

## An Efficient CT Image Denoising using WT-GAN Model

Hae Chan Jeong\*, Dong Hoon Lim\*\*

\*Undergraduate graduate, Dept. of Information and Statistics, Gyeongsang National University, Jinju, Korea

\*\*Professor, Dept. of Information and Statistics, RINS, Gyeongsang National University, Jinju, Korea

## [Abstract]

Reducing the radiation dose during CT scanning can lower the risk of radiation exposure, but not only does the image resolution significantly deteriorate, but the effectiveness of diagnosis is reduced due to the generation of noise. Therefore, noise removal from CT images is a very important and essential processing process in the image restoration. Until now, there are limitations in removing only the noise by separating the noise and the original signal in the image area. In this paper, we aim to effectively remove noise from CT images using the wavelet transform-based GAN model, that is, the WT-GAN model in the frequency domain. The GAN model used here generates images with noise removed through a U-Net structured generator and a PatchGAN structured discriminator. To evaluate the performance of the WT-GAN model proposed in this paper, experiments were conducted on CT images damaged by various noises, namely Gaussian noise, Poisson noise, and speckle noise. As a result of the performance experiment, the WT-GAN model is better than the traditional filter, that is, the BM3D filter, as well as the existing deep learning models, such as DnCNN, CDAE model, and U-Net GAN model, in qualitative and quantitative measures, that is, PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index Measure) showed excellent results.

▶ **Key words:** WT-GAN, deep learning, CT image, noise reduction

## [요 약]

CT 촬영 시 방사선량을 줄이면 피폭 위험성을 낮출 수 있으나, 영상 해상도가 크게 저하 될 뿐 아니라 잡음(noise) 발생으로 인해 진단의 효용성이 떨어진다. 따라서, CT 영상에서의 잡음제거는 영상복원 분야에 있어 매우 중요하고 필수적인 처리 과정이다. 영상 영역에서 잡음과 원래 신호를 분리하여 잡음만을 제거하는 것은 한계가 있다. 본 논문에서는 웨이블릿 변환 기반 GAN 모델 즉, WT-GAN(wavelet transform-based GAN) 모델을 이용하여 CT 영상에서 효과적으로 잡음 제거하고자 한다. 여기서 사용된 GAN 모델은 U-Net 구조의 생성자와 PatchGAN 구조의 판별자를 통해 잡음제거 영상을 생성한다. 본 논문에서 제안된 WT-GAN 모델의 성능 평가를 위해 다양한 잡음, 즉, 가우시안 잡음(Gaussian noise), 포아송 잡음(Poisson noise) 그리고 스펙클 잡음(speckle noise)에 의해 훼손된 CT 영상을 대상으로 실험하였다. 성능 실험 결과, WT-GAN 모델은 전통적인 필터 즉, BM3D 필터뿐만 아니라 기존의 딥러닝 모델인 DnCNN, CDAE 모형 그리고 U-Net GAN 모형보다 정성적이고, 정량적인 척도 즉, PSNR (Peak Signal-to-Noise Ratio) 그리고 SSIM (Structural Similarity Index Measure) 면에서 우수한 결과를 보였다.

▶ **주제어:** WT-GAN, 딥러닝, CT 영상, 잡음제거

- 
- First Author: Hae Chan Jeong, Corresponding Author: Dong Hoon Lim
  - \*Hae Chan Jeong (zz0909zz@naver.com), Dept. of Information and Statistics, Gyeongsang National University
  - \*\*Dong Hoon Lim (dhl@gnu.ac.kr), Dept. of Information Statistics and RINS, Gyeongsang National University
  - Received: 2024. 03. 22, Revised: 2024. 05. 20, Accepted: 2024. 05. 20.

## I. Introduction

CT(Computed Tomography, 컴퓨터 단층촬영) 영상 검사는 X-선을 인체에 투과시켜 그 흡수 차이를 이용하여 인체 내 구조를 단층 영상이나, 3차원 입체 영상으로 재구성하는 영상 검사로서 질병의 조기 진단 및 병변의 이상 유무를 판별하는데 널리 사용되고 있다. CT는 촬영하고자 하는 부위에 방사선을 이용하는 것으로, 반복적인 CT 촬영은 인체의 세포 조직 변형 및 암 유발 등의 위험성을 내재하고 있다. 방사선량이 적은 저선량 CT 촬영으로 CT의 피폭 위험성을 낮출 수 있으나, 영상 해상도가 크게 저하될 뿐 아니라 잡음(noise) 발생으로 인해 진단의 효용성이 떨어진다. 따라서, CT 영상에서의 잡음제거는 영상복원 분야에 있어 매우 중요하고 필수적인 처리 과정이다[1].

CT 영상에 존재하는 잡음에는 가우시안 잡음(Gaussian noise), 포아송 잡음(Poisson noise), 그리고 스펙클 잡음(speckle noise) 등이 있다[2-3].

지금까지 전통적인 잡음제거 방법으로는 비지역적 평균 필터(Non-local Mean Filter)를 사용한 방법들이 주로 사용되었다. BM3D 방법이 대표적인 필터링 방법에 해당된다[2].

최근 딥러닝을 이용한 영상 처리 기술이 급격히 발달하여 영상인식, 물체 인식, 영상복원 등의 분야에서 활발히 연구가 이루어지고 있다. 특히, CNN(Convolution Neural Network) 모델과 AE(AutoEncoder, 오토인코더) 모델이 잡음 제거에 좋은 성능을 보이고 있다[3-4]. CNN 모델 중에서 DnCNN(Denoising Convolution Neural Network)은 잔차 학습(Residual Learning)과 배치 정규화(Batch Normalization)를 결합하여 학습속도를 증가시켜 잡음 제거 성능을 높였고, DAE(Denoising AutoEncoder)는 기존 AE 모델과 마찬가지로 인코더(Encoder)와 디코더(Decoder) 두 개의 신경망으로 구성된다. 잡음 영상에 대해 인코더는 차원 축소를 통해 잡음이 제거되고, 디코더는 차원 확장을 통해 원 영상과 같은 영상을 재구성한다. 또한, CDAE(Convolution Denoising AutoEncoder) 모델은 CNN과 DAE의 결합 형태로 역시 잡음 제거 영역에서 좋은 성능을 보였다[5].

