CO2: Consistent Contrast for Unsupervised Visual Representation Learning

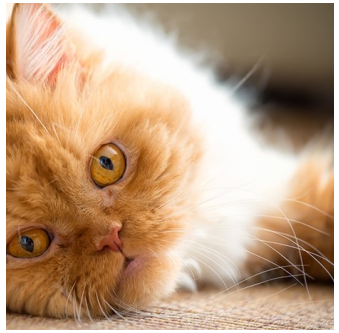
ICLR 2021

Chen Wei, Huiyu Wang, Wei Shen, Alan Yuille

Johns Hopkins University, Shanghai Jiao Tong University

Instance Discrimination

query



positive key

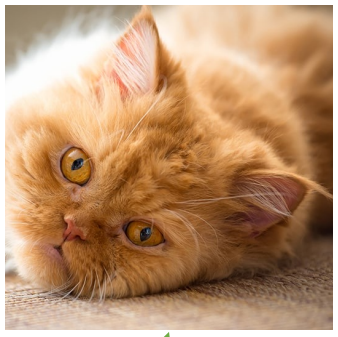
negative keys

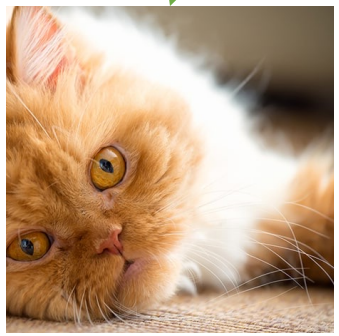
Wu et al., CVPR 2018

Instance Discrimination

…from the same instance

query





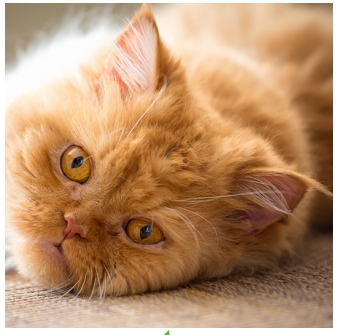
positive key

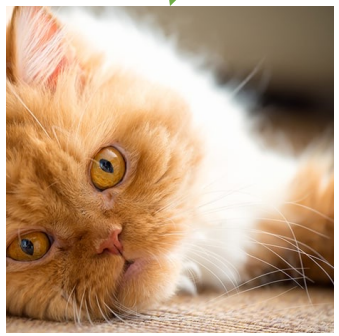
negative keys

Instance Discrimination

…from the same instance

query





positive key

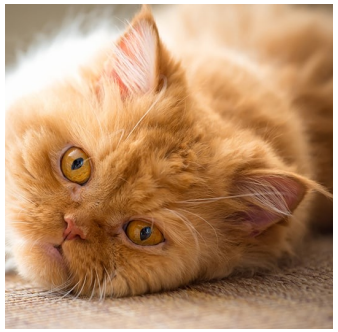
randomly sampled…

negative keys

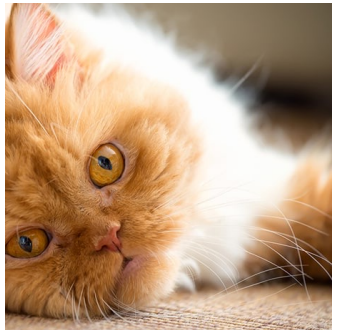
Cross-Entropy Loss

…from the same instance

query



1 0 0 0 0 0



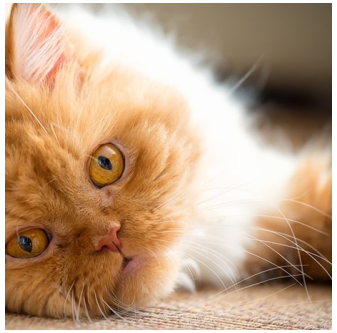
positive key

negative keys (randomly sampled)

