High-Resolution Image Synthesis with Latent Diffusion Models

latent diffusion model 을 사용한 고해상도 이미지 합성

✅ 기존 diffusion 모델 = **픽셀 공간에서 직접 작동**하기 때문에, 강력한 DM을 최적화하는 데에는 종종 **수백 GPU-days**가 소요되며, **순차적 평가**로 인해 추론 또한 비용이 큼

⬇️

LDM = **잠재 공간(latent space)** 에서 확산 모델을 적용. **복잡도 감소와 세부 보존 간의 준최적 지점**에 도달할 수있게 됨.

➡️ 인페인팅과 클래스 조건 이미지 합성에서 새로운 최신 성능을 달성하며, 텍스트-투-이미지 합성, 무조건 이미지 생성, 초해상도(super-resolution) 등 다양한 작업에서도 매우 경쟁력 있는 성능을 보인다

Intro

기존 모델과 비교

- **AR Transformer**
  - 장점: 표현력
  - 단점: 수십억 파라미터, 확장성 최악
- **GAN**
  - 장점: 빠름
  - 단점: mode collapse, multimodal 분포에 약함
- **Diffusion**
  - 장점: 안정적, mode-covering

- 단점: **너무 비쌈**

→ Diffusion 이 성능적으로 가장 좋으나 현실적으로 너무 비싸서 쓰기 힘듦.

Democratizing High-Resolution Image Synthesis

- Democratizing : 민주화 = **연구실/기업 독점 탈피**
- diffusion model 의 **모드 커버링 특성(mode-covering behavior)** 은 데이터의 지각 불가능한(imperceptible) 세부 정보까지 모델링하는 데 과도한 용량(즉, 계산 자원)을 소비하게 만든다.
  - reweighted variational objective 의 경우도 초기 잡음 제거 단계의 샘플링을 줄임으로써 이 문제를 완화하려 하지만, 확산 모델은 여전히 **RGB 이미지의 고차원 공간에서 반복적인 함수 평가와 그래디언트 계산**이 필요하기 때문에 계산적으로 매우 부담

→ 모델의 접근성을 높이면서 동시에 자원 소비를 줄이기 위해서는, **학습과 샘플링 모두에서 계산 복잡도를 줄이는 방법**이 필요.

Departure to Latent Space

모든 우도 기반 모델과 마찬가지로

1. perceptual compression - 고주파 세부 정보를 제거하면서도 의미적 변동을 학습
2. 실제 생성 모델이 데이터의 **의미적·개념적 구성(semantic compression)** 을 학습

> 목표 - **지각적으로 동등하면서도 계산적으로 훨씬 효율적인 공간**을 먼저 찾고, 그 공간에서 고해상도 이미지 합성을 위한 확산 모델을 학습하는 것

먼저, 데이터 공간과 지각적으로 동등한 **저차원 표현 공간**을 제공하는 오토인코더를 학습시키기

- 이전 연구 [23, 66]와 달리, **과도한 공간 다운샘플링에 의존할 필요가 없다**. 학습된 잠재 공간에서의 확산 모델은 공간 차원에 대해 더 나은 스케일링 특성을 보이기 때문

→ 잠재 공간에서 **단일 네트워크 패스만으로도 효율적인 이미지 생성**을 가능하게 함
→ 이러한 결과 모델을 **잠재 확산 모델(Latent Diffusion Models, LDMs)** 이라고 부름

⇒ **보편적 오토인코더를 한 번만 학습하면**, 이를 여러 확산 모델 학습이나 전혀 다른 작업 탐색에 재사용할 수 있다.

Contributions

(i) 순수 트랜스포머 기반 접근법 [23, 66]과 달리, 본 방법은 고차원 데이터로 더 우아하게 확장되며,

(a) 이전 연구보다 더 충실하고 세밀한 복원을 제공하는 압축 수준에서 작동하고 (그림 1 참조),

(b) 효율적으로 학습될 수 있다.

(ii) 이미지 합성, 인페인팅, 확률적 초해상도 등 여러 작업과 데이터셋에서 경쟁력 있는 성능을 달성하면서도 계산 비용을 크게 낮춘다. 픽셀 기반 확산 접근법과 비교할 때, 추론 비용 역시 현저히 감소한다.

