HARMONIC UNPAIRED IMAGE-TO-IMAGE TRANSLATION

ICLR 2019

Rui Zhang
Google Cloud AI &
Chinese Academy of Sciences
Beijing, China
zhangrui@ict.ac.cn

Tomas Pfister
Google Cloud AI
Sunnyvale, USA
tpfister@google.com

Jia Li
Google Cloud AI
Sunnyvale, USA
lijiali@google.com

한양대학교 AI Lab 유재창

Concepts

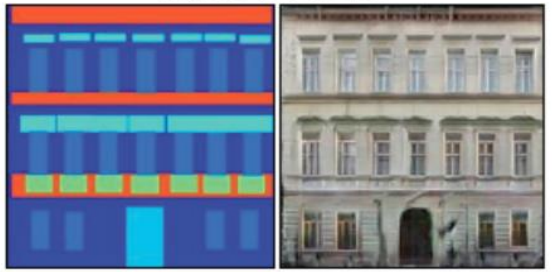
Image to Image translation

Source image를 target image로 변환시켜주는 기술

BW to Color

Labels to Street Scene

Labels to Facade



Edges to Photo

Day to Night

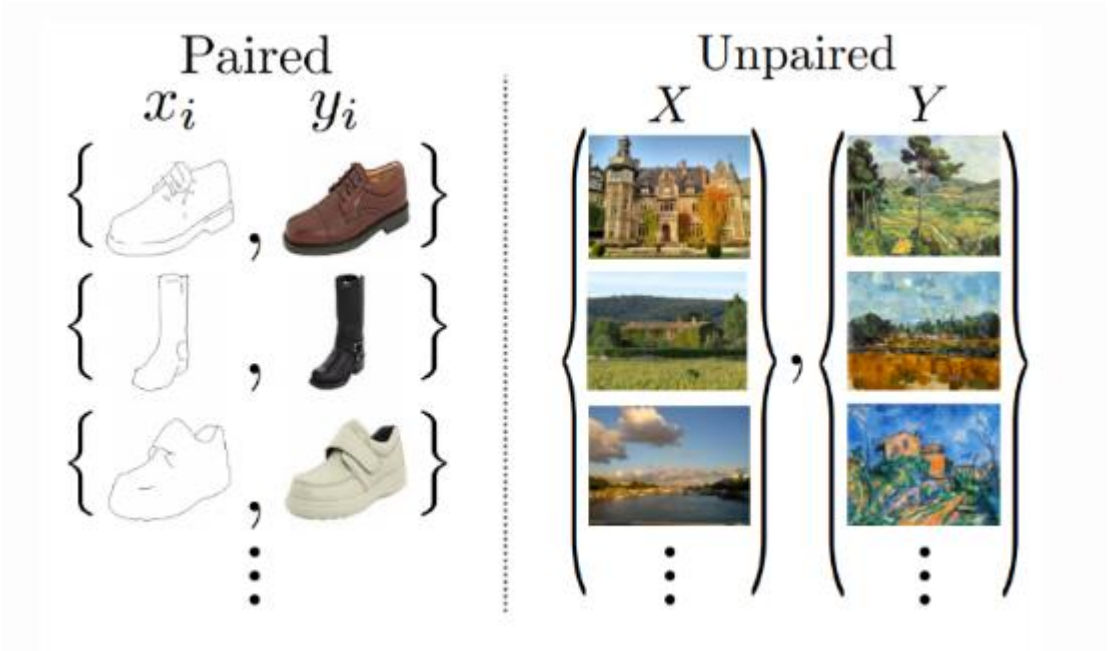
Concepts

Paired / Unpaired image

초기 I2I모델의 경우 Pair의 이미지 쌍이 존재해야만 학습가능.

Concepts

Paired / Unpaired image



초기 I2I모델의 경우 Pair의 이미지 쌍이 존재해야만 학습가능.

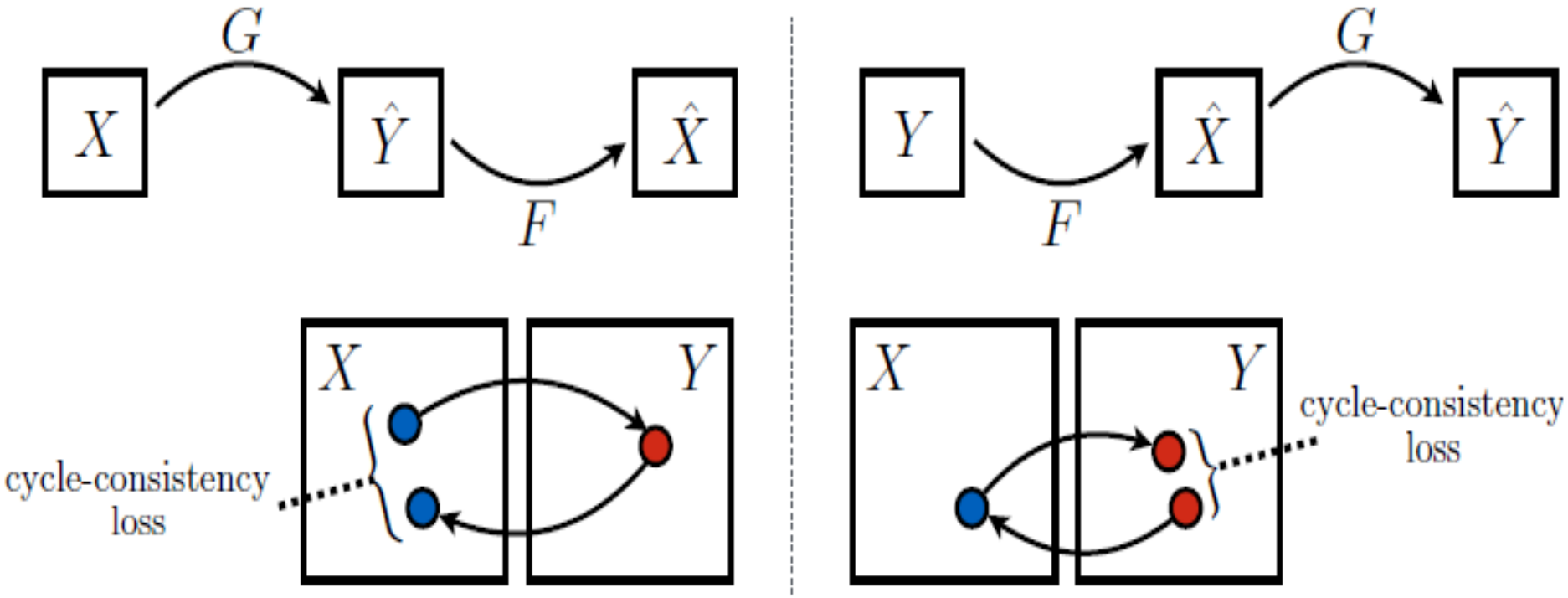
Style transfer와 같은 경우 Pair의 이미지쌍을 구하기 힘듦.

