Muscle Segmentation in Ultrasound Images via Convolutional Neural Networks

Sai Mandava, Yonatan Nozik,
Daniel Ho

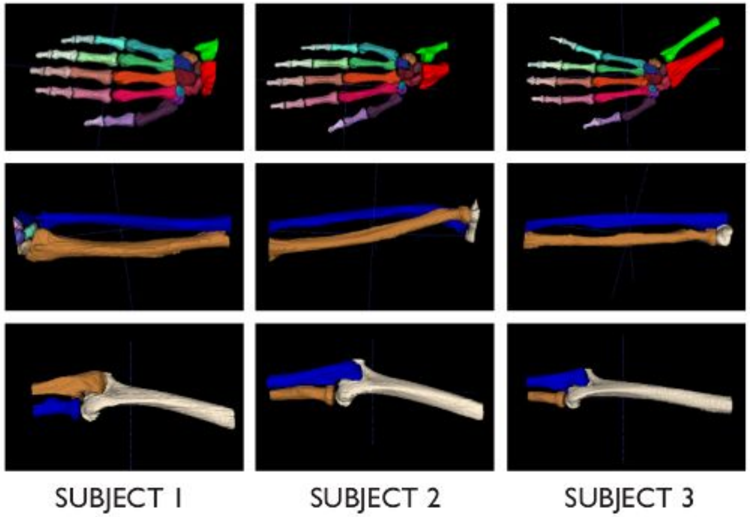
Advised by:
Laura Hallock
Ruzena Bajcsy
2018.08.06

HART Lab
Human-Assistive Robotic Technologies

Berkeley
UNIVERSITY OF CALIFORNIA

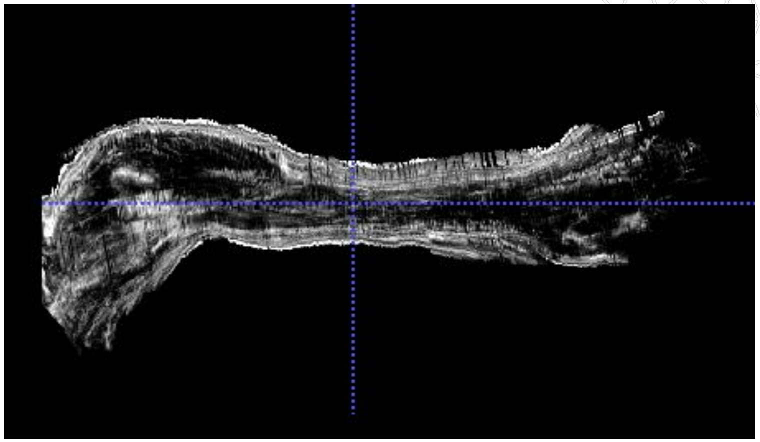
Why Segment the Arm?

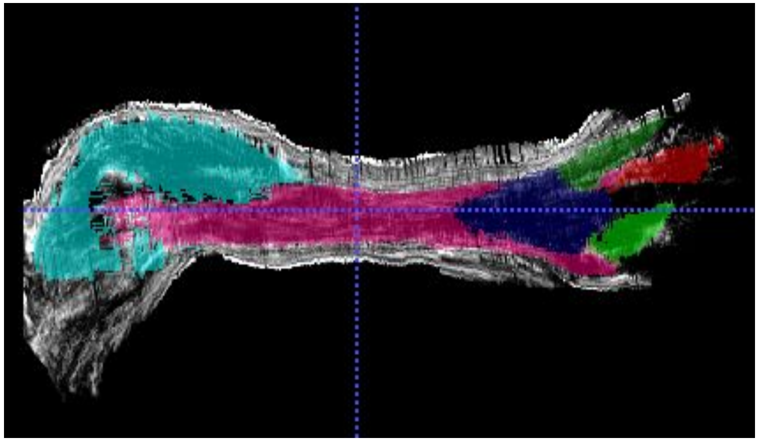
- The HART lab seeks to create safe and effective prosthetics/exoskeletal devices.
  - To do this, must have some system that can predict forces and torques produced by body
- There are some frameworks to model human dynamics
  - Many simplifying assumptions about how human muscles work
  - We don't know how accurate these models are



Why Segment the Arm?