(iii) 이전 연구 와 달리, 본 접근법은 인코더/디코더 아키텍처와 score-based prior를 동시에 학습하지 않으며, 복원 성능과 생성 성능 사이의 섬세한 가중 조절이 필요하지 않다.

- 이는 매우 충실한 복원을 보장하며, 잠재 공간에 대한 정규화 요구도 매우 작다.

(iv) 초해상도, 인페인팅, 의미 합성과 같이 조밀한 조건이 필요한 작업에서도, 본 모델은 **합성곱 방식으로 적용 가능**하며 약 **1024^2 픽셀** 크기의 크고 일관된 이미지를 생성할 수 있다.

(v) 우리는 cross-attention을 기반으로 한 **범용 조건 메커니즘**을 설계하여, 멀티모달 학습을 가능하게 한다. 이를 이용해 클래스 조건, 텍스트-투-이미지, 레이아웃-투-이미지 모델을 학습한다.

(vi) 마지막으로, 우리는 **사전학습된 잠재 확산 및 오토인코딩 모델**을 공개하며, 이는 DM 학습 외의 다양한 작업에도 재사용될 수 있다

Methods

고해상도 이미지 합성 시 확산 모델이 대응되는 손실 항을 언더샘플링함으로써 지각적으로 무관한 세부 사항을 무시할 수 있게 됨. 그럼에도 불구하고 여전히 **픽셀 공간에서 비용이 큰 함수 평가**를 필요로 하며, 이것이 막대한 계산 시간과 에너지 자원 요구를 초래.

> **압축 학습 단계와 생성 학습 단계를 명시적으로 분리하자**

(i) 고차원 이미지 공간을 벗어남으로써, 샘플링이 저차원 공간에서 수행되므로 계산적으로 훨씬 효율적인 확산 모델을 얻을 수 있음

(ii) U-Net 아키텍처 에서 상속된 확산 모델의 **귀납적 편향(inductive bias)** 을 활용하는데, 이는 공간 구조를 가진 데이터에 특히 효과적이며, 이전 접근법들 에서 필요했던 공격적이고

품질을 저하시킬 수 있는 압축 수준의 필요성을 완화한다.

(iii) **범용 압축 모델** 을 얻게 되며, 이 모델의 잠재 공간은 여러 생성 모델을 학습하는 데 사용될 수 있을 뿐 아니라, 단일 이미지 CLIP 기반 합성과 같은 다른 다운스트림 응용에도 활용될 수 있다.

3.1 Perceptual Image Compression

Perceptual Image Compression model = perceptual loss + 패치 기반 적대적 목적 함수 의 결합으로 학습된 오토인코더

- 픽셀 공간 손실(L_2 또는 L_1 목적 함수)에만 의존할 때 발생하는 블러 현상을 피하고, **지역적 사실성(local realism)** 을 강제함으로써 복원이 이미지 매니폴드 내에 머물도록 보장.

보다 구체적으로, RGB 공간에서 이미지 $x \in \mathbb{R}^{H\times W\times 3}$ 가 주어지면, 인코더 $\mathcal{E}$ 는 이를 잠재 표현

$$z = \mathcal{E}(x)$$

으로 인코딩하고, 디코더 $\mathcal{D}$ 는 잠재 공간으로부터 이미지를 복원하여

$$\tilde{x} = \mathcal{D}(z) = \mathcal{D}(\mathcal{E}(x))$$

를 얻는다. 여기서

$$z \in \mathbb{R}^{h\times w\times c}$$

이다.

중요하게도, 인코더는 이미지를 비율

$$f = H/h = W/w$$

만큼 다운샘플링하며, 우리는 $f = 2^m$ ($m \in \mathbb{N}$) 형태의 다양한 다운샘플링 비율을 실험한다.

- 정규화 방식
  1. **KL-reg : 학습된 잠재 공간이 표준 정규 분포를 따르도록 약한 KL 패널티를 부과. VAE 와 유사**
  2. VQ-reg: 디코더 내부에 벡터 양자화 계층 을 사용하며, 이 모델은 VQGAN 으로 해석될 수 있다.
- 이후 학습되는 확산 모델은 잠재 공간 z=E(x) 의 **2차원 구조** 를 활용하도록 설계되었기 때문에, 비교적 완만한 압축 비율만으로도 매우 우수한 복원 성능을 달성할 수 있다.

