

DDPM: Denoising Diffusion Probabilistic Models 요약

(디퓨전 모델의 중심논문)

1. 서론

본 논문에서는 확산(diffusion) 기반 생성 모델의 새로운 접근법인 Denoising Diffusion Probabilistic Models (DDPM)을 제안합니다. 기존 GAN이나 VAE와 같은 생성 모델들이 가진 학습 불안정성 및 모드 붕괴 문제를 극복하기 위해, 연구진은 데이터에 점진적으로 노이즈를 추가하는 forward diffusion process와 이를 역전시키는 reverse diffusion process를 도입하였습니다. 이 방식은 단순한 확률적 프로세스를 활용하면서도 안정적이고 고품질의 샘플을 생성할 수 있음을 실험적으로 입증하였습니다.

2. 배경 및 관련 연구

2.1 기존 생성 모델의 한계

기존 생성 모델들은 복잡한 데이터 분포를 직접 모사하려는 시도에서 학습의 불안정성, 수렴 문제, 그리고 생성된 샘플의 다양성 부족 등의 문제를 보였습니다. 특히, GAN은 적대적 학습 구조로 인한 학습 불안정성이, VAE는 낮은 해상도와 과도한 분산으로 인한 품질 저하 문제가 대표적입니다.

2.2 확산 모델의 등장

확산 모델(Diffusion Models)은 데이터에 점진적으로 노이즈를 추가한 후 이를 역전시켜 원본 데이터를 복원하는 방식을 취합니다. 이러한 접근은 자연 이미지의 연속적인 변환 과정을 모방함으로써, 보다 자연스럽게 다양한 샘플을 생성할 수 있는 가능성을 보여주었습니다. 본 논문에서는 확률론적 프레임워크 내에서 forward diffusion과 reverse diffusion 과정을 명확히 수학적으로 모델링하여, 안정적 학습과 뛰어난 생성 성능을 도모하였습니다.

3. 모델 개요 및 방법론

3.1 Forward Diffusion Process (노이즈 생성)

논문에서는 우선 forward diffusion process를 다음과 같이 정의합니다.

- 실제 데이터 x_0 에 대해, 단계 t 마다 가우시안 노이즈를 점진적으로 추가합니다.
- 각 단계 t 에서 데이터 x_t 는 다음과 같이 표현됩니다:

$$q(x_t | x_{t-1}) = N(x_t; \sqrt{(1 - \beta_t)} * x_{t-1}, \beta_t * I)$$

여기서 β_t 는 t 단계에서 추가되는 노이즈의 분산을 의미하며, 이 값은 일정한 스케줄이나 학습된 스케줄을 사용할 수 있습니다. 이 과정은 Markov chain으로 모델링되며, t 가 커짐에 따라 $q(x_T | x_0)$ 는 점점 단순한 가우시안 분포로 수렴합니다.

3.2 Reverse Diffusion Process (디노이저)

forward process에서 추가된 노이즈를 제거하는 reverse diffusion process는 학습 가능한 모델 $p(\theta)$ 에 의해 근사됩니다. 이 과정은 다음과 같이 정의됩니다:

$$p(\theta)(x_{t-1} | x_t) = N(x_{t-1}; \mu(\theta)(x_t, t), \Sigma(\theta)(x_t, t))$$

즉, 각 단계에서 x_t 로부터 x_{t-1} 를 생성하는 확률적 모델이며, 여기서 평균 $\mu(\theta)$ 와 공분산 $\Sigma(\theta)$ 은 네트워크 파라미터 θ 에 의해 결정됩니다. 연구진은 네트워크가 예측하는 “노이즈 항” 또는 “잔여값”을 활용하여 $\mu(\theta)$ 를 계산하는 방식으로 접근하였으며, 이는 Denoising Autoencoder의 개념과 유사합니다.

3.3 학습 목표 및 최적화

논문의 주요 기여 중 하나는 변분 하한(variational lower bound, VLB)를 활용한 학습 목표입니다. 전체 모델의 로그 가능도($\log p(\theta)(x_0)$)에 대해 변분 하한을 도출하고, 이를 최소화하는 방향으로 학습을 진행합니다.

- 각 시간 단계 t 에 대해 KL 발산(KL divergence)을 최소화하는 것이 목표입니다.
- 이로부터 도출되는 손실 함수는 다음과 같이 표현됩니다:

$$L = \sum_{t=1}^T E(q(x_t | x_0)) [KL(q(x_{t-1} | x_t, x_0) || p(\theta)(x_{t-1} | x_t))]$$

또한, 연구진은 특정 조건 하에서 이 손실 함수를 단순화하여, 모델이 직접 노이즈를 예측하도록 재구성할 수 있음을 보였습니다. 이 접근은 학습을 보다 효율적이고 안정적으로 만듭니다.

3.4 모델의 장점 및 차별성

- **안정적인 학습:** 적대적 학습 구조가 없으므로 학습 과정에서 발생할 수 있는 불안정성이 현저히 줄어듭니다.
- **고품질 샘플 생성:** 역확산 과정을 통해 점진적으로 노이즈를 제거하며 데이터의 세밀한 특성을 복원함으로써, 생성된 이미지의 품질이 우수합니다.
- **모델 해석 용이성:** 확산 과정과 역확산을 명확한 수학적 프레임워크로 설명하여, 모델 내부 작동 원리를 이해하기에 용이합니다.

