What are Unseen Objects?
- Never observed during training

### 3D Orientation Estimation
- Given a previously unseen object, we predict the category label and 3D orientation by using a retrieval-based method.

**Network Architecture**
- Given a pair of images, we use a siamese network to extract multi-scale local features.
- We present an adaptive fusion module to convert local feature similarities to a single image similarity score.

**Fast Retrieval**
- A naive image retrieval strategy compares query with every reference. Given $N$ objects with $R$ references each, the cost of $O(NR)$ quickly becomes unaffordable as $N$ and $R$ increase.
- We design a fast retrieval strategy, which is around 60 times faster than the naïve one.

**Algorithm 1: Fast Retrieval**

```
Input: $I_{Q}, I_{1}, I_{2}, \ldots, I_{R}$
Output: $L_{s}$, $R_{m}$
1. Sample $k_{Q}$ anchors from $I_{Q}$ using PFS;
2. Estimate similarities using Eq. 2;
3. Initialize $L_{s}$ as the most similar anchor;
4. $j := 1$;
5. repeat
6. Define a search space around $I_{j}$, with a radius of $R/R'$;
7. Compute anchors using PFS;
8. Estimate similarities using Eq. 2;
9. Update $L_{s}$ to the most similar anchor;
10. $j := j + 1$;
11. until $L_{s}$ converges;
12. Determine $R_{m}$ as $R_{m} := L_{s}$. 
```

**Quantitative Results**
- We conduct experiments on three datasets, LineMOD, LineMOD-Occluded, and T-LESS.
- For LineMOD and LineMOD-O, we split images to three groups according to the contained objects. We use two groups as training data and the other one as testing data.
- For T-LESS, we test the methods using the models pretrained on LineMOD.

**Ablation Studies**
- The accuracy significantly decreases when local similarities are replaced by the global similarity in our framework.

**Contact Info**
E-mail: chen.zhao@epfl.ch
Homepage: https://sailor-z.github.io