

특집논문 (Special Paper)

방송공학회논문지 제26권 제6호, 2021년 11월 (JBE Vol.26, No.6, November 2021)

<https://doi.org/10.5909/JBE.2021.26.6.725>

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

## Wavelet 기반의 영상 디테일 향상 잡음 제거 네트워크

정 군<sup>a)</sup>, 위 승 우<sup>a)</sup>, 정 제 창<sup>a)†</sup>

### WDENet: Wavelet-based Detail Enhanced Image Denoising Network

Jun Zheng<sup>a)</sup>, Seungwoo Wee<sup>a)</sup>, and Jechang Jeong<sup>a)†</sup>

#### 요 약

현재 카메라 성능이 점점 발전해 왔지만 카메라로부터 얻은 디지털 영상에는 잡음 (Noise)이 존재하고 이는 높은 해상도의 영상을 획득하는 데 있어서 방해요소로 작용한다. 전통적으로 잡음을 제거하기 위하여 필터링 방법을 사용해 왔고 최근 딥러닝 기법의 하나인 합성곱 신경망 (Convolutional Neural Network)은 영상 잡음 제거 분야에서 전통적인 기법보다 좋은 성능을 나타내고 있어 많은 연구가 진행되고 있다. 하지만 합성곱 신경망으로 학습하는 과정에서 영상 내 디테일한 부분이 손실될 수 있는 문제점이 있다. 본 논문에서는 웨이블릿 변환 (Wavelet Transform)을 기반으로 영상 내 디테일 정보도 같이 학습하여 영상 디테일을 향상하는 잡음 제거 합성곱 신경망 네트워크를 제안한다. 제안하는 네트워크는 디테일 향상 서브 네트워크 (Detail Enhancement Subnetwork)와 영상 잡음 추출 서브 네트워크 (Noise Extraction Subnetwork)를 이용하게 된다. 실험은 가우시안 잡음과 실제 카메라 잡음을 통해 진행했고 제안하는 방법은 기존 알고리듬보다 디테일 손실 문제를 효과적으로 해결할 수 있었고 객관적 품질 평가와 주관적 품질 비교에서 모두 우수한 결과가 나온 것을 확인하였다.

#### Abstract

Although the performance of cameras is gradually improving now, there are noise in the acquired digital images from the camera, which acts as an obstacle to obtaining high-resolution images. Traditionally, a filtering method has been used for denoising, and a convolutional neural network (CNN), one of the deep learning techniques, has been showing better performance than traditional methods in the field of image denoising, but the details in images could be lost during the learning process. In this paper, we present a CNN for image denoising, which improves image details by learning the details of the image based on wavelet transform. The proposed network uses two subnetworks for detail enhancement and noise extraction. The experiment was conducted through Gaussian noise and real-world noise, we confirmed that our proposed method was able to solve the detail loss problem more effectively than conventional algorithms, and we verified that both objective quality evaluation and subjective quality comparison showed excellent results.

Keyword : Image Denoising, Convolutional Neural Network, Wavelet Transform, Detail Enhancement

## I. 서 론

현재 카메라 성능의 발전과 스마트폰의 보급화에 의해 영상은 누구나 쉽게 획득할 수 있게 되었다. 하지만 카메라로부터 얻은 디지털 영상 내에는 코딩, 전송, 수집, 처리하는 과정 또는 야간이나 악천후 등의 환경적 요인으로 인해 열화와 잡음 (Noise)이 생성되기도 한다. 따라서 선명한 영상을 얻기 위해서는 잡음을 최소화할 수 있는 영상 처리 기술이 필수적이고 이러한 영상 내에 존재하는 잡음을 제거해 영상 화질을 개선하는 과정을 영상 잡음 제거 (Image Denoising)라고 한다. 영상 잡음 제거 기술에서 카메라 센서 잡음을 포함한 일반적인 영상에서의 잡음을 모델링할 때 가산성 백색 가우시안 잡음 (Additive White Gaussian Noise, AWGN)을 많이 사용한다. 이것은 모든 주파수 대역에 존재하는 잡음을 의미하며 가우시안 형태의 확률 변수를 가지고 있기 때문에 자연계에서 흔히 볼 수 있는 잡음과 유사한 특성을 가지고 있다. 즉 잡음 제거 알고리듬은 잡음이 있는 영상  $x$ 로부터 잡음이 없는 깨끗한 영상  $x_{dean}$ 를 복원하는 것을 목적으로 하며 아래 수식으로 나타날 수 있다.  $n$ 은 잡음을 의미한다.

$$x = x_{dean} + n \quad (1)$$

지난 수십 년간 필터링 기반의 방법부터 딥 러닝 기반의 방법까지 다양한 방법들이 AWGN을 사용하여 영상 잡음 제거 연구를 진행하였다. 전통적인 방법으로는 필터링 기반의 알고리듬을 이용하여 영상 잡음 제거를 진행하였고 특히 비지역적 평균 필터 (Non-local Mean Filter)<sup>[1]</sup>를 사용하는 BM3D<sup>[2]</sup>, LSSC<sup>[3]</sup>, NCSR<sup>[4]</sup>와 WNNM<sup>[5]</sup> 알고리듬들

a) 한양대학교 응집전자공학부(Hanyang University)

‡ Corresponding Author : 정제창(Jechang Jeong)

E-mail: jjeong@hanyang.ac.kr

Tel: +82-2-2220-4370

ORCID: <https://orcid.org/0000-0002-3759-3116>

\* 이 논문의 연구 결과 중 일부는 한국방송·미디어공학회 “2021년 하계 학술대회에서 발표한 바 있음.”

\* 이 연구는 2021년도 산업통상자원부 및 한국산업기술평가관리원 (KEIT) 연구비 지원에 의한 연구임 (20013726).

\* This work was supported by the Technology Innovation Program (20013726, Development of Industrial Intelligent Technology for Smart Factory) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea).

• Manuscript received September 6, 2021; Revised November 11, 2021; Accepted November 11, 2021.

이 대표적이다. 예를 들어 BM3D는 잡음이 있는 영상에서 블록을 추출하여 전체 영역에서 비슷한 블록들을 찾아 3차원 그룹으로 구성하고, 그룹에 협업 필터링 (Collaborative Filtering)을 실행한 다음 얻어진 추정치를 원래 위치에 반환하는 알고리듬이다. 이러한 방법들은 잡음 제거 기술에서 혁신적인 성과를 보였지만 크게 2가지의 단점이 있다. 첫째, 테스트 단계에서 복잡한 최적화 문제와 결과를 풀기 위해 많은 계산량이 필요하다. 둘째, 최적화 문제들이 대부분 비불록 행렬 (Non-Convex)이기 때문에 파라미터 값을 수동적으로 정해야 한다는 단점이 있다.

최근에는 기존 필터링 기반의 잡음 제거 기술에 이어 수동으로 파라미터 값을 정하는 방법과 달리 합성곱 신경망 네트워크 (Convolutional Neural Network, CNN)를 이용한 딥 러닝 기반의 알고리듬들이 컴퓨터 비전 분야에서 높은 성능을 보이면서 이를 이용한 잡음 제거 기술이 많이 개발되고 있다. 즉, 네트워크로 얻은 결과와 원본의 차이가 최소화되도록 학습하여 잡음을 제거하는 방법이다. 대표적인 것이 잔차 학습 (Residual Learning)<sup>[6]</sup>, 합성곱, 배치 정규화 (Batch Normalization)<sup>[7]</sup>와 ReLU<sup>[8]</sup>를 이용하여 17개의 층으로 AWGN을 제거하는 신경망인 DnCNN<sup>[9]</sup>이 있다. DnCNN은 잡음 있는 영상에서 원본 영상을 뺀 결과 영상으로 학습하기 때문에 상대적으로 출력 층 (Layer)에서 얻은 값의 범위가 기존 합성곱 신경망보다 작다. 이로써 학습의 정확도가 높아졌고 잡음 제거에 있어서 큰 효과를 보였다. 이후로는 합성곱 신경망 기반의 영상 잡음 제거 방법에 더 빠르고 유연한 해결책을 위해 FFDNet<sup>[10]</sup>이 등장하게 됐다. 이 방법은 잡음이 있는 영상과 잡음 레벨 (Noise Level) 정보를 함께 입력으로 사용하여 DnCNN의 네트워크와 동일하게 이용하면서 속도와 성능을 향상하였다. 이들뿐만 아니라 기존 모델 기반의 최적화 방법과 합성곱 신경망 네트워크를 결합하여 영상 잡음 문제를 해결하는 IRCNN<sup>[11]</sup> 알고리듬도 개발이 됐다. 하지만 이러한 단일 합성곱 신경망 네트워크로 잡음을 제거하는 과정에서 영상 내 디테일이 손실되고 AWGN과 같은 합성 잡음 기반의 잡음 제거 모델은 잡음 레벨이 학습 시와 동일한 경우에는 좋은 성능을 보여주지만 실제 카메라를 이용해 촬영된 실제 잡음 (Real-world Noise)에 대해서는 성능이 크게 저하되는 문제점이 있다.

