[Supplementary Material] House-GAN++: Generative Adversarial Layout Refinement Network towards Intelligent Computational Agent for Professional Architects

Nelson Nauata
Simon Fraser University
nnauata@sfu.ca

Sepidehsadat Hosseini
Simon Fraser University
sepidh@sfu.ca

Kai-Hung Chang
Autodesk Research
kai-hung.chang@autodesk.com

Hang Chu
Autodesk Research
hang.chu@autodesk.com

Chin-Yi Cheng
Autodesk Research
chin-yi.cheng@autodesk.com

Yasutaka Furukawa
Simon Fraser University
furukawa@sfu.ca

The supplementary document provides 1) a description of the graph edit distance metric computation, 2) additional details on the user study, 3) details on the vectorization algorithm using Floor-SP, 4) discussion on rule based methods and windows generation and 5) ablation on the input graph information plus, additional generated layout samples.

1. Modified graph edit distance (GED)

The original House-GAN computes the graph edit distance (GED) metric using the NetworkX library, which is computationally intensive as it can take any two graphs as input. In order to speed up the GED computation, we use the information about nodes correspondences between the input graph and estimated graph from generated layouts. For additional simplification, we change the set of action candidates for computing the graph edit distance, we drop node/edge substitution and allow only node/edge insertion and node/edge deletion. When constructing the graph from a generated layout, multiple components may be generated for a node or an edge. For the former, we keep the largest one and create a node for each additional component. For the latter, we keep the largest one and discard the rest. If a node/edge mask is empty, we do not add the node/edge.

2. Additional details on the user study

This section provides a complete description of the user study and a feedback from the participating architects. The realism scores are presented in the main paper as the average over the architects an the amateurs. Figures 1 and 2 show the scores separately.

2.1. Complete description

Figures 3 and 4 show screenshots of our user study for realism on pixelwise segmentation masks and vector-floorplans, respectively. A subject is presented 1) a legend containing the room types and their associated colors, 2) a set of ground-truth house layouts as reference and 3) a pair of floorplans for each question. A subject is asked to score 75 pairs of layouts represented as pixelwise segmentation masks and 30 pairs as vector-floorplans. The segmentation masks are sampled from the three competing methods, ours (i.e. Ours$^{50\%}_{heur}$ and Ours$^{100\%}_{static}$) or ground-truth. The vector-floorplans are sampled from House-GAN, ours (i.e. Ours$^{100\%}_{static}$ and Ours$^{100\%}_{static*}$) or ground-truth. A pair of layout is scored as one of the three possible choices: "A is better" $(+1/-1)$, "B is better" $(-1/+1)$, "Similar" $(0/0)$. We enforce that each possible pair of model is selected exactly 5 times during the entire session, which takes around 20-30 minutes to be completed. We conducted the user study on 10 amateurs and 10 architects, where each pair of model (or GT) is scored 50 times by amateurs and 50 times by architects in total. The average scores for each pair of methods (or GT) are shown in figures 1 and 2 for amateurs and architects.

2.2. Architects feedback

We interviewed architects after the user study sessions for collecting more detailed feedback. Overall, architects were very impressed with the quality of our layouts and mentioned that "many of the designs, were actually ready to use" (in their words). We asked what aspects of the layouts allow them to rate the layouts. The answers are categorized into three groups.

Functional soundness: The house layouts were considered

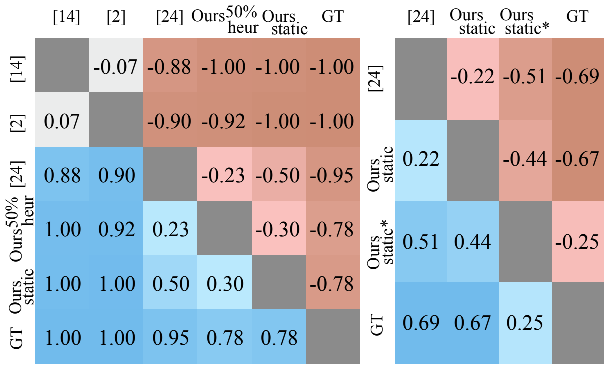

Figure 1. Realism scores for amateurs based on the user study for each pair of methods (or GT). The tables are to be read row-by-row: The bottom row shows that the GT receives positive scores against all the other methods. The left is the evaluation with raw segmentation masks, and the right is with the vector-floorplan images.

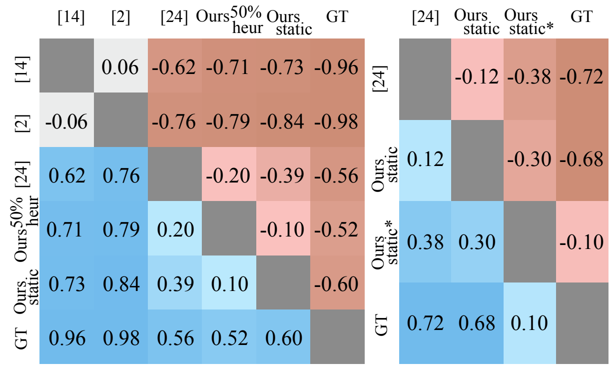

Figure 2. Realism scores for architects based on the user study for each pair of methods (or GT). The left is the evaluation with raw segmentation masks, and the right is with the vector-floorplan images.

unrealistic if they were not functional, for instance, if rooms were not accessible through doors, frontal door was missing or multiple rooms were needed to be crossed for reaching a living room or kitchen. Architects also noted that a layout should not contain tiny rooms and should contain essential rooms such as bathrooms.