최근 주목받고 있는 생성형 AI 기술 중 하나인 GAN(Generative Adversarial Network, 적대적 생성 신경망)은 잡음 제거 연구에도 좋은 결과를 도출하고 있다. GAN은 두 개의 인공 신경망인 생성자(Generator)와 판별자(Discriminator)로 구성되어 있으며, 학습 과정에서 두 개의 인공 신경망이 서로 경쟁을 한다. GAN은 기술적으로 고해상도 이미지를 생성할 수 없다는 점과 학습이 불

안정하다는 점을 한계점으로 가지고 있다. 이러한 한계점들을 극복하고자 DCGAN[6], WGAN [7], WGAN-GP [8] 등의 GAN 모델이 개발되어 왔다. 그러나 기존 GAN은 생성된 영상이 전역적/지역적으로 일관성(coherence, consistency)이 떨어지는 점이 있었다. 이를 해소하기 위해 GAN의 판별자에 U-Net 구조를 도입하였다[9].

영상 영역에서 잡음과 원래 신호를 분리하여 잡음만을 제거하는 것이 어렵기 때문에, 주파수 영역(frequency domain)에서 잡음을 제거하려는 연구가 이루어져 왔다 [10-11]. 웨이블릿 변환(wavelet transform)은 영상을 다양한 주파수 대역으로 분해하여 분석할 수 있으므로 이를 이용하여 에지(edge) 성분을 보존하면서 잡음 제거하는데 널리 사용되고 있다.

본 논문에서는 웨이블릿 변환기반 GAN 즉, WT-GAN(wavelet transform-based GAN) 모델을 이용하여 CT 영상에서 잡음 제거하고자 한다. 여기서 사용된 GAN 모델은 U-Net 구조의 생성자와 PatchGAN 구조의 판별자를 통해 잡음제거 영상을 생성한다. PatchGAN 판별자는 전체의 영상에서 진위를 판별하는 일반 GAN 구조에 비해 패치 단위 별로 진위를 판별함으로써 적은 파라미터 사용으로 인해 학습 속도가 빠르다는 장점을 갖고 있다.

본 논문에서는 제안된 WT-GAN 모형의 성능을 평가하기 위해 다양한 잡음, 즉, 가우시안 잡음, 포아송 잡음, 그리고 스펙클 잡음 하에서 전통적인 공간 필터인 BM3D 과 기존의 딥러닝의 DnCNN, CDAE 모델과 비교하였다 [12-13]. 성능실험에서 비교적도로서 직접 눈으로 평가하는 정성적인 비교와 PSNR (Peak Signal-to-Noise Ratio)와 SSIM(Structural Similarity Index Measure) 척도를 통한 정량적인 비교를 수행한다[14].

본 논문은 다음과 같이 구성되어 있다. 제 II 장에서는 관련 연구로서 GAN 모델과 이산 웨이블릿 변환(Discrete Wavelet Transform, DWT)에 대해 논의하고, 제 III 장에서는 WT-GAN 모델을 이용한 잡음제거 방법에 대해 논의한다. 제 IV 장에서는 영상 실험을 통해 정성적이고, PSNR과 SSIM 척도에 의한 정량적인 비교를 수행하고 제 V 장에서 결론을 맺는다.

## II. Preliminaries

### 1. GAN

GAN은 2014년 구글의 이안 굿펠로우 (Ian J. Goodfellow) 등이 제안한 비지도 학습 방식의 모델이다

[15]. DCGAN(Deep Convolutional GAN)은 2016년 Radford 등에 의해 고안된 GAN의 확장 모델로서 최근 대부분의 GAN 모델들은 DCGAN 기반으로 개발되고 있다 [16]. DCGAN은 GAN의 생성자와 판별자 구조에 CNN을 적용한 것이다. 좀더 구체적으로 설명하면, DCGAN은 GAN의 생성자와 판별자 구조에서 완전 연결 층 (fully connected layers) 대신 영상 특징을 추출하기 위해 사용되는 컨볼루션 층(convolutional layer)을 사용함으로써 기본 GAN보다 성능이 탁월하다. DCGAN은 Fig. 1에서 보는 것처럼 생성자와 판별자에 각각 역컨볼루션 층 (deconvolutional layer, transposed convolution layer)과 컨볼루션 층을 사용한다.

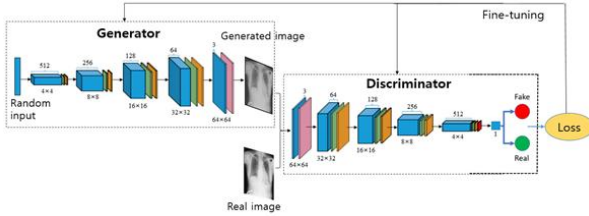


Fig. 1. Architecture of general DCGAN

Fig. 1.에서 생성자는 임의의 노이즈 벡터를 입력받아 여러 번의 업샘플링 (up-sampling)과 역컨볼루션을 반복 사용하여 가짜 영상을 생성한다. 그리고 판별자는 여러 번의 다운샘플링 (down-sampling)과 컨볼루션을 반복 사용하여 생성된 영상이 진짜 영상인지 가짜 영상인지 판별하고 손실함수(loss function)를 계산한다. 학습은 오차 역전파 (Backpropagation)를 통해 생성자가 손실함수를 최소화하도록, 판별자는 손실함수를 최대화하도록 미세조정 (fine-tuning)을 거쳐 진행되며, 최종적으로는 판별자가 진짜 영상과 가짜 영상을 판별하는 능력이 한계에 도달하여 판별 확률이 1/2로 수렴하게 되어 생성자가 진짜와 거의 구별할 수 없는 높은 품질의 가짜 영상을 생성하게 된다.

