

# 基于元维度掩码的自监督表示学习方法

MetaMask: Revisiting Dimensional Confounder for Self-Supervised Learning

李江梦, 强文文, 张雅楠, 莫文怡, 郑昌文, 苏冰, 熊辉  
NeurIPS 2022, **Spotlight**, jiangmeng2019@iscas.ac.cn

## 研究动机

首先, 进行了探索性实验, 并实验结果中发现一些容易被研究者们忽略的维度层次自监督方法特性。

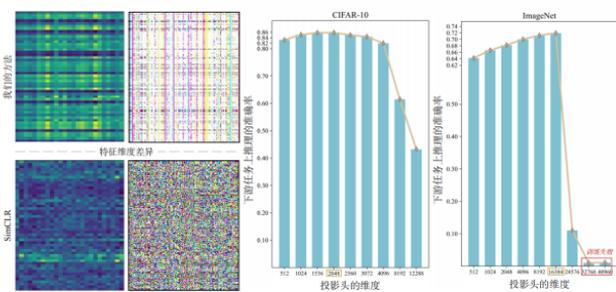


图 1: (a) 基于传统自监督学习方法 (SimCLR) 以及使用冗余减少技术的方法学习到的表示的可视化。(b) 使用一种特定的冗余减少技术的自监督表示学习方法 (Barlow Twins) 进行实验。实验增加了投影头的维度, 使表示维度能够拟合更多不同的信息。

(a) 观察: 使用维度冗余减少的模型学习到的表示比传统的自监督模型更加“有序”。

结论: 维度冗余减少鼓励学习表示的维度来对解耦信息进行建模, 自然地避免了维度冗余问题。

(b) 观察: Barlow Twins的性能随着维度的增加而逐渐下降, 直至达到峰值后崩溃。这一现象与维数灾难一致。

结论: 与任务相关的语义信息是有限的, 随着不同维度捕获的信息逐渐解耦, 更多与任务无关的噪声信息被编码到表示中, 即维度混杂因素增加, 因此, 模型的性能会变差。

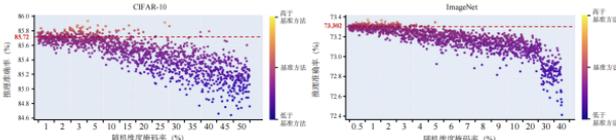
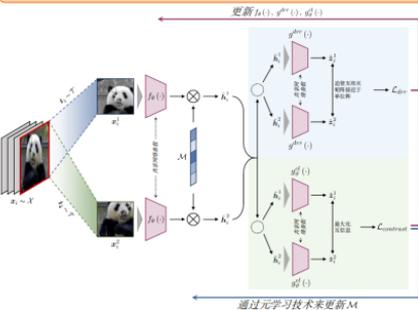


图 2: Barlow Twins 获得的随机屏蔽尺寸的实验散点图, 其中红色虚线表示未屏蔽表示所实现的性能。每个点表示通过使用特定掩码率随机掩码原始表示而获得的独立实验结果。在实验过程中, 原始表示是固定的。

观察: 具有某些特定屏蔽维度的表示确实可以比未屏蔽的表示获得更好的性能。

结论: 进一步证明了维度混杂因素的存在, 因为这种维度混杂因素可能会干扰模型的推理。

## 方法描述



为了得到能够兼具维度去冗以及对维度混杂因素鲁棒的自监督表示, 采取了基于元学习的自适应维度掩码以及维度去冗正则项来进行自监督表示学习模型的训练。

## 理论分析

基于对泛化误差界的分析, 证明了相比于典型对比学习方法, MetaMask可以得到更紧致的下游分类误差界。

$$\begin{aligned} \mathcal{L}_{contrast}(g_{\theta}^{cl}(\bar{h})) - \sqrt{\Phi(g_{\theta}^{cl}(\bar{h})|y)} - \frac{1}{2} \cdot \sum_{k \in [1, D]} \sqrt{\Phi(|g_{\theta}^{cl}(\bar{h})|^k|y)} - \mathcal{O}(M^{-\frac{1}{2}}) \\ \leq \mathcal{L}_{crossentropy}(h) + \log\left(\frac{M}{D}\right) \leq \mathcal{L}_{contrast}(g_{\theta}^{cl}(\bar{h})) + \sqrt{\Phi(g_{\theta}^{cl}(\bar{h})|y)} + \mathcal{O}(M^{-\frac{1}{2}}) \end{aligned}$$

## 实验验证

基于经验归纳实验证明了MetaMask在多样的数据集上对不同自监督模型具有一致性能提升, 且对维数灾难鲁棒。

| 模型                           | CIFAR-10               | CIFAR-100              | STL-10                 | IN-200                 |
|------------------------------|------------------------|------------------------|------------------------|------------------------|
| Barlow Twins <sup>[64]</sup> | 85.72                  | 60.31                  | 73.62                  | 45.10                  |
| + MetaMask <sub>-1.25</sub>  | 87.53 <sub>+1.81</sub> | 61.42 <sub>+1.11</sub> | 73.97 <sub>+0.35</sub> | 46.81 <sub>+1.71</sub> |
| SimCLR <sup>[33]</sup>       | 81.73                  | 58.27                  | 70.09                  | 44.46                  |
| + MetaMask <sub>-3.33</sub>  | 86.01 <sub>+4.28</sub> | 61.03 <sub>+2.76</sub> | 74.90 <sub>+4.81</sub> | 45.87 <sub>+1.41</sub> |
| BYOL <sup>[25]</sup>         | 88.05                  | 60.94                  | 72.04                  | 46.72                  |
| + MetaMask <sub>-0.28</sub>  | 87.53 <sub>+0.52</sub> | 61.42 <sub>+0.48</sub> | 73.97 <sub>+1.93</sub> | 46.81 <sub>+0.11</sub> |
| DCL <sup>[54]</sup>          | 85.61                  | 59.29                  | 71.18                  | 45.77                  |
| + MetaMask <sub>+1.65</sub>  | 86.47 <sub>+0.86</sub> | 60.82 <sub>+1.53</sub> | 74.79 <sub>+3.61</sub> | 46.35 <sub>+0.58</sub> |
| HCL <sup>[53]</sup>          | 85.27                  | 61.21                  | 71.92                  | 46.90                  |
| + MetaMask <sub>-0.95</sub>  | 85.97 <sub>+0.70</sub> | 61.63 <sub>+0.42</sub> | 74.03 <sub>+2.11</sub> | 46.06 <sub>+0.55</sub> |
| NNCLR <sup>[184]</sup>       | 81.44                  | 61.64                  | 71.89                  | 47.10                  |
| + MetaMask <sub>+1.63</sub>  | 85.78 <sub>+4.34</sub> | 61.89 <sub>+0.25</sub> | 74.02 <sub>+2.13</sub> | 46.90 <sub>+0.20</sub> |