Previous Work

Paired I2I – Pix2Pix…
Unpaired I2I – **CycleGAN**, DiscoGAN, DistanceGAN….

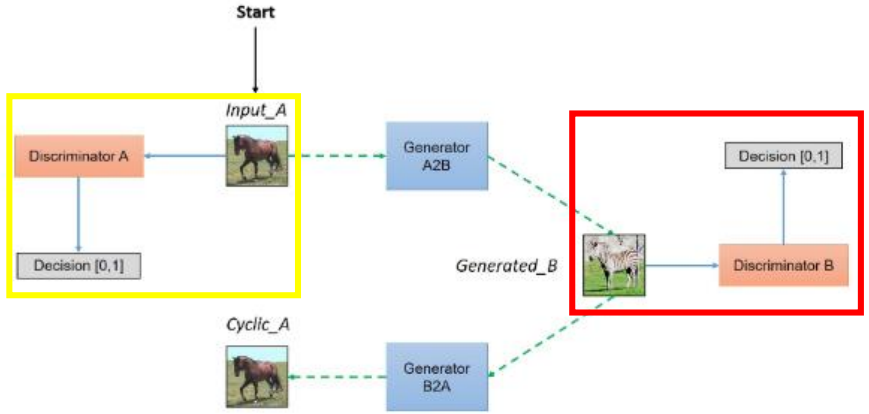
Previous Work

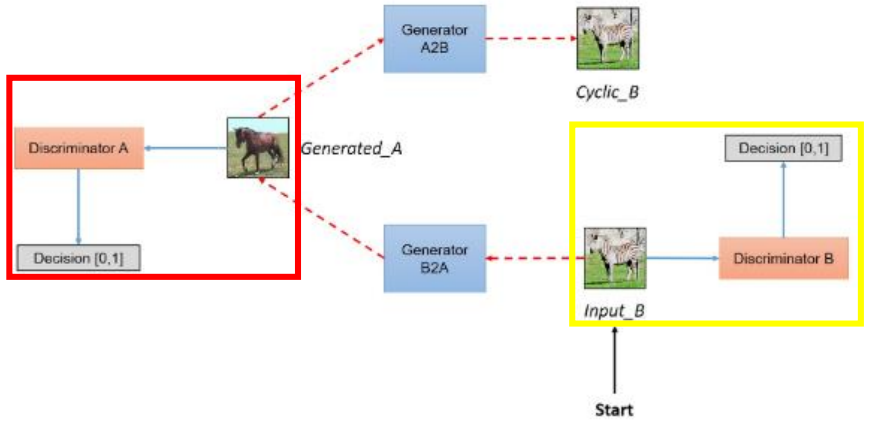
Cycle GAN



Previous Work

Cycle GAN





Loss function

$$\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) = \mathbb{E}_{\mathbf{y}\in Y}[\log D_Y(\mathbf{y})] + \mathbb{E}_{\mathbf{x}\in X}[\log(1 - D_Y(G(\mathbf{x})))]$$

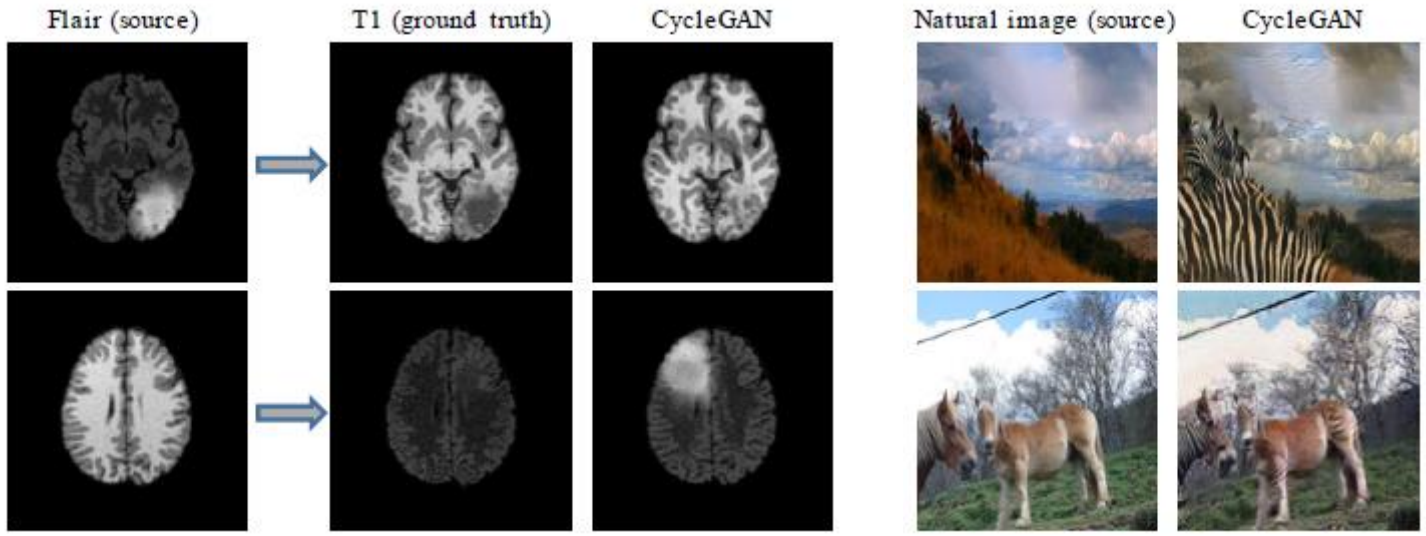
$$\mathcal{L}_{\text{GAN}}(F, D_X, X, Y) = \mathbb{E}_{\mathbf{x}\in X}[\log D_X(\mathbf{x})] + \mathbb{E}_{\mathbf{y}\in Y}[\log(1 - D_X(F(\mathbf{y})))]$$