이러한 문제를 해결하기 위해 본 논문에서는 웨이블릿 변

환 (Wavelet Transform)과 U-Net<sup>[12]</sup>구조를 이용한 WDENet (Wavelet-based Detail Enhanced Image Denoising Network) 모델을 제안한다. 이를 통해 영상 내 디테일 정보를 추가적으로 학습함으로써 원본 영상 복원 시 디테일이 손실되는 문제를 개선할 수 있었다. 그중 웨이블릿 변환은 주파수 영역에서 잡음과 원본 영상을 분리하여 디테일 정보를 학습하는 데에 효과적이고 U-Net 구조는 경계선 검출 (Edge Detection)에 예민하다. 제안하는 네트워크는 기존 딥 러닝 기반의 단일 합성곱 신경망 네트워크 모델과 달리 다중 경로 학습 (Multi Path Learning)으로 디테일 향상 서브 네트워크 (Detail Enhancement Subnetwork)와 영상 잡음 추출 서브 네트워크 (Noise Extraction Subnetwork)를 이용한다.

실험은 잡음 레벨이 각각 15, 25, 50인 AWGN을 추가한 그레이 잡음 영상 (Grayscale noisy image)과 실제 잡음 영상에서 많이 알려진 SIDD (Smartphone Image Denoising Dataset)<sup>[13]</sup> 컬러 잡음 영상 (Color noisy image)을 대상으로 진행했다. 제안하는 모델은 두 가지 상황에서 모두 기존의 방법보다 향상된 객관적 및 주관적 성능을 입증하였다. 특히 잡음 레벨이 50인 AWGN에서 객관적 품질 평가 지표 값 PSNR(Peak Signal-to-Noise Ratio)과 SSIM (Structural Similarity Index Map)는 각각 최대 0.26dB, 0.01 향상시켰고 실제 카메라 잡음에 대해서는 각각 최대 0.20dB, 0.0516 향상시켰으며, 주관적 품질 비교에서도 모델의 우수성을 보여주면서 디테일 부분의 보존 능력을 입증하였다.

본 논문의 구성은 다음과 같다. 2장에서는 제안한 방법에서 사용되는 이론적 지식인 잔차 학습과 2D 영상신호에 대한 이산 웨이블릿 변환 (Discrete Wavelet Transform, DWT)을 살펴본 후, 3장에서는 제안하는 방법의 전체적인 구조에 대해 설명한다. 4장에서는 가우시안 잡음과 실제 카메라 잡음에 대한 실험 결과를 통해 기존의 영상 잡음 제거 방법과 제안하는 방법을 비교하여 성능을 확인하고 마지막으로 5장에서는 결론을 맺는다.

## II. 이론적 배경

### 1. 잔차 학습

합성곱 신경망 네트워크는 영상을 분석하기 위해 영상

특징 (Feature)을 추출하는 데 유용한 알고리듬으로 영상 데이터를 직접 학습하고 특징을 이용하여 영상을 분류한다. 그래서 네트워크의 깊이는 모델의 성능을 결정하는 데 있어서 중요한 요소 중의 하나이고<sup>[14]</sup> 네트워크의 층 (Layer)을 깊게 쌓아 학습할수록 영상 데이터에 존재하는 대표적인 특징을 효과적으로 추출할 수 있어 좋은 학습 결과를 얻게 된다<sup>[15]</sup>. 하지만 합성곱 신경망 네트워크를 통해 영상의 특징을 추출하는 과정에서 네트워크의 층을 너무 깊게 쌓으면 파라미터 수의 증가에 의해 에러가 커지게 된다. 이러한 문제로 과적합 (Overfitting)의 문제와 달리 기울기가 소실 및 증폭 (Gradient Vanishing/Exploding) 현상이 발생하여 성능이 저하되는 문제가 존재하며 모델의 성능에 치명적인 영향을 끼치게 된다. 즉, 네트워크의 층이 많이 깊어지게 되면 오차 역전파 (Back Propagation)를 진행함에 따라 미분의 횟수가 많아져, 입력 층과 가까울수록 미분 값이 가중치  $w$ 의 영향을 많이 받게 된다. 이로써 기울기가 너무 작아지거나 커지는 상황이 발생하게 된다.

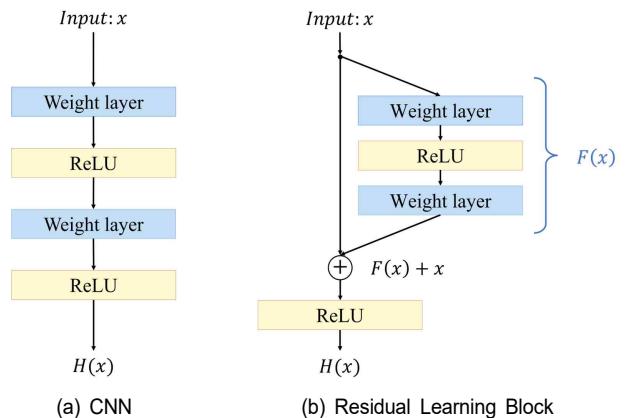


그림 1. 기존 합성곱 신경망 네트워크 구조와 잔차 학습  
 Fig. 1. Architecture of CNN and Residual Learning Block

이러한 문제점을 해결하기 위해 잔차 학습이 제안되었고 기울기의 영향을 줄여줌으로써 깊은 신경망 네트워크에서도 학습을 효과적으로 할 수 있도록 하였다. 잔차 학습은 입력을 그대로 출력에 더하는 속도 연결 (Shortcut Connection)이 추가되는 구조를 적용하였고 추가적인 파라미터가 많이 필요하지 않다. 즉 잔차 학습 블록 (Residual Learning Block)은 그림 1 (a)처럼 기존 합성곱 신경망 네트워크의

출력 값  $H(x)$ 를 최적으로 학습하는 대신 그림 1 (b)에서 입력 값  $x$ 와 출력 값  $H(x)$ 의 차이 값  $F(x)$ 를 최적으로 학습하는 것이 목표이다. 따라서 잔차 학습 블록의 출력 값  $H(x)$ 를 다음과 같이 재정의한다.

$$F(x) = H(x) - x \quad (2)$$

$$H(x) = F(x) + x \quad (3)$$

이러한 구조로 사용하게 되면 위의 솟컷 연결을 합성곱에 추가하여 역전파 과정에서 출력값  $H(x)$ 를  $x$ 로 미분하였을 때  $F'(x) + 1$ 이 되게 함으로써 적어도 1 이상의 값으로 최소한의 기울기를 만들어서 네트워크가 깊더라도 기존 방법보다 최적화하기 더 쉽고 안정적인 학습이 가능하며 층에 대해서 잔차를 학습한다는 구조로 성능 저하되는 문제를 해결할 수 있다.

