

특집논문 (Special Paper)

방송공학회논문지 제24권 제5호, 2019년 9월 (JBE Vol. 24, No. 5, September 2019)

<https://doi.org/10.5909/JBE.2019.24.5.703>

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

RDB 및 웨이블릿 예측 네트워크 기반 단일 영상을 위한 심층 학습기반 초해상도 기법

응우엔 휴중^{a)}, 김 응 태^{a)‡}

Deep Learning-based SISR (Single Image Super Resolution) Method using RDB (Residual Dense Block) and Wavelet Prediction Network

HUU DUNG NGUYEN^{a)} and Eung-Tae Kim^{a)‡}

요 약

단일 영상 초해상도 (Single Image Super-Resolution - SISR) 기법은 카메라로 획득된 저해상도 영상에 필터 기반의 연산을 적용하여 좋은 화질의 고해상도 영상을 복원하는 과정이다. 최근에 심층 합성곱 신경망 학습의 발전에 따라 단일 영상 초해상도에 적용되는 심층 학습 기법들은 좋은 성과를 보여 주고 있다. 그 대표적인 방법으로 영상의 특징 맵 기반 웨이블릿 계수 학습을 통해 고해상도 영상을 복원하는 WaveletSRNet이 있다. 하지만 복잡한 알고리즘으로 인해 계산량이 증대되어 처리 속도가 늦고 특징 추출할 때 특징 맵을 효율적으로 활용하지 못 한다는 단점을 가지고 있다. 이를 개선하기 위해 본 논문에서는 단일 영상 초해상도 RDB-WaveletSRNet 기법을 제안한다. 제안된 기법은 잔여밀집블록(Residual Dense Block)을 사용하여 저해상도의 특징 맵을 효과적으로 추출하여 초해상도의 성능을 향상시키고 적절한 성장률을 설정하여 복잡한 계산량 문제까지 해결하였다. 또한 웨이블릿 패킷 분해를 사용하여 확대율에 맞게 웨이블릿 계수를 획득하므로 높은 확대율의 단일 영상 초해상도를 얻게 하였다. 다양한 영상에 대한 실험을 통하여, 제안하는 기법이 기존 기법보다 수행시간이 빠르며 영상 품질도 우수함을 입증하였다. 제안하는 방법은 기존 방법보다 화질은 PSNR 0.1813dB만큼 우수하며 속도는 1.17배 빠른 것을 실험을 통해 확인하였다.

Abstract

Single image Super-Resolution (SISR) aims to generate a visually pleasing high-resolution image from its degraded low-resolution measurement. In recent years, deep learning - based super - resolution methods have been actively researched and have shown more reliable and high performance. A typical method is WaveletSRNet, which restores high-resolution images through wavelet coefficient learning based on feature maps of images. However, there are two disadvantages in WaveletSRNet. One is a big processing time due to the complexity of the algorithm. The other is not to utilize feature maps efficiently when extracting input image's features. To improve this problems, we propose an efficient single image super resolution method, named RDB-WaveletSRNet. The proposed method uses the residual dense block to effectively extract low-resolution feature maps to improve single image super-resolution performance. We also adjust appropriated growth rates to solve complex computational problems. In addition, wavelet packet decomposition is used to obtain the wavelet coefficients according to the possibility of large scale ratio. In the experimental result on various images, we have proven that the proposed method has faster processing time and better image quality than the conventional methods. Experimental results have shown that the proposed method has better image quality by increasing 0.1813dB of PSNR and 1.17 times faster than the conventional method.

Keyword : Super Resolution, Deep Learning, Wavelet Coefficient, Residual Dense Block, WaveletSRNet

1. 서론

단일 영상 초해상도 (Single Image Super-Resolution - SISR)는 획득되는 저해상도 영상에 필터 기반의 연산을 적용하여 좋은 화질의 고해상도 영상을 복원하는 기법이다. 고해상도 영상은 획득되는 저해상도 영상보다 화소 밀도가 높고 더 많은 정보를 제공한다. 단일 영상 초해상도 기법은 디지털 TV, 모바일, 보안 카메라 등 다양한 분야에서 활용되고 있다. 게다가 최근 컴퓨터 성능 향상으로 인해 패턴 인식 및 이미지 분석 기법, 영상 부호화/복호화, 영상 잡음 제거, 의료 영상 분석 등 다양한 기법을 초해상도 기법과 결합해서 성능을 향상시키는데 사용하고 있다.

단일 영상의 해상도를 높이는 기법은 보간법 (Interpolation)^[1-4], 재구성법 (reconstruction)^[6-8] 및 심층 학습법 (Deep Learning)^[10,12,14,15] 3가지로 분류할 수 있다. 보간법 기반 초해상도는 해상도를 높이기 위해 기존 화소들 사이에 적절한 화소를 삽입하는 기술이며 간단하고 처리 속도가 비교적 빠르다는 장점이 있어 실시간 응용에 적합하다. 그러나 이러한 기법은 저주파수 대역필터 특성을 가지고 있으므로 계단 (Jagging) 현상 및 겹침 왜곡 (Aliasing), 파문 (Ringing) 현상에 의해 고주파수 성분이 효과적으로 복원될 수 없다는 단점이 있다. 재구성법 기반 초해상도는 Multi-Scale Expected Patch Log Likelihood (EPLL)^[7] 및 Adaptive High-Dimensional Non-Local Total Variation (AHNLTV)^[8], Steering Kernel Regression (SKR)과 Non-local Means (NLM)^[6] 등 다양한 기법의 발달에 의해 많이 개선되어 보간법보다 계단 현상 및 겹침 왜곡, 파문 현상이

많이 감소되었다. 하지만, 보간법에 비해 상대적으로 복잡한 계산량을 가지는 단점이 있다.

