

이미지 디블러링을 위한 지역적 적응형 블러 주의 기반 신경망

이민영*, 이호섭**¹, 이하림**²

Local Adaptive Blur Attention-based Neural Network for Image Deblurring

Minyoung Lee*, Ho Sub Lee**¹, and Harim Lee**²

이 연구는 국립금오공과대학교 대학 연구과제비로 지원되었음(2023년~2024년)

요약

기존의 딥러닝 기반 이미지 디블러링 기법들은 블러가 다양한 크기와 방향으로 발생하는 특성을 충분히 고려하지 못하며, 블러의 공간적 연관성과 방향성을 효과적으로 학습하는 데 한계를 보였다. 이러한 문제를 해결하기 위해, 본 논문에서는 LABA-Net(Local Adaptive Blur Attention-Based Neural Network)을 제안한다. LABA-Net은 SSM(Spatial-wise Square Module)을 활용하여 다양한 방향성을 가진 블러의 공간 정보를 학습하고, 적응적으로 위치 간 상관관계를 포착한다. 또한, SSDO(Symmetrical Square Depth-wise Overparameterized Convolution)을 적용하여 공간 정보를 반영하고, 다양한 dilation 값을 활용하여 receptive field를 확장함으로써 블러 제거 성능을 향상시켰다.

Abstract

Existing deep learning-based methods have difficulty effectively handling blur with diverse sizes and directions, which leads to a degradation of image deblurring performance. To address this issue, this paper proposes a novel Local Adaptive Blur Attention-Based Neural Network(LABA-Net). The proposed Spatial-wise Square Module(SSM) learns the spatial properties of blur across multiple directions and adaptively captures positional correlations. Additionally, the Symmetrical Square Depth-wise Overparameterized Convolution(SSDO) effectively incorporates spatial information and expands the receptive field by utilizing multiple dilation rates, thereby improving deblurring performance.

Keywords

image deblurring, motion blur, convolutional neural network, CNN, deep learning, attention

* 국립금오공과대학교 반도체시스템공학부 석사과정
- ORCID: <https://orcid.org/0009-0001-2766-0864>

** 국립금오공과대학교 전자공학부 반도체시스템 전공 조교수
(공동교신저자)
- ORCID¹: <https://orcid.org/0009-0009-5830-8730>
- ORCID²: <https://orcid.org/0000-0001-8758-3250>

• Received: Jan. 23, 2025, Revised: Feb. 10, 2025, Accepted: Feb. 13, 2025
• Corresponding Authors: Ho Sub Lee and Harim Lee
Dept. of EE, Kumoh National Inst. of Tech., Korea
Tel.: 82+54-478-7451, 82+54-478-7428
Email: hslee8634@kumoh.ac.kr, hrlee@kumoh.ac.kr

1. 서 론

모션 블러(Motion blur)는 사진이나 비디오에서 빠른 움직임으로 인해 발생하는 현상으로, 카메라의 흔들림, 물체의 이동, 초점 조절 문제 등 다양한 요인에 의해 나타난다. 이 현상은 영상에 왜곡을 유발하며, 세부 정보의 손실, 선명도 저하 등 영상 품질을 크게 저하시킨다. 이미지 디블러링(Image deblurring)은 이러한 문제를 해결하기 위해 모호하거나 흐릿한 이미지를 복원하여 원래의 선명한 상태로 되돌리는 기술을 의미한다. 특히, 컴퓨터 비전 분야에서는 블러로 인한 영상 품질 저하를 극복하고 고품질 영상 복원을 위해 이미지 디블러링 기술에 대한 연구가 활발히 진행되고 있다. 디블러링은 이미지 및 비디오 데이터의 전처리 과정으로, 세그멘테이션[1], 객체 검출[2], 3D 재구성[3], 타겟 추적[4] 등 다양한 응용 분야에서 핵심적인 역할을 한다.

전통적인 디블러링 방법[5]-[7]은 주로 Deconvolution 기반 기법에 의존하였다. 이 방법은 블러 커널(Blur kernel)을 추정 후, 이를 역변환하여 원본 이미지를 복원하는 방식으로 이루어진다. 전통적인 이미지 디블러링 기법들은 블러 커널을 정확히 추정하기 어렵다는 한계와 반복적인 연산으로 인한 높은 계산 비용 문제가 지속적으로 제기되어 왔다.

이러한 문제를 해결하기 위해 최근에는 DNN(Deep Neural Network)를 활용한 학습 기반 디블러링 방식[8]-[18]이 제안되고 있다. 학습 기반 디블러링은 블러 커널을 추정하지 않고, 블러가 발생한 이미지를 직접 입력으로 받아 원본 이미지를 복원하는 접근 방식을 채택한다. 이 방식은 블러링된 이미지를 입력으로 받아 디블러링된 이미지를 바로 출력할 수 있어 효율적이다. 또한, 저수준 특성에서 고수준 특성에 이르기까지 계층적으로 학습하여 이미지의 복잡한 패턴을 효과적으로 이해하고 처리할 수 있어, 전통적인 방식에 비해 더욱 우수한 복원 성능을 제공한다. 이러한 특성 덕분에 학습 기반 디블러링은 다양한 블러 환경에서 강건한 복원 결과를 도출하는 데 있어 중요한 역할을 하고 있다. 최근, 다양한 학습 기반 디블러링 기법들이 제안되며 이미지 디블러링 성능이 크게 향상되고 있다. [8]은

