CT super-resolution using deep Convolutional Neural Network

Abstract

- Data : training – 52set, test – 13set
- Validation : 5-fold cross-validation
- Experiment for resolution, contrast, noise properties
- Improvement :
  - deblurring of boundaries of bone structures and air cavities
  - 10% higher peak SNR, lower normalized root mean square than input
  - Output noise level : lower than ground truth, equivalent to iterative reconstruction result

Introduction

- CT : immense amount of acquired data -> longer scan range & slower operating speed

⇒ Using small angle views with a thicker slice thickness -> recon

- Reconstructed slice thickness ↓
  - Z-axis resolution ↑
  - Noise ↑
  - Dose ↑

⇒ Without dose ↑ , noise ↑ , enhance CT img resolution (similar to using thin slice)

- CNN : good at super-resolution, de-noising
- CNN base super-resolution approach -> some drawbacks in biomedical image processing
  1. Network run for each patch -> redundancy -> slow computation
  2. Trade off between localization accuracy and use of context
     - Large patch : use large context but max pooling ↑ -> localization accuracy ↓
     - Small patch : use little context -> localization accuracy ↑

Introduction

⇒ U-Net : provide end to end mapping architecture

- Contracting path : capture the context
- Expanding path : precise localization with large receptive field

- **Main purpose : CNN is also useful for CT image super-resolution**

⇒ end-to-end mapping between low and high resolution imgs using modified U-Net

⇒ Performance : resolution↑, contrast↑ & noise↓

Materials and methods

- Data : 65 clinical PET/CT studies for brain with suspected Parkinson's disease
  - CT : 512 X 512 X (60 - 80)
  - Training dataset : 5 slices of a 3mm thickness -> averaged into 1 slice of 15mm thickness / 52 patients
  - Ground truth : middle slice of five slices (high resolution) / 13 patients

- 5 – fold cross-validation
- Input : thick slice(low resolution) /Output : thin slice(high resolution)
  - 2D slice
  - Total number of slice for training : 7670

