arXiv:2602.01046v1 [cs.CV] 1 Feb 2026

ReLayout：通过关系感知设计重建实现多功能且保留结构的设计布局编辑

Jiawei Lin [1] **Shizhao Sun** [2] **Danqing Huang** [2] **Ting Liu** [1] **Ji Li** [2] **Jiang Bian** [2]

摘要

无需手动调整即可自动重新设计，是设计工作流程向前迈出的重要一步。在这项工作中，我们聚焦于一项基础的重新设计任务，即设计布局编辑，其目标是根据用户意图自主修改设计的几何构成。为克服以自然语言表达的用户需求所存在的歧义，我们提出了四种基本且重要的编辑动作，并规范了编辑操作的格式。这项尚未得到充分探索的任务面临着独特挑战：既要满足指定的编辑操作，又要同时保留未编辑元素的布局结构。此外，三元组（原始设计、编辑操作、编辑后的设计）样本的匮乏也构成了另一项重大挑战。为此，我们提出了 ReLayout，这是一种无需三元组数据即可实现多功能且保留结构的设计布局编辑新框架。具体而言，ReLayout 首先引入关系图，其中包含未编辑元素之间的位置和尺寸关系，作为保持布局结构的约束。随后，我们提出关系感知设计重建（RADR），以绕过数据挑战。RADR 通过学习根据设计元素、关系图和合成的编辑操作重建设计，以自监督方式有效模拟编辑过程。RADR 以多模态大语言模型为骨干网络，在单个模型中统一多种编辑动作，从而通过微调实现多功能编辑。定性和定量结果以及用户研究表明，在编辑质量、准确性和布局结构保持方面，ReLayout 的表现显著优于基线模型。

1. 引言

平面设计是一个迭代过程，包含多次重新设计，以探索不同方案或更新设计偏好。在各种重新设计中，设计布局编辑专注于调整输入设计的布局，发挥着基础而关键的作用。例如，可以扩大文本框以突出重要信息，或重新定位产品图像以提升视觉效果（图1a）。在设计布局编辑过程中，输入设计的几何构成会根据用户需求进行调整。

然而，执行此类编辑需要考虑多个维度，对用户而言门槛较高。编辑后的设计应具备良好的视觉质量，并符合指定的编辑操作。此外，应尽可能保留未编辑元素之间的布局结构，以维持原始设计意图、阅读顺序等。图1b 展示了一些常见的不理想结果。重新定位产品图像时，如果没有相应调整周围元素，可能会出现遮挡问题，从而降低整体质量。对于未编辑的标题“FRUIT”，它在原始设计中位于画布顶部，因此最好不要在编辑后的设计中将其移到底部。

在本研究中，我们旨在自动化设计布局编辑过程，为更高效、易用的设计工作流程铺平道路。给定一个原始设计和一项元素级编辑操作（图1a），目标输出是编辑后的设计（图1c）。为提高清晰度并消除歧义，我们不使用自然语言描述编辑操作，而是采用标准化格式，明确指定编辑动作、目标元素及相应参数（例如，将元素 1 移动到 (100, 100)）。对于编辑动作，我们考虑四种最实用的操作：添加、删除、移动和调整大小。请注意，我们不包括内容相关操作（例如，更改产品图像中的物体），因为这属于图像编辑的范畴，而非布局编辑。

尽管其十分重要，自动设计布局编辑受到的关注仍然有限。目前，最相关的工作是布局生成（Jyothi 等人，2019年；Kong 等人，2022年；Li 等人，2020b年；Kikuchi 等人，2021年；Jiang 等人，2023；

[1]中国西安，西安交通大学 [2]微软研究院。通讯作者：Jiawei Lin < kylelin@stu.xjtu.edu.cn>，Shizhao Sun < shizsu@microsoft.com>。

预印本。*2026*年*2*月*3*日。

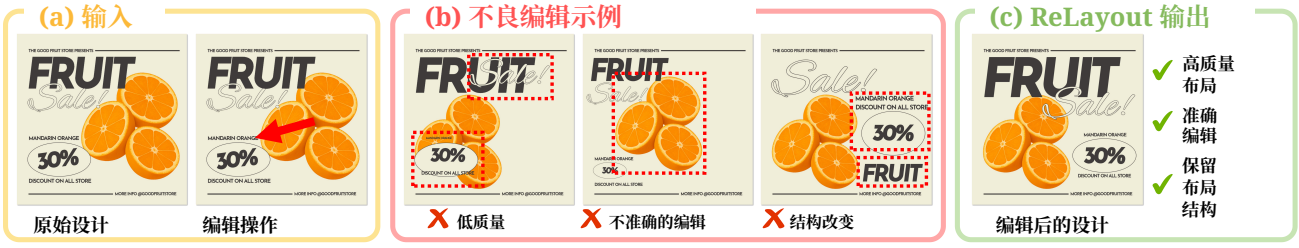


图 *1.* (a) 设计布局编辑以原始设计和编辑操作（例如，移动）作为输入。红色箭头指示移动的元素（橙色图像）及目标位置。(b) 设计布局编辑是一项具有挑战性的任务。我们展示了一些典型的不理想结果，包括布局质量低、编辑不准确以及布局结构遭到破坏。这些问题用红色虚线框标出。(c) 所提出的 ReLayout 能够实现有效的设计布局编辑，其效果体现在编辑后的设计具有较高的布局质量、准确的编辑效果以及得到良好保留的布局结构。

Inoue 等人，2023b；Lin等人，2023）以及设计生成（Inoue 等人，2023a；Cheng 等人，2024；Shabani 等人，2024；Lin等人，2024）。它们将布局/设计表示为一组元素属性（例如 x、y、宽度、高度、颜色、字体等），并探索用于可控生成的各种条件。然而，它们都无法以现有设计作为输入条件来生成编辑后的版本。另一类相关工作是图像编辑方法（Brooks 等人，2023；Kawar 等人，2023；Shi 等人，2024；Zhang 等人，2023；2025），这些方法随着文生图扩散模型的发展取得了显著进展（Podell 等人，2023；Esser 等人，2024；Peebles 和 Xie，2023；Labs，2024）。然而，将这些像素级编辑模型直接应用于设计布局编辑存在问题，因为它们不可避免地会改变内容细节。此外，像素空间输出也会阻碍进一步编辑。

设计布局编辑面临着独特的挑战。首先，它需要保留未编辑元素的布局结构，即它们的相对尺寸和位置关系。这些关系传达了设计意图中的阅读顺序、视觉焦点和层级，因此编辑后应尽可能予以保留。其次，编辑数据（原始设计、编辑操作、编辑后的设计）的缺乏从根本上阻碍了这项任务的发展。尽管一些数据集（Deka 等人，2017；Yamaguchi，2021年）提供了高质量的平面设计，却不包含这类三元组。此外，手动构建此类样本成本极高，且难以扩展。

为此，我们提出*ReLayout*，这是一种旨在应对上述挑战并有效实现设计布局编辑的新框架。首先，我们将关系图作为 ReLayout 中明确的编辑约束。通过将所有未编辑元素和画布表示为节点，并提取它们的尺寸和位置关系作为边，关系图忠实地捕捉了原始布局结构，并可作为额外输入来控制编辑过程，从而确保布局结构保持。基于关系图，我们提出关系感知设计重建（RADR）来解决数据稀缺问题。我们的核心思路是，可以通过自监督目标从单个设计布局中学习编辑知识，从而无需三元组训练数据。具体而言，RADR 提出根据设计布局中的元素、关系图以及随机合成的编辑操作重建设计布局；这与设计布局编辑的目标相呼应，并能直接泛化到目标任务。借鉴先前方法的成功实践（Lin等人，2024；Cheng 等人，2024），我们采用多模态大语言模型作为 RADR 的骨干网络，并对其进行微调，使其能够处理多种编辑操作。最终，即使缺少显式的编辑数据集，也能得到一种灵活且保留结构的设计布局编辑模型。

我们从不同角度评估性能：编辑质量、编辑准确性和布局结构保持。由于 ReLayout 是首个针对该任务的方法，我们使用 GPT-4o（Hurst 等人，2024）建立基线，并在输出质量方面将 ReLayout 与最先进的设计生成模型进行比较。定性结果、定量结果和用户研究表明，ReLayout 在所有评估维度上均优于基线。此外，我们将 ReLayout 扩展到其他实际任务，包括语言引导的设计布局编辑和复合设计布局编辑。

