DOI：10.19789/j.1004-9398.2026.04.007　　**Open Access**
文献引用：郭凯文，马明，夏立坤．基于渐进式难样本衰减的噪声标签学习方法[J]．首都师范大学学报（自然科学版），2026，47（4）：60-69. Guo Kaiwen，Ma Ming，Xia Likun. Noisy label learning through progressive attenuation of hard samples[J]. Journal of Capital Normal University（Natural Science Edition），2026，47（4）：60-69.

基于渐进式难样本衰减的噪声标签学习方法*

郭凯文[1]，马　明[2]，夏立坤[1**]

（1. 首都师范大学信息工程学院，北京　100048；
2. 叶史瓦大学卡茨科学与健康学院，纽约　10016）

摘要：针对标签噪声导致深度神经网络性能下降的问题，本文提出了一种新的难样本自适应加权方法，通过在每个训练周期中多次重新加权训练模型，从难样本的不同子集中学习并迭代预测伪标签，以提高噪声标签修正的准确性。实验表明，该方法在噪声标签学习任务上优于现有方法，展现出良好的性能提升和泛化能力。本研究为解决实际场景中标签噪声问题提供了一种新思路。

关键词：噪声标签；难样本学习；自适应重加权；半监督学习
中图分类号：TP18　　**文献标志码**：A

Noisy label learning through progressive attenuation of hard samples*

Guo Kaiwen[1]，Ma Ming[2]，Xia Likun[1**]

（1. Information Engineering College，Capital Normal University，Beijing　100048；
2. Katz School of Science and Health，Yeshiva University，New York　10016）

Abstract：To address the performance degradation of deep neural networks caused by noisy labels，we propose hard sample adaptive labeling with optimal reweighting for noisy labels，a novel hard sample adaptive weighting method. By reweighting and retraining the model multiple times in each epoch，learning from different subsets of hard samples，and iteratively predicting pseudo-labels，HEALON improves the accuracy of noisy label correction. Experiments demonstrate that our method outperforms existing approaches on noisy label learning tasks，showing significant performance gains and better generalization. This research provides a new perspective for tackling the label noise problem in real-world scenarios.

Keywords：noisy labels；hard sample learning；adaptive reweighting；semi-supervised learning.
CLC：TP18　　**DC**：A

0　引　言

近年来，深度神经网络在计算机视觉、语音识别和自然语言处理等各个领域获得了显著成功。决定深度神经网络训练性能的一个关键因素是具有高质量标签的充足数据集的可用性。然而，获取大型数据集的正确标签既耗时又费力。在实践中，噪声标签经常被引入，对深度神经网络的性能产生

收稿日期：2024-11-08
*北京市自然科学基金面上项目（4242033）
**通信作者：xlk@cnu.edu.cn

不利影响,阻碍泛化。特别是在大型语言模型的进步中,预训练数据集中存在噪声标签会降低下游任务的性能。在这些预训练数据集中,噪声标签的非可忽略存在可能会降低大型语言模型应用于不同下游任务时的性能。

为了解决从噪声标签中学习的问题,已经开发了许多方法[1]:样本加权方法[2-5],例如Jiang等的MentorNet[2,6],专注于在训练期间为深度网络提供样本权重方案,以解决噪声标签问题;样本选择方法[7-9],例如Huang等[10]的噪声净化框架和Wu等[11]的TopoFilter方法,旨在过滤掉噪声样本;与半监督学习技术相结合[12],如Han等[13]的共同教学方法和Li等[14]的DivideMix方法,利用损失较小的样本子集或对损失分布进行建模来划分训练数据。本文提出的方法与上述工作相关,但有重要创新,在相关工作部分将详细讨论这些方法的技术细节及其与本文方法的关系。

上述基于半监督学习的方法假设模型具有良好的记忆效应[15]。尽管完全信任模型对所有噪声标签预测的记忆效应可以防止模型受到大多数噪声标签的影响[16-18],但可能存在一些标签校正失败的困难样本,一旦被模型信任,就会导致最终结果的泛化偏差。在学习过程中,具有相似输出概率的困难样本更接近模型的决策边界,导致记忆效应准确性降低。因此,依赖记忆效应进行标签校正可能会对困难样本产生负面影响。

为了实证验证这一假设,本研究分析了样本预测概率复杂度对标签校正准确性的影响。训练不足、训练充分和过度训练3个训练阶段样本难度与伪标签准确率之间的关系如图1所示,对应分别训练第8、40和100轮。在每个子图中,橙色实线表示不同难度水平下的伪标签准确率;蓝色直方图显示了样本在这些难度水平上的分布;橙色虚线表示总体伪标签准确率。表明无论数据中是否存在噪声,预测概率具有高确定性的简单样本都可以实现高噪声过滤准确性和校正准确性,而预测概率相对复杂的困难样本的噪声过滤准确性和校正准确性较低。可以推断,同时校正所有难度级别的样本将导致模型逐渐积累错误记忆,从而导致最终泛化性能下降。

然而,在没有先验知识的情况下区分困难样本的真实类别几乎是不可能的[19]。此外,如图1(a和c)所示,受噪声标签影响的困难样本的进一步学习会导致记忆效应的消失。简单地增加训练迭代次数无法解决这个问题。受样本加权和重采样的启发[20-21],本研究假设在训练过程中,困难样本可以通过样本加权来影响,这对错误标签的影响较小。此外,利用样本权重部分影响样本难度不需要整体进一步训练。因此,通过这种方法部分优化困难样本不会导致整体记忆效应的消失。因此,为了改善困难样本的标签校正准确性,本研究提出了一个新的渐进式困难样本衰减框架,在每个周期中多次运行,使框架从简单样本和困难样本中逐步学习尽可能多的有用信息,从而最大限度地利用噪声数据集中的所有样本来解决噪声标签学习中困难样本的问题。所提出框架的主要贡献总结如下:

(1)本研究提出了一个创新的框架,称为HEALON(用于噪声标签的困难样本自适应标记与最优重新加权),可以稳健地纠正困难样本的噪声标签。HEALON将迭代估计与先进的半监督方法相结合。在伪标签生成阶段,提出HEAL-Predict,其结合样本难度加权信息,以提供更高精度的伪标

