



## 基于K-means聚类和特征空间增强的噪声标签深度学习算法

吕佳, 邱小龙

引用本文:

吕佳,邱小龙. 基于K-means聚类和特征空间增强的噪声标签深度学习算法[J]. *智能系统学报*, 2024, 19(2): 267-277.

LYU Jia, QIU Xiaolong. A noisy label deep learning algorithm based on K-means clustering and feature space augmentation[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(2): 267-277.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202303014>

## 您可能感兴趣的其他文章

### 一种自训练框架下的三优选半监督回归算法

Three-optimal semi-supervised regression algorithm under self-training framework  
*智能系统学报*. 2020, 15(3): 568-577 <https://dx.doi.org/10.11992/tis.201905033>

### 一种双优选的半监督回归算法

A dual-optimal semi-supervised regression algorithm  
*智能系统学报*. 2019, 14(4): 689-696 <https://dx.doi.org/10.11992/tis.201805010>

### 结合谱聚类的标记分布学习

Label distribution learning based on spectral clustering  
*智能系统学报*. 2019, 14(5): 966-973 <https://dx.doi.org/10.11992/tis.201809019>

### SUCE:基于聚类集成的半监督二分类方法

SUCE: semi-supervised binary classification based on clustering ensemble  
*智能系统学报*. 2018, 13(6): 974-980 <https://dx.doi.org/10.11992/tis.201711027>

### 一种具有迁移学习能力的RBF-NN算法及其应用

A RBF-NN algorithm with transfer learning ability and its application  
*智能系统学报*. 2018, 13(6): 959-966 <https://dx.doi.org/10.11992/tis.201705021>

### 应用k-means算法实现标记分布学习

Label distribution learning based on k-means algorithm  
*智能系统学报*. 2017, 12(3): 325-332 <https://dx.doi.org/10.11992/tis.201704024>

DOI: 10.11992/tis.202303014

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20231222.1054.002>

# 基于 K-means 聚类 and 特征空间增强的 噪声标签深度学习算法

吕佳<sup>1,2</sup>, 邱小龙<sup>1,2</sup>

(1. 重庆师范大学计算机与信息科学学院, 重庆 401331; 2. 重庆市数字农业服务工程技术研究中心, 重庆 401331)

**摘要:** 深度学习中神经网络的性能依赖于高质量的样本, 然而噪声标签会降低网络的分类准确率。为降低噪声标签对网络性能的影响, 噪声标签学习算法被提出。该算法首先将训练样本集划分成干净样本集和噪声样本集, 然后使用半监督学习算法对噪声样本集赋予伪标签。然而, 错误的伪标签以及训练样本数量不足的问题仍然限制着噪声标签学习算法性能的提升。为解决上述问题, 提出基于 K-means 聚类 and 特征空间增强的噪声标签深度学习算法。首先, 该算法利用 K-means 聚类算法对干净样本集进行标签聚类, 并根据噪声样本集与聚类中心的距离大小筛选出难以分类的噪声样本, 以提高训练样本的质量; 其次, 使用 mixup 算法扩充干净样本集和噪声样本集, 以增加训练样本的数量; 最后, 采用特征空间增强算法抑制 mixup 算法新生成的噪声样本, 从而提高网络的分类准确率。并在 CIFAR10、CIFAR100、MNIST 和 ANIMAL-10 共 4 个数据集上试验验证了该算法的有效性。

**关键词:** 噪声标签学习; 深度学习; 半监督学习; 机器学习; 神经网络; K-means 聚类; 特征空间增强; mixup 算法  
**中图分类号:** TP181 **文献标志码:** A **文章编号:** 1673-4785(2024)02-0267-11

中文引用格式: 吕佳, 邱小龙. 基于 K-means 聚类 and 特征空间增强的噪声标签深度学习算法 [J]. 智能系统学报, 2024, 19(2): 267-277.

英文引用格式: LYU Jia, QIU Xiaolong. A noisy label deep learning algorithm based on K-means clustering and feature space augmentation[J]. CAAI transactions on intelligent systems, 2024, 19(2): 267-277.

## A noisy label deep learning algorithm based on K-means clustering and feature space augmentation

LYU Jia<sup>1,2</sup>, QIU Xiaolong<sup>1,2</sup>

(1. College of Computer and Information Sciences, Chongqing Normal University, Chongqing 401331, China; 2. Chongqing Digital Agriculture Service Engineering Technology Research Center, Chongqing 401331, China)

**Abstract:** The performance of neural networks in deep learning relies on high-quality samples. However, the presence of noisy labels reduces the classification accuracy of the network. To reduce the impact of noisy labels, we propose a learning algorithm that categorizes training samples into clean and noisy subsets, assigning pseudo-labels to the noisy samples using a semisupervised learning algorithm. Despite these measures, the performance of the noisy label learning algorithm can be hindered by inaccurate pseudo-labels and a lack of sufficient training samples. To address the aforementioned problems, we propose a noisy label deep learning algorithm that leverages K-means clustering and feature space augmentation. First, the algorithm applies the K-means clustering algorithm to cluster the clean samples based on their labels. It then selects noisy samples that are difficult to classify according to the distance between the noisy samples and the cluster center. This process enhances the quality of the training samples. Second, the mix-up algorithm is used to expand both the clean and noisy samples, thereby increasing the number of training samples. Finally, a feature space augmentation algorithm is used to suppress the noise samples generated by the mix-up algorithm, leading to improved network classification accuracy. The effectiveness of the proposed algorithm has been validated on four data sets: CIFAR10, CIFAR100, MNIST, and ANIMAL-10.

**Keywords:** noisy label learning; deep learning; semisupervised learning; machine learning; neural network; K-means clustering; feature space augmentation; mix-up algorithm

收稿日期: 2023-03-07. 网络出版日期: 2023-12-25.

**基金项目:** 国家自然科学基金重大项目 (11991024); 重庆市教委“成渝地区双城经济圈建设”科技创新项目 (KJCX2020024); 重庆市高校创新研究群体资助项目 (CXQT20015); 重庆市教委科研项目重点项目 (KJZD-K202200511).

