

특집논문 (Special Paper)

방송공학회논문지 제29권 제6호, 2024년 11월 (JBE Vol.29, No.6, November 2024)

<https://doi.org/10.5909/JBE.2024.29.6.832>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## 개선된 지식 증류 기법을 활용한 효율적인 ESRGAN 압축

백 보 성<sup>a)</sup>, 김 아 로<sup>a)</sup>, 장 재 윤<sup>b)</sup>, 황 효 석<sup>b)</sup>, 최 재 열<sup>c)</sup>, 류 영 일<sup>c)</sup>, 류 은 석<sup>c)</sup>, 박 상 효<sup>a)†</sup>

### Efficient ESRGAN Compression Using Enhanced Knowledge Distillation Method

Bosung Baek<sup>a)</sup>, Aro Kim<sup>a)</sup>, Jaeyun Jang<sup>b)</sup>, Hyoseok Hwang<sup>b)</sup>, Jaeyeol Choi<sup>c)</sup>, Yeongil Ryu<sup>c)</sup>, Eunseok Ryu<sup>c)</sup>, and Sang-hyo Park<sup>a)†</sup>

#### 요 약

최근 초해상화 연구에서 생성적 적대 신경망을 활용한 모델들이 높은 성능을 보이며 주목받고 있다. 그러나 이러한 모델들은 복잡한 구조와 다수의 파라미터로 인해 제한된 환경에서의 적용이 어려운 문제가 있다. 이러한 문제를 해결하기 위해 본 연구에서는 생성적 적대 신경망 기반 초해상화 모델의 경량화를 목표로 가지치기와 지식 증류 기법을 결합한 새로운 방법을 제안한다. 제안된 방법은 교사 모델에 가지치기를 적용하여 불필요한 가중치를 제거하고, 중요한 지식만을 학생 모델에 전달함으로써 모델 크기를 교사 모델의 39%까지 줄이면서도 성능 저하를 최소화한다. 실험 결과, 테스트 데이터셋에서 PSNR에서 최대 1.2dB, SSIM에서 최대 0.04의 향상을 기록하였으며, FID와 LPIPS 지표에서도 기존 모델과 유사한 성능을 유지하였다. 본 연구는 제한된 자원 환경에서도 고품질의 초해상화 이미지를 생성할 수 있는 생성적 적대 신경망 모델 경량화의 실용적 가능성을 제시한다.

#### Abstract

Recent super-resolution research has highlighted the effectiveness of generative adversarial networks (GANs) in producing high-quality images. However, the complex architecture and large number of parameters in these models pose significant challenges for deployment in resource-constrained environments, such as mobile devices and embedded systems. To address these issues, this study proposes a novel approach to compress GAN-based SR models by combining pruning and knowledge distillation techniques. The proposed method applies pruning to the teacher model to remove unnecessary weights, then transfers only the essential knowledge to the student model, reducing the model size by 39% compared to teacher model while minimizing performance loss. Experimental results show improvements of up to 1.2dB in PSNR and 0.04 in SSIM, with the Student Model maintaining similar performance levels to the teacher model in FID and LPIPS metrics. This research demonstrates the practical potential of deploying compressed GAN models in resource-limited environments, generating high-quality super-resolution image.

Keyword: Super-resolution, GAN, Pruning, Knowledge distillation

Copyright © 2024 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

## I. 서론

초해상화(Super-Resolution, SR) 기술은 저해상도 이미지를 고해상도로 변환하는 이미지 처리 기술로, 의료 영상 분석, 보안 및 감시 시스템 등의 분야에서 그 활용도가 높아지고 있다. 또한, 생성적 적대 신경망(generative adversarial networks, GAN)을 기반으로 한 초해상화 모델이 보간 기반 초해상화 모델에 비해 뛰어난 성능을 보이고 있다. 그러나 이러한 모델들은 복잡한 네트워크 구조와 많은 수의 파라미터로 인해 높은 연산 자원을 요구한다. 생성적 적대 신경망 기반 초해상화 모델이 높은 성능을 유지하면서도 실제 응용 가능성을 갖추기 위해서는, 모델의 크기와 연산량을 효과적으로 줄이는 방안이 필요하다. 이를 위해 다양한 경량화 접근법이 제안되어 왔으며<sup>[1][2]</sup>, 그중 대표적인 방법으로 가지치기(Pruning)<sup>[3]</sup>와 지식 증류(Knowledge Distillation)<sup>[4]</sup>가 있다. 가지치기는 모델 내에서 중요도가 낮은 가중치나 노드를 제거하여 모델의 크기와 연산량을 줄이는 기법이며, 지식 증류는 성능이 우수한 교사 모델(Teacher Model)의 지식을 작은 학생 모델(Student Model)로 전달하여 성능 저하를 최소화하면서 모델을 경량화하는 방법이다.

본 논문에서는 이러한 필요성에 부응하여 생성적 적대 신경망 기반 초해상화 모델의 경량화를 위해 가지치기와 지식 증류를 결합한 새로운 프레임워크를 제안한다. 이 프레임워크는 먼저 교사 모델에 가지치기를 적용하여 불필요한 가중치를 제거하고, 이후 교사 모델의 은닉층에서 출력되는 특징 정보를 학생 모델에 전달함으로써 성능 저하를 최소화하는 것을 목표로 한다. 제안된 방법은 기존의 생성적 적대 신경망 기반 초해상화 모델이 가지고 있는 연산 자원 문

제를 해결하고, 경량화된 모델이 기존 모델과 유사하거나 더 나은 성능을 보일 수 있음을 실험적으로 검증한다.