$$\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{\mathbf{x}\in X}\|F(G(\mathbf{x})) - \mathbf{x}\|_1 + \mathbb{E}_{\mathbf{y}\in Y}\|G(F(\mathbf{y})) - \mathbf{y}\|_1$$

$$\mathcal{L}_{\text{CycleGAN}}(G, F) = \lambda_{\text{GAN}} \times \mathcal{L}_{\text{GAN}}(G, F) + \lambda_{\text{cyc}} \times \mathcal{L}_{\text{cyc}}(G, F)$$

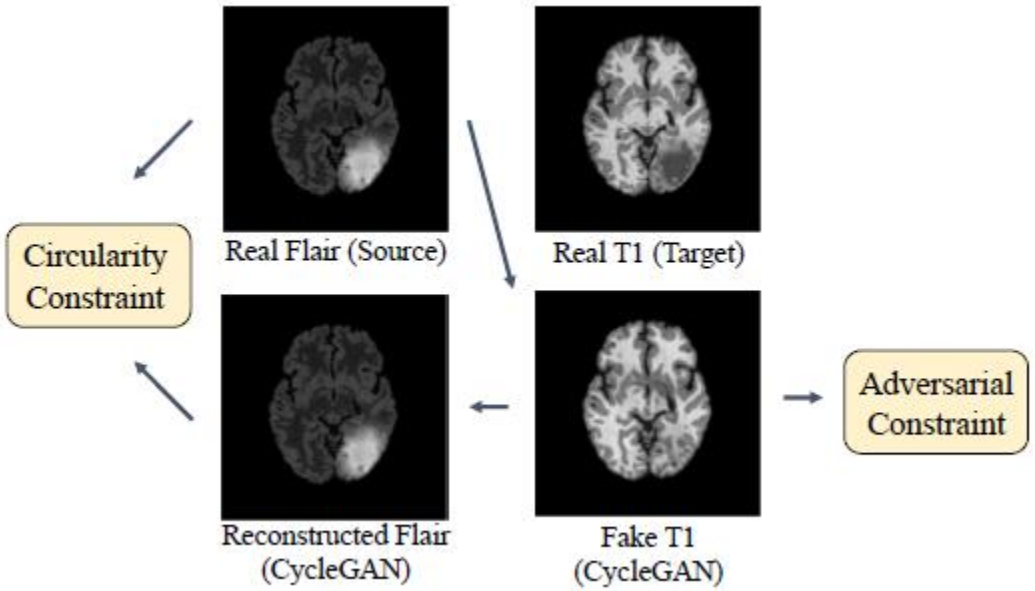
Previous Work

Cycle GAN의 한계점



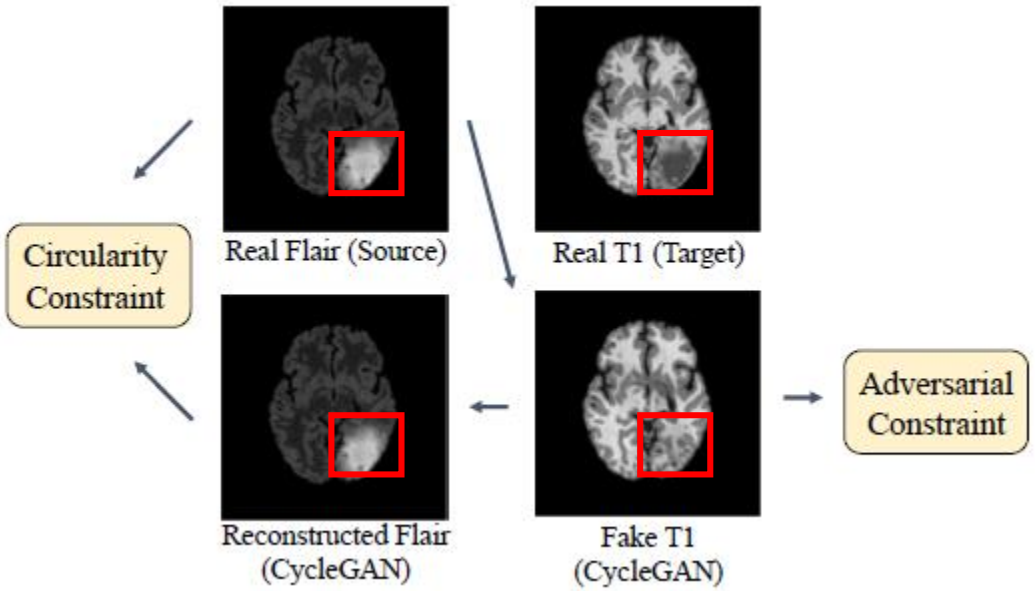
Previous Work

Cycle GAN의 한계점



Previous Work

Cycle GAN의 한계점

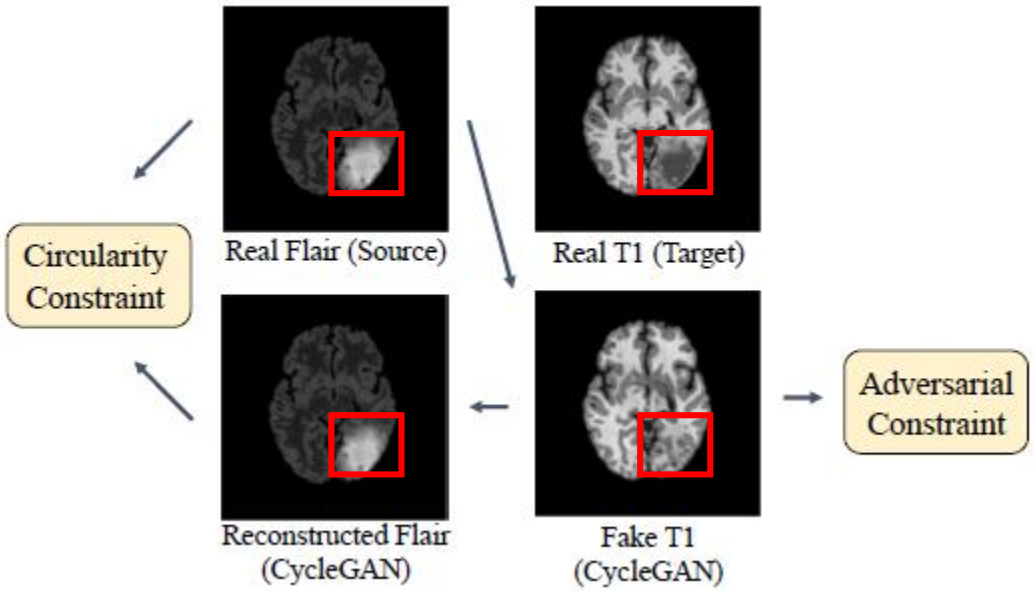


