

Cross-CNN 기반의 초해상도 기술

박장수, 이종석, 박시내, 심동규
 광운대학교
 {jangsoopark, suk2080, psea1118, dgsim}@kw.ac.kr

Cross-CNN based Image Super Resolution

Jangsoo Park, Jongseok Lee, SeaNae Park, Donggyu Sim
 Kwangwoon University

요 약

본 논문은 초해상도 영상 생성을 위한 CNN 구조를 제안한다. 제안하는 Cross-CNN 은 2 차원 필터의 분리성과 활성화 함수의 비선형성을 바탕으로 VDSR 구조의 시작과 마지막 층을 제외한 중간 층들에 교차하는 1 차원 필터를 적용한다. 제안하는 방법은 기존의 방법보다 적은 가중치를 사용하여 실행 시간을 단축하였다. 실험은 VDSR 실험에 사용된 291 개의 영상과 B100 영상을 이용하였다. 제안하는 방법은 네트워크 중간층에서 기존 방법 대비 약 1/3 의 가중치를 사용하여 20%의 속도 향상을 보였다.

1. 서론

저해상도(Low Resolution) 영상으로부터 고해상도(High Resolution) 영상을 생성하기 위하여 사용되는 초해상도 기술(Super Resolution)은 컴퓨터 비전, 영상 압축 등 다양한 영상처리 분야에 사용된다. 컴퓨터 비전 분야에서는 의료 영상, 위성 영상 등에서 관심 영역의 자세한 정보 추출을 위하여 사용하며, 영상 압축 분야에서는 압축 효율 향상을 위해 다운 샘플링(down sampling)하여 압축된 영상을 복호화 할 때 원본 영상의 크기로 복원하기 위하여 사용한다[1, 2]. 초해상도 기술은 하나의 저해상도 영상이 다수의 고해상도 영상이 존재할 수 있다. 즉 초해상도 기술은 명확한 해답이 없는 문제(ill-posed problem)로 알려져 있다. 따라서 최적의 고해상도 영상을 찾기 위한 많은 연구들이 선행되어왔다[1, 3, 4].

최근 컴퓨터 비전, 영상 압축 등 다양한 영상처리 분야에서 딥 러닝(Deep Learning)을 이용한 많은 기술들이 연구되어 왔다[5, 6]. 딥 러닝 기반 영상처리 기술들은 활성화 함수의 비선형성을 통해 깊은 층의 네트워크를 구성하여 기존 방법들 대비 좋은 성능을 보여왔으며 초해상도 기술 또한 딥 러닝을 이용한 연구들이 활발히 진행되고 있다[3, 4]. 딥 러닝 기반 초해상도 기술은 초해상도 영상 생성을 위하여 저해상도 입력 영상과 고해상도 정답(Ground-Truth) 영상의 관계를 학습한다. SRCNN[3]은 3 개의 층을 갖는 CNN(Convolutional Neural Network)을 통해 저해상도 영상과 고해상도 영상의 대응 관계를 직접 학습하여 고해상도 영상을 예측하였다. VDSR[4]은 20 개의 층을 갖는 CNN 에 잔여 학습(residual learning)을 적용하여 저해상도 영상과 고해상도 영상 간의 차이를 학습하여 저해상도 영상과 네트워크의 출력을 더함으로 고해상도 영상을 예측하였으며 기존 방법들 대비 높은 PSNR 성능을 보였다. 하지만 VDSR 은 깊은 층으로 구성된 CNN 을

사용함에 따라 높은 계산 복잡도를 가진다는 단점이 있다. 따라서 본 논문에서는 기존 VDSR 의 계산 복잡도 개선을 위한 1 차원 필터 기반의 CNN 구조인 Cross-CNN 을 제안한다. 제안하는 Cross-CNN 은 영상의 분리성(Separability) 및 활성화 함수(activation function)의 비선형성을 바탕으로 네트워크 시작 및 마지막 층을 제외한 중간에 존재하는 각층들에 교차하는 1 차원 필터를 사용한다. 제안하는 구조는 적은 수의 파라미터(Parameter)를 사용함으로 VDSR 대비 짧은 시간에 초해상도 영상 생성이 가능하다.

본 논문의 구성은 다음과 같다. 2 장에서는 본 논문에서 제안하는 방법을 설명한다. 3 장에서는 실험 결과 및 분석, 4 장에서는 결론을 맺는다.

