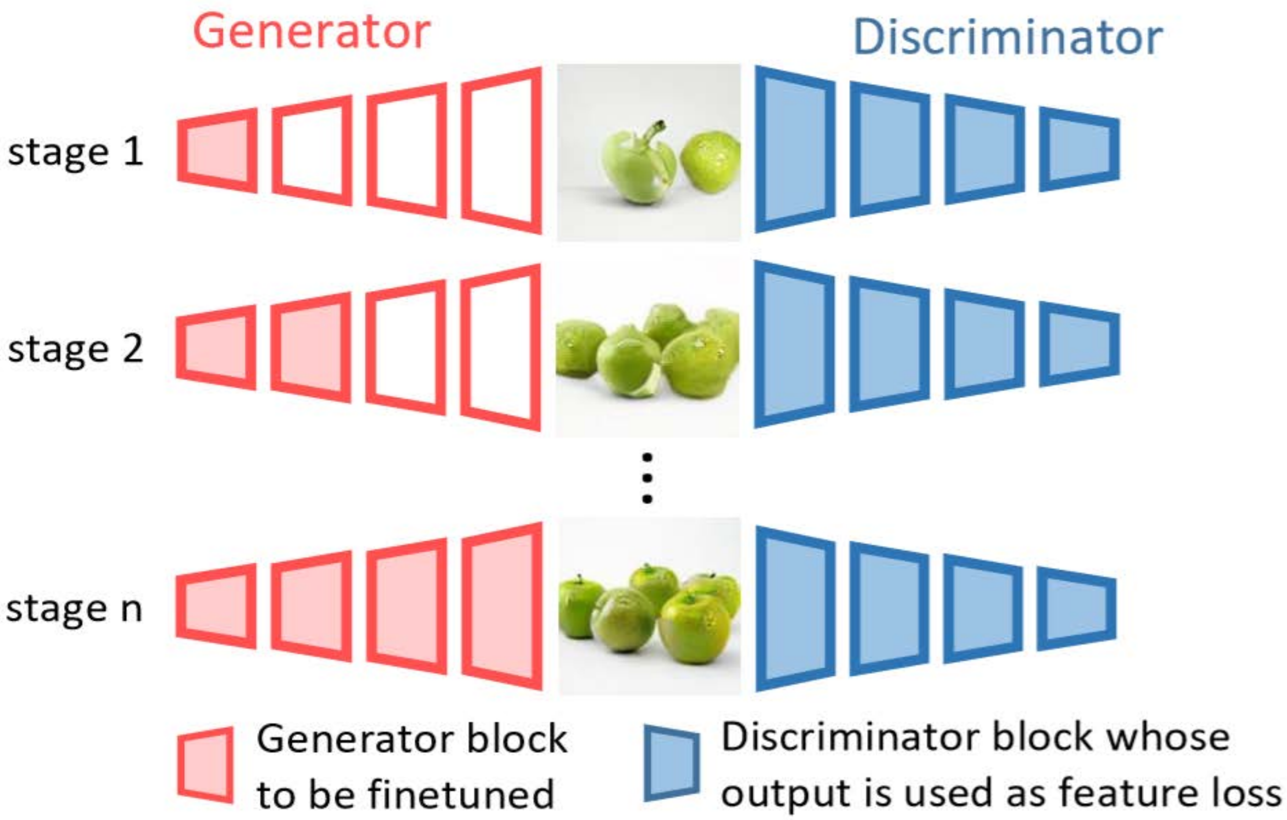
MSE

Perceptual Loss

Discriminator Loss

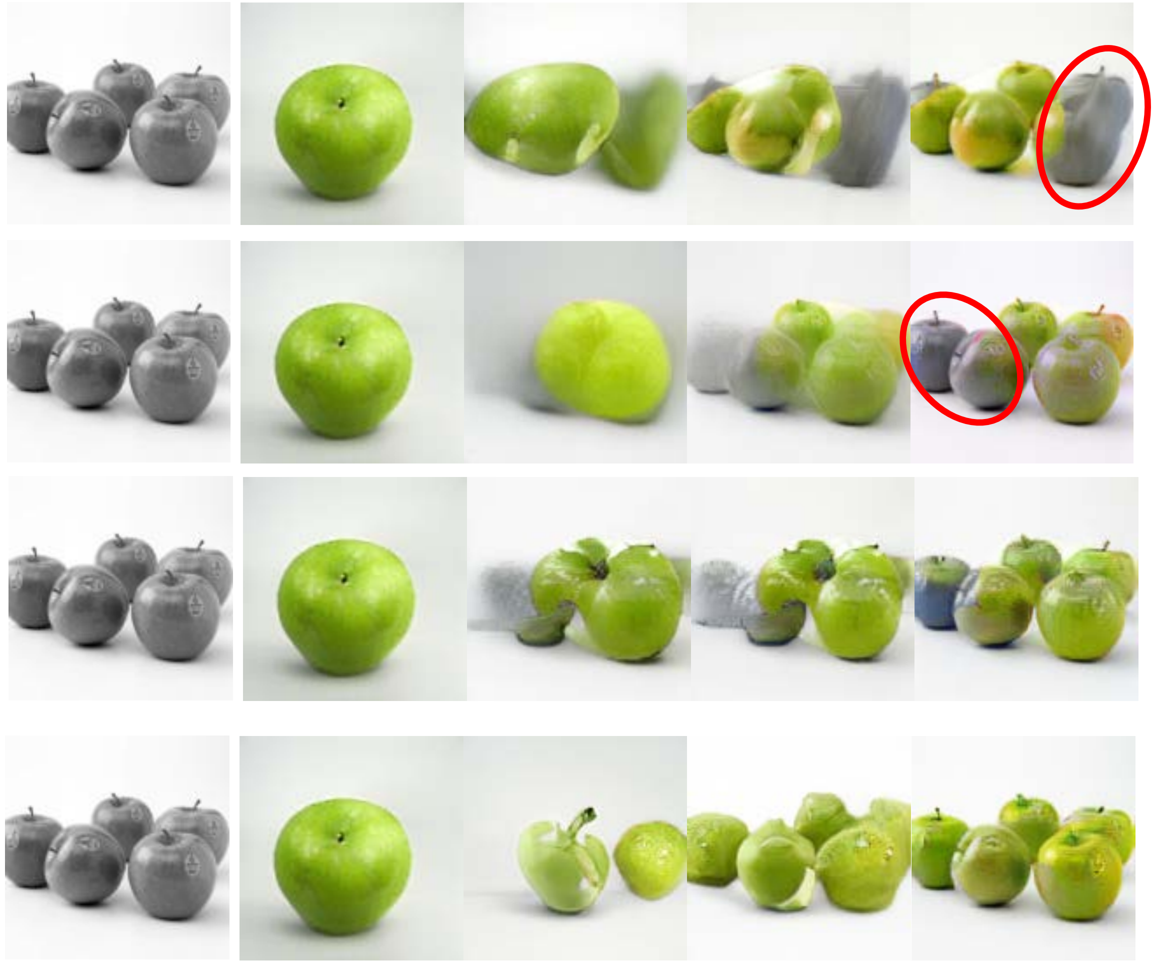


Solution 3: Progressive Reconstruction



Conventional Loss

Target $\hat{\mathbf{x}}$ —— $G(\mathbf{z};\theta)$ during finetuning →



MSE

Perceptual Loss

Discriminator Loss

Discriminator Loss
+
Progressive Reconstruction

Applications

Degradation transform $\phi(\mathbf{x})$ for different tasks:

Colorization: $\phi(\mathbf{x}) = 0.2989\mathbf{x}_r + 0.5870\mathbf{x}_g + 0.1140\mathbf{x}_b$

Inpainting: $\phi(\mathbf{x}) = \mathbf{x} \odot \mathbf{m}$ $\mathbf{m}$ is inpainting mask

Super-resolution: $\phi(\mathbf{x})$ is Lanczos downsampling operator

Model: BigGAN trained on ImageNet training set (Brock et al. ICLR2018)

