Registration of the T2 and DWI b1200 MRI Sequences of the Prostate

Rohan Sikand[1]
[1]Urologic Cancer Innovation Lab, Stanford University School of Medicine
Canary CREST Program 2021

Introduction

Prostate cancer is the most commonly diagnosed cancer as well as the second leading cause of death in men (American Cancer Society 2020). MRI is used for diagnosis, prognosis evaluation, and treatment planning. Image registration enables scans taken at different times and with different modalities to be aligned so that medical professionals can better interpret patient scans. That is, medical image registration enables one to be able to view different scans possibly taken at different times all in one. The clinical significance is the ability to view disease progression overtime as well as to view anatomical contents in different modalities all in one.

Research Goal

The goal of this research is to develop and test methods for registering the T2 and DWI b1200 sequences of Prostate MRI scans. Within this goal, we aim to test classical approaches and a deep learning-based approach for this registration task.

Image Registration

The goal of medical image registration is to transform the desired anatomical contents of one image (moving image) into the coordinate system of another image (fixed image).

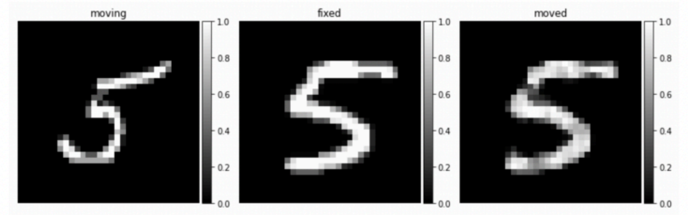
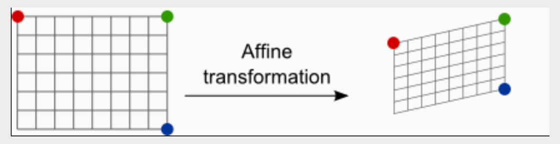
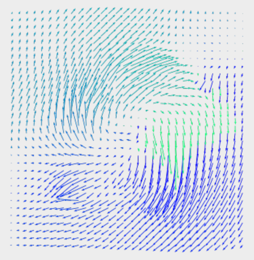


Figure 1: Image registration example showing the co-registration of two handwritten digits. The main contents of the moving image are transformed to fit the geometric space of the fixed image and its contents.

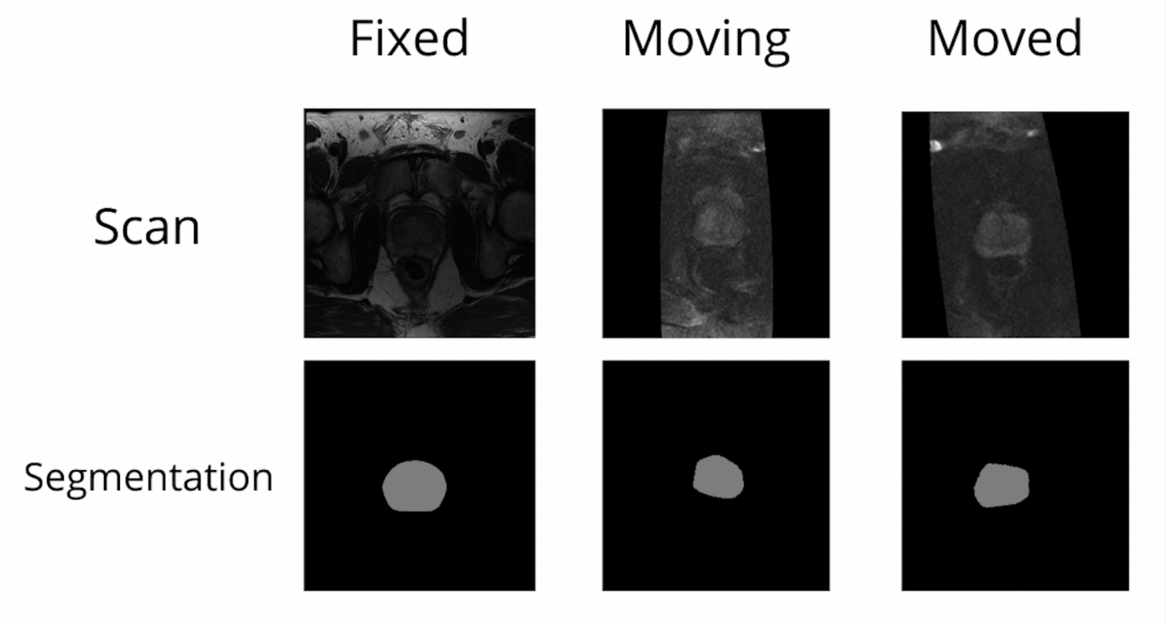
Linear transformation



Deformation field:



Below is an example of medical image registration. In this case, the goal is to align the prostate in the DWI sequence (moving) with the prostate in the T2 sequence (fixed). Segmentations are shown for a visualization of this process.