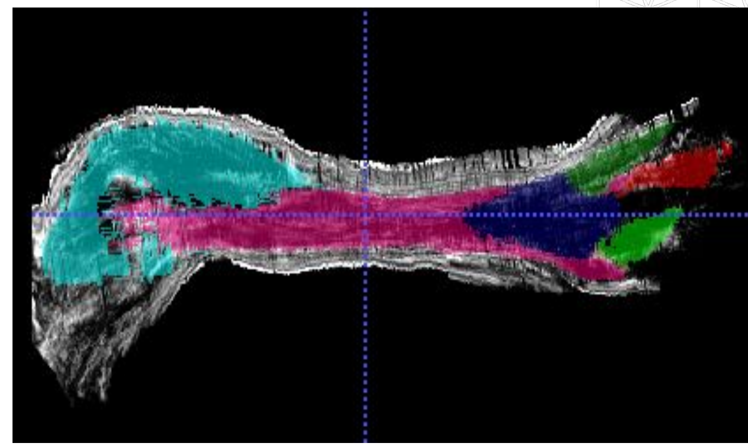
- Analyze changes in morphology across subjects
  - How do current frameworks handle these differences?
  - How does this affect ability to predict dynamics?
- Collect relevant data e.g. volume and length of muscle and bone
  - Compare across subjects

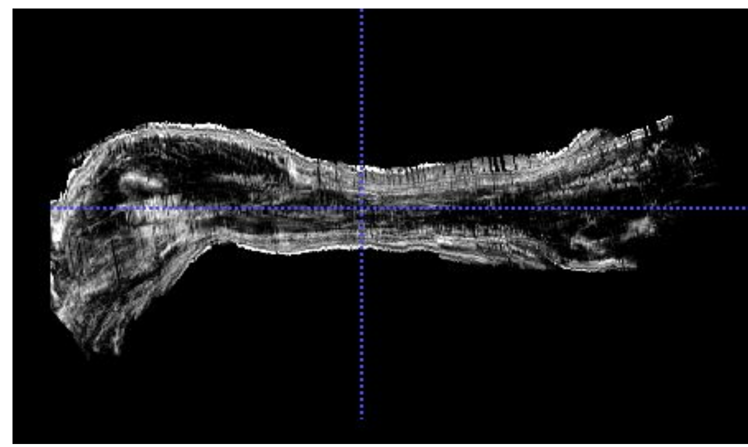




Segmenting is time-intensive

Top: Labeled cross section.
Bottom: Unlabeled "volume".

