

특집논문 (Special Paper)

방송공학회논문지 제27권 제1호, 2022년 1월 (JBE Vol.27, No.1, January 2022)

<https://doi.org/10.5909/JBE.2022.27.1.3>

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

HD 해상도에서 실시간 구동이 가능한 딥러닝 기반 블러 제거 알고리즘

심 규 진^{a)}, 고 강 옥^{a)}, 윤 성 준^{a)}, 하 남 구^{b)}, 이 민 석^{b)}, 장 현 성^{b)}, 권 구 용^{b)}, 김 은 준^{c)}, 김 창 익^{a)†}

A Deep Learning-based Real-time Deblurring Algorithm on HD Resolution

Kyujin Shim^{a)}, Kangwook Ko^{a)}, Sungjoon Yoon^{a)}, Namkoo Ha^{b)}, Minseok Lee^{b)},
Hyunsung Jang^{b)}, Kuyong Kwon^{b)}, Eunjoon Kim^{c)}, and Changick Kim^{a)†}

요 약

영상 블러 제거(deblurring)는 피사체의 움직임, 카메라의 흔들림, 초점의 흐림 등으로 인해 촬영 도중 발생한 영상 블러(blur)를 제거하는 것을 목표로 한다. 최근 스마트폰이 보급되며 휴대용 디지털카메라를 들고 다니는 것이 일상인 시대가 오면서 영상 블러 제거 기술은 그 필요성을 점점 더해가고 있다. 기존의 영상 블러 제거 기술들은 전통적인 최적화 기법을 활용하여 연구되어 오다가 최근에는 딥러닝이 주목받으며 합성곱 신경망 기반의 블러 제거 방법들이 활발하게 제안되고 있다. 하지만 많은 방법들이 성능에 먼저 초점을 맞추어 개발되어 알고리즘의 속도로 인하여 현실에서 실시간 활용이 어렵다는 문제점을 안고 있다. 이를 해결하고자 본 논문에서는 여러 신경망 설계 기법을 활용하여 HD 영상에서도 30 FPS 이상의 실시간 구동이 가능한 딥러닝 기반 블러 제거 알고리즘을 설계하여 이를 제안한다. 또한 학습 및 추론 과정을 개선하여 속도에 별다른 영향 없이 신경망의 성능을 높이고 동시에 성능에 별다른 영향 없이 신경망의 속도를 높였다. 이를 통해 최종적으로 1280x720 해상도에서 초당 33.74장의 프레임을 처리하며 실시간 동작이 가능함을 보여주었고 GoPro 데이터 세트를 기준으로 PSNR 29.79, SSIM 0.9287의 속도 대비 우수한 성능을 보여주었다.

Abstract

Image deblurring aims to remove image blur, which can be generated while shooting the pictures by the movement of objects, camera shake, blurring of focus, and so forth. With the rise in popularity of smartphones, it is common to carry portable digital cameras daily, so image deblurring techniques have become more significant recently. Originally, image deblurring techniques have been studied using traditional optimization techniques. Then with the recent attention on deep learning, deblurring methods based on convolutional neural networks have been actively proposed. However, most of them have been developed while focusing on better performance. Therefore, it is not easy to use in real situations due to the speed of their algorithms. To tackle this problem, we propose a novel deep learning-based deblurring algorithm that can be operated in real-time on HD resolution. In addition, we improved the training and inference process and could increase the performance of our model without any significant effect on the speed and the speed without any significant effect on the performance. As a result, our algorithm achieves real-time performance by processing 33.74 frames per second at 1280x720 resolution. Furthermore, it shows excellent performance compared to its speed with a PSNR of 29.78 and SSIM of 0.9287 with the GoPro dataset.

Keyword : Deblurring, Real-time, HD resolution, Deep-learning, Neural Network

Copyright © 2022 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

1. 서론

영상 블러 제거(deblurring)는 촬영 도중 카메라의 흔들림, 피사체의 움직임, 초점의 흐림 등으로 인해 발생한 영상 블러(blur)를 제거하는 기술을 말한다. 최근 스마트폰과 같은 휴대용 디지털카메라가 대중화됨에 따라 일상에서 영상을 촬영하는 일이 흔해지고 동시에 블러한 영상이 촬영되는 일 또한 빈번해지면서 영상 블러 제거 기술은 그 필요성을 더해가고 있다. 특히 재현할 수 없는 특정 순간을 촬영하였을 때 영상 블러가 발생한 경우에는 이를 제거하여 선명한 영상으로 만들 수 있는 블러 제거 기술이 필수라고 할 수 있다.

이러한 영상 블러 제거 기술은 전통적인 최적화 기법을 활용하여 활발하게 연구되었다. 예를 들어, Pan 등^[1]은 밝기와 경사도(gradient) 기반의 단순하면서도 효과적인 L0 정규화된 prior와 새로운 최적화 기법을 제안하며 이를 통해 복잡한 장면, 어두운 조명, 비균일한 블러 현상을 포함한 자연 영상에서도 효과적으로 블러를 제거할 수 있음을 보였다. 또한 Kim과 Lee^[2]는 강인한 total variation (TV)-L1 모델을 기반으로 움직임 흐름(motion flow)과 잠상(latent image)을 동시에 예측하는 새로운 에너지 모델을 설계하여 블러 제거를 수행하였다. 하지만 이러한 전통적인 최적화 기법을 사용하는 블러 제거 기술들은 수행 시간이 매우 느리다는 단점이 있다.

