

Method	Editing Quality			
	Quality	Adherence	Consistency	Overall
Gemini 2	3.021	3.387	2.868	3.215
GPT-4o	3.184	3.450	3.059	3.503
Ours	3.712	3.487	3.603	3.465

CA-PAG		DARM		AUC(%)
Initial	Refine	DA	Update	
✓				94.9
✓	✓			96.0
✓			✓	97.6
✓	✓		✓	97.7
✓		✓		96.0
✓	✓	✓		96.8
✓	✓	✓	✓	98.2

Model	FG-OVD			
	Hard	Medium	Easy	Trivial
CLIP	15.4	25.3	25.7	38.8
EVA-CLIP	18.3	38.4	35.2	62.7
LongCLIP	9.6	19.7	16.0	39.8
FineCLIP	22.8	46.0	46.0	73.6
FG-CLIP [‡]	48.4	69.5	71.2	89.7

β -CLIP <i>CE</i>				
K=6, $\beta=0.5$	19.3	37.9	41.1	62.2
K=36, $\beta=0.5$	<u>28.3</u>	<u>48.8</u>	53.3	70.9
K=6, $\beta=0.75$	20.3	41.1	44.3	67.9
K=36, $\beta=0.75$	28.7	49.1	<u>52.1</u>	<u>72.2</u>

β -CLIP <i>BCE</i>				
K=6, $\beta=0.5$	15.1	34.6	29.2	32.4
K=36, $\beta=0.5$	14.4	29.4	23.7	45.6
K=6, $\beta=0.75$	15.0	36.3	33.1	36.8
K=36, $\beta=0.75$	15.4	31.5	23.9	45.9

Models	CRR		WRR		LRR	
	Val	Test	Val	Test	Val	Test
Proposed HWR	94.80	95.20	83.60	84.70	72.78	73.24

Method	ImageNet-100			CUB		
	All	Old	New	All	Old	New
SimGCD	83.0	93.1	77.9	60.3	65.6	57.7
+ LSP	85.4	94.2	81.0	64.6	71.6	61.1
+ DAS	84.8	92.9	80.7	66.5	67.0	66.2
+ LSP + DAS (Ours)	87.1	93.9	83.7	69.0	69.1	68.9