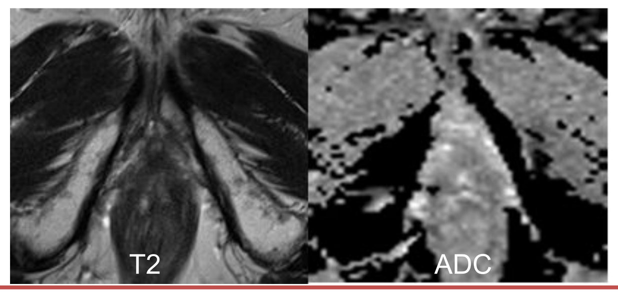
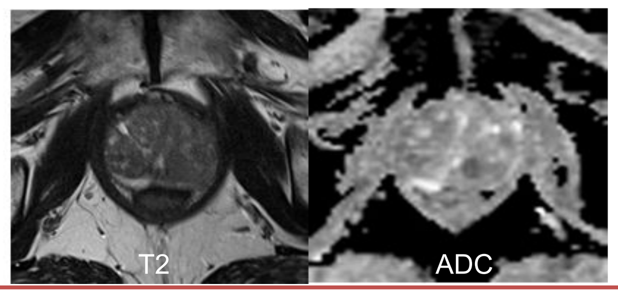
Unpaired Magnetic Resonance Image-to-Image Translation for Prostate using Cycle-Consistent Adversarial Networks

Steve Burke sfburke@stanford.edu
Vidush Mukund vmukund@stanford.edu
Brian Wilcox wilcoxee@stanford.edu

Motivation

Early detection of prostate cancer has a significant impact on the prognosis for patients. Histologies, which involve taking samples of tissue, can be high cost and high risk for patients, so MRI scans are useful for clinicians to first determine whether a patient needs a biopsy. MR image acquisition can be time consuming and involves taking both T2-weighted images and ADC images, each providing different structural information for analysis. We aim to reduce the scan time for acquiring MR images as well as the number of scans needed per patient, while still providing clinicians with all of the information needed for an accurate diagnosis.