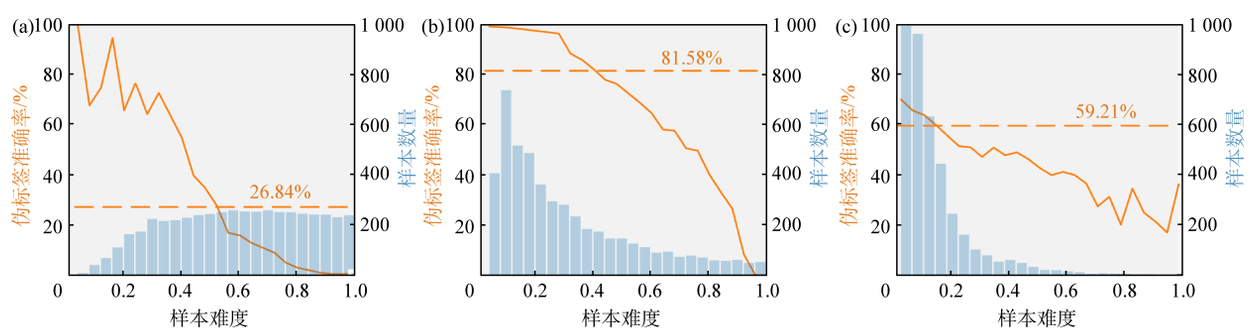


图1　不同训练阶段样本难度与伪标签准确率之间的关系

(a)训练不足(第8轮);(b)训练充分(第40轮);(c)过度训练(第100轮)

Figure 1　relationship between sample difficulty and pseudo-label accuracy at different training stages

(a) under-trained (epoch 8);(b) well-trained (epoch 40);(c) over-trained (epoch 100)

签更新,从而减少了困难样本导致的错误累积,这可能会阻碍模型的记忆能力。

(2)在迭代估计过程中,引入HEAL-Weight,计算优化的样本权重,以自适应地调整困难样本的半监督损失。通过将HEAL-Weight纳入半监督学习框架,逐步执行困难样本衰减,从而减少困难样本的全局影响,提高伪标签估计的整体准确性。

(3)基于样本难度与伪标签质量高度相关的有意义假设,理论上将标签校正过程与伪标签和真实标签差异联系起来,然后求解困难样本的最优训练权重。还基于信息熵量化样本难度,为所提出的框架奠定了基础。

(4)在标准监督学习设置下,对各种噪声数据集上的最先进方法进行广泛地实验和性能比较,以评估所提出的框架。实验结果证明了本研究提出的HEALON在噪声标签校正准确性、噪声稳健性和整体性能方面的优越性。

1 相关工作

1.1 基于样本选择的噪声标签学习方法

Mirzasoleiman等[7]引入了一种选择干净样本子集的算法,以防止过拟合噪声标签;Pleiss等[8]测量了指定类别与其最高非指定类别之间的输出概率值的平均差异,以划分干净样本和噪声样本;Cheng等[9]使用了一种新的置信正则化项,以逐步过滤掉损坏的样本;Kim等[22]专注于每个数据点的潜在表示动力学,并使用其特征向量过滤噪声标签,以测量样本差异;Wu等[11]提出了一个新的最优传输滤波方法,考虑了几何上有意义的距离和保留分布模式来测量样本差异。

1.2 半监督学习和样本选择方法的结合

Han等[13]提出了共同教学方法,基于2个网络,从每个网络中提取损失较小的样本子集,然后使用另一个网络过滤的样本子集分别训练每个网络;Yu等[12]结合了不一致的训练过程与共同教学,提出了co-teaching+,仅选择2个网络预测结果不同的样本。

1.3 样本加权方法

样本加权是一种用于解决数据不平衡、标签噪声或指导模型训练的技术。例如,标签分布感知边缘损失[2],减少训练数据重要性权重对深度网络的影响[3],以及多相似性损失的通用配对加权框架[4]。Shu等[23]提出了一种从数据中自适应学习显式加权函数的方法。数据加权也被应用于解决标签噪声问题;Jiang等[6]提出了MentorNet,为深度网络的训练提供了一种样本加权方案。然而,MentorNet没有明确考虑样本难度,也没有进行标签校正。相反,所提出的HEALON框架结合了样本难度信息,以自适应地调整困难样本的权重,并采用迭代标签校正来细化噪声标签。这种方法有效地减少了噪声标签的影响,保留了模型的记忆能力,并最大限度地利用了可用数据。

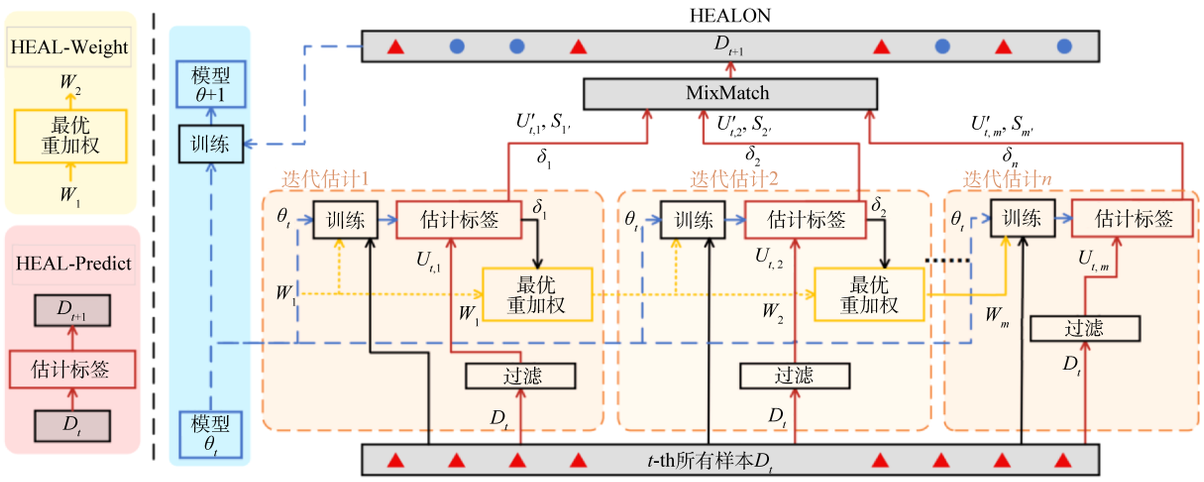
2 方 法

在本节中,为了避免减少困难样本的数量和避免记住噪声之间的矛盾,提出了HEALON,其迭代多次以最优地估计困难样本,而无需减少其数量。HEALON的过程如图2所示,在每个周期中,HEALON使用不同的权重重新训练模型多次。每次迭代使用HEAL-Predict进行标签校正;在每次迭代之后,HEAL-Weight计算最优权重。在每次迭代中,首先通过输出概率计算样本难度,使用权重更新估计的伪标签,并使用难度信息优化权重;然后,使用样本难度信息优化伪标签的权重。这样,每次迭代都引导模型使用不同的样本权重进行学习,仅减少部分样本的难度,使模型能够准确估计困难样本的真实标签,而不会记住噪声。在这里,本研究首先定义符号。