블러의 다양한 특성을 효과적으로 학습하기 위해 scale 비율이 0.5인 3개의 scale을 사용한 Deep Multi-Scale CNN 모델을 제안하여 실제 블러에 대한 높은 복원 성능을 달성하였다. [9]는 기존 scale에서의 블러를 복원한 뒤, scale을 점진적으로 축소하며 작은 크기의 블러를 순차적으로 복원할 수 있도록 설계된 순환 구조 네트워크를 제안하여 다양한 크기의 블러에 적응적으로 대응하였다. [10]은 이미지를 크기별로 자른 다중 패치를 입력으로 받아들여 전역적 및 지역적 블러 특성을 동시에 학습하도록 설계하였다. 또한, 인코더-디코더 구조를 4개 쌓아 세밀한 부분부터 전체적인 선명도를 점진적으로 복원하는 Fine-to-Coarse 학습 방식을 도입하여 복원 정확도를 높였다. [11]은 3개의 인코더와 3개의 디코더로 구성된 네트워크를 통해 파라미터를 선택적으로 공유하고, 다층 특성 맵 간 skip connection을 중첩하여 파라미터 효율성을 높이는 동시에 디블러링 성능을 향상시켰다. [12]와 [13]은 GAN(Generative Adversarial Networks)를 활용한 Blind Motion 디블러링 기법을 제안하였다. [12]는 GAN 기반 end-to-end 학습 방식을 사용한 반면, [13]은 [12]의 방식이 전역 정보를 고려하지 못하는 문제를 해결하기 위해 FPN(Feature Pyramid Network)를 활용하여 전역적 및 지역적 특징을 동시에 학습하도록 네트워크 구조를 개선하였다. 또한, 두 개의 Discriminator를 도입하여 Multi-scale에서의 복원 성능을 평가하였다. [14]는 기존의 업샘플링/다운샘플링 방식 대신 dilated convolution을 변형한 Straight Dilated Convolution을 적용하여 필터 크기를 확장하지 않고 receptive field를 늘려 다양한 블러 특성을 효과적으로 포착하였다. 또한, wavelet transformation을 전처리와 후처리 단계에 추가하여 블러 특성을 더 잘 표현하도록 하였다. [15]는 U-Net 구조를 기반으로 coarse-to-fine 방식이 아닌 coarse와 fine 정보를 동시에 처리할 수 있는 비대칭적 단일 단계 네트워크를 제안하여 효율성을 높였다. [16]은 SPPCARB(Spatial Pyramid Pooling Channel Attention Residual Block)을 적용하여 지역적 정보와 채널 간 상호 의존성을 결합하고, 패치 단위로 이미지를 분할하여 다양한 scale의 특징을 추출함으로써 디블러링 성능을 개선하였다.

[17]은 Multi-scale 구조에서 발생하는 정보 손실 문제를 해결하기 위해 1개의 인코더와 2개의 디코더로 구성된 네트워크를 제안하였다. 두 번째 디코더는 90도 회전된 convolution 커널을 사용하고, 이를 통해 수직 및 수평 방향의 블러 특성을 효과적으로 모델링하였다. 이러한 설계를 통해 기존 U-Net 구조와 동일한 파라미터 수로 더 나은 디블러링 성능을 달성하였다. [18]은 반복적인 프레임 구조로 인해 발생하는 추론 시간 및 메모리 사용 문제를 해결하기 위해 단일 순방향 네트워크를 설계하였다. 이 네트워크는 다중 커널 스트립 풀링을 활용하여 self-attention 구조에서 발생할 수 있는 공간 정보 손실 문제를 해결하고, 다양한 크기와 방향의 흐림 패턴을 분리하였다. 또한, [30]은 물체 경계와 세부 정보를 복원하는 기존 방법의 한계를 극복하기 위해 채널 간 상호 의존성을 학습하는 지역적 attention 기반 네트워크를 제안하여 디블러링 성능을 효과적으로 향상시켰다. 하지만 기존의 이미지 디블러링 방법들은 여전히 성능 향상에 한계를 보인다. 첫째, 블러의 공간적 정보를 충분히 활용하지 못해 효과적인 디블러링이 어렵다. 둘째, 픽셀 간 상관관계를 효과적으로 고려하지 못해 상호 연관된 블러를 정확하게 복원하는 데 한계가 있다. 마지막으로, 다양한 크기의 블러 특성을 반영하지 못해 전반적인 복원 성능이 제한되는 문제가 존재한다.

본 논문에서는 다양하고 불규칙적인 방향과 크기를 가진 블러들의 공간 정보를 효과적으로 고려하여 디블러링 성능을 향상시키기 위한 LABA-Net (Local Adaptive Blur Attention-Based Neural Network)을 제안한다. 기존 self-attention 구조는 각 위치 간의 상관관계만을 고려하여 공간 정보 손실을 초래하고, 이는 디블러링 성능 저하로 이어진다. 이를 해결하기 위해 LABA-Net은 블러 크기에 따른 공간적 정보를 적응적으로 추출할 수 있는 SSM(Spatial-wise Square Module)을 적용하였다. SSM은 SSA(Spatial-wise Square Attention)와 SSDO (Symmetrical Spatial-wise Depth-wise Overparameterized convolution)로 구성되며, 두 구성 요소는 병렬적으로 동작한다. 또한 LABA-Net은 다양한 커널 크기를 활용한 풀링 기법을 적용함으로써, 다양한 크기와 방향의 블러를 적응적으로 학습

하고 블러의 공간적 특성과 이들 간의 상관관계를 효과적으로 포착할 수 있도록 설계되었다.

본 논문의 주요 기여는 다음과 같다.

1. 기존 self-attention 기법에서 발생하는 공간 정보 손실 문제를 극복하기 위해, Multi-scale 적응형 풀링 어텐션 메커니즘을 도입한 LABA-Net을 제안한다. 이 메커니즘은 추출된 공간 정보를 바탕으로 여러 크기의 커널을 사용하여 적응형 풀링을 통해 다양한 방향과 크기를 가진 블러의 특징을 더욱 효과적으로 포착하고, 추출된 특징을 결합하여 블러의 전역적 및 지역적 특징을 학습하여 제안하는 방법이 공간적 연관성과 방향성을 반영할 수 있도록 설계하였다.