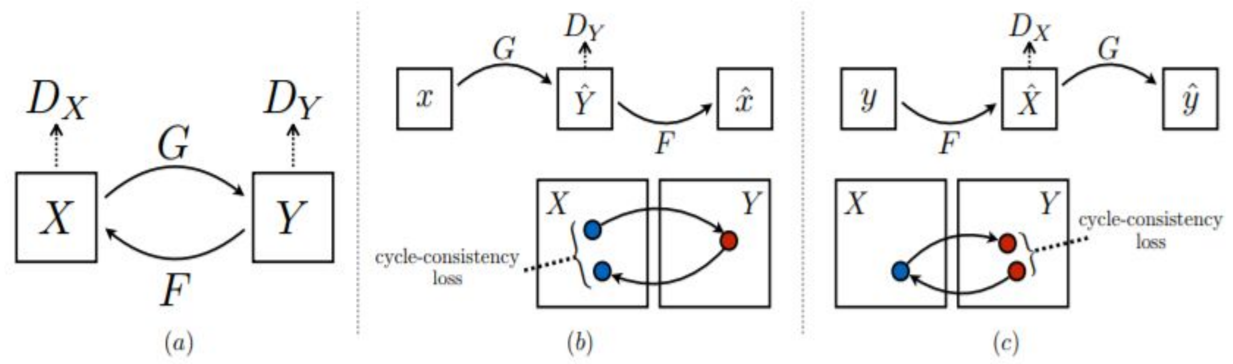




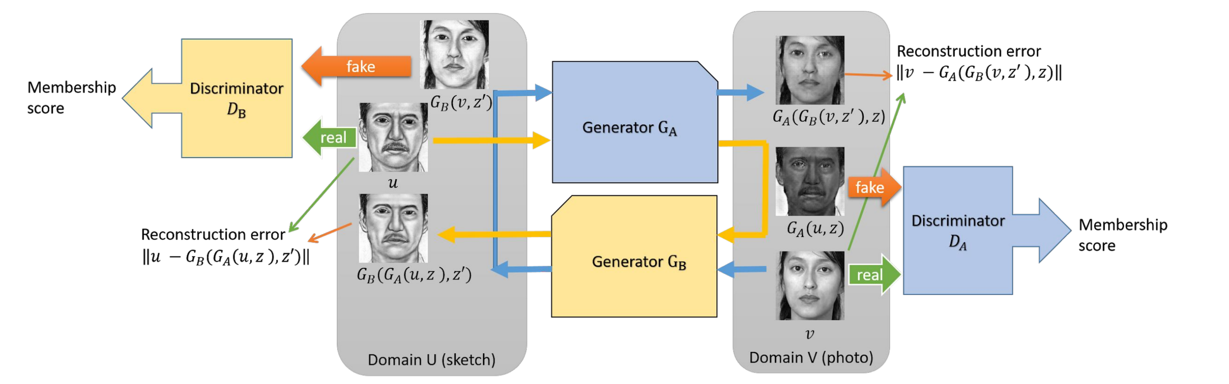
Data and Features

MRI scans were originally stored as DICOM (Digital Imaging and Communications in Medicine) images [3]. All images were normalized between 0 and 1, then scaled to the PNG range of 0 to 255. Image slices containing the prostate, bladder, and colon were cropped and resized to 256x256 pixels.In total there were 345 patients with at least 18 T2-weighted and 18 ADC images each. 300 patients' scans were grouped as the train data and 45 patients' scans were grouped as the test data. The train set was augmented by flipping the images over the vertical axis. To ensure that all images in the T2 set had corresponding images in the ADC set and vice versa for each patient, images without corresponding pairs were removed from the dataset. This pairing of images allows for comparison of original scans with generated scans of the same type of contrast.

Models



a) The CycleGANs architecture constructs two mappings, G: X → Y and F: Y → X. The discriminators are D{x} and D{y} respectively. b,c) Two cycle-consistency losses are introduced to ensure that translation from one domain to another is invertible, i.e. F(G(x)) ≅ x and G(F(x)) ≅ x. [1]



The DualGAN architecture constructs two mappings, $G_B(A)$ to B and $G_A(B)$ to A. The discriminators are D_A and D_B respectively. Two reconstruction losses are introduced to ensure that translation from one domain to another is invertible, i.e. $G_A(G_B(A)) \cong A$ and $G_B(G_A(B) \cong B$. [2]

Methods - CycleGAN

$$\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) = \mathbb{E}_{y\sim p_{\text{data}}(y)}[\log D_Y(y)] + \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log(1 - D_Y(G(x))],$$

GAN Loss for Cycle GAN

$$\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\|F(G(x)) - x\|_1] + \mathbb{E}_{y\sim p_{\text{data}}(y)}[\|G(F(y)) - y\|_1]$$

Cycle Loss for Cycle GAN

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) + \lambda\mathcal{L}_{\text{cyc}}(G, F),$$

Objective Function for Cycle GAN

Methods - DualGAN

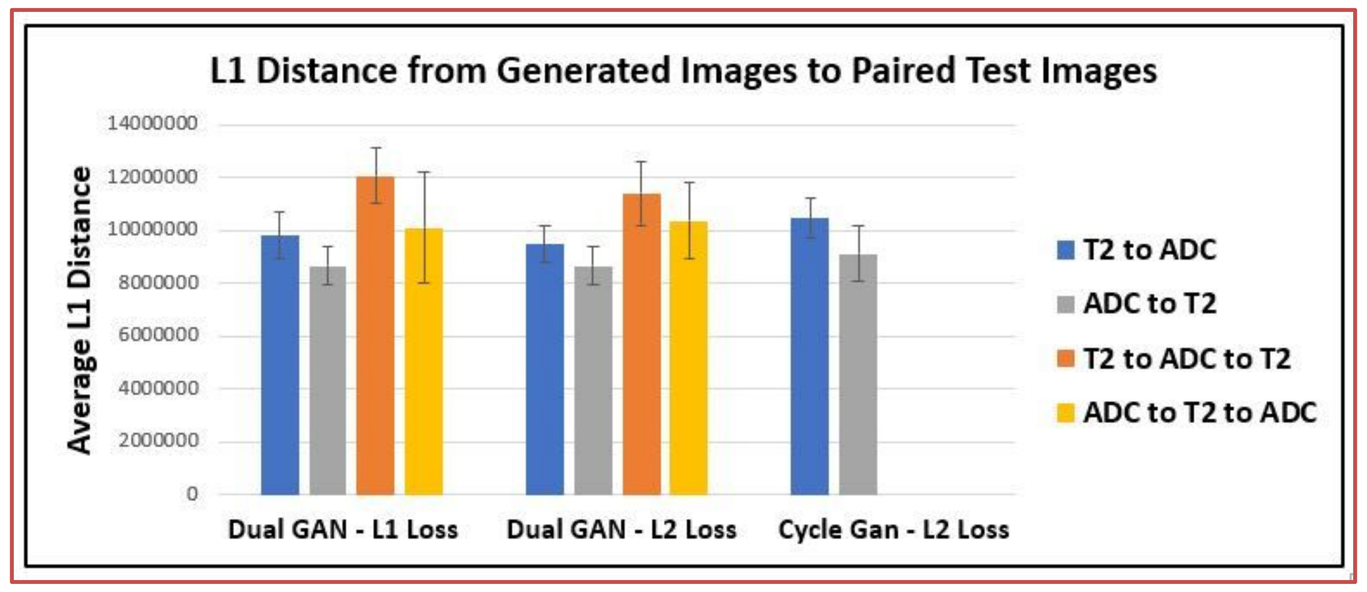
$$l^d_A(u, v) = D_A(G_A(u, z)) - D_A(v),$$
$$l^d_B(u, v) = D_B(G_B(v, z')) - D_B(u),$$