2.1 符号

设一个大小为N的数据集表示为$\mathcal{D}^{*}=\left(x_i,y_i^{*}\right)$,$i\in N$,其中$x_i\in\mathbf{R}^d$表示第$i$个样本的$d$维特征,$y_i^{*}\in\mathbf{R}^K$表示对应的真值标签。在许多实际问题中,从分布$\mathcal{D}$独立抽取的示例是不可用的。在观察之前,清洁标签通常被随机损坏为噪声标签。对于噪声标签学习,只能访问一个噪声样本集$\mathcal{D}=\left(x_i,y_i\right)$,$i\in\mathbf{N}$,其中$y_i$表示噪声标签的变量。本研究使用交叉熵损失$\mathcal{L}\left(y_i^{*},y_i\right)=\sum_K-\left(y_i\right)_K\log\left(y_i^{*}\right)_K$来测量真实标签和预测伪标签之间的差异。目标是学习一个鲁棒的分类器,该分类器可以从噪声样本集中准确分类测试实例。

为了逐步校正噪声样本集$\mathcal{D}$,利用样本过滤技术,将噪声数据过滤为无标签噪声样本集和有标签干净样本集。为了与其他最先进的方法进行比较,选择使用DivideMix中的样本过滤方法,这是一种常用的半监督学习方法,用于处理噪声标签。Divi-



注:HEALON框架通过在每个训练周期中迭代执行的3个关键阶段来解决带噪声标签的学习挑战:首先,当进行到第 m 次迭代,噪声样本过滤器根据预测置信度将原始数据集 D_t 分出未标记集 $U_{t,m}$;然后,通过加权训练期间多个模型 $B(U,\theta_m,w_m)$ 执行推理以生成预测,这些预测通过预测 logits 的信息熵用于估计样本难度 δ_m,并通过聚合跨模型的输出产生加权集成预测 $U'_{t,m}$;最后,根据当前性能 S_m 和难度值 δ_m 计算最优权重 w_m。

图2　HEALON迭代过程示意

Figure 2　schematic diagram of the HEALON iterative process

deMix[14]采用期望最大化算法,将2个分量的高斯混合模型拟合到损失中,以过滤噪声样本。定义样本过滤方法如下:

定义1　(噪声样本过滤)将噪声样本过滤器表示为 F,将噪声样本集 $\mathcal{D}$ 分为有标签样本集 $\mathcal{X}=\{(x_i,y_i)|(x_i,y_i)\in\mathcal{D},y_i=y_i^*\}$ 和无标签样本集 $\mathcal{U}=\{y_i|(x_i,y_i)\in\mathcal{D},y_i\neq y_i^*\}$ 带有标签的样本集,即

$$\mathcal{X},\mathcal{U}=F(\mathcal{D})。\tag{1}$$

2.2　HEAL-Predict:基于样本难度的伪标签聚合

此处详细介绍基于样本难度的伪标签估计方法。旨在实现最终预测伪标签的最高预测准确性。

定义2　(样本难度)假设有 $m=m_{\max}$ 个模型,不同的参数初始化 θ^m 和 N 个噪声样本 $\mathcal{U}$ 带有标签。设 $\mathcal{U}_j^i$ 表示第 i 次运行中的第 j 个噪声样本,$\mathcal{U}^i$ 和 $\mathcal{U}_j$ 表示第 i 次运行中的所有噪声样本和所有运行中的第 j 个噪声样本。设 $\bar{y}^m=\mathcal{B}(\mathcal{U},\theta^m,w^m)$ 为模型 $\mathcal{B}$ 的输出,参数为 θ^m 和输入 $\mathcal{U}$,其中 w^m 是模型训练过程中的样本权重。记 δ^i 为第 i 次运行的难度值,通过以下公式计算:

$$\delta_j^i=-\sum_k \bar{y}_{jK}^i\log\left(\bar{y}_j^i\right)_K。\tag{2}$$

最后,记 $\mathcal{S}^i=\sum_{j=0}^{i}(1-\delta^j)$ 为前 i 次运行的样本难度之和。

定义3　(加权平均)基于模型输出的概率分布 $\bar{y}^m$,对无标签样本集 $\mathcal{U}$ 的真实标签 $\bar{Y}^m$ 进行加权平均估计,即:

$$\bar{Y}^m=\frac{\sum_{i=1}^{m}\left(1-\delta_j^i\right)\bar{y}_j^i}{\sum_{i=1}^{m}\left(1-\delta_j^i\right)}。\tag{3}$$

MixMatch[25]通过结合一致性正则化来鼓励对扰动无标签数据的一致预测,通过熵最小化来鼓励对无标签数据自信的预测,并通过 MixUp 增强来鼓励样本之间的线性行为[24],辅助模型更好地利用无标签数据,提升模型的泛化能力和鲁棒性。因此,为了生成高置信率的伪标签,并且充分利用模型的记忆效应,将 MixMatch 与样本加权结合。整个算法详细描述如下:

输入:训练数据集 $\mathcal{D}$,模型参数 θ^t,最大估计迭代次数 $m_{\max}$,
输出:训练后的模型参数 θ^{t+1}。

初始化样本权重 $w^{t,0}$,伪标签 $\bar{Y}^{t,0}$;
使用公式(1)中的 F 将数据集 $\mathcal{D}$ 分为有标签数据 $\mathcal{X}^t$ 和无标签数据 $\mathcal{U}^t$;
从 $m=1$ 到 $m_{\max}$ 循环;
使用加权数据 $(\mathcal{D},w^{t,m-1})$ 训练 $\mathcal{B}(\theta^t)$ 得到 $\mathcal{B}(\theta^{t,m})$;
计算样本难度 $\delta^{t,m}$;
$\bar{y}^{t,m}=\mathcal{B}(\mathcal{U}^t,w^{t,m-1},\theta^{t,m})$,
$\bar{Y}^{t,m}=U_{\mathrm{l}}(\bar{Y}^{t,m-1},\bar{y}^{t,m},\delta^{t,m})$,
$w^{t,m}=U_{\mathrm{w}}(\delta^{t,m})$,
结束循环。
对 $\bar{Y}^{t,m_{\max}}$ 进行数据增强得到 $\mathcal{X}''$ 和 $\mathcal{U}''$ 打乱增强后的数据 $\mathcal{X}''$ 和 $\mathcal{U}''$;
使用 MixUp 混合有标签数据 $\mathcal{X}''$ 和伪标签数据 $\mathcal{U}''$;
使用 $\mathcal{X}''$ 和 $\mathcal{U}''$ 的混合数据及平均权重训练 $\mathcal{B}(\theta^t)$ 得到 $\mathcal{B}(\theta^{t+1})$。