## 2. DWT

영상에서 사용되는 DWT에서 이산 웨이블릿  $\Psi_{j,k}(t)$ 는 다음과 같이 기본 웨이블릿 기저함수(mother wavelet basis function)  $\psi$ 에 대한 시간 축 방향으로 확대 및 축소, 그리고 평행 이동을 통해 얻어진다.

$$\Psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k)$$

여기서  $j$ 는 스케일 파라미터(scale parameter)이고,  $k$ 는 이동 파라미터(shift parameter)이다. 따라서 웨이블릿 계수  $c(j,k)$ 는 다음과 같이 신호  $f(t)$ 와 웨이블릿

$\Psi_{j,k}(t)$ 들의 선형결합에 의해 얻어진다.

$$c(j,k) = \sum_t f(t) \Psi_{j,k}(t) \quad (2.1)$$

또한, 임의의 신호  $f(t)$ 는 다음과 같이 웨이블릿  $\Psi_{j,k}(t)$ 과 웨이블릿 계수  $c(j,k)$ 들의 선형결합으로 표현된다.

$$f(t) = \sum_k \sum_j c(j,k) \Psi_{j,k}(t) \quad (2.2)$$

DWT에 많이 사용되는 기본 웨이블릿 기저함수에는 Haar와 Daubechies 등이 있다. Haar 기저함수는 불연속 이면서 계단 함수와 유사한 형태를 갖는다. 따라서 Haar 기저함수는 모양이 단순하기 때문에 간단하며, 시간 정보에 대한 해석이 용이하다. Daubechies 함수는 최근 널리 사용되는 기저함수로 연속이면서 계수의 길이가 최소 2 이상이고 따라서 더욱 지역적이고 스무딩한 값을 나타낼 수 있다[17].

DWT는 식 (2.1)과 식 (2.2)에서 보는 것처럼 가역성을 가지고 있기 때문에 영상 내의 모든 정보를 다운 샘플링하고, 역 웨이블릿 변환(Inverse Wavelet Transform, IWT)으로 업샘플링하여 에지 성분에 손실없이 복원할 수 있는 특성을 가지고 있다. Fig. 2는 2차원 영상에서 DWT와 IWT 과정을 보여주고 있다.

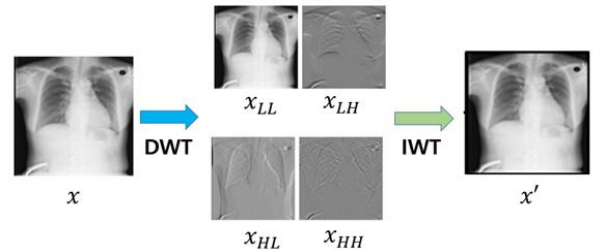


Fig. 2. Transform process of DWT and IWT in 2D image

Fig. 2에서, 영상에 저주파 필터와 고주파 필터를 각각 적용시키고 반으로 다운 샘플링을 수행하면 4개의 서브밴드(sub band) 영상  $x_{LL}$ ,  $x_{LH}$ ,  $x_{HL}$  그리고  $x_{HH}$ 로 나누어지고 4개의 서브 밴드 영상에 IWT를 적용하면 원본 영상으로 복원이 가능하다.

## III. WT-GAN model for Image Denoising

### 1. WT-GAN model Structure

본 논문에서 제안한 WT-GAN 모델의 전반적인 구조를 도식화하면 Fig. 3과 같다.

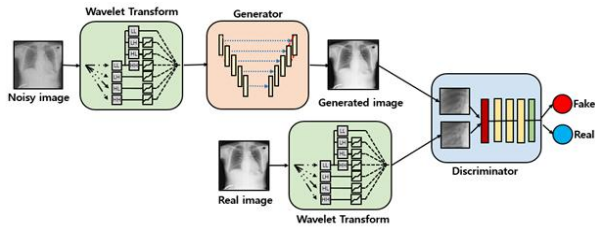


Fig. 3. General architecture of WT-GAN

Fig. 3에서 보면, 잡음 영상에 대해 웨이블릿 변환을 수행한 후 U-Net 구조의 생성자를 적용하여 생성된 가짜 영상을 생성하고, 실제 영상에 대해서도 웨이블릿 변환을 수행한 후, 이 둘 영상에 대해 Patch 기반 판별자에 의해 진짜 영상과 가짜 영상인지 구분한다. 여기서 보통 GAN에 웨이블릿 변환을 적용하는 경우에는 잡음 영상에 대해서만 적용하는데 여기서는 잡음영상과 실제 영상 모두 적용하였다. 그 이유는 실제 영상에 대해서도 잡음이 존재할 수 있기 때문이다.

WT-GAN 모델에서 손실함수는 기존 GAN 손실함수와 L1 손실함수의 가중합 형태로 식 (3.1)과 같다.

$$\min_G \max_D V(D, G) = L_{GAN}(G, D) + \lambda L_1(G), \quad (3.1)$$

여기서  $L_{GAN}(G, D)$ 는 기존 GAN의 손실함수, 여기서는 BCE (Binary Cross Entropy) 손실함수이고,  $L_1(G)$ 는 L1 손실함수를 나타낸다. 여기서  $G$ 와  $D$ 는 각각 생성자와 판별자를 나타낸다. 그리고  $\lambda$ 는 GAN 손실함수와 L1 손실함수 간의 trade-off를 조절하는 하이퍼 파라미터 (hyper-parameter)이다[18].

### 2. U-Net Generator Model

제안된 WT-GAN 모델에서 U-Net 구조의 생성자를 도식화하면 그림 Fig. 4와 같다.

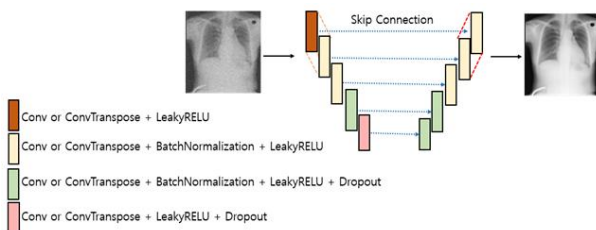


Fig. 4. U-Net based generator structure

생성자에 사용된 U-Net 구조는 영상의 특징 추출을 위한 축소 경로(Contracting Path)와 원래의 영상으로 되돌리는 복원을 위한 확대 경로(Expanding Path)로 이루어져 있으며, FCN (Fully Convolutional Network) 모델을

토대로 축소 경로와 확대 경로 사이에 각층마다 스킵 연결(Skip Connection)이 존재해 시간적인 면에서 상당히 단축되고, 영상의 세부적인 특징과 더불어 전체적인 정보도 간직할 수 있도록 하는 장점이 있다.

