

일반논문 (Regular Paper)

방송공학회논문지 제25권 제2호, 2020년 3월 (JBE Vol. 25, No. 2, March 2020)

<https://doi.org/10.5909/JBE.2020.25.2.242>

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

CARDB를 이용한 반복적인 업-다운 샘플링 네트워크 기반의 단일 영상 초해상도 복원

김 인 구^{a)}, 유 송 현^{a)}, 정 제 창^{a)†}

Single Image Super-Resolution Using CARDB Based on Iterative Up-Down Sampling Architecture

Ingu Kim^{a)}, Songhyun Yu^{a)}, and Jechang Jeong^{a)†}

요 약

최근 단일 영상 초해상도에 깊은 합성곱 신경망을 적용한 알고리듬이 많이 연구되었다. 현존하는 딥러닝 기반 초해상도 기법들은 네트워크의 후반부에 해상도를 업샘플링 하는 구조를 가진다. 이러한 구조는 저해상도에서 고해상도로 한 번에 매핑을 하기에 많은 정보를 예측하는 높은 확대율에서 비효율적인 구조를 가진다. 본 논문에서는 반복적인 업-다운 샘플링 구조를 기반으로 하여 채널 집중하여 밀집 블록을 이용한 단일 영상 초해상도 기법을 제안한다. 제안한 알고리듬은 저해상도와 고해상도의 매핑 관계를 효율적으로 예측하여 높은 확대율에서 기존의 알고리듬에 비해 최대 0.14dB 성능 향상과 개선된 주관적 화질을 보여준다.

Abstract

Recently, many deep convolutional neural networks for image super-resolution have been studied. Existing deep learning-based super-resolution algorithms are architecture that up-samples the resolution at the end of the network. The post-upsampling architecture has an inefficient structure at large scaling factor result of predicting a lot of information for mapping from low-resolution to high-resolution at once. In this paper, we propose a single image super-resolution using Channel Attention Residual Dense Block based on an iterative up-down sampling architecture. The proposed algorithm efficiently predicts the mapping relationship between low-resolution and high-resolution, and shows up to 0.14dB performance improvement and enhanced subjective image quality compared to the existing algorithm at large scaling factor result.

Keyword : Super-Resolution, Deep Learning, Iterative Up-Down Sampling, Channel Attention Residual Dense Block

a) 한양대학교 전자컴퓨터통신공학과(Department of Electronics and Computer Engineering, Hanyang University)

† Corresponding Author : 정제창(Jechang Jeong)

E-mail: jjeong@hanyang.ac.kr

Tel: +82-2-2220-4372

ORCID: <http://orcid.org/0000-0002-3759-3116>

· Manuscript received December 20, 2019; Revised February 24, 2020; Accepted February 24, 2020.

I. 서 론

최근 디스플레이 기술이 나날이 발전하면서 디스플레이 장치의 해상도가 높아지고 있다. 이로 인해 고해상도 디스플레이 장치에서 해상도에 맞는 콘텐츠를 제공해야 고품질의 영상과 여백이 없는 스크린을 체감할 수 있지만 기존에 출시된 콘텐츠 및 영상들은 대부분 저해상도를 가져 콘텐츠가 매우 부족한 상황이다. 수학적 모델링을 통한 영상의 해상도를 높이는 방법인 보간법은 예측할 화소의 주변 화소 값들을 활용하여 해상도를 높인다. 보간법은 계산량이 적다는 장점이 있지만 저주파수 필터 특성 때문에 영상을 전반적으로 블러시켜 디테일을 보존하는 데에 한계가 있다. 이를 해결하는 방법으로 저해상도의 영상을 고해상도로 높이며 영상의 세세한 부분을 개선해주는 초해상도 (SR: super-resolution) 기법이 연구되었다^[1,2].

현재 SR 기법은 딥러닝을 적용하지 않은 기법과 적용한 기법으로 크게 두 가지로 나뉜다. 먼저, 딥러닝을 적용하지 않은 SR은 영상의 내부와 외부 정보의 사용 여부에 따라 두 가지로 나눌 수 있다. Glasner 등은 복수 영상의 초해상도 기법과 단일 영상의 복수 패치 방법을 통합한 coarse-to-fine 구조를 기반으로 영상의 내부 유사성을 사용하여 재구성한다^[2]. Freeman 등은 외부의 데이터베이스에 저해상도와 고해상도 패치 쌍을 저장하고 저장된 정보를 참조하여 입력 데이터와 가장 유사한 저해상도의 패치를 찾아 대응되는 고해상도 패치를 활용하여 영상을 재구성한다^[1]. 외부의 정보를 활용한 방식은 내부 정보를 활용한 방법에 비해 참조할 수 있는 유의미한 고주파수 정보들이 많이 있어 보다 나은 성능 향상을 보여주었다.

