Relational Embedding for Few-Shot Classification
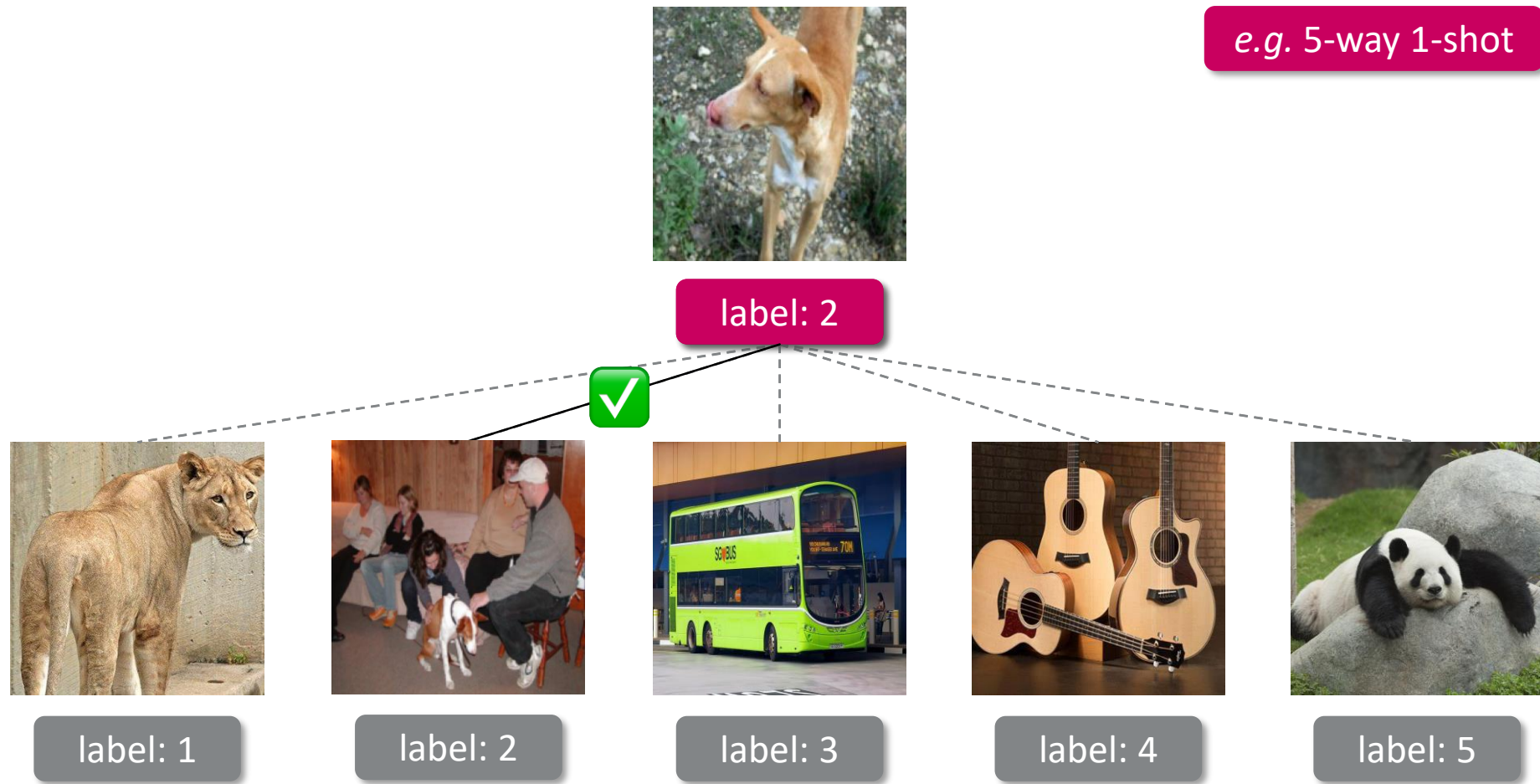
Dahyun Kang   Heeseung Kwon   Juhong Min   Minsu Cho

Pohang University of Science and Technology (POSTECH)

POSTECH

POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Few-shot image classification

e.g. 5-way 1-shot

label: 2

✅

label: 1   label: 2   label: 3   label: 4   label: 5

Few-shot image classification setup: train/test classes are disjoint

train classes	test classes (**target** classes)
bald eagle boxer bath towel …	shopping cart pretzel pineapple …

- Classifying an image into target classes that are *unseen during training*.