최근에는 CNN (Convolution Neural Networks) 기술의 발달에 힘입어 단일 영상 초해상도 분야에서도 CNN 기술을 적용한 방법들이 좋은 성능을 보여 주고 있다. 최초로 초해상도 분야에 CNN을 적용하는 SRCNN^[13]로부터 최근에 제안된 방법^[14,17]들은 화소 단위 평균 자승 오차 (Mean Squared Error) 손실 함수 기반으로 학습을 수행한다. 학습을 통해 출력 영상 화소 단위로 점차 실측정보(ground-truth) 고해상도 영상에 근접되고 있다. 그러나 이러한 방법은 뭉롱화(blur) 현상과 과도한 평탄화 현상에 의해 출력 영상의 일부 정보, 특히 고주파수 정보 손실이 많이 발생된다. 게다가 이러한 방법은 한정된 확대율 (2x 혹은 4x) 일 때만 잘 수행되고 큰 확대율 (8x 혹은 16x) 일 경우에는 성능이 급격하게 저하된다. 최근에는 GAN을 적용하는 기법^[15,18]들은 단일 영상 초해상도의 성능을 많이 개선하였으나 이러한 방법들은 학습시 정량적 성능 측정 기준이 미비한 문제가 존재하고 진동(oscillation)문제가 나타나 오래 학습해도 더 이상 수렴하지 못 해 학습이 불안정하게 된다.

웨이블릿 변환 (Wavelet transform)은 다중 해상도 (multi-resolution) 영상을 표현 및 복원할 때 효율적인 기법이다^[9]. 웨이블릿 변환은 영상의 정보를 다중 레벨 (different levels)에서 표현할 수 있어 높은 확대율 초해상도를 수행 가능하고 고주파수 성분을 잘 나타내는 장점이 있어 학습 기반 초해상도 분야에 적용 되었다^[10,11]. 그러나 이러한 방법은 복잡한 계산량 때문에 학습 진행에 어려움이 있고 저해상도 영상의 특징 (feature)을 잘 찾아 내지 못 하는 단점이 있다.

본 논문에서는 이러한 문제를 해결하기 위해 단일 영상 초해상도 네트워크인 RDB-WaveletSRNet를 제안한다. 제안된 방식은 저해상도 영상의 특징 맵을 잘 추출해내기 위해 특징 추출 네트워크에 Residual Dense Block(RDB)기법을 적용한다. RDB의 성장률(growth rate) 계수를 적절히 조절하여 복잡한 계산량 문제를 개선하도록 하였다. 또한 영상의 특징 맵 (feature map)으로 부터 웨이블릿 영역의 계수

a) 한국산업기술대학교 전자공학부(Department of Electronics Engineering, Korea Polytechnic University)

‡ Corresponding Author : 김응태(Eung-Tae Kim)

E-mail: etkim@kpu.ac.kr

Tel: +82-31-8041-0488

ORCID: <https://orcid.org/0000-0001-5984-0045>

※ 이 논문의 연구 결과 중 일부는 한국방송·미디어공학회 “2019년 하계 학술대회”에서 발표한 바 있음.

· Manuscript received July 22, 2019; Revised September 4, 2019; Accepted September 4, 2019.

를 학습하고 확대율에 맞게 다중 레벨의 웨이블릿 계수를 구해서 초해상도 영상을 구현하게 된다.

본 논문의 구성은 다음과 같다. 2장에서 단일 영상 초해상도에 관련한 기존 연구들을 설명한다. 3장에서 제안하는 RDB-WaveletSRNet 네트워크 구조를 설명한다. 4장에서는 모의실험을 통해 기존의 기법들과 성능을 비교한다. 5장에서 결론은 맺는다.

II. 기존 초해상도 관련 연구

초해상도 분야와 관련하여 많은 웨이블릿 기반 기법들이 제안되었다. 그러나 저해상도 시퀀스 영상으로부터 초해상도 영상을 추정하거나, 단일 영상 초해상도에서는 보간법과 재구성법 기반 초해상도를 보조해 주는 데만 웨이블릿이 적용된 기법이 대부분이다. Gao et al.^[5]은 하이브리드 웨이블릿 합성곱 네트워크(hybrid wavelet convolution network)를 제안했다. 이 기법은 웨이블릿을 활용하여 sparse code 후보자 (candidates)를 생성하고 이를 CNN을 통해 고해상도 영상을 추출해내는 기법이다.

Huang et al.^[10]은 그림 1에서 보는 바와 같이 저해상도 영상의 특징 맵으로부터 심층 학습으로 웨이블릿 계수를 학습시켜 고해상도 영상을 추출하는 WaveletSRNet 네트워크

를 제안했다. 높은 확대율 (8x와 16x)에서도 다른 기법과 비교할 때 상대적으로 좋은 성능을 보여 주었다.

WaveletSRNet에서는 저해상도 영상의 특징 맵을 추출할 때 제안된 Embedding Net을 사용한다. Embedding Net은 잔여블록(residual block)^[19]기반 심층 학습 네트워크이다. 다양한 분야에서 성능을 개선하기 위해 더 많은 계층의 심층 학습을 이용하는 경향이 있다. 그러나 전통적으로 많이 쓰는 활성화 함수의 기울기(gradient) 범위는 0에서 1까지이며 네트워크가 깊어질수록 기울기가 0으로 수렴하게 되는 소실(vanishing) 문제가 발생한다. 잔여블록은 shortcut connection과 element-wise addition 기법으로 이 소실 문제를 잘 해결하였다. ILSVRC 2015에서, 3.57%의 가장 낮은 top-5 오차율 가진 ResNet^[19]은 잔여블록의 좋은 성능을 보여 주었다. 그러나 잔여블록은 마지막 계층에만 과거 특징 정보만 도입되고 다른 계층에는 과거 특징 정보가 활용되지 못하는 단점이 있다. 게다가 콘볼루션의 크기가 증가함에 따라 잔여블록의 계산량은 급격하게 증가되는 문제가 있다.

