

심층 신경망 검색 기법을 통한 이미지 고해상도화

안준영, *조남익

서울대학교 뉴미디어 통신 공동 연구소(INMC)

sohwa360@ispl.snu.ac.kr, *nicho@snu.ac.kr

Image Super Resolution Using Neural Architecture Search

Joon Young Ahn *Nam Ik Cho

Seoul national university, Institute of New Media and Communications

요 약

본 논문에서는 심층 신경망 검색 방법을 사용하여 이미지 고해상도화를 위한 심층 신경망을 설계하는 방법을 구현하였다. 일반적으로 이미지 고해상도화, 잡음 제거 및 번짐 제거를 위한 심층신경망 구조는 사람이 설계하였다. 최근에는 이미지 분류 등 다른 영상처리 기법에서 사용하는 심층 신경망 구조를 검색하기 위한 방법이 연구되었다. 본 논문에서는 강화학습을 사용하여 이미지 고해상도화를 위한 심층 신경망 구조를 검색하는 방법을 제안하였다. 제안된 방법은 policy gradient 방법의 일종인 REINFORCE 알고리즘을 사용하여 심층 신경망 구조를 출력하여 주는 제어용 RNN(recurrent neural network)을 학습하고, 최종적으로 이미지 고해상도화를 잘 실현할 수 있는 심층 신경망 구조를 검색하여 설계하였다. 제안된 심층 신경망 구조를 사용하여 이미지 고해상도화를 구현하였고, 약 36.54dB 의 피크 신호 대비 잡음 비율(PSNR)을 가지는 것을 확인할 수 있었다.

1. 서론

이미지 고해상도화는 figure1 과 같이 주어진 이미지를 더 높은 해상도의 동일한 이미지로 변환하는 기술로써 영상처리, 패턴 인식, 의료영상 분석 등 다양한 분야에서 더 좋은 성능을 내기 위해 필요로 하는 중요한 기술이다. 최근 들어 심층 신경망 학습방법의 발전으로 기존의 이미지 고해상도화 방법에 비해 성능이 뛰어난 방법들이 많이 제안되었다[1, 2, 3].

지금까지의 심층 신경망 기반 이미지 고해상도화 방법들은 우선 사람이 학습할 심층 신경망 구조를 설계한 후, 해당 신경망을 학습하여 성능을 측정하는 방식을 취했다. 그러나 심층 신경망을 사람이 설계하기 위해서는 어떤 연산을 수행할지, 연산 결과를 몇 층으로 쌓을지, 어떤 연산끼리 연결할 것인지 등 많은 사전지식이 필요하며, 설계한 신경망이 최적의 성능을 내리라는 보장 또한 할 수 없다. 반면에 심층 신경망 검색 기법을 사용할 경우, 사람이 신경망 구조를 직접 설계하지 않기 때문에 전문 지식이 필요하지 않고, 설계되는 구조 또한 최적의 성능을 보장할 수 있다.

본 논문에서는 주로 이미지 분류 분야에서 많이 연구되어 온 심층 신경망 검색 기법을 응용하여 이미지 고해상도화에 적용할 수 있는 심층 신경망 검색 기법을 제안한다. 구체적으로 figure 2 와 같이 제어 RNN 에서 심층 신경망 구조를 예측하면, 예측된 심층 신경망 구조를 구현하여 성능을 측정한 후, 해당 성능을 보상 값으로 하여 제어 RNN 을 학습하고 더 좋은 심층 신경망 구조를 예측하도록 하는 방법을 제안한다. 또한 이미지 분류와 다른 특성을 가지는 이미지 고해상도화에 신경망 검색 기법을 적용하기 위해 dilated convolution 등 새로운 연산을

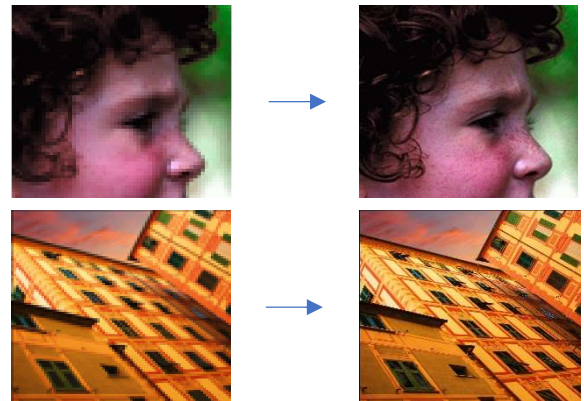


Figure 1. 이미지 고해상도화

검색 후보에 추가하였고, 필요 없는 연산을 제거하였다.