2. 다양한 dilation 값으로 구성된 Depth-wise Overparameterized convolution을 병렬적으로 통과 후 결합하여 Receptive field를 확장한다. 또한, 공간 정보를 전후로 강조하여 블러의 공간적 특징과 지역적 정보의 관계를 학습한다. 이를 통해 다양한 크기와 방향을 가진 블러를 효과적으로 복원할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 지금까지 제안된 기존 디블러링 기법과 관련된 선행 연구들을 검토한다. 3장에서는 제안하는 방법의 구체적인 구조와 학습에 사용된 손실 함수에 대해 상세히 설명한다. 4장에서는 실험 결과를 통해 제안하는 방법의 성능을 평가하며, 기존 방법들과의 정량적 및 정성적 비교를 진행한다. 마지막으로 5장에서는 연구의 결론을 정리하고 주요 연구 결과를 요약한다.

II. 관련 연구

2.1 최적화 기반 기법

기존 디블러링 연구는 주로 최적화 기반 방식에 의존하였다. [35]는 패치 기반 커널 추정 전략을 도입하여 이미지의 가장자리 및 모서리를 모델링할 수 있는 사전 정보를 활용해 블러 커널을 추정하고 디블러링을 수행하였다. [36]은 정규화된 컬러 라인 사전 정보를 사용하여 블러 커널 추정을 위한 알고리즘을 제안하였으며, 이를 통해 노이즈를 억제하면서 선명한 엣지를 복원하였다.

[37]은 흐린 이미지의 어두운 픽셀 희소성을 이용하여 어두운 채널의 희소성을 강제하는 방식으로 블러 커널을 추정하였다. 또한, [38]은 빛줄기를 활용하여 흐림 커널을 최적화 프레임워크 내에서 추정하고 디블러링을 수행하는 방식을 제안하였다. 그러나 이러한 최적화 기반 방법들은 정규화 항목 선정, 가중치 설정, 최적화 알고리즘 선택 등 다양한 하이퍼파라미터 설정이 요구되며, 수렴 속도가 느리고 최적해를 찾기 어려운 한계를 가지고 있다.

2.2 딥러닝 기반 기법

최적화 기반 방법의 한계를 극복하기 위해 딥러닝 기반 디블러링 방법이 제안되었으며, 현재까지 빠르게 발전하고 있다. [23]은 패치 계층적 어텐션 구조 [31]-[33]를 사용하여 입력 이미지에서 공간적 블러 변화를 암시적으로 발견하고, 전역 및 로컬 정보를 조절하는 픽셀 적응형 및 피쳐 어텐션 설계를 제안하였다. 이를 통해 공간 위치에 따른 큰 블러 변화를 처리하고, 각 테스트 이미지를 적응적으로 복원할 수 있었다. [39]는 흐린 이미지 패치를 입력으로 받아 선명한 이미지 패치를 복원하는 신경망을 설계하였으며, 이를 통해 초기 선명한 이미지 추정값과 관측 이미지 사이에서 단일 전역 블러 커널을 명시적으로 추정할 수 있었다. 또한, [34]는 U-Net 기반 Multi-scale 특징 cross-fusion 전략과

self-mechanism을 결합하여 비균일 블러 이미지를 복원하였으며, wavelet transform을 depth residual network에 통합하여 세부 texture 정보를 복원하였다.

III. 제안하는 방법

3.1 LABA-Net의 구조

본 논문에서 제안하는 LABA-Net은 디블러링에 사용하는 DNN의 고질적인 문제인 공간 정보 손실 문제를 해결하고자 제안되었다. 이미지를 디블러링하여 우수하게 복원하기 위해서는, 블러의 공간 정보를 손실시키지 않고 학습하는 것이 중요하다. 기존에 사용되었던 방식들은 공간 정보를 학습하는 데에는 한계가 존재했다. LABA-Net은 이러한 문제를 해결하고자, 공간 정보에 집중하기 위해 제안한 새로운 구조의 네트워크이다. 그림 1에 제시된 것처럼, LABA-Net은 앞 뒤의 convolution 레이어들과 여러 개 쌓인 SSM들로 이루어져 있다. SSM은 SSA과 SSDO이 병렬로 구성되어 있다. Input 이미지는 convolution을 통과하면서 채널 개수를 늘리고, 특징 맵의 크기는 줄어든다. 이는 SSM을 모두 통과한 뒤, Transposed convolution을 통해 원래의 크기로 원상복구한다. Input은 SSM들을 통과하면서 블러들의 저수준부터 고수준의 공간 정보를 포착하고 복원하는 역할을 한다.

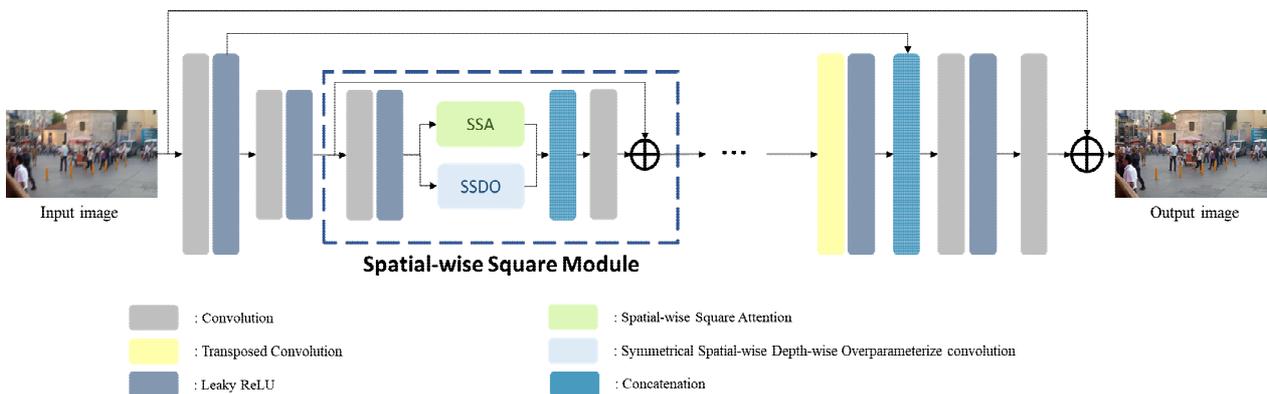


그림 1. LABA-Net의 구조
Fig. 1. Overview of LABA-Net

3.2 Spatial-wise Square Attention(SSA)

SSA는 그림 2 와 같은 구조로 이루어져 있다. Convolution layer들을 통과한 input feature는 먼저 1×1 convolution을 통해 특징을 추출하고 차원 수를 조절한다. 이를 식으로 표현하면 식 (1)과 같다.

