



MEDISYSLAB

연구기간 : 2024.09.01  
~ 2024.12.01

연구자 : 김남훈

# REDCNN and UNet with Attention Mechanisms on the Low Dose CT Grand Challenge

**ksouth0413/Low-Dose-CT-Grand-Challenge**



REDCNN and UNet with Attention Mechanisms on the Low Dose CT Grand Challenge

1 Contributor 0 Issues 1 Star 0 Forks



**ksouth0413/Low-Dose-CT-Grand-Challenge: REDCNN and UNet with Attention Mechanisms on the Low Dose CT...**

REDCNN and UNet with Attention Mechanisms on the Low Dose CT Grand Challenge - ksouth0413/Low-Dose-CT-Grand-Challenge

GitHub

## Task: Image Denoising



**Low-dose CT**



**Normal-dose CT**

- **Model**

- REDCNN
- UNet
- +attention

- **Evaluation**

- RMSE, PNSR, SSIM

(목표) LDCT 이미지 노이즈 제거

선량을 감축하면서도 명확하고 정확한 이미지를 얻어 진단의 효율성 향상

CT 이미지의 HU 단위는 Hounsfield Unit의 약자로, 의료 영상에서 조직 밀도를 측정하는 데 사용되는 단위임

CT 스캔에서 나오는 각 픽셀의 값을 HU 단위로 표현하며, 이는 X선이 조직을 통과할 때 얼마나 감쇠되는지를 나타냄

기준값:

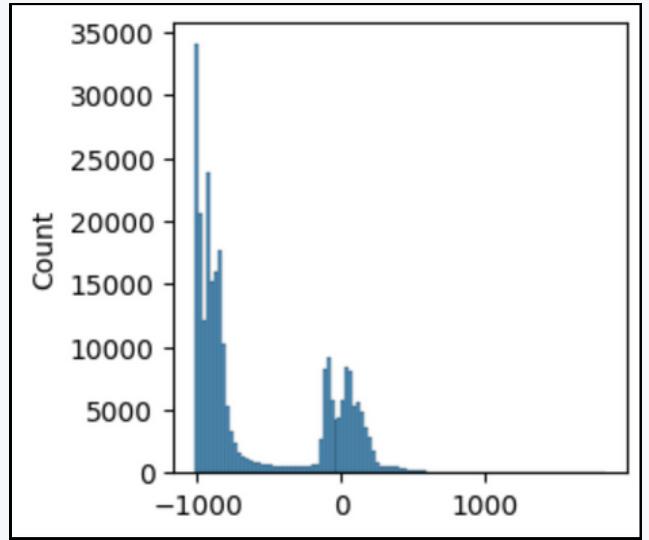
- 물(Water): 0 HU로 정의.
- 공기(Air): -1000 HU로 정의.
- 뼈, 지방 등: HU 값은 밀도와 X선 감쇠율에 따라 달라짐.(물보다 높음)

-> 따라서 주어진 데이터셋 이미지의 픽셀값을 HU 단위로 바꾸는 전처리 과정이 필요함

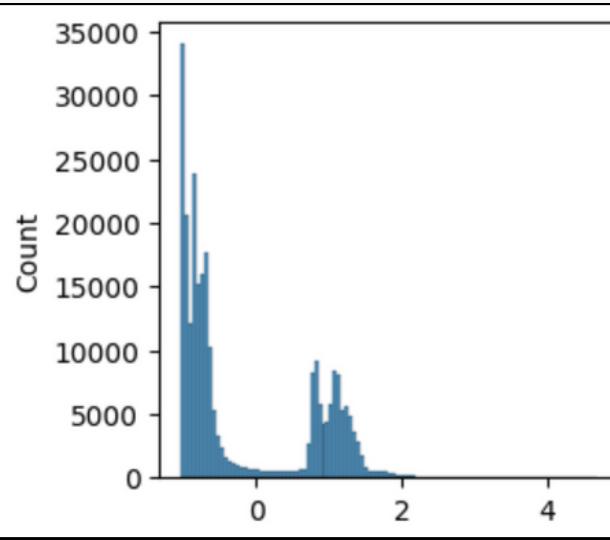
```
❶ # 원본값으로 변환(픽셀화, 스케일링, 실수화 진행)
img = dicom_data.pixel_array.astype(np.float32) * dicom_data.RescaleSlope + dicom_data.RescaleIntercept
img

❷ array([[-1006., -1010., -1010., ..., -1017., -1015., -1011.],
       [-1000., -997., -993., ..., -1013., -1011., -1005.],
       [-999., -998., -989., ..., -1007., -1004., -1002.],
       ...,
       [-1002., -1013., -1012., ..., -1014., -1015., -1015.],
       [-995., -1008., -1013., ..., -1012., -1015., -1012.],
       [-1002., -1009., -1013., ..., -1004., -1011., -1006.]],
      dtype=float32)
```

