# Exploring Incompatible Knowledge Transfer in Few-shot Image Generation

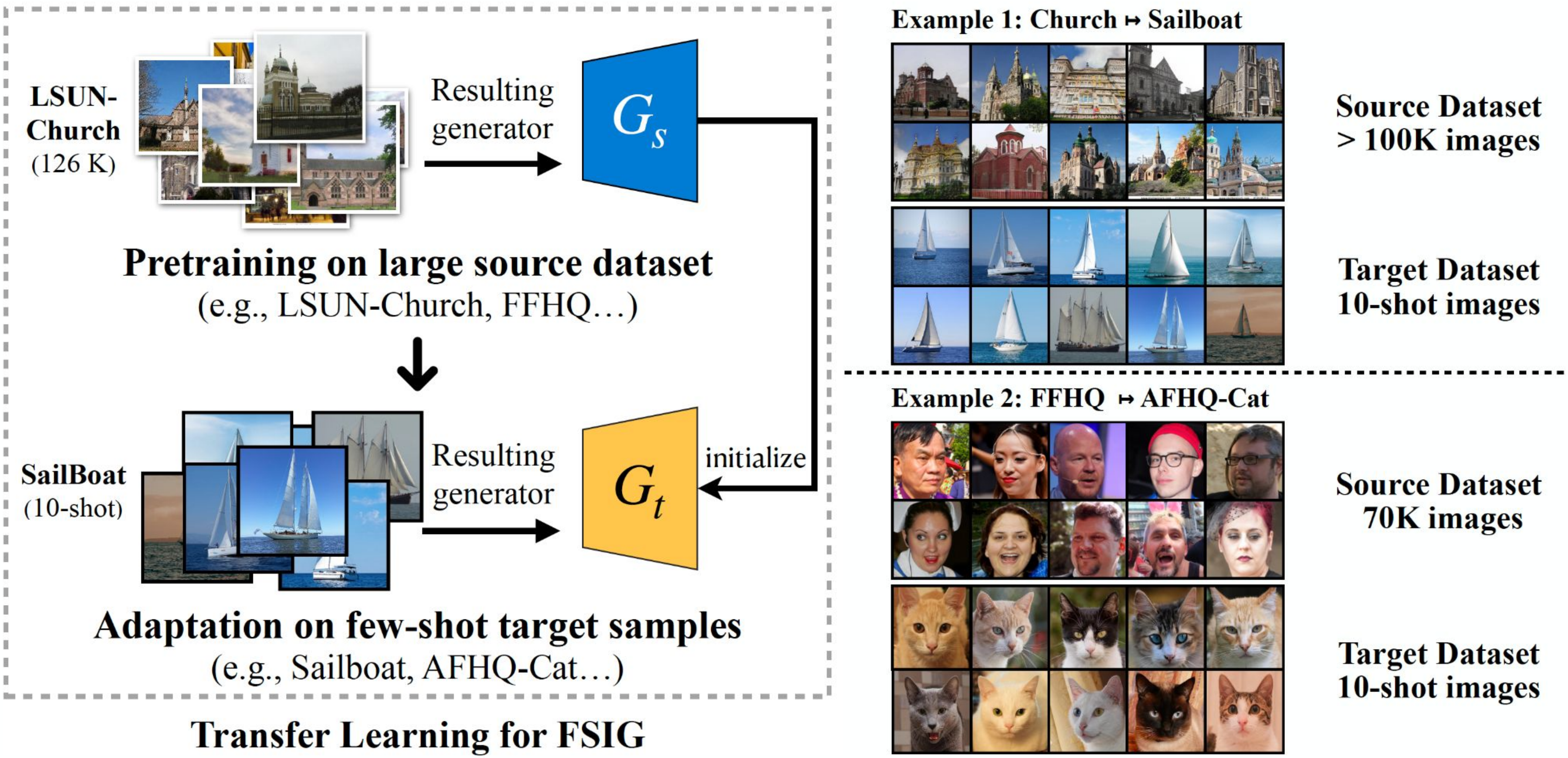
Yunqing Zhao[1] Chao Du[2] Milad Abdollahzadeh[1] Tianyu Pang[2]
Min Lin[2] Shuicheng Yan[2] Ngai-Man Cheung[1]

[1]Singapore University of Technology and Design (SUTD)

[2]Sea AI Lab (SAIL)

# Brief: Few-shot Image Generation (FSIG)

# Brief: Few-shot Image Generation (FSIG)



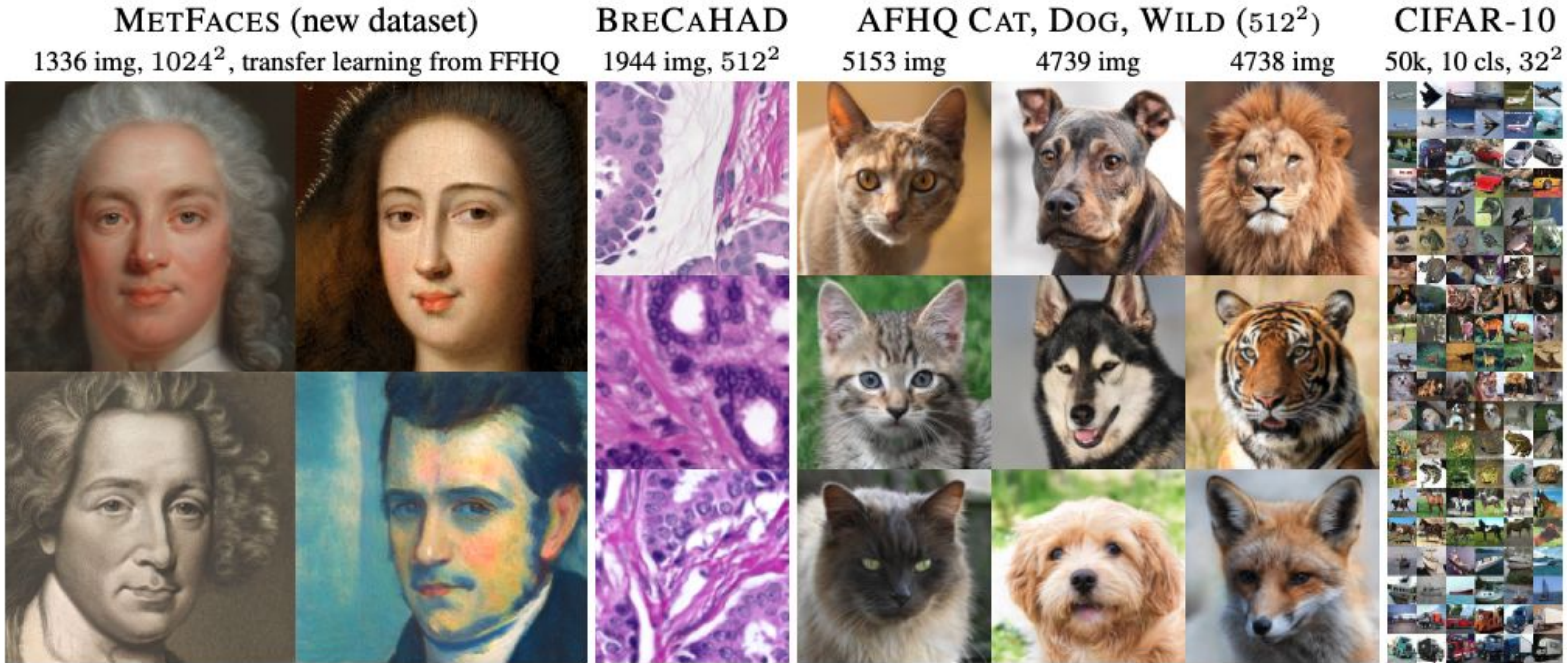
# Background: Training GANs requires large datasets


