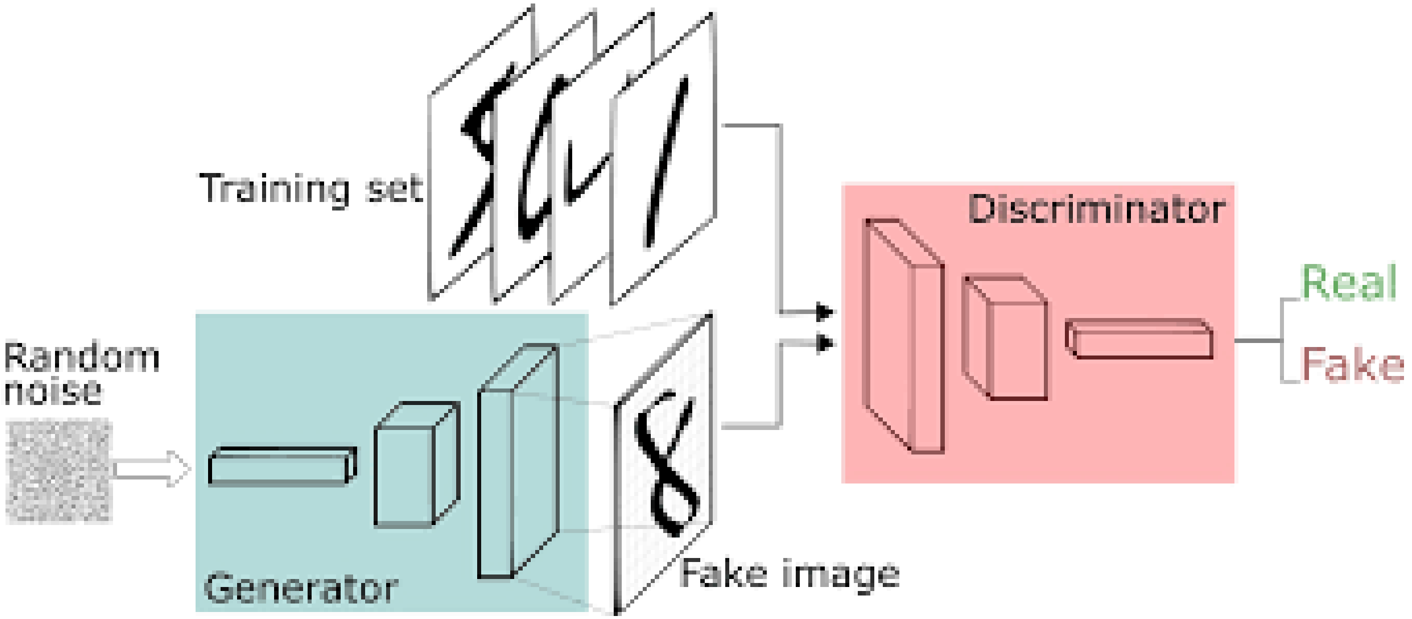
Transferring GANs: generating images from limited data

