

임의배율 초해상도를 위한 하이브리드 도메인 고주파 집중 네트워크

윤준석¹ · 이성진¹ · 유석봉^{2*} · 한승희^{3*}

Hybrid-Domain High-Frequency Attention Network for Arbitrary Magnification Super-Resolution

Jun-Seok Yun¹ · Sung-Jin Lee¹ · Seok Bong Yoo^{2*} · Seunghwoi Han^{3*}

¹Graduate Student, Department of AI Convergence, Chonnam National University, Gwangju, 61186 Korea

^{2*}Assistant Professor, Department of AI Convergence, Chonnam National University, Gwangju, 61186 Korea

^{3*}Assistant Professor, School of Mechanical Engineering, Chonnam National University, Gwangju, 61186 Korea

요약

최근 이미지 초해상도는 정수배율만 가능한 모델에만 집중적으로 연구되고 있다. 하지만 관심 객체 인식, 디스플레이 화질 개선 등 실제 초해상도 기술의 대표 적용 분야에서는 소수 배율을 포함하는 임의배율 확대 필요성이 대두되고 있다. 본 논문에서는 기존 정수배율 모델의 가중치를 활용하여 임의배율을 실행할 수 있는 모델을 제안한다. 이 모델은 정수배율에 의해 우수한 성능을 가진 초해상도 결과를 DCT 스펙트럼 도메인으로 변환하여 임의배율을 위한 공간을 확장한다. DCT 스펙트럼 도메인에 의한 확장으로 인해 발생하는 이미지의 고주파 정보 손실 문제를 줄이기 위해 고주파 스펙트럼 정보를 적절히 복원할 수 있는 모델인 고주파 집중 네트워크를 제안한다. 제안된 네트워크는 고주파 정보를 제대로 생성하기 위해서 RGB 채널간의 상관관계를 학습하는 레이어인 channel attention을 활용하고, 잔차 학습 구조를 통해 모델을 깊게 만들어 성능을 향상시켰다.

ABSTRACT

Recently, super-resolution has been intensively studied only on upscaling models with integer magnification. However, the need to expand arbitrary magnification is emerging in representative application fields of actual super-resolution, such as object recognition and display image quality improvement. In this paper, we propose a model that can support arbitrary magnification by using the weights of the existing integer magnification model. This model converts super-resolution results into the DCT spectral domain to expand the space for arbitrary magnification. To reduce the loss of high-frequency information in the image caused by the expansion by the DCT spectral domain, we propose a high-frequency attention network for arbitrary magnification so that this model can properly restore high-frequency spectral information. To recover high-frequency information properly, the proposed network utilizes channel attention layers. This layer can learn correlations between RGB channels, and it can deepen the model through residual structures.

키워드 : 이미지 초해상도, 임의배율, 고주파 집중, 스펙트럼 도메인

Keywords : Image super-resolution, Arbitrary magnification, High-frequency attention, Spectral domain

Received 18 August 2021, Revised 23 August 2021, Accepted 10 September 2021

* Corresponding Author Seok Bong Yoo(E-mail:sbyoo@jnu.ac.kr, Tel:+82-62-530-3437), Seunghwoi Han(E-mail:shan@jnu.ac.kr, Tel:+82-62-530-5361)
Assistant Professor, Department of AI Convergence, Chonnam National University, Gwangju, 61186 Korea
Assistant Professor, School of Mechanical Engineering, Chonnam National University, Gwangju, 61186 Korea