최근에는 합성곱 신경망을 초해상도 기법에 적용하여 뛰어난 성능을 보여주고 있다^[3,4,5,6,7,10,11,12,13,16]. 최초로 SR에 딥러닝을 적용한 SRCNN은 사전 업샘플링 구조로 3개의 합성곱 계층을 통해 영상을 재구성하였다^[3]. 이러한 사전 업샘플링 구조는 학습 시 한 번에 다양한 배율을 동시에 학습할 수 있지만 네트워크의 입력으로 보간법을 통해 확대된 저화질의 고해상도를 사용하기에 연산량이 많아지는 문제가 발생하여 네트워크를 깊게 구성하지 못하는 한계가 있다. 이후 연산량 문제를 고려하여 대부분의 SR 네트워크의 구조는 사후 업샘플링 구조를 기반으로 연구가 진행되

고 있다. 대표적인 사후 업샘플링 구조인 EDSR은 ResNet에서 제안한 잔차 블록을 기반으로 한 end-to-end 구조이다^[4]. 이러한 사후 업샘플링 구조는 연산량을 줄여 네트워크를 깊게 구성할 수 있는 장점이 있지만 단 한 번에 저해상도에서 고해상도로 매핑을 하기에 적은 정보를 예측하는 낮은 배율 ($\times 2$)에서만 효율을 가지는 구조적 한계가 있다. 최근에는 학습 시 높은 배율에 대한 효율성을 가지기 위해 반복적인 업-다운 샘플링 네트워크 구조가 연구되었다. DBPN은 반복적인 업-다운 샘플링 네트워크 구조를 가지는 대표적인 알고리듬으로 서로 다른 고해상도 특정 정보를 활용하여 높은 배율에서 성능 향상을 보여주었다^[5]. 그러나 이러한 방법은 각각의 샘플링 모듈에서 단순 반복된 구조만을 통해 저해상도와 고해상도 간의 매핑 연관성을 예측하기에 이전에 학습된 유의미한 특징 정보들을 재활용하지 못하는 단점이 있다.

본 논문에서는 높은 확대율에서 비효율적인 구조를 가지는 문제를 해결하기 위해 채널 집중 잔여 밀집 블록 (CARDB: channel attention residual dense block)을 제안하고 이를 이용한 반복적인 업-다운 샘플링 네트워크 구조를 구성한다. 제안된 방식은 이전의 학습된 유의미한 특징 정보를 재활용하기 위한 잔여 밀집 블록 (RDB: residual dense block)과 특징별로 중요도에 따른 가중치를 주기 위해 채널 집중 (channel attention) 기법을 적용하여 고해상도와 저해상도의 매핑 연관성을 효율적으로 추론하도록 하였다.

본 논문은 다음과 같이 구성된다. 본 논문의 2장에는 관련 이론을 소개하고 3장에서 제안한 알고리듬에 대해 상세히 설명한다. 4장에서는 기존의 초해상도 알고리듬과 제안한 알고리듬을 비교하고 성능을 평가한다. 마지막 5장에서는 결론을 맺는다.

II. 관련 이론

기존의 초해상도 기법들은 합성곱 신경망을 기반으로 한 많은 알고리듬이 제안되었다^[3,4,5,6,7,10,11,12,13,16]. 기존의 신경망 네트워크 구조는 한 번에 저해상도 영상에서 고해상도로 높이는 구조를 대부분 가진다.

Kim 등은 사전 업샘플링 구조를 가진 매우 깊은 초해상도 기법 (VDSR)을 제안했다^[6]. 이 기법은 VGGNet^[15]에서 영감을 받아 20개의 깊은 합성곱 계층을 통해 개선된 영상을 재구성하였으며 잔차 학습 (residual learning)을 통해 고해상도로 확대된 입력 정보를 출력 정보에 전달하여 네트워크가 고주파수 정보만을 학습하도록 구성한 기법이다. Lim 등은 사후 업샘플링 구조를 가진 개선된 깊은 잔차 초해상도 기법 (EDSR)을 제안했다^[4]. 이 기법은 low-level 비전에 맞게 개선된 잔차 블록 구조를 활용하여 매우 깊은 합성곱 신경망을 통해 디테일이 개선된 고해상도 영상을 복원하는 기법이다. Zhang 등은 ResNet^[8]의 초기 특징 정보의 활용성과 DenseNet^[9]의 이전 특징 정보 재활용성을 융합한 잔여 밀집 블록 (RDB)을 기반으로 초해상도 기법을 제안했다^[7]. RDB는 초기의 특징 정보를 마지막 계층에 전달하며 이전의 계층들이 이후의 계층들의 입력정보로 사용하는 방법이다. 이러한 방법은 학습된 다양한 특징 정보를 재활용을 통해 정보의 흐름을 유연하게 함으로써 성능 향상을 보여주었다. Zhang 등은 채널 집중 (channel attention) 네트워크를 제안했다^[10]. 이 기법은 합성곱 계층이 영상의 지역적 영역 (local area)을 활용한 방법을 벗어나 영상의 전체 영역 (global area)을 활용하여 영상의 특징맵에서 중요도에 따라 0과 1 사이의 웨이트를 할당하는 방법을 사용했다.

