

Hierarchical Similarity Learning For Language-Based Product Image Retrieval

- 研究动机

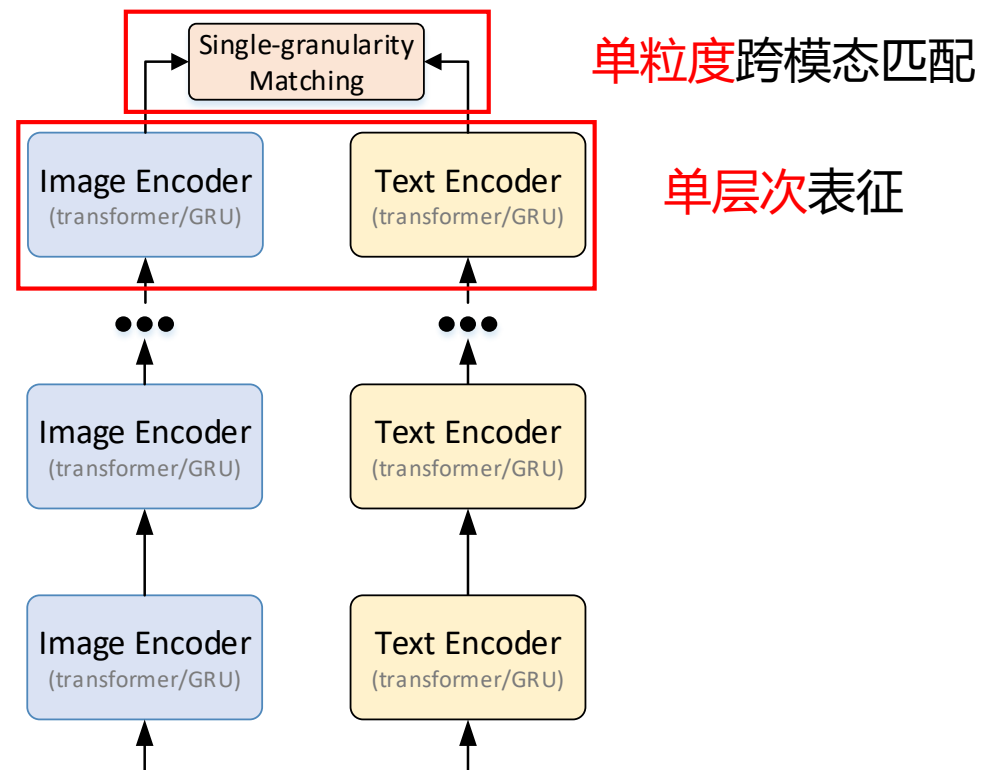
产品图像检索更为复杂

Query: Korean small fresh women's blouse



- 产品种类丰富，内容多样
- 物体尺度差异大，伴随着变形遮挡
- 服装可以单独展示（折叠）或是穿着展示
- 背景可以十分复杂（直播）
-

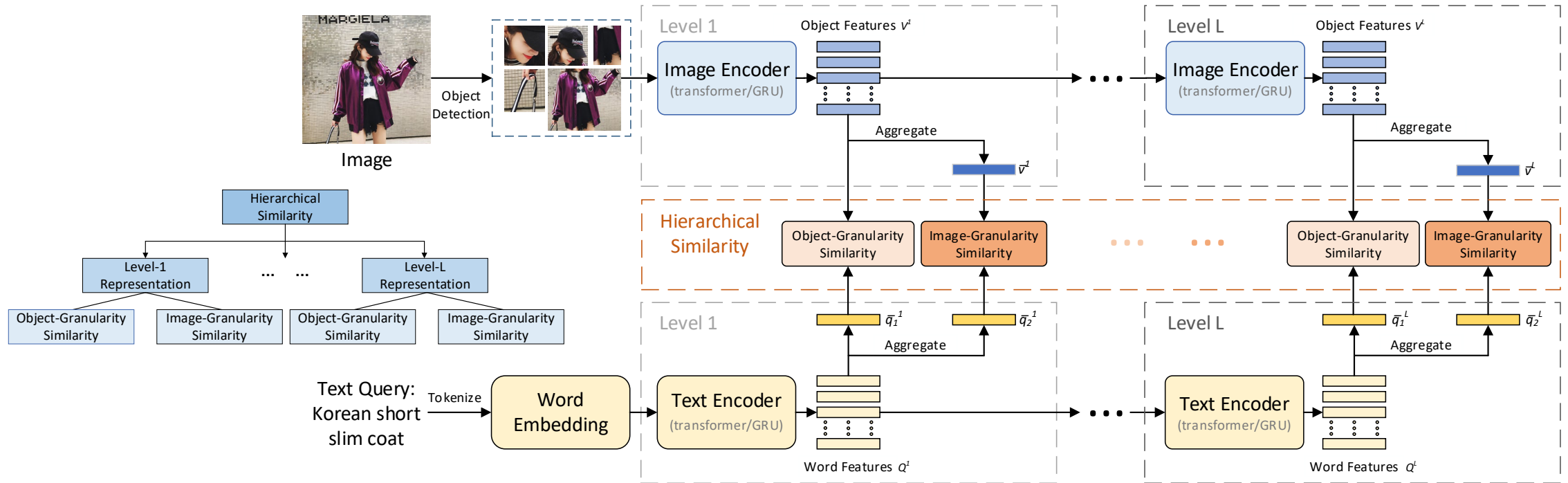
一般的跨模态检索方法难以胜任





Hierarchical Similarity Learning For Language-Based Product Image Retrieval

- 层级相似性学习



$$s(I, T) = \sum_{l=1}^L \lambda_l (\sigma_{obj}^l(I, T) + \sigma_{img}^l(I, T))$$

Hierarchical Similarity Learning For Language-Based Product Image Retrieval

• 实验结果

- 检索结果准确合理
- 在不同的召回样本数目下能够得到一致的性能
- 性能远超一般跨模态检索方法, 优于同类基于自然语言的商品图像检索方法
