Network slimming for compressed-sensing cardiac cine MRI

Seong Jae Park, Chae Guk Lim, and Chang-Beom Ahn
Department of Electrical Engineering, Kwangwoon University, Seoul, Korea

Correspondence
Chang-Beom Ahn, Department of Electrical Engineering, Kwangwoon University, Seoul, Korea. Email: cbahn@kw.ac.kr

Introduction: Neural networks (NNs) have been successfully applied to various applications in medical imaging [1]. The U-net [2] is a convolutional NN that was originally developed for biomedical image segmentation. It won the IEEE international symposium on biomedical imaging (ISBI) Cell Tracking Challenge in 2015. Several applications have been developed using the U-net, for example, segmentation [3] and reconstruction [4]. Despite the excellent performance of the U-net, its heavy structure often restricts its applications in an environment with limited computational resources. In this study, the U-net is attempted to be slimmed for compressed-sensing (CS) cardiac cine magnetic resonance imaging (MRI) [5], without performance degradation. The training time and the number of trainable parameters are evaluated as the figures of merit for slimming.

Data: Cardiac cine data of eight healthy volunteers were measured using the balanced steady-state free precession (SSFP) sequence in a 3.0T Siemens system with the repetition and echo times of 3.88 and 1.94 ms, respectively. A five-channel phased array coil with a sensitivity encoding (SENSE) factor of 2 was used. The k-space data without undersampling were reconstructed to prepare the ‘ground truth’ (GT) image using inverse Fourier transform and SENSE reconstruction. Undersampled data were generated by undersampling the channel data along the phase encoding (PE) axis with the acceleration factors (AFs) of 2, 3, 4, and 8. The missing data were first interpolated from the data at adjacent cardiac frames and reconstructed to prepare the ‘initial reconstructed (IR) image’. The measured data for each volunteer had 12 slices and 16–24 cardiac frames. The reconstructed image has a size of 256 \times 256 in the transverse plane. The difference image between the GT and IR is normalised to be transverse plane. The difference image between the GT and IR is normalised to be

Result: The slimming efficacy was evaluated using the number of trainable parameters and training time. The evaluation was performed with a computer equipped with an Intel® Xeon® CPU E5-2620 v4 @ 2.1 GHz with a memory of 128 GB, and an NVIDIA TITAN RTX GPU system
The use of the U-net for reconstructing CS cardiac cine construction of CS cine MRI with diagnostic image quality is another challenging task. The use of the U-net for reconstructing CS cardiac cine MRI has been investigated [6]. The main difficulties in using the U-net are its heavy structure and long training time.

We focused on three components to make the U-net slim: Multiple convolutions, transpose convolution, and deep layers. As multiple convolutions are redundant for some applications, we reduced them to a single convolution. Moreover, we adopted a larger kernel size (5 × 5) to compensate for the reduced range by single convolution. Although the transpose convolution has trainable kernels, it may not be very useful from the viewpoint of CS reconstruction. The transpose convolution mixes data over channels, unlike max pooling, whose operation is confined to a 2 × 2 array within the channel. Thus, the transpose convolution may also be replaced by upsampling with bilinear interpolation. As shown in Figure 3, the upsampling results in a greater reduction in the NMSE than the transpose convolution. The number of layers can be reduced to simplify the network. This may be dependent on the application of the NN. The range affected by the convolution may be expressed as $K-2^L-1$, where $K$ denotes the kernel size and $L$ is the number of layers of the network. The range of the slim-net with $K=5$ and $L=4$ is 40, similar to that of the U-net with $K=3$ and $L=5$.

Although we tested the slim-net for CS cardiac cine MRI application, the network may be generally applicable to other applications in which computational resources are limited.

**Conclusion:** We proposed the slim-net for CS cardiac cine MRI. Through quantitative analysis, the proposed network successfully realised slimming with one-eighth of the trainable parameters and approximately half of the training time of the original U-net with a slightly improved performance. The slim-net is useful for applications executed on devices or platforms with limited computational resources.

**Acknowledgements:** This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2019R1A2C2005660). The present research was also conducted by the Research Grant of Kwangwoon University in 2019.

© 2021 The Authors. Electronics Letters published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Received: 29 September 2020 Accepted: 31 December 2020
doi: 10.1049/el2.12084

**References**

1. Schlemper, J., et al.: A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans. Med. Imaging* 37(2), 491–503 (2018)
2. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015), Munich, Germany, pp. 234–241 (2015)
3. Norman, B., Pedoia, V., Majumdar, S.: Use of 2D U-net convolutional neural networks for automated cartilage and meniscus segmentation of knee MR imaging data to determine relaxometry and morphometry. *Radiology* 288(1), 177–185 (2018)

| AF = 2 | AF = 3 | AF = 4 | AF = 8 | Average |
|---|---|---|---|---|
| U-net | NMSE | 2.81 | 4.29 | 5.33 | 8.83 | 5.31 |
| PSNR | 50.8 | 49.8 | 49.3 | 49.3 | 49.8 |
| SSIM | 0.93 | 0.91 | 0.91 | 0.90 | 0.91 |
| Slim-net | NMSE | 2.80 | 4.24 | 5.25 | 8.67 | 5.24 |
| PSNR | 51.0 | 49.7 | 49.4 | 49.5 | 49.9 |
| SSIM | 0.93 | 0.91 | 0.92 | 0.91 | 0.92 |

**Table 1. Normalised mean square error (NMSE), peak signal-to-noise ratio (PSNR), and structural similarity (SSIM) of the reconstructed images using the U-net and slim-net as a function of the acceleration factor (AF). The scale of NMSE is $10^{-3}$**

**Fig. 3** Ratios of the number of trainable parameters, training time, and normalised mean square error of the slim-net and other slim networks with respect to those of the U-net in TensorFlow 2.1.0.
4 Hyun, C.M., et al.: Deep learning for undersampled MRI reconstruction. *Phys. Med. Biol.* 63(13), 135007 (2018)

5 Park, J., et al.: Fast cardiac CINE MRI by iterative truncation of small transformed coefficients. *Invest. Magn. Reson. Imaging* 19(1), 19 (2015)

6 Park, S.J., Yoon, J.H., Ahn, C.B.: Transfer learning for compressed-sensing cardiac CINE MRI. In: Proceedings of International Society for Magnetic Resonance in Medicine, virtual, pp. 2223 (2020)