2. 相关工作

布局生成。自动布局生成是计算机视觉中的一个新兴课题。现有方法开发了各种布局生成模型，涵盖无条件布局生成（Arroyo 等人，2021年；Li 等人，2020a；Jiang 等人，2022年），以及基于元素类型（Jyothi 等人，2019年；Kong 等人，2022年；Li 等人，2020b）、尺寸（Kong 等人，2022年）、关系（Kikuchi 等人，2021年；Lee 等人，2020年）或文本描述（Lin等人，2023）的条件布局生成。一些近期研究进一步在单个模型中统一了多种生成任务（Jiang 等人，2023；Tang 等人，2023；Inoue 等人，2023b）。然而，这些方法都没有探讨布局编辑任务。据我们所知

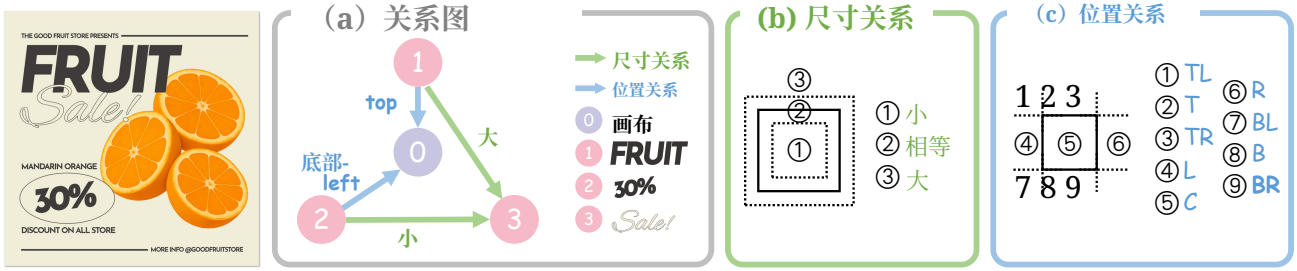


图 2。（a）关系图示例。为简化起见，这里仅可视化了部分节点和边。每条边的非箭头端节点表示源节点，箭头端节点表示目标节点。（b）和（c）用于检测尺寸和位置关系的启发式规则。实线框表示目标节点。L、R、T、B 和 C 分别表示左侧、右侧、顶部、底部和中心。

，ReLayout 是首个用于设计布局编辑的框架，能够有效减少重新设计过程中的工作量。

设计生成。 与仅生成布局（即元素位置和尺寸）的布局生成模型相比，设计生成模型能够创建完整的设计。为此，FlexDM（Inoue 等人，2023a）将每个元素表示为一组多模态字段，并采用掩码字段预测策略来填充被掩码的元素属性。VLC（Shabani 等人，2024）将图像扩散模型与矢量扩散模型相结合，用于设计生成，发挥两个领域各自的优势以提升质量。Graphist（Cheng 等人，2024）将该任务定义为序列到序列问题。它对多模态大语言模型（MLLM）进行微调，使其根据元素内容序列预测 JSON 格式的属性序列。LaDeCo（Lin等人，2024）进一步引入分层设计原则，以提升生成质量。尽管这些方法展现出良好前景，但仍无法执行设计布局编辑。

其他领域的编辑。编辑是各种内容创作领域中的一项关键任务，例如图像（Brooks 等人，2023；Kawar 等人，2023；Shi 等人，2024；Zhang 等人，2023；2025）、室内场景（Zheng 等人，2024；Yang 等人，2024b）和三维物体（Haque 等人，2023；Chen 等人，2023；Yuan 等人，2025）。与这些领域相比，设计布局编辑面临着独特的挑战。具体而言，编辑一个元素往往会引发对其他未编辑元素的调整，但其布局结构应尽可能得到保留。而其他领域的编辑任务通常更加局部化，例如为三维人脸添加胡子（Haque 等人，2023）。

3. 任务定义

令 $\mathcal{D}$ 表示一个平面设计，$\mathcal{O}$ 表示一项编辑操作。设计布局编辑可形式化地写为函数 $f:(\mathcal{D}_{\text{in}},\mathcal{O})\mapsto\mathcal{D}_{\text{out}}$，该函数以原始设计 $\mathcal{D}_{\text{in}}$ 和编辑操作 $\mathcal{O}$ 作为输入，并生成编辑后的设计 $\mathcal{D}_{\text{out}}$。f 应满足若干关键要求：（1）保持未编辑 $\mathcal{D}_{\text{in}}$中元素的布局结构，（2）准确执行指定的编辑操作 $\mathcal{O}$，以及（3）相应地输出高质量的编辑后的设计 $\mathcal{D}_{\text{out}}$。为阐明该任务，下文将详细阐述 $\mathcal{D}$ 和 $\mathcal{O}$，以解释图形设计如何构成以及编辑操作如何指定。

对于一个设计 $\mathcal{D}$，它由多个独立元素$\{E_i\}_{i=1}^{N}$ 组成（例如标题、图标），其中 N 表示元素的数量，i 表示元素索引。每个元素都有自己的内容和属性集，对于第i个元素，分别记为 C_i 和 A_i。在本文中，我们考虑两种元素模态：图像和文本。对于图像模态，内容为 RGB 像素值，属性则是用于定位其位置和尺寸的四个坐标，即中心 x 坐标、中心 y 坐标、宽度和高度。对于文本模态，内容为原始文本（例如“生日快乐”），并带有字体、颜色、文本对齐方式等其他属性。

我们将 $\mathcal{O}$ 的格式标准化为包含三个组成部分：编辑动作、该动作的目标元素，以及任何必要参数。本文考虑了四种最实用的动作：移动、调整大小、添加和删除。对于移动操作，应将新的中心 x 坐标和中心 y 坐标指定为参数。对于调整大小操作，参数为新的宽度和高度。添加和删除动作不需要额外参数。

4. ReLayout

4.1. 概述

令 $\mathcal{C}:=\{C_i\}_{i=1}^{N}$ 和 $\mathcal{A}:=\{A_i\}_{i=1}^{N}$，其中 C_i 和 A_i已在第 3 节中定义。我们有 $\mathcal{D}_{\text{in}}=(\mathcal{C},\mathcal{A}_{\text{in}})$,$\mathcal{D}_{\text{out}}=(\mathcal{C},\mathcal{A}_{\text{out}})$，其中 $\mathcal{A}_{\text{in}}$ 和 $\mathcal{A}_{\text{out}}$ 分别是输入设计布局和编辑后的设计布局的完整元素属性。$\mathcal{C}$ 是包含所有元素内容的集合，在两种设计布局中均相同。因此，我们可以将第 3 节中的函数 f 改写为 f:$(\mathcal{C},\mathcal{A}_{\text{in}},\mathcal{O})\mapsto\mathcal{A}_{\text{out}}$。然而，由于缺乏可用数据集，训练这种映射并不可行。在 ReLayout 中，我们首先根据 $\mathcal{A}_{\text{in}}$构建关系图 $\mathcal{G}$，以捕捉 $\mathcal{D}_{\text{in}}$中未编辑元素的布局结构（第 4.2 节）。接下来，提出关系感知设计重建

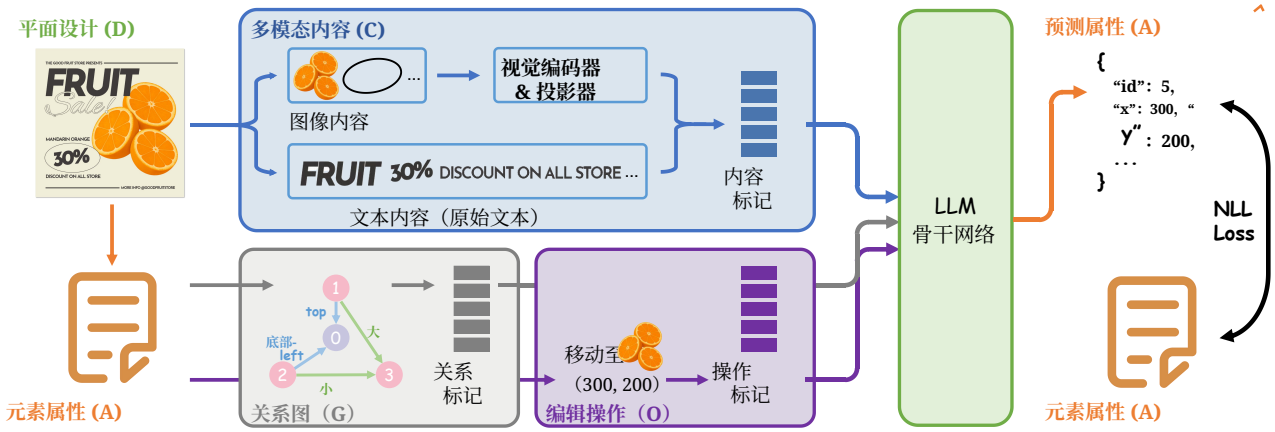


图 3. RADR 示意图。RADR 通过自监督目标实现设计布局编辑：根据元素 $\mathcal{C}$、关系图 $\mathcal{G}$和编辑操作 $\mathcal{O}$，重建设计 $\mathcal{D}(\mathcal{A})$。多模态大语言模型充当骨干网络，并对属性 $\mathcal{A}$ 使用负对数似然损失。

