On Leveraging Pretrained GANs for Generation with Limited Data

Miaoyun Zhao, Yulai Cong, Lawrence Carin

Duke University

August 11, 2020

Table of Contents

Motivation


Generated from BigGAN


Generated from StyleGAN

GANs can generate highly realistic synthetic (“fake”) images

- Can augment training data, with new & realistic samples
- Useful in settings with limited training data

However, training the GAN itself is challenging with limited data

- Training GANs with limited data may yield overfitting or training/mode collapse

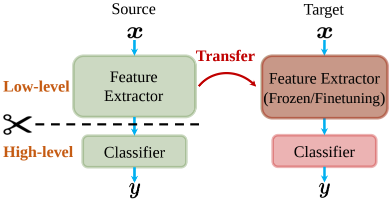
Propose to transfer additional information to facilitate GAN training with limited data

- Leverage valuable generalizable knowledge within GANs trained on different large-scale datasets

Motivation

Key observations associated with generalizable knowledge:

- For classification models pretrained on large-scale datasets
  - lower-level filters (those close to the observation x) are fairly general/transferable (Gabor-like)
  - higher-level filters are more task-specific



- For pretrained GAN generators
  - lower-level layers portray generally-applicable local patterns
  - higher-level layers represent more specific semantic objects or object parts
- It's data-demanding to train well-behaved low-level filters
  - transfer often delivers better efficiency and performance

Our Contributions

To better transfer common knowledge for generators, for design of generators based on limited data

- From GANs pretrained on large-scale source datasets

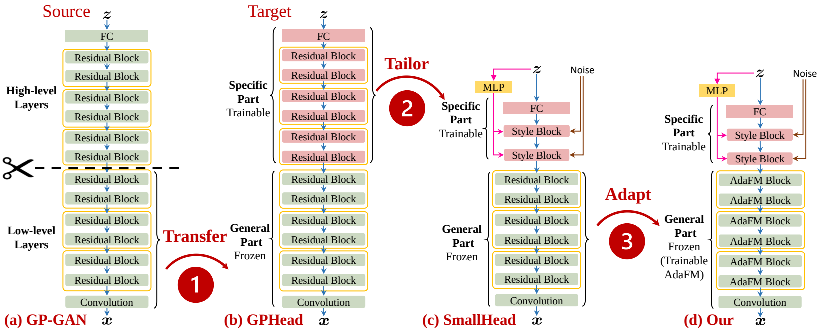
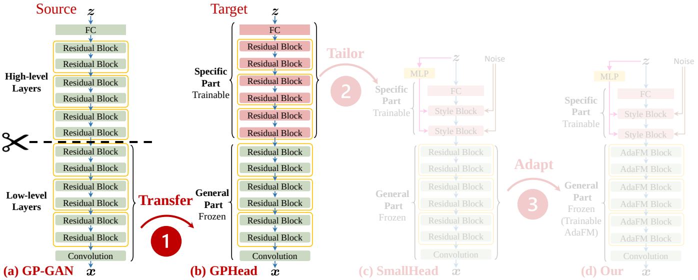


Table of Contents

Notation

- Within a GAN, there is a generator (actor) and a discriminator (critic)
- “General-Part” of either the generator or discriminator is composed of those model layers that are generally applicable across a wide range of images
- “Specific-Part” of generator or discriminator composed of layers that are specifically associated with a class of images
- Seek to transfer General-Part from GANs learned in data-rich settings, to those for which there are limited data
- The General-Part tends to be at and near layers that touch the input (discriminator) or output (generator) image

1. On Specifying the General-Part for Transfer



Source model: the GP-GAN[1] pretrained on ImageNet

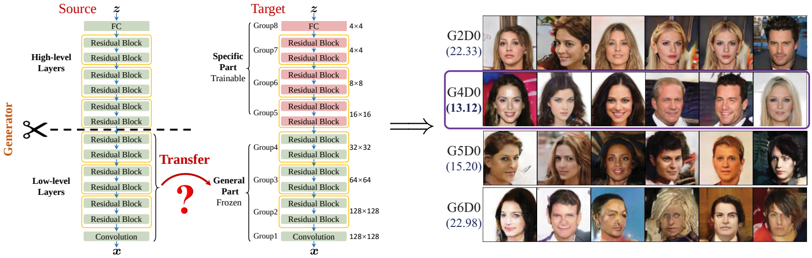
Target dataset: the perceptually-distinct CelebA