이 문제를 해결하기 위해 Zhang et al.^[12]은 잔여밀집블록(residual dense block, RDB)을 제안하였다. RDB은 잔여블록의 장점을 여전히 이어받고 dense connection^[20]으로 각 계층들은 그 전의 모든 계층의 정보를 활용하여 영상의 특징 맵을 추출하는데 적합한 기법이다.

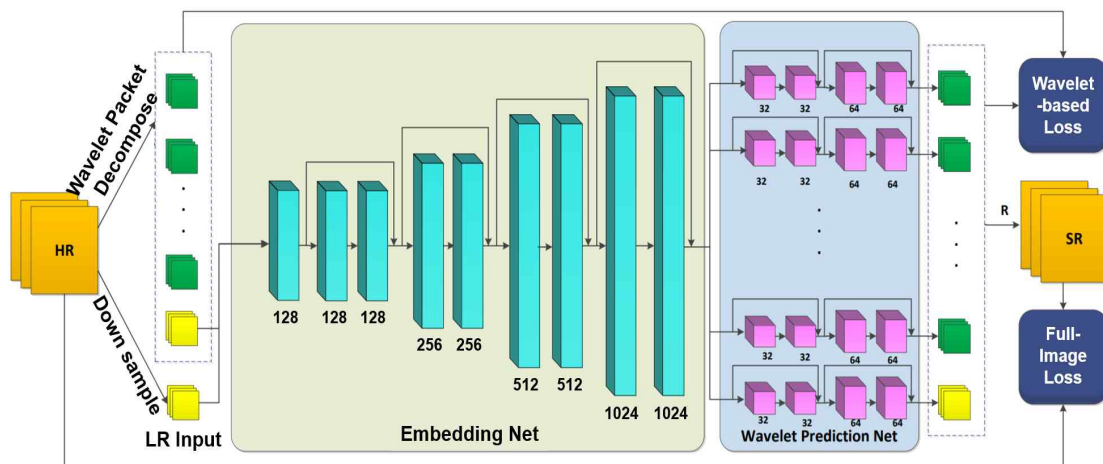


그림 1. WaveletSRNet 네트워크 구조
 Fig. 1. Architecture of WaveletSRNet

III. 제안된 RDB-WaveletSRNet기반 초고해상도 기법

본 논문은 그림 1의 WaveletSRNet의 성능을 개선하기 위해 RDB를 이용하는 WaveletSRNet를 제안한다. 그림 2는 RDB-WaveletSRNet의 구조를 보여준다.

RDB-WaveletSRNet은 주어진 원본 고해상도 영상을 웨이블릿 패킷 분해(Wavelet Packet Decomposition)을 통해 확대율에 맞는 다양한 레벨의 웨이블릿 계수들과 저해상도 영상으로 분해한다. 저해상도 영상은 특징 추출망(Feature Extraction Net, FEN)의 입력이 되고, 이 네트워크를 통해 저해상도 영상의 특징 맵을 추출하게 된다. 웨이블릿 예측망(Wavelet Prediction Net)은 저해상도 영상의 특징 맵과 분해된 웨이블릿 계수기반 고해상도의 웨이블릿 계수를 학습시킨다. 복원단계에서는 학습을 통해 얻어진 웨이블릿 계수를 저해상도 영상의 초해상도 영상으로 복원하는 단계이다.

망(Wavelet Prediction Net)은 저해상도 영상의 특징 맵과 분해된 웨이블릿 계수기반 고해상도의 웨이블릿 계수를 학습시킨다. 복원단계에서는 학습을 통해 얻어진 웨이블릿 계수를 저해상도 영상의 초해상도 영상으로 복원하는 단계이다.

1. Wavelet Packet Decomposition

제안된 방식에서 2D 고속 웨이블릿 변환 (fast wavelet transform, FWT)^[16]을 사용하여 입력 고해상도 영상의 N_w 개의 웨이블릿 계수를 분해한다. 웨이블릿 변환은 적용되는 웨이블릿 원형함수 (wavelet function/ wavelet mother

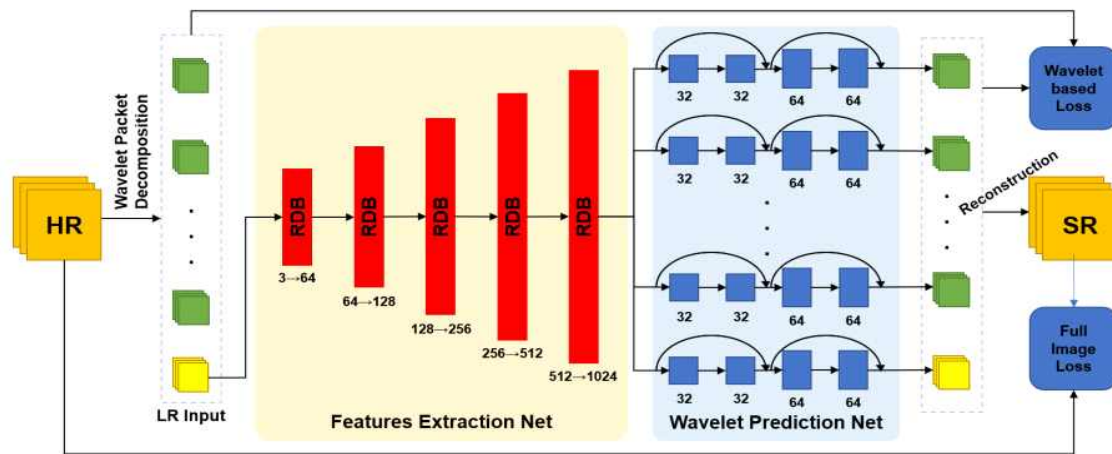


그림 2. 제안된 RDB-WaveletSRNet 네트워크의 구조
Fig. 2. The architecture of proposed network RDB-WaveletSRNet