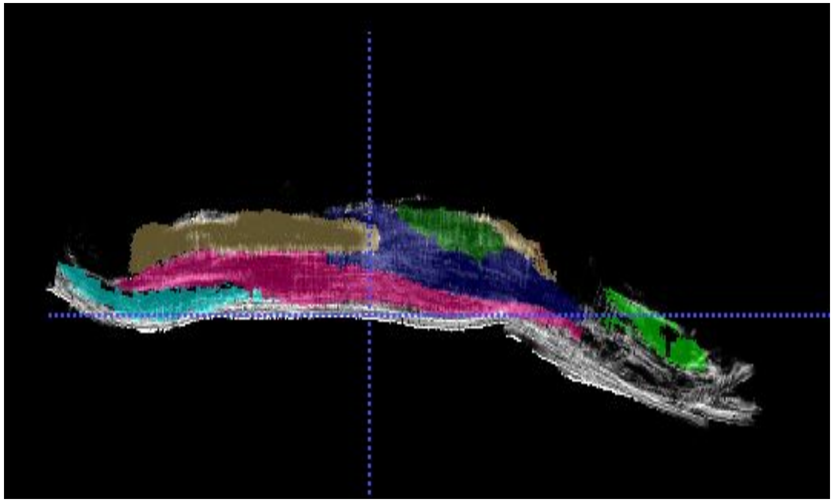


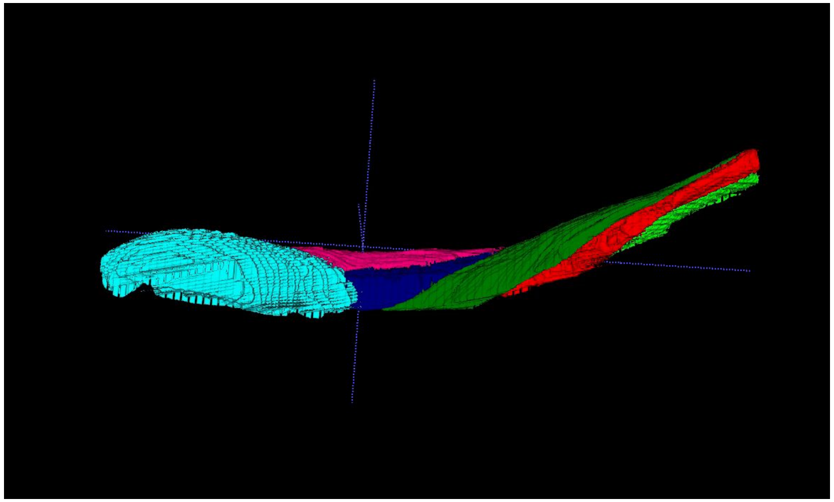


- Problem with collecting this data is that manual segmentation is very slow and labor/time intensive
- Would like to automate this process
    - General medical segmentation methods don't work very well
    - We attempt to solve this problem via convolutional neural networks (CNNs)

Our Data

- Began with 9 pairs of ultrasound volumes, labeled and unlabeled
- Three angle conditions each with three weight conditions
- All from the same subject





Why Convolutional Networks

- Convolutional networks have become the state-of-the-art approach to automatic image recognition and classification
- Recent years have shown significant progress in convolutional networks
  - Huge progress both in image classification and **segmentation**

Images from the PASCAL VOC dataset.

Primary approach: the U-net

- Began by implementing the U-net architecture
- Upsides of the U-net:
  - The U-net has been shown to yield good results in various medical imaging segmentation benchmarks (making it very popular for biomedical image segmentation)
  - Relatively simple architecture to implement and train