GAN Loss for DualGAN

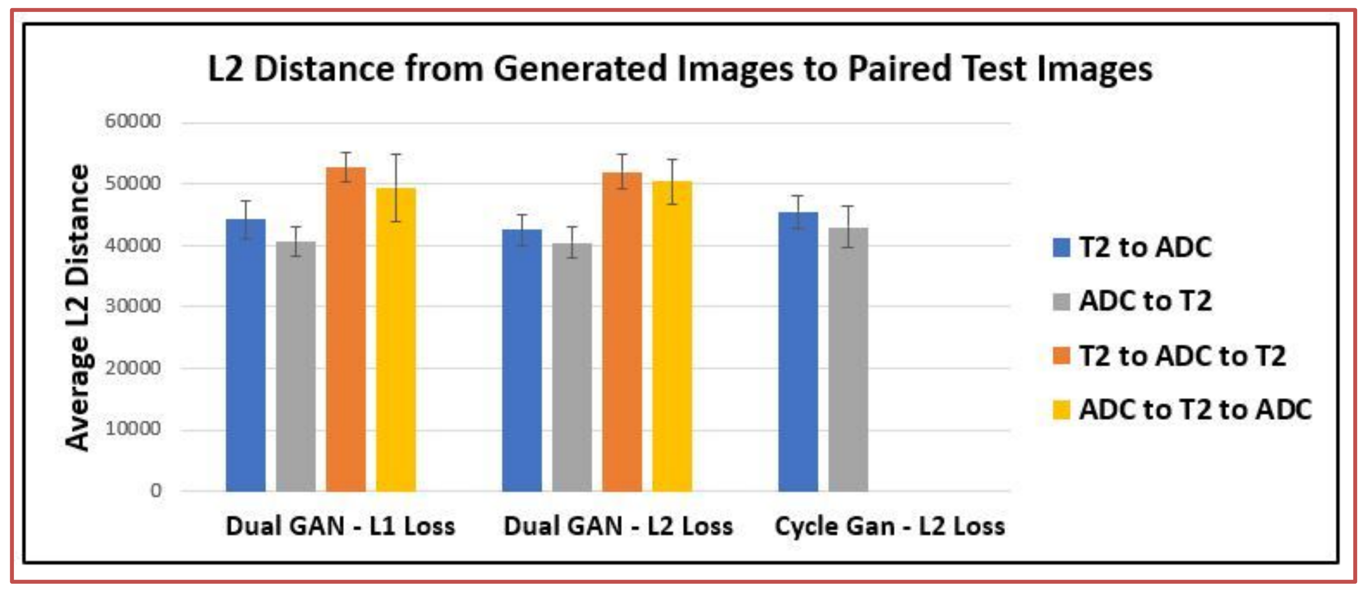
$$l^g(u, v) = \lambda_U\|u - G_B(G_A(u, z), z')\| + \lambda_V\|v - G_A(G_B(v, z'), z)\| - D_A(G_B(v, z')) - D_B(G_A(u, z)),$$