Cross-Entropy Loss

query

1 0 0 0 0 0





Cross-Entropy Loss

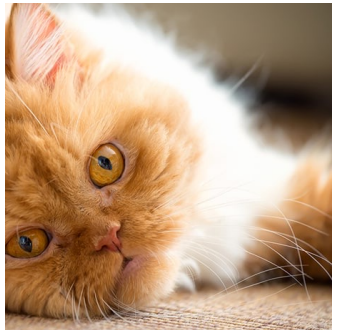
query

self-**supervised**:

artificially created task

artificial **labels**

1 0 0 0 0 0

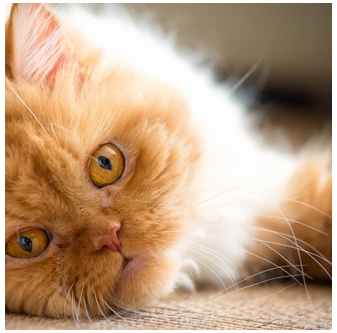




Negative?

query

1 0 0 0 0 0



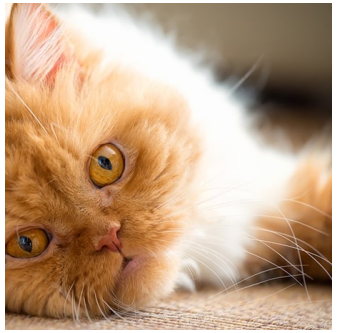
semantically similar!

Negative?

query

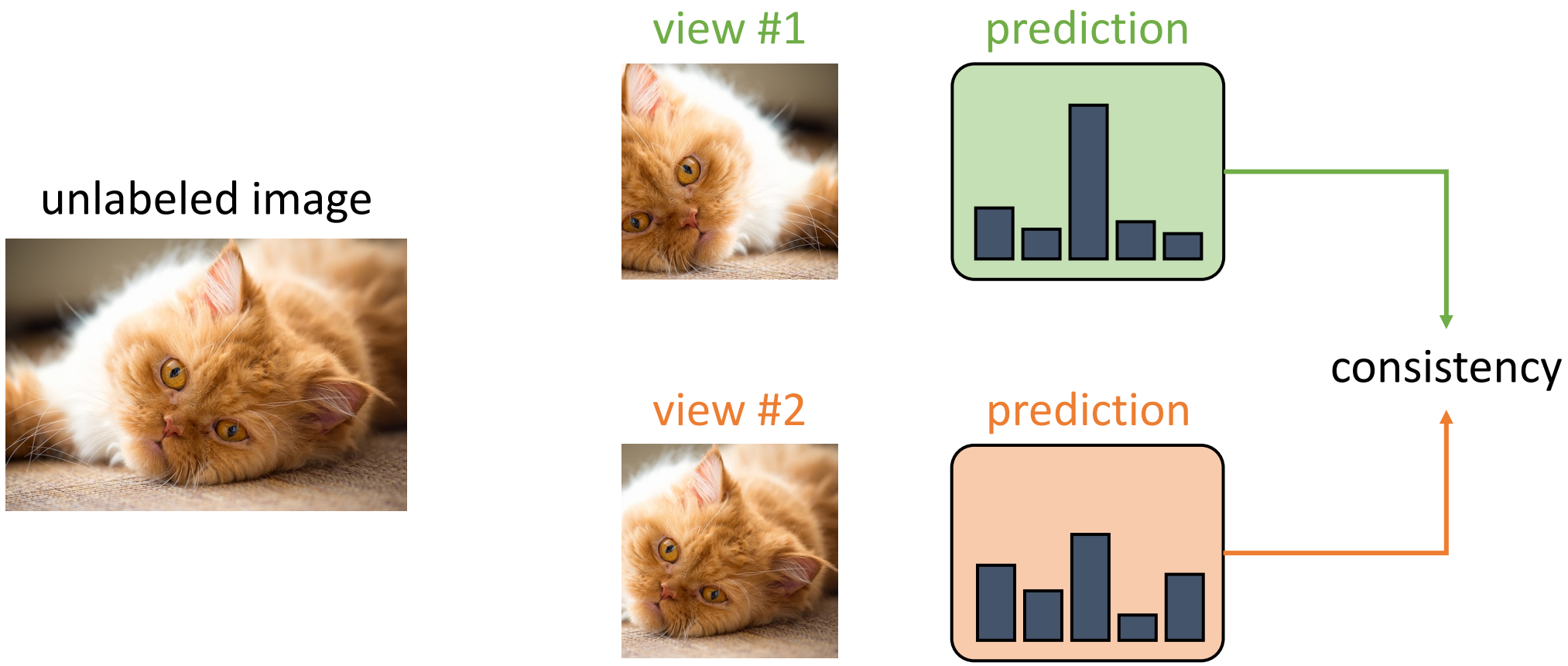
but we have no labels for this query-negative similarity…

