Pacific Graphics 2025
M. Christie, N. Pietroni, and Y.-S. Wang
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 44 (2025), Number 7

Supplemental Material of LayoutRectifier: An Optimization-based Post-processing for Graphic Design Layout Generation

I-Chao Shen[1] Ariel Shamir[2] Takeo Igarashi[1]

[1] The University of Tokyo, Japan
[2] Reichman University, Israel

1. Method Details

1.1. Layout-specific criteria

In this section, we list the layout-specific design criteria ϕ for all three datasets used in our paper.

PubLayNet dataset - document
- child set (ϕ_{child}): $\emptyset$
- parent set (ϕ_{parent}): $\emptyset$
- other set (ϕ_{others}): {“text”, “title”, “list”, “table”, “figure”}.

Magazine dataset - magazine
- child set (ϕ_{child}): {“text-over-image”, “headline-over-image”}.
- parent set (ϕ_{parent}): {“image”}.
- other set (ϕ_{others}): {“text”, “headline”}.

CGL dataset - poster
- child set (ϕ_{child}): {“embellishment”, “logo”, “text”}.
- parent set (ϕ_{parent}): {“underlay”}.
- other set (ϕ_{others}): $\emptyset$

2. All Alignment Categories

In Figure 1, we show the illustrations of all alignment categories used in our method. Meanwhile, in Figure 2, we illustrate how these alignment categories are detected. If two boxes are both left and right aligned, they will be categorized as left-right alignment. Similarly, if two boxes are both top and bottom aligned, they will be categorized as top-bottom alignment.

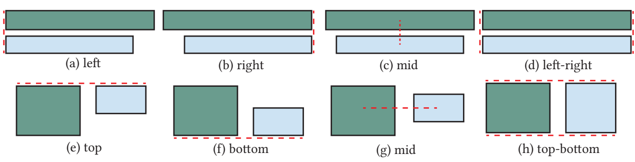

Figure 1: *We extract eight different alignment categories between pair of element boxes.*

3. Search-and-snap

3.1. All Snap Options

In Figure 3, we illustrate all possible snap options used in our method. In Algorithm 1, we describe the process of our search-and-snap operation.

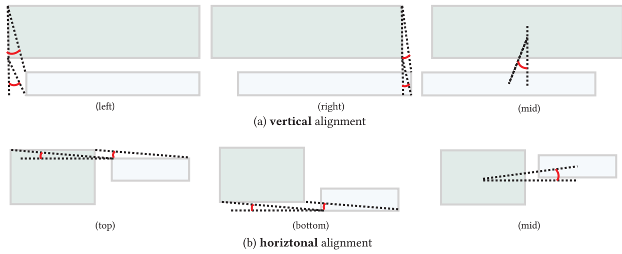

Figure 2: *The criteria we used to detect (a) vertical and (b) horizontal alignments.*

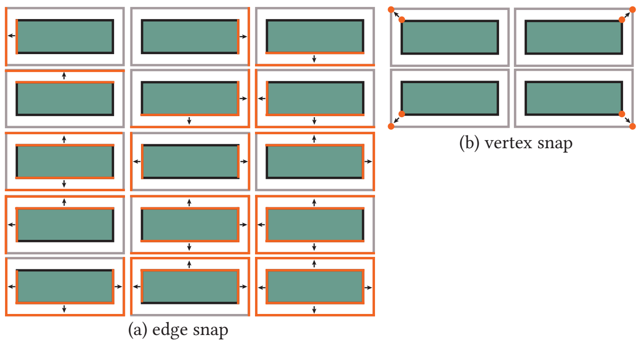

Figure 3: *Our method allows both (a) edge snap and (b) vertex snap. (a) In edge snap, the highlighted edge of the element box will snap to the highlighted edge of the grid. (b) Similarly, in vertex snap, the highlighted vertex of the element box will snap to the highlighted vertex of the grid.*

3.2. Additional Energy Functions

Box distance term ($\mathcal{E}_{\text{dist}}$) This term aims to enforce each box do not deviate from its original location too much. Specifically, for each element box $b_i = [x_i, y_i, w_i, h_i]$, where (x_i, y_i) represents its center location, the box distance term is defined as:

$$\mathcal{E}_{\text{dist}} = \sum_{i=1}^{N} ((\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2)^2, \tag{1}$$

Algorithm 1 Search-and-snap

Input: current layout L, a example grid system G^e, all snap options S.