## 1. 정규화

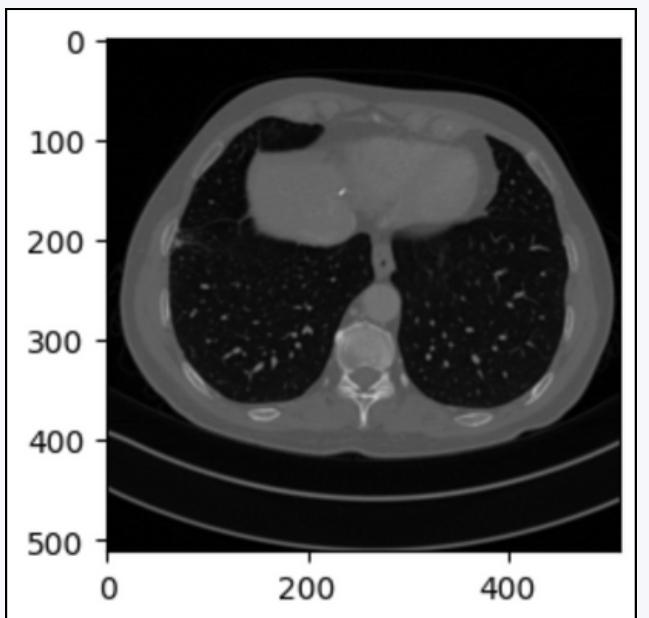


▲ (-1000,2000) 범위

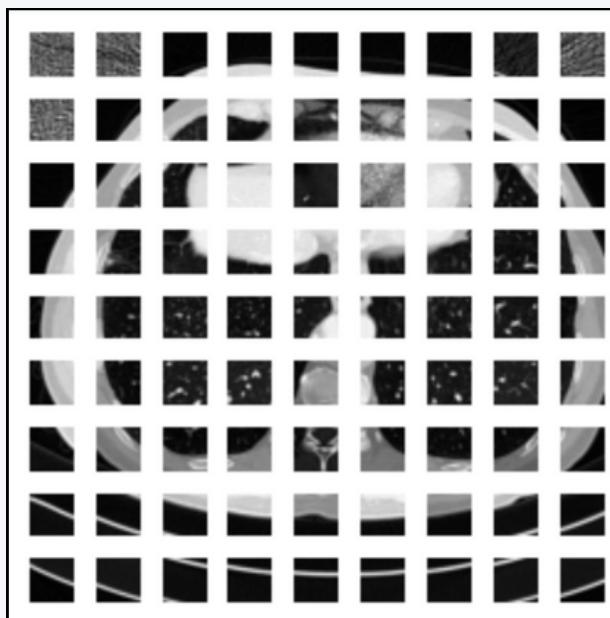


▲ (-1~2) 범위

## 2. patch로 쪼갬

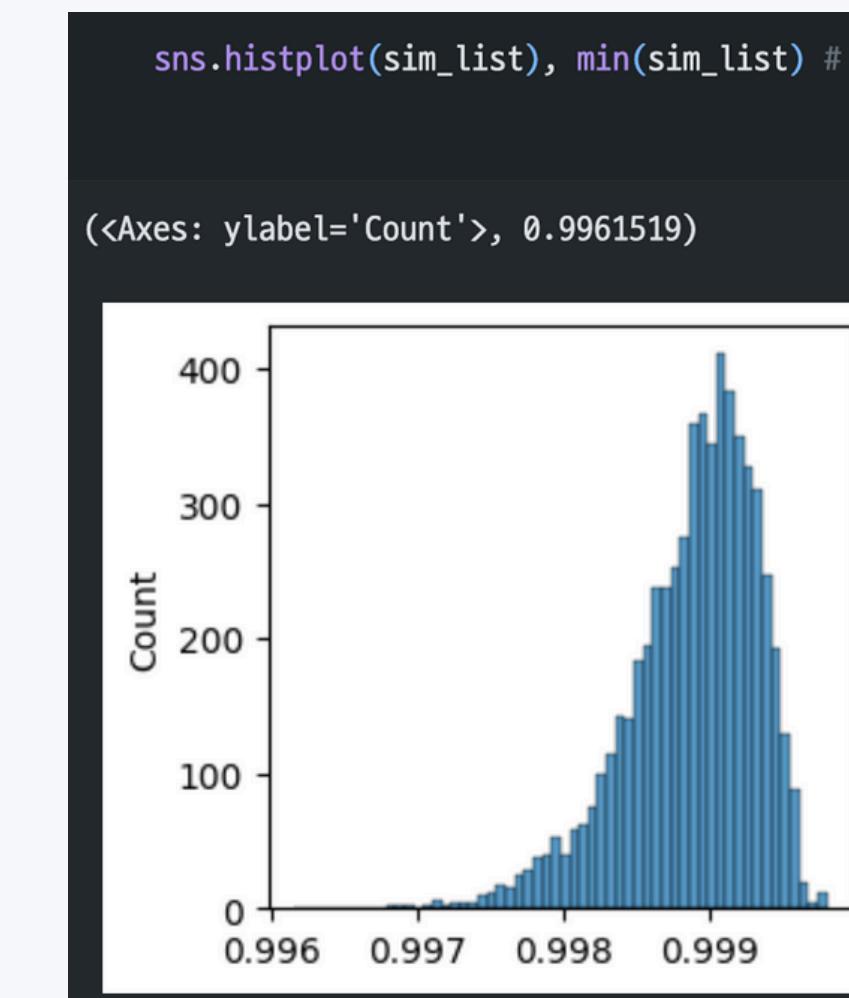


▲ 원본 512x512 (1개)



▲ patch 55x55 (81개)

## 3. Cosine Similarity(low-dose & normal-dose)



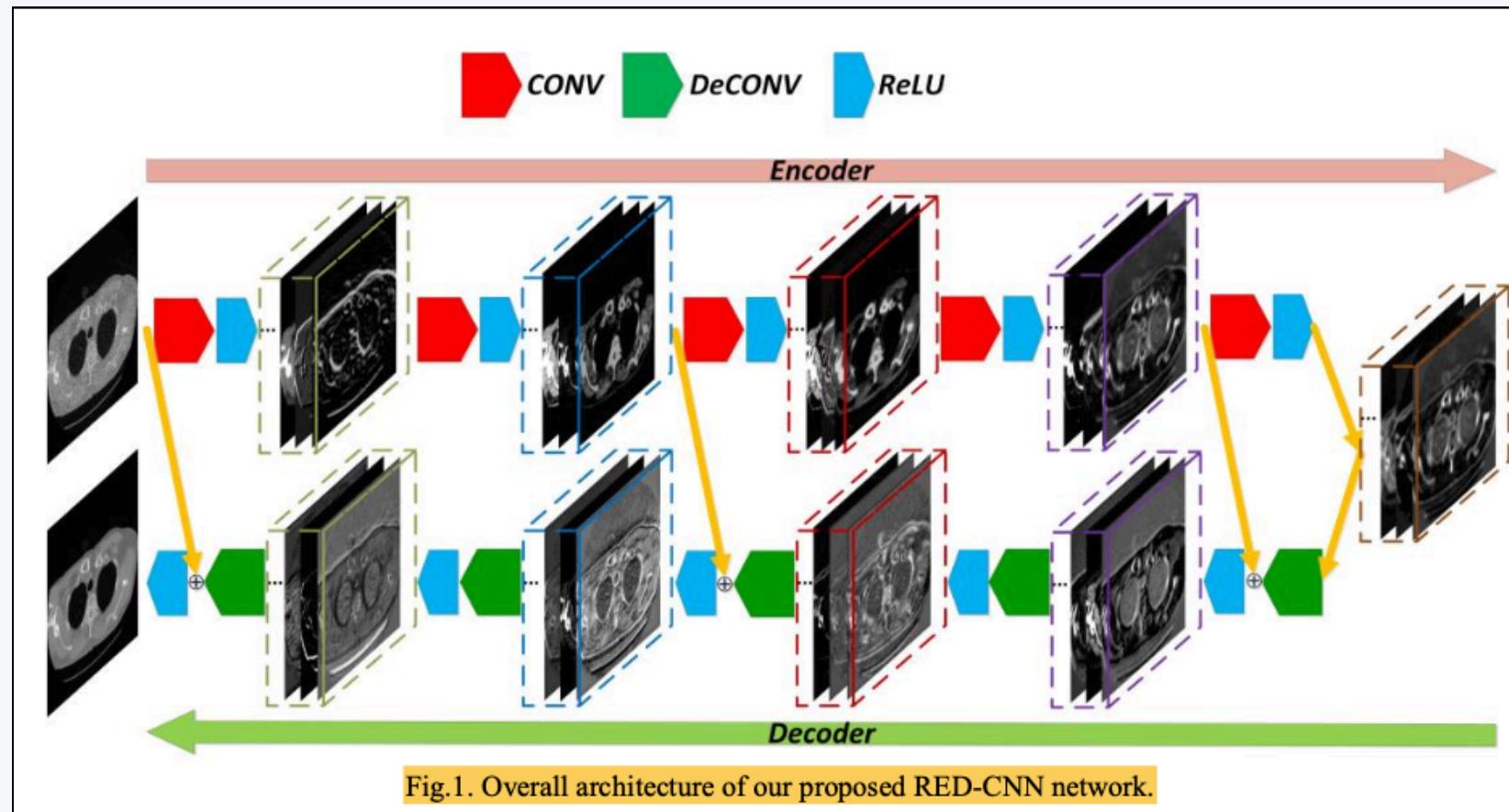
**(step1) ResNet 50을 가져와서 벡터(2048)로 특징 추출**

- weights = models.ResNet50\_Weights.IMAGENET1K\_V1

**(step2) cosine값으로 유사도 계산**

- 최솟값이 0.996일 정도로 두 이미지는 CNN으로 특징추출시 비슷
- CNN 적용의 가능성이 생김

## What is REDCNN ?



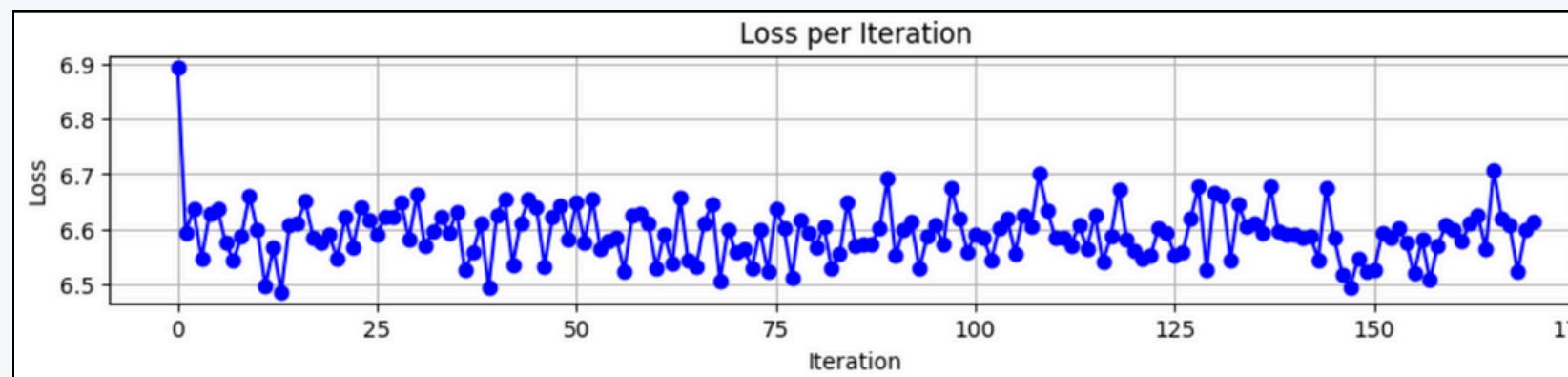
Low-Dose CT with a Residual Encoder-Decoder  
Convolutional Neural Network (RED-CNN) (2017)

