Super Resolution in Medical Imaging using a Generative Adversarial Network

Rewa Sood, Stanford University; Binit Topiwala, MS; Karthik Choutagunta Nagaraj, MS; Rohit Sood, MD, PhD; Mirabela Rusu, PhD, MS

Introduction
Acquiring high-resolution (HR) magnetic resonance (MR) images requires the patient to remain still for long periods of time, which causes patient discomfort and increases the probability of motion induced image artifacts. A possible solution is the Super Resolution Generative Adversarial Network (SRGAN) (Ledig et al, CVPR 2016). GANs are deep learning frameworks that contain a generator network, which generates images from LR input and a discriminator network, which discriminates between real and generated images. In this work, we apply the SRGAN framework to MR images of the prostate to improve the in-plane resolution by factors of 4 and 8.

Methods
The generator has the difficult task of creating new images from the LR images while the discriminator has a simple classification task. Thus, to ensure that the discriminator does not dominate and prevent the generator from learning, we pretrain the generator component of the SRGAN for 20 epochs followed by training both the generator and discriminator networks for 50 epochs. In order to evaluate the quality of the proposed model, we compare the SRGAN against nearest neighbor and bicubic interpolation, Sparse Representation, SRCNN (SR Convolutional Neural Network), SRRResNet (SR Residual Network), and the original HR image.

Results
The outputs from the SRGAN are visually closer to the original HR ground truth images (Figure 1). The quantitative metrics used in this study are PSNR (Peak Signal to Noise Ratio), SSIM (Structural SIMilarity) and MOS (Mean Opinion Score). The SSIM ranges from -1 to 1, with 1 meaning that the two compared images are structurally identical. The MOS ranges from 1 to 5 with 1 being bad quality and 5 being excellent quality. In this case, quality was measured with respect to edge fidelity. Interestingly, the SRGAN output perceptual accuracy is not reflected in the PSNR and SSIM metrics. However, the MOS obtained from a blinded evaluation by a radiologist quantitatively supports the perceptual similarity between the SR and HR images (Table 1).

(Figure 1: Example of input and output of SRGAN compared to the ground truth image)
Table 1: Comparison of metrics for all models

<table>
<thead>
<tr>
<th></th>
<th>Bicubic</th>
<th>Sparse Rep.</th>
<th>SRCNN</th>
<th>SRResNet</th>
<th>SRGAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>21.68</td>
<td>21.82</td>
<td>24.02</td>
<td>21.03</td>
<td>21.27</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.71</td>
<td>0.74</td>
<td>0.68</td>
<td>0.70</td>
<td>0.66</td>
</tr>
<tr>
<td>MOS</td>
<td>2.6</td>
<td>2</td>
<td>2.6</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Discussion
The SRGAN output images contain high frequency information that is similar to the HR ground truth images while the Sparse Representation and SRCNN models tend to smooth the images out to achieve a high PSNR. The SRCNN is especially biased toward smoothing the image because the network backpropagation only depends on MSE loss. The SRResNet has both MSE and perceptual loss, just as the SRGAN does, yet fails to outperform the SRGAN. Clearly, the discriminator network seeks out the high frequency information that differentiates HR and LR images, thus forcing the SRGAN output to have far more high frequency details than the output of the SRResNet.

Conclusion
Overall, the SRGAN output images contain far more high-resolution details than any of the other models’ output images and produces images that appear visually closer to the HR ground truth images than the images produced by the other models, quantified by each model’s MOS.

Keywords
medical imaging, super resolution, machine learning, generative adversarial networks