$$F_1 = (W_1 * F_{in}), \quad (1)$$

여기서 F_{in} 은 convolution layer들을 통과한 input feature를 의미하며, W_1 는 1×1 convolution을 통해 생성된 가중치를 의미한다. 이후, F_1 은 max pooling을 통과하여 중요한 공간적 특징을 추출한다. 이후 sigmoid를 사용하여 특징의 중요도에 따라 가중치를 부여하도록 하는 어텐션 기능을 수행하도록 한다. 이를 수식으로 표현하면 식 (2)와 같다.

$$F_2 = UP(W_g(P_{\max}(W_2 * F_1))), \quad (2)$$

여기서 W_2 는 stride = 2인 convolution을 통과한 가중치이고, P_{\max} 는 max pooling을 의미한다. 여기서의 kernel size는 7이다. W_g 는 convolution group을 통과한 가중치이며, convolution group은 conv-ReLU-conv-ReLU-conv 형태로 구성되어 있다. UP은 업샘플링을 의미하며, 이 구조에서는 보간법을 사용한다. 이후로는 sigmoid 함수인 σ_s 를 적용하여 중요한 공간 정보를 강조하고 skip connection을 추가하여 저수준에서 잃을 수 있는 정보를 보충한다. 이를 통한 결과물은

식 (3)과 같이 표현할 수 있다:

$$F_{out} = \sigma_s(W_3(F_2 + F_1)) \times F_{in} \quad (3)$$

여기서 W_3 과 σ_{sig} 는 각각 1×1 convolution의 가중치, sigmoid 함수를 의미한다. 이 과정은 SSA의 전/후에 추가되어 지역적 공간 정보를 더욱 집중할 수 있도록 한다.

이를 통과한 특징맵들은 ASP(Adaptive Square Pooling)을 수행한다. Adaptive Square Pooling은 특징맵들을 $N \times N \times C$ ($N=1,3,5,7$) 로 각각 풀링을 진행한 후 Depth-wise Overparameterized convolution을 통과시킨 뒤, 이를 concatenate 하는 구조로 이루어져 있다. 기존의 풀링 방식은 kernel 사이즈가 고정값이기 때문에, 다양한 크기로 발생한 블러들을 모두 고려할 수 없었다. 하지만 ASP를 수행함으로써 다양한 사이즈에서 풀링을 진행하며 크고 작은 특징들까지 세밀하게 학습할 수 있다. 이후 Depth-wise Overparameterized convolution을 사용하여 receptive field를 확장시켜 더 많은 정보를 학습할 수 있도록 한다. 이후 이 결과물들을 $H \times W \times C$ 로 만들도록 보간법을 수행하고, 이를 concatenation 하며 각 픽셀간의 상관관계를 알아낼 수 있다. 이는 이미지를 디블러링 할 때에 블러의 수평/수직 정보만 고려할 수 있던 기존의 방법에 비해, 정사각형 모양으로 결과물이 나오게 Adaptive Average Pooling을 진행하여 다양한 방향으로 흩어져 있는 블러들의 방향들을 학습하여 블러들을 복원할 수 있다.

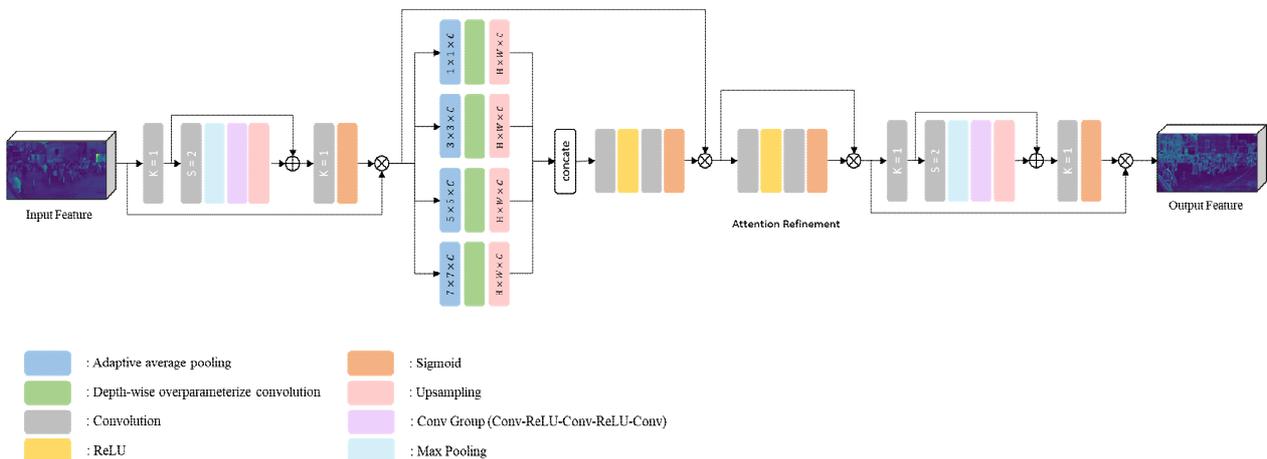


그림 2. SSA의 구조
Fig. 2. Architecture of SSA

이후에는, convolution layer와 ReLU activation function σ_r 를 통과시켜 픽셀간의 상관관계를 표현하는 어텐션 마스크를 생성한다. 이를 통해 중요한 부분에 대한 특징만을 강조할 수 있다. 이를 수식으로 표현하면 식 (4)와 같다.