在每个周期中,重新初始化和多次使用新的权重 w^m 重新训练网络,修正标签 $\bar{y}$ 和真值标签 y^* 在不同难度 δ 的样本上的损失,在 CIFAR-10 数据集上进行训练,噪声率为 10%、30%、50% 和 70% 对称噪声。在给定 w^m 的情况下,使用 HEAL-Predict 估计所有无标签样本 $\mathcal{U}$ 的最优伪标签后,结合 MixMatch 和上述基于难度的加权平均,线性组合 m 次迭代的估计标签 $\bar{y}^m$ 及其概率,得到改进的伪标签,对有标签和无标签数据进行数据增强,输出 $\mathcal{X}'$ 和 $\mathcal{U}'$。

除了将加权平均方法集成到 MixMatch 中,为了考虑伪标签估计和训练困难样本,改进了 MixMatch 的损失函数,使其能够自适应地学习和估计困难样本的真实标签。为了使多次估计能够尽可能地关注困难样本,改进了 DivideMix 的损失函数,结合了 m-th 训练的权重 w^m。改进的损失函数包括交叉熵损失 $\mathcal{L}_{\mathcal{X}}^m$、均方误差损失 $\mathcal{L}_{\mathcal{U}}^m$ 和正则化项 $\mathcal{L}_{\text{reg}}$ 的计算公式为:

$$\mathcal{L}_{\mathcal{X}}^m = -\frac{1}{|\mathcal{X}'|}\sum_{x,y^*\in\mathcal{X}'} w_x^m \sum_K \left(y^*\right)_K \log\left(\mathcal{B}\left(x;\theta\right)^K\right), \quad (4)$$

$$\mathcal{L}_{\mathcal{U}}^m = \frac{1}{|\mathcal{U}'|}\sum_{x,y^*\in\mathcal{U}'} w_x^m \left\| y^* - \mathcal{B}\left(x;\theta\right)\right\|_2^2, \quad (5)$$

$$\mathcal{L}_{\text{reg}} = \sum_K \pi_K \log\left(\pi_K / \frac{1}{|\mathcal{X}'|+|\mathcal{U}'|}\sum_{x\in X'+U'} B\left(x;\theta\right)^K\right)。 \quad (6)$$

通过将样本权重 w^m 添加到监督损失 $\mathcal{L}_{\mathcal{X}}^m$,一致性损失 $\mathcal{L}_{\mathcal{U}}^m$ 的 MixMatch 和正则化项 $\mathcal{L}_{\text{reg}}$ 引入样本权重,模型可以更多地关注难以学习的样本和减少噪声影响的关键样本。用于训练的最终损失函数为 $\mathcal{L}^m$ 的计算公式为

$$\mathcal{L}^m = \mathcal{L}_{\mathcal{X}}^m + \lambda_u \mathcal{L}_{\mathcal{U}}^m + \mathcal{L}_{\text{reg}}。 \quad (7)$$

2.3 HEAL-Weight:最优样本权重计算方法

为了设计一个鲁棒和自适应的加权方法,本文研究了样本难度对伪标签估计准确性的影响,并利用样本难度与伪标签估计准确性之间的密切关系。在任何噪声水平和训练阶段,样本难度与伪标签和真值标签之间的准确性高度相关(图3)。这一现象在不同数据集和噪声率下普遍存在,并且直接关系到任何训练时刻可以测量的状态量与模型预测和真实无噪声标签之间的关系。这为利用样本难度和自适应加权方法优化带噪声标签的学习过程提供了可能,因为带噪声标签的学习无法在训练中利用真实无噪声标签。

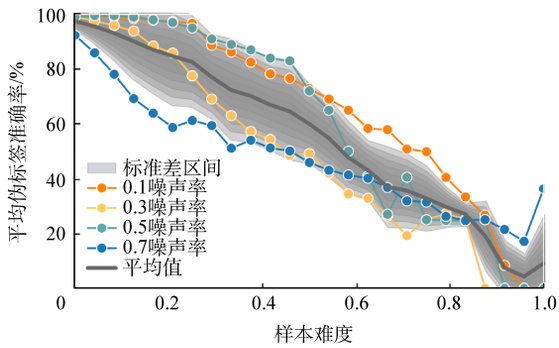


图3　算法详细描述

Figure 3　detailed description of the algorithm

受上述观察的启发,本研究尝试通过利用样本难度和自适应加权方法来优化模型训练和伪标签预测。下面提供了最小化伪标签和真值标签之间损失的最优样本加权方法,以前的工作[14,26]中进行了简单噪声标签设置。主要在这里介绍结论及其解释。首先引入符号,然后建立理论结论并提供由改进算法在某些假设下求解的最优样本权重。证明的细节可以在附录中找到。

假设 1　存在一个函数 $C\left(\delta^m\right)$ 使得当条件 $\sum_{i=0}^N \delta_i^m N\delta_{\min}$ 满足时,$\mathcal{L}\left(\bar{y}^m, y^*\right) \approx C\left(\delta^m\right)$ 并且关于 δ 的导数近似恒定。

假设 2　样本难度 $\delta^m = \sum_K - \mathcal{B}\left(\mathcal{U},\theta^m,w^m\right)_K \log \mathcal{B}\left(\mathcal{U},\theta^m,w^m\right)_K$,则 δ^m 随 w^m 单调变化,当模型 $\mathcal{B}$ 使用 w^m 训练时。

