

# 基于对齐自修正的鲁棒跨模态检索

郭金一<sup>1</sup>, 丁洁玉<sup>2</sup>

<sup>1</sup>南京理工大学计算机科学与工程学院, 中国南京市, 210094

<sup>2</sup>青岛大学数学与统计学院, 中国青岛市, 266071

**摘要:** 跨模态检索通过为不同模态数据建立一致的对齐方式来实现模态间的相互检索。目前多种跨模态检索方法已被提出并取得良好性能。这些方法使用干净对齐的跨模态数据进行训练。虽然这些数据在语义上是匹配的, 但相较于互联网上容易获得的噪声对齐的数据 (即成对但在语义上不匹配), 标注成本很高。当用噪声对齐的数据训练这些模型时, 它们的性能会急剧下降。因此, 本文提出一种对齐自修正的鲁棒跨模态检索算法 (RCAR), 显著降低了噪声数据对模型的影响。具体来说, RCAR首先进行多任务学习, 减缓模型对噪声数据的过拟合, 使数据分离。然后, 利用两成分的贝塔混合模型将数据分为干净数据和噪声数据, 并根据后验概率修正对齐标签。此外, 在噪声对齐范式中定义两种噪声类型: 部分噪声数据和完全噪声数据。实验结果表明, 与当下流行的跨模态检索方法相比, RCAR在两种类型的噪声下都能取得更稳健的性能。

**关键词:** 跨模态检索; 鲁棒学习; 对齐修正; 贝塔混合模型

<https://doi.org/10.1631/FITEE.2200514>