## 2. 2D 영상신호에 대한 이산 웨이블릿 변환

이산 웨이블릿 변환 (DWT)은 입력 영상을 계층 구조로 재구성하며, 이를 통해 피처 맵 (Feature Map)으로 주파수 영역과 공간 영역에서의 정보를 파악할 수 있으므로 영상 내 디테일한 부분을 강조하는 데 도움이 된다. 따라서 본 논문에서는 2D 이산 웨이블릿 변환을 사용하게 되고 그중 대표적이고 계산량이 적은 Haar 웨이블릿<sup>[16]</sup>을 선택하였다. 그리고 DWT는 가역성을 가지고 있기 때문에 영상 내의 모든 정보를 다운 샘플링하고 역 웨이블릿 변환 (Inverse Wavelet Transform, IWT)으로 에지 (Edge) 성분에 순실 없

이 복원할 수 있는 특성을 가지고 있다.

그림 2처럼 DWT는 영상의 크기가 가로, 세로 각  $\frac{1}{2}$ 씩 줄어든 입력 영상이 저역 필터 (Low-pass Filter)  $f_{LL}$ 와 고역 필터 (High-pass Filter)  $f_{LH}$ ,  $f_{HL}$ ,  $f_{HH}$  총 4개의 필터를 거쳐 4개의 서브 밴드 (Sub Band) 영상  $x_{LL}$  (평균),  $x_{LH}$  (수평),  $x_{HL}$  (수직),  $x_{HH}$  (대각선)로 나누어지고 4개의 서브 밴드 영상에 IWT을 적용하게 되면 원본 영상으로 복원이 가능하다. Haar 웨이블릿의 경우 4개의 필터는 아래와 같이 정의된다.

$$\begin{aligned} f_{LL} &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, f_{LH} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \\ f_{HL} &= \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, f_{HH} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \end{aligned} \quad (4)$$

## III. 제안하는 방법

이번 장에서는 웨이블릿 변환과 U-Net 구조를 기반으로 디테일 정보를 같이 학습하는 합성곱 신경망 네트워크의 전체적인 구조를 소개하고, 다음 제안하는 WDENet 모델에 두 서브 네트워크를 사용하는 이유에 대해 설명한다.

### 1. WDENet 모델 구조

그림 3은 제안하는 WDENet 모델의 입력부터 출력까지 전체 네트워크 구조를 보여준다. 제안하는 WDENet 모델은 그림 3과 같이 디테일을 향상하는 서브 네트워크와 영상

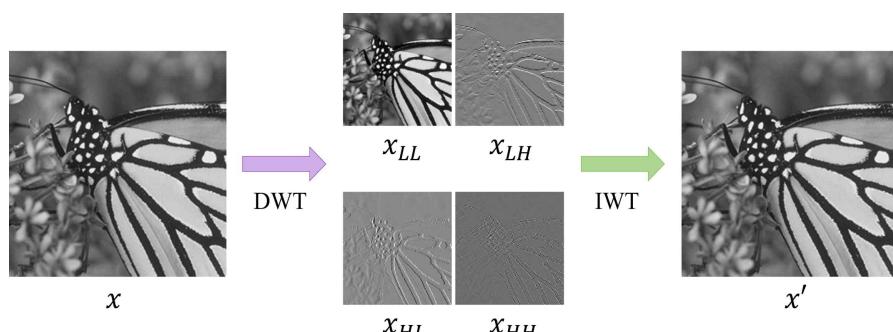


그림 2. 2D 영상에서 DWT와 IWT의 변환 과정 (Haar 웨이블릿)  
Fig. 2. Conversion process of DWT and IWT in 2D image (Haar Wavelet)

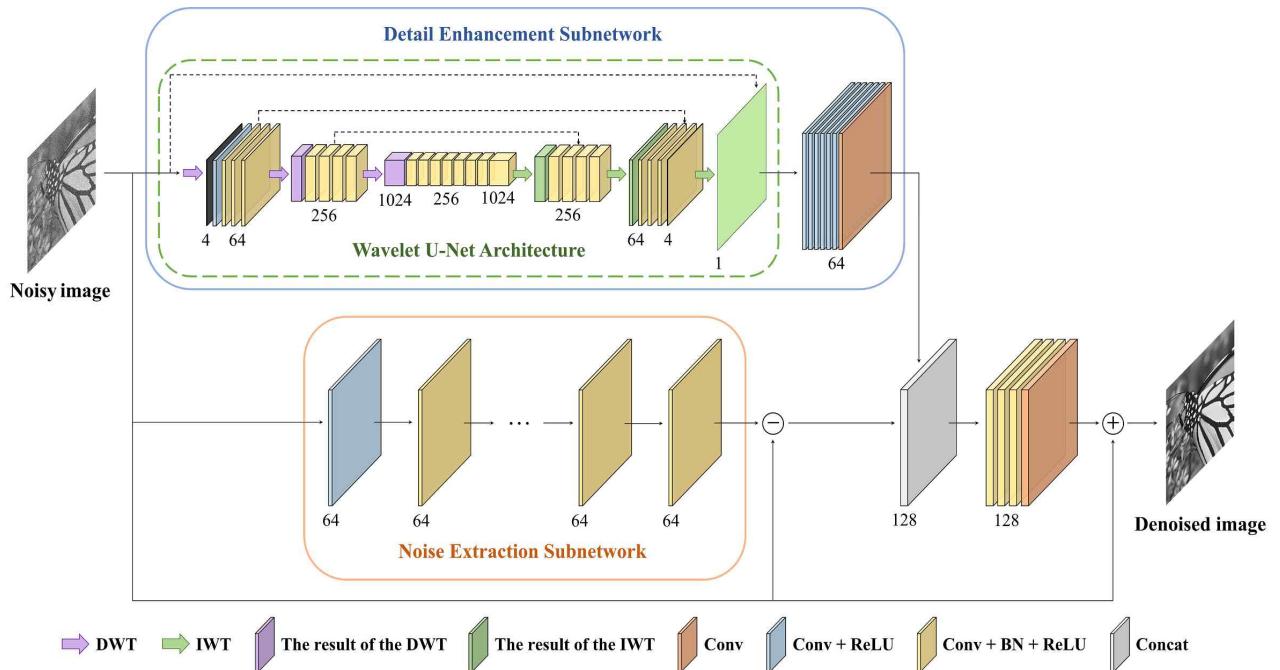


그림 3. 제안하는 WDENet 모델의 전체 네트워크 구조  
 Fig. 3. Overall architecture of the proposed WDENet network

잡음을 추출하는 서브 네트워크로 나눌 수 있다.

WDENet 모델의 전체 네트워크 구조에서 사용하는 활성화 함수 (Activation Function)는 정류된 선형 유닛 (Rectified Linear Unit, ReLU) 함수로 구성되어 있다. 그리고 합성곱 필터를 적용하기 전 패딩 (Padding)으로 데이터를 채워주고 풀링 층 (Pooling Layer)은 존재하지 않는다. 그 이유는 딥러닝 기반의 영상 잡음 제거에서 합성곱 신경망 네트워크로 추출한 영상 특징이 압축으로 인해 잃을 수 있기 때문이다.

WDENet 모델에서는 먼저 입력 영상인 잡음 영상으로부터 다중 경로 학습을 통해 두 개의 서브 네트워크를 거친다. 디테일 향상 서브 네트워크는 웨이블릿 변환과 U-Net 구조를 사용하여 영상 내 디테일을 강조시키는 역할로 디테일 맵 (Detail Map)을 추출하고, 영상 잡음을 추출 서브 네트워크는 DnCNN 모델과 비슷한 구조로 잡음을 추출하여 디테일이 손실된 잡음 제거 영상 피처 맵을 추출한다. 다음으로 디테일 정보 특징과 잡음 정보를 동시 파악하기 위해 서로 다른 관점에서 추출한 두 결과를 연결 (Concatenation)한다. 마지막으로 세 번의 깊이 (Depth)가 128인  $3 \times 3$  합성곱, 배치 정규화와 ReLU의 결합 구조와 한 번의 깊이가 128인

$3 \times 3$  합성곱을 사용하고, 깊은 신경망 네트워크에서도 안정적인 학습 정확도를 위해 잔차 학습을 추가해 최종적으로 디테일이 향상된 잡음 제거 영상을 얻을 수 있도록 한다.

