FloorPlanCAD: A Large-Scale CAD Drawing Dataset for Panoptic Symbol Spotting
(Supplementary Material)

Zhiwen Fan[1]† Lingjie Zhu[1]† Honghua Li[1] Xiaohao Chen[1] Siyu Zhu[1] Ping Tan[1,2]
[1]Alibaba A.I. Labs [2]Simon Fraser University †Equal contribution

1. Introduction

Due to space limitations in the paper, this supplemental material contains more descriptions about the dataset and more quantitative and qualitative results of the proposed methods.

2. Dataset

2.1. Entity Distribution among Classes

In Table 1, we provide statistics of graphical entities of 30 object classes, including 28 thing classes and 2 stuff classes. We can see that the *wall* class takes up a large portion of the whole dataset.

2.2. Visualizations on Train Set

We visualize several samples in train set of FloorPlanCAD dataset in Figure 1 and Figure 2 to demonstrate the variety of the proposed dataset.

3. Per-Class Evaluation

3.1. Semantic Symbol Spotting

The third column of Table 1 shows semantic symbol spotting results of DeepLabv3+ [1] and the proposed GCN-based method on all object classes. Here, we use *weighted F1 score* as the metric which use the entity length $\log(1 + L(e_i))$ to weight the *TP*, *FP* and *FN*. We can see the GCN-based methods significantly outperforms DeepLabv3+ in *wall* class since *wall* class always mixes with thing classes.

3.2. Instance Symbol Spotting

We provide the class-wise mAP for Faster R-CNN [3], FCOS [4] and YOLOv3 [2] in the fourth column of Table 1. The results includes 28 thing classes. We can notice Faster R-CNN is comparable with FCOS and both outperform YOLOv3 which may caused by our dataset contains various scenes, 28 possible symbol classes and complex background.

3.3. Panoptic Symbol Spotting Results

In the fifth column of Table 1, we provide the detailed evaluation results of panoptic quality(PQ), segmentation quality(SQ) and recognition quality(RQ). Additional visualization results of PanCADNet on FloorPlanCAD dataset are shown in Figure 3, Figure 4 and Figure 5. The results for 2 stuff classes are obtained by GCN head while 28 thing classes are obtained by detection head [3].

4. Limitations and Future Works

The proposed PanCADNet benefits from the GCN head which takes the vectorized data as input, utilizing both the geometric feature and aligned CNN features, aggregating neighbour information by graph topology, resulting in a good results for the two key stuff classes (i.e. *wall* and *parking*). With the help of predicted bounding box using a detection head, we can distinguish each instance in thing classes easily.

Although the proposed method can solve the panoptic symbol spotting problem, some limitations still exist: as is pointed out in Figure 6, some background elements may be mis-classified by the predicted box. The future works include generating instance proposals in vector space which can propose object instances in a more flexible way.

