

한국정보처리학회 ASK 2023(춘계)

주파수 영역을 활용한 GAN

GAN using Frequency Domain

한성대학교 AI응용학과 석사과정 이채은
한성대학교 AI응용학과 교수 정성훈

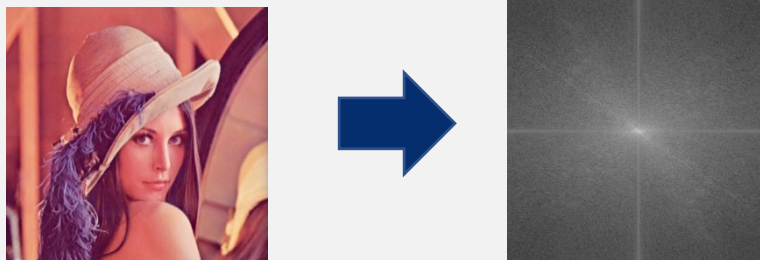
- 연구 소개
- 관련 연구
- 제안 방법
- 실험 환경
- 실험 결과
- 추가 실험
- 결론

- 연구 배경

- 2014년 GAN이 발표된 후로 매우 많은 파생 버전이 발표되었고 이미지 생성 모델로서 이미지 공간에서 좋은 결과를 보임
- GAN은 이미지 생성 뿐 아니라 이미지 변환, 해상도 향상 등 다양한 이미지 처리 분야에 활용
- 그러나, GAN은 학습이 잘 안되는 문제와 다양한 이미지를 생성하는데 한계를 보임
- GAN의 성능을 향상시키기 위한 다양한 방법이 제안되었으나 기존 방법은 모두 이미지 공간에서 처리함
- 이미지 공간보다 주파수 공간에서의 이미지 처리는 다양한 장점을 제공

- 제안 방법

- **GAN의 능력을 더욱 향상하기 위하여 주파수 영역에서 이미지를 학습하고 생성**하는 새로운 방법을 제안
- 주파수 성분으로 분해된 정보는 위치 별 주파수 정보를 포함하고 있으며 이미지의 주요한 구조와 특징을 제공
 - 이미지 공간에서의 정보보다 더욱 다양한 정보를 제공 → 주파수 영역에서 성능이 향상될 것으로 기대