또한 GAN 모델의 경량화가 제한된 연산 자원 환경에서도 고품질의 초해상화 이미지를 생성할 수 있는 가능성을 제시하며, 특히 모바일 기기와 같은 응용 분야에서의 실용적 적용을 가능하게 한다. 본 논문에서는 관련 연구를 통해 기존 경량화 기법들의 한계와 문제점을 살펴본 후, 제안된 프레임워크의 세부 구조와 그 효과를 설명한다. 또한, 실험을 통해 제안된 방법의 성능을 검증하고, 그 결과를 분석하여 GAN 모델 경량화의 새로운 방향을 제시한다.

## II. 관련 연구

### 1. 초해상화 모델에 대한 연구

딥러닝 기술의 발전에 따라, 컨볼루션 신경망(Convolutional Neural Network, CNN)을 이용한 초해상화 기법이 큰 주목을 받기 시작했다. Super-Resolution Convolutional Neural Network(SRCNN)<sup>[5]</sup>은 저해상도 이미지를 고해상도로 변환하는 최초의 CNN 기반 모델 중 하나로, 전통적인 방법에 비해 PSNR 지표에서 평균 3.6dB 이상의 성능 향상을 이루었다. 이 모델은 3개의 간단한 컨볼루션 레이어를 사용하여 저해상도 이미지를 처리하며, CNN의 특징 추출 능력을 통해 기존 방식들보다 뛰어난 성능을 보였다. 이후 Very Deep Super-Resolution(VDSR)<sup>[6]</sup>과 Enhanced Deep Super-Resolution(EDSR)<sup>[7]</sup> 모델이 제안되었으며, 이들 모델은 네트워크 깊이를 증가시켜 더 높은 품질의 이미지를 생성할 수 있게 되었다.

VDSR은 깊은 네트워크 구조를 통해 복잡한 이미지 특징을 학습함으로써 성능을 높였고, EDSR은 이러한 VDSR의 성능을 개선하기 위해 잔차 블록(Residual Blocks)을 사용하고, 불필요한 네트워크 요소를 제거하여 효율성을 높였다. 이와 같은 모델들은 높은 성능을 제공하지만, 네트워크의 깊이와 복잡성이 증가함에 따라 연산 비용이 크게 증가하게 되었다.

또한 생성적 적대 신경망(GAN)이 등장함에 따라, 이를 초해상화에 적용한 Super-Resolution GAN(SRGAN)<sup>[8]</sup>은 기존의 CNN 기반 모델과 달리 고주파 세부 정보를 잘 복원

a) 경북대학교(Kyungpook National University)

b) 경희대학교(Kyung Hee University)

c) 성균관대학교(Sungkyunkwan University)

‡ Corresponding Author : 박상호(Sang-hyo Park)

E-mail: s.park@knu.ac.kr

Tel: +82-53-950-6373

ORCID: <https://orcid.org/0000-0002-7282-7686>

※ 이 논문의 연구 결과 중 일부는 한국방송-미디어공학회 2024년 하계 학술대회에서 발표한 바 있음.

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임. (No.RS-2022-00167169, 이동형 로봇 기반 실사 메타버스 실감형 비디오의 획득 및 처리 기술 개발)

· Manuscript September 4, 2024; Revised October 21, 2024; Accepted October 22, 2024.

하는 특징을 가지고 있다. SRGAN은 저해상도 이미지를 입력으로 받아 고해상도 이미지를 생성하며, 생성망(Generator)과 판별망(Discriminator)으로 구성되어 있다. 이 모델은 고주파 영역에서 시각적으로 더 자연스러운 이미지를 생성한다는 특징이 있다.

ESRGAN(Enhanced SRGAN)<sup>[9]</sup>은 SRGAN의 성능을 개선한 모델로, Residual-in-Residual Dense Block(RRDB)을 사용하여 네트워크의 학습 안정성을 높이고, Perceptual Loss와 GAN Loss를 결합하여 더욱 자연스러운 이미지를 생성할 수 있다. 이 모델은 PSNR과 SSIM 지표에서 소폭의 성능 향상과 더불어, 시각적으로도 자연스러운 결과를 도출하였다. 그러나 RRDB에서 사용되는 완전 연결층으로 인해 깊고 복잡한 구조를 지니게 되는 ESRGAN은 SRGAN에 비해 많은 연산량을 요구하고, 이를 경량화하는 연구가 필수적이다.

## 2. 초해상화 모델에서의 경량화 적용 연구

### 2.1. 가지치기(Pruning)

가지치기(Pruning)는 신경망에서 중요도가 낮은 가중치나 노드를 제거하여 모델의 크기와 연산량을 줄이는 기법으로, 주로 경량화를 목표로 하는 연구에서 활용된다. 초해상화 분야에서도 가지치기 기법이 적용되어 모델의 효율성을 높이려는 시도가 있어 왔다<sup>[10]</sup>. 그러나 고해상도 이미지 복원 시 세부 정보 손실이 발생하여 PSNR 및 SSIM 성능이 저하되는 한계를 보였다. 이러한 결과는 가지치기 기법이 SR 모델의 경량화에 유용할 수 있지만, 중요한 구조적 요소를 제거할 위험이 있다는 점에서 신중한 설계가 필요함을 시사한다.