**通信作者:** 吕佳. E-mail: lvjia@cqu.edu.cn.

神经网络的性能依赖于高质量的样本, 但获取高质量的样本标签需要投入大量的时间与人力成本。为了降低样本标注的时间和成本, 通常会借助众包技术<sup>[1]</sup>, 通过雇佣众包工人来标注样

本。然而该方式会产生噪声标签,其在训练过程中会降低网络的分类准确率<sup>[2]</sup>。为了解决该问题,研究者在噪声标签学习算法中采用样本划分策略<sup>[3-5]</sup>和半监督学习算法<sup>[6-9]</sup>来提高网络的抗噪性。

噪声标签学习算法采用样本划分策略来移除噪声样本,从而降低噪声标签对网络训练的影响。Han 等<sup>[3]</sup>提出协同教学 (co-teaching, CT) 算法,该算法通过双网络模型来划分训练样本,利用网络之间的差异性来过滤噪声样本。随着网络的不断更新,双网络过滤出的噪声样本的质量会逐渐降低。为了更好地利用双网络的差异性, Yu 等<sup>[4]</sup>提出 co-teaching+算法,该算法选择双网络预测分歧的样本作为训练样本,然而分歧样本的数量会不断减少,使得网络的分类准确率降低。为了解决该问题, Wei 等<sup>[5]</sup>提出联合正则化训练算法 (joint training method with co-regularization, JOCOR), 该算法使用联合正则化损失来保持 2 个网络的一致性,通过正则化损失选择干净样本集来训练网络。然而上述算法仅使用干净样本集进行训练,忽略了噪声样本中潜在的语义信息。

噪声标签学习算法利用半监督学习的思想,将干净样本集视为有标签样本集,将噪声样本集视为无标签样本集,通过少量的有标签样本集和大量的无标签样本集来训练网络。Li 等<sup>[6]</sup>提出的 dividemix 算法首次在噪声标签算法中应用半监督学习算法。该算法首先利用高斯混合模型将训练样本集划分为干净样本集和噪声样本集,随后使用半监督学习算法纠正噪声样本集的标签,最后通过 mixup 算法扩充有效的训练样本数量。为了提高干净样本集的质量, Cordeiro 等<sup>[7]</sup>对样本进行多次预测,选取多次被预测为高置信度的样本作为干净样本,再利用 mixup 算法扩充有效的样本数量来进行半监督训练。除了静态数据存在噪声标签外,流式数据中也有类似的问题。为了解决流式数据中的噪声标签, Karim 等<sup>[8]</sup>首先建立样本缓冲区,将干净数据流和噪声数据流分别存放在干净样本缓冲区和噪声样本缓冲区中,然后同时使用弱增强算法<sup>[9]</sup>和 mixup 算法来增加 2 个缓冲区中有效训练样本数量,最后使用半监督学习算法训练样本。上述算法均使用 mixup 算法扩充训练样本数量,然而新生成的样本会带有噪声信息,使得网络在学习过程中易受到错误样本的干扰。

综上所述,噪声标签学习算法通过样本划分策略和半监督学习算法来提升网络的抗噪性,但上述工作仍存在以下问题:

1) 由于初始时网络的分类性能较弱,在使用样本划分策略时,易将边界样本错误地划分到干净样本集中,从而降低了网络的泛化性。

2) 由于 mixup 算法生成的新样本会包含新的噪声信息,使网络易受到错误样本的干扰,从而降低网络的分类准确率。

为了解决上述问题,本研究提出一种基于 K-means 聚类 and 特征空间增强的噪声标签深度学习算法 (noisy label deep learning algorithm based on K-means clustering and feature space augmentation, NLDLBKCFSA)。针对问题 1), 对噪声样本进行二次划分,采用 K-means 聚类算法对干净样本集进行聚类,根据噪声样本集与聚类中心的距离大小,筛选出难以分类的噪声样本;针对问题 2), 本研究提出了样本组合与矩交换增强 (mixup and moment exchange augmentation, MMEA) 算法,该算法首先使用 mixup 算法进行样本组合增强,增加训练样本数量,再采用特征空间增强算法生成样本的组合特征,帮助网络学习新生成样本的特征信息,从而提高网络的分类准确率。

## 1 理论基础

### 1.1 变量说明

本研究涉及到的变量以及相关说明如表 1 所示。其中,  $D = \{(x_i, y_i)\}_i^N = 1$ ,  $x_i$  表示第  $i$  个样本,  $y_i$  表示第  $i$  个样本的噪声标签,  $N$  为样本数量,  $y_i \in \{1, 2, \dots, C\}$ ,  $C$  为样本类别数。

表 1 变量描述  
Table 1 Description of variables

变量	变量说明
$D$	训练样本集
$X$	从训练样本集 $D$ 中划分出的带有正确标签的干净样本集
$U$	从训练样本集 $D$ 中划分出的带有错误标签的噪声样本集
$U_s$	从噪声样本集中 $U$ 中划分出的易分类的噪声样本集
$U_{uns}$	从噪声样本集 $U$ 中划分出来的难以分类的边界噪声样本集
$u_{uns}^s$	对边界噪声样本集 $U_{uns}$ 中的样本进行强增强后的样本
$u_{uns}^w$	对边界噪声样本集 $U_{uns}$ 中的样本进行弱增强后的样本

### 1.2 样本划分策略

通过样本划分策略比较样本损失值与噪声阈值的大小,从而筛选出噪声样本。

基于固定阈值的样本划分策略首先设置一个

阈值用于筛选噪声样本, 再使用下式计算  $x_i$  的损失值。

$$L_{CE} = - \sum_i y_i \log f(x_i; \theta) + (1 - y_i) \log(1 - f(x_i; \theta)) \quad (1)$$

其中,  $f(x_i; \theta)$  表示网络对样本  $x_i$  的预测值。

并根据损失值的大小筛选出噪声样本。然而, 过大或过小的固定阈值会使得网络产生错误的样本划分, 从而降低网络的分类准确率。

为了解决固定阈值导致的错误分类问题, Cheng 等<sup>[10]</sup> 提出动态阈值划分策略。

$$L_{CR}(x_i) = -\beta \cdot E[L_{CE}] \quad (2)$$

$$\alpha_{i,t} = \frac{1}{N} \sum L_{CE}^{(t)} + L_{CR}(x_i) \quad (3)$$

$$L = L_{CE}^{(t)} + L_{CR}(x_i) \quad (4)$$

$$v_i^{(t)} = \mathbb{1}(L < \alpha_{i,t}) \quad (5)$$

式中:  $\beta$  为样本损失的权重因子;  $\mathbb{1}(\cdot)$  为指示函数; 当  $v_i^{(t)} = 0$  时表示样本  $i$  为噪声样本,  $v_i^{(t)} = 1$  时表示样本  $i$  为干净样本。

该策略使用式 (2) 中的正则化损失  $L_{CR}$  来提高网络对干净样本的置信度, 再通过式 (3) 调整当前第  $t$  轮的噪声阈值  $\alpha_{i,t}$ , 并按照式 (4) 计算当前样本的整体损失值  $L$ , 最后根据式 (5) 比较损失值与当前噪声阈值的大小, 从中筛选出大于阈值的样本作为噪声样本。