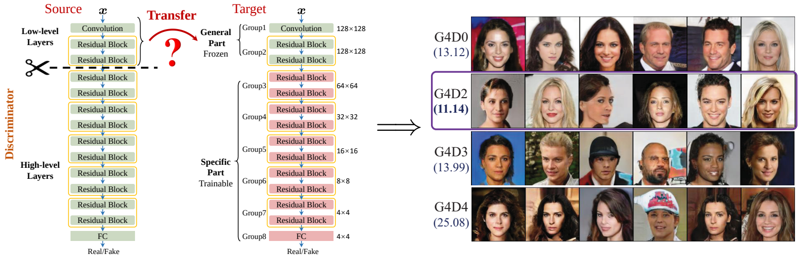

ImageNet


CelebA

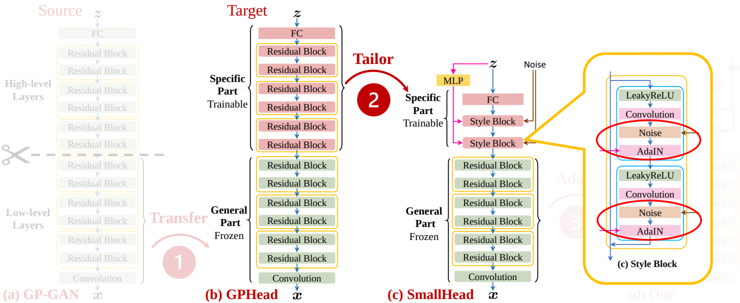
[1] Which training methods for GANs do actually converge? ICML 2018.

1. On Specifying the General-Part for Transfer





2. On Tailoring the High-Level Specific-Part

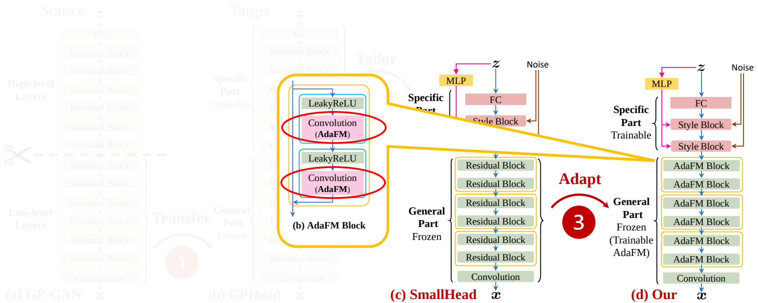


Even with the G4D2 general-part, mode collapse may still happen on small data (Flowers 8,189).

Style blocks deliver

- disentangled high-level attributes ⋙ efficient exploration of underlying data manifold ⋙ better generative quality
- style mixing
- cheaper computation

3. On Better Adaption of the Transferred General-Part



We introduce the **adaptive filter modulation (AdaFM)** to

- better adapt the transferred general-part to target domains
- relax the requirements for the general-part

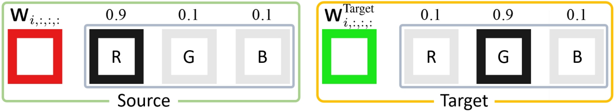
Given a Conv filter $\mathbf{W} \in \mathbb{R}^{C_{\text{out}} \times C_{\text{in}} \times K_1 \times K_2}$, AdaFM uses learnable $\gamma \in \mathbb{R}^{C_{\text{out}} \times C_{\text{in}}}$ and $\beta \in \mathbb{R}^{C_{\text{out}} \times C_{\text{in}}}$ to modulate its statistics

$$\mathbf{W}^{\text{AdaFM}}_{i,j,:,:} = \gamma_{i,j} \mathbf{W}_{i,j,:,:} + \beta_{i,j} \quad (1)$$

3. On Better Adaption of the Transferred General-Part

The underlying assumption is

- basic shape/pattern within $\mathbf{W}_{i,j,:,:}$ $\Rightarrow$ generally applicable
- statistics/correlation among i-,j-channels $\Rightarrow$ target-specific
- empirically verified in the experiments