최근에는 합성곱 신경망(convolutional neural network, CNN)을 다양하게 활용하여 영상의 블러를 제거하고자 하는 연구들이 활발히 이루어지고 있다. Sun 등^[3]은 합성곱 신경망을 통해 각 영상 패치의 움직임 커널을 예측하고 이를 마르코프 랜덤 필드(markov random field) 모델로 전체 영상에 대한 움직임 블러 필드로 통합하여 블러를 제거하는 알고리즘을 제안하였다. 또한 Noroozi 등^[4]은 지름길 연

결을 통해 더욱 효과적으로 블러를 제거할 수 있는 신경망을 설계하였고, Nah 등^[5]은 총 120개의 합성곱 층을 사용하여 매우 깊은 잔여 신경망(residual network)을 사용하는 블러 제거 신경망을 발표하였다. 이에 더해 순환 신경망(recurrent network)을 통해 동영상의 특성을 잘 활용하여 블러를 제거할 수 있는 알고리즘들^[6,7,8,9] 또한 제안되었다. 하지만 이러한 합성곱 신경망 기반의 알고리즘들은 성능은 뛰어나지만, HD 해상도에서 30 FPS 수준의 실시간 구동이 대체로 불가능하여 현실적인 활용이 어렵다는 문제들이 꾸준히 제기되었다.

이를 해결하기 위해 최근에는 속도를 고려한 블러 제거 알고리즘들 또한 제안되고 있다^[10,11]. 특히 DMPHN^[11]은 조세 단계(Coarse-to-Fine) 방식의 계층적 표현을 통해 블러 제거 수행하고 다중 패치 신경망을 적층하여 신경망의 심층적 계층적 깊이를 늘림으로써 좋은 성능을 달성하면서도 기존의 다중 스케일(multi-scale) 기반의 블러 제거 알고리즘들에 비해 더욱 빠른 속도를 보여주었다. 하지만 실험 결과의 표 4에서와 같이 충분한 성능과 동시에 HD 해상도에서 30 FPS 이상의 실시간 구동을 보여주는 부분에서는 아직 부족함이 있는 것을 확인할 수 있다.

본 논문에서는 HD 영상에서도 30 FPS 이상의 실시간 구동이 가능하면서도 이전 방법들에 비해 더욱 뛰어난 속도 대비 성능을 보여줄 수 있는 딥러닝 기반 블러 제거 알고리즘을 제안한다. 그림 1에서 볼 수 있듯이 제안된 신경망은 인코더와 디코더 구조를 기본으로 입력 영상을 네트워크의 출력에 더해주는 전역 지름길 연결(global shortcut connection)^[12], 신경망의 마지막 부분에서 디코더의 특징들을 모아주는 다중 스케일 특징 집합(multi-scale feature aggregation, MSFA), 픽셀 셔플(pixel shuffle)^[13]을 통한 특징 업샘플링 등을 적용하여 더 적은 계산량으로도 효과적인 블러 제거가 가능하도록 설계하였다. 또한, 학습 과정에서도 워업(warmup) 훈련^[12,14], 코사인 학습률 감소(cosine learning rate decay) 함수^[15,16], L1 손실 함수를 활용하여 기존에 많이 사용되는 일반적인 학습 프로세스에 비해 속도에 아무런 영향 없이도 블러 제거의 성능을 더욱 개선하였다. 마지막으로 추론 단계에서 신경망의 기본 자료형을 32-bit에서 16-bit 부동소수점으로 변환해 성능의 저하 없이

a) 한국과학기술원 전기 및 전자공학과(KAIST EE)

b) LIG 넥스원 전자광학연구소(LIG Nex1 EO/IR R&D)

c) 국방과학연구소(ADD)

‡ Corresponding Author : 김창익(Changick Kim)

E-mail: changick@kaist.ac.kr

Tel: +82-42-350-7421

ORCID:https://orcid.org/0000-0001-9323-8488

※ This work was supported by LIG Nex1.Co.,Ltd, originally funded by DAPA and ADD(UC190031FD).

· Manuscript received November 10, 2021; Revised December 21, 2021; Accepted December 21, 2021.