**FFHQ dataset (credit)**

**ImageNet Dataset (credit)**

1. Training GANs requires abundant training data (e.g., **FFHQ-70k** (left), **ImageNet** (right))
2. In real world, collecting images could be expensive and difficult (e.g., rare birds species)

# Background: Training GANs with limited Data

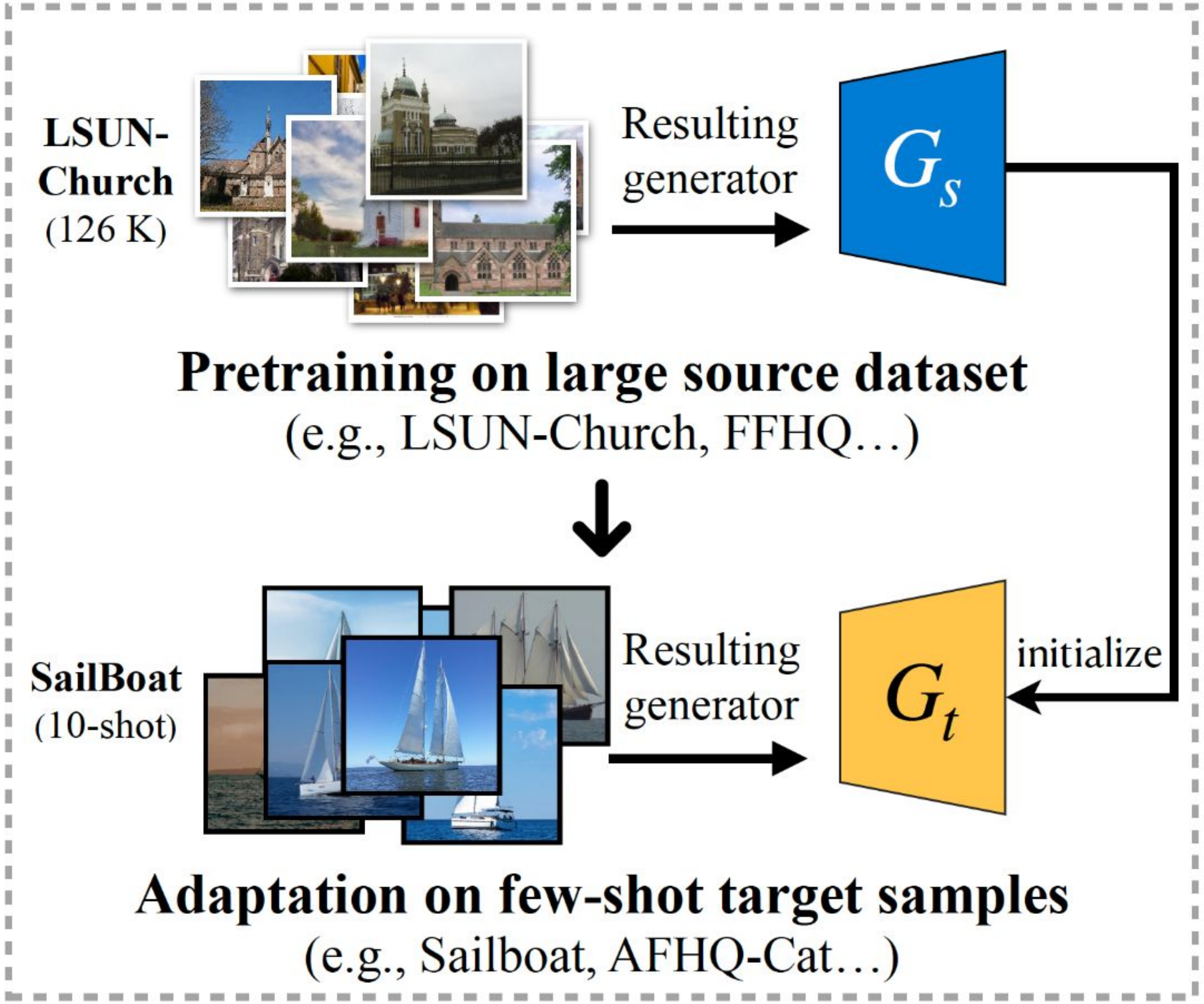



**Small datasets used in Adaptive Data Augmentation (Karras et al., credit)**

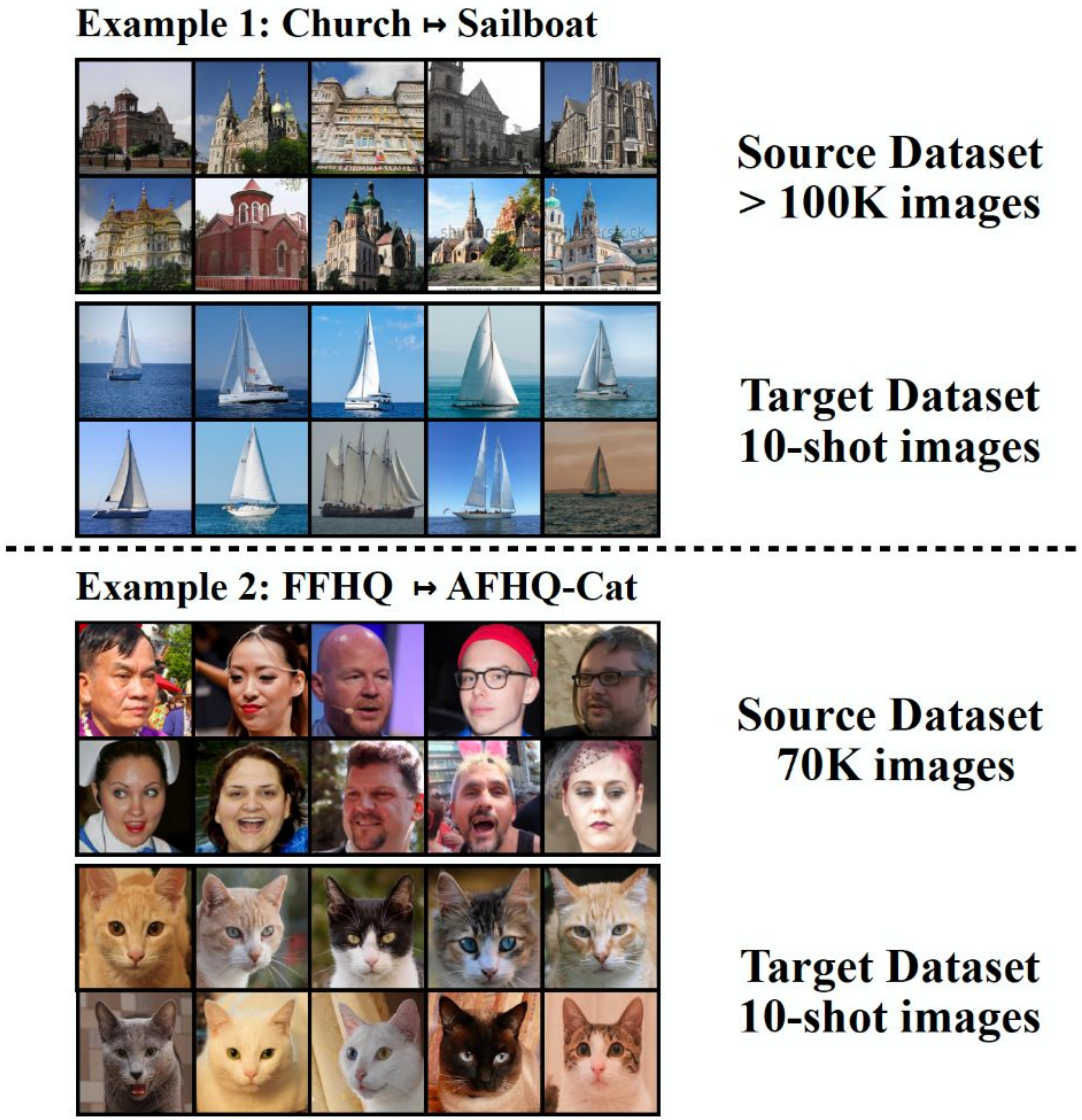
1. Different works propose to training GAN with limited data (e.g., **1K ~ 5K images** )
2. **Approach:**
   a. Training from scratch (e.g., Karras et al., 2020)
   b. Transfer learning (e.g., Wang et al., 2018)

Wang, Yaxing, et al. "Transferring gans: generating images from limited data." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
Karras, Tero, et al. "Training generative adversarial networks with limited data." Advances in Neural Information Processing Systems 33 (2020): 12104-12114.

# Preliminary: Few-shot Image Generation (FSIG)




Transfer Learning for FSIG



# Related Works of FSIG

## State-of-the-art methods:

- **EWC (Li et al., 2020):**
