

# 효율적 생성모델 압축을 위한 가지치기된 가중치 특이값 조정

김현진<sup>o</sup>, 유재준

울산과학기술원 인공지능대학원

hyeonjin.kim@unist.ac.kr, jaejun.yoo@unist.ac.kr

## 요약

생성 모델 압축을 위한 가지치기 기법은 추가 훈련 비용 없이 압축된 모델의 성능을 유지할 수 있지만, 사전 학습 모델의 주요 연결 보존에만 초점을 맞추어 가지치기된 가중치가 미세 조정 과정에 미치는 영향을 간과하는 경우가 많다. 또한, 대부분의 생성 모델 압축 기법은 StyleGAN 과 같은 특정 모델에 맞춰 개발되었으며, 확산 모델 압축에 대한 연구는 초기 단계에 머물러 있다. 본 논문에서는 생성적 적대 신경망과 확산 모델 모두에 적용 가능한 범용적인 가지치기된 가중치 정제 기법인 특이값 조정 (Singular Value Scaling)을 제안한다. 분석 결과, 가지치기된 가중치에서 발생하는 지배적인 특이벡터가 미세 조정의 효율성을 저해하는 주요 원인임을 확인하였다. 이를 해결하기 위해, 특이값 조정 기법은 가지치기된 가중치의 특이값 격차를 줄여 가중치 초기화를 개선하고, 압축된 모델이 최적화된 솔루션으로 빠르게 수렴할 수 있도록 유도한다. StyleGAN2 와 DDPM 을 대상으로 한 실험에서 특이값 조정이 다양한 모델에서 압축 성능을 효과적으로 개선할 수 있음을 보였다.

## 1. 서론

생성 모델, 특히 생성적 적대 신경망 (GANs)[1] 과 확산 모델 (Diffusion Models)[3]은 이미지 생성 [2,16], 편집[21,22], 비디오 생성[23,24], 3D 생성 [25,26] 등 다양한 컴퓨터 비전 작업에서 뛰어난 성능을 보여 왔다. 하지만 생성 모델의 인상적인 성능은 높은 메모리와 연산 자원을 필요로 하여 실제 응용에서의 한계를 초래하고 있다. 특히, 모바일 기기나 엣지 디바이스 (edge device)와 같이 자원이 제한된 환경에서는 이러한 모델을 활용하기 어렵다.

이러한 문제를 해결하기 위해 모델 압축 기법 [4,5,6,7]이 활발히 연구되고 있으며, 보통 가지치기 (pruning)와 미세 조정 (fine-tuning)으로 이루어져 있다. 가지치기는 사전 학습 모델의 가중치 (weights)에서 상대적으로 불필요한 연결을 제거하여 사전 학습된 지식을 유지하고, 이후의 미세 조정 과정에서 향상된 성능으로의 빠른 수렴을 가능하게 한다. 그러나 기존의 압축 기법들은 가지치기된 가중치 (pruned weights)가 미세 조정에 미치는 영향을 충분히 고려하지 않았고, 이는 비효율적인 미세 조정으로 이어졌다. 이 문제는 모델의 용량 (capacity)이 감소할수록 더욱 심각해지며, 경우에 따라 가지치기된 가중치가 무작위 초기화보다도 느린 수렴 속도와 낮은 성능을 보이는 결과를 초래할 수 있다. 따라서 이러한 요인을 해결하는 것은 더 효율적인 생성 모델을 얻기 위해 필요하다.

우리의 주요 관찰은 가지치기된 가중치에서 지배적인 특이벡터가 나타나며, 이로 인해 큰 특이값과 작은 특이값 간의 격차가 커진다는 점이다. 이러한 지배적인 특이벡터는 모델의 순방향 전파를 주도하며, 작은 특이벡터의 기여를 약화시켜 미세 조정 효율성을 저하시킨다. 이를 해결하기 위해 초기화 단계에서 특이벡터 간의 균형 있는 기여를 보장하는 것이 필요하다.

본 논문에서는 이러한 문제를 해결하기 위해 특이값 조정 (Singular Value Scaling, SVS)이라는 가지치기된 가중치 정제 기법을 제안한다. 특이값 조정은 가지치기된 가중치의 가장 큰 특이값과 작은 특이값 간의 격차를 줄이고, 모든 특이벡터가 초기화 단계에서 균형 있게 기여할 수 있도록 한다. 이 과정은 가지치기된 가중치의 초기화를 개선하여 효율적이고 효과적인 미세 조정이 이루어지도록 돕는다. 또한, 특이값 조정은 특정 모델 구조에 의존하지 않으며, 생성적 적대 신경망과 확산 모델 모두에 적용 가능한 범용적 기법이다.

