Group Equivariant GANs

Dey et al. 2021

Presentation by Sacha Morin

Outline

1. Cohen, Taco, and Max Welling. **"Group equivariant convolutional networks."** International conference on machine learning. PMLR, 2016.
2. Dey, Neel, Antong Chen, and Soheil Ghafurian. "**Group equivariant generative adversarial networks.**" arXiv preprint arXiv:2005.01683 (2020).

Cohen, Taco, and Max Welling. "**Group equivariant convolutional networks**." International conference on machine learning. PMLR, 2016.

Group Equivariant CNNs : Motivation

- CNNs work great thanks to **translation equivariance.**
- CNNs are **not equivariant** to other simple image transforms, such as reflections and rotations.

Group Equivariant CNNs : Equivariance

Commute function (layer, model) with group action.

$$\Phi(T_g\, x) = T'_g\, \Phi(x)$$

Group Equivariant CNNs : G-CNN

Convolution/Correlation

$$[f \star \psi^i](x) = \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K^l} f_k(y) \psi_k^i(y - x)$$

Group Convolution :

$$[f \star \psi](g) = \sum_{y \in \mathbb{Z}^2} \sum_{k} f_k(y) \psi_k(g^{-1} y)$$

Feature maps are functions of group elements.

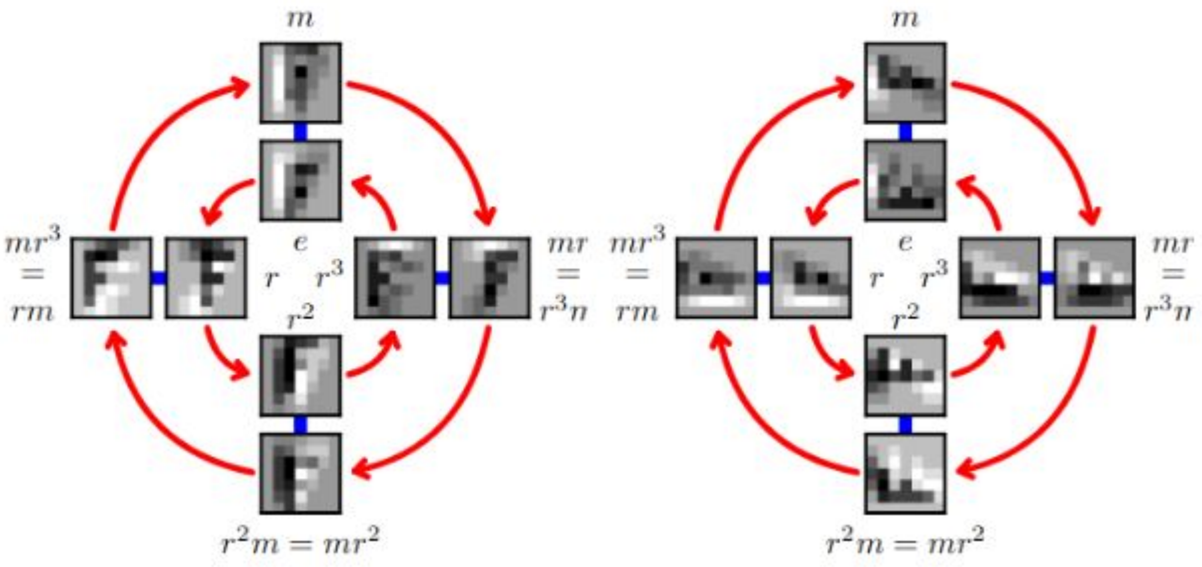
Group Equivariant CNNs : Groups

p4 :

- Translations
- 90° Rotations
  - 4 rotations

p4m :

- Translations
- 90° Rotations
- Reflections
  - 8 “roto-flips”



Group Equivariant CNN: G-CNN Feature map

- **Input :** 28 x 28 image
- **Augmented Filter Bank :** Rotate/Flip K learnable filters
- Run standard CNN planar routine with augmented filter bank
- **Output:** K x 8 x 28 x 28 feature map
    - K filters
    - 8 roto-flips
    - 28 x 28 translations

Group Equivariant CNNs : Findings

- Competitive results on CIFAR-10 (at the time) with substantially fewer parameters.
- CIFAR-10 results are interesting because objects are typically upright
  - Unclear why it would benefit from rotations

Dey, Neel, Antong Chen, and Soheil Ghafurian. "**Group equivariant generative adversarial networks.**" ICLR 2021.