Open Access <http://doi.org/10.6109/jkiice.2021.25.11.1477>

print ISSN: 2234-4772 online ISSN: 2288-4165

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

최근 컴퓨터 비전 분야 이미지 초해상도 관련 연구에서 심층 합성곱 신경망(Deep Convolutional Neural Network, DCNN)을 통해서 저해상도 이미지에 대한 우수한 고해상도 복원 능력을 보여주고 있다. 또한 초해상도 기술을 활용하여 저해상도의 이미지에서 다양한 컴퓨터 비전의 응용 분야의 성능을 향상시키는 연구가 진행되고 있다. 예를 들면, 이미지에서 관심 객체를 인식하기 위해서는 객체 검출 신경망을 통해 이미지에서 객체 영역을 검출한 후, 검출 영역을 객체 속성 인식 신경망에 크기에 알맞게 조정하는 과정이 필수적이다. 하지만 CCTV, 블랙박스, 드론 등을 사용하여 촬영된 현실 세계 이미지는 중, 원거리의 작은 객체 영역을 가지고 있고, 해당 이미지를 일반적인 보간법으로 크기 조절을 할 경우 블러 현상을 일으켜 객체 인식의 성능을 저하시킨다. 이에 대한 해결책으로써, 초해상도를 작은 객체 영역의 이미지에 적용하여 선명도를 향상시키며 크기를 확대함으로써 객체 인식 정확도 향상을 가져올 수 있다[1]. 한편, 기존 대부분의 초해상도 방법들은 다양한 크기의 작은 객체 영역을 목표크기로 변환할 수 없는 한계점을 가진다. 일반적인 초해상도 방법들은 정수배율(x2, x4)로만 복원하거나, 입력 이미지를 기본 보간법을 이용하여 소수 배율로 확대, 축소한 후 초해상도 신경망을 통해 임의배율을 수행하게 된다. 이 방법은 입력 이미지의 변형을 통한 복원 능력 소실과 연산량 증가 원인이 된다. 따라서 실제 초해상도 응용 환경에서는 임의배율이 가능한 초해상도 신경망이 필요하고, 이에 본 논문에서는 임의배율이 가능한 초해상도 네트워크를 다음과 같이 제안한다. 먼저 초해상도 신경망을 통해 정수배율을 진행하고 잔여 배율을 위한 공간을 Discrete Cosine Transform(DCT) 스펙트럼 도메인에서 확장시킨다. 이 때 확장시킨 DCT 공간들은 고주파 영역의 부분이다. 따라서 고주파 집중 네트워크를 통하여 공간 도메인에서 부족한 고주파를 채워주어 임의배율을 수행한다. 본 논문에서 제시한 임의배율 모델은 정수배율 모델의 우수함을 유지하고 추가적인 임의배율을 진행함으로써 다른 임의배율 방법들보다 우수한 복원 성능을 가질 수 있다. 본 논문의 핵심 내용은 다음과 같이 세 가지로 요약할 수 있다.

1. DCT 스펙트럼 도메인에서 공간 확장을 통하여 목표 해상도의 이미지로 확대한다.
2. 스펙트럼 도메인 공간 확장으로 인하여 부족한 고주파를 공간 도메인 고주파 집중 네트워크를 통해 고주파를 복원한다.
3. 기존 초해상도 모델의 우수성을 보존하며, 기존 모델에 대한 수정 및 추가 학습 없이 간편하게 하이브리드 도메인 잔여 소수 배율 모델을 추가함으로써, 임의배율 복원 성능을 보장한다.

II. 관련 연구 동향

2.1. 전통적인 초해상도

초해상도의 목적은 저해상도의 이미지를 입력으로 받아 대응되는 고해상도 이미지를 예측하는 것이다. 그러나 고해상도로부터 해상도 저하를 발생시키는 다양한 원인이 있기 때문에 이는 해결하기 어려운 문제이고, 이를 해결하기 위한 다양한 연구들이 진행되어 왔다. SRCNN[2]은 최초로 초해상도를 CNN을 활용하여 혁신적인 복원 성능을 보여주었다. 이를 발판 삼아 더 나은 초해상도 성능 확보를 위한 연구들이 진행되었다. VDSR[3]은 잔차 학습 전략을 통해 모델을 더욱 깊이 설계하였고, ESPCNN[4]은 입력으로 목표 배율의 이미지를 넣어주는 한계점을 극복하기 위해 학습할 수 있는 upsampling 모듈인 pixel-shuffling 층을 구현하였다. 입력의 이미지를 보간법으로 확대해서 넣어주지 않아도 되기 때문에 컴퓨팅 비용이 절약되고, 작은 크기의 필터들을 활용함에 따라 더 깊은 네트워크를 쌓을 수 있게 되었다. DBPN[5]은 이미지 업스케일링, 다운스케일링 하는 층을 반복적으로 쌓고 입력 이미지의 크기를 축소 및 확대를 반복하는 모델 구조를 만들어 보다 나은 성능을 보여 주었다. RCAN[6]은 깊은 모델을 만들기 위해 채널 집중 메커니즘을 도입하였고, DRN[7]은 모델 내부에 폐회로를 구성하여 기존의 고해상도 도메인 손실 함수 외에도 초해상도 결과 이미지를 다운스케일링하여 입력 이미지와의 차이를 계산하는 저해상도 도메인 손실 함수를 추가함으로써 복원 능력을 향상시켰다. RDN[8]은 잔차 밀집 구조를 통하여 모든 특징 맵의 계층적 표현을 학습하였고, SAN[9]은 이미지 특징 맵의 표현을 강력하게 향상시키고 특징 맵 간의 상호 의존성

