

Generative Advanced

21기 이소영

대표적인 이미지 생성 모델들

1. GAN
2. VAE
3. Flow-base model
4. Diffusion

각각의 개념, 구조, 차이점을 정리해서 작성하기

1. GAN (Generative Adversarial Network) 생성적 적대 신경망 :

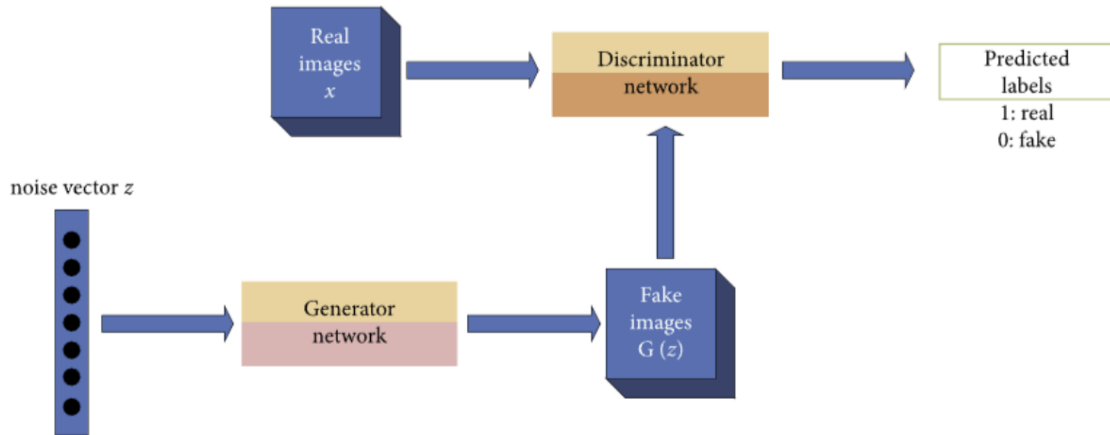
- 개념 : 이미지 생성을 위한 모델로, 가짜 이미지를 만들어내는 생성자(Forger)와 진위 여부를 판별하는 감별자(Detective)가 서로 경쟁하며 학습하는 방식으로 작동하는 모델. 매우 현실적인 이미지를 생성할 수 있도록 함.
- 작동 방식: 두 개의 딥러닝 모델을 이용하여 작동

1) 생성자 : 무작위 노이즈 벡터 z 와 같은 잠재 벡터를 입력으로 받아 가짜 이미지 데이터를 생성

2) 판별자 : 생성된 데이터와 실제 데이터를 입력으로 받아, 실제 이미지와 생성된 이미지를 구별해 내는 것이 목표

3) 두 모델이 서로 경쟁 관계를 맺어, 생성자는 판별자를 속이기 위해 더 실제같은 데이터를 생성하고, 판별자는 생성자가 만든 가짜 데이터를 구별해 내기 위해 학습.

4) 이 경쟁 과정을 반복하면서 생성자는 점점 더 현실적인 데이터를 생성할 수 있게 되고, 판별자는 가짜 데이터와 실제 데이터를 구별하는 능력을 향상시킴.



- 기존 딥러닝 분류모델과의 차이점: 기존의 딥러닝 분류 모델은 학습 데이터를 이용하여 입력 데이터를 분류하는 데에 목적이 있지만 GAN은 새로운 데이터를 생성하는 것이 목적. 또한, 생성자와 판별자라는 두 개의 모델을 이용하여 경쟁적으로 학습하는 방식을 사용

2. VAE (Variational Autoencoder) 변이형 오토인코더:

- 개념: VAE는 잠재 변수를 사용하여 데이터를 생성하는 확률적 생성 모델입니다. 잠재 변수는 데이터의 특징을 나타내는 확률 분포를 학습합니다.