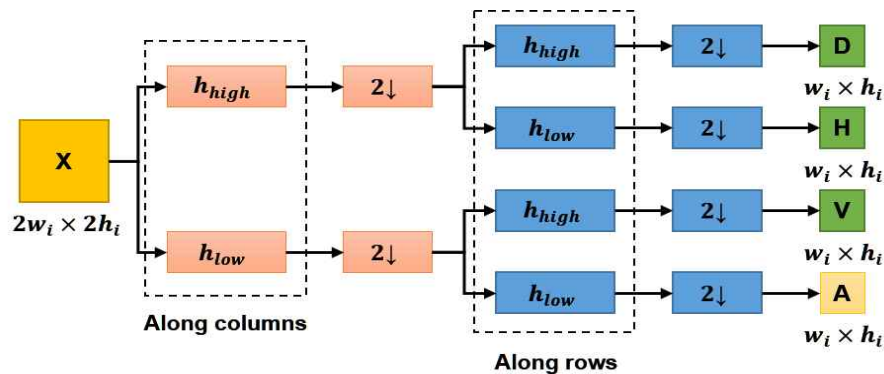


그림 3. i 번째 레벨의 FWT
Fig. 3. i level FWT

function)로 분류된다. 따라서 그 중에 가장 간단하고 계산량이 낮으며 에지 성분에 아무 영향도 없이 완벽히 복원할 수 있는 특징을 가지는 Haar 웨이블릿을 선택하였다.

제안된 방식에서 확대율 $r=8$ 인 경우, $l=\log_2(r)=3$ 개의 FWT 레벨로 N_w 개의 웨이블릿 계수를 분해할 수 있다. 여기서, $N_w=4^l=64$ 이다. 이전 웨이블릿 계수들을 그림 3에서 나타내는 FWT 3레벨까지 반복하여 통과시킨 후 N_w 개의 웨이블릿 계수를 계산한다.

2. Features Extraction Net (FEN)

웨이블릿 패킷 분해를 통해서 얻어진 웨이블릿 계수들 중에 가장 낮은 주파수 성분 가진 웨이블릿 계수는 바로 주어진 고해상도 영상의 확대를 만큼 다운 스케일된 저해상도 영상이다. FEN을 통해 이 저해상도 영상의 특징 세트를 추출한다. $3 \times w \times h$ 크기인 입력영상을 FEN을 통해 $w \times h$ 크기를 가지는 1024개의 특징 맵을 추정한다.

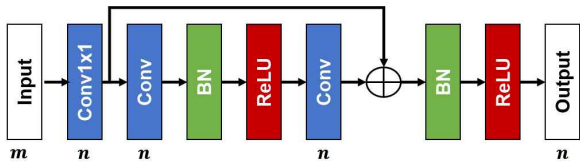


그림 4. Residual Block 구조
 Fig. 4. Architecture of Residual Block

FEN은 서로 크기가 다른 병렬로 연결된 5개의 RDB로 구성되어 있다. RDB는 연속 메모리 구조 (Contiguous Memory, CM mechanism)와 국소특징결합 (Local Feature Fusion), 국소잔여학습 (Local Residual Learning) 등을 통

하여 블록 안의 모든 층에서 전의 모든 층의 정보를 취득하고 전 블록의 정보까지 도입되는 형태로써 RB보다 더 효율적으로 학습이 진행되고 더 좁게 설계 가능하여 계산량이 개선되었다. 그림 5는 본 논문에서 사용하는 RDB 구조이다. 각 RDB를 통해 저해상도 영상의 특징 개수를 확대시키므로 RDB의 입력 채널 개수 m 과 출력 채널 개수 n 은 서로 다르다. RDB의 addition 연산을 하기 위해 m 개 채널 입력을 n 개로 확대해 주는 팽창 합성곱(expand convolution, Conv_ex)을 사용하는 것을 제안한다. g 는 RDB의 성장률(growth rate)이며 잔여블록의 복잡한 계산량 문제를 해결하기 위해 적절한 $g=32$ 를 사용한다.

잔여블록(RB) 구조로부터 RB의 파라미터 수를 계산하는 식은 다음과 같다.

$$Params_{RB} = mn + 2nnk_xk_y \quad (1)$$

여기서 m 과 n 은 입력 및 출력 채널의 크기이고 k_x 과 k_y 는 convolution의 커널 크기이며 RB와 RDB는 동일하게 3×3 커널을 쓴다.

본 논문에서 사용하는 잔여밀집블록(RDB)의 파라미터 수를 계산하는 식은

$$Params_{RDB} = mnk_xk_y + (Dg+n)n + \sum_{i=0}^{D-1} (n+ig)k_xk_y \quad (2)$$

이다. 여기서 D 는 RDB의 dense 계층 수이며, 본 논문에서는 $D=3$ 을 적용한다. g 는 RDB의 성장률이다^[12]. 표 1은 식 (1)과 (2)로부터 채널 크기에 따라 RB와 RDB의 파라미

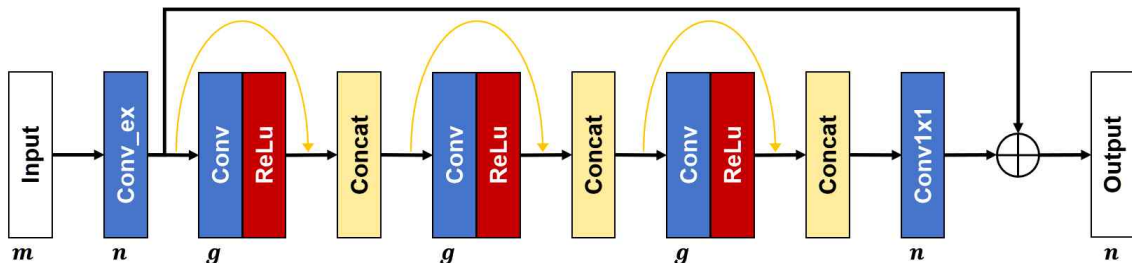


그림 5. 제안된 방식에서 사용되는 개선된 RDB 구조
 Fig. 5. The architecture of proposed RDB

터수를 나타낸다. 표 1을 통해 낮은 계산량으로 RDB는 폭이 넓고 깊은 학습에 적합함을 보여주었다.