을 학습함으로써 좋은 성능을 보여주었다. 또한 SRGAN [10]은 초해상도 성능을 향상시키기 위해 적대적 학습을 이용하였다. 이 모델은 생성자와 판별자 네트워크로 구성되어 있고, 생성자는 판별자가 판별할 수 없는 초해상도 출력을 만드는 것을 목적으로 한다. 최근에는 정규 분포에서 복잡한 확률분포를 예측하는 normalizing flow를 활용한 SRFlow[11]가 주목 받고 있다. 고해상도 이미지를 복잡한 확률분포로 변형한 뒤 야코비안 행렬 기반 손쉬운 미분이 가능한 점을 이용하여 저해상도 이미지의 정규분포로 점차 확률분포를 미분해 나간다. 또한 SRFlow에는 affine coupling 층이 존재하는 데, 이 층은 입력 값의 차원을 두 개로 나누어, 한 개 차원은 그대로 두고, 나머지 차원은 shift 및 affine 변환을 한다. 이 변환을 통해 역변환 및 야코비안 행렬식의 계산이 쉬워진다. 이러한 미분 과정을 학습하게 되고, 저해상도 이미지의 정규분포가 입력으로 들어오면 복잡한 확률분포로 만들어 줄 수 있다. 복잡한 확률분포는 flow-based generative model을 활용하여 확률분포에서 이미지를 생성하게 된다. SRFlow의 장점으로는 Generative Adversarial Network(GAN)에서 부족한 저해상도 이미지에서 고해상도 이미지를 만들 수 있는 다양성이 크다는 것이다. 또한 log likelihood loss를 사용하여 학습 시 발산을 방지하여 안정성을 보장하고, GAN의 생성자와 판별자를 따로 학습하는 것 대비 학습이 쉽다고 주장하였다. SRFlow-DA[12]는 SRFlow 모델의 receptive field를 확장하기 위해 합성곱 층을 6개 더 추가하고, 초해상도 구조에 맞지 않는 normalization 층을 제거함으로써 더 향상된 성능을 보였다. NCSR[13]은 학습 시 저해상도 이미지와 고해상도 이미지에 노이즈를 삽입하고, 노이즈로 인한 아티팩트를 제거하는 noise conditional layer를 제안하여 flow 모델의 다양성을 확장했다.

2.2. 임의배율 초해상도

대부분의 초해상도는 저해상도 이미지를 소수 배율로 확대할 때 비 학습 기반 보간법에 의존하였다. Shi et al[14]은 pixel-shuffling 층을 제안함으로써 학습이 가능한 확대방법을 제안했다. 이를 이용하여 VDSR은 저해상도 이미지를 목표 해상도로 확대한 후 초해상도 모델에 넣어서 임의배율을 할 수 있다. 그러나 이미지를 확대하여 신경망을 거치게 되면 많은 컴퓨팅 비용이 들고, 배율마다 큰 모델의 가중치를 가지고 있어야 하므로 그

성능은 정수배율에 더 특화되어 있고 볼 수 있다. Meta-SR[15]은 기존 초해상도 모델에서 확대하는 부분을 upscale-module로 대체함으로써 하나의 모델로 임의 배율을 진행할 수 있다고 주장한다. Meta-SR은 학습할 수 있는 weight prediction 층을 두어 정수배율을 확대할 수 있는 가중치를 예측한다. 이 가중치는 upscale-module에 적용되고 이미지를 정수배율로 확대한다. 확대한 이미지에 임의배율에 맞는 픽셀 mask를 이용하여 정수배로 늘린 이미지에서 임의배율에 맞는 사이즈에 맞게 픽셀값을 선택한다. 이는 k-neighborhood 알고리즘을 딥러닝 분야에 적용함으로써, 기존 알고리즘이 갖는 한계를 극복하였다. SRWarp[16]은 확대, 축소, 뒤틀림 등으로 warping 된 이미지를 입력값으로 받는다. 입력 이미지에 대하여 backbone에서 배율(x1, x2, x4) 별로 특징 맵을 추출하고, adaptive warping 층을 통해 특징 맵 이미지를 원래 형태로 복원할 수 있는 transform function을 예측한다. 이후 warping이 되지 않은 형태로 복원된 배율별 특징 맵은 각 배율(x1, x2, x4) 별로 가지고 있는 풍부한 정보를 이용한 multi-scale blending을 통해 결합한다. 이와 같이 SRWarp은 임의배율로 확대된 이미지를 multi-scale blending을 통해 임의배율이 가능한 신경망을 제안했다.

2.3. 주파수 도메인 초해상도

이미지는 다양한 주파수 도메인으로 변환할 수 있고, 이런 점에 착안하여 고해상도 이미지를 표현할 수 있는 다양한 주파수 정보를 CNN을 통해 예측하는 연구들이 진행되었다. Kumar et al[17]은 고해상도 이미지의 Wavelet 계수를 예측하는 Wavelet domain Super-Resolution (CNNWSR)을 제안했다. 예측된 Wavelet 계수는 2차원 inverse Discrete Wavelet Transform(DWT)를 이용해 고해상도 이미지를 재구성하는 데 사용한다. 단일 이미지를 재구성하는 기존의 초해상도 기법과는 달리, CNNWSR[18]은 3장의 이미지가 가지고 있는 Wavelet 계수를 합성곱 층을 통해 예측하여 이미지를 복원하였다. FNNSR[19]과 IFNNSR[20]은 Fourier domain에서 초해상도 문제를 해결하였다. FNNSR은 Rectified Linear Unit(ReLU) 활성화함수를 근사시키기 위해 단일 합성곱 층을 사용하여 스펙트럼 도메인에서 point-wise multiplication으로 파라미터 화하는 신경망을 공식화하였다. IFNNSR은 Fourier 변환 대신 Hartley 변환을 사용