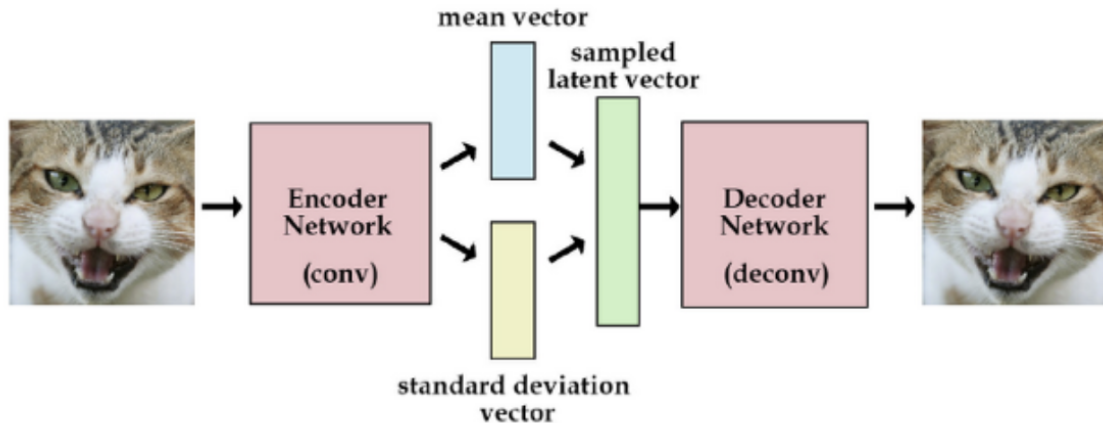
- 작동 방식:

1) 디코더는 인코더가 만들어낸 z 의 평균, 분산을 모수로 하는 정규분포를 전제하고 데이터의 사후확률을 학습

2) 이 사후확률 계산을 위해 다루기 쉬운 분포로 근사하는 variational inference (변분추론)을 사용

*변분추론 : $p(z|x)$ 와 $q(z)$ 사이의 KL-Divergence를 계산하고, KLD가 줄어드는 쪽으로 $q(z)$ 를 조금씩 업데이트 해서 $q(z)$ 를 얻어나는 과정

VAE (Variational Auto-Encoder)



- input x 가 만들어지는 확률 분포를 찾는 것이 목적

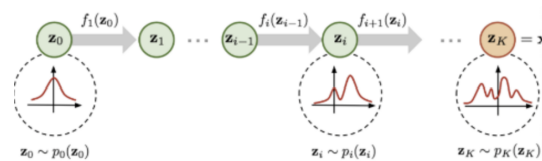
- 차이점:

- 1) 확률적인 접근 : 데이터의 특징을 나타내는 잠재 변수를 확률 분포로 모델링하여 데이터 생성 과정을 확률적으로 처리하는 것을 의미. 이는 데이터에 대한 불확실성을 고려할 수 있도록 함.
- 2) 잠재 변수 학습 : 데이터에 대한 잠재 변수의 확률 분포를 학습함으로써 데이터의 특징을 효과적으로 추출. 이는 데이터의 잠재적인 구조를 탐색하고 더욱 유의미한 특징을 추출하도록 함.
- 3) 생성과 인코딩을 동시에 수행 : 입력 데이터를 잠재 공간으로 encoding하는 과정과 잠재 공간의 샘플을 decoding하여 원본 데이터를 생성하는 과정을 동시에 수행함. 이는 데이터의 잠재적인 표현을 학습하는 동시에 데이터를 생성할 수 있는 장점을 제공.
- 4) 지도 및 비지도 학습 가능 : 비지도 학습에서는 입력 데이터만을 사용하여 잠재 변수를 학습하고 데이터를 생성할 수 있으며, 지도 학습에서는 입력 데이터와 관련된 라벨 정보를 함께 사용하여 잠재 변수를 학습하고 데이터를 생성 가능.

3. Flow-based model:

- 개념: 잠재변수 z 를 활용하여 $x = f(z)$ 와 잠재변수 z 에 대한 확률밀도함수 $p(z)$ 가 주어졌을 때, $p(x)$ 에 대해 최대한 비슷한 함수로 접근하는 것.
- 작동 방식 :