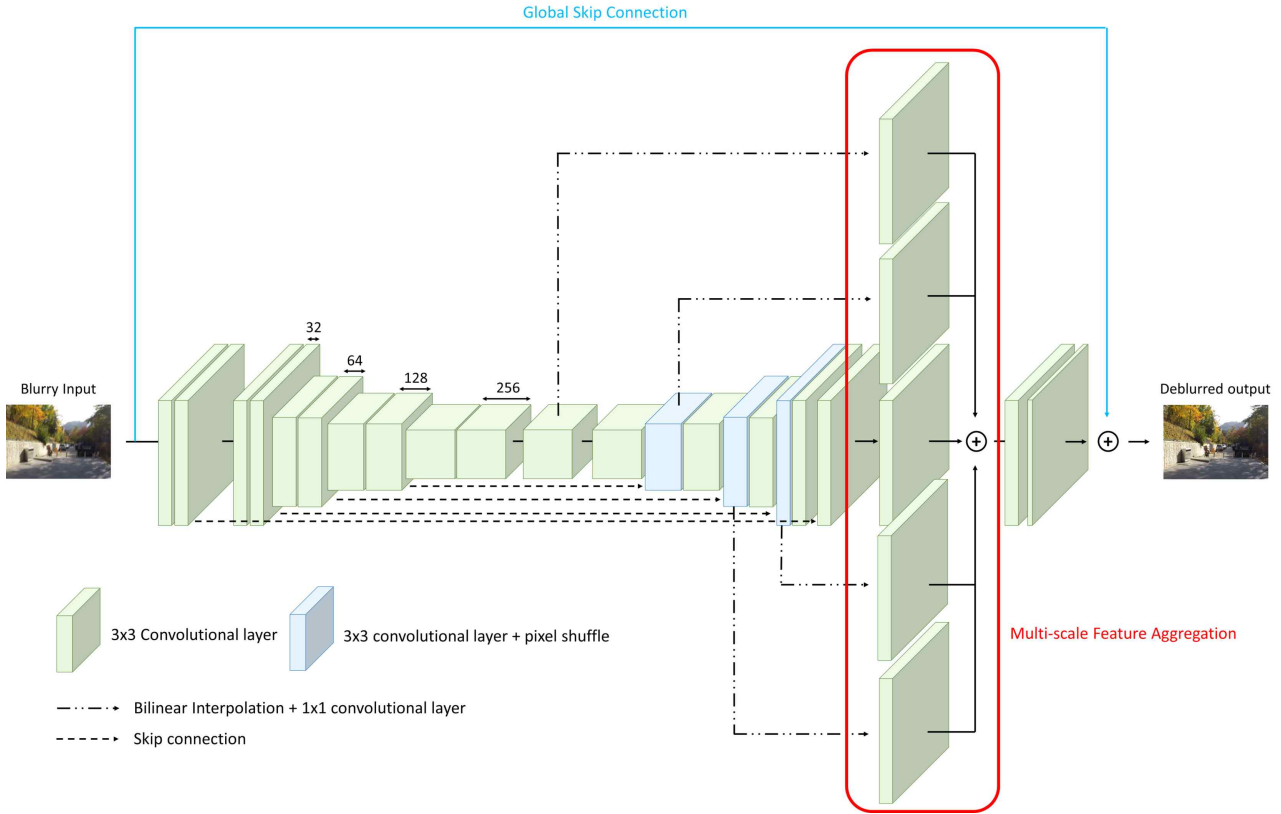


그림 1. 제안하는 블러 제거 신경망의 전체 구조도. 각 초록색 블록은 하나의 합성곱 층을 뜻하고, 각 파란색 블록은 하나의 합성곱 층과 픽셀 셔플이 결합한 것을 뜻한다.

Fig. 1. The overall structure of the proposed deblurring network. Each green block represents a convolutional layer, and each blue block represents a combination of convolutional layer and pixel shuffle.

알고리즘을 더욱 가속하였다.

이후 본문에서는 우리가 제안하는 알고리즘에 대해 구체적으로 설명하고 실험에서는 데이터 세트, 실험 세부 설정, 실험 결과에 관해 설명한다. 마지막으로 결론에서 논의의 요지를 정리한다.

II. 본 론

1. 신경망 설계

그림 1에서 볼 수 있듯이 우리의 신경망은 인코더와 디코더 구조를 기본으로 설계되었다. 인코더는 점진적으로 특징 맵을 축소하며 다양한 크기의 특징을 추출한다. 이때 더 넓은 수용영역(receptive field)을 위해 최대 풀링(max pool-

ing) 혹은 평균 풀링(average pooling) 대신 보폭(stride) 2의 3x3 합성곱을 통해 특징 맵을 축소하며 그렇게 특징 맵의 크기가 반으로 축소될 때마다 32에서 시작하여 두 배씩 특징 맵의 채널 수를 늘려준다. 반대로 디코더에서는 3x3 합성곱과 픽셀 셔플을 통해 특징 맵의 크기가 기존의 입력 영상과 같아지도록 확장하며 그렇게 특징 맵의 크기가 두 배로 늘어날 때마다 256에서 시작하여 절반씩 채널 수를 줄여준다. 이때 픽셀 셔플은 초해상화 분야에서 영상 복원에 매우 효과적이라는 것이 보고되었었고^[13] 본 연구에서도 특징 맵의 업샘플링을 위해 기존에 많이 사용되는 이중 선형 보간법 혹은 전치 합성곱(transposed convolution) 대신 3x3 합성곱과 픽셀 셔플을 활용하였다. 우리의 신경망은 총 26개의 합성곱 층으로 이루어져 있고 제일 마지막을 제외한 모든 합성곱 층 다음에는 Leaky ReLU^[17] 활성화 함수를

배치하였다. 이때 배치 정규화의 경우 픽셀 단계의 상관(pixel-level correlation)을 해치는 경향이 있어^[10] 최근 합성곱 신경망 기반의 영상 처리 알고리즘들에서 점점 사용되지 않고 있기 때문에 마찬가지로 우리의 방법에서도 사용하지 않았다.