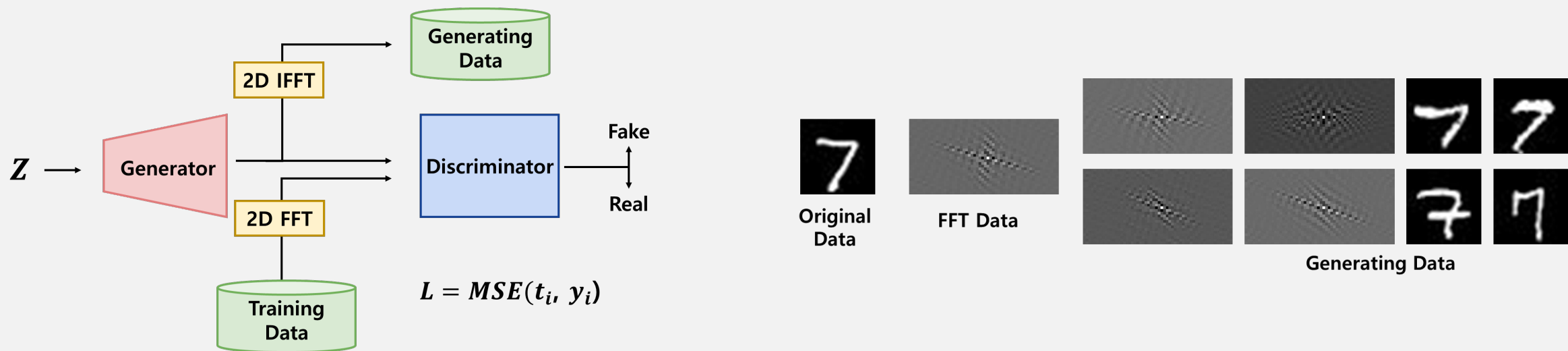


<이미지를 2D FFT를 통해 주파수 영역으로 변환>

이미지 공간보다 이미지의 주파수 영역에서의 분석이나 실험이 더 유용하다는 많은 연구

- Xu, Kai, et al. Learning in the frequency domain. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.2020.
 - 주파수 영역에 포함된 풍부한 정보가 모델의 **이미지 학습을 도울 뿐 아니라 이미지 정보를 더 잘 보존하고 모델의 정확도를 향상**
- Rippel, Oren, Jasper Snoek, and Ryan P. Adams. "Spectral representations for convolutional neural networks." Advances in neural information processing systems 28 (2015).
 - 이산 푸리에 변환(Fourier Transformation)이 딥러닝 컨볼루션 계산에서 **상당한 속도를 제공한다**는 연구 결과
 - 다양한 CNN 모델에서 이산 푸리에 변환을 수행하여 모델을 학습시키면 더 빠른 속도로 높은 정확도에 모델이 수렴
- Too, J., et al. "Classification of EMG signal based on time domain and frequency domain features." International Journal of Human and Technology Interaction (IJHaTI) 1.1 (2017)
 - 시간 영역(Time Domain)에서와 주파수 영역 각각의 데이터로 EMG 신호를 분류하는 모델 비교
 - **주파수 영역 데이터를 이용한 모델이 신호 분류에서 가장 높은 성능**
- 주파수 대역에 따라 **이미지의 정보 표현을 더욱 세밀하게 다룰 수 있음**
