Learning Co-segmentation by Segment Swapping for Retrieval and Discovery

Xi Shen¹  Alexei A. Efros²  Armand Joulin³  Mathieu Aubry¹

¹LIGM (UMR 8049), École des Ponts ParisTech  ²UC Berkeley  ³Meta AI

1. Introduction

Motivation
Discover influences via repetition of visual details.
Place recognition with repeated segments.

Challenges
• No training data available.

Key idea: training with synthetic pairs

Main contributions
• A method of generating synthetic pairs to learn co-segmentation
• A transformer-based architecture for co-segmentation producing competitive performances on art detail retrieval and place recognition
• Spectral clustering on a correspondence graph for discovery in image collections.

2. Learning co-segmentation

Objective function
Cross-entropy loss on the predicted mask
Cross-entropy loss on the transported mask
Regression loss on the correspondences

3. Experiments

Art detail retrieval on Brueghel dataset, predicted masks are shown in transparency.
Place recognition on Tokyo24/7 (Top) and Pitts30K (Bottom) dataset, predicted masks are shown in transparency.

Score between a pair of images
\[ S(F, T) = \sum \left[ \sum \left( \text{Feature at coordinate } (i,j) \text{ in the source} \right) \cdot \text{Feature at warped coordinate } (i,j) \text{ in the target} \right] \]
Where \( M_{\text{source}}(i,j) \) is the product of the source and the transported target mask.

4. Object discovery

Discovered clusters on Brueghel dataset.
Co-segmentation results in Internet dataset for the Horse, Airplane and Car categories.

5. Training pairs

Project page (code, more experimental and visual results):
http://imagine.enpc.fr/~shenx/SegSwap/

Validation of unsupervised segments on unsupervised saliency detection benchmarks: ECSSD, DUTS, DUT-OMRON.