Latent Diffusion Models

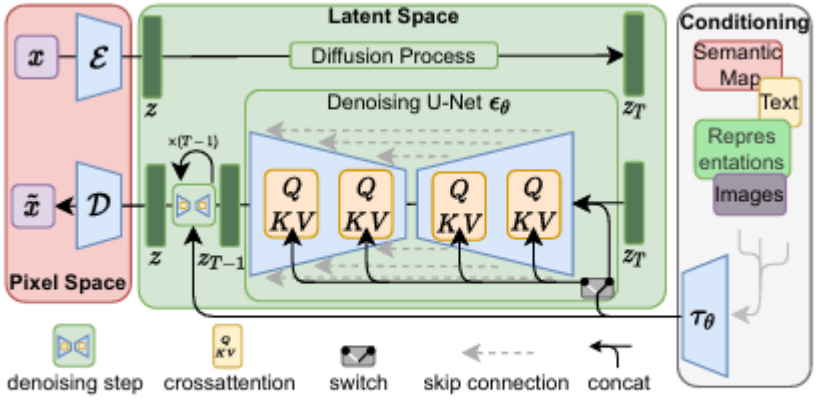

Figure 3. We condition LDMs either via concatenation or by a more general cross-attention mechanism. See Sec. 3.3

diffusion model = 정규 분포를 따르는 변수로부터 점진적으로 노이즈를 제거함으로써 데이터 분포 p(x) 를 학습하는 확률적 모델

- 이미지 합성에서 가장 성공적인 모델들 [15, 30, 72]은 p(x) 에 대한 변분 하한의 재가중 버전을 사용하며, 이는 **score matching** 과 동일한 구조를 가진다.
- 입력 x_t 에 대해 노이즈가 제거된 출력을 예측하도록 학습된 동일 가중치의 잡음 제거 오토인코더

$$\epsilon_\theta(x_t, t), \quad t = 1, \ldots, T$$

  - 의 연속으로 해석 가능
- 이에 대응하는 목적 함수는 (Sec. B에서) 다음과 같이 단순화될 수 있다

$$L_{DM} = \mathbb{E}_{x,\epsilon\sim\mathcal{N}(0,1),t}\left[\|\epsilon - \epsilon_\theta(x_t, t)\|_2^2\right],$$

  - t = {1,....,T} 에서 균등 샘플링

Conditioning Mechanisms

확산 모델은 원칙적으로 p(z ∣ y) 형태의 조건부 분포를 모델링할 수 있다.

- 조건부 잡음 제거 오토인코더

$$\epsilon_\theta(z_t, t, y)$$

를 사용하여 구현할 수 있으며, 텍스트 , 시맨틱 맵, 또는 기타 이미지-투-이미지 변환 작업 과 같은 입력 y 를 통해 합성 과정을 제어할 수 있다

cross-attention 메커니즘 을 통해 U-Net 백본을 확장함으로써, 확산 모델을 더 유연한 조건부 이미지 생성기로 만들기

- **다양한 모달리티의 조건 y** 를 전처리하기 위해

  도메인 특화 인코더 τ_θ 를 도입하여

$$\tau_\theta(y) \in \mathbb{R}^{M \times d_\tau}$$

  형태의 중간 표현을 생성하고, 이를 cross attention 계층을 통해 U-NET 의 중간 레이어에 주입

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d}}\right) V,$$

$$Q = W_Q^{(i)} \cdot \varphi_i(z_t), \quad K = W_K^{(i)} \cdot \tau_\theta(y), \quad V = W_V^{(i)} \cdot \tau_\theta(y).$$

→ **조건부 LDM 학습**

$$L_{LDM} := \mathbb{E}_{\mathcal{E}(x), y, \epsilon \sim \mathcal{N}(0,1), t}\left[\|\epsilon - \epsilon_\theta(z_t, t, \tau_\theta(y))\|_2^2\right],$$

