

Model	DB	HR	LE	NR	RR	SR	SNR	WR	CT	DP	OR	ORP	BRP	CA	SA	RA
HIVE	4.34	4.80	4.18	5.05	4.70	4.71	5.81	4.11	5.80	5.72	5.29	5.61	5.78	5.86	5.89	5.71
InstructDiffusion	3.67	4.72	4.70	5.16	4.40	4.19	5.67	5.05	5.51	5.44	4.82	5.46	5.60	5.82	5.57	5.47
InstructPix2Pix	3.82	4.66	3.61	4.50	4.85	4.09	6.03	4.50	5.69	5.74	5.26	5.68	5.75	5.86	5.70	5.68
MagicBrush	3.74	4.56	3.80	4.23	4.75	3.93	5.80	4.42	5.76	5.77	5.47	5.58	6.01	5.96	5.82	5.73
MGIE	3.86	4.72	4.83	4.92	4.73	4.12	5.59	4.68	5.86	5.73	5.31	5.45	5.55	5.90	5.77	5.63
InstructEdit	3.83	4.61	3.90	4.79	4.94	4.11	6.41	4.85	5.75	5.65	5.93	5.74	5.65	5.95	5.66	5.85
InstructAny2Pix	5.74	5.87	5.84	6.22	6.13	5.56	6.47	5.95	6.29	6.26	6.24	6.21	6.49	6.36	6.54	6.30
HQ-Edit	4.71	5.25	4.82	5.82	5.23	4.49	6.00	5.23	5.85	5.84	5.92	6.00	5.89	6.12	5.72	5.83

Table I: Aesthetic Predictor’s Score of edited images using **original instructions**. Abbreviations: DB = Deblurring, HR = Haze Removal, LE = Lowlight Enhancement, NR = Noise Removal, RR = Rain Removal, SR = Shadow Removal, SNR = Snow Removal, WR = Watermark Removal, CT = Counting, DP = Direction Perception, OR = Object Removal, ORP = Object Replacement, BRP = Background Replacement, CA = Color Alteration, SA = Style Alteration, RA = Region Accuracy.

Model	DB	HR	LE	NR	RR	SR	SNR	WR	CT	DP	OR	ORP	BRP	CA	SA	RA
HIVE	4.17	4.71	4.45	5.02	4.98	4.18	6.02	4.51	5.85	5.72	5.76	5.70	5.78	5.94	5.60	5.99
InstructDiffusion	3.65	4.68	4.45	5.11	4.63	4.01	5.69	4.97	5.65	5.57	5.52	5.37	5.73	5.81	5.65	5.71
InstructPix2Pix	3.82	4.49	4.23	4.19	4.88	3.67	6.23	4.29	5.68	5.72	5.75	5.72	5.71	5.92	5.53	5.87
MagicBrush	3.87	4.59	4.41	4.07	4.82	3.63	6.07	4.41	5.76	5.72	5.68	5.63	6.09	5.93	5.77	5.83
MGIE	3.91	4.63	4.57	4.80	4.74	4.19	6.04	4.77	5.76	5.62	5.58	5.48	5.73	5.85	5.73	5.79
InstructEdit	3.79	4.57	3.82	4.77	4.87	4.06	6.46	4.77	5.73	5.63	5.93	5.74	5.68	5.94	5.62	5.85
InstructAny2Pix	5.79	5.85	5.74	6.26	6.04	5.60	6.46	5.88	6.29	6.09	6.35	6.19	6.35	6.40	6.48	6.41
HQ-Edit	4.98	5.08	5.55	5.81	5.31	4.31	6.04	5.02	5.81	5.91	5.83	5.98	6.06	6.02	5.66	6.04

Table II: Aesthetic Predictor’s Score of edited images using **diverse instructions**.

Model	Animal	Object	Scenery	Plant	Human	Global
HIVE	32.72	36.93	42.53	35.74	30.55	38.69
InstructDiffusion	33.86	37.59	40.91	38.70	33.59	38.82
InstructPix2Pix	30.99	33.56	39.87	32.59	26.25	38.42
MagicBrush	34.53	39.94	43.82	39.90	32.37	38.15
MGIE	31.82	39.11	45.48	40.22	33.08	37.62
InstructEdit	24.83	32.27	40.18	31.54	29.05	33.97
InstructAny2Pix	28.17	29.34	38.45	25.62	27.43	38.06
HQ-Edit	26.74	30.53	39.98	24.91	24.66	26.91

Table III: Quantitative results of different categories using **original instructions**.

Model	Animal	Object	Scenery	Plant	Human	Global
HIVE	28.50	34.31	45.31	34.52	29.06	37.63
InstructDiffusion	25.04	32.44	40.54	31.98	29.10	34.11
InstructPix2Pix	27.44	35.67	39.58	33.65	25.74	36.87
MagicBrush	32.82	37.49	45.60	35.95	32.26	37.39
MGIE	32.58	35.77	44.02	36.29	35.22	37.76
InstructEdit	32.16	36.01	40.45	35.94	31.11	36.15
InstructAny2Pix	28.24	31.65	39.69	28.57	27.07	37.24
HQ-Edit	27.33	30.96	41.47	26.07	23.22	27.05

Table IV: Quantitative results of different categories using **diverse instructions**.

Metric	DB	HR	LE	NR	RR	SR	SNR	WR	CT	DP	OR	ORP	BRP	CA	SA	RA
CLIP Sim	88.65	86.21	88.23	88.15	84.89	87.03	90.10	90.12	91.35	93.88	93.21	92.88	86.69	92.81	86.91	93.28
Jaccard Sim	16.78	14.68	13.98	15.29	16.87	21.41	27.36	28.20	45.92	59.67	42.32	36.62	35.62	30.13	22.88	37.11
TF-IDF	15.41	10.46	6.09	18.71	12.16	20.56	26.27	16.50	60.53	73.10	60.16	66.94	30.27	65.53	22.19	62.21
Word2Vec	52.21	46.41	49.86	55.23	51.98	59.42	65.55	68.54	74.64	80.98	74.08	77.22	69.37	75.89	70.59	75.23
FastText	74.94	69.65	60.21	72.94	72.03	84.68	82.74	77.55	79.84	93.07	77.41	74.35	90.96	72.52	74.34	74.40

Table V: Text similarity between original instructions and diverse instructions using different similarity metrics.