본 논문에서 입력 영상의 크기는  $128 \times 128 \times 3$ 이고 생성자의 수축 경로에서 일련의 컨볼루션 블록 즉, [컨볼루션-배치 정규화-Leaky Relu 층] 사용과  $2 \times 2$  맥스 풀링 연산을 통해 다운 샘플링할 때 마다 출력 영상의 크기는 작아지고 채널(channel)의 수는 2배 증가한다. 즉,  $C_k$ 을  $k$ 채널의 수를 갖는 컨볼루션 블록이라 하면 처음 입력 채널의 수가 64개에서 최종 1024 개까지 다음과 같이 2배씩 증가한다.

$$C_{64} \rightarrow C_{128} \rightarrow C_{256} \rightarrow C_{512} \rightarrow C_{1024}$$

여기서 단,  $C_{64}$ 에는 배치 정규화가 적용되지 않는다.

확대 경로에서는 축소 경로에서 추출된 의미 정보를 바탕으로  $2 \times 2$  역 컨볼루션 층을 통해 채널의 수를 절반으로 줄이고, 영상의 크기를 2배로 만든 후에 일련의 컨볼루션 블록 즉, [컨볼루션-배치 정규화-Leaky Relu 층]의 사용을 축소 경로에서의 과정과 동일한 수만큼 진행하여 출력영상이 입력영상과 같은 크기의 영상으로 복원할 수 있게 한다.

### 3. PatchGAN Discriminator Model

PatchGAN 판별자 구조는 영상의 전체 영역이 아니고 특정 크기의 패치 단위로 진위 여부를 판단한다. Fig. 5는 본 논문에서 사용되는 PatchGAN 판별자 구조를 나타내고 있다.

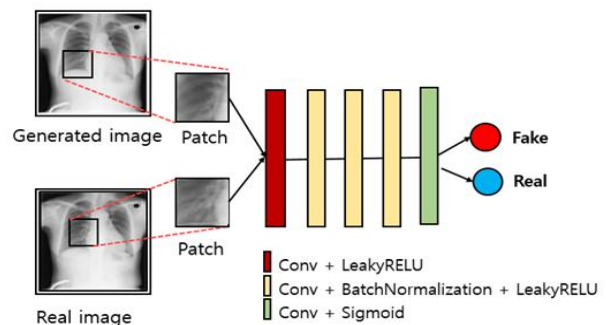


Fig. 5. PatchGAN discriminator structure

기존 DCGAN에서는 영상의 전체 영역에 대해 진위를 판별했다면, PatchGAN에서는  $N \times N$  크기의 패치 단위로 진위 여부를 판별한다. Fig. 5에서는  $N \times N$  크기의 패치에 대해 일련의 모듈 즉, [컨볼루션-Leaky Relu 층] 모듈, 그리고 [컨볼루션-배치 정규화-Leaky Relu 층] 모듈, 마

지막으로 [컨볼루션-시그모이드] 모듈을 사용하여 특징 추출을 바탕으로 진위를 판단한다. PatchGAN에서 패치의 크기는 하나의 픽셀 크기인  $1 \times 1$  패치, 즉 PixelGAN인 경우 적은 파라미터 수에 의해 연산속도가 빠르지만, 너무 작은 특징 맵으로 인해 전체 영상의 특징 파악이 힘들고, 패치의 크기가 영상 전체 영역인 ImageGAN인 경우 많은 파라미터 수에 의해 연산속도가 느리고, 너무 큰 특징 맵으로 인해 영상의 질이 낮아질 수 있다. 따라서 패치의 크기가 작을수록 국부적으로 선명한 결과를 얻을 수 있지만 인공 부산물(artifacts) 발생 가능성이 높고, 패치의 크기가 클수록 비록 인공 부산물 가능성은 낮지만 영상 품질이 다소 떨어질 수 있다. 대부분 연구에서 PatchGAN에서  $70 \times 70$  크기의 패치를 추천하고 있다. PatchGAN 판별자와 일반 GAN 판별자와의 주된 차이점은 일반 GAN 판별자는 전체 영상에 대해 진위를 나타내기 위해  $[0, 1]$  사이 하나의 값을 출력하는 대신에 PatchGAN 판별자는 각각의 패치에 대한 진위를 하나의 값으로 나타내고, 전체 영상에 대해서는 이들 값으로 구성된 배열 값을 출력한다.

## IV. Performance experiment and results

### 1. Performance experiment

본 논문에서 제안된 WT-GAN 모델의 잡음제거 성능을 평가하기 위해 전통적 잡음제거 방법인 BM3D 방법, 기존 딥러닝 모형인 CDAE 모델, DnCNN 그리고 U-Net GAN과 비교하고자 한다.

본 논문에서 제안된 WT-GAN 모델의 웨이블릿 변환은 2단계의 웨이블릿 분해 과정과 복원 과정을 거쳤고, 임곗값 처리는 소프트 임곗값 처리(Soft Thresholding)를 하였고, 기저 웨이블릿 함수로는 Daubechies 4(db4)를 사용하였다[19].

WT-GAN 모델을 학습하는 과정에서 손실함수는 식 (3.1)에서  $\lambda = 100$ 이고 최적화 알고리즘은  $\alpha = 2 \times 10^{-3}$ ,  $\beta_1 = 0.5$ ,  $\beta_2 = 0.999$ 인 Adam을 사용하였다.