Reconstruct된 image는 input image space의 consistency가 지켜지지만 translation된 image는 inconsistency함.

Translation 과정에서 어느정도 변환 전의 data에 대한 특성을 유지할 수 있는 consistent한 mapping이 필요.

Harmonic GAN

Method



Manifold learning의 관점으로 문제를 해결.

같은 그림의 비슷한 부분들은 비슷한 translation이 일어난다.

비슷한 부분들은 Target image에서도 비슷한 의미를 가질 것이다.

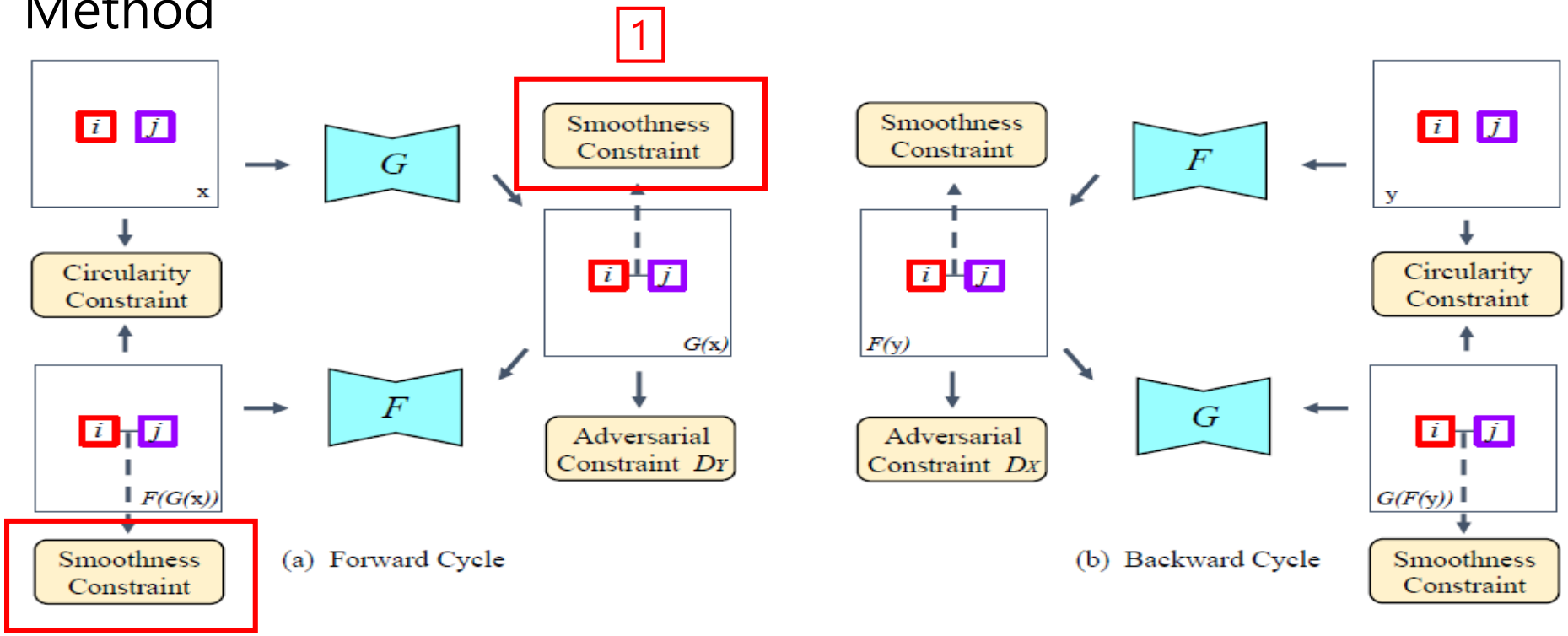
Harmonic GAN

$$w_{ij}(X) = \exp\{-Dist[\vec{x}(i), \vec{x}(j)]/\sigma^2\}$$

$\vec{x}_i$ is referred to as the feature vector of the i-th image patch in $\mathbf{x} \in X$