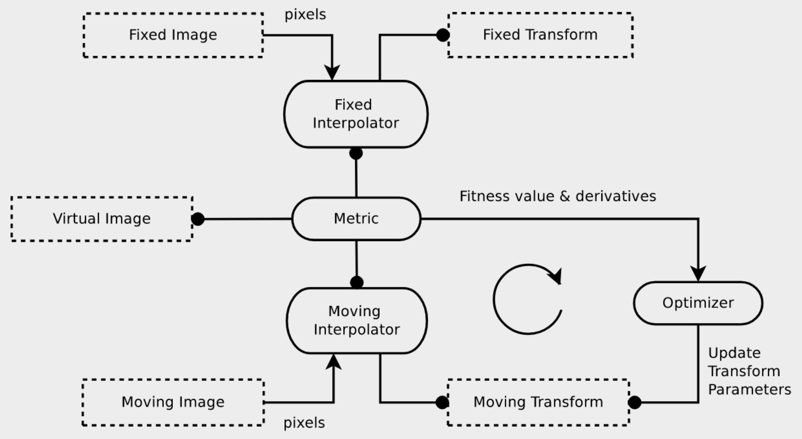


Methods

Classical registration:

Classical registration consists of defining the registration problem as an optimization task. That is, for each image pair, a loss function is defined that measures the similarity between the output and the fixed image. The optimizer then produces the moved image by selecting voxel values that result in the minimum value of the loss function.

The figure to the right (SimpleITK) shows an overview of the classical image registration process. There are four main components:
- Initialization transformation
- Similarity metric
- Optimizer
- Interpolator



For each of the experiments performed in this project, it was determined that the similarity metric, optimizer, and interpolator should be held constant (controls, details below). The independent variable is the initialization transformation. There are two main types of **transformations**: linear and deformable. In this project, we focused solely on linear transformations. The transformations used in each experiment are denoted below (right).

- **Optimizer**: Gradient Descent (100 epochs, 1e-4)
- **Interpolator**: Linear
- **Similarity metric**: Mattes mutual information

- CR-R1: Affine transformation
- CR-R2: Versor transformation
- CR-R3: Versor 3D transformation
- CR-R4: Euler 3D transformation
- CR-R5: Similarity 3D transformation
- CR-R6: Scaled Versor 3D transformation
- CR-R7: Scaled transformation
- CR-R8: Scaled Skew Versor 3D transformation

Deep learning-based registration:

As opposed to classical registration, the goal of DL-based registration is to optimize a global loss function by showing a neural network training data. We used an auto-encoder network which was trained in an unsupervised setting. The decoder outputs the registered image. The loss is calculated between the network output and fixed image. Several loss functions were tested including Dice Coefficient and Cosine Similarity.

