A Simple Framework for Contrastive Learning of Visual Representations

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A Simple Framework for Contrastive Learning of Visual Representations.

arXiv:2002.05709v3 [cs.LG] 1 Jul 2020

Primary Thesis

- presents ***SimCLR*:** a simple framework for contrastive learning of visual representations.
- contrastive self-supervised learning
- In order to understand what enables the contrastive prediction tasks to learn useful representations

Hypotheses

- data augmentations plays a critical role in defining effective predictive tasks
- learnable nonlinear transformation between the representation and the contrastive loss substantially improves the quality of the learned representations
- contrastive learning benefits from larger batch sizes and more training steps

Methods and Measures used

- **The Contrastive Learning Framework**
    - stochastic data augmentation module
    - base encoder f (·)
    - projection head g(·)
    - **contrastive loss function**:sim(u, v) = $u^\top v/\|u\|\|v\|$ de- note the dot product between l_2 normalized u and v
    - base network structure
- **Training with Large Batch Size**

Key Findings

- Composition of data augmentation operations is crucial for learning good representations
    - no single transformation suffices to learn good representations
    - One composition of augmentations *stands out*: **random cropping** and **random color distortion**
- Contrastive learning needs stronger data augmentation than supervised learning
- Unsupervised contrastive learning benefits (more) from bigger models
- A nonlinear projection head improves the representation quality of the layer before it

- Normalized cross entropy loss with adjustable temperature works better than alternatives
- Contrastive learning benefits (more) from larger batch sizes and longer training

Dataset

- ImageNet ILSVRC-2012
- CIFAR-10