Semantic validity of door types: Inconsistencies in the door types were also pointed as an important factor for evaluating layouts, meaning the front door should not be connecting two rooms nor an internal door should be connecting to the outside area.

Plausibility of door placements: House layouts were also punished when containing implausible door locations such as doors floating in the space or intersecting walls.

As future suggestions, including more layout elements such as windows could further help to evaluate the quality of the layouts. In addition, considering the site boundary, neighbourhood conditions could potentially impact the design of house layouts, for instance, designing layouts for smaller sites may impose further challenges in generating diverse layouts.

3. Details on vectorization using Floor-SP

We performed floorplan vectorization by applying Floor-SP, which was originally designed for receiving RGBD scans as input. We use the original code for Floor-SP with a few modifications: 1) we extract corners, edges and regions masks from generated pixelwise segmentation masks by a model, instead of point-density/normal maps in top-down view, 2) we invoke Manhattan-world assumption on the output vector-floorplans and 3) we apply a 3×3 Gaussian ($\sigma = 1$) blur on the Floor-SP edge masks to allow corners and edges to move by a narrow margin.

Table 1. Ablation study on the input graph information. We first drop room type information ("Types") by setting all room types to be the same, then the connectivity information ("Conn.") by making the graph fully-connected. The results are computed on floorplans with 8 rooms and the methods were trained on floorplans with 5, 6 and 7 rooms. At test time, we run the $50\%_{heur}$ scheme five times and report the average and standard deviation. "Divers." and "Compat." indicate FID and the graph edit distance metrics.

	✗ Types ✗ Conn.	✓ Types ✗ Conn.	✓ Types ✓ Conn.
Divers. (↓)	48.9±2.0	46.2±8.0	32.9±4.9
Compat. (↓)	8.2±0.1	8.2±0.1	3.9±0.5

4. Additional discussion on rule based methods and windows generation

Traditional rule based methods (RBMs) exhibit a few critical drawbacks against data-driven approaches: 1) RBMs need to hand-design rules depending on the input architectural styles or culture; and 2) RBMs takes minutes in optimizing a single layout, making it impractical for interactive design. On discussing the system design with professional architects, we learned that window placements highly depend on the environmental factors (e.g., daylighting and neighboring buildings). Proper evaluation would be difficult, without such information in the current public datasets. Therefore, this paper focuses on generating rooms and doors, while the window generation is an interesting future work.

5. Additional results

Table 1 shows additional ablation study on the input graph information. Figures 5-12 present additional generated layout samples by our system. Each row shows an input bubble diagram followed by eight generated house layouts. The samples are divided into four groups depending on the numbers of rooms and two pages of results are presented for each group, in ascending order of room numbers.

Figure 3. Screenshot of our user study for the realism evaluation on pixelwise segmentation masks. The legend appears on the top, followed by a set of ground-truth samples in the middle, and a pair of generated sample at the bottom for each question.

Figure 4. Screenshot of our user study for the realism evaluation on vector-floorplans. The legend appears on the top, followed by a set of ground-truth samples in the middle, and a pair of generated sample at the bottom for each question.