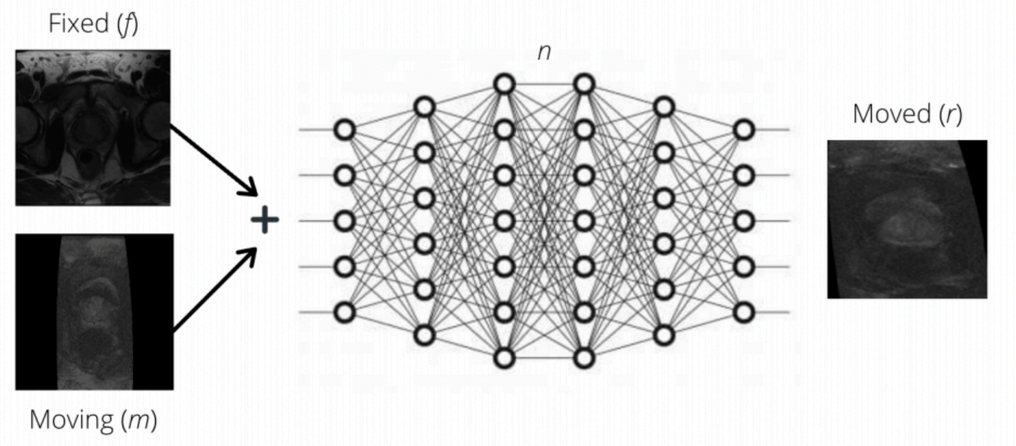


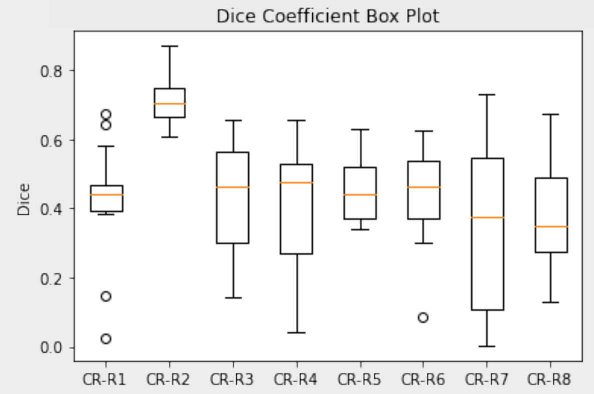
Figure 2: Visualization of the learning-based image registration base model. For each image pair, the fixed sample (f) and the moving sample (m) are concatenated together forming $f+m$. This result is then evaluated by the network n, $n(f+m)$, to obtain the moved output (r).

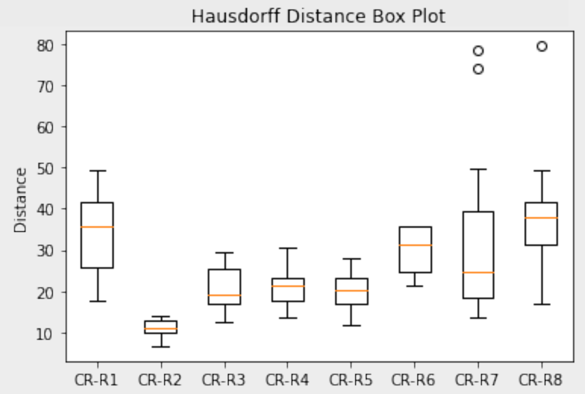
Results

Average Dice Coefficient and Hausdorff Distances between the fixed and moved segmentation masks are shown below. The Dice Coefficient measures the similarity between two images whereas the Hausdorff Distance is the longest distance between two images.

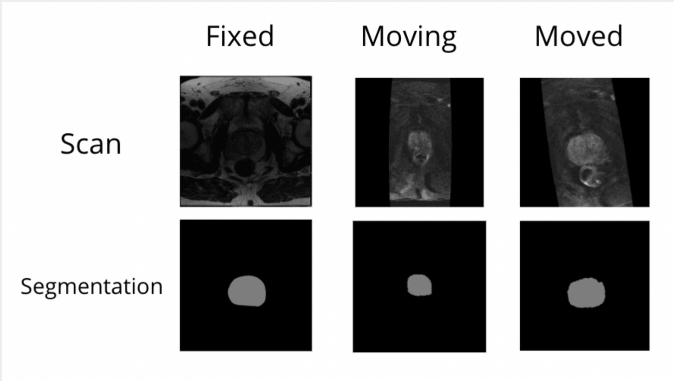
Experiment	Dice Coefficient	Hausdorff Distance
CR-R1	0.424	33.437
CR-R2	**0.709**	**11.087**
CR-R3	0.430	20.225
CR-R4	0.414	20.831
CR-R5	0.457	20.051
CR-R6	0.427	29.899
CR-R7	0.354	34.053
CR-R8	0.378	38.864

Table 1: Average scores for the metric calculations of each experiment. Each method described in the respective experiment was tested on the same subset of the data. From here, the average for each metric is calculated.

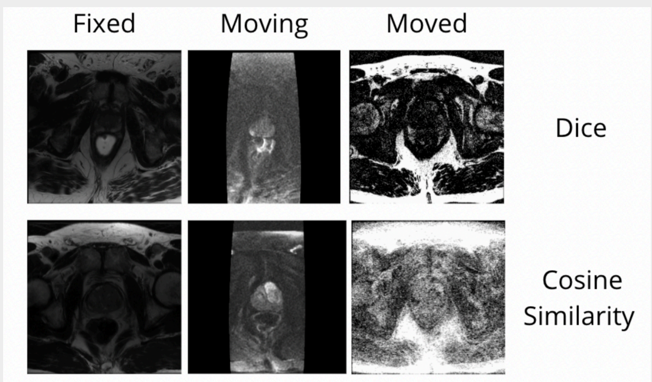




Classical registration result:



Deep learning-based registration result:



The network was unable to learn the correct pixel intensities. Moreover, DL-based registration is generally used more for the registration of images of the same modality rather than different modalities like we have here. Future research could be done to include a spatial transformation layer to preserve the pixel intensities.

Acknowledgments

I would like to thank The Canary Center at Stanford Medicine for inviting me to participate in the CREST program. I would like to acknowledge my faculty mentor Dr. Geoffrey Sonn as well as my lab mentor Dr. Richard Fan. In addition, I would like to acknowledge the following members of the lab that helped me conduct my research: Mirabela Rusu, Ph.D., Ahmed Alsinian, Ph.D., Simon John Christoph Soerenson, and Wei Shao, PhD.