Group Equivariant GANs : Motivation

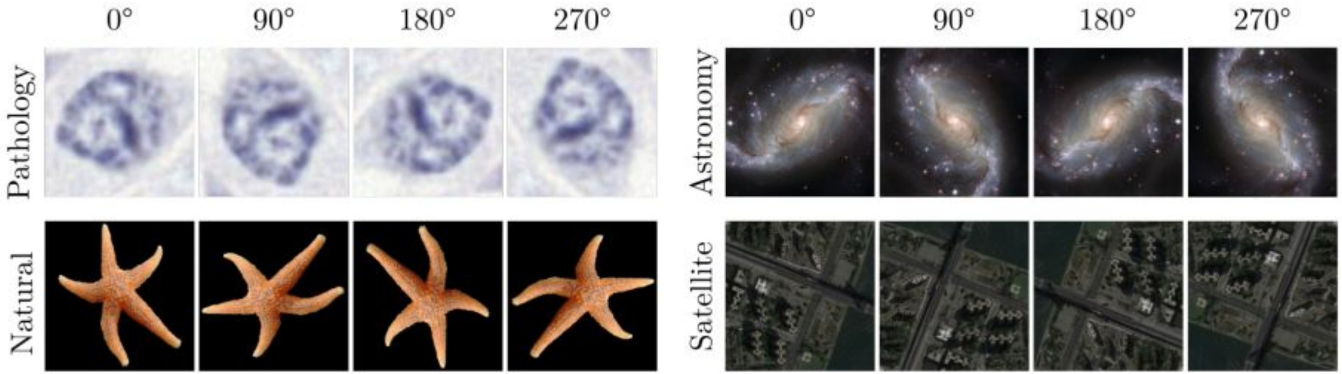


Figure 1: Several image modalities have no preferred orientation for tasks such as classification. We improve their generative modeling by utilizing image symmetries within a GAN framework.

Group Equivariant GANs : Motivation

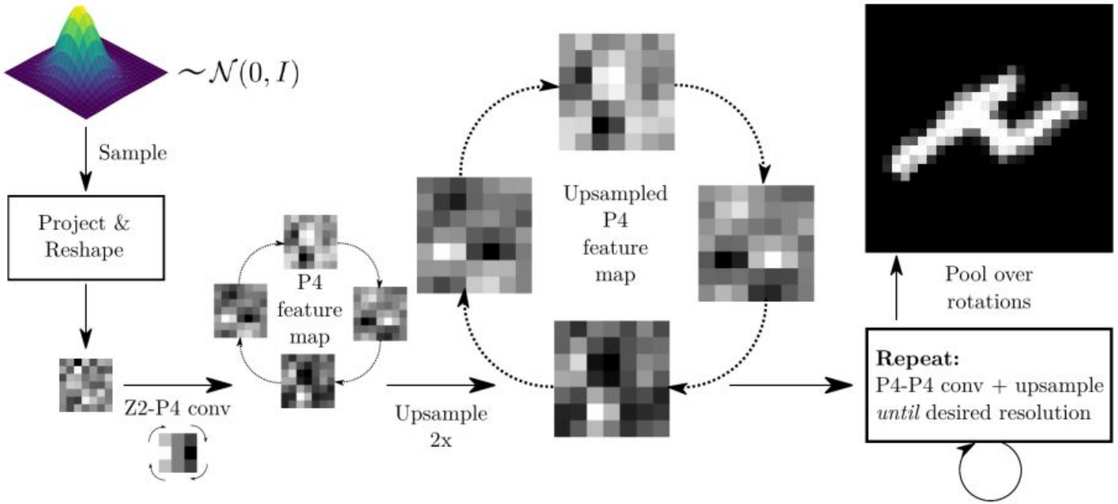
Why not simply use augmentations from the desired group?

Built-in equivariance outperforms augmentations.

- Veeling, Bastiaan S., et al. "Rotation equivariant CNNs for digital pathology." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2018.
- Lafarge, Maxime W., et al. "Roto-translation equivariant convolutional networks: Application to histopathology image analysis." Medical Image Analysis 68 (2021): 101849.

