

基于结构感知的可视化检索复现

摘要

随着数据可视化的广泛应用，在线上创建和分享的可扩展矢量图形（SVG）可视化数量激增。如何从海量的可视化中检索出感知上相似的可视化成为一个关键问题，因其在可视化推荐等下游应用中具有重要意义。现有方法主要关注将可视化视为位图图像，仅考虑其视觉外观，忽略了 SVG 本身固有的结构信息。本文复现了 Haotian Li 等人发表于 CHI 的《Structure-aware Visualization Retrieval》方法，该方法通过结合视觉信息和 SVG 结构信息，提升了可视化检索的性能。

关键词：数据可视化；可视化检索；结构感知；对比学习；图神经网络

1 引言

1.1 选题背景

现代网络信息传递都是基于大数据，数据可视化成为数据分析和决策支持的重要工具。在 Plotly 和 Observable 等在线平台上，用户可以创建并分享大量基于 SVG 的可视化。这些可视化不仅具有良好的信息表达，还能在不同分辨率下保持高质量。然而，面对海量的可视化资源，如何高效地检索出在感知上相似的可视化成为一个亟待解决的问题。相似的可视化不仅有助于知识共享和可视化推荐，还能提升用户的数据探索和分析效率。

1.2 选题依据

现有的可视化检索方法主要通过将可视化图视为位图图像，然后通过计算机视觉的相关技术来衡量图片的相似性。然而，对基于 SVG 的可视化，这些方法忽略了 SVG 本身固有的结构信息，如视觉元素之间的空间和层级关系。结构信息能够更准确地描述可视化的组成和组织方式，对于区分不同类型或具有不同数据趋势的可视化具有重要作用。因此，结合视觉信息和结构信息的可视化检索方法具有更高的潜力和实际应用价值。

1.3 选题意义

通过复现 Haotian Li [3] 等人的结构感知可视化检索方法，我们不仅验证了该方法的有效性，还深入理解了结构信息在可视化相似性评估中的关键作用。这一复现工作为后续的可视化检索研究提供了实证基础，并为实际应用中开发更智能的可视化推荐系统提供了参考。此外，复现过程中对方法的细节优化和创新也为进一步提升检索性能提供了可能。

2 相关工作

2.1 可视化检索查询方法

可视化检索旨在从海量的可视化集合中找到与查询可视化相似的结果，支持用户在数据分析和可视化设计中的需求。根据查询方式的不同，现有方法主要分为基于定义的检索和基于示例的检索。

2.1.1 基于定义的检索

基于定义的检索允许用户通过编程语言或自然语言明确指定检索条件。例如，Hoque 和 Agrawala [2] 提供了一种 JSON-like 的规范，用户可以通过指定编码类型等特征来检索目标可视化。

2.1.2 基于示例的检索

基于示例的检索则通过提供一个或多个示例可视化，系统返回与之相似的可视化。近年来，许多研究如 ScatterNet [4] 和 ChartNavigator [8] 利用 CNN 提取可视化的视觉特征，通过计算特征向量之间的距离来衡量相似性。

2.2 可视化相似性评估

2.2.1 数据相似性

数据相似性方法侧重于可视化所呈现的数据的分布或元数据，如 SeeDB [5] 和 ShapeSearch [6]，但这些方法极其依赖于原始数据的可用性，因而限制了其应用场景。

2.2.2 感知相似性

感知相似性方法则不依赖于原始数据，通过提取可视化图的视觉特征来衡量相似性。尽管这些方法在一定程度上反映了数据的相似性，但由于忽略了结构信息，其在区分具有不同结构的可视化方面存在不足。

2.3 可视化存储方法

可视化的存储格式多样，包括光栅图像、矢量图形（如 SVG）和基于程序的格式（如 D3、Vega-Lite）。光栅图像兼容性高，但难以编辑且可能丢失部分可视化特定信息。矢量图形如 SVG 能够保留更多结构信息，支持交互和高质量缩放，是在线可视化平台的常用格式。

