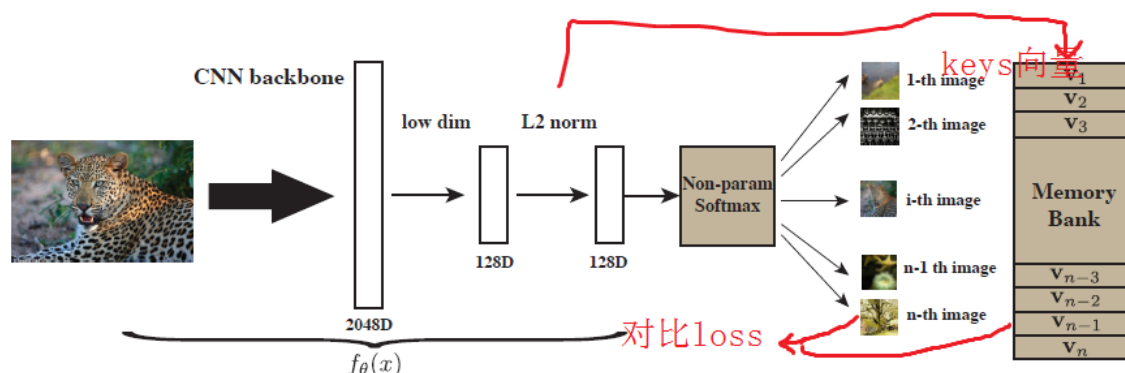


Momentum Contrast for Unsupervised Visual Representation Learning

Paper <https://arxiv.xilesou.top/abs/1911.05722>

无监督视觉表示学习



无监督视觉表示学习的核心是要从样本中学习得到特征向量集，相当于NLP中embedding后的字典。基本的模型框架：CNN网络将图片编码成query向量，在同Keys向量集比较，计算对比loss，从而更新网络参数和keys向量集；目标是学习到无标签图片本身的特征，使其显著区别于其他类别的图片。

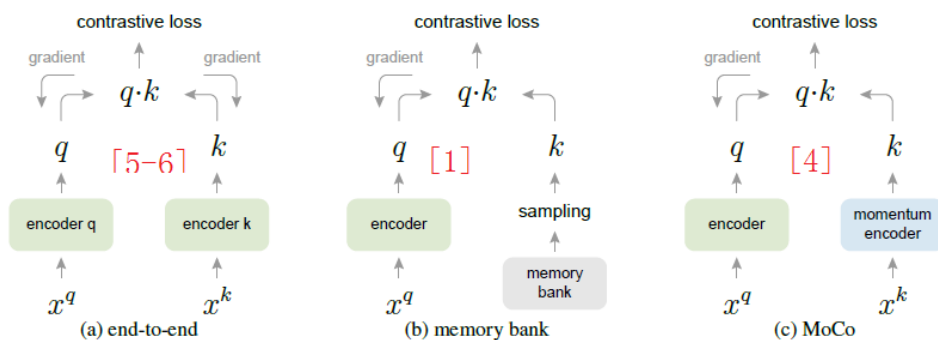
近3年来该方向比较重要的论文有：

- [1] **Unsupervised Feature Learning via Non-Parametric Instance Discrimination**-2018
- [2] CPC-Representation Learning with Contrastive Predictive Coding-2019
- [3] Contrastive Multiview Coding-2020
- [4] **MoCo-Momentum Contrast for Unsupervised Visual Representation Learning**-2020
- [5] **SimCLR-A Simple Framework for Contrastive Learning of Visual Representations**-2020
- [6] **SimCLR_v2-Big Self-Supervised Models are Strong Semi-Supervised Learners**-2020

可以归纳为三个技术点：1) 丰富同源图片的视图；2) 优化对比loss；3) **更好的学习keys向量集**

MoCo方法

聚焦于“更好的学到keys向量集”的问题，探讨了keys向量更新的机制，提出了基于动量更新（MoCo）keys向量的方法，有效解决了keys向量“滞后”导致特征不一致的问题。