Haris 등은 그림 1과 같이 반복적으로 저해상도 영상과 고해상도 영상의 매팽 연관성을 학습을 통해 서로 다른 고해상도 특징 정보들을 활용하여 개선된 고해상도 영상을 복원하는 깊은 역투영 초해상도 기법 (DBPN)을 제안했다^[5]. DBPN은 반복적으로 저해상도와 고해상도 간의 투영 및 역투영을 학습할 때 업샘플링 모듈과 다운샘플링 모듈을 사용한다. 각 모듈의 합성곱 계층은 해상도 축소를 목적으로 사용이 되며 해상도를 높이기 위해 디컨볼루션 (deconvolution)을 사용하였다. 제안된 각 모듈은 내부적으로 고해상도로 높이고 저해상도로 낮추는 연관성에 대해 학습하며 학습된 정보와 초기 정보 간의 차이 성분인 에러를 피드백하는 메커니즘을 가진다. 이러한 방법은 에러 정보를 활용하여 외부적으로 밀집 연결 (dense connection) 된 도메인 정보를 사용함으로써 높은 배율에서 향상된 성능을 보여주며 기존의 초해상도 기법들과 비교할 때 높은 배율의 결과 ($\times 4$)에서 적은 파라미터를 가지면서 뛰어난 성능을 보여 주었다.

III. 제안하는 알고리듬

제안하는 알고리듬은 기본적으로 DBPN과 같이 반복적

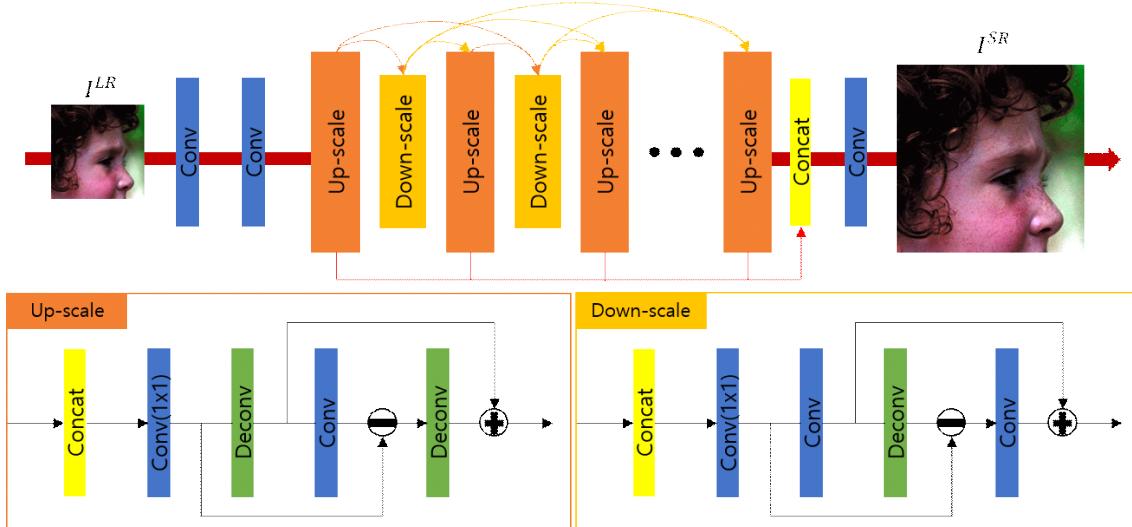


그림 1. DBPN 네트워크 구조
Fig. 1. Architecture of DBPN

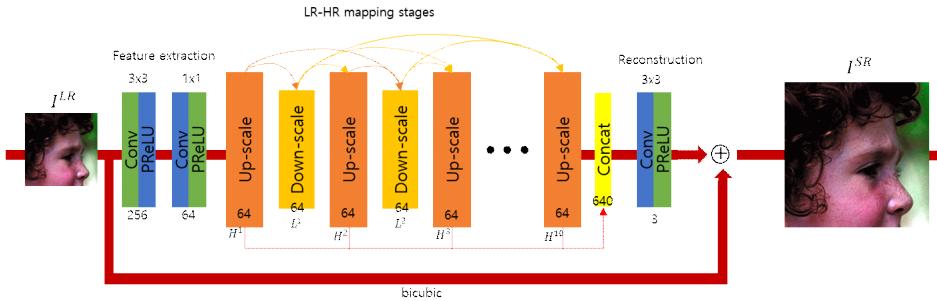


그림 2. 제안하는 네트워크의 구조
 Fig. 2. The architecture of proposed network

업-다운 샘플링 네트워크 구조를 가진다. 성능을 개선하기 위해 잔여 밀집 블록과 채널 집중을 결합한 채널 집중 잔여 밀집 블록 (CARDB)을 활용한 반복적인 업-다운 샘플링 네트워크를 제안한다. 그림 2는 제안하는 네트워크 구조를 도식화하였다.

제안하는 알고리듬은 3단계로 특징 추출, 반복적인 저해상도-고해상도 매핑, 재구성 단계로 구성된다. 첫째, 저해상도의 영상이 특징 추출 (feature extraction) 단계에 입력되어 저해상도 영상의 특징을 추출한다. 둘째, 저해상도-고해상도 매핑 단계에서 서로 다른 고해상도 특징 맵들을 추출한다. 마지막으로 추출된 고해상도 특징 맵들을 종합하여 개선된 고해상도 영상으로 복원한다. 네트워크의 효율적인 학습을 위해 입력 영상을 썽 입방 보간법 (bicubic interpolation)을 통해 확대 후 마지막 단계에 전달함으로써 영상의 미세한 부분을 학습하도록 초점을 맞추었다.