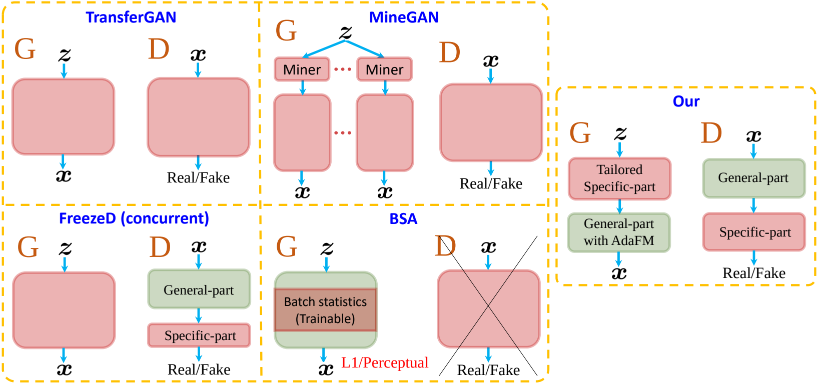


- Source and target filters share the same basic shape/pattern but with different among-channel correlations.
- AdaFM learns $\gamma_{i,:} = [1/9, 9, 1]$ to adapt source $\mathbf{W}_{i,:,:,:}$ to target $\mathbf{W}^{\text{Target}}_{i,:,:,:}$.

Table of Contents

Related Work

Exploit GANs to transfer knowledge for limited-data generation.



TransferGAN: Transferring GANs: generating images from limited data. ECCV 2018.
BSA: Image generation from small datasets via batch statistics adaptation. ICCV 2019.
MineGAN: MineGAN: effective knowledge transfer from GANs to target domains with few images. CVPR 2020.
FreezeD: Freeze discriminator: A simple baseline for fine-tuning GANs. arXiv 2020.

Table of Contents

Experiments

Comparisons with existing/naive methods on

1. moderate or small datasets
2. limited datasets with 1,000 images
3. extremely limited datasets with 25 images

Analysis of the proposed techniques

1. ablation study of our method
2. modulations from AdaFM
3. style augmentation/mixing with the tailored specific-part

Comparisons with Existing/Naive Methods

1. On moderate or small datasets

CelebA (202,599), Flowers (8,189), Cars (8,144), Cathedral (7,350)

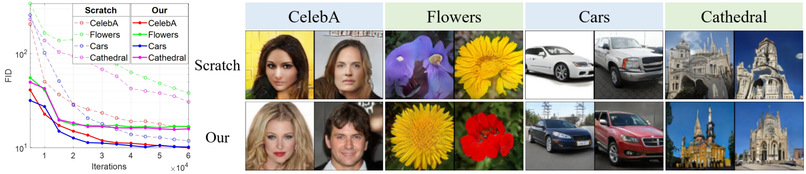


Figure 8. FID scores (left) and generated images (right) of Scratch and Our method on 4 target datasets. The transferred general-part dramatically accelerates the training, leading to better performance.

Table 2. FID scores of the compared methods after 60,000 training iterations. Lower is better. "Failed" means training/mode collapse.

Method\Target	CelebA	Flowers	Cars	Cathedral
TransferGAN	18.69	failed	failed	failed
Scratch	16.51	29.65	11.77	30.59
Our	**9.90**	**16.76**	**10.10**	**15.78**

- TransferGAN vs Scratch/Our: tailored specific-part $\ggg$ overfitting
- Scratch vs Our: (i) the transferred general-part, (ii) AdaFM

Comparisons with Existing/Naive Methods

2. On limited datasets with 1,000 images

Random selection ⋙ CelebA-1K, Flowers-1K, and Cathedral-1K

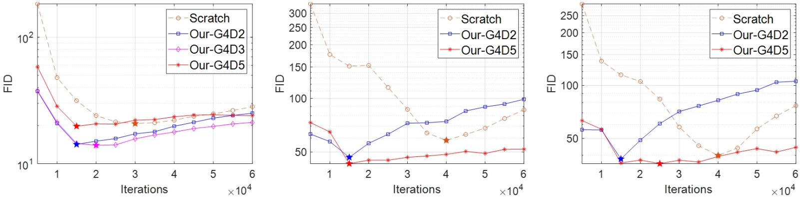


Figure 10. FID scores on CelebA-1K (left), Flower-1K (center), and Cathedral-1K (right). The best FID achieved is marked with a star.

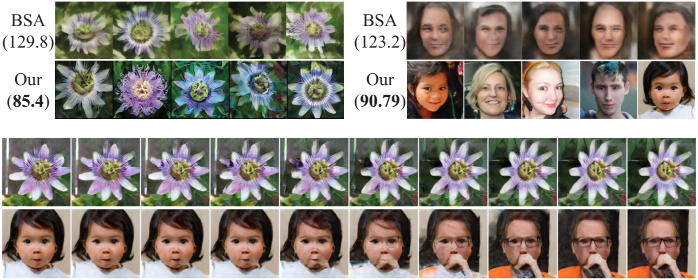
Table 3. The best FID achieved within 60,000 training iterations on the limited-1K datasets. Lower is better.