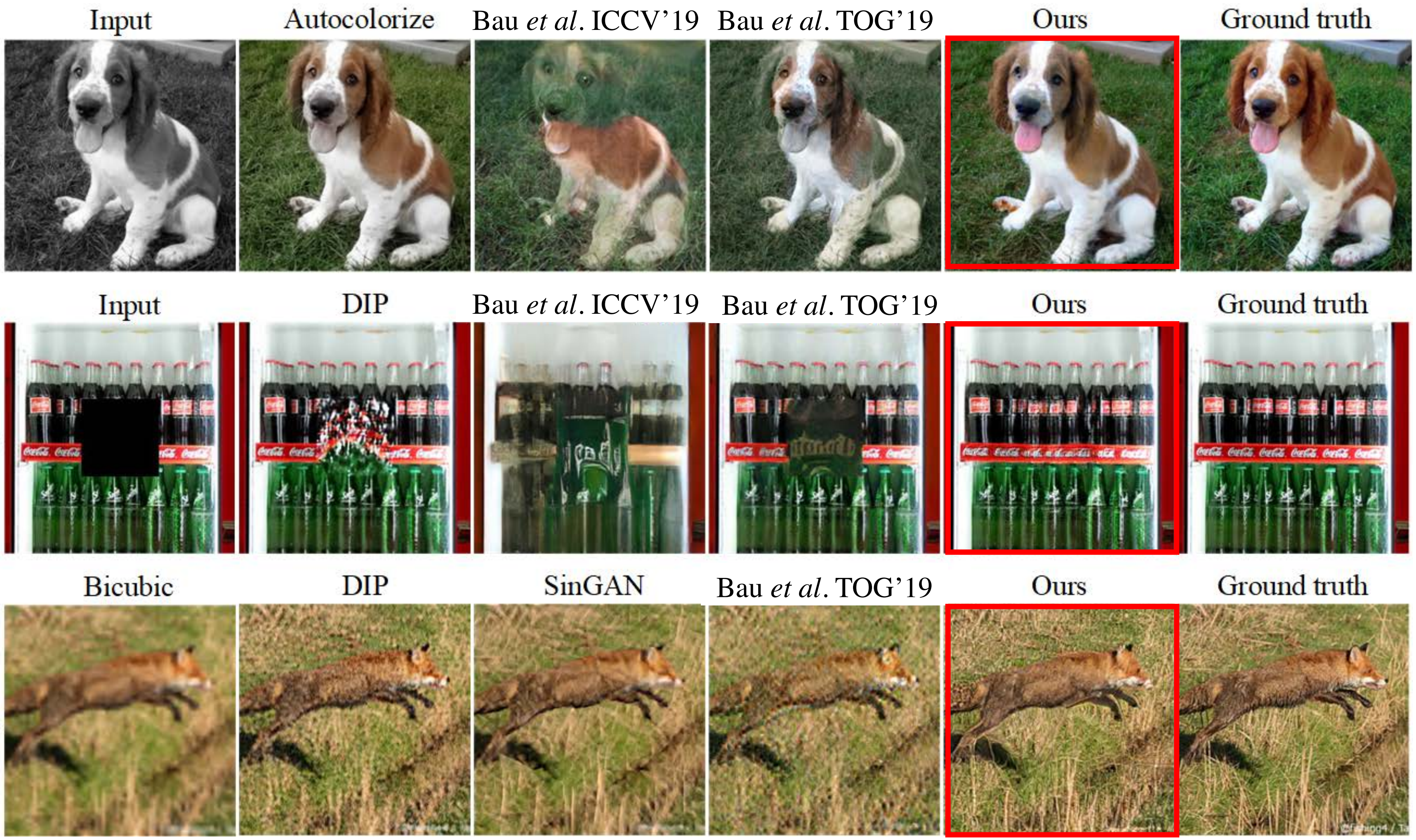
Evaluation: ImageNet validation set

Application: Image Restoration

Colorization

Inpainting

Super-resolution



Demo: Image Restoration

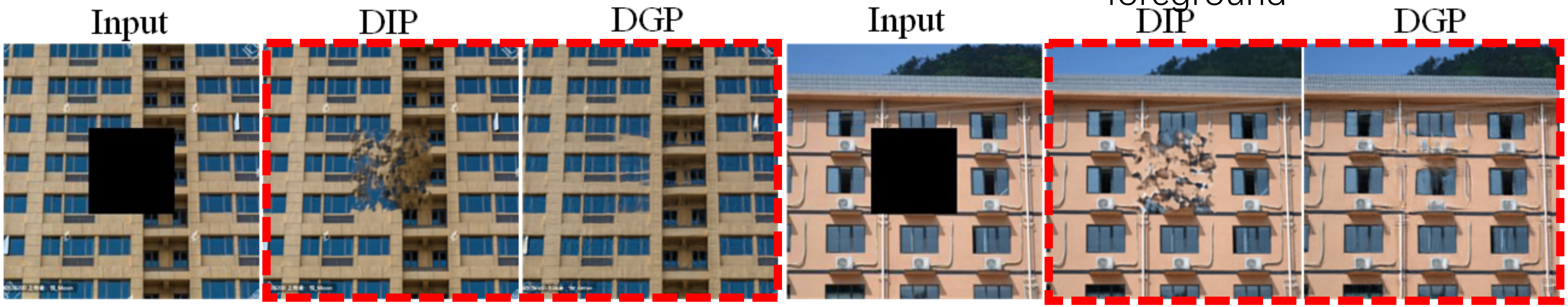
Colorization

Inpainting

Super-resolution

Generalization to non-ImageNet Images

(a) Raccoon (b) Places (c) No foreground

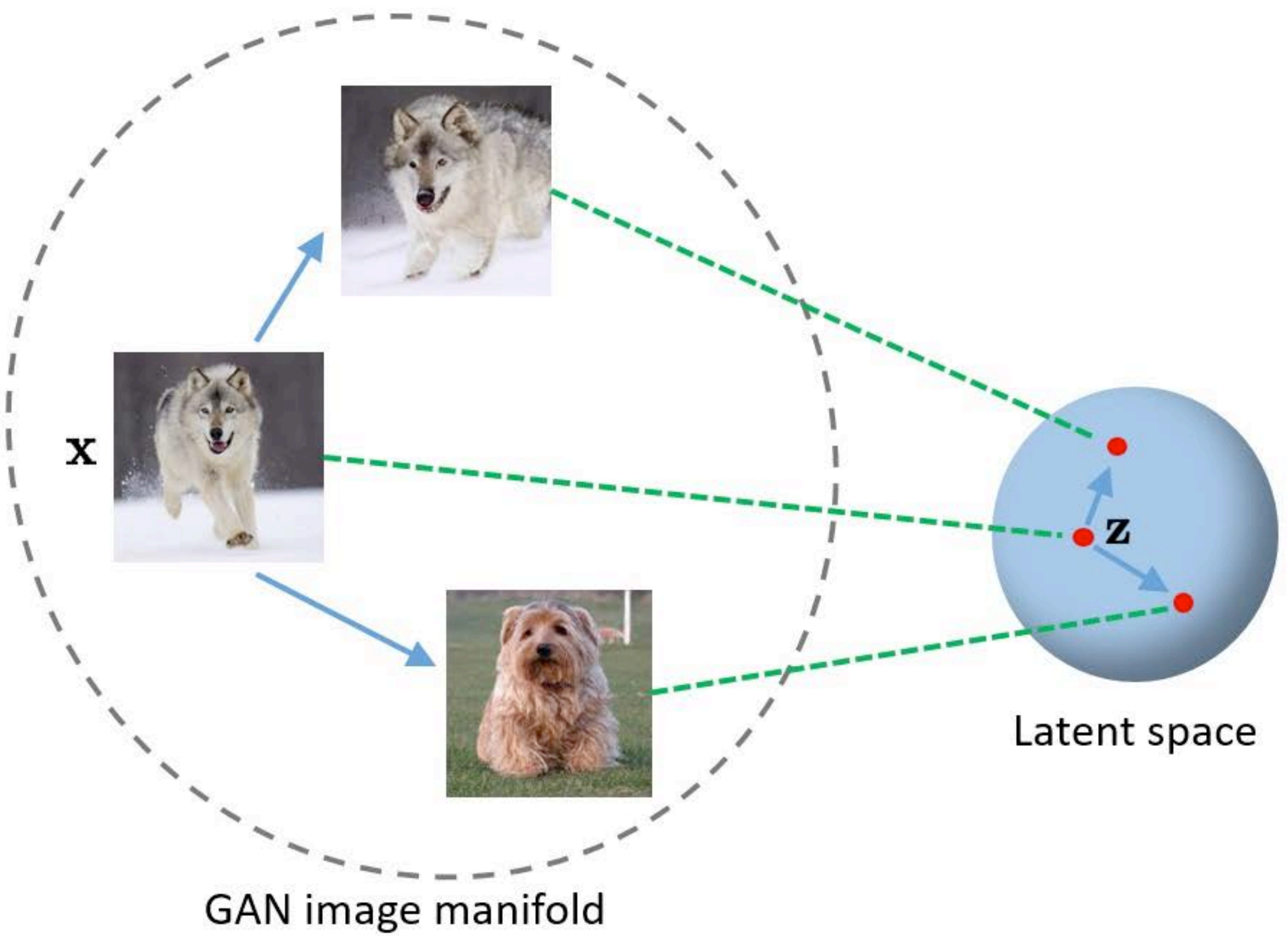


(d) Windows

Flexibility of DGP

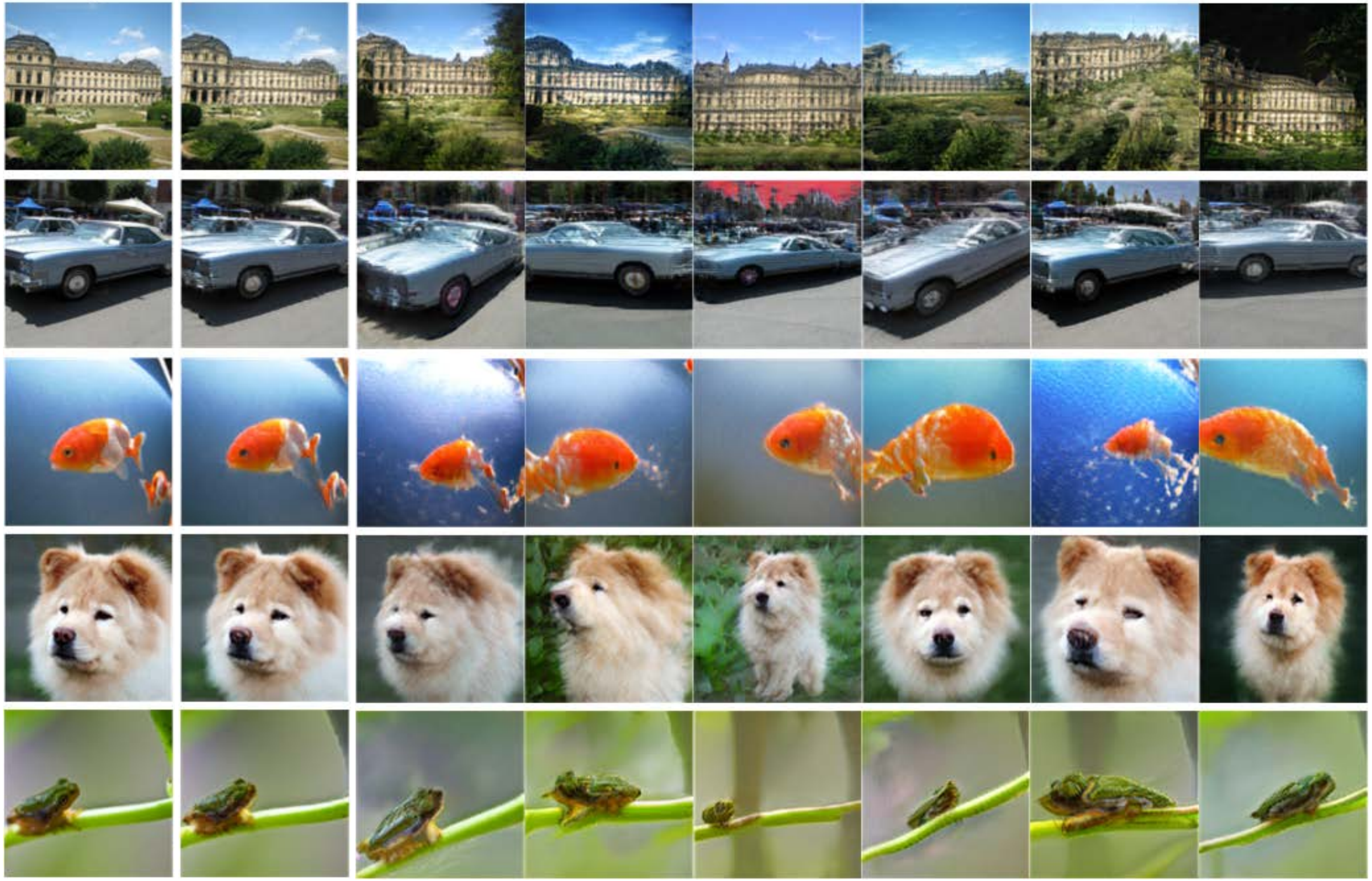
Hybrid restoration

Application: Image Manipulation

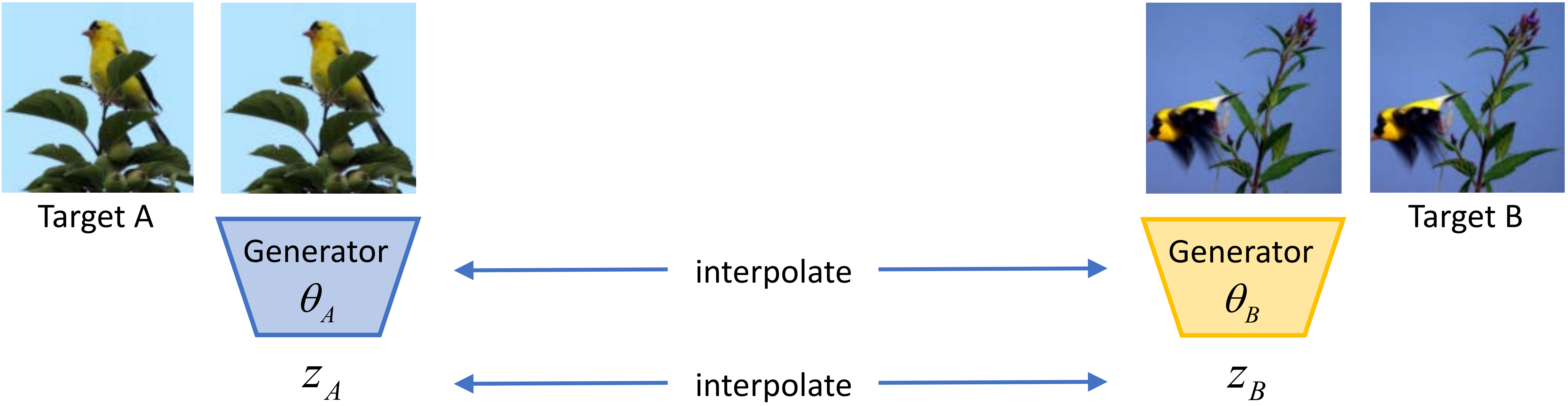


Application: Random Jittering

target reconstruction random jittering effects

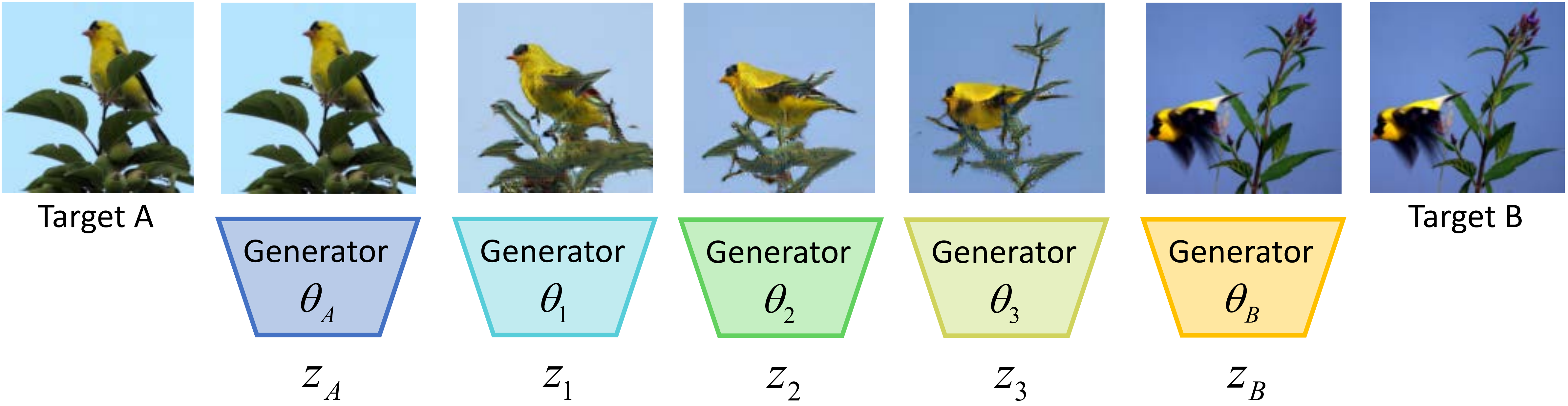


Application: Image Morphing



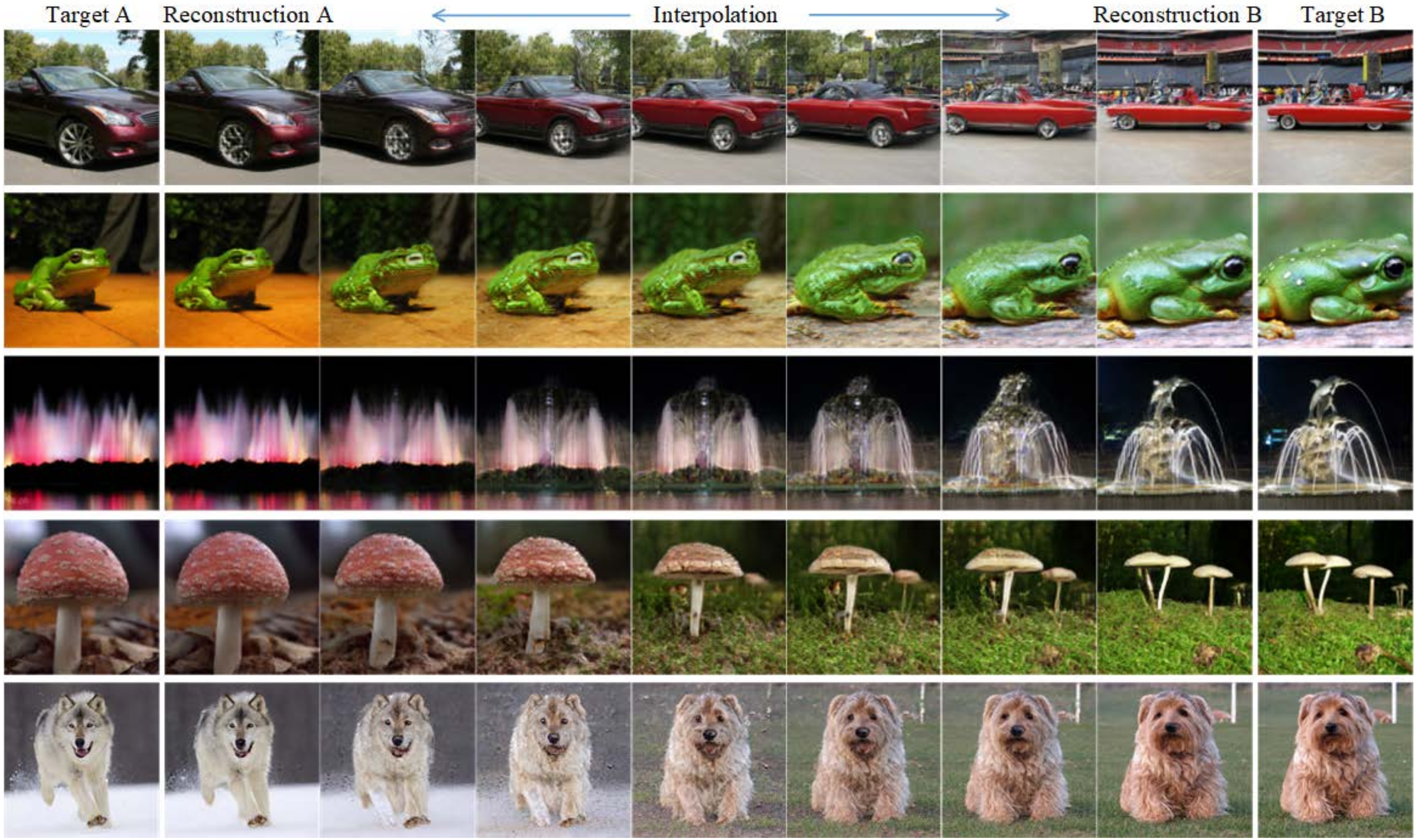
Xintao Wang et al. Deep Network Interpolation for Continuous Imagery Effect Transition， CVPR 2019

Application: Image Morphing



Application: Image Morphing

Application: Image Morphing



Improved GAN Inversion

Relaxed GAN Inversion

$$\theta^*, \mathbf{z}^* = argmin_{\theta,\mathbf{z}} \mathcal{L}(\hat{\mathbf{x}}, \phi(G(\mathbf{z};\theta)))$$

Reconstruction strategies

- Feature matching loss from the coupled discriminator
- Progressive reconstruction

Optimization is conducted in an iterative manner for each image at runtime.

Outline

1. How to exploit generic image prior in pretrained GANs?
2. How to exploit pretrained GAN as a latent bank for large-factor image super-resolution?
3. Do 2D GANs model 3D geometry? Shape-from-X where X=GAN?