추가로, 그림 1에서와 같이 인코더와 디코더 사이에는 지름길 연결이 적용되어 있고 입력 영상과 신경망의 출력 사이에는 전역 지름길이 연결되어 있다. 지름길 연결은 심층 신경망 내부 합성곱 층들 사이에서의 나머지 학습(residual learning)을 통해 신경망 앞에 자리한 매개변수까지 경사도 흐름(gradient flow)을 원활히 전달하여 신경망 훈련이 더욱 효과적으로 진행되도록 도와준다^[12]. 또한 디코더에서 인코더의 정보를 활용할 수 있게 하여 결과적으로 전체 신경망이 더욱 효과적으로 블러 제거를 수행할 수 있도록 도와준다. 전역 지름길 연결은 나머지 학습을 통해 신경망의 부담을 줄여 잡음 제거, 초해상화 등의 영상 처리 알고리즘들에서 큰 효과를 보여주었고^[18,19], 특히 속도에 별다른 영향 없이도 성능을 높일 수 있다는 점에서 큰 장점이 있다. 본 연구에서도 이를 신경망의 설계에 포함하여 그 효과를 확인하였다.

마지막으로 디코더의 끝부분에는 다양한 크기의 디코더의 특징 맵들을 병합하는 다중 스케일 특징 집합을 적용하였다. 구체적으로는 각 디코더의 특징 맵들을 입력 영상의 크기와 동일하게 이중 선형 보간법을 통해 확대하고 1x1 합성곱을 사용하여 적절한 특징만을 남기며 채널 수를 일치시켜 모두 더해준다. 이때 특징 맵의 확대 배율이 최대 8배인 경우도 있어 신경망의 속도 향상을 위해 전치 합성곱 혹은 픽셀 셔플 대신 이중 선형 보간법을 사용하였다. 이러한 다중 스케일 특징 집합은 디코더의 다양한 크기의 특징들로부터 모든 정보를 모아 더욱 효과적으로 블러 제거를 수행할 수 있게 한다.

우리의 신경망은 기존의 다른 블러 제거 방법들과는 달리 굉장히 깊은 신경망^[5]을 사용하거나 여러 장의 연속된 영상을 함께 사용하는 순환 신경망^[6,7,8,9]을 사용하지 않고 상대적으로 얇은 신경망을 통해 단일 입력 영상을 처리한다. 또한 다양한 스케일의 합성곱^[10], 확장된 합성곱(dilated convolution)^[10], 입력 영상을 여러 패치로 나누어 각각 처리하고 그 과정에서의 정보들을 활용하는 다중 패치 신경

망^[10,11]을 사용하는 기존의 방법들과는 달리 3x3 합성곱만을 활용하여 전체 영상을 한 번에 처리한다. 이를 통해 HD 해상도에서도 30 FPS 이상의 속도를 달성하였고 동시에 위에서 설명한 여러 가지 요소들을 포함한 신경망의 설계를 통해 속도 대비 충분한 성능을 보여주었다.

2. 신경망 학습

신경망의 학습은 L1 손실 함수를 통해 워업 훈련과 코사인 학습률 감소 전략을 적용한 본 훈련으로 나누어 진행하였다. 이때 손실 함수는 신경망의 훈련에 있어 굉장히 중요한 요소 중 하나로 딥러닝 기반의 블러 제거, 초해상화 등의 영상 처리 알고리즘에서는 L1, L2, SSIM 손실 함수 등이 많이 사용된다. 그중에서도 L2 손실 함수나 SSIM 손실 함수보다 L1 손실 함수를 사용하였을 때 더 성능이 높게 나온다는 것이 보고된 바 있어^[20] 본 연구에서도 아래의 식과 같이 L1 손실 함수로 훈련을 진행하였다.

$$L = \frac{1}{N} \sum_i |f(I(i)) - I'(i)|, \quad (1)$$

이때, $f(\cdot)$ 는 블러 제거 신경망 모델, I 는 블러한 입력 영상, I' 는 정답 영상, i 는 영상의 픽셀 인덱스, 그리고 N 은 입력 및 정답 영상의 총 픽셀 수이다.

워업 훈련은 모든 매개변수가 임의의 값으로 시작하여 학습이 진행되는 합성곱 신경망의 특성상 처음부터 큰 학습률을 적용하면 수렴 과정이 불안정해지는 것을 방지한다. 본 연구에서는 초기에 작은 학습률을 사용하고 수렴 과정이 안정되면 초기 학습률로 다시 전환하는 워업 휴리스틱^[12]과 학습률을 0에서 초기 학습률까지 선형적으로 증가시키는 점진적 워업 전략^[14] 등의 다양한 방법 중에서 점진적 워업 전략을 적용하였다.