표 1. 입/출력 채널수에 따라 RB와 RDB의 파라미터 수 비교
Table 1. Comparing number of parameters of RB and RDB according to the number of input/ output channels

Number of Input → Output channels ($m \rightarrow n$)	Number of Parameters	
	RB - Residual Block	RDB - Residual Dense Block
3→64	39,040	94,912
64→128	229,888	240,640
128→256	918,528	633,856
256→512	3,672,064	1,960,960
512→1024	14,684,160	6,777,856
Total	19,543,680	9,708,224

3. Wavelet Prediction Network

복구할 초해상도 영상의 N_w 개의 웨이블릿 계수를 N_w 개의 독립적인 서브넷(subnet)으로 웨이블릿 예측 네트워크로 학습한다. 각 서브넷은 2개의 잔여블록으로 구성되며 $1024 \times w \times h$ 특징 맵 입력으로 해당되는 웨이블릿 계수를 생성한다. 저해상도 입력 영상과 모든 N_w 개의 웨이블릿 계수의 크기는 동일하게 $3 \times w \times h$ 로 유지하기 위해 본 논문에서 WPN에 3×3 커널, stride 1, padding 1 가지는 convolution을 사용한다. Haar 웨이블릿의 웨이블릿 계수들은

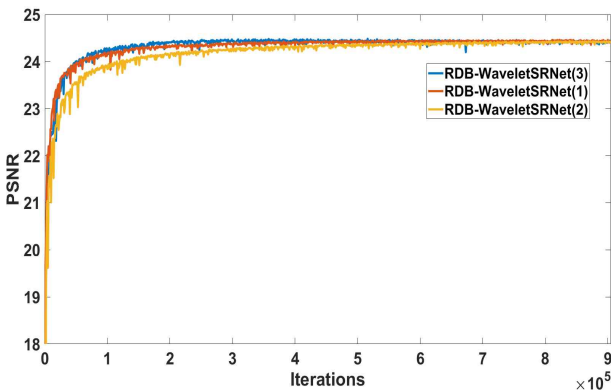


그림 6. 학습 반복에 따른 RDB-WaveletSRNet(1),(2),(3)의 PSNR 변환
Fig. 6. PSNR convergence analysis on RDB-WaveletSRNet(1),(2),(3)

서로 독립적이기 때문에 N_w 개의 서브넷을 병렬로 독립적으로 설계한다.

4. Reconstruction

복원단계에서는 단순히 *invest-FWT*를 사용하여 N_w 개의 $3 \times w \times h$ 웨이블릿 계수를 원래 $3 \times (r \times w) \times (r \times h)$ 영상으로 복원하는 단계이다.

IV. 모의실험 결과

본 모의실험에서는 CelebA^[21]의 얼굴 영상 데이터 셋을 기반으로 성능을 분석하였다. CelebA 얼굴 영상 데이터 셋은 원본영상 (In-The-Wild)과 얼굴 가운데 정렬된 크기인 데이터 셋이 각 202,599개의 영상이 있다. 본 실험에서는 얼굴 가운데 정렬된 데이터 셋을 쌍입방 보간법을 통해 다운 스케일하여 사용되는 고해상도 영상을 획득한다. 획득된 고해상도 영상의 202,599개 영상 중에 162,770개 학습 영상, 19,867개 validation 영상, 나머지 19,962개 테스트 영상으로 사용했다. 학습 데이터 셋을 (0번째 epoch 포함)151개의 epoch (768k iterations)을 학습한 후 19,962개의 영상으로 bicubic 보간법, WaveletSRNet^[10]과 PSNR 및 SSIM을 비교하여 제안된 RDB-WaveletSRNet의 성능을 검토하였다.

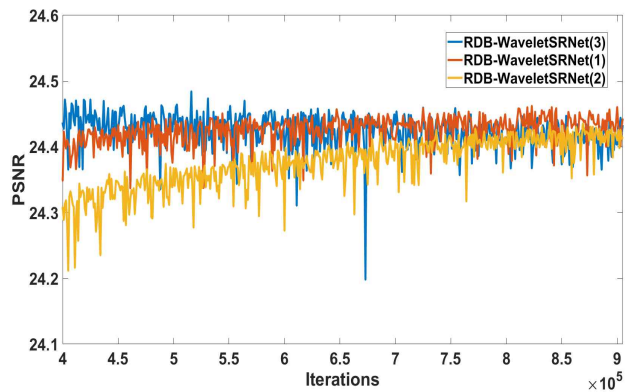


그림 7. 학습 반복에 따른 RDB-WaveletSRNet(1),(2),(3)의 PSNR 변환 (확대)
Fig. 7. PSNR convergence analysis on RDB-WaveletSRNet(1),(2),(3) models (zoom)

그림 6과 그림 7은 실험 대상인 3개 RDB-WaveletSRNet의 학습 반복에 따른 PSNR 변환을 나타낸다. RDB-WaveletSRNet(1)은 다음과 같은 조건하에 학습한다.