Experiment

1. On Perceptual Compression Tradeoffs

- 다운샘플링 계수 를 달리한 LDM들의 동작을 분석한다
  - f∈{1,2,4,8,16,32}

f	$\|\mathcal{Z}\|$	c	R-FID ↓	R-IS ↑	PSNR ↑	PSIM ↓	SSIM ↑
16 *VQGAN* [23]	16384	256	4.98	–	19.9 ±3.4	1.83 ±0.42	0.51 ±0.18
16 *VQGAN* [23]	1024	256	7.94	–	19.4 ±3.3	1.98 ±0.43	0.50 ±0.18
8 *DALL-E* [66]	8192	-	32.01	–	22.8 ±2.1	1.95 ±0.51	0.73 ±0.13
32	16384	16	31.83	40.40 ±1.07	17.45 ±2.90	2.58 ±0.48	0.41 ±0.18
16	16384	8	5.15	144.55 ±3.74	20.83 ±3.61	1.73 ±0.43	0.54 ±0.18
8	16384	4	1.14	201.92 ±3.97	23.07 ±3.99	1.17 ±0.36	0.65 ±0.16
8	256	4	1.49	194.20 ±3.87	22.35 ±3.81	1.26 ±0.37	0.62 ±0.16
4	8192	3	0.58	224.78 ±5.35	27.43 ±4.26	0.53 ±0.21	0.82 ±0.10
4†	8192	3	1.06	221.94 ±4.58	25.21 ±4.17	0.72 ±0.26	0.76 ±0.12
4	256	3	0.47	223.81 ±4.58	26.43 ±4.22	0.62 ±0.24	0.80 ±0.11
2	2048	2	0.16	232.75 ±5.09	30.85 ±4.12	0.27 ±0.12	0.91 ±0.05
2	64	2	0.40	226.62 ±4.83	29.13 ±3.46	0.38 ±0.13	0.90 ±0.05
32	KL	64	2.04	189.53 ±3.68	22.27 ±3.93	1.41 ±0.40	0.61 ±0.17
32	KL	16	7.3	132.75 ±2.71	20.38 ±3.56	1.88 ±0.45	0.53 ±0.18
16	KL	16	0.87	210.31 ±3.97	24.08 ±4.22	1.07 ±0.36	0.68 ±0.15
16	KL	8	2.63	178.68 ±4.08	21.94 ±3.92	1.49 ±0.42	0.59 ±0.17
8	KL	4	0.90	209.90 ±4.92	24.19 ±4.19	1.02 ±0.35	0.69 ±0.15
4	KL	3	0.27	227.57 ±4.89	27.53 ±4.54	0.55 ±0.24	0.82 ±0.11
2	KL	2	0.086	232.66 ±5.16	32.47 ±4.19	0.20 ±0.09	0.93 ±0.04

Table 8. Complete autoencoder zoo trained on OpenImages, evaluated on ImageNet-Val. † denotes an attention-free autoencoder.

- LDM들이 사용하는 1단계 모델의 하이퍼파라미터와 재구성 성능

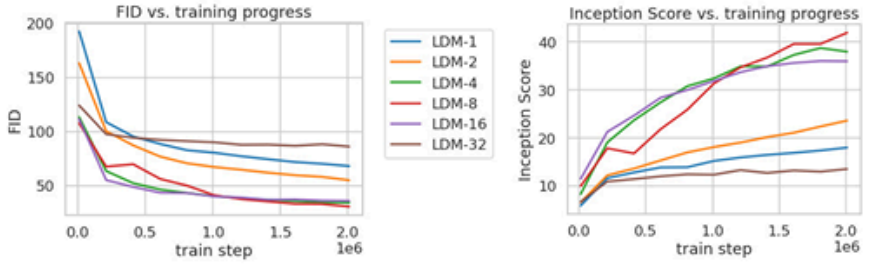

Figure 6. Analyzing the training of class-conditional *LDMs* with different downsampling factors f over 2M train steps on the ImageNet dataset. Pixel-based *LDM-1* requires substantially larger train times compared to models with larger downsampling factors (*LDM-*$\{$*4-16*$\}$). Too much perceptual compression as in *LDM-32* limits the overall sample quality. All models are trained on a single NVIDIA A100 with the same computational budget. Results obtained with 100 DDIM steps [84] and $\kappa = 0$.

- ImageNet [12] 데이터셋에서 클래스 조건 LDM의 학습 진행(2M 스텝)에 따른 샘플 품질

- 작은 다운샘플링 계수(LDM-{1,2})는 학습 속도가 느린 반면, 과도하게 큰 다운샘플링 계수는 동일한 학습 스텝 수 이후 **충실도 저하(stagnating fidelity)** 를 초래한다.