하고, ReLU 활성화함수를 잘 근사시키기 위해 다중 합성곱 층을 사용하였으며, 새로운 가중 유클리드 손실을 제안하여 고주파 성분의 오류를 강조하였다. AYDIN et al[21]은 입력 이미지를 목표배율로 보간법을 통해 확장한 후 DCT 스펙트럼 도메인에서 Fully Connected(FC) 층을 통해 고해상도 이미지를 복원할 수 있는 DCT 계수를 예측한다. 이때 사용된 loss는 대응되는 High-Resolution (HR) 이미지에 대한 DCT 계수와와의 평균 제곱 오차로 정의되며, 이는 DCT 스펙트럼 도메인에서의 CNN 학습 가능성을 보여주었다.

III. 제안 방법

이번 섹션에서는 임의배율 초해상도를 위해 본 연구에서 제안한 방법에 대해 설명한다. 하위섹션 3.1에서는 DCT 개요에 대해서 먼저 설명하고, 섹션 3.2에서는 제안된 하이브리드 도메인 고주파 집중 네트워크를, 섹션 3.3에서는 본 네트워크에서 정의한 손실함수에 대해 설명한다.

3.1. 이산 코사인 변환 (DCT)

공간 도메인은 스펙트럼 도메인으로 변환할 수 있고, 반대로 스펙트럼 도메인에서 공간 도메인으로 변환도 가능하다. 가장 많이 이용하는 변환은 이산 푸리에 변환 (Discrete Fourier Transform, DFT)이다. DFT는 입력 신호가 실수인 경우에도 변환 결과가 복소수를 포함하게 된다. 복소수가 포함된 값을 산술을 통하여 계산할 수 있지만, 복잡한 산술 때문에 저비용 장치에 이를 적용할 때 문제가 발생한다. 따라서 저비용 장치에는 신호를 코사인 함수로 분해하고 스펙트럼 표현 시 실숫값만 산출

하는 이산 코사인 변환(DCT)을 널리 사용한다. 크기가 $N \times M$ 인 2차원의 공간 도메인 이산 신호 입력인 f 를 DCT를 통해 주파수 도메인 F 로 표현할 수 있고, 2차원 DCT는 다음의 수식으로 표현된다.

$$F[x,y] = a(x)a(y) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f[i,j] \gamma(i,j,x,y), \quad (1)$$

$$\gamma(i,j,x,y) = \cos\left(\frac{\pi(2i+1)x}{2N}\right) \cos\left(\frac{\pi(2j+1)y}{2M}\right), \quad (2)$$

$$a(x) = \begin{cases} \sqrt{\frac{1}{N}} & x = 0 \\ \sqrt{\frac{2}{N}} & x \neq 0 \end{cases}. \quad (3)$$

식 (1)에서 $f[i,j]$ 는 입력 이미지의 (i,j) 위치의 픽셀 값이며, $F[x,y]$ 는 (x,y) 위치에서의 DCT 계수 값이다. 식 (2)와 (3)은 코사인 기저함수와 정규화 상수의 정의를 각각 보여준다. 한편, 주파수 영역으로 변환된 신호는 식 (4)에 제시된 것과 같이, 2차원 Inverse DCT (IDCT)를 이용해서 공간 영역으로 변환할 수 있다.

$$f[i,j] = \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} a(x)a(y)F[u,v] \gamma(i,j,x,y). \quad (4)$$

그림 1은 샘플 이미지와 그 이미지에 대하여 2차원 DCT를 실행한 결과를 보여준다. 그림에서 볼 수 있듯이 DCT 스펙트럼의 이미지를 보면 공간 구조의 변형 때문에 직관적이지 않을지라도, 다양한 성분의 주파수 정보를 쉽게 관찰할 수 있다. DCT 도메인에서 이미지의 공간을 확장한 후 IDCT를 수행함으로써, 목표로 하는 공간 크기를 가지는 결과 이미지를 생성할 수 있다. 이때, DCT 공간 확장 시, 이미지를 좌상단 방향 혹은 우하단 방향으로 확장을 할 수 있는데, 좌상단 방향은 저주파 성분, 우하단 방향은 고주파 성분이 많이 몰려있기 때문에 확장하는 영역에 따라 이미지는 해당 주파수 정보가 부족한 이미지가 된다. 이미지 초해상도의 목표는 흐릿한 이미지를 선명한 이미지로 향상 시키는 것이고, 이는 이미지를 선명하게 만드는 고주파 성분을 복원시키는 것으로 볼 수 있다. 이에 따라, 본 논문에서는 DCT 도메인에서 고주파 영역으로 목표로 하는 배율로 이미지를 확장하고, IDCT를 통해 공간 도메인으로 복원한 뒤 손실된 고주파를 복원하는 하이브리드 도메인 고주파 집중 네트



Fig. 1 Example of 2D DCT.

