

Prompt	Model	Prompt-Image Alignment \uparrow			Safe Generation Rate \uparrow						ABSS
		CLIPScore	VQAScore	A-Avg.	GPT-4o	DeepSeek-V3	Qwen2.5-VL	Q16	FLUX-Filter	S-Avg.	
Misbinding (Ours)	SD1.5	0.8351	0.7764	0.8058	0.7476	0.7366	0.7248	0.7955	0.6866	0.7402	0.6032
	SDXL	0.8921	0.8249	0.8585	<u>0.7221</u>	0.6968	0.7012	0.7891	0.6778	0.7174	0.6234
	FLUX.1-dev	0.8674	0.7935	0.8305	0.7153	0.7075	<u>0.7088</u>	0.8055	0.6868	<u>0.7248</u>	<u>0.6130</u>
	UniPortrait	0.7810	0.7324	0.7567	0.6946	0.5979	0.5993	0.7669	0.6533	0.6424	0.5537
	PhotoMaker	0.7873	0.7524	0.7699	0.6976	0.6645	0.6197	0.6566	0.6477	0.6572	0.5636
	InfiniteYou	0.8050	0.7630	0.7840	0.6718	0.6569	0.6693	0.7677	0.6521	0.6835	0.5828
	PuLID	<u>0.8722</u>	<u>0.8013</u>	<u>0.8368</u>	0.6466	0.6447	0.6399	0.7741	0.6659	0.6742	0.5983
	FLUX.1-Kontext	0.7657	0.7146	0.7402	0.7013	<u>0.7186</u>	0.6983	<u>0.7813</u>	0.7039	0.7207	0.5771

Table 5: Evaluating model vulnerability to Misbinding Prompts. **Bold** and underline indicate the best and second-best, respectively (for generative models, higher Prompt-Image Alignment and higher Safe Generation Rate indicate better generation performance).