- 通过维持一个队列来存储keys字典，解耦字典大小和mini-batches；
- 动量更新keys向量： $m=0.999$ 说明缓慢的迭代key编码向量很重要

$$\theta_k \leftarrow m\theta_k + (1 - m)\theta_q.$$

对比损失：

$$\mathcal{L}_q = -\log \frac{\exp(q \cdot k_+ / \tau)}{\sum_{i=0}^K \exp(q \cdot k_i / \tau)}$$

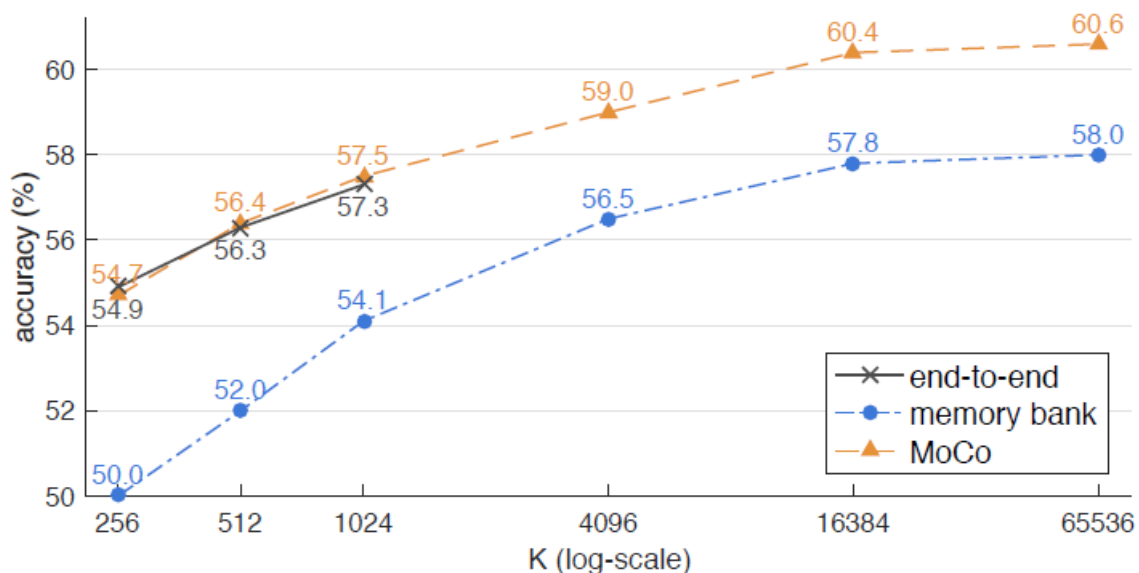
BN细节：

在编码key向量时，分配给多gpu前shuffle样本的顺序，编码后还原。query向量样本顺序不变，保证了batch统计时，计算query和positive key的样本是不同的数据子集。

试验结果

比较了三种对比机制的效果，K是负样本数量

- 端到端的机制在K比较小时和MoCo差不多
- memory bank的方式支持大字典，性能比MoCo差2.6%



动量超参数m的试验：K的大小取4096

- 在0.99~0.9999之间表现好，说明MoCo缓慢进化keys很重要

momentum m	0	0.9	0.99	0.999	0.9999
accuracy (%)	fail	55.2	57.8	59.0	58.9

特征迁移到下游任务的试验：

- fine-tune中做特征归一化；
- 设计好的fine-tune schedule可以到达有监督的预训练模型一样的效果

pre-train	AP ₅₀				AP MoCo	AP ₇₅	
	RelPos, by [14]	Multi-task [14]	Jigsaw, by [26]	LocalAgg [66]		Multi-task [14]	MoCo
super. IN-1M	74.2	74.2	70.5	74.6	74.4	44.3	42.7
unsup. IN-1M	66.8 (-7.4)	70.5 (-3.7)	61.4 (-9.1)	69.1 (-5.5)	74.9 (+0.5)	43.9 (-0.4)	50.1 (+7.4)
unsup. IN-14M	-	-	69.2 (-1.3)	-	75.2 (+0.8)	-	50.2 (+7.5)
unsup. YFCC-100M	-	-	66.6 (-3.9)	-	74.7 (+0.3)	-	49.0 (+6.3)
unsup. IG-1B	-	-	-	-	75.6 (+1.2)	-	51.7 (+9.0)