Automated Contouring Using Neural Networks

Hanlin Wan, PhD

MIM Software Inc., Cleveland, OH, USA

Introduction

Automated segmentation of structures on medical images has always been challenging. In clinical practice, much of the segmentation is performed manually, which is highly time-consuming. Many semi- and fully-automated algorithms have been developed to aid physicians in contouring, such as thresholding[1], edge detection based methods[2], deformable models[3], and atlas-based registration[4].

However, many structures have various textural and intensity patterns, making it almost impossible for any single computer vision-based algorithm to work effectively in a universal fashion. In recent years, research on the use of neural networks has grown rapidly. The performance of these networks on many computer vision tasks has often surpassed that of a human, such as in image classification tasks like ImageNet[5].

Contour ProtégéAI+™ provides a neural network framework for automated contouring of normal structures on CT and MR images.

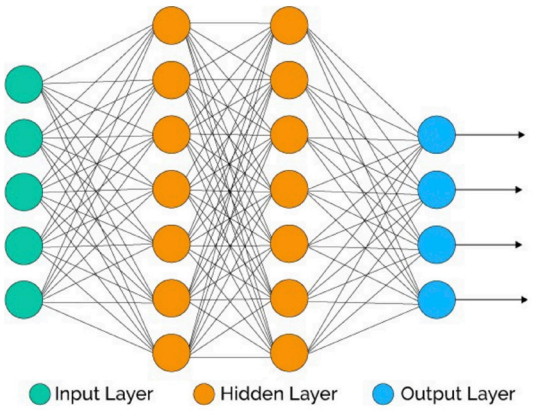

FIGURE 1 *Schematic of a basic neural network*

Neural networks attempt to mimic how the human brain works. The brain consists of billions of neurons. Each neuron receives multiple signals from other neurons and sends out a signal based on those inputs. These neurons are often organized into layers, allowing the brain to use simple building blocks to process complicated input signals. Figure 1 shows an example of a basic neural network. Each line has a weight and an associated bias.

Therefore, the nodes calculate a weighted sum of its inputs and then applies some activation function. The output is then sent to the next node until the final output layer is reached. When training a neural network, pairs of inputs and desired outputs are shown. The network learns to adjust the network parameters to minimize the difference between its outputs and the desired outputs. Essentially, neural networks learn to recognize patterns much in the same way the human brain does. This pattern learning is what makes neural networks so powerful, allowing them to exceed the performance of traditional methods.

Contour ProtégéAI+'s neural network model is based on the U-Net architecture, which has been used for segmentation in numerous different applications. The model consists of many layers of weights and biases as mentioned above to transform the input image to a segmentation mask for each structure at the final output layer. This output is then post-processed to keep the single, largest connected component. Appropriate image visualization software must be used to review and, if necessary, edit results automatically generated by Contour ProtégéAI+.

Training and Validation

A large, multi-institution dataset was assembled for training, along with a separate, large multi-institution dataset for validation. None of the validation data came from any of the institutions from the training pool. The mean and standard deviation of the model performance on this validation set was then calculated.

Results

Five different sets of metrics were used to assess the performance of the neural network segmentations:

1. Dice coefficient - a measure of the spatial overlap between the ground truth contours and the neural network segmentations.
2. Mean distance to agreement (MDA) - the mean symmetric surface distance between the ground truth contours and the neural network segmentations.
3. Qualitative user feedback score - assessment of contour quality by experienced users on a scale of 1-3 (none, moderate, signficant time savings compared to contouring from scratch).
4. Localization success - percentage of images where the structure was correctly localized by the neural network segmentation.
5. Added path length (APL) - a measure of the cumulative amount of editing in mm needed to match the ground truth contour.

Tables 1-3 tabulate the Dice, MDA, user feedback score, and localization success for each of the models, and Table 4 shows the APL per model. Figures 2-7 compare the mean Dice and mean MDA of the neural network and MIM Software's atlas using majority vote 5. In all cases, Contour ProtégéAI+'s neural network segmentations were proportionate or superior to atlas-based segmentation.