Method\Target	CelebA-1K	Flowers-1K	Cathedral-1K
Scratch	20.75	58.18	39.97
Our-G4D2	14.19	46.68	38.17
Our-G4D3	**13.99**	-	-
Our-G4D5	19.77	**43.05**	**35.88**

Comparisons with Existing/Naive Methods

3. On extremely limited datasets with 25 images

Random selection ⋙ Flowers-25 and FFHQ-25, following BSA.[2]



Our: G4D6 general-part, GP on both real and fake samples

- More realistic generation
- Smooth interpolations on the learned data manifold

[2] Image generation from small datasets via batch statistics adaptation. ICCV 2019.

Analysis of the Proposed Techniques

1. Ablation Study of Our Method

- GP-GAN: no filters are transferred; baseline for GPHead
- GPHead: GP-GAN architecture + transferred general-part
- SmallHead: transferred general-part + tailored specific-part
- Our: SmallHead + the proposed AdaFM

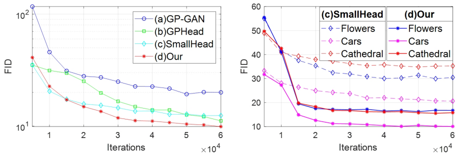


Figure 9. FID scores from the ablation studies of our method on CelebA (left) and the 3 small datasets of Flower, Cars, and Cathedral (right).

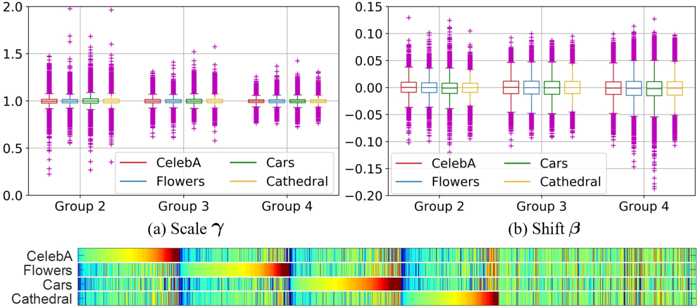
Table 1. FID scores from ablation studies on our method after 60,000 training iterations. Lower is better.

Method\Target	CelebA	Flowers	Cars	Cathedral
(a)GP-GAN	19.48	failed	failed	failed
(b)GPHead	11.15	failed	failed	failed
(c)SmallHead	12.42	29.94	20.64	34.83
(d)Our	**9.90**	**16.76**	**10.10**	**15.78**

Analysis of the Proposed Techniques

2. Modulations from AdaFM

Boxplots of the learned scale γ and shift β on target datasets



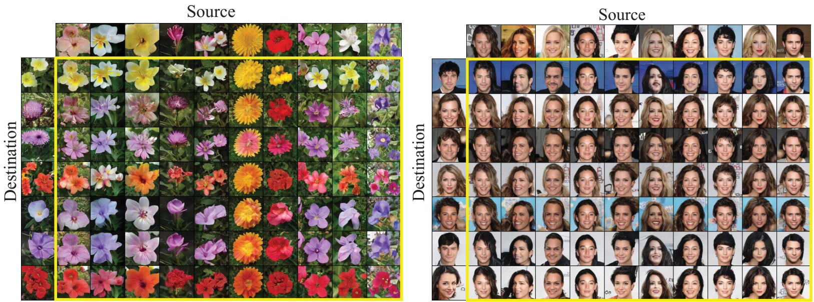
(a) Scale γ (b) Shift β

(c) γ in Group2

- All filters are used in target domains but with modulations
- Different target datasets prefer different modulations

Analysis of the Proposed Techniques

3. Style Mixing/Augmentation with the Tailored Specific-Part



Style mixing is extremely appealing for limited-data applications

- Vast novel generation via style/attribute combinations
- Diverse synthetic augmentation

Conclusions

- For lifelong learning, important to appropriately transfer knowledge from the past to new tasks
- Such transfer critical for performing model learning with limited data
- Have developed a novel means of performing lifelong learning with GAN models
- Allows generation of realistic synthetic data based on limited training data
- By style augmentation, allows significant expansion of training data, generating new and realistic data for training other models (e.g., supervised models)