### 2.2. 지식 증류(Knowledge Distillation)

지식 증류(Knowledge Distillation)는 큰 모델(교사 모델, Teacher Model)의 성능을 작은 모델(학생 모델, Student Model)에 전달하여, 학생 모델이 적은 파라미터로도 높은 성능을 유지할 수 있도록 하는 기법이다. SR 분야에서 지식 증류 기법은 주로 CNN 기반 초해상화 모델에 적용되어 경량화를 시도한 사례가 있다<sup>[11]</sup>. 그러나 지식 증류 과정에서 학생 모델이 교사 모델의 성능을 완전히 모방하지 못하여 세부 이미지 품질에서 미세한 손실이 발생하는 한계가 있

다. 또한 전통적인 지식 증류 방식인 응답-기반 지식 증류 방식은 깊은 네트워크에서는 효과적으로 지식을 전달하지 못하는 한계가 존재한다<sup>[12]</sup>.

본 연구는 SR 분야에서 가지치기와 지식 증류 기법이 각각 가진 한계를 극복하기 위해 두 기법을 결합한 새로운 방법을 제안한다. 가지치기를 통해 교사 모델에서 불필요한 가중치를 제거하고, 지식 증류를 통해 핵심적인 지식만을 학생 모델에 전달함으로써 경량화와 성능 유지 간의 균형을 달성하고자 한다. 특히, 본 연구에서는 다양한 가지치기 비율과 은닉층 구성을 조합하여 최적의 경량화 성능을 도출하기 위한 실험을 수행하였다. 이는 기존 연구들이 가지치기와 지식 증류를 각각 독립적으로 적용한 것과 달리, 두 기법의 시너지를 극대화하는 접근으로, SR 모델 경량화의 새로운 가능성을 제시한다. 본 연구는 SR 분야에서 가지치기와 지식 증류 기법이 각각 가진 한계를 극복하기 위해 두 기법을 결합한 새로운 방법을 제안한다. 가지치기를 통해 교사 모델에서 불필요한 가중치를 제거하고, 지식 증류를 통해 핵심적인 지식만을 학생 모델에 전달함으로써 경량화와 성능 유지 간의 균형을 달성하고자 한다. 특히, 본 연구에서는 다양한 가지치기 비율과 은닉층 구성을 조합하여 최적의 경량화 성능을 도출하기 위한 실험을 수행하였다. 이는 기존 연구들이 가지치기와 지식 증류를 각각 독립적으로 적용한 것과 달리, 두 기법의 시너지를 극대화하는 접근으로, SR 모델 경량화의 새로운 가능성을 제시한다.

## III. 제안 방법

생성적 적대 신경망 기반 초해상화 모델은 PSNR과 SSIM 등의 지표에서 우수한 성능을 보이며, 시각적으로도 자연스러운 결과를 도출한다. 그러나 높은 연산량과 복잡한 구조를 가지고 있어 임베디드 또는 모바일 환경에서의 적용에 어려움이 있다. 이를 경량화하기 위한 노력으로 본 논문에서는 교사 모델로 생성적 적대 신경망 기반 초해상화 모델로 ESRGAN을 사용하였다. 제안하는 프레임워크는 그림 1과 같은 구조를 지니고 있다. 교사 모델은 각각 0.0, 0.3, 0.5, 0.8의 비율로 가지치기 기법을 적용하였고, 학생 모델의 성능 저하를 최소화하면서 교사 모델의 지식