Structure	Mean ± Std Dice	MDA (Mean ± Std mm)	Beta User Feedback Score	Localization Success on Relavant FOV CTs	Localization Success on Whole Body CTs
Bone_Mandible	0.86 ± 0.07	0.64 ± 0.31	2.86	100	100
BrachialPlex_L	0.22 ± 0.10	2.82 ± 2.61	2.60	100	91
BrachialPlex_R	0.19 ± 0.09	2.73 ± 2.38	2.60	100	95
Brain	0.98 ± 0.01	0.52 ± 0.36	2.71	100	100
Brainstem	0.82 ± 0.08	1.20 ± 0.65	2.71	100	100
Cavity_Oral	0.76 ± 0.13	3.20 ± 1.93	2.71	100	100
Cochlea_L	0.30 ± 0.18	1.17 ± 0.87	2.29	95	95
Cochlea_R	0.32 ± 0.20	1.05 ± 0.68	2.29	96	95
Eye_L	0.87 ± 0.06	0.65 ± 0.51	2.57	100	100
Eye_R	0.86 ± 0.07	0.66 ± 0.48	2.57	100	100
Glnd_Lacrimal_L	0.40 ± 0.16	0.68 ± 0.35	2.71	98	100
Glnd_Lacrimal_R	0.45 ± 0.14	0.71 ± 0.52	2.71	100	100
Glnd_Submand_L	0.77 ± 0.11	0.76 ± 0.32	3.00	100	100
Glnd_Submand_R	0.75 ± 0.11	0.81 ± 0.34	3.00	100	100
Glnd_Thyroid	0.71 ± 0.19	1.38 ± 3.18	2.71	98	95
Lens_L	0.61 ± 0.17	0.56 ± 0.47	2.29	97	100
Lens_R	0.63 ± 0.16	0.55 ± 0.33	2.29	97	95
Lips	0.37 ± 0.15	5.26 ± 3.41	2.83	100	95
OpticChiasm	0.13 ± 0.13	2.46 ± 1.99	2.00	80	91
OpticNrv_L	0.53 ± 0.14	0.77 ± 0.85	2.57	99	100
OpticNrv_R	0.52 ± 0.13	0.80 ± 0.83	2.57	99	100
Parotid_L	0.79 ± 0.09	1.30 ± 0.62	3.00	99	100
Parotid_R	0.80 ± 0.06	1.26 ± 0.43	3.00	100	100
Pituitary	0.54 ± 0.15	0.87 ± 0.54	3.00	100	86
SpinalCord	0.65 ± 0.16	0.73 ± 0.41	2.86	99	100
LN_Neck_IA	0.60 ± 0.14	0.47 ± 0.45	2.75	100	100
LN_Neck_IB_L	0.79 ± 0.04	0.73 ± 0.24	2.75	100	100
LN_Neck_IB_R	0.79 ± 0.05	0.73 ± 0.20	2.75	100	100
LN_Neck_IIA_L	0.75 ± 0.05	1.54 ± 0.44	2.25	100	100
LN_Neck_IIA_R	0.76 ± 0.04	1.41 ± 0.49	2.25	100	95
LN_Neck_IIB_L	0.80 ± 0.06	0.63 ± 0.31	3.00	100	100
LN_Neck_IIB_R	0.80 ± 0.06	0.58 ± 0.15	3.00	100	95
LN_Neck_III_L	0.75 ± 0.07	0.92 ± 0.44	3.00	100	100
LN_Neck_III_R	0.75 ± 0.07	0.99 ± 0.37	3.00	100	100
LN_Neck_IV_L	0.69 ± 0.08	1.13 ± 0.37	2.75	100	100
LN_Neck_IV_R	0.71 ± 0.09	1.06 ± 0.33	2.75	100	95
LN_Neck_V_L	0.58 ± 0.09	1.13 ± 0.59	3.00	100	100
LN_Neck_V_R	0.58 ± 0.11	1.43 ± 0.98	3.00	100	91
LN_Neck_VIA	0.36 ± 0.11	0.80 ± 0.35	2.75	100	100

Continued on next page.