Using GANs as Latent Banks

GAN Inversion

Generator of pretrained GANs

Need to solve an expensive optimization problem

$$\theta^*, \mathbf{z}^* = argmin_{\theta,\mathbf{z}} \mathcal{L}(\hat{\mathbf{x}}, \phi(G(\mathbf{z};\theta)))$$

Encoder-Decoder Structure

Encoder | Decoder

A common architecture

It is typically trained from scratch using a combined objective function consisting of a fidelity term and an adversarial loss

The generator is responsible for both capturing the natural image characteristics and maintaining the fidelity to the ground-truth. This inevitably limit its capability of approximating the natural image manifold.

Encoder-Bank-Decoder Structure

Encoder | Generator of pretrained GANs | Decoder

Lifts the burden of learning both fidelity and texture generation simultaneously

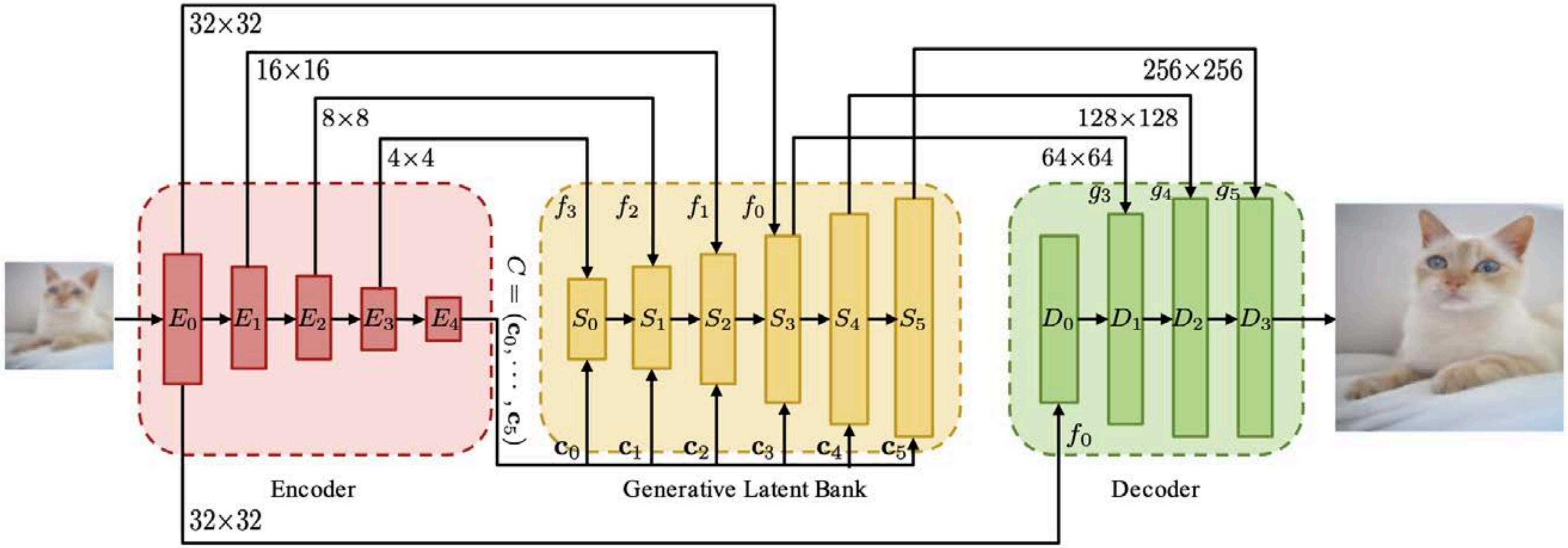
Does not involve image-specific optimization at runtime

Needs a single forward pass to perform image restoration

Inspired by the classic notion of dictionary but exploit GAN as a more effective way for storing priors

GLEAN: Generative LatEnt BANk for Large-Factor Image Super-Resolution
Arxiv preprint, 2020

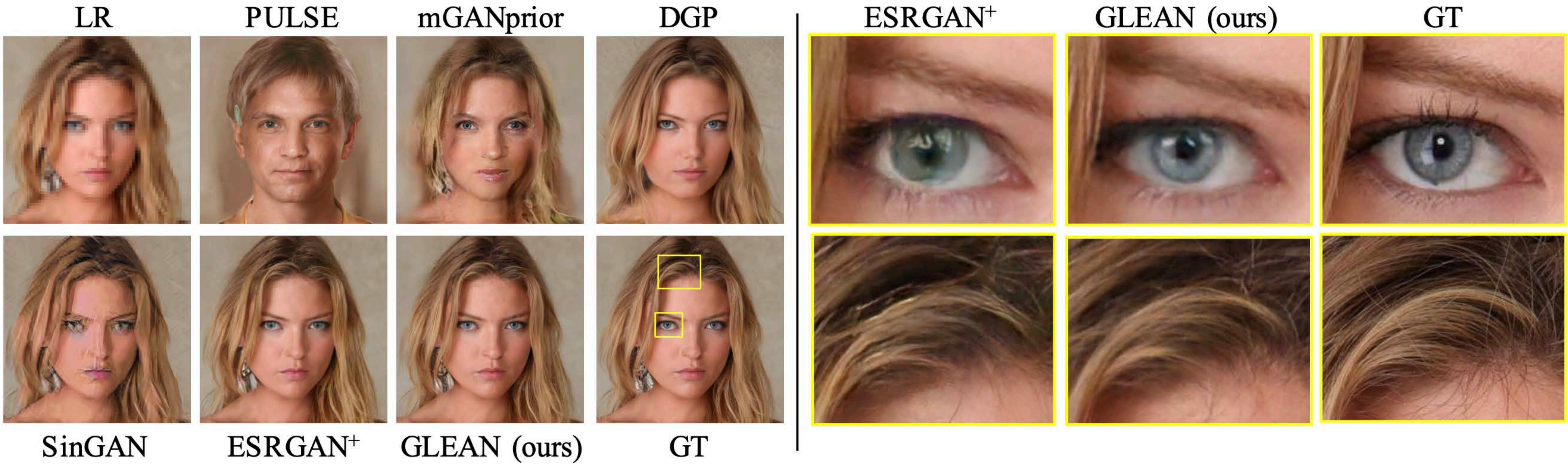
Using GANs as Latent Banks



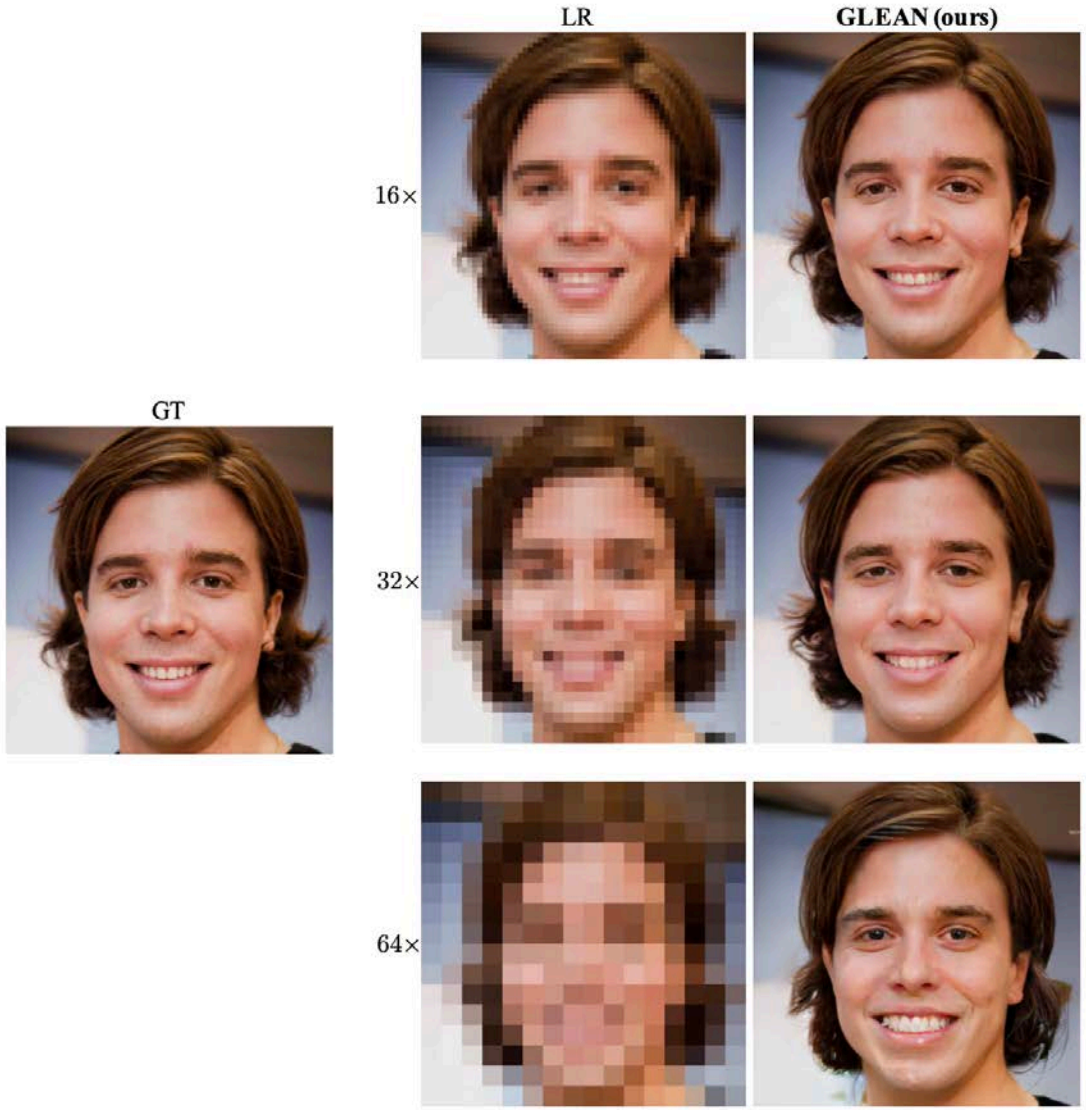
Condition the bank by passing both the latent vectors and multi-resolution convolutional features from the encoder to achieve high-fidelity results. Symmetrically, multi-resolution cues need to be passed from the bank to the decoder.