- 이미지를 필터링 하여 이미지의 잡음 제거 및 경계 검출 등을 통해 이미지 향상이 가능
- 이미지 경계 검출을 위한 Sobel 필터를 사용하면, 주파수 영역에서 컨볼루션을 사용하여 더욱 빠르게 처리가 가능하다는 장점

이에 본 논문에서는 주파수 영역을 생성모델 GAN에 적용시킬 수 있는 방법을 제안



• 학습 및 생성과정

- 모든 학습 데이터를 2D FFT를 사용하여 주파수 영역으로 변환
- 주파수 영역으로 변환된 데이터는 이미지 크기와 동일한 크기로 진폭(amplitude)과 위상(phase) 정보를 갖는 배열을 얻음
 - MNIST(28*28) 데이터는 진폭 28*28 데이터와 위상 28*28 데이터로 표현되며, 이를 합쳐 28*56 데이터로 저장
 - 컬러데이터의 경우, RGB 각각의 채널별로 주파수 데이터를 변환
- 주파수 영역으로 변환된 학습데이터를 학습한 후 새로운 주파수 영역의 데이터를 생성
- 생성된 데이터를 2D IFFT 하면 이미지 영역의 데이터를 얻을 수 있음

- **실험에 사용한 GAN: LSGAN**

- LSGAN 이후에 특화된 GAN은 주파수 영역에서의 효과를 보기 더 어려움

- **실험 데이터셋**

- MNIST, Fashion MNIST, MNIST-M, SVHN

- **성능평가**

- 품질을 평가하기 위하여 4개의 평가지표(FID, IS, P&R, D&C)를 사용하여 평가
- 모델의 초기 가중치를 바꿔가며 5번의 실험결과를 평균 낸 것
- GAN 학습 파라미터
 - Batch size = 128
 - learning rate = 0.0002
 - epochs = 1500
 - optimizer = Adam

• 생성된 이미지 비교

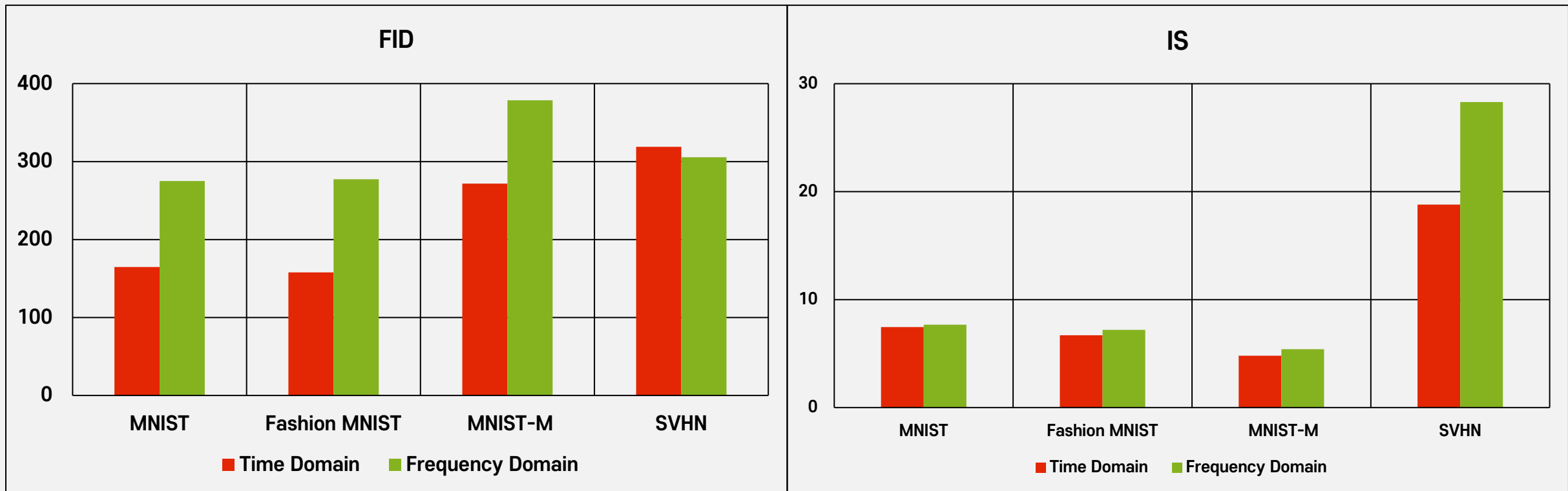
- LSGAN: 이미지 영역에서 학습 후 생성
- FDGAN: 주파수 도메인에서 학습 후 생성