1 0 0 0 0 0



semantically similar!

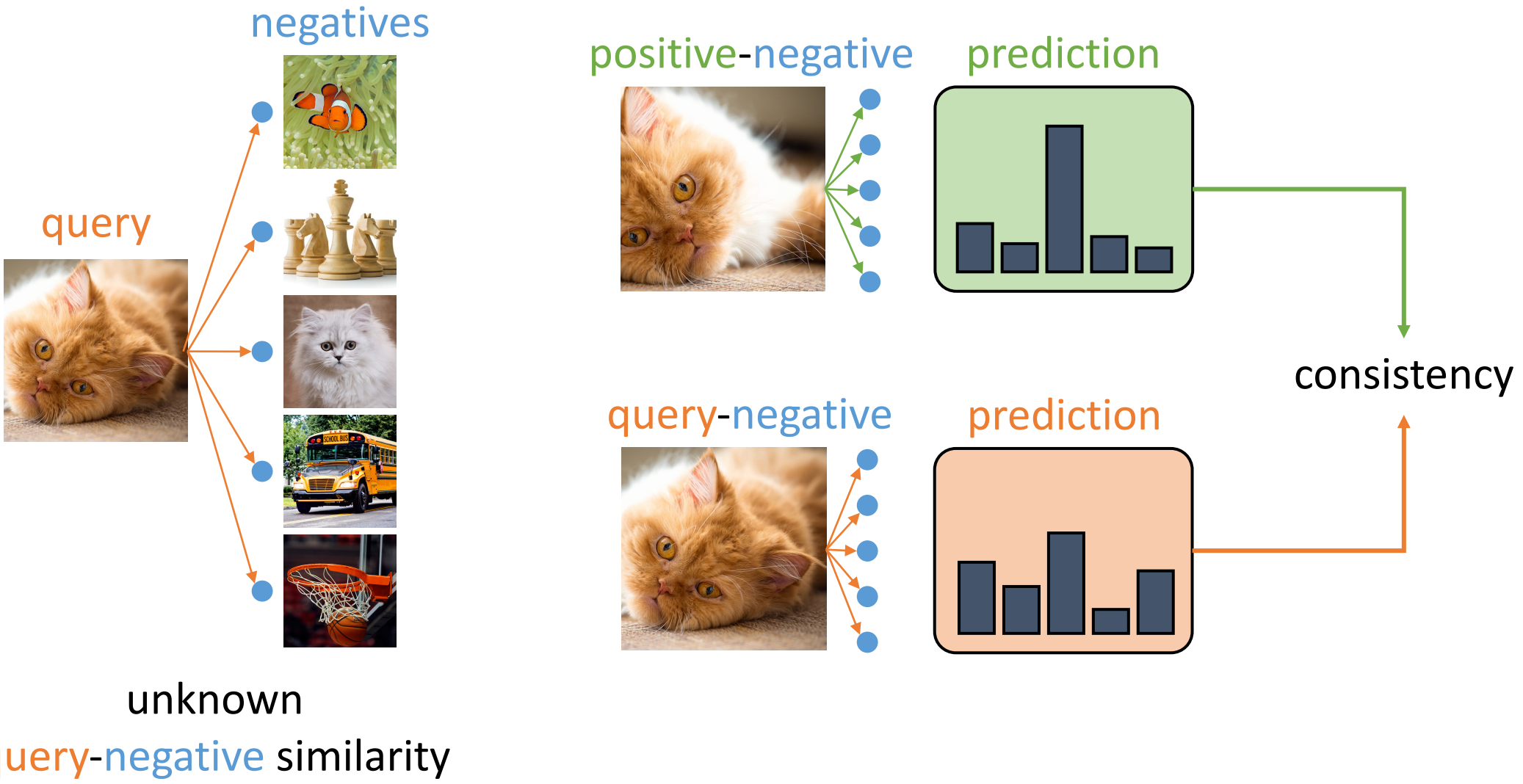
Inspiration: Consistency