

基于对比学习的标签带噪图像分类^①



李俊哲, 曹 国

(南京理工大学 计算机科学与工程学院, 南京 210094)

通信作者: 曹 国, E-mail: caoguo@njust.edu.cn

摘 要: 标签噪声会极大地降低深度网络模型的性能. 针对这一问题, 本文提出了一种基于对比学习的标签带噪图像分类方法. 该方法包括自适应阈值、对比学习模块和基于类原型的标签去噪模块. 首先采用对比学习最大化一幅图像的两个增强视图的相似度来提取图像鲁棒特征; 接下来通过一种新颖的自适应阈值过滤训练样本, 在模型训练过程中根据各个类别的学习情况动态调整阈值; 然后创新性地引入基于类原型的标签去噪模块, 通过计算样本特征向量与原型向量的相似度更新伪标签, 从而避免标签中噪声的影响; 在公开数据集 CIFAR-10、CIFAR-100 和真实数据集 ANIMAL10 上进行对比实验, 实验结果表明, 在人工合成噪声的条件下, 本文方法实验结果均高于常规方法, 通过计算图像鲁棒的特征向量与各个原型向量的相似度更新伪标签的方式, 降低了噪声标签的负面影响, 在一定程度上提高模型的抗噪声能力, 验证了该模型的有效性.

关键词: 深度学习; 标签噪声; 对比学习; 类原型; 自适应阈值

引用格式: 李俊哲, 曹国. 基于对比学习的标签带噪图像分类. 计算机系统应用, 2023, 32(12): 104-111. <http://www.c-s-a.org.cn/1003-3254/9339.html>

Label Noisy Image Classification Based on Contrastive Learning

LI Jun-Zhe, CAO Guo

(School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: Label noise can greatly reduce the performance of deep network models. To address this problem, this study proposes a contrastive learning-based label noisy image classification method. The method includes an adaptive threshold, contrastive learning module, and class prototype-based label denoising module. Firstly, the robust features of the image are extracted by maximizing the similarity between two augmented views of the same image using contrastive learning. Then, a novel adaptive threshold filtering training sample is used to dynamically adjust the threshold based on the learning status of each class during model training. Finally, a class prototype-based label denoising module is introduced to update pseudo-labels by calculating the similarity between sample feature vectors and prototype vectors, thus avoiding the influence of label noise. Comparative experiments are conducted on the publicly available datasets CIFAR-10 and CIFAR-100 and the real dataset ANIMAL10. The experimental results show that under the condition of artificially synthesized noise, the proposed method outperforms conventional methods. By updating pseudo-labels based on the similarity between the robust feature vector of the image and various prototype vectors, the negative impact of noisy labels is reduced, and the anti-noise ability of the model is improved to certain extent, verifying the effectiveness of the proposed model.

Key words: deep learning; label noise; contrastive learning; class prototype; adaptive threshold

① 基金项目: 国家自然科学基金 (62201282); 江苏省自然科学基金 (BK20231456)

收稿时间: 2023-06-21; 修改时间: 2023-07-19; 采用时间: 2023-07-27; csa 在线出版时间: 2023-10-20

CNKI 网络首发时间: 2023-10-23

1 引言

深度神经网络的成功很大程度上取决于正确标记的大规模训练数据集。然而,这些大规模数据集的收集成本和时间成本都很高。为了缓解这一问题,人们常通过以下方式获取标签数据:非专业标注人员标注、众包平台或其他来源获取标签。通过这些方法获取的数据标签可以降低成本,但不可避免地会带来标签噪声。如图1所示,对标注人员来说从巨嘴鸟中识别犀鸟是非常困难的。



图1 巨嘴鸟与犀鸟

因此如何从带噪声标签的数据集中学习逐渐成为深度学习的一项重要研究^[1]。目前基于深度学习的噪声标签算法可以归纳为以下4类:基于数据的噪声标签算法、基于损失函数的噪声标签算法、基于训练方式的标签噪声学习方法、基于模型的标签噪声算法。

首先是基于数据的标签噪声学习算法。最直观的想法就是找出疑似标注错误的样本,通过更改样本权重或者修改样本标签的方式来改善模型学习效果。如张增辉等^[2]提出的动态概率抽样方法,通过统计样本的标签置信度对样本分区。陈庆强等^[3]采用欧氏距离度量样本分布的密度来划分不同的区域,按照分区采用不同的规则过滤标签噪声。这类方法主要通过分区对样本进行重加权的方式过滤噪声,但是需要手动设计权重函数和超参,因此很难实际应用。标签重修正通过对样本标签进行一定程度地修改来提高分类器的性能。例如余孟池等^[4]利用训练好的基网络预测噪声标签。基于教师-学生网络的架构对标签进行软化,使得学生网络能够在带噪学习中更加鲁棒^[5-8]。但是这类算法在高噪声率的情况下,容易将标签拟合到错误样本上,导致分类器效果不理想。