워크 기반 임의배율 초해상도(Hybrid-domain High-frequency Attention network for Arbitrary magnification Super-Resolution, H^2A^2-SR)를 제안한다.

3.2. 하이브리드 도메인 고주파 집중 네트워크

이번 섹션에서는 하이브리드 도메인 고주파 집중 네트워크(H^2A^2-SR)에 대해 설명한다. 제안하는 H^2A^2-SR 모델의 구조는 그림 2와 같다. 입력으로 받은 저해상도 이미지는 기존 초해상도 네트워크를 통해 목표배율과 근접한 정수배율로 확대된다. 예컨대, 목표 배율이 2.5일 경우 정수배율 네트워크는 2배율을 진행하고, 목표 배율이 3.5일 경우 3배율을 진행한다. 정수배율로 확대된 이미지는 DCT를 통하여 공간 도메인에서 스펙트럼 도메인으로 변환된다. DCT의 특성인 좌상단의 저주파, 우하단 방향 고주파가 밀집되어 있는 특성을 이용하고, 우하단 방향으로 목표 배율에서 정수배율을 나눈 잔여 배율만큼 공간영역을 확장한다. DCT 스펙트럼에서 확장된 공간영역은 공간 도메인에서도 같은 공간 크기를 가지는 원리에 따라 IDCT를 통해 공간영역으로 재변환했을 때, 목표 배율로 조정된 이미지 결과를 얻을 수 있다. 이러한 과정으로 획득된 이미지는 고주파 영역 부분을 임의로 확장했기 때문에, 고주파가 부족한 이미지가 되고, 이를 극복하기 위해서 고주파를 정확히 복원하는데 초점을 맞추어 설계된 고주파 집중 네트워크를 통과시킨다. 고주파 집중 네트워크는 고주파 정보를 잘 만들

기 위해 RGB 채널간의 상관관계를 잘 학습할 수 있는 channel attention 층을 사용하며, 잔차 학습 구조를 통해 모델을 깊게 만들어 그림 3에 제시된 것과 같이 Residual Channel Attention Block(RCAB) 이라는 블록 단위 channel attention을 구성한다. 제안된 모델은 3개의 RCAB을 층으로 쌓아 구성하고, 각 블록 간의 상관관계를 위해 블록별로 잔차 학습을 적용한다. 기존 임의배율 확대 방법에서는 bicubic 보간법을 통해 목표 배율로 확대한 후 초해상도 신경망을 통과함으로써 컴퓨팅 비용과 모델의 용량이 늘어나는 문제를 가진다. 또한 실 사용시 임의배율마다 모든 신경망을 학습해야 하기 때문에 큰 용량을 가지고 있어야 한다. 반면에 제안하는 임의배율 확대 방법에서는 정수배율 가중치를 그대로 보존함에 따라 정수배율의 성능을 보존하고, 상대적으로 적은 용량인 H^2A^2-SR 을 추가로 부차하여 고성능의 임의배율 확대를 달성할 수 있다. 또한 이미지의 전 주파수 대역을 한꺼번에 복원하는 기존 방법들과 달리 타겟으로 하는 고주파 성분을 집중적으로 복원함으로써, 더 좋은 성능을 이끌어낼 수 있다.

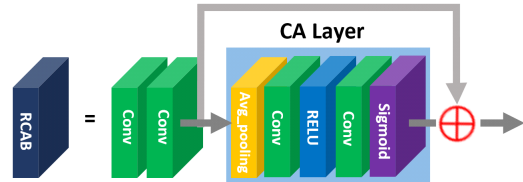


Fig. 3 Configuration of RCAB structure.

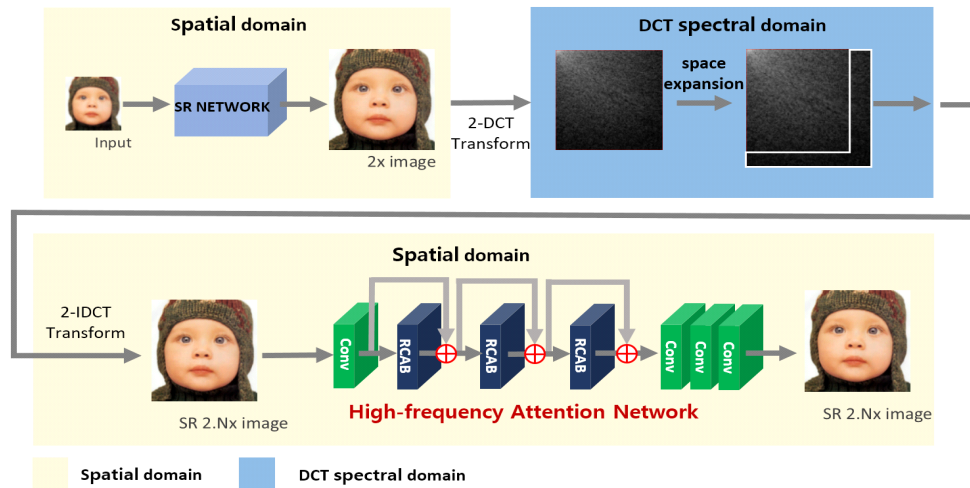


Fig. 2 Overall organization of the proposed H2A2-SR model.

