Accuracy enhancement of CT kernel conversions using convolutional neural net for super-resolution with Squeeze-and-Excitation blocks and progressive learning among smooth and sharp kernels

Ilsang Woo **Beomhee Park** **Joon Beom Seo** **Sang Min Lee** **Namkug Kim**
Asan Medical Center
University of Ulsan College of Medicine
{bidfore, beomheep, joonbeom.seo, sangmin.lee.md, namkugkim}@gmail.com

Abstract

Reconstructing smooth to sharp images with several CT kernels have been widely used in radiology, because images reconstructed with standard and high frequency kernels are more suitable for radiologic reading, but smoother kernels of higher SNR for quantitative software. However, it is very difficult to sustain several datasets with different kernels, due to limited storage and maintenance issues. We proposed accuracy enhancement of CT kernel conversion method using convolutional neural net for super resolution with SE block and progressive learning among smooth and sharp kernels. Our CT kernel conversion method showed significant enhancement and could be applicable to actual clinical environment for radiologic reading and quantitative SW without duplicated datasets.

1 Introduction

Computed tomography (CT) images have been widely utilized and are reconstructed by various kinds of kernels from smooth to sharp according to various kinds of clinical purposes. For example, when radiologists reading on chest CT, sharp kernel is preferred. However, when measuring on chest CT, smooth kernel is much better. In this way, reconstruction kernels are important to determine noise, signal-to-noise ratio and resolution of CT images. We proposed, therefore, a kernel conversion method based on super resolution (SR) convolutional neural networks (CNN).

Image super-resolution (SR) aims to reconstruct high resolution (HR) images from corresponding low resolution (LR) images and it is ill-posed problem in the way that space of HR images can be very large. Enhanced deepresidual networks for single image super-resolution (EDSR) [1] trained a model with simply modified residual blocks by removing batch normalization.

In medical imaging, CNN for SR has been mainly proposed to deal with low-dose chest CT reconstruction. In this study, we, therefore, proposed a kernel conversion method based on CNN with squeeze-and-excitation (SE) blocks [2] and progressive learning (PL) among smooth and sharp kernels.

2 Materials and Methods

2.1 Subjects

The dataset included CT images from 10 patients with chronic obstructive pulmonary disease (five men and five women; mean age, 63.0 $\pm$ 8.6 years) obtained with Somatom Sensation 16 (Siemens

1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

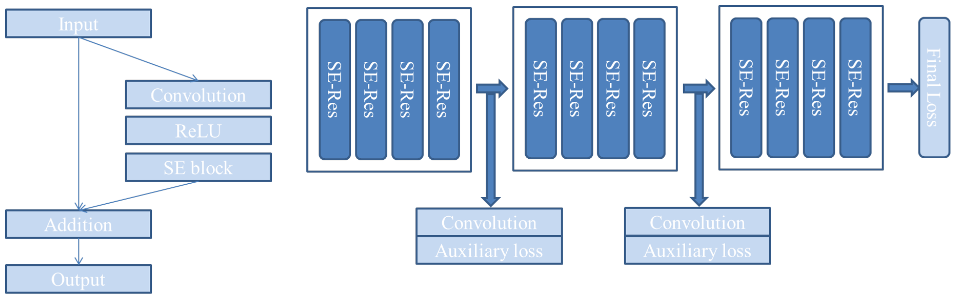

Figure 1: Network architecture with SE residual block and auxiliary losses

Medical Systems, Forchheim, Germany) in May 2016. These CT images were reconstructed with B10f, B30f, B50f, and B70f kernels, from smooth to sharp kernels.

2.2 Network Architecture

CT images generated by kernels from smooth to sharp could be represented by progressive representation. Our model, therefore, recursively stacked residual modules enhanced by SE blocks and additional auxiliary losses. These losses tune features in intermediate layers to CT images from intermediated kernels. In recent, super-resolution with SE blocks [3] is similar to our study in terms of introduction of SE blocks and PL, but our network additionally used auxiliary losses and was applied to CT kernel conversion.

2.3 Squeeze-and-Excitation Residual Block

SE blocks [2] capture dynamic and nonlinear channel relationships using *squeeze* and *excitation* operations. The *squeeze* operation generates channel-wise descriptors that represent global information of each channel in feature maps by using a global average pooling. The *excitation* operation composed of 1×1 convolutional layers and a ReLU adaptively strengthens informative features, which means that the SE blocks are self-gating mechanism.