### 1.3 数据增强

数据增强通过对样本添加微弱的标签扰动或生成新样本来丰富样本的多样性。传统的数据增强包括图像平移、转换、颜色变换等方式。Zhang 等<sup>[11]</sup> 提出了 mixup 算法。

$$\mathbf{x}' = \lambda \mathbf{x}_i + (1 - \lambda) \mathbf{x}_j \quad (6)$$

$$\mathbf{y}' = \lambda \mathbf{y}_i + (1 - \lambda) \mathbf{y}_j \quad (7)$$

其中,  $\lambda$  用于控制数据增强的比例。

mixup 算法依据式 (6) 和式 (7) 对任意 2 个样本  $(\mathbf{x}_i, \mathbf{y}_i)$  和  $(\mathbf{x}_j, \mathbf{y}_j)$  进行线性增强, 从而生成新样本  $(\mathbf{x}', \mathbf{y}')$ 。

然而, mixup 算法在增加训练样本的过程中易产生新的噪声样本, 从而误导网络学习错误的样本信息。为了降低新样本中噪声的干扰, Li 等<sup>[12]</sup> 提出空间增强算法, 该算法通过组合样本的特征矩来进行特征增强。

$$\mathbf{u} = \frac{1}{N} \sum \mathbf{x} \quad (8)$$

$$\boldsymbol{\sigma} = \sqrt{\frac{1}{N} \sum (\mathbf{x} - \mathbf{u})^2 + \varepsilon} \quad (9)$$

$$\mathbf{x}_i^{(j)} = \boldsymbol{\sigma}_j \frac{\mathbf{x}_i - \boldsymbol{\mu}_i}{\boldsymbol{\sigma}_i} + \boldsymbol{\mu}_j \quad (10)$$

式中:  $\varepsilon$  为常数;  $\mathbf{x}_i^{(j)}$  是生成的样本组合特征。首先

根据式 (8) 和式 (9) 分别提取  $\mathbf{x}_i$  和  $\mathbf{x}_j$  的一阶特征矩  $\mathbf{u}_i$ 、 $\mathbf{u}_j$  和二阶特征矩  $\boldsymbol{\sigma}_i$ 、 $\boldsymbol{\sigma}_j$ , 再根据式 (10) 将  $\mathbf{x}_i$  的特征矩与  $\mathbf{x}_j$  的特征矩进行融合以实现特征矩之间的组合, 从而协助网络学习新样本的组合特征。

## 2 本研究算法

### 2.1 NLDLBKCFSA 算法

本研究提出了一种基于 K-means 聚类和特征空间增强的噪声标签深度学习算法。由于同一类别的样本应归属于同一簇, 故利用 K-means 聚类算法对  $X$  进行聚类, 并根据  $U$  与聚类中心的距离大小筛选出难以分类的噪声样本, 从而提高训练样本的质量。为了降低 mixup 算法生成的噪声样本对网络的干扰, 使用特征空间增强算法对样本特征表示进行增强, 以提高网络对新生成样本的预测准确率。

本算法采用双网络模型来交叉更新网络参数。在每一轮迭代中, 动态调整噪声阈值进行样本划分。首先, 利用 K-means 聚类算法对  $X$  进行聚类, 从  $U$  中筛选出难以分类的边界噪声样本, 将其加入到  $U_{\text{uns}}$  中, 并将剩余的噪声样本加入到  $U_s$  中。随后使用半监督学习算法纠正  $U_s$  中的噪声标签, 并利用 MMEA 算法增强  $X$  和  $U_s$  中的样本。为了进一步利用样本中的潜在信息, 本研究对  $U_{\text{uns}}$  中的样本分别使用弱增强算法和强增强算法, 再计算增强后样本的 KL 散度值, 最后通过一致性正则化策略引导网络对同一样本的不同形态产生相近的预测值。算法 NLDLBKCFSA 的整体框架如图 1 所示。

### 2.2 边界样本筛选策略

网络初始分类性能较弱的问题易导致边界噪声样本被错分, 使得网络学习到错误的样本信息。图 2 为筛选误分类样本示意图, 其中三角形和圆形分别表示  $D$  中的 2 个类别, 实心图形表示  $X$  中的样本, 空心图形表示  $U$  中的样本。图 2(a) 为样本的初始分类状态, 虚线表示样本空间中的真实分类边界, 样本 A 和样本 B 分别为处于分类边界周围的 2 个噪声样本。由于网络的初始分类准确率较差, 在对 A 和 B 进行分类时, 网络会将 A 和 B 划分到不正确的类别中, 如图 2(b) 所示。为了解决该问题, 本研究提出了边界样本筛选策略。该策略使用 K-means 聚类算法对  $X$  进行聚类, 生成对应的聚类簇 (实心五角星为聚类中心), 再计算 A、B 与聚类中心的距离大小来筛选出  $U$  中难以分类的噪声样本, 从而提高训练样本的质量, 图 2(c) 为筛选边界样本的过程。