마지막으로, 학습률을 초기에 설정한 값으로부터 훈련을 진행하며 꾸준히 감소시키는 학습률 감소 전략 중에서 코사인 감소 전략^[15]이 학습률을 특정 에포크 수마다 정해진 비율로 감소시키는 스텝 감소 전략보다 더 나은 성능을 보여주는 것으로 알려져,^[16] 본 연구에서도 이를 적용하여 그 효과를 확인하였다. 이때 코사인 감소 전략은 총 배치(batch)의 크기를 N 이라고 가정하고 초기 학습률이 r_0 일 때

n 번째 배치에서의 학습률 r_n 을 다음의 식과 같이 계산한다.

$$r_n = \frac{1}{2} (1 + \cos(\frac{n\pi}{N})) r_0, \quad (2)$$

코사인 감소 전략은 훈련 초반에 학습률을 서서히 감소 시키다가 중반에 거의 선형적으로 감소시키고 마지막에 다시 서서히 감소시킨다. 즉 훈련의 후반부에서도 상대적으로 학습률을 크게 유지하여 훈련의 효과를 향상한다고 할 수 있다. 본 연구에서는 이러한 학습 방법들을 통해 기존에 많이 활용되던 학습 설정에 비해 속도에 아무런 영향 없이도 성능을 크게 개선할 수 있음을 확인하였다.

3. 추론 과정 개선

대부분의 인공 신경망 프레임워크는 32-bit 혹은 64-bit 부동소수점을 사용한다. 이때 인공 신경망의 훈련 혹은 추론 과정에서 16-bit 부동소수점을 사용하면 성능에 별다른 영향 없이도 메모리 사용량과 처리 속도 등에서 큰 이득을 볼 수 있다. 본 연구에서도 신경망 모델의 가속화를 위해 이를 적용해보고 성능 및 속도를 확인해보았다.

III. 실험

1. 데이터 세트

본 연구에서는 대표적인 공개 블러 제거 데이터 세트인 GoPro 데이터 세트^[5]에서 실험을 진행하였다. 이 데이터 세트는 GOPRO4 Hero Black 카메라를 통해 촬영한 1280x720 해상도의 240 fps 동영상에서 7 ~ 13장 사이의 주변 프레임을 평균하여 블러 영상을 제작하고 평균한 주변 프레임 중 중간 프레임을 정답 영상으로 골라 블러 - 정답 영상 페어를 구성하였다. 총 3,214개의 페어로 구성되어 있고 이 중 2,103개가 학습용으로, 나머지 1,111개가 테스트용으로 이루어져 있다.

2. 실험 세부 설정

신경망의 훈련은 기본적으로 300 에포크의 점진적 워업

훈련과 300 에포크의 본 훈련으로 구성되어 신경망의 출력과 그에 대응되는 정답 영상 간의 L1 손실을 통해 진행되었다. 워업 훈련 동안은 학습률을 매 에포크마다 10^{-6} 에서 10^{-4} 까지 선형적으로 증가시켰고 본 훈련 동안은 학습률을 10^{-4} 에서 0까지 코사인 감소시켰다. 배치 크기는 6으로 설정하였고 훈련 영상은 원본 영상에서 무작위로 256x256 패치로 잘라 생성하였다. 또한 각 패치를 무작위로 좌우 반전시키고 모든 픽셀의 값이 -0.5~0.5 범위 사이가 되도록 정규화하였다. 이때 전체 학습 과정에서 무작위 시드 (random seed)는 10000으로 고정하였다. 우리의 알고리즘은 Tensorflow 2.4 프레임워크를 통해 구현되었고 Intel i9-7920X CPU, 64GB RAM, NVIDIA RTX 3090 GPU의 환경에서 구동되었다. 추론 단계에서는 배치 크기를 1로 설정하여 1280x720 크기의 원본 영상 전체를 입력으로 넣어주었으며 FPS 또한 같은 조건으로 측정되었다. 이때 FPS는 다른 알고리즘들과의 공평한 비교를 위해 모두 PyTorch 프레임워크로 통일하여 측정하였다. 또한 성능 결과는 1,111장의 전체 테스트용 데이터 세트에 대한 평균으로 계산되었다.

3. 정량적 실험 결과

표 1은 신경망의 구조 중에서 다중 스케일 특징 집합, 전역 지름길 연결, 픽셀 셔플의 효과를 확인하기 위하여 각 방법을 하나씩 제외하거나 다른 일반적인 방법으로 대체하였을 때의 실험 결과를 보여준다. 표 1에서 볼 수 있듯이 다중 스케일 특징 집합과 전역 지름길 연결 방법을 제외하

표 1. 신경망 구조 설계 과정에서 각 방법을 하나씩 제외하거나 다른 일반적인 방법으로 대체하였을 때의 실험 결과. MSFA는 다중 스케일 특징 집합, global skip은 전역 지름길 연결을 뜻한다.

Table 1. Experimental results when each method is excluded one by one from the neural network structure design process or replaced with another method. "MSFA" stands for the multi-scale feature set, and "global skip" stands for global skip connection.