L1 distance as the Dist function

Method



$$\mathcal{L}_{\text{Smooth}}(G,X,Y)=\mathbb{E}_{\mathbf{x}\in X}\left[\sum_{i,j} w_{ij}(X)\times Dist[G(\vec{x})(i),G(\vec{x})(j)]+\sum_{i,j} w_{ij}(G(X))\times Dist[F(G(\vec{x}))(i),F(G(\vec{x}))(j)]\right]$$

1 2

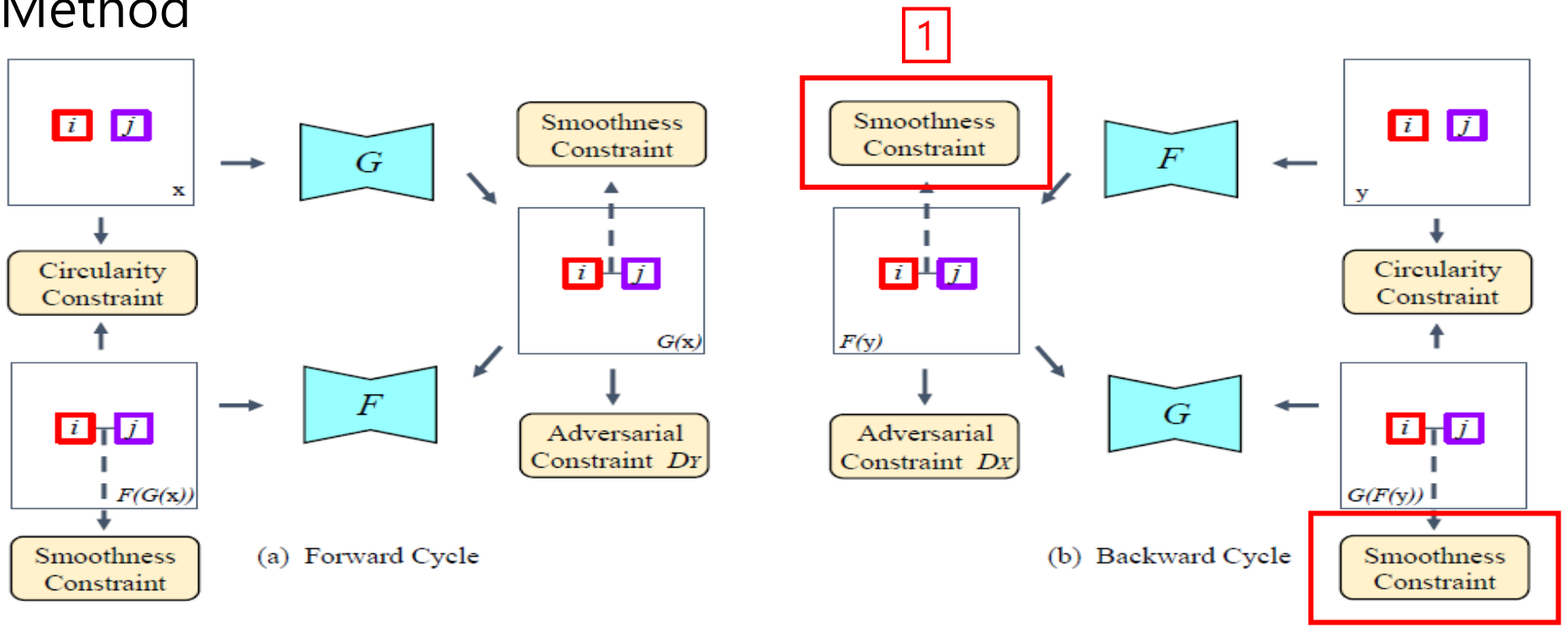
Harmonic GAN

$$w_{ij}(X) = \exp\{-Dist[\vec{x}(i), \vec{x}(j)]/\sigma^2\}$$

$\vec{x}_i$ is referred to as the feature vector of the i-th image patch in $\mathbf{x} \in X$

L1 distance as the Dist function

Method



$$\mathcal{L}_{\text{Smooth}}(F,Y,X)=\mathbb{E}_{\mathbf{y}\in Y}\left[\sum_{i,j} w_{ij}(Y)\times Dist[F(\vec{y})(i),F(\vec{y})(j)]+\sum_{i,j} w_{ij}(F(Y))\times Dist[G(F(\vec{y}))(i),G(F(\vec{y}))(j)]\right]$$

1 2

Harmonic GAN

Method

$$\mathcal{L}_{\text{Smooth}}(G,X,Y)=\mathbb{E}_{\mathbf{x}\in X}\left[\sum_{i,j} w_{ij}(X)\times Dist[G(\vec{x})(i),G(\vec{x})(j)]+\sum_{i,j} w_{ij}(G(X))\times Dist[F(G(\vec{x}))(i),F(G(\vec{x}))(j)]\right]$$