Group Equivariant GANs : Method

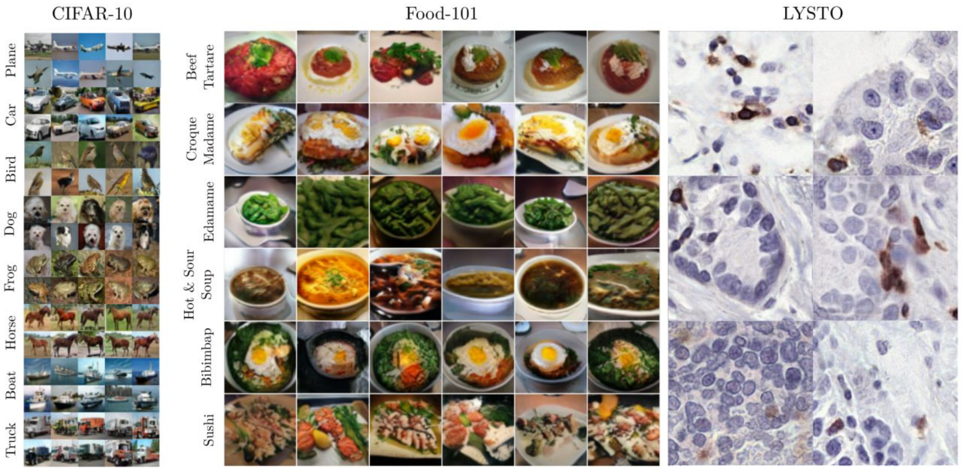
Add group convolutions in generator and/or discriminator!



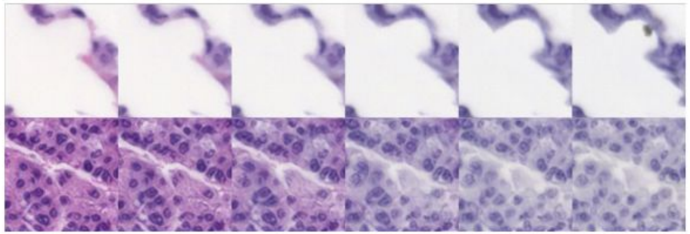
Group Equivariant GANs : Method

- Focus on conditional image generation
  - Both generator and discriminator are conditioned on class
- Class conditional batch normalization :
  - Learn scale and shift features from class labels
  - **Equivariance :** Learn transformation features for each group-feature map instad of each spatial feature.

Group Equivariant GANs : Five Benchmarks



(a) RotMNIST



(b) ANHIR

Group Equivariant GANs : Fréchet Inception Distance

1. Extract Inception features from generated and ground truth images
2. Fit two Gaussians :
    a. One for generated features
    b. One for ground truth features
3. Compute Fréchet distance (closed form in the case of Gaussians)

$$d_F(\mathcal{N}(\mu,\Sigma),\mathcal{N}(\mu',\Sigma'))^2 = \|\mu-\mu'\|_2^2 + \mathrm{tr}\left(\Sigma+\Sigma'-2\left(\Sigma^{\frac{1}{2}}\cdot\Sigma'\cdot\Sigma^{\frac{1}{2}}\right)^{\frac{1}{2}}\right)$$

Quantitative Results

RotMNIST with different % of training data.

Ablation over using group convolutions in generator, discriminator or in both.

Equivariance is generally helpful.

Unclear if it should be used in the discriminator, generator or both.

Table 2: Min. & mean Fréchet distances (lower is better) of generated RotMNIST samples, evaluated at every 1K generator iterations. All evaluations are visualized in Appendix A Figure 6.

Loss	Setting	Min. & Mean Fréchet Distance Available Training Data 10%	33%	66%	100%
-	Real data	0.6854	0.3208	0.1324	0.1296
RAGAN	CNN in G & D	(2.04, 11.40)	(1.42, 11.65)	(1.20, 11.10)	(1.36, 11.68)
RAGAN	CNN in G & G-CNN in D	(1.84, 4.26)	(0.88, **3.26**)	(0.52, 2.85)	(0.53, 3.12)
RAGAN	G-CNN in G & CNN in D	(1.49, 9.75)	(1.08, 9.29)	(0.90, 8.70)	(0.95, 9.62)
RAGAN	G-CNN in G & D	(1.61, **4.25**)	(0.76, 3.40)	(0.54, 2.92)	(0.53, **2.90**)
NSGAN	CNN in G & D	(**1.00**, 7.02)	(0.74, 8.25)	(0.84, 8.07)	(0.97, 8.49)
NSGAN	CNN in G & G-CNN in D	(2.77, 5.48)	(1.02, 3.51)	(0.55, **2.85**)	(0.54, 3.08)
NSGAN	G-CNN in G & CNN in D	(**1.00**, 7.00)	(0.96, 7.42)	(0.87, 6.83)	(0.94, 7.52)
NSGAN	G-CNN in G & D	(2.85, 5.67)	(1.04, 4.24)	(0.82, 3.27)	(0.64, 3.32)
WGAN	CNN in G & D	(3.42, 16.21)	(3.90, 18.32)	(3.87, 17.81)	(4.88, 19.40)
WGAN	CNN in G & G-CNN in D	(2.87, 5.98)	(0.76, 4.11)	(**0.50**, 3.57)	(**0.39**, 3.51)
WGAN	G-CNN in G & CNN in D	(2.67, 16.02)	(3.40, 17.03)	(3.77, 17.76)	(3.74, 17.82)
WGAN	G-CNN in G & D	(2.51, 5.67)	(**0.58**, 3.32)	(0.56, 3.52)	(0.54, 3.76)