## 2. 영상 디테일 향상 서브 네트워크

제안하는 WDENet 모델에서 영상 디테일 향상 서브 네트워크는 딥러닝 기반의 영상 잡음 제거 과정에서 디테일 한 부분이 손실되는 문제를 개선하기 위해 사용되고 웨이블릿 변환과 U-Net 구조를 사용한다.

웨이블릿 변환은 주파수 영역에서 잡음과 원본 영상을 분리하여 디테일 정보를 학습하는 데에 효과적이고 DWT는 가역성을 가지고 있기 때문에 영상 내의 모든 정보를 손실 없이 복원할 수 있는 특성이 있다. U-Net 구조는 생체 의학 분야에서 영상 분할 (Image Segmentation)을 목적으로 제안된 구조로 경계선 검출에 예민하다. 이러한 장점 때문에 기준보다 우수한 디테일 추출을 위해 U-Net 구조에 풀링 층 대신 웨이블릿 변환을 포함하여 WDENet 모델에 사용한다. 즉, 그림 4처럼 U-Net 구조의 인코더 (Encoder)

모듈에서 DWT로 다운 샘플링 (Down-sampling)하여 해상도가 절반으로 줄어든 결과를 연결하고, 디코더 (Decoder) 모듈에서 IWT로 업 샘플링 (Up-sampling)하여 영상의 원래 크기로 점차 복원하도록 해상도를 두 배씩 늘어나게 한다. WDENet 모델의 U-Net 구조는 6개의 합성곱 블록으로 이루어진다. 각 블록에서는 4번의  $3 \times 3$  합성곱 계층을 사용하고 첫 번째 블록의 깊이는 64이고, 두 번째와 세 번째 블록의 깊이는 256으로 설정하고 나머지 블록의 깊이는 대응되게 설정해 준다. 추가로 네 번째 블록의 마지막 합성곱 깊이는 업 샘플링할 때 인코더 모듈과 동일한 깊이로 복원할 수 있도록 1024로 설정해 준다. 그리고 학습을 돋기 위해 블록마다 배치 정규화와 ReLU를 추가한다. 마지막으로는 그림 3처럼 웨이블릿 U-Net 구조로 얻어진 1채널(컬러 영상인 경우 3채널) 디테일 정보 영상으로부터 6번의 깊이가 64인  $3 \times 3$  합성곱과 ReLU의 결합 구조와 한 번의 깊이가 64인  $3 \times 3$  합성곱 계층을 통해 디테일 맵이 추출된다.

따라서 기존 U-Net 구조가 인코더 모듈에서 Max 풀링 연산, 디코더 모듈에서 전치 합성곱 (Transposed Convolution)을 사용하는 반면 제안하는 WDENet 모델은 DWT로 다운 샘플링하고, IWT로 업 샘플링하여 디테일 추출에 적

극적인 효과를 주었고 웨이블릿 변환으로 다운 샘플링을 수행하면서 피쳐 맵 깊이가 증가하여 더 많은 특징을 추출하는 데 유리한 장점이 있다.

### 3. 영상 잡음 추출 서브 네트워크

영상 잡음 추출 서브 네트워크는 DnCNN 모델과 유사하게 17번의 깊이가 64인  $3 \times 3$  합성곱 계층으로 구성되고 첫 번째 계층은 합성곱과 ReLU의 결합 구조이고 나머지 계층은 합성곱, 배치 정규화와 ReLU의 결합 구조를 사용하여 영상 디테일 향상 서브 네트워크에서 추출한 결과와 같은 깊이를 얻게 된다. 여기서는 잡음과 영상 내 객체 위주의 피쳐 맵을 추출한다.

## IV. 실험 결과 및 분석

### 1. 실험 환경 및 학습 (Training)을 위한 세팅

본 논문의 실험은 Python 3.7 기반의 Pytorch로 모델을

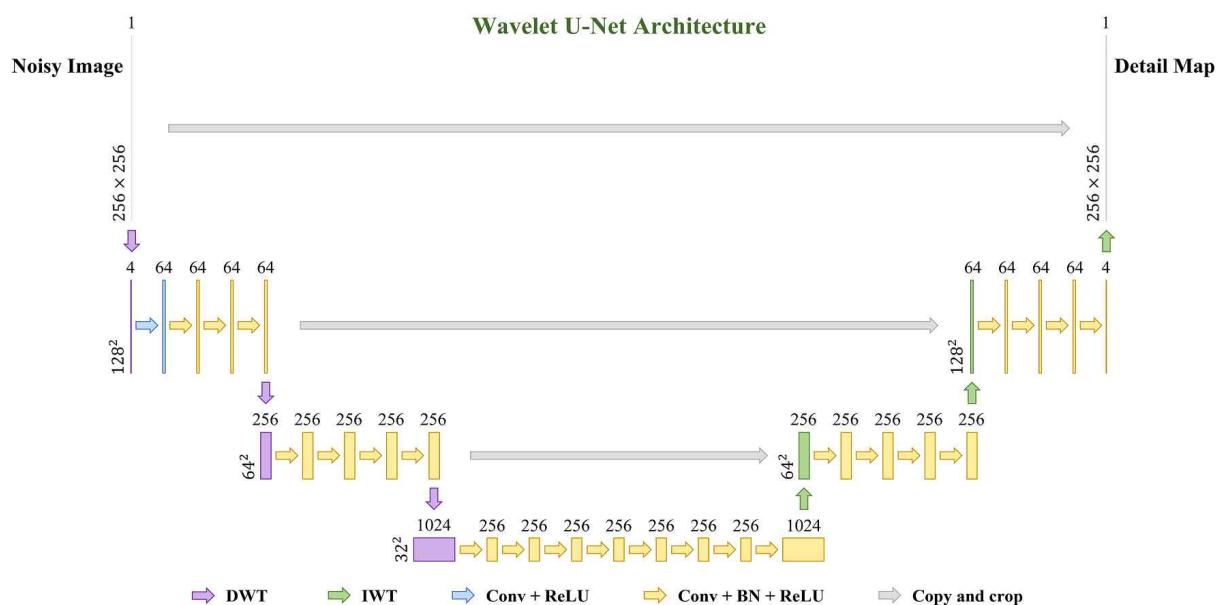


그림 4. 웨이블릿 U-Net 구조

Fig. 4. Architecture of the Wavelet U-Net network

구현하였고 CPU는 Inter(R) Core(TM) i7-4790 CPU 3.60GHz, GPU는 GeForce GTX 1080 Ti를 사용하여 성능을 평가했다. AWGN 합성 잡음에 대해서 패치 크기 (Patch Size)와 배치 크기 (Batch Size)는 각각  $96 \times 96$ , 16으로 설정하여 사용했고, 실제 카메라 잡음에 대해서 패치 크기와 배치 크기는 각각  $128 \times 128$ , 12로 설정하여 사용했다. 에폭 (Epoch)은 55로 설정했고, Loss함수로는 제곱 오차 함수 (Mean Squared Error, MSE)를 이용하였고, 옵티마이저 (Optimizer)로는 아담 옵티마이저 (Adaptive Moment Estimation-optimizer, Adam-optimizer)를 사용하였다. 옵티마이저에 사용하는 학습률 (Learning Rate)은 0.0001로 시작하고 30에폭 이후에는 0.00001로 감소한 학습률을 사용하여 학습시킨다. 실험 결과는 객관적인 수치인 PSNR과 SSIM 결과와 주관적인 품질 및 디테일을 관찰하여 기존 방법과 비교한다.

## 2. 가우시안 잡음 제거 학습과 성능 평가에 사용된 데이터셋 (Dataset)

AWGN 합성 잡음에 대해서 학습에 필요한 데이터셋으로 영상 복구 (Image Restoration) 분야에서 널리 쓰이고 있는 DIV2K<sup>[17]</sup> 데이터셋을 사용하고 입력 영상인 잡음 영상은 깨끗한 영상에 AWGN을 추가해 주었다. 그리고 학습된 모델의 성능을 평가하기 위해 테스트 셋 (Test Set)은 테스트 단계에서 많이 사용되는 Set12 데이터셋과 68개의 자연 이미지로 이루어진 Berkeley 분할 (Segmentation) 데이터셋 BSD68<sup>[18]</sup>을 사용했다.