2. Cross-CNN

2.1 네트워크 구조

합성 곱(Convolution) 연산의 2 차원 필터는 분리 가능할 때, 두 개의 1 차원 필터로 표현 가능하며 딥 러닝 모델은 활성화 함수를 통해 비 선형적인 관계를 학습함으로 복잡한 데이터를 처리할 수 있다. 본 논문에서는 $n \times 1$ 및 $1 \times n$ 크기를 갖는 1 차원 필터 각각에 활성화 함수를 추가함으로 VDSR 의 계산 복잡도를 개선한다. 그림 1 은 본 논문에서 제안하는 Cross-CNN 의 네트워크 구조를 나타낸다. 네트워크 구조에서 첫 번째 와 마지막 층은 2 차원 필터를 이용하며 각각의 중간 층들은 하나의 1 차원 필터를 이용한다. 중간 층은 수직 방향 필터와 수평 방향의 필터의 모양이 교차하는 형태로 구성이 되어 있으며 각 1 차원 필터의 결과에 활성화 함수를 사용함으로 비선형성을 추가하였다. 제안하는 방법은 각각의 1 차원 필터 결과에 활성화 함수를 적용함으로 단순 2 차원

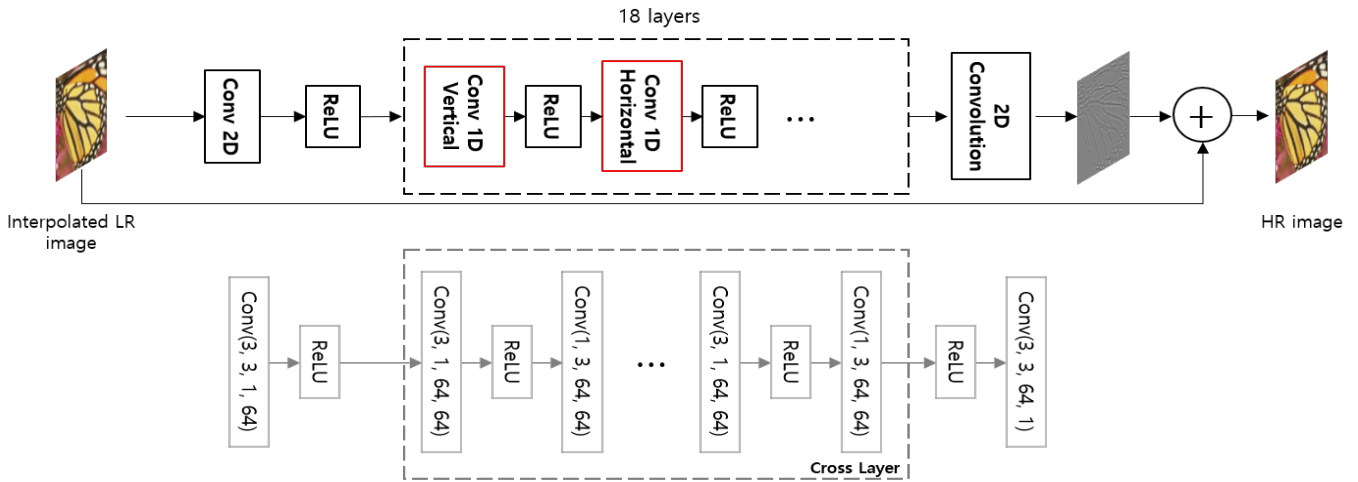


그림 1. 제안하는 Cross-CNN 구조

필터를 분리 가능 필터를 사용하는 방법 대비 더 적은 가중치 및 계산 량을 필요로 한다.

2.2 학습 방법

제안하는 Separable-CNN 구조는 학습을 위하여 MSE(Mean Squared Error)를 사용하였으며 최적화 함수로 Adam Optimizer[6]를 사용하였다. 학습 율(learning rate)은 초기 값을 설정한 후 지수적 감소(exponential decay)를 이용한 학습 율 감소(learning rate decay) 방법을 이용하였다.

2.3 학습 데이터 생성

본 논문은 [3, 4]과 같은 방법으로 학습 데이터를 생성하였다. 그 과정은 입력 영상에 대하여 다운 샘플링(down sampling)을 적용하여 저해상도 영상을 생성한 후 이를 다시 원 영상 크기로 업 샘플링(up sampling)한다. 이 과정의 입력으로 사용된 영상(original image)는 학습을 위한 정답으로 사용하며 과정을 통해 생성된 영상(interpolated image)는 학습의 입력 영상으로 사용하게 된다. 다운 샘플링 과정은 입력 영상의 크기가 ($W \times H$) 일 때, bicubic 보간을 적용하여 ($\frac{W}{scale\ factor} \times \frac{H}{scale\ factor}$) 크기의 저해상도 영상을 생성하며 업 샘플링 과정에서는 저해상도에 대하여 bicubic 보간을 적용하여 ($W \times H$) 크기의 보간 된 저해상도 영상(interpolated low resolution image)를 생성한다.