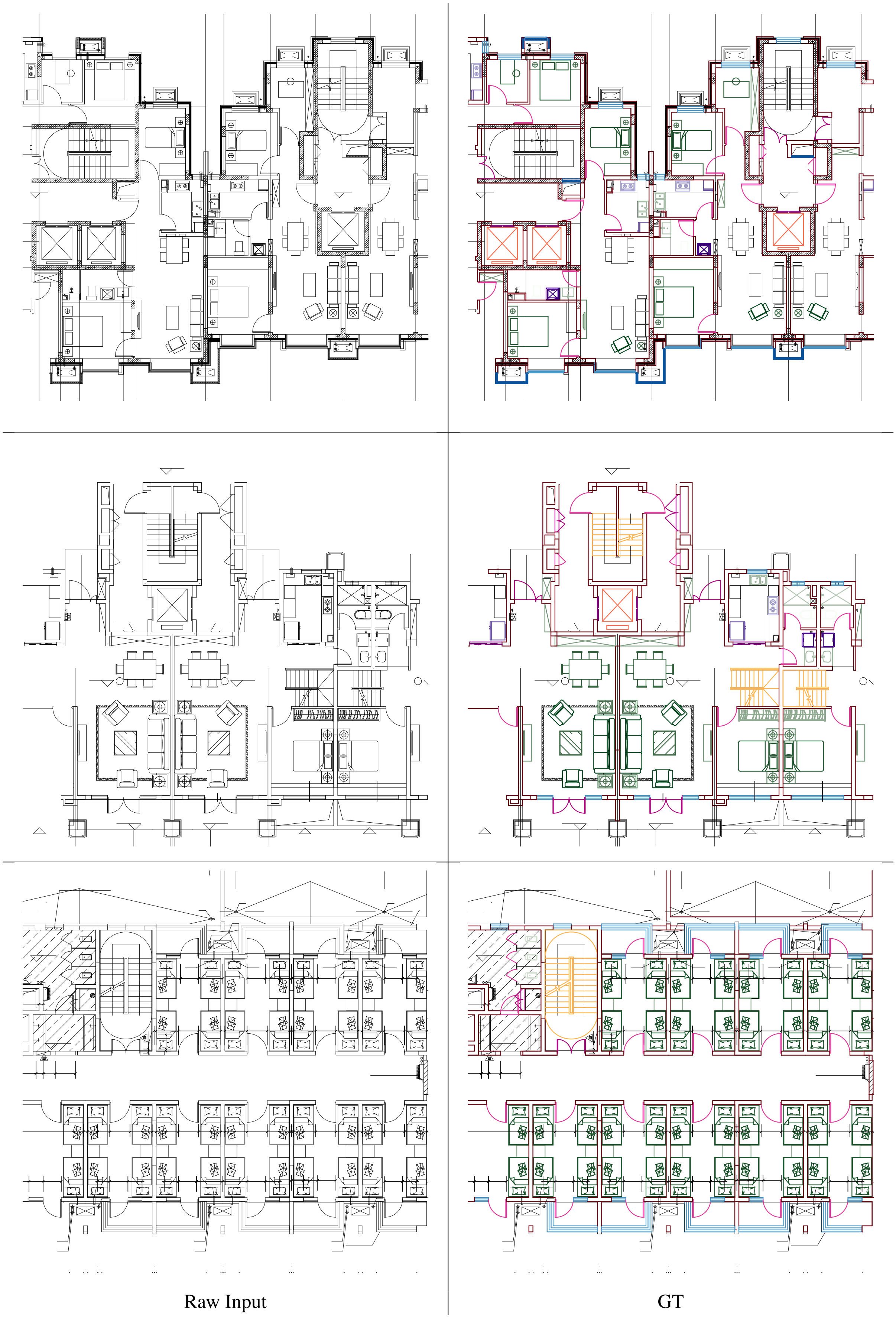

Figure 1: Exemplary raw inputs and annotations in FloorPlanCAD, see the main manuscript for annotation details. The images are part of our *train* set of *residential building* and *school* CAD drawings.

