CT image denoising using patch-based reinforcement learning with frequency-selective actions

Gwangsoo Han[1], Seungwan Lee[1,2*]

[1]Dept. Medical Science, Konyang Univ., Daejeon, South Korea

[2]Dept. Radiological Science, Konyang Univ., Daejeon, South Korea

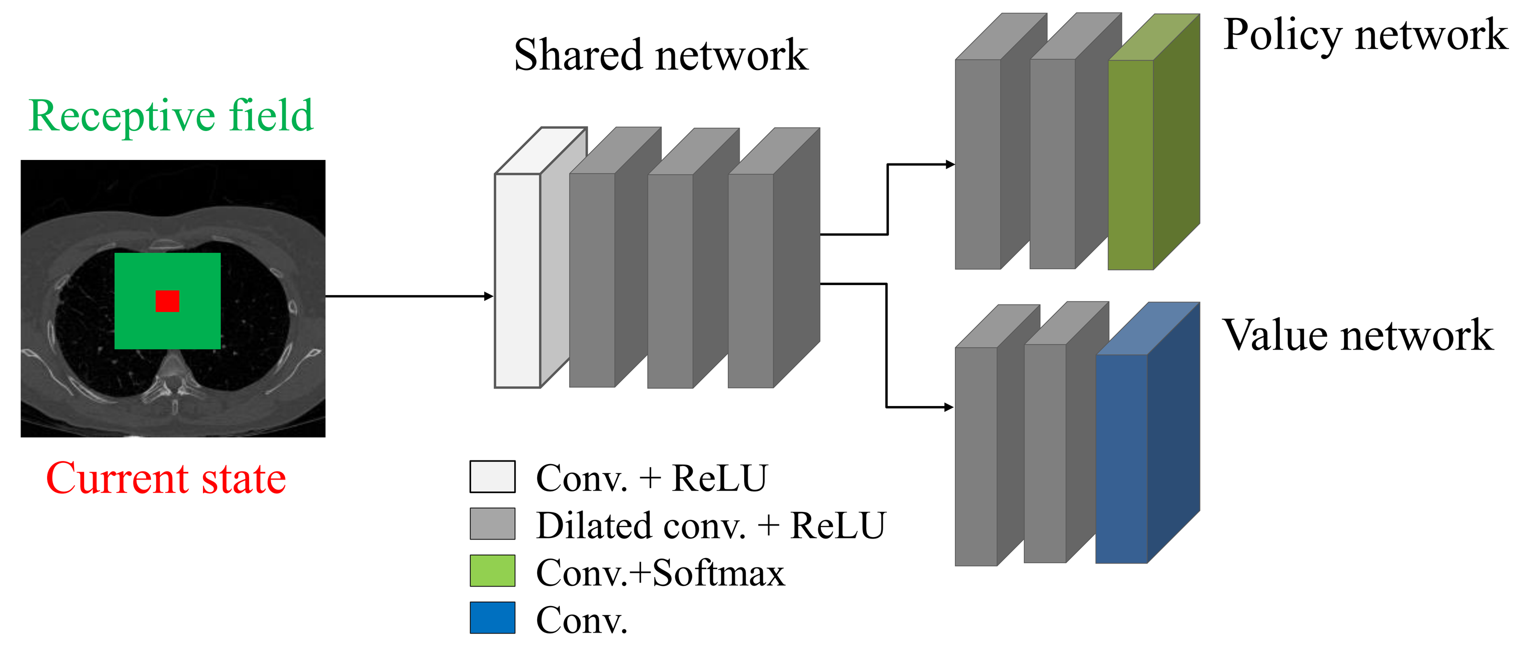
Radiation Physics & Imaging Lab.

INTRODUCTION

- **Pixel-wise reinforcement learning (PixelRL)**
  - Having a sequential decision-making strategy using current states
  - Assigning agents in each image pixel [1]
  - Enabling the pixel-wise manipulation using a fully-connected network
- **Low-dose CT**
  - Reducing radiation dose to minimize patient exposure
  - Leading to noise increment and anatomical structure distortion
- **CNN- and Transformer-based denoising methods**
  - Causing an over-smoothing effect, eventually degrading spatial resolution
- **Aims of this study**
  - Proposing a PixelRL-based CT denoising framework.
  - Extending an action set by adding frequency-selective filters.
  - Expecting noise suppression and structural preservation through pixel-wise policies.