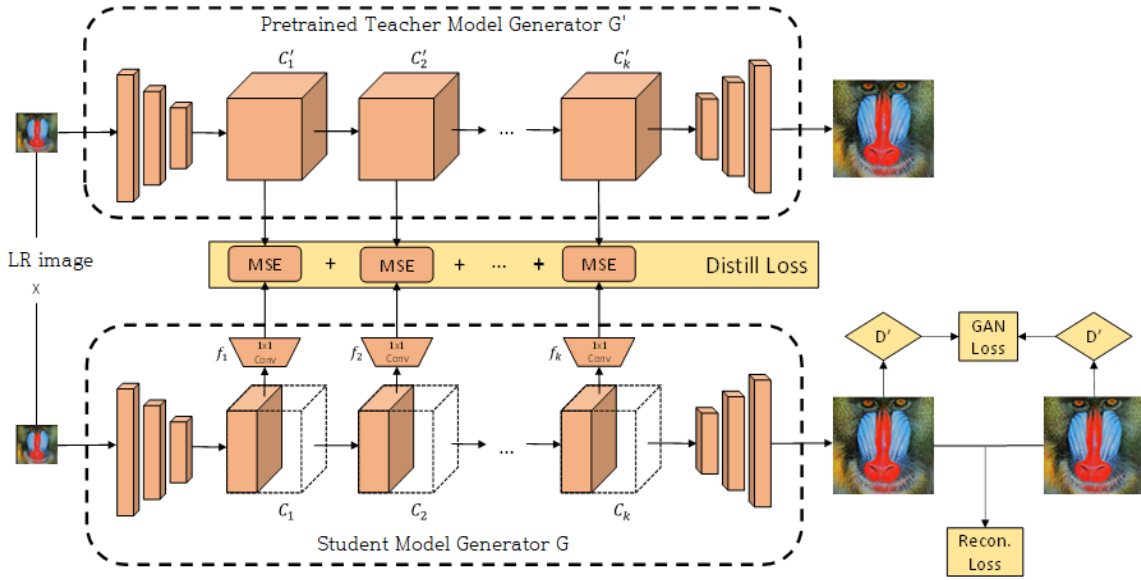


그림 1. 제안하는 교사-학생 모델의 지식 증류 과정  
 Fig. 1. Proposed Teacher-Student Model Knowledge Distillation Process

을 효과적으로 전달하기 위해 교사 모델의 각 층의 특징을 추출하여 학생 모델에 지식을 증류한다. 학생 모델은 교사 모델과 동일한 구조를 가지며, 학생 모델은  $n$ 개의 필터 파라미터를 받게 된다. 기존의 교사 모델은 64개의 필터 파라미터를 입력으로 받으며 학생 모델은 각각 16, 32, 48의 필터 파라미터를 받도록 하여 경량화를 하였다. 학생 모델로 효과적으로 지식을 전달하기 위해서는 다양한 목적 함수를 필요로 하고 복원 목적 함수, 생산적 적대 신경망 목적 함수와, 특징 증류 목적 함수로 구분된다. 특징 증류 목적 함수는 교사 모델과 학생 모델의 각 층의 특징을 비교하기 위해 설계한 목적 함수이다. 3개의 목적 함수에 대한 손실 함수를 구한 뒤 모두 합하여 최종 손실 함수를 정의한다.

## 1. 학습 목적 함수

### 1.1 복원 목적 함수

생성적 적대 신경망 모델<sup>[13]</sup>은 일반적으로 목표값인  $y$ 와 입력값인  $x$ 를 매핑하는 것에 대한 학습을 진행한다. 초해상화 모델을 학습할 때도 마찬가지로 학생 모델의 출력인  $G(x)$ 와 목표값  $y$ 의 차이를 기반으로 한 복원 목적 함수 (reconstruction objective function)를 계산하고, 다음과 같

은 식으로 표현할 수 있다.

$$\mathcal{L}_{\text{recon}} = \mathbb{E}_{x,y} \|G(x) - y\|_1 \quad (1)$$

여기서  $\|\cdot\|_1$ 은 L1 norm이다.

이를 통해 목표값과  $G(x)$ 의 손실 함수를 구하고, 학생 모델이 더욱 안정적으로 학습할 수 있도록 한다.

### 1.2 생성적 적대 신경망 목적 함수

지식 증류 과정에서는 미리 학습된 교사 모델의 판별망  $D'$ 이 사전 학습 모델로서 학생 모델의 판별망으로 사용된다. 지식 증류가 진행되면서 해당 판별망은 미세 조정을 통해 학생 모델의 생성망이 효과적으로 학습할 수 있도록 한다. 전통적인 생성적 적대 신경망의 목적 함수와 같이, 다음과 같은 식으로 목적 함수( $\mathcal{L}_{\text{GAN}}$ )를 표현한다.

$$\mathcal{L}_{\text{GAN}} = \mathbb{E}_{x,y} [\log D'(x,y)] + \mathbb{E}_x [\log(1 - D'(x, G(x)))] \quad (2)$$

학생 모델의 판별망은 교사 모델의 판별망  $D'$ 을 그대로 사용하였다. 교사 모델의 판별망  $D'$ 는 학생 모델의 생성망  $G$ 와 함께 표준 최소-최대 정규화를 통해 학습한다.

### 1.3 특징 증류 목적 함수

교사 모델의 지식을 출력값에서만 가져오는 것이 아닌 중간 특징 맵에서부터 1x1 합성곱 계층인  $f_i$ 로 교사 모델의 각 계층인  $C'_i$ 과 대응하는 학생 모델의 계층  $C_i$ 를 서로 연결한다. 교사 모델의 가중치를 단순히 전달하는 것이 아닌 학습가능한 계층을 통해 전달함으로써 효과적으로 지식을 전달할 수 있도록 한다. 교사 모델과 학생 모델 각각의 계층에 대하여 증류 목적 함수( $\mathcal{L}_{distill}$ )를 구할 수 있고 다음과 같은 식으로 표현할 수 있다.

$$\mathcal{L}_{distill} = \sum_{i=1}^T \|f_i(C_i(x)) - C'_i(x)\|_2 \quad (3)$$

1x1 학습 가능한 합성곱 계층  $f_i$ 는 학생 모델의 특징 계층  $C_i$ 의 채널 수를 대응하는 교사 모델의 특징 계층  $C'_i$ 의 채널 수와 같게 만들어 목적 함수를 구할 수 있는 형태로 변환한다.

