

GENERATIVE ADVERSARIAL NETWORKS

GAN

Deep Shark Lab 김태빈

목차

01. 생성 모델이란?

02. GAN 구조

03. GAN 문제점

04. GAN 종류

05. GAN 활용 및 사례

01 | 생성 모델이란?

Generative Adversarial Network,
GAN(생성적 적대 신경망)은 **생성 모델**의 한 예시

생성 모델

역할

데이터가 생성되는 확률 분포
 p_{data} 와 최대한 비슷한 모델
분포 p_{model} 을 생성하는 것

목표

판별 모델 속이는것, 즉 판별 모델이
생성된 샘플을 '실제(real)' 데이터
라고 판단하도록 만드는 것

판별 모델

역할

입력 데이터를 두 가지 클래스
real or fake로 분류

목표

입력 데이터가 실제 데이터인지, 아
니면 생성 모델이 만든 가짜 데이터
인지를 구분

사용 이유

고차원 분포 정복

이미지나 음성 데이터등 수많은 특징이 서로 복잡하게 연결되어 있는 데이터의 분포를 학습하여 생성

강화학습 고도화

생성 모델을 이용하면 실제 환경을 직접 반복해서 실행하지 않아도, 미래 상태나 가상의 환경을 만들어 시뮬레이션할 수 있음

이를 통해 에이전트가 다양한 상황을 미리 경험하면서 더 효율적으로 학습

준지도 학습

실제 데이터는 많지만, 정답이 붙어 있는 라벨 데이터는 부족한 경우,, 생성 모델은 라벨이 없는 데이터의 특징까지 학습하여, 적은 양의 라벨 데이터만으로도 모델의 성능을 높이는 데 도움

다양한 결과 생성

하나의 입력에 여러 개의 정답이 가능한 경우, 단순 평균으로 결과물을 만드는 대신 다양한 정답지 후보를 생성 가능

분류

확률 분포를 어떻게 다루느냐

Explicit Density Models

실제 데이터 분포를 수학적인 확률 밀도 함수

$p_{model}(x; \theta)$ 형태로 정의

Implicit density model

확률 밀도 함수를 직접 계산 X

무작위 노이즈 z 를 입력받아 어떤 연산(함수 G)을 거쳐 샘플 x 를 뱉어내는 과정

1. **마르코프 체인 방식**: 이전 샘플을 조금씩 수정하며 여러 단계를 거쳐 최종 샘플을 얻는 방식

2. **단일 단계(Single-step) 방식**: 노이즈 z 를 한 번의 통과만으로 바로 샘플 x 를 만드는 방식

"**GAN**은 암시적 모델 중에서도 **단일 단계**"

GAN 구조

데이터를 학습하여 **두 개의 신경망**(생성자와 판별자)을 서로 경쟁시켜,
실제 데이터와 구분이 안 되는 **가짜 데이터**를 만들어내는 학습 알고리즘

(Generator) G 가짜 데이터를 생성하는 생성자

(Discriminator) D 데이터가 실제 데이터인지 생성자가 만든 가짜 데이터인지 판별하는 판별자

z 랜덤 노이즈

x 실제 데이터

GAN 구조

판별자 D

생성자 G

$$x \rightarrow D(x)$$

$$D(x) \rightarrow 1$$

$$z \rightarrow G(z) \rightarrow D(G(z))$$

$$D(G(z)) \rightarrow 0$$

$$D(G(z)) \rightarrow 1$$

→ 실제 데이터일 확률

GAN 학습 과정

목적 함수

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

실제 데이터를 실제라고 판별 가짜 데이터를 가짜라고 판별

이상적인 학습 상태

$$p_g = p_{data}$$

$$D(x) = 0.5$$

GAN 문제점

1. Training Instability

판별자가 지나치게 강해짐 → 생성자의 gradient 감소 → 학습 불안정

2. Mode collapse

$z_1, z_2, z_3 \rightarrow \text{비/슷 한 } G(z)$

- 생성 품질은 괜찮아 보일 수 있지만 생성 결과의 다양성이 부족
- 실제 데이터 분포 전체를 학습하지 못함

3. Evaluation Difficulty

GAN의 loss만으로 생성 품질을 판단하기 어려움

- 손실값과 이미지 품질이 직접 비례하지 않음
- 사람의 시각적 평가는 주관적일 수 있음

GAN 발전 흐름

Vanilla GAN → DCGAN → cGAN → WGAN → CycleGAN → StyleGAN

주요 발전 방향을 대표

Vanilla GAN (Original GAN, 2014)