우리는 대표적 생성모델인 StyleGAN2 와 DDPM 을 CIFAR10[8], CelebA-HQ[9], FFHQ[10], 그리고 LSUN Cat[11]의 다양한 데이터셋을 대상으로 실험을 진행하였다. 실험 결과는 제안된 기법이 추가 훈련 비용 없이 기존 생성 모델 압축 기법 대비 더 빠른 수렴 속도와 높은 성능을 달성할 수 있음을 보여준다.

## 2. 관련 연구

### 2.1 생성 모델 압축

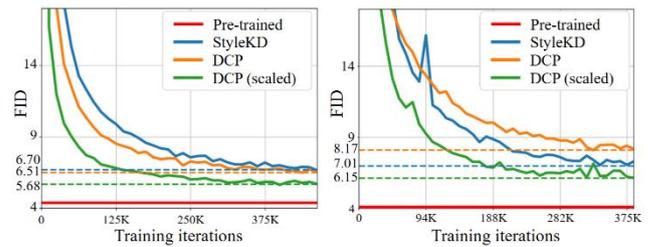
StyleGAN[2,10,16]과 같은 생성 모델의 압축을 위한 다양한 연구가 진행되고 있다. CAGAN[4]은 사전 학습된 StyleGAN 모델의 가지치기와 지식 증류를 결합한 첫 번째 방법으로, 픽셀 및 특징 (feature) 수준의 지식 증류를 통해 가지치기 이후의 미세 조정을 수행했다. 특히, CAGAN 은 콘텐츠 인식 가지치기 (Content-aware Pruning) 기법을 도입하여 생성 이미지의 의미론적 정보 (semantic part)를 보존하는 데 중요한 연결을 유지했다. StyleKD[5]는 기존 모델의 매핑 네트워크 (mapping network)만을 상속하고 합성 네트워크 (synthesis network)는 무작위 초기화하는 것이 낫다는 것을 발견하였다. 또한, 관계 손실 (relation loss)을 활용하여 미세 조정 단계를 더욱 강화했다. DCP-GAN[6]은 생성 다양성을 고려한 채널 가지치기 기법(Diversity-aware Channel Pruning)을 도입하여, 합성 네트워크에서 샘플 다양성에 기여하는 연결을 보존함으로써 생성 이미지의 품질을 유지하는 동시에 학습 속도를 크게 향상시켰다.

확산 모델에서도 가지치기를 활용한 효율성 향상이 주목받고 있다. 확산 모델은 안정적인 학습과 뛰어난 생성 성능을 보이지만, 반복적인 추론 과정으로 인한 높은 계산 비용이 단점으로 지적된다. 최근 제안된 Diff-Prune[7]은 중요도가 낮은 확산 단계 (diffusion step)는 배제하고, 중요도가 높은 확산 단계를 기준으로 가지치기를 수행하여, 성능 손실을 최소화하고, 미세 조정 과정에서 필요한 학습 반복 횟수를 크게 감소시킬 수 있음을 보였다.

### 2.2 가중치 초기화와 학습 가능성

가중치 초기화는 딥러닝에서 효율적인 학습을 위해 필수적인 요소이다. Glorot 초기화 (Glorot initialization)[12]와 He 초기화 (He initialization)[13]는 활성화 (activation) 값의 분산을 유지하고 학습의 안정성을 보장하기 위해 널리 사용되며, 각각 sigmoid 와 tanh 활성화 함수, 그리고 ReLU 활성화 함수를 사용하는 전방 전달 신경망 (feed-forward network)에서 효과적이다. 또한, 심층 선형 모델의 학습에 대한 연구[14]에서는 가중치의 모든 특이값을 1 로 설정하는 정규 직교 초기화(orthogonal initialization)를 사용하여 학습 가능성 (trainability)을 극대화하는 이상적인 상태인 동적 등형성 (dynamical isometry)을 달성하는 것이 중요함을 보였다. 최근 분류기 가지치기 연구인 TPP[15]에서 학습 가능성을 유지하는 것이 가지치기된 네트워크의 성능 복원을 지원하는데 매우 중요하다는 점을 강조했다.

## 3. 실험 결과 및 분석



(a) FID graph of StyleGAN2 (base) (b) FID graph of StyleGAN2 (small)

그림 1. 서로 다른 모델 용량에서 각 압축 기법별 StyleGAN2 FID 수렴 속도 그래프. 실선은 각 학습 반복 횟수별 FID 를, 점선은 각 압축 기법의 최고 FID 를 나타낸다. “scaled”는 특이값 조정 기법이 가지치기된 가중치에 적용된 것을 뜻한다.

