

GAN Evaluation Metrics

1. Introduction
2. Related Works
3. Proposed Idea
4. Experiments
5. Conclusion and Discussion

Image quality assessment: from error visibility to structural similarity (2004, Transactions on Image Processing)

Structure SIMilarity

사람 시각 시스템은 이미지에서 구조 정보를 도출하는데 특화되어 있기 때문에 구조 정보의 왜곡 정도가 시각 품질에 가장 큰 영향을 미친다.

가장 영향을 많이 미치는 3요소로 각 두 이미지를 비교하여 비율로서 "얼마나 손상되었는가"에 대한 점수를 냄

휘도(평균 밝기), 명암비, 구조 세 가지로 구성되어 있음

$$SSIM(x,y)=l(x,y)c(x,y)s(x,y)$$

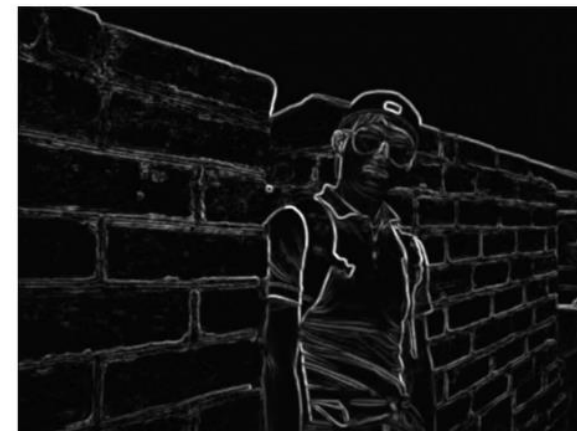
SSIM



휘도(brightness)



명암비(contrast)



구조 (structure)

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad \text{휘도 (평균밝기, Luminance)}$$

$$C_1 = (0.01 * L)^2$$

$$C_2 = (0.03 * L)^2$$

$$C_3 = C_2/2$$

두 이미지간의 평균밝기를 비교

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad \text{명암비 (Contrast)}$$

두 이미지의 표준편차 값을 이미지의 명암비로 정하고 값을 비교

$$\frac{A - \mu_A}{\sigma_A}, \frac{B - \mu_B}{\sigma_B}$$

이미지 구조는 이미지에서 평균 휘도를 뺀 후
표준편차를 나눔

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad \text{구조 (Structure)}$$

x,y 상관관계를 구하면서 비교함

SSIM



SSIM

SSIM



$$\text{IS}(G) = \exp \left(\mathbb{E}_{\mathbf{x} \sim p_g} D_{KL} (p(y|\mathbf{x}) \parallel p(y)) \right),$$

- Image의 퀄리티를 평가하는 metric
- 두 가지의 평가기준으로 평가함
 - 이미지가 얼마나 다양한가
 - 각 이미지가 뚜렷한 어떤 것을 형성하고 있다 라는 직관적인 관점

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. In Advances in Neural Information Processing Systems, pp. 2234-2242, 2016.