Relational Embedding Network (RENet)

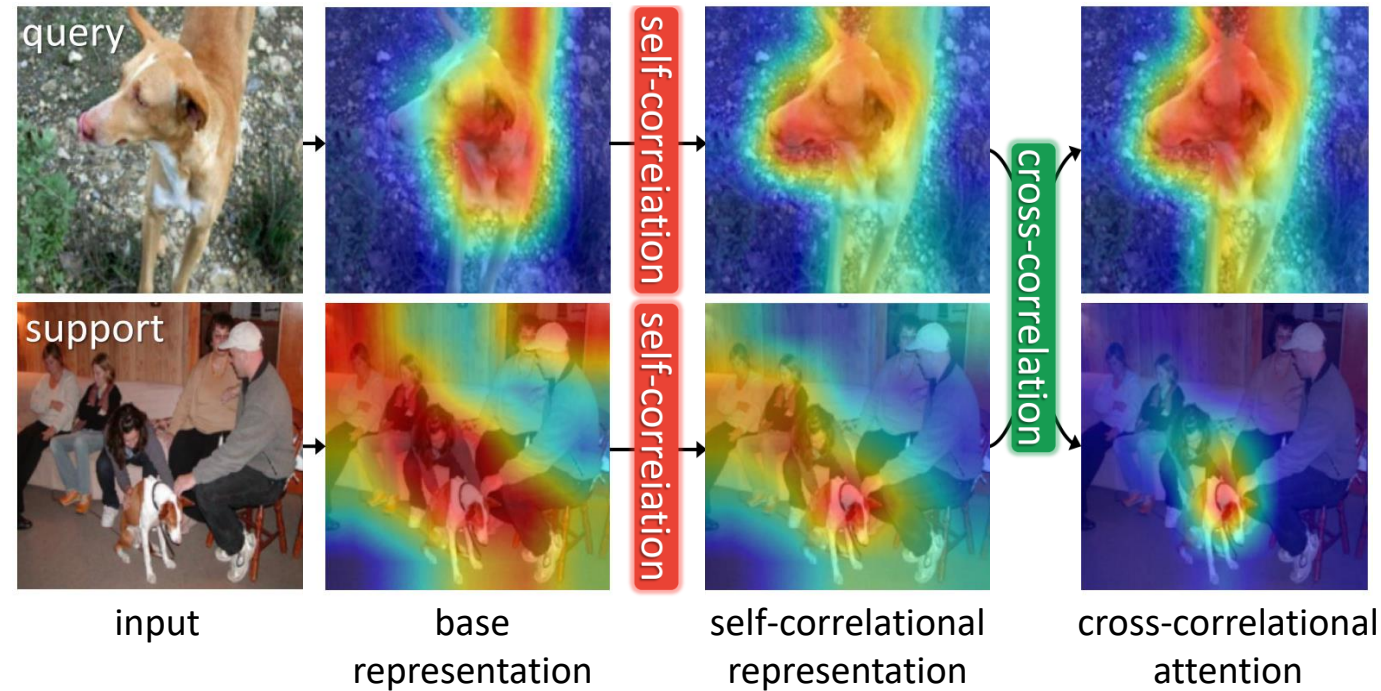input · base representation · self-correlational representation · cross-correlational attention

- Relational embedding for a metric-based classifier
  - Learning *what to observe* via self-correlational representation (SCR) module
  - Learning *where to attend* via cross-correlational attention (CCA) module