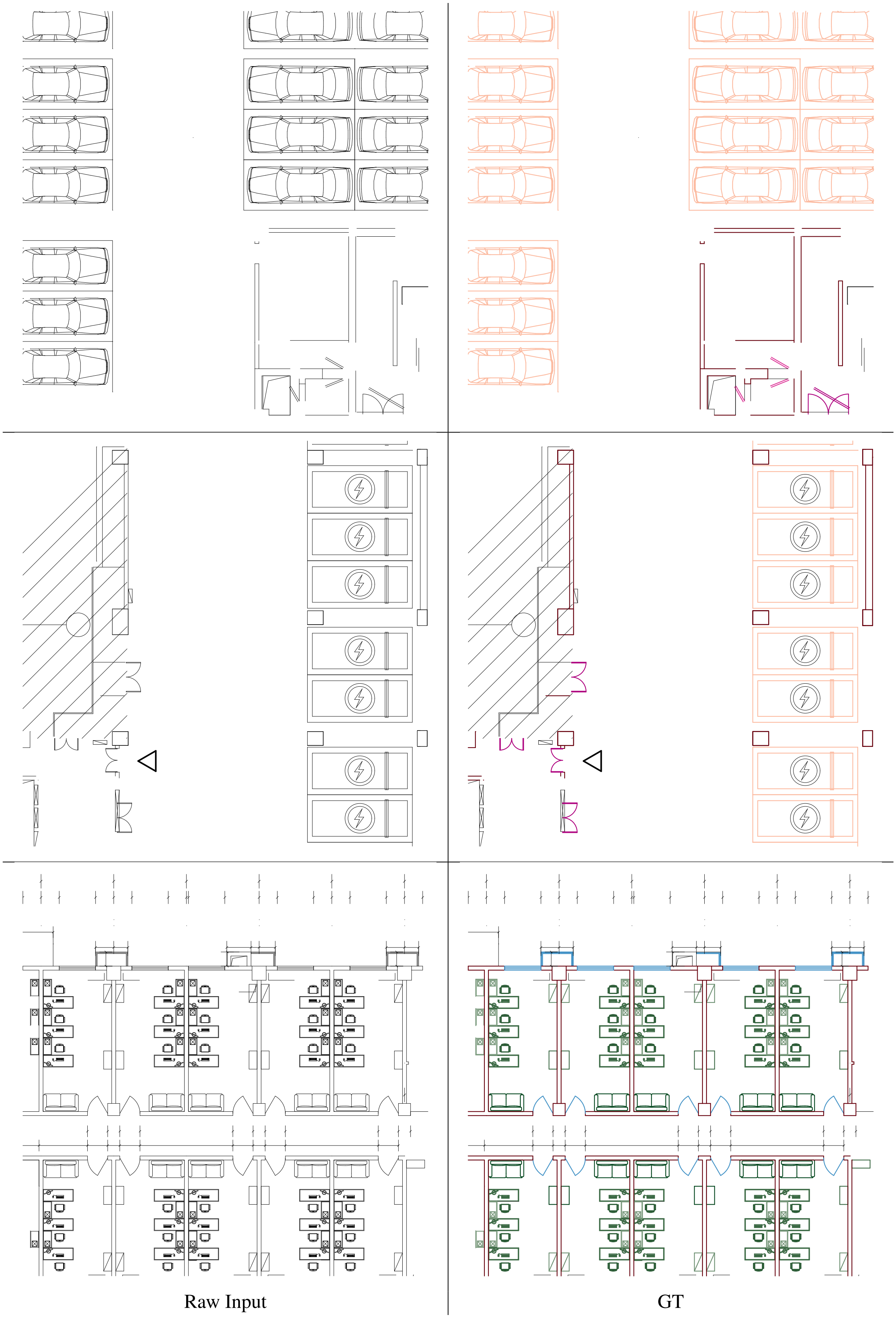

Figure 2: Exemplary raw inputs and annotations in FloorPlanCAD, see the main manuscript for annotation details. The images are part of our *train* set of *underground parking lot* and *office building* CAD drawings.