基于损失函数的标签噪声算法中关于损失函数的设计分为两个方面:一方面单独针对损失函数进行改进。例如损失校正法。损失校正的方法需要首先估计标签转移矩阵^[9,10],然后通过修改训练期间的损失函数来提高鲁棒性。前向传播和后向校正是两种基于标签转

移矩阵的损失校正方法^[11]。但是这种方法对于标签转移矩阵非常敏感,现实中的标签转移矩阵也很难估计。另一方面是向损失函数中加入正则化来提高鲁棒性。

基于训练方式的标签噪声算法通过将半监督方法引入标签噪声模型中,改变模型的训练方法来缓解标签噪声对模型分类的影响。Self-training 算法^[12]可以提高模型的鲁棒性。如 Jiang 等^[13]提出的 Mentor 是利用数据驱动来训练基础网络的一种方法,但是在样本不均衡且含有噪声的数据集上效果不理想。Self-training 算法同时容易出现错误累加^[14]。Co-training 算法可以避免这种情况。在 Li 等^[15]提出的 DivideMix 中,采用高斯混合模型来判别噪声样本,运用两个网络划分数据集其性能比利用单个网络划分数据集的效果更佳。但是随着训练次数的增加,Co-training 算法在效果上会退化成 Self-training 算法中的 MentorNet。

基于模型的标签噪声算法,通过修改模型架构来更好地建模噪声数据的转移矩阵。Sukhbaatar 等^[16]提出了自顶向下和自底向上两种噪声模型。通过这两种模型可以较为准确地从数据中学习噪声分布,提高神经网络的性能。Goldberger 等^[17]通过使用附加的 Softmax 层模拟期望最大算法优化的似然函数,提高了模型的普适性。但是这种方法在每次预测后都要重新训练模型。Han 等^[18]用人类的经验为先验知识,利用人类对无效标签转换的认知,可以对噪声进行约束建模。由于是人工设计的,因此不具备噪声适配层的可扩展性。对模型进行正则化处理,在不改变模型参数量的前提下能够有效地提高模型的泛化能力。Li 等^[19]提出一种噪声容忍训练算法,首先进行元学习更新,之后在进行传统的梯度更新。Jindal 等^[20]添加了一个应用了 dropout 正则化的额外的 Softmax 层,认为它可以提供更鲁棒的训练,以此防止由于 dropout^[21]随机性而产生的记忆噪声。但是对模型进行正则化可能会降低模型的表达能力,使其无法充分拟合数据集。并且正则化技术通常需要调整超参数,需要花费大量的时间和精力。

目前研究表明,对比学习模型可以获取图像鲁棒的特征表示,不需要人工参与且泛化性好,在高噪声率下依旧拥有良好的性能。目前使用的对比学习方法^[22-25]需要通过最小化正样本对之间的距离和最大化负样本对之间的距离来实现。利用对比学习拉近相似样本拉远相斥样本的特性,Wang 等^[26]将对比学习引入噪声标签图像分类的工作中,产生了很好的分类效果。Yao

等^[27]、Zhang 等^[28]也基于对比学习进行了相关研究. 这些对比学习方法得益于大量的负样本. 但是这些方法区分正负样本对的途径是不可靠的, 可能会引入错误的样本对, 从而影响对比学习得到的特征表示. 本文采用一种较新的对比学习方法 SimSiam^[29], 该方法直接最大化图像两个视图之间的相似性, 而不使用存储负样本的方式, 保证对比学习得到的特征表示不受影响. 本文基于 SimSiam 框架, 引入基于类原型的标签去噪策略来逐步更新伪标签. 使用自适应阈值方法根据模型学习情况自适应地调整阈值. 在多个数据集上进行对比实验, 验证了本文所提模型的有效性.

本文的主要工作如下.

(1) 本文提出一种新的对标签噪声鲁棒的图像分类方法, 该方法使用对比学习与类原型模块结合, 可以对带有标签噪声的数据进行准确分类.

(2) 使用自适应阈值方法过滤训练样本. 避免由于固定阈值设置过高, 导致模型初期训练通过阈值的样本数量过少, 参数更新缓慢, 以及对于较难学习的样本分类效果不好的问题.

(3) 在 3 个公开数据集上进行实验, 验证了本文方法的性能高于同类方法.

2 相关工作

最近的研究表明对比学习获取图像的鲁棒特征可以提高模型的抗噪声能力, 主要通过吸引正样本对, 排斥负样本对实现, 如图 2 所示.

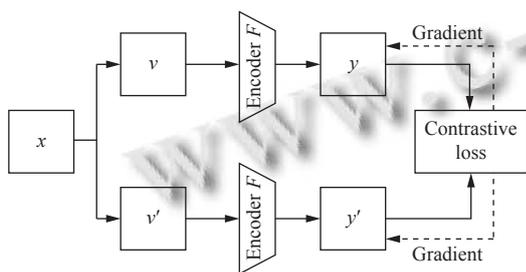


图 2 对比学习流程图

图 2 中 x 时输入图像, v 和 v' 是图像经过两次数据增强后得到的视图. 使用经过各自编码器后得到对应的特征向量计算对比学习损失. 这类方法的正样本对由同一个样本的两个视图组成, 其他数据对为负样本对. MOCO^[30]在此基础上通过维护一个负样本队列, 增大负样本的数量使模型性能得到了比较明显的提高.

Chen 等^[31]提出的 SimCLR 直接使用当前批次中共存的负样本, 取得较好的实验结果. 田浩江等^[32]在 SimCLR 的基础上对特征提取部分进行改进提高了分类精度. 这两种方法都需要较大的批次才能正常工作. 这种通过增大负样本数量的方式会使模型的训练时间增长, 并且会增加模型超参数数量. 不仅如此, 这类方法引入标签噪声学习中还会带来样本类别判断错误问题. 如 PiCO 将 MOCO 方法引入标签噪声学习中, 正负样本对的区分通过分类器的分类结果来判断, 分类器错误的分类结果无疑会影响对比学习得到的特征表示. 本文使用的 SimSiam 框架无需使用大量的负样本, 通过最大化一个图像两个视图之间的相似性来学习图像的特征表示, 网络结构如图 3 所示.

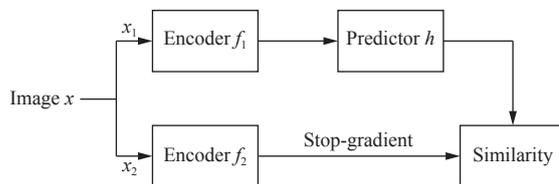


图 3 SimSiam 网络结构

图 3 中 x_1 、 x_2 为图像经过两次数据增强后的视图. 本文在此基础上加入了基于类原型的标签去噪模块, 使用自适应阈值过滤训练样本, 提高了模型的抗噪声能力.

3 CLPAT 模型

本文的标签噪声问题使用以下方式定义: 设 X 为输入样本空间, $\mathcal{Y} = \{1, 2, \dots, C\}$ 为输出标签空间. 本文考虑一个训练样本集 $D = \{(x_i, Y_i)\}_{i=1}^n$, 其中每一个元组有一个图像 x_i 和一个候选标签集 $Y_i \subset \mathcal{Y}$ 组成. 与监督学习任务相同, 本模型学习的任务是获得一个函数映射, 预测输入图像的真实标签. 但不同的是, 在标签中设置了噪声增加了更多的不确定性. 本文假设真实标签 y_i 隐藏在它的候选集中, 即 $y_i \in Y_i$, 但这对模型来说是不可见的. 因此与监督学习相比, 学习的过程存在固有的模糊性.

在本节中将详细描述基于对比学习和类原型的标签噪声学习框架, 主要有两个关键组件, 分别为对比学习提取特征模块和基于类原型的标签去噪模块, 这两部分耦合为一个整体共同训练, 取得了较好的实验结果.

3.1 对比学习模块

本模块的架构如图4所示,对图像 x 进行两次随机数据增强,得到增强视图 x_1 、 x_2 . 将这两个视图分别输入到对应的编码器 f_1 、 f_2 中处理. 这两个编码器分别由一个骨干网络 (ResNet) 组成. 使用动量编码器的方

式更新两个分支之间的参数,如式(1)所示:

$$\theta_{f_2} = m\theta_{f_1} + (1 - m)\theta_{f_2} \quad (1)$$

其中, θ 为编码器的参数. 编码器 f_1 输出的特征向量经过两个 MLP 模块后得到输出向量 p , m 为人工设置的超参数.

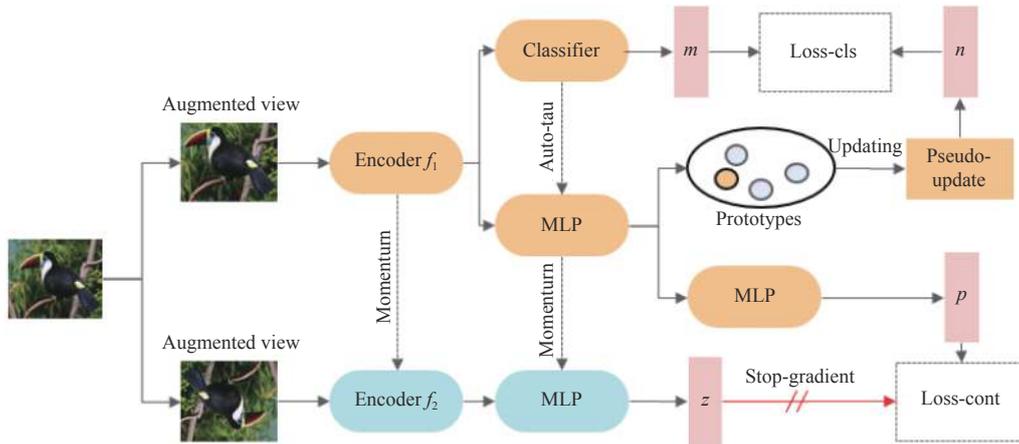


图4 CLPAT 网络结构图

编码器 f_2 的输出通过 MLP 模块后得到向量 z , 通过极小化这向量 z 和向量 p 的负余弦相似度, 来进行对比学习, 提取图像鲁棒的特征表示. 如式(2)所示:

$$S(p, z) = -\frac{p}{\|p\|_2} \times \frac{z}{\|z\|_2} \quad (2)$$

其中, $\|\cdot\|_2$ 是 L2 范数. 该模块的任务是用输出向量 p 来预测输出向量 z , 在无需使用标签的情况下从图像中学习有意义的表示. 在编码器 f_2 这一条分支使用到了 stop-gradient 操作^[29]. 实验表明使用 stop-gradient 操作, 可以有效避免对比学习出现崩溃解的问题. 因此对比学习的损失函数定义如式(3)所示:

$$L_{cont} = S(p, stopgrad(z)) \quad (3)$$

网络中的 MLP 模块以全连接层-BatchNorm 层-ReLU 激活函数-全连接层形式构成. 如图5所示.

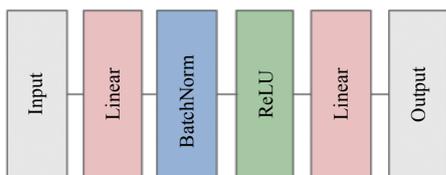


图5 MLP 模块

3.2 基于类原型的标签去噪模块

本文创新性地引入基于类原型的标签去噪模块.

对每一个类别保留一个原型向量 μ_c , $c \in \{1, 2, \dots, C\}$. 该向量视为每一个类别最具有代表性的特征向量. 开始训练时会为每一个样本初始化伪标签. 通过计算从 MLP 模块得到的特征向量与每一个类原型的相似度来判断当前样本的类别. 并通过移动平均的方式逐步更新当前样本的伪标签. 通过伪标签的加入, 降低原始标签中噪声的影响.

3.2.1 伪标签更新

首先对每个样本初始化其对应的均匀分布的伪标签, 使用式(4)进行初始化操作:

$$s_j = \frac{1}{|Y_i|} \prod (j \in Y_i) \quad (4)$$

其中, s_j 为样本第 j 个类别的初始化值, Y_i 为该样本的带噪标签向量. 然后根据从 MLP 中得到的特征向量与每个类别的原型向量的相似度计算结果来逐步更新伪标签, 伪标签的更新通过式(5)移动平均的方式实现:

$$\begin{cases} s = \eta s + (1 - \eta)v \\ v_c = \begin{cases} 1, & \text{if } c = \arg \max_{j \in Y} q^T \mu_j \\ 0, & \text{else} \end{cases} \end{cases} \quad (5)$$

其中, q 是从 MLP 中得到的特征向量, μ 为原型向量, v 是与该样本最相似的类原型所属类别的独热向量, η 是人工设置的超参数. 对每个输入样本 x 而言, 本文认为在计算相似度时, 与该样本相似度最高的原型向

量所属类别是该样本的真实标签. 通过使用移动平均的方式更新伪标签, 可以保证在模型训练不稳定的初期, 平滑地将伪标签向正确的方向更新. 当一个样本的特征向量始终与某个原型相似度最高, 随着训练的逐渐进行, 该样本的伪标签几乎可以收敛成一个独热向量 n . 使用交叉熵损失来训练分类器, 对每一个样本的损失使用式 (6) 计算:

$$\begin{cases} L_{cls}(f; x_i; Y_i) = \sum_{j=1}^C -s_{i,j} \log(f^j(x_i)) \\ \text{s.t. } \sum_{j \in Y_i} s_{i,j} = 1 \quad \text{and} \quad s_{i,j} = 0, \forall j \notin Y_i \end{cases} \quad (6)$$

其中, j 表示标签的索引, $s_{i,j}$ 代表样本 x_i 的伪标签中第 j 个类别的概率. f^j 表示输出样本结果中第 j 个标签的概率. 本文将分类器和对比学习模块一起训练, 将二者损失加权求和来组成最终的损失函数, 如式 (7) 所示, 其中 λ 是人工设置的超参数.

$$L = L_{cls} + \lambda L_{cont} \quad (7)$$

3.2.2 类原型更新

类原型是每一个类别最具有代表性的特征向量. 随着训练的进行, 为了使每个类别的特征向量能够更加准确地判断与哪个类原型更加相似, 类原型也需要不断地更新. 为了减少计算代价, 同样采用类似移动平均的方式来更新类原型, 如式 (8) 所示:

$$\begin{cases} \mu_c = \text{Normalize}(\gamma \mu_c + (1 - \gamma)q), \\ \text{if } c = \text{argmax}_{j \in Y} f^j(\text{Aug}_q(x)) \end{cases} \quad (8)$$

其中, γ 是人工设置的超参数, μ_c 是第 c 个类别的原型向量, q 是 MLP 模块输出的特征向量, f^j 表示输出样本结果中第 j 个标签的概率. 模型根据分类器的分类结果, 更新对应类别的类原型向量 μ_c .

3.3 自适应阈值

在分类任务中预置信度越高说明模型判断的越准确. 当出现较低的预测值时, 往往任务是模型分类出现错误. 通常情况下会采用设计固定阈值的方案来将预测值较低的样本过滤掉, 不参与模型最终的损失计算. 但是通过这种方式提高性能也存在着问题. 固定的阈值没有考虑不同类别学习难度不同的问题, 而且会忽略掉相当一部分训练样本. 因此本文采用一种新颖的自适应阈值方式过滤训练样本, 根据学习在各个类别的学习情况动态调整阈值. 过滤掉明显分类错误的样

本而只留下预测置信度高的样本, 这样可以有效地减少确认偏差. 因此, 本文使用式 (9) 来更新各类别的阈值:

$$T_t(c) = \beta_t(c) \times \tau \quad (9)$$

其中, $T_t(c)$ 表示在时刻 t 处 c 类别的阈值. $\beta_t(c)$ 为时刻 t 处 c 类别的缩放因子, τ 是人工设置的超参数. 当该类别学习效果较差时, 将会导致较低的阈值. 本文使用计数器 σ_t 来评估各个类别的学习效果, 并通过归一化 σ_t 来得到 β_t , 如式 (10) 和式 (11) 所示:

$$\sigma_t(c) = \sum_{n=1}^N \prod (\max(f(x)) > T_t(c)) \times \prod (\text{argmax}(f(x)) = c) \quad (10)$$

$$\beta_t(c) = \frac{\sigma_t(c)}{\max_c \sigma_t} \quad (11)$$

其中, 式 (10) 中的 $f(x)$ 为分类器输出结果. 这样进行归一化有一个好处: 对于学习效果最好的类别, 其缩放因子在归一化后等于 1, 这表示当前对于该类别的阈值设置的比较合理不需要改动. 对于学习效果不好的类别则降低阈值, 增加该类别的学习样本. 最终所有的类别都达到可靠的精度. 本文整体执行流程图如图 6 所示.

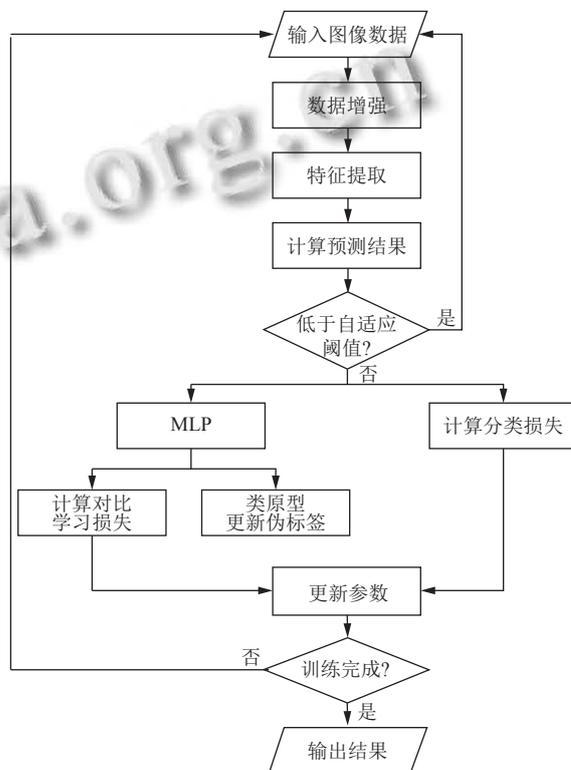


图 6 执行流程图

4 实验结果

4.1 对比实验结果

本文在两个常用公开数据集 CIFAR-10 和 CIFAR-100 以及一个真实数据集 ANIMAL10 上进行实验. CIFAR-10 和 CIFAR-100 分别包含 10 个类别和 100 个类别, 分别包含 5 万张训练图片和 1 万张测试图片. ANIMAL10 中包含 10 个类别, 每个类别的样本数量如表 1 所示. 每个类别取 10% 的样本作为测试样本. 在数据集中以概率 $q=P(\bar{y} \in Y|\bar{y} \neq y)$ 修改标签, 生成带有噪声的候选标签数据集. 在实验中对于 CIFAR-10 数据集中 q 取 {0.1, 0.3, 0.5}. 在 CIFAR-100 数据集中 q 取 {0.01, 0.05, 0.1}. 使用 PreAct ResNet18 作为 CIFAR 数据集的编码器. 对比学习输入的图像为对一张图片使用两次强数据增强生成的两个数据表示, 强数据增强方法包括随机反转、随机裁剪, 高斯模糊和颜色失真. 分类器输入图像仅使用一次弱图像增强生成, 确保最大程度保留样本信息, 包括随机图像翻转和随机图像裁剪.