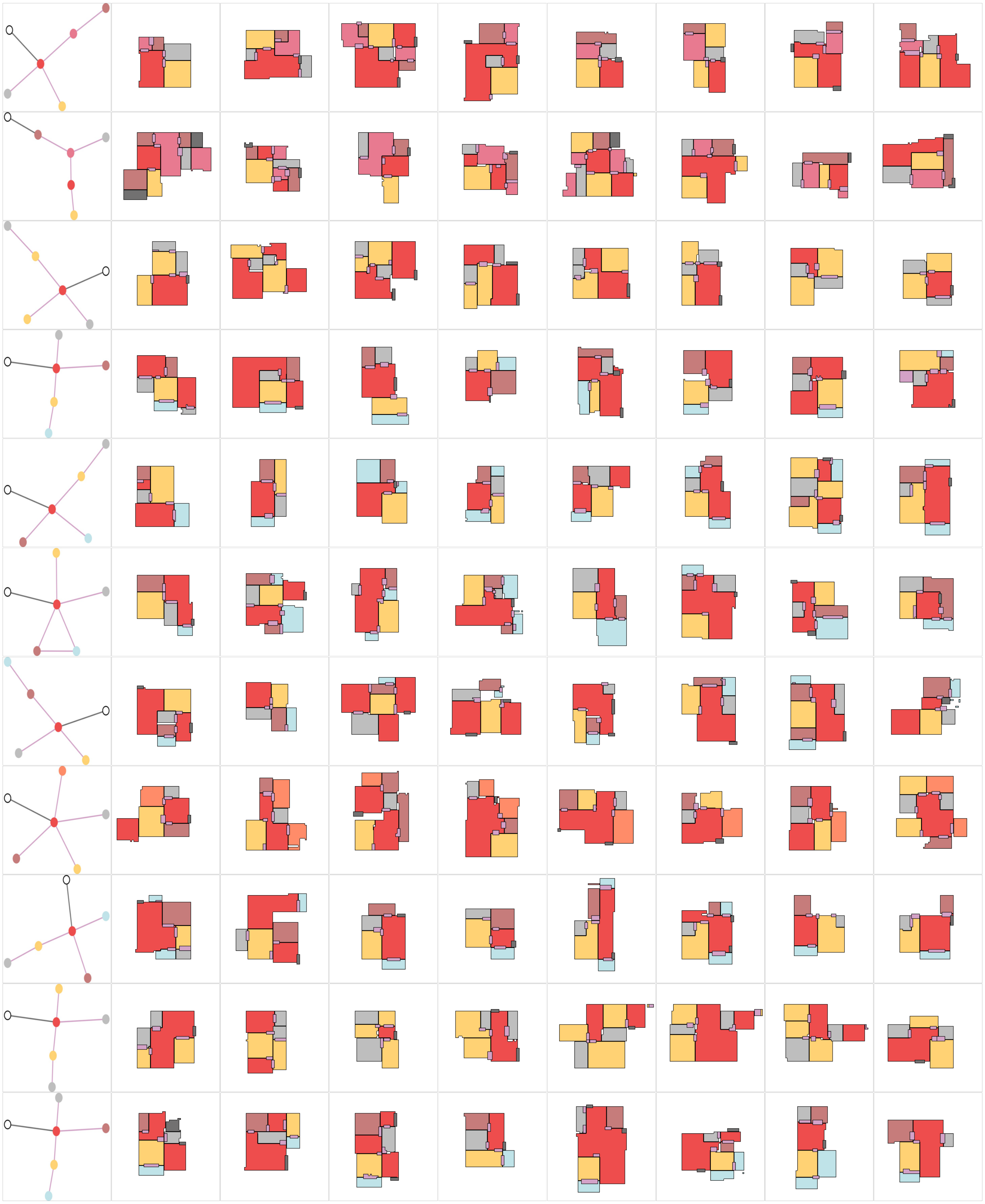

Figure 5. Additional qualitative results. Each row shows eight generated layouts for the same graph presented in the firs column. Model was trained on graphs with (6, 7, 8) rooms and tested on graphs with 5 rooms.