## 주요 특징

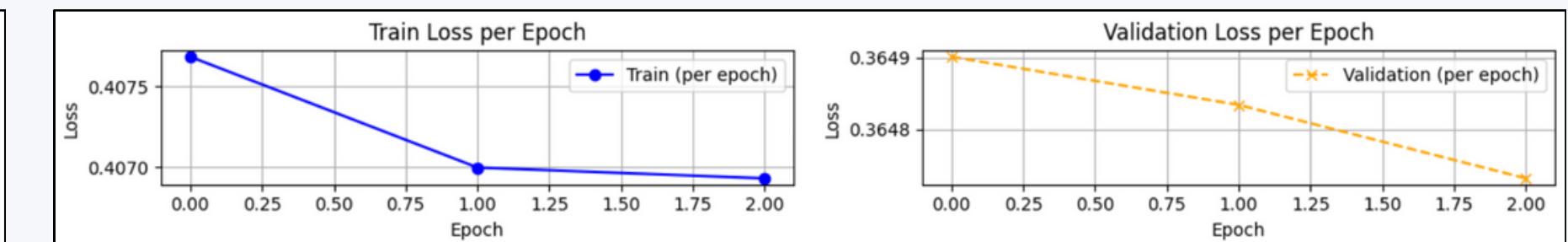
- 1. Feature map 사이즈가 조금씩 변화 (+-4 each)**
  - use filter = 5x5, stride = 1
  - 55x55 → 51x51 → 47x47 → ... → 39x39
- 2. Channel 수는 96으로 고정**
  - nn.Conv2d(96, 96, 5, 1, 0)
  - nn.ConvTranspose2d(96, 96, 5, 1, 0)
- 3. Skip connections**

-> 저레벨(edge/texture) 강조, 고레벨(추상적이고 의미론적인 정보)  
이 약함.

## 학습 결과

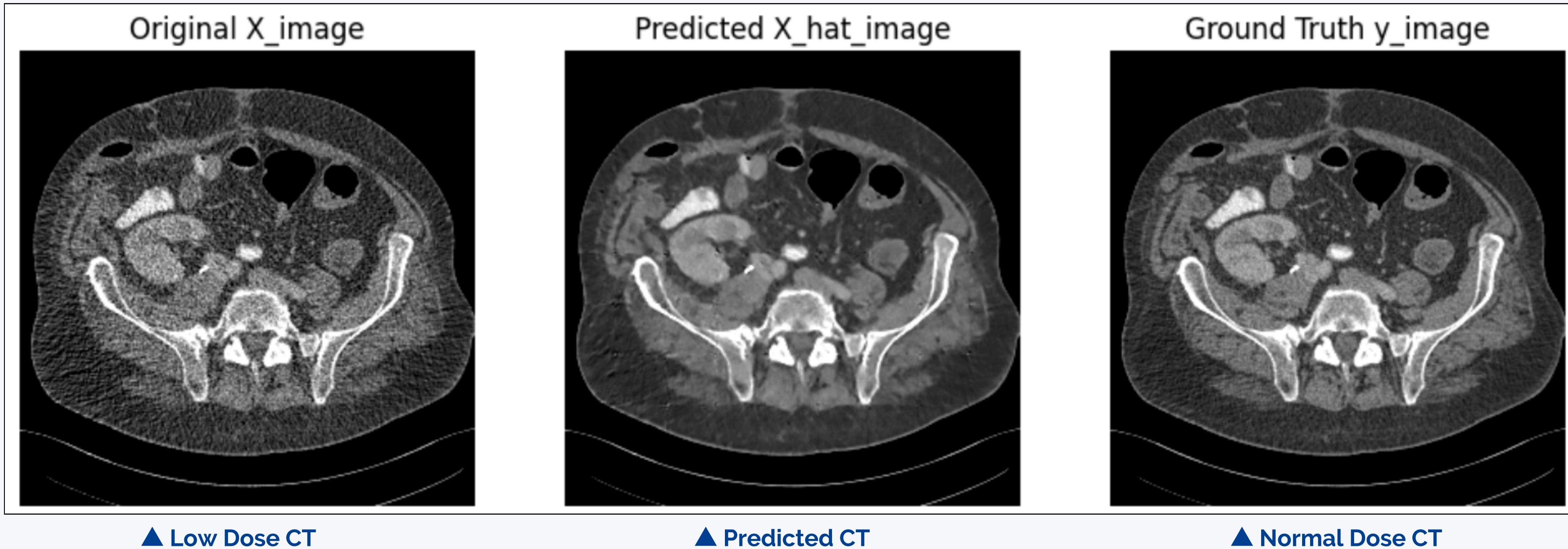


▲ loss 추이(1 iter = 1000개 당 평균)



▲ loss 추이(1 epoch = 57만개 당 평균)

## 결과



Low Dose의 noise가 많이 줄어들었다. 하지만 원본 사진의 디테일도 같이 사라져버렸다.

-> 추가 모델 개발 필요성

## 4가지 모델 비교

### 1. base(논문에서 제시된 기준모델)

### 2. Kaiming He Weight Initialization

- relu에 적합
- nn.init.kaiming\_normal\_(self.conv1.weight, mode='fan\_out', nonlinearity='relu')

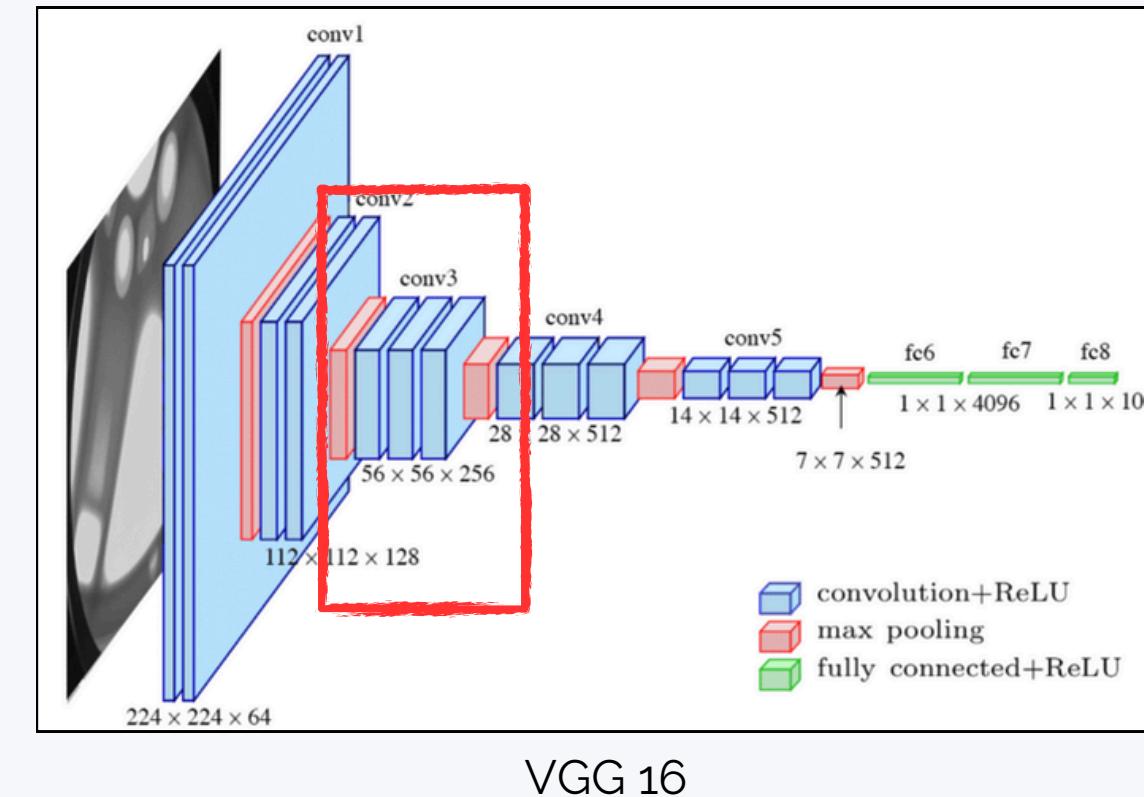
### 3. Batch Norm

- conv-layer → batch-norm → relu
- self.relu(self.bn1(self.conv1(x)))