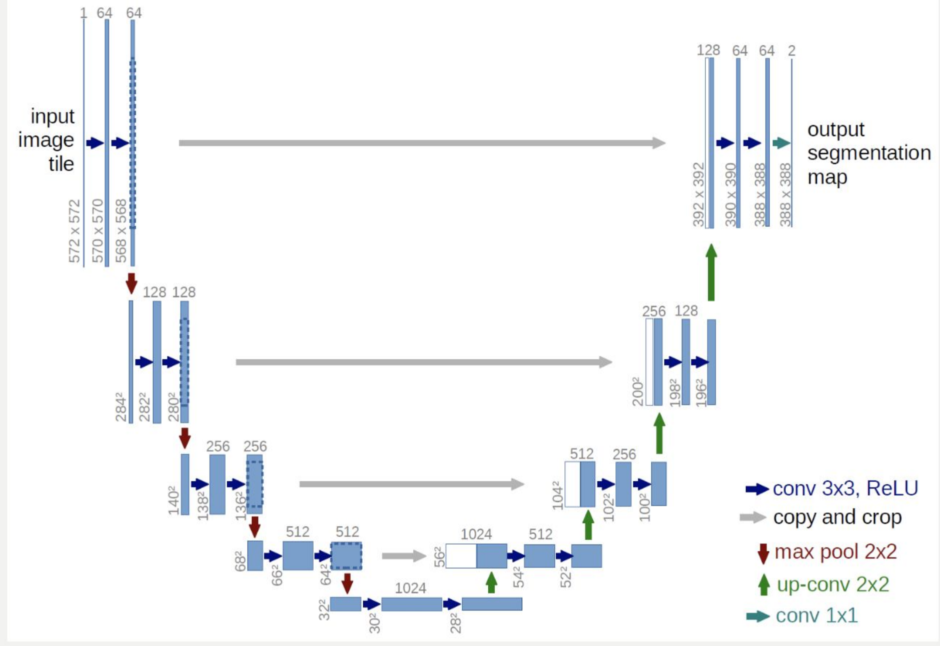


Image from Ronneberger et al 2015

Primary approach: the U-net

- Downsides of the U-net:
  - 2-dimensional architecture: although the U-net learns from and predicts 2-dimensional images, our problem deals with 3-dimensional volumes
    - (3D version does exist, however, will return to this)

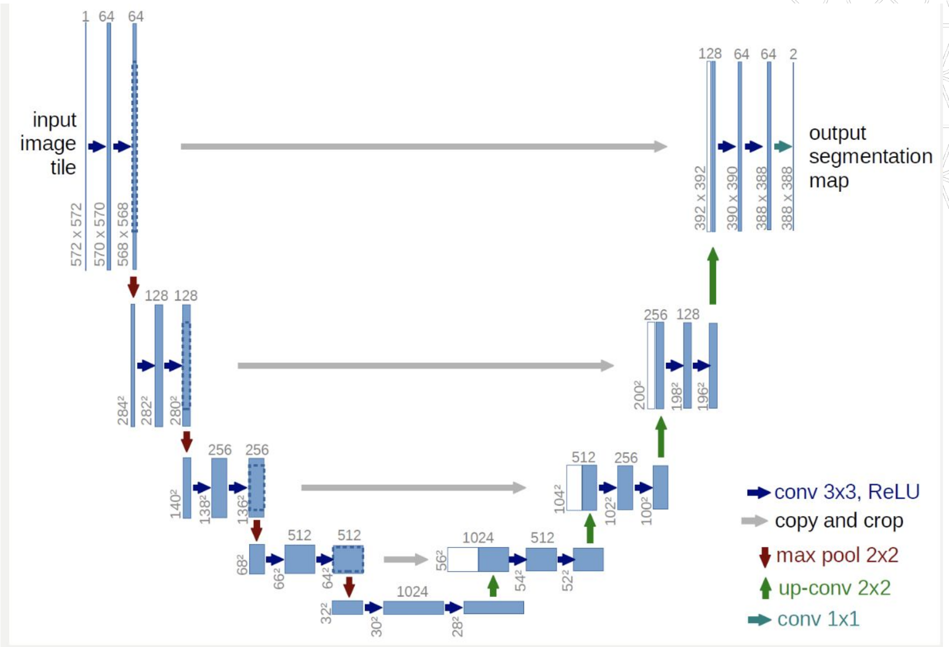
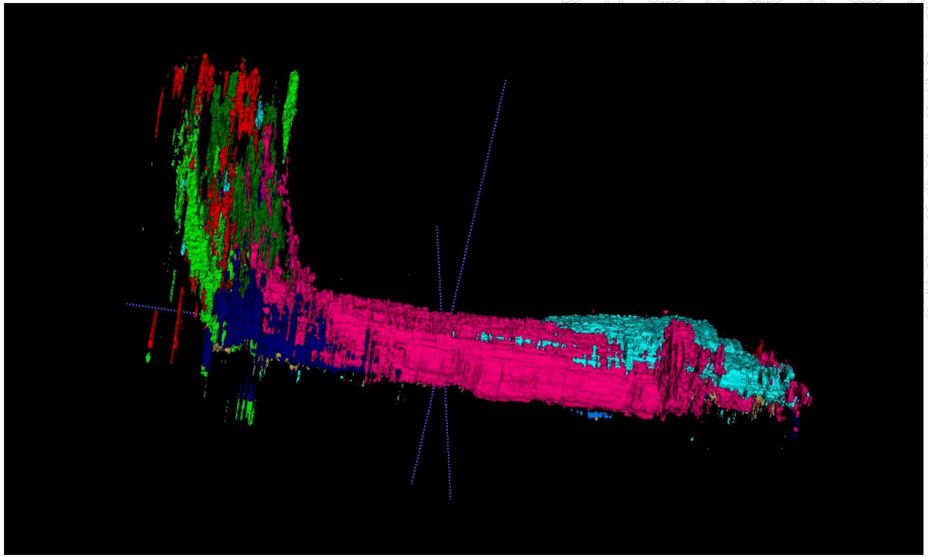
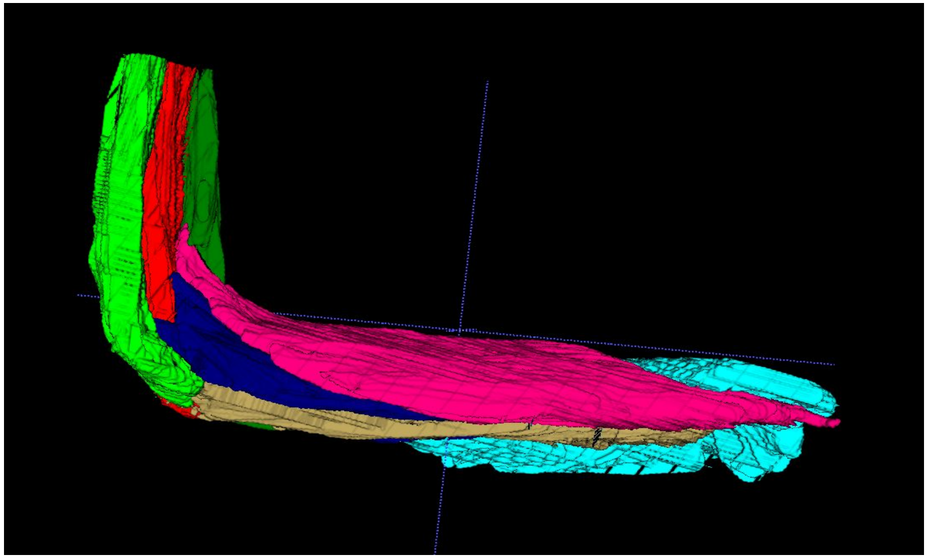