## 3. 실제 카메라 잡음 제거 학습과 성능 평가에 사용된 데이터셋

실제 카메라 잡음에 대해서 SIDD Dataset의 sRGB 영상을 사용했다. SIDD는 10개의 장면 (Scene)에 대해서 서로 다른 5개의 스마트폰 카메라로 다른 조명 및 ISO Level에서 촬영하여 다양한 실제 잡음 영상을 제공하고 있다.

학습에 필요한 데이터셋으로 SIDD-Medium Dataset을 사용했고 SIDD-Medium Dataset은 320쌍의 실제 카메라 잡음 영상과 이에 대응하는 잡음이 거의 없는 영상을 포함하고

있다. 그리고 검증용 (Validation) 데이터 셋도  $256 \times 256$ 의 패치 크기로 1280쌍이 제공되고 있다.

## 4. 가우시안 잡음 제거 모델

가우시안 잡음 제거 모델에서는 기존에 AWGN 합성 잡음에 대해서 좋은 성능을 보여준 방법과 비교한다. 예를 들어 전통적인 방법인 BM3D<sup>[2]</sup>과 딥러닝 기반의 알고리듬인 DnCNN<sup>[9]</sup>, FFDNet<sup>[10]</sup>, IRCNN<sup>[11]</sup>, ADNet<sup>[19]</sup>가 있다. 추가로 제안하는 모델에서 기존 U-Net 구조의 업샘플링과 다운샘플링을 사용하는 모델의 성능도 같이 확인해 본다. 즉, Max 폴링 연산으로 업샘플링하고 전치 합성곱으로 다운샘플링하는 것이고 모델 이름을 WDENet-Pool로 설정한다. 모든 알고리듬에 대해서는 각각 학습을 진행했고 알고리듬마다 AWGN 잡음 레벨을 15, 25, 50으로 나눠서 학습했다.

표 1은 잡음 레벨이 15, 25, 50인 경우 기존 방법과 제안한 방법의 잡음 제거 성능을 객관적 수치인 PSNR과 SSIM 결과로 보여준다. 제안하는 방법은 데이터셋, 잡음 레벨과 관계없이 모든 상황에서 기존 알고리듬 및 네트워크에 비해 높은 수치를 보여주고 있다. 특히 잡음 레벨이 15인 경우 최고의 PSNR과 SSIM 결과를 달성하였고 잡음 레벨이 커질수록 PSNR과 SSIM 결과 값이 줄어드는 것을 알 수 있다. 이는 잡음이 심해짐에 따라 영상 정보가 점차 손실되고 이런 정보가 손실된 영상을 복원하기 어렵기 때문이다.

Set12 데이터셋의 경우 제안한 WDENet은 15의 잡음 레벨에서 가장 좋은 결과를 얻은 기존 방법보다 PSNR 결과가 0.08만큼 향상하고, 25의 잡음 레벨에서 가장 좋은 결과를 얻은 기존 방법보다 0.16만큼 향상하며 50의 잡음 레벨에서 가장 좋은 결과를 얻은 기존 방법보다 0.26만큼 더 높은 수치를 보였다. BSD68 데이터셋의 경우 15의 잡음 레벨에서 가장 좋은 결과를 얻은 기존 방법보다 0.07만큼 향상하고, 25의 잡음 레벨에서 가장 좋은 결과를 얻은 기존 방법보다 0.13만큼 향상하며 50의 잡음 레벨에서 가장 좋은 결과를 얻은 기존 방법보다 0.13만큼 더 높은 수치를 보였다. 따라서 위의 결과를 통해 제안하는 방법은 잡음 레벨의 영향을 받아 낮은 잡음 레벨보다 높은 잡음 레벨에 대해 더 나은 성능을 발휘하는 것을 알 수 있다.

표 1. 다양한 잡음 레벨에 따라 기존 방법과 제안한 방법에 대한 회색 영상 잡음 제거의 평균 PSNR(dB)/SSIM 결과 비교(기준 방법 중 가장 좋은 결과를 초록색으로, 제안한 방법은 빨간색으로, 두 결과의 차이를 파란색으로 강조하여 표시한다.)

Table 1. The average PSNR(dB)/SSIM results of gray image denoising for conventional methods and the proposed method with different noise levels(The best result among conventional methods is highlighted in green, the proposed method is highlighted in red, and the difference between the two results is highlighted in blue.)

Methods	Set12			BSD68		
	Noise Level			Noise Level		
	15	25	50	15	25	50
Noisy	24.60 / 0.5008	20.17 / 0.3190	14.15 / 0.1421	24.61 / 0.5369	20.17 / 0.3505	14.15 / 0.1582
BM3D <sup>[2]</sup>	32.40 / 0.8952	30.00 / 0.8507	26.74 / 0.7677	31.10 / 0.8723	28.51 / 0.8015	25.65 / 0.6870
IRCNN <sup>[11]</sup>	32.73 / 0.9004	30.36 / 0.8598	27.08 / 0.7792	31.63 / 0.8880	29.14 / 0.8250	26.18 / 0.7165
DnCNN <sup>[9]</sup>	32.82 / 0.9019	30.38 / 0.8611	27.12 / 0.7812	31.72 / 0.8900	29.21 / 0.8281	26.23 / 0.7187
FFDNet <sup>[10]</sup>	32.73 / 0.9020	30.42 / 0.8632	27.29 / 0.7892	31.63 / 0.8897	29.19 / 0.8289	26.29 / 0.7239
ADNet <sup>[19]</sup>	32.96 / 0.8997	30.55 / 0.8636	27.37 / 0.8003	31.74 / 0.8834	29.24 / 0.8248	26.29 / 0.7338
WDENet-Pool	32.99 / 0.9019	30.65 / 0.8666	27.55 / 0.8080	31.79 / 0.8901	29.35 / 0.8299	26.35 / 0.7404
WDENet	33.04 / 0.9022	30.71 / 0.8679	27.63 / 0.8103	31.81 / 0.8905	29.37 / 0.8312	26.42 / 0.7423
Difference	0.08 / 0.0002	0.16 / 0.0043	0.26 / 0.0100	0.07 / 0.0005	0.13 / 0.0023	0.13 / 0.0085

추가로 기존 U-Net 구조를 사용하는 WDENet-Pool에 대해서도 데이터 셋, 잡음 레벨과 관계없이 모든 상황에서 기존 방법에 비해 높은 PSNR과 SSIM 결과 수치를 보여주면서 U-Net 구조와 다중 경로 학습의 성능을 입증했다. 그리고 웨이블릿 U-Net 구조를 사용하는 WDENet은 WDENet-Pool보다 항상 PSNR과 SSIM 결과가 높았고 이를 통해 웨이블릿 변환으로 U-Net의 업 샘플링과 다운 샘플링을 대체하는 것이 성능 향상에 효과적인 것을 알 수 있다.

그림 5는 잡음 레벨이 15인 경우 Set12 데이터 셋의 “Lena” 영상으로 기존 잡음 제거 방법과 제안한 방법을 비교하기 위한 결과 영상이다. 영상 중 모자 깃털과 머리카락 부분 및 모자 라인 부분을 비교했을 때 제안하는 WDENet

모델로 형성한 그림 5 (g)와 WDENet-Pool 모델로 형성한 그림 5 (f)는 기존 알고리듬에 비해 얇은 선들이 가장 잘 표현되었고 그림 5 (g)는 그림 5 (f)에 비해 머리 결과 모자 접힌 부분이 조금 선명한 것을 확인할 수 있다.