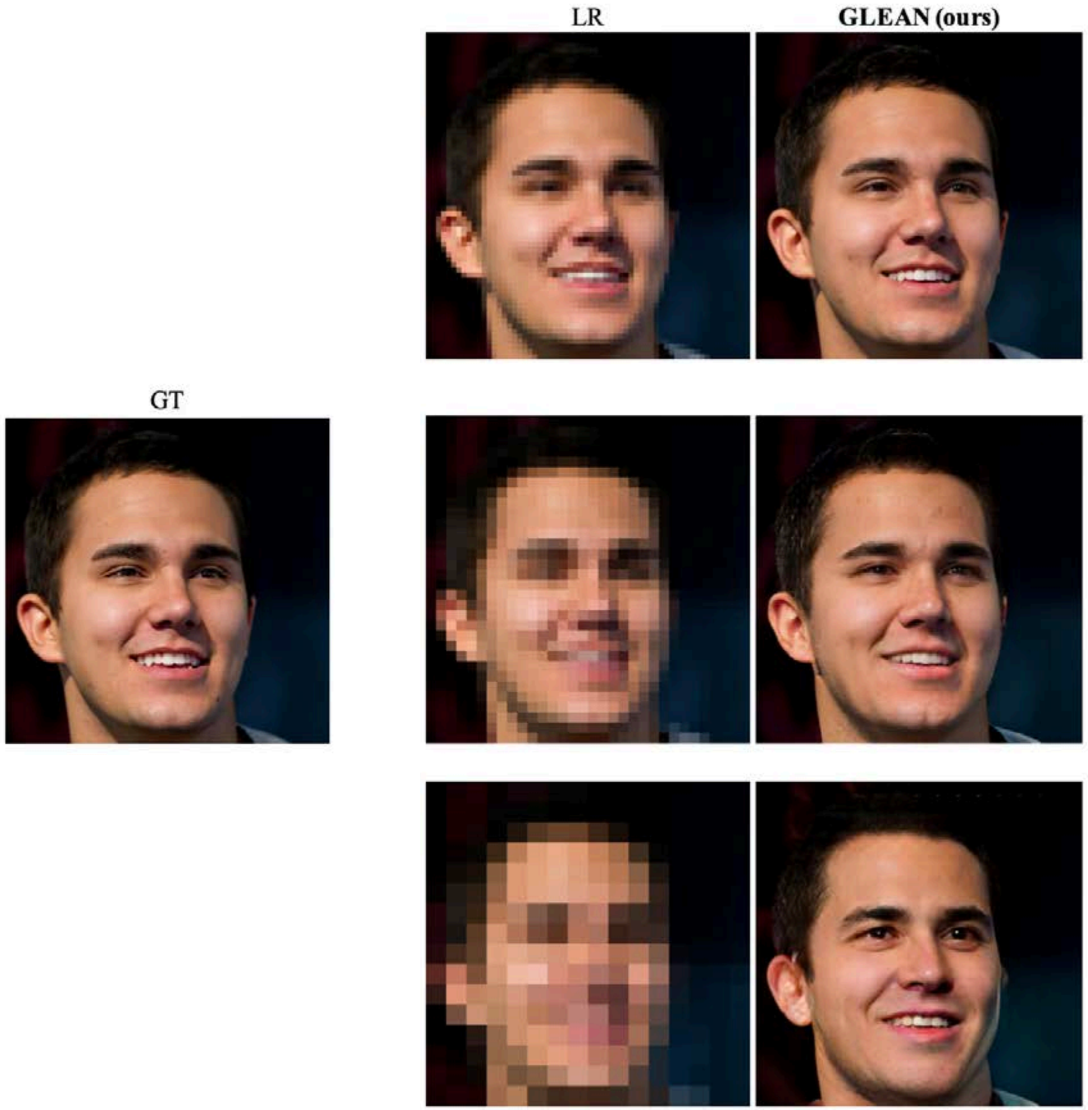
GLEAN: Generative Latent Bank for Large-Factor Image Super-Resolution
Arxiv preprint, 2020

Using GANs as Latent Banks



Using GANs as Latent Banks





Using GANs as Latent Banks

Using GANs as Latent Banks

Pretrained GANs can be exploited in many ways

- GAN inversion
- Encoder-bank-decoder

More results on
https://ckkelvinchan.github.io/projects/GLEAN

1. How to exploit generic image prior in pretrained GANs?

2. How to exploit pretrained GAN as a latent bank for large-factor image super-resolution?

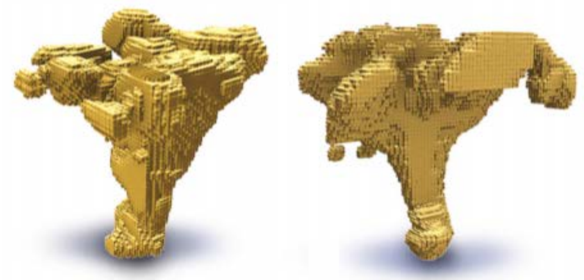
3. **Do 2D GANs model 3D geometry? Shape-from-X where X=GAN?**

Outline

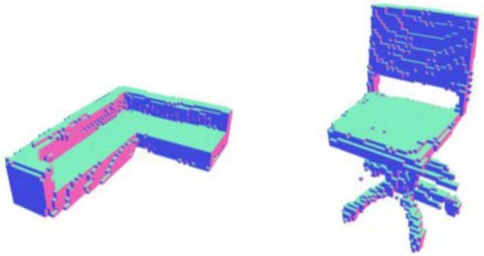
Prior Work

Learning 3D GANs from 2D images: *Need explicit 3D representation for GANs*

Generative3D
(Szabo et al. 2019)

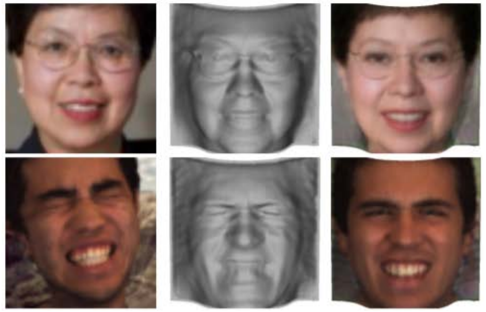


PlatonicGAN
(Henzler et al. ICCV2019)



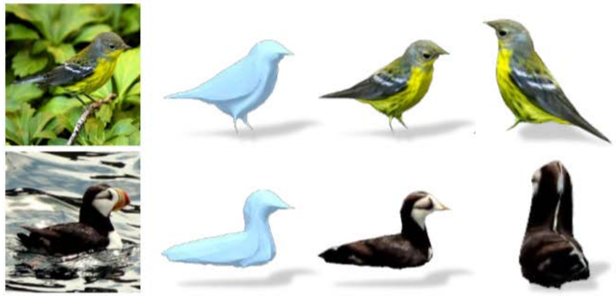
Inverse Graphics GAN
(Lunz et al. 2020)

Unsupervised 3D Shape Learning: *Rely on the symmetry assumption on shapes*



Unsup3d
(Wu et al. CVPR2020)

U-CMR
(Goel et al. ECCV2020)



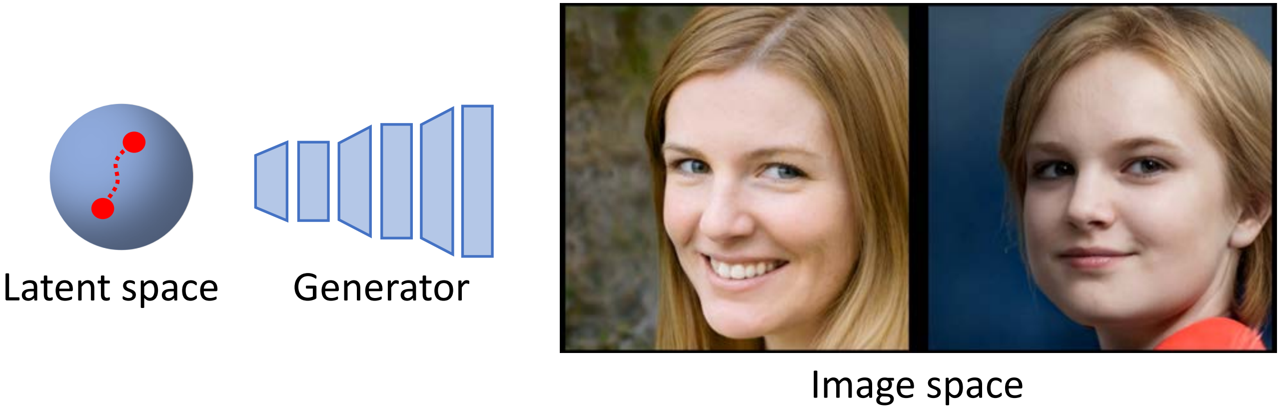
UMR
(Li et al. ECCV2020)

Do 2D GANs Model 3D geometry?

Natural images are projections of 3D objects on a 2D image plane.

An ideal 2D image manifold (e.g., GAN) should capture 3D geometric properties.

The following example shows that there is a direction in the GAN image manifold that corresponds to viewpoint variation.