Regularization



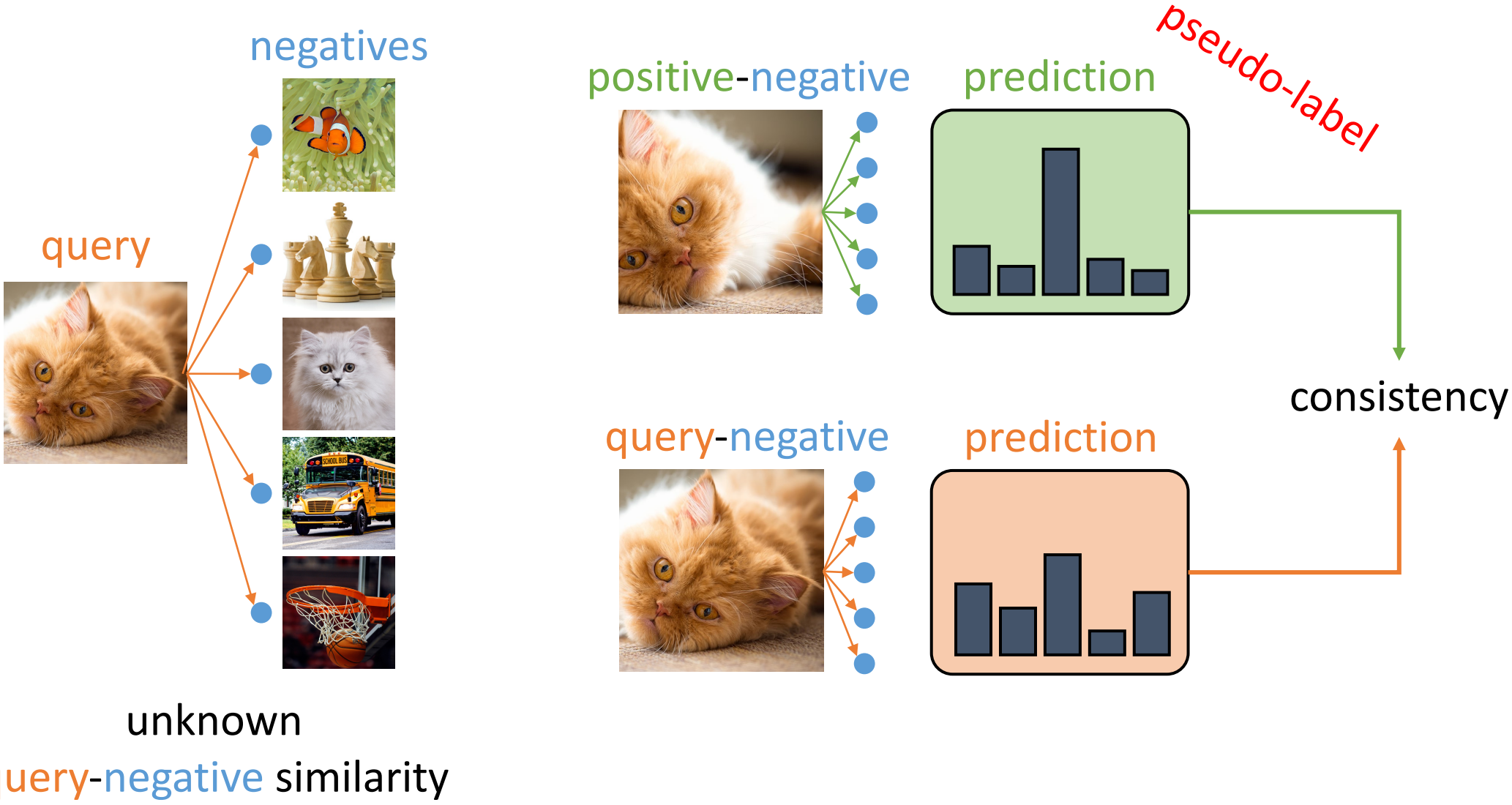
…effective in **semi**-supervised learning