- 多层次表征学习与多粒度相似度效果明显, 模型性能增量式提升

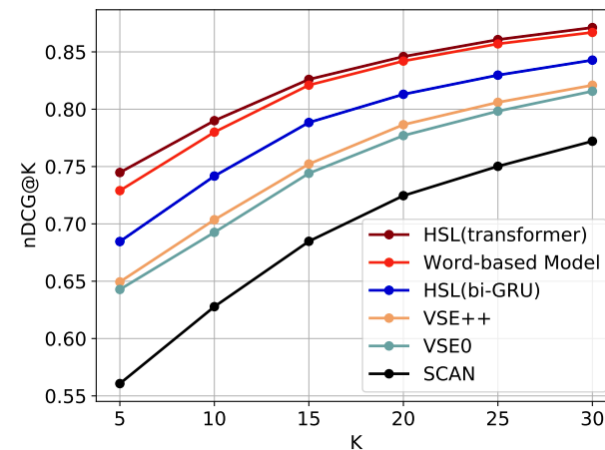
Query 1: baby toddler sandals



Query 2: Japan and Korea style flowers earrings



Query 3: nordic children's floor mats



Method	nDCG@5	Multi-level Representation?	Multi-granularity Similarity ?	nDCG@5
SCAN[23]	0.5609		✗(Object)	0.7275
VSE0[24]	0.6381	✗(Level 1)	✗(Image)	0.7240
VSE++[24]	0.6494		✓	0.7339
LXMERT+LightGBM [8]	0.6200		✗(Object)	0.7351
MCAN[10, 7]	0.6900	✗(Level 2)	✗(Image)	0.7273
VisualBERT[11, 7]	0.7100		✓	0.7362
Word-based Model[9]	0.7290		✗(Object)	0.7412
HSL(bi-GRU)	0.6846	✓	✗(Image)	0.7381
HSL(transformer)	0.7448		✓	0.7448