  - *Preserve* knowledge important for source domain;
- **CDC (Ojha et al., 2021):**
  - *Preserve* the distance between generated images, before and after adaptation;
- **DCL (Zhao et al., 2022):**
  - *Preserve* multi-level source knowledge when adapting the source models to the target domain;
- **AdAM (Zhao et al., 2022):**
  - *Preserve* source knowledge that is deemed important to the target domain;

## Observation:

- Existing SOTA methods focus on knowledge preservation

Li, Yijun, et al. "Few-shot Image Generation with Elastic Weight Consolidation." *Advances in Neural Information Processing Systems* 33 (2020): 15885-15896.
Ojha, Utkarsh, et al. "Few-shot image generation via cross-domain correspondence." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
Zhao, Yunqing, et al. "A Closer Look at Few-shot Image Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

# Related Works of FSIG

## Questions

- Is the knowledge **preserved from source proper** for target domain?
- Note that, the target domain could be **semantically distant** to the source domain.

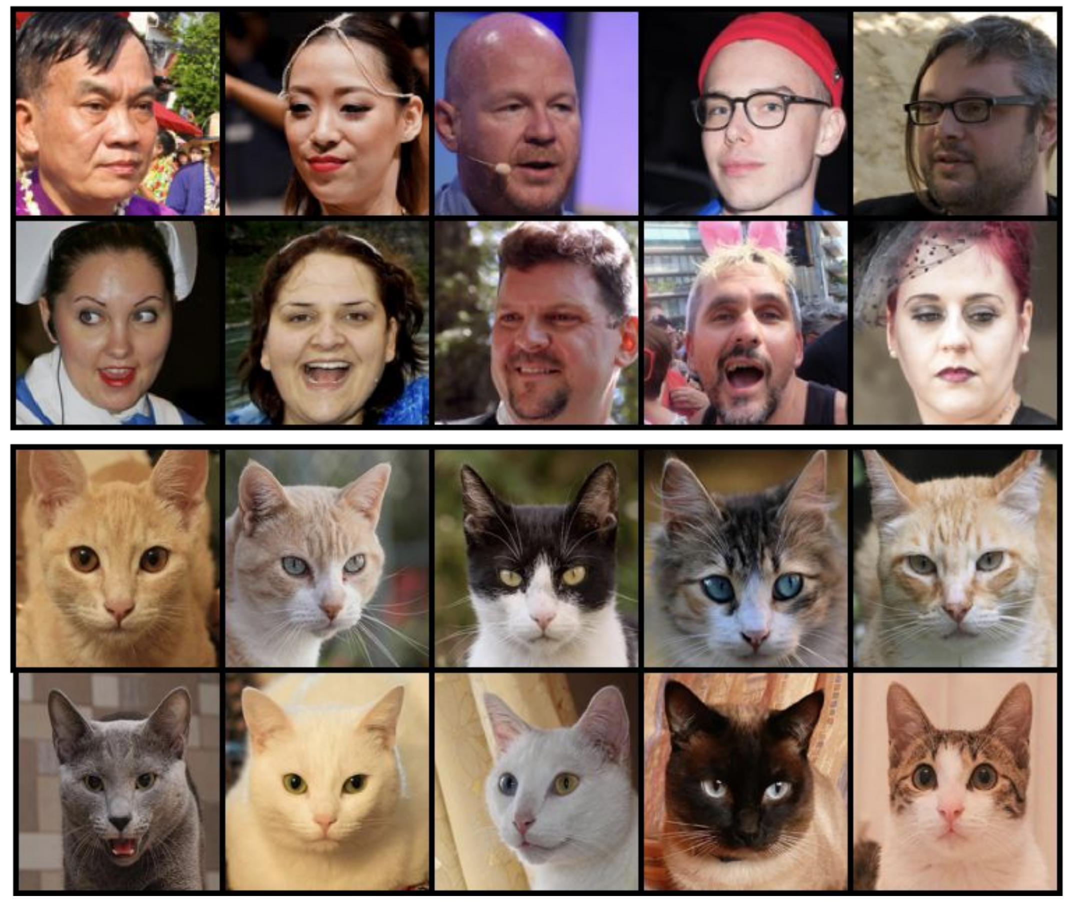
## Examples of adaptation:

**Example 1: Church ↦ Sailboat**

**Source Dataset > 100K images**

**Target Dataset 10-shot images**

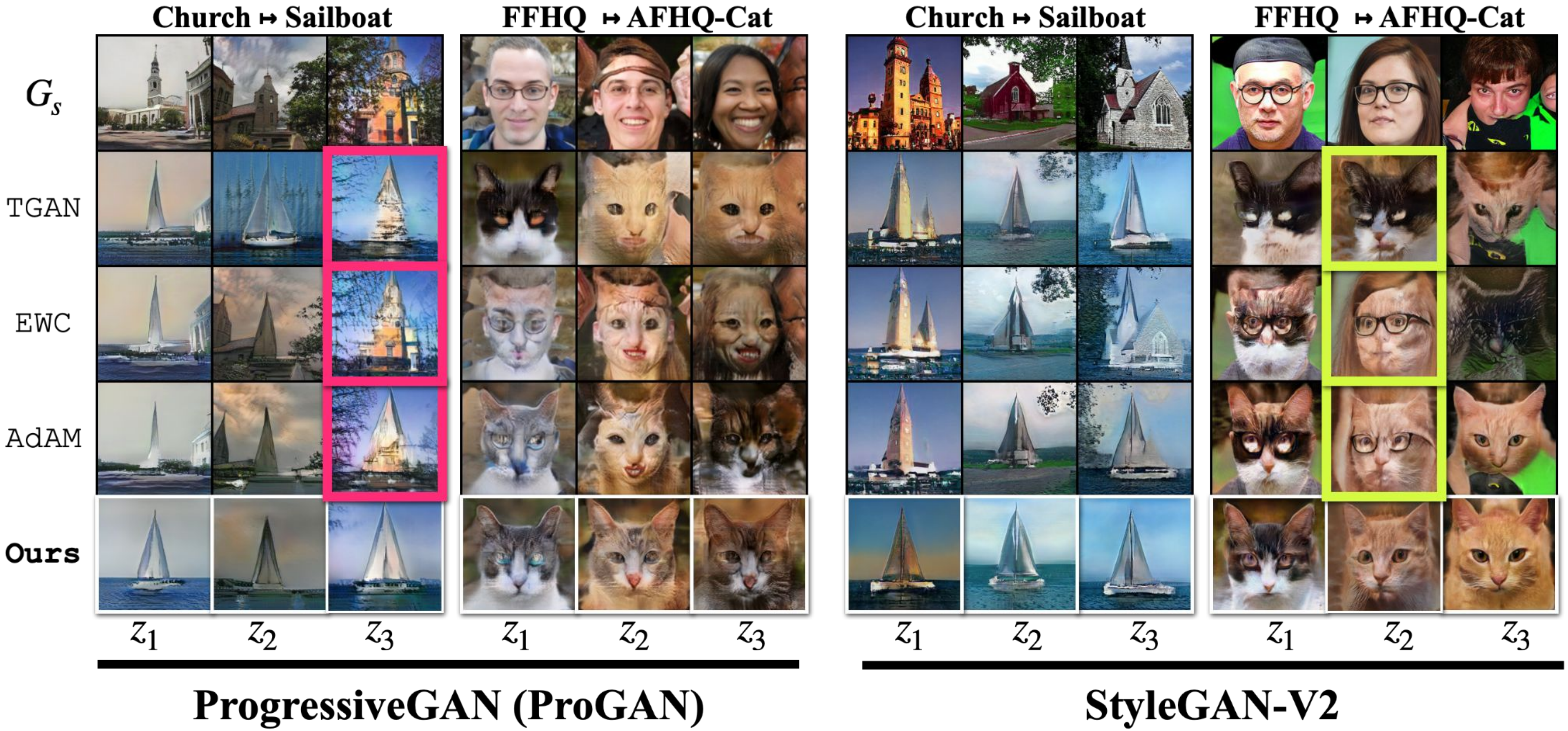
**Example 2: FFHQ ↦ AFHQ-Cat**



**Source Dataset 70K images**

**Target Dataset 10-shot images**

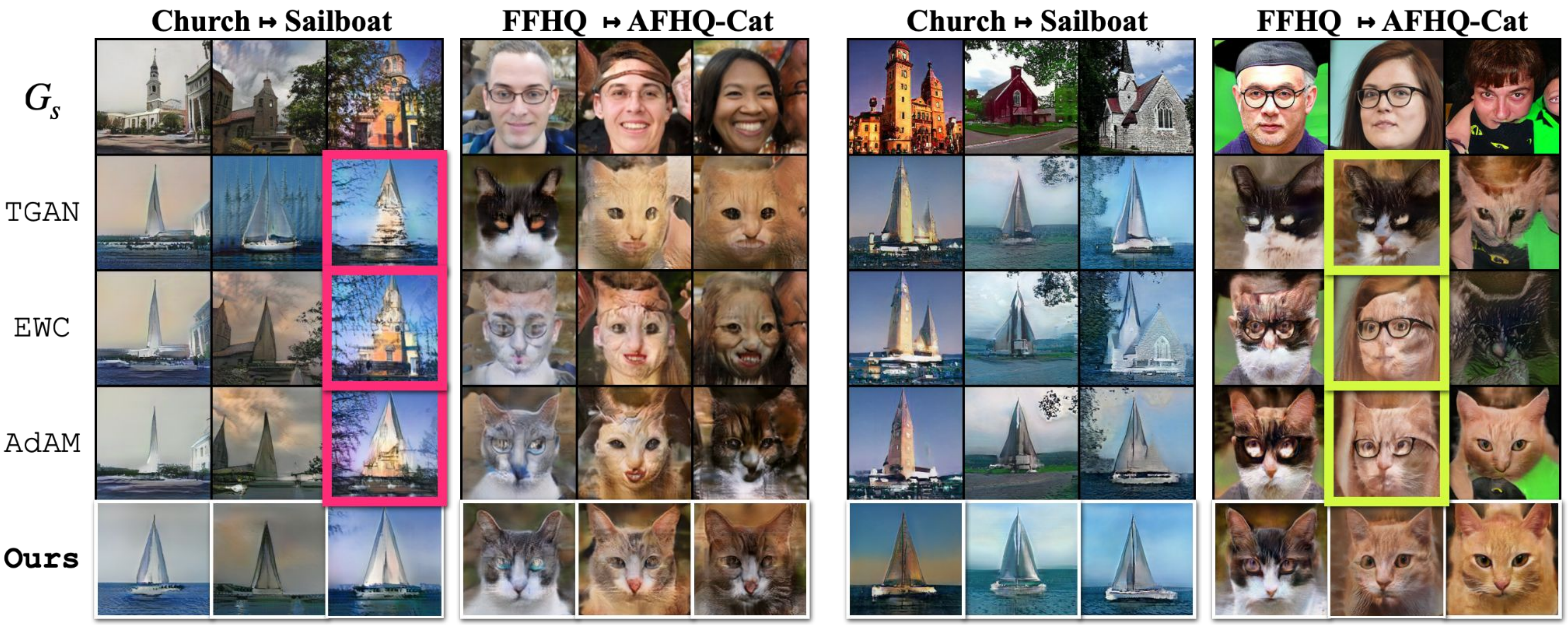
# Discover: Incompatible Knowledge Transfer for FSIG




**We design experiments and compare with SOTA methods:**

1. G_s is the source generator, target domains are semantically distant
2. Each column represents a fixed noise input (z)
3. FSIG with different types of GAN architectures

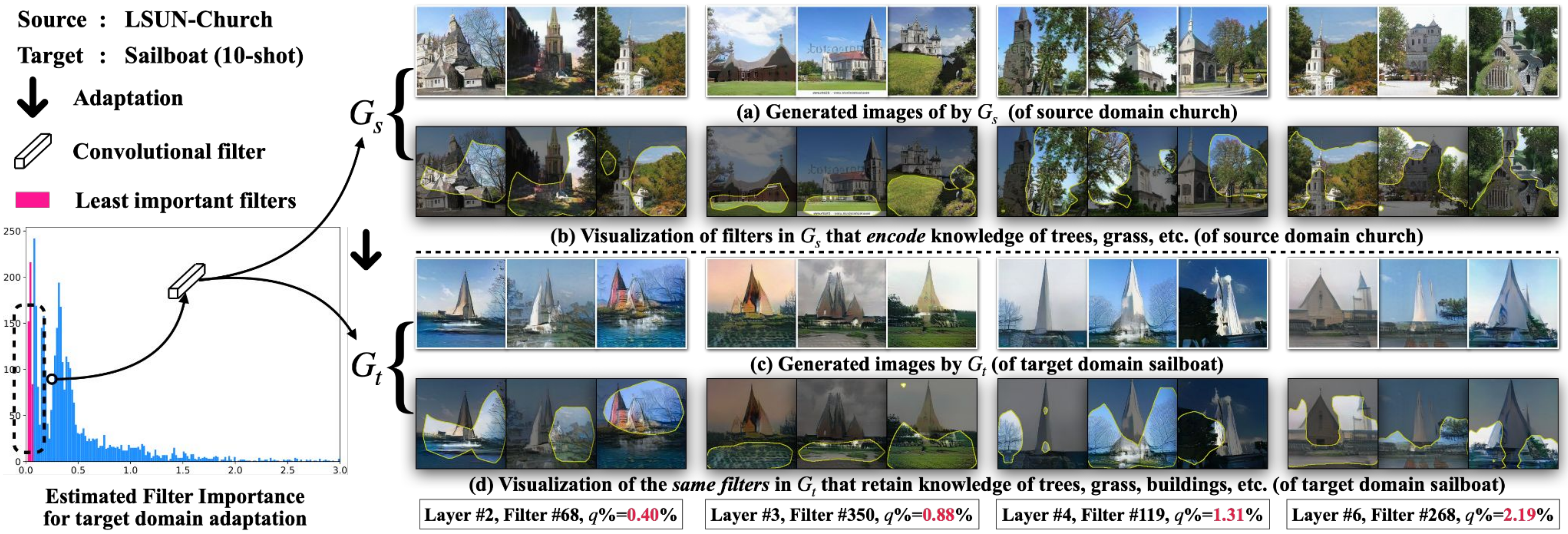
# **Discover:** Incompatible Knowledge Transfer for FSIG




**Empirical study:**

1. **Problem setup**: Prior knowledge of $G_s$ is selected, and transferred to learn the target generator $G_t$
2. **Dataset:** the target domain could be **semantically distant** from the source domain, e.g., Church ↦ Sailboat
3. **Observation:** existing methods focus on knowledge preservation, while failing to prevent the transfer of knowledge that is incompatible with the target domain.
4. This gives rise to **unrealistic samples** generated by $G_t$: "*Trees/Buildings on the Sea*" (**Red** frames), or "*Cats with Glasses*" (**Green** frames)