Structure	Mean ± Std Dice	MDA (Mean ± Std mm)	Beta User Feedback Score	Localization Success on Relavant FOV CTs	Localization Success on Whole Body CTs
LN_Retropharynx_L	0.28 ± 0.10	1.54 ± 1.47	2.75	100	100
LN_Retropharynx_R	0.28 ± 0.09	1.49 ± 1.36	2.75	95	100
LN_Retrostyloid_L	0.64 ± 0.09	1.23 ± 0.48	2.75	100	100
LN_Retrostyloid_R	0.66 ± 0.11	1.13 ± 0.66	2.75	100	100

TABLE 1 *Performance statistics for the* CT Head & Neck *model*

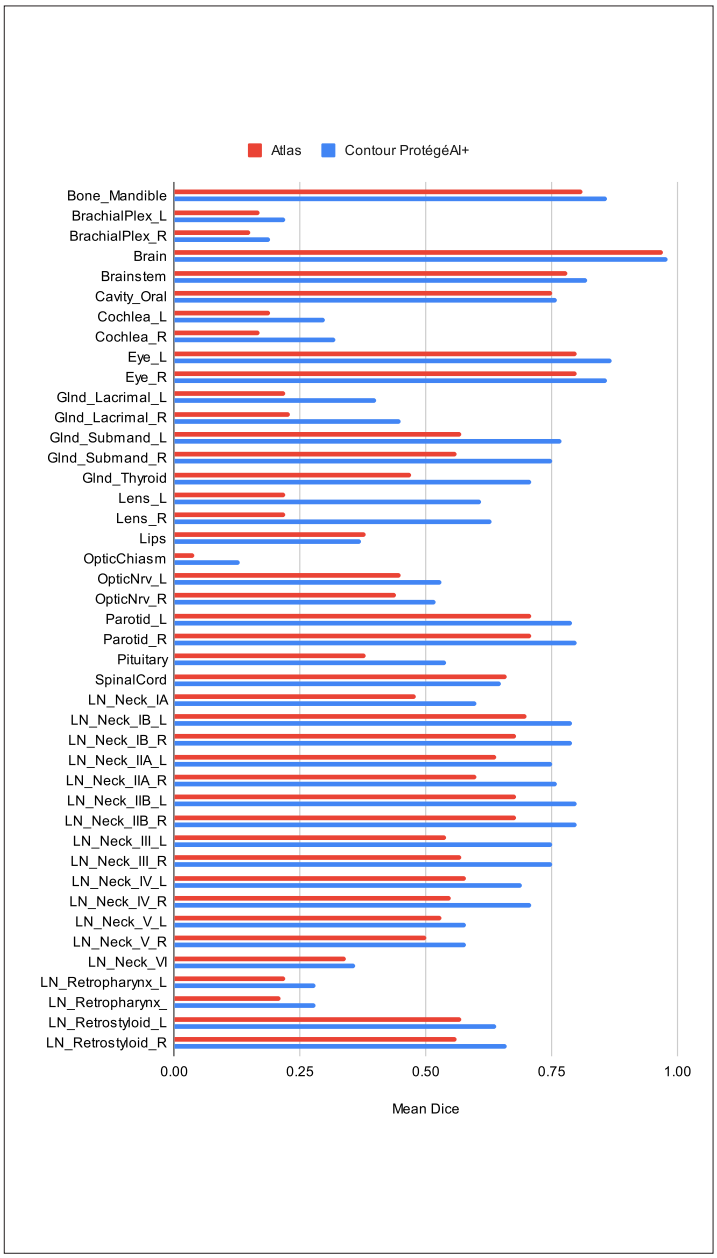

FIGURE 2 *Comparison of atlas and Contour ProtégéAI+ Dice on the* CT Head & Neck *model*