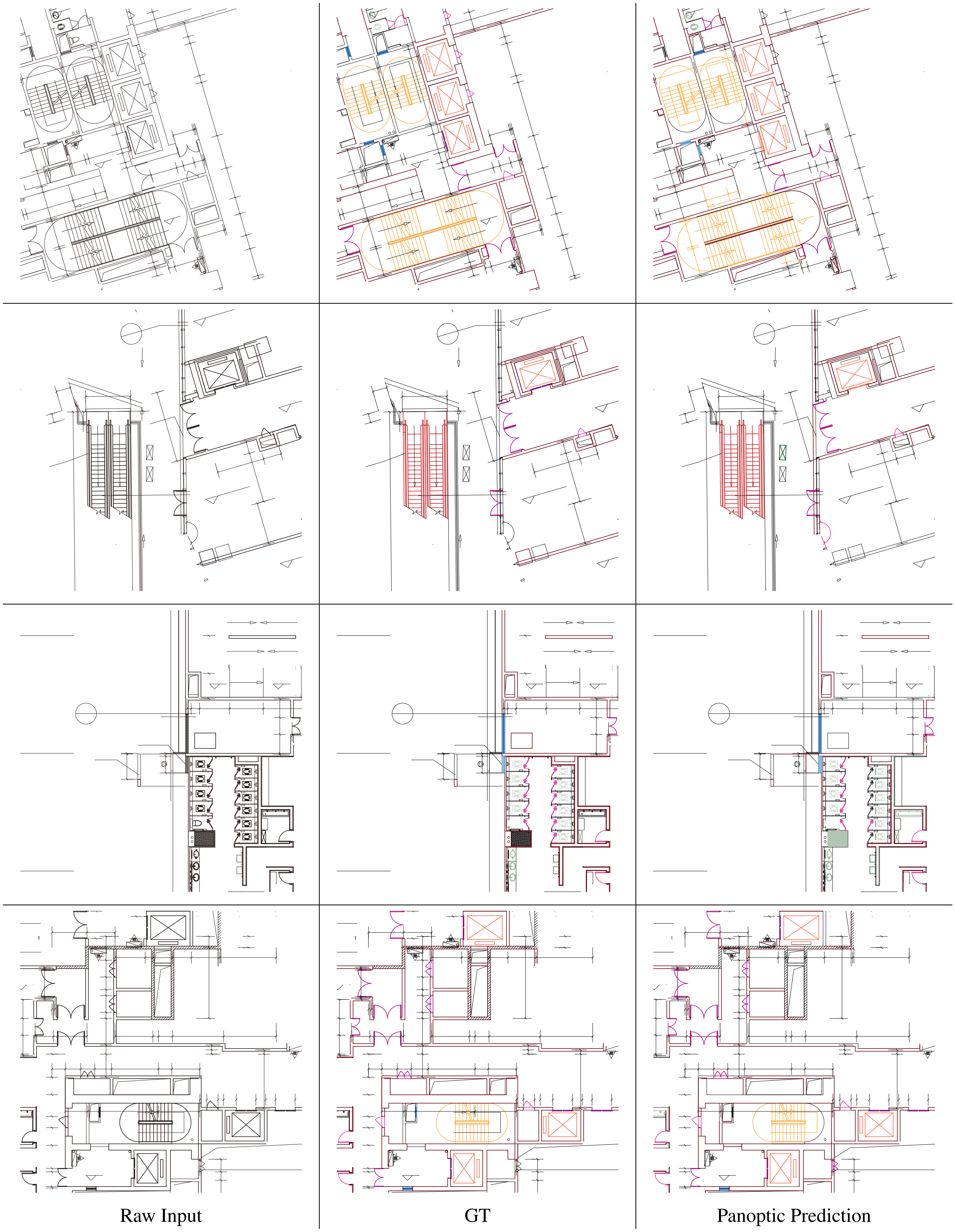

Figure 3: Results of PanCADNet on FloorPlanCAD, see the main manuscript for annotation details. The images are part of our *test* set of *large shopping mall* CAD drawings.