- 1) **역함수 변환** : 주어진 데이터의 확률 분포를 변환하는 함수를 학습. 이 함수는 주로 역함수로 구성되며, 입력 공간에서 출력 공간으로의 변환을 수행.
- 2) **연속적인 변환** : 연속적인 변환을 적용하여 입력 데이터를 잠재 공간으로 매핑. 이러한 변환은 단순하고 역함수가 쉽게 계산될 수 있는 함수들의 조합으로 이루어짐. (ex. Affine Coupling Layer 등의 변환 함수 이용)
- 3) **확률 밀도 함수 학습** : 변환된 데이터의 확률 분포를 학습함으로써 생성 모델로 사용. 이를 위해 변환된 데이터의 확률 밀도 함수를 모델링하고, 이 함수를 통해 샘플링된 데이터를 생성.
- 4) **훈련** : 모델은 학습 데이터셋에 대해 주어진 데이터의 확률 밀도 함수를 근사하는 것을 목표로 학습되므로, 확률 밀도 함수의 로그 우도를 최대화하는 방향으로 모델 파라미터를 업데이트. 학습 과정에서 역전파 알고리즘을 사용하여 모델의 파라미터를 업데이트.
- 5) **데이터 생성** : 잠재 공간에서 샘플링된 벡터를 디코딩하여 원본 데이터를 생성. 이 과정은 역변환 함수를 사용하여 변환된 데이터를 원래의 공간으로 복원하는 것을 포함.



- 차이점:

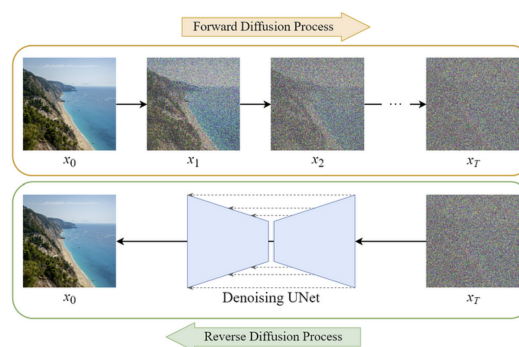
- 1) **확률적 생성**: 기존의 딥러닝 모델은 대부분 결정론적인 방식으로 데이터를 생성하는 반면, 데이터를 생성할 때 확률 밀도 함수를 모델링하고, 이를 통해 샘플링된 데이터를 생성. 이는 데이터의 확률적 특성을 고려하여 생성하는 것으로, 생성된 데이터가 확률적으로 다양한 특성을 가질 수 있음.
- 2) **연속적인 변환**: 기존의 딥러닝 모델은 비선형 함수를 사용하여 데이터를 생성하거나 변환하는 반면, 이 모델은 연속적인 변환으로 데이터를 생성하므로, 단순하며 역함수가 쉽게 계산될 수 있는 함수들의 조합으로 이루어짐.
- 3) **모델 구조**: Flow-based model은 연속적인 변환으로 이루어진 네트워크 구조를 사용함. 이러한 네트워크는 단순하며 역함수를 쉽게 계산할 수 있는 구조를 가지고 있음. 반면에 기존의 딥러닝 모델은 주로 다층 신경망과 같은 복잡한 구조를 사용하여 데이터를 생성함.

4. Diffusion:

- 개념: Diffusion은 이미지 생성을 위해 반복적으로 정규분포를 가진 Noise를 여러 단계에 걸쳐 추가하고, 여러 단계에 거쳐 제거함으로써, 입력 이미지와 유사한 확률 분포를 가진 결과 이미지를 생성하는 모델
- 작동 방식 :

Forward Diffusion Process : 이미지에 고정된(fixed) 정규 분포(=Gaussian분포)로 생성된 Noise가 더해짐

Reverse Diffusion Process : 이미지를 학습된(learned) 정규 분포로 생성된 Noise이미지로 빼짐



Forward → Reverse 단계를 거친 '결과 이미지'를 '입력 이미지'의 확률 분포와 유사하게 만드는 것. 이를 위해 Reverse단계에서, Noise 생성 확률 분포 Parameter인 평균과 표준편차를 업데이트하며 학습이 진행됨.

- 차이점:

- 1) 확률 분포를 모델링 : 입력과 출력 간의 직접적인 매핑을 학습하는 딥러닝 모형과 달리 데이터의 확률 분포를 모델링하여 샘플을 생성함.
- 2) 목적의 차이 : 기존 딥러닝 모델은 주로 분류, 회귀 등의 작업에 사용되지만, 이는 데이터 생성을 목표로 하며, 입력 데이터와 유사한 샘플을 생성함.
- 3) 활용 분야의 차이 : 주로 생성 모델링을 필요로 하는 이미지 생성, 자연어 생성, 음악 생성 등에 활용됨.

출처 :

<https://ai-inform.tistory.com/entry/생성모델-GAN이란-생성적-적대-신경망-GAN-쉬운-설명>

<https://xoft.tistory.com/32>