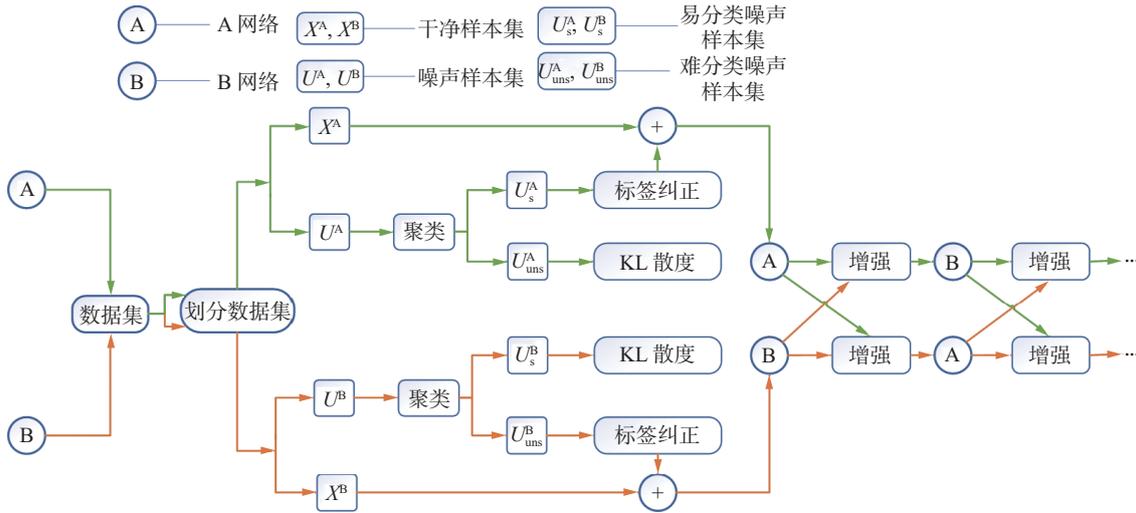


图 1 NDLBKCFS 算法框架  
Fig. 1 Framework of NDLBKCFS algorithm

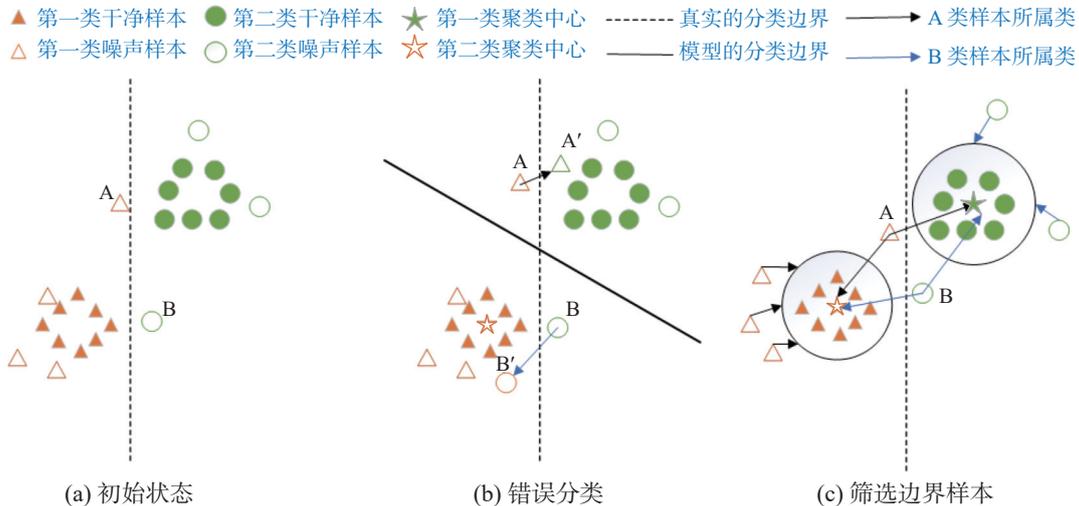


图 2 筛选误分类样本示意

Fig. 2 Schematic diagram of filtering misclassified sample

$$\widehat{D} = \frac{\sum_{i=1}^N \sum_{j=1, j \neq i}^M (x_i \times c_j)}{\sqrt{\sum_{i=1}^N (x_i)^2} \times \sqrt{\sum_{j=1, j \neq i}^M (c_j)^2}} \quad (11)$$

式 (11) 用于计算噪声样本  $(x_i, y_i)$  与聚类中心  $(c_j, y_j)$  的距离  $\widehat{D}$ 。

$$d = |\widehat{D}_1 - \widehat{D}_2| \quad (12)$$

在计算出噪声样本与聚类中心的距离后, 边界样本筛选策略会根据式 (12) 计算最近距离  $\widehat{D}_1$  和次近距离  $\widehat{D}_2$  的差的绝对值  $d$ , 然后将其与预设的边界阈值进行比较, 若  $d$  小于边界阈值, 则说明该样本会导致分类结果不可靠, 从而筛选出该噪声样本以提高训练样本质量。

### 2.3 特征空间增强

为了增加有效的训练样本数量, 本研究使用

mixup 算法对  $D$  进行增强。该算法根据式 (6) 和式 (7) 对  $D$  中任意 2 个样本进行线性组合, 生成新样本用于网络训练。图 3 为 mixup 算法过程图, 其中图 3(a) 和图 3(b) 分别为  $D$  中的 2 个训练样本, 利用 mixup 算法生成图 3(c) 的新样本。然而, Berthelot 等<sup>[13]</sup> 研究发现, mixup 算法生成的新样本是原样本间像素的简单组合, 该方式会不可避免地产生部分噪声信息, 使得网络对新样本产生错误的类别预测。例如图 3(c) 是由类别马和类别骆驼组合生成的新样本, 其真实标签为马和骆驼的组合类别, 然而网络在学习过程中易将其识别为其他类别, 从而降低了网络的分类准确率。为了解决该问题, 本研究提出了 MMEA 算法。首先使用 mixup 算法扩充训练样本数量, 接着利用特征空间增强算法进行特征组合增强, 根据式 (8) 和式 (9) 分别提取图 3(a) 和图 3(b) 的特征一阶

矩和特征二阶矩, 并按照式 (10) 对特征矩进行组合增强, 生成新的样本特征。网络通过学习 2 个样本及其组合生成的特征信息, 能够提高对新生成样本的预测, 图 4 为特征空间增强算法的过程。MMEA 算法通过 mixup 算法生成新的训练样本, 并采用特征空间增强算法生成样本的组合特征。网络通过学习该组合特征来提高对新样本的预测准确率, 从而减少新样本中的噪声信息对网络的干扰, 进而提高网络的分类准确率。



图 3 mixup 算法图像生成示意

Fig. 3 Schematic diagram of mixup algorithm image generation

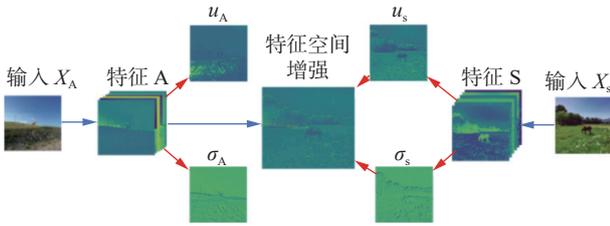


图 4 特征空间增强算法示意

Fig. 4 Schematic diagram of feature space enhancement algorithm

$$p' = \frac{1}{2} \sum_{x_i \in X} (f(x_i; \theta^{(1)}) + f(x_i; \theta^{(2)})) \quad (13)$$

$$y' = (1-r)y + rp' \quad (14)$$

式中:  $\theta^{(1)}$  和  $\theta^{(2)}$  分别为网络 1 和网络 2 的训练参数;  $r$  为样本标签的错误比例。

为了降低潜在噪声标签对网络的影响, 根据式 (13) 对  $X$  中的样本进行预测, 得到预测值  $p'$ , 并利用式 (14) 将  $p'$  与  $y$  进行组合, 得到软标签值  $y'$ , 从而提高  $X$  中样本标签的可信度。

#### 2.4 损失函数

NLDLBKCFSA 算法属于有监督学习算法, 其整体损失  $L_{\text{all}}$  分别由  $X$  的损失  $L_x$ ,  $U$  的损失  $L_u$  以及一致性正则化损失  $L_{\text{KL}}$  组成。

$$L_{\text{all}} = L_x + \lambda_u L_u + \lambda_k L_{\text{KL}} \quad (15)$$

式中:  $\lambda_u$  和  $\lambda_k$  为调整  $L_x$  和  $L_u$  的权重因子,  $L_x$  和  $L_u$  的表达式分别为

$$L_x = -\frac{1}{|X|} \sum_{i=1}^{|X|} y_i \log(f(v_i; \theta)) \quad (16)$$

$$L_u = \frac{1}{|U_s|} \sum_{j=1}^{|U_s|} \|y_j - f(u_j; \theta)\|_2^2 \quad (17)$$