$$M_{ASP} = \sigma_s(\text{conv}(\sigma_r(\text{conv}(y_1 \oplus y_3 \oplus y_5 \oplus y_7)))) \quad (4)$$

여기서 M_{ASP} 은 ASP로 생성된 어텐션 마스크를 의미한다. 이를 통해 중요한 전역적 및 지역적 정보와 특징 간의 상관관계를 포착한 이 모델은, 이후 Attention Refinement를 진행하여 비선형성을 강조시킨다. Attention Refinement의 구조를 수식으로 표현하면 식 (5)와 같다.

$$F_{AR} = \sigma_{sig}(\text{conv}(\sigma(\text{conv}(\cdot)))) \quad (5)$$

여기에 F_{out} 과 M_{ASP} 를 Element-wise multiplication한 결과물을 \tilde{x} 라 할 때, Attention Refinement를 통과한 결과물은 식 (6)과 같다.

$$F_{AR}(\tilde{x}) \otimes \tilde{x} \quad (6)$$

여기서 \otimes 는 element-wise multiplication을 의미한다.

3.3 Symmetrical Spatial-wise Depth-wise Overparameterize convolution(SSDO)

SSDO는 기존의 convolution보다 receptive field를 더욱 넓혀, 특징 정보를 폭넓게 추출하기 위해 그림 3과 같은 구조로 제안한다. 이를 위해, 앞서 언급된

식 (1), (2), (3)을 SSDO의 앞 뒤로 통과시킨다. 이를 통해서 본 네트워크는 중요한 공간적 특징 정보를 우선적으로 강조하고, 추출된 이후 다시 한 번 강조하여 공간 정보의 손실을 최소화하고 블러 복원 능력을 향상시키는데 기여한다. 이후 Depth-wise overparameterized convolution을 통해 dilation값을 사용하여 receptive field를 넓힌다. 이러한 방법으로 제안하는 방법은 보다 넓게 특징들을 포착하여 디블러링 성능을 향상시킬 수 있도록 하였다. SSDO에서는 이러한 구조를 바탕으로 Dilation 비율을 여러 가지로 설정하여, 결과물을 concatenation 하여 block을 두 번 반복하도록 구성하였다. 하지만 이러한 구조를 사용한다면 channel 개수가 기존에 비해 과도하게 증가하기 때문에 이러한 문제를 해결하고자, 두 block 사이에 Dilation 값이 1인 3×3 convolution을 추가하였다. 이를 통해 SSDO는 디블러링에 필요한 중요한 정보들을 효율적으로 추출하여 결합할 수 있다.

3.4 손실 함수

본 논문에서 제안하는 LABA-Net은 다음과 같은 Loss Function을 사용한다. 수식은 식 (7)과 같다.

$$L = L_{char} + \lambda L_{FFT} \quad (7)$$

λ 값은 0.01로 설정하였으며, L_{char} 의 수식은 식 (8)과 같다.

$$L_{char} = \sqrt{\|R - Y\|^2 + \epsilon^2} \quad (8)$$

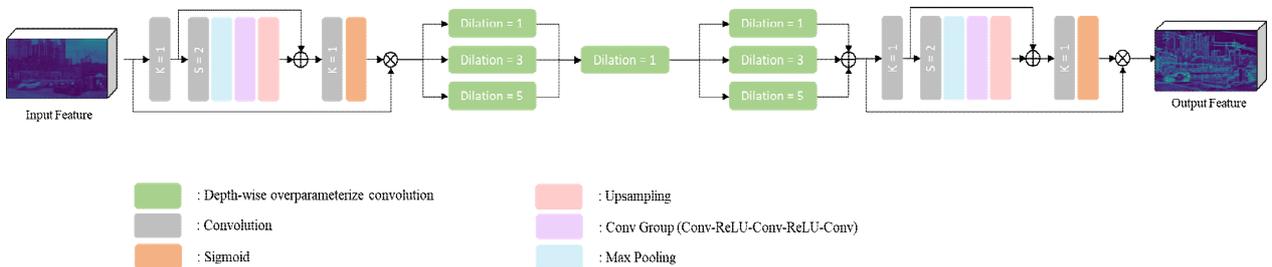


그림 3. SSDO의 구조
Fig. 3. Architecture of SSDO

수식에서 R와 Y는 각각 복원된 이미지와 Ground Truth를 의미하며, 이다. L_{FFT} 은 FFT Loss[15]를 의미하며 수식은 식 (9)와 같다.

$$L_{FFT} = \| F(R) - F(Y) \|_1 \quad (9)$$

여기서 F 는 Fourier Transform Function을 의미한다.

IV. 실험 결과

4.1 실험 조건

본 논문에서 제안한 네트워크는 파이토치를 사용하여 구현되었다. 모델의 사전 학습은 2개의 NVIDIA GeForce RTX 3090 GPU에서 수행되었으며, 총 3,000 epoch 동안 학습하였다. 이후, 모델의 성능을 더욱 향상시키기 위해 파인튜닝을 진행하였다. 파인튜닝은 4개의 NVIDIA GeForce RTX 3090 GPU에서 수행되었으며, 총 1,000 epoch 동안 학습하였다. 모든 실험에서 배치 크기는 8로 설정되었으며, 초기 학습률은 0.001로 시작하여, 학습이 진행됨에 따라 0.000001씩 점진적으로 감소시켰다. 모델의 최적화를 위해 Adam 옵티마이저를 활용하였다. 제안된 방법의 성능을 비교하기 위해 다음 벤치마크 모델들을 사용

하였다: DeepDeblur[8], DeblurGAN[12], SRN[9], PSS-NSC[11], DMPHN[10], MTRNN[29], DBGAN[40], MSCAN[16], MIMO-UNet[15], XYDeblur[17], BANet[18], LACA-SSN[30], MRDNet[34]. 성능 평가는 다음의 5가지 성능 지표를 사용하여 수행되었다: PSNR(Peak Signal-to-Noise Ratio)[24], SSIM(Structural Similarity)[24], MS-SSIM(Multi-Scale Structural Similarity)[25], VIF(Visual Information Fidelity)[26], MSE(Mean Square Error)[27].