### 4. VGG Loss

- self.feature\_extractor =  
nn.Sequential(\*list(vgg.features)[:16]).eval()  
# relu3\_3까지 사용
- self.criterion = nn.L1Loss()
- total\_loss = loss\_mse + 0.01 \* loss\_vgg

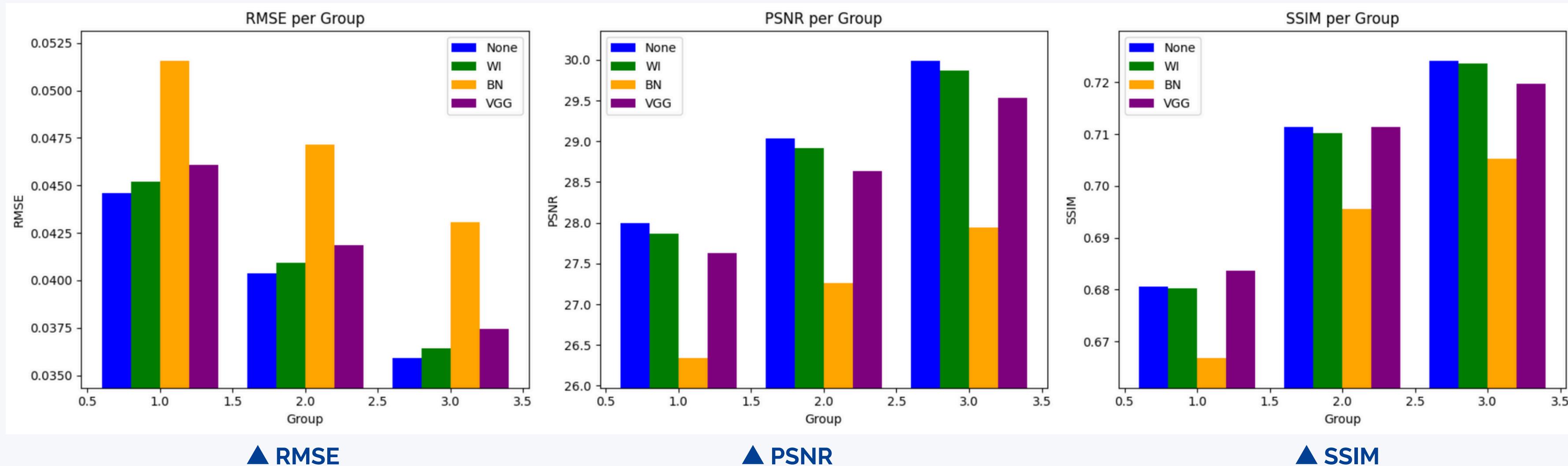
## VGG16 Loss



## 평가지표(미리보기)

지표	설명	해석 기준	특징
<b>RMSE</b>	픽셀 차이의 평균 제곱 뿐만 아니라 크기	낮을수록 좋음	절대적인 픽셀 값 오차를 평가.
<b>PSNR</b>	RMSE를 로그 스케일로 변환 한 유사도	높을수록 좋음	상대적 품질 비교에 유리.
<b>SSIM</b>	구조적, 대비, 밝기 측면에서 유사도 평가	1에 가까울수록 좋음	인간 시각과 유사한 품질 평가.

## 결과



## 해석

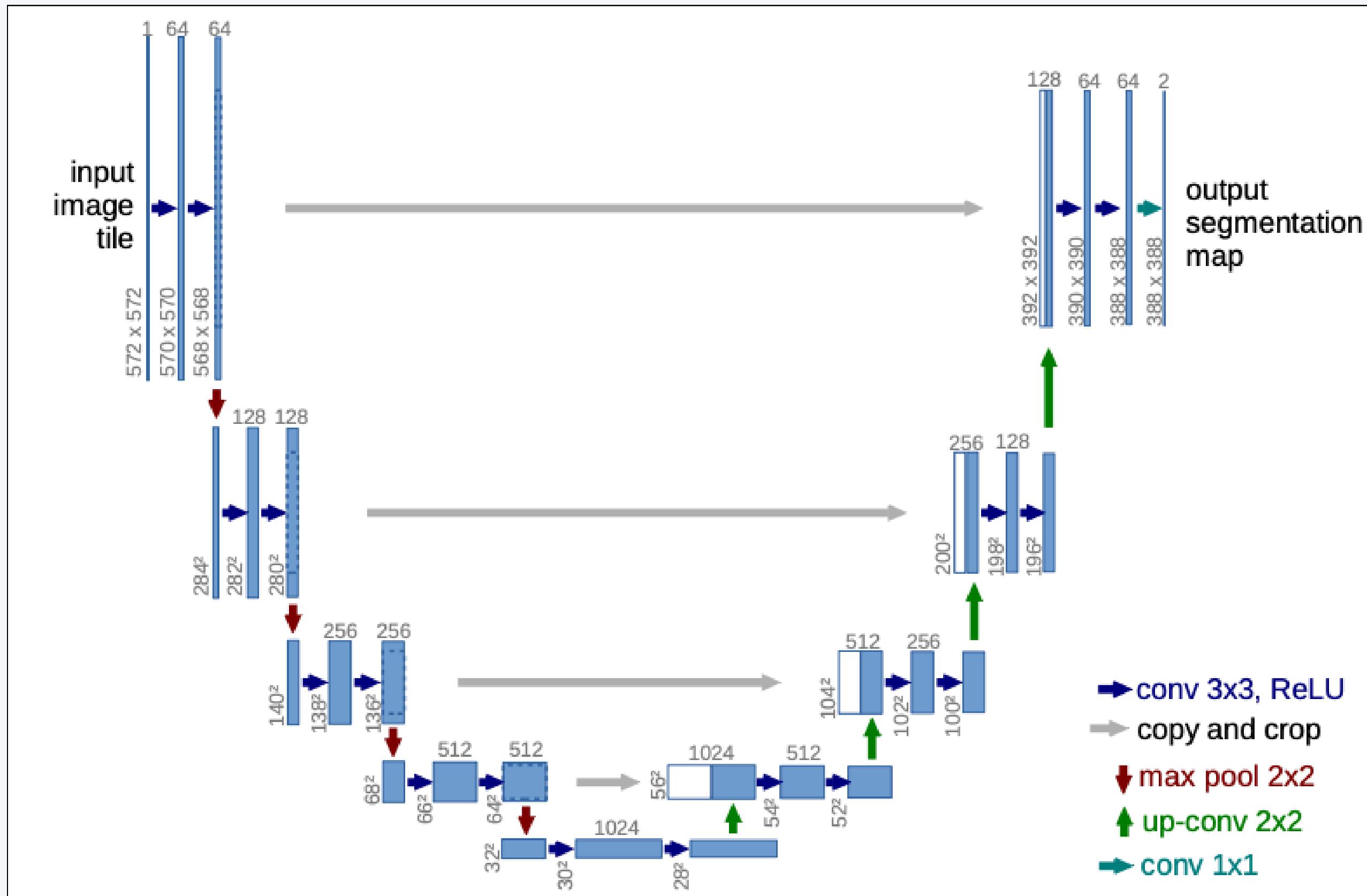
### 1. 학습이 잘 이루어졌다.