본 논문의 구성은 다음과 같다. 2 절에서는 이미지 고해상도화 및 심층 신경망 검색 기법에 대해 살펴본 후, 3 절에서는 본 논문에서 제안하는 이미지 고해상도화를 위한 심층 신경망 검색 기법을 설명한다. 4 절에서는 제안한 검색 기법의 성능을 실험을 통해서 확인하고 마지막으로 5 절에서는 본 논문에 대한 결론을 맺는다.

2. 관련 연구

2.1 이미지 고해상도화

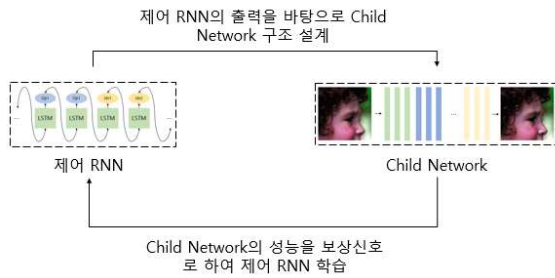


Figure 2. 강화학습을 이용한 신경망 검색 기법

심층 신경망 학습 기법은 이미지 고해상도화를 포함한 다양한 이미지 보정 및 영상처리 분야에서 좋은 성능을 나타냈다[1, 2, 3, 4, 5, 6, 7]. 일반적으로 이미지 고해상도화 분야에서는 저해상도의 이미지를 입력으로 받아 심층 신경망을 통과한 후 고해상도의 이미지를 직접 출력하는 방식이 많이 쓰였다. Dong et al. [1]은 3 개의 층을 쌓은 심층 신경망을 처음 설계하여 이미지 고해상도화를 구현하였고, 좋은 성능을 보였다. Kim et al. [2]은 16 개의 층을 가지고 residual 학습 방법을 사용하였다. Lim et al. [6]은 더 깊고 넓은 residual 블록들을 쌓아서 성능을 향상 시켰다. RDN[7]은 residual 블록 대신 더 중간 연결이 빈번한 dense 블록을 사용하였고, 과 RCAN[3]은 채널간 관심도(attention)를 예측하여 이미지 고해상도화에 사용하였다.

2.2 심층 신경망 검색

기존의 신경망 검색 방법들은 mutation 이나 crossover 등을 사용하여 신경망 구조를 변화시켜 좋은 구조는 후보로 남겨두고 나쁜 구조는 버리는 진화 알고리즘으로 최적화된 구조를 찾았다[8, 9]. 최근에는 이미지 분류 분야에서 강화학습을 신경망 검색에 사용하는 방법들이 제안되었다[10, 11]. Zoph et al. [10]에서는 ‘controller’ 라고 명명된 RNN 을 사용하여 학습하고자 하는 신경망 구조를 출력하였다. 출력된 신경망을 학습하여 성능을 계산하고, 이 성능을 보상신호로 간주하여 controller RNN 을 강화학습을 통해 최적화하였다. Zoph et al. [11]은 controller RNN 으로 신경망 구조 전체를 예측하는 것이 아니라 cell 이라고 불리는 신경망 단위 구조를 예측하고 이 cell 구조를 쌓아서 신경망 검색 성능을 높이고자 하였다.

하지만 이 방법들은 필요 연산 량이 높고 검색이 오래 걸리는 단점이 있어서 이를 보완하기 위한 효율적인 알고리즘들이 제안되었다[12, 13, 14]. PNAS[12]는 cell 의 복잡도를 높여 나가면서 이전 구조의 성능을 가지고 다음 구조의 성능을 예측하는 점진적인 방법을 사용하였다. ENAS[13]는 NAS[10]와 같은 구조로 controller RNN 을 사용하고 강화학습을 통해 신경망 구조 검색을 진행하였는데, 진행 시 기존에 학습한 파라미터들을 버리지 않고 다음 번 학습에 공유하여 사용함으로써 신경망 구조 검색 속도를 비약적으로 상승시켰다. DARTS[14]는 신경망 내부 파라미터들과 신경망 층간 연결을 모두 변수로 생각하여 직접 gradient 를 구해 최적화하여 신경망 검색을 효율적으로 진행하였다.