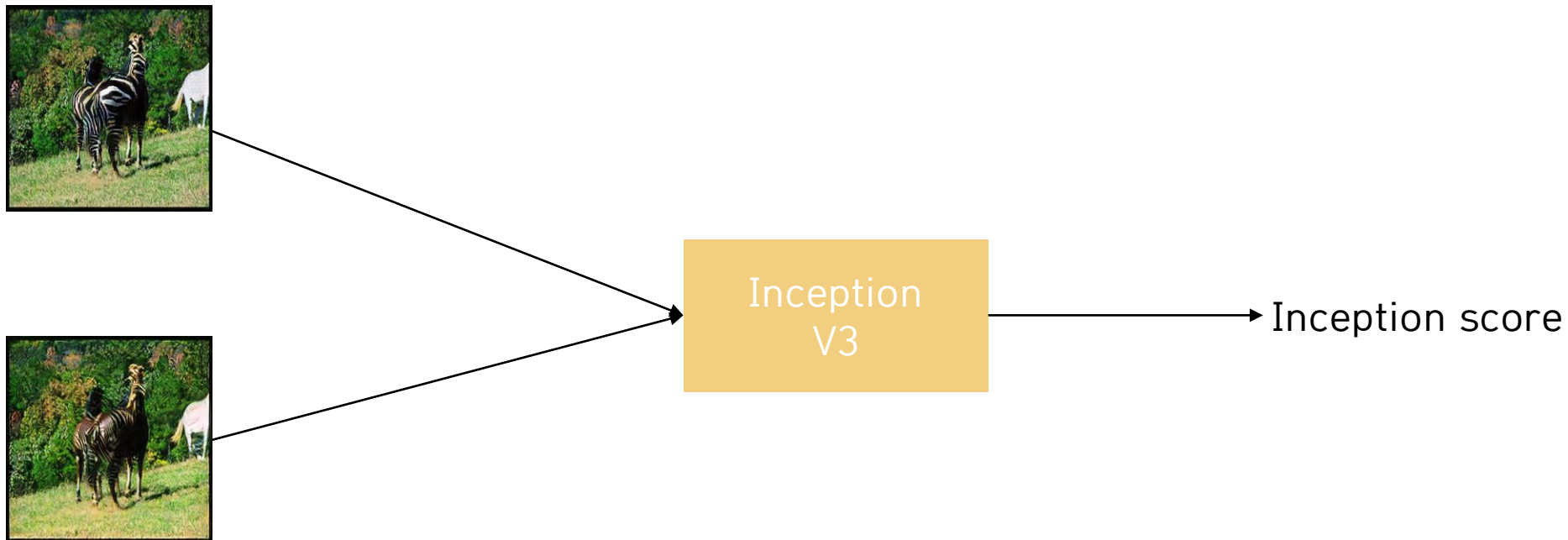
$$\text{IS}(G) = \exp \left(\mathbb{E}_{\mathbf{x} \sim p_g} D_{KL}(\underline{p(y|\mathbf{x})} \parallel p(y)) \right),$$



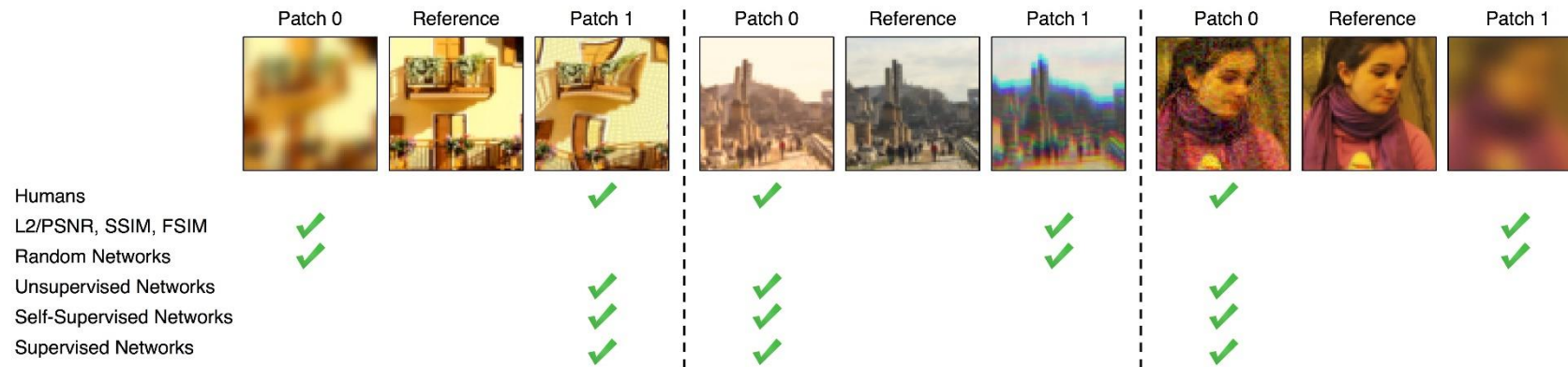
- Inception v3 를 이용하여 Classification 문제를 풀 때, 얼마나 뚜렷하고 정확하게 구분이 되는지를 판별하여 점수를 구분

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. In Advances in Neural Information Processing Systems, pp. 2234-2242, 2016.