采用以下设置进行训练: 学习率设置为 0.015, 并且采用余弦衰减的形式随着训练过程逐步改变学习率. 使用 SGD 优化器, 权重衰减设置为 0.0001, 动量设置为 0.9. batchsize 大小设置为 128, 这在单 GPU 的情况下也可以很好地运行. 本文选择了 5 种主流方法: 1) PiCO^[26]通过拉近相似样本距离及迭代更新伪标签的方式降低噪声标签影响; 2) LWS^[33]通过权衡候选标签上的损失和剩余标签上的损失来对风险函数进行加权; 3) PRODEN^[34]以自我训练的方式迭代更新标签分布; 4) CC^[35]假设数据生成过程是均匀的, 并且能够使得分类器表现一致; 5) MSE^[36]和 EXP^[36]是两种简单的基准方法, 它们使用均方误差和指数损失作为评估标准. 表 2、表 3 是与机器学习顶会发表的 PiCO、LWS、PRODEN、CC 模型以及 MSE 和 EXP 主流方法在 CIFAR-10 和 CIFAR-100 数据集上的对比试验结果.

表 1 ANIMAL10 数据集

类别	数量	类别	数量
狗	4863	鸡	3098
猫	1668	牛	1866
马	2623	羊	1820
大象	1446	蜘蛛	4821
蝴蝶	2112	松鼠	1862

本文采用相同的实验设置, 并在 3 次试验结果中取最优结果. 在 3 种不同噪声率上测试模型的性能. 从表 2、表 3 中结果可以看出, 本文方法在 CIFAR-10 和

CIFAR-100 数据集上的分类结果均超过其他方法. 在 CIFAR-10 数据集上, 相比于目前取得最高分类准确率的 PiCO, 在噪声率为 0.1、0.3、0.5 的情况下分别提升 0.25%、0.27%、0.24%. 在 CIFAR-100 数据集这类样本空间较大的数据集上, 本文方法的优势更加明显. 相较于取得最高分类准确率的 PiCO 模型而言, 本文方法并未采用分类器主导对比学习中正负样本对选择的方式, 在噪声率为 0.01、0.05、0.1 的情况下分别提升 2.47%、1.64%、1.49%. 表明本文方法在样本空间较大的数据集上可以产生分离更良好的聚类以及更好区分的特征表示, 具有良好的抗噪声能力.

表 2 CIFAR-10 数据集实验结果 (%)

Model	$q=0.1$	$q=0.3$	$q=0.5$
PiCO ^[26]	94.39±0.18	93.58±0.12	93.28±0.06
LWS ^[33]	90.30±0.60	88.99±1.43	86.16±0.65
PRODEN ^[34]	90.24±0.32	89.38±0.31	87.78±0.07
CC ^[35]	82.30±0.30	79.08±0.06	74.05±0.36
MSE ^[36]	79.97±0.45	75.64±0.28	67.09±0.66
EXP ^[36]	79.32±0.12	75.79±0.21	70.34±1.32
CLPAT	94.64±0.21	93.85±0.08	93.52±0.04

表 3 CIFAR-100 数据集实验结果 (%)

Model	$q=0.01$	$q=0.05$	$q=0.1$
PiCO ^[26]	73.09±0.34	72.74±0.30	69.91±0.24
LWS ^[33]	678±0.02	59.56±0.34	53.53±0.08
PRODEN ^[34]	62.60±0.02	60.73±0.03	56.80±0.29
CC ^[35]	49.76±0.05	47.62±0.08	35.72±0.37
MSE ^[36]	49.17±0.05	46.02±1.82	43.81±0.49
EXP ^[36]	44.45±1.50	41.05±1.40	29.27±1.81
CLPAT	75.56±0.25	74.38±0.28	71.41±0.18

进一步验证本文方法的有效性, 在 ANIMAL10 数据集上的对比试验. 表 4 是与 PiCO、LWS、PRODEN、CC 模型以及 MSE 和 EXP 方法在 ANIMAL10 真实数据集上的对比试验结果. 在真实数据集 ANIMAL10 上本文设置学习率为 0.2, 其余试验设置与 CIFAR 数据集实验相同. 从结果可以看出本文模型实验结果超过其他模型, 相比于效果最好的 PiCO 模型提高了 0.24%.

表 4 ANIMAL10 数据集实验结果 ($q=0.1$) (%)

Model	准确率
PiCO ^[26]	93.39±0.34
LWS ^[33]	90.16±0.42
PRODEN ^[34]	89.94±0.32
CC ^[35]	81.72±0.26
MSE ^[36]	78.43±0.34
EXP ^[36]	78.12±0.13
CLPAT	93.63±0.14

4.2 消融实验结果

为验证模型效果, 将本模型与3个变体进行比较: 1) CLPAT w/o th, 是模型指使用固定阈值过滤样本; 2) CLPAT w/o pseudo, 在第1个变体的基础上去除伪标签更新, 使伪标签保持初始化时的值; 3) CLPAT w/o cont, 进一步去除对比学习模块, 只保留单独的分类器. 本节所有的消融实验均在 CIFAR-10 数据集上进行, 噪声率设为 0.5, 其余设置与训练过程一致.

从图7中可以看出, 自监督对比学习优化了数据分布, 使得相似的样本更加接近, 最终得到鲁棒的特征表示, 变体3相较于未加入对比学习的变体2提高了9.17% 可以看出对比学习的加入提高了模型的性能. 基于类原型的标签去噪模块通过逐渐更新样本的伪标签来降低带噪标签的影响, 变体1相较于未使用去噪模块的变体2提高了7.60%. 自适应阈值主要用于根据模型学习情况改变阈值, 过滤掉置信度较低的样本, 提升相对较少, CLPAT 相较于未使用自适应阈值的变体1提高了0.69%. 由此可以说明3个小方法都可以提高模型抗噪声的能力.

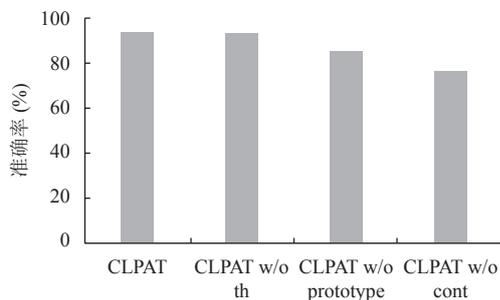


图7 消融实验结果

5 结论与展望

针对图像分类模型训练时需要大规模准确标注数据, 而数据来源不可靠导致获得数据标签包含噪声的问题, 本文提出了一种基于对比学习的标签噪声学习框架, 从网络特征提取能力出发, 采用对比学习网络来提取图像鲁棒的特征表示, 并结合基于类原型的标签去噪模块逐步更新样本的伪标签, 提高模型抗噪声能力. 采用自适应阈值来过滤训练样本, 进一步提高模型的性能, 学习了一个具有良好抗噪声能力的模型. 在 CIFAR-10、CIFAR-100 和 ANIMAL10 数据集上的对比试验证明了本文模型的有效性. 在类别更多的数据集上, 模型效果并不理想, 原因在于当类别达到一定数

量时, 在训练初期原型向量之间的区别并不明显. 后续将采用预训练的方式处理原型向量, 提高模型在多类别噪声数据集上的分类效果.

参考文献

- 1 王晓莉, 薛丽. 标签噪声学习算法综述. 计算机系统应用, 2021, 30(1): 10–18. [doi: 10.15888/j.cnki.csa.007776]
- 2 张增辉, 姜高霞, 王文剑. 基于动态概率抽样的标签噪声过滤方法. 计算机应用, 2021, 41(12): 3485–3491.
- 3 陈庆强, 王文剑, 姜高霞. 基于数据分布的标签噪声过滤. 清华大学学报(自然科学版), 2019, 59(4): 262–269. [doi: 10.16511/j.cnki.qhdxxb.2018.26.059]
- 4 余孟池, 牟甲鹏, 蔡剑等. 噪声标签重标注方法. 计算机科学, 2020, 47(6): 79–84.
- 5 Cho JH, Hariharan B. On the efficacy of knowledge distillation. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 4794–4802.
- 6 Xie QZ, Luong MT, Hovy E, *et al.* Self-training with noisy student improves ImageNet classification. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 10687–10698.
- 7 Yang CL, Xie LX, Qiao SY, *et al.* Training deep neural networks in generations: A more tolerant teacher educates better students. Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu: AAAI Press, 2019. 5628–5635.
- 8 Yim J, Joo D, Bae J, *et al.* A gift from knowledge distillation: Fast optimization, network minimization and transfer learning. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 4133–4141.
- 9 Patrini G, Rozza A, Krishna Menon A, *et al.* Making deep neural networks robust to label noise: A loss correction approach. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 1944–1952.
- 10 Gui XJ, Wang W, Tian ZH. Towards understanding deep learning from noisy labels with small-loss criterion. Proceedings of the 30th International Joint Conference on Artificial Intelligence. Montreal: IJCAI.org, 2021. 2469–2475.
- 11 Shu J, Xie Q, Yi LX, *et al.* Meta-weight-net: Learning an explicit mapping for sample weighting. Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver, 2019. 172.
- 12 Li M, Zhou ZH. SETRED: Self-training with editing. Proceedings of the 9th Pacific-Asia Conference on