4.2 성능 평가

본 섹션에서는 제안한 방법이 기존 방법들보다 디블러링 성능의 우수함을 입증하기 위해, 13가지 벤치마크 방법들과 제안한 방법의 디블러링 성능을 비교하였다. 벤치마크로 사용한 방법들과 제안한 방법의 성능을 GoPro dataset에서 정량적으로 평가한 결과는 Table 1에 제시되어 있다. 표 1에 나타난 바와 같이, 제안된 LABA-Net은 최신 기법인 [34]보다 PSNR 1.17dB와 SSIM 0.096 더 높은 성능을 보였다. 또한, 벤치마크 중 가장 성능이 높은 모델보다도 PSNR 0.42dB, SSIM 0.033 더 높은 성능을 달성하였다. 이러한 성능 개선은 그림 4의 시각적 비교를 통해 보다 명확하게 확인할 수 있다.

표 1. GoPro dataset[8]에 대한 기존방법과 제안하는 방법의 정량적 비교 표
Table 1. Quantitative comparison of GoPro dataset[8] and proposed method

Methods	Year	PSNR [dB] [24]	SSIM [24]	MS-SSIM [25]	VIF [26]	MSE [27]
DeepDeblur [8]	2017	27.83	0.9151	0.9377	0.4709	134.99
DeblurGAN [12]	2018	25.02	0.8474	0.8739	0.3544	251.90
SRN [9]	2018	30.24	0.9348	0.9574	0.5171	77.72
PSS-NSC [11]	2019	30.96	0.9420	0.9641	0.5325	64.39
DMPHN [10]	2019	31.39	0.9477	0.9684	0.5486	56.87
MTRNN [29]	2020	31.12	0.9447	0.9665	0.5482	62.70
DBGAN [40]	2020	31.18	0.9462	0.9652	0.5454	61.51
MSCAN [16]	2021	31.23	0.9452	0.9655	0.5499	61.29
MIMO-UNet [15]	2021	31.72	0.9503	0.9686	0.5621	54.78
XYDeblur [17]	2022	30.92	0.9436	0.9638	0.5394	65.43
BANet [18]	2022	32.54	0.9573	0.9744	0.5887	44.31
LACA-SSN [30]	2023	31.43	0.9469	0.9664	0.5535	59.01
MRDNet [34]	2024	31.71	0.9499	0.9682	0.5620	55.56
LABA-Net	2024	32.96	0.9606	0.9769	0.6029	39.61

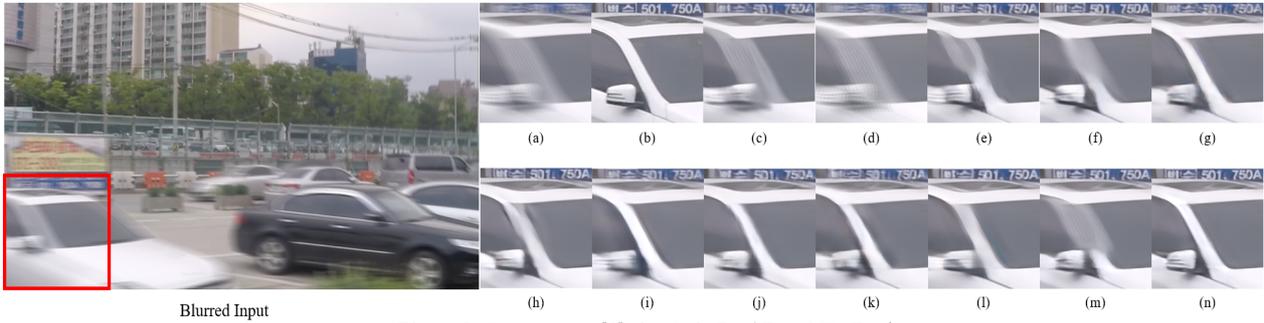


그림 4. GoPro dataset[8]에 대한 정성적 비교 결과

- (a) 블러 패치, (b) 정답 이미지, (c) DeepDeblur (PSNR : 18.24dB, SSIM : 0.7642), (d) DeblurGAN (18.04dB, 0.7354), (e) XYDeblur (20.64dB, 0.8236), (f) DMPHN (22.40dB, 0.8615), (g) PSS-NSC (24.58dB, 0.8665), (h) MIMO-UNet (24.54dB, 0.8705), (i) MSCAN [(24.39, 0.8688), (j) MTRNN (25.56dB, 0.8860), (k) SRN (24.35dB, 0.8630), (l) LACA-SSN (23.41dB, 0.8607), (m) BANet (20.68dB, 0.8420), (n) LABA-Net (27.84dB, 0.9207)

Fig. 4. Qualitative comparisons on GoPro dataset[8]

- (a) Blur patch, (b) Ground truth, (c) DeepDeblur (PSNR : 18.24dB, SSIM : 0.7642), (d) DeblurGAN (18.04dB, 0.7354), (e) XYDeblur (20.64dB, 0.8236), (f) DMPHN (22.40dB, 0.8615), (g) PSS-NSC (24.58dB, 0.8665), (h) MIMO-UNet (24.54dB, 0.8705), (i) MSCAN [(24.39, 0.8688), (j) MTRNN (25.56dB, 0.8860), (k) SRN (24.35dB, 0.8630), (l) LACA-SSN (23.41dB, 0.8607), (m) BANet (20.68dB, 0.8420), (n) LABA-Net (27.84dB, 0.9207)

V. 결 론

본 논문에서는 self-attention의 공간 정보 손실 문제를 해결하고, 효과적인 이미지 디블러링을 위해 LABA-Net을 제안하였다. LABA-Net은 어텐션 메커니즘을 통해 중요한 블러 영역을 강조하고 이를 병렬적으로 처리함으로써 효율적인 디블러링을 수행한다. 실험 결과, 제안된 LABA-Net은 기존 self-attention 기반 기법들의 단점을 극복하고 블러의 공간적 정보를 효과적으로 보존하여 우수한 성능을 보였다.