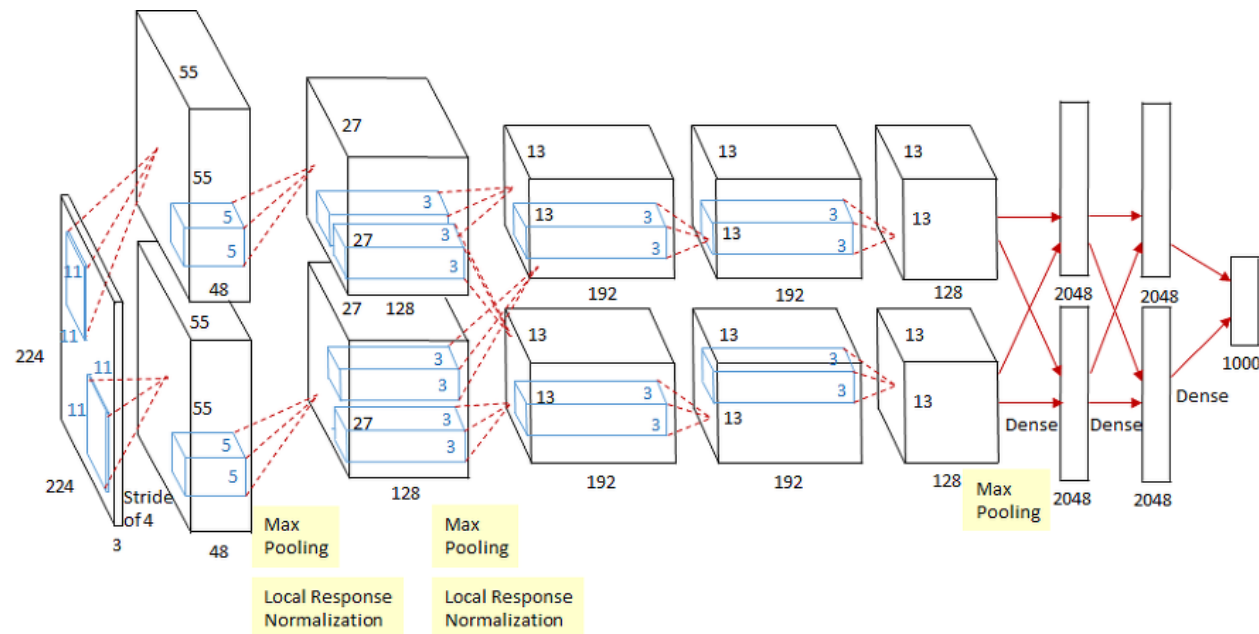
Improved techniques for training gans, 2016 NIPS



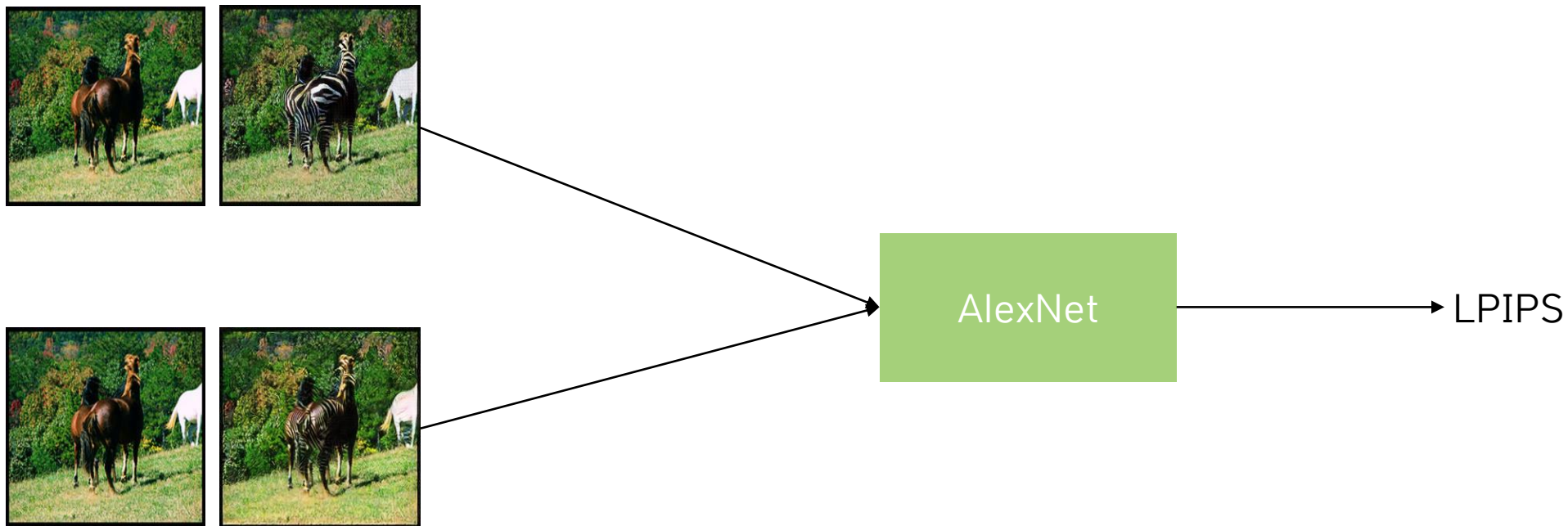
Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. In Advances in Neural Information Processing Systems, pp. 2234-2242, 2016.