本研究选择利用 SVG 中的结构信息，通过图神经网络（GNN）和对比学习方法，结合视觉信息，实现更准确的可视化检索。

3 本文方法

3.1 本文方法概述

本文复现了 Li [3] 等人提出的结构感知可视化检索方法。该方法通过结合 SVG 中的结构信息和可视化的视觉信息，生成高维嵌入向量，用于衡量可视化之间的相似性。具体步骤如图 1 所示：

- 结构信息提取：将 SVG 转换为图结构，提取视觉元素的特征，构建图神经网络（GNN）模型进行编码，生成结构嵌入向量。
- 视觉信息提取：将 SVG 渲染为位图图像，利用卷积神经网络（CNN）提取视觉特征，生成视觉嵌入向量。
- 特征融合与检索：将结构嵌入向量和视觉嵌入向量进行归一化和拼接，形成最终的嵌入表示。通过计算嵌入向量之间的余弦相似度，实现相似可视化的检索。

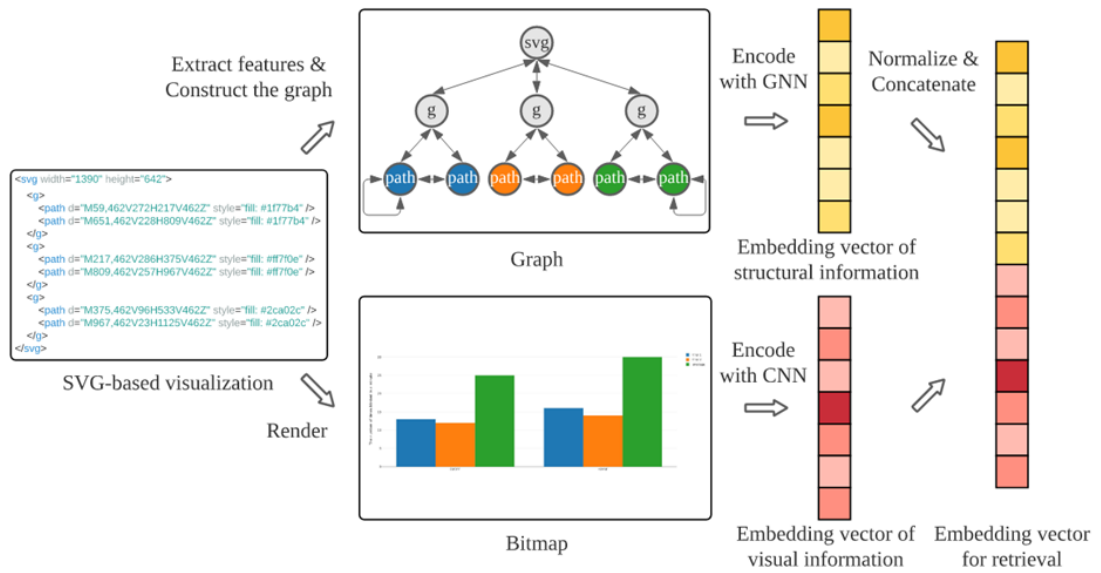


图 1. 方法示意图

3.2 特征提取模块

3.2.1 结构信息提取

- 图构建：将 SVG 中的视觉元素（如 `<path>`、`<rect>`）作为图的节点，根据元素的层级关系和空间位置构建边。
- 特征提取：为每个节点提取类型、样式、形状和位置等特征。类型通过独热编码表示，样式包括颜色、线宽、透明度等，形状特征包括面积、宽度、高度等。
- 图神经网络编码：利用 InfoGraph [7]，一种基于对比学习的图级表示学习方法，编码图结构信息，生成结构嵌入向量。

3.2.2 视觉信息提取

- 位图渲染：将 SVG 渲染为固定尺寸的位图图像，保持可视化的原始趋势。
- 特征提取：采用 SimSiam [1] 对比学习框架，利用 ResNet-50 作为编码器，提取视觉特征，生成视觉嵌入向量。