최종적으로, 본 논문에서는 주로 이미지 분류 영역에서 쓰이던 신경망 검색 방법 중 ENAS[13]를 이미지 고해상도화

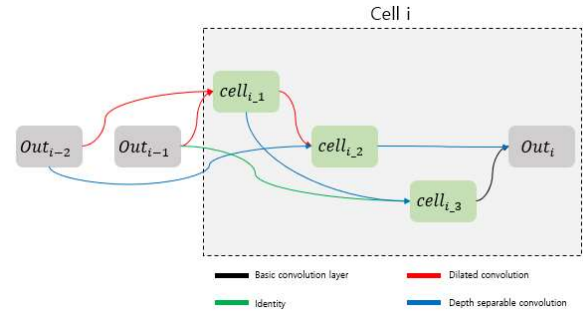


Figure 3. Cell 구조 설계 예시

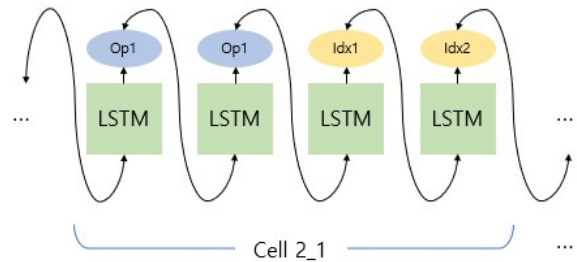


Figure 4. Child network 예측을 위한 제어 RNN 설계 분야로 확장하여 이미지 고해상도화 성능이 좋은 구조를 검색하는 방법을 제안하였다.

3. 실험 방법

본 논문에서 사용한 전체적인 방법은 ENAS[13]을 참조하여 구성하였다. Controller RNN 을 사용하여 이미지 고해상도화를 구현할 심층 신경망 후보를 추출한 후, 해당 신경망을 학습 후 그 성능을 보상신호로 하여 강화학습을 통해 Controller RNN 의 파라미터를 학습하였다. 이 때 학습 시 사용한 심층 신경망 파라미터를 제거하지 않고 다음 학습에 사용함으로써 효율적인 신경망 검색을 구현하였다.

3.1 검색 범위 구성 및 신경망 구조

본 논문에서 제안한 신경망 구조는 ENAS[13]와 같이 cell 로 구성하였다. 1 개의 Cell 구조는 5 개의 신경망 층을 쌓아서 만들었다. Cell 의 각 신경망 층은 figure 3 과 같이 해당 cell 이전의 두개 cell 의 출력과 동일 cell 의 아래층 신경망 중 두개를 입력으로 받을 수 있다. 또한 각 입력에 대해 신경망 층이 취할 수 있는 신경망의 종류는 다음과 같이 총 4 가지이다.

- 3*3 convolution layer
- 3*3 depth separable convolution layer
- 3*3 dilated convolution layer with rate 3
- Identity layer

각 층에 연결할 두개의 입력과 연산이 정해지면 두 결과를 합하여 해당 층의 결과로 만든다. 각 cell 의 최종 출력은 cell 내부 층 중에서 한번도 쓰이지 않은 층들을 합쳐서 만든다. 이에 따라 하나의 cell 구조를 만들기 위해 검색해야 할 경우의

Table 1. 이미지 고해상도화 성능 비교

Dataset	Scale	Bicubic	SRCNN	Ours
Set5	X2	33.66	36.66	36.54
	X3	30.39	32.75	32.32

수는 약 5.95×10^{10} 가지이다. 각 Cell1의 연결과 사용할 연산은 controller RNN의 출력을 통해 선택하게 된다. Controller RNN은 2층의 LSTM으로 이루어져 있으며, figure 4와 같이 4개의

출력으로 cell의 한 층을 예측하게 되고, cell 내부 신경망 층이 5개 이므로 총 20개의 출력을 통해 cell 구조를 예측한다.

3.2 학습 방법

본 논문이 제안한 신경망 검색 방법의 최종 목표는 주어진 저해상도 이미지와 고해상도 이미지에 대해 성능이 좋은 심층 신경망을 찾는 것이다. 즉, controller가 예측한 연결과 연산들을 $a_{1:T}$ 라고 하였을 때, controller의 출력으로 만든 신경망의 기대 성능 $J(\theta_c)$ 가 최대가 되도록 하는 것이고 이를 식으로 나타내면 다음과 같다

$$J(\theta_c) = E_{p(a_{1:T}, \theta_c)}[R] \quad \dots (1)$$

위 식에서 θ_c 는 controller의 파라미터이고 R은 보상 신호로써, 예측한 심층 신경망의 성능 (PSNR)이다. 이 때 보상 신호 R이 미분 불가능하므로, controller의 파라미터를 학습하기 위해서 ENAS[13]에서와 같이 REINFORCE[14] 알고리즘을 사용하였다. REINFORCE 알고리즘에 의해 식 1을 최적화하기 위한 미분 값을 구하면 아래와 같다

$$\begin{aligned} \nabla_{\theta_c} J(\theta_c) &= \sum_{t=1}^T E_{p(a_{1:T}, \theta_c)}[\nabla_{\theta_c} \log P(a_t | a_{(t-1:1)}; \theta_c) R] \\ &\approx \frac{1}{m} \sum_{k=1}^m \sum_{t=1}^T \nabla_{\theta_c} \log P(a_t | a_{(t-1:1)}; \theta_c) R_k \quad \dots (2) \end{aligned}$$

최종적으로, R_k 를 그냥 사용할 경우 분산이 커지므로 이를 줄이기 위해 R_k 의 지수이동평균(exponential moving average) b 를 뺀 값($R_k - b$)를 대신 사용하였다.