References

- [1] J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, and H. Lu, "Dual attention network for scene segmentation", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, pp. 3146-3154, Jun. 2019. <https://doi.org/10.1109/CVPR.2019.00326>.
- [2] S. Zheng, Y. Wu, S. Jiang, C. Lu, and G. Gupta, "Deblur-YOLO: Real-Time Object Detection with Efficient Blind Motion Deblurring", 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, pp. 1-8, Jul. 2021. <https://doi.org/10.1109/IJCNN52387.2021.9534352>.
- [3] J. Zhang, K. Yu, Z. Wen, X. Qi, and A. Paul, "3D reconstruction for motion blurred images using deep learning-based intelligent systems", Computers, Materials & Continua, Vol. 66, No. 2, pp. 2087-2104, Jan. 2021. <http://dx.doi.org/10.32604/cmc.2020.014220>.
- [4] Q. Guo, W. Feng, R. Gao, Y. Liu, and S. Wang, "Exploring the Effects of Blur and Deblurring to Visual Object Tracking", IEEE transactions on image processing, Vol. 30, pp. 1812-1824, Jan. 2021. <https://doi.org/10.1109/TIP.2020.3045630>.
- [5] W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration", Journal of the Optical Society of America, Vol. 62, No. 1, pp. 55-59, Jan. 1972. <https://doi.org/10.1364/JOSA.62.000055>.
- [6] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph", ACM Transactions on Graphics (TOG), Vol. 25, No. 3, pp. 787-794, Jul. 2006. <https://doi.org/10.1145/1141911.1141956>.
- [7] A. Levin, Y. Weiss, F. Durand, and W. T.

- Freeman, "Understanding Blind Deconvolution Algorithms", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 12, pp. 2354-2367, Dec. 2011. <https://doi.org/10.1109/TPAMI.2011.148>.
- [8] S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring", IEEE conference on computer vision and pattern recognition, pp. 3883-3891, Jul. 2017.
- [9] X. Tao, H. Gao, X. Shen, J. Wang, and J. Jia, "Scale-recurrent network for deep image deblurring", IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 8174-8182, Jun. 2018. <https://doi.org/10.1109/CVPR.2018.00853>.
- [10] H. Zhang, Y. Dai, H. Li, and P. Koniusz, "Deep stacked hierarchical multi-patch network for image deblurring", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, pp. 5978-5986, Jun. 2019. <https://doi.org/10.1109/CVPR.2019.00613>.
- [11] H. Gao, X. Tao, X. Shen, and J. Jia, "Dynamic scene deblurring with parameter selective sharing and nested skip connections", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, pp. 3848-3856, Jun. 2019. <https://doi.org/10.1109/CVPR.2019.00397>.
- [12] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks", IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 8183-8192, Jun. 2018. <https://doi.org/10.1109/CVPR.2018.00854>.
- [13] O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, "Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better", IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, pp. 8878-8887, Oct. 2019. <https://doi.org/10.1109/ICCV.2019.00897>.
- [14] W. Zou, M. Jiang, Y. Zhang, L. Chen, Z. Lu, and Y. Wu, "SDWNet: A straight dilated network with wavelet transformation for image deblurring", IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, pp. 1895-1904, Oct. 2021. <https://doi.org/10.1109/ICCVW54120.2021.00216>.
- [15] S.-J. Cho, S.-W. Ji, J.-P. Hong, S.-W. Jung, and S.-J. Ko, "Rethinking coarse-to-fine approach in single image deblurring", IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, pp. 4641-4650, Oct. 2021. <https://doi.org/10.1109/ICCV48922.2021.00460>.
- [16] S. Wan, S. Tang, X. Xie, J. Gu, R. Huang, B. Ma, and L. Luo, "Deep convolutional-neural-network-based channel attention for single image dynamic scene blind deblurring", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 31, No. 8, pp. 2994-3009, Aug. 2021. <https://doi.org/10.1109/TCSVT.2020.3035664>.
- [17] S. W. Ji, J. Lee, S. W. Kim, J. P. Hong, S. J. Baek, S. W. Jung, and S. J. Ko, "XYdeblur: Divide and conquer for single image deblurring", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, pp. 17421-17430, Jun. 2022. <https://doi.org/10.1109/CVPR52688.2022.01690>.
- [18] F. J. Tsai, Y. T. Peng, C. C. Tsai, Y. Lin, and C.-W. Lin, "BANet: A Blur-Aware Attention Network for Dynamic Scene Deblurring", IEEE transactions on image processing, Vol. 31, pp. 6789-6799, Oct. 2022. <https://doi.org/10.1109/TIP.2022.3216216>.
- [19] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 2117-2125, Jul.