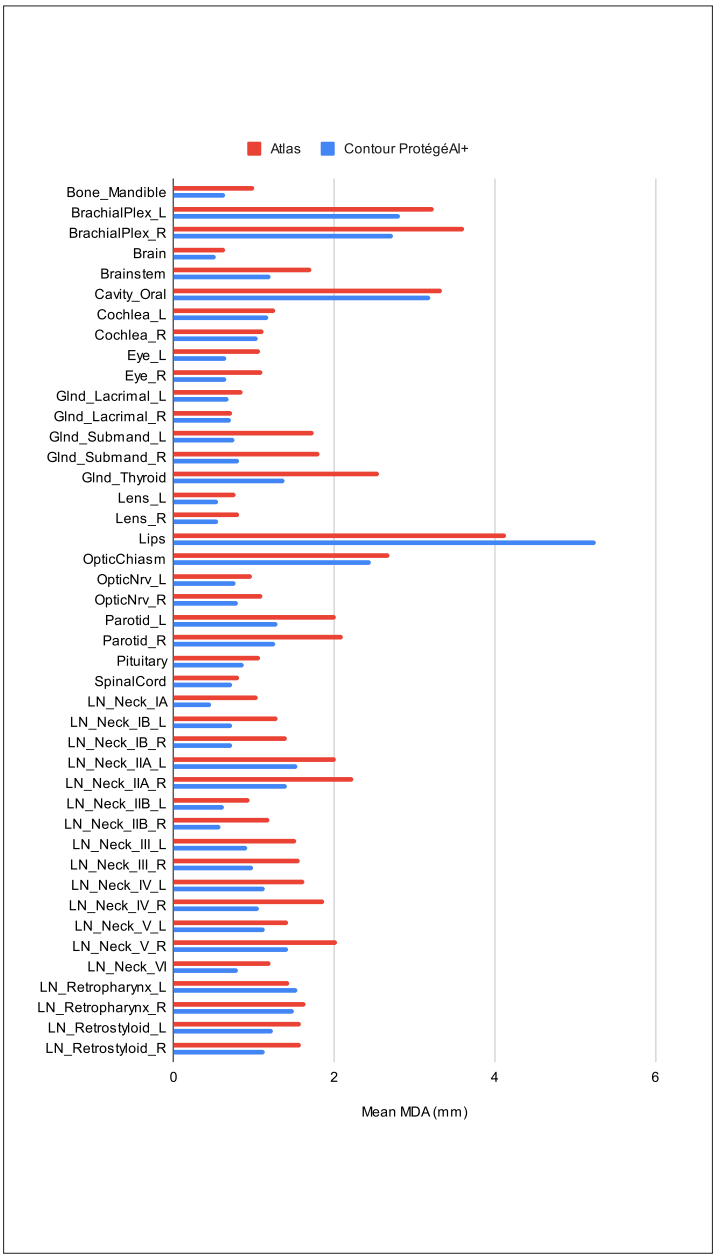

FIGURE 3 *Comparison of atlas and Contour ProtégéAI+ Mean MDA on the* CT Head & Neck *model*

