LOCL: Learning Object-Attribute Composition using Localization

Background: The problem of unseen O-A associations has been well studied in the field of Composition Zero-shot Learning (CZSL); however, the performance of existing methods is limited in challenging scenes.

LOCL generalizes CZSL to objects in cluttered/more realistic settings.

Approach:

**Localized Feature Extractor (LFE)**
- Generates proposals that are likely to contain objects.
- A proposal classifier takes the proposal features and refines them with object-attribute semantics.
- The refined features are used to make the final object attribute prediction.

**Results**

- Comparison with SOTA methods on simple datasets (MIT-states & UT-Zappos) and challenging dataset (CGQA).
- Table below shows effectiveness of the proposed methods compared to existing methods.

**Authors:**
SATISH KUMAR, ASM Iftekhar, Ekta Prashnani, B.S. Manjunath

This research was partially supported by NSF STC-S2S award 1725814 and US Army Research Laboratory (ARL) under agreement # W911NF2020157.