- 기존 방법들은 Human perceptual과는 거리가 있다.
- 딥-러닝사용시 어느정도 일치하는 것을 볼 수 있다.



- Pretrained AlexNet을 이용해서 추출된 feature를 통해 L1 Distance를 구함
- 낮은 값은 두 이미지가 유사하다는 의미를 갖는다
- 높은 값은 두 이미지가 다르다는(diversity) 의미를 갖는다

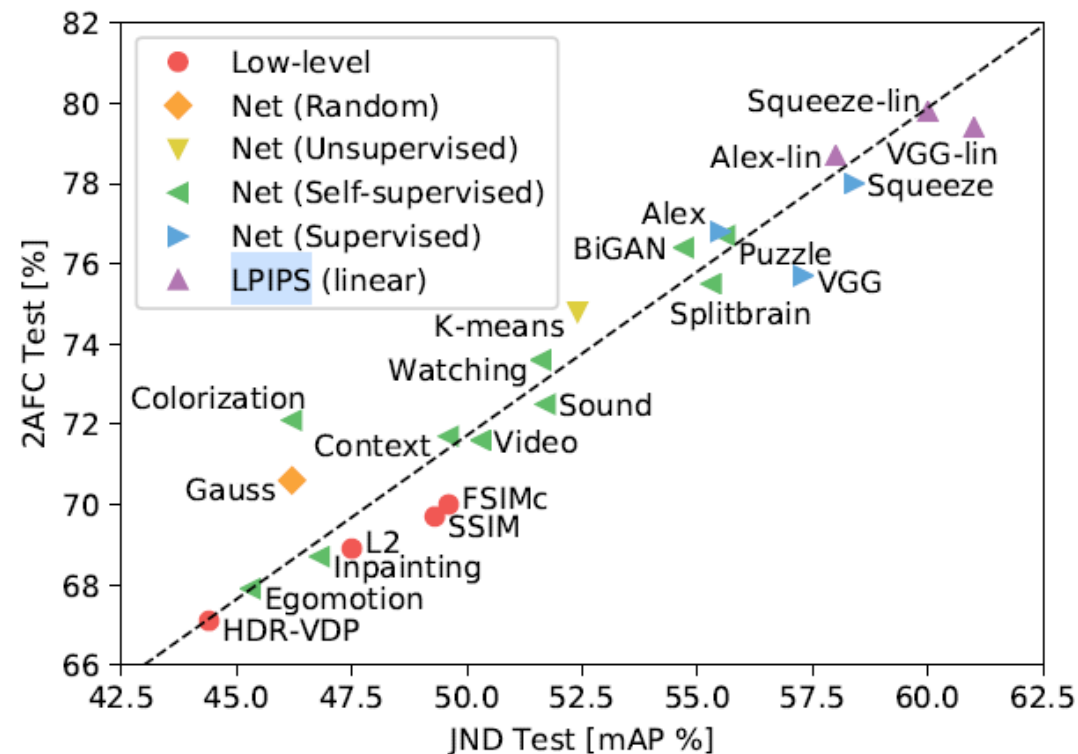
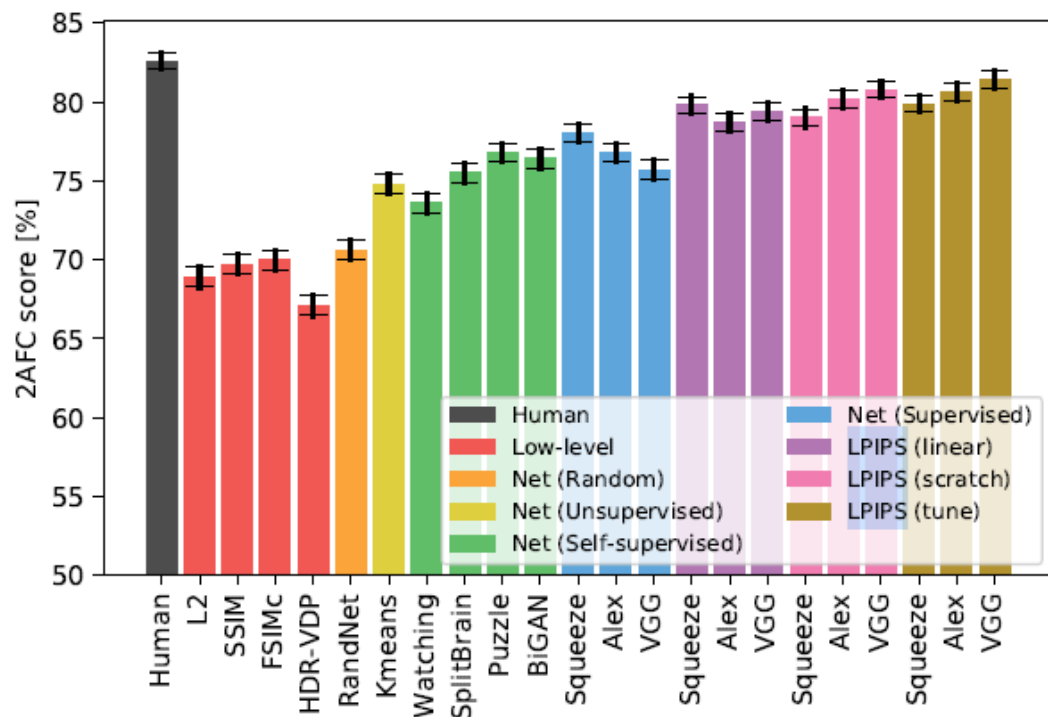