  → (i) 지각적 압축의 대부분을 확산 모델에 맡긴 경우와 (ii) 지나치게 강한 1단계 압축으로 인해 정보 손실이 발생한 경우에 기인

- LDM-{4–16}은 효율성과 지각적으로 충실한 결과 사이에서 좋은 균형을 이루며, 픽셀 기반 확산(LDM-1) 대비 2M 학습 스텝 후 **FID가 약 38 포인트 개선**되는 결과

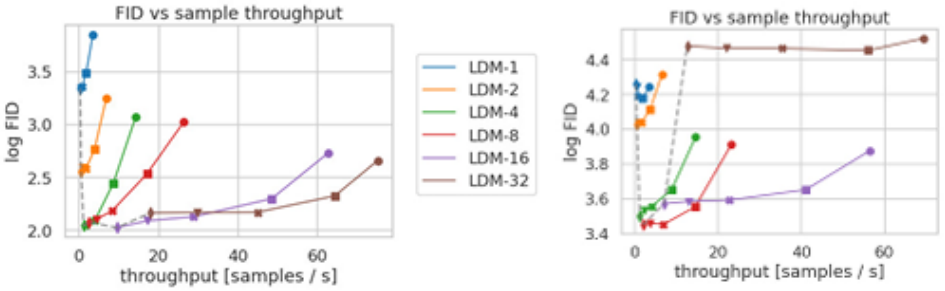

Figure 7. Comparing *LDMs* with varying compression on the CelebA-HQ (left) and ImageNet (right) datasets. Different markers indicate $\{10, 20, 50, 100, 200\}$ sampling steps using DDIM, from right to left along each line. The dashed line shows the FID scores for 200 steps, indicating the strong performance of *LDM-*$\{4\text{-}8\}$. FID scores assessed on 5000 samples. All models were trained for 500k (CelebA) / 2M (ImageNet) steps on an A100.

- CelebA-HQ [39] 및 ImageNet 데이터셋에서, DDIM 샘플러 [84]의 서로 다른 노이즈 제거 스텝 수에 따른 샘플링 속도와 FID를 비교
  - **LDM-{4–8}은 지각 압축과 생성 압축 비율이 부적절한 모델들을 일관되게 능가**
  - 픽셀 기반 LDM-1과 비교하면, 훨씬 낮은 FID를 달성하면서도 샘플 처리량을 크게 향상

⇒ **LDM-4 및 LDM-8**이 고품질 합성을 달성하기 위한 최적 조건을 제공한다.

2. Image Generation with Latent Diffusion

CelebA-HQ , FFHQ , LSUN-Churches 및 LSUN-Bedrooms 에서 **256^2 해상도의 무조건 이미지 생성 모델**을 학습하고,

(i) 샘플 품질과

(ii) 데이터 매니폴드 커버리지

를 각각 FID 와 Precision-and-Recall 로 평가

CelebA-HQ 256 × 256				FFHQ 256 × 256			
Method	FID ↓	Prec. ↑	Recall ↑	**Method**	FID ↓	Prec. ↑	Recall ↑
DC-VAE [63]	15.8	-	-	ImageBART [21]	9.57	-	-
VQGAN+T. [23] (k=400)	10.2	-	-	U-Net GAN (+aug) [77]	10.9 (7.6)	-	-
PGGAN [39]	8.0	-	-	UDM [43]	5.54	-	-
LSGM [93]	7.22	-	-	StyleGAN [41]	4.16	0.71	0.46
UDM [43]	7.16	-	-	ProjectedGAN [76]	**3.08**	0.65	0.46
LDM-4 (ours, 500-s†)	**5.11**	0.72	0.49	*LDM-4* (ours, 200-s)	4.98	**0.73**	**0.50**
LSUN-Churches 256 × 256				LSUN-Bedrooms 256 × 256			
Method	FID ↓	Prec. ↑	Recall ↑	**Method**	FID ↓	Prec. ↑	Recall ↑
DDPM [30]	7.89	-	-	ImageBART [21]	5.51	-	-
ImageBART [21]	7.32	-	-	DDPM [30]	4.9	-	-
PGGAN [39]	6.42	-	-	UDM [43]	4.57	-	-
StyleGAN [41]	4.21	-	-	StyleGAN [41]	2.35	0.59	0.48
StyleGAN2 [42]	3.86	-	-	ADM [15]	1.90	**0.66**	**0.51**
ProjectedGAN [76]	**1.59**	0.61	0.44	ProjectedGAN [76]	**1.52**	0.61	0.34
*LDM-8** (ours, 200-s)	4.02	**0.64**	**0.52**	*LDM-4* (ours, 200-s)	2.95	**0.66**	0.48

