

| Model | Metrics (training) | | | Metrics (validation) | | | Metrics (testing) | | |
|-------------------------------|--------------------|-------------|--------|----------------------|-------------|--------|-------------------|-------------|--------|
| | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ |
| Mod. CycleGAN (our model) | 0.63 | 0.12 | 17 | 0.60 | 0.16 | 15 | 0.59 | 0.16 | 15 |
| Pix2Pix | 0.24 | 0.15 | 16 | 0.21 | 0.17 | 14 | 0.22 | 0.22 | 15 |
| Std. CycleGAN + inv. CT input | 0.27 | 0.25 | 13 | 0.31 | 0.25 | 13 | 0.33 | 0.25 | 12 |
| Std. CycleGAN | 0.15 | 0.44 | 5 | 0.18 | 0.43 | 5 | 0.20 | 0.43 | 5 |

Table 1: Evaluation of models with median values of the metrics SSIM, LPIPS and PSNR for training (from 40 WSI samples), validation (10 WSI samples) and testing (3 WSI samples). The top values are highlighted with bold text.