3.3. 고주파 집중 네트워크 손실함수

제안된 모델에서는 DCT에 의해 확장된 영역에 있는 성분인 고주파를 잘 복원하기 위해 식 (5)와 같이 손실함수(L)를 정의한다.

$$L = \frac{1}{N} \| F_{H^2A^2-SR}(F_{sr}(x^{hr})) - x^{hr} \|^2. \quad (5)$$

여기서 N은 이미지 배치사이즈를 의미하며, $F_{sr}(x^{hr})$ 은 저해상도 이미지 x^{hr} 을 초해상도 네트워크에 통해 정수배율로 확대하는 모델을 의미하고, $F_{H^2A^2-SR}$ 은 잔여 소수 배율 확대 모델을 의미한다. 임의배율 초해상도 모델 결과와 대응되는 고해상도 이미지 x^{hr} 과의 평균제곱 오차를 이용하여 네트워크 학습을 위한 손실함수가 계산된다.

IV. 실험 결과

4.1. 네트워크 학습

전통적인 초해상도는 학습 시 입력 이미지인 저해상도 이미지와 목표 이미지인 고해상도 이미지에서 각각 패치 단위로 불러와 비교를 통하여 학습하게 된다. 예컨대, 60x60 크기의 고해상도 패치에 대하여 2배 확대 모델에서는 30x30 크기의 저해상도 입력 패치를 가져와 네트워크 학습을 진행한다. 이러한 패치 기반 학습법은 임의배율을 학습하는데 픽셀 이동 현상의 원인이 되기 때문에 부적절할 수 있다. 따라서 제안하는 학습 방법에서는 패치 기반 대신 입력으로 저해상도 이미지 전체를 받고, 목표 이미지 또한 고해상도 이미지 전체를 받아 임의배율별로 학습을 수행한다.

4.2. Meta-SR과 제안 방법의 성능 비교

이 섹션에서는 임의배율이 가능한 모델인 Meta-SR과 제안하는 H^2A^2-SR 의 성능을 비교하고자 한다. 기존 Meta-SR은 임의배율시 배율에 따른 특화된 가중치 모델을 활용하지 않기 때문에 성능에 한계가 있으므로, 배율 별 각각 학습을 진행하는 H^2A^2-SR 과의 동등 조건 비교를 위해 Meta-SR 네트워크에 대해 배율 별 학습을 진행하여 실험하였다. 임의배율 모델에 있어 정수배율을 진행하는 정수배율 모델이 필요하므로, 본 논문에서는

DRN을 직접 x2, x3, x4 배율에 대하여 학습시켜 정수배율 모델로 활용하였다. 데이터 셋으로는 Celebrate[21]를 사용하였고, 학습용 40,920장, 검증용 5,060장으로 각각 구성하여 학습을 진행했다. Meta-SR은 정수배율 늘린 이미지에서 임의배율에 맞는 사이즈에 맞게 픽셀 값을 선택하는 반면에 H^2A^2-SR 은 신경망에 고주파 복원이라는 목적성을 집중적으로 부여함으로써 고주파 성분과 연관되어 있는 엣지와 텍스처 등을 더욱 잘 복원함을 표 1의 이미지에서 볼 수 있다. H^2A^2-SR 이 다른 임의배율 모델보다 이미지의 노이즈가 적고, 얼굴 인식률과 큰 연관이 있는 눈과 코 부분을 더욱 선명하게 표현함을 볼 수 있다. 또한, 표 1의 정량적인 PSNR 평가지표에서 H^2A^2-SR 이 기존 방법 대비 0.5~0.8 dB 정도의 높은 수치를 보인다.

4.3. 기존 임의배율 방법과 제안방법의 성능 비교

기존 임의배율 방법들과의 추가적인 성능 비교를 위해 기존 임의배율 방법에서 사용한 학습용 데이터 셋 [15]과 테스트 데이터 셋을 제안 방법에도 그대로 사용하여 실험을 진행하였다. DIV2K[22] 데이터 셋을 이용하여 제안 네트워크를 훈련시켰고, 학습된 모델을 테스트하기 위한 데이터 셋으로 B100[23]을 사용하였다. 임의배율 입력 이미지를 만들기 위해 각 임의배율별로 bicubic으로 줄여서 사용하였다. 기존 정수배율이 가능한 State-Of-The-Art(SOTA) 신경망과 비교하기 위해 입력 이미지를 소수 배율을 위해 bicubic으로 확장하여 넣어주었고, 우리 모델은 임의배율별 이미지를 그대로 넣어주었다. 정수배율을 위해서 RDN 모델을 사용하였고 임의배율을 위해 H^2A^2-SR 을 추가하여 실험하였다. 참고로 SOTA 중 하나인 SRWarp은 현재 임의배율 테스트를 위한 소스 코드가 이용 가능하지 않기 때문에 실험이 불가하였다. 표 2에서 볼 수 있듯이 2.2배율, 3.2배율 같이 이미지에서 적은 범위만을 확장하여도 PSNR 수치에 큰 손실을 준다. 하지만 제안하는 우리 모델은 소수 배율을 위한 확장에 대해 강인하므로 평균 PSNR 측면에서 약 3dB 정도 우위를 보여준다.