Table 1. Evaluation metrics for unconditional image synthesis. CelebA-HQ results reproduced from [43, 63, 100], FFHQ from [42, 43]. †: N-s refers to N sampling steps with the DDIM [84] sampler. *: trained in KL-regularized latent space. Additional results can be found in the supplementary.

→ CelebA-HQ에서 **FID 5.11의 새로운 SOTA**를 달성하여 기존 우도 기반 모델 및 GAN 모두를 능가함

3. Conditional Latent Diffusion

1. Transformer Encoders for LDMs

   Cross-attention 기반 조건 메커니즘을 LDM에 도입함으로써, 우리는 기존 확산 모델로는 다루기 어려웠던 다양한 조건 모달리티를 가능하게 한다.

   - **텍스트-투-이미지 합성**을 위해, LAION-400M 데이터셋에서 언어 프롬프트를 조건으로 하는 **14.5억 파라미터의 KL-정규화 LDM**을 학습

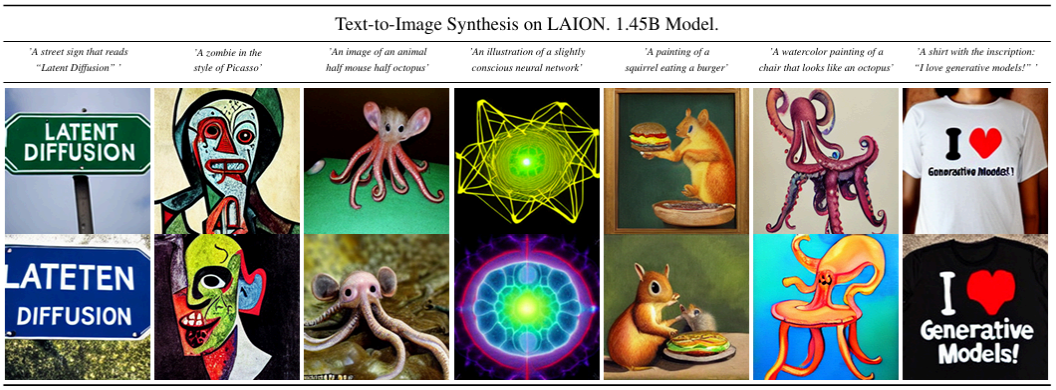

Figure 5. Samples for user-defined text prompts from our model for text-to-image synthesis, *LDM-8 (KL)*, which was trained on the LAION [78] database. Samples generated with 200 DDIM steps and $\eta = 1.0$. We use unconditional guidance [32] with $s = 10.0$.

Figure 8. Layout-to-image synthesis with an *LDM* on COCO [4], see Sec. 4.3.1. Quantitative evaluation in the supplement D.3.

→ 복잡한 사용자 정의 텍스트 프롬프트에 대해서도 잘 일반화되는 강력한 모델 생성

2. Convolutional Sampling Beyond 256^2
   - 공간적으로 정렬된 조건 정보를 ϵ_θ 의 입력에 연결함으로써, LDM은 **범용 이미지-투-이미지 변환 모델**로 활용.
     - 시멘틱 합성 - 시맨틱 맵과 풍경 이미지 쌍을 사용하고, 시맨틱 맵을 잠재 이미지 표현과 함께 연결

       → 입력 해상도는 256^2 로 학습하지만 합성곱 방식으로 평가할 경우 **메가픽셀 해상도까지 일반화됨을 확인**

Figure 9. A *LDM* trained on 256^2 resolution can generalize to larger resolution (here: 512×1024) for spatially conditioned tasks such as semantic synthesis of landscape images. See Sec. 4.3.2.