Output: refined layout $\tilde{L}$.

```
for (c, b) ∈ L do
    bestSnapIndex ← −1
    bestScore ← 10000
    for s ∈ S do
        b̄ ← snap(b, s)
        score(b̄) ← E_all(L \ b ∪ b̄)
        if score(b̄) < bestScore then
            bestScore ← score(b̄)
            bestScoreIndex ← s
        end if
    end for
    ▷ Update layout set L using the snapped box.
    L ← L \ b ∪ snap(b, bestSnapIndex)
end for
```

where $(\hat{x}_i, \hat{y}_i)$ represents the center location after this optimization step.

Occlusion term ($\mathcal{E}_{\text{occ}}$) This term computes the average saliency value in the overlapping region between the saliency map S and the layout elements. Specifically, for each element box $b_i = [x_i, y_i, w_i, h_i]$, where (x_i, y_i) represents its center location and (w_i, h_i) represents its width and height, the occlusion term is defined as:

$$\mathcal{E}_{\text{occ}} = \frac{1}{N}\sum_{i=1}^{N}\sum_{q=y_i-0.5*h_i}^{y_i+0.5*h_i}\sum_{p=x_i-0.5*w_i}^{x_i+0.5*w_i} S[p,q]. \tag{2}$$

Blank space term ($\mathcal{E}_{\text{blank}}$) This term aims to minimize the blank spaces in the resulting layout. Specifically, it is defined as:

$$\mathcal{E}_{\text{blank}} = |C| - \overbrace{\sum_{i=1}^{N}|b_i|}^{\text{Individual}} + \overbrace{\sum_{\forall(i,j)}|b_i \cap b_j|}^{\text{Overlap area}}, \tag{3}$$

where C is the minimum bounding box of all element boxes $[b_1, ..., b_N]$, and $|C|$ is the area of C.

4. Definition of Evaluation Metric

We use the alignment and overlay metrics from Li *et al.* [LYZ*20]. We include the details of the definition here only for self-containiness purpose, please refer to their paper for more details.

4.1. Alignment Metric

For each layout element, we use the alignment loss defined as:

$$L_{\text{align}} = \sum_{i=1}^{N}\min(g(\Delta x_i^L), g(\Delta x_i^C), g(\Delta x_i^R), g(\Delta y_i^T), g(\Delta y_i^C), g(\Delta y_i^B)), \tag{4}$$

where N is the total number of elements in the layout, $g(x) = -log(1-x)$, and Δx_i^* (*=L,C,R) and Δy_i^* (*=T,C,B) are computed as:

$$\Delta x_i^* = \min_{\forall j \neq i}|x_i^* - x_j^*|, \tag{5}$$

$$\Delta y_i^* = \min_{\forall j \neq i}|y_i^* - y_j^*|, \tag{6}$$

4.2. Overlap Metric

Given an input layout with N elements, the overlap score is defined as:

$$L_{\text{over}} = \sum_{i=1}^{N}\sum_{\forall j \neq i}\frac{s_i \cap s_j}{s_i}, \tag{7}$$

where s_i denotes the area of element i and $s_i \cap s_j$ denotes the overlapping area between element i and j.

5. Ablation Study

The effectiveness of two-stage optimization. We showed the full quantitative results of the ablation study on the effectiveness of the stage A in Table 1 and stage B in Table 2.

The effectiveness of preservation terms. In Figure 4, we showed comparisons between the layout generated with and without the size and aspect ratio preservation terms. Without preservation terms, many elements are altered, whereas those refined with preservation terms maintain their aspect ratio and size.

Figure 4: ***Qualitative comparison of element preserving ablation study.*** *We compared the refined layouts generated with and without the size and aspect ratio preservation terms. The size and aspect ratio of elements are better preserved in the layouts generated with the preservation terms.*

Gen \ Refine	Ove↓ Ori	Ove↓ w/o Stage A	Ove↓ Ours	Align (x100)↓ Ori	Align (x100)↓ w/o stage A	Align (x100)↓ Ours
LGAN++	0.108	0.021	0.0016	0.18	0.036	0.031
BLT	1.049	0.423	0.18	0.139	0.083	0.061
LDM	0.155	0.048	0.006	0.12	0.043	0.025
LF++	0.009	0.003	0.003	0.023	0.003	0.002