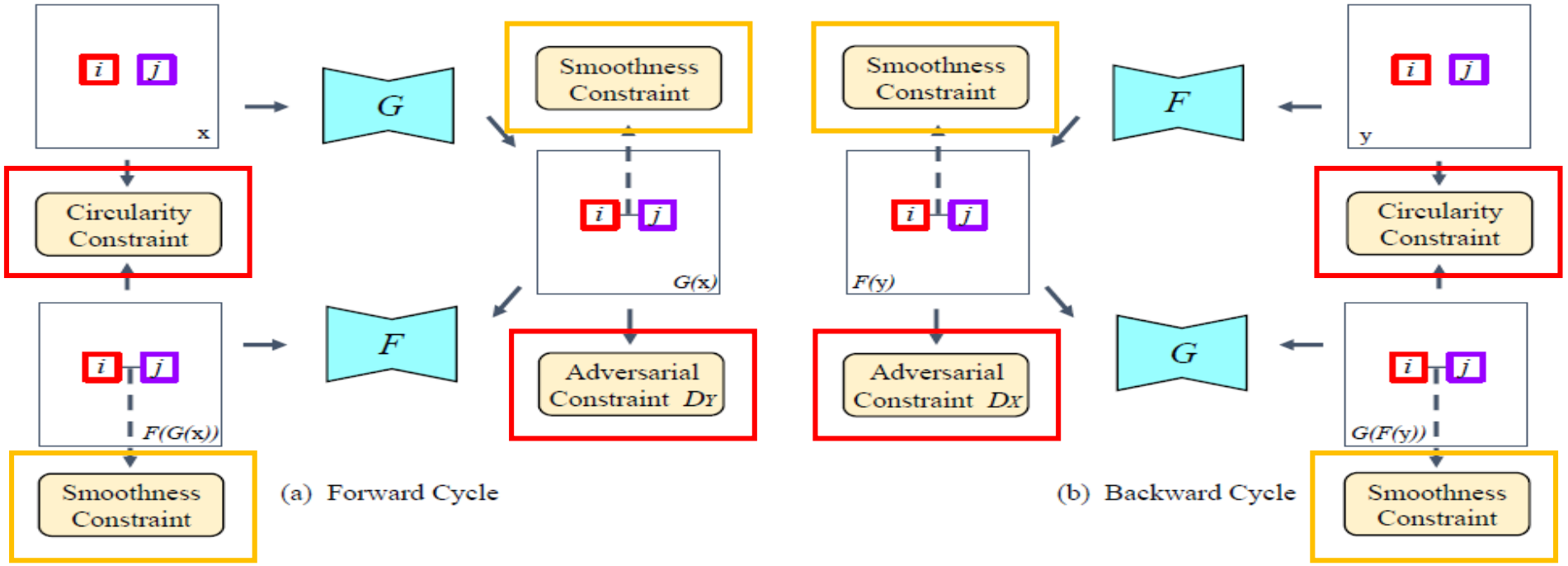
$$\mathcal{L}_{\text{Smooth}}(F,Y,X)=\mathbb{E}_{\mathbf{y}\in Y}\left[\sum_{i,j} w_{ij}(Y)\times Dist[F(\vec{y})(i),F(\vec{y})(j)]+\sum_{i,j} w_{ij}(F(Y))\times Dist[G(F(\vec{y}))(i),G(F(\vec{y}))(j)]\right]$$

$$\mathcal{L}_{\text{Smooth}}(G,F)=\mathcal{L}_{\text{Smooth}}(G,X,Y)+\mathcal{L}_{\text{Smooth}}(F,Y,X)$$

Harmonic GAN

- Adversarial constraint는 translation된 이미지가 real 이미지와 구분이 어렵게 만듦
- Circularity constraint는 cycle consistency를 보장
- Smoothness constraint는 patch들 사이의 similarity consistency를 보장

Method



$$\mathcal{L}_{\text{HarmonicGAN}}(G,F) = \mathcal{L}_{\text{CycleGAN}}(G,F) + \lambda_{\text{Smooth}} \times \mathcal{L}_{\text{Smooth}}(G,F)$$