실험 환경으로는 NVIDIA GPU인 GeForce RTX 3090을 사용하였으며 파이토치(Pytorch)를 사용하여 구현하였다. 실험에 사용한 데이터 셋은 COVID-19 Chest CT 데이터 셋으로, 총 10,192개의 CT 영상들로 이루어져 있다[20-21]. 전체 영상을 80% : 20% 비율로 8,154개의 학

습용 데이터와 2,038개의 테스트용 데이터로 분할하여 실험하였으며 모든 영상의 픽셀 값은 -1과 1 사이로 정규화하였다.

실험에 사용된 3 가지 잡음 영상, 즉, 가우시안 잡음 영상은 정규 분포(Normal distribution), 포아송 잡음 영상은 포아송 분포(Poisson distribution) 그리고 스펙클 잡음 영상은 레일리 분포(Rayleigh distribution)로부터 얻어진대[22-23]. 여기서는 잡음 비율  $\alpha = 0.1, 0.3, 0.5$ 을 고려한다. Fig. 6은 원 영상과 잡음비율  $\alpha = 0.1$ 인 대표적인 가우시안 잡음 영상, 포아송 잡음 영상 그리고 스펙클 잡음 영상을 나타낸다.

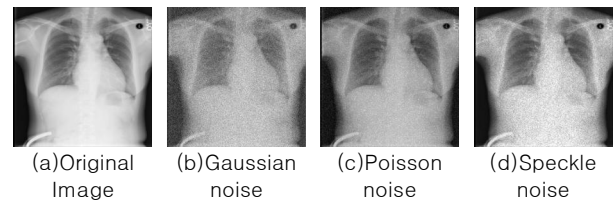
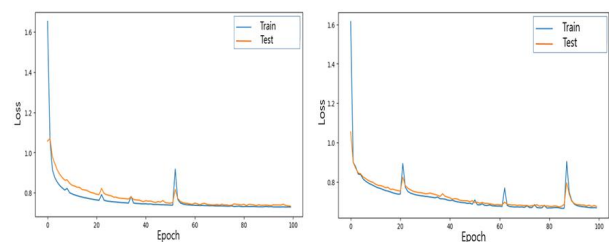


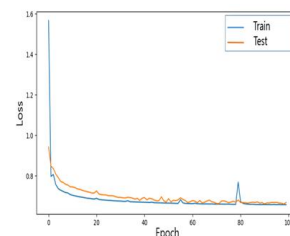
Fig. 6. Original image and noisy images from COVID-19 chest CT dataset

제안된 WT-GAN 모델의 학습 과정을 손실함수 그래프를 통해 확인하고자 한다. 여기서는 에포크(epoch, 학습 반복 수)가 증가함에 따라 손실함수의 변화추이를 살펴보았다.



(a) Gaussian noise

(b) Poisson noise



(c) Speckle noise

Fig. 7. WT-GAN model convergence for noisy images from COVID-19 Radiography Dataset

Fig. 7은 COVID-19 방사선 데이터셋에서 고려되는 3 가지 잡음영상에 대해 에포크에 따른 WT-GAN 모델의 수



럼상태를 보여주고 있다. 여기서 Train 과 Test는 각각 훈련 데이터와 테스트 데이터에서 모델의 손실함수를 나타낸다.

본 실험에서는 정량적인 평가를 위해 두 가지 평가 척도인 PSNR과 SSIM을 사용하였다. PSNR은 원 영상과 잡음 제거 영상의 전체적인 픽셀 값 차이를 수치로 재는 척도이고, SSIM은 PSNR과 다르게 수치적인 차이가 아닌 사람의 눈으로 보았을 때 두 영상 간 시각적인 차이를 평가하는 척도이다. PSNR을 수식으로 표현하면 다음과 같다.

$$\text{PSNR} = 10 \times \log\left(\frac{\text{MAX}^2}{\text{MSE}}\right).$$

위 식에서 MSE란 평균 제곱 오차(Mean Square Error)로 두 영상의 픽셀 값 차이를 의미하며, MAX는 영상의 픽셀 값 중 최댓값을 말한다. MSE가 작을수록 PSNR이 크고 이 값이 클수록 좋은 품질을 나타낸다. 다음으로, SSIM을 수식으로 표현하면 다음과 같다.

$$\text{SSIM}(x, y) = [\ell(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma,$$

여기서  $\ell$ 은 휘도 (Luminance),  $c$ 는 대비 (Contrast),  $s$ 는 구조 (Structure)를 나타내고 이를 함수식으로 표현하면 다음과 같다.

$$\ell(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad c(x, y) = \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2},$$

$$s(x, y) = \frac{2\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3},$$

위 식에서  $\mu_x$ 와  $\mu_y$ 는 두 영상( $x, y$ )의 픽셀값의 평균,  $\sigma_x$ 와  $\sigma_y$ 는 픽셀값의 분산, 그리고  $\sigma_{xy}$ 는 공분산을 뜻한다. 그리고  $\alpha, \beta, \gamma > 0$  이면 휘도, 대비, 구조에 대한 상대적인 중요도를 설정할 수 있으나 보통  $\alpha = \beta = \gamma = 1$ 로 두고,  $C_1 = (K_1L)^2$ ,  $C_2 = (K_2L)^2$ ,  $C_3 = C_2/2$ 이다. 여기서  $K_1 = 0.01$ ,  $K_2 = 0.03$ ,  $L = 255$ 이다[14]. SSIM의 최종 결과는 0 ~ 1 사이이며 1에 가까울수록 두 영상의 차이가 작고 원 영상과 잡음제거 영상이 유사함을 의미한다.

## 2. Experiment results

### 2.1 Qualitative Comparison

여러 잡음 제거 방법들의 성능을 가시적인 방법으로 육안으로 평가하는 정성적인 비교를 수행하고자 한다.

Fig. 8은 COVID-19 Chest CT 데이터 셋에서 여러 잡음 제거 방법들을 Fig. 6(b)의 가우시안 잡음 영상에 적용하여 얻은 결과 영상을 보여주고 있다.