（RADR），它将 $\mathcal{G}$ 作为显式约束来控制编辑过程。我们使用多模态大语言模型（MLLM）实现 RADR，并对其进行微调，使其能够同时处理四种动作（第 4.3节）。由此得到的布局编辑模型兼具灵活性并能保留设计布局结构。

4.2. 关系图

设 $\mathcal{G} := (\mathcal{V}, \mathcal{R})$ 为设计 $\mathcal{D}$的关系图，其中$\mathcal{V}$ 表示节点，$\mathcal{R}$ 表示边。其节点包含 $\mathcal{D}$中的所有元素。由于画布 E_0 在定义布局结构时起着至关重要的作用，因此我们将其作为一个额外节点加入，即 $\mathcal{V} = \{E_i\}_{i=0}^{N}$（图 2（a））。每条边 $R_{ij} \in \mathcal{R}$表示源节点 E_i 与目标节点 E_j之间的位置关系或尺寸关系。我们定义了三种尺寸关系（图 2（b））和九种位置关系（图2（c））。这两类关系均通过启发式规则从 $\mathcal{A}_{\text{in}}$中识别得出。

尺寸关系。 对于 E_i 中的源节点E_j 和目标节点 $\mathcal{G}$，其尺寸关系 R_{ij}^{s} 根据面积比（AR）确定（图 2（b））：$\text{AR} = w_i h_i / w_j h_j$，其中$w$ 和 h 分别表示宽度和高度。尺寸关系分为以下几类：（1） $R_{ij}^{s} =$ 小，当 $\text{AR} < 1 - \alpha$时；（2） $R_{ij}^{s} =$ 相等，当 $1 - \alpha \le \text{AR} \le 1 + \alpha$时；（3） $R_{ij}^{s} =$ 大，当 $\text{AR} > 1 + \alpha$时。这里，α 是预定义的容差因子。值得注意的是，画布节点与其他元素节点之间不建立尺寸关系，因为画布通常远大于各个元素，因此无需定义此类关系。

位置关系。 位置关系分为两种情况 R_{ij}^{p}。首先，当源节点和目标节点都是元素时，可以根据目标节点的边界框，通过网格划分来定义它们的位置关系 3×3。具体而言，我们沿边界框的四条边延伸，将周围空间划分为九个区域。随后，根据源节点的中心点所在区域来确定位置关系（图 2(c)）。其次，当源节点是元素而目标节点是画布时，我们在画布内采用3×3 等分网格划分 来 确定元素相对于画布的位置。值得注意的是，位置关系中没有以画布为源节点的边。这是因为画布是其他元素的定位参照，仅用作目标节点。

由于元素节点之间的关系是双向的（即边 1→2和 2→1），我们随机选择其中一条作为有效边，以避免信息冗余。

4.3. 关系感知设计重建

由于 $\mathcal{G}$ 是从 $\mathcal{A}_{\text{in}}$中提取的，我们可以将f中的 $\mathcal{A}_{\text{in}}$：$(\mathcal{C}, \mathcal{A}_{\text{in}}, \mathcal{O}) \mapsto \mathcal{A}_{\text{out}}$（第 4.1 节）替换为 $\mathcal{G}$，从而得到一个近似映射 f：$(\mathcal{C}, \mathcal{G}, \mathcal{O}) \mapsto \mathcal{A}$。为简便起见，我们去掉 $\mathcal{A}_{\text{out}}$的下标。新的映射使我们能够以自监督方式训练布局编辑模型，从而克服数据问题（图 3）。由于可以直接从设计 $\mathcal{D}$中提取 f 中的 $\mathcal{C}$ˆ和 $\mathcal{A}$，且 $\mathcal{G}$已在第 4.2 节介绍，下面我们将讨论如何获得 $\mathcal{O}$。

编辑操作。 第一步是从 {添加、删除、移动、调整大小} 中采样一个动作，并从 $\mathcal{D}$ 中采样一个元素作为目标元素。其次，根据采样的动作，确定编辑操作的参数。如果动作是移动，参数就是目标元素的中心 x 坐标和中心 y 坐标。如果动作是调整大小，参数就是目标元素的宽度和高度。如果动作是添加或删除，则不需要额外参数。将采样的动作、目标元素和参数组合起来，我们就能合成一个有效的编辑操作 $\mathcal{O}$。值得注意的是，我们会删除 $\mathcal{G}$ 中与目标元素相关的边，因为编辑后它与其他节点之间的关系

方法	LLaVA-OV评分 ↑ (i)	(ii)	(iii)	(iv)	(v)	Ove ↓	Ali ↓	尺寸关系 ↑	位置关系↑	Op↑
PosterLLaVa(Yang等人,2024a)	-	-	-	-	-	0.0972	0.0013	0.8822	0.8458	-
FlexDM (Inoue 等人, 2023a)	5.34	5.29	5.41	5.09	4.54	0.3242	0.0016	-	-	-
LaDeCo(Lin等人,2024)	8.08	7.92	8.00	7.82	6.98	**0.0865**	0.0013	-	-	-
GPT-4o（Hurst 等人，2024）	7.41	7.31	7.54	7.09	6.37	0.1942	**0.0011**	0.8386	0.5544	0.9983
Ours	**8.25**	**8.10**	**8.24**	**7.87**	**7.10**	0.0996	0.0013	**0.9150**	**0.8684**	**0.9991**

表*1*。 在重建设置下与基线方法进行定量比较。每项指标的最佳结果以**粗体**标出，次佳结果加下划线。请注意，对于不考虑元素内容的基线方法（即 PosterLLaVA），LLaVA-OV评分不适用。对于忽略元素关系的基线方法（即 FlexDM 和 LaDeCo），尺寸关系和位置关系不适用。对于不考虑编辑操作的基线方法（即 PosterLLaVA、FlexDM 和 LaDeCo），操作不适用。

可能会发生变化。

序列表示。 我们为每个组件设计了序列表示方式 f。（1） $\mathcal{C}$。本文考虑两种元素模态：图像和文本。对于图像内容，我们使用视觉编码器将其转换为图像词元。文本内容天然由语言词元构成，无需额外处理。（2） $\mathcal{G}$。我们按照“节点 1，关系，节点 2”的格式描述所有边，并将这些边串联起来，序列化为一个序列。（3） $\mathcal{O}$。我们用名称指定编辑动作（例如，添加、移动）。目标元素由整数索引（例如，元素 5），所需参数四舍五入为整数。（4） $\mathcal{A}$。遵循先前的工作（Lin等人，2024），我们采用 JSON 格式表示元素属性（见图 3 ）。图像元素有四个预测属性：中心 x 坐标、中心 y 坐标、宽度和高度。对于文本元素，我们还预测三个属性：字体大小、文本对齐和角度。我们不预测文本的字体类型和颜色，因为用户通常不希望在这四种动作中修改它们。

模型架构。我们将设计布局编辑表述为序列到序列变换，并微调一个多模态大语言模型（MLLM）来实现该任务（图3）。该多模态大语言模型由三个关键组件组成：视觉编码器、投影器和大语言模型主干。视觉编码器用于对图像内容进行编码，以生成图像词元。随后，投影器将其投影到符合大语言模型主干维度要求的表示。其他输入，包括文本内容、关系图和编辑操作，均直接馈入大语言模型主干，因为它们已表示为离散词元。大语言模型主干对所有输入的联合分布进行建模，并以自回归方式生成属性。

5. 实验

5.1. 设置

数据集。我们的实验使用了公开可用的数据集 Crello（Ya-maguchi，2021年）v4。该数据集包含 23,421 个图形设计，涵盖多个领域，例如卡片、海报和 Instagram帖子。每个设计都包含一组元数据，包括画布大小、元素类型、元素位置等。文本元素还具有字体、行高、颜色等其他属性。该数据集划分为训练集、验证集和测试集，分别包含 19,095、1,951 和 2,375 个样本。我们遵循 LaDeCo（Lin等人，2024）的做法，过滤掉包含 25 个以上元素的训练样本。为展示 ReLayout 的有效性，我们在实验中引入了两种测试设置。(1) 重建。在此设置中，以第4.3节所述的相同方式，从设计中提取编辑操作。此设置用于测试直接从RADR 获得的重建能力。(2)泛化。这一设置更值得关注，也代表了 ReLayout 的目标。它用于测试使用 RADR 训练的模型能否泛化到间接设计布局编辑任务。我们模拟用户为此设置合成编辑操作。最终，两种设置的测试样本总数均为 2,375（Crello测试集大小）* 2（每个测试设计有 2 个编辑操作） = 4,750。更多详情请参阅补充材料。