Harmonic GAN

Method

Image patch -> feature vector

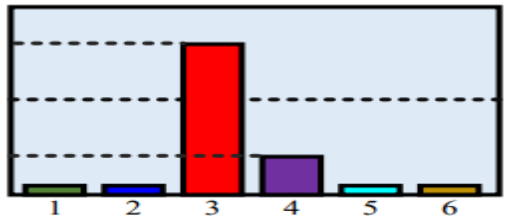
1. VGG-16 network

Pre-trained VGGNet

2. Low-level soft histogram

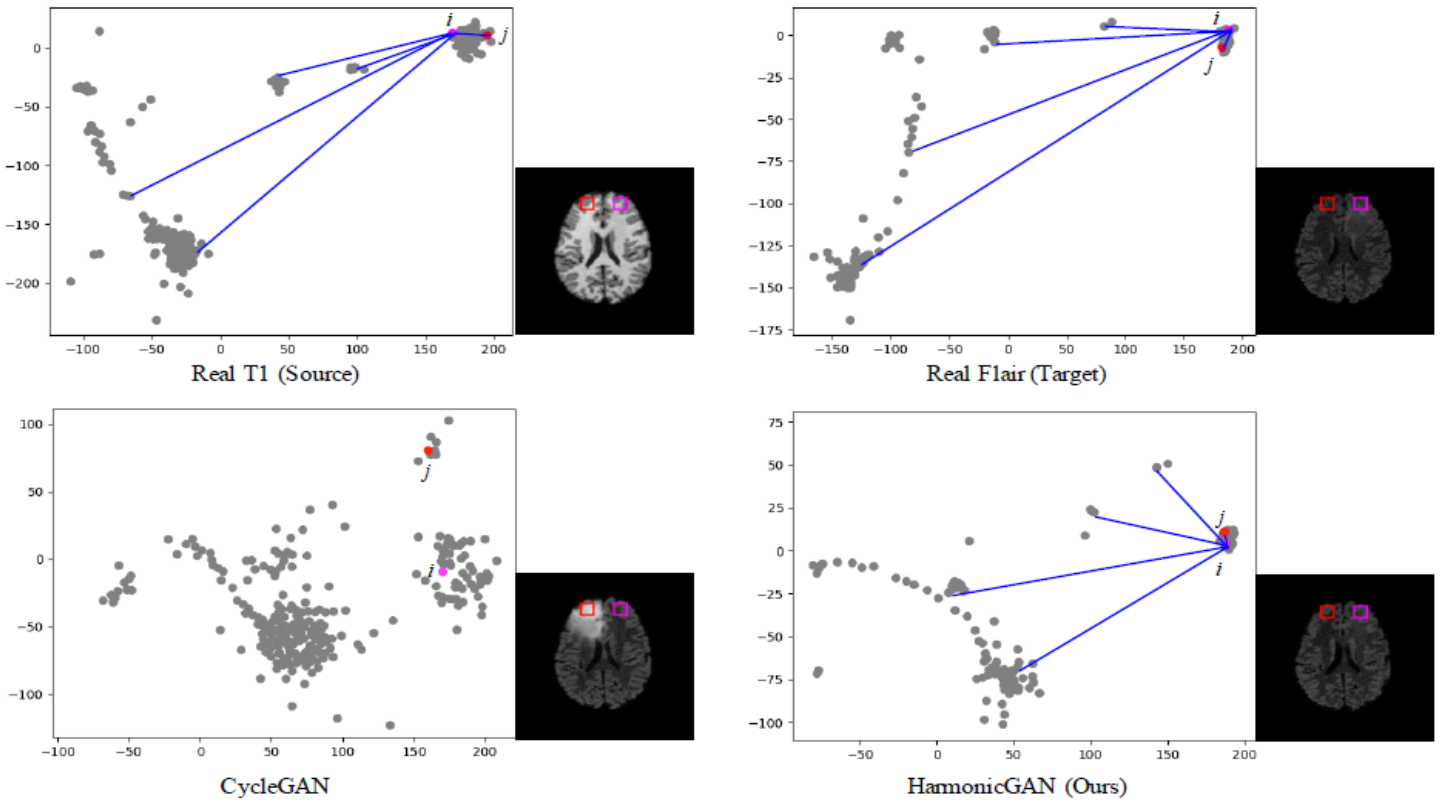
$$\psi_b(\vec{x}_i(j)) = \max\{0, 1 - |\vec{x}_i(j) - \mu_b| \times w_b\}$$