MSCV2, University of Bourgogne

Student: Yunke LI, Chen ZHU, Qi YANG

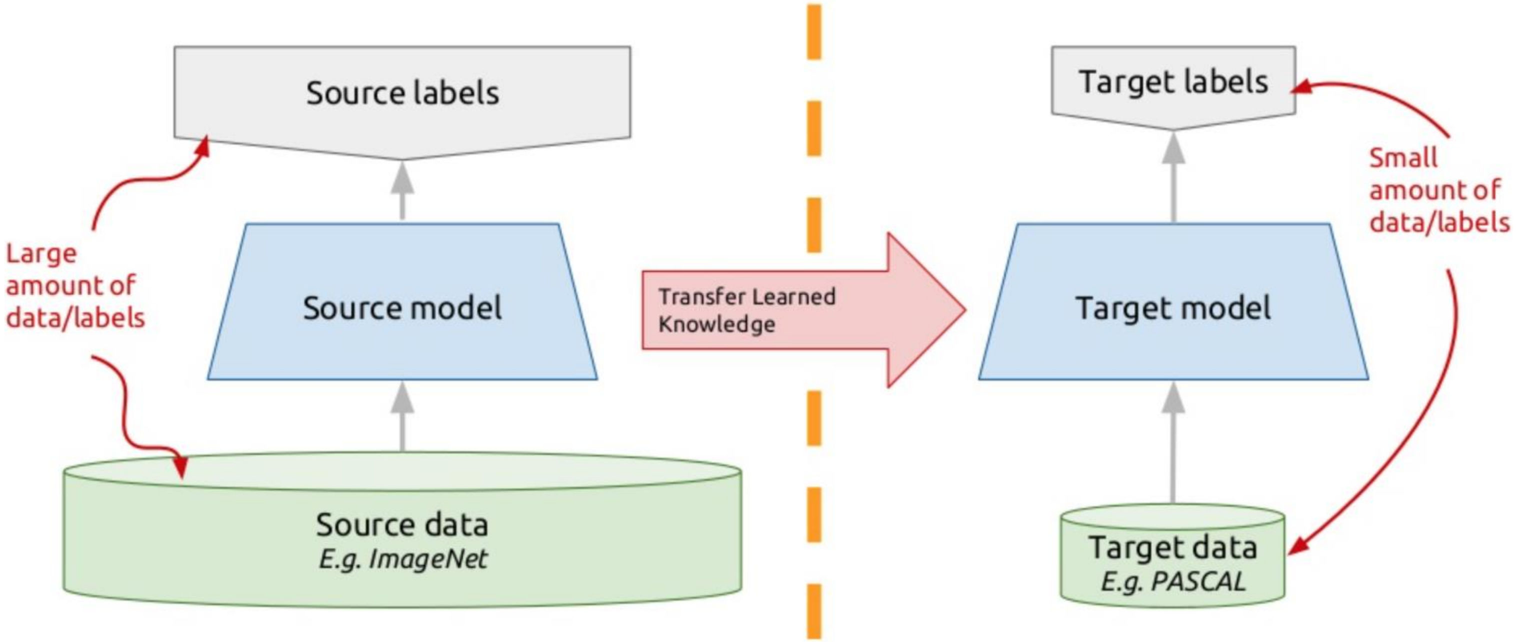
Supervisor: Abdul QAYYUM

1 Introduction



1 Introduction

Transfer learning: idea



Contribution

- Evaluation of several transfer configurations
- How the relation between source and target domains impacts the results
- Evaluation of the transfer from unconditional GANs to conditional GANs on two commonly used methods

4 Transferring GAN representations

- Evaluation Metrics

1.Frechet Inception Distance

$$\mathrm{FID}(\mathbb{P}_r, \mathbb{P}_g) = \|\mu_r - \mu_g\| + \mathrm{Tr}(\mathbf{C}_r + \mathbf{C}_g - 2(\mathbf{C}_r\mathbf{C}_g)^{1/2}),$$

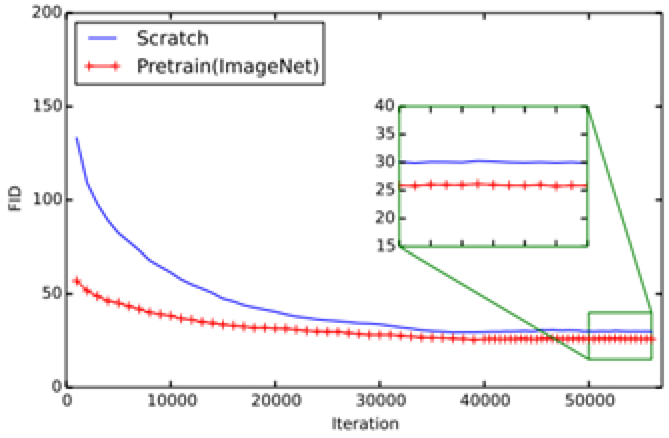
2. Independent Wasserstein (IW) critic

$$\mathrm{IW}(\mathcal{X}_1, \mathcal{X}_2) = \mathbb{E}_{x\sim\mathcal{X}_1}\left(\hat{D}(x)\right) - \mathbb{E}_{x\sim\mathcal{X}_2}\left(\hat{D}(x)\right)$$

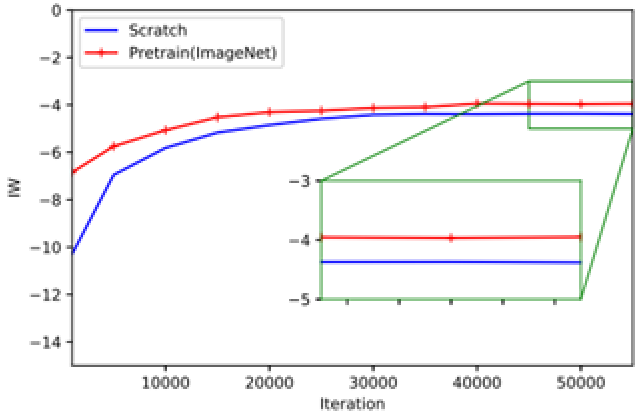
4 Transferring GAN representations

- Generator/discriminator transfer configuration

Generator	Scratch		Pre-trained	
Discriminator	Scratch	Pre-trained	Scratch	Pre-trained
FID $(\mathcal{X}_{data}^{tgt}, \mathcal{X}_{gen}^{tgt})$	32.87	30.57	56.16	**24.35**
IW $(\mathcal{X}_{val}^{tgt}, \mathcal{X}_{gen}^{tgt})$	-4.27	-4.02	-6.35	**-3.88**



(a) Unconditional GAN(FID)



(b) Unconditional GAN(IW)

4 Transferring GAN representations

- Size of the target dataset

Table 2: FID/IW for different sizes of the target set (LSUN Bedrooms) using ImageNet as source dataset.