- 특징 추출 망에서 잔여밀집블록의 expand convolution 크기는 3×3
- 잔여밀집블록은 그림 4-8에서 보는 바와 같이 동일 성장률 $g=32$ 사용
- 학습 미니배치 b_l 와 테스트 미니배치 b_t 는 32를 사용

RDB-WaveletSRNet(2)는 특징 추출 망에서 잔여밀집블록의 팽창합성곱의 크기를 1×1 형태로 사용한다. RDB-WaveletSRNet(3)은 고정 성장률이 아닌 특징 추출 망의 5개 잔여밀집블록에서 블록 크기의 증가함에 따라 적응적으로 성장률을 32씩 증가를 시켜 사용하여 실험을 진행하였다. 즉, #1 RDB의 성장률은 32, #2 RDB의 성장률은 64,... #5 RDB의 성장률은 160으로 사용한다.

표 2는 실험 대상인 RDB-WaveletSRNet(1), (2)와 (3) 3개의 네트워크의 성능 비교를 보여 준다. 잔여밀집블록에서 3×3 크기인 팽창 합성곱과 성장률 $g=32$ 를 가지는 RDB-WaveletSRNet(1)은 성능이 가장 높은 방법으로 볼 수 있다.

실험 결과를 통해서 기본 네트워크 RDB-WaveletSRNet(1)에 비해 잔여밀집블록의 1×1 크기인 팽창 합성곱을 사용하는 RDB-WaveletSRNet(2)는 PSNR과 SSIM 성능이 저하되지만 계산 량과 학습 시간은 각 86%와 88%로 줄어든 것을 확인하였다. 또한 그림 7을 보면 700k iteration부터 RDB-WaveletSRNet(1)의 PSNR은 거의 변화하지 않는 반면에 RDB-WaveletSRNet(2)의 PSNR은 계속 상승하는 것

이 보인다. 즉, 더 많은 epoch로 학습할 때 RDB-WaveletSRNet(2)는 효율적인 방법이 될 수도 있다.

그림 6에서 보는 바와 같이 적응적 성장률을 사용하는 RDB-WaveletSRNet(3)은 다른 2개의 기법에 비해 더 빨리 최솟값 (global minimum)을 찾을 수 있지만 500k iteration부터 과대적합 (overfitting)이 발생하여 학습 성능이 저하되는 것을 그림 7에서 확인할 수 있다.

그림 8은 150개의 epoch (905k iterations)을 통해 1000 iterations마다 WaveletSRNet과 제안된 RDB-WaveletSRNet의 PSNR을 측정 한 학습 결과이다. 첫 번째 150k iteration은 RD가 파라미터 수가 많아서 RDB보다 학습이 더 잘 되었으나 200k iteration부터 WaveletSRNet은 과대적합(overfitting)이 된다. 반면에 제안하는 RDB-WaveletSRNet은 계속 잘 학습되는 것을 볼 수 있다.

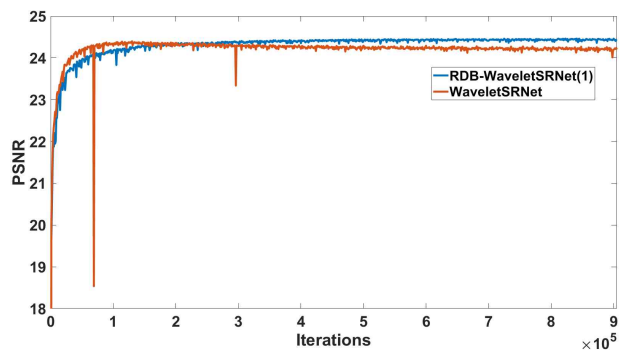


그림 8. 학습 반복에 따른 RDB-WaveletSRNet(1)과 WaveletSRNet의 PSNR 결과

Fig. 8. The PSNR result of RDB-WaveletSRNet(1) and WaveletSRNet according to the learning iterations

그림 9에서 보는 바와 같이 CelebA 데이터셋의 얼굴 가운데 정렬된 178×218 크기인 데이터셋으로부터 128×128

표 2. RDB-WaveletSRNet(1), (2)와 (3)의 PSNR, SSIM, 파라미터 수와 학습시간 비교 (확대율 $r=8$)
 Table 2. Benchmark results with RDB-WaveletSRNet(1), (2) and (3) models. PSNR, SSIM, number of parameters, training time for scaling factor $r=8$

Method	PSNR	SSIM	FEN #Parameter	Total #Parameter	Training Time (hours)
Ours(1)	24.7903	0.8521	9.7M	40M	82
Ours(2)	24.7434	0.8503	4.1M	34.4M	72
Ours(3)	24.7752	0.8514	11.5M	41.8M	86

의 고해상도 영상을 얻었다. 그림 9에서 (b) 16×16 입력 저해상도를 확대율 $r=8$ 경우 (c) bicubic 보간법, (d) WaveletSRNet^[10]과 제안된 RDB-WaveletSRNet의 결과를 보여 준다. 그림 9에서 제안된 방식은 원본 영상(HR)과 가장 비슷하게 고해상도 영상을 얻음을 보여 준다. 특히 3번째와 4번째 영상의 뒤 배경 글씨나 패턴 등이 더 원 영상에 비슷하고, 전체적으로 얼굴 윤곽에 대해서 좋은 효과를 보이고 있다.

표 3에서 19,962개의 테스트 영상을 대상으로 bicubic 보간법과 WaveletSRNet, RDB-WaveletSRNet(1)을 통해 측정된 PSNR, SSIM, 파라미터 수와 수행시간을 보여준다. GTX 1070 GPU를 사용하여 학습을 진행한 결과, 기존 WaveletSRNet보다 PSNR이나 시간측면에서 개선된 효과를 얻음을 알 수 있었다. 제안하는 방법은 기존 방법보다 화질은 PSNR 0.18dB만큼 우수하며 속도는 1.17배 빠른 것을 모의실험을 통해 확인하였다.