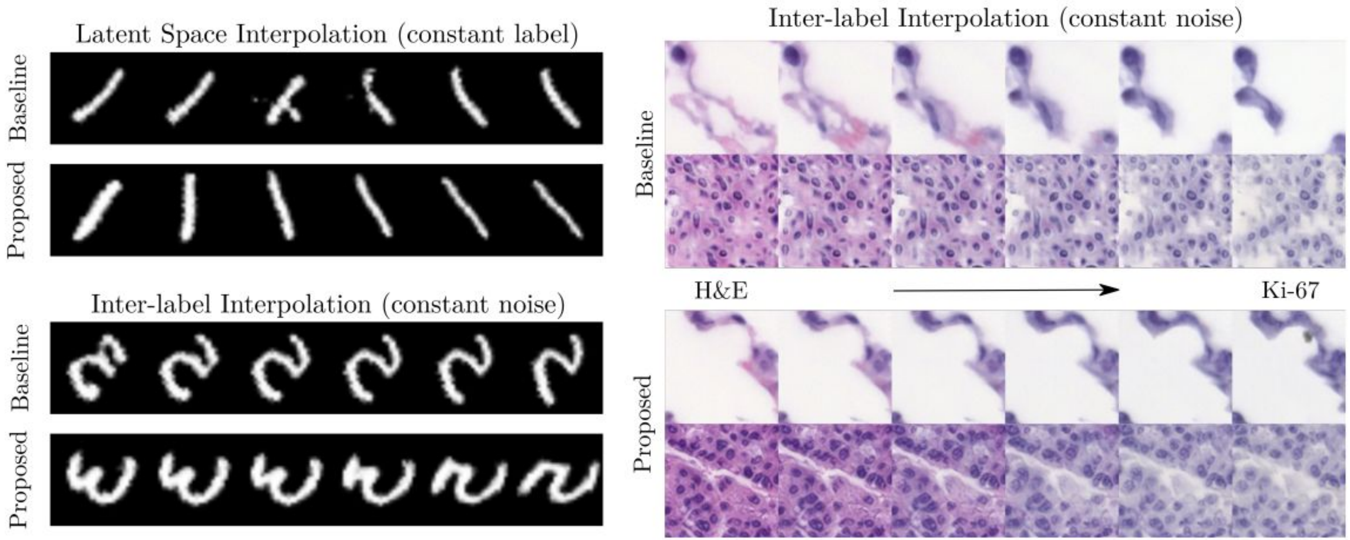
Quantitative Results

Real-world datasets. G-CNN appear to be particularly useful in the discriminator.

Table 3: FID evaluation (lower is better) of all real-world datasets across ablations and augmentation-based baseline comparisons. - indicates an inapplicable setting for the method.

	Setting	ANHIR	LYSTO	CIFAR-10	Food-101
Ablation	CNN in G & D	7.32	7.27	20.89	27.34
	G-CNN in G; CNN in D	6.93	6.68	21.20	24.16
	CNN in G; G-CNN in D	5.56	5.02	**17.09**	**16.91**
	G-CNN in G & D	**5.54**	**3.90**	17.49	17.73

Qualitative Results : Interpolations



(a) RotMNIST (b) ANHIR

Qualitative Results : Map translation

Pros

- Careful study of components to ensure equivariance throughout the architecture (e.g., batch normalization).
- Motivation for some datasets is clear (maps, LYSTO, ANHIR).
- Compelling qualitative results.

Cons

- Unclear if group convolutions are always helpful or should only be used in generator or discriminator.
- They claim sample efficiency, but only test the claim on RotMNIST.

Future work

- Current approach is restricted to discrete groups of a moderate size.
- Continuous equivariance (e.g., SE(2))?
- Potential diminishing returns of continuous equivariance.
  - Lafarge et al., 2020a.

Thank you!