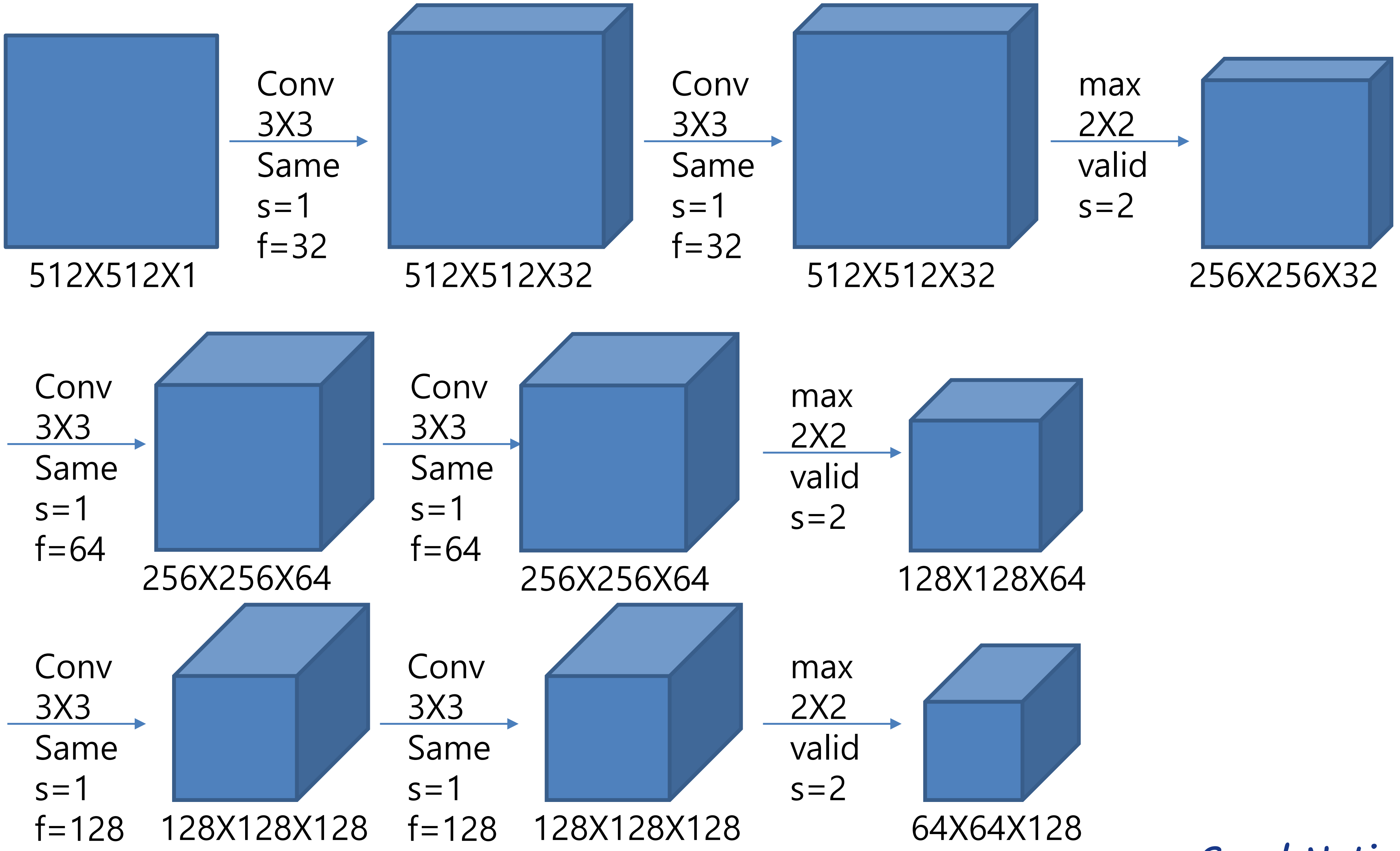
Materials and methods

- Network architecture
  - end to end mapping between thinner and thicker slice images
  - Architecture : contracting path & expanding path (symmetrical, each path : 5 sequential layer)

- Contracting path : capture the context
  - Typical CNN architecture
  - Each layer : two [conv(3X3) → BN → ReLU] => max pooling(2X2, s = 2)
  - mini-batch size : size of six images

- Expanding path : enable precise localization
  - Each layer : up-sampling(s = 2) => two[conv(3X3) → BN → ReLU]
    → Propagate context information to higher resolution layer
  - Last layer : 1X1 conv for scaling
  - Skip connection between conv layer and deconv layer