1. 특징 추출

특징 추출 단계에서 신경망 네트워크에 입력된 저해상도 영상의 특징 맵을 추출한다. 첫 번째 합성 곱에서 256개의 특징 맵을 추출하고 반복적인 업-다운 샘플링 단계의 고정

된 채널 수를 위해 Szegedy 등^[14]이 제안한 1×1 병목 계층 (bottleneck layer)을 통해 64개의 특징 맵을 추출한다.

2. 저해상도-고해상도 매핑 단계

높은 확대율에서 더욱 섬세한 영역을 학습하기 위해 제안하는 반복적 업-다운 샘플링 모듈을 사용하였다 (그림 3). 기존의 업-다운 샘플링 모듈은 이전에 학습된 특징 정보를 제대로 활용하지 않는 비효율적인 구조를 가진다. 제안한 업-다운 샘플링 모듈은 이전 특징 맵 정보들을 효율적으로 사용하기 위해 채널 집중과 잔여 밀집 블록을 결합한 구조 채널 집중 잔여 밀집 블록 (CARDB)을 사용하였다.

그림 4는 제안하는 채널 집중 잔여 밀집 블록을 나타낸다. 하나의 블록은 7개의 3×3 합성 곱과 1개의 1×1 병목 계층으로 구성된다. 각 합성 곱 계층은 서로 크기가 다른 특징 맵들을 추출하며 각 계층에서 이전의 학습된 모든 특징 정보들을 재활용한다. 이러한 모듈 구조는 연속 메모리 구조 및 잔여 학습 그리고 특징 결합을 통해 효율적으로 신경망 학습을 할 수 있다. 성장률 (growth rate)은 32로 설정을 하였고 블록의 입력과 출력 채널을 64로 설정하였다. 네트워크가 깊어지면 수용 영역이 커져 학습 시 합성 곱

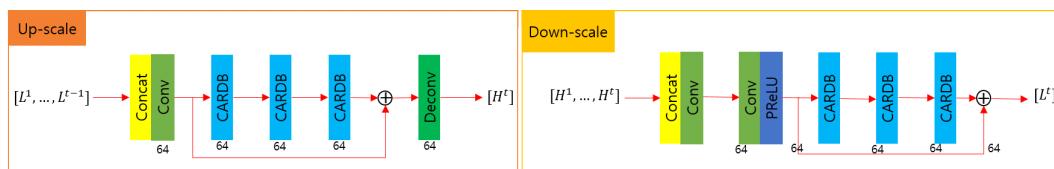


그림 3. 제안하는 알고리듬에서 사용한 업-다운 샘플링 모듈 구조
 Fig. 3. The architecture of proposed up-down sampling module

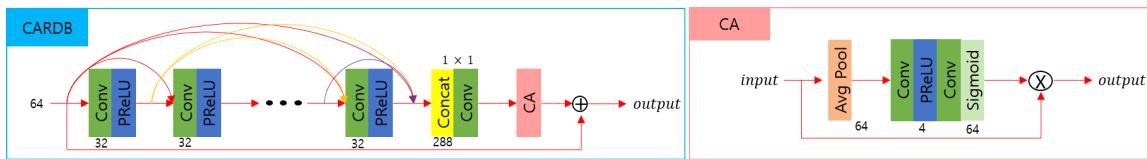


그림 4. CARDB 구조

Fig. 4. The Architecture of CARDB

계층을 통한 결과는 국소적 영역 (local area)에 대한 정보만을 추출하여 한계가 있다. 전체 영역 (global area)의 정보를 활용하기 위해 채널 집중 기법을 사용했다. 채널 집중은 각 특징 맵마다 중요도에 따른 웨이트 값을 할당하기 위해 전역 평균 풀링 (global average pooling), 시그모이드 (sigmoid)를 사용하였다.

3. 복원

최종 복원단계에서 서로 다른 고해상도 특징들을 입력 채널로 하여 3×3 합성곱을 통해 RGB 영상으로 복원한다.

IV. 실험 결과

제안한 알고리듬을 실험하기 위해 고해상도 영상과 이에 상응하는 저해상도 영상이 존재해야 한다. 학습을 위한 데이터셋으로 DIV2K 800장과 Flickr2K 2650장을 사용하였

고 검증셋 (validation set)으로 5장의 영상으로 이루어진 Set5^[18]을 이용하였다. 영상을 임의의 위치에서 48×48 크기의 패치 (patch) 단위로 나누어 학습에 사용했다. 또한 무작위 회전과 반전을 포함한 데이터 확장 방법 (data augmentation)을 사용하였다. 결과를 실험하기 위해 테스트셋으로 Set5^[18], Set14^[19], B100^[20], Urban100^[21]을 사용하여 비교 및 분석을 하였다. 확대 배율은 2배, 3배, 4배로 설정하였다. 영상을 시험하는 과정에서 좌우 반전 및 0도, 90도, 180도, 270도 회전하여 얻은 8개의 결과 영상을 평균화하는 자기 앙상블 (self ensemble) 기법을 적용하였다. PSNR 측정시 L1 loss가 L2 loss보다 영상 복원에서 더 좋은 학습 성능을 보였다^[17]. 손실 함수는 L1 loss를 사용했다.