其中,  $(v_i, y_i) \in X$ ,  $(u_j, y_j) \in U_s$ 。

对于  $U_{\text{uns}}$  样本集, 本研究采用一致性正则化策略<sup>[14]</sup>。

$$L_{\text{KL}} = \text{KL}(u_{\text{uns}}^s \| u_{\text{uns}}^w) = \sum u_{\text{uns}}^s \log \frac{u_{\text{uns}}^s}{u_{\text{uns}}^w} \quad (18)$$

该策略首先利用弱增强算法和强增强算法<sup>[15]</sup>对  $U_{\text{uns}}$  中的样本进行增强, 分别生成  $u_{\text{uns}}^s$  和  $u_{\text{uns}}^w$ , 再根据式 (18) 计算 KL 散度值来衡量网络对不同增强形态下样本的预测概率分布差异, 并引导网络产生一致性预测。

### 3 试验结果与分析

#### 3.1 试验设置与数据集描述

本研究选择 2 种噪声类型进行研究: 对称噪声和非对称噪声。对称噪声以  $r$  的概率将样本的真实标签随机替换成其他类型的标签, 非对称噪声以  $r$  的概率将样本的真实标签替换成指定类型的标签。本研究参照文献 [6,16] 所使用的网络训练参数和相应的  $r$  进行试验, 分别使用 20%、40%、60% 和 80% 的对称噪声率以及 10%、20%、30% 和 40% 的非对称噪声率。CIFAR10 和 CIFAR100 均使用网络 PreAct ResNet18 进行训练, 并采用随机梯度下降法, 动量为 0.9, 权重衰减为 0.0005, batch-size 大小为 64, 训练轮次为 300 epochs, 初始学习率为 0.01, 在第 150 epoch 时降低为 0.001。CIFAR10 和 CIFAR100 分别采用 10 epochs 和 30 epochs 进行预热 (warm-up) 操作。ANIMAL-10N 采用 PreAct ResNet34 进行训练, 网络参数与 CIFAR10、CIFAR100 共享, batch-size 大小设置为 32, 训练轮数共 200 epochs, 初始学习率为 0.001, 并在第 100 epochs 后降低为 0.0001, warm-up 训练轮数设为 5。超参数  $\lambda_u$  从 {0, 25, 50} 中进行选择,  $\lambda_k$  值设置为 0.4。算法采用 K-means 聚类算法进行噪声样本的二次划分, 其中聚类个数  $K$  为  $X$  的类别总数。为验证本研究所提 NLDLBKCFSA 算法的有效性, 共选取 4 个数据集进行试验, 详细的数据集信息如表 2 所示, 其中前 3 个数据集均采用人工合成的方式将正确的样本标签转换成对称噪声标签和非对称噪声标签, ANIMAL-10N 为真实收集的噪声样本集, 该样本集的噪声率为 8%。

本研究在 Pytorch 环境详细的配置环境见表 3。

表 2 数据集描述

数据集	样本数	大小	类别数
CIFAR10 <sup>[17]</sup>	60000	32×32	10
CIFAR100 <sup>[17]</sup>	60000	32×32	100
MNIST <sup>[18]</sup>	70000	32×32	10
ANIMAL-10N <sup>[19]</sup>	60000	64×64	10

表 3 试验配置环境

Table 3 The environment of experimental configuration

配置环境	版本
编程语言	Python 3.8
CPU	Intel(R)Core(TM)i5-12400
RAM	16 GB
操作系统	Windows 10 专业版
显卡	NVIDIA 3060

3.2 对比试验

为了证明 NLDLBKCFSA 算法在对称噪声和非对称噪声下的分类准确率, 本研究选取了 13 个噪声标签学习算法进行对比试验。

1) Cross-Entropy: 使用传统的交叉熵损失进行噪声标签学习。

2) Co-teaching<sup>[3]</sup>: 训练 2 个网络进行噪声标签学习, 并且交叉更新 2 个网络的参数。

3) F-correction<sup>[20]</sup>: 使用标准的网络估计噪声转移矩阵, 并利用该矩阵纠正网络预测。

4) GCE<sup>[21]</sup>: 改进交叉熵损失来提升网络的抗噪性, 并使用改进后的损失函数进行噪声标签学习。

5) M-correction<sup>[22]</sup>: 通过拟合 BMM 模型划分  $X$  和  $U$ , 并使用划分后的样本进行噪声标签训练。

6) DivideMix<sup>[6]</sup>: 使用 GMM 模型对每个 batch 的训练样本进行划分, 并使用半监督学习算法对划分后的样本进行训练。

7) DivideMix+<sup>[23]</sup>: 使用 GMM 模型对每个 mini-batch 的训练样本进行划分, 并使用半监督学习算法对划分后的样本进行训练。

8) F-Class2Sim<sup>[16]</sup>: 将噪声样本类别转换为噪声标签之间的相似度关系, 并使用相似度关系进行噪声标签学习。

9) GCE+SR<sup>[24]</sup>: 使用空间正则化损失函数训练网络, 提高网络对噪声样本的抗噪性。

10) Taylor-CE<sup>[25]</sup>: 对交叉熵损失进行 Taylor 分解, 并使用 Taylor 损失函数进行噪声标签学习, 提高网络的抗噪性。

