Unsupervised Pre-Training of Image Features on Non-Curated Data

Summary

Research Objective

- Bridge the gap between unsupervised learning on the curated dataset and on the raw dataset.
- Capture complementary statistics from large scale of data via combining classification and clustering

Problem Statement(What is the problem to be solved?)

- Convnets on pretrained data perform well but collecting large curated dataset is effort-costing
- Simply discarding labels doesn't undo the effect of the effort of collecting curated dataset
- Previous unsupervised learning are trained on curated dataset
- Cluster relying on inter-image similarities are sensitive to data distribution

Self-supervised Learning via cluster

Methods

- Automatically generates targets by clustering the features of the entire dataset, under constraints derived from self-supervision.
- propose a hierarchical formulation that is suitable for distributed training.
-

Evaluation(How to evaluate this method?)

Conclusion(Strong or weak conclusion)

References

1. Mathilde Caron, Piotr Bojanowski, Armand Joulin, and Matthijs Douze. Deep clustering for unsupervised learning of visual features. In Proceedings of the European Confer- ence on Computer Vision (ECCV), 2018

2. Spyros Gidaris, Praveer Singh, and Nikos Komodakis. Un- supervised representation learning by predicting image rota- tions. In International Conference on Learning Representa- tions (ICLR), 2018.

3. Mehdi Noroozi, Ananth Vinjimoor, Paolo Favaro, and Hamed Pirsiavash.
   Boosting self-supervised learning via knowledge transfer. In Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), 2018.