2.4 Progressive Learning

We introduced auxiliary losses to our model trained with multiple scale at the same time. This structure is shown as Figure 1. In conversion from B10f to B70f, for example, the total loss is

$$\mathcal{L}_{Total} = \mathcal{L}_{B30f} + \mathcal{L}_{B50f} + \mathcal{L}_{B70f}. \tag{1}$$

and it is simultaneously minimized by ADAM optimizer. Comparing single-scale reconstruction models, PL in multi-scale methods could lead to better quality due to sharing information extracted by auxiliary losses.

Table 1: Effects of SE block and progressive learning. Maximum error means that RMSE between input and target images.

Conversion	Maximum error	VDSR [4]	EDSR [1]	PEDSR	EDSR+SE	PEDSR+SE
B10f - B30f	15.67(1.46)	5.59(0.70)	5.64(0.88)	4.79(1.87)	4.72(0.86)	4.28(1.51)
B10f - B50f	56.02(8.05)	30.96(6.62)	30.87(7.41)	28.56(10.78)	29.02(6.93)	27.24(10.27)
B10f - B70f	114.49(21.09)	79.58(19.50)	78.16(20.53)	74.47(30.52)	75.57(20.12)	**71.96(29.92)**
B70f - B50f	63.33(13.94)	15.29(3.33)	11.02(2.24)	10.00(3.86)	9.31(2.2)	8.77(3.36)
B70f - B30f	102.99(20.62)	19.85(3.77)	13.45(2.55)	11.08(4.93)	10.46(2.42)	9.22(3.92)
B70f - B10f	114.49(21.09)	22.41(4.35)	11.99(2.58)	10.93(5.15)	10.05(2.22)	8.64(3.99)

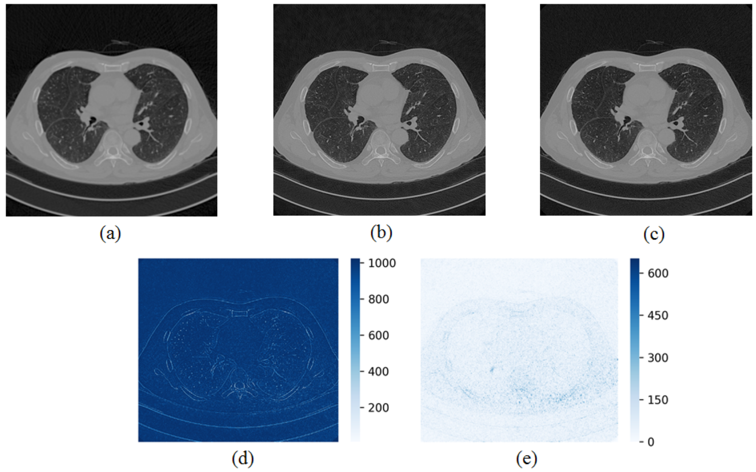

Figure 2: Results of kernel conversion from a B10f to B70f kernel. (a) B10f, (b) output, (c) B70f, (d) difference map between B10f and B70f, (e) difference map between output and B70f.

3 Experimental Results

In order to fairly compare accuracies between single-scale and multi-scale models, layer number of single-scale model is same to multi-scale model. For instance, layer number of single-scale EDSR is 4, 8 and 12 for B10f-B30f, B10f-B50f, and B10f-B70f, respectively. The performance of the proposed network was evaluated by root mean squared error (RMSE) (Table 1). EDSR+SE and PEDSR (EDSR with PL) were evaluated. SE blocks and PL improved the overall performance. In case of B10f to B70f, accuracy of PEDSR+SE is significantly better than that of EDSR. Figure 2 illustrates results of PEDSR+SE.

4 Discussion and Conclusion

In this study, we proposed a multi-scale SR model for kernel conversion with SE blocks and PL, which is more efficient and accurate than single-scale models. This study has limitations to experiment on CT images from a single CT vendor. For further study, this model need to be tested on CT images from multi-vendor from multi-center study.

Acknowledgments

This work was supported by the Industrial Strategic technology development program (10072064) funded by the Ministry of Trade Industry and Energy (MI, Korea) and by Kakao and Kakao Brain corporations.

References

[1] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.

[2] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. *arXiv preprint arXiv:1709.01507*, 2017.

[3] Xi Cheng, Xiang Li, Ying Tai, and Jian Yang. SESR: single image super resolution with recursive squeeze and excitation networks. *arXiv preprint arXiv:1801.10319*, 2018.

[4] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*, 2016.