11) CTRR<sup>[26]</sup>: 使用对比正则化函数学习噪声样本的特征表示。

12) OVA<sup>[15]</sup>: 采用多个二进制分类器联合进行噪声标签学习任务。

13) PHuber-CE<sup>[27]</sup>: 从梯度裁剪的角度对交叉熵损失进行线性优化, 从而提高损失函数的抗噪性。

以上 13 个算法和本研究所提出的算法在 CIFAR10、CIFAR100 和 MNIST 共 3 个数据集上进行训练。试验结果如表 4-9 所示。

表 4 在对称噪声率下, 各算法在 CIFAR10 上的平均准确率

Table 4 Average accuracy of each algorithm on CIFAR10 under symmetric noise rate

%

算法	噪声率			
	20%	40%	60%	80%
Cross-Entropy	83.48±0.17	68.49±0.40	48.65±0.06	27.56±0.43
Co-teaching <sup>[3]</sup>	67.73±0.71	62.83±0.72	48.81±0.78	27.56±2.71
F-correction <sup>[20]</sup>	83.27±0.04	73.67±0.30	77.64±0.11	63.95±0.32
GCE <sup>[21]</sup>	89.72±0.10	87.75±0.05	84.11±0.26	63.95±0.32
M-correction <sup>[22]</sup>	89.72±0.10	90.09±0.68	85.90±0.22	70.57±0.85
DivideMix <sup>[6]</sup>	94.82±0.09	93.95±0.14	92.28±0.08	89.30±0.17
DivideMix+ <sup>[23]</sup>	<u>94.84±0.12</u>	<u>94.03±0.20</u>	<u>93.08±0.19</u>	<b>91.91±0.07</b>
F-Class2Simi <sup>[16]</sup>	91.38±0.19	88.22±0.19	79.45±0.53	32.95±0.01
GCE+SR <sup>[24]</sup>	87.93±0.27	84.82±0.06	77.65±0.05	51.97±1.13
Taylor-CE <sup>[25]</sup>	85.96±0.09	80.51±0.11	78.18±0.36	33.48±0.44
CTRR <sup>[20]</sup>	93.05±0.32	92.16±0.31	78.18±0.36	33.48±0.44
OVA <sup>[15]</sup>	65.38±0.42	60.84±0.68	54.55±0.82	34.42±0.61
PHuber-CE <sup>[27]</sup>	90.73±0.07	86.05±0.37	74.06±0.92	32.97±0.32
NLDLBKCFSA(本研究)	<b>95.86±0.40</b>	<b>95.27±0.13</b>	<b>94.64±0.20</b>	<u>91.83±0.48</u>

注: 加粗的为最优分类准确率, 下划线为次优分类准确率, 下同。

表 5 在对称噪声率下, 各算法在 CIFAR100 上的平均准确率

Table 5 Average accuracy of each algorithm on CIFAR100 under symmetric noise rate

%

算法	噪声率			
	20%	40%	60%	80%
Cross-Entropy	60.93±0.40	46.24±0.74	29.00±0.38	11.42±0.19
Co-teaching <sup>[3]</sup>	60.93±0.40	60.00±0.60	48.30±0.10	16.10±1.10

续表 5

算法	噪声率			
	20%	40%	60%	80%
F-correction <sup>[20]</sup>	60.49±0.29	48.93±0.21	48.74±0.41	22.93±0.78
GCE <sup>[21]</sup>	69.20±0.10	65.90±0.25	57.33±0.18	18.19±1.15
M-correction <sup>[22]</sup>	67.96±0.17	64.48±0.76	55.37±0.72	24.21±1.06
DivideMix <sup>[6]</sup>	73.17±0.28	71.01±0.16	66.61±0.18	43.25±0.82
DivideMix+ <sup>[23]</sup>	<u>73.17±0.28</u>	<u>71.03±0.32</u>	<u>67.52±0.19</u>	<b>58.07±0.71</b>
F-Class2Simi <sup>[16]</sup>	60.26±0.18	54.85±0.60	40.38±0.58	21.86±0.01
GCE+SR <sup>[24]</sup>	67.03±0.46	60.68±0.90	44.66±0.84	17.35±0.42
Taylor-CE <sup>[25]</sup>	59.11±0.11	50.99±0.09	38.31±0.12	15.96±0.31
CTRR <sup>[20]</sup>	70.09±0.45	65.32±0.20	54.20±0.34	43.69±0.28
OVA <sup>[15]</sup>	59.29±0.01	51.61±0.04	33.85±0.07	18.26±0.09
PHuber-CE <sup>[27]</sup>	57.90±0.31	52.36±0.77	37.93±0.86	13.83±0.25
NLDLBKCFSA(本研究)	<b>78.15±0.16</b>	<b>77.28±0.23</b>	<b>71.56±0.12</b>	<u>52.53±0.09</u>

表 6 在对称噪声率下, 各算法在 MNIST 上的平均准确率

Table 6 Average accuracy of each algorithm on MNIST under symmetric noise rate

%

算法	噪声率			
	20%	40%	60%	80%
Cross-Entropy	86.16±0.34	70.39±0.59	50.35±0.51	23.41±0.96
Co-teaching <sup>[3]</sup>	91.20±0.03	90.02±0.02	83.21±0.71	25.33±0.84
F-correction <sup>[20]</sup>	93.93±0.10	84.30±0.43	65.06±0.64	29.81±0.63
GCE <sup>[21]</sup>	94.36±0.11	93.61±0.17	92.46±0.20	85.04±0.66
M-correction <sup>[22]</sup>	97.25±0.03	96.63±0.04	95.07±0.08	86.19±0.42
DivideMix <sup>[6]</sup>	96.80±0.08	96.53±0.06	96.47±0.04	95.15±0.25
DivideMix+ <sup>[23]</sup>	96.83±0.06	96.79±0.06	96.69±0.03	95.91±0.10
F-Class2Simi <sup>[16]</sup>	<u>99.26±0.07</u>	<u>99.18±0.06</u>	<u>98.91±0.09</u>	80.80±0.16
GCE+SR <sup>[24]</sup>	99.13±0.07	99.06±0.02	98.84±0.09	<b>98.37±0.26</b>
Taylor-CE <sup>[25]</sup>	98.99±0.14	98.46±0.12	97.53±0.23	91.14±0.18
CTRR <sup>[20]</sup>	96.98±0.49	98.73±0.41	98.05±0.41	96.17±0.41
OVA <sup>[15]</sup>	98.10±0.02	97.48±0.01	95.88±0.13	74.21±0.59
PHuber-CE <sup>[27]</sup>	98.46±0.08	97.60±0.11	95.37±0.07	90.08±0.15
NLDLBKCFSA(本研究)	<b>99.48±0.03</b>	<b>99.25±0.04</b>	<b>99.24±0.02</b>	<u>98.21±0.15</u>

表 7 在非对称噪声率下, 各算法在 CIFAR10 上的平均准确率

Table 7 Average accuracy of each algorithm on CIFAR10 under asymmetric noise rate

%

算法	噪声率			
	10%	20%	30%	40%
Cross-Entropy	90.85±0.06	87.23±0.40	81.92±0.32	76.23±0.45
Co-teaching <sup>[3]</sup>	62.85±2.20	61.04±1.31	54.50±0.39	51.68±1.66
F-correction <sup>[20]</sup>	89.79±0.33	86.79±0.67	83.34±0.30	76.81±1.08
GCE <sup>[21]</sup>	90.40±0.09	89.30±0.13	86.89±0.22	82.60±0.17
M-correction <sup>[22]</sup>	92.28±0.12	92.13±0.17	91.38±0.11	90.43±0.23
DivideMix <sup>[6]</sup>	93.61±0.15	92.99±0.21	<u>92.82±0.28</u>	90.57±0.31
DivideMix+ <sup>[23]</sup>	<u>94.27±0.23</u>	<u>93.92±0.20</u>	91.79±0.36	<b>91.91±0.24</b>
F-Class2Simi <sup>[16]</sup>	90.50±0.02	91.24±0.27	88.28±0.02	87.79±0.36
GCE+SR <sup>[24]</sup>	89.14±0.02	87.55±0.08	84.69±0.46	79.01±0.18
Taylor-CE <sup>[25]</sup>	87.34±0.12	85.02±0.11	85.63±0.07	72.65±0.11

