1

Table 1. Fréchet Inception Distance (FID) of the distribution of ground-truth untruncated vs. outpainted slices by various models, stratified by vertebral levels. CT-Palette achieves the lowest overall FID.

Method	Overall (n=997)	T5 (n=339)	T8 (n=348)	T10 (n=310)
RFR-Net	17.505	25.277	21.571	19.471
S-EFOV	4.809	8.106	6.416	5.405
S-EFOV (FT)	3.968	6.714	5.427	**4.921**
CT-Palette (SI)	**3.521**	**5.579**	5.218	5.111
CT-Palette (MI)	3.567	5.625	**5.189**	5.191

Table 2. The mean and standard deviation of the Dice Similarity Coefficient (DSC) between the muscle and SAT segmentation masks of ground-truth untruncated and outpainted slices by various models. The top row indicates the DSC between untruncated and truncated slices for reference. CT-Palette achieves the highest DSC overall, for both muscle and SAT.

Method	Overall (n=997)		T5 (n=339)		T8 (n=348)		T10 (n=310)	
	Muscle	SAT	Muscle	SAT	Muscle	SAT	Muscle	SAT
Truncated	0.940 ± 0.083	0.865 ± 0.132	0.932 ± 0.078	0.862 ± 0.114	0.933 ± 0.096	0.861 ± 0.141	0.956 ± 0.070	0.874 ± 0.140
RFR-Net	0.957 ± 0.057	0.901 ± 0.078	0.963 ± 0.037	0.906 ± 0.068	0.946 ± 0.072	0.894 ± 0.086	0.962 ± 0.055	0.902 ± 0.078
S-EFOV	0.976 ± 0.040	0.948 ± 0.060	0.977 ± 0.029	0.944 ± 0.059	0.969 ± 0.052	0.944 ± 0.066	$\mathbf{0.982 \pm 0.033}$	$\mathbf{0.958} \pm 0.052$
S-EFOV (FT)	0.974 ± 0.043	0.947 ± 0.057	0.979 ± 0.024	0.947 ± 0.051	0.965 ± 0.054	0.944 ± 0.059	0.978 ± 0.044	0.950 ± 0.062
CT-Palette (SI)	0.977 ± 0.038	0.950 ± 0.048	0.982 ± 0.022	0.953 ± 0.045	0.971 ± 0.047	0.947 ± 0.052	0.979 ± 0.040	0.952 ± 0.047
CT-Palette (MI)	$\mathbf{0.979 \pm 0.036}$	$\mathbf{0.954 \pm 0.045}$	$\mathbf{0.984 \pm 0.019}$	$\mathbf{0.957 \pm 0.038}$	$\mathbf{0.972 \pm 0.045}$	$\mathbf{0.951 \pm 0.050}$	0.980 ± 0.038	0.954 ± 0.045

Table 3. PSNR and SSIM between ground-truth untruncated slices and outpainted slices by various models, stratified by vertebral levels. The top row indicates the PSNR and SSIM between untruncated and truncated slices for reference.

Method	PSNR				SSIM			
	Overall (n=997)	T5 (n=339)	T8 (n=348)	T10 (n=310)	Overall (n=997)	T5 (n=339)	T8 (n=348)	T10 (n=310)
Truncated	34.052 ± 9.029	32.599 ± 7.631	33.729 ± 8.801	36.008 ± 10.265	0.974 ± 0.025	0.972 ± 0.025	0.973 ± 0.026	0.977 ± 0.025
RFR-Net	33.569 ± 7.339	32.119 ± 7.177	33.678 ± 7.433	35.032 ± 7.101	0.969 ± 0.031	0.965 ± 0.031	0.968 ± 0.032	0.974 ± 0.029
S-EFOV	$\mathbf{39.390 \pm 12.657}$	$\mathbf{36.793 \pm 10.389}$	$\mathbf{39.470 \pm 12.820}$	$\mathbf{42.147 \pm 14.066}$	$\mathbf{0.976 \pm 0.026}$	$\mathbf{0.973 \pm 0.026}$	$\mathbf{0.976 \pm 0.027}$	$\mathbf{0.981 \pm 0.026}$
S-EFOV (FT)	38.594 ± 11.957	35.982 ± 10.695	38.700 ± 11.848	41.332 ± 12.734	0.975 ± 0.027	0.972 ± 0.027	0.975 ± 0.027	0.979 ± 0.026
CT-Palette (SI)	36.655 ± 8.363	35.518 ± 8.170	36.497 ± 8.259	38.076 ± 8.479	0.974 ± 0.028	0.972 ± 0.027	0.974 ± 0.028	0.978 ± 0.027
CT-Palette (MI)	36.907 ± 8.467	35.702 ± 8.319	36.830 ± 8.317	38.311 ± 8.583	0.975 ± 0.027	$\mathbf{0.973 \pm 0.026}$	0.974 ± 0.028	0.979 ± 0.026