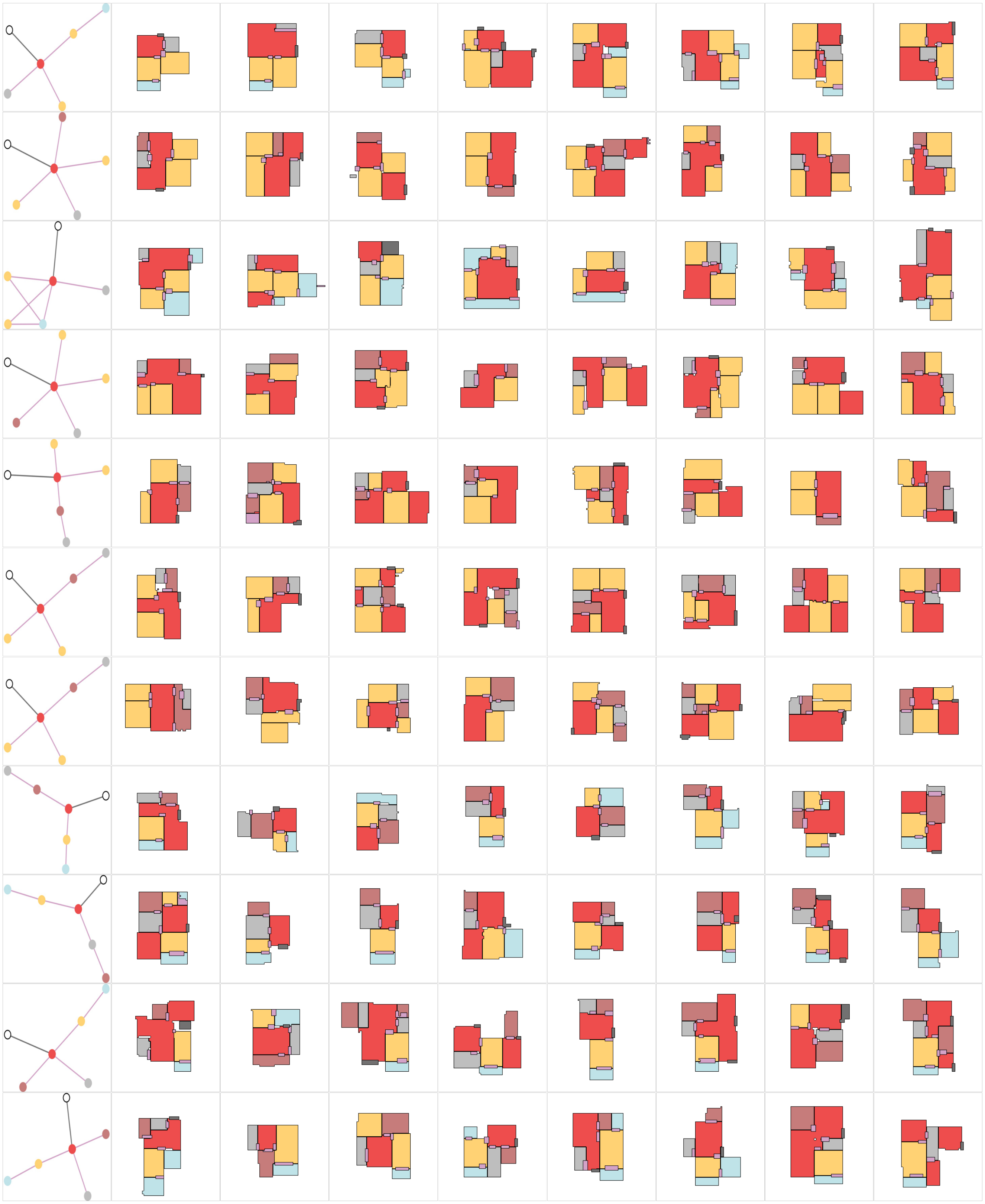

Figure 6. Continued.

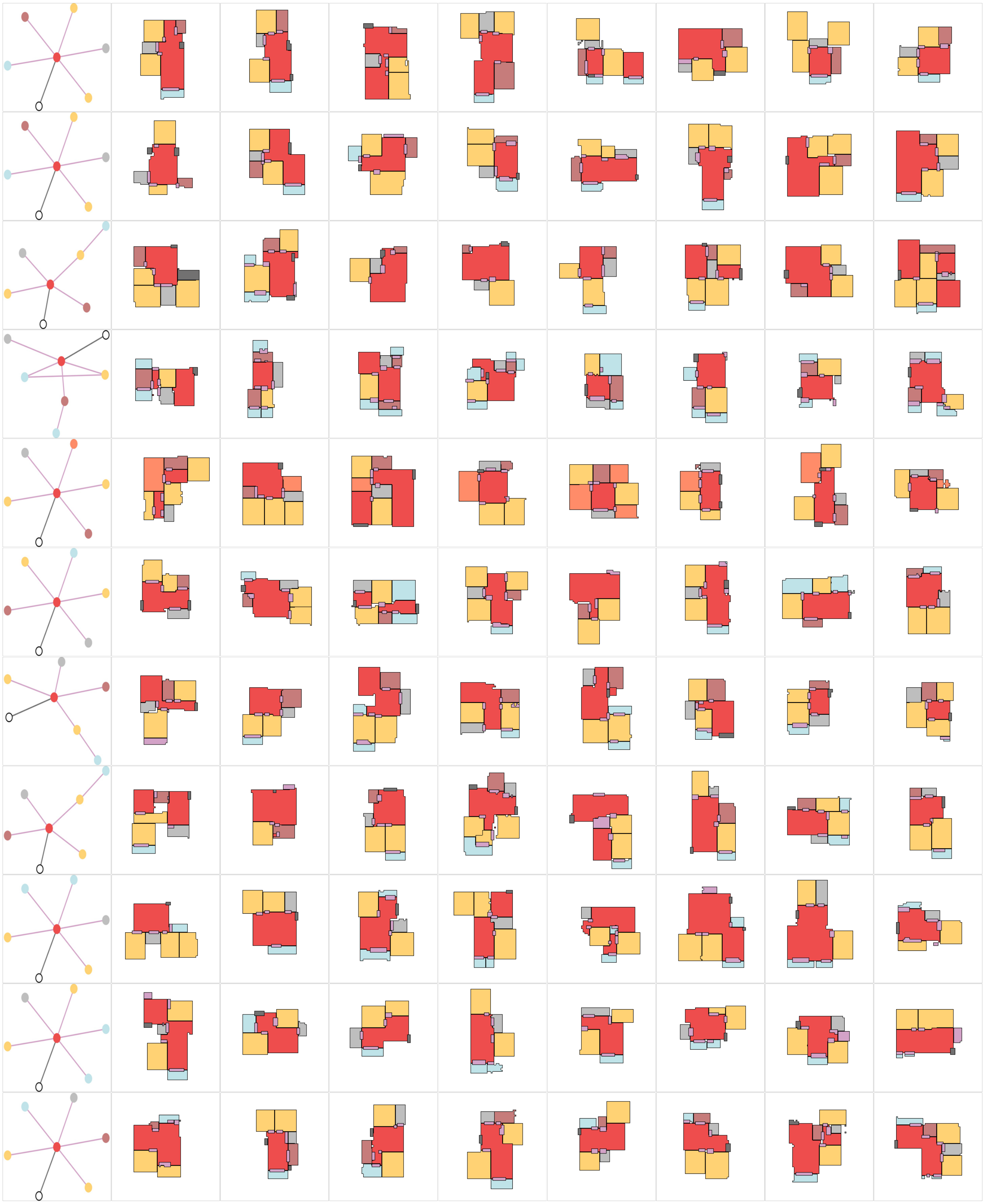

Figure 7. Additional qualitative results. Each row shows eight generated layouts for the same graph presented in the firs column. Model was trained on graphs with (5, 7, 8) rooms and tested on graphs with 6 rooms.