- X축의 Group은 1 iter=1000개 당 지표값을 의미(group=1 : 첫번째 1000개, group=2 : 두번째 1000개)
- 학습이 진행됨에 따라 RMSE(감소) / PSNR(증가) / SSIM(증가)

### 2. 성능 순위 : None(기존논문모델) > WI(가중치초기화) > VGG >> BN(배치정규화)

- BN : feature map을 정규화하면 저레벨 특징의 미묘한 차이를 희석시켜 성능 저하를 초래
- VGG : 고레벨 특징에 대한 loss를 추가해도 성능이 개선되지 않음

## What is UNet ?



## 주요 특징

1. Feature map 사이즈가 절반씩 감소
2. Channel수가  $1 \rightarrow 64 \rightarrow 128 \rightarrow \dots \rightarrow 1024 \rightarrow \dots \rightarrow 1$ 로 커졌다가 작아짐
3. Skip connections

## 저레벨에 특화된 REDCNN

1. Feature map 사이즈가 조금씩 변화 (+-4 each)
2. Channel수는 96으로 고정
3. Skip connections

-> 저레벨(edge/texture) 강조, 고레벨(추상적이고 의미론적인 정보)이 약함.

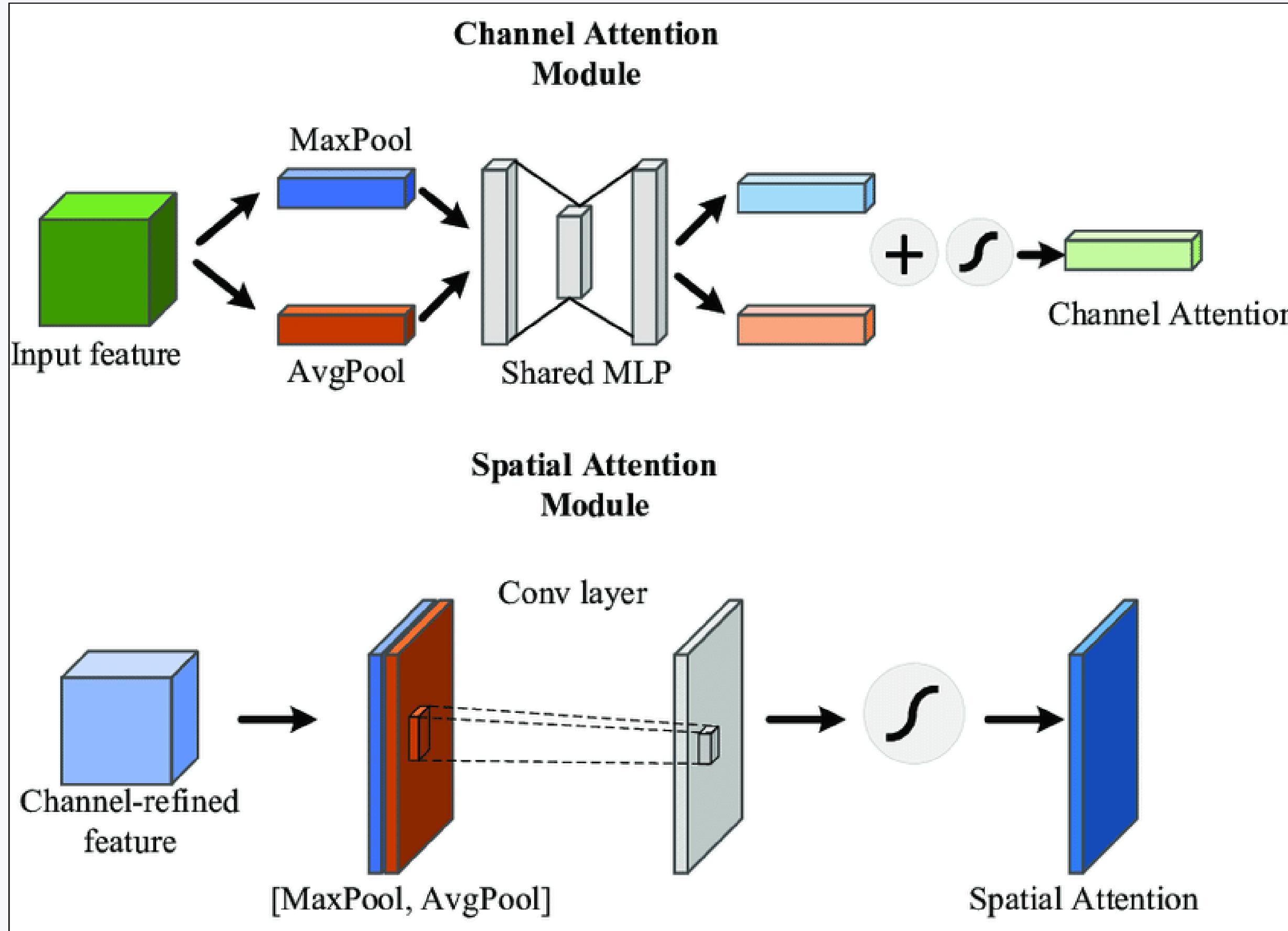
## 고레벨도 같이보는 UNet

1. Feature map 사이즈가 절반씩 감소
2. Channel수가  $1 \rightarrow 64 \rightarrow 128 \rightarrow \dots \rightarrow 1024 \rightarrow \dots \rightarrow 1$ 로 커졌다가 작아짐
3. Skip connections

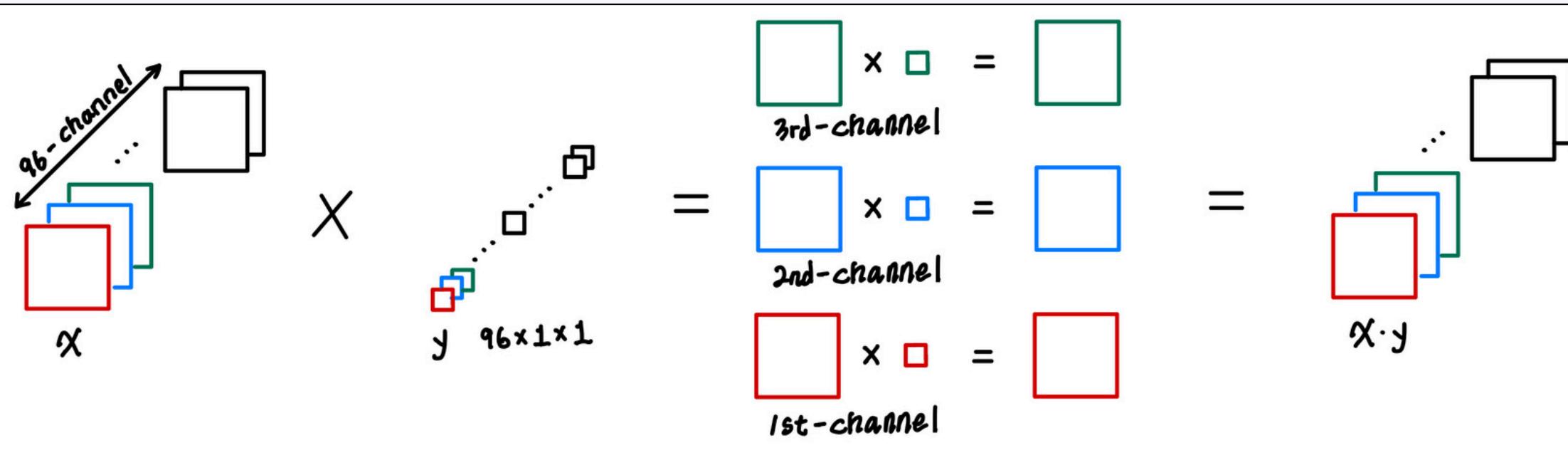
-> 저레벨(edge/texture) 뿐 아니라, 고레벨(추상적이고 의미론적인 정보도 학습에 사용함).