아무 확장도 하지 않은 가장 기본적인 원조 GAN

DCGAN

등장 배경

Vanilla GAN의 기본 구조(MLP) → 이미지 처리에 강한 CNN 구조를 적용

특징

- 생성자 : 랜덤 노이즈 → 작은 특징맵 → 더 큰 특징맵 → 최종 이미지
- 판별자 : 입력 이미지 → 특징 추출 → 크기 축소 → 실제/가짜 판단
- Pooling 대신 stride convolution 사용 → 신경망이 직접 학습하게 함

장단점

- Vanilla GAN보다 이미지의 공간적 특징을 잘 학습
- 생성 이미지뿐 아니라 비지도 특징 표현으로도 활용 가능
- 학습 불안정성이 완전히 해결되지는 않음
- 고해상도·복잡한 이미지 생성에는 한계

cGAN

등장 배경

Vanilla GAN은 랜덤 노이즈만 받음 → 조건 정보 y 를 추가해 원하는 클래스나 속성에 맞는 데이터를 생성

특징

기존 GAN $G(z)$ cGAN $G(z, y)$

y 는 클래스 라벨이나 추가 정보

- 생성자: 노이즈 z 와 조건 y 를 함께 입력
- 판별자: 데이터 x 또는 $G(z, y)$ 를 함께 입력
- 주어진 조건과 생성 결과가 일치하는지 학습

장단점

- 원하는 클래스나 속성의 데이터 생성 가능
- 이미지 라벨, 텍스트, 다른 이미지 등 다양하게 확장 가능
- 조건 라벨이나 추가 정보가 필요함
- 조건 종류가 많거나 복잡하면 학습 난도가 증가함

WGAN

등장 배경

기존 GAN의 Training Instability, Mode collapse, Evaluation Difficulty 문제

특징

Critic은 확률을 출력하지 않고 데이터에 점수 부여 후 얼마나 가까운지 판별

Weight clipping:
Critic의 가중치가 일정 범위를 벗어나지 않도록 강제로 제한하는 방법

장단점

- Training Instability 향상
- Mode collapse 완화
- Loss와 생성 품질의 관계가 비교적 명확
- Weight clipping 필요 → 표현력 ↓
- 적절한 clipping 범위 설정이 어려움
- Critic 반복 학습으로 계산량 증가

CycleGAN

등장 배경

기존 이미지 변환 모델은 입력 이미지와 정답 이미지가 짝지어진 한 쌍의 데이터가 필요
→ 대응하는 이미지 쌍 없이도, 두 이미지 집합만으로 변환 관계를 학습

특징

$$G : X \rightarrow Y$$
$$F : Y \rightarrow X$$

$$x \rightarrow G(x) \rightarrow F(G(x)) \approx x$$

Cycle Consistency Loss 이용
원본 이미지를 복원하도록 학습

장단점

- Paired data 불필요
- 데이터 수집 부담 감소
- 큰 형태 변화에 취약
- 다양한 출력 생성에 한계
- 학습 구조와 계산량 증가

StyleGAN

등장 배경

기존 GAN도 사람 얼굴 같은 복잡한 이미지를 생성 할수 있었지만,
이미지 품질 한계 , 세부 속성 조절 난도, 잠재 벡터 z 와 이미지 특징 해석 어려움

특징

랜덤 벡터 z

→ Mapping Network

→ 중간 잠재벡터 w

→ 각 생성층에 스타일 입력

장단점

- 고해상도·사실적인 이미지 생성
- 큰 구조와 세부 특징을 제어
- 서로 다른 이미지 특징 조합 가능
- 구조와 학습 과정이 복잡함
- 높은 계산량과 많은 학습 데이터 필요

GAN 모델별 주요 활용

활용 분야	관련 모델	활용 사례
이미지 생성	DCGAN	학습 데이터와 유사한 새로운 이미지 생성
조건부 생성	cGAN	클래스·라벨 등 지정된 조건에 맞는 이미지 생성
안정적인 생성 학습	WGAN	Wasserstein distance를 활용해 GAN의 학습 안정성 개선
이미지 변환	CycleGAN	말↔얼룩말, 계절·화풍 등 서로 다른 이미지 영역 변환
고품질 이미지 생성	StyleGAN	사실적인 가상 인물 생성 및 이미지 속성 제어

GAN의 대표적인 실제 활용 사례

NVIDIA Canvas

cGAN → GauGAN

사용자가 입력한 스케치나 영역 정보를
조건으로 사실적인 풍경 이미지를 생성

간단한 그림 → 사실적 풍경

GFPGAN

StyleGAN 얼굴 생성 사전지식 → GFPGAN

학습한 얼굴 구조 정보를 활용해
흐릿하거나 손상된 얼굴을 복원

저화질 얼굴 → 고화질로 복원 얼굴



GENERATIVE ADVERSARIAL NETWORKS

발표를 마치겠습니다

Deep Shark Lab 김태빈