续表 7

算法	噪声率			
	10%	20%	30%	40%
CTRR <sup>[20]</sup>	86.82±0.38	86.42±0.39	80.97±0.39	76.89±0.41
OVA <sup>[15]</sup>	66.36±0.06	64.62±0.50	59.10±0.16	44.28±0.66
PHuber-CE <sup>[27]</sup>	87.91±0.13	84.87±0.26	79.01±0.38	72.65±0.11
NLDLBKCFSA(本研究)	<b>94.45±0.40</b>	<b>94.02±0.13</b>	<b>92.93±0.23</b>	<u>91.43±0.50</u>

表 8 在非对称噪声率下, 各算法在 CIFAR100 上的平均准确率

Table 8 Average accuracy of each algorithm on CIFAR100 under asymmetric noise rate

%

算法	噪声率			
	10%	20%	30%	40%
Cross-Entropy	68.58±0.34	68.82±0.22	53.99±0.50	44.31±0.23
Co-teaching <sup>[3]</sup>	—	63.40±0.90	57.60±0.30	49.20±0.30
F-correction <sup>[20]</sup>	68.87±0.06	64.11±0.37	56.45±0.59	46.44±0.50
GCE <sup>[21]</sup>	70.77±0.14	69.22±0.15	64.60±0.25	51.72±1.17
M-correction <sup>[22]</sup>	69.44±0.52	67.25±0.81	63.16±1.55	52.90±1.79
DivideMix <sup>[6]</sup>	<u>74.00±0.29</u>	73.28±0.42	<u>72.84±0.36</u>	54.33±0.69
DivideMix+ <sup>[23]</sup>	73.49±0.31	<u>73.30±0.22</u>	72.36±0.43	<b>55.63±0.60</b>
F-Class2Simi <sup>[16]</sup>	63.53±0.01	59.10±0.13	58.63±0.03	52.99±0.78
GCE+SR <sup>[24]</sup>	68.99±0.03	64.35±0.78	57.22±0.80	49.51±1.31
Taylor-CE <sup>[25]</sup>	60.96±0.21	55.45±0.12	45.81±0.19	35.45±0.25
CTRR <sup>[20]</sup>	56.24±0.40	49.58±0.42	43.02±0.38	36.23±0.37
OVA <sup>[15]</sup>	61.50±0.01	57.93±0.02	54.76±0.02	44.29±0.05
PHuber-CE <sup>[27]</sup>	60.07±0.09	53.30±0.10	44.39±0.14	35.36±0.13
NLDLBKCFSA(本研究)	<b>77.71±1.40</b>	<b>76.18±0.21</b>	<b>72.93±1.03</b>	<u>55.56±0.38</u>

表 9 在非对称噪声率下, 各算法在 MNIST 上的平均准确率

Table 9 Average accuracy of each algorithm on MNIST under asymmetric noise rate

%

算法	噪声率			
	10%	20%	30%	40%
Cross-Entropy	95.78±0.19	91.15±0.26	86.01±0.25	79.92±0.32
Co-teaching <sup>[3]</sup>	90.32±0.02	89.03±0.02	79.80±0.27	64.94±0.02
F-correction <sup>[20]</sup>	96.39±0.04	94.27±0.21	89.33±0.94	81.61±0.42
GCE <sup>[21]</sup>	94.61±0.13	94.43±0.07	94.00±0.12	93.42±0.12
M-correction <sup>[22]</sup>	96.74±0.03	96.70±0.10	96.67±0.07	94.85±0.40
DivideMix <sup>[6]</sup>	96.17±0.06	96.11±0.09	95.88±0.05	95.83±0.05
DivideMix+ <sup>[23]</sup>	96.67±0.04	96.66±0.07	96.50±0.04	96.46±0.04
F-Class2Simi <sup>[16]</sup>	<u>99.33±0.01</u>	<u>99.25±0.05</u>	<u>99.18±0.02</u>	<u>99.08±0.07</u>
GCE+SR <sup>[24]</sup>	99.28±0.04	99.22±0.02	99.13±0.05	<b>99.09±0.02</b>
Taylor-CE <sup>[25]</sup>	99.01±0.06	98.84±0.09	97.53±0.23	91.14±0.18
CTRR <sup>[20]</sup>	99.19±0.38	99.10±0.45	98.71±0.39	97.12±0.43
OVA <sup>[15]</sup>	98.23±0.02	97.80±0.10	94.31±0.54	80.25±1.25
PHuber-CE <sup>[27]</sup>	98.65±0.07	96.97±0.11	93.98±0.09	87.37±0.19
NLDLBKCFSA(本研究)	<b>99.36±0.07</b>	<b>99.25±0.13</b>	<b>99.20±0.05</b>	98.16±0.23

表 4-6 分别给出了 NLDLBKCFSA 算法与上述对比算法在对称噪声下的分类准确率。通过分析可以发现, 当噪声率分别为 20%、40% 和 60% 时, NLDLBKCFSA 算法的分类准确率均优于其

他噪声标签学习算法。其原因在于网络在进行样本选择的过程中, 使用 K-means 聚类筛选出难以分类的边界噪声样本, 以提高训练样本的质量。然而, 在噪声率为 80% 的情况下, 本研究提出的

算法的准确率低于次优算法。这是因为在高噪声率条件下,  $X$  中的样本类别数量会减少, 因此当使用 K-means 算法对  $X$  进行聚类时,  $K$  的数量亦减少, 这会导致部分边界噪声样本所属的聚类簇不存在, 从而无法筛选出干扰网络分类的边界噪声样本。

表 7-9 分别给出了 NLDLBKCFSA 与上述对比算法在非对称噪声下的分类准确率。分析表 7-9 发现, 当噪声率分别为 10%、20% 和 30% 时, 本算法的分类准确率优于其他噪声标签学习算法。在 CIFAR10 数据集上, 相较于其余算法中的最优算法, 本算法的分类准确率分别提高了 0.19%, 0.11% 和 0.12%, 在 CIFAR100 数据集上, 与其余算法中的最优算法相比, 本算法的分类准确率分别提高了 5.01%, 3.93% 和 0.12%, 在 MNIST 数据集上, 本算法的分类准确率相比于其余算法中的最优算法, 分别提高了 0.03%, 0.08% 和 0.09%。然而当非对称噪声率达到 40% 时, 网络的分类准确率较低于其他噪声标签学习算法。一方面, 在高噪声率条件下,  $X$  的数量会变少, 使得网络无法学习到充足的样本信息。另一方面, 随着  $U$  的数量不断增加, mixup 算法会选择 2 个噪声样本进行组合增强, 生成带有噪声标签的新样本, 如图 5 所示, 样本的错误标签会误导网络学习无关特征信息, 从而降低网络的分类准确率。