4. Super-Resolution with Latent Diffusion

LDM은 저해상도 이미지를 입력으로 직접 조건화(concatenation)함으로써 초해상도 학습을 효율적으로 수행할 수 있다.

- LDM-SR은 현실적인 질감 재현에서 강점을 보이며, SR3는 더 구조적으로 일관된 결과를 생성한다. FID 측면에서는 LDM-SR이 SR3를 능가하지만, PSNR과 SSIM은 단순 회귀 모델이 가장 높다. 그러나 이러한 지표들은 인간 지각과 잘 일치하지 않는다
- 픽셀 기반 모델과 LDM-SR을 비교하는 사용자 연구

	SR on ImageNet		Inpainting on Places	
User Study	Pixel-DM ($f1$)	*LDM-4*	LAMA [88]	*LDM-4*
Task 1: Preference vs GT ↑	16.0%	**30.4%**	13.6%	**21.0%**
Task 2: Preference Score ↑	29.4%	**70.6%**	31.9%	**68.1%**

Table 4. Task 1: Subjects were shown ground truth and generated image and asked for preference. Task 2: Subjects had to decide between two generated images. More details in E.3.6

→ LDM-SR 의 성능이 우수함.

- Bicubic 열화가 일반적인 이미지에 잘 일반화되지 않는 문제를 해결하기 위해, 보다 다양한 열화 모델을 사용하는 **LDM-BSR**도 학습

5. Inpainting with Latent Diffusion

- 인페인팅 = 이미지의 손상되거나 제거하고자 하는 영역을 새로운 콘텐츠로 채우는 작업

- Fast Fourier Convolutions 기반의 특수 아키텍처를 사용하는 최신 인페인팅 모델 LaMa 와 비교 평가

  → 픽셀 기반 모델 대비 **최소 2.7배의 속도 향상**과 함께 **FID는 최소 1.6배 개선**됨

Method	40-50% masked		All samples	
	FID ↓	LPIPS ↓	FID ↓	LPIPS ↓
LDM-4 (ours, big, w/ ft)	**9.39**	0.246± 0.042	**1.50**	0.137± 0.080
LDM-4 (ours, big, w/o ft)	12.89	0.257± 0.047	2.40	0.142± 0.085
LDM-4 (ours, w/ attn)	11.87	0.257± 0.042	2.15	0.144± 0.084
LDM-4 (ours, w/o attn)	12.60	0.259± 0.041	2.37	0.145± 0.084
LaMa [88]†	12.31	**0.243**± 0.038	2.23	**0.134**± 0.080
LaMa [88]	12.0	**0.24**	2.21	0.14
CoModGAN [107]	10.4	0.26	1.82	0.15
RegionWise [52]	21.3	0.27	4.75	0.15
DeepFill v2 [104]	22.1	0.28	5.20	0.16
EdgeConnect [58]	30.5	0.28	8.37	0.16

Table 7. Comparison of inpainting performance on 30k crops of size 512×512 from test images of Places [108]. The column *40-50%* reports metrics computed over hard examples where 40-50% of the image region have to be inpainted. †recomputed on our test set, since the original test set used in [88] was not available.

  → attention을 포함한 우리 모델이 LaMa 대비 FID와 LPIPS 측면에서 전반적인 이미지 품질을 개선함

- VQ-정규화 잠재 공간에서 attention 없이 더 큰 확산 모델을 추가 학습

  → 이는 BigGAN 스타일의 residual block을 사용하며 총 **3.87억 파라미터**를 가지고, 이후 512^2 해상도에서 반 에폭 파인튜닝을 수행하여 새로운 FID SOTA를 달성함.

Figure 11. Qualitative results on object removal with our *big, w/ ft* inpainting model. For more results, see Fig. 22.

Conclusion

> **잠재 확산 모델(latent diffusion models):**
> 노이즈 제거 확산 모델의 학습 및 샘플링 효율을, 품질 저하 없이 크게 향상시키는 단순하고 효율적인 방법

→ cross-attention 기반 조건 메커니즘을 바탕으로 수행한 실험들에서, 작업별로 특화된 아키텍처를 사용하지 않고도, 광범위한 조건부 이미지 합성 과제에서 SOTA 방법들과 비교해도 우수한 성능을 달성할 수 있음을 보여주었음.