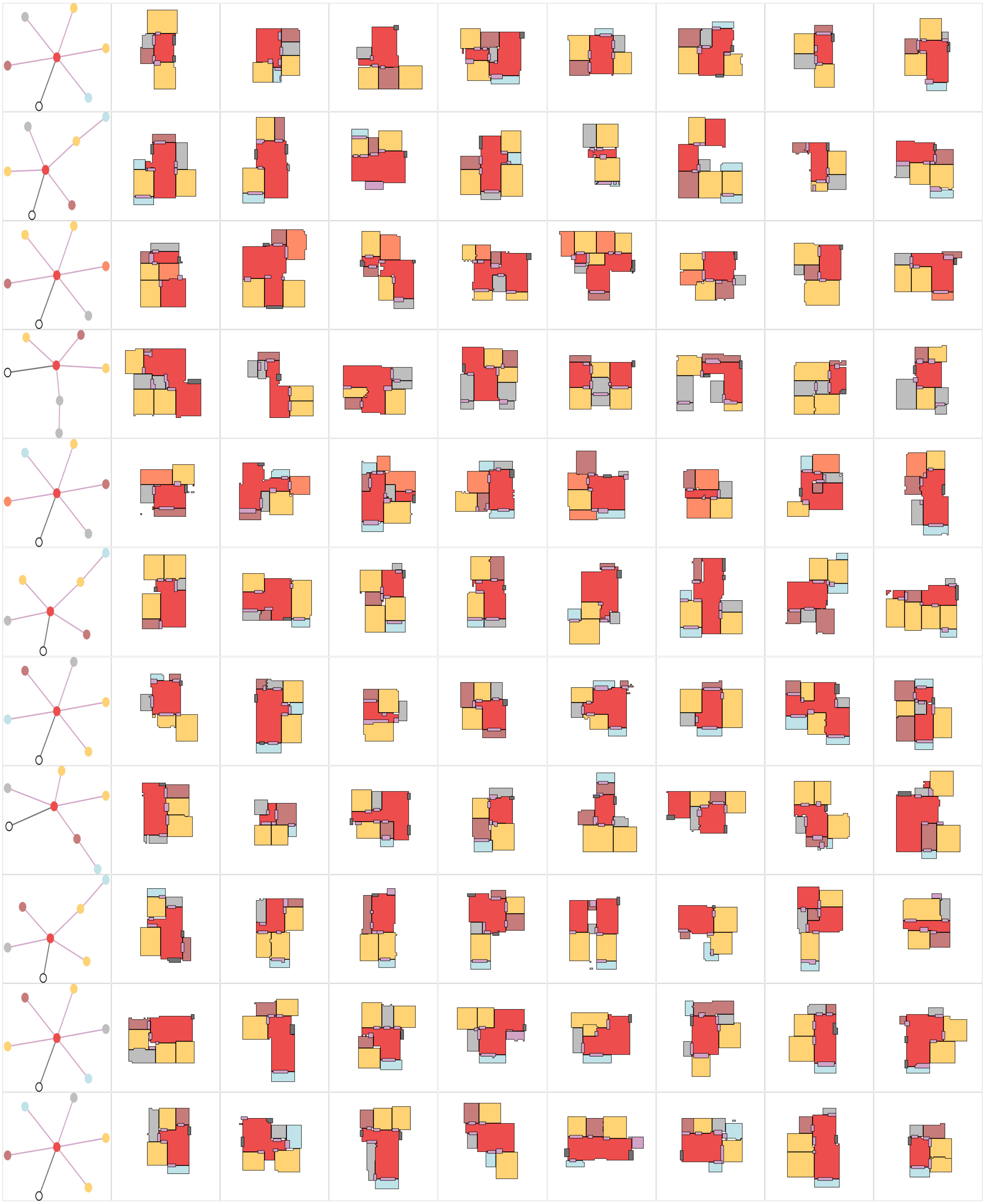

Figure 8. Continued.