Image from Ronneberger et al 2015

Initial Baseline

Top: Predicted segmentation.
Bottom: Ground truth reference.





- Began by constructing some "easiest possible" cases for the U-net
  - Only expose to one angle condition
- Elbow and forearm has lowest accuracy
- In general low performance even in the "easy areas"

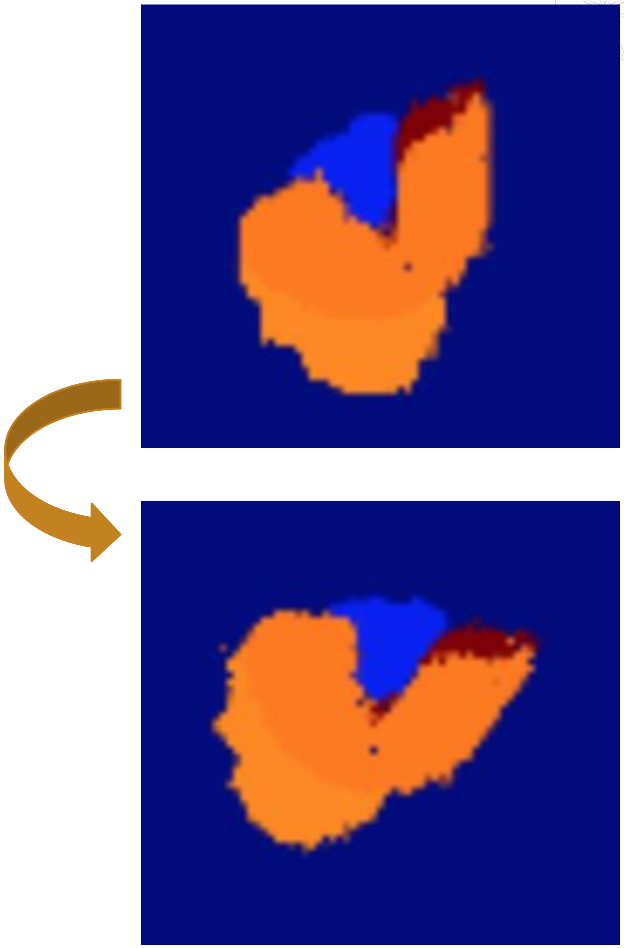
Improve through Data Augmentation

With limited available data, want to artificially create more to improve the network performance. First seek to do this through