Overall architecture

Relational Embedding for Few-Shot Classification
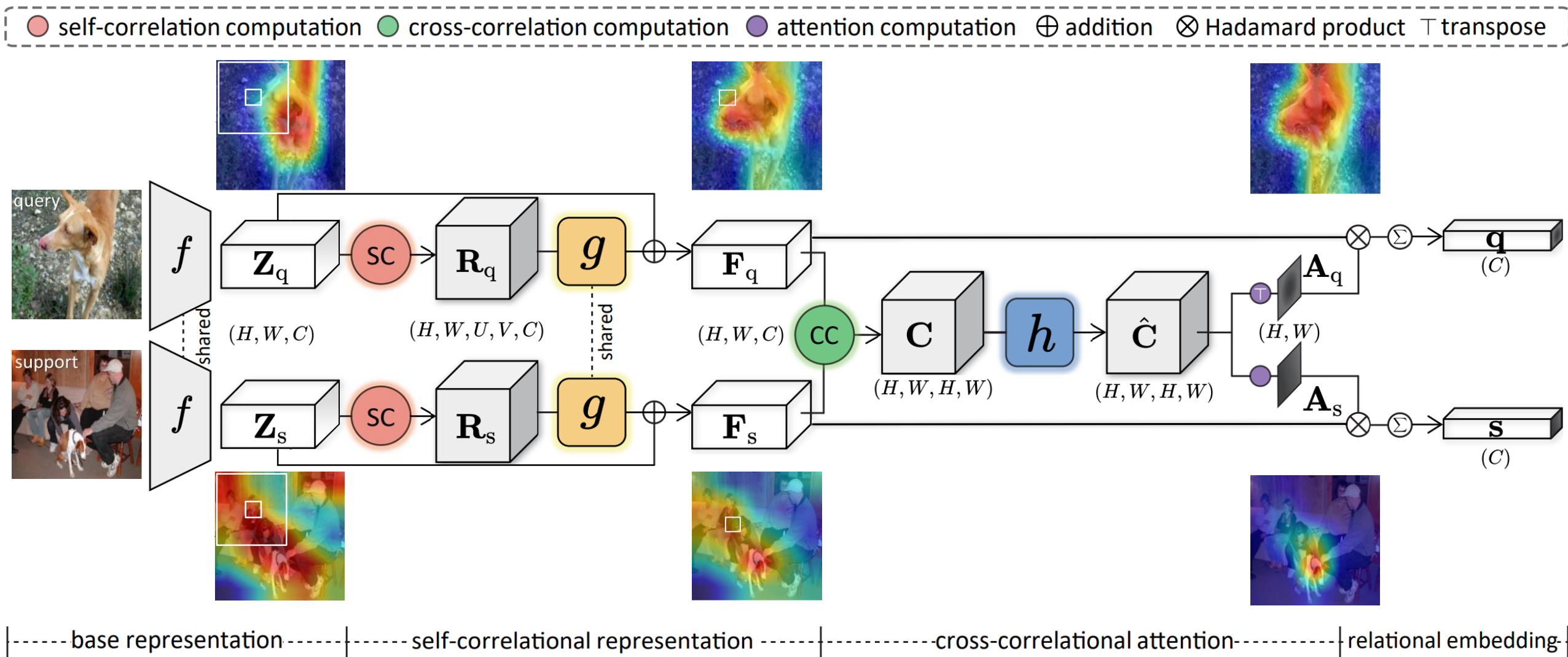
Comparison with state of the arts

method	venue	5-way 1-shot	5-way 5-shot
CTM (Li *et al.*)	ICCV'19	64.12 ± 0.82	80.51 ± 0.13
CAN (Hou *et al.*)	Neurips'19	63.85 ± 0.48	79.44 ± 0.34
FEAT (Ye *et al.*)	CVPR'20	66.78 ± 0.20	82.05 ± 0.14
DeepEMD (Zhang *et al.*)	CVPR'20	65.91 ± 0.82	82.41 ± 0.56
RENet (ours)	ICCV'21	$\mathbf{67.60 \pm 0.44}$	$\mathbf{82.58 \pm 0.30}$

miniImageNet

method	venue	5-way 1-shot	5-way 5-shot
NegMargin (Liu *et al.*)	ECCV'20	72.66 ± 0.85	89.40 ± 0.43
S2M2 (Mangla *et al.*)	WACV'20	72.92 ± 0.83	86.55 ± 0.51
FEAT (Ye *et al.*)	CVPR'20	73.27 ± 0.22	85.77 ± 0.14
DeepEMD (Zhang *et al.*)	CVPR'20	75.65 ± 0.83	88.69 ± 0.50
RENet (ours)	ICCV'21	$\mathbf{79.49 \pm 0.44}$	$\mathbf{91.11 \pm 0.24}$