비교 알고리듬은 쌍 입방 보간법, Dong 등의 SRCNN^[3], Kim 등의 VDSR^[6], Haris 등의 DBPN^[5], Lim 등의 EDSR^[4]으로, 총 5개의 기존 알고리듬과 비교하였다. 결과 영상의 성능을 비교하기 위해 최대 신호 대 잡음 비 (PSNR: peak signal-to-noise ratio)와 구조적 유사성 (SSIM: structural similarity)을 사용하였다. 표 1과 표 2의 “+”는 자기 앙상블

표 1. 시험 데이터셋에 대한 최대 신호 대 잡음비 (dB)

Table 1. The PSNR of test dataset

Dataset	Scale	Bicubic	SRCCN ^[3]	VDSR ^[6]	DBPN ^[5]	DBPN+ ^[5]	EDSR ^[4]	EDSR+ ^[4]	Proposed	Proposed+
Set5 ^[18]	$\times 2$	33.69	36.72	37.53	38.09	38.15	38.11	38.19	38.12	38.18
	$\times 3$	30.41	32.78	33.67	-	-	34.65	34.71	34.68	34.76
	$\times 4$	28.43	30.50	31.35	32.47	32.52	32.46	32.52	32.61	32.67
Set14 ^[19]	$\times 2$	30.25	32.51	33.05	33.85	33.95	33.92	33.97	33.86	33.92
	$\times 3$	27.55	29.32	29.78	-	-	30.52	30.57	30.56	30.63
	$\times 4$	26.01	27.52	28.02	28.82	28.87	28.80	28.87	28.89	28.97
B100 ^[20]	$\times 2$	29.57	29.57	31.90	32.27	32.33	32.32	32.35	32.29	32.34
	$\times 3$	27.22	27.22	28.83	-	-	29.25	29.29	29.28	29.32
	$\times 4$	25.97	25.97	27.29	27.72	27.76	27.71	27.76	27.77	27.82
Urban100 ^[21]	$\times 2$	26.89	26.89	30.77	32.55	32.69	32.93	33.06	32.60	32.80
	$\times 3$	24.47	24.47	27.14	-	-	28.80	28.94	28.84	29.00
	$\times 4$	23.15	23.15	25.18	26.38	26.55	26.64	26.77	26.77	26.94

표 2. 시험 데이터셋에 대한 구조적 유사성

Table 2. The SSIM of test dataset

Dataset	Scale	Bicubic	SRCNN ^[3]	VDSR ^[6]	DBPN ^[5]	DBPN+ ^[5]	EDSR ^[4]	EDSR+ ^[4]	Proposed	Proposed+
Set5 ^[18]	×2	0.9312	0.9552	0.9590	0.9600	0.9629	0.9602	0.9628	0.9622	0.9625
	×3	0.8687	0.9091	0.9213	-	-	0.9280	0.9315	0.9313	0.9319
	×4	0.8106	0.8627	0.8838	0.8980	0.9012	0.8968	0.9013	0.9014	0.9022
Set14 ^[19]	×2	0.8698	0.9083	0.9124	0.9190	0.9233	0.9195	0.9236	0.9221	0.9226
	×3	0.7747	0.8229	0.8314	-	-	0.8462	0.8501	0.8503	0.8516
	×4	0.7044	0.7534	0.7679	0.7860	0.7936	0.7876	0.7932	0.7940	0.7955
B100 ^[20]	×2	0.8435	0.8889	0.8960	0.9000	0.9047	0.9013	0.9048	0.9040	0.9044
	×3	0.7412	0.7878	0.7990	-	-	0.8093	0.8150	0.8154	0.8161
	×4	0.6701	0.7122	0.7261	0.7400	0.7491	0.7420	0.7501	0.7510	0.7519
Urban100 ^[21]	×2	0.8413	0.8960	0.9140	0.9324	0.9347	0.9351	0.9360	0.9335	0.9347
	×3	0.7369	0.8011	0.8290	-	-	0.8653	0.8685	0.8678	0.8697
	×4	0.6601	0.7254	0.7540	0.7946	0.8007	0.8033	0.8094	0.8098	0.8126

표 3. 파라미터 수, 실행시간에 대한 성능 비교

Table 3. Performance comparison of number of parameters, running time

	DBPN ^[5]	EDSR ^[4]	Proposed (growth rate=16)	Proposed stage4	Proposed stage7	Proposed stage10
Parameter	10M	43M	13M	9M	17M	25M
Running time (sec)	0.194	1.218	0.872	0.101	0.582	0.872
PSNR (dB)	26.38	26.64	26.59	26.52	26.73	26.77

(self ensemble)을 적용한 결과를 나타낸다.

표 1과 표 2는 각 시험 데이터셋에 대한 기존 모델들과 제안하는 모델의 정량적 화질 평가 결과를 비교한다. 기존 기법들과 비교한 결과, ×4의 확대 배율에서 최대 0.14dB 향상된 결과를 보여주며 자기 양상률을 적용한 결과에서 최대 0.17dB만큼 향상된 결과를 보여준다.