定理 1　(难样本的最优加权)让模型 $\bar{y}^m = \mathcal{B}\left(\mathcal{U},\theta^m,w^m\right)$ 使用 m 不同的初始化 θ^m 和不同的权重 w^m 训练,对伪标签进行预测。当上述假设满足时,最小化 $\mathcal{L}\left(\bar{Y},y^*\right)$ 的 w^m 的更新满足

$$w_i^{m+1} = \frac{\mathcal{S}_i^m \hat{\mathcal{L}}\left(y_i^*, \bar{Y}_i^m\right)}{\left(\mathcal{S}_i^m + 1\right)\sum_j^N w_j^m}, \quad (8)$$

其中

$$\hat{\mathcal{L}}\left(y^*,\bar{Y}^m\right) = \frac{\mathcal{S}^{m-1}\hat{\mathcal{L}}\left(y^*,\bar{Y}^{m-1}\right)}{\mathcal{S}^{m-1}+1-\delta^m} + \frac{\delta^m - \left(\delta^m\right)^2}{\mathcal{S}^{m-1}+1-\delta^m}。 \quad (9)$$

定理1表明 $(m+1)$-th 迭代的最佳样本加权取决于当前难度系数 δ^m,所有先前难度的和 $\mathcal{S}^{m-1}$,以及先前估计的预测损失 $\hat{\mathcal{L}}\left(y^*,\bar{Y}^{m-1}\right)$。基于定理1的自适应加权方法在每次迭代中基于模型预测置信度计算当前样本的难度,并将其与历史难度进行累积,以反映模型的长期学习状态,同时通过惩罚持续高误差样本来抑制其权重膨胀,最终借助比例系

数平衡难度与损失的贡献,实现下一轮迭代的样本权重自适应生成。

随着所有样本的平均难度下降,即 $\mathcal{S}^m$ 增加,估计标签 $\bar{Y}^m$ 和真值标签 y^* 之间的交叉熵损失可以保证很小。因此,通过迭代更新权重 w^m 和总体置信度 $\mathcal{S}^m$,可以最小化估计标签和真值标签之间的差异。

3 实　验

本研究在这部分中对提出的 HEALON 进行了彻底的分析,验证了 HEAL-Predict 和 HEAL-Weight 在样本选择和估计过程中的有效性,并与现有的最先进方法在2个数据集上进行了比较。在2个基准数据集上进行了广泛的实验:CIFAR-10和CIFAR-100[27]使用2种类型的标签噪声,分为对称噪声和非对称噪声。对称噪声是通过随机替换训练数据的一部分标签来生成的。在非对称噪声中,标签仅被替换为相似的类别。

3.1 实验设置

本研究使用 18 层 ResNet[28]作为骨干网络基础模型,并使用随机梯度下降(SGD)以 0.9 的动量和 5×10^{-4} 的权重衰减,以及 128 的批量大小进行训练,总共训练 600 个周期。

本研究遵循 DivideMix[14]的架构和超参数设置,除了自监督学习方法。初始阈值 t_0 设置为 0.5。预热周期设置为 CIFAR-10 的 15 个周期和 CIFAR-100 的 30 个周期。当噪声率< 0.5 时,最大迭代估计次数 $m_{\max}$ 设置为 3;当噪声率>0.5 时,设置为 5。训练过程在超过 530 个总周期后停止,并使用所有可信样本进行最终训练迭代。所有实验都在 Windows OS 上使用 Nvidia RTX8000,40 GB 图形内存实现。

3.2 对难样本加权伪标签聚合方法的验证

在这一部分,主要验证了 HEAL-Weight。通过比较第一次校正(确保在相同条件下跨方法的一致验证准确性)的效果,评估了 Cross-Entropy[22]、DivideMix[14]和 HEALON 的性能。对于 CIFAR-10,在对称噪声率设置为 0.1、0.3、0.5、0.7、0.9 下进行测试,而对于 CIFAR-100,使用对称噪声率为 0.3、0.5、0.7 的设置。为了确保公平比较,所有实验都使用固定的迭代估计次数 $m = 3$ 进行,详细结果如图 4 所示。与传统校正方法相比,困难样本选择方法可以帮助模型在大多数校正中实现更高的校正准确性和更低的错误校正率。具体来说,在较高的噪声条件下,本研究提出的的框架显著获得了更高的伪标签准确性。在 CIFAR-10 上,当噪声率为 0.9 时,该框架比 DivideMix 高出约 7.1%,比 Cross-Entropy 高出约 19.2%。在 CIFAR-100 上,当噪声率为 0.7 时,该框架比 DivideMix 高出约 5.8%,比 Cross-Entropy 高出约 15.6%。这些结果与在2个数据集上噪声率为 0.3 时的平均改进 0.25% 相比更为显著。本研究相信这是因为在高噪声率条件下,存在更多对模型影响较大的困难样本,因此,针对困难样本的方法带来的改进更为明显。

值得指出的是,具有更高校正准确性和更低错误校正率的模型可以帮助模型更快地执行样本校正并减少最终的泛化误差,这对于标签校正至关重要。本研究提出的框架在困难样本的伪标签推理准确性方面显示出显著的改进,正如事实所表明的

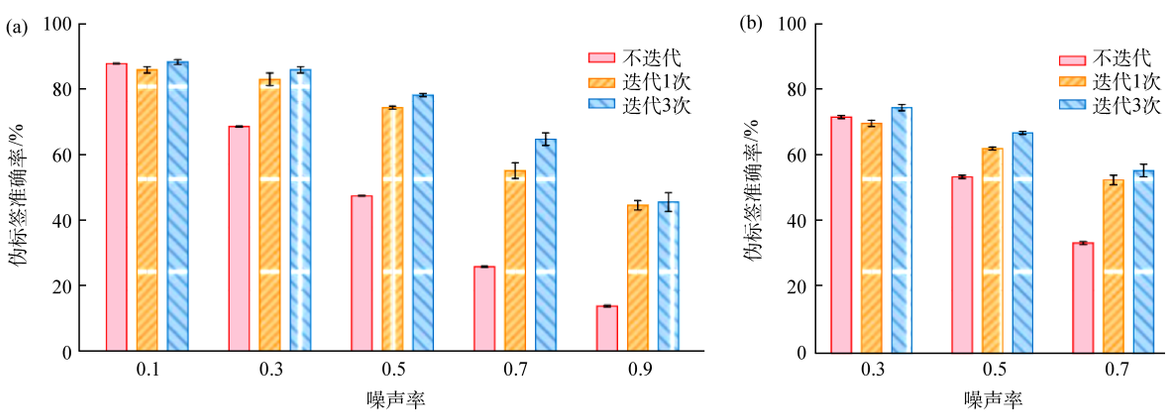


图 4　不同迭代次数下2种数据集上对称噪声设置下的伪标签准确性比较

(a)CIFAR-10;(b)CIFAR-100

Figure 4　comparison of pseudo-label accuracy under symmetric noise settings on two datasets with different iteration counts

(a)CIFAR-10;(b)CIFAR-100

那样,在仅3次迭代后,该框架在任何噪声比率下都实现了伪标签准确性的增加。这表明该框架在困难样本的推理准确性方面有显著的改进,最终导致比所有其他标签校正方法更好的最终分类性能。

3.3　对迭代困难样本衰减方法的验证

本节重点研究HEAL-Weight。困难样本的伪标签估计准确性较低严重影响了其标签校正。因此,本研究设计了2项实验来展示HEALON对困难样本影响的实际衰减能力和困难样本衰减对实际校正过程的影响。

首先,为了评估困难样本的衰减能力,使用累积难度值S^m>1的样本比例来表示在m次估计后难度较低的样本。接下来,将其与并行困难样本衰减方法在0.1、0.3、0.5和0.7对称噪声条件下的衰减效果进行比较。在CIFAR-100上,使用0.1、0.3和0.5的噪声率,同时保持其余设置和超参数与之前相同。其次,为了评估困难样本衰减对实际校正过程的影响,观察了过滤伪标签U和校正伪标签U'在每个epoch的准确性差异变化。2项实验在每个epoch的迭代次数不同,即$m_{\max}$分别设置为1和3。相同颜色的线表示2次运行的相同值,2条线之间的区域表示校正的数据量。

在本实验中,没有使用MixMatch。相反,直接使用模型的预测结果来生成伪标签。通过这样做,U'和U之间的收敛现象变得更加明显。

高置信度校正样本和细化结果如图5所示。对于第1个实验,在CIFAR-10数据集,随着迭代次数的增加,总训练数据中难度较低样本的比例迅速增加。在所有噪声条件下,当迭代次数仅为3次时,困难样本的数量减少到非常低的水平。在CIFAR-100数据集,困难样本的减少率小于CIFAR-10,但在0.5的高噪声条件下,困难样本的比例在4次迭代后从40%下降到低于10%。在第2个实验中,可以明显观察到在$m_{\max}=3$的条件下,U'和U的准确性线之间的区域大于$m_{\max}=1$的条件下,2条线相交的epoch增加6。这表明HEAL-Weight延长

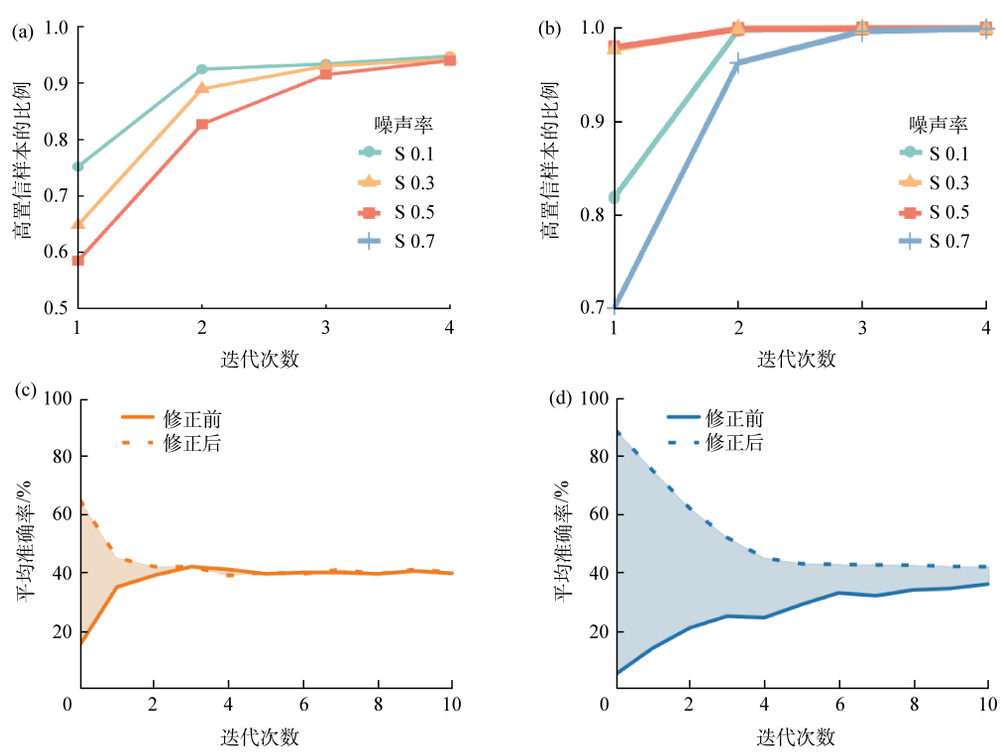


图5　高置信度校正样本和细化

(a)CIFAR-10上高置信度校正样本的比例;(b)CIFAR-100上高置信度校正样本的比例;(c)m=1时CIFAR-10上50%对称噪声下过滤和预测集差异;(d)m=3时CIFAR-10上50%对称噪声下过滤和预测集差异

Figure 5　high-confidence corrected samples and refinement

(a) proportion of high-confidence corrected samples on CIFAR-10; (b) proportion of high-confidence corrected samples on CIFAR-100; (c) m=1, differences between filtered and predicted sets under 50% symmetric noise on CIFAR-10; (d) m=3, differences between filtered and predicted sets under 50% symmetric noise on CIFAR-10

了有效标签校正的阶段,允许更多的噪声样本被校正。

因此,本研究提出的方法可以通过迭代估计最优权重来显著衰减困难样本,减少困难样本的低伪标签预测准确性对模型学习过程的影响。此外,该实验为后续参数设置提供了指导。在后续实验中,当噪声率<0.7时,选择迭代3次;否则,选择5次迭代。

3.4 与先进方法的比较

本研究与以下样本校正和样本选择方法进行了比较:(1) Co-teaching+[13],(2) DivideMix[14],(3) FINE[22],(4) TopoFilter[11],(5) TCL[29],(6) Late-Stopping[19],不同样本选择方法在各种噪声分布和数据集上的性能在2.2节均已表述。

本研究观察到HEALON在不同噪声条件下比SOTA方法表现更好。与类似的半监督噪声标签学习框架相比,DivideMix,HEALON在较高噪声条件下比其高出10%。实验结果表明,对困难样本的特殊优化,特别是在由高噪声引起的高难度情况下,可以减少学习过程中累积误差的影响,显著提高伪标签预测的准确性,从而使模型在高度嘈杂的环境中更加稳健。

不同样本选择方法在CIFAR-10和CIFAR-100数据集上,针对不同类型和比例的噪声标签的性能比较见表1。其中,CIFAR-10的噪声类型包括对称噪声和非对称噪声,含对称噪声标签的样本比例分别为20%、50%和90%,含非对称噪声的样本比例为40%。CIFAR-100仅包含对称噪声,含对称噪声标签样本比例为20%、50%和80%。表1的数值代表各方法在测试集上的准确率。

4 结束语

本研究引入HEALON,假设样本难度与伪标签估计准确性高度相关。利用这一假设指导标签校正,结合样本加权方法,迭代地细化伪标签。大多数现有的噪声标签校正方法直接依赖于模型预测,但由于困难样本的低预测准确性,模型在校正过程中可能会产生记忆偏差。本研究提出的框架通过HEAL-Predict和HEAL-Weight缓解了这一问题,减少了困难样本对噪声校正过程的影响。本研究通过实验证明,迭代难度加权伪标签聚合(HEAL-Predict)可以显著提高困难样本的标签校正效果,并减少错误校正对模型学习的影响。通过优化迭代权重求解,困难样本衰减(HEAL-Weight)可以逐步减少困难样本的数量,全局减少噪声标签对数据学习的影响,增强模型对噪声标签的鲁棒性。

尽管本方法通过迭代估计和减少噪声标签对困难样本的影响来提高鲁棒性,但在预测高度挑战性的样本时仍然遇到困难。一些与大多数样本显著不同的样本可能仍然被分配错误的伪标签,或者未被噪声过滤机制检测到,导致错误估计的困难样本。这些样本通常表现出与其真实类别中的主要样本不同的模式。即使在多次迭代估计之后,这也使得模型难以推断出真实标签。在CIFAR-10上进行4次迭代估计后,仍有63个样本的伪标签估计置信度较低。为了限制提出方法所消耗的计算资源,将大规模数据集的迭代次数($m_{\max}$)设置为相对较低的值。因此,本方法可能无法完全实现对高难度样本的准确伪标签估计,这可能导致少量此类样本被

表1 CIFAR-10和CIFAR-100在不同噪声类型和分数下的测试准确率 单位:%

Table 1 test accuracy of CIFAR-10 and CIFAR-100 under different noise types and ratios unit:%

实验设置	CIFAR-10数据集				CIFAR-100数据集		
	20%[a]	50%[a]	90%[a]	40%[b]	20%[a]	50%[a]	80%[a]
CE	82.7	57.9	16.8	85.0	61.8	37.3	18.1
TopoFilter	90.2	87.2	62.4	87.5	65.6	62.0	35.2
Co-teaching+	88.2	84.1	30.1	86.5	62.0	37.9	27.9
Late-Stopping	91.1	88.9	—	87.4	68.6	64.1	—
FINE	91.0	87.3	69.4	89.5	70.3	64.2	25.6
DivideMix	95.4	94.4	74.2	93.4	76.9	74.2	57.6
DivideMix+FINE	96.1	94.8	89.4	93.8	78.8	74.3	60.1
TCL	95.0	93.9	89.4	78.0	78.0	73.3	65.0
Ours	96.0	94.6	86.2	94.2	78.6	75.1	62.7
Ours+FINE	96.1	94.9	90.1	94.5	78.9	75.3	63.5

注:[a]指对称噪声样本比;[b]指非对称噪声样本比;—指无数据。

错误地学习这一问题。

后续计划在以下2个方面解决这个问题:一方面,将研究更高效的样本加权利用方法,以优化每次迭代的效率;另一方面,将尝试将本研究的框架与一次学习或少量学习相结合。

©The editorial department of Journal of Capital Normal University (Natural Science Edition) 2024. This is an open access article under the CC BY-NC-ND 4.0 license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

参 考 文 献

[1] Tu Y P, Zhang B S, Li Y X, et al. Learning from noisy labels with decoupled meta label purifier [PP/OL]. [2024-11-01].https://api.semanticscholar.org/CorpusID:256846984.

[2] Cao K D, Wei C, Gaidon A, et al. Learning imbalanced datasets with label-distribution-aware margin loss [PP/OL]. V2. arXiv (2019-10-27)[2024-11-01]. https://doi.org/10.48550/arXiv.1906.07413.

[3] Byrd J, Lipton Z. What is the effect of importance weighting in deep learning International conference on machine learning [J].Proceedings of Machine Learning Research, 2019(36):872-881.

[4] Wang X, Han X T, H W L, et al. Multi-similarity loss with general pair weighting for deep metric learning [C]// Proceedings of Machine Learning Research, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 15-20, 2019. Long Beach, CA, USA: IEEE, 2019: 5017-5025. DOI: 10.1109/cvpr.2019.00516.

[5] Feng C W, Ren Y L, Xie X K. Ot-filter: an optimal transport filter for learning with noisy labels [PP/OL]. [2024-11-01]. https://openaccess. thecvf. com/content/CVPR2023/papers/Feng_OT-Filter_An_Optimal_Transport_Filter_for_Learning_With_Noisy_Labels_CVPR_2023_paper.pdf.

[6] Jiang L, Zhou Z Y, Leung T, et al. MentorNet: learning data-driven curriculum for very deep neural networks on corrupted labels[PP/OL]. V2.arXiv (2018-08-13)[2024-11-01].https://doi.org/10.48550/arXiv. 1712.05055.

[7] Mirzasoleiman B, Cao K D, Leskovec J. Coresets for robust training of deep neural networks against noisy labels [J]. Advances in Neural Information Processing Systems, 2020, 33: 11465-11477.

[8] Pleiss G, Zhang T Y, Elenberg E, et al. Identifying mislabeled data using the area under the margin ranking [J]. Advances in Neural Information Processing Systems, 2020, 33: 17044-17056.

[9] Cheng H, Zhu Z W, Li X Y, et al. Learning with instance-dependent label noise: a sample sieve approach [PP/OL]. V2.arXiv (2021-03-22)[2024-11-01].https://doi.org/10.48550/arXiv.2010.02347.

[10] Huang J C, Qu L, Jia R F, et al. O2U-Net: a simple noisy label detection approach for deep neural networks [C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27-November 2, 2019. Seoul, Korea: IEEE, 2019: 3325-3333. DOI: 10.1109/iccv. 2019.00342.

[11] Wu P X, Zheng S Z, Goswami M, et al. A topological filter for learning with label noise [PP/OL]. V2. arXiv (2022-01-28) [2024-11-01]. https://doi. org/10.48550/arXiv.2012.04835.

[12] Yu X R, Han B, Yao J C, et al. How does disagreement help generalization against label corruption[PP/OL]. V3. arXiv(2019-05-12)[2024-11-01]. https://doi.org/10.48550/arXiv.1901.04215.

[13] Han B, Yao Q M, Yu X R, et al. Co-teaching: robust training of deep neural networks with extremely noisy labels [PP/OL]. V3. arXiv (2018-10-30) [2024-11-01]. https://doi.org/10.48550/arXiv.1804.06872.

[14] Li J N, Socher R, Hoi S C H. DivideMix: learning with noisy labels as semi-supervised learning [PP/OL]. V1. arXiv (2020-02-18) [2024-11-01]. https://doi. org/10.48550/arXiv.2002.07394.

[15] Song H, Kim M, Park D, et al. Learning from noisy labels with deep neural networks: a survey [J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(11): 8135-8153. DOI: 10.1109/TNNLS. 2022.3152527.

[16] Yuan S Q, Feng L, Liu T L. Early stopping against label noise without validation data[PP/OL]. V1.arXiv (2025-02-11)[2024-11-01].https://doi.org/10.48550/arXiv.2502.07551.

[17] Liu S, Niles-Weed J, Razavian N, et al [J]. Early-learning regularization prevents memorization of noisy labels [J]. Advances in Neural Information Processing Systems, 2020, 33: 20331-20342.

[18] Bai Y L, Yang Erkun, Han B, et al. Understanding and improving early stopping for learning with noisy labels [J]. Advances in Neural Information Processing Systems, 2021, 34: 24392-24403.

[19] Yuan S Q, Feng L, Liu T L. Late stopping: avoiding confidently learning from mislabeled examples [C]// 2023 IEEE/CVF International Conference on Computer

Vision（ICCV）. October 1-6，2023. Paris，France：IEEE，2023：16033-16042. DOI：10.1109/iccv51070.2023.01473.

[20] Liu T L，Tao D C. Classification with noisy labels by importance reweighting [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2016，38(3)：447-461. DOI：10.1109/TPAMI.2015.2456899.

[21] Hastie T，Rosset S，Zhu J，et al. Multi-class AdaBoost [J]. Statistics and Its Interface，2009，2(3)：349-360. DOI：10.4310/sii.2009.v2.n3.a8.

[22] Kim T，Ko J，Choi J，et al[J]. Fine samples for learning with noisy labels. Advances in Neural Information Processing Systems，2021，34：24137-24149.

[23] Shu J，Xie Q，Yi L X，et al [J]. Meta-weight-net：Learning an explicit mapping for sample weighting[J]. Advances in Neural Information Processing Systems，2019(32)：1919-1930.

[24] Zhang H Y，Cisse M，Dauphin Y N，et al. Mixup：beyond empirical risk minimization[PP/OL]. V2. arXiv（2018-04-27）[2024-11-01]. https：//doi. org/10.48550/arXiv.1710.09412.

[25] Berthelot D，Carlini N，Goodfellow I，et al. Mixmatch：A holistic approach to semi-supervised learning [J]. Advances in Neural Information Processing Systems，2019(32)：5049-5059.

[26] Tanaka D，Ikami D，Yamasaki T，et al. Joint optimization framework for learning with noisy labels [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23，2018. Salt Lake City，UT：IEEE，2018：5552-5560. DOI：10.1109/cvpr.2018.00582.

[27] Krizhevsky A. Learning multiple layers of features from tiny images [PP/OL]. [2024-11-01]. http：//www. cs. toronto.edu/~kriz/learning-features-2009-TR.pdf.

[28] He K M，Zhang X Y，Ren S Q，et al. Deep residual learning for image recognition [PP/OL]. V1. arXiv（2015-12-10）[2024-11-01]. https：//doi. org/10. 48550/arXiv.1512.03385.

[29] Huang Z Z，Zhang J P，Shan H M. Twin contrastive learning with noisy labels [PP/OL]. V1. arXiv（2023-03-13）[2024-11-01]. https：//doi. org/10. 48550/arXiv.2303.06930.

（责任编辑：马田田）

（上接第25页）

[13] Li C L，Li Y B，Delaunay J J. A novel method to synthesize highly photoactive Cu_2O microcrystalline films for use in photoelectrochemical cells [J]. ACS Applied Materials & Interfaces，2014，6(1)：480-486. DOI：10.1021/am404527q.

[14] Kuo C H，Yang Y C，Gwo S，et al. Facet-dependent and Au nanocrystal-enhanced electrical and photocatalytic properties of Au-Cu_2O core-shell heterostructures [J]. Journal of the American Chemical Society，2011，133(4)：1052-1057. DOI：10.1021/ja109182y.

[15] Lu B，Liu A P，Wu H P，et al. Hollow Au-Cu_2O core-shell nanoparticles with geometry-dependent optical properties as efficient plasmonic photocatalysts under visible light [J]. Langmuir，2016，32(12)：3085-3094. DOI：10.1021/acs.langmuir.6b00331.

[16] Zhang W W，Wang B S，Hao C C，et al. Au/Cu_2O Schottky contact heterostructures with enhanced photocatalytic activity in dye decomposition and photoelectrochemical water splitting under visible light irradiation [J]. Journal of Alloys and Compounds，2016，684：445-452. DOI：10.1016/j.jallcom.2016.05.192.

[17] Xiong J Y，Li Z，Chen J，et al. Facile synthesis of highly efficient one-dimensional plasmonic photocatalysts through Ag@Cu_2O core-shell heteronanowires [J]. ACS Applied Materials & Interfaces，2014，6(18)：15716-15725. DOI：10.1021/am502516s.

[18] Li J T，Cushing S K，Bright J，et al. Ag@Cu_2O core-shell nanoparticles as visible-light plasmonic photocatalysts [J]. ACS Catalysis，2013，3(1)：47-51. DOI：10.1021/cs300672f.

（责任编辑：兰丽丽）