实现细节。 遵循 LaDeCo（Lin等人，2024），我们采用 Llama-3.1-8B [1] 作为骨干网络，并采用 CLIP ViT-L/14 [2] 作为视觉编码器。带有 GELU（Hendrycks 和 Gimpel，2016）激活函数的两层多层感知机用作投影器，连接视觉编码器与骨干网络。我们在八张 A100-40G GPU 上采用 LoRA（Hu等人，2021年）微调来更新骨干网络和投影器（视觉编码器保持冻结），并使用 AdamW（Loshchilov 和 Hutter，2017）作为优化器。全局批大小设为 64，学习率为 2e-4，LoRA 秩为 32。

基线。 由于此前没有针对该任务的研究，我们选择 GPT-4o（Hurst 等人，2024）作为基线，因为它具备处理多模态输入的强大能力。GPT-4o 的提示词收录于补充材料中。此外，我们将 ReLayout 与两个相关类别中的现有方法进行比较：（1）设计生成。由于设计布局编辑最终会生成一个新的设计，因此我们选择了几种先进的设计生成方法，即 FlexDM（

[1] https://huggingface.co/meta-llama/Llama-3.1-8B [2] https://huggingface.co/openai/clip-vit-large-patch14-336

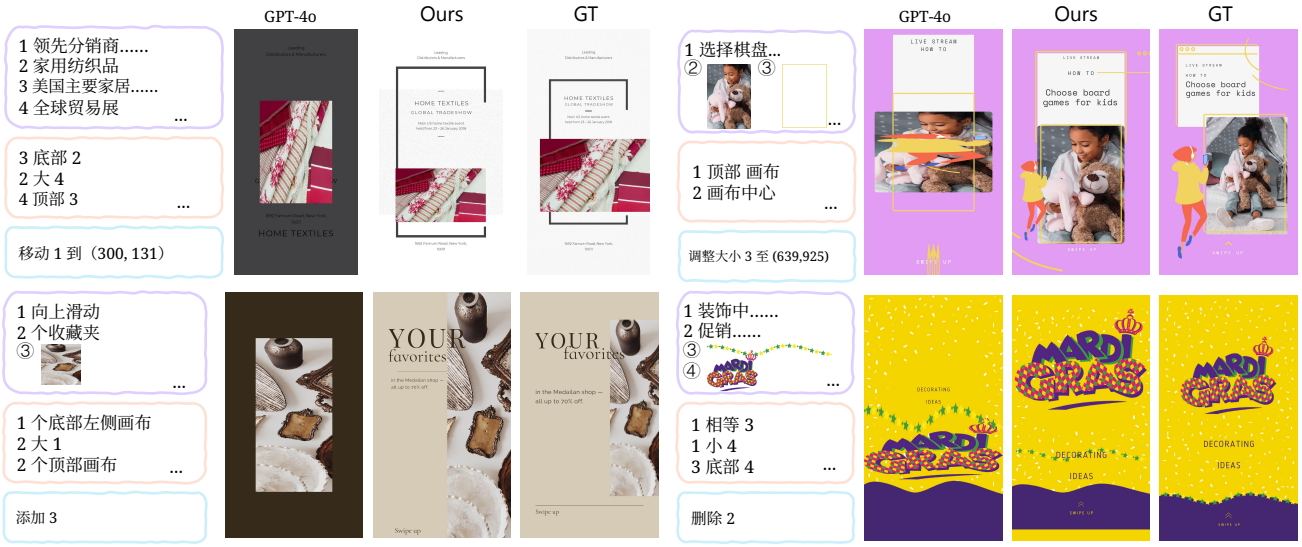


图*4*。 可视化展示了重建设计，以定性比较我们的方法与 GPT-4o。图中还展示了真实标注（GT）设计。每个案例左侧的三个框是模型输入，即 元素内容、 元素关系和 编辑操作。由于元素数量较多，我们仅可

设计布局编辑最终会生成一个新的设计，因此我们选择了几种先进的设计生成方法，即 FlexDM（Inoue 等人，2023a）和 LaDeCo（Lin等人，2024）作为基线，以比较输出质量。（2）关系条件布局生成。鉴于 ReLayout 将元素关系作为生成条件，我们还将其与 PosterLLaVA（Yang等人，2024a）进行比较。PosterLLaVA 是一种专门用于关系条件布局生成的方法，我们据此评估各方法遵循指定关系的能力。

评估指标。 定量指标涉及三个维度。(1) 质量。我们遵循先前的工作（Lin等人，2024），使用LLaVA-OV评分（Li 等人，2024）、元素重叠（Ove）和对齐（Ali）来评估设计质量。LLaVA-OV评分从五个角度评估质量：(i) 设计和布局；(ii) 内容相关性；(iii) 字体排印和颜色；(iv) 图形和图像；以及 (v) 创新性和原创性。(2) 布局结构保持。为评估编辑后的设计是否保留未编辑元素的布局结构，我们提出关系满足率（Rel）。具体而言，对于关系图中的每个关系，我们检查其在输出设计中是否得到满足：满足记为 (1)，否则记为 (0)。Rel指标计算为满足的关系数量与关系总数之比。根据关系类别，我们进一步将其细分为尺寸相关关系满足率（Size Rel）和位置相关关系满足率（Pos Rel）。(3) 编辑准确性。该维度采用编辑操作满足率（Op）。与Rel指标类似，如果编辑后的设计满足指定的编辑操作，则Op值为1，否则为0。

5.2. 定量结果

重建设置的定量结果如表 1 所示。ReLayout在几乎所有评估指标上都显著优于GPT- 4o。这表明我们的方法能够执行高质量且准确的设计布局编辑，同时有效保持未编辑元素的布局结构。此外，ReLayout也展现出相较于专用基线方法的明显优势。具体而言，ReLayout生成的设计甚至优于最先进的设计生成方法LaDeCo（Lin等人，2024）。与关系条件布局生成模型相比，ReLayout在两个Rel指标上超过了PosterLLaVA（Yang等人，2024a），这表明其在布局结构保持方面更具优势。关于泛化设置的定量结果，请参阅补充材料。

5.3. 定性结果

重建设置下的定性比较见图4。我们还展示了每个案例的模型输入和真实标注（GT）设计。就设计质量和关系满足度而言，ReLayout优于GPT-4o基线，彰显了我们模型强大的重建能力。泛化设置下的定性结果见图5、6、7和8。从结果可以看出，ReLayout能够灵活执行各种编辑操作。此外，结果进一步表明，在真实的编辑场景中，ReLayout能够生成高质量的编辑后设计，同时保留原有的布局结构。相比之下，GPT-4o要么存在严重的