그림 6은 잡음 레벨이 25인 경우 Set12 데이터 셋의 “Barbara” 영상으로 기존 잡음 제거 방법과 제안한 방법을 비교하기 위한 결과 영상이다. 영상 중 목도리 부분 및 테이블 보 부분을 비교하였을 때 제안하는 WDENet 모델로 형성한 그림 6 (g)와 WDENet-Pool 모델로 형성한 그림 6 (f)는 기존 알고리듬에 비해 잡음이 가장 깨끗하게 제거됐고 목도리의 구겨짐과 테이블 보의 라인 패턴이 원본 영상인 그림 6 (a)에 가장 가깝게 표현되었다. 그리고 그림 6 (g)는

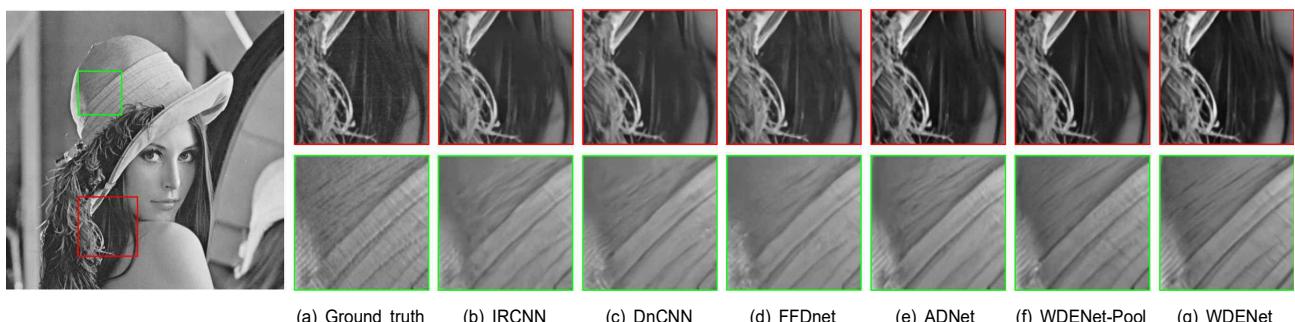


그림 5. 가우시안 잡음이 제거된 결과 화질 비교 (Set12,  $\sigma = 15$ )

Fig. 5. Gaussian denoising results of the grayscale image “Lena” from Set12 with noise level 15

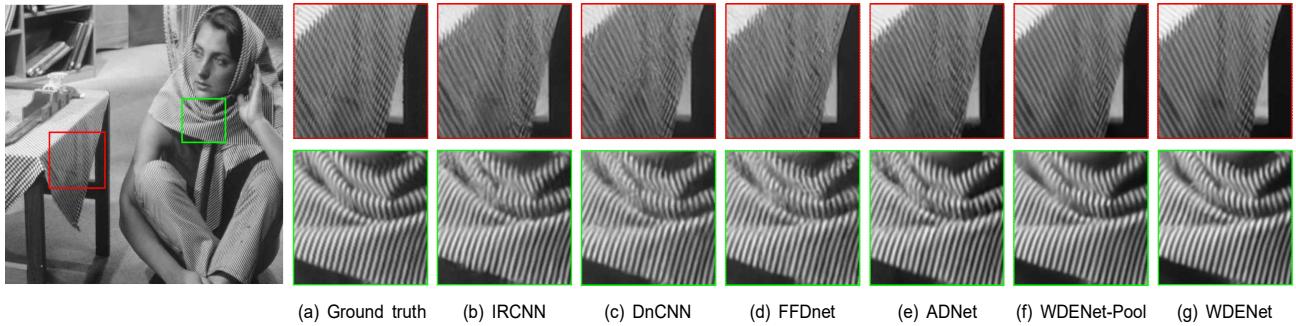


그림 6. 가우시안 잡음이 제거된 결과 화질 비교 (Set12,  $\sigma = 25$ )

Fig. 6. Gaussian denoising results of the grayscale image "Barbara" from Set12 with noise level 25

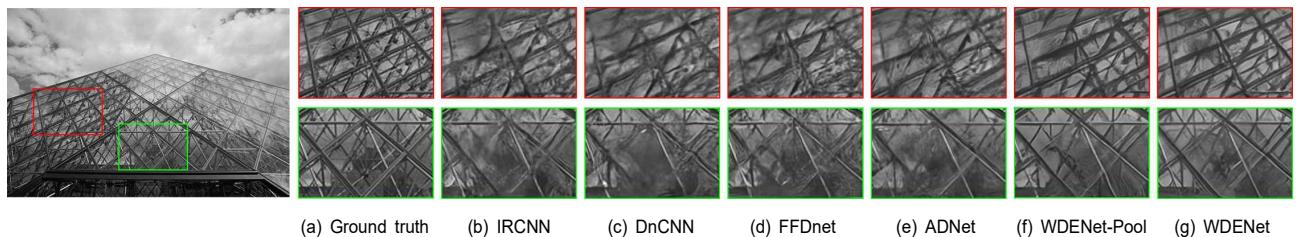


그림 7. 가우시안 잡음이 제거된 결과 화질 비교 (BSD68,  $\sigma = 50$ )

Fig. 7. Gaussian denoising results of the grayscale image "Test044" from BSD68 with noise level 50

그림 6 (f)보다 라인 패턴과 구거진 부분의 안쪽에 있는 디테일 정보가 향상되었고 웨이블릿 U-Net 구조가 경계선 보존에 효과적인 것을 확인할 수 있었다.

그림 7은 잡음 레벨이 50인 경우 BSD68 데이터 셋으로 기존 잡음 제거 방법과 제안한 방법을 비교하기 위한 결과 영상이다. 영상 내 복잡한 경계선들이 모여 있는 부분을 비교해 봤을 때 기존 알고리듬들의 결과 영상은 경계선들이 모두 뭉개졌지만 제안하는 WDENet 모델과 WDENet-Pool 모델은 잡음 레벨이 50임에도 불구하고 제안한 모델과 WDENet-Pool 모델로 얻은 그림 7 (g)와 그림 7 (f)는 기존 알고리듬에 비해 디테일 정보의 보존이 눈에 띄게 향상되었고, 블러 (Blur) 현상이 적어 영상을 디테일하게 복원할 수 있었다. 그리고 그림 7 (f)는 그림 7 (g)에 비해 뭉개진 부분이 있고 경계선들이 비뚤어지는 경향이 있어 웨이블릿 U-Net 구조의 디테일 보존 능력을 확인할 수 있었다.

## 5. 실행 시간

실행 시간의 경우 표 2에 표시된 것과 같이 AWGN 잡음

표 2. AWGN 잡음 레벨이 25인 경우 그레이 잡음 영상의 크기가  $256 \times 256$ ,  $512 \times 512$ 와  $1024 \times 1024$ 에 따라 다양한 방법의 실행 시간(초) 비교  
 Tabel 2. Run time (in seconds) of different methods on grayscale images of size  $256 \times 256$ ,  $512 \times 512$  and  $1024 \times 1024$  with noise level 25

Methods	Device	$256 \times 256$	$512 \times 512$	$1024 \times 1024$
BM3D <sup>[2]</sup>	CPU	0.63	2.61	11.33
IRCNN <sup>[11]</sup>	GPU	0.00787	0.03312	0.10907
DnCNN <sup>[9]</sup>	GPU	0.01183	0.03766	0.12675
FFDNet <sup>[10]</sup>	GPU	0.00457	0.00948	0.03528
ADNet <sup>[19]</sup>	GPU	0.03932	0.11028	0.44896
WDENet	GPU	0.03437	0.10598	0.40851

레벨이 25인 경우 그레이 잡음 영상의 크기가  $256 \times 256$ ,  $512 \times 512$ 와  $1024 \times 1024$ 인 잡음 영상에 대해서 각 잡음 제거 모델의 실행 시간을 실험한다. 비교 대상은 제안하는 방법 외에 5개의 대표적인 잡음 제거 알고리듬으로 선택했다. 제안하는 방법은 최근에 제안된 ADNet에 비해 PSNR과 SSIM 값이 모두 더 좋은 결과를 얻었고 실행 시간도 단축시켰다. 그리고 제안하는 방법은 기존의 일반적인 딥

러닝 알고리듬에 비해 실행 시간이 조금 더 걸리지만 적당한 실행 시간을 갖추었고 CPU로 결과 영상을 계산하는 BM3D에 비해 영상 크기가  $256 \times 256$  기준 약 18배 정도 시간을 단축시켰고, 영상 크기가  $512 \times 512$  기준 약 25배 정도 시간을 단축시켰으며, 영상 크기가  $1024 \times 1024$  기준 약 28배 정도 시간을 단축시켰다. 따라서 이런 정량 및 정성 분석 측면에서 제안하는 모델은 잡음 제거 기술에 경쟁력이 크다는 것을 알 수 있다.