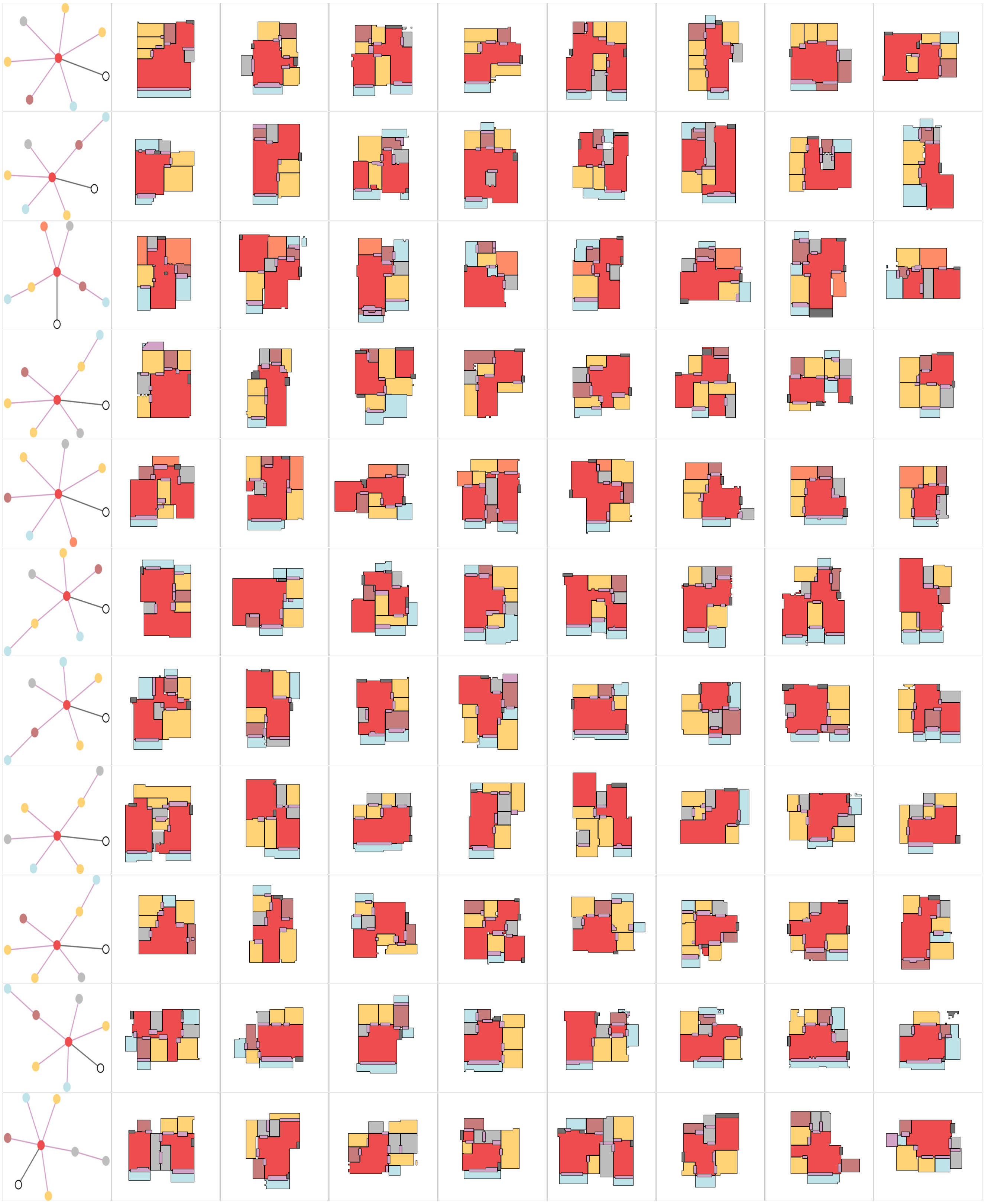

Figure 9. Additional qualitative results. Each row shows eight generated layouts for the same graph presented in the firs column. Model was trained on graphs with (5, 6, 8) rooms and tested on graphs with 7 rooms.