최근 제안된 생성 모델 압축 연구[6,7]에 따르면, 잘 설계된 가지치기 기법은 무작위 초기화보다 빠르게(StyleGAN2 약 2.5 배, DDPM 약 8 배) 동일한 성능에 도달하며, 초기 성능도 더 우수함을 보인다. 이는 가지치기된 가중치가 사전 학습 모델의 성능을 유지하며, 더 나은 초기 상태에서 미세 조정을 시작할 수 있기 때문이다. 그러나 그림 1 에서 확인할 수 있듯이, 미세 조정이 진행됨에 따라 가지치기된 가중치로 초기화된 모델 (주황색 선, DCP-GAN)이 무작위 초기화된 모델 (파란색 선, StyleKD)보다 수렴 속도가 느려지는 경향이 있다. 이 문제는 모델 크기가 작아질수록 더욱 심화되며, 이는 가지치기된 가중치 자체에 미세 조정을 저해하는 요인이 포함되어 있음을 시사한다.

우리는 가지치기된 가중치가 사전 학습 모델로부터 상속받은 학습된 사전 정보 (prior)를 분석하기 위해 특이값 분해 (Singular Value Decomposition, SVD)를 활용했다. 가중치 행렬  $W \in \mathbb{R}^{m \times n}$  ( $m < n$ ) 는 SVD 를 통해 아래와 같이 분해된다:

$$W = U\Sigma V^T = \sum_{i=1}^m \sigma_i u_i v_i^T$$

이때  $U = [u_1, u_2, \dots, u_m] \in \mathbb{R}^{m \times m}$ ,  $V = [v_1, v_2, \dots, v_n] \in \mathbb{R}^{n \times n}$  는 직교 행렬이고,  $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_m) \in \mathbb{R}^{m \times n}$  은 대각 성분에 특이값을 가진 대각 행렬이다. 합성곱 신경망 (convolutional layer)는 가중치를 평탄화 (flatten)한 후 특이값 분해를 시행한다. 각 특이값은 해당 특이벡터가 가중치 내에서 차지하는 영향을 나타낸다.

가장 주목할 만한 관찰은 가지치기된 가중치의 가장 큰 특이값과 가장 작은 특이값 사이의 큰 격차 (약 100 배)이다. 이는 가중치의 순방향 및 역방향 전파 (forward and backward propagation)가 이러한 지배적인 특이벡터에 의해 크게 영향을 받는다는 것을 의미한다. 이로 인해 학습 과정에서 압축된 모델이 이러한 특이벡터에 편향될 가능성이 높아지고, 가중치 공간에서의 다양한 탐색이 심각하게 제한될 수 있다.[14]