## 6. 실제 카메라 잡음 제거 모델

실제 카메라 잡음 제거 모델에서는 잡음 레벨을 정하지 않아도 잡음 제거가 가능한 AWGN 블라인드 (Blind) 모델, 그리고 최근에 실제 카메라 잡음에 대해서 좋은 성능을 보여준 방법과 비교한다. 예를 들어 DnCNN의 블라인드 모델인 DnCNN-B<sup>[9]</sup>와 최근 실제 카메라 잡음에 우수한 성능을 보여준 CBDNet<sup>[20]</sup>, VDN<sup>[21]</sup>, AINDNet<sup>[22]</sup>가 있다. 모든 알고리듬에 대해서는 각각 학습을 진행하여 실험했고 SIDD 벤치마크 (Benchmark) 데이터 셋으로 성능 테스트했다.

표 3은 SIDD 벤치마크 데이터 셋을 이용하여 다양한 방법의 sRGB 영상 잡음 제거 성능을 객관적 수치인 PSNR과 SSIM 결과로 보여준다. 제안하는 방법은 기존의 모든 방법보다 더 좋은 수치를 얻었다. 하지만 DnCNN-B와 CBDNet는 별로 안 좋은 성능을 보여주고 있다. 이는 이들이 SIDD 가 아닌 다른 데이터 셋으로 학습시켰기 때문이다. 따라서 더욱 공정하고 정확한 성능 차이를 판단하기 위해 DnCNN-B와 CBDNet에 대해 재학습을 진행했고 DnCNN-B\*와 CBDNet\*로 표시한다. 그 결과 제안하는 방법은 DnCNN-B와 CBDNet\*의 객관적 수치 차이 값은 원래의 PSNR 13.27, SSIM 0.5276에서 PSNR 1.15, SSIM 0.0592로 향상하였고 CBDNet\*의 객관적 수치 차이 값은 원래의 PSNR 5.46, SSIM 0.1695에서 PSNR 0.88, SSIM 0.0565로 향상하였으며 제안한 WDENet는 여전히 DnCNN-B와 CBDNet보다 뛰어난 성능을 보여줬다. 그리고 제안한 WDENet은 가장 좋은 결과를 얻은 기존 방법보다 PSNR 결과가 0.20만큼 향상하였고, SSIM 결과가 0.0516만큼 향상하였으며 제안한 모델의 높은 영상 복원 성능을 확인할 수 있었다.

표 3. SIDD 벤치마크 데이터 셋으로 기존 방법과 제안한 방법에 대한 sRGB 영상 잡음 제거의 평균 PSNR(dB)/SSIM 결과 비교(기존 방법 중 가장 좋은 결과를 초록색으로, 제안한 방법은 빨간색으로, 두 결과의 차이를 파란색으로 강조하여 표시한다.)

Table 3. The average PSNR(dB)/SSIM results of sRGB image denoising for conventional methods and proposed method with SIDD benchmark dataset(The best result among conventional methods is highlighted in green, the proposed method is highlighted in red, and the difference between the two results is highlighted in blue.)

Methods	PSNR	SSIM
Noisy	23.66	0.3335
DnCNN-B <sup>[9]</sup>	26.21	0.4407
DnCNN-B*	38.33	0.9091
CBDNet <sup>[20]</sup>	34.02	0.7988
CBDNet*	38.60	0.9118
AINDNet <sup>[22]</sup>	39.02	0.9121
VDN <sup>[21]</sup>	39.28	0.9167
WDENet	39.48	0.9683
Difference	+0.20	+0.0516

그림 8은 SIDD 벤치마크 데이터 셋의 11번 예에 대해 기존의 실제 잡음 제거 방법과 제안한 방법의 실제 잡음 제거 성능을 비교하기 위한 결과 영상이다. 영상 중 DnCNN-B 모델로 형성한 그림 8 (b)는 표 3에서 보여준 결과처럼 잡음이 거의 제거가 안 됐고 CBDNet 모델로 형성한 그림 8 (c)는 실제 카메라 잡음이 어느 정도 제거됐지만, 여전히 심한 잡음들이 남아 있다는 것을 확인할 수 있다. 그리고 재학습 된 DnCNN-B\*와 CBDNet\* 모델과 기존의 실제 잡음 제거 방법과 비교해보면 제안하는 WDENet 모델은 잡음을 가장 깨끗하게 제거된 것을 확인할 수 있다. 게다가 SSID에 사용하는 잡음 제거 목표 (Target)영상인 GT (Ground truth) 영상에는 어느 정도 잡음이 남아 있지만 제안하는 WDENet 모델은 그림 8 (i)에 남아 있는 잡음도 제거하여 뛰어난 잡음 제거 성능을 보여줬다.

그림 9는 SIDD 벤치마크 데이터 �셋의 36번 예에 대해 기존의 실제 잡음 제거 방법과 제안한 방법의 실제 잡음 제거 성능을 비교하기 위한 결과 영상이다. 영상 중 DnCNN-B 모델로 형성한 그림 9 (b)는 그림 8처럼 실제 카메라 잡음에 대해 큰 효과가 없고 재학습 된 DnCNN-B\*



그림 8. SIDD 벤치마크 데이터 셋의 11번 예에 대해 실제 잡음이 제거된 결과 화질 비교 (sRGB)  
 Fig. 8. Real image denoising results of the example 11 from the SIDD benchmark dataset (sRGB)

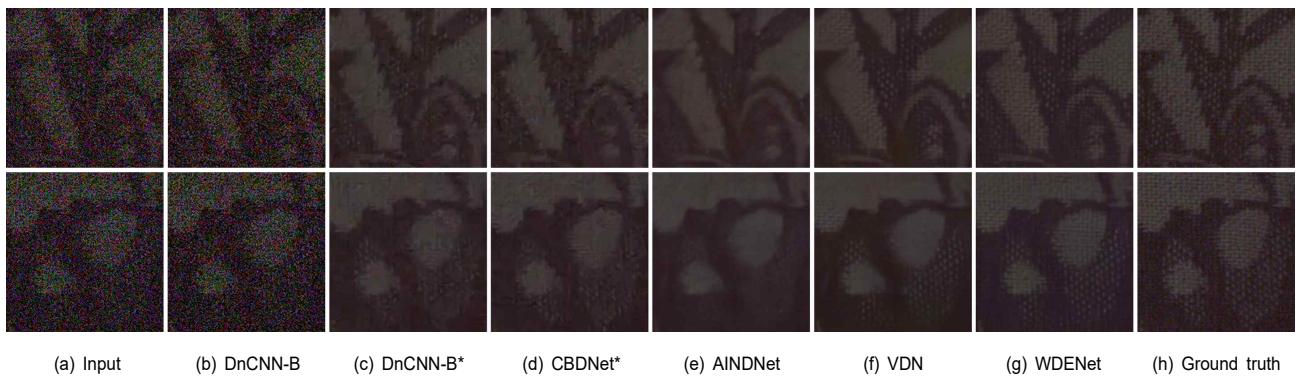


그림 9. SIDD 벤치마크 데이터 셋의 36번 예 실제 잡음이 제거된 결과 화질 비교 (sRGB)  
 Fig. 9. Real image denoising results of the example 36 from the SIDD benchmark dataset (sRGB)

와 CBDNet\* 모델은 심한 잡음은 제거됐지만 작은 잡음이 아직 존재하고 디테일도 손실됐다. 최근에 실제 잡음 제거에 좋은 성능을 보여준 AINDNet 모델과 VDN 모델로 형성한 그림 9 (e)와 그림 9 (f)는 잡음은 제거가 됐지만, 디테일 정보가 손실되는 문제는 아직 해결하지 못했고 그림 9 (f)는 색 정보도 잃어버리는 경우가 존재한다. 하지만 제안하는 WDENet 모델은 잡음을 깨끗하게 제거하는 동시 디테일 정보도 살려주는 특성이 있어 목표 이미지인 그림 9 (h)와 가장 가깝게 복원되었다.