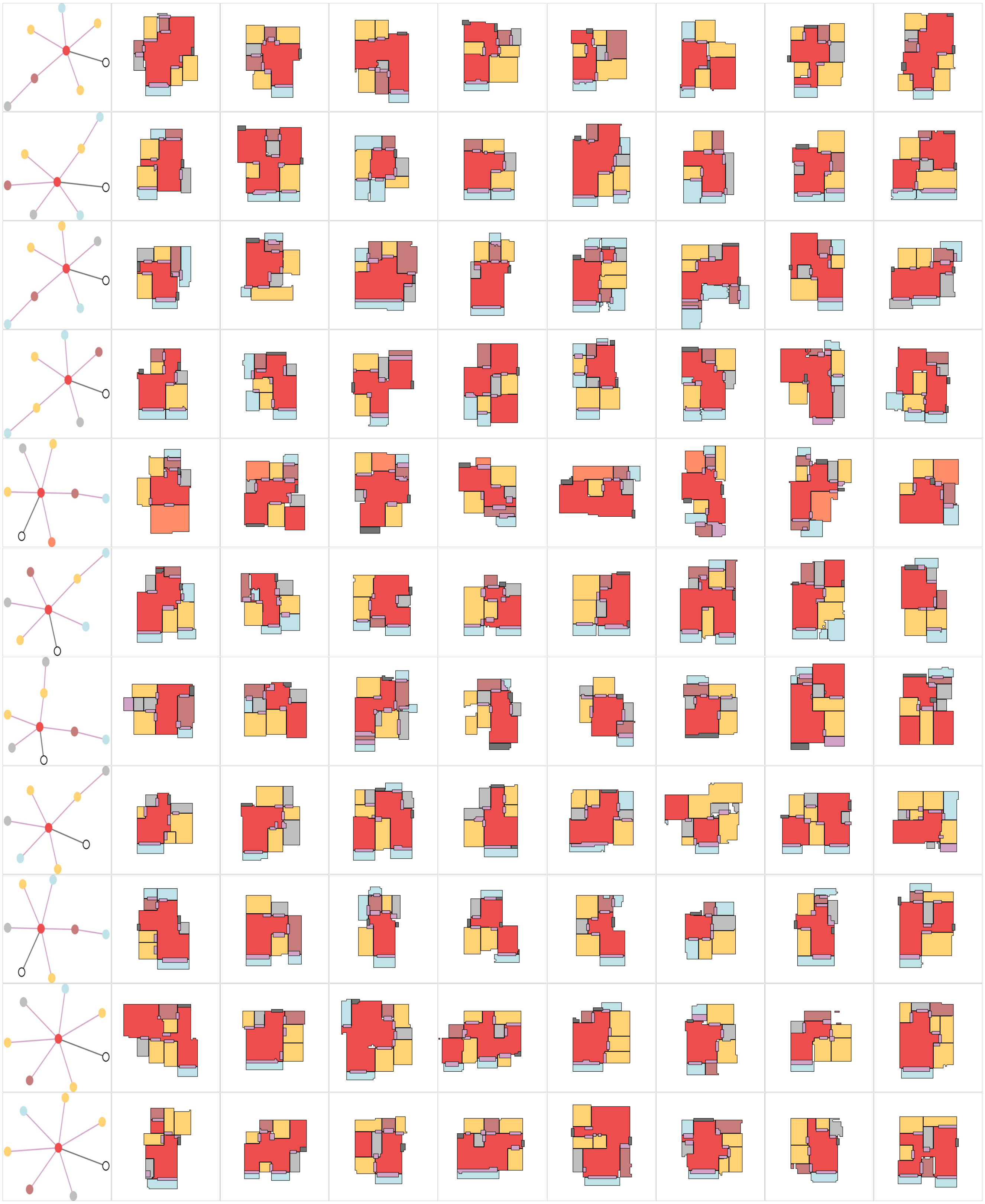

Figure 10. Continued.