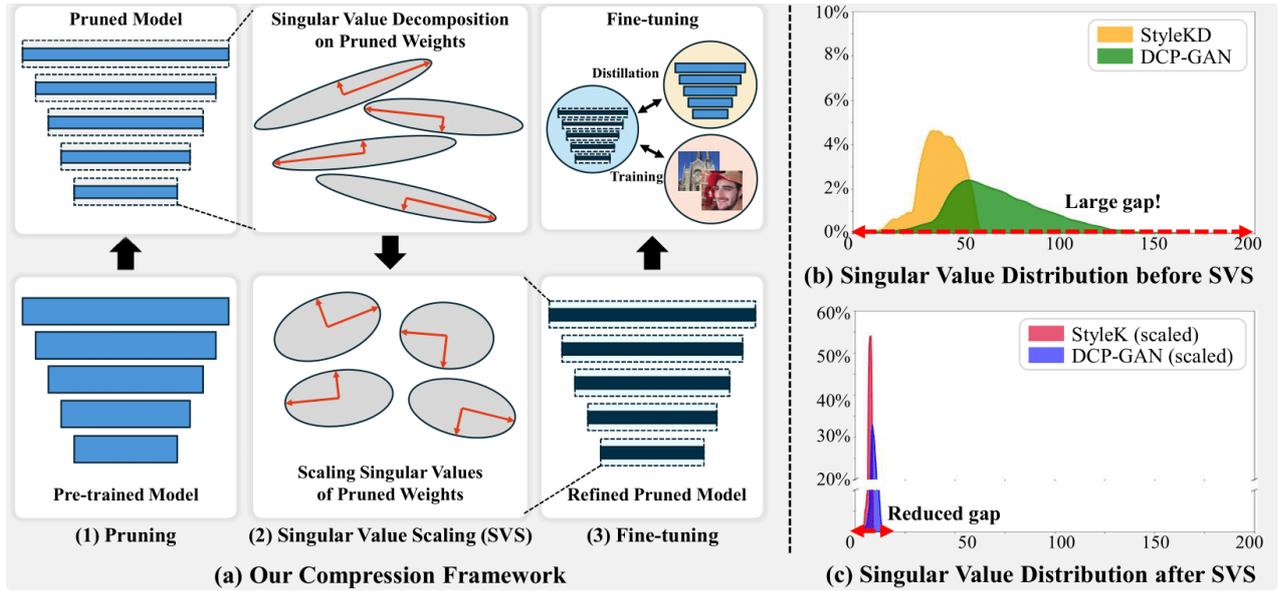


그림 2. (a) 우리의 압축 프레임워크의 개요도. 기존 모델 압축 기법과 달리, 효율적인 모델 미세 조정을 위해 추가적으로 가지치기된 가중치 정제 단계를 수행한다. 먼저, 사전 학습 모델을 가지치기한 후, 가지치기된 모델의 가중치에 대한 특이값을 계산한다. 이후, 특이값 조절을 통해 크기가 큰 특이값은 감소시키고, 크기가 작은 특이값은 증가시켜 가지치기된 가중치를 정제한다. 마지막으로, 정제된 가지치기된 모델을 미세 조정한다. (b), (c) DCP-GAN 및 Diff-Prune 에 의해 가지치기된 가중치의 특이값 조절 적용 전후의 특이값 분포를 나타낸 그래프. x 축은 특이값을, y 축은 밀도를 나타낸다. 특이값 조절 적용시 특이값 간 격차가 줄어들며, 가중치 내 각 특이벡터의 기여도가 균형을 이루게 된다.

#### 4. 방법

우리의 관찰을 바탕으로, 미세 조정 효율성을 향상시키기 위해 가지치기된 가중치를 정제하는 "특이값 조절 (Singular Value Scaling, SVS)"을 제안한다. 특이값 조절의 주요 목표는 가지치기된 가중치의 특이값 간 격차를 줄이는 것이다. 가지치기된 가중치에서는 지배적인 특이벡터가 상대적으로 매우 큰 특이값을 가진다. 가지치기된 가중치의 모든 특이벡터는 사전 학습 모델의 중요한 지식을 담고 있으므로, 특정 특이벡터가 지나치게 지배적이지 않도록 하여 학습 초기 단계에서 모든 특이벡터가 균형 있게 기여하도록 유도한다. 이를 위해 간단히 "제공근 함수"를 이용해 특이값을 조정한다.

$$W_{scaled} = U \Sigma_{scaled} V^T = \sum_{i=1}^m \sqrt{\sigma_i} u_i v_i^T$$

여기서  $U$  와  $V$  는 변하지 않고,  $\Sigma_{scaled} = \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \dots, \sqrt{\sigma_m}) \in \mathbb{R}^{m \times n}$  이다. 제공근 함수는 다음과 같은 장점을 가진다. 1) 특이값이 항상 비음수(non-negative)이므로, 결과값도 비음수를 유지한다. 2)  $\sigma_i < 1$ 은 증가시키고,  $\sigma_i > 1$ 인 경우 값이 클수록 더욱 크게 감소시킨다. 3) 양수 영역에서 단조 증가 함수로, 특이값의 상대적 순서를 유지한다. 이와 같은 방식으로 특이값을 조정하면, 원래의 특이벡터를 유지하면서 각 특이벡터의 상대적 기여를 균형 있게 조정할 수 있다. 이러한 균형 조절은 압축된 모델이 사전 학습 모델의 지식을 효과적으로 활용하

도록 돕고, 최적 상태로 더 효율적으로 수렴할 수 있도록 한다.

가중치와 편향 (bias)은 동시에 학습되기 때문에, 가중치의 특이값 조절시 편향의 영향도 고려해야 한다. 순방향 전파  $y = Wx + b$  ( $W \in \mathbb{R}^{m \times n}$ 는 가중치,  $x \in \mathbb{R}^n$ 는 입력,  $b \in \mathbb{R}^m$ 는 편향, 그리고  $y \in \mathbb{R}^m$ 는 출력)을  $W$ 에 대해 표현하면 다음과 같다:

$$y = W(x + W^\dagger b) = W(x + V \Sigma^{-1} U^T b)$$

이때,  $W^\dagger \in \mathbb{R}^{n \times m}$  은 Moore-Penrose 역행렬,  $\Sigma^{-1} = \text{diag}(\sigma_1^\dagger, \sigma_2^\dagger, \dots, \sigma_m^\dagger) \in \mathbb{R}^{n \times m}$ , 그리고

$$\sigma_i^\dagger = \begin{cases} \frac{1}{\sigma_i}, & \sigma_i \neq 0 \\ 0, & \sigma_i = 0 \end{cases}$$