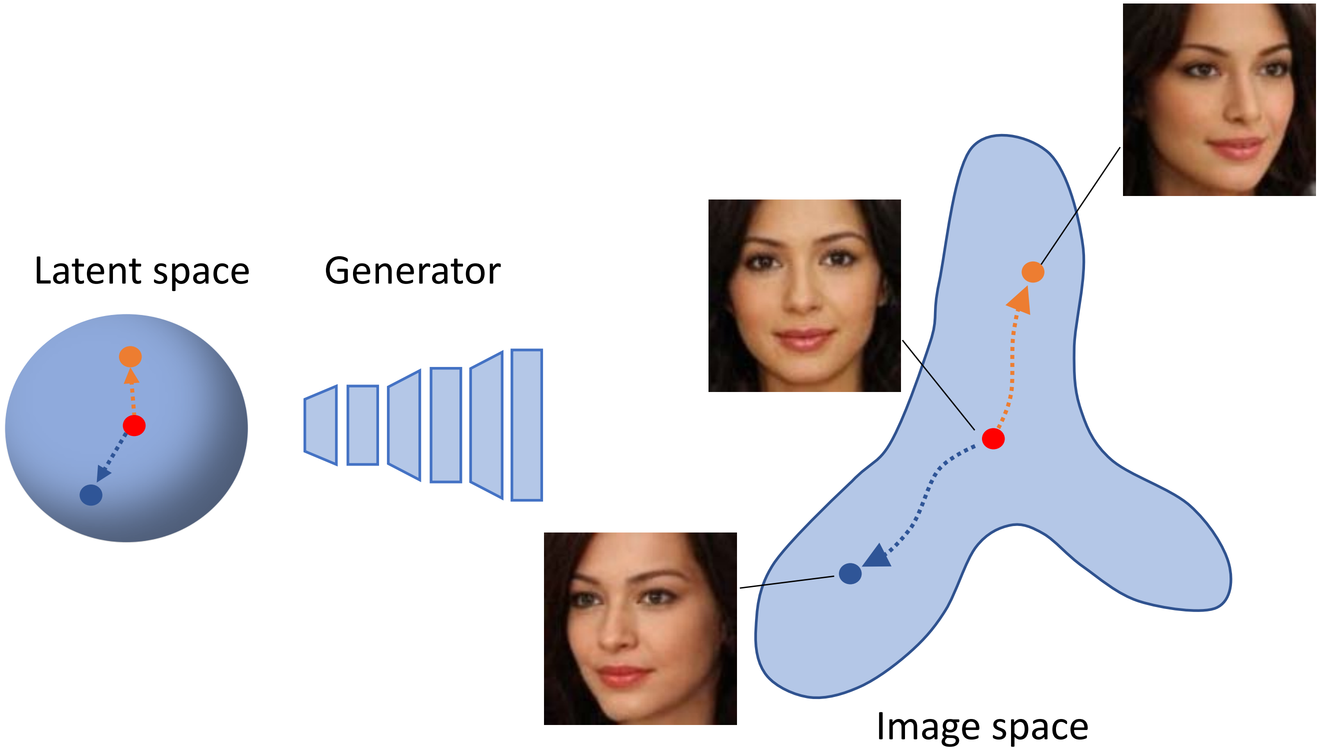
Can we Make Use of such Variations?

Can we make use of such variations for 3D reconstruction?

If we have multiple viewpoint and lighting variations of the same instance, we can infer its 3D structure.

Let's create these variations by exploiting the image manifold captured by 2D GANs!

Challenge



It is non-trivial to find well-disentangled latent directions that control viewpoint and lighting variations in an unsupervised manner.

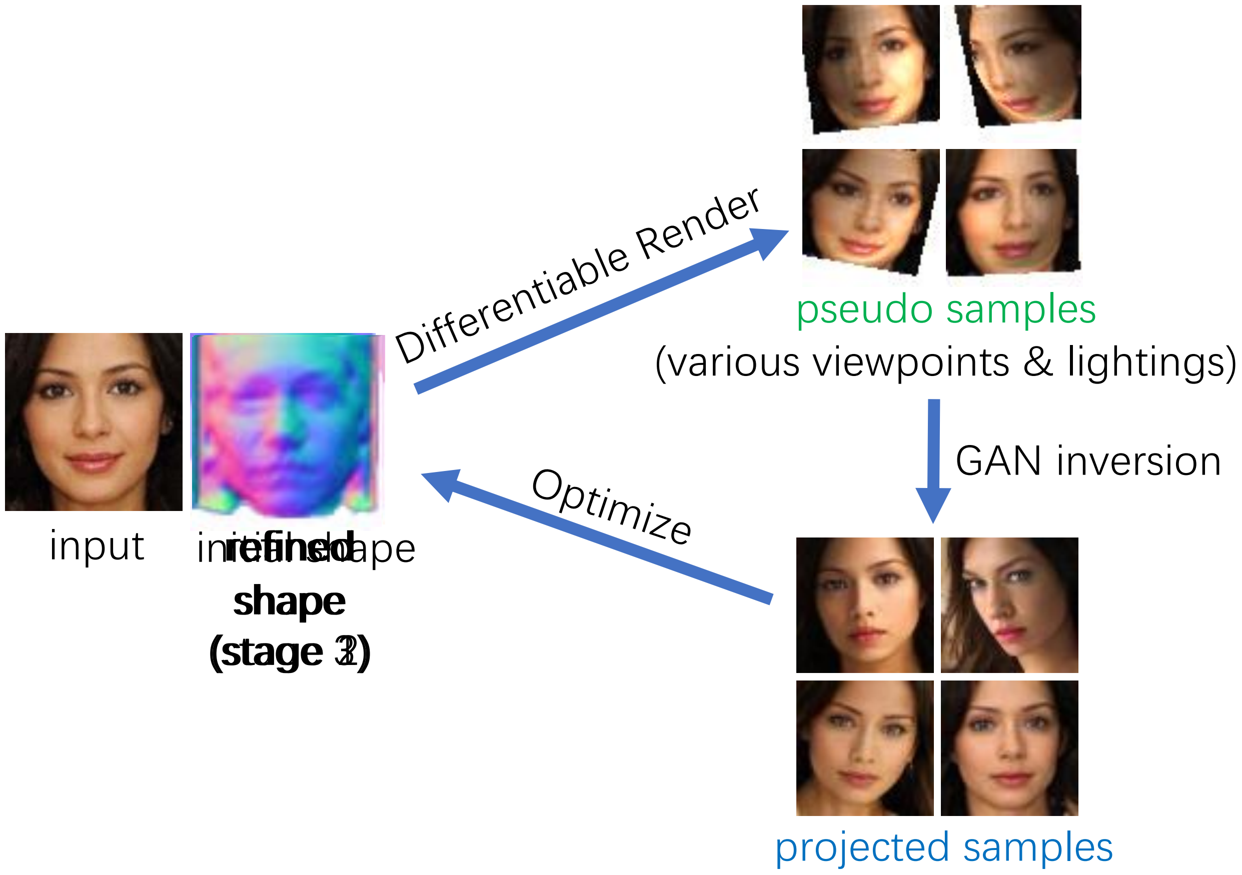
Our Solution

Idea 1: Many objects such as faces and cars, a convex shape prior like ellipsoid could provide a hint on the change of their viewpoints and lighting conditions

Idea 2: Use GAN inversion constrained by this prior to “find” the latent directions.

Steps

- Initialize the shape with ellipsoid.
- Render ‘*pseudo samples*’ with different viewpoints and lighting conditions.
- GAN inversion is applied to these samples to obtain the ‘*projected samples*’
- ‘*Projected samples*’ are used as the ground truth of the rendering process to optimize the 3D shape.
- Iterative training to progressively refine the shape.



Methodology

Step1:

Initialize shape with ellipsoid.

Optimize albedo network *A*.

Step2:

Render 'pseudo samples' with various viewpoints & lightings.

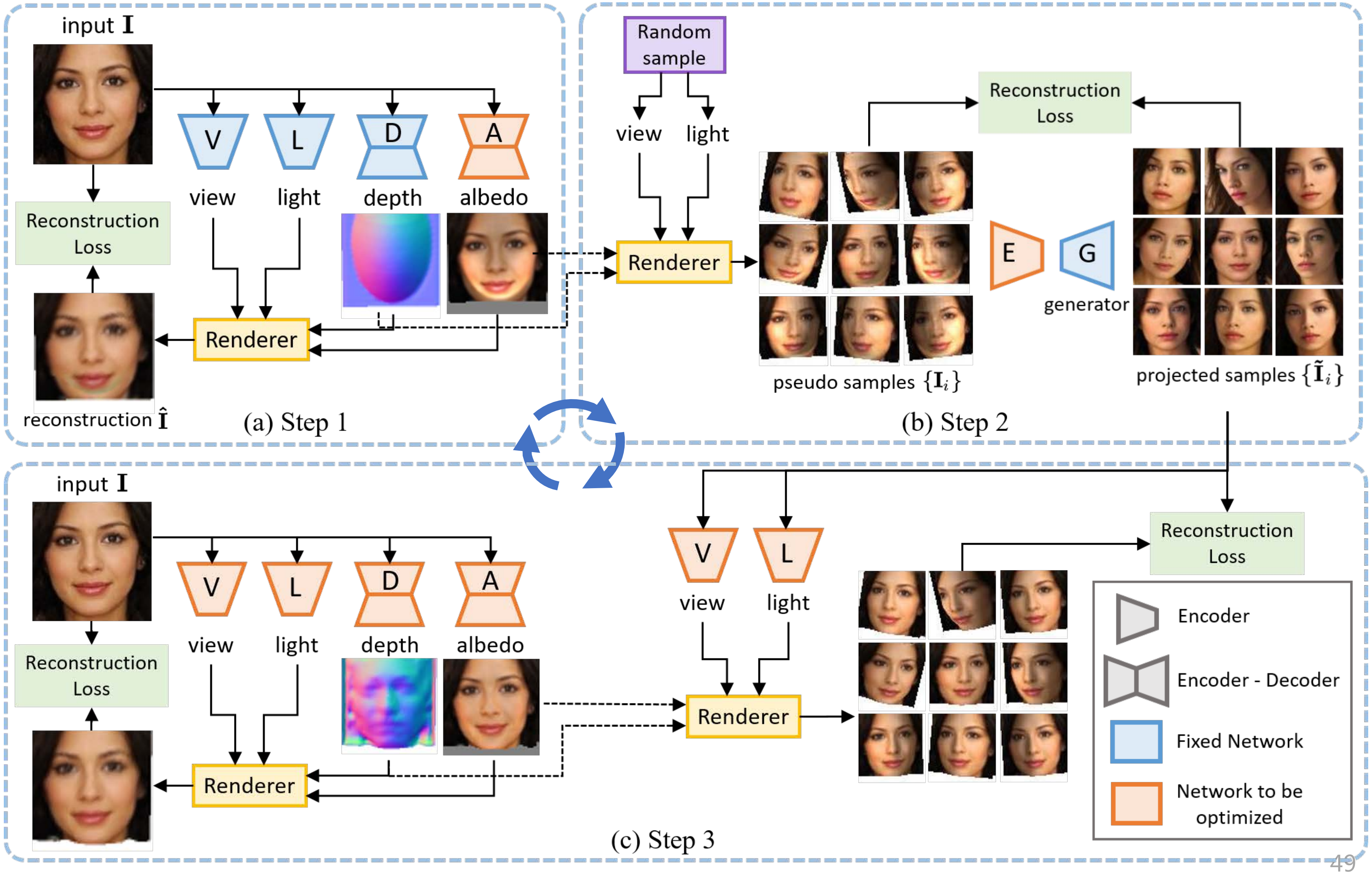
Perform GAN inversion to the pseudo samples to obtain the 'projected samples'.