MATERIALS AND METHODS

- **Data preparation**
  - Chest CT images of 70 patients from the publicly available dataset [2]
  - Adding the Gaussian noise with zero mean and 20 standard deviation for simulating a low-dose condition
  - Extracting random patches from a training image
  - Data augmentation with horizontal flipping and random rotation
- **Network architecture and action set** (Fig.1, Table 1)
  - Shared backbone: Convolution layer & dilated convolution layers for allowing a large receptive field while preserving spatial resolution
  - Sharing the extracted features in policy and value networks for the action selection and state evaluation, respectively
  - Policy & value networks: Dilated convolution layers & convolution layer
  - A total of 11 pre-defined actions: pixel-wise intensity adjustments (3), spatial-domain filters (6) and frequency-selective filters (2)



Shared network	Policy network	Value network
Layer (Kernel size, Dilation factor)	Layer (Kernel size, Dilation factor)	Layer (Kernel size, Dilation factor)
Conv. + ReLU (3×3, 1)	Dilated conv. + ReLU (3×3, 3)	Dilated conv. + ReLU (3×3, 3)
Dilated conv. + ReLU (3×3, 2)	Dilated conv. + ReLU (3×3, 2)	Dilated conv. + ReLU (3×3, 2)
Dilated conv. + ReLU (3×3, 3)	Conv. + Softmax (3×3, 1)	Conv. (3×3, 1)
Dilated conv. + ReLU (3×3, 4)		

Fig. 1. Network architecture of the proposed PixelRL framework.

- Frequency-selective filters

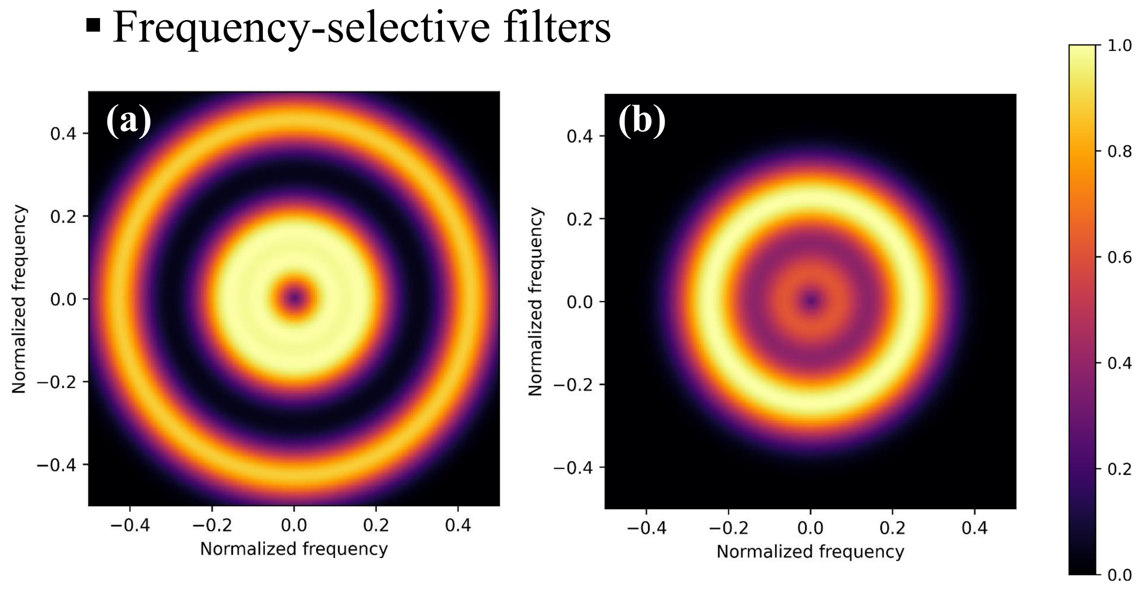


Fig. 2. Visualization of the frequency-selective filters used in the proposed PixelRL framework. (a) includes higher frequency peaks than (b), while (b) mainly includes low- and mid-frequency peaks.

Table. 1. Actions for noise suppression and structural preservation.