Table. 1 Quantitative comparison between our H^2A^2 -SR and Meta-SR.

PSNR(dB) on Celebrate[21] with arbitrary scale factors						
Method	x2.2	x2.5	x2.8	x3.2	x3.5	x3.8
DRN + Meta-SR	30.32	31.45	30.64	29.45	29.35	28.89
DRN + H^2A^2 -SR(ours)	30.32	32.13	31.46	29.95	30.13	29.58









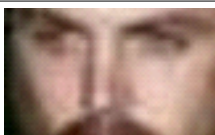
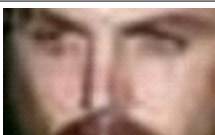
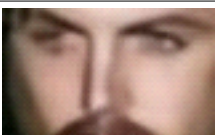
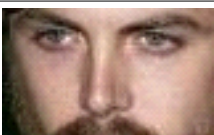

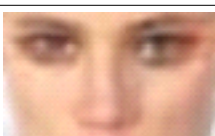
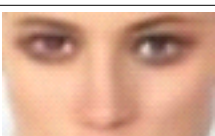
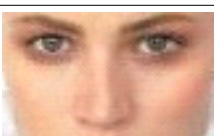
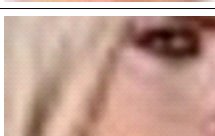
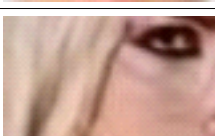
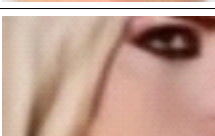
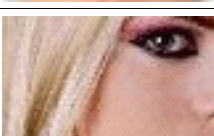
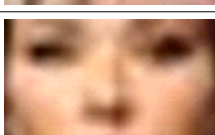
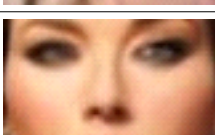
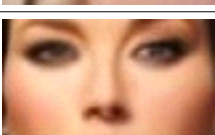
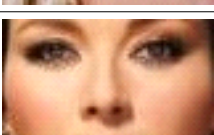
Scale	(a) Bicubic	(b) DRN + Meta-SR	(c) DRN + H^2A^2 -SR	(d) Ground truth
x2.2				
x2.5				
x2.8				
x3.2				
x3.5				
x3.8				

Table. 2 Quantitative comparison of the state-of-the-art SR methods.

PSNR(dB) on B100 [23] with arbitrary scale factors						
Method	x2.2	x2.5	x2.8	x3.2	x3.5	x3.8
RCAN [6]	24.60	28.20	25.31	25.66	25.16	25.40
DRN [7]	24.65	28.31	25.31	25.69	25.28	25.51
RDN [8]	24.61	28.24	25.31	25.67	25.18	25.44
SAN [9]	24.61	28.21	25.32	25.67	25.17	25.41
RDN + H^2A^2 -SR	30.07	30.26	28.27	28.85	27.36	27.38

V. 결론 및 향후연구

본 논문에서는 공간 및 스펙트럼 하이브리드 도메인을 이용하여 고주파 성분을 복원하는 임의배율 초해상도 방법을 제안하였다. DCT 스펙트럼 도메인에서 공간 확장을 통하여 목표 해상도의 이미지로 유연하게 확대할 수 있고, 확장된 이미지의 부족한 고주파 성분을 제안하는 고주파 집중 네트워크를 통해 복원한다. 이를 통해 기존 정수배율 초해상도 모델의 정확도를 임의의 소수 배율 확대 시에도 보존하며, 기존 모델의 수정 및 재학습 없이 제안하는 임의배율 모델을 추가함으로써, 고성능 소수 배율 결과를 얻을 수 있다. 실험 결과는 기존 임의배율 초해상도 방법 대비 제안하는 방법이 정량 및 정성적으로 모두 우수

한 복원 성능을 가지는 것을 보여준다. 향후 연구로는 정수배율 간 가중치 공유 방법과 knowledge distillation 기법을 적절히 결합하여 네트워크를 경량화하는 연구가 가능하다. 또한 임의배율 초해상도 네트워크와 객체 인식 네트워크를 통합하여 저해상도 이미지에 대한 객체 인식률을 향상시키는 연구가 수행 될 수 있다.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(NRF-2020R1A4A1019191) and Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No.2020-0-00004, Development of Previsional Intelligence based on Long-term Visual Memory Network).