표 3은 제안하는 모델의 업-다운 샘플링의 반복 횟수와 성장률에 따른 성능을 비교한다. 각 단계는 성장을 32를 기본 설정으로 가지며 업-다운 샘플링 모듈을 하나의 쌍으로 하여 4번, 7번, 10번 시행하였다.

Urban100 시험 영상 데이터셋의 256×256 해상도의 영상을 1024×1024로 확대할 때, 제안하는 기법의 4번 시행 한 결과는 DBPN과 비교하면 1M 적은 파라미터 수를 가지면서 0.093초 빠른 결과를 보여주면서 0.14dB 향상된 결과를 보여주었다. 7번 시행한 결과와 10번 시행한 결과와 EDSR을 비교하면 2배 이상 적은 파라미터 수를 가지면서 0.636초, 0.346초 빠른 실행 시간과 0.09dB, 0.13dB만큼 높은 PSNR 수치를 확인하였다. 성장률을 16으로 설정한 결

과는 성장을 32인 10번 시행한 결과와 비교하면 약 2배 정도의 적은 파라미터를 가지지만 같은 실행 시간을 가지며 0.18dB만큼 적은 수치를 확인하였다. 실험에 사용된 사용된 하드웨어 환경은 GTX 1080Ti GPU와 Intel i5-6600K CPU를 사용하였다.

표 4는 제안된 기법에 사용된 채널 집중과 전역 잔차 학습의 사용 여부에 따른 성능 비교를 나타낸다. Set5 시험 영상 데이터셋으로 부터 모든 기법을 적용한 경우 각 기법을 사용하지 않았을 때보다 향상된 성능을 보여준다.

표 4. 채널 집중과 전역 잔차 학습 사용에 따른 성능 비교

Table 4. Performance comparison by using channel attention and global residual learning

	Original	Channel Attention (X)	Global Residual Learning (X)
PSNR (dB)	32.61	32.53	32.59

그림 5는 ×2 확대 배율에 대한 기존의 기법들과의 주관적 화질 비교 결과를 나타낸다. Set14의 ‘barbara’, ‘ppt3’을

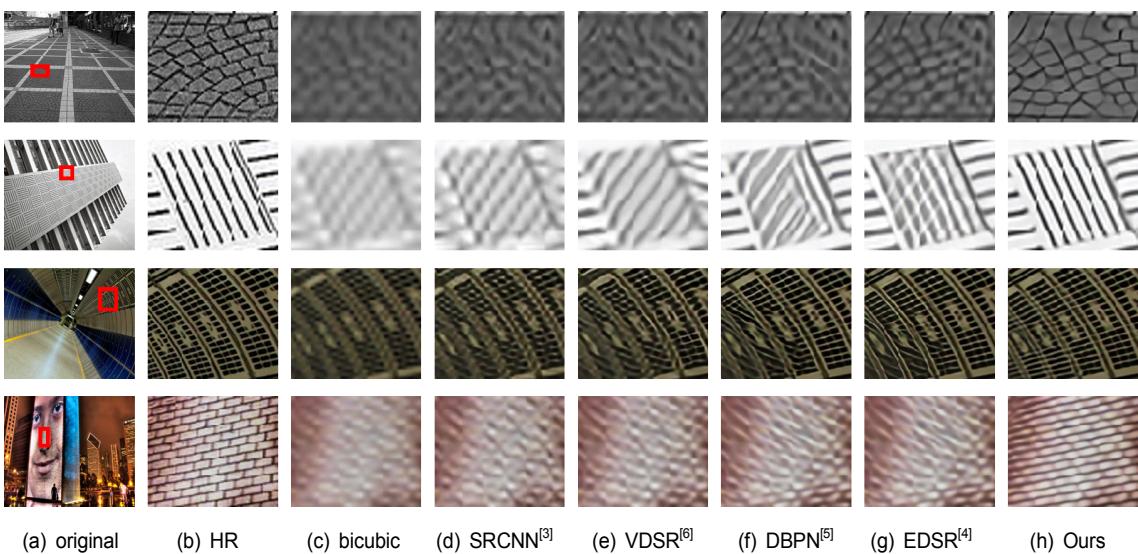
그림 5. 기존 기법들과의 $\times 2$ 주관적 화질 비교Fig. 5. $\times 2$ Subjective image quality comparison with conventional methods

사용한 결과 영상이다. 그림 5의 결과에서 (c) bicubic은 영상을 전체적으로 부드럽게 만들기에 블러링이 심한 화질을 보여준다. (d)와 (e)는 기존 고해상도 이미지가 가지는 옷무늬 사선과 다르게 체커보드 형태의 무늬와 깨진 글자 패턴을 보여준다. (f)와 (g) 또한 기존의 사선 방향과 반대의 사선 무늬와 정확하지 않은 알파벳 ‘e’를 복원하였다. 반면에 제안한 (h)는 비슷한 옷 무늬 패턴과 방향과 식별 가능하게 덜 깨진 알파벳 ‘e’를 복원한 것을 보여 준다.