Target samples	1K	5K	10K	50K	100K	500K	1M
From scratch	256.1/-33.3	86.0/-18.5	73.7/-15.3	45.5/-7.4	32.9/-4.3	24.9/-3.6	21.0/-2.9
Pre-trained	93.4/-22.5	74.3/-16.3	47.0/-7.0	29.6/-4.56	24.4/-4.0	21.6/-3.2	18.5/-2.8

4 Transferring GAN representations

- Source and target domains

Table 3: Datasets used in the experiments.

Source datasets	ImageNet [38]	Places [48]	Bedrooms [45]	CelebA [27]
Number of images	1M	2.4M	3M	200K
Number of classes	1000	205	1	1
Target datasets	Flower [31]	Kitchens [45]	LFW [19]	Cityscapes [6]
Number of images	8K	50K	13K	3.5K
Number of classes	102	1	1	1

Table 4: Distance between target real data and target generated data FID/IW $(\mathcal{X}^{tgt}_{data}, \mathcal{X}^{tgt}_{gen})$.

Source → Target ↓	Scratch	ImageNet	Places	Bedrooms	CelebA
Flowers	71.98/-13.62	**54.04/-3.09**	66.25/-5.97	56.12/-5.90	67.96/-12.64
Kitchens	42.43/-7.79	34.35/-4.45	34.59/**-2.92**	**28.54**/-3.06	38.41/-4.98
LFW	19.36/-8.62	9.65/-5.17	15.02/-6.61	7.45/-3.61	**7.16/-3.45**
Cityscapes	155.68/-9.32	**122.46**/-9.00	151.34/-8.94	123.21/-8.44	130.64/**-6.40**

Source datasets:
ImageNet, Places,
LSUN Bedrooms , CelebA
Target:
Oxford Flowers, LSUN
Kitchens , LFW CityScapes

4 Transferring GAN representations

- Selecting the pre-trained model

Table 4: Distance between target real data and target generated data FID/IW $(\mathcal{X}^{tgt}_{data}, \mathcal{X}^{tgt}_{gen})$.

Source → Target ↓	Scratch	ImageNet	Places	Bedrooms	CelebA
Flowers	71.98/-13.62	**54.04/-3.09**	66.25/-5.97	56.12/-5.90	67.96/-12.64
Kitchens	42.43/-7.79	34.35/-4.45	34.59/**-2.92**	**28.54**/-3.06	38.41/-4.98
LFW	19.36/-8.62	9.65/-5.17	15.02/-6.61	7.45/-3.61	**7.16/-3.45**
Cityscapes	155.68/-9.32	**122.46**/-9.00	151.34/-8.94	123.21/-8.44	130.64/**-6.40**

Table 5: Distance between source generated data $\mathcal{X}^{src}_{gen}$ and target real data $\mathcal{X}^{tgt}_{data}$, and distance between source real $\mathcal{X}^{src}_{data}$ and generated data $\mathcal{X}^{src}_{gen}$.

	Source → Target ↓	ImageNet	Places	Bedrooms	CelebA
FID $(\mathcal{X}^{src}_{gen}, \mathcal{X}^{tgt}_{data})$	Flowers	**237.04**	251.93	278.80	284.74
	Kitchens	183.27	180.63	**70.06**	254.12
	LFW	333.54	333.38	329.92	**151.46**
	Cityscapes	233.45	**181.72**	227.53	292.66
FID $(\mathcal{X}^{src}_{gen}, \mathcal{X}^{src}_{data})$	Source	63.46	55.66	17.30	75.84

5 Transferring to conditional GANs

- Conditional GAN adaptation
- Auxiliary Classifier GAN (AC-GAN)

generator， discriminator

Cond Concat

- Input： connects condition and input noise
- Randomly initialize： the weights of the layer connected to 'cond concat'

Cond BNorm

- the conditioning prior is embedded in the batch normalization layers of the generator
- copying the values from the unconditional GAN to all classes