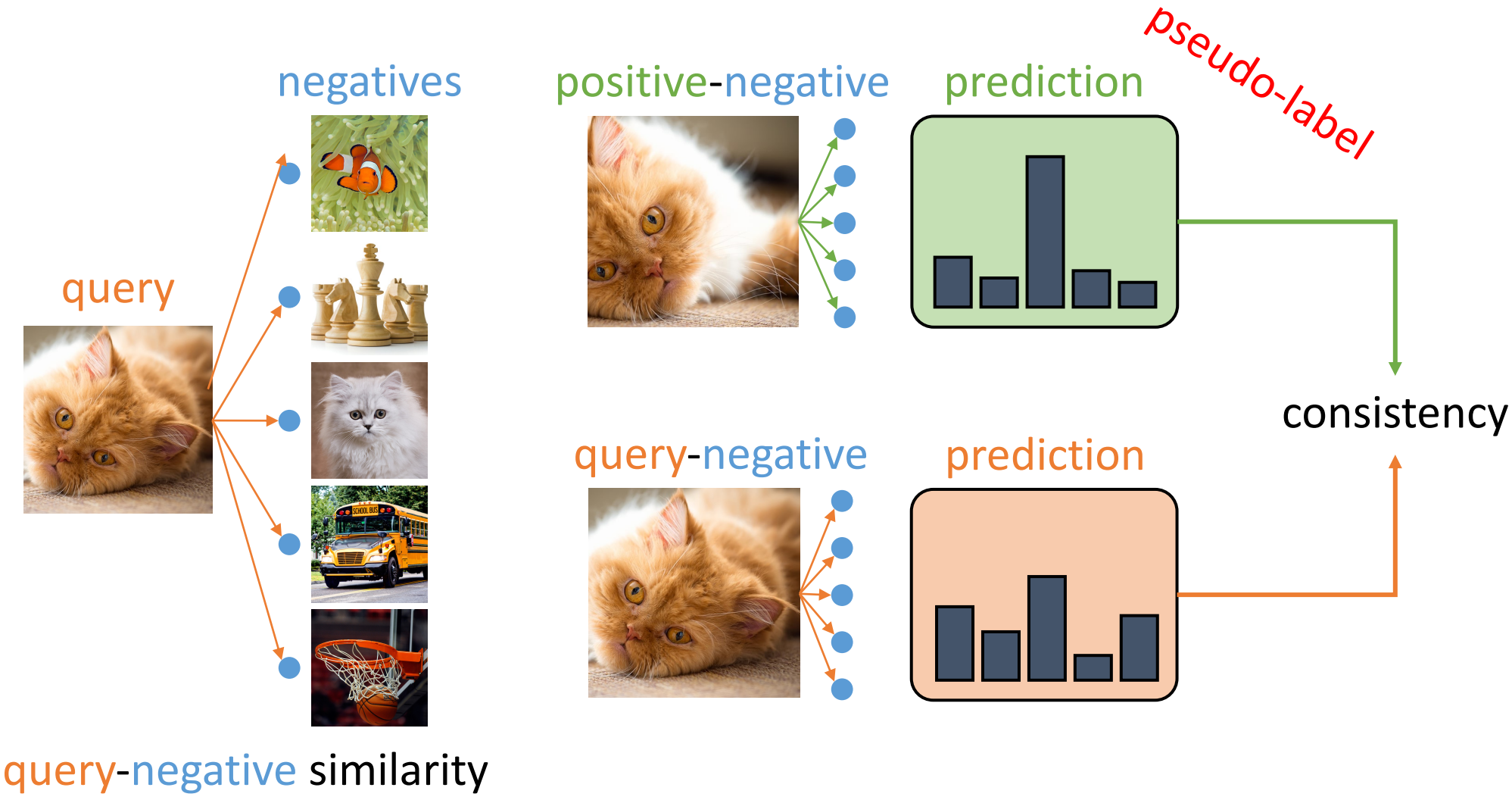
Sohn et al., NeurIPS 2020…

Our Method: Consistent Contrast (CO2)