- Rotations
- Elastic Deformation

Data Augmentation: Rotation

- Transformed data through arbitrary rotations
- Created augmented data that was rotated along all three axes
    - However for now because we use a 2D U-net, we only trained with data rotated along a single axis

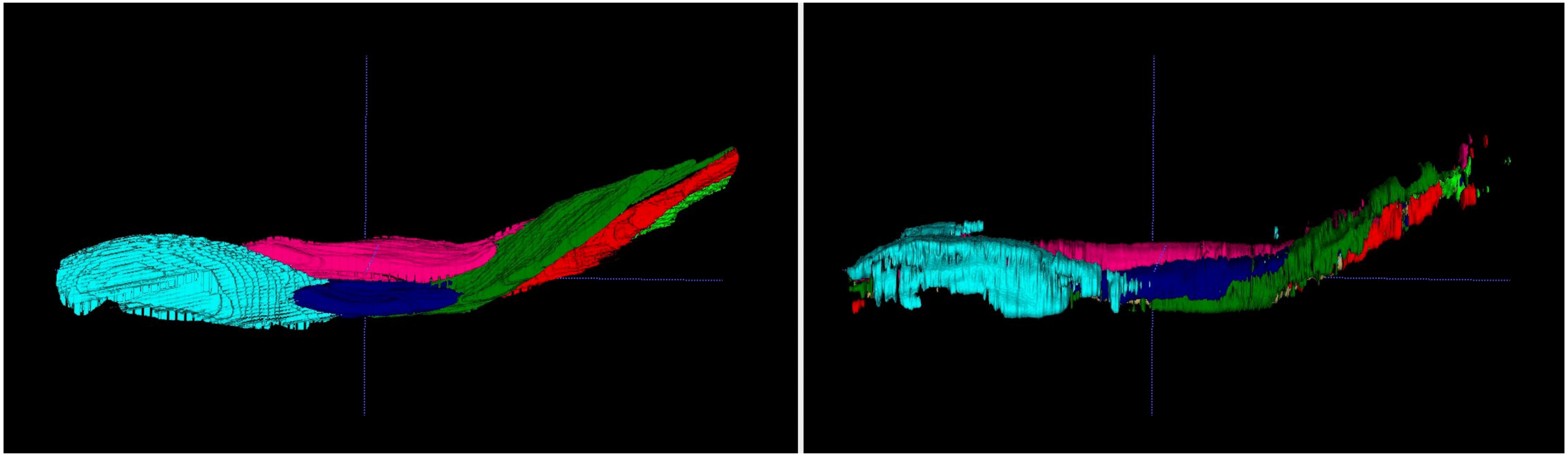


Preliminary Results with Rotation

Including a number of augmented scans during the training process yielded far better accuracy on unseen data.

Issues/limitations:

- Artifacts at both far ends
- Forearm is still a weak point
- Still “easiest case” - all data same angle



Left: Ground truth reference. Right: Prediction generated with augmented data.