이다. 편향  $b$ 를  $b = |b| \vec{\beta}$  ( $| \cdot |$ 는 벡터 노름,  $\vec{\beta}$ 는 정규화된 벡터)로 나타내면, 위 식은 다음으로 나타낸다:

$$y = W(x + V(|b| \Sigma^{-1}) U^T \vec{\beta})$$

여기서,  $|b| \Sigma^{-1} = \text{diag}(|b| \sigma_1^\dagger, |b| \sigma_2^\dagger, \dots, |b| \sigma_m^\dagger)$ 이다. 만약  $\sigma_i$ 가  $\sqrt{\sigma_i}$ 로 조정된다면,  $b$ 와  $\sigma_i$ 의 관계는  $\frac{|b|}{\sigma_i}$ 에서  $\frac{|b|}{\sqrt{\sigma_i}}$ 로 변환된다. 우리가 가지치기된 가중치의 특이값을 제공근 방식으로 조정하면 편향과 특이값의 관계도 제공근 방식으로 조정하고자 한다. 구체적으로, 편향  $b$ 를  $b_{scaled} = \frac{b}{\sqrt{|b|}}$ 로 조정하여 다음과 같은  $b$ 와  $\sigma_i$ 의 관계를 유도한다.

$$\frac{|b_{scaled}|}{\sigma_{i,scaled}} = \sqrt{\frac{|b|}{\sigma_i}}$$

표 1. StyleGAN2 압축에 대한 정량적 결과. FFHQ-256 과 LSUN Cat-256 데이터셋에서 이전 압축 기법을 이용한 압축된 StyleGAN2 학습 성능을 보여준다. **Bold 체**는 최고 성능을, 밑줄은 차선의 성능을 뜻한다. 채널 압축률은 70%로 설정되었다. “파라미터”와 “연산량”은 모델 내 파라미터 수와 FLOPs (FLoating-point OPERations) 를 의미한다. “P”, “R”, “D”, 그리고 “C”는 각각 precision, recall, density, 그리고 coverage 점수를 뜻한다. “base” 와 “small”은 서로 다른 용량의 StyleGAN2 구조를 뜻한다. “SVS”는 제안된 특이값 조정이 적용되었음을 의미한다.

데이터셋	기법	구조	파라미터	연산량	FID (↓)	P (↑)	R (↑)	D (↑)	C (↑)
FFHQ	사전 학습	StyleGAN2 (base)	30.0M	45.1B	4.29	0.762	0.561	1.061	0.845
	StyleKD		5.6M	4.1B	6.70	0.711	<u>0.551</u>	0.879	<u>0.786</u>
	DCP-GAN				<u>6.51</u>	<u>0.712</u>	<b>0.556</b>	<u>0.884</u>	0.785
	DCP-GAN+SVS (Ours)			<b>5.68</b>	<b>0.740</b>	0.530	<b>0.981</b>	<b>0.806</b>	
	사전 학습	StyleGAN2 (small)	24.7M	14.9B	4.02	0.769	0.555	1.095	0.854
	StyleKD		4.9M	1.3B	<u>7.01</u>	<u>0.707</u>	<b>0.543</b>	<u>0.900</u>	<u>0.783</u>
DCP-GAN	8.17				0.680	<u>0.539</u>	0.787	0.741	
DCP-GAN+SVS (Ours)			<b>6.15</b>	<b>0.731</b>	0.524	<b>0.952</b>	<b>0.794</b>		
LSUN Cat	사전 학습	StyleGAN2 (base)	30.0M	45.1B	6.60	0.622	0.521	0.788	0.463
	StyleKD		5.6M	4.1B	9.20	<b>0.608</b>	0.442	<u>0.723</u>	<u>0.413</u>
	DCP-GAN				<u>9.08</u>	0.590	<u>0.462</u>	0.681	0.404
	DCP-GAN+SVS (Ours)			<b>8.09</b>	<u>0.604</u>	<b>0.471</b>	<b>0.736</b>	<b>0.428</b>	
	사전 학습	StyleGAN2 (small)	24.7M	14.9B	8.25	0.623	0.448	0.778	0.423
	StyleKD		4.9M	1.3B	11.93	<u>0.573</u>	0.390	<u>0.657</u>	<u>0.368</u>
DCP-GAN	<u>11.58</u>				0.568	<u>0.411</u>	0.644	0.364	
DCP-GAN+SVS (Ours)			<b>10.71</b>	<b>0.585</b>	<b>0.412</b>	<b>0.666</b>	<b>0.372</b>		

표 2. DDPM 압축에 대한 정량적 결과. CIFAR10 32 × 32와 CelebA-HQ 64 × 64 데이터셋에서 이전 압축 기법을 이용한 압축된 DDPM 학습 성능을 보여준다. **Bold 체**는 최고 성능을 뜻한다. “파라미터”와 “연산량”은 모델 내 파라미터 수와 MACs (Multiply-ACcumulate)를 의미한다. “P”, “R”, “D”, 그리고 “C”는 각각 precision, recall, density, 그리고 coverage 점수를 뜻한다. “SVS”는 제안된 특이값 조정이 적용되었음을 의미한다.