### 1.4 종합 목적 함수

제안 방법의 최종 목적 함수는 다음과 같이 표현한다.

$$\mathcal{L} = \lambda_{recon} \mathcal{L}_{recon} + \mathcal{L}_{GAN} + \lambda_{distill} \mathcal{L}_{distill} \quad (4)$$

복원 비율 조정 파라미터와 증류 비율 조정 파라미터인  $\lambda_{recon}$ 과  $\lambda_{distill}$ 을 조정하여 각 목적 함수의 중요도에 가중치를 부여한다.

## 2. 가지치기 기법

본 논문에서 교사 모델에 가지치기 기법을 적용하였다. 가지치기 기법은 모델 경량화에 주로 사용되는 방식으로, 필요 없는 가중치를 제거하여 모델의 학습 시간과 연산량을 감소시킨다. 하지만 지식 증류 기법 이전에 교사 모델에 가지치기 기법을 적용하는 것은 또 다른 의미가 있다. 본 논문에서는 지식 증류를 위해 학생 모델의 각 계층에 1x1 합성곱 계층을 대응하는 교사 모델의 계층에 연결하여 목적 함수를 구한다고 언급했다. 이때 교사 모델의 불필요한 가중치를 제거한 상태로 목적 함수를 구해 더욱 양질의 지

식만 학생 모델에게 전달할 수 있도록 하였다. 따라서 지식 증류 이전에 가지치기 기법을 적용하는 것은 전체적인 학습 시간을 줄여줄 뿐만 아니라, 학생 모델의 학습의 질에도 영향을 미친다.

## IV. 모델 실험

### 1. 모델 설계

교사 모델은 ESRGAN 모델의 구조에서 가지치기 비율을 0.0, 0.3, 0.5, 0.8로 적용하였고, 해당 교사 모델을 기준으로 학생 모델에 지식 증류 기법을 적용하였다. 그 과정에서 학생 모델의 생성망의 은닉층 필터 수를 기존 교사 모델이 가지는 64개에서 각각 16, 32, 48개로 두어 실험하였다. 교사 모델의 파라미터 수는 가지치기가 적용되지 않은 모델이 16.7 Million(M)개이고, 학생 모델의 경우 필터 수 변수에 따라 각각 6.5M, 9.6M, 12.9M개로 가지치기를 하지 않은 교사 모델의 39%, 57%, 77% 크기로 경량화하였다.

### 2. 실험 환경

실험을 위해 학습 데이터로 DIV2K<sup>[14]</sup> 데이터셋을 활용하였고 테스트 데이터로는 BSD100<sup>[15]</sup>, Set14<sup>[16]</sup>, MIV<sup>[17]</sup>, KHU RGBD Dataset<sup>[18]</sup>, VRroom<sup>[19]</sup>을 사용하였다. 학습 데이터셋들은 고해상도 이미지인 원본 이미지와 Bicubic 보간법으로 4배 다운샘플링된 저해상도 이미지로 구성되어 있다. 실험은 Ubuntu 22.04에서 RTX 3090 1대를 사용해 교사 모델 학습, 지식 증류 각각 100,000 이터레이션으로 진행하였다.

### 3. 실험 결과

본 연구에서는 제안된 모델의 성능을 평가하기 위해 PSNR, SSIM<sup>[20]</sup>, Fréchet Inception Distance(FID)<sup>[21]</sup>, Learned Perceptual Image Patch Similarity(LPIPS)<sup>[22]</sup> 네 가지 지표를 사용하였다. 각 지표는 모델이 생성한 초해상화 이미지의 품질을 객관적으로 평가하는 데 사용되며, 이를

표 1. 학생 모델의 은닉층 필터 파라미터 수가 16일 때의 교사 모델의 가지치기 비율에 따른 평균 성능 비교  
 Table 1. Comparison of average performance based on the pruning ratio of the teacher model when the number of hidden layer filter parameters in the student model is 16

Baseline	Student_16_	Student_16_	Student_16_	Student_16_
Metric	ESRGAN	Teacher_0.3	Teacher_0.5	Teacher_0.8
SSIM ↑	0.78	0.78	0.78	0.77
FID ↓	126.79	104.99	127.48	123.16

통해 교사 모델과 학생 모델의 성능을 비교하였다.

표 1은 가지치기 기법을 적용하지 않은 ESRGAN 모델과 가지치기 비율을 각각 0.3, 0.5, 0.8로 적용한 교사 모델에 대해 각각 은닉층 필터 수를 16으로 설정한 학생 모델에 지식 증류를 적용해 성능 비교를 한 표이다. 전체적으로 학생 모델의 평균 SSIM 수치는 비슷하였으나, 평균 FID 수치는 0.3 비율로 적용했을 때 104.99로 가장 낮은 수치를 보여주었다. 따라서 30% 가지치기 기법을 적용한 교사 모델에 지식 증류를 적용한 학생 모델을 제안 모델로서 제시한다. 표 2는 30% 가지치기 기법을 적용한 교사 모델과 해당 모