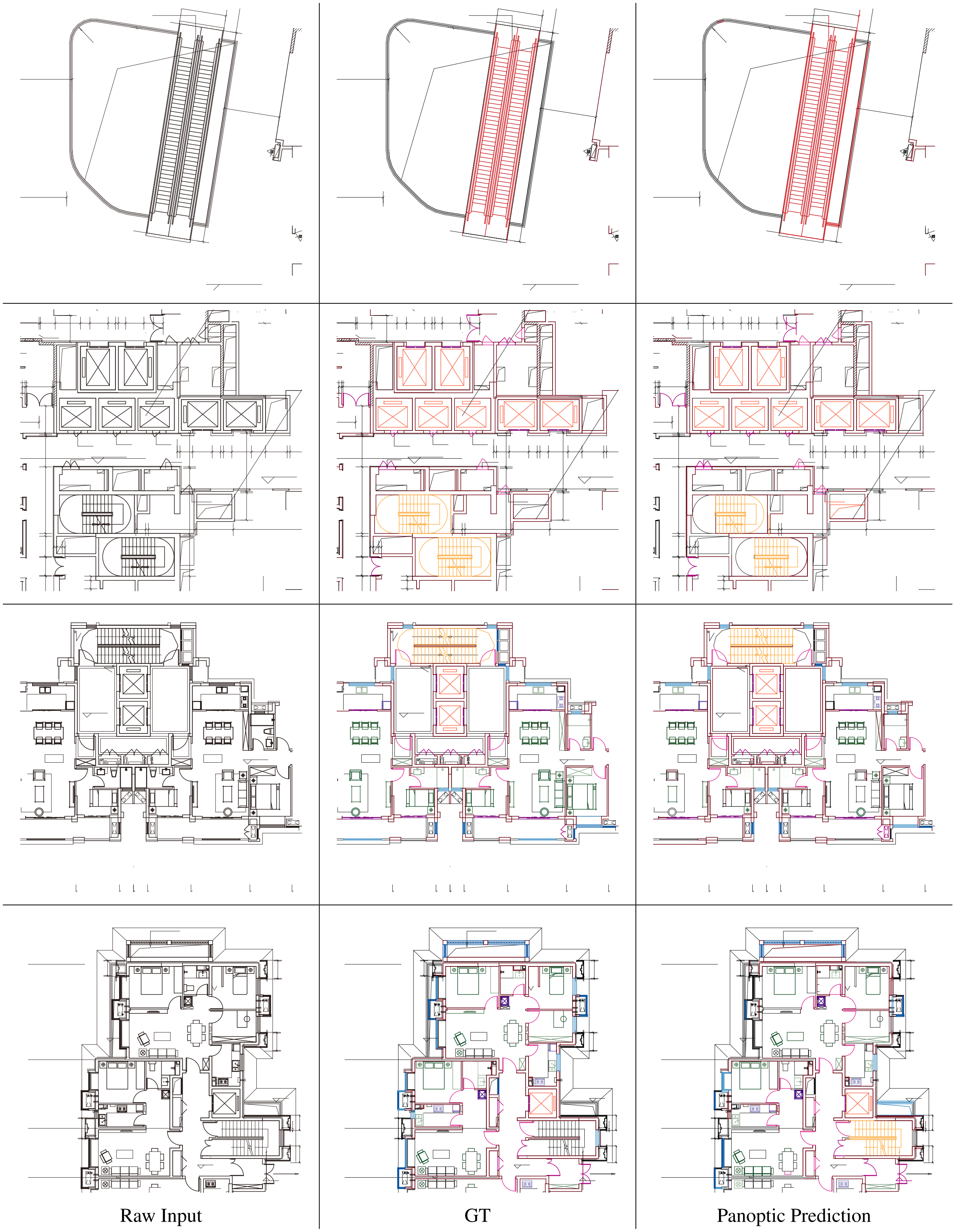

Figure 4: Results of PanCADNet on FloorPlanCAD, see the main manuscript for annotation details. The images are part of our *test* set of *large shopping mall* and *residential building* CAD drawings.

Figure 5: Results of PanCADNet on FloorPlanCAD, see the main manuscript for annotation details. The images are part of our *test* set of *underground parking lot*, *residential building* and *hotel* CAD drawings.

Class	Property	Semantic Symbol Spotting		Instance Symbol Spotting			Panoptic Symbol Spotting		
	#Entity($\times 10^4$)	weighted F1		mAP			PQ	SQ	RQ
		GCN-based	DeepLabv3+ [1]	Faster R-CNN [3]	FCOS [4]	YOLOv3 [2]	PanCADNet		
single door	301	0.885	0.827	0.843	0.859	0.829	0.763	0.878	0.869
double door	285	0.796	0.831	0.779	0.771	0.743	0.748	0.845	0.885
sliding door	122	0.874	0.876	0.556	0.494	0.481	0.763	0.895	0.852
window	266	0.691	0.603	0.518	0.465	0.379	0.459	0.795	0.577
bay window	15.1	0.050	0.163	0.068	0.169	0.062	0.154	0.595	0.260
blind window	98.6	0.833	0.856	0.614	0.520	0.322	0.706	0.869	0.813
opening symbol	2.68	0.451	0.721	0.496	0.542	0.168	0.455	0.945	0.481
stairs	197	0.857	0.853	0.464	0.487	0.370	0.608	0.784	0.775
gas stove	175	0.789	0.847	0.503	0.715	0.601	0.743	0.957	0.776
refrigerator	55.0	0.705	0.730	0.767	0.774	0.723	0.769	0.888	0.866
washing machine	115	0.784	0.569	0.379	0.261	0.374	0.430	0.719	0.599
sofa	105	0.606	0.674	0.160	0.133	0.435	0.252	0.928	0.272
bed	1480	0.893	0.908	0.713	0.738	0.664	0.805	0.909	0.886
chair	176	0.524	0.543	0.112	0.087	0.132	0.481	0.802	0.600
table	77.9	0.354	0.496	0.175	0.109	0.173	0.228	0.811	0.282
bedside cupboard	68.0	0.509	0.770	0.231	0.363	0.310	0.600	0.825	0.727
TV cabinet	32.8	0.581	0.611	0.231	0.187	0.247	0.426	0.800	0.533
half-height cabinet	4.18	0.144	0.179	0.133	0.108	0.110	0.009	0.970	0.009
high cabinet	20.1	0.325	0.426	0.271	0.188	0.296	0.287	0.820	0.351
wardrobe	502	0.462	0.426	0.325	0.354	0.354	0.433	0.821	0.527
sink	512	0.825	0.844	0.468	0.470	0.384	0.778	0.895	0.870
bath	254	0.540	0.432	0.422	0.446	0.430	0.413	0.720	0.573
bath tub	45.8	0.476	0.637	0.259	0.248	0.215	0.817	0.856	0.955
squat toilet	139	0.842	0.904	0.836	0.821	0.599	0.901	0.989	0.911
urinal	118	0.866	0.923	0.780	0.762	0.622	0.921	0.981	0.938
toilet	298	0.875	0.864	0.666	0.599	0.664	0.831	0.906	0.917
elevator	78.7	0.948	0.900	0.767	0.816	0.750	0.838	0.897	0.935
escalator	10.0	0.744	0.864	0.115	0.190	0.129	0.439	0.718	0.612
parking	163	0.529	0.667	-	-	-	0.251	0.661	0.380
wall	1880	0.814	0.634	-	-	-	0.451	0.661	0.682
total	7600	0.798	0.714	0.452	0.453	0.413	0.561	0.838	0.660

Table 1: Dataset entities number and quantitative results for semantic symbol spotting, instance symbol spotting and panoptic symbol spotting of each category. Entity length weighted F1 is used for semantic symbol spotting evaluation, mAP is used for instance symbol spotting evaluation, panoptic quality is used for panoptic symbol spotting evaluation.

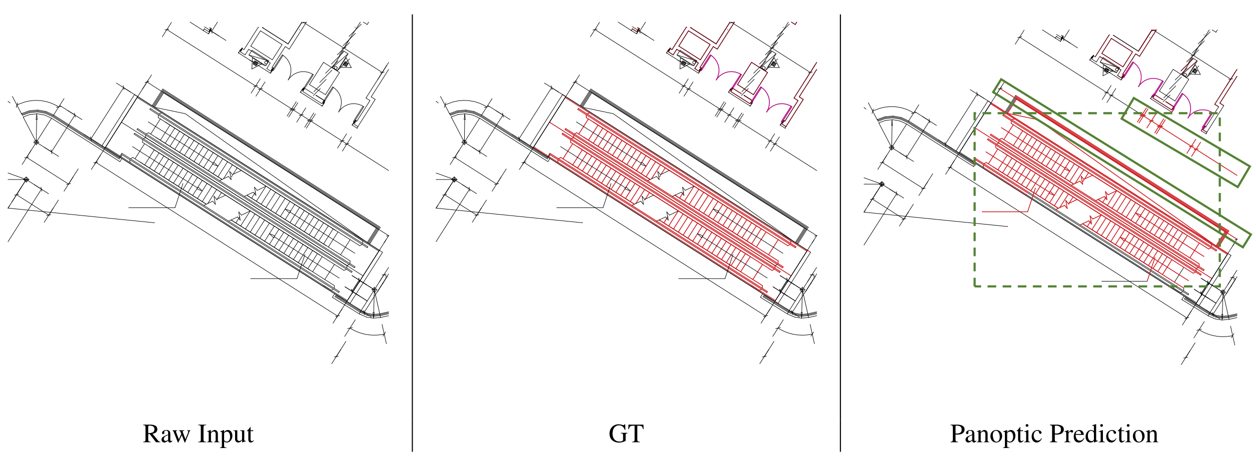

Figure 6: A failure case of PanCADNet, where the predicted box of the stairs is shown using a dotted green box, the entities mis-classified by the predicted box are highlighted by solid green boxes.

References

[1] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 801–818, 2018.

[2] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[3] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6):1137–1149, 2016.

[4] Zhi Tian, Chunhua Shen, Hao Chen, and Tong He. Fcos: Fully convolutional one-stage object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9627–9636, 2019.