데이터셋	기법	압축률	파라미터	연산량	FID (↓)	P (↑)	R (↑)	D (↑)	C (↑)
CIFAR10	사전 학습	0%	35.7M	6.1B	4.19	0.672	0.759	0.762	0.897
	Diff-Prune	30%	19.8M	3.4B	5.49	<b>0.671</b>	<b>0.742</b>	<b>0.772</b>	0.888
	Diff-Prune+SVS (Ours)				<b>5.14</b>	0.668	0.740	0.768	<b>0.889</b>
	Diff-Prune	50%	8.96M	1.5B	7.77	0.670	<b>0.720</b>	0.805	0.861
Diff-Prune+SVS (Ours)	<b>7.41</b>				<b>0.671</b>	0.718	<b>0.817</b>	<b>0.868</b>	
CelebA-HQ	사전 학습	0%	78.7M	23.9B	6.48	0.587	0.812	0.488	0.755
	Diff-Prune	30%	43.7M	13.3B	6.17	0.591	<b>0.803</b>	0.502	0.760
	Diff-Prune+SVS (Ours)				<b>5.57</b>	<b>0.619</b>	0.793	<b>0.573</b>	<b>0.785</b>
	Diff-Prune	50%	19.7M	6.0B	5.32	0.618	<b>0.773</b>	0.594	0.793
Diff-Prune+SVS (Ours)	<b>4.66</b>				<b>0.634</b>	0.766	<b>0.626</b>	<b>0.812</b>	

## 5. 실험 및 분석

### 5.1 실험 환경

우리는 대표적인 생성 모델인 StyleGAN2[2]와 DDPM[3]에서 실험을 진행하였다. StyleGAN2 의 경우, 가장 최신 가지치기 기법인 DCP-GAN[6]에 특이값 조정 기법을 적용하였다. 모델 모델 용량의 경

우 기존 기법들이 실험한 기본 구조 (StyleGAN2 (base)라고 지칭)와 더 작은 용량을 가진 구조 (StyleGAN2 (small)라고 지칭)을 사용하였다. 채널 압축률은 70%로 설정하였고, 데이터셋은 FFHQ[10]와 LSUN Churc[11]에서 실험을 진행하였다. 현재 가장 최고 성능을 보이는 StyleKD[5]와 DCP-GAN[6]과 비교하였다. 학습률 (learning rate) 및 배치 크기 (batch size) 등의 학습 환경은 공식 StyleGAN2[2,16] 학습