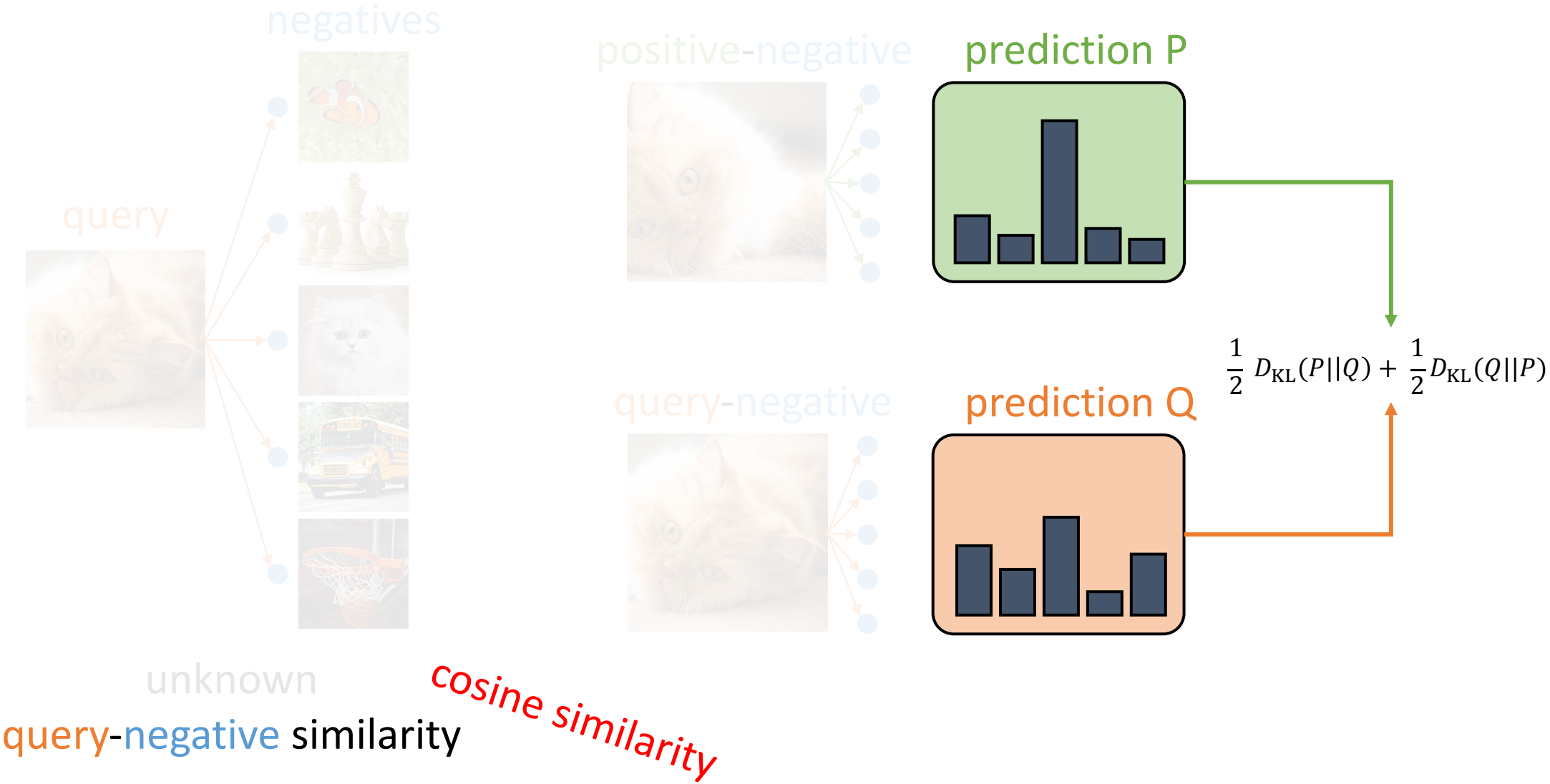


Our Method: Consistent Contrast (CO2)

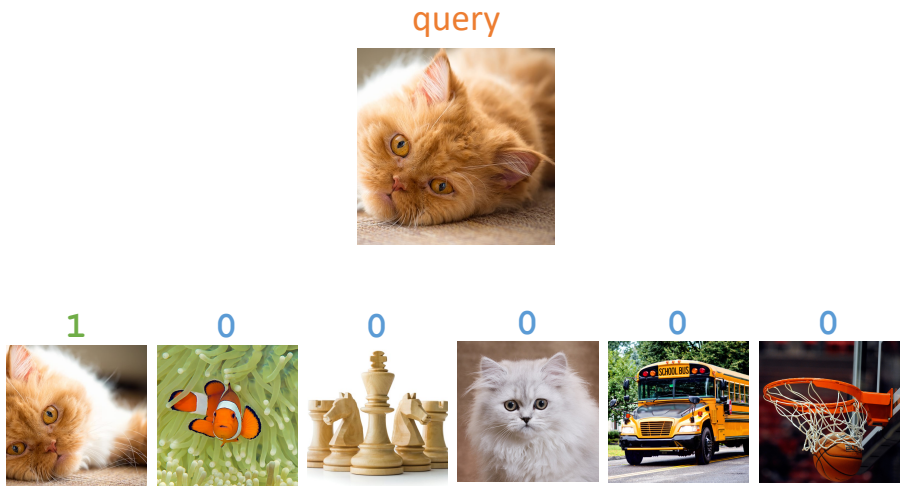



negatives
positive-negative
prediction
pseudo-label
query
consistency
query-negative
prediction
query-negative similarity