# Investigation: the Cause of Incompatible Knowledge Transfer



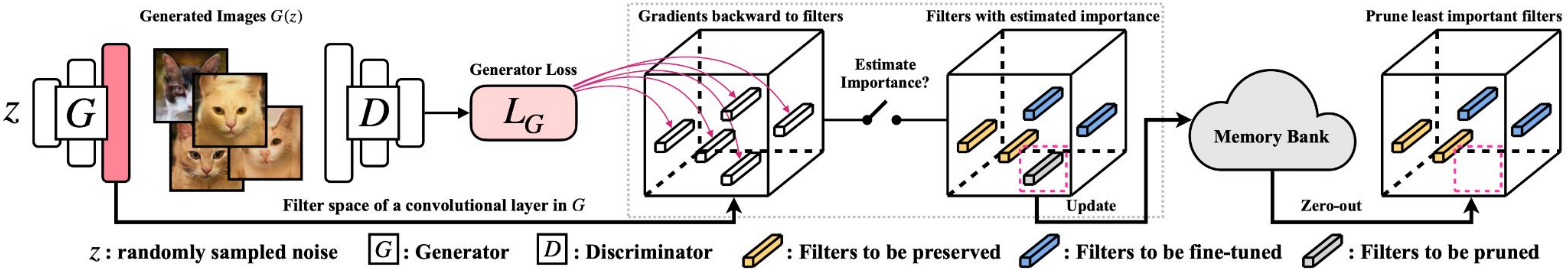

## Our investigation:

1. **We know that:** fine-tuning is inadequate for preventing incompatible knowledge transfer (shown in observation)
2. **GAN Dissection:** identify interpretable filters that are highly correlated to the region of an object class (e.g., trees)
3. **Finding-1 - Before Adaptation:** Incompatible knowledge (e.g., trees, grass and buildings to Sailboat) is correlated to $G_s$ filters that are deemed to be unimportant / irrelevant for target domain
4. **Finding-2 - After Adaptation:** Similar knowledge (tree, building, grass) remains in the **same filters** in $G_t$ after fine-tuning
5. **Conclusion:** Filters with least importance for target domain degrades the realisctiness of generated images' quality.

**rules to determine importance/unimportance of filters, other implementation details, are in our paper**

# Method: Remove InCompatible Knowledge (RICK) for FSIG

## Pipeline of RICK:




**Overview of RICK:**

1. **Prior works: In CNN,** each filter can be viewed as an encoding of a specific part of knowledge.
2. **We propose** *knowledge truncation*, a novel concept for FSIG, to remove the knowledge that is incompatible to the target.
3. **Training:** We estimate the filter importance every certain iterations, via Fisher Information (**FI**)
   - **A.** Then, we use a fixed threshold to determine the importance/unimportance for each filter.
   - **B.** Least important filters: **Zero-out**; Most important filters: **Freeze**; Others: **Fine-tuning**.
   - **C.** The FI of each filter is dynamically updated in a lightweight memory bank, size ~= 5000 (chars) for G.
   - **D.** Validate the method on StyleGAN-V2, and ProGAN.

# Method: Remove InCompatible Knowledge (RICK) for FSIG

## Importance measurement:

To measure the importance, we directly compute the **Fisher Information (FI)** to each filter (additional metrics are explored in our paper)

$$\mathrm{FI}(\Theta) = \mathbb{E}\Big[-\frac{\partial^2}{\partial \boldsymbol{\Theta}^2}\mathscr{L}(x\,|\,\Theta)\Big]$$

$$\begin{cases} \boldsymbol{\Theta}: \textbf{parameters of G/D} \\ \mathscr{L}: \textbf{loss of G/D} \end{cases}$$

In practice, use the first-order approximation to estimate FI.

# Method: Remove InCompatible Knowledge (RICK) for FSIG

**Algorithm 1** Proposed Knowledge Truncation for Few-shot Image Generation

**Input:** target samples $\mathbf{X}$, number of total iterations $\mathbf{N}$, number of warmup iterations $\mathbf{N}_w$, interval for filter importance estimation $\mathbf{T}$, Conv filters of the model $\mathbf{W}$, FI of the filter $\mathcal{F}(\mathbf{W})$, quantile of filter importance $q(\mathbf{W})$, quantile threshold for filter that is selected for pruning $t_l$, threshold for filter that is selected for preserving $t_h$, total pruning rate $p\%$.

**Initialize with pretrained GAN**: $G_t \leftarrow G_s$ and $D_t \leftarrow D_s$.

**for** $iteration = 1; iteration \leqslant \mathbf{N}_w; iteration++$ **do**

  Update the discriminator $D_t$ given $\mathbf{X}$, via Eqn. 1 # warmup for $D_t$

**end for**

**for** $iteration = \mathbf{N}_w + 1; iteration \leqslant \mathbf{N}; iteration++$ **do**

  **If** $iteration \% \mathbf{T} = 0$ :

    Estimate $\mathcal{F}(\mathbf{W})$ via Eqn. 2;

    Preserve $\mathbf{W}$ if $t_h \leqslant q(\mathbf{W})$;

    Fine-tune $\mathbf{W}$ if $t_l \leqslant q(\mathbf{W}) \leqslant t_h$;

    Prune $\mathbf{W}$ if $q(\mathbf{W}) \leqslant t_l$;

    Update the operation of $\mathbf{W}$ in Memory Bank. # size of the Memory Bank $\sim$ 5,000

  **ELSE**

    Extract the operation of $\mathbf{W}$ in Memory Bank;

    Update model via Eqn. 1.