Structure	Mean ± Std Dice	MDA (Mean ± Std mm)	Beta User Feedback Score	Localization Success on Relavant FOV CTs	Localization Success on Whole Body CTs
BrachialPlex_L	0.38 ± 0.14	2.94 ± 1.84	3.00	100	100
BrachialPlex_R	0.37 ± 0.16	3.28 ± 2.60	3.00	100	100
Breast_L	0.76 ± 0.15	3.64 ± 1.85	2.80	100	100
Breast_R	0.78 ± 0.15	3.45 ± 1.51	2.60	96	91
Bronchus	0.66 ± 0.13	1.90 ± 1.22	2.60	100	100
Carina	0.49 ± 0.12	1.03 ± 0.47	2.40	98	100
Cricoid	0.04 ± 0.05	5.71 ± 1.24	2.40	67	100
Esophagus	0.70 ± 0.15	0.97 ± 0.54	2.80	100	100
Glnd_Thyroid	0.65 ± 0.17	1.75 ± 1.78	2.75	100	100
GreatVes	0.74 ± 0.10	2.67 ± 1.73	3.00	100	100
Heart	0.90 ± 0.09	2.55 ± 1.40	3.00	100	100
Humerus_Head_L	0.95 ± 0.02	0.40 ± 0.31	2.80	100	95
Humerus_Head_R	0.96 ± 0.02	0.31 ± 0.26	2.80	100	100
Kidney_L	0.92 ± 0.05	0.94 ± 0.80	3.00	97	91
Kidney_R	0.91 ± 0.07	0.98 ± 0.66	3.00	97	95
Larynx	0.52 ± 0.14	3.06 ± 1.08	2.40	99	95
Liver	0.93 ± 0.06	2.00 ± 1.79	3.00	100	100
Lung_L	0.96 ± 0.02	0.77 ± 0.37	3.00	100	100
Lung_R	0.96 ± 0.02	0.89 ± 0.43	3.00	100	100
Musc_Constrict	0.45 ± 0.16	2.07 ± 1.52	2.00	100	91
Pancreas	0.41 ± 0.24	8.68 ± 11.12	2.25	54	91
SpinalCord	0.66 ± 0.16	1.41 ± 0.63	2.75	100	100
Stomach	0.79 ± 0.19	2.33 ± 2.12	3.00	100	100
Trachea	0.74 ± 0.16	0.88 ± 0.54	3.00	99	100
LN_Ax_L1_L	0.65 ± 0.10	2.58 ± 1.17	3.00	100	100
LN_Ax_L1_R	0.60 ± 0.14	3.19 ± 1.61	3.00	100	100
LN_Ax_L2_L	0.55 ± 0.15	2.10 ± 0.96	2.67	100	100
LN_Ax_L2_R	0.51 ± 0.16	2.70 ± 1.77	2.67	100	100
LN_Ax_L3_L	0.52 ± 0.17	2.15 ± 1.52	3.00	100	100
LN_Ax_L3_R	0.51 ± 0.17	2.53 ± 1.41	3.00	100	100
LN_IMN_L	0.44 ± 0.17	1.63 ± 0.96	2.67	100	100
LN_IMN_R	0.40 ± 0.16	1.55 ± 1.60	2.67	100	100
LN_Sclav_L	0.58 ± 0.13	1.87 ± 0.89	2.67	100	100
LN_Sclav_R	0.54 ± 0.11	2.60 ± 0.85	2.67	100	100