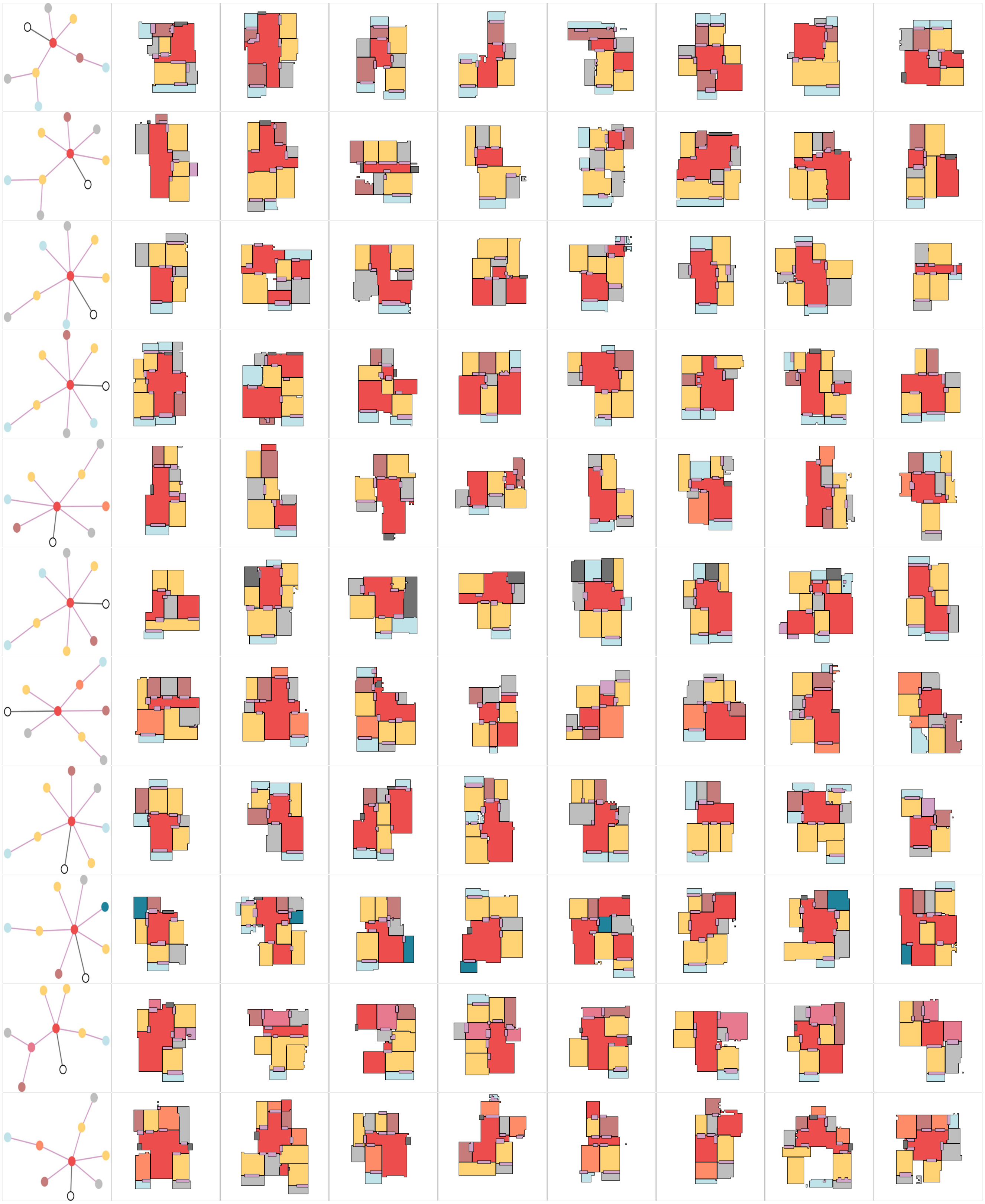

Figure 11. Additional qualitative results. Each row shows eight generated layouts for the same graph presented in the firs column. Model was trained on graphs with (5, 6, 7) rooms and tested on graphs with 8 rooms.

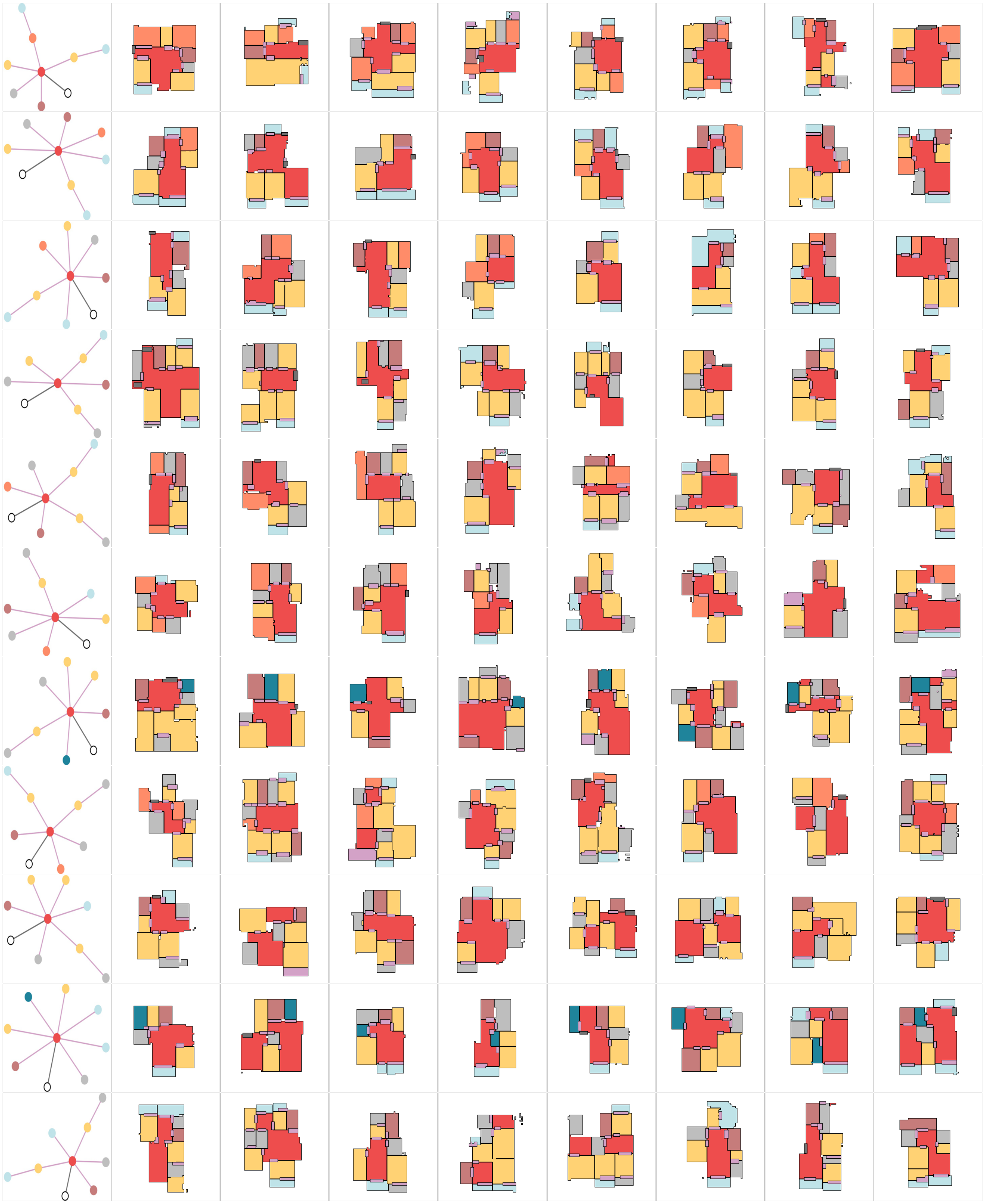

Figure 12. Continued.