## 저레벨 vs 고레벨

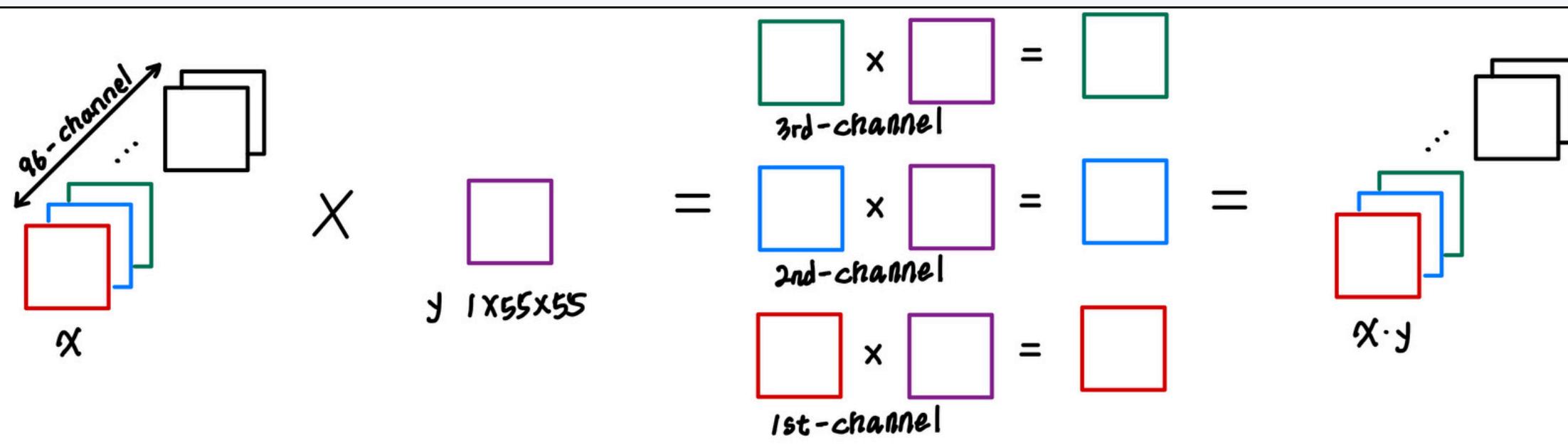
level 종류	설명	feature map 깊이	Channel 수	기타
<b>low-level</b>	고해상도 / high-freq	얕은 층	변화 x	skip connection으로 high-freq 정보 전달
<b>high-level</b>	저해상도 / low-freq	깊은 층	점점 깊어짐	feature map 사이즈를 크게 감소시켜 정보 축약



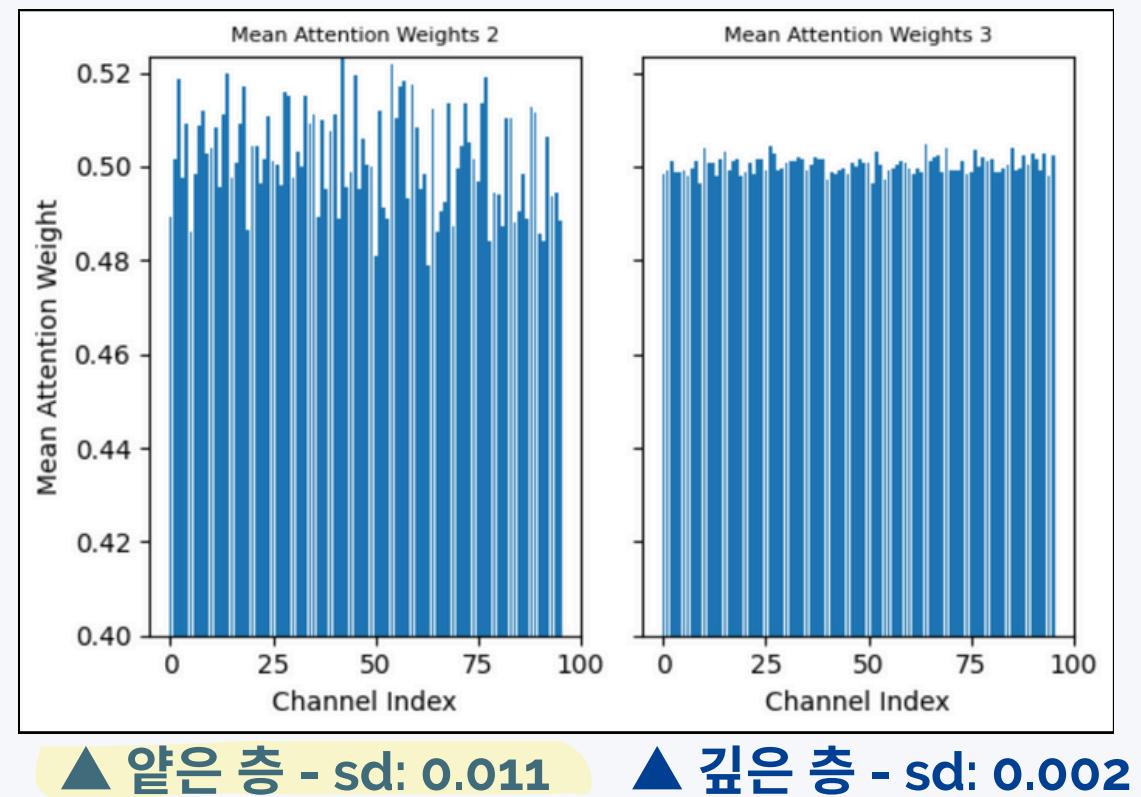
## Channel Attention



## Spatial Attention



## REDCNN



**얕은 층이 중요한 REDCNN vs 깊은 층이 중요한 UNet**

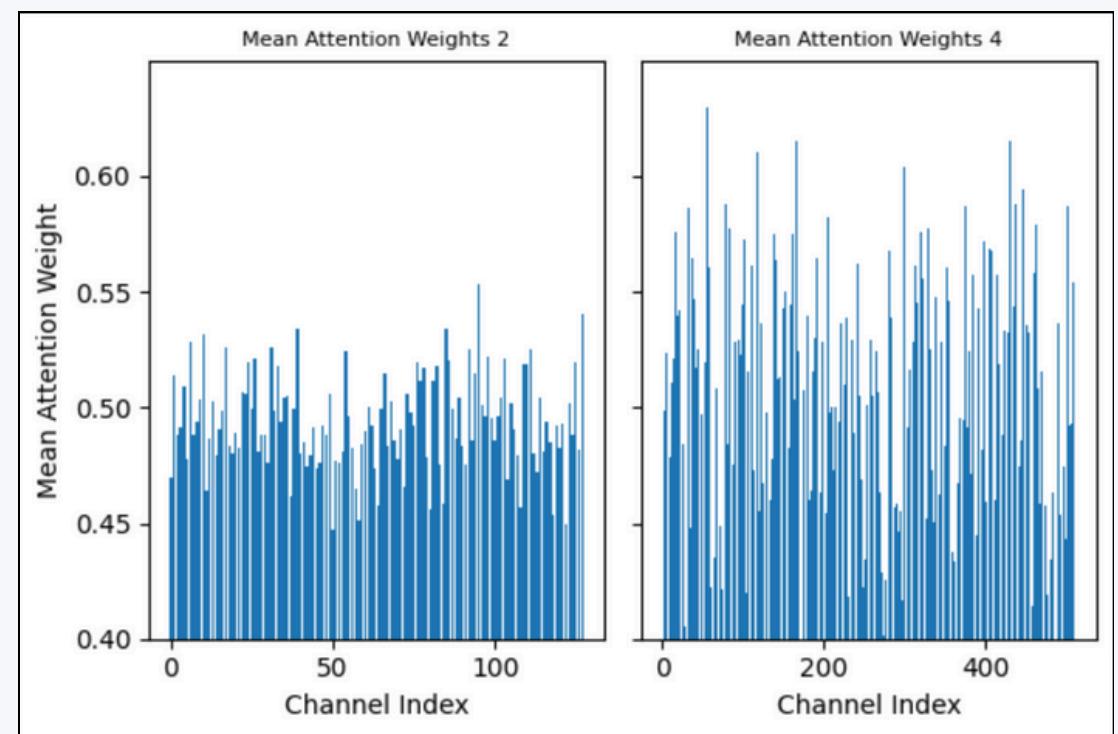
1. **attention weights가 특정 channel에 몰려있다**면, 그 층이 중요한 층이다
  - channel별로 attention의 표준편차가 클수록, 특정 channel정보가 증폭/감쇠 하여 중요한 정보를 선별한다