그림 6는 $\times 4$ 확대 배율에 대한 기존의 기법들과의 주관적 화질 비교 결과를 나타낸다. Urban100의 ‘img076’, ‘img078’, ‘img092’, ‘img095’를 사용한 결과 영상이며 각 해상도는 256×256 에서 1024×1024 로 4배 확대한 결과이

다. 픽셀 단위로 개선 정도를 육안으로 확인하기 위해 관심 영역을 확대하였다. 그림 6의 결과에서 (c) bicubic은 영상을 부드럽게 하여 전체적으로 블러링이 심해 화질이 좋지 못하다. 딥러닝을 활용한 기존의 기법들의 경우 원본 (b) HR과 비교하였을 때 영상의 패턴을 제대로 복원하지 못하였다. 반면에 제안한 방법 (h)는 도로의 블록, 건물의 외벽들을 원복과 비슷하게 복원을 한 것을 보여 준다. 전체적으로 제안한 방식은 높은 배율의 결과에서 뛰어난 성능을 보인다.

그림 7은 표 1, 표 2, 표 4의 결과에서 보여준 자기 앙상블과 채널 집중 및 전역 잔차 학습에 따른 주관적 화질 비교 결과를 나타낸다. Urban100 시험 영상 데이터셋의

그림 6. 기존 기법들과의 $\times 4$ 주관적 화질 비교Fig. 6. $\times 4$ Subjective image quality comparison with conventional methods

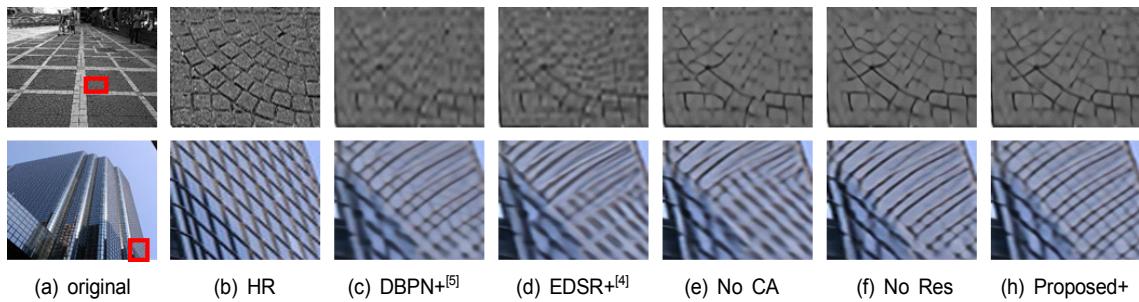


그림 7. 채널 집중 및 전역 잔차 학습과 자기 양상블에 대한 $\times 4$ 주관적 화질 비교

Fig. 7. $\times 4$ Subjective image quality comparison for channel attention, global residual learning and self ensemble

‘img095’, ‘img074’를 사용하여 4배 확대한 결과 영상이다. 그림 7의 결과에서 (c) DBPN+과 (d) EDSR+는 블러링이 심해 왜곡된 도로의 블록과 건물의 창문을 복원하였다. 채널 집중을 사용하지 않은 (e)또한 (c)와 (d)와 같이 패턴을 제대로 복원하지 못하였다. 반면에 전역 잔차 학습을 사용하지 않은 (f)는 덜 왜곡된 영상으로 복원을 하였지만 창문의 가로 성분만을 복원하였다. 반면에 제안한 방법의 자기 양상블 (g)는 건물 창문의 가로 및 세로 성분을 가장 비슷하게 복원하여 기존의 기법에 비해 뛰어는 성능을 보여 준다.

결과적으로, 제안하는 기법의 CARDB는 이전의 의미 있는 특징 정보를 효율적으로 활용하고 특징 맵의 전역적 정보를 사용함으로써 각 실험의 단계별로 객관적, 주관적 화질 비교 결과에서 우수한 성능을 보여준다.

V. 결 론

본 논문에서는 채널 집중 잔여 밀집 블록을 이용한 반복적 업-다운 샘플링 기반의 초해상도 기법을 제안하였다. 기존의 반복적 업-다운 샘플링 네트워크인 DBPN은 이전의 유의미한 특징 정보를 효율적으로 활용하지 못하는 단점과 네트워크가 깊어지면 수용 영역이 커지지만 합성곱 계층에 의해 국소적 영역의 정보만을 활용하는 단점을 가진다. 문제를 개선하고 성능 향상을 위해 이전의 학습된 특징 맵 정보들을 각 계층에서 재활용하여 비효율적인 매핑 구조를 해결하는 RDB와 특징 맵의 전역 정보를 중요도에 따라 서로 다른 웨이트 값을 매핑하여 깊은 신경망 구조에서 효율적으로 학습하는 채널 집중 기법을 결합한 CARDB 기반의 반복적 업-다운 샘플링 초해상도 기법을 제안하였다.

제안한 방법은 채널 집중 기법과 잔차 학습의 순차적 실험을 통해 높은 배율에서 효율성을 갖도록 구성하였다. 파라미터 비교 및 실행 시간 그리고 PSNR 결과를 통해 4번 시행한 결과에서 DBPN보다 1.92배 빠르며 PSNR은 0.14dB만큼 우수하며 7번, 10번 시행한 결과에서 EDSR과 비교하면 2.09배, 1.4배 빠르며 0.09dB, 0.13dB만큼 뛰어난 것을 확인하였다.

