	FID	$\sqrt{L_2}$	LPIPS	PSNR	SSIM
ILVR	28.85	118.1	0.557	11.81	0.326
SDEdit	28.31	94.00	0.516	16.69	0.399
EGSDE	32.26	72.68	0.466	16.00	0.430
CycleDiff	33.25	79.47	0.443	15.19	0.460
DMD	40.40	107.0	0.503	16.69	0.398
RDMD(0.1)	30.81	62.40	**0.379**	**21.67**	**0.564**
EGSDE† (p)	30.93	**53.44**	0.441	18.32	0.510
EGSDE (p)	43.57	**42.04**	0.390	20.35	0.574
RDMD(0.15)	32.11	54.75	**0.339**	**21.96**	**0.606**

Table 1: Comparison of RDMD with diffusion-based baselines on 256×256 CelebA-HQ *Male* $\rightarrow$ *Female* in latent space. EGSDE(p) models operate in pixel space.

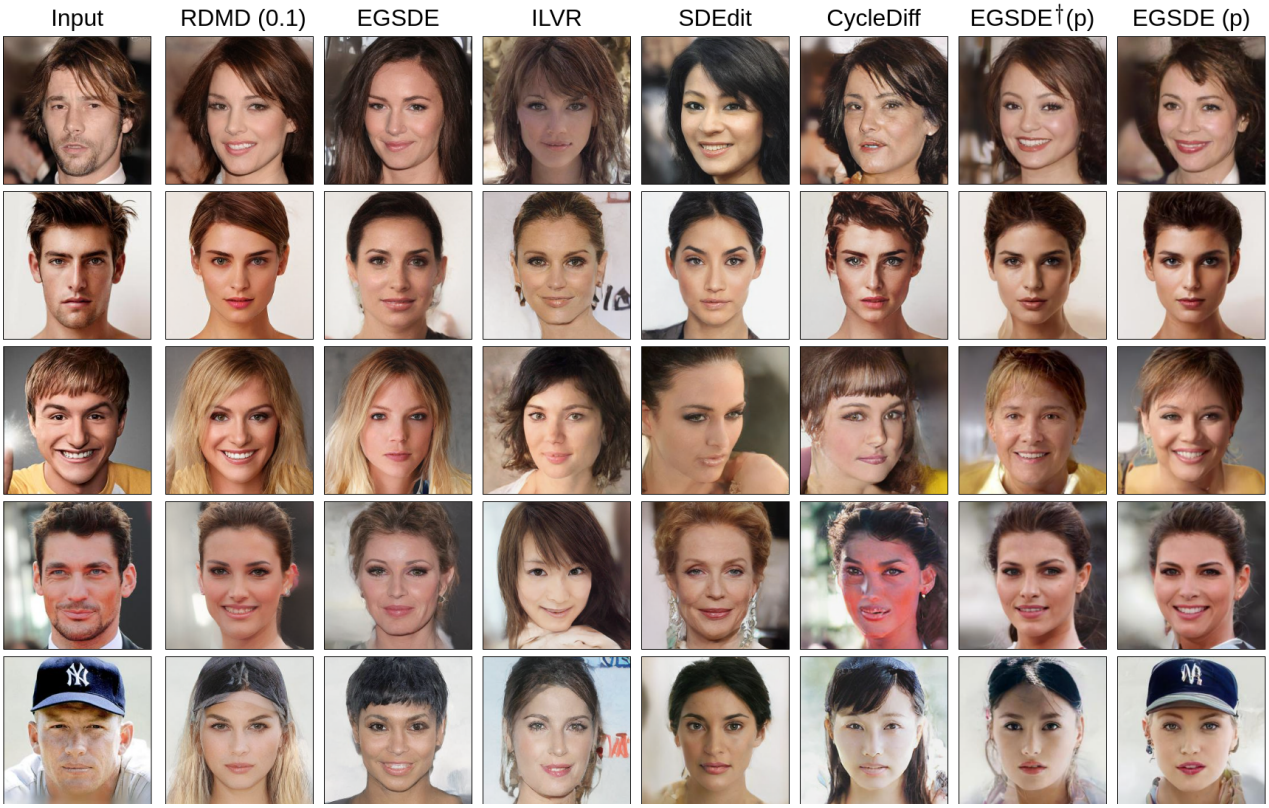

Figure 1: Comparison of RDMD in 256×256 *Male*$\rightarrow$*Female* translation problem in latent space. EGSDE models, marked with "(p)", operate in the pixel space. DDIB is excluded due to poor image quality (FID > 70.0).