2017. <https://doi.org/10.1109/CVPR.2017.106>.
- [20] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions", arXiv preprint, arXiv:1511.07122, Nov. 2015. <https://doi.org/10.48550/arXiv.1511.07122>.
- [21] C. Min, G. Wen, B. Li, and F. Fan, "Blind deblurring via a novel recursive deep cnn improved by wavelet transform", IEEE Access, pp. 69242-69252, Nov. 2018. <https://doi.org/10.1109/ACCESS.2018.2880279>.
- [22] Q. Hou, L. Zhang, M.-M. Cheng, and J. Feng, "Strip Pooling: Rethinking Spatial Pooling for Scene Parsing", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, pp. 4003-4012, Jun. 2020. <https://doi.org/10.1109/CVPR42600.2020.00406>.
- [23] M. Suin, K. Purohit, and A. N. Rajagopalan, "Spatially-attentive patchhierarchical network for adaptive motion deblurring", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, pp. 3606-3615, Jun. 2020. <https://doi.org/10.1109/CVPR42600.2020.00366>.
- [24] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", IEEE transactions on image processing, Vol. 13, No. 4, pp. 600-612, Apr. 2004. <https://doi.org/10.1109/TIP.2003.819861>.
- [25] Z. Wang, E. P. Simoncelli, and A. C. Bovik. "Multiscale structural similarity for image quality assessment", The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, Vol. 2, Nov. 2003. <https://doi.org/10.1109/ACSSC.2003.1292216>.
- [26] H. R. Sheikh and A. C. Bovik, "Image information and visual quality", IEEE Transactions on image processing, pp. 430-444, Feb. 2006. <https://doi.org/10.1109/TIP.2005.859378>.
- [27] H. L. Tan, Z. Li, Y. H. Tan, S. Rahardja, and C. Yeo, "A perceptually relevant MSE-based image quality metric", IEEE Transactions on Image Processing, Vol. 22, No. 11, pp. 4447-4459, Nov. 2013. <https://doi.org/10.1109/TIP.2013.2273671>.
- [28] Z. Shen, W. Wang, X. Lu, J. Shen, H. Ling, T. Xu, and L. Shao, "Human-Aware Motion Deblurring", IEEE/CVF international conference on computer vision, Seoul, Korea, pp. 5572-5581, Oct. 2019. <https://doi.org/10.1109/ICCV.2019.00567>.
- [29] D. Park, D. U. Kang, J. Kim, and S. Y. Chun, "Multi-Temporal Recurrent Neural Networks For Progressive Non-Uniform Single Image Deblurring With Incremental Temporal Training", European Conference on Computer Vision, Glasgow, United Kingdom, Vol. 12351, pp. 327-343, Aug. 2020. https://doi.org/10.1007/978-3-030-58539-6_20.
- [30] H. S. Lee and S. I. Cho, "Locally Adaptive Channel Attention-based Spatial-Spectral Neural Network for Image Deblurring", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 33, No. 10, pp. 5375-5390, Oct. 2023. <https://doi.org/10.1109/TCSVT.2023.3250509>.
- [31] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks", IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 7132-7141, Jun. 2018. <https://doi.org/10.1109/CVPR.2018.00745>.
- [32] X. Zhang, T. Wang, J. Wang, G. Tang, and L. Zhao. "Pyramid channel-based feature attention network for image dehazing", Computer Vision and Image Understanding, Vol. 197-198, pp. 103003, Aug. 2020. <https://doi.org/10.1016/j.cviu.2020.103003>.
- [33] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image Super-Resolution Using Very Deep Residual Channel Attention Networks", European Conference on Computer Vision (ECCV), Munich, Germany, pp. 286-301, sep. 2018. https://doi.org/10.1007/978-3-030-01234-2_18.
- [34] B. Zhang, J. Sun, F. Sun, F. Wang, and B. Zhu, "Image deblurring method based on self-attention

and residual wavelet transform", Expert Systems with Applications, Vol. 244, pp. 123005, Jun. 2024. <https://doi.org/10.1016/j.eswa.2023.123005>.

- [35] L. Sun, S. Cho, J. Wang, and J. Hays, "Edge-based blur kernel estimation using patch priors", IEEE International Conference on Computational Photography (ICCP), Cambridge, MA, USA, pp. 1-8, Apr. 2013. <https://doi.org/10.1109/ICCPHOT.2013.6528301>.
- [36] W.-S. Lai, J.-J. Ding, Y.-Y. Lin, and Y.-Y. Chuang, "Blur kernel estimation using normalized color-line priors", IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, pp. 64-72, Jun. 2015. <https://doi.org/10.1109/CVPR.2015.7298601>.
- [37] J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Blind image deblurring using dark channel prior", IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 1628-1636, Jun. 2016. <https://doi.org/10.1109/CVPR.2016.180>.
- [38] Z. Hu, S. Cho, J. Wang, and M.-H. Yang, "Deblurring low-light images with light streaks", IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, pp. 3382-3389, Jun. 2014. <https://doi.org/10.1109/CVPR.2014.432>.
- [39] A. Chakrabarti, "A neural approach to blind motion deblurring", arXiv preprint, arXiv:1603.04771, Mar. 2016. <https://doi.org/10.48550/arXiv.1603.04771>.

저자소개

이 민 영 (Minyoung Lee)



2023년 8월 : 국립금오공과대학교
전자공학부 반도체시스템전공
(공학사)

2023년 8월 ~ 현재 :
국립금오공과대학교
반도체시스템공학부 석사과정
관심분야 : 영상처리, 이미지

디블러링, 이미지 초해상화

이 호 섭 (Ho Sub Lee)



2014년 2월 : 경북대학교 IT 대학
전자공학부 (공학사)

2016년 2월 : 포항공과대학교
전기전자공학 (공학석사)

2020년 2월 : 포항공과대학교
전기전자공학 (공학박사)

2020년 3월 ~ 2021년 2월 :

경남대학교 전자공학부 조교수

2021년 3월 ~ 2023년 2월 : 대구대학교 전자공학부
조교수

2023년 3월 ~ 현재 : 국립금오공과대학교 전자공학부
반도체시스템전공 조교수

관심분야 : 이미지 분석 및 향상, 비디오 처리, 컴퓨터
비전, 딥러닝

이 하 림 (Harim Lee)



2013년 2월 : 경북대학교 IT 대학
전자공학부(공학사)

2015년 2월 : 포항공과대학교 IT
융합공학(공학석사)

2020년 2월 : UNIST 전기전자공학
(공학박사)

2020년 3월 ~ 2021년 8월 :

포항공과대학교 박사 후 연구원

2021년 9월 ~ 현재 : 국립금오공과대학교 전자공학부
반도체시스템전공 조교수

관심분야 : 딥러닝 기반 지능형 시스템, 목소리 감성
생성 모델, 목소리-얼굴 잠재 벡터 정렬 연구, 차세대
통신