## V. 결 론

본 논문에서는 기존 U-Net 구조에서 풀링 연산으로 다운샘플링하고 전치 합성곱으로 업 샘플링을 하는 대신 이산 웨이블릿 변환과 역 이산 웨이블릿 변환을 U-Net 구조에 포함하여 경계선 특징을 추출하는 웨이블릿 U-Net 구조를 사용한다. 이 구조를 기반으로 영상 잡음 제거 과정에서 손실되는 디테일 정보를 보완하는 합성곱 신경망 네트워크

WDENet 모델을 제안한다. 기존의 단일 합성곱 신경망 네트워크 모델과 달리 제안하는 WDENet 모델은 다중 경로 학습을 사용하여 입력 영상인 잡음 영상을 디테일 향상 서브 네트워크와 잡음 추출 서브 네트워크를 통해 다른 관점에서 얻은 두 결과를 연결하여 같이 학습하는 구조로 잡음과 함께 디테일 정보가 손실되는 문제를 개선할 수 있다. 실험 결과에서 볼 수 있듯이 WDENet 모델은 기존 잡음 제거 알고리듬 및 네트워크보다 객관적 품질 평가 수치인 PSNR과 SSIM 결과와 주관적 품질 비교에서 모두 뛰어난 결과가 나타났고 결과 영상의 디테일 정보가 블러 되는 현상을 완화할 수 있어 디테일 향상에 좋은 성능을 보여주었다. 그리고 WDENet-Pool 모델도 기존의 방법에 비해 우수한 성능을 보여줬지만 WDENet 모델에 웨이블릿 U-Net 구조를 사용함으로써 영상의 디테일 부분과 물체의 경계선에서 잡음이 더 효과적으로 제거된 것을 확인할 수 있었다. 하지만 제안하는 WDENet 모델의 영상 디테일 향상 서브 네트워크에서 추출된 디테일 맵이 입력 영상인 잡음 영상으로부터 추출되기 때문에 원본 영상에서 직접 분리한 것처럼 완벽한 디테일 맵 추출은 아직까지 한

계가 있고 또한 사람 얼굴 부분에 대한 완벽한 복원도 지속적인 연구 진행이 필요하다. 따라서 향후 연구에서는 영상 초 고해상도화 기법 (Super- Resolution), 객체 인식 (Object Detection), 생성적 적대 신경망 (Generative Adversarial Network, GAN) 등의 기법을 접목하여 더욱 정확하고 선명한 잡음 제거 영상이 얻을 수 있게 연구를 진행할 예정이다.

### 참 고 문 헌 (References)

- [1] Buades, B. Coll, and J.M. Morel, “A non-local algorithm for image denoising,” Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, San Diego, CA, USA, pp. 60–65, 2005.
- [2] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering,” IEEE Transactions on Image Processing, Vol. 16, No. 8, pp. 2080–2095, August 2007.
- [3] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, “Non-local Sparse Models for Image Restoration,” IEEE International Conference on Computer Vision, Kyoto, Japan, pp. 2272–2279, 2009.
- [4] W. Dong, L. Zhang, G. Shi, and X. Li, “Nonlocally Centralized Sparse Representation for Image Restoration,” IEEE Transactions on Image Processing, Vol. 22, No. 4, pp. 1620–1630, April 2013.
- [5] S. Gu, L. Zhang, W. Zuo, and X. Feng, “Weighted Nuclear Norm Minimization with Application to Image Denoising,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, pp. 2862–2869, 2014.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 770–778, 2016.
- [7] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” International Conference on Machine Learning, Lille, France, pp. 448–456, 2015.
- [8] V. Nair and G. E. Hinton, “Rectified Linear Units Improve Restricted Boltzmann Machines,” International Conference on Machine Learning, Haifa, Israel, pp. 807–812, 2010.
- [9] 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, Vol. 26, No. 7, pp. 3142–3155, July 2017.
- [10] K. Zhang, W. Zuo, and L. Zhang, “FFDNet: Toward a Fast and Flexible Solution for CNN based Image Denoising,” IEEE Transactions on Image Processing, Vol. 27, No. 9, pp. 4608–4622, September 2018.
- [11] K. Zhang, W. Zuo, S. Gu, and L. Zhang, “Learning Deep CNN Denoiser Prior for Image Restoration,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 3929–3938, 2017.
- [12] O. Ronneberger, P. Fischer, and T. Brox, “Unet: Convolutional Networks for Biomedical Image Segmentation,” Proceeding of Medical Image Computer-Assisted Intervention, Berlin, Heidelberg, pp. 234–241, 2015.
- [13] A. Abdelhamed, S. Lin, and M. S. Brown, “A high-Quality Denoising Dataset for Smartphone Cameras,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 1692–1700, 2018.
- [14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going Deeper With Convolutions,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, pp. 1–9, 2015.
- [15] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-scale Image Recognition,” Proceeding of the International Conference on Learning Representations, San Diego, CA, USA, pp. 1–14, 2015.
- [16] K. H. Talukder and K. Harada, “Haar Wavelet Based Approach for Image Compression and Quality Assessment of Compressed Image,” arXiv preprint arXiv:1010.4084, 2010.
- [17] R. Timofte, E. Agustsson, L. Van Gool, M. H. Yang, and L. Zhang, “Ntire 2017 Challenge on Single Image Super-Resolution: Methods and Results,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, pp. 114–125, 2017.
- [18] S. Roth, and M. J. Black, “Fields of Experts: A Framework for Learning Image Priors,” 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, San Diego, CA, USA, pp. 860–867, 2005.
- [19] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei, and H. Liu, “Attention-guided CNN for Image Denoising,” Neural Networks, Vol. 124, pp. 117–129, April, 2020.
- [20] S. Guo, Z. Yan, K. Zhang, W. Zuo, and L. Zhang, “Toward Convolutional Blind Denoising of Real Photographs,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 1712–1722, 2019.
- [21] Z. Yue, H. Yong, Q. Zhao, L. Zhang, and D. Meng, “Variational Denoising Network: Toward Blind Noise Modeling and Removal,” arXiv preprint arXiv:1908.11314, 2019.
- [22] Y. Kim, J. W. Soh, G. Y. Park, and N. I. Cho, “Transfer Learning from Synthetic to Real-Noise Denoising with Adaptive Instance Normalization,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, pp. 3482–3492, 2020.

---

### 저자 소개

---

#### 정 군



- 2020년 7월 : 대련민족대학교 통신학과 학사
- 2020년 8월 : 한림대학교 스마트 IoT학과 학사
- 2020년 9월 ~ 현재 : 한양대학교 융합전자공학과 석사과정
- ORCID : <https://orcid.org/0000-0003-1642-8880>
- 주관심분야 : 영상처리, 딥 러닝

#### 위 승 우



- 2015년 2월 : 한양대학교 전자통신공학과 학사
- 2015년 9월 ~ 현재 : 한양대학교 융합전자공학과 석박사통합과정
- ORCID : <https://orcid.org/0000-0002-2182-2746>
- 주관심분야 : 영상처리, 딥 러닝

#### 정 제 창



- 1980년 2월 : 서울대학교 전자공학과 학사
- 1982년 2월 : KAIST 전기전자공학과 석사
- 1990년 : 미국 미시간대학 전기공학과 공학박사
- 1980년 ~ 1986년 : KBS 기술연구 소선임연구원 (디지털 및 뉴미디어 연구)
- 1990년 ~ 1991년 : 미국 미시간대학 전자컴퓨터공학부 연구교수 (영상 및 신호처리 연구)
- 1991년 ~ 1995년 : 삼성전자 HDTV 연구개발 담당 수석연구원
- 1995년 ~ 현재 : 한양대학교 융합전자공학부 교수 (영상통신 및 신호처리 연구실)
- ORCID : <https://orcid.org/0000-0002-3759-3116>
- 주관심분야 : 영상처리, 영상압축