	FID	$\sqrt{L_2}$	LPIPS	PSNR	SSIM
ILVR	26.43	105.37	0.620	12.76	0.236
SDEdit	38.49	82.73	0.543	14.93	0.330
RDMD	32.57	**55.55**	**0.322**	**18.77**	**0.562**

Table 2: Comparison of RDMD with baselines on 256×256 *Outdoor* $\rightarrow$ *Church* in latent space.

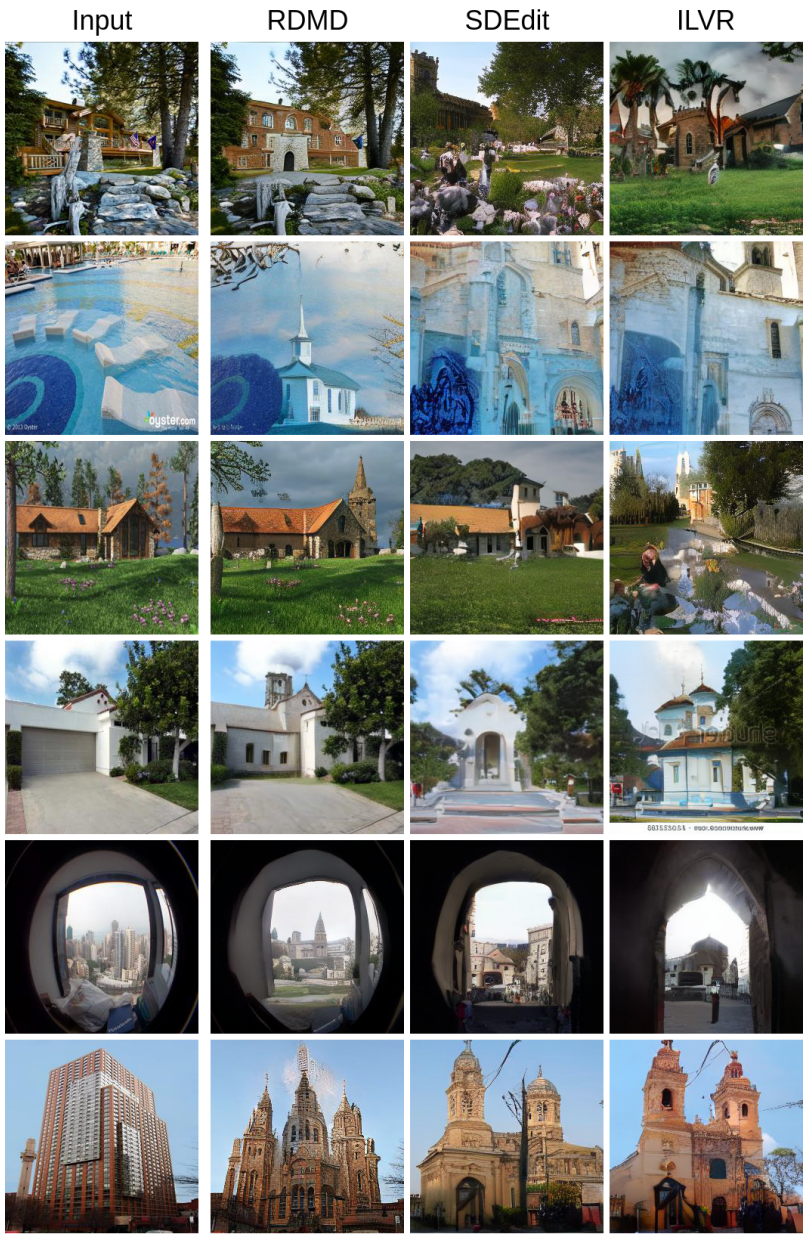

Figure 2: Comparison of RDMD in 256×256 *Outdoor*→*Church* translation problem in latent space.