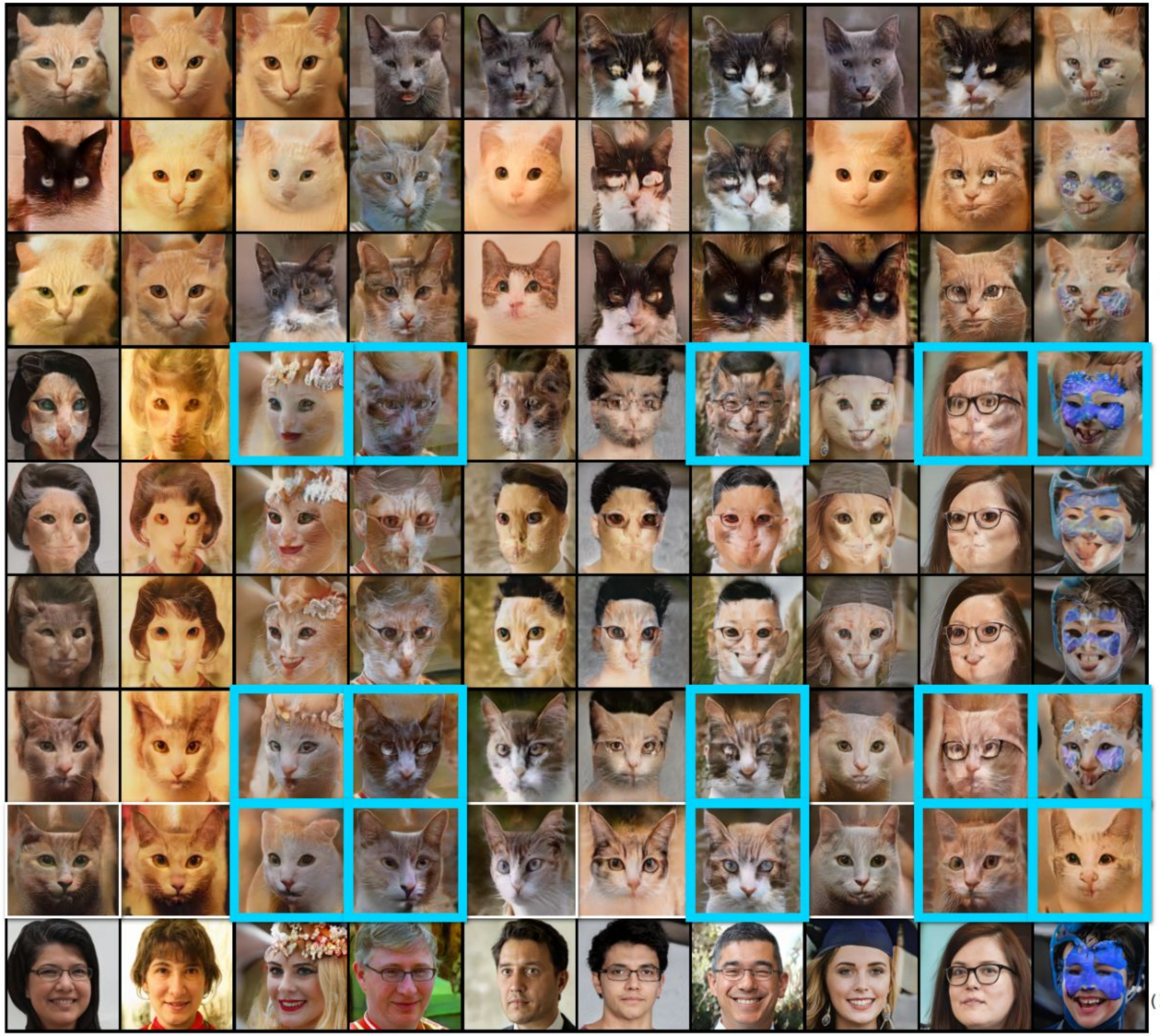
**end for**

**Output:** The adapted GAN

# Experiment Results

**Each col: Fixed noise**

10-Shot Real Cats

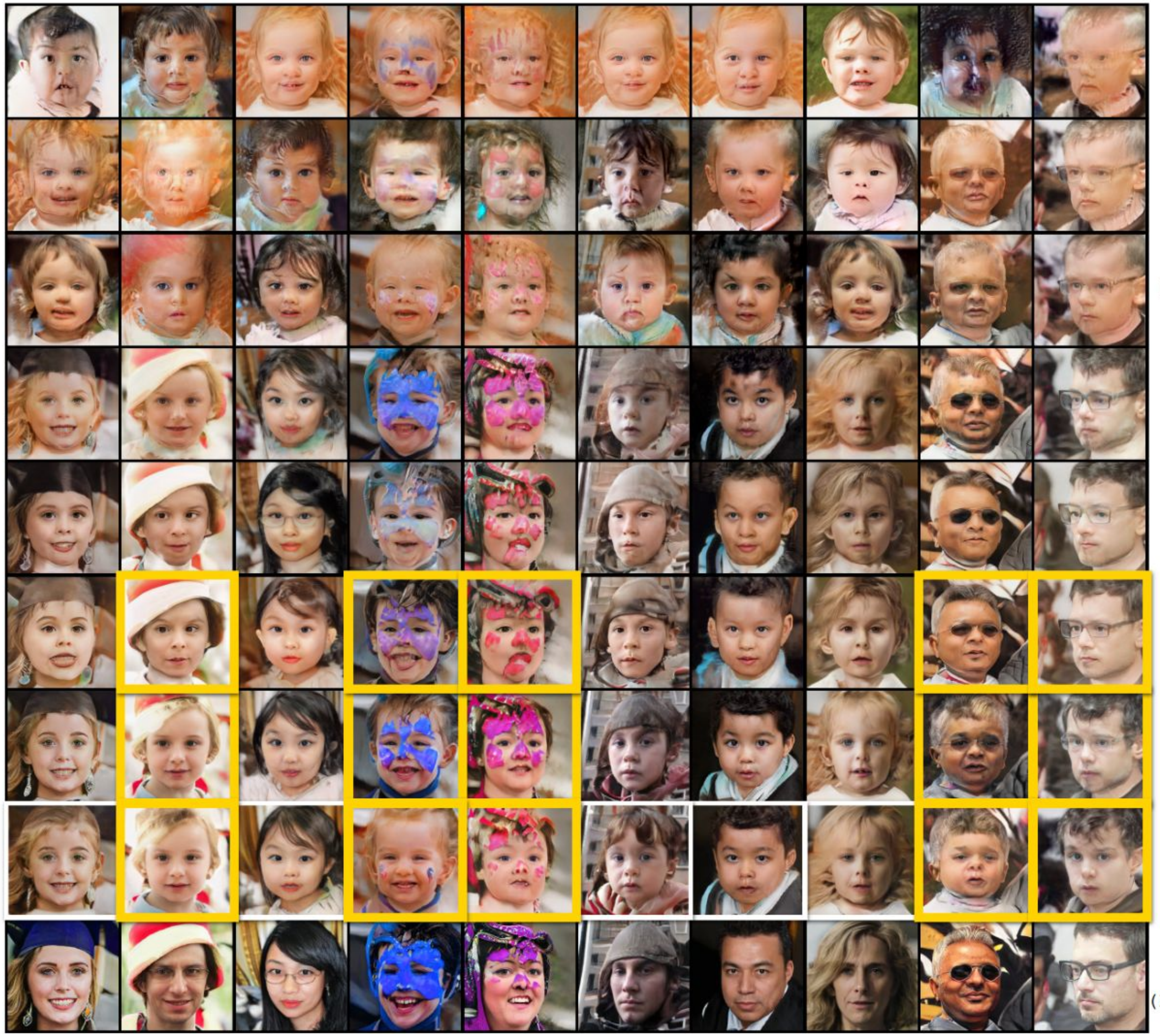


| | FID (↓) | Intra-LPIPS (↑) |
|---|---|---|
| TGAN | 64.68 | 0.490 |
| TGAN + ADA | 80.16 | 0.513 |
| FreezeD | 63.60 | 0.492 |
| EWC | 74.61 | 0.587 |
| CDC | 176.21 | 0.629 |
| DCL | 156.82 | 0.616 |
| AdAM | 58.07 | 0.557 |
| **Ours** | **53.27** | 0.569 |
| $G_s$ (Pretrained) | | |

