Using multiple frame input U-net for automated segmentation of spinal ultrasound images

Victoria Wu, Tamas Ungi, Kyle Sunderland, Grace Pigeau, Abigael Schonewille,
Gabor Fichtinger

Laboratory for Percutaneous Surgery, School of Computing, Queen's University, Kingston, Canada

INTRODUCTION: Scoliosis, an excessive curvature of the spine, is diagnosed in approximately 2-3% of the population, mostly in children. One common method to reduce the need for vertebral fusion surgery is bracing of the spine. Effective bracing requires early referral, which is prevented due to the lack of a safe and affordable instrument for spinal curvature measurement. Currently, x-rays are used, but this exposes young children to harmful levels of radiation. Spinal ultrasound is a viable alternative to X-ray [1], but manual annotation of images is difficult. We propose using deep learning through a U-net neural network to generate automatic segmentations of the spine. We develop a variation of the U-net that takes consecutive images per individual input, as an enhancement over using single input images. This is based on the user looking back on previous ultrasound images to acquire general information about the anatomy.

METHODS: Data was collected from nine healthy volunteers, using tracked ultrasound. Images were manually segmented using the open source software 3D Slicer (www.slicer.org). In total, seven of the scans were used for training data, and the additional two scans were used for testing data. To accommodate for consecutive input images, the ultrasound images were exported along with previous images stacked to serve as input for a modified U-net. All the images were converted into numpy arrays for network input. Five separate series of numpy arrays were created based on the exported images, containing one to four preceding images, giving arrays that contained two to five channels in total. Resulting output segmentations were evaluated based on the percentage of true negative and true positive pixel predictions.

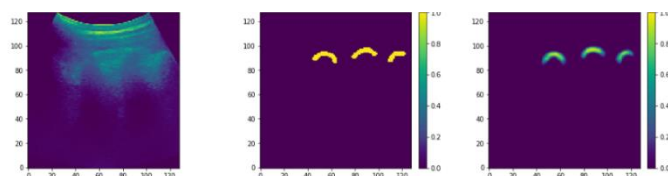

Figure 1: *Left.* Ultrasound image. *Middle.* Input segmentation. *Right.* Generated prediction

Figure 2: Comparison of average true positive (blue) and true negative (orange) rates between single image (left) and three-image input (right)

RESULTS: After comparing the single to five-image input arrays, the three-image input had the best performance in terms of true positive value. The single input and three-input images were then further tested. The single image input neural network had a true negative rate of 99.79%, and a true positive rate of 63.56%. The three-image input neural network had a true negative rate of 99.75%, and a true positive rate of 66.64%.

CONCLUSION: The three-image input network outperformed the single input network in terms of the true positive rate by 3.08%. This is a marginal improvement, suggesting that the capabilities of this methodology of examining consecutive images is limited. However, despite the modest improvement, segmentations generated from the single input and three-input networks are both quite accurate, suggesting that U-net is a viable architecture for spine ultrasound segmentation.

ACKNOWLDGEMENTS: V. Wu is supported by CIHR Undergraduate Summer Studentship Award. G. Fichtinger is supported as a Canada Research Chair. This work was funded, in part, by NIH/NIBIB and NIH/NIGMS (via grant 1R01EB021396-01A1 - Slicer+PLUS: Point-of-Care Ultrasound) and by CANARIE's Research Software Program.

REFERENCES:

[1] Ungi T, King F, Kempston M, Keri Z, Lasso A, Mousavi P, Rudan J, Borschneck DP, Fichtinger G. "Spinal curvature measurement by tracked ultrasound snapshots," Ultrasound Med Biol 40(2), 447-54 (2014).