4. 실험 결과

실험은 DIV2K 학습 데이터를 64*64 크기로 잘라 예측된 신경망 학습에 사용하였고, DIV2K 검증 데이터를 보상신호 R 계산에 사용하였다. 총 200epoch의 학습을 진행하였고, 예측된 신경망을 1epoch 학습할 때 마다 controller RNN의 REINFORCE 알고리즘을 이용한 학습을 40번씩 진행하였다.

실험 결과 Set5 실험 데이터에서 2배 고해상도화 기준 36.54dB의 PSNR을 가지는 것을 확인할 수 있었다. Table 1에서 확인할 수 있듯 기존의 방법들과 비교할 만한 성능을 나타내고 있다. Figure 5는 실제로 신경망 검색 결과 찾아낸 cell 구조이고, figure 6은 Set5 데이터에서의 실제 고해상도화 결과이다.

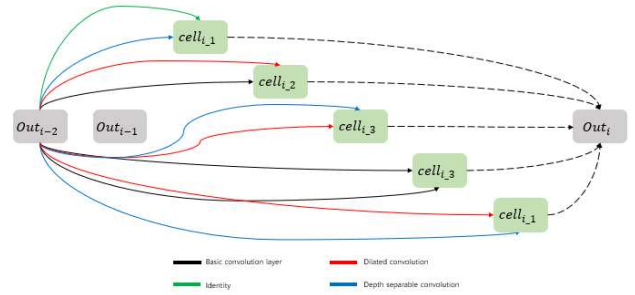


Figure 5. 고해상도화를 위한 신경망 검색 결과



Figure 6. 이미지 고해상도화 결과

5. 결론

본 논문에서는 신경망 검색 기법을 사용하여 이미지 고해상도화를 위한 신경망 구조를 찾는 방법에 대하여 제안하였다. 실험 결과 본 논문에서 제안한 신경망 검색 방법을 통해 기존의 신경망 성능과 비슷한 심층 신경망을 구성할 수 있음을 확인하였다. 추후 더 복잡한 신경망 검색 범위 및 구성을 통하여 더 성능이 뛰어난 신경망을 검색할 수 있도록 발전시킬 수 있을 것으로 보인다.

감사의 글

이 논문은 경찰청과 치안과학기술연구개발사업단의 지원을 받아 수행된 치안과학기술 연구개발사업임(PA-C000001).

참고문헌

[1] Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." *European conference on computer vision*. Springer, Cham, 2014.

[2] Kim, Jiwon, Jung Kwon Lee, and Kyoung Mu Lee. "Accurate image super-resolution using very deep convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[3] Zhang, Yulun, et al. "Image super-resolution using very deep residual channel attention networks." *Proceedings of the European Conference*

- on *Computer Vision (ECCV)*. 2018.
- [4] Liu, Wei, et al. "Ssd: Single shot multibox detector." *European conference on computer vision*. Springer, Cham, 2016.
- [5] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.
- [6] Lim, Bee, et al. "Enhanced deep residual networks for single image super-resolution." *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017.
- [7] Zhang, Yulun, et al. "Residual dense network for image super-resolution." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
- [8] Angeline, Peter J., Gregory M. Saunders, and Jordan B. Pollack. "An evolutionary algorithm that constructs recurrent neural networks." *IEEE transactions on Neural Networks* 5.1 (1994): 54–65.
- [9] Stanley, Kenneth O., and Risto Miikkulainen. "Evolving neural networks through augmenting topologies." *Evolutionary computation* 10.2 (2002): 99–127.
- [10] Zoph, Barret, and Quoc V. Le. "Neural architecture search with reinforcement learning." *International Conference on Learning Representations*. 2017.
- [11] Zoph, Barret, et al. "Learning transferable architectures for scalable image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
- [12] Liu, Chenxi, et al. "Progressive neural architecture search." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
- [13] Pham, Hieu, et al. "Efficient Neural Architecture Search via Parameter Sharing." *International Conference on Machine Learning*. 2018.
- [14] Williams, Ronald J. "Simple statistical gradient-following algorithms for connectionist reinforcement learning." *Machine learning* 8.3–4 (1992): 229–256..