TABLE 2 *Performance statistics for the* CT Thorax *model*

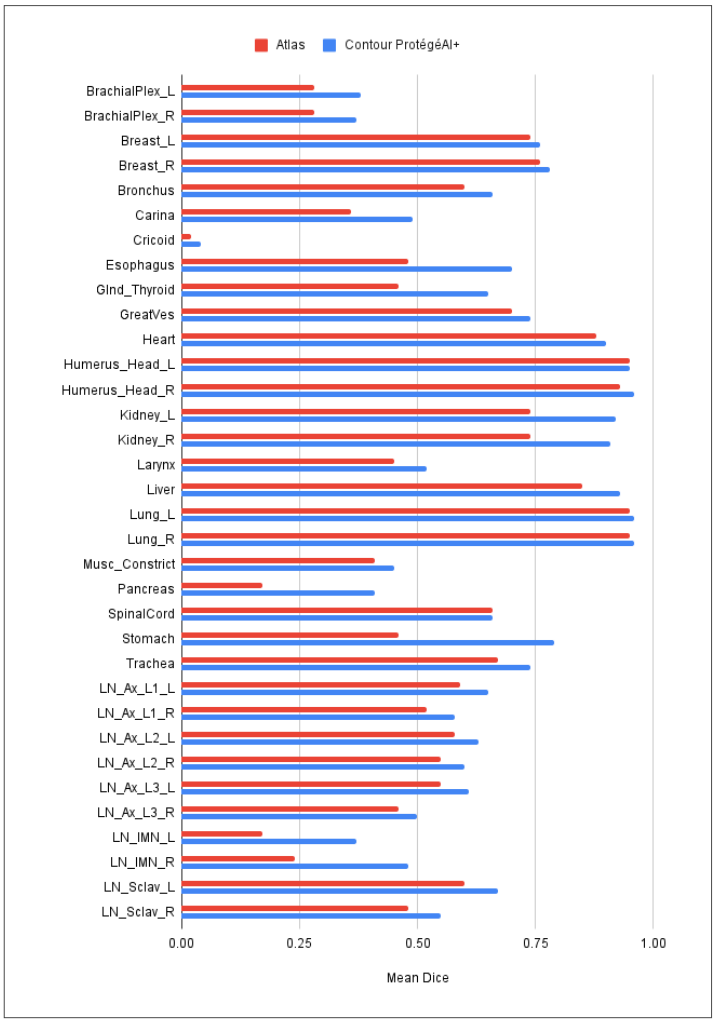

FIGURE 4 *Comparison of atlas and Contour ProtégéAI+ Dice on the* CT Thorax *model*

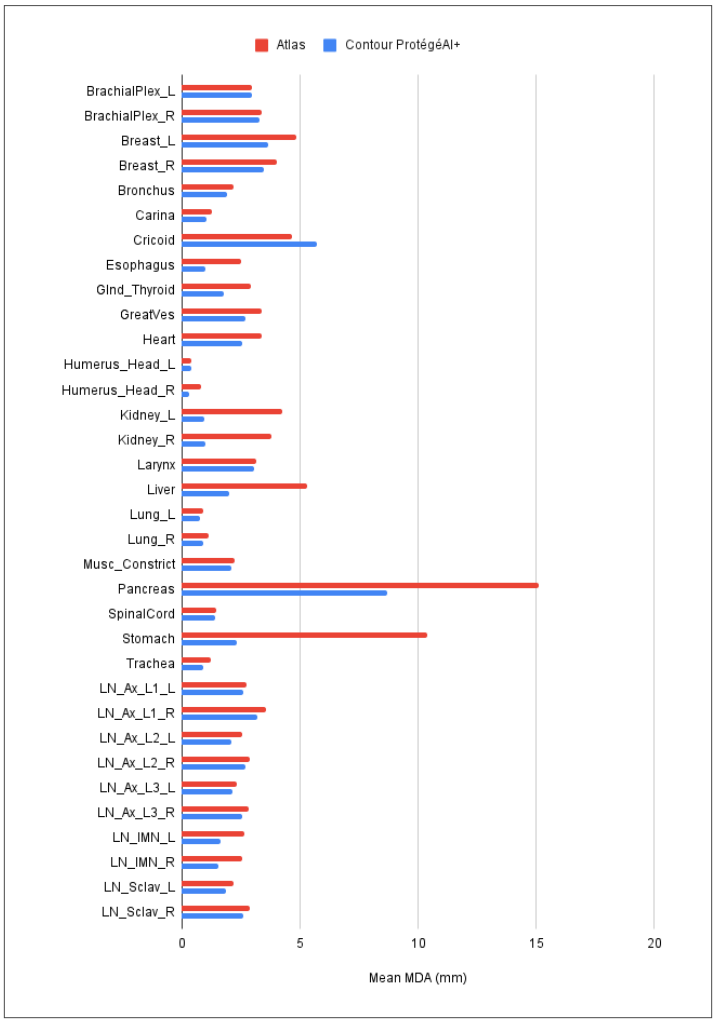

FIGURE 5 *Comparison of atlas and Contour ProtégéAI+ Mean MDA on the* CT Thorax *model*