References

- [1] S. J. Lee, T. J. Kim, C. H. Lee, and S. B. Yoo, "Image super-resolution for improving object recognition accuracy," in *Journal of the Korea Institute of Information and Communication Engineering*, vol. 25, no. 6, pp. 774-784, Jun. 2021.
- [2] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295-307, Feb. 2016.
- [3] J. W. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1646-1654, 2016.
- [4] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proceeding of the European Conference on Computer Vision*, pp. 1874-1883, Sep. 2016.
- [5] M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," in *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1664-1673, Mar. 2018.
- [6] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proceeding of the European Conference on Computer Vision*, pp. 286-301, Mar. 2018.
- [7] Y. Guo, J. Chen, J. Wang, Q. Chen, J. Cao, Z. Deng, Y. Xu, and M. Tan, "Closed-loop matters: Dual regression networks for single image super-resolution," in *Proceeding of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5407-5416, May. 2020.
- [8] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2472-2481, Mar. 2018.
- [9] T. Dai, J. Cai, Y. Zhang, S. T. Xia, and L. Zhang, "Second-order attention network for single image super-resolution," in *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 11065-11074, 2019.
- [10] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, and Z. Wang, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceeding of the European Conference on Computer Vision*, pp. 4681-4690, 2017.
- [11] A. Lugmayr, M. Danelljan, L. V. Gool, and R. Timofte, "SrfLOW: Learning the super-resolution space with normalizing flow," in *Proceeding of the European Conference on Computer Vision*, pp. 715-732, 2020.
- [12] Y. H. Jo, S. J. Yang, and S. J. Kim, "SrfLOW-da:

Super-resolution using normalizing flow with deep convolutional block,” in *Proceeding of the European Conference on Computer Vision Workshops*, pp. 364-372, June. 2021.

[13] Y. G. Kim and D. H. Son, “Noise conditional flow model for learning the super-resolution,” in *Proceeding of the European Conference on Computer Vision Workshops*, pp. 424-432, June. 2021.

[14] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, “Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network,” in *Proceeding of the European Conference on Computer Vision*, pp. 1874-1883, 2016.

[15] X. Hu, H. Mu, X. Zhang, Z. Wang, T. Tan, and J. Sun, “Meta-SR: A magnification-arbitrary network for super-resolution,” in *Proceeding of the European Conference on Computer Vision*, pp. 1575-1584, 2019.

[16] S. H. Son and K. M. Lee, “SRWarp: Generalized image super-resolution under arbitrary transformation,” in *Proceeding of the European Conference on Computer Vision*, pp. 7782-7791, June. 2021.

[17] N. Kumar, R. Verma, and A. Sethi, “Convolutional neural networks for wavelet domain super resolution,” in *Pattern Recognition Letters*, vol. 90, pp. 65-71, 2017.

[18] J. Li, S. You, and A. Robles-Kelly, “A frequency domain neural network for fast image super-resolution,” in *International Joint Conference on Neural Networks*, pp. 1-8, 2018.

[19] S. Xue, W. Qiu, F. Liu, and X. Jin, “Faster image super-resolution by improved frequency-domain neural networks,” *Signal, Image and Video Processing*, vol. 14, no. 2, pp. 257-265, 2020.

[20] O. Aydin and R. G. Cinbis, “Single-image super-resolution analysis in DCT spectral domain,” in *Balkan Journal of Electrical & Computer Engineering*, vol. 8, no. 3, pp. 209-217, 2020.

[21] Z. Liu, P. Luo, X. Wang, and X. Tang, “Large-scale CelebFaces Attributes (CelebA) Dataset,” in *Retrieved August*, vol. 2018.15, no. 11, 2018.

[22] R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, L. Zhang, B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, “NTIRE 2017 challenge on single image super-resolution: Methods and results,” in *Proceeding of the European Conference on Computer Vision Workshops*, July. 2017.

[23] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to

evaluating segmentation algorithms and measuring ecological statistics,” in *International Conference on Computer Vision*, pp. 416-423, 2001.



윤준석(Jun-Seok Yun)

조선대학교 전자공학과 공학사
전남대학교 인공지능융합학과 석사과정
※관심분야: 딥러닝, 컴퓨터비전



이성진(Sung-Jin Lee)

전남대학교 IoT인공지능융합전공 공학사
전남대학교 인공지능융합학과 석사과정
※관심분야: 딥러닝, 컴퓨터비전



유석봉(Seok Bong Yoo)

연세대학교 전기전자공학 공학사
한국과학기술원 전기및전자공학과 공학석사
한국과학기술원 전기및전자공학과 공학박사
전남대학교 인공지능융합학과 조교수
※관심분야: 시각지능, 컴퓨터비전, 실감미디어



한승회(Seunghwoi Han)

한국과학기술원 기계공학과 공학사
한국과학기술원 기계공학과 공학석사
한국과학기술원 기계공학과 공학박사
전남대학교 기계공학부 조교수
※관심분야: 인공지능, 자율주행