Materials and methods



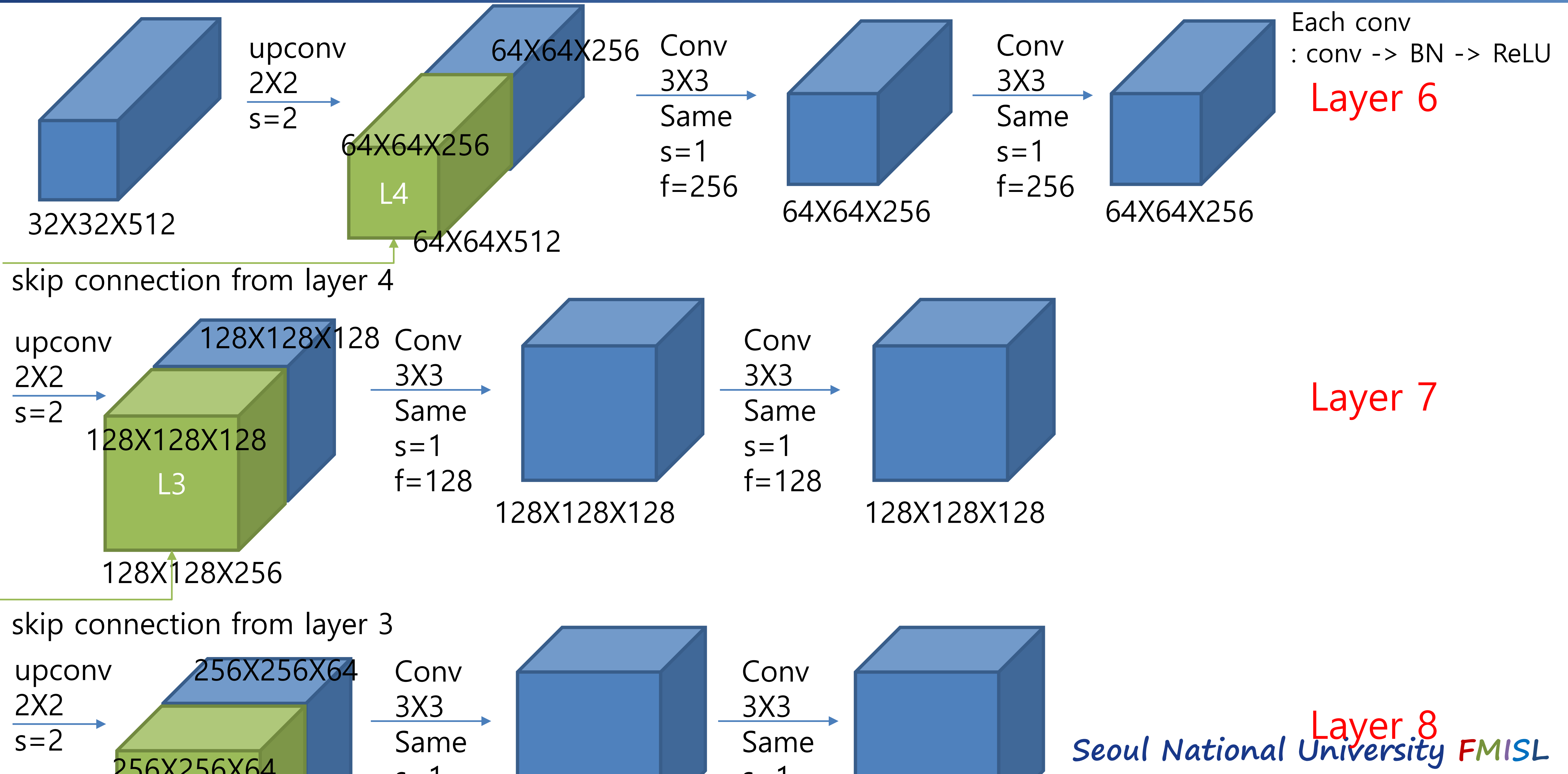
Each conv
: conv -> BN -> ReLU

Layer 1

Layer 2

Layer 3

Materials and methods



Data training and loss function

- Weight : truncated normal distribution
- Loss function : L2 Loss
  - F(Yi) : reconstructed image from low-resolution image Yi
  - Xi : high resolution label image

$$L(v) = \frac{1}{n}\sum_{i=0}^{n} \| F(Y_i; v) - X_i \|^2 , v = \{W_1, W_2, \cdots, W_m\}$$

n : #imgs in mini-batch
m : #layer

- Activation function : ReLU
- Optimizer : Adam

$$\Delta_{t+1} = \Delta_t - \frac{\eta}{\sqrt{\widehat{v_t}+\epsilon}}\widehat{m_t}, \quad \eta = \frac{\tau\sqrt{1-\beta_2^t}}{1-\beta_1^t}, \quad \widehat{m_t} = \frac{m_t}{1-\beta_1^t}, \quad \widehat{v_t} = \frac{v_t}{1-\beta_2^t}$$

t : iteration number
Δ_t : update parameter(weight/bias)
τ : learning rate
m_t, v_t : moment vector (initialized to zero)
β_1 : 0.9 β_2 : 0.999 ϵ : 10^{-8}

- Epoch : 18

Image analysis

- Evaluation : 13 patients data (5 – fold cross validation among 65 patient data sets)
- Image quality metrics
  - Peak signal-to-noise ratio : $PSNR = 10 \cdot \log_{10} \frac{MAX_I^2}{(\sum_{ij}^{PQ}[X(i,j)-Y(i,j)]^2)/N}$
  - Normalized root mean square error : $NRMSE = \frac{1}{MAX_I - MIN_I}\sqrt{\frac{\sum_{ij}^{PQ}[X(i,j)-Y(i,j)]^2}{N}}$

X(i,j) : high resolution image
Y(i,j) : output image
MAX_I, MIN_I : max, min intensity of image

- Comparison : 3D Richardson-Lucy deblurring algorithm
  - Kernel size of the point spread function : 3X3X3(pixel)
  - sigma = 1

  Highest PSNR, lowest NRMSE
  (kernel size from 0.5 to 3.0 in step of 0.5)

  → larger kernel size & larger sigma : too noisy

Image analysis

- Compare PSNR, NRMSE of original input / RL deblurred input / network output
- ROI based SD compare to assess the noise characteristics of soft tissue
- Additional verification for clinical practice : recover 1mm slice from 5mm slice
  - Data : training - 56 patients(9302 slices) / test – 14 patients(2338 slices)
  - CT : 512 X 512 X (180-200)
  - Verification : ROI based SD comparison & 5 – fold cross-validation

