Unsupervised Image Reconstruction using Deep Generative Adversarial Networks

Elizabeth Cole 1, Frank Ong 1, John Pauly 1, Sheyas Vasanavada 2
1 Electrical Engineering, 2 Radiology, Stanford University

Purpose: Many techniques exist for reconstruction in accelerated imaging, such as compressed sensing (CS), parallel imaging, and various deep learning (DL) methods. Most of these DL techniques require fully-sampled ground truth data for training. This poses a problem for applications such as dynamic contrast enhancement (DCE), 4D flow, etc. This paper addresses this issue using a Needleman-Wasserstein distance to calculate the Wasserstein GAN with gradient penalty (WGAN-GP) with modified penalty and data gradients to backpropagate small negative gradients into the generator. Instead, we compare to CS reconstruction and qualitatively evaluate the sharpness of the vessels and other anatomical structures in the generated images.

Methods: We have created a DL framework to reconstruct MR images using only undersampled datasets for training. The input to this network is an undersampled two-dimensional image and the output is a reconstructed two-dimensional image (Figure 1). Complex images from zero-filled reconstruction of undersampled data are the input to the generator. This reconstruction network attempts to generate an image, $X_1$, of higher quality than the input. Next, a sensing matrix comprised of coil sensitivity maps generated using ESPIRiT, an FFT, an undersampling mask, and an IFFT is applied to $X_1$ to generate measured image $Y$. A different image from the same dataset of undersampled images acts as the real measured image $Y$. Finally, the discriminator differentiates between generated and real measured images.

The loss functions of the generator and discriminator originate from the Wasserstein GAN with gradient penalty (WGAN-GP). Here, $D_{loss} = D(fake\ measurement) - D(real\ measurement) + \text{gradient penalty and}\ G_{loss} = D(fake\ measurement)$. An unrolled network based on the Iterative Shrinkage-Thresholding Algorithm (ISTA) is used as the generator architecture (Figure 2a), and the discriminator is shown in Figure 2b. We tested the framework in two scenarios. The first was fully sampled 3T knee images acquired using a 3D FSE CUBE sequence with proton density weighting including fat saturation. Subjects were used for training; each subject had a complex-valued volume of size 320x320x256 that was split into axial slices. Because a fully-sampled ground truth exists for this scenario, we can quantitatively validate our results. We created undersampled images by applying pseudo-random Poisson disc variable-density sampling masks to the fully-sampled k-space. Although we initially used fully-sampled datasets to create sub-sampled datasets, the generator and discriminator are never trained with fully-sampled data. The second scenario consists of dynamic contrast enhanced (DCE) acquisitions of the abdomen, with a fat-suppressed butterfly-navigated free-breathing SPGR acquisition with an acceleration factor of 5. 886 subjects were used for training. Because DCE is inherently undersampled, we have no ground truth to assess performance. Instead, we compare to CS reconstruction and qualitatively evaluate the sharpness of the vessels and other anatomical structures in the generated images.

Results: Representative results in the knee scenario are shown in Figure 3 with an undersampling factor of 2 in both $k$, and $k_c$. The generator markedly improves the image quality by recovering vessels and structures that were not visible before but uses no ground truth data in the training. Figure 4 displays a comparison between our results and CS with L1-wavelet regularization on our test dataset. Representative DCE results are shown in Figure 5. The generator improves the image quality by sharpening and adding more structure to the input images.

Discussion: In the knee scenario, the generated images are quite similar to the ground truth. In the DCE application, the generated images are sharper than those reconstructed by CS and have higher diagnostic quality. The main advantage of this method over existing DL reconstruction methods is obviating the need for fully-sampled data. Additionally, this method produces better quality reconstructions compared to baseline CS methods. While the method has been demonstrated here for reconstructing undersampled fast spin echo and DCE datasets, the discriminator can act on any simulated loss measurement as long as the measurement process is known. Therefore, this method could also be useful for high noise environments where the acquisition of high SNR data is difficult. Other adverse situations where ground truth data are precluded include real-time imaging due to motion and arterial spin labeling due to low SNR. Further applications where it is hard to fully sample are time-resolved MR angiography, cardiac cine, low contrast agent imaging, EPI-based sequences, diffusion tensor imaging, and fMRI.

Conclusion: Our method has applications in cases where fully-sampled datasets are difficult to obtain or unavailable. We will continue refining the quality of the framework as applied to these scenarios and other applications.