Dataset	LSGAN	FDGAN
Mnist (28*28)		
Fashion Mnist (28*28)		
Mnist M (28*28)		
SVHN (32*32)		

SVHN의 경우 FDGAN이 월등히 좋음

- 평가지표(FID, IS)에서의 평가 결과

- FID 실험에서 MNIST, Fashion MNIST, MNIST-M 결과를 제외한 모든 실험에서 주파수 영역에서 GAN을 적용한 것이 더 좋은 성능

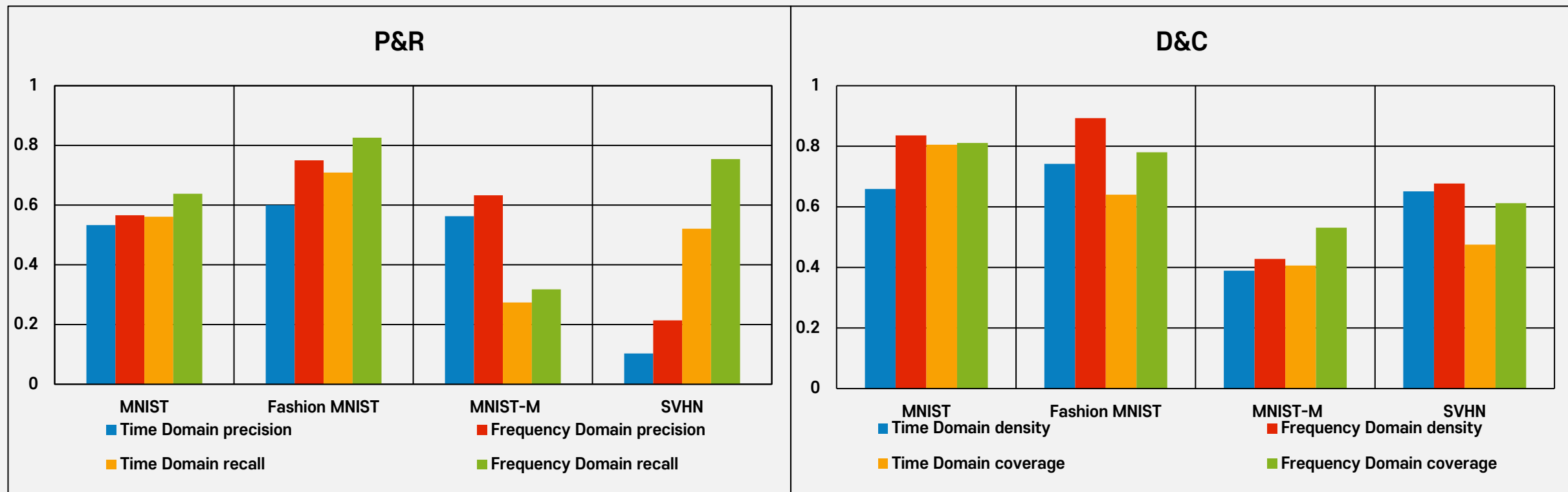


낮을수록 좋음

높을수록 좋음

• 평가지표(P&R, D&C)에서의 평가 결과

- 모든 실험에서 주파수 영역에서 GAN을 적용한 것이 더 좋은 성능

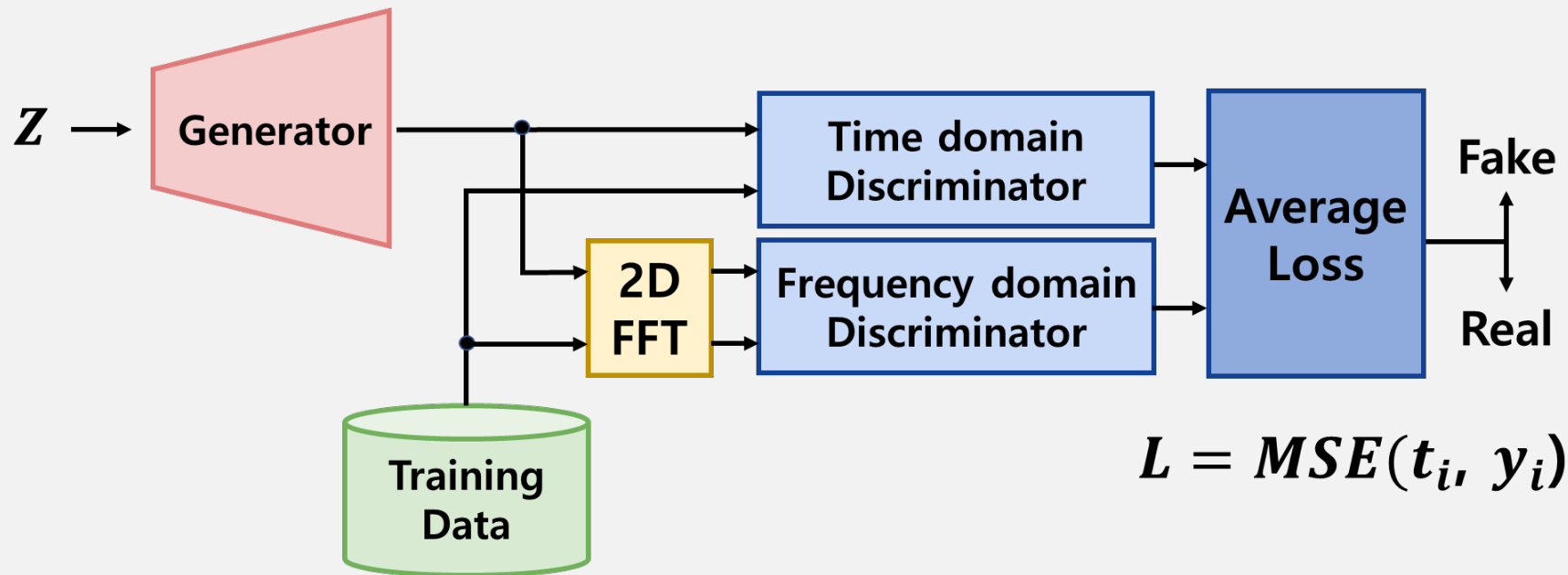


높을수록 좋음

높을수록 좋음

- 이미지 공간 판별자와 주파수 공간 판별자를 함께 사용하여 추가 실험

- Discriminator는 이미지 공간과 주파수 영역 모두에서 Real과 Fake를 판별
- 각 도메인에서의 판별 결과를 모두 합쳐 평균내어 학습에 사용



- GAN 학습 이미지를 2D FFT를 적용시켜 주파수 영역으로 변환하여 학습하고 주파수 영역에서 데이터를 생성한후 2D IFFT로 역변환 하여 이미지를 생성하는 방법 제안
- 4개의 데이터셋 (MNIST, Fashion MNIST, SVHN, MNIST-M)에서 4개의 평가지표(FID, IS,P&R,D&C) 로 평가한 결과,
FID의 MNIST, Fashion MNIST, MNIST-M 의 결과를 제외한 모든 실험에서 주파수 공간에서 생성된 이미지가 더 좋은 성능을 보임
- **이미지 영역 판별자와 주파수 영역 판별자를 동시에 사용하는 방법 초기 실험**
 - 두 판별자의 출력을 평균해서 손실을 계산
 - 이미지를 생성하는 것은 성공했으나 평가지표 실험은 완료하지 못함
 - 추가적으로 다양한 방법을 실험해서 더 좋은 방법을 고안하고 실험해야 함!!!