델에 지식 증류 기법을 적용한 학생 모델의 PSNR, SSIM, FID, LPIPS값을 나타내고 있다. 각 학생 모델은 은닉층 필터 수를 각각 16, 32, 48로 조절하여 지식 증류를 적용하였다. 학생 모델은 은닉층 필터 수에 따라 각각 Student\_16, Student\_32, Student\_48로 지칭한다. 표 1을 보면 전체적으로 교사 모델보다 학생 모델이 1dB 이상 더 좋은 PSNR을 가지고 있는데, 그중 평균 PSNR이 가장 높은 Student\_48과 교사 모델을 비교했을 때 PSNR과 SSIM 각각 1.23dB, 0.04만큼 학생 모델이 더 증가한 것을 볼 수 있고, LPIPS값은 상승하였지만 FID값은 최소 4에서 크게는 70 이상 차이

표 2. 교사 모델의 가지치기 비율이 0.3일 때의 학생 모델의 은닉층 필터 파라미터 수에 따른 성능 비교  
 Table 2. Performance comparison based on the number of hidden layer filter parameters in the student model when the teacher model's pruning ratio is 0.3

Metric		BSD100	Set14	KHU	MIV	VRroom	Urban100	Average
Baseline	Dataset							
Teacher_0.3(11.7M)	PSNR ↑	24.24	24.81	27.95	22.03	28.78	21.39	24.87
	SSIM ↑	0.66	0.70	0.83	0.69	0.86	0.67	0.74
	FID ↓	63.66	6.42	194.11	74.06	135.72	71.83	90.97
	LPIPS ↓	0.16	0.12	0.09	0.13	0.08	0.15	0.12
Student_16(6.5M)	PSNR ↑	25.34	25.57	29.73	23.20	29.78	22.24	25.98
	SSIM ↑	0.71	0.75	0.87	0.74	0.90	0.70	0.78
	FID ↓	64.71	9.80	160.66	154.91	114.95	124.91	104.99
	LPIPS ↓	0.33	0.24	0.16	0.28	0.12	0.26	0.23
Student_32(9.6M)	PSNR ↑	25.37	25.58	29.69	23.20	29.72	22.25	25.97
	SSIM ↑	0.70	0.74	0.86	0.74	0.90	0.70	0.77
	FID ↓	96.71	11.26	176.85	155.26	129.16	201.91	128.53
	LPIPS ↓	0.33	0.24	0.17	0.28	0.13	0.25	0.23
Student_48(12.9M)	PSNR ↑	25.34	25.67	29.64	23.19	30.00	22.28	26.02
	SSIM ↑	0.70	0.75	0.87	0.74	0.90	0.70	0.78
	FID ↓	59.41	10.16	147.34	149.33	134.40	113.05	102.28
	LPIPS ↓	0.32	0.23	0.16	0.27	0.12	0.25	0.23



그림 2. 생성 이미지별 PSNR/SSIM값

Fig. 2. PSNR/SSIM values for each generated image

나는 것을 볼 수 있다. 이에 대한 시각적 결과는 그림 2에서 확인해 볼 수 있다. 또한 세 학생 모델 모두 기존 모델의 39%, 57%, 77% 크기로 편차가 큰 파라미터 개수를 가지지만, 학생 모델 각각의 PSNR, SSIM, FID, LPIPS 수치는 차이가 미미하다. 오히려 파라미터 수가 가장 적은 Student\_16 모델의 경우 특정 데이터셋(BSD100, KHU, MIV)에서 파라미터 수가 더 많은 다른 학생 모델보다도 더 높은 PSNR과 SSIM값을 가진다는 결과를 나타낸다. 이는 적은 파라미터를 가지는 학생 모델이 특정 부분에서는 초해상화가 잘 이루어진다는 것을 의미하며, 데이터 종류에 따라 최적의 파라미터 수가 존재할 가능성도 있다는 것을 보여준다.

## V. 결론

본 논문에서는 생성적 적대 신경망 기반 초해상화 모델에서의 지식 증류 및 가지치기 기법을 제시하고 교사 모델과 다양한 필터 파라미터를 가진 학생 모델에 해당 기법을 적용함으로써 모델에 대한 평가를 진행하였다. 제안된 방법은 먼저 교사 모델에 가지치기 기법을 적용하여 불필요한 가중치를 제거하고, 이후 지식 증류 기법은 기존에 주로 사용하던 학생 모델과 교사 모델의 출력값을 기준으로 목적 함수를 정의하는 것이 아닌 각 모델의 은닉층의 출력값을 직접적으로 비교하는 방식으로써, 모델의 파라미터 수를 크게 줄이면서도 성능 저하를 최소화하는 것을 목표로 하였다. 또한 이를 다양한 데이터셋에 대해 평가하여 실험

적으로 검증하였다. 이는 초해상화 모델이 모바일 기기나 임베디드 시스템과 같은 연산 자원이 제한된 환경에서도 활용될 수 있는 가능성을 열어주며, 실용적 응용에 중요한 기여를 할 것으로 기대된다. 향후 연구로는 제안된 방법을 ESRGAN 외의 다른 생성적 적대 신경망 기반 초해상화 모델에도 적용하여 그 효용성을 검증하는 것이 필요하다. 또한, 다양한 데이터셋과 응용 환경에서 최적의 가지치기 비율과 학생 모델 구조를 찾아내는 작업도 진행될 예정이다. 이를 통해, 더욱 다양한 상황에서 적용 가능한 경량화 초해상화 모델을 개발할 수 있을 것이다.