Optimize latent encoder *E*.

Step3:

Reconstruct 'projected samples' with shared depth & albedo and independent view & light.

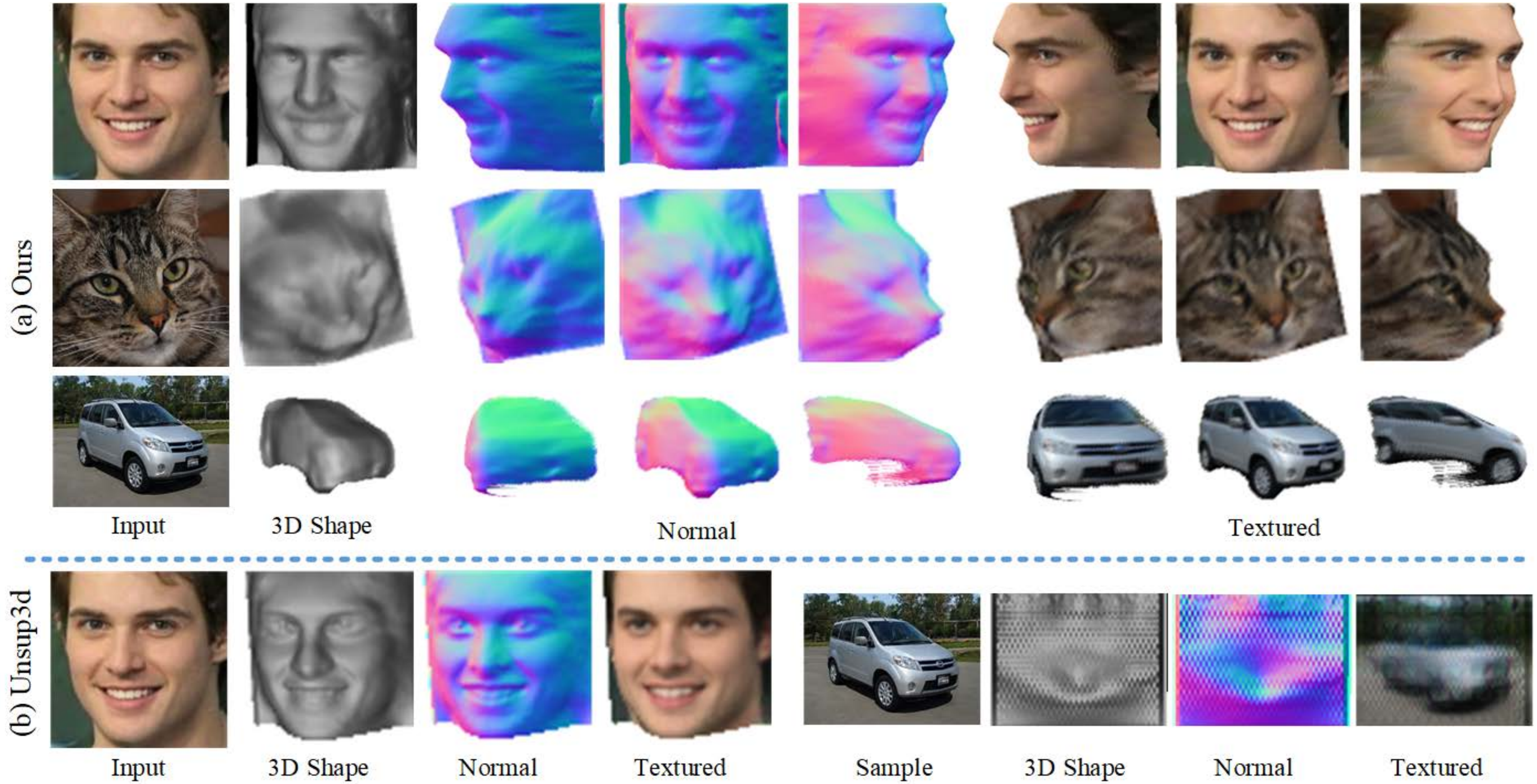
Optimize network *V, L, D, A*.



3D Reconstruction Results

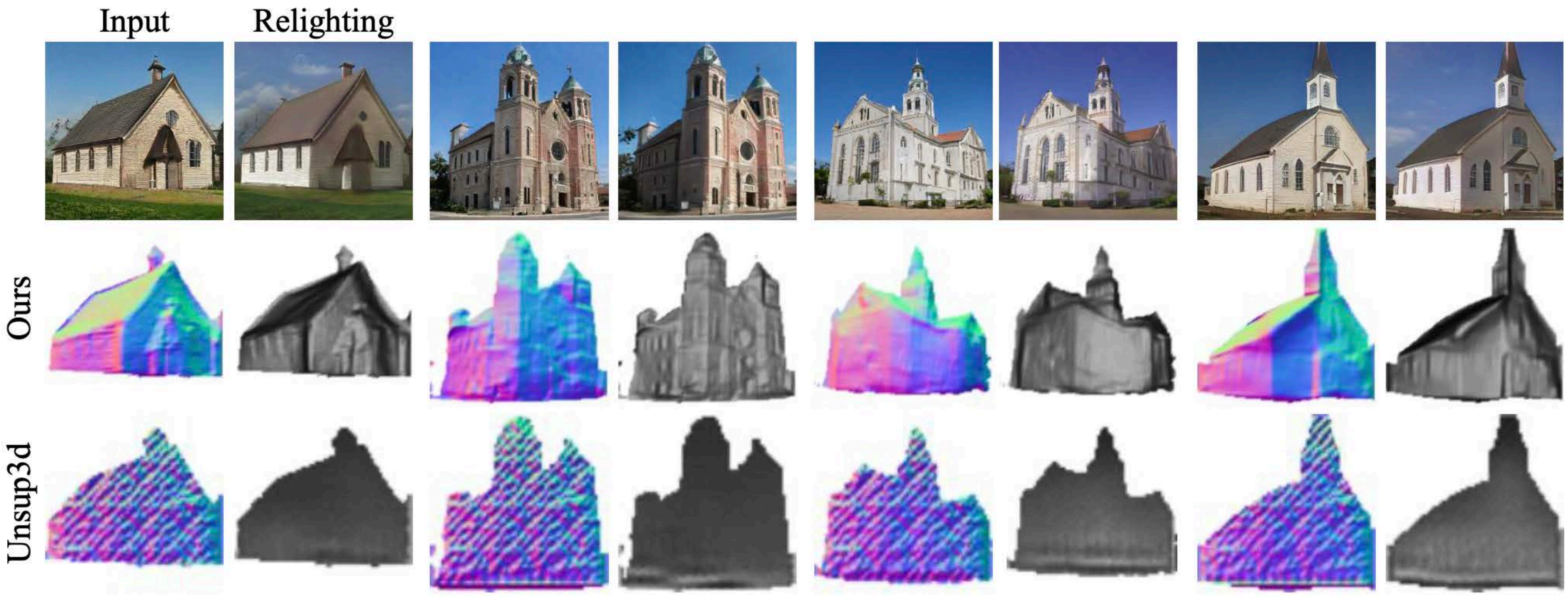
Without any 2D keypoint or 3D annotations
Unsupervised 3D shape reconstruction from unconstrained 2D images

Without symmetry assumption
Work on many object categories such as human faces, cars, buildings, *etc*.



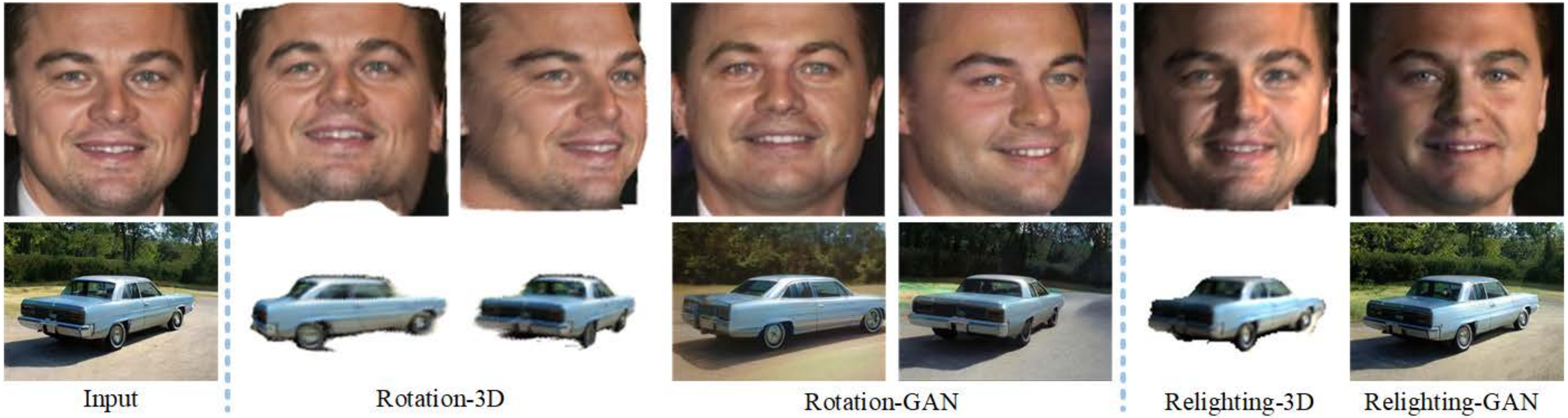
Shangzhe Wu, Christian Rupprecht, and Andrea Vedaldi. Unsupervised learning of probably symmetric deformable 3D objects from images in the wild. In *CVPR*, 2020

