SFT-KD-Recon: Learning a Student-friendly Teacher for Knowledge Distillation in Magnetic Resonance Image Reconstruction

Author name(s) withheld EMAIL(S) WITHHELD
Address withheld

Editors: Under Review for MIDL 2023

Abstract

Deep cascaded architectures for magnetic resonance imaging (MRI) reconstruction have shown remarkable success in providing high-quality reconstruction. However, as the number of cascades increases, the improvements in reconstruction tend to become marginal, indicating possible excess model capacity, resulting in a need to potentially compress the cascaded models. Knowledge distillation (KD) is an emerging technique to compress models, in which a trained deep 'teacher' network is used to distill knowledge to a smaller 'student' network, and the student network learns to mimic the behavior of the teacher network. Most KD methods focus on effectively training the student with a pre-trained teacher that is unaware of the student model. We propose SFT-KD-Recon, a student-friendly training approach that involves training the teacher model with the student as a prior step to KD to make the teacher aware of the student's structure and capacity and offer scope for aligning representations with the student. For KD in MRI reconstruction, the teacher DC-CNN with deeper blocks collaboratively learns block-wise with the student DC-CNN with shallower blocks to ensure consistency between the teacher and the student features. This mutual learning helps the teacher to provide better transferable knowledge to the student in the KD stage. Extensive experiments on brain and cardiac datasets for five KD methods show that the proposed approach SFT-KD (i) exhibits consistent reconstruction performance and improved image quality compared to conventional KD approaches, (ii) reduces the performance difference between the teacher and the student from 1.32% in existing KD methods to 0.076%. Our framework also improves the performance consistently for MRI super-resolution.

Keywords: Knowledge Distillation (KD), Student-Friendly Teacher KD framework (SFT-KD Recon), MRI, deep cascaded convolutional neural networks (DC-CNN).

© 2023 Anonymous.