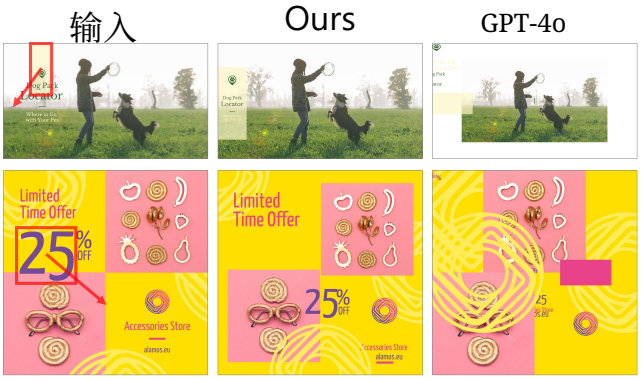


图5. 移动操作的定性比较。移动的元素以方框标记。箭头指向目标位置。

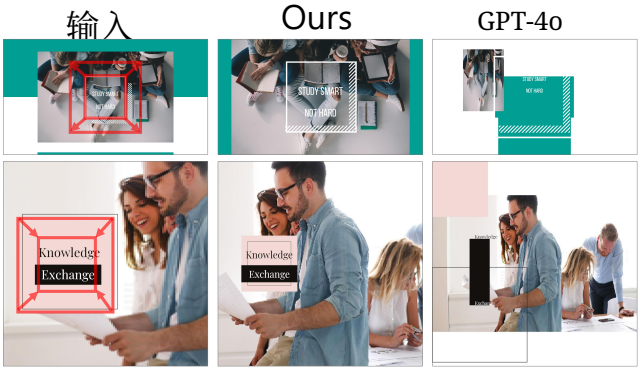


图6. 调整大小操作的定性比较。调整大小后的元素以方框标记。箭头指向目标尺寸。

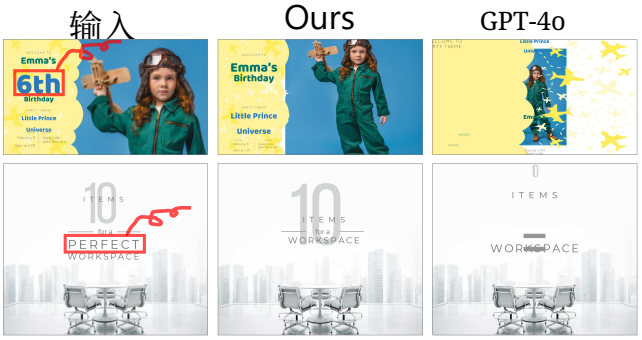


图7. 删除操作的定性比较。被删除的元素以红框标出。

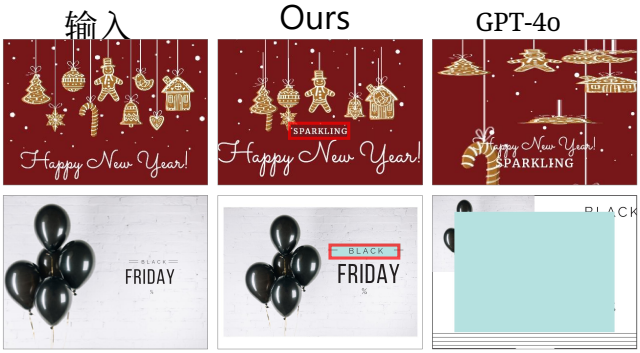


图8. 添加操作的定性比较。被添加的元素以红框标出。

质量问题（例如图8中的第二个案例、图7中的第一个案例），要么会改变布局结构（例如图6中的第一个案例、图7中的第一个案例）。此外，我们发现ReLayout能够妥善调整未指定的元素，以保持视觉协调。例如，在图5中的第二个案例里，当文本“25”移动到新位置时，百分号和文本“OFF”也会随之移动，从而保留原有语义。补充材料中提供了与更多基线方法的定性比较。

5.4. 人工评估

我们还进行了人工评估，将 ReLayout 与 GPT-4o 进行比较。为此，我们随机抽取了 200 组三元组（原始设计、ReLayout 编辑后的设计、GPT-4o 编辑后的设计），并将其均匀分配给四位众包工作者进行评估。对于每组三元组，他们需要回答两个问题：（1）两种编辑后的设计中，哪一种在视觉上更具吸引力？（2）两种编辑后的设计中，哪一种更好地保留了原始设计的布局结构？我们在两种设置下进行评估，并在表 2 中展示了各项结果所占的百分比。从结果可以看出，ReLayout 在布局结构保持和生成质量方面均显著优于 GPT-4o，进一步证明了 ReLayout 的有效性。

问题	ReLayout	Tie	GPT-4o
重建设置			
(1)	79.5	15.5	5.0
(2)	86.0	13.0	1.0
泛化设置			
(1)	71.5	21.5	7.0
(2)	86.0	10.5	3.5

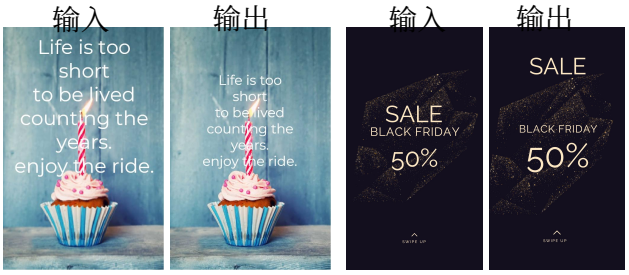
表2. 人工评估结果。问题(1)衡量输出质量，问题(2)反映布局结构的保留情况。在这两种设置下，ReLayout在这两个维度上均显著优于GPT-4o。

5.5. 消融研究

关系感知设计重建。 由于缺少三元组数据，我们将关系感知设计重建（RADR）引入 ReLayout，以实现设计布局编辑。作为一种潜在的变体，我们通过随机采样元素属性来合成原始设计（例如，在 0 到画布宽度之间均匀采样元素宽度），从而构成（元素内容、随机设计、编辑操作、良好设计）样本。其定量指标如表 3所示，记为*w/o RADR*。结果

	LLaVA-OV评分 ↑					Ove ↓	Ali ↓	尺寸关系 ↑	位置关系↑	Op↑
	(i)	(ii)	(iii)	(iv)	(v)					
不使用 RADR	8.12	8.03	8.18	7.82	7.01	0.1007	0.0011	0.8008	0.3750	0.9978
不使用关系图	8.21	8.08	8.25	7.85	7.07	0.0980	0.0013	0.7943	0.3751	0.9987
使用矩阵	8.17	8.07	8.20	7.84	7.04	0.0954	0.0012	0.8892	0.8529	0.9909
Ours	8.25	8.10	8.24	7.87	7.10	0.0996	0.0013	0.9150	0.8684	0.9991

表3. 消融研究的定量结果。所有实验均在重建设置下进行并评估。



编辑指令：将文本“生活太……”缩小

编辑指令：移动文本“促销”至顶部

图9. ReLayout 还可以通过集成 GPT-4o，根据文本指令执行设计布局编辑。



两项添加：添加“欢迎”，添加“日期：12月25日”

移动和调整大小：移动到 (538, 1554)，调整大小为 (775, 775)

图10. ReLayout 可以同时执行多个编辑操作，从而增强其实用性。

表明两个 Rel 指标均显著下降，这说明该训练策略在维持布局结构方面不如我们的 RADR。

关系图。 在本研究中，我们使用关系图表示布局结构，并指导编辑过程，以保持结构。为探究其作用，我们将其移除，从而得到用于模型训练的（元素内容、编辑操作、编辑后的设计）样本。结果如表 3 不使用关系图所示。同样，两个 Rel 指标也有所下降，这表明没有关系图时，模型无法很好地保持布局结构。

图表示。 除了采用序列表示外，我们还使用邻接矩阵来表示关系图（在表中记为 使用矩阵，见表3），并研究不同表示方法对模型性能的影响。从结果可以看出，序列表示的 Rel 指标略优于矩阵表示，这表明紧凑的序列格式更适合表示关系图。此外，两种表示方法的性能相近，也凸显了 ReLayout 的稳健性。

5.6. 应用

语言引导的设计布局编辑。对于自然语言编辑指令，我们利用 GPT-4o（Hurst 等人，2024）将其转换为相应的编辑操作。在转换过程中，原始设计也用作输入。随后，ReLayout 执行转换后的操作，实现语言引导的设计布局编辑。

图9 展示了一些结果。这显著提升了 ReLayout 的易用性，因为用自然语言表达编辑意图更加直观、友好。

复合设计布局编辑。 尽管 ReLayout 是基于单个操作训练的，但在测试时可以轻松扩展以支持复合操作（见图 10）。这种灵活的组合使 ReLayout 能够处理训练期间未曾遇到的复杂编辑需求。

6. 结论

在这项工作中，我们提出了 ReLayout，以解决尚未得到充分探索的设计布局编辑任务。通过引入用于结构表示的关系图，以及一种新颖的关系感知设计重建策略，我们无需三元组数据即可实现灵活且保留结构的布局设计。大量定量和定性实验、用户研究以及支持的应用均证明了该模型的有效性。尽管取得了这些进展，仍存在一些局限。具体而言，随着输入设计中的元素数量增加，在编辑过程中保留整体布局结构并确保设计质量会变得越来越困难。典型的失败案例包括：（1）违反指定的关系；（2）意料之外的重叠等质量问题。未来的工作将扩展系统功能，涵盖字体、颜色和内容编辑，并开发更完善的语言引导界面。我们相信，ReLayout 是迈向更普及、更智能的设计工作流的重要基础。

参考文献

Arroyo, D. M., Postels, J., and Tombari, F. 用于布局生成的变分Transformer网络。载于*IEEE/CVF*计算机视觉与模式识别会议（*CVPR*）论文集，第 13642–13652 页，2021年6月。

Brooks, T., Holynski, A., and Efros, A. A. Instructpix2pix：学习遵循图像编辑指令。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 18392–18402 页，2023。