No.	Action	Standard deviation
1	Pixel value + = 1	–
2	Pixel value − = 1	–
3	Do nothing	–
4	Frequency-selective filter 1	–
5	Frequency-selective filter 2	–
6	Box filter	–
7, 8	Bilateral filter 1, 2	$\sigma = 1.0, \sigma = 0.1$
9	Median filter	–
10, 11	Gaussian filter 1, 2	$\sigma = 1.5, \sigma = 0.5$

RESULTS

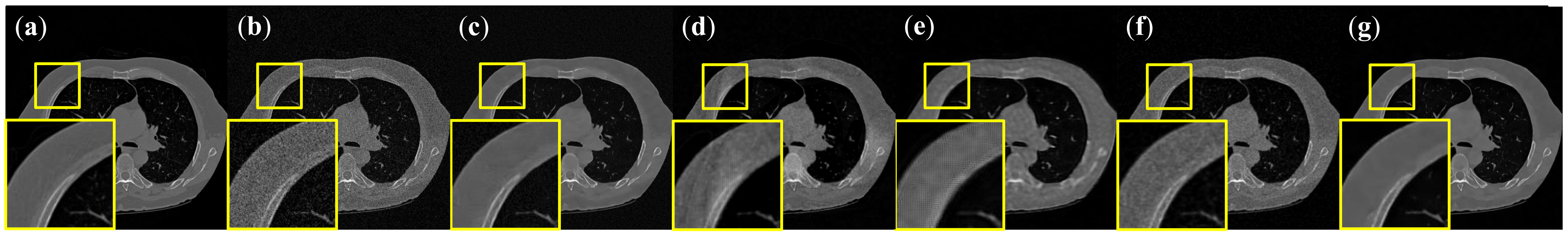


Fig. 3. Output CT images with a 32 × 32 patch. (a) GT and (b) input noisy CT images. (c)–(g) show the results from DnCNN, UNet, FFDNet, SwinIR and the proposed PixelRL, respectively.

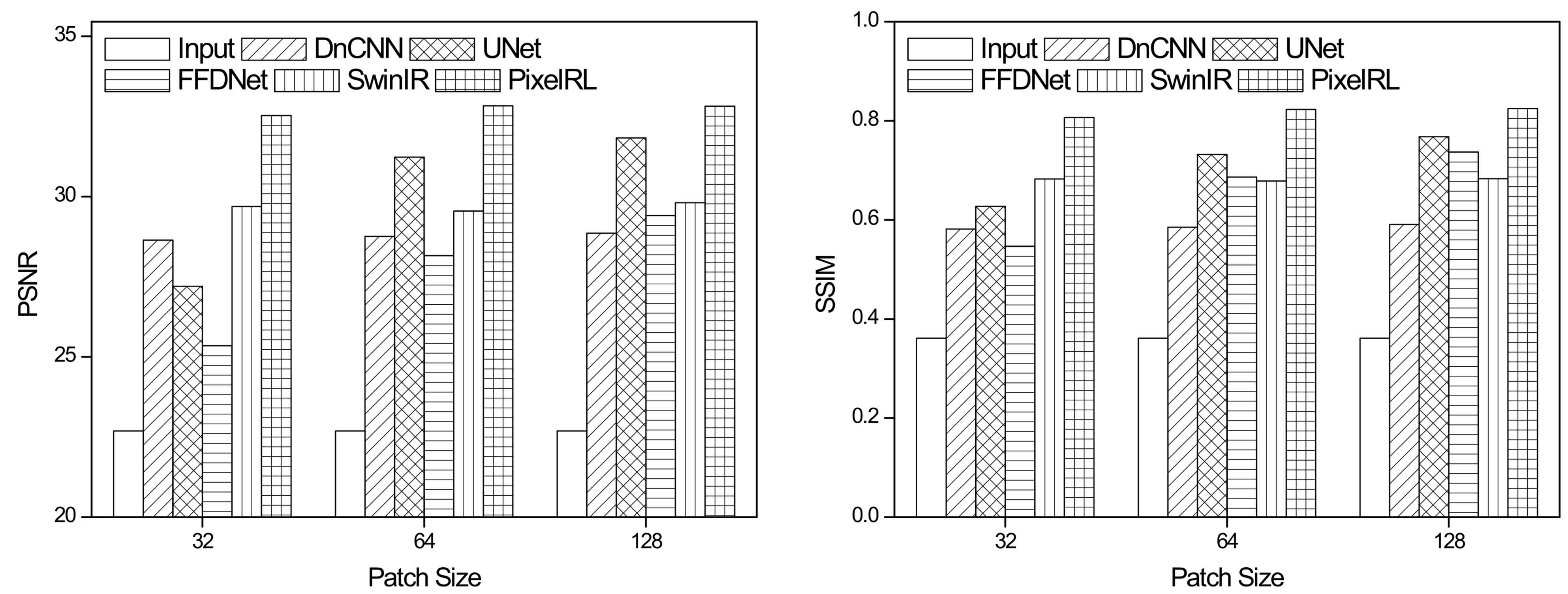


Fig. 4. Comparison of PSNR and SSIM for the output CT images obtained by different trained models using 32 × 32, 64 × 64, and 128 × 128 patch sizes.