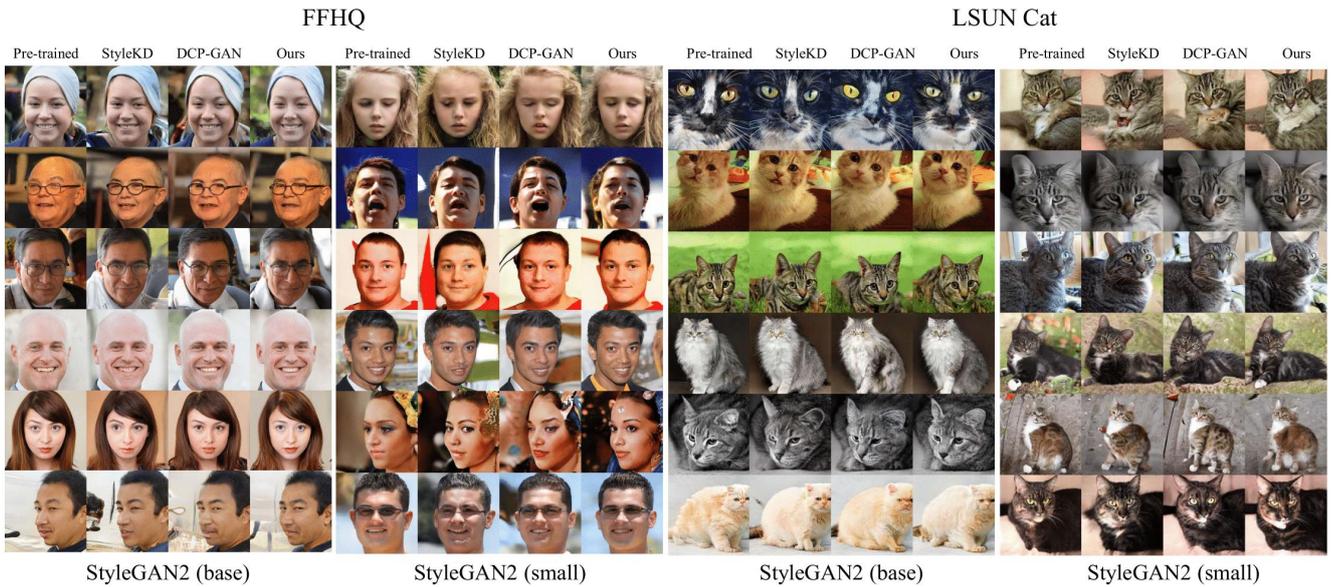


그림 3. FFHQ 및 LSUN Cat 데이터셋에 대한 정성적 결과. 각 행의 샘플은 동일한 노이즈 벡터를 사용하여 생성되었다. 서로 다른 용량의 StyleGAN2가 최신 압축 기법을 활용해 채널 압축률 70%로 압축되었다. "Ours"는 특이값 조정으로 정제된 DCP-GAN을 사용하여 압축된 모델을 나타낸다.

환경을 따랐으며, 손실 함수 (loss function) 및 하이퍼 파라미터 (hyper parameter)는 DCP-GAN 및 StyleKD를 따랐다 (DCP-GAN과 StyleKD는 동일한 학습 환경을 사용하고, 가중치 초기화만 다르다).

DDPM[3]의 경우, 최신 확산 모델 가지치기 기법인 Diff-Prune[7]에 특이값 조정 기법을 적용하였다. 채널 압축률은 30%와 50%로 설정하였고, 데이터셋은 CIFAR10[8]과 CelebA-HQ[9]를 사용하였다. 학습 환경은 Diff-Prune을 따랐다.

### 5.2 평가 지표

모델 압축 성능 평가를 위해 Frechet Inception Distance (FID)[17], Precision과 Recall (P&R)[18], 그리고 Density과 Coverage (D&C)[19]를 사용하였다. FID는 생성된 샘플의 분포가 실제 데이터 분포와 얼마나 일치하는지를 측정한다. Precision과 Density는 생성 샘플의 품질 (정밀도)을, Recall과 Coverage는 생성 샘플의 다양성을 평가한다. DDPM의 경우 샘플링에는 100 단계 DDIM 샘플러[20]를 사용하였다.

### 5.3 실험 분석

그림 1은 FFHQ 데이터셋에서 StyleGAN2 (base)와 StyleGAN2 (small)의 FID 수렴 그래프를 보였다. 그림 1 (a)에서 DCP-GAN은 StyleKD보다 더 나은 성능에 도달하지만, 학습 후반부에서 수렴 속도가 느려지는 문제가 나타난다. 이 문제는 모델 용량이 감소할수록 (그림 2 (b)) 더욱 심각해지며, StyleKD가 초기 미세 조정 단계에서 DCP-GAN을 추월하고 DCP-GAN이 최적화되지 못한 성능에 머물게 된다.

반면, 특이값 조정으로 정제된 가지치기된 가중치를 사용하면 미세 조정이 용이해지고, 대부분의 구간에서 빠르게 적합하며 이전 방법들보다 더 우수한 성능을 달성하였다.

표 1은 FFHQ와 LSUN Cat 데이터셋에서 StyleGAN2 압축 실험 결과를 보여준다. FID 기준, 제안된 방법은 모든 실험에서 이전 압축 방법보다 우수한 성능을 보였으며, Precision과 Density 지표에서도 대체로 뛰어난 성능을 나타냈다. Recall과 Coverage 지표에서는 이전 방법과 비슷하거나 더 나은 성능을 보였다. 특히, StyleGAN2 (base)에서는 DCP-GAN이 StyleKD보다 일관되게 우수했지만, StyleGAN2 (small)에서는 DCP-GAN의 성능이 StyleKD에 뒤처지는 경향을 보였다. 이는 DCP-GAN의 가지치기된 가중치가 미세 조정하기 더 어렵다는 것을 시사한다. 반면, 제안된 방법은 효과적인 초기화를 제공해 최종 성능을 개선했음을 보여준다.

표 2은 DDPM 압축 결과를 보여준다. 제안된 방법은 미세 조정 과정에서 모든 압축률에서 더 나은 FID 점수를 달성하며, 효과적인 분포 근사능력을 보였다. 또한, Precision, Recall, Density, 그리고 Coverage 점수에서도 대부분 최고 성능을 달성하여 제안된 방법이 미세 조정을 더욱 용이하게 하여 더 나은 최적화 성능으로 이끈다는 것을 입증한다. 위 결과로부터 제안된 특이값 조정 기법이 압축된 모델의 학습가능성을 회복하여 효과적으로 미세 조정되어 최적 지점으로 수렴할 수 있다는 것을 보여준다. 각 StyleGAN2 압축 기법을 통해 압축된 모델로부터 생성된 영상은 그림 3에서 확인할 수 있다.