(a) *PubLayNet*

Gen \ Refine	Ove↓ Ori	Ove↓ w/o stage A	Ove↓ Ours	Align (x100)↓ Ori	Align (x100)↓ w/o stage A	Align (x100)↓ Ours	Cont↑ Ori	Cont↑ w/o stage A	Cont↑ Ours
LGAN++	0.387	0.253	0.204	0.924	0.717	0.273	0.409	0.424	0.455
BLT	0.739	0.466	0.353	0.601	0.389	0.305	0.269	0.284	0.375
LDM	0.633	0.411	0.213	0.468	0.848	0.269	0.307	0.328	0.404
LF++	0.742	0.360	0.219	0.467	0.715	0.269	0.371	0.108	0.282

(b) *Magazine*

Table 1: ***Quantitative result of ablation study on stage A.***

Gen \ Refine	Ove↓ Ori	Ove↓ w/o Stage B	Ove↓ Ours	Align (x100)↓ Ori	Align (x100)↓ w/o stage B	Align (x100)↓ Ours
LGAN++	0.108	0.14	0.0016	0.18	0.025	0.031
BLT	1.049	1.03	0.18	0.139	0.057	0.061
LDM	0.155	0.136	0.006	0.12	0.023	0.025
LF++	0.009	0.006	0.003	0.023	0.002	0.002

(a) *PubLayNet*

Gen \ Refine	Ove↓ Ori	Ove↓ w/o stage B	Ove↓ Ours	Align (x100)↓ Ori	Align (x100)↓ w/o stage B	Align (x100)↓ Ours	Cont↑ Ori	Cont↑ w/o stage B	Cont↑ Ours
LGAN++	0.387	0.371	0.204	0.924	0.263	0.273	0.409	0.418	0.455
BLT	0.739	0.698	0.353	0.601	0.287	0.305	0.269	0.294	0.375
LDM	0.633	0.646	0.213	0.468	0.277	0.269	0.307	0.331	0.404
LF++	0.742	0.587	0.219	0.467	0.231	0.269	0.371	0.149	0.282

(b) *Magazine*

Table 2: ***Quantitative result of ablation study on stage B.***

6. User study details

6.1. User Study Interface

We conducted our crowdsourced user study through an web interface illustrated in Figure 5. In the beginning of the study, we provide an explanation on the common flaws occurs in the layouts. And we describe the criteria of the two aspects we asked the participants, i.e., visual appealing and similarity to the input layout. Regarding visual appealing, we instruct participants to choose the layout with least flaws. Regarding similarity to the input layout, we ask participants to consider the following aspects:

- layout element relationships, e.g., relative positions.
- layout element aspect ratio and area.

For each worker, we duplicated three randomly selected comparison for consistency check.

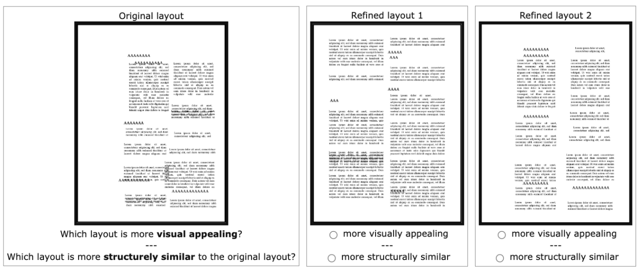

Figure 5: ***User interface of our user study.***

6.2. User Evaluation Examples

We showed all document layouts used in the user evaluation in Figure 6 and all magazine layouts in Figure 7.

7. Additional Discussion

7.1. Integrate with other 2D arrangement cost functions.

Our method effectively rectifies layout by combining the grid system with our box containment cost function. Meanwhile, similar promising functions such as Minkowski penalty [MENC24] serve similar purposes as our box containment function. Thus, we conducted a comparison by replacing the overlay term in the original objective function with it. As shown in Figure 8, while the Minkowski penalty reduces unwanted overlaps, it does not effectively remove misalignment and maintain similarity to the input layout. We plan to explore better integrations with such promising 2D arrangement cost functions in the future.