2. **REDCNN은 얕은 층, UNet은 깊은 층의 attention weight가 특정 channel에 몰려있다**

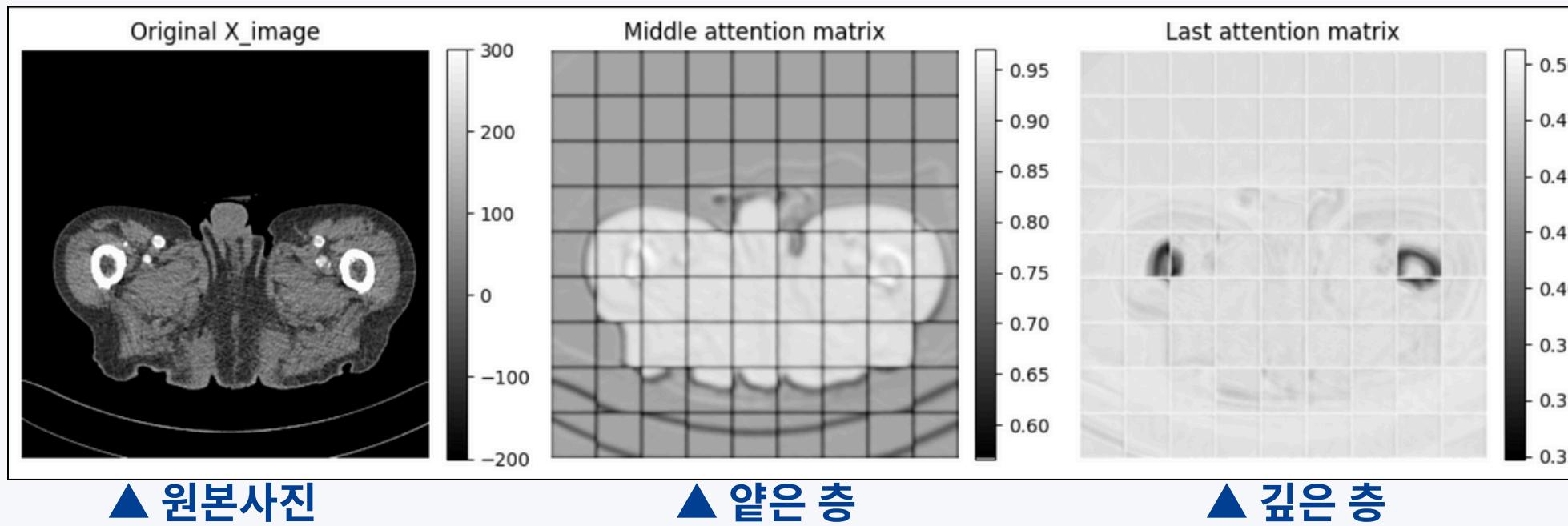
- REDCNN은 얕은 층의 channel이, UNet은 깊은 층의 channel이 모델 성능향상에 중요하다

-> **REDCNN은 저레벨의 정보를, UNet은 고레벨의 정보를 중요시한다.**

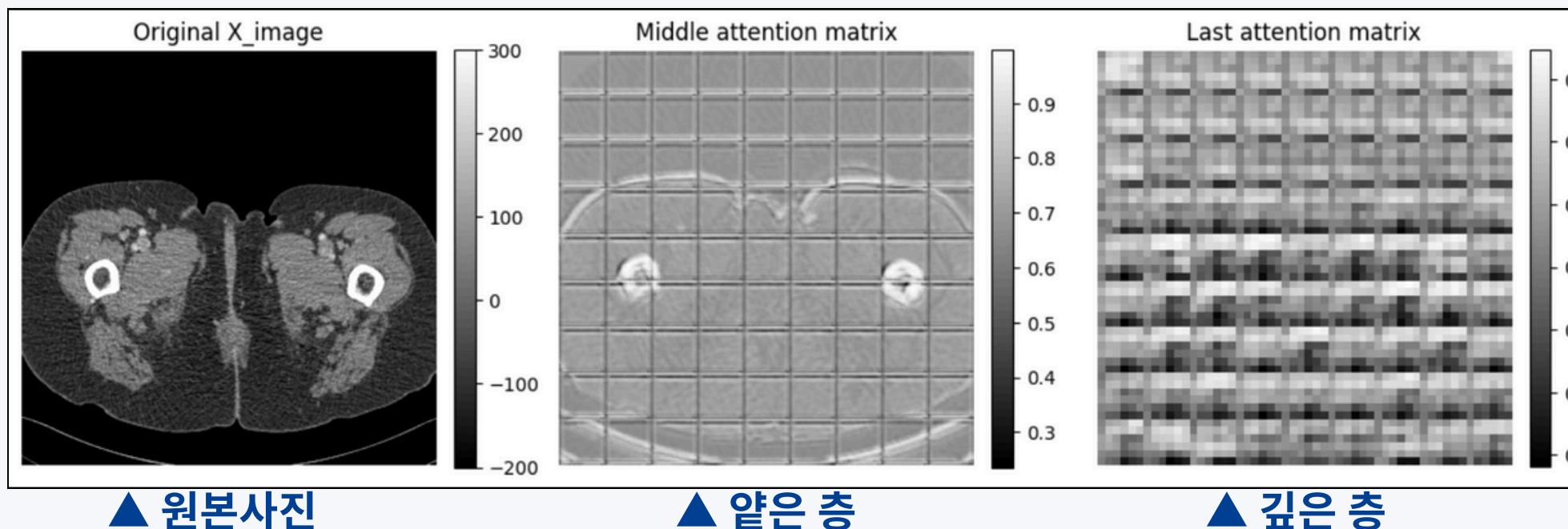
## UNet



## REDCNN



## UNet



**저레벨을 중시하는 REDCNN vs 저레벨,고레벨을 중시하는 UNet**

1. REDCNN의 경우, 얕은 층에서는 테두리(edge) 정보를  
깊은 층에서는 안에 결/질감(texture) 정보를 획득한다
  - 깊은 층이더라도 feature map 사이즈가 원본 이미지 크기와 비슷하게 유지되면서, 저레벨 정보를 학습하는데에 사용된다
2. UNet의 경우, 얕은 층에서는 테두리(edge)/texture 정보를 깊은 층에서는 추상적인 정보를 획득한다

- feature map 사이즈가 절반씩 줄어들기에, 깊은 층은 보다 추상적인 고레벨 정보를 담고 있다

-> REDCNN은 저레벨만을, UNet은 저레벨 뿐 아니라 고레벨의 정보도 중요시한다.