Dual Loss forGAN

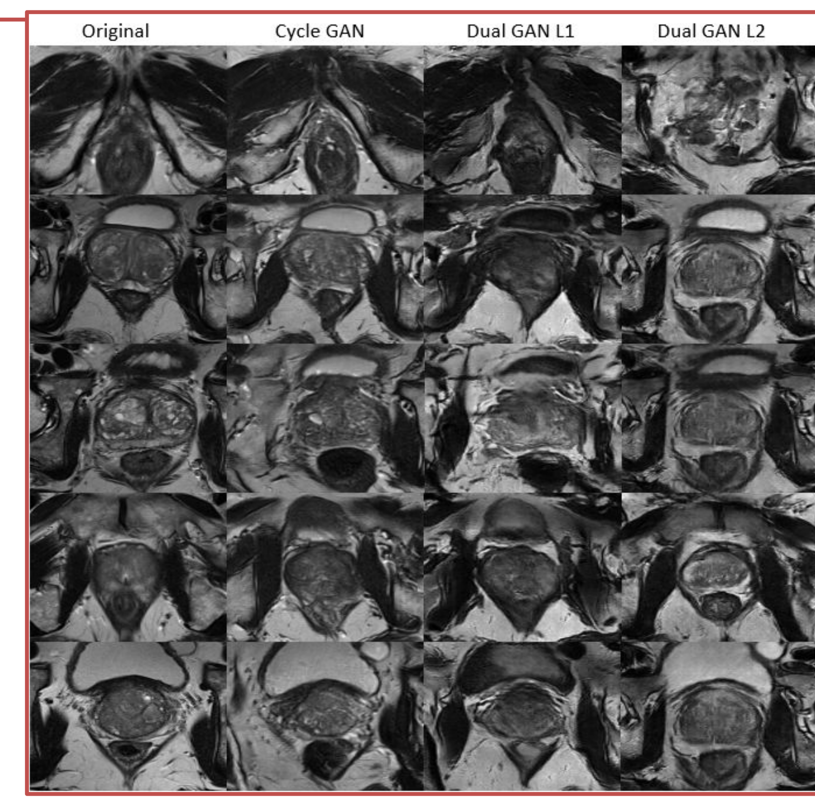
Results



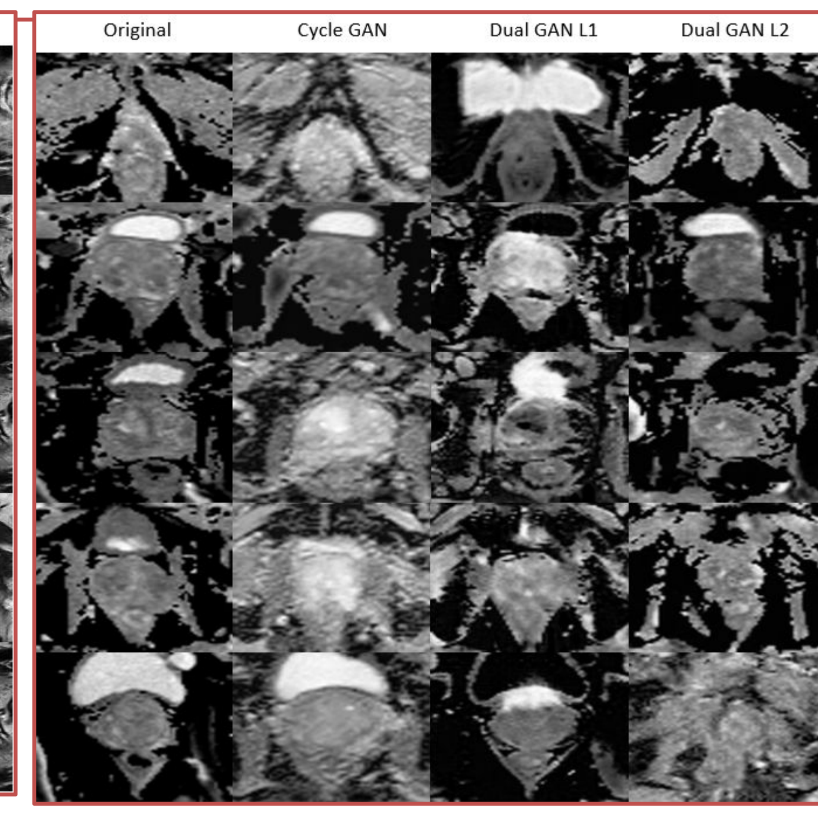
Pixelwise L1 Loss calculations for 256 x 256 Original versus Generated Images



Pixelwise L2 Loss calculations for 256 x 256 Original versus Generated Images



T2-Weighted Original Image Results



ADC Original Image Results

Discussion

The distance results show that in general the distances for the single direction cases are lower than for the dual direction cases. This is somewhat surprising, as we might expect the distance to be lower for the dual direction case, as the GAN should be able to effectively map back to the original modality that was input to the network. Additionally, the distances for the DualGAN cases appear to be slightly lower than those for the CycleGAN, although we do not have data on the dual cases for the CycleGAN.

Conclusion/Future Work

- While the L1 and L2 reconstruction losses are lower for the DualGAN, from a qualitative perspective the CycleGAN model produces images most similar to ground truth images.
- A ResNet architecture was used for the generators in the CycleGAN model. Moving forward, we are looking into using a U-Net architecture for the generators and preliminary results are promising.
- The Prostatex dataset contains some scans that do not have the prostate present. For a future iteration, the dataset should be cleaned of any images that do not have the prostate in the field of view.

References

[1] Jun-Yan Zhu*, Taesung Park*, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", in IEEE International Conference on Computer Vision (ICCV), 2017.

[2] Zili Yi, Hao Zhang, Ping Tan, and Minglun Gong "DualGAN: Unsupervised Dual Learning for Image-to-Image Translation", in IEEE International Conference on Computer Vision (ICCV), 2017.

[3] Prostatex data: https:wiki.cancerimagingarchive.net/plugins/servlet/mobile?contentId=23691656#content/view/23691656

Acknowledgements

Dr. Mirabela Rusu, The Laboratory for Integrative Personalized Medicine (PIMed)