5 Transferring to conditional GANs

- Results

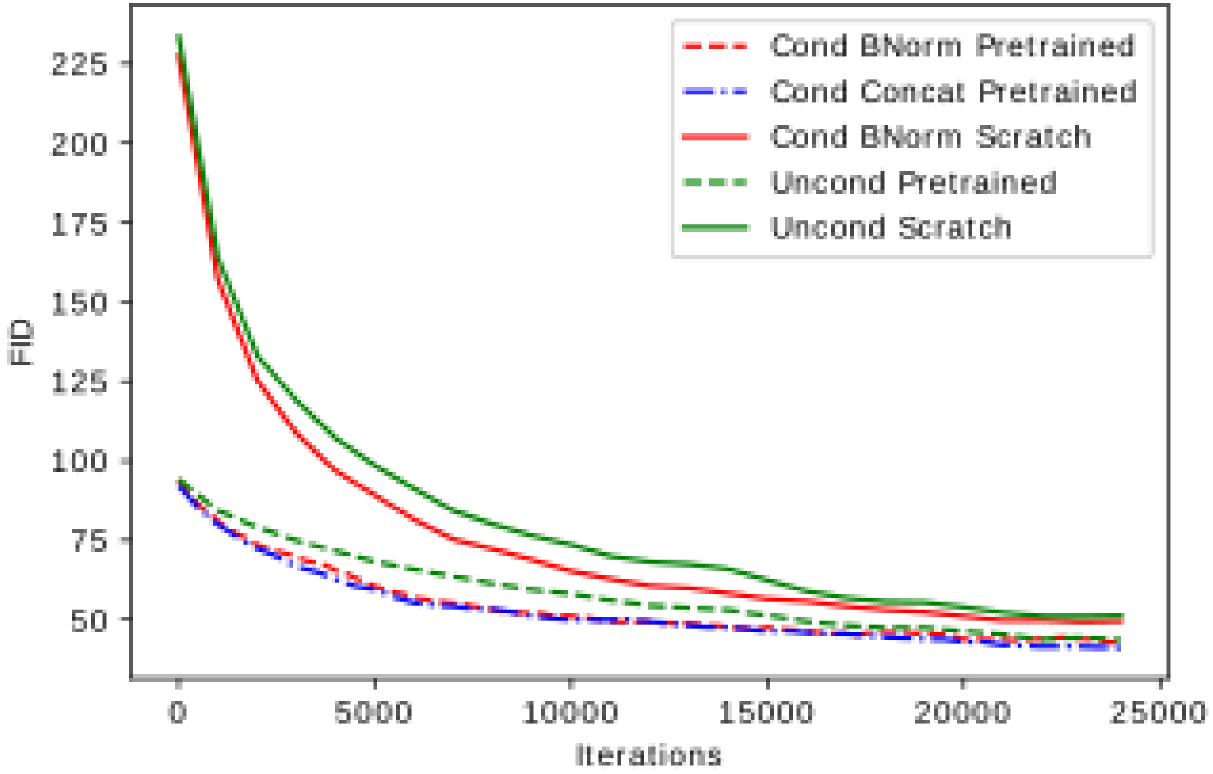
Table 6: Per-class and overall FID for AC-GAN. Source: Places, target: LSUN

Init	Iter	Bedr	Bridge	Church	Classr	Confer	Dining	Kitchen	Living	Rest	Tower	Avg.	All
Scratch	250	298.4	310.3	314.4	376.6	339.1	294.9	314.2	316.5	324.4	301.0	319.0	352.4
	2500	195.9	135.0	133.0	218.6	185.3	173.9	167.9	189.3	159.5	125.6	168.4	137.3
	25000	72.9	78.0	52.4	106.7	76.9	40.1	53.9	56.1	74.7	59.8	67.2	49.6
Pre-trained	250	**168.3**	**122.1**	**148.1**	**145.0**	**151.6**	**144.2**	**156.9**	**150.1**	**113.3**	**129.7**	**142.9**	**107.2**
	2500	**140.8**	**96.8**	**77.4**	**136.0**	**136.8**	**84.6**	**85.5**	**94.9**	**77.0**	**69.4**	**99.9**	**74.8**
	25000	**59.9**	**68.6**	**48.2**	**79.0**	**68.7**	**35.2**	**48.2**	**47.9**	**44.4**	**49.9**	**55.0**	**42.7**

- 10K images per class
- for 25K iterations

5 Transferring to conditional GANs

- Results



6 Conclusions

- GAN and cGAN use **pre-trained models** to achieve good results in a **shorter time**(shorter iterations) or with **fewer data**.
- In the selection of pre-training models, **density** is more important than diversity.
- The effect of using transfer learning is the best in **both G and D**. Only D: not obvious, only G: worse.

Thank you!