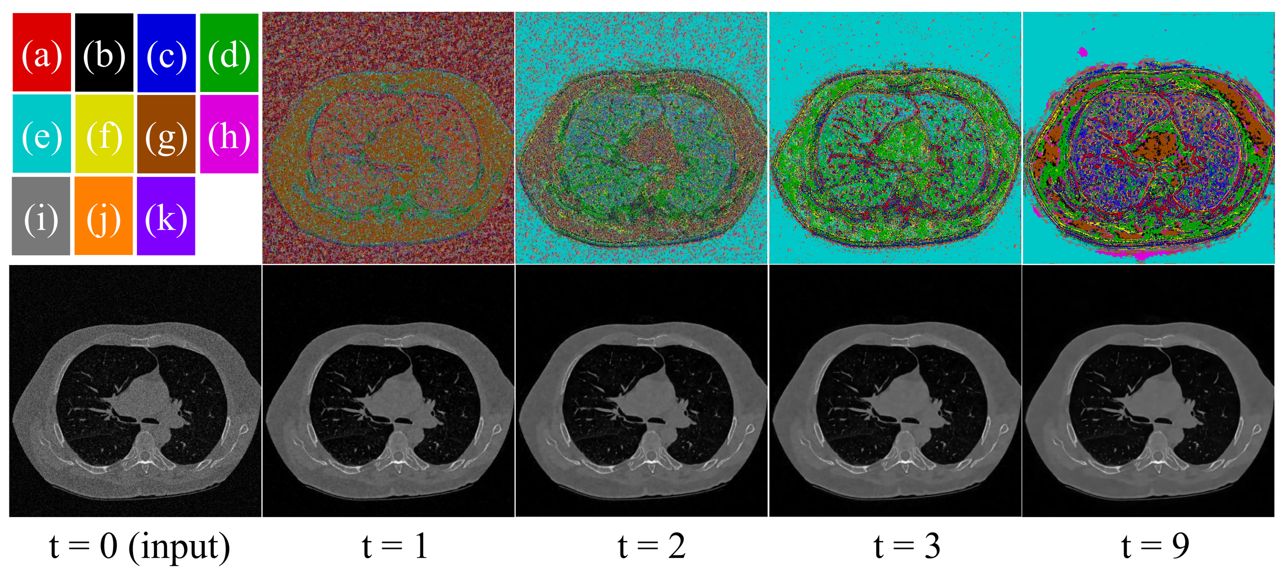


Fig. 5. A top row visualizes the pixel-wise actions used in the PixelRL framework at each episode *t*. A bottom row showed the result CT images at each episode *t*. The action set included (a) pixel value increment (+1), (b) no operation (do nothing), (c) pixel value decrement (−1), (d) Gaussian filter with the standard deviation σ = 0.5, (e) bilateral filter with the standard deviation σ = 0.1, (f) median filter, (g) Gaussian filter with the standard deviation σ = 1.5, (h) bilateral filter with the standard deviation σ = 1.0, (i) box filter, (j) frequency-selective filter 1 and (k) frequency-selective filter 2.

CONCLUSIONS

- This study demonstrated that the proposed PixelRL framework with frequency-selective filters effectively reduces noise while preserving anatomical structures in low-dose CT images.
- The proposed method consistently achieved the highest image quality for the various patch sizes and demonstrated robust the denoising performance even with a 32 × 32 patch size, highlighting the efficiency of a reinforcement learning framework for the patch-based CT image denoising.

REFERENCES

[1] R. Furuta, N. Inoue and T. Yamasaki, IEEE Trans. Multimedia, vol. 22, pp. 1704–1719, 2020.

[2] S. G. Armato III, et al., J. Med. Imaging, vol. 2, Art. no. 020103, 2015.

***Corresponding author**
Seungwan Lee
Email : slee1@konyang.ac.kr