3D Reconstruction Results



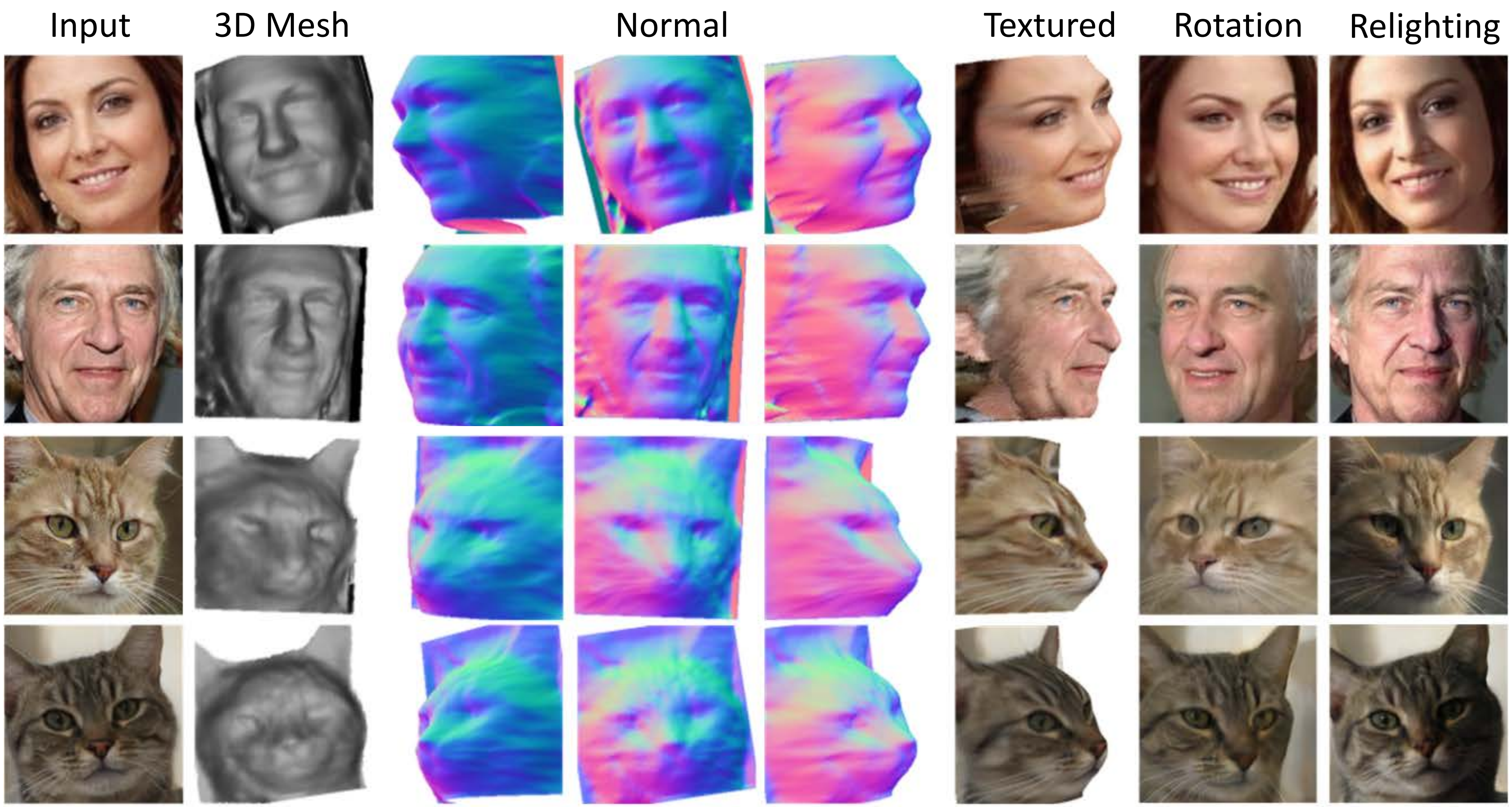
Shangzhe Wu, Christian Rupprecht, and Andrea Vedaldi. Unsupervised learning of probably symmetric deformable 3D objects from images in the wild. In *CVPR*, 2020

3D-aware Image Manipulation

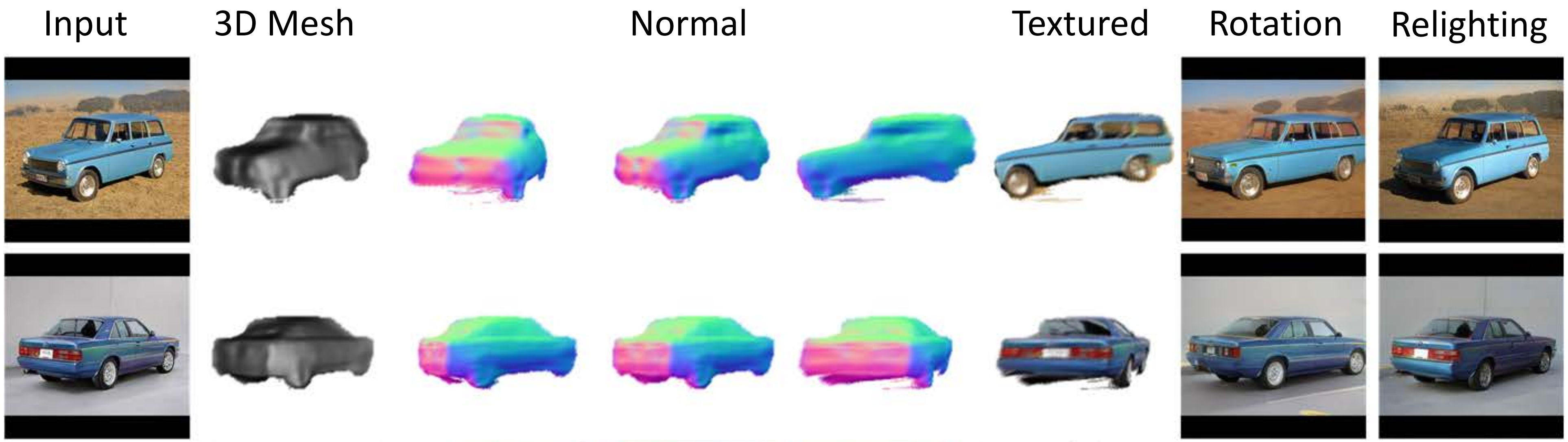


- Effect-3D: Rendered using the reconstructed 3D shape and albedo.
- Effect-GAN: Obtained by performing GAN inversion to "Effect-3D" using the trained encoder *E*.

More Results

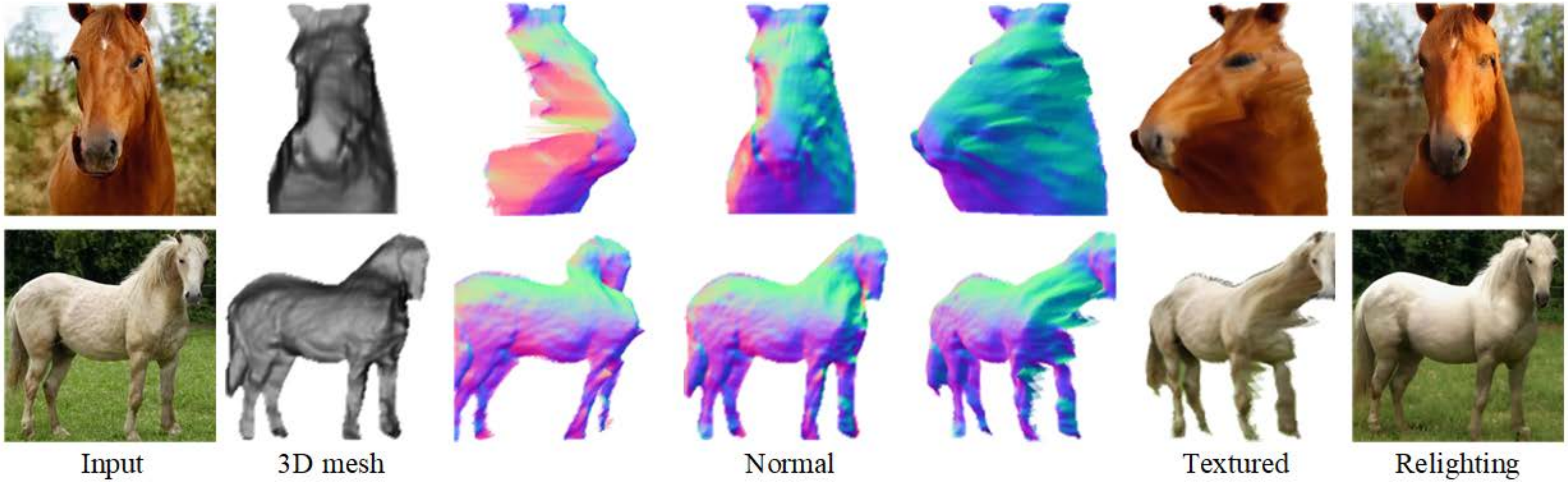


More Results



More Results

More Results



Summary

A GAN generator trained on massive natural images

- Can be used as a generic image prior for many image restoration and manipulation tasks
- Inherently captures the underlying 3D geometry of objects

We show approaches

- Restore the missing information of a degraded image by progressively reconstructing it under the discriminator metric
- Use GAN as a latent bank for single forward pass restoration
- Shape-from-GAN i.e., recovering 3D shape from unconstrained 2D images without relying on the symmetry assumption or external 3D models