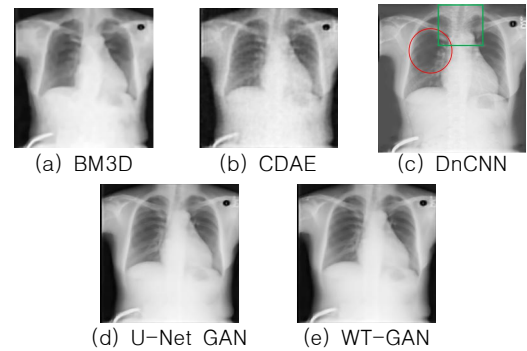


Fig. 8. Experiment results for Gaussian noisy image in Fig. 6(b)

Fig. 8에서 보면, BM3D 모델은 전체적으로 흐릿한 결과 영상을 얻었고, CDAE 모델은 잡음이 덜 제거된 듯 보이거나 형태는 유지하고 있고, DnCNN 모델은 늑골 부근에서 일부 갈비뼈의 형태가 보이지 않음을 알 수 있다. GAN 모델 즉, U-Net GAN과 WT-GAN 모델은 서로 비슷한 결과를 내고 상대적으로 다른 모델과 비교하여 선명한 결과 영상을 얻었다.

Fig. 9는 Fig. 6(c)의 포아송 잡음영상에 여러 잡음제거 방법들을 적용하여 얻은 결과 영상을 나타내고 있다.

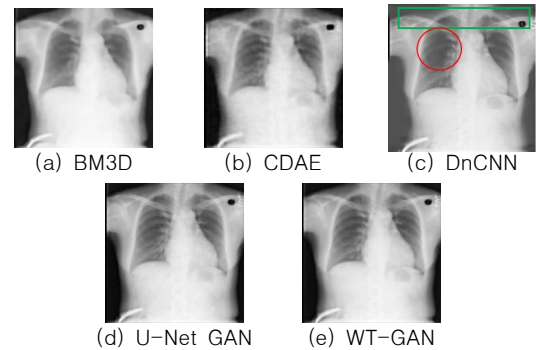


Fig. 9. Experiment results for Poisson noisy image in Fig. 6(c)

Fig. 9에서도 Fig. 8의 결과와 비슷한 결과를 얻었다. BM3D와 CDAE 모델의 결과 영상의 화질이 상대적으로 부영고, DnCNN 모델은 어깨선 부근이 약간 흐릿하고 늑골 부근에서 갈비뼈 형태가 보이지 않음을 확인할 수 있다. 제안된 WT-GAN 모델은 U-Net GAN 모델과 비교하여 늑골 부근에서 약간 미세하게나마 더 선명한 결과 영상을 얻었다.

Fig. 10은 Fig. 6(d)의 스펙클 잡음 영상에 여러 잡음 제거 모델을 적용하여 얻은 결과 영상을 보여주고 있다.

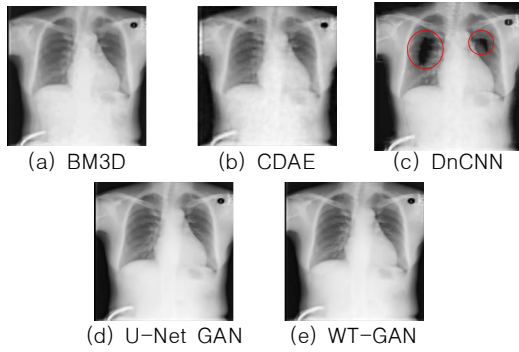


Fig. 10. Experiment results for Speckle noisy image in Fig. 6(d)

Fig. 10에서 보면, CDAE 모델은 잡음이 남아 있는 것처럼 흐린 영상을 보이고, DnCNN 모델은 늑골 부분이 마치 동공이 있는 것처럼 검게 나타났다. 제안된 모델은 U-Net GAN과 같이 여전히 선명한 결과 영상으로 유지함을 알 수 있다.

## 2.2 Quantitative Comparison

모델의 성능 평가를 육안으로 비교하는 정성적인 평가는 한계가 있다. 본 실험에서 모델의 정량적인 평가를 위해 에포크가 100인 경우 PSNR과 SSIM을 계산한다.

Table 1은 잡음 비율  $\alpha = 0.1$ 인 Fig. 6(b)과 잡음 비율  $\alpha = 0.3, 0.5$ 에 의해 생성된 가우시안 잡음영상에 여러 잡음 제거 방법들을 적용하여 얻은 성능 수치를 나타내고 있다.

Table 1. Performance comparison among several denoising models for Gaussian noisy images

Methods	PSNR			SSIM		
	0.1	0.3	0.5	0.1	0.3	0.5
BM3D	32.6332	24.9778	17.0195	0.8781	0.7066	0.1873
CDAE	32.8743	26.1956	18.6529	0.9141	0.7899	0.6611
DnCNN	34.3522	29.1538	26.6514	0.8905	0.7614	0.6779
U-Net GAN	34.1031	30.3567	27.9056	0.9020	0.8288	0.7724
WT-GAN	35.2297	30.5640	28.3088	0.9216	0.8291	0.7762

Table 1의 가우시안 잡음영상에서 제안된 WT-GAN 모델이 두 개의 척도상에서 모두 가장 높은 수치를 얻었다. 다른 딥러닝 모델, 즉 CDAE, DnCNN, U-Net GAN 모델은 PSNR과 SSIM 척도에 따라 성능 순위가 다르게 나타났다. BM3D는 다른 딥러닝 모델에 비해 낮은 수치를 보였다.

Table 2는 잡음 비율  $\alpha = 0.1$ 인 Fig. 6(c)과 잡음 비율  $\alpha = 0.3, 0.5$ 에 의해 생성된 포아송 잡음영상에 여러 잡음 제거 방법들을 적용하여 얻은 성능 수치를 나타내고 있다.