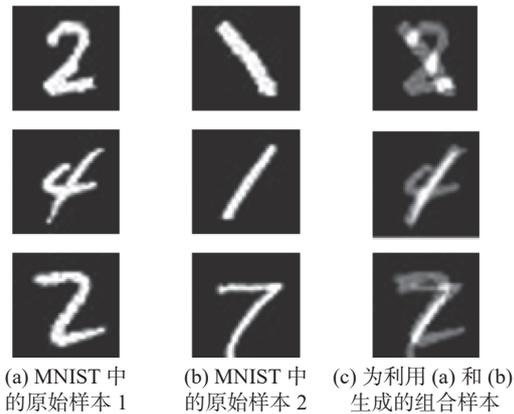


图 5 MNIST 中生成的组合样本  
Fig. 5 Mixed generated samples in MNIST

### 3.3 真实数据集对比试验

通过上述试验可知, NLDLBKCFSA 算法在人工合成的噪声样本集上表现较好。为了进一步验证本算法对真实数据集是否有效, 本研究在 ANIMAL-10 数据集上进行试验, 并将试验结果与以下 4 种噪声标签学习算法进行对比。

1) Cross-Entropy: 使用传统的交叉熵损失进行噪声标签学习。

2) SELFIE<sup>[28]</sup>: 选择高置信度的噪声样本进行标签纠正, 并将纠正后的样本与  $X$  共同进行噪声标签学习。

3) PLC<sup>[29]</sup>: 利用网络迭代的纠正噪声标签进行训练。

4) Nested-CE<sup>[30]</sup>: 对 Nested-Dropout<sup>[31]</sup> 算法施加正则化方法来对抗噪声标签。

表 10 给出了不同算法在 ANIMAL-10 上的分类准确率, 由表 10 可知, NLDLBKCFSA 算法的分类准确率优于其他噪声标签学习算法。这是因为本研究不仅使用 K-means 聚类算法获取高质量样本, 还利用 MMEA 算法提高了网络对 mixup 算法新生成样本的预测准确率, 从而引导网络具有良好的抗噪性。因此, NLDLBKCFSA 算法在真实世界的噪声标签训练集上具有较好的分类准确率。

表 10 在真实数据集 ANIMAL-10 上的分类准确率  
Table 10 Accuracy on the real data ANIMAL-10 %

算法	准确率
Cross-Entropy	79.4
SELFIE <sup>[28]</sup>	81.8
PLC <sup>[29]</sup>	83.4
Nested-CE <sup>[30]</sup>	84.1
NLDLBKCFSA(本研究)	<b>87.7</b>

### 3.4 消融试验

为了验证 NLDLBKCFSA 算法中 MMEA 算法模块的有效性, 本研究在 CIFAR10 数据集上, 分别采用 20%、40%、60% 和 80% 的对称噪声率进行消融试验, 试验结果见表 11, 其中 Baseline 为不采用 mixup 算法和空间增强算法的噪声标签学习算法。通过表 11 可知, MMEA 算法相较于 Baseline、Baseline+mixup 算法及 Baseline+feature space augmentation 均有明显的提升。原因在于本研究不仅使用 mixup 算法对训练样本集进行有效地扩充, 还使用特征空间增强算法来帮助网络学习 2 个样本的组合特征, 缓解了新样本中的噪声对网络学习的干扰, 从而进一步提高网络的分类准确率。

表 11 不同数据增强部分对网络准确率的影响  
Table 11 The impact of different data augmentation parts on network accuracy %

算法	噪声率			
	20%	40%	60%	80%
Baseline	88.21	85.18	78.61	64.69
Baseline + mixup	95.13	94.83	93.87	93.16
Baseline + feature space augmentation	92.81	87.48	85.96	85.08
NLDLBKCFSA (Baseline + MMEA)	<b>96.06</b>	<b>95.32</b>	<b>94.76</b>	<b>94.38</b>

## 4 结束语

本研究针对噪声样本在训练过程中会降低网络分类准确率的问题,提出了一种基于 K-means 聚类和特征空间增强的噪声标签深度学习算法。该算法首先对  $X$  进行 K-means 聚类,接着计算  $U$  中的样本与聚类中心的距离大小来筛选出难以分类的边界噪声样本,从而提高训练样本的质量,然后使用 mixup 算法增加有效的训练样本数量,最后使用特征空间增强算法来协助网络学习样本的组合特征,以减少新生成样本中噪声信息对网络训练的干扰,从而提高网络的分类准确率。

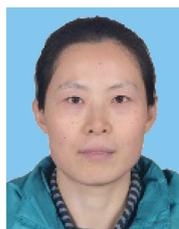
然而随着样本噪声率的提高,  $X$  和  $U$  的数量会产生显著差异,导致网络无法从  $X$  中学习到的有效的样本信息来纠正  $U$  中的噪声标签,从而降低网络的分类准确率。因此在后续工作中,将继续研究如何在高噪声条件下,利用少量的干净样本来筛选出难分类样本的噪声样本,从而提升网络的泛化性。

## 参考文献:

- [1] ZHANG Jing, WU Xindong, SHENG V S. Learning from crowdsourced labeled data: a survey[J]. *Artificial intelligence review*, 2016, 46(4): 543–576.
- [2] 伏博毅, 彭云聪, 蓝鑫, 等. 基于深度学习的标签噪声学习算法综述 [J]. *计算机应用*, 2023, 43(3): 674–684.  
FU Boyi, PENG Yuncong, LAN Xin, et al. Survey of label noise learning algorithms based on deep learning[J]. *Journal of computer applications*, 2023, 43(3): 674–684.
- [3] HAN Bo, YAO Quanming, YU Xingrui, et al. Co-teaching: robust training of deep neural networks with extremely noisy labels[C]//*Advances in Neural Information Processing Systems*. Montreal: NIPS, 2018: 1602–1613.
- [4] YU Xingrui, HAN Bo, YAO Jiangchao, et al. How does disagreement help generalization against label corruption [EB/OL]. (2019-01-14)[2022-12-25]. <https://arxiv.org/abs/1901.04215.pdf>.
- [5] WEI Hongxin, FENG Lei, CHEN Xiangyu, et al. Combating noisy labels by agreement: a joint training method with co-regularