### APPENDIX: Quantitative Evaluation

|                  | LDA $\sigma \in [1-6]$ (n=200) | MDA $\sigma \in [7-12]$ (n=200) | HDA $\sigma \in [13-18]$ (n=200) | Control $\sigma \in [19-24]$ (n=40) |
|------------------|---------------------------------|---------------------------------|---------------------------------|-----------------------------------|
|                  | Original (n=100) | Modified (n=100) | $p$ value | Original (n=100) | Modified (n=100) | $p$ value | Original (n=100) | Modified (n=100) | $p$ value |
| Grader 1         | 85 | 77 | 0.21 | 85 | 76 | 0.15 | 91 | 75 | $4^{*}$-4 | 20 | 13 | $8^{*}$-3 |
| Grader 2         | 75 | 71 | 0.63 | 73 | 65 | 0.28 | 81 | 61 | $3^{*}$-3 | 17 | 4 | $4^{*}$-3 |
| Grader 3         | 76 | 76 | 1 | 80 | 63 | 0.01 | 83 | 50 | $1^{*}$-4 | 15 | 4 | $1^{*}$-3 |

**Supplementary Table 1.** Quantitative evaluation of the realistic value of the images by the three reviewers for the four categories of levels of the incorporated elastic deformation: Low-, Medium-, High- data augmentation and control.

|                  | Deformation intensity $\sigma \in [7-9]$ (n=200) | Deformation intensity $\sigma \in [10-11]$ (n=200) |
|------------------|---------------------------------|---------------------------------|
|                  | Original (n=100) | Modified (n=100) | $p$ value | Original (n=100) | Modified (n=100) | $p$ value |
| Grader 1         | 89 | 93 | 0.43 | 73 | 58 | 0.037 |
| Grader 2         | 93 | 89 | 0.47 | 48 | 27 | 0.003 |
| Grader 3         | 99 | 93 | 0.55 | 80 | 50 | $<1^{*}$-3 |

**Supplementary Table 2.** Results of the quantitative evaluation of realistic value of the images by the three reviewers for the sub range $\sigma = 7-9$ and $\sigma = 10-11$. 
Supplementary Figure 1. Evaluation of the influence of elastic deformation magnitude on a segmentation model. All results report the dice score of the validation set for a u-net architecture averaged over 10 independent runs.