CUB-200-2011

method	venue	5-way 1-shot	5-way 5-shot
CTM (Li *et al.*)	ICCV'19	68.41 ± 0.39	84.28 ± 1.73
CAN (Hou *et al.*)	Neurips'19	69.89 ± 0.51	84.23 ± 0.37
FEAT (Ye *et al.*)	CVPR'20	70.80 ± 0.23	84.79 ± 0.16
DeepEMD (Zhang *et al.*)	CVPR'20	71.16 ± 0.87	$\mathbf{86.03 \pm 0.58}$
RENet (ours)	ICCV'21	$\mathbf{71.61 \pm 0.51}$	85.28 ± 0.35

tieredImageNet

| method	venue	5-way 1-shot	5-way 5-shot
MetaOptNet (Lee *et al.*)	CVPR'19	72.6 ± 0.7	84.3 ± 0.5
Boosting (Gidaris *et al.*)	ICCV'19	73.6 ± 0.3	86.0 ± 0.2
S2M2 (Mangla *et al.*)	WACV'20	62.77 ± 0.23	75.75 ± 0.13
RFS (Tian *et al.*)	ECCV'20	71.5 ± 0.8	86.0 ± 0.5
RENet (ours)	ICCV'21	$\mathbf{74.51 \pm 0.46}$	$\mathbf{86.60 \pm 0.32}$

CIFAR-FS

RENet achieves ~3.8%p improvement from the recent state-of-the-arts models

Effects of SCR and CCA

SCR	CCA	*mini-*ImageNet	CUB
✗	✗	65.33	77.54
✓	✗	66.66 (+1.33)	78.69 (+1.15)
✗	✓	65.90 (+0.57)	78.49 (+0.95)
✓	✓	**67.60** (+2.27)	**79.49** (+1.95)

- The two proposed modules consistently improve the baseline model.

Qualitative results
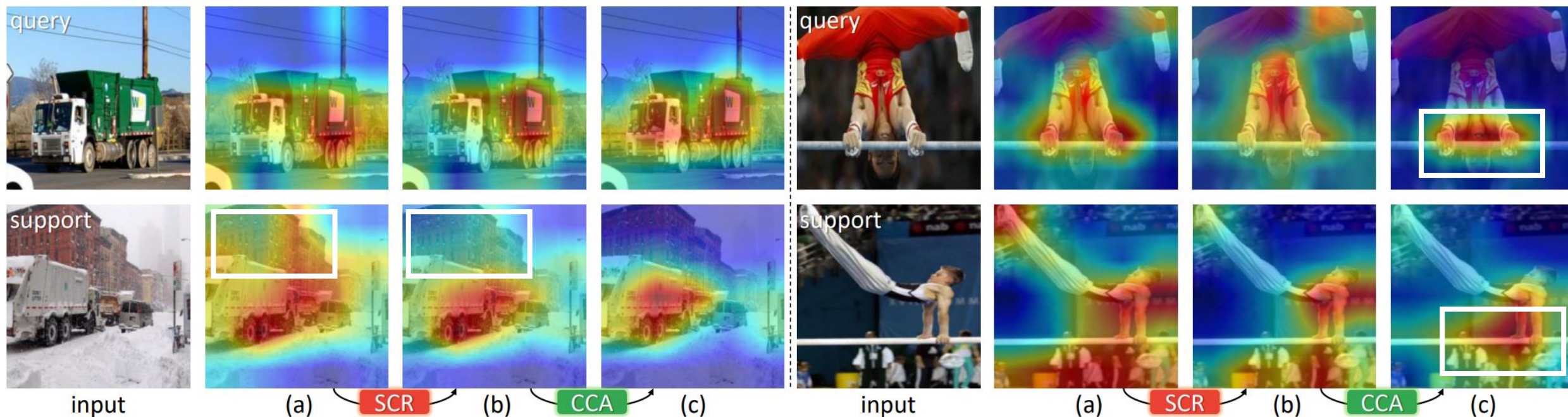
(a): Channel activation of base representation.

(b): Channel activation of self-correlational representation.

(c): Cross-correlational attention maps.

Visit our poster #4102

Paper	Code	Project homepage	Author's homepage
arxiv.org/abs/2108.09666	github.com/dahyun-kang/renet	cvlab.postech.ac.kr/research/RENet	dahyun-kang.github.io

Relational Embedding for Few-Shot Classification