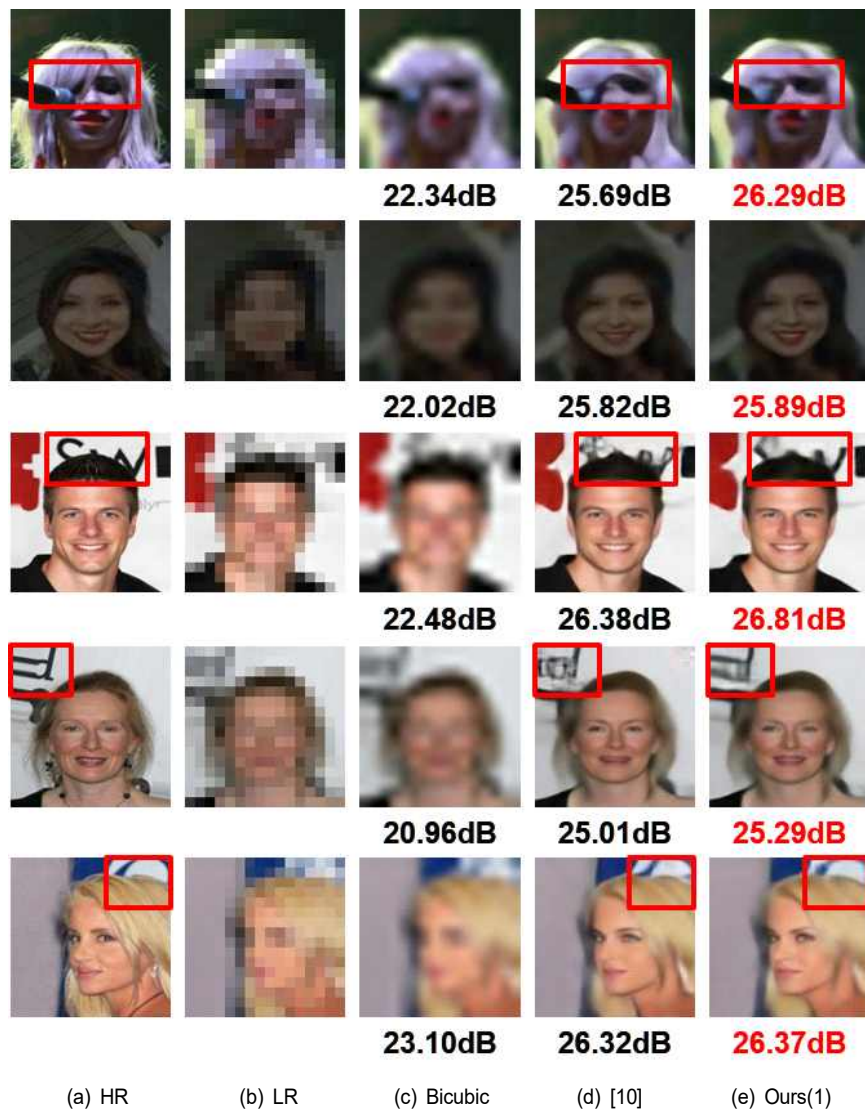


그림 9. 16×16 입력 저해상도, 확대율 $r=8$ 경우 기존 방식들과 비교

Fig. 9. Visual comparison results with the conventional methods when input image size is 16×16 and scaling factor $r=8$

표 3. CelebA 데이터 세트 기반 PSNR, SSIM, 파라미터 수와 학습시간 비교
 Table 3. Benchmark results compared with the state-of-the-art methods

Method	PSNR	SSIM	FEN #Parameter	Total #Parameter	Training Time (hours)
Bicubic	21.6975	0.7787	-	-	-
WaveletSRNet	24.6090	0.8496	19.5M	50M	96
Ours	24.7903	0.8521	9.7M	40M	82

V. 결론

최근 고해상도 디스플레이 기술이 발전함에 따라서 기존에 획득된 저해상도 영상을 고해상도 영상으로 복원하는 기술을 필요로 하는 영상장치가 증가하고 있다. 이를 위해 획득된 저해상도 영상의 해상도나 화질을 향상시키는 초해상도 방법이 연구되어져 왔으며 특히 최근에는 심층학습 기반 초해상도 방법들이 활발하게 연구되며 신뢰성 있고 높은 성능을 보이고 있다. 그 대표적인 방법으로 영상의 특징 맵 기반 웨이블릿 계수 학습을 통해 고해상도 영상을 복원하는 WaveletSRNet이 있다. 하지만 복잡한 알고리즘으로 인해 계산량이 증대되어 처리 속도가 늦고 특징 추출할 때 이전 특징 맵을 효율적으로 활용을 못 한다는 단점을 가지고 있다.

이 문제를 개선하기 위해 본 논문에서는 저해상도의 특징 맵을 효과적으로 추출하여 초해상도의 성능을 향상시키고 적절한 성장률 설정하여 복잡한 계산량 문제까지 해결하는 잔여밀집블록 (RDB)과 웨이블릿 예측 네트워크 기반 초해상도 RDB-WaveletSRNet 기법을 제안하였다. 또한 웨이블릿 패킷 분해 (Wavelet Packet Decomposition)를 사용하여 확대율에 맞게 웨이블릿 계수를 획득하므로 높은 확대율의 단일 영상 초해상도 수행하도록 하였다.

다양한 영상에 대한 실험으로 제안하는 기법에 최적 성장률과 가장 효율적인 특징 추출 망을 설계하였다. 실험 결과를 통하여, 기존 기법보다 수행시간이 빠르며 영상 화질도 우수함을 입증하였다. 제안하는 방법은 기존 방법보다 화질은 PSNR 0.1813dB만큼 우수하며 속도는 1.17배 빠른 것을 실험을 통해 확인하였다.