Table 2. Performance comparison among several denoising models for Poisson noisy images

Methods	PSNR			SSIM		
	0.1	0.3	0.5	0.1	0.3	0.5
BM3D	20.2547	12.9192	10.4972	0.5954	0.2523	0.1953
CDAE	34.8270	29.2855	26.3242	0.9181	0.8325	0.8297
DnCNN	29.7921	28.1529	25.3064	0.9120	0.8091	0.7700
U-Net GAN	36.1663	35.8065	30.9837	0.9513	0.9531	0.8602
WT-GAN	38.1648	36.0197	31.7271	0.9659	0.9551	0.8773

Table 2의 포아송 잡음 영상에서도 제안된 WT-GAN 모델이 PSNR, SSIM 모두 가장 높은 수치를 보였고, 다음으로 U-Net GAN, CDAE 순으로 나타났다. BM3D는 가장 낮은 수치를 보였다.

Table 3은 잡음 비율  $\alpha = 0.1$ 인 Fig. 6(d)과 잡음 비율  $\alpha = 0.3, 0.5$ 에 의해 생성된 스펙클 잡음영상에 여러 잡음 제거 방법들을 적용하여 얻은 성능 수치를 나타내고 있다.

Table 3. Performance comparison among several denoising models for Speckle noisy images

Methods	PSNR			SSIM		
	0.1	0.3	0.5	0.1	0.3	0.5
BM3D	33.9454	30.2178	25.7437	0.8879	0.8444	0.6864
CDAE	35.3599	18.6529	19.9444	0.9299	0.6611	0.4224
DnCNN	37.7742	32.9347	30.6612	0.9424	0.8628	0.8121
U-Net GAN	37.4263	33.6334	30.7133	0.9456	0.8977	0.8576
WT-GAN	38.5386	33.9492	32.1480	0.9594	0.8974	0.8634

Table 3의 스펙클 잡음 영상에도 Table 1과 Table 2에서처럼 WT-GAN 모델이 가장 좋은 성능을 보였고, 다른 딥러닝 모델, 즉 DnCNN, U-Net GAN 모델은 PSNR과 SSIM 척도에 따라 성능 순위가 다르게 나타났고, 다음 CDAE 모델, BM3D 순으로 성능 수치를 보였다.

## V. Conclusions and Future Research

CT 촬영 시 방사선량을 줄이면 상대적으로 안전하지만 잡음으로 인해 영상의 질이 떨어져 정확한 진단을 위해서는 일정 선량 이하로 낮출 수 없다. 저선량 CT 영상에서 잡음 제거는 환자 보호와 고품질 영상 도출이란 두 마리 토끼를 잡을 수 있게 해준다. 최근에 CT의 방사선 피폭을 줄이면서도 영상의 질을 진단적으로 유지하기 위해 인공지능을 사용해 한층 효율적인 영상잡음 제거가 가능하게 됐고, 이는 Siemens, GE, Philips, Canon 등 세계적으로 널리 알려진

CT 제조사의 상용 CT 제품에 이미 적용되고 있다.

현재까지 영상에서 잡음제거 문제를 해결하기 위하여 다양한 연구가 진행되어 왔다. 크게는 필터링 방법과 딥러닝 방법으로 나눈다. BM3D와 같은 필터링 방법들은 우수한 성능을 보였지만, 도메인 지식을 필요로 하기 때문에 특정한 잡음에 제한적으로 적용 가능하다는 한계를 갖고 있다. 최근 딥러닝 방법으로 CNN과 AE 모델 등이 잡음제거에 주로 사용되어 왔다. CNN 모델 중에서 DnCNN 모델은 몇 번의 컨볼루션 층을 거치면서 원 영상과의 오차가 최소가 되는 잡음제거 영상을 얻을 수 있고, AE 모델은 잡음이 추가된 영상을 인코더에 넣어 학습하고 디코더에서 잡음 없는 영상과의 오차를 최소화하는 잡음 제거 영상을 얻는다. 또한, CDAE 모델은 CNN과 DAE의 결합 형태로 역시 잡음 제거 영역에서 좋은 성능을 보였다.

오늘날 딥러닝 분야에서 생성형 AI에 대한 관심이 뜨겁다. 생성형 AI에 대한 대중적인 접근 방식 중 하나인 GAN은 새로운 콘텐츠를 생성하는 생성자와 콘텐츠를 평가하여 진위 여부를 결정하는 판별자의 두 가지 신경망으로 구성된다. 두 신경망은 생성자에 의해 생성적 모델, 즉 잡음제거 영상의 품질을 향상하기 위해 서로 경쟁한다.

지금까지 잡음제거를 위한 딥러닝 방법은 전통적인 방법에 비해 잡음제거 능력이 뛰어났지만, 잡음제거 과정에서 영상 내 디테일이 손실되고 또한, 인공물과 같은 선명하지 않은 결과물을 생성하는 결과를 보이는 문제점이 있다.

본 논문에서는 웨이블릿 변환을 이용하여 입력 영상을 고주파와 저주파 성분으로 대역 분할하고, 잡음이 포함된 고주파 영역에서 1차적인 잡음 제거가 이루어지고, 더욱 상세한 잡음 제거는 GAN 모델을 통해 더욱 향상된 결과를 얻고자 하였다. 여기서 사용된 GAN 모델은 U-Net 구조의 생성자와 PatchGAN 구조의 판별자를 통해 잡음제거 영상을 생성하였다. 성능실험에서 정성적인 평가를 해 보면, 제안된 WT-GAN 모델은 BM3D 모델, 그리고 다른 딥러닝 모델인 DnCNN과 CDAE 모델보다 확연히 선명한 영상을 얻을 수 있었고, U-Net GAN 보다 약간 선명한 영상을 얻을 수 있었다. 정량적인 평가에서 보면, WT-GAN 모델은 모든 잡음에서 두 가지 척도 PSNR과 SSIM에 대해 가장 성능 수치를 보였다.

비록 제안된 WT-GAN 모델은 영상을 주파수 대역으로 분해하여 분석할 수 있는 웨이블릿 변환과 딥러닝의 GAN 모델을 결합하여, 웨이블릿 변환 기반 GAN 모델을 통해 주어진 CT 영상에서 최적의 잡음 제거 방법을 찾았으나, 향후 연구에서는 최신의 웨이블릿 변환과 최신의 GAN 모델의 결합을 통해 특정 CT 영상에 국한되지 않고 모든 영

상에서 에지 성분을 보존하면서 잡음만을 완벽하게 제거할 수 있는 효율적인 잡음제거 방법을 개발하고자 한다.

## ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2022R1A2C1011140).

## REFERENCES

- [1] M. Ameen, and S. A. Ahmed, "An Extensive Review of Medical Image Denoising Techniques," *International Journal of Electronics and Communication Engineering and Technology*, Vol. 7, No. 6, pp. 85-90, December 2016.
- [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, July 2007, DOI:10.1109/TIP.2007.901238.
- [3] K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," *IEEE Transactions on Image Processing*, Vol. 26, No. 7, July 2017, pp 3142-3155, DOI:10.1109/TIP.2017.2662206.
- [4] W. H. Lee, M. Ozger, U. Challita, and K. W. Sung. "Noise Learning-Based Denoising Autoencoder," *IEEE Communications Letters*, Vol. 25, No. 9, pp. 2983 - 2987, September 2021. DOI: 10.1109/LCOMM.2021.3091800.
- [5] Y. Farooq, and S. Savaş, "Noise Removal from the Image Using Convolutional Neural Networks-Based Denoising Auto Encoder," *Journal of Emerging Computer Technologies*, Vol. 3, No. 1, pp. 21-28. December 2023, Doi: 10.57020/ject.1390428
- [6] Q. Yan and W. Wang, "DCGANs for image super-resolution, denoising and deblurring," *Advances in Neural Information Processing Systems*, pp. 487-495, 2017.
- [7] M. Arjovsky, S. Chintala and L. Bottou, "Wasserstein Generative Adversarial Networks," *Proceedings of Machine Learning Research*, Vol. 70, 2017, pp. 214-223, DOI:10.48550/arXiv.1701.07875.
- [8] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. C. Courville, "Improved training of wasserstein gans," in *Advances in Neural Information Processing Systems*, pp. 5767-5777, 2017.
- [9] E. Schönfeld, B. Schiele, and A. Khoreva, "A U-Net Based Discriminator for Generative Adversarial Networks," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* pp.8204 - 8213, February 2020, DOI: 10.1109/CVPR42600.2020.



- 00823.
- [10] Z. Sheng, X. Liu, S.Y. Cao, and H.L. Shen, "Frequency-Domain Deep Guided Image Denoising", IEEE Transactions on Multimedia, Vol. 25, pp.6767- 6781, 2023, DOI: 10.1109/TMM.2022.3214375.
- [11] Z. Li, H. Liu, L. Cheng, and X. Jia, "Image Denoising Algorithm Based on Gradient Domain Guided Filtering and NSST", IEEE Access, Vol. 11, pp.11923 - 11933, February 2023, DOI: 10.1109/ACCESS.2023.3242050
- [12] R. N. Kumar, and U. J. Kumar, "A Spatial Mean and Median Filter for Noise Removal in Digital Images," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 4, No. 1, pp. 246-253, January 2015. DOI:10.15662/ijareeie.2015.0401037.
- [13] S. Anitha, L. Kola, P. Sushma, and S. Archana, "Analysis of filtering and novel technique for noise removal in MRI and CT images," International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques, pp. 815-827, February 2017. DOI: 10.1109/ICEECCOT.2017.8284618.
- [14] 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, 2004. DOI: 10.1109/TIP.2003.819861.
- [15] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S.Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets", Advances in Neural Information Processing Systems, Vol. 27, pp. 2672-2680, 2014, DOI:10.48550/arXiv.1406.2661.
- [16] A. Radford, L. Metz and S. J. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," ICLR, November 2016, DOI:10.48550/arXiv.1511.06434.
- [17] I. Sharif and S. Khare, "Comparative Analysis of Haar and Daubechies Wavelet for Hyper Spectral Image Classification", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-8, 2014, pp. 937-941, DOI:10.5194/isprsarchives-XL-8-937-2014.
- [18] P. Isola, J. Y. Zhu, T. Zhou, and A.A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks", IEEE Conference on Computer Vision and Pattern Recognition, July 2017, DOI:10.48550/arXiv.1611.07004.
- [19] M. Talbi, R. Baazaoui, and B. Nasraoui, "A novel method of image denoising based on 2D dual-tree DWT and SWT", International Journal of Wavelets, Multiresolution and Information Processing, 2024, DOI:10.1142/S0219691324500097.
- [20] M. E. H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al-Emadi, M. B. I. Reaz, and M. T. Islam, "Can AI help in screening viral and COVID-19 pneumonia?", IEEE Access, Vol. 8, pp.132665-132676, July 2020, DOI: 10.1109/ACCESS.2020.3010287.
- [21] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. B. A., Kashem, M. T. Islam, S. A. Maadeed, S. M. Zughaier, M. S. Khan, and M. E. Chowdhury, "Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images," Computers in Biology and Medicine, Vol. 132, pp. 1- 16, May 2021, DOI:10.1016/j.combiomed.2021.104319.
- [22] Y. Li, J. Lu, L. Wang, and T. Yahagi, "Noise Removal for Degraded Images with Poisson Noise Using M-Transformation and BayesShrink Method", Electronics and Communications in Japan(Part III: Fundamental Electronic Science), Vol. 90, pp. 11-20, May 2007, DOI: 10.1002/ecjc.20357.
- [23] E. S. Seong, S. H. Han, J. H. Heo, and D. H. Lim, "Efficient CT Image Denoising Using Deformable Convolutional AutoEncoder Model", Journal of The Korea Society of Computer and Information Vol. 28 No. 3, pp. 25-33, March 2023, DOI:10.9708/jksci.2023.28.03.025.

## Authors



Hae Chan Jeong received the B.S. degree in Information and Statistics from Gyeongsang National University, Korea, in 2024.

Mr. Jeong is interested in information statistics and deep learning.



Dong Hoon Lim received the B.S. and M.S. degrees in Computer Science and Statistics, and Ph.D. degree in Statistics from Busan National University, Korea, respectively.

Dr. Lim is currently a Professor in the Department of Information and Statistics, Gyeongsang National University. He is interested in image processing, information statistics, big data analysis and deep learning.