Structure	Mean ± Std Dice	MDA (Mean ± Std mm)	Beta User Feedback Score	Localization Success on Whole Body CTs
Bladder	0.92 ± 0.16	0.91 ± 1.02	3.00	95
Bone	0.87 ± 0.05	0.48 ± 0.13	3.00	100
Bowel	0.74 ± 0.10	7.69 ± 5.07	3.00	100
BowelBag	0.75 ± 0.15	17.18 ± 5.24	3.00	100
Brain	0.97 ± 0.03	0.91 ± 1.17	3.00	100
Cavity_Oral	0.78 ± 0.12	3.19 ± 1.66	3.00	100
Glnd_Lacrimal_L	0.37 ± 0.16	0.72 ± 0.43	3.00	100
Glnd_Lacrimal_R	0.46 ± 0.14	0.81 ± 0.79	3.00	100
Glnd_Submand_L	0.79 ± 0.06	0.81 ± 0.35	3.00	100
Glnd_Submand_R	0.77 ± 0.06	0.81 ± 0.38	3.00	100
Heart	0.91 ± 0.04	2.65 ± 0.93	3.00	100
Kidney_L	0.93 ± 0.04	0.90 ± 0.51	3.00	95
Kidney_R	0.92 ± 0.07	0.85 ± 0.44	3.00	100
Liver	0.93 ± 0.07	2.26 ± 3.82	3.00	100
LN_Iliac	0.71 ± 0.03	1.59 ± 0.33	3.00	64
Parotid_L	0.81 ± 0.05	1.30 ± 0.47	3.00	100
Parotid_R	0.82 ± 0.05	1.34 ± 0.56	3.00	100
Prostate	0.85 ± 0.05	1.55 ± 0.61	3.00	100
Spleen	0.95 ± 0.01	0.61 ± 0.38	3.00	100

TABLE 3 *Performance statistics for the* Whole Body - Physiological Uptake Organs *model*

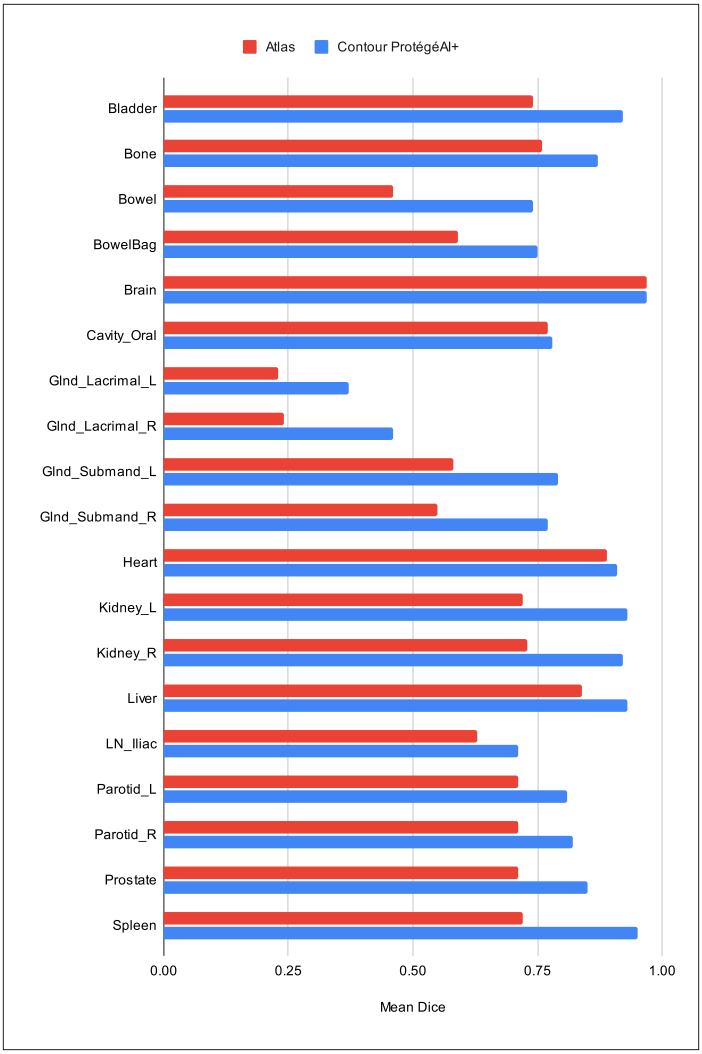

FIGURE 6 *Comparison of atlas and Contour ProtégéAI+ Dice on the* Whole Body - Physiological Uptake Organs *model*