	MSFA	global skip	Up-sampling	PSNR	SSIM
Ours 1	O	O	pixel shuffle	29.79	0.9287
Ours 2	X	O	pixel shuffle	29.74	0.9281
Ours 3	X	X	pixel shuffle	29.66	0.9274
Ours 4	X	X	transposed convolution	29.50	0.9253
Ours 5	X	X	bilinear interpolation	29.44	0.9244

였을 때 유의미한 성능의 저하를 확인할 수 있고 이는 즉 다중 스케일 특징 집합과 전역 지름길 연결이 성능 개선에 효과적이라는 것을 보여준다. 픽셀 셔플 또한 이중 선형 보 간법과 3×3 전치 합성곱으로 대체하였을 때 마찬가지로 유의미한 성능 저하를 확인할 수 있었다.

표 2는 신경망의 학습 과정에서 워업 훈련, 코사인 학습률 감소 전략, L1 손실 함수의 효과를 확인하기 위하여 각 방법을 차례대로 제외하거나 다른 일반적인 방법으로 대체하며 블러 제거 성능을 확인한 결과를 보여준다. 비교 실험을 위하여 워업 훈련의 경우 단순히 제외하여 학습을 진행하였고 코사인 학습률 감소 전략은 총 3,000 에폭의 훈련 과정에서 각 1,000 에폭마다 0.1 비율로 줄여가는 스텝 감소 전략으로 대체하였으며 L1 손실 함수는 L2 손실 함수로 대체하였다. 표 2에서 볼 수 있듯이 각 방법을 제외하거나 다른 방법으로 대체할 때마다 상당한 성능의 저하를 확인할 수 있고 이는 우리가 활용한 방법들의 효과를 확인시켜준다.

표 2. 신경망 학습 과정에서 각 방법을 하나씩 제외하거나 다른 일반적인 방법으로 대체하였을 때의 실험 결과. 워업은 워업 훈련, cos는 코사인 학습률 감소 전략 L1은 L1 손실 함수를 뜻한다.

Table 2. Experimental results when each method is excluded one by one in the neural network learning process or replaced with another method. "Warm-up" stands for warm-up training, "cos" stands for cosine learning rate decay strategy, and "L1" stands for L1 loss function.

	워업	cos	L1	PSNR	SSIM
Ours	O	O	O	29.79	0.9287
Ours	X	O	O	29.74	0.9278
Ours	X	X	O	29.57	0.9257
Ours	X	X	X	29.38	0.9219

표 3은 신경망 추론 과정에서 기본 자료형으로 16-bit 부동소수점을 사용했을 때의 성능 변화와 속도 개선 효과를 보여준다. 표에서 볼 수 있듯이 별다른 성능의 변화 없이 알고리즘의 속도가 1.56배 개선된 것을 볼 수 있다. 표 4와 그림 2는 이전의 다른 블러 제거 알고리즘들과의 성능 비교를 보여준다. 이때 공정한 비교를 위해 다른 방법들 또한 16-bit 부동소수점을 기본 자료형으로 사용하여 속도를 측정하였다. 표와 그림에서 볼 수 있듯이 우리의 방법은 이전의 다른 방법들에 비해 속도는 더욱 빠르면서도 더 우수하거나 충분한 성능을 보여준다. 특히 성능이 가장 비슷한

DMPHN(1-2)^[10]과 비교했을 때 우리의 방법이 1.4배 정도 더욱 향상된 속도를 보여주는 것을 확인할 수 있다. 또한 STRCNN^[6] 등과 같이 속도도 더 느리지만 성능 또한 더 부족한 방법들이 존재하고 우리의 방법보다 성능이 더 우수한 방법들도 30 FPS에는 굉장히 못 미치는 것을 볼 수 있다. 따라서 우리의 방법이 실시간 구동이 필요한 블러 제거 애플리케이션에서 좋은 대안이 될 수 있음을 시사한다.

표 3. 신경망 추론 과정에서 기본 자료형을 32-bit 부동소수점과 16-bit 부동소수점을 사용하였을 때의 블러 제거 성능과 FPS

Table 3. Deblurring performance and FPS when 32-bit floating-point and 16-bit floating-point are used as basic data types in the neural network inference process

	PSNR	SSIM	FPS
Ours (FP32)	29.79	0.9287	21.68
Ours (FP16)	29.79	0.9287	33.74

표 4. 이전의 다른 블러 제거 방법들과의 성능 비교

Table 4. Performance comparison with other previous deblurring methods

	PSNR	SSIM	FPS
Kim and Lee ^[2]	23.64	0.8239	0.0008
Sun et al ^[3]	24.64	0.8429	0.0003
Nah et al ^[5]	30.40	0.9018	4.96
STRCNN ^[6]	28.74	0.8465	4.10
SRN ^[7]	30.10	0.9323	1.32
IFIRNN(c2h3) ^[8]	29.97	0.8859	3.23
DMPHN(1-2) ^[11]	29.77	0.9286	24.21
ESTRNN(B9C80) ^[9]	30.79	0.9016	4.82
Ours	29.79	0.9287	33.74

4. 질적 실험 결과

그림 3은 위에서부터 차례대로 입력 블러된 영상, Nah et al^[5], ESTRNN(B9C80)^[9], 우리의 방법으로 블러를 제거한 결과 영상들을 보여준다. 두 번째 영상의 숫자 부분에서는 ESTRNN이 어떤 숫자인지 더 알아보기 쉽게 영상을 복원한 것을 볼 수 있다. 하지만 그 외의 다른 영역과 영상에서는 우리의 방법이 시각적으로 더욱 탁월한 것을 확인할 수 있다. 특히 3번째 영상의 경우 우리의 방법이 Nah et al^[5], ESTRNN(B9C80)^[9]의 방법들 보다 사람의 다리를 훨씬 현실적으로 복원한 것을 볼 수 있다.