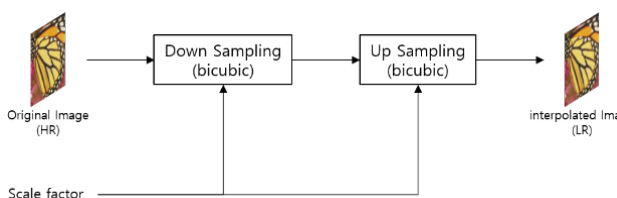


그림 2. 학습 데이터 생성 과정

3. 실험 결과

본 논문은 제안하는 방법의 실험을 위하여 TensorFlow[7]를 이용하여 구현하였으며 GPU 는 GTX1060 을 사용하였다. 학습 영상은 VDSR[4]의 291 개의 영상을 사용하였으며 테스트 영상은 B100 을 사용하였다. 학습 및 테스트는 스케일 팩터(scale factor) 2 에서 진행하였다. 본 논문에서는 기존 방법과 비교를 위하여 가중치 개수에 기반한 계산 복잡도 비교, 100 개의 영상에 대한 예측 속도 및 PSNR 성능을 측정하였다. 표 1 은 기존 방법과 제안하는 방법의 가중치 개수에 기반한 계산 복잡도이다. 제안하는 방법의 첫번째와 마지막 층은 VDSR 과 같은 크기의 필터를 사용하였으며 중간층은 1 차원 필터를 사용함으로 제안하는 방법은 기존 방법 대비 약 2/3 의 가중치를 감소하였다. 표 2 는 기존 방법과 제안하는 방법의 평균 PSNR 성능 비교이며, 표 3 은 기존 방법과 제안하는 방법의 평균 실행 속도 비교이다. 실험 결과를 통해 제안하는 방법은 VDSR 대비 더 적은 가중치를 사용하여 기존 방법 대비 20%의 속도 향상이 있음을 확인 하였다.

4. 결론

본 논문은 초해상도 영상 생성을 위한 Cross-CNN 구조를 제안하였다. 제안하는 구조는 2 차원 필터의 분리성과 활성화 함수의 비선형성을 바탕으로 기존 VDSR 구조의 시작과 마지막 층을 제외한 중간 층들에 교차하는 1 차원 필터를 적용하여 VDSR 대비 1/3 의 가중치를 사용함으로 계산 복잡도를 개선하였다. 제안하는 Cross-CNN 구조는 스케일 팩터 2 의 초해상도 영상 생성 실험에서 기존 방법 대비 약 20%의 속도 향상을 보였다.

표 1. VDSR 과 제안하는 방법의 계산 복잡도 비교

Layer Index	VDSR	Ours
First	576	576
Intermediate	663,552	221,184
Last	576	576

표 2. VDSR 과 제안하는 방법의 PSNR 성능 비교

VDSR	CCNN
31.90	30.11

표 3. VDSR 과 제안하는 방법의 평균 실행 속도 비교

VDSR	CCNN
0.10	0.08

감사의 글

"본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학 ICT 연구센터육성지원사업의 연구결과로 수행되었음" (IITP-2018-2016-0-00288)

참고 문헌

- [1] S. C. Park, M. K. Park, M. G. Kang, "Super-resolution image reconstruction: a technical overview", in IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 21–36, May. 2003
- [2] S. Ma, L. Zhang, X. Zhang and W. Gao, "Block Adaptive Super Resolution Video Coding", in Pacific-Rim Conference on Multimedia, Springer, Berlin, Heidelberg, pp. 1048–1057, 2009.
- [3] C. Dong, C. C. Loy, K. He and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks", in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295–307, Feb. 1 2016.
- [4] J. Kim, J. K. Lee and K. M. Lee, "Accurate Image Super-Resolution Using Very Deep Convolutional Networks", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Las Vegas, NV, pp. 1646–1654, 2016.
- [5] N. Yan, D. Liu, H. Li, B. Li, L. Li and F. Wu, "Convolutional Neural Network-Based Fractional-Pixel Motion Compensation," in IEEE Transactions on Circuits and Systems for Video Technology.
- [6] 박장수, 이종석, 박시내, 심동규, "Unified-CNNIF 구조를 이용한 영상 보간", 영상처리 및 이해에 관한 워크샵, 2018.
- [7] D. Kingma, B. Jimmy, "Adam: A method for stochastic optimizer", arXiv preprint, arXiv:1412.6980, 2014.
- [8] M. Abadi, et al. "Tensorflow: Large-scale machine learning on heterogeneous distributed systems", arXiv preprint, arXiv:1603.04467, 2016.