본 논문에서 제안하는 CARDB를 이용한 반복적 업-다운 샘플링 네트워크 기반의 초해상도 기법은 높은 배율에서 뛰어난 성능을 보여주기에 영상 인식 분야의 전처리, CCTV, TV, 의료 영상 등에 다양한 산업 분야에서 활용될 수 있다.

참 고 문 헌 (References)

- [1] W. T. Freeman, T. R. Jones, and E. C. Pasztor, “Example-based super resolution,” In Computer Graphics and Applications, pp. 56-65, 2002.
- [2] D. Glasner, S. Bagon, and M. Irani, “Super-resolution from a single image,” In International Conference on Computer Vision, pp. 349-356, 2009.
- [3] C. Dong, C. C. Loy, K. He, and X. Tang. “Image Super-Resolution Using Deep Convolutional Networks,” In Transactions on Pattern Analysis and Machine Intelligence, pp. 295-307, 2016.
- [4] B. Lim, S. Son, H. Kim, S. Nah and K. M. Lee. “Enhanced Deep Residual Networks for Single Image Super-Resolution,” In Conference Workshop on Computer Vision and Pattern Recognition, pp. 1132-1140, 2017.
- [5] M. Haris, G. Shakhnarovich, and N. Ukita, “Deep Back-Projection Networks for Super-Resolution,” In Conference on Computer Vision and Pattern Recognition, pp. 1664-1673, 2018.
- [6] J. Kim, J. K. Lee and K. M. Lee. “Accurate Image Super-Resolution Using Very Deep Convolutional Networks,” In Conference on

- Computer Vision and Pattern Recognition, pp. 1646-1654, 2016.
- [7] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual Dense Network for Image Super-Resolution," In Conference on Computer Vision and Pattern Recognition, pp. 2472-2481, 2018.
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," In Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.
- [9] G. Huang, Z. Liu, L. v. d. Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," In Conference on Computer Vision and Pattern Recognition, pp. 2261-2269, 2017.
- [10] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image Super-Resolution Using Very Deep Residual Channel Attention Networks," In European Conference on Computer Vision, pp. 286-301, 2018.
- [11] C. Dong, C. C. Loy, and X. Tang, "Accelerating the Super-Resolution Convolutional Neural Network," In European Conference on Computer Vision, pp. 391-407, 2016.
- [12] Y. Tai, J. Yang, and X. Liu, "Image Super-Resolution via Deep Recursive Residual Network," In Conference on Computer Vision and Pattern Recognition, pp. 3147-3155, 2017.
- [13] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-Recursive Convolutional Network for Image Super-Resolution," In Conference on Computer Vision and Pattern Recognition, pp. 1637-1645, 2016.
- [14] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," In Conference on Computer Vision and Pattern Recognition, pp. 2818-2826, 2016.
- [15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," In International Conference on Learning Representations, 2015.
- [16] 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 Conference on Computer Vision and Pattern Recognition, pp. 1874-1883, 2016.
- [17] H. Zhao, O. Gallo, J. Frosio, J. Kautz, "Loss functions for image restoration with neural networks," In Transactions on Computational Imaging, pp. 47-57, 2017.
- [18] M. Bevilacqua, A. Roumy, C. Guillemot and M.-L. A. Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," In Proceedings British Machine Vision Conference, 2012.
- [19] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," In International conference on Curves and Surfaces, pages 711-730, 2012.
- [20] 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, 2001.
- [21] J.-B. Huang, A. Singh, and N. Ahuja, "Single Image Super-Resolution From Transformed Self-Exemplars," In Conference on Computer Vision and Pattern Recognition, pp. 5197-5206, 2015.

저자소개

김 인 구



- 2018년 2월 : 영남대학교 정보통신공학과 학사
- 2018년 3월 ~ 현재 : 한양대학교 전자컴퓨터통신공학과 석사과정
- ORCID : <https://orcid.org/0000-0001-9687-3073>
- 주관심분야 : 영상처리, 딥 러닝

유 송 현



- 2015년 2월 : 한양대학교 융합전자공학부 학사
- 2015년 3월 ~ 현재 : 한양대학교 전자컴퓨터통신공학과 석박사통합과정
- ORCID : <https://orcid.org/0000-0003-0725-9202>
- 주관심분야 : 영상처리, 딥 러닝

저자소개



정제창

- 1980년 2월 : 서울대학교 전자공학과 학사
- 1982년 2월 : KAIST 전기전자공학과 석사
- 1990년 : 미국 미시간대학 전기공학과 공학박사
- 1980년 ~ 1986년 : KBS 기술연구소 선임연구원 (디지털 및 뉴미디어 연구)
- 1990년 ~ 1991년 : 미국 미시간대학 전자컴퓨터공학부 연구교수 (영상 및 신호처리 연구)
- 1991년 ~ 1995년 : 삼성전자 HDTV 연구개발 담당 수석연구원
- 1995년 ~ 현재 : 한양대학교 전자컴퓨터통신공학과 교수 (영상통신 및 신호처리 연구실)
- ORCID : <http://orcid.org/0000-0002-3759-3116>
- 주관심분야 : 영상처리, 영상압축