## 저레벨에 특화된 REDCNN

1. Feature map 사이즈가 조금씩 변화 (+-4 each)
2. Channel 수는 96으로 고정
3. Skip connections

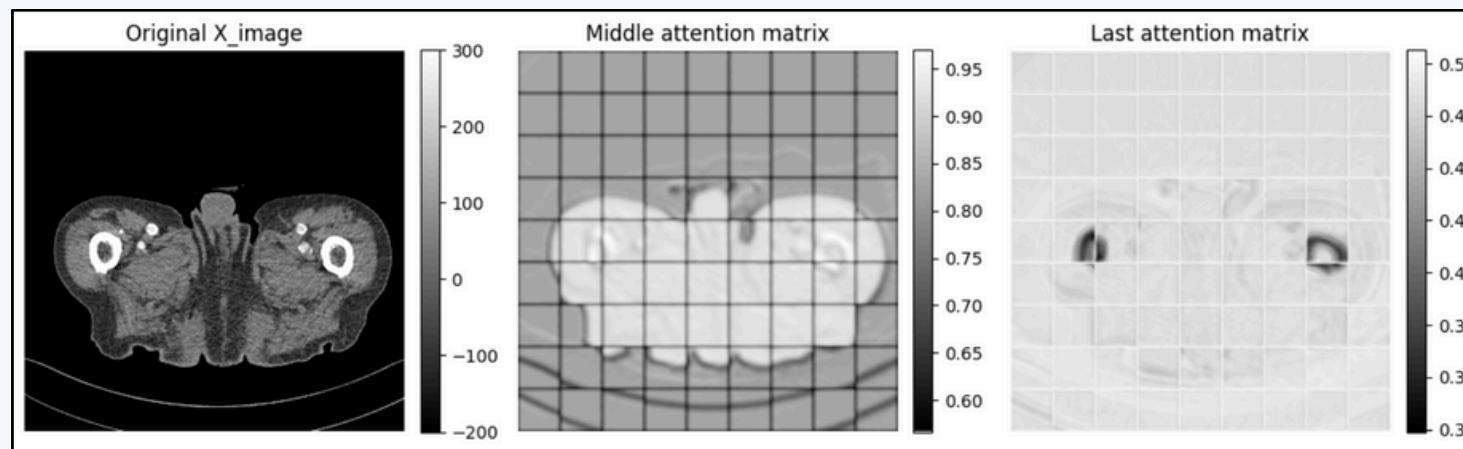
-> 저레벨(edge/texture) 강조, 고레벨(추상적이고 의미론적인 정보)이 약함.

## 고레벨도 같이보는 UNet

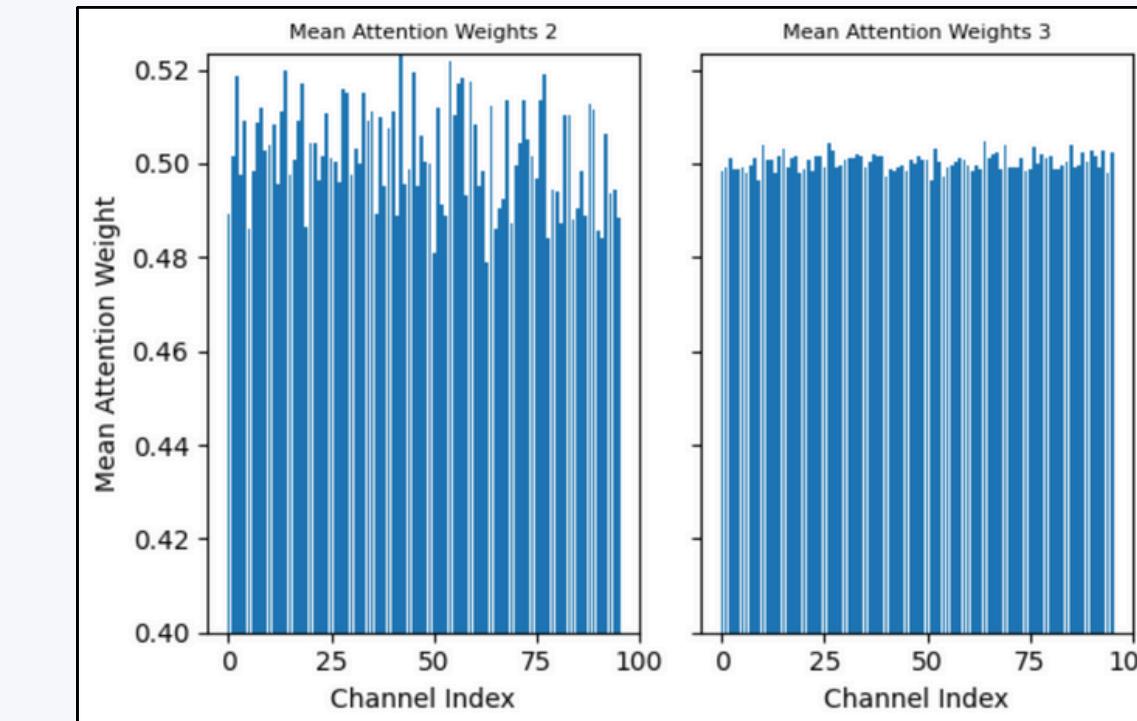
1. Feature map 사이즈가 절반씩 감소
2. Channel 수가 1->64->128->...->1024->...->1로 커졌다가 작아짐
3. Skip connections

-> 저레벨(edge/texture) 뿐 아니라, 고레벨(추상적이고 의미론적인 정보)도 학습에 사용함.

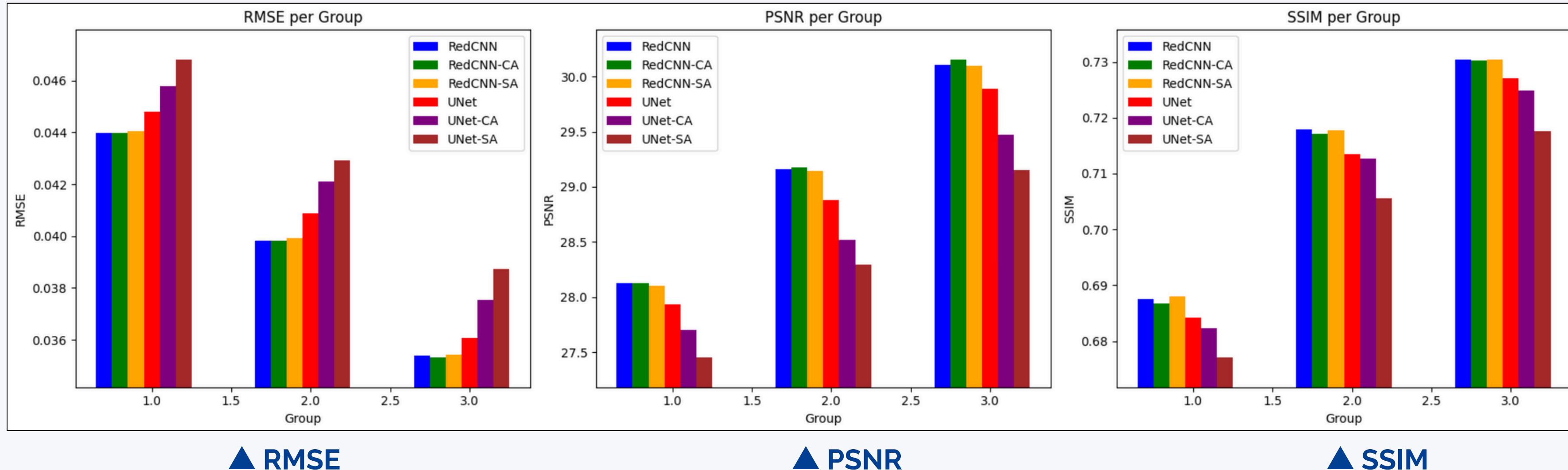
## Feature map -> Spatial Attention



## Channels -> Channel Attention



## Experiment Summary



저레벨에 집중한 REDCNN이 저레벨과 고레벨 둘 다에 집중한 UNet보다 성능이 우수하다  
-> REDCNN 모델구조에 대한 타당성 검증

## References

- [1] Chen, H., Zhang, Y., Kalra, M.K., Lin, F., Chen, Y., Liao, P., Zhou, J., Wang, G. & Zhang, J. (2017). Low-Dose CT with a Residual Encoder-Decoder Convolutional Neural Network (RED-CNN). *IEEE Transactions on Medical Imaging*, 36(12), 2524–2535.
- [2] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Lecture Notes in Computer Science, 9351, 234–241. Springer, Cham. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)

## Figure

- [1] Figure source: Adapted from “Diagram of the channel attention module and spatial attention module,” by authors in [ResearchGate](#).