

# Supplementary Material for: Jointly Discriminative and Frequent Visual Representation Mining

Qiannan Wang, Ying Zhou, Zhaoyan Zhu, Xuefeng Liang (✉), and Yu Gu

School of Artificial Intelligence, Xidian University, China  
xliang@xidian.edu.cn

## 1 Attention Maps and Discovered Visual Representations

We illustrate the attention maps and discovered instances of visual representations of all categories in VOC2012-10 and Travel in Fig.1 and Fig.2, respectively. Two instances are shown for each category. The top row lists the original images, the middle row shows the attention maps after joint optimization, and the bottom row demonstrates the discovered instances of visual representation. Best viewed in color.

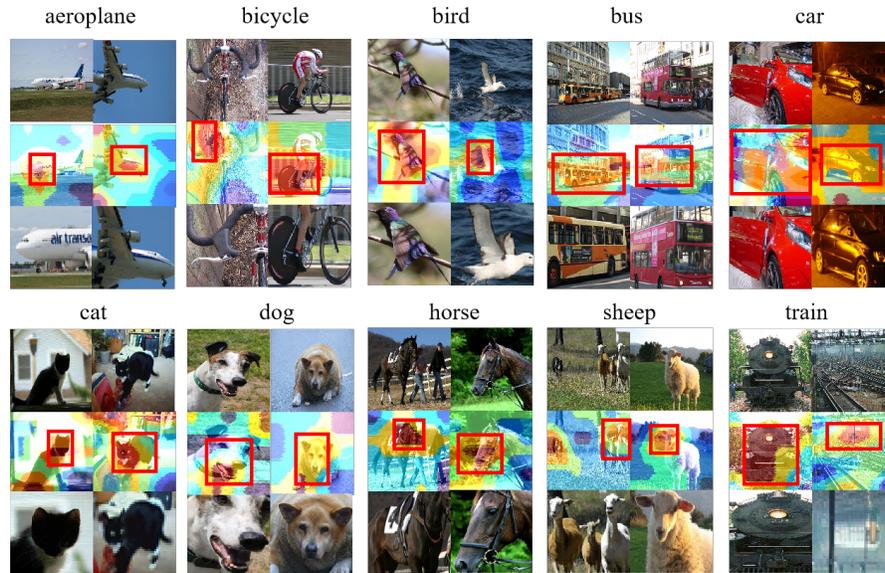


Fig. 1: Instances of discovered visual representations by JDFR from VOC2012-10.

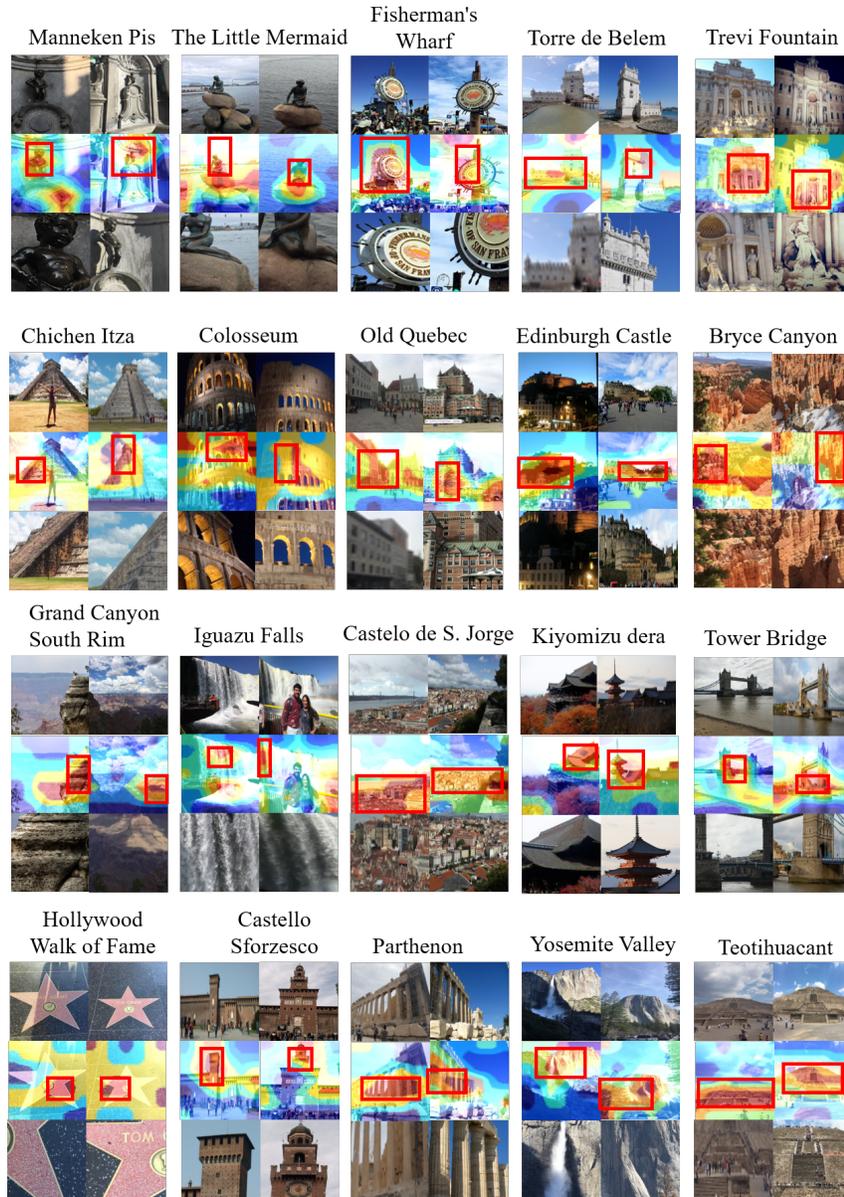


Fig. 2: Instances of discovered visual representations by JDFA from Travel.