Chen, Y., Chen, Z., Zhang, C., Wang, F., Yang, X., Wang, Y., Cai, Z., Yang, L., Liu, H., and Lin, G. GaussianEditor：使用高斯泼溅进行快速且可控的三维编辑，2023。

Cheng, Y., Zhang, Z., Yang,M., Nie, H., Li, C., Wu,X., and Shao, J. 使用大型多模态模型进行平面设计。*arXiv* 预印本 *arXiv:2404.14368*，2024。

Deka, B., Huang, Z., Franzen, C., Hibschman, J., Afergan, D., Li, Y., Nichols, J. 与 Kumar, R. Rico：用于构建数据驱动设计应用的移动应用数据集。见第*30*届*ACM*用户界面软件与技术年度研讨会论文集，第845–854页，2017。

Esser, P., Kulal, S., Blattmann, A., Entezari, R., M¨uller, J., Saini, H., Levi, Y., Lorenz, D., Sauer, A., Boesel, F. 等。面向高分辨率图像合成的整流流Transformer规模化。见第四十一届国际机器学习会议，2024。

Haque, A., Tancik, M., Efros, A. A., Holynski, A. 与 Kanazawa, A. Instruct-NeRF2NeRF：基于指令编辑三维场景。见 *IEEE/CVF*国际计算机视觉会议（*ICCV*）论文集，第19740–19750页，2023年10月。

Hendrycks, D. 和 Gimpel, K. 高斯误差线性单元（GELU）。*arXiv* 预印本 *arXiv:1606.08415*，2016年。

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L. 和 Chen, W. LoRA：大语言模型的低秩适配。*arXiv* 预印本 *arXiv:2106.09685*，2021年。

Hurst, A., Lerer, A., Goucher, A. P., Perelman, A., Ramesh, A., Clark, A., Ostrow, A., Welihinda, A., Hayes, A., Radford, A. 等。GPT-4o 系统卡。*arXiv* 预印本 *arXiv:2410.21276*，2024。

Inoue, N., Kikuchi, K., Simo-Serra, E., Otani, M. 和 Yamaguchi, K. 迈向灵活的多模态文档模型。载于 *IEEE/CVF*计算机视觉与模式识别会议（*CVPR*）论文集，第 14287–14296 页，2023a。

Inoue, N., Kikuchi, K., Simo-Serra, E., Otani, M. 和 Yamaguchi, K. LayoutDM：用于可控布局生成的离散扩散模型，2023b。网址 https://arxiv.org/abs/2303.08137。Jiang, Z., Sun, S., Zhu, J., Lou, J.-G. 和 Zhang, D. 图形布局的由粗到细生成建模。载于 *AAAI*人工智能会议论文集，第 36 卷，第 1096–1103 页，2022年。

Jiang, Z., Guo, J., Sun, S., Deng, H., Wu, Z., Mijovic, V., Yang, Z. J., Lou, J.-G., 和 Zhang, D. Layoutformer++：通过约束序列化和解码空间限制实现条件图形布局生成，2023。网址https://arxiv.org/abs/2208.08037。

Jyothi, A. A., Durand, T., He, J., Sigal, L. 和 Mori, G. LayoutVAE：从标签集生成随机场景布局。载于*IEEE/CVF*国际计算机视觉会议论文集，第 9895–9904 页，2019年。

Kawar, B., Zada, S., Lang, O., Tov, O., Chang, H., Dekel, T., Mosseri, I. 和 Irani, M. Imagic：使用扩散模型进行基于文本的真实图像编辑。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 6007–6017 页，2023。

Kikuchi, K., Simo-Serra, E., Otani, M. 和 Yamaguchi, K. 通过潜在操作优化实现受约束的图形布局生成。载于*ACM*国际多媒体会议，MM '21，第 88–96 页，2021年。doi：10.1145/3474085. 3475497。

Kong, X., Jiang, L., Chang, H., Zhang, H., Hao, Y., Gong, H., 和 Essa, I. BLT：用于可控布局生成的双向布局变换器。载于欧洲计算机视觉会议，第 474–490 页。Springer，2022年。

Labs, B.F. Flux。https://github.com/black-forest-labs/flux，2024。

Lee, H.-Y., Jiang, L., Essa, I., Le, P. B., Gong, H., Yang, M.-H., 和 Yang, W. 神经设计网络：带约束的图形布局生成，2020年。网址https://arxiv.org/abs/1912.09421。

Li, B., Zhang, Y., Guo, D., Zhang, R., Li, F., Zhang, H., Zhang, K., Li, Y., Liu, Z., 和 Li, C. Llava-onevision：轻松实现视觉任务迁移。*arXiv* 预印本 *arXiv:2408.03326*，2024。

Li, J., Yang, J., Hertzmann, A., Zhang, J., andXu, T. Layoutgan：使用矢量线框对抗网络合成图形布局。*IEEE Transactions on PatternAnal-ysis and Machine Intelligence*，43(7):2388–2399，2020a。

Li, J.、Yang, J.、Zhang, J.、Liu, C.、Wang, C. 和 Xu, T. 用于自动图形生成的属性条件布局生成对抗网络

设计。 *IEEE Transactions on Visualization and Computer Graphics*，27(10):4039–4048，2020年b。

Lin, J.、Guo, J.、Sun, S.、Xu, W.、Liu, T.、Lou, J.-G.与Zhang, D. 一种用于从文本描述生成图形布局的先解析后放置方法。见*IEEE/CVF*国际计算机视觉会议论文集，第23622–23631页，2023。

Lin, J.、Sun, S.、Huang, D.、Liu, T.、Li, J.与Bian, J. 从元素到设计：面向自动图形设计构图的分层方法。*arXiv*预印本 *arXiv:2412.19712*，2024。

Loshchilov, I.与Hutter, F. 解耦权重衰减正则化。*arXiv*预印本 *arXiv:1711.05101*，2017。

Peebles, W.与Xie, S. 基于变换器的可扩展扩散模型。见*IEEE/CVF*国际计算机视觉会议论文集，第4195–4205页，2023。Podell, D.、English, Z.、Lacey, K.、Blattmann, A.、Dockhorn, T.、M¨uller, J.、Penna, J.与Rombach, R. SDXL：改进用于高分辨率图像合成的潜在扩散模型。*arXiv*预印本 *arXiv:2307.01952*，2023。Shabani, M. A.、Wang, Z.、Liu, D.、Zhao, N.、Yang, J.与Furukawa, Y. 视觉布局合成器：用于设计布局生成的图像-矢量双扩散模型。见*IEEE/CVF*计算机视觉与模式识别会议（*CVPR*）论文集，2024年6月。Shi, Y.、Wang, P.与Huang, W. SeedEdit：将图像再生成对齐到图像编辑。*arXiv*预印本 *arXiv:2411.06686*，2024。Tang, Z.、Wu, C.、Li, J.与Duan, N. LayoutNUWA：揭示大语言模型中隐藏的布局专长，2023。URLhttps://arxiv.org/abs/2309.09506。Yamaguchi, K. CanvasVAE：学习生成矢量图形文档。见*IEEE/CVF*国际计算机视觉会议论文集，第5481–5489页，2021年。Yang, T.、Luo, Y.、Qi, Z.、Wu, Y.、Shan, Y.与Chen, C. W. PosterLLaVA：利用大语言模型构建统一的多模态布局生成器，2024a。URL https://arxiv.org/abs/2406.02884。Yang, Y.、Lu, J.、Zhao, Z.、Luo, Z.、Yu, J. J. Q.、Sanchez, V.与Zheng, F. LLPlace：通过大语言模型进行三维室内场景布局生成与编辑，2024b。URL https://arxiv.org/abs/2406.03866。

Yuan, Y., Sun, S., Liu, Q., and Bian, J. Cad-editor：一种用于基于文本的 CAD 编辑、结合自动化训练数据合成的先定位后填充框架，2025。网址 https://arxiv.org/abs/2502.03997. Zhang, K., Mo, L., Chen, W., Sun, H., and Su, Y. Magicbrush：一个用于指令式图像编辑的人工标注数据集。神经信息处理系统进展 , 36:31428–31449, 2023. Zhang, Z., Xie, J., Lu, Y., Yang, Z., and Yang, Y. 上下文内编辑：通过大规模扩散变换器中的上下文内生成实现指令式图像编辑。 *arXiv* 预印本 *arXiv:2504.20690*, 2025. Zheng, K., Chen, X., He, X., Gu, J., Li, L., Yang, Z., Lin, K., Wang, J., Wang, L., and Wang, X. E. Editroom：用于可组合的 3D 房间布局编辑的、由大语言模型参数化的图扩散方法，2024。网址 https://arxiv.org/abs/2410.12836。