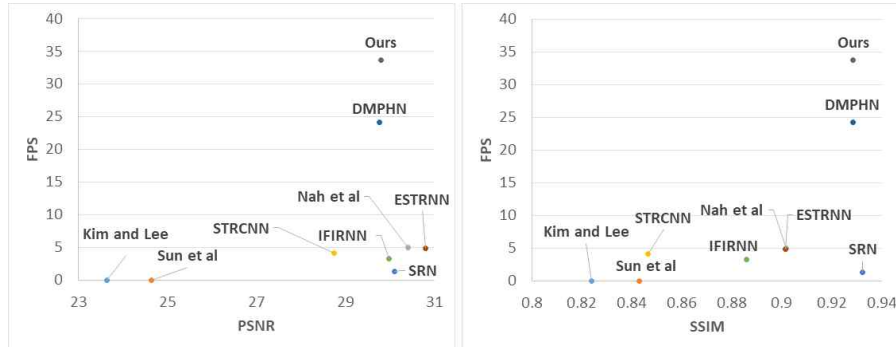


그림 2. 블러 제거 알고리즘의 성능 비교
Fig. 2. Performance comparison of deblurring algorithms



그림 3. 블러 제거 알고리즘의 비교. 위에서부터 차례대로 입력 블러된 영상, Nah et al^[6], ESTRNN(B9C80)^[9], 우리의 방법으로 블러를 제거한 영상
Fig. 3. Comparison of deblurring algorithms. Input blurred images, deblurred images with Nah et al^[6], ESTRNN(B9C80)^[9] and our method shown sequentially from the top

IV. 결 론

본 연구에서는 최근 스마트폰과 같은 휴대용 디지털카메라가 대중화되면서 일상에서 점점 더 필요성을 더해가고 있는 영상 블러 제거 기술에 대해 다루었다. 기존의 블러 제거 연구는 전통적인 최적화 기법을 활용하여 많이 진행되었고, 최근에는 합성곱 신경망 기반의 블러 제거 알고리즘들이 많이 제안되고 있다. 하지만 많은 방법들이 성능은 뛰어나지만 속도가 느려 실시간 활용이 어렵다는 문제점을 안고 있다. 이를 해결하고자 본 논문에서는 HD 영상에서도 실시간 구동이 가능한 딥러닝 기반 블러 제거 알고리즘을 제안하고 다양한 방법들을 활용하여 학습 및 추론 과정을 개선하였다. 최종적으로 제안된 방법은 1280x720 해상도에서 실시간 동작이 가능하고 GoPro 데이터 세트를 기준으로 PSNR 29.79, SSIM 0.9287의 속도 대비 우수한 성능을 보여주었다.

참 고 문 헌 (References)

