Appendix

A ABLATION RESULTS

We conducted an ablation study to check the influence of the underlay design element. We labelled all the underlay design elements in each poster image of our dataset, and duplicated our dataset but removed all the underlay design elements (Poster101K*) to figure out how the underlay affected the canvas overlay result. The experimental results are given in Tab. 1. The table shows lower (i.e., better) overlap results when all the underlay elements are ignored (the 2nd metric of each task). Besides, the variation of the alignment score explains that the appearance of the underlay elements makes the canvas look neater (the 3rd metric of each task).

CLG-LO	Beau. Constrains Generation				Rela. Constraints Generation				U-Gen			
Dataset	mIOU	Overlap	Align.	FID	mIOU	Overlap	Align.	FID	mIOU	Overlap	Align.	FID
Poster101K*	0.29	1.85	0.24	0.06	0.31	22.01	1.00	0.02	0.31	22.92	0.88	0.02
Poster101K	0.21	0.80	0.18	0.28	0.26	40.80	0.40	0.07	0.25	43.76	0.29	0.15
DLT	C-Gen				CS-Gen				U-Gen			
Dataset	pIOU	Overlap	Align.	FID	pIOU	Overlap	Align.	FID	pIOU	Overlap	Align.	FID
Poster101K*	11.22	36.9	0.80	6.06	13.5	48.99	0.77	0.6	8.16	27.6	0.84	3.48
Poster101K	17.33	50.34	0.70	5.68	19.58	62.59	0.67	0.5	14.47	40.52	0.71	3.85
LayoutFormer++	C-Gen				CS-Gen				U-Gen			
Dataset	mIOU	Overlap	Align.	FID	mIOU	Overlap	Align.	FID	mIOU	Overlap	Align.	FID
Poster101K*	0.268	0.503	0.0014	0.374	0.363	0.345	0.012	0.027	0.527	0.345	0.056	4.959
Poster101K	0.260	0.668	0.0025	0.479	0.368	0.457	0.008	0.022	0.488	0.603	0.017	1.951
LACE	C-Gen				CS-Gen				U-Gen			
Dataset	mIOU	Overlap	Align.	FID	mIOU	Overlap	Align.	FID	mIOU	Overlap	Align.	FID
Poster101K*	0.131	8.026	1.025	1.641	0.170	11.434	1.402	1.628	-	37.895	1.404	1.676
Poster101K	0.109	22.819	0.899	1.327	0.155	31.068	1.051	1.296	-	38.036	0.665	1.455

Table 1: Ablation on Underlay labels (Poster101K* represents the filtered dataset without underlay labels)

B IMPLEMENTATION SPECIFICATIONS

All models for testing the Poster101K dataset are implemented in PyTorch1.9. All the experiments are conducted on 4 NVIDIA GeForce RTX 3090 GPUs. The LACE model is trained with a 1e-6 learning rate, and the others are trained with 1e-4. The batch size is 256 for the LACE model and 64 for the others.

C QUALITATIVE COMPARISON

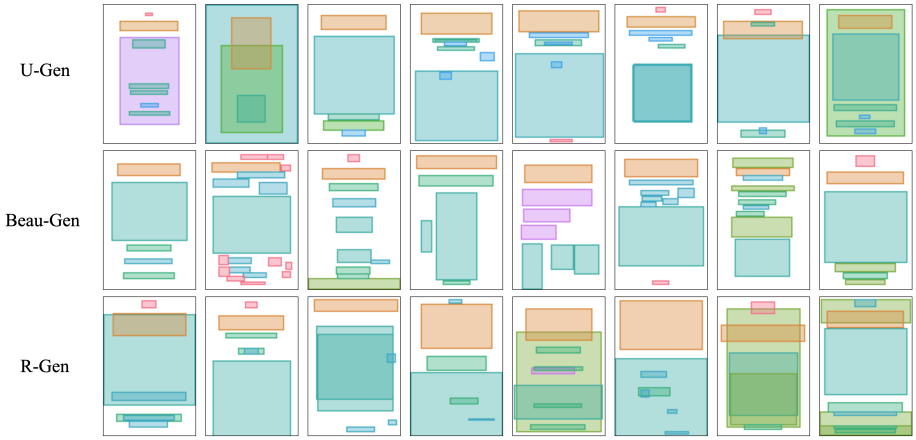

Figure 1: Qualitative results of CGL-LO on Poster101K

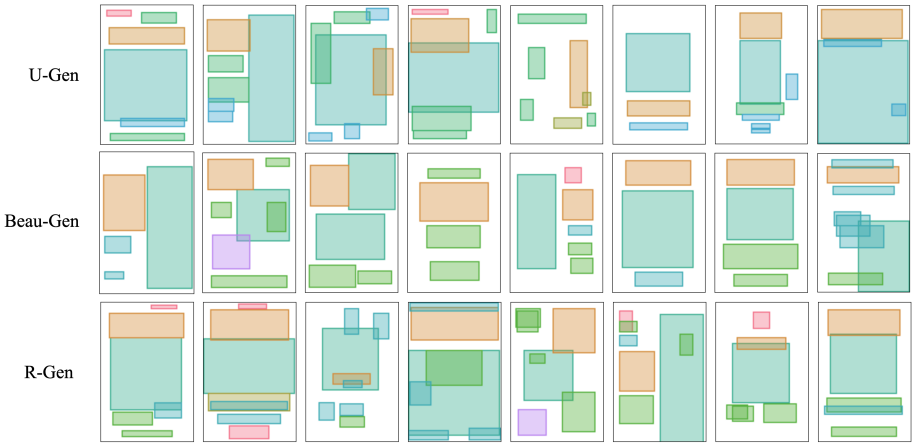

Figure 2: Qualitative results of CGL-LO on Poster101K*

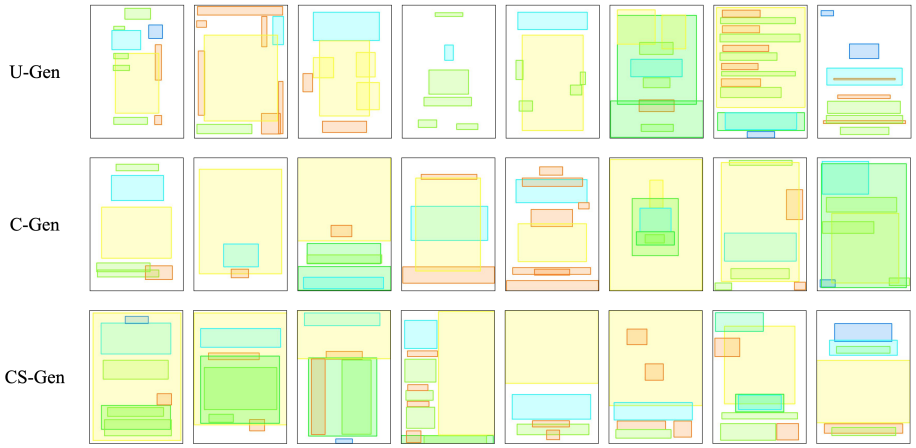

Figure 3: Qualitative results of DLT on Poster101K

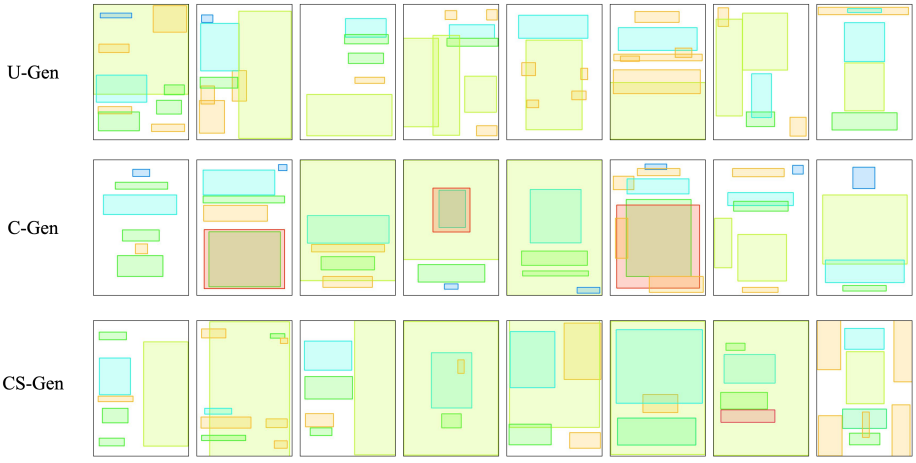

Figure 4: Qualitative results of DLT on Poster101K*

Figure 5: Qualitative results of LayoutFormer++ on Poster101K

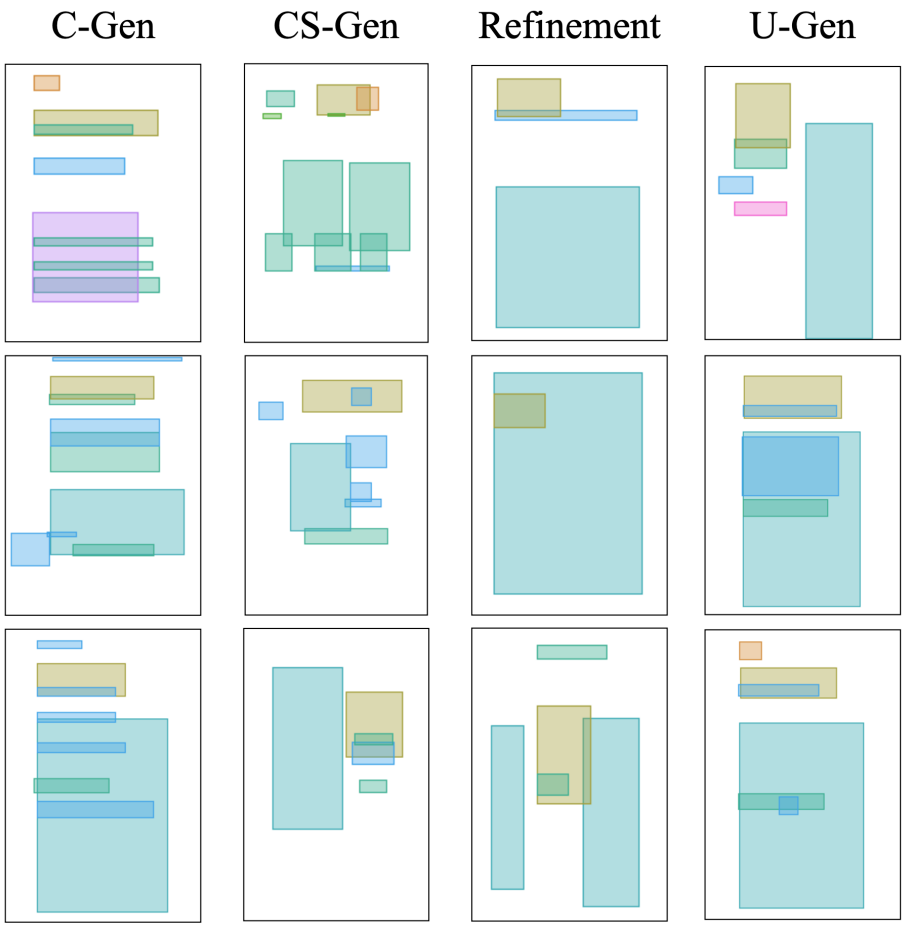

Figure 6: Qualitative results of LayoutFormer++ on Poster101K*

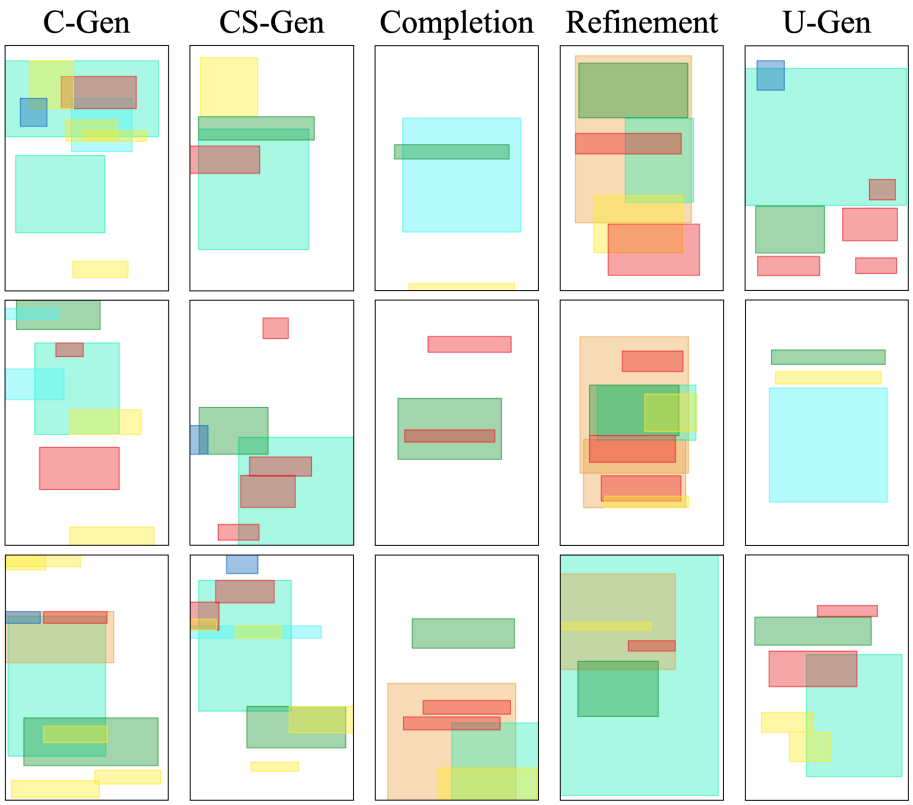

Figure 7: Qualitative results of LACE on Poster101K

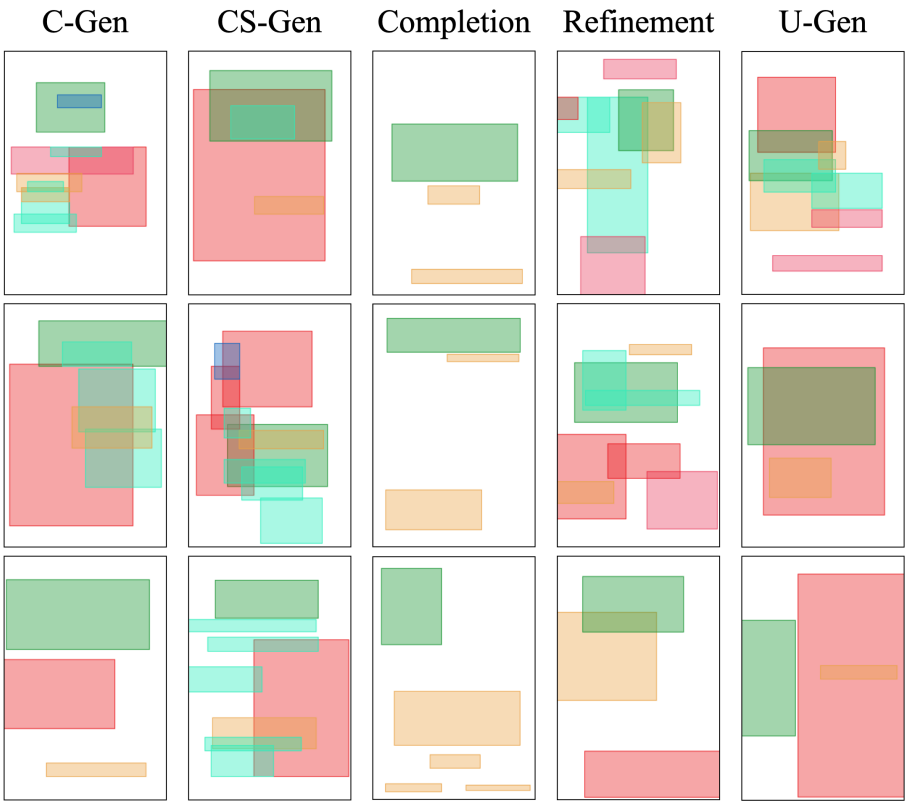

Figure 8: Qualitative results of LACE on Poster101K*