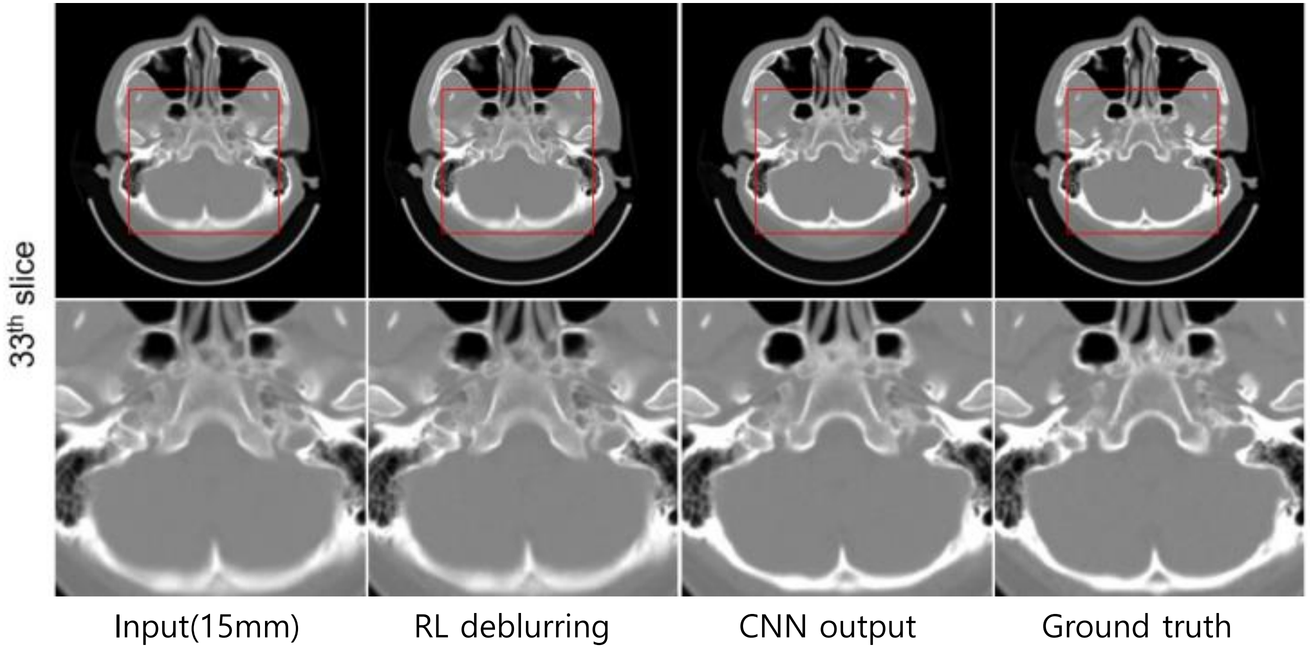
Results

1. Recovery from 3mm slice from 15mm slices

   1) image

   – Remarkable improvement : **Deblurring of boundaries of bone structures and air cavities**

   – Computation time : 30ms per slice



Results

2) Quantitative analysis of PSNR, NRMSE

- 40% higer PSNR, lower NRMSE than input(thicker slice)

