

# 计算机视觉中的提示学习：综述

雷一鸣<sup>1</sup>, 李婧琦<sup>1</sup>, 李子龙<sup>1</sup>, 曹原<sup>1</sup>, 单洪明<sup>2,3,4</sup>

<sup>1</sup>上海市智能信息处理重点实验室, 计算机科学技术学院, 复旦大学, 中国上海市, 200438

<sup>2</sup>类脑智能科学与技术研究院, 复旦大学, 中国上海市, 200433

<sup>3</sup>脑科学前沿科学中心, 复旦大学, 中国上海市, 200433

<sup>4</sup>上海脑科学与类脑研究中心, 中国上海市, 201210

**摘要:** 自大型预训练视觉-语言模型 (VLM) 爆发以来, 提示学习已在计算机视觉领域引发广泛关注。基于VLM构建的视觉和语言信息之间的密切关系, 提示学习成为许多重要应用领域 (如人工智能内容生成 (AIGC)) 中的关键技术。本综述循序渐进且全面地总结了与AIGC相关的视觉提示学习。首先介绍了VLM, 它是视觉提示学习的基础。然后, 回顾了视觉提示学习方法和提示引导生成模型, 并讨论了如何提高将AIGC模型适用于下游特定任务的效率。最后, 提供了一些有前景的关于提示学习的研究方向。

**关键词:** 提示学习; 视觉提示微调; 图像生成; 图像分类; 人工智能内容生成 (AIGC)

<https://doi.org/10.1631/FITEE.2300389>