AnchorNet: A Weakly Supervised Network to Learn Geometry-sensitive Features For Semantic Matching

**Introduction**

**The task:**

**Semantic matching**

Given a pair of semantically related objects => estimate matches between corresponding parts

**Motivation:**

Fully supervised approaches [4,6] require expensive annotations / synthetic datasets => we target weak supervision

**Proposed approach - overview:**

- Given large dataset with object category image level labels
- Learn distinct features of the object categories
- Use the features within a matching algorithm

**Experiments**

**Semantic matching:**

Given a pair of images of the same object category => estimate matches between corresponding parts

**Evaluated approach**

Step 1. Extract pixel-wise descriptors
Step 2. Use a matching algorithm

**Benchmarks**

- Pascal Parts [7]
- Animal Parts [8]
- ImageNet [9]

**Evaluation procedure**

- Target image
- Source image
- Target mask
- Source mask
- Animal domains

**Results**

- **Pascal VOC - qualitative results**
- **Cross-class semantic matching:**
  Given a pair of images of related object categories => estimate matches between corresponding parts

**References**

- [4] Novotny et al. "I have seen enough: Transferring parts across categories.".
- [5] Long et al. "Do convnets learn correspondence?".