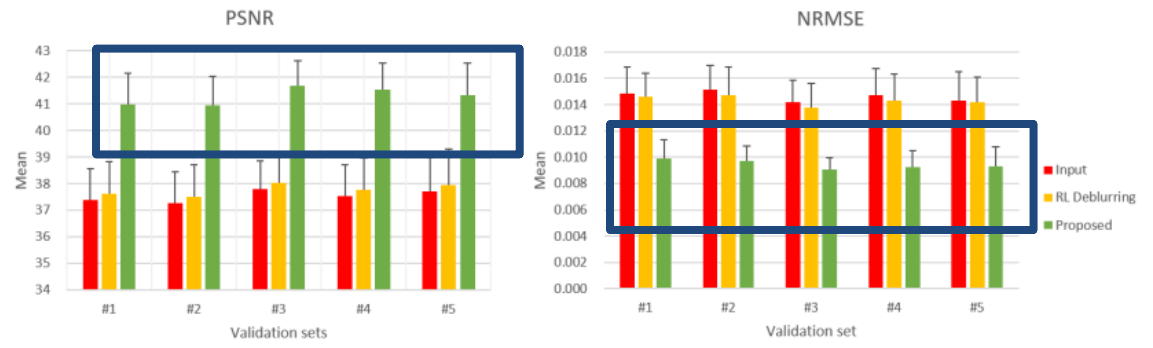
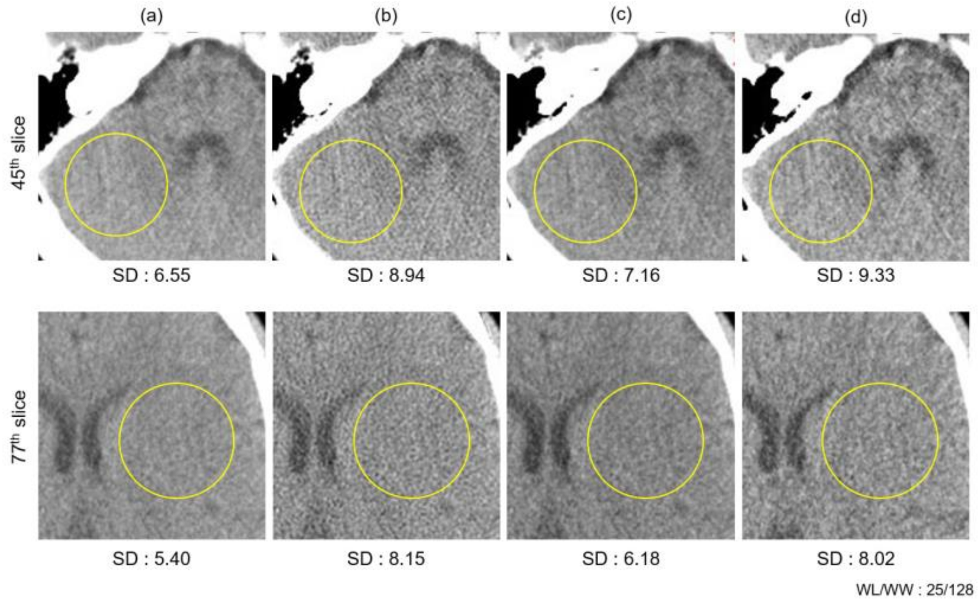


Figure 5. Cross-validation result with mean and standard deviation for PSNR and NRMSE in recovering the 3-mm CT data from 15 mm slices.

3) ROI based standard deviation

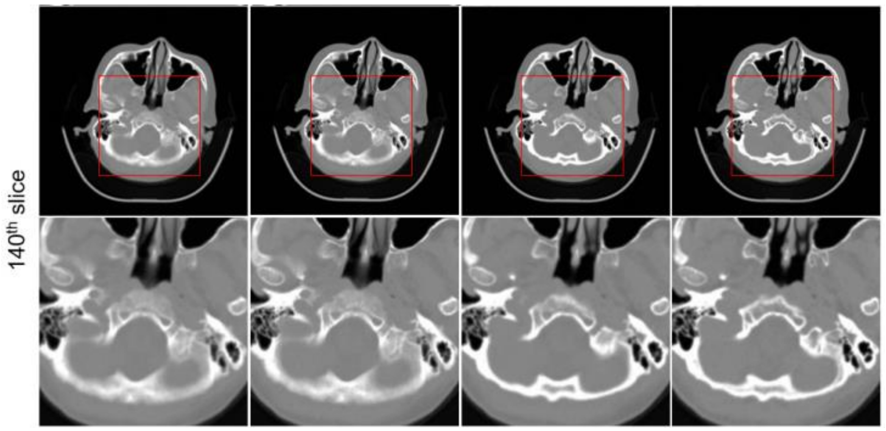
- RL deblurring increase noise level



Results

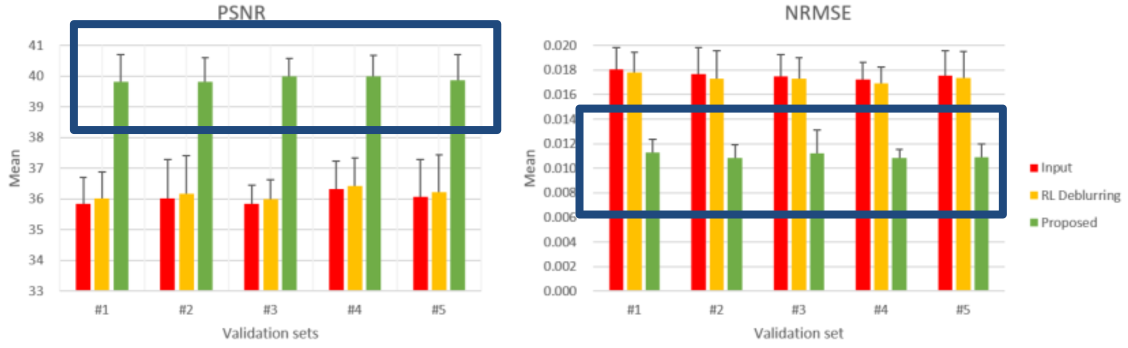
1. Recovery from 1mm slice from 5mm slices