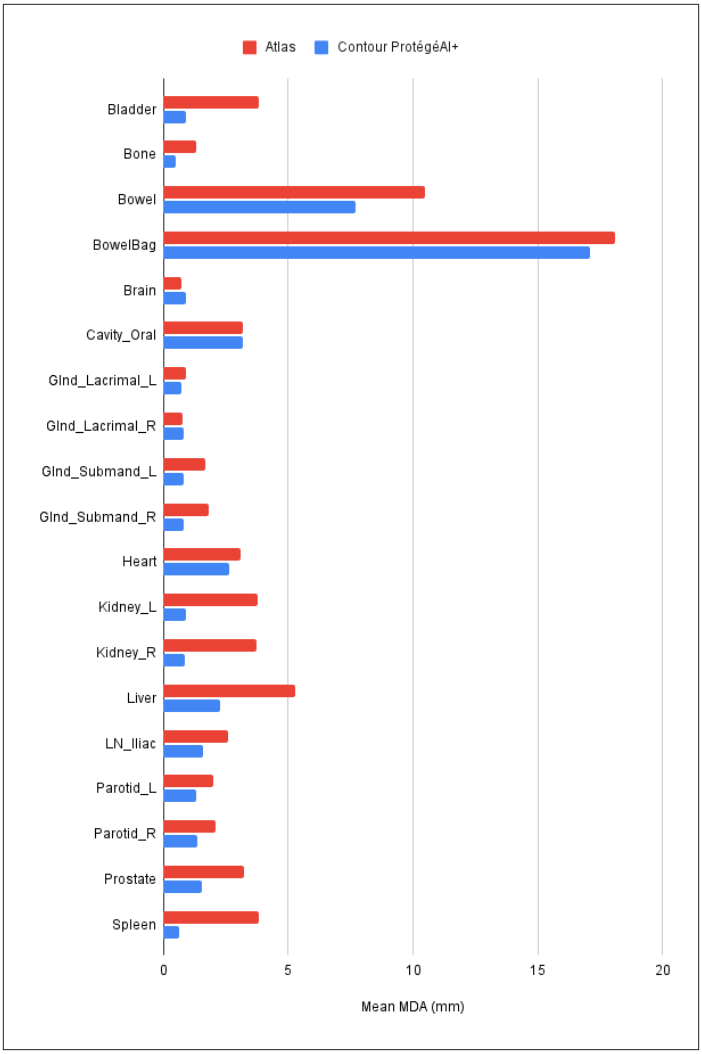

FIGURE 7 *Comparison of atlas and Contour ProtégéAI+ Mean MDA on the* Whole Body - Physiological Uptake Organs *model*

Model	Atlas	Contour ProtégéAI+
Head & Neck	38.69 ± 33.36	28.61 ± 29.59
Thorax	89.24 ± 82.73	65.44 ± 68.85
Whole Body - Physiological Uptake Organs	138.06 ± 142.42	98.20 ± 127.11

TABLE 4 *Mean added path length, atlas vs. Contour ProtégéAI+*

References

1. Weszka, J. S. (1978). A survey of threshold selection techniques. Computer Graphics and Image Processing, 7(2), 259-265.

2. Senthilkumaran, N., & Rajesh, R. (2009). Edge detection techniques for image segmentation-a survey of soft computing approaches. International journal of recent trends in engineering, 1(2), 250.

3. McInerney, T., & Terzopoulos, D. (1996). Deformable models in medical image analysis: a survey. Medical image analysis, 1(2), 91-108.

4. Klein, S., Staring, M., Murphy, K., Viergever, M. A., & Pluim, J. P. (2010). Elastix: a toolbox for intensity-based medical image registration. IEEE transactions on medical imaging, 29(1), 196-205.

5. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., & Berg, A.C. (2015). Imagenet large scale visual recognition challenge. International journal of computer vision, 115(3), 211-252.