# Experiment Results

**Each col:**
**Fixed noise**

10-Shot
Real Babies



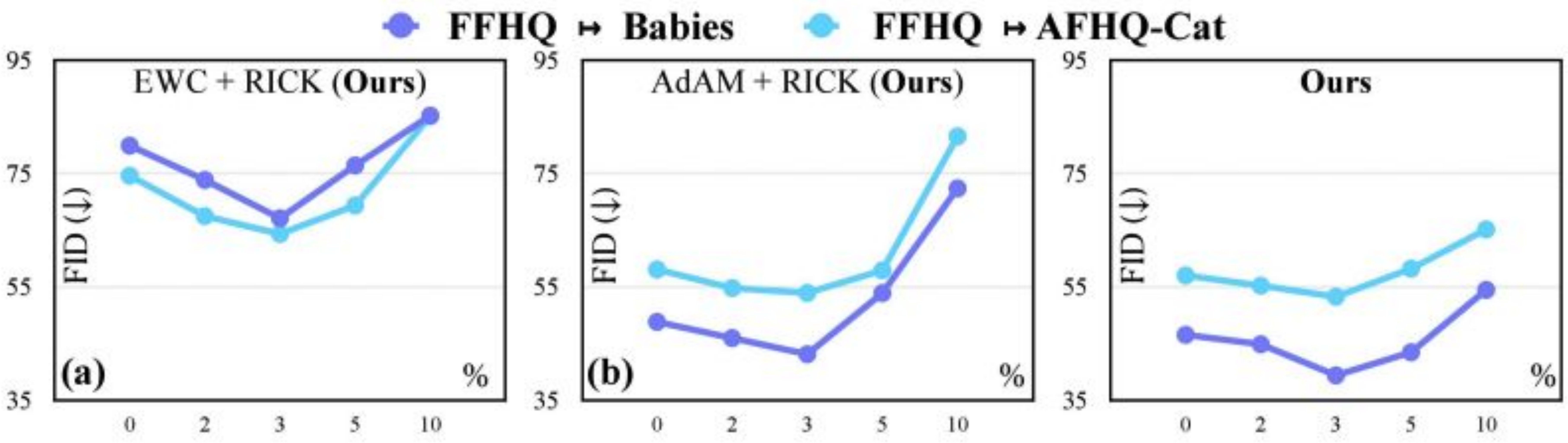
| | FID (↓) | Intra-LPIPS (↑) |
|---|---|---|
| TGAN | 101.58 | 0.517 |
| TGAN + ADA | 97.91 | 0.511 |
| FreezeD | 96.25 | 0.518 |
| EWC | 79.93 | 0.521 |
| CDC | 69.13 | 0.578 |
| DCL | 56.48 | 0.580 |
| AdAM | 48.83 | 0.590 |
| **Ours** | **39.39** | **0.608** |
| $G_s$ (Pretrained) | | |

# Quantitative Experiment Results

Table 1. We report FID (↓) as quantitative results for FSIG (10-shot), FFHQ is the source domain. We compare our proposed with other baseline and SOTA methods over six target datasets, including the challenging setups that target domains are distant to the source (*e.g.* AFHQ datasets). We emphasize that, for SOTA methods that focus only on knowledge preservation (*e.g.*, EWC [41], CDC [50], DCL [77], AdAM [75]), incompatible source knowledge is still transferred and therefore it curtails the quality of generated images. In contrast, our methods can remove the knowledge incompatible for the target and preserve the knowledge important for the target, therefore achieve improved quality of generated images.

| **Target Domain** | **Babies** [50] | **Sunglasses** [50] | **MetFaces** [26] | **AFHQ-Cat** [8] | **AFHQ-Dog** [8] | **AFHQ-Wild** [8] |
|---|---|---|---|---|---|---|
| TGAN [65] | 101.58 | 55.97 | 76.81 | 64.68 | 151.46 | 81.30 |
| TGAN+ADA [26] | 97.91 | 53.64 | 75.82 | 80.16 | 162.63 | 81.55 |
| FreezeD [48] | 96.25 | 46.95 | 73.33 | 63.60 | 157.98 | 77.18 |
| CDC [50] | 69.13 | 41.45 | 65.45 | 176.21 | 170.95 | 135.13 |
| DCL [77] | 56.48 | 37.66 | 62.35 | 156.82 | 171.42 | 115.93 |
| EWC [41] | 79.93 | 49.41 | 62.67 | 74.61 | 158.78 | 92.83 |
| EWC + RICK (**Ours**) | 68.22(**−11.71**) | 39.53(**−9.88**) | 54.7(**−7.97**) | 64.35(**−10.26**) | 124.50(**−34.28**) | 56.83(**−36.00**) |
| AdAM [75] | 48.83 | 28.03 | 51.34 | 58.07 | 100.91 | 36.87 |
| AdAM + RICK (**Ours**) | 43.12(**−5.71**) | 26.25(**−1.78**) | 49.47(**−1.87**) | 53.94(**−4.13**) | 100.35(**−0.56**) | 35.54(**−1.33**) |
| **Ours** | **39.39** | **25.22** | **48.53** | **53.27** | **98.71** | **33.02** |

# Quantitative Experiment Results




**Ablation Study:**

1. Apply RICK to different SOTA methods;
2. Prune different amount of filters.

# Conclusion

## Our analysis & main findings

- **F**ine-tuning based SOTA methods lead to incompatible knowledge transfer;
- **W**ith GAN Dissection, we find that those incompatible knowledge is highly correlated to filters with least importance;
- **T**his lead to generated images are with poor quality (e.g., “trees on sea, etc.”)

## Our contributions

- **W**e propose **RICK**:
    - A knowledge truncation method, complementary to SOTA methods
    - Dynamically measure the importance of each filter during adaptation
    - Preserve, transfer and Remove source GAN knowledge to the target domain
- **W**e achieve new SOTA performance for FSIG with different source/target setups.

# Exploring Incompatible Knowledge Transfer in Few-shot Image Generation

Yunqing Zhao[1] Chao Du[2] Milad Abdollahzadeh[1] Tianyu Pang[2]
Min Lin[2] Shuicheng Yan[2] Ngai-Man Cheung[1]

[1]**Singapore University of Technology and Design (SUTD)**

[2]**Sea AI Lab (SAIL)**

[ *We use materials compiled from various sources such as textbooks, lecture materials, conference talks, web resources and are shared for research purposes only. In the interest of brevity, every source is not cited. The compiler of these materials gratefully acknowledges all such sources. Please contact yunqing_zhao@mymail.sutd.edu.sg for any queries.* ]