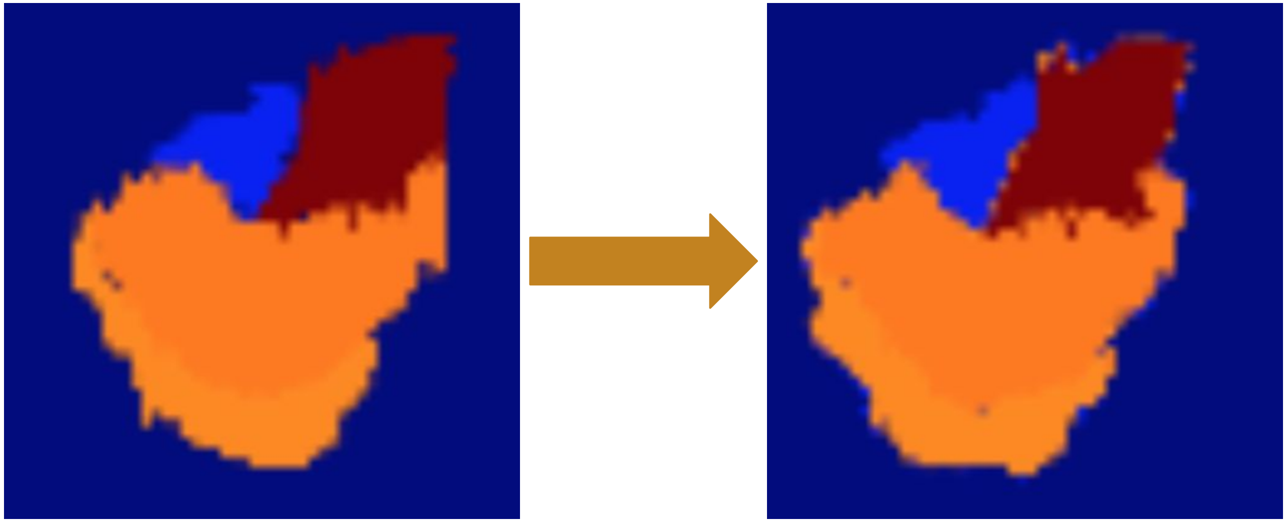
Data Augmentation: Elastic Deformation

- Elastic deformation is a random warping transformation
- Plausible approximation of the kind of natural variation present in muscles
- Done by generating a random displacement field then smoothing it out

Images from Simard (2003).

Data Augmentation: Elastic Deformation

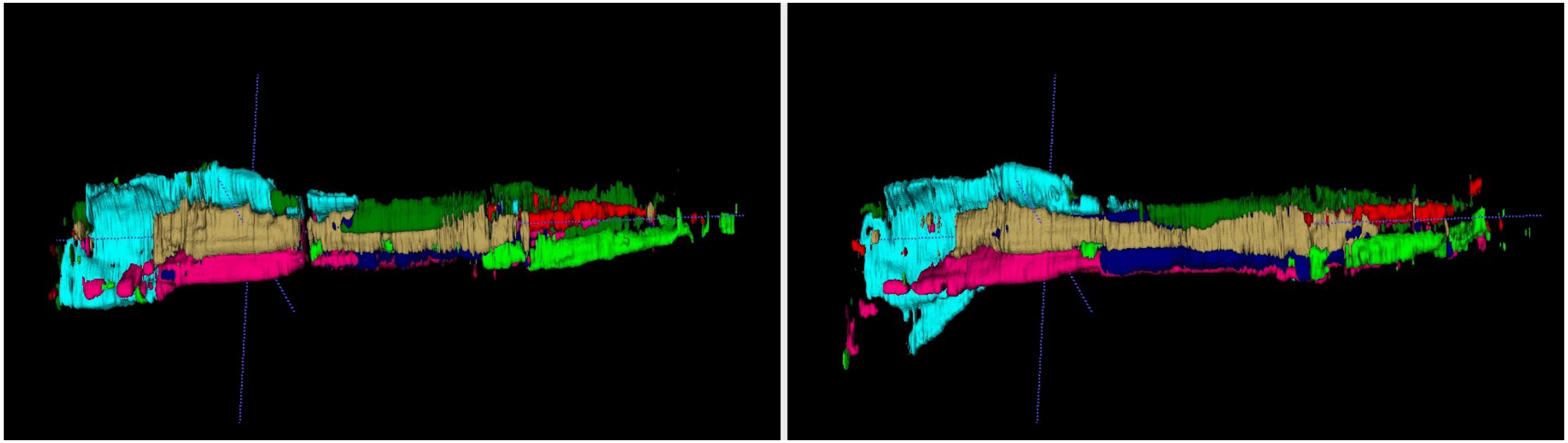
- Need to maintain labels, introduces some error
- Requires time consuming postprocessing



Inclusion of Elastic Deformation

Replacing some rotation augmented data with elastic deformation yielded mixed results.

