A comparison of representation evaluation scores

1 Background

Representation learning is the task of extracting features from a dataset that can retain most of the information while being simpler to manipulate, supporting thus ideally multiple downstream tasks.

One of the most common approaches in the recent years to representation learning is to use self-supervised learning, as demonstrated by the seminal works of SIMCLR (Chen et al., 2020) and DINO (Caron et al., 2021). Visualisation methods, e.g. t-SNE (Maaten and Hinton, 2008) or UMAP (McInnes et al., 2018), can also be seen as representation learning methods because they provide an alternative projection of the dataset.

However, and particularly in the case of self-supervised learning, assessing the quality of the learnt features is a challenging task. It is often done by judging how well downstream tasks. While providing some insights, this evaluation approach is limited because (i) it relies on a (stochastic) model to assess another model, and (ii) the relevance of the downstream tasks is somewhat subjective to the dataset, which can raise implicit biases upon evaluation.

To judge fairly representations, lines of research for the evaluation of the representations *per se* have emerged. Their idea is to directly score the representation. Examples of such scores include: RankMe (Garrido et al., 2023), CLID (Lu et al., 2023), and the alignment/uniformity scores (Wang and Isola, 2020).

The goal of the thesis would be to explore how this different scores behave and apply on different image datasets and pre-trained models. Ideally, managing to create a competitive representation scoring method would be a plus.

2 Data description

The datasets will mainly be the classical deep learning datasets, e.g. CIFAR10, MNIST, ImageNet, etc. More can be explored depending on their accessibility and the availability of matching pre-trained models.

Smaller datasets could be used for comparing and assessing these scores on visualisation techniques.

3 Research questions

First sketches of research questions are, though not limited to:

1. Are the representation scores good predictors of the usefulness of representations for downstream tasks?
2. Are representation scores adapted to evaluate visualization techniques?

4 Eligibility criteria

- Strong background in statistics, linear algebra and high-dimensional geometry.
- Proficiency in Python is a must in order to cover the thesis in due time.
- A strong background in deep learning is highly recommended.
- Knowledge of self-supervised learning algorithms is a merit.

5 Contact person

Please reach out `louis.ohl@liu.se` if interested.

References

Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9650–9660, 2021.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A Simple Framework for Contrastive Learning of Visual Representations. In *International Conference on Machine Learning*, pages 1597–1607. PMLR, 2020.

Quentin Garrido, Randall Balestriero, Laurent Najman, and Yann Lecun. RankMe: Assessing the Downstream Performance of Pretrained Self-Supervised Representations by Their Rank. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 10929–10974. PMLR, July 2023. URL `https://proceedings.mlr.press/v202/garrido23a.html`.

Yuchen Lu, Zhen Liu, Aristide Baratin, Romain Laroche, Aaron Courville, and Alessandro Sordoni. Using Representation Expressiveness and Learnability to Evaluate Self-Supervised Learning Methods. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856. URL `https://openreview.net/forum?id=BxdrpnRHNh`.

Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-SNE. *Journal of machine learning research*, 9(Nov):2579–2605, 2008.

Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

Tongzhou Wang and Phillip Isola. Understanding Contrastive Representation Learning through Alignment and Uniformity on the Hypersphere. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 9929–9939. PMLR, July 2020. URL `https://proceedings.mlr.press/v119/wang20k.html`.