1) Image

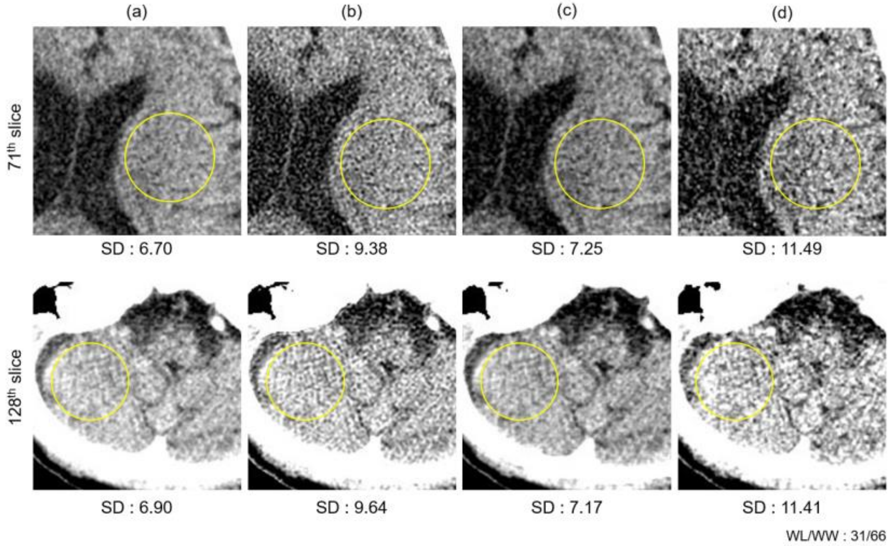


Input(15mm) RL deblurring CNN output Ground truth

2) PSNR & NRMSE



3) ROI based standard deviation



Discussion

- Show the potential of deep-learning based super-resolution of x-ray CT images
  - Thick CT image : higher SNR (higher number of x-ray quanta per voxel)
  - → CNN generated output : indistinguishable from ground truth

- Not analyze the impact of super-resolution on lesion detectability
  - Useful for previewing CT image reconstruction

- Thick CT slice : low soft-tissue contrast, high noise level
  - Reduce artifacts(insufficient angular sampling)

- Limitation
  - The proposed method use 2D CNN
  - CNN super-resolution : expect to reproduce small structure

Discussion

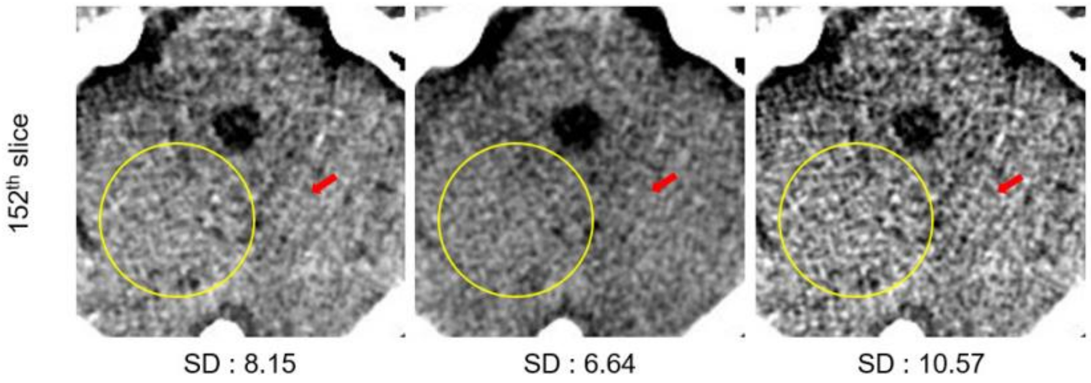
- Spatial resolution of CT image : determined by the size of detector elements
  - Development of new detector : time intensive, high cost
  - Smaller detector element : increase noise
  - sharpening (HPF) / deblurring (Laplacian filter / Richardson-Lucy algorithm) → noise↑, not as effective as proposed method
- Focus on image resolution and contrast recovery in the degraded images

→ proposed method : kind of deconvolution operation for axial spatial resolution recovery

⇒ CNN based super resolution technique : improve medical image quality & reduce scan time

- Noise reduction in CT : important for reducing radiation dose & improving diagnostic power

→ proposed method : effective in noise reduction



SAFIRE | Proposed method | Ground truth