Details

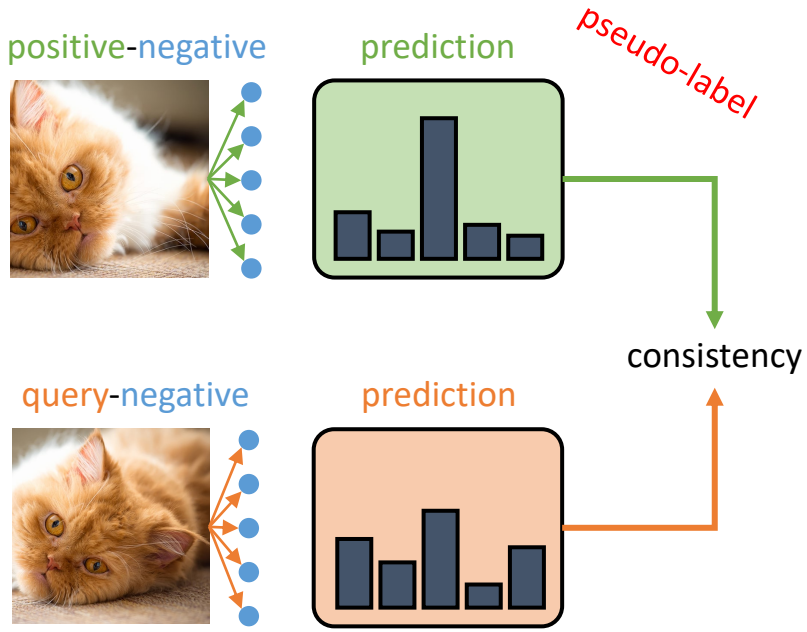

negatives
query
positive-negative
prediction P
query-negative
prediction Q
$\frac{1}{2} D_{\mathrm{KL}}(P||Q) + \frac{1}{2} D_{\mathrm{KL}}(Q||P)$
unknown
cosine similarity
query-negative similarity

Comparison



Instance Discrimination

- Artificial one-hot label
- Ignoring query-negative similarity
- Self-supervised



Consistent Contrast

- Soft pseudo-label
- Learning query-negative similarity by consistency
- Self-semi-supervised

Combination

Instance Discrimination

- Artificially created one-hot label
- Ignoring query-negative similarity
- Self-supervised

+

α · Consistent Contrast

- Soft pseudo-label
- Learning query-negative similarity by consistency
- Self-semi-supervised

Results: Linear Classifier

Pretext Task	Arch.	Head	#epochs	Top-1 Acc. (%)
MoCo	R50	Linear	200	60.6
MoCo + CO2	R50	Linear	200	63.5 (**+2.9**)
MoCo v2	R50	MLP	200	67.5
MoCo v2 + CO2	R50	MLP	200	68.0 (**+0.5**)

ResNet50
ImageNet-1K

Pretext Task	Arch.	Head	#epochs	Top-1 Acc. (%)
MoCo	R18	MLP	200	63.1
MoCo + CO2	R18	MLP	200	69.7 (**+6.6**)
SimCLR	R18	MLP	200	68.9
SimCLR + CO2	R18	MLP	200	72.3 (**+3.4**)

ResNet18
ImageNet-100

He et al., CVPR 2020

Results: Linear Classifier

Pretext Task	Arch.	Head	#epochs	Top-1 Acc. (%)
MoCo	R50	Linear	200	60.6
MoCo + CO2	R50	Linear	200	63.5 (**+2.9**)
MoCo v2	R50	MLP	200	67.5
MoCo v2 + CO2	R50	MLP	200	68.0 (**+0.5**)

ResNet50
ImageNet-1K

Pretext Task	Arch.	Head	#epochs	Top-1 Acc. (%)
MoCo	R18	MLP	200	63.1
MoCo + CO2	R18	MLP	200	69.7 (**+6.6**)
SimCLR	R18	MLP	200	68.9
SimCLR + CO2	R18	MLP	200	72.3 (**+3.4**)

ResNet18
ImageNet-100

Chen et al., ICML 2020

Results: Transfer Learning

	Image Classification	Object Detection	Semantic Segmentation
Pretext Task	mAP	AP_{50}	mIoU
MoCo	79.7	81.6	72.6
MoCo + CO2	82.6 (**+2.9**)	81.9 (**+0.3**)	73.3 (**+0.7**)
MoCo v2	85.0	82.4	74.2
MoCo v2 + CO2	85.2 (**+0.2**)	82.7 (**+0.3**)	74.7 (**+0.5**)

ResNet50 ImageNet-1K

↓

VOC 2007

Results: Semi-Supervised Learning

Pretext Task	1% Labels (Top-5)	10% Labels (Top-5)
MoCo	62.4	84.1
MoCo + CO2	66.2 (**+3.8**)	85.2 (**+1.1**)
MoCo v2	69.5	85.1
MoCo v2 + CO2	70.6 (**+1.1**)	85.4 (**+0.3**)

ResNet50
ImageNet-1K

More broadly…

- Rethinking the use of artificial labels…
  - Do they have to be known and clean?
  - Or they can also be noisy, partially-available…