3.3 损失函数定义

采用对比学习中的对比损失，通过最大化相似样本对的相似性，最小化不同样本对的相似性。结构嵌入和视觉嵌入分别通过 GNN 和 CNN 进行编码，并通过余弦相似度计算损失，实现无监督的特征学习。

4 复现细节

4.1 与已有开源代码对比

使用 InfoGraph [7] 对构建好的 svg 图生成结构信息的嵌入向量，源码可见<https://github.com/dmlc/dgl/tree/master/examples/pytorch/infograph>. 该方法的核心思想是：

- 使用 GNN 作为编码器
- 最大化整个图与其子结构之间的互信息
- 最小化不同图之间子结构的互信息

使用 SimSiam [1] 来对可视化位图生成视觉信息的嵌入向量，源码可见<https://github.com/facebookresearch/simsiam>. 该方法的核心思想是：

- 对比学习
- 适配可视化图的数据增强

4.2 实验环境搭建

- GPU: Tesla P100 (16GB)
- 操作系统: Ubuntu 20.04
- 编程语言: Python 3.8
- 深度学习框架: PyTorch 1.90
- 其它依赖库: DGL etc.

4.3 创新点

在复现过程中，我们对原方法进行了以下创新：

- 特征优化：在结构特征提取中，增加了对 `<text>` 元素长度的特征提取，增强了对带有文本注释的可视化的区分能力。
- 融合策略改进：尝试了多种特征融合方法，包括加权平均和主成分分析（PCA），最终发现简单拼接方法在性能上表现最佳，但为后续研究提供了多种融合思路。
- 扩展支持：扩展了对更多 SVG 元素类型的支持，如 `<ellipse>` 和 `<polygon>`，提升了方法的适用范围。
- 数据增强：在视觉特征提取模块，进行了适应可视化图的数据增强方式

5 实验结果分析

如图 2 所示，是对输入可视化图的检索结果，从上至下分别是只使用视觉特征、只使用结构特征、结合两种特征检索得到的 top-5 结果，从感知上可知，第三个各结果更加符合输入图携带的可视化信息