- Real data와 Fake data의 페어를 비교하면 스코어가 나옴
- 일반적으로 이미지의 완성도를 보는것이 아닌, 이미지가 얼마나 변했는지(diversity)를 검사
- 보통 multimodal에서 이미지의 다양성을 판단하기 위해 많이 쓰임



- 보통 multimodal에서 이미지의 다양성을 판단하기 위해 많이 쓰임

Distortions



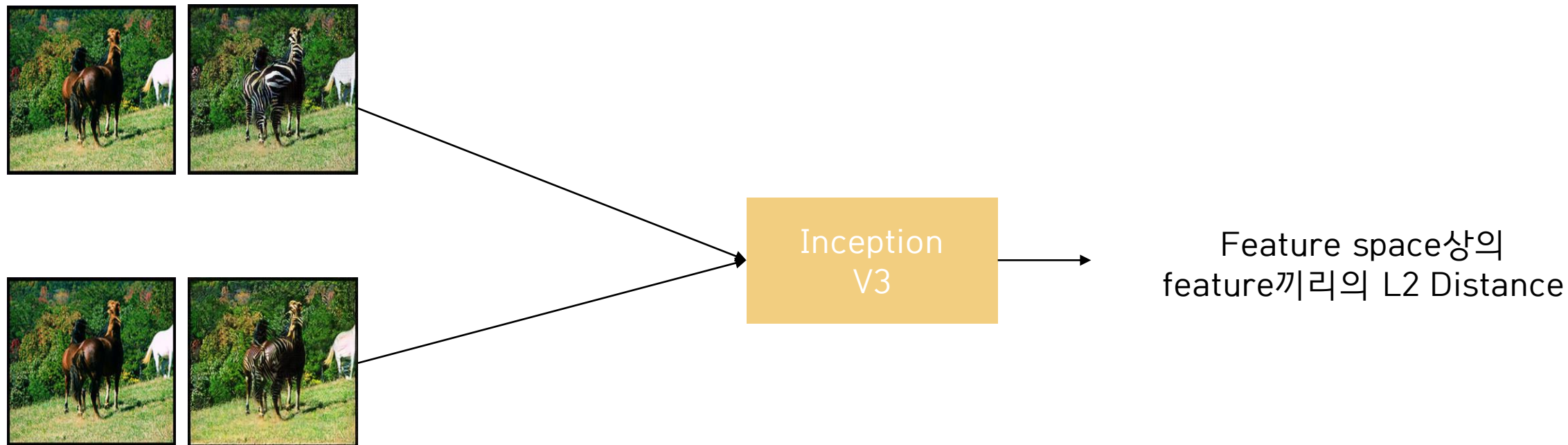
- 기존 방법들보다 Human Perceptual에 가까움

$$d^2((m, C), (m_w, C_w)) = \|m - m_w\|_2^2 + \text{Tr}(C + C_w - 2(CC_w)^{1/2}) .$$

- Inception score와 마찬가지로 Inception V3를 사용해 Feature를 추출함.
- 두 이미지를 비교해 얼마나 유사한지 비교 (Similarity)

M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," in NIPS, 2017.

Frechet Inception Distance



FID

GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium

Frechet Inception Distance

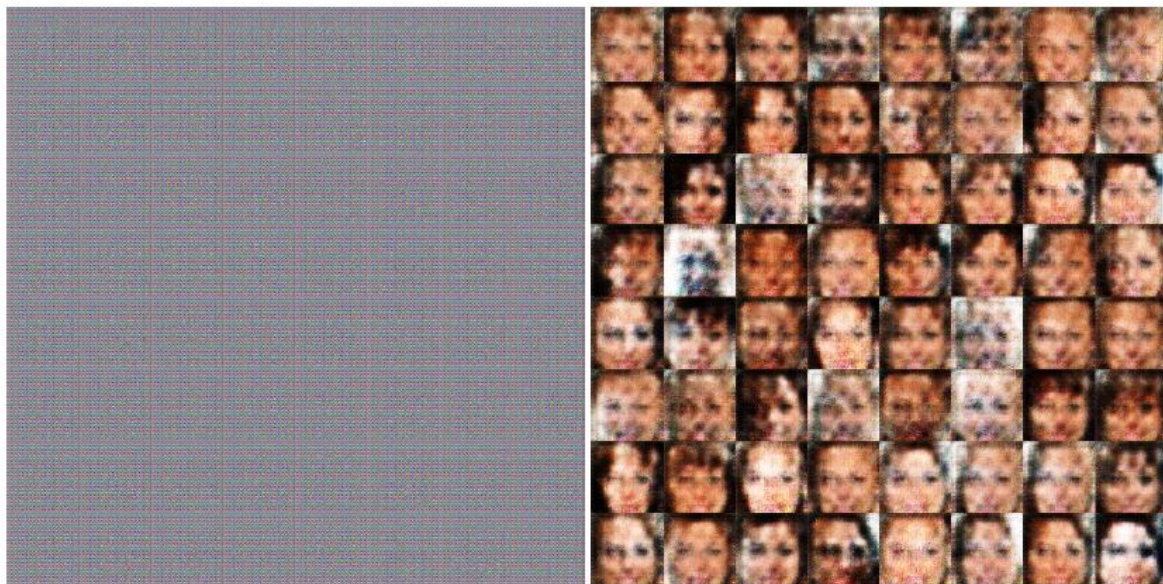


Figure A9: Samples generated from DCGAN trained on CelebA with different FIDs. **Left:** FID 500 and **Right:** FID 300.

FID 500

FID 300



Figure A10: Samples generated from DCGAN trained on CelebA with different FIDs. **Left:** FID 133 and **Right:** FID 100.

FID 133

FID 100

FID

GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium

Frechet Inception Distance



Figure A11: Samples generated from DCGAN trained on CelebA with different FIDs. **Left:** FID 45 and **Right:** FID 13.

FID 45

FID 13



Figure A12: Samples generated from WGAN-GP trained on CelebA with a FID of 3.

FID 3

LPIPS↑	FID ↓	SSIM↑	IS↑
Evaluate only multimodal	210.48	0.7898	1.3771
	195.04	0.8495	1.4354
	221.09	0.4392	1.5045
	226.16	0.4166	1.4620
0.114	17.61	0.8221	3.2183
0.109	19.36	0.8252	3.1949
0.047	105.94	0.3344	1.8457
0.044	105.81	0.3348	1.8322