



北京大学

博士研究生学位论文

题目： 面向大规模机器学习的
深度表示学习方法

姓 名： 张博
学 号： 2001234567
院 系： 计算机学院
专 业： 计算机软件与理论
研究方向： 机器学习
导 师： 李明 教授

☒ 学术学位 ☐ 专业学位

二〇二六年六月

版权声明

任何收存和保管本论文各种版本的单位和个人，未经本论文作者同意，不得将本论文转借他人，亦不得随意复制、抄录、拍照或以任何方式传播。否则，引起有碍作者著作权之问题，将可能承担法律责任。

摘要

深度表示学习是现代机器学习的基础。随着数据规模和模型规模的增长，如何在有限的计算资源下高效地学习可迁移的表示，成为一个重要问题。本文研究大规模场景下的表示学习方法，主要贡献有三点：第一，提出一种基于记忆库的对比学习方法，在不增大批量的情况下获得大量负样本；第二，提出一种分层梯度压缩的分布式训练方案，将通信量降低约百分之六十；第三，在图像和文本两类任务上系统评估了所学表示的迁移性能。实验表明，在相同计算预算下，本文方法的线性评估准确率比基线提高了二点七个百分点。

关键词：表示学习，对比学习，大规模训练，分布式优化

Deep Representation Learning Methods for Large-Scale Machine Learning

Zhang Bo (Computer Software and Theory)

Directed by Prof. Li Ming

ABSTRACT

Deep representation learning underlies modern machine learning. As data and models grow, learning transferable representations efficiently under a fixed compute budget becomes a central problem. This thesis studies representation learning at scale and makes three contributions: a contrastive method with a memory bank that supplies many negatives without large batches; a distributed training scheme with hierarchical gradient compression that reduces communication by about 60%; and a systematic evaluation of transfer to image and text tasks. At equal compute, the proposed method improves linear-evaluation accuracy by 2.7 percentage points over the baseline.

KEY WORDS: representation learning, contrastive learning, large-scale training, distributed optimization

目录

第一章 绪论	1
第二章 相关工作	3
第三章 方法	5
第四章 实验	7
第五章 总结与展望	9
附录 A 实验设置	13
致谢	15
北京大学学位论文原创性声明和使用授权说明	17

第一章 绪论

表示学习的目标是从原始数据中学习有用的特征。自监督方法利用数据自身构造学习任务，不依赖人工标注，近年来在图像和文本上都取得了显著进展^[1-2]。然而，这些方法往往需要很大的批量或很长的训练时间，在计算资源有限时难以使用。本文围绕“在有限预算下高效学习可迁移表示”这一问题展开研究。

第二章 相关工作

对比学习通过拉近同一样本不同视图的表示、推远不同样本的表示来学习特征。常用的损失函数为

$$\mathcal{L} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_k \exp(\text{sim}(z_i, z_k)/\tau)}, \quad (2.1)$$

其中 τ 为温度参数。负样本的数量对效果影响很大，动量对比方法使用队列维护大量负样本^[2]。

第三章 方法

本文使用一个按类别簇组织的记忆库提供负样本，并在分布式训练中对梯度进行分层压缩：节点内先做全精度规约，节点间只传输压缩后的梯度。

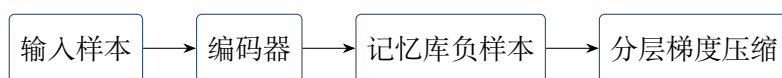


图 3.1 本文方法的整体流程

第四章 实验

表 4.1 给出了相同计算预算下的线性评估结果。

表 4.1 相同计算预算下的线性评估准确率 (%)

方法	准确率	通信量 (相对值)
基线方法	71.4	1.00
本文方法	74.1	0.41

第五章 总结与展望

本文提出的记忆库对比学习和分层梯度压缩，使得在有限资源下训练高质量表示成为可能。未来工作包括将方法扩展到多模态数据。

参考文献

- [1] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations[C]//Proceedings of the 37th International Conference on Machine Learning. 2020: 1597-1607.
- [2] HE K, FAN H, WU Y, et al. Momentum contrast for unsupervised visual representation learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 9729-9738.

附录 A 实验设置

所有实验均在 32 张加速卡上进行，训练 200 轮，使用余弦学习率调度。

致谢

感谢导师李明教授在研究和生活上给予的悉心指导。感谢实验室的各位同学在讨论和实验中提供的帮助。感谢家人一直以来的理解与支持。

北京大学学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文作者签名： 日期： 年 月 日

学位论文使用授权说明

(必须装订在提交学校图书馆的印刷本)

本人完全了解北京大学关于收集、保存、使用学位论文的规定，即：

- 按照学校要求提交学位论文的印刷本和电子版本；
- 学校有权保存学位论文的印刷本和电子版，并提供目录检索与阅览服务，在校园网上提供服务；
- 学校可以采用影印、缩印、数字化或其它复制手段保存论文；
- 因某种特殊原因须要延迟发布学位论文电子版，授权学校 ☐ 一年 / ☐ 两年 / ☐ 三年以后，在校园网上全文发布。

(保密论文在解密后遵守此规定)

论文作者签名： 导师签名： 日期： 年 月 日