A. 更多实现细节

A.1. 数据示例

我们展示一个示例，以说明 ReLayout 中的模型输入和输出。对于图像元素（例如元素 0、1、2），我们的模型会预测其中心 x 坐标、中心 y 坐标、宽度和高度。对于文本元素（例如元素 3），模型还会预测其角度、字体大小和文本对齐方式。

ReLayout 的数据示例。

输入:
画布大小: {“宽度”: 940, “高度”: 788}元素内容: [{“索引”: 0, “内容”: “<图像>”}, {“索引”: 1, “内容”: “<图像>”}, {“索引”: 2, “内容”: “<图像>”}, {“索引”: 3, “内容”: “停止空想，开始行动”}]尺寸关系: [“元素 0 大于元素 1”, “ 元素 2 小于元素 1”, ...] 位置关系 : [“ 元素 1 位于元素 0 中心 ”, “ 元素 2 位于画布顶部 ”, ...] 编辑操作 : 将元素 3 移动到 {“x”: 583, “y”: 394}

输出:

[
{"index": 0, "x": 470, "y": 394, "width": 940, "height": 806}, {"index": 1, "x": 470, "y": 394, "width": 873, "height": 721}, {"index": 2, "x": 598, "y": 147, "width": 443, "height": 418}, {"index": 3, "x": 583, "y": 394, "width": 441, "height": 336, "angle": 0, "font_size": 70, "text_align": "left"}]

A.2. GPT-4o 提示词

我们还展示了 GPT-4o 基线所使用的提示词。模型输入部分与第A.1 节中的相同，以便进行公平比较。

GPT-4o 基线的提示词。

你是一名自主人工智能助手，通过根据画布大小、元素内容、元素尺寸关系、元素位置关系以及编辑操作预测元素属性，帮助设计师进行设计布局编辑。

请确保输出设计满足指定的元素关系。例如，“element 1 large element 2”表示元素 1 应大于元素 2，“element 1 top-left canvas”表示元素 1 应放置在画布的左上角。此外，还应考虑编辑操作。

对于图像元素，你需要预测中心 x 坐标、中心 y 坐标、宽度和高度属性。对于文本元素，你还应额外预测其角度、字体大小和文本对齐方式。仅以 JSON 格式作答，不要提供其他信息。以下是一个输出示例：（一个输出示例）注意：其格式与第A.1 节中的相同。

输入:
（模型输入）注：其格式与第A.1节相同。

输出:
（模型输出）

A.3. 模型测试

为测试其性能 对于ReLayout，我们在Crello的测试集上整理了两种评估设置：重建 an d 泛化能力，如下文所述 。

重建。由于 ReLayout 利用关系感知设计重建（RADR）策略，在不使用三元组数据的情况下实现设计布局编辑，因此有必要评估其重建能力。在此设置中，测试样本的获取方式与第 4.3 节所述相同。也就是说，给定一个输入设计，我们将其元素内容、关系图以及从中提取的编辑操作作为模型输入，并观察输出设计在布局结构方面与该设计（即真实值）的相似程度（通过尺寸关系和位置关系衡量）。此外，我们还会评估输出设计的质量。为便于理解，下面详细说明从每种编辑动作中提取的编辑操作：

• 移动：通过从输入设计中随机抽取一个元素来选定目标元素。参数为目标元素实际的中心 x 坐标和中心 y 坐标。

• 调整大小：通过从输入设计中随机抽取一个元素来选定目标元素。参数为目标元素实际的宽度和高度。

• 添加：目标元素通过从输入设计中采样一个元素来选定。此动作没有参数。

• 删除：由于从给定设计中删除一个元素会改变元素数量，因此我们改为从 Crello 的验证集中采样一个元素作为目标元素。此动作没有参数。

泛化。 在此设定中，我们通过模拟用户指定的编辑操作来测试 ReLayout 的泛化能力。为此，我们对编辑操作的参数进行采样，而不是从输入设计中提取它们。因此，在此设定下不存在真实值。具体而言，

• 移动：目标元素通过从输入设计中采样一个元素来选定。其新的中心 x 坐标从范围 [0, 画布宽度] 中均匀采样。类似地，新的中心 y 坐标从 [0,画布高度]中均匀采样。

• 调整大小：目标元素通过从输入设计中采样一个元素来选定。缩放因子从范围 [0.5, 2.0]中均匀采样。编辑操作中该元素的新宽度通过将其原始宽度乘以缩放因子计算得到。类似地，其新高度通过将其原始高度乘以同一因子得到。

• 添加：目标元素从 Crello 的验证集中采样。此动作没有参数。

• 删除：从输入设计中采样一个元素作为目标元素。此动作没有参数。

B. 更多实验结果

B.1. 泛化设置的定量结果

我们已在正文中列出重建设置的定量结果。此处，我们在表4中展示泛化设置的定量结果。ReLayout 在几乎所有评估指标上都显著优于 GPT-4o。

方法	LLaVA-OV评分 ↑ (i)	(ii)	(iii)	(iv)	(v)	Ove ↓	Ali ↓	尺寸关系 ↑	位置关系 ↑	Op ↑
GPT-4o	7.23	7.20	7.39	6.92	6.21	0.1789	0.0013	0.8447	0.5444	0.9979
Ours	8.19	8.14	8.24	7.88	7.11	0.0949	0.0009	0.9475	0.9157	0.9960

表4。 泛化设置下与 GPT-4o 的定量比较。

	LLaVA-OV评分 ↑					Ove ↓	Ali ↓	尺寸关系 ↑	位置关系↑	Op↑
	(i)	(ii)	(iii)	(iv)	(v)					
替换	8.13	8.05	8.18	7.85	7.04	0.0978	0.0012	0.8840	0.8410	1.0000
双重添加	8.10	8.03	8.13	7.78	7.04	0.1126	0.0013	0.8911	0.8375	1.0000
移动与调整大小	8.20	8.06	8.22	7.85	7.11	0.0956	0.0009	0.8744	0.8467	0.9850
我们的（泛化）	8.19	8.14	8.24	7.88	7.11	0.0949	0.0009	0.9475	0.9157	0.9960

表5. ReLayout 在新编辑动作上的定量结果。

B.2. 扩展至新动作

为评估 ReLayout 的多功能性和泛化能力，我们将其应用范围从四种动作扩展到更复杂的多动作场景。具体而言，我们通过组合现有动作定义了三种全新的复合动作：

- 替换：删除和添加的组合。
- 双重添加：同时添加两个元素。
- 移动与调整大小：同时调整元素的位置和尺寸。

我们整理了一个测试集，其中每种复合动作包含 200 个样本。输入设计从 Crello测试集中随机抽取。值得注意的是，动作组合非常灵活，并不限于这三种。例如，“添加 + 移动”表示将一个元素添加到特定位置，“添加 + 调整大小”表示添加一个具有特定尺寸的元素。

ReLayout 的一项关键优势是能够在推理时处理这些复合动作，而无需额外训练或进行架构修改。如表 5 所示，我们的模型在三种新颖应用中均取得了有竞争力的表现，无需进行特定动作训练，同时保持了较高的编辑质量和关系保持能力。定量结果表明，ReLayout 对未见过的动作组合具有稳健的泛化能力。

总之，ReLayout 是首个支持多种编辑操作的设计布局编辑方法。除了论文中讨论的四种动作外，多种动作的组合也可扩展至其他新颖且未见过的动作。实验结果突显了我们方法的关键优势：强大的泛化能力、广泛的适用性、对多种编辑操作的支持，以及出色的灵活性。

B.3. 泛化设置的更多定性比较

我们在图 11、12、13 和 14 中展示了泛化设置下更多的定性结果，以证明 ReLayout 的有效性。结果表明：（1）ReLayout 是一种灵活的方法，能够使用单个模型准确执行不同类型的编辑操作；（2）ReLayout 能够在编辑后的设计中良好地保留输入设计的布局结构；（3）ReLayout 能够生成高质量的编辑后设计。

B.4. 质量比较

如图所示，我们对 ReLayout 与现有设计生成模型（包括 LaDeCo 和 FlexDM）进行了定性比较 15。在视觉质量方面，我们的方法生成的设计布局与最先进模型 LaDeCo 不相上下。此外，与真实值（GT）设计布局的比较凸显了 ReLayout 强大的重建能力。

C. 效率分析

我们的方法平均每个样本的推理时为 14.8 秒。该速度慢于 FlexDM（0.3 秒/样本）这一非自回归生成模型，也慢于 PosterLLaVA（12.0 秒/样本）；后者不会将所有图像内容编码到提示词中，因此输入序列更短。另一方面，我们的方法比 LaDeCo（17.5 秒/样本）更高效，后者包含中间渲染过程；也比受 API调用速率限制约束的 GPT-4o 更高效。



图 *11.* 更多定性　　移动操作的比较。被编辑元素和目标位置在输入设计中标出。　　.

