

# UI 图层合并器：基于计算机视觉与边界先验的 UI 图层合并方法

陈云农<sup>1,5</sup>, 甄焱鲲<sup>4</sup>, 施楚宁<sup>2</sup>, 李佳智<sup>2</sup>, 陈柳青<sup>2,3,5</sup>,  
李泽健<sup>1,3,5</sup>, 孙凌云<sup>2,3,5</sup>, 周婷婷<sup>4</sup>, 常艳芳<sup>4</sup>

<sup>1</sup>浙江大学软件学院, 中国杭州市, 310027

<sup>2</sup>浙江大学计算机科学与技术学院, 中国杭州市, 310027

<sup>3</sup>阿里巴巴-浙江大学前沿技术联合研究中心, 中国杭州市, 310027

<sup>4</sup>阿里巴巴集团, 中国杭州市, 311121

<sup>5</sup>浙江-新加坡人工智能与创新设计联合实验室, 中国杭州市, 310027

**摘要:** 随着互联网行业图形用户界面 (GUI) 开发工作量的快速增长, 国内外学者试图从 GUI 图片生成可维护的前端代码。该任务更适合从包含有用户界面元数据的 UI 设计稿来实现代码生成。然而, 碎片图层不可避免地会出现在 UI 设计稿中, 这大大降低了生成代码的质量。目前尚未有自动化 GUI 技术来检测并合并碎片图层以提高生成代码的可维护性。本文提出 UI 图层合并器 (UILM), 一种基于计算机视觉的方法, 可以自动检测并合并碎片图层为完整的 UI 组件。本文提出的 UILM 包含合并区域检测器和碎片图层合并算法。合并区域检测器结合边界先验知识来准确地检测出 UI 组件的边界。基于此, 碎片图层合并算法可以搜索出边界内的相关碎片层并将其合并为整体。此外, 我们提出一种动态数据增强算法来提高合并区域检测器的性能。我们还构建了用于训练合并区域检测器和测试 UILM 的大规模 UI 数据集。实验结果表明, 本文提出的方法在合并区域检测方面优于最佳基线, 并达到了有效的图层合并准确率。最后, 对真实应用程序的用户研究也证实了本文方法的有效性。

**关键词:** 用户界面 (UI) 生成代码; UI 设计规范; UI 碎片图层合并; 目标检测

<https://doi.org/10.1631/FITEE.2200099>