Table 4. p-values from the Wilcoxon signed-rank test comparing the distribution of muscle/SAT areas of ground-truth untruncated and outpainted slices by various models. Higher p-values indicate more similar distributions. p-values below 0.05 are considered significantly different and are marked with *. CT-Palette has the most similar muscle/SAT area distribution with the ground-truth.

Method	Overall (n=997)		T5 (n=339)		T8 (n=348)		T10 (n=310)	
	Muscle	SAT	Muscle	SAT	Muscle	SAT	Muscle	SAT
RFR-Net	2.0×10^{-63}*	8.9×10^{-33}*	3.0×10^{-20}*	3.7×10^{-12}*	5.2×10^{-28}*	5.7×10^{-16}*	5.5×10^{-18}*	1.0×10^{-7}*
S-EFOV	4.5×10^{-70}*	4.1×10^{-17}*	1.7×10^{-34}*	1.5×10^{-13}*	1.9×10^{-29}*	6.3×10^{-7}*	3.8×10^{-9}*	0.144
S-EFOV (FT)	5.0×10^{-104}*	2.3×10^{-95}*	2.2×10^{-39}*	5.7×10^{-29}*	1.6×10^{-38}*	1.1×10^{-37}*	2.0×10^{-29}*	7.8×10^{-34}*
CT-Palette (SI)	**0.116**	0.057	0.026	**0.406**	0.198	0.876	0.042	0.005
CT-Palette (MI)	0.065	**0.866**	**0.303**	0.348	**0.681**	**0.895**	**0.050**	**0.248**
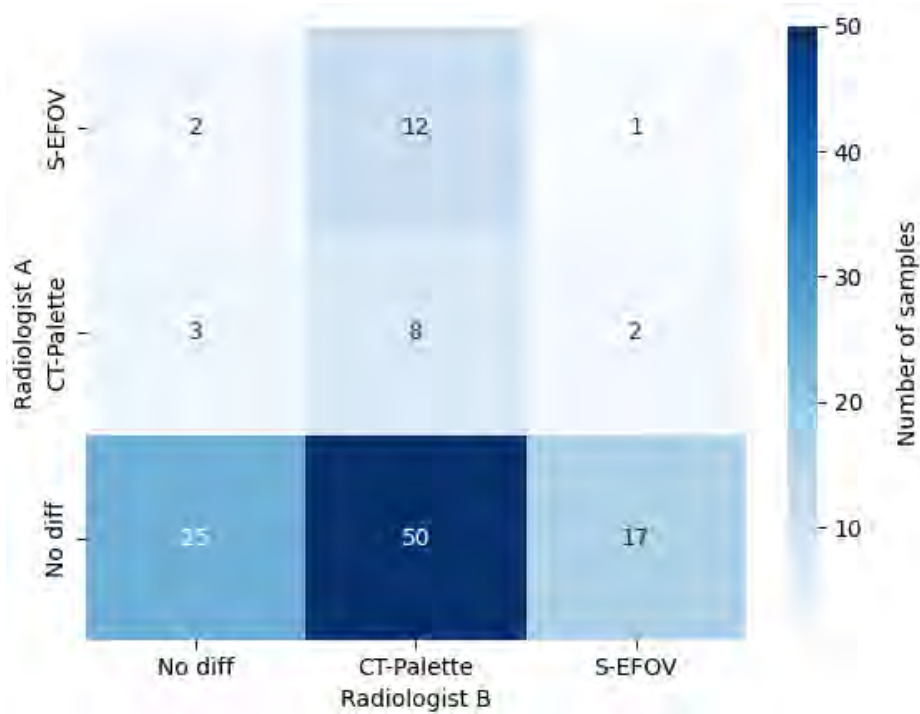
Fig. 1. The distribution of choices of two trained radiologists (A and B) on 120 real-world truncated slices (40 slices for each vertebral level T5, T8, and T10) recovered by CT-Palette (MI) and S-EFOV.