본 논문에서 제안된 RDB-WaveletSRNet 기법은 실제 어플리케이션에 보다 빠른 처리 속도와 우수한 화질로 디지털 TV, 보안 카메라, 의료 영상, 영상 편집 분야 등에 널리 이용될 수 있다.

참고 문헌 (References)

- [1] X. Zhang and X. Wu, "Image Interpolation by Adaptive 2-D Autoregressive Modeling and Soft-Decision Estimation," in IEEE Trans. Image Process., 2008, <https://doi.org/10.1109/tip.2008.924279>
- [2] Y. Romano, M. Protter and M. Elad, "Single Image Interpolation Via Adaptive Nonlocal Sparsity-Based Modeling," in IEEE Trans. Image Process., 2014, <https://doi.org/10.1109/tip.2014.2325774>
- [3] W. Ye and K. Ma, "Convolutional Edge Diffusion for Fast Contrast-guided Image Interpolation," in IEEE Signal Process. Lett., 2016, <https://doi.org/10.1109/lsp.2016.2571738>
- [4] C. Ji, X. He, Y. Fu, Z. Liang, and L. Qing, "An edge directed interpolation algorithm based on regularization," J. Electron. Inf. Technol., vol. 36, no. 2, pp. 293 - 297, 2014.
- [5] X. Gao and H. Xiong, "A hybrid wavelet convolution network with sparse-coding for image super-resolution," in ICIP, 2016, <https://doi.org/10.1109/icip.2016.7532596>
- [6] K. Zhang, X. Gao, D. Tao and X. Li, "Single Image Super-Resolution With Non-Local Means and Steering Kernel Regression," in IEEE Trans. Image Process., 2012, <https://doi.org/10.1109/tip.2012.2208977>
- [7] V. Pappyan and M. Elad, "Multi-Scale Patch-Based Image Restoration," in IEEE Trans. Image Process., 2016, <https://doi.org/10.1109/tip.2015.2499698>
- [8] C. Ren, X. He and T. Q. Nguyen, "Single Image Super-Resolution via Adaptive High-Dimensional Non-Local Total Variation and Adaptive Geometric Feature," in IEEE Trans. Image Process., 2017, <https://doi.org/10.1109/tip.2016.2619265>
- [9] S. Mallat, "Wavelets for a vision," in Proc. IEEE, 1996, <https://doi.org/10.1109/5.488702>
- [10] H. Huang, R. He, Z. Sun and T. Tan, "Wavelet-SRNet: A Wavelet-Based CNN for Multi-Scale Face Super Resolution," in ICCV, 2017, <https://doi.org/10.1109/iccv.2017.187>
- [11] Y. Matsuo and S. Sakaida, "Super-resolution for 2K/8K television using wavelet-based image registration," in IEEE GlobalSIP, 2017, <https://doi.org/10.1109/globalsip.2017.8308668>
- [12] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in CVPR, 2018, <https://doi.org/10.1109/cvpr.2018.00262>
- [13] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a Deep Convolutional Network for Image Super-Resolution," in ECCV, 2014, https://doi.org/10.1007/978-3-319-10593-2_13
- [14] J. Bruna, P. Sprechmann, and Y. LeCun, "Super-Resolution with Deep

- Convolutional Sufficient Statistics,” in ICLR, 2016, <https://arxiv.org/abs/1511.05666>
- [15] S. Hitawala, Y. Li, X. Wang, D. Yang, "Image Super-Resolution Using VDSR-ResNeXt and SRCGAN," in CVPR, 2018, <https://arxiv.org/abs/1810.05731>
- [16] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," in TPAMI, 1989, <https://doi.org/10.1109/34.192463>
- [17] H. Chang, D.-Y. Yeung and Y. Xiong, "Super-resolution through neighbor embedding," in CVPR, 2004, <https://doi.org/10.1109/CVPR.2004.1315043>
- [18] W. Zhang, Y. Liu, C. Dong, Y. Qiao, "RankSRGAN: Generative Adversarial Networks with Ranker for Image Super-Resolution," in CVPR, 2019, <https://arxiv.org/abs/1908.06382>
- [19] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in CVPR, 2015, <https://doi.org/10.1109/cvpr.2016.90>
- [20] T. Tong, G. Li, X. Liu and Q. Gao, "Image Super-Resolution Using Dense Skip Connections," in ICCV, 2017, <https://doi.org/10.1109/iccv.2017.514>
- [21] Z. Liu, P. Luo, X. Wang and X. Tang, "Deep Learning Face Attributes in the Wild," in ICCV, 2015, <https://doi.org/10.1109/iccv.2015.425>
- [22] G. Huang, Z. Liu, L. v. d. Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," in Computer Vision and Pattern Recognition, 2017, <https://arxiv.org/abs/1608.06993>

저 자 소 개

응우엔 휴중



- 2017년 : 한국산업기술대학교 전자공학부 공학사
- 2019년 8월 : 한국산업기술대학교 전자공학부 공학석사
- 2019년 9월 ~ 현재 : ㈜탐엔지니어링 연구원
- ORCID : <https://orcid.org/0000-0002-1889-6197>
- 주관심분야 : 딥러닝기반 영상처리, 사물인터넷, 컴퓨터 비전, Blockchain

김 응 태



- 1991년 : 인하대학교 전자공학과 공학사
- 1993년 : KAIST 전기및전자공학과 공학석사
- 1999년 : KAIST 전기및전자공학과 공학박사
- 1998년 3월 ~ 2004년 2월 : ㈜LG전자 DTV연구소 책임연구원
- 2004년 3월 ~ 현재 : 한국산업기술대학교 전자공학부 교수
- ORCID : <https://orcid.org/0000-0001-5984-0045>
- 주관심분야 : 멀티미디어 신호처리, DTV SOC, 지능형 영상감시 시스템, 딥러닝기반 영상처리