## 2 Qualitative comparison

We illustrate ten instances of visual representation discovered by six approaches from two categories in Travel, as shown in Fig.3 and Fig.4, and two categories in VOC2012-10, as shown in Fig.5 and Fig.6. The six approaches are (a) MDPM, (b) CBMS, (c) P-CNN, (d) PatternNet, (e) MFCN and (f) JDFR (ours), respectively.



Fig. 3: Instances of visual representation of Fisherman’s Wharf category in Travel.



Fig. 4: Instances of visual representation of Torre de Belem category in Travel.

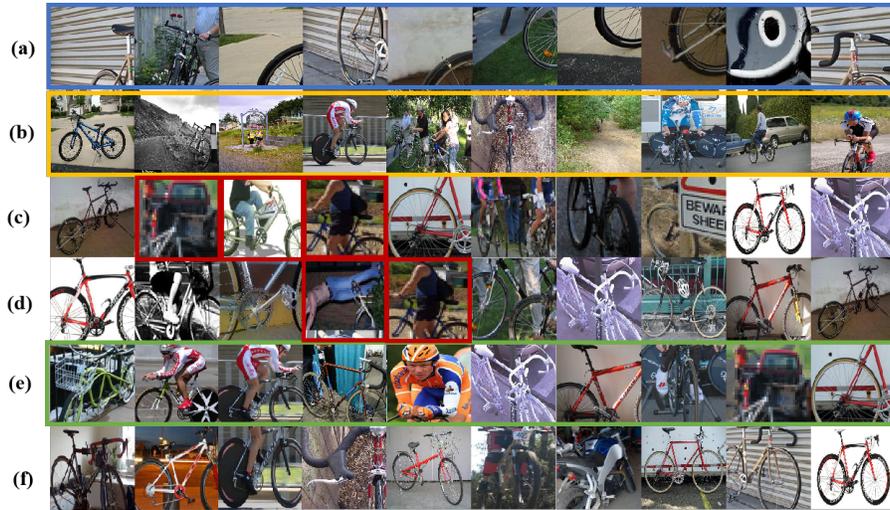


Fig. 5: Instances of visual representation of bicycle category in VOC2012-10.

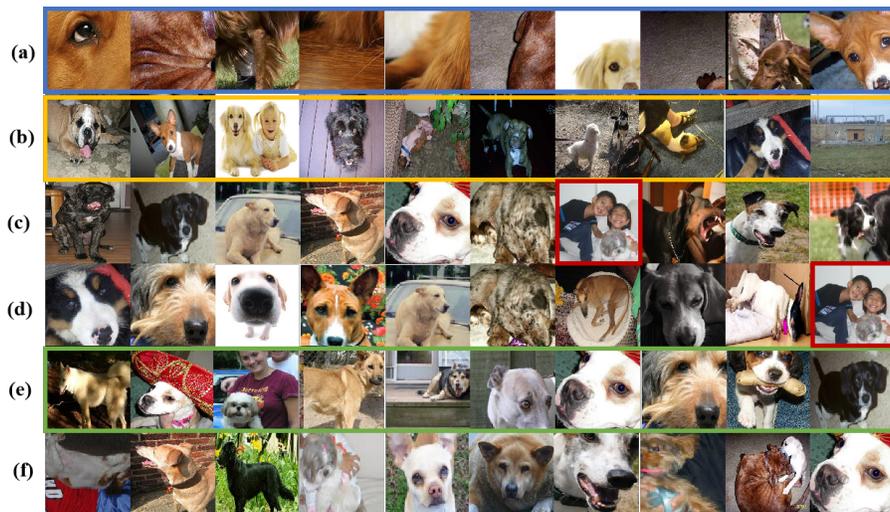


Fig. 6: Instances of visual representation of dog category in VOC2012-10.

### 3 Instances of Visual Representation in CIFAR100-20

The images in CIFAR100-20 have a low resolution  $32 \times 32$ . And, each image only contains one object that occupies the most space. Therefore, we only list four discovered instances of visual representation of a category in Fig.7.

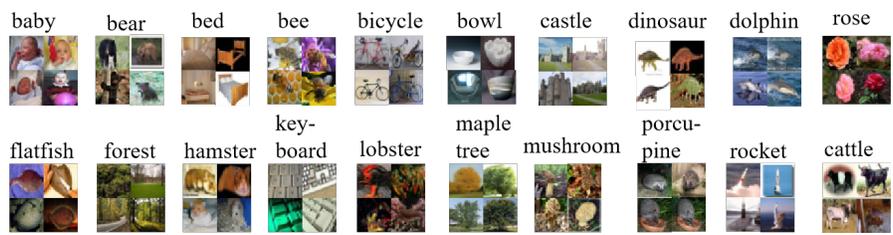


Fig. 7: Instances of CIFAR100-20.