- Knowledge Discovery and Data Mining. Hanoi: Springer, 2005. 611–621.
- 13 Jiang L, Zhou ZY, Leung T, *et al.* MentorNet: Learning data-driven curriculum for very deep neural networks on corrupted labels. Proceedings of the 35th International conference on machine learning. Stockholm: PMLR, 2018. 2309–2318.
 - 14 Blum A, Mitchell TM. Combining labeled and unlabeled data with co-training. Proceedings of the 11th Annual Conference on Computational Learning Theory. Madison: ACM, 1998. 92–100.
 - 15 Li JN, Socher R, Hoi SCH. DivideMix: Learning with noisy labels as semi-supervised learning. Proceedings of the 8th International Conference on Learning Representations. Addis Ababa: OpenReview.net, 2020. 1–14.
 - 16 Sukhbaatar S, Fergus R. Learning from noisy labels with deep neural networks. Proceedings of the 3rd International Conference on Learning Representations. San Diego, 2015. 4–14.
 - 17 Goldberger J, Ben-Reuven E. Training deep neural-networks using a noise adaptation layer. Proceedings of the 5th International Conference on Learning Representations. Toulon: OpenReview.net, 2017. 1–9.
 - 18 Han B, Yao JC, Niu G, *et al.* Masking: A new perspective of noisy supervision. Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal: Curran Associates Inc., 2018. 5841–5851.
 - 19 Li JN, Wong YK, Zhao Q, *et al.* Learning to learn from noisy labeled data. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 5051–5059.
 - 20 Jindal I, Nokleby M, Chen XW. Learning deep networks from noisy labels with dropout regularization. Proceedings of the 16th IEEE International Conference on Data Mining. Barcelona: IEEE, 2016. 967–972.
 - 21 Srivastava N, Hinton G, Krizhevsky A, *et al.* Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 2014, 15(1): 1929–1958.
 - 22 Ciortan M, Dupuis R, Peel T. A framework using contrastive learning for classification with noisy labels. Data, 2021, 6(6): 61. [doi: [10.3390/data6060061](https://doi.org/10.3390/data6060061)]
 - 23 Ghosh A, Lan A. Contrastive learning improves model robustness under label noise. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 2703–2708.
 - 24 Li JN, Xiong CM, Hoi SCH. Learning from noisy data with robust representation learning. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 9485–9494.
 - 25 Li JN, Xiong CM, Hoi SCH. MoPro: Webly supervised learning with momentum prototypes. Proceedings of the 9th International Conference on Learning Representations. OpenReview.net, 2021. 1–15.
 - 26 Wang HB, Xiao RX, Li YX, *et al.* PiCO: Contrastive label disambiguation for partial label learning. Proceedings of the 10th International Conference on Learning Representations. OpenReview.net, 2022. 1–18.
 - 27 Yao YZ, Sun ZR, Zhang CY, *et al.* Jo-SRC: A contrastive approach for combating noisy labels. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 5192–5201.
 - 28 Zhang X, Liu ZX, Xiao KW, *et al.* CoDiM: Learning with noisy labels via contrastive semi-supervised learning. arXiv:2111.11652, 2021.
 - 29 Chen XL, He KM. Exploring simple siamese representation learning. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 15750–15758.
 - 30 He KM, Fan HQ, Wu YX, *et al.* Momentum contrast for unsupervised visual representation learning. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 9729–9738.
 - 31 Chen T, Kornblith S, Norouzi M, *et al.* A simple framework for contrastive learning of visual representations. Proceedings of the 37th International Conference on Machine Learning. PMLR, 2020. 149.
 - 32 田浩江, 路娜, 崔二洋. 基于改进对比学习的道路裂缝图像分类. 计算机系统应用, 2023, 32(2): 310–315. [doi: [10.15888/j.cnki.csa.008944](https://doi.org/10.15888/j.cnki.csa.008944)]
 - 33 Wen HW, Cui JY, Hang HY, *et al.* Leveraged weighted loss for partial label learning. Proceedings of the 38th International Conference on Machine Learning. PMLR, 2021. 11091–11100.
 - 34 Lv JQ, Xu M, Feng L, *et al.* Progressive identification of true labels for partial-label learning. Proceedings of the 37th International Conference on Machine Learning. PMLR, 2020. 603.
 - 35 Feng L, Lv JQ, Han B, *et al.* Provably consistent partial-label learning. Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2020. 919.
 - 36 Feng L, Kaneko T, Han B, *et al.* Learning with multiple complementary labels. Proceedings of the 37th International Conference on Machine Learning. PMLR, 2020. 288.

(校对责编: 孙君艳)