$$\phi_h(X, i, b) = \phi_h(\vec{x}_i, b) = \sum_j \psi_b(\vec{x}_i(j))$$



Harmonic GAN

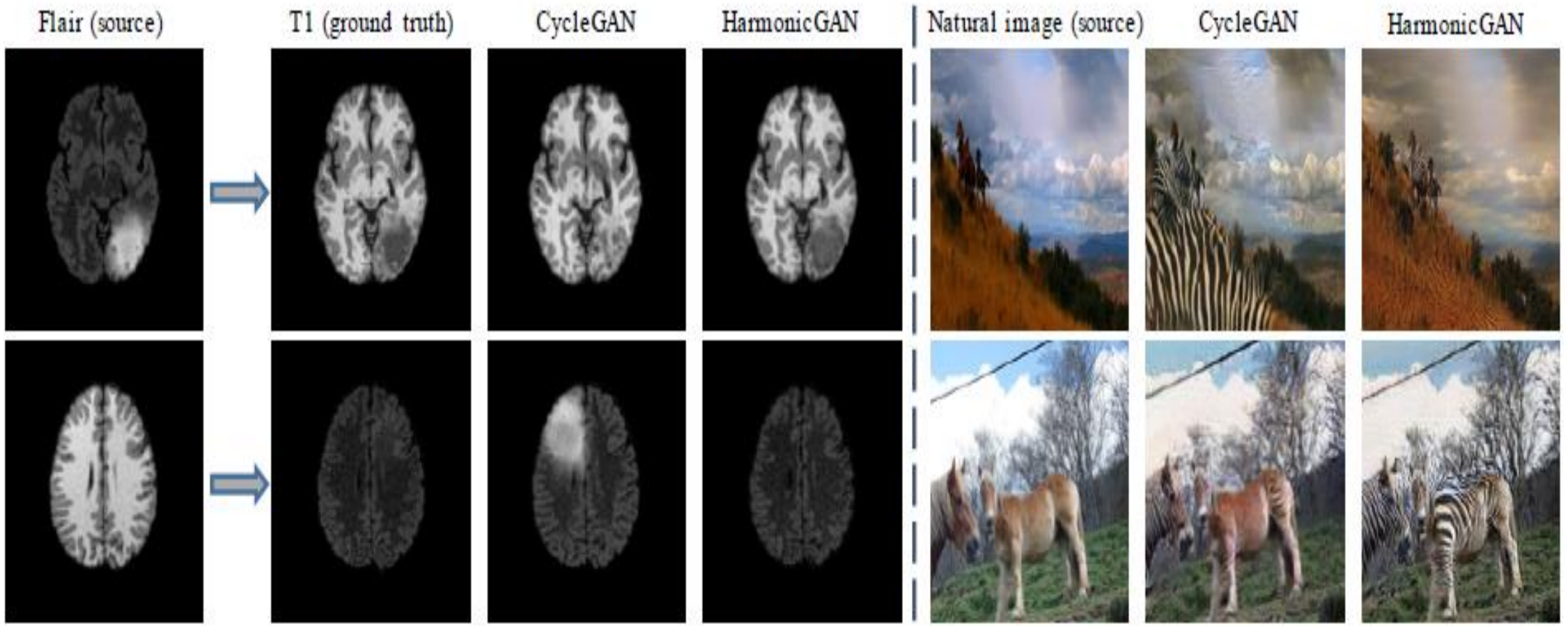
Method



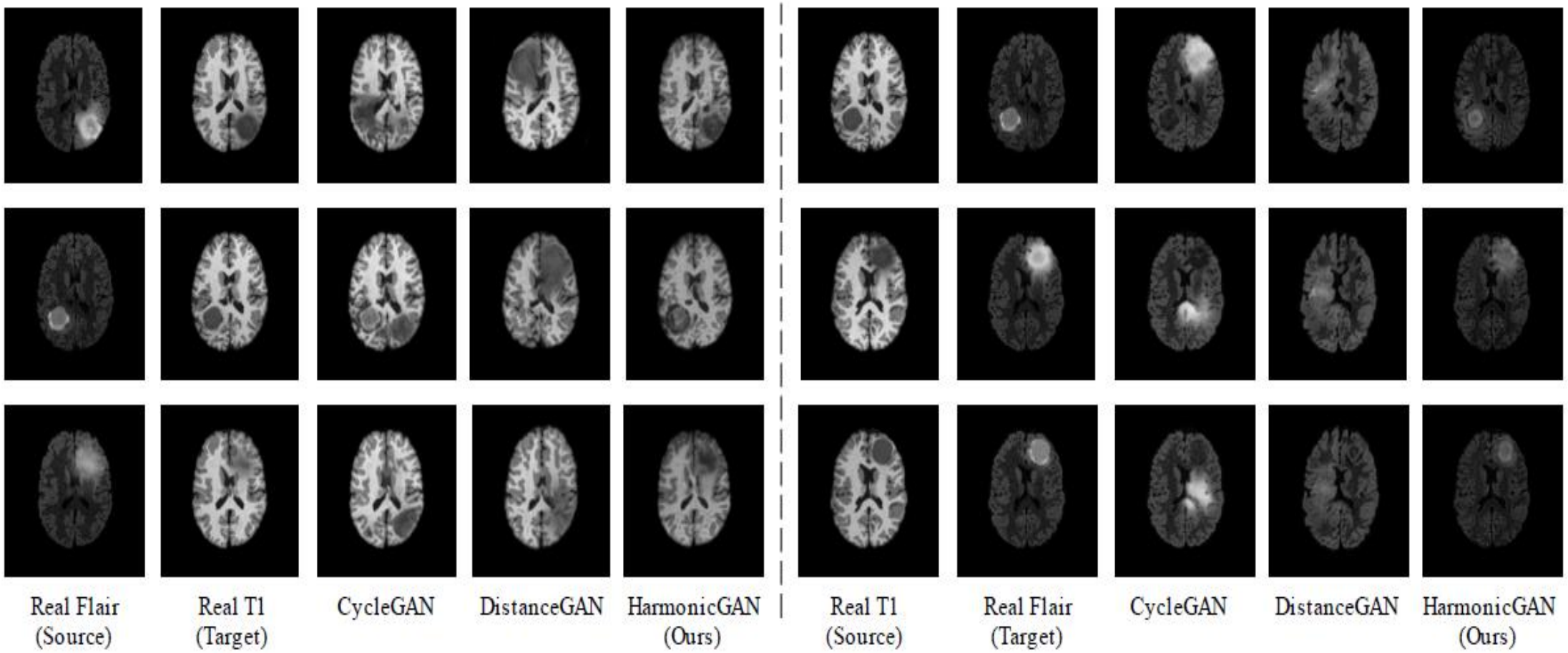
Visualization using t-SNE

- 데이터 차원 축소 및 시각화 방법론.
- Bottom-right : source 에서 두 위치상의 patch -> target space 상의 distance ↓
- Bottom-left : source 에서 두 위치상의 patch -> target space 상의 distance ↑

Experiments Harmonic GAN



Experiments Harmonic GAN



Experiments Harmonic GAN

Method	Flair → T1				T1 → Flair			
	MAE ↓	MSE ↓	PSNR ↑	SSIM ↑	MAE ↓	MSE ↓	PSNR ↑	SSIM ↑
CycleGAN	10.47	674.40	22.35	0.80	11.81	1026.19	18.73	0.74
DiscoGAN	10.63	641.35	20.06	0.79	10.66	839.15	19.14	0.69
DistanceGAN	14.93	1197.64	17.92	0.67	10.57	716.75	19.95	0.64
UNIT	9.48	439.33	22.24	0.76	6.69	261.26	25.11	**0.76**
HarmonicGAN (ours)								
Histogram	**6.38**	**216.83**	**24.34**	**0.83**	5.04	163.29	26.72	0.75
VGG	6.86	237.94	24.14	0.81	**4.69**	**127.84**	**27.22**	**0.76**

Experiments Harmonic GAN

Method	Label → Photo			Photo → Label		
	Pixel Acc. ↑	Class Acc. ↑	Class IoU ↑	Pixel Acc. ↑	Class Acc. ↑	Class IoU ↑
CycleGAN	52.7	15.2	11.0	57.2	21.0	15.7
DiscoGAN	45.0	11.1	7.0	45.2	10.9	6.3
DistanceGAN	48.5	10.9	7.3	20.5	8.2	3.4
UNIT	48.5	12.9	7.9	56.0	20.5	14.3
HarmonicGAN (ours)						
Histogram	52.2	14.8	10.9	56.6	20.9	15.7
VGG	**55.9**	**17.6**	**13.3**	**59.8**	**22.1**	**17.2**

Experiments Harmonic GAN

Table 3: User study on the BRATS dataset.

Metric	CycleGAN	DistanceGAN	HarmonicGAN
Prefer [%] ↑	5	0	**95**
Mean Likert ↑	1.68	1.62	**4.00**
Std Likert	0.99	0.95	0.88

Table 4: User study on the horse to zebra dataset.

Metric	CycleGAN	DistanceGAN	HarmonicGAN
Prefer[%]↑	28	0	**72**
Mean Likert ↑	3.16	1.08	**3.60**
Std Likert	0.81	0.23	0.78

끝