Class-Agnostic Object Counting Robust to Intraclass Diversity
Supplementary Material

Shenjian Gong[1], Shanshan Zhang[*,1], Jian Yang[1], Dengxin Dai[2], and Bernt Schiele[2]

[1] PCA Lab, Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, and Jiangsu Key Lab of Image and Video Understanding for Social Security, School of Computer Science and Engineering, Nanjing University of Science and Technology {shenjiangong,shanshan.zhang,csjyang}@njust.edu.cn
[2] MPI Informatics {ddai,schiele}@mpi-inf.mpg.de

S1 Qualitative results of augmented exemplars.

To understand exactly how our EFA module generates new examplars, We train a image reconstruction network and reconstruct augmented exemplars. Qualitative results of augmented exemplars are shown in Fig. S1. We can find that the generated exemplars mix the color, shape, scale, texture of the original exemplars. EFA is like creating new samples of different color, shape, scale, texture in the feature space via combining the provided exemplars. We need all the exemplars we generate to be able to match all of the objects in the same class correctly, so that the matching will be robust to intraclass diversity after training.

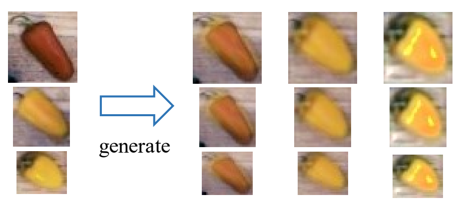

Fig. S1. We train an image reconstruction network with encoder-decoder structure. At test time, we mix the encoded features of different exemplars, and then reconstruct the corresponding image information from the mixed features.

S2 Impact of the additional branch.

Since the additional branch of EM increases the number of parameters, we further analyze whether its gain comes from more parameters or edge matching. As shown in Tab. S1, we replace the edge input of the 2nd branch with RGB and grayscale images and find that an additional stream always brings some improvement; however, using edge images brings the largest gain.

*Corresponding author

Table S1. Effect of using edge images at the 2nd branch.

2nd branch	N/A	RGB	Gray	Edge (Our EM)
MAE	23.08	22.42	21.67	20.54

S3 Impact of test time exemplar feature augmentation.

In our previous experiments, we use the same exemplar feature augmentation strategy at both training and test time. In order to improve the efficiency w.r.t. memory and running time, we investigate the impact of test time exemplar feature augmentation. As shown in Tab. S2, when we remove the exemplar feature augmentation module at test time, the result drops slightly from 20.54 to 20.94 w.r.t. MAE. Thus, our method can run with lower costs while keeping comparable performance. These results demonstrate that our method learns how to obtain robust features by employing the exemplar feature augmentation module at training time.

Table S2. Impact of test time exemplar feature augmentation.

Augmentation Quantity N at Test Time	MAE	RMSE
0	20.94	61.88
7	20.54	60.78

S4 Application of EFA in image classification.

Essentially, EFA is a method for data augmentation, which generates new samples from features of the same class of objects. We can consider proposed EFA as a plug-in module and combine it with prior methods. Here, we employ EFA to image classification on CIFAR10 dataset and analyze its impact. From S3, we find that it is not very effective, we think the reason is that the examplars belong to the same image for counting. EFA can work when data augmentation is performed on the same type of objects in the same domain.

Table S3. Impact of EFA on image classification.

Method	Acc
Resnet18 (w/o data augmentation)	77.52
Resnet18 (w/ EFA)	77.78