## 참고 문헌 (References)

- [1] Hu, T., Lin, M., You, L., Chao, F., & Ji, R. "Discriminator-Cooperated Feature Map Distillation for GAN Compression," *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vol. 20351, No. 20360, pp. 20351-20360, June 2023.  
doi: <https://doi.org/10.1109/cvpr52729.2023.01949>
- [2] Li, S., Wu, J., Xiao, X., Chao, F., Mao, X., & Ji, R. "Revisiting Discriminator in GAN Compression: A Generator-Discriminator Cooperative Compression Scheme," *Advances in Neural Information Processing Systems*, Vol. 34, pp. 28560-28572, 2021.
- [3] He, Y., Zhang, X., & Sun, J. "Channel Pruning for Accelerating Very Deep Neural Networks," *IEEE International Conference on Computer Vision*, Vol. 1389, No. 1397, pp. 1389-1397, 2017.  
doi: <https://doi.org/10.1109/iccv.2017.155>
- [4] Gou, J., Yu, B., Maybank, S. J., & Tao, D. "Knowledge Distillation: A Survey," *International Journal of Computer Vision*, Vol. 129, No. 6, pp. 1789-1819, June 2021.  
doi: <https://doi.org/10.1007/s11263-021-01453-z>
- [5] Dong, C., Loy, C. C., He, K., & Tang, X. "Learning a Deep



- Convolutional Network for Image Super-Resolution,” *Computer Vision -ECCV 2014: 13th European Conference, Springer International Publishing, Zurich, Switzerland*, pp. 184-199, 2014.
- [6] Kim, J., Lee, J. K., & Lee, K. M. “Accurate Image Super-Resolution Using Very Deep Convolutional Networks,” *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1646, No. 1654, pp. 1646-1654, 2016.  
doi: <https://doi.org/10.1109/cvpr.2016.182>
- [7] Lim, B., Son, S., Kim, H., Nah, S., & Mu Lee, K. “Enhanced Deep Residual Networks for Single Image Super-Resolution,” *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Vol. 136, No. 144, pp. 136-144, 2017.  
doi: <https://doi.org/10.1109/cvprw.2017.151>
- [8] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network,” *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 4681, No. 4690, pp. 4681-4690, 2017.  
doi: <https://doi.org/10.1109/cvpr.2017.19>
- [9] Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., ... & Change Loy, C. “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks,” *European Conference on Computer Vision (ECCV) Workshops*, pp. 0-0, 2018.  
doi: [https://doi.org/10.1007/978-3-030-11021-5\\_5](https://doi.org/10.1007/978-3-030-11021-5_5)
- [10] Jiang, X., Wang, N., Xin, J., Xia, X., Yang, X., & Gao, X. “Learning Lightweight Super-Resolution Networks with Weight Pruning,” *Neural Networks*, Vol. 144, pp. 21-32, August 2021.  
doi: <https://doi.org/10.1016/j.neunet.2021.08.002>
- [11] Gao, Q., Zhao, Y., Li, G., & Tong, T. “Image Super-Resolution Using Knowledge Distillation,” *Asian Conference on Computer Vision, Springer International Publishing*, pp. 527-541, December 2018.
- [12] Gou, J., Yu, B., Maybank, S. J., & Tao, D. Knowledge distillation: A survey. *International Journal of Computer Vision*, 129(6), 1789-1819. 2021.  
doi: <https://doi.org/10.1007/s11263-021-01453-z>
- [13] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. “Generative Adversarial Nets,” *Advances in Neural Information Processing Systems*, Vol. 27, 2014.  
doi: [https://doi.org/10.1007/978-3-658-40442-0\\_9](https://doi.org/10.1007/978-3-658-40442-0_9)
- [14] Agustsson, E., & Timofte, R. “NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study,” *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Vol. 126, No. 135, pp. 126-135, 2017.  
doi: <https://doi.org/10.1109/cvprw.2017.150>
- [15] Martin, D., Fowlkes, C., Tal, D., & Malik, J. “A Database of Human Segmented Natural Images and Its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics,” *IEEE International Conference on Computer Vision*, Vol. 416, No. 423, pp. 416-423, July 2001.  
doi: <https://doi.org/10.1109/iccv.2001.937655>
- [16] Zeyde, R., Elad, M., & Protter, M. “On Single Image Scale-Up Using Sparse-Representations,” *Curves and Surfaces: 7th International Conference, Springer Berlin Heidelberg, Avignon, France*, pp. 711-730, 2012.  
doi: [https://doi.org/10.1007/978-3-642-27413-8\\_47](https://doi.org/10.1007/978-3-642-27413-8_47)
- [17] ISO/IEC JTC 1/SC 29/WG 04 N 0307. “Common Test Conditions for MPEG Immersive Video”, February 2023.
- [18] KHU airtlab “Multi-agent based RGBD Dataset”. December 2023.
- [19] ISO/IEC JTC 1/SC 29/WG 04 m64721, “Report with New Natural INVR Video Contents: SKKU\_VRroom”. October 2023.
- [20] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. “Image quality assessment: from error visibility to structural similarity”. *IEEE transactions on image processing*, 13(4), 600-612. 2004.  
doi: <https://doi.org/10.1109/tip.2003.819861>
- [21] Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. “Gans trained by a two time-scale update rule converge to a local nash equilibrium”. *Advances in neural information processing systems*, 30. 2017.  
doi: <https://doi.org/10.18034/ajase.v8i1.9>
- [22] Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. “The unreasonable effectiveness of deep features as a perceptual metric”. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 586-595). 2018.  
doi: <https://doi.org/10.1109/cvpr.2018.00068>