图 2. 实验结果示意

如图 3 所示，是对两个可视化结果的相似度计算，分别是基于两种特征、只基于视觉特征、只基于结构特征。

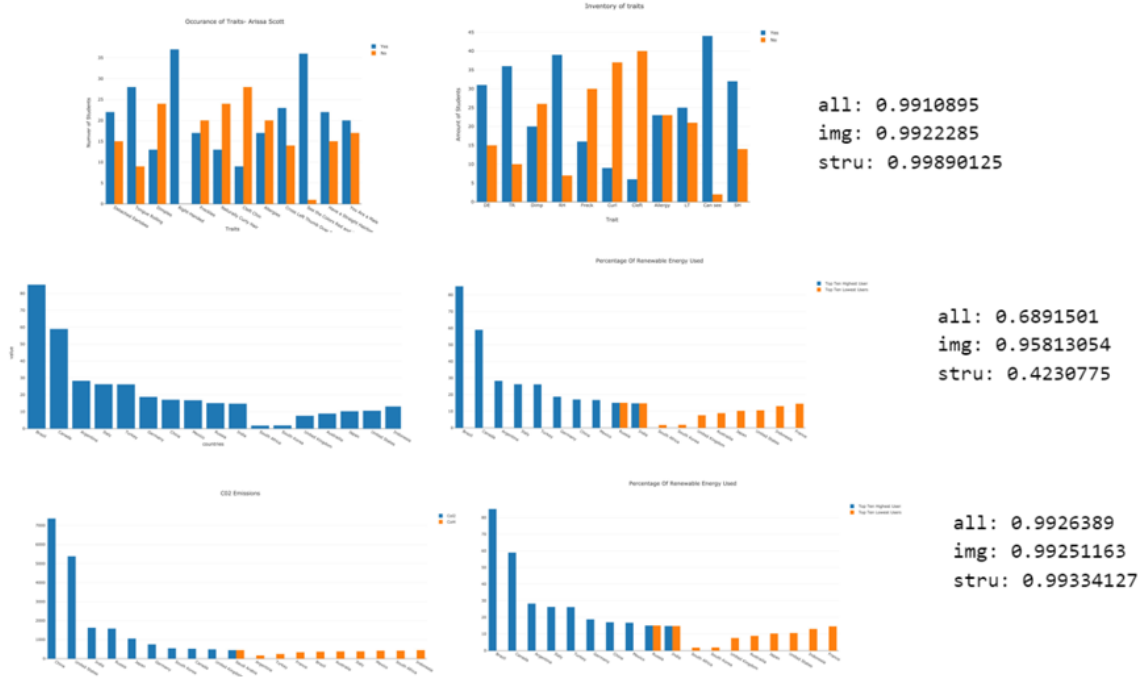


图 3. 实验结果示意

6 总结与展望

本文通过复现 Li [3] 等人提出的结构感知可视化检索方法，验证了结合 SVG 中的结构信息和视觉信息在提升可视化检索性能方面的有效性。通过定量评估和用户研究，我们证明了结构信息在保持可视化类型一致性和视觉元素数量一致性方面的重要作用。此外，在复现过程中，我们对特征提取和融合策略进行了优化和扩展，进一步提升了方法的适用性和性能。

未来的研究可以从以下几个方面进一步拓展：

- 特征提取优化：探索更高级的结构特征和视觉特征提取方法，如引入注意力机制，提升特征表示能力。
- 融合策略改进：尝试更复杂的多模态特征融合方法，如基于图注意力网络的融合策略，以更有效地结合结构与视觉信息。
- 跨平台适应：扩展方法以适应不同可视化创建工具生成的 SVG 格式，提升方法的通用性和适用范围。
- 实时检索优化：优化嵌入向量的生成和检索算法，提升系统在大规模数据集上的实时检索能力。

参考文献

- [1] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020.

- [2] Enamul Hoque and Maneesh Agrawala. Searching the visual style and structure of d3 visualizations. *IEEE transactions on visualization and computer graphics*, 26(1):1236–1245, 2019.
- [3] Haotian Li, Yong Wang, Aoyu Wu, Huan Wei, and Huamin Qu. Structure-aware visualization retrieval. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2022.
- [4] Yuxin Ma, Anthony KH Tung, Wei Wang, Xiang Gao, Zhigeng Pan, and Wei Chen. Scatter-net: A deep subjective similarity model for visual analysis of scatterplots. *IEEE transactions on visualization and computer graphics*, 26(3):1562–1576, 2018.
- [5] Tarique Siddiqui, Albert Kim, John Lee, Karrie Karahalios, and Aditya Parameswaran. Effortless data exploration with zenvisage: an expressive and interactive visual analytics system. *arXiv preprint arXiv:1604.03583*, 2016.
- [6] Tarique Siddiqui, Paul Luh, Zesheng Wang, Karrie Karahalios, and Aditya Parameswaran. Shapesearch: A flexible and efficient system for shape-based exploration of trendlines. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, pages 51–65, 2020.
- [7] Fan-Yun Sun, Jordan Hoffmann, Vikas Verma, and Jian Tang. Infograph: Unsupervised and semi-supervised graph-level representation learning via mutual information maximization. *arXiv preprint arXiv:1908.01000*, 2019.
- [8] Tianye Zhang, Haozhe Feng, Wei Chen, Zexian Chen, Wenting Zheng, Xiaonan Luo, Wenqi Huang, and Anthony Tung. Charnavigator: an interactive pattern identification and annotation framework for charts. *IEEE Transactions on Knowledge and Data Engineering*, 35(2):1258–1269, 2021.