References

[LYZ*20] LI J., YANG J., ZHANG J., LIU C., WANG C., XU T.: Attribute-conditioned layout gan for automatic graphic design. *IEEE Transactions on Visualization and Computer Graphics 27*, 10 (2020), 4039–4048. 2

[MENC24] MINARČÍK J., ESTEP S., NI W., CRANE K.: Minkowski penalties: Robust differentiable constraint enforcement for vector graphics. In *Proc. SIGGRAPH* (New York, NY, USA, 2024), SIGGRAPH '24, Association for Computing Machinery. URL: https://doi.org/10.1145/3641519.3657495. 3

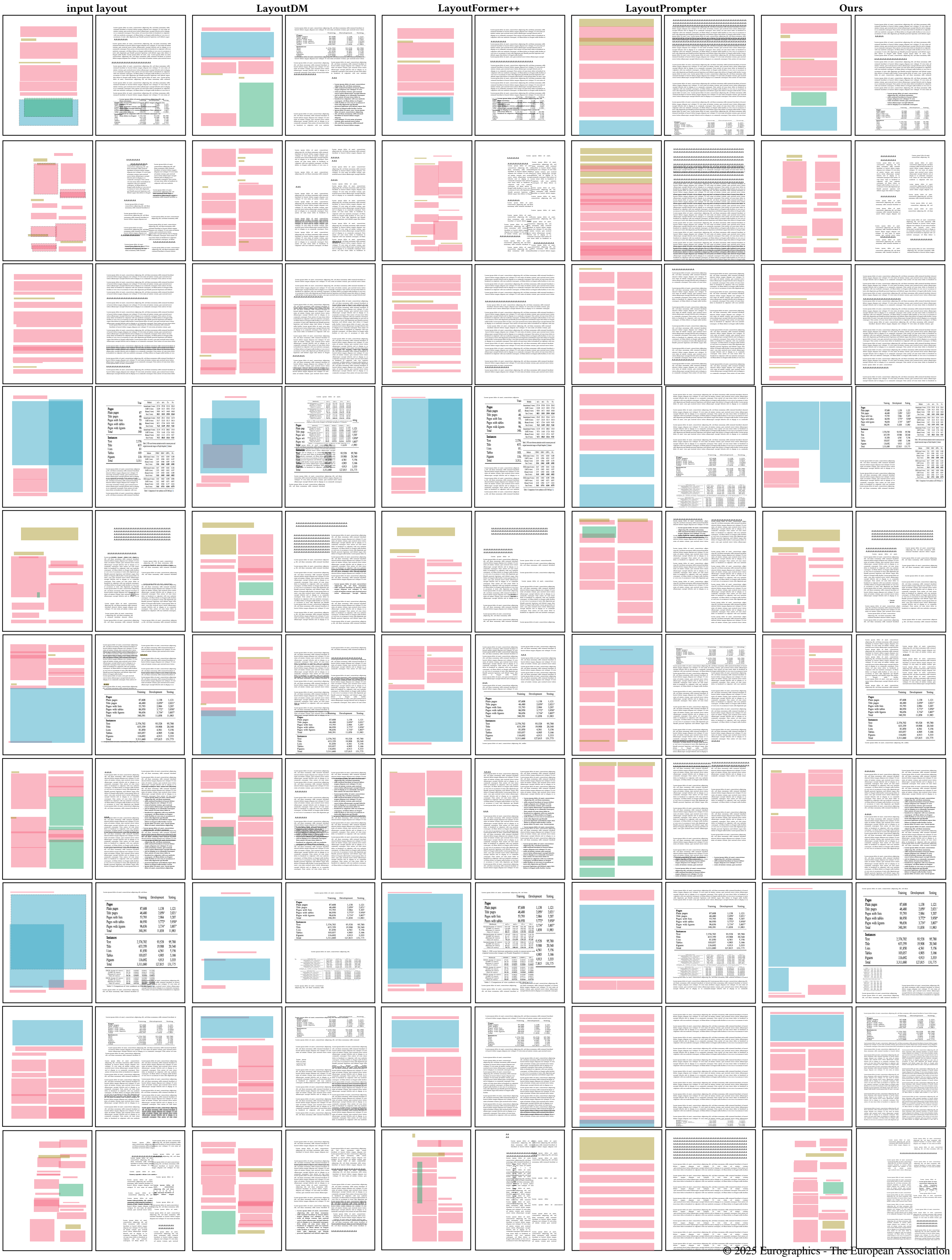

Figure 6: ***Document layouts used in user study.***

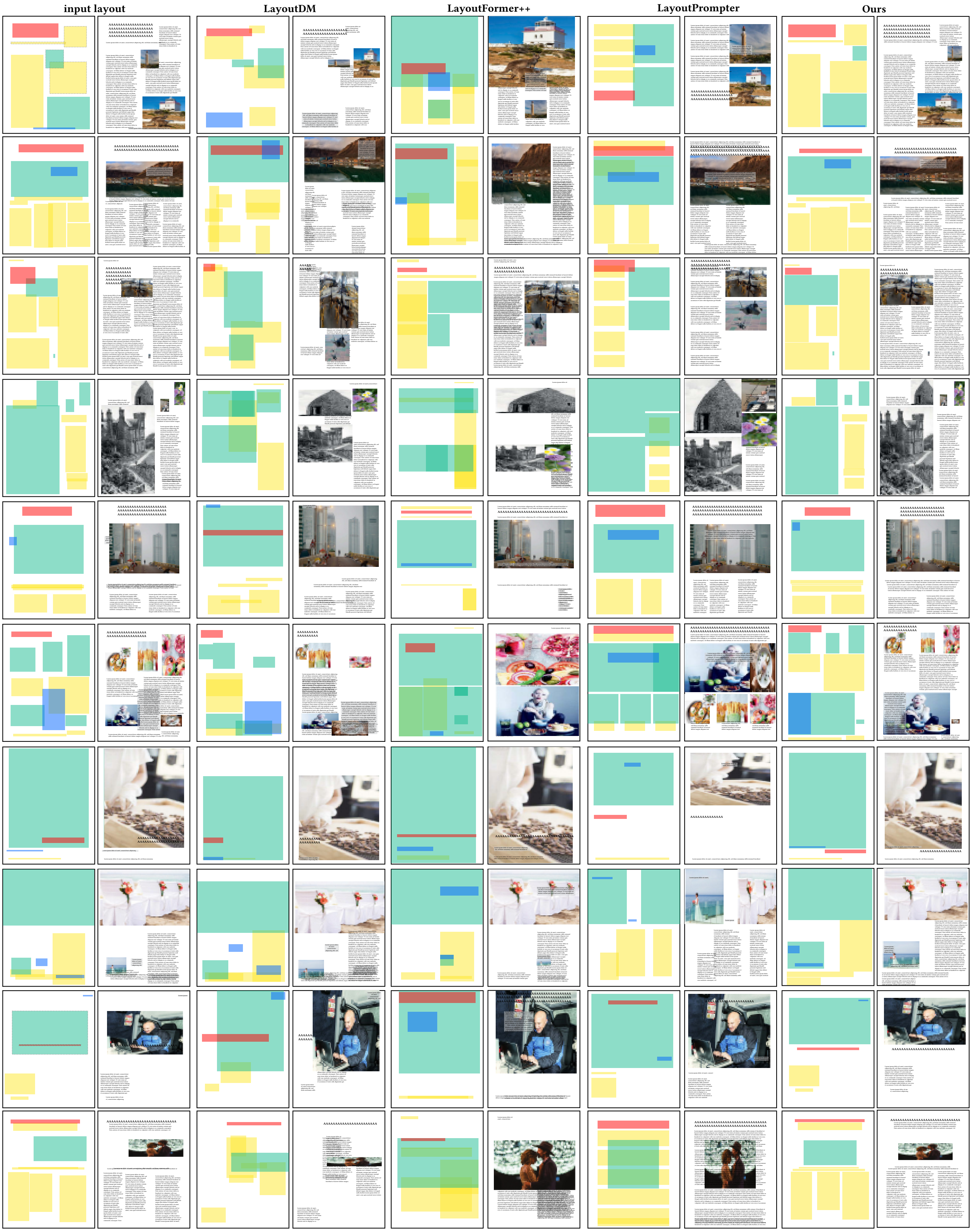

Figure 7: ***Magazine layouts used in user study.***

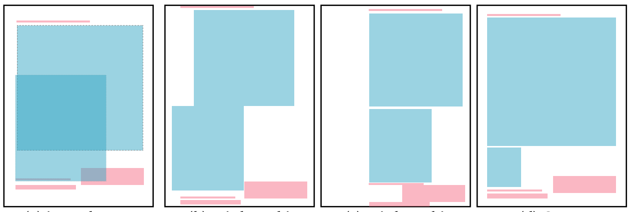

Figure 8: ***Minkowski penalty comparison.*** *Given the (a) input layout, Minkowski penalty successfully mitigate unwanted overlaps (b) with and (c) without using a reference grid but cannot resolve the misalignment. Moreover, it fails to preserve the similarity to the input layout.*