4. 실험 및 평가

연구진은 CIFAR-10, LSUN(Large-scale Scene UNDERstanding Challenge)등 다양한 데이터셋을 대상으로 DDPM의 성능을 평가하였습니다.

4.1 성능 비교



Figure 3: LSUN Church samples. FID=7.89



Figure 4: LSUN Bedroom samples. FID=4.90

Algorithm 3 Sending x_0

```

1: Send  $x_T \sim q(x_T|x_0)$  using  $p(x_T)$ 
2: for  $t = T - 1, \dots, 2, 1$  do
3:   Send  $x_t \sim q(x_t|x_{t+1}, x_0)$  using  $p_\theta(x_t|x_{t+1})$ 
4: end for
5: Send  $x_0$  using  $p_\theta(x_0|x_1)$ 

```

Algorithm 4 Receiving

```

1: Receive  $x_T$  using  $p(x_T)$ 
2: for  $t = T - 1, \dots, 1, 0$  do
3:   Receive  $x_t$  using  $p_\theta(x_t|x_{t+1})$ 
4: end for
5: return  $x_0$ 

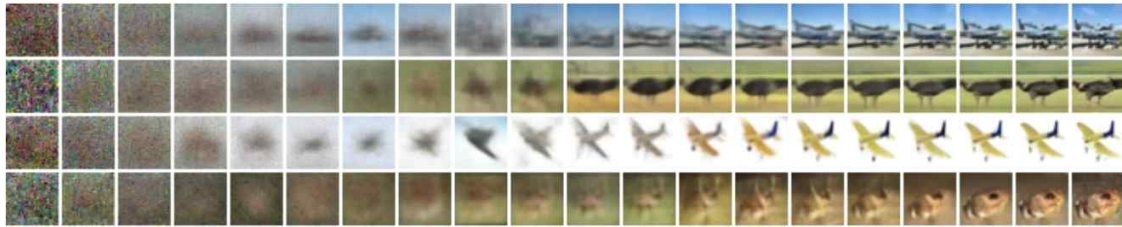
```

- **비교 대상:** 기존의 GAN, VAE 및 기타 score-based generative models와 비교할 때, DDPM은 샘플의 다양성과 세밀한 구조 보존 측면에서 우수한 성능을 보였습니다.
- **정량적 평가:** FID(Frechet Inception Distance) 등의 지표에서 경쟁력 있는 결과를 나타내었으며, 점진적 노이즈 제거 과정 덕분에 생성된 이미지의 품질이 뚜렷하게 개선됨을 확인할 수 있었습니다.

4.2 모델의 효율성

- **샘플링 시간:** 역확산 과정을 거쳐 여러 단계에 걸쳐 샘플을 생성해야 하므로, 연산 비용이 다소 높은 단점이 있습니다.
- **최적화 및 개선 가능성:** 후속 연구에서는 샘플링 단계를 단축할 수 있는 가속화 기법 및 학습된 스케줄 조정 등 다양한 개선 전략을 모색하고 있습니다.

4.3 조건부 생성 실험



(무조건적 CIFAR10 진행형 생성(시간 경과에 따른 \hat{X}_0 (reverse diffusion에 따른 결과), 왼쪽에서 오른쪽으로).

논문에서는 조건부 생성 작업에 대해서도 모델의 효과를 입증하였습니다. 예를 들어, 특정 클래스나 속성을 지정한 조건 하에서 샘플을 생성할 때도 안정적인 성능을 보임으로써, 다양한 응용 분야로의 확장이 가능함을 보여주었습니다.

5. 결론 및 향후 연구

논문의 결론에서는 DDPM이 기존 생성 모델의 한계를 극복할 수 있는 효과적인 대안임을 강조합니다.

- **모델의 강점:** 확산과 역확산 과정을 통해 데이터의 세밀한 구조와 분포를 효과적으로 복원하여 고품질 샘플을 생성할 수 있습니다.
- **학습 안정성:** 적대적 학습 구조가 배제된 변분 추론 기반 학습 방식으로 인해, 학습 과정의 안정성이 크게 향상되었습니다.
- **향후 연구 방향:** 샘플링 속도 개선, 다양한 도메인(영상, 음성 등)으로의 확장, 조건부 생성 작업에 대한 추가 연구가 필요하며, 확산 모델과 다른 생성 모델 간의 하이브리드 접근법도 모색될 수 있습니다.

종합 평가

본 논문은 생성 모델 분야에 새로운 연구 방향을 제시하며, 기존 모델들의 한계를 효과적으로 극복할 수 있는 학률론적 접근 방식을 도입하였습니다.

- **이론적 기여:** 명확한 수학적 모델링을 통해 확산 및 역확산 과정을 통합적으로 설명함으로

써, 모델 내부 메커니즘에 대한 깊은 이해를 돕습니다.

- **실험적 검증:** 다양한 데이터셋에 대한 정량적, 정성적 평가를 통해 모델의 우수한 샘플 품질과 안정성을 입증하였습니다.
- **응용 가능성:** 조건부 생성 및 다양한 도메인으로의 확장이 가능함을 보여주어, 실무 적용 및 후속 연구에 중요한 참고 자료가 될 것으로 판단됩니다.

<https://arxiv.org/pdf/2006.11239>

Denoising Diffusion Probabilistic Models (DDPM): 디퓨전모델 논문

<https://jang-inspiration.com/ddpm-2>

해당 논문 리뷰, 요약 블로그