## 6. 결론

본 논문에서는 가지치기된 가중치의 특이값 격차를 줄여 생성 모델 압축의 효율성과 효과를 높이는 "특이값 조정" 기법을 제안한다. 이 방법은 모델이 더 다양한 가중치 공간을 탐색하도록 돕고, 생성적 적대 신경망과 확산 모델 등 다양한 모델에 적용 가능하다. 실험 결과, 제안된 기법은 보다 효과적인 초기화를 제공하며, StyleGAN2 과 DDPM 압축에서 성능 향상을 입증했다.

## 감사의 글

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원과 한국연구재단의 지원을 받아 수행된 연구임. (No.2020-0-01336, 인공지능대학원지원(울산과학기술원), No.2022-0-00959, No.RS-2022-II220959 (Part 2), 의사결정 지원을 위한 퓨샷 학습 기반 시각 및 언어에 대한 인과관계 추론기술개발, No.RS-2022-II220264, 지식기반 심층논리 신경망을 활용한 통합적 비디오 해석과 생성 연구, No.2021-0-02068, 인공지능 혁신 허브 연구 개발, No.2022R1C1C100849612)

## 참고문헌

- [1] Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems* 27 (2014).
- [2] Karras, Tero, et al. "Analyzing and improving the image quality of stylegan." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.
- [3] Ho, Jonathan, et al. "Denoising diffusion probabilistic models." *Advances in neural information processing systems* 33 (2020).
- [4] Liu, Yuchen, et al. "Content-aware gan compression." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.
- [5] Xu, Guodong, et al. "Mind the gap in distilling StyleGANs." *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 2022.
- [6] Chung, Jiwoo, et al. "Diversity-aware Channel Pruning for StyleGAN Compression." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.
- [7] Fang, Gongfan, et al. "Structural Pruning for Diffusion Models." *Advances in Neural Information Processing Systems* 36 (2023).
- [8] Krizhevsky, Alex, et al. "Learning multiple layers of features from tiny images." 2009.
- [9] Liu, Ziwei, et al. "Deep learning face attributes in the wild." *Proceedings of the IEEE international conference on computer vision*. 2015.
- [10] Karras, et al. "A Style-Based Generator Architecture for Generative Adversarial Networks." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.
- [11] Fisher, Yu, et al. "LSUN: Construction of a Large-scale Image Dataset using Deep Learning with Humans in the Loop". *CoRR* abs/1506.03365. (2015).
- [12] Glorot, Xavier and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks." *Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings*, 2010.
- [13] He, Kaiming, et al. "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification." *Proceedings of the IEEE international conference on computer vision*. 2015.
- [14] Saxe, Andrew M., et al. "Exact solutions to the nonlinear dynamics of learning in deep linear neural networks." *Proceedings of the International Conference on Learning Representations*. 2014.
- [15] Wang, Huan, and Yun Fu. "Trainability Preserving Neural Pruning." *The Eleventh International Conference on Learning Representations*. 2023.
- [16] Karras, Tero, et al. "Training generative adversarial networks with limited data." *Advances in neural information processing systems* 33 (2020).
- [17] Heusel, Martin, et al. "Gans trained by a two time-scale update rule converge to a local nash equilibrium." *Advances in neural information processing systems* 30 (2017).
- [18] Kynkäänniemi, Tuomas, et al. "Improved precision and recall metric for assessing generative models." *Advances in neural information processing systems* 32 (2019).
- [19] Naeem, Muhammad Ferjad, et al. "Reliable fidelity and diversity metrics for generative models." *International Conference on Machine Learning*. PMLR, 2020.
- [20] Song, Jiaming, et al. "Denoising Diffusion Implicit Models." *International Conference on Learning Representations*. 2021.
- [21] Baykal, Ahmet Canberk, et al. "CLIP-guided StyleGAN Inversion for Text-driven Real Image Editing." *ACM Transactions on Graphics* 42.5 (2023).
- [22] Kavar, Bahjat, et al. "Imagic: Text-based real image editing with diffusion models." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023.
- [23] Skorokhodov, Ivan, et al. "Stylegan-v: A continuous video generator with the price, image quality and perks of stylegan2." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022.
- [24] Ho, Jonathan, et al. "Video diffusion models." *Advances in Neural Information Processing Systems* 35 (2022).
- [25] Chan, Eric R., et al. "Efficient geometry-aware 3d generative adversarial networks." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022.
- [26] Ntavelis, Evangelos, et al. "Autodecoding latent 3d diffusion models." *Advances in Neural Information Processing Systems* 36 (2023).