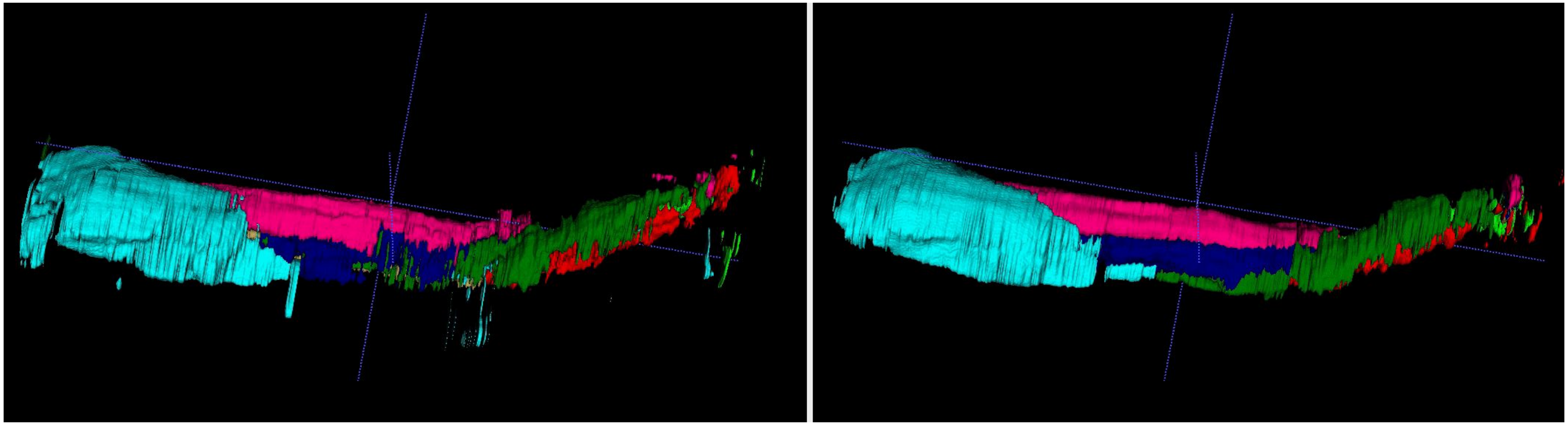
- Reduced far-end artifacts
- All-rotation does slightly better overall
- Need to determine optimal intensity of deformation, improve error-correction



Left: Prediction with mix of elastic data, rotation. Right: Prediction with rotation only. Equal number of total scans.

Multiple Angle Conditions

- Initially used data from one angle condition
- Turns out that using multiple angles likely more effective
- Including multiple angle conditions without any augmented data makes scans on par with previous best results of single angle, lots of augmentation



Left: Prediction with multiple angle conditions, no augmented data. Right: Multiple angles, with augmented data.

Future Directions

1. Improve quality and quantity of augmented data
2. Implement CNN architecture compatible with 3D image data
3. Post-processing images
4. For 2D U-net: train multiple models from different slice planes
5. Experiment with other architectures: FCN, FCN with LSTM architecture, Faster RCNN
6. Experimenting with GANS for more data augmentation

References

U-Net: Convolutional Networks for Biomedical Image Segmentation: https://arxiv.org/abs/1505.04597

3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation: https://arxiv.org/abs/1606.06650

Laura Hallock, Akira Kato, and Ruzena Bajcsy. Empirical quantification and modeling of muscle deformation: Toward ultrasound-driven assistive device control. In *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018.

FCN LSTM for Joint 4D: https://www.cs.utah.edu/~jeffp/papers/ISBI18.pdf

Questions?