---

## 저 자 소 개



### 백 보 성

- 2020년 3월 ~ 2024년 2월 : 경북대학교 전자공학부 학사
- 2024년 3월 ~ 현재 : 경북대학교 컴퓨터학부 석사과정
- ORCID : <https://orcid.org/0009-0004-4762-0079>
- 주관심분야 : 컴퓨터비전, 인공지능, 영상처리



---

저 자 소 개

---



**김 아 로**

- 2023년 2월 : 경북대학교 응용생명과학과 졸업
- 2023년 3월 ~ 현재 : 경북대학교 컴퓨터학부 석사과정
- ORCID : <https://orcid.org/0009-0001-8585-5993>
- 주관심분야 : 컴퓨터비전, 인공지능, 영상처리



**장 재 윤**

- 2024년 2월 : 경희대학교 소프트웨어융합학과 공학사
- 2024년 3월 ~ 현재 : 경희대학교 소프트웨어융합학과 석사과정
- ORCID : <https://orcid.org/0009-0001-5832-3879>
- 주관심분야 : 강화학습, 로보틱스



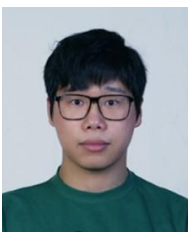
**황 호 석**

- 2004년 : 연세대학교 기계공학과 공학사
- 2009년 : KAIST 로보틱스 프로그램 공학석사
- 2017년 : KAIST 전기및전자공학과 공학박사
- 2009년 ~ 2018년 : 삼성전자 삼성종합기술원 Multimedia Processing Lab 선임연구원
- 2018년 ~ 2021년 : 가천대학교 IT대학 소프트웨어학과 조교수
- 2021년 ~ 현재 : 경희대학교 소프트웨어융합대학 소프트웨어융합학과 부교수
- ORCID : <https://orcid.org/0000-0003-3241-8455>
- 주관심분야 : 컴퓨터 비전, 로보틱스



**최 재 열**

- 2018년 3월 ~ 2024년 2월 : 성균관대학교 컴퓨터교육학과 학사
- 2024년 3월 ~ 현재 : 성균관대학교 인공지능융합학과 석사과정
- 2023년 1월 ~ 2023년 2월 : 한국전자통신연구원 학생연구원
- ORCID : <https://orcid.org/0000-0009-2923-1252>
- 주관심분야 : 실감미디어, 인공지능, 그래픽스, 멀티미디어 통신 및 시스템



**류 영 일**

- 2016년 2월 : 가천대학교 컴퓨터공학과 학사
- 2018년 2월 : 가천대학교 컴퓨터공학 석사
- 2021년 3월 ~ 현재 : 성균관대학교 컴퓨터교육학과 박사과정
- ORCID : <http://orcid.org/0000-0002-9210-6245>
- 주관심분야 : 멀티미디어 통신 및 시스템, 비디오 압축 표준

---

저 자 소 개

---



**류 은 석**

- 1999년 8월 : 고려대학교 컴퓨터학과 학사
- 2001년 8월 : 고려대학교 컴퓨터학과 석사
- 2008년 2월 : 고려대학교 컴퓨터학과 박사
- 2004년 3월 ~ 2008년 8월 : 고려대학교 연구교수
- 2008년 8월 ~ 2010년 12월 : 조지아공대 박사후과정
- 2011년 2월 ~ 2014년 2월 : InterDigital Labs Staff Engineer
- 2014년 3월 ~ 2015년 2월 : 삼성전자 수석연구원/파트장
- 2015년 3월 ~ 2019년 8월 : 가천대학교 컴퓨터공학과 조교수
- 2019년 9월 ~ 2023년 6월 : 성균관대학교 컴퓨터교육과 부교수
- 2023년 7월 ~ 현재 : 성균관대학교 실감미디어공학과 부교수
- ORCID : <https://orcid.org/0000-0003-4894-6105>
- 주관심분야 : 멀티미디어 통신 및 시스템, 비디오 코딩 및 국제 표준, HMD/VR 응용분야



**박 상 호**

- 2011년 2월 : 한양대학교 컴퓨터전공 학사
- 2017년 8월 : 한양대학교 컴퓨터·소프트웨어학과 박사
- 2017년 5월 ~ 2018년 2월 : 전자부품연구원 지능형영상처리센터 Post-doc
- 2018년 3월 ~ 2018년 12월 : 연세대학교 바른ICT연구소 연구원
- 2019년 2월 ~ 2020년 1월 : 이화여자대학교 전자전기공학과 박사후연구원
- 2020년 3월 ~ 현재 : 경북대학교 컴퓨터학부 부교수
- ORCID : <https://orcid.org/0000-0002-7282-7686>
- 주관심분야 : VVC, Encoding/Decoding Complexity, Omnidirectional Video, Deep Learning, Generative Model