- [1] J. Pan, Z. Hu, Z. Su, and M. Yang, "Deblurring text images via L0-regularized intensity and gradient prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 39, no. 2, pp. 342-355, Feb 2017.
- [2] T. Kim and K. Lee, "Segmentation-Free Dynamic Scene Deblurring," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, USA, pp. 2766-2773, 2014.
- [3] J. Sun, W. Cao, Z. Xu, and J. Ponce, "Learning a convolutional neural network for non-uniform motion blur removal," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 769-777, 2015.
- [4] M. Noroozi, P. Chandramouli, and P. Favaro, "Motion deblurring in the wild," *Proceedings of the 39th German Conference on Pattern Recognition (GCPR)*, Basel, Switzerland, pp. 65-77, 2017.
- [5] S. Nah, T. Kim, and K. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, Hawaii, USA, pp. 3883-3891, 2017.
- [6] T. Kim, K. Lee, B. Scholkopf, and M. Hirsch, "Online Video Deblurring via Dynamic Temporal Blending Network," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 4058-4067, 2017.
- [7] X. Tao, H. Gao, X. Shen, J. Wang, and J. Jia, "Scale-recurrent network for deep image deblurring," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, Utah, USA, pp. 8174-8182, 2018.
- [8] S. Nah, S. Son, and K. Lee, "Recurrent Neural Networks With Intra-Frame Iterations for Video Deblurring," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 8094-8103, 2019.
- [9] Z. Zhong, Y. Gao, Y. Zheng, and B. Zheng, "Efficient spatio-temporal recurrent neural network for video deblurring," *Proceedings of the 16th European Conference on Computer Vision (ECCV)*, pp. 191-207, 2020.
- [10] X. Hu, W. Ren, K. Yu, K. Zhang, X. Cao, W. Liu, and B. Menze, "Pyramid Architecture Search for Real-Time Image Deblurring," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4298-4307, 2021.
- [11] H. Zhang, Y. Dai, H. Li, and P. Koniusz, "Deep stacked hierarchical multi-patch network for image deblurring," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 5978-5986, 2019.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, Nevada, USA, pp. 770-778, 2016.
- [13] 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," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, Nevada, USA, pp. 1874-1883, 2016.
- [14] P. Goyal, P. Dollár, R. B. Girshick, P. Noordhuis, L. Wesolowski, A. Kyrola, A. Tulloch, Y. Jia, and K. He, "Accurate, large minibatch SGD: training imagenet in 1 hour," *CoRR*, abs/1706.02677, 2017.
- [15] I. Loshchilov and F. Hutter, "Sgdr: stochastic gradient descent with restarts," *CoRR*, abs/1608.03983, 2016.
- [16] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, and M. Li, "Bag of tricks for image classification with convolutional neural networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 558-567, 2019.
- [17] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," *Proceedings of the 30th International Conference on Machine Learning Workshop on Deep Learning for Audio, Speech and Language Processing*, Atlanta, GA, USA, 2013.
- [18] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing (TIP)*, vol. 26, no. 7, pp. 3142-3155, July 2017.
- [19] J. Kim, J. K. Lee, and K. Lee, "Accurate image super-resolution using very deep convolutional networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, Nevada, USA, pp. 1646-1654, 2016.
- [20] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for neural networks for image processing," *IEEE Transactions on Computational Imaging (TCI)*, vol. 3, no. 1, pp. 47-57, March 2016.

저 자 소 개



심 규 진

- 2017년 : 한국과학기술원 전기 및 전자공학부 학사 졸업
- 2017년 ~ 현재 : 한국과학기술원 전기 및 전자공학부 석박통합과정
- ORCID : <https://orcid.org/0000-0003-3015-8725>
- 주관심분야 : 영상 디블러, 다중 물체 추적



고 강 욱

- 2021년 : 한국과학기술원 전기 및 전자공학부 학사 졸업
- 2021년 ~ 현재 : 한국과학기술원 전기 및 전자공학부 석사과정
- ORCID : <https://orcid.org/0000-0002-5101-9751>
- 주관심분야 : 영상처리, 컴퓨터 비전



윤 성 준

- 2019년 : 한국과학기술원 수리과학과 학사
- 2019년 ~ 현재 : 한국과학기술원 전기 및 전자공학부 석사과정
- ORCID : <https://orcid.org/0000-0003-4410-2650>
- 주관심분야 : 컴퓨터 비전, 딥러닝



하 남 구

- 2008년 : 경북대학교 컴퓨터공학과 박사
- 2009년 ~ 현재 : LIG넥스원 수석연구원
- ORCID : <https://orcid.org/0000-0001-5146-4794>
- 주관심분야 : 감시정찰 시스템, 컴퓨터비전, 딥러닝



이 민 석

- 2018년 : 한국항공대학교 소프트웨어학과 학사
- 2020년 : 한국항공대학교 항공전자정보공학부 석사
- 2020년 ~ 현재 : LIG 넥스원 선임연구원
- ORCID : <https://orcid.org/0000-0002-7100-4354>
- 주관심분야 : 딥러닝, 영상처리, 전자광학시스템

저 자 소 개



장 현 성

- 2007년 : 부산대학교 컴퓨터공학과 학사
- 2009년 : 부산대학교 컴퓨터공학과 석사
- 2019년 ~ 현재 : 연세대학교 전기전자공학과 박사과정
- 2011년 ~ 현재 : LIG넥스원 수석연구원
- ORCID : <https://orcid.org/0000-0002-5797-7264>
- 주관심분야 : 컴퓨터비전, EdgeAI, 양자컴퓨팅



권 구 용

- 2014년 : 한국항공대학교 항공전자및정보통신공학부 학사
- 2014년 ~ 현재 : LIG넥스원 선임연구원
- ORCID : <https://orcid.org/0000-0002-3687-6771>
- 주관심분야 : 영상처리, 인공지능, 임베디드 시스템



김 은 준

- 2011년 : 한양대학교 미디어통신공학과 학사 졸업
- 2014년 : 한국과학기술원(KAIST) 전기 및 전자 공학과 석사 졸업
- 2014년 ~ 현재 : 국방과학연구소 선임연구원
- ORCID : <https://orcid.org/0000-0001-8668-448X>
- 주관심분야 : 영상개선, 영상추적



김 창 익

- 2000년 12월 : 워싱턴주립대학교 전기공학과 박사 졸업
- 2000년 12월 ~ 2005년 1월 : Epson Palo Alto Lab. 책임연구원
- 현재 : 한국과학기술원 전기 및 전자공학부 교수
- ORCID : <https://orcid.org/0000-0001-9323-8488>
- 주관심분야 : 딥러닝, 최적화, 컴퓨터 비전, 그래프 신호 처리, 그래프 CNNs