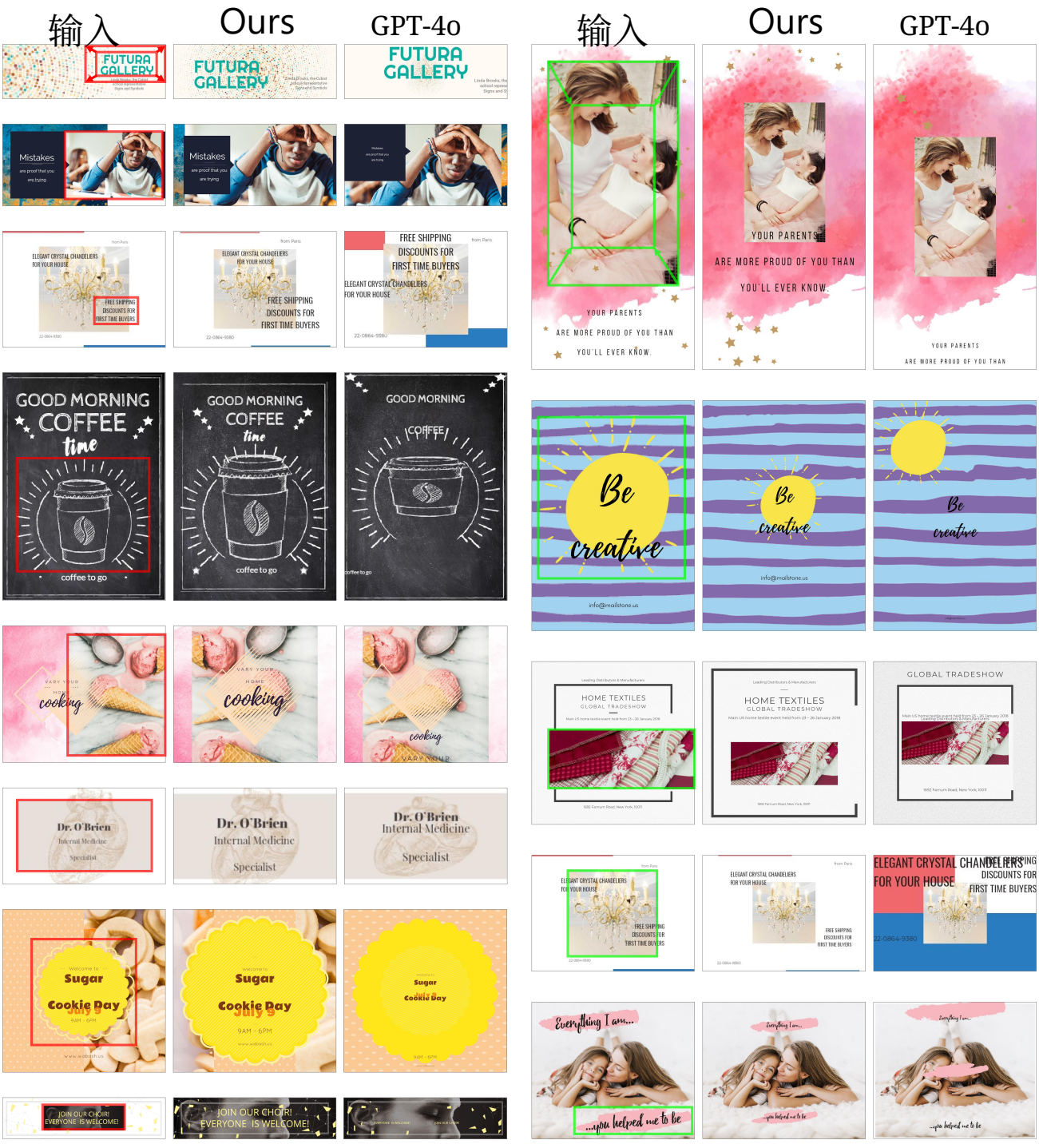


图*12*。 调整大小操作的更多定性比较。输入设计中标出了被编辑元素，其中红色 表示放大调整操作（左侧），绿色 表示缩小操作（右侧）。

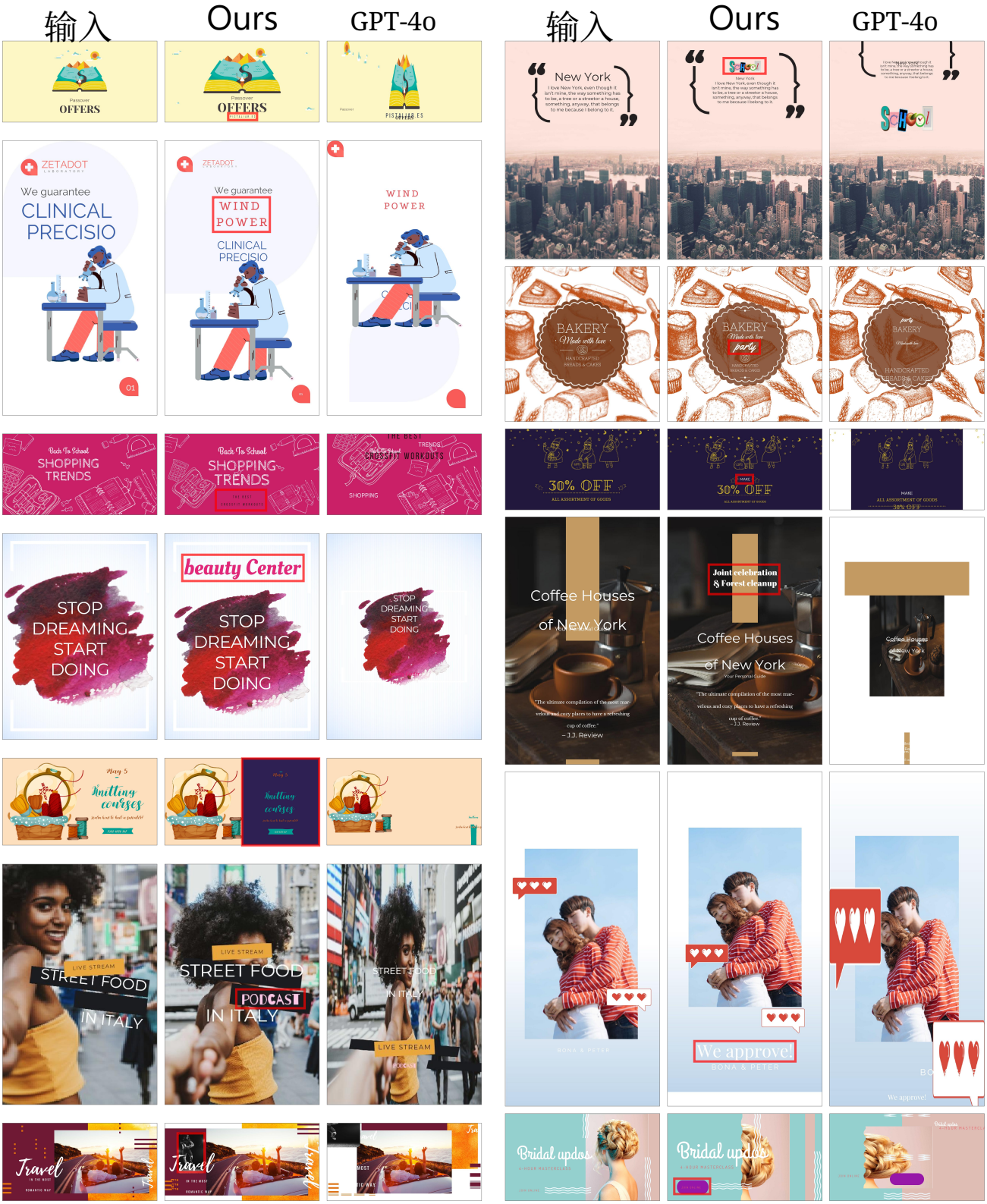


图13。更多定　　性的添加操作对比。我们的方法在输出设计中标记了添加的元素。



图*14.* 更多删除操作的定性比较。输入设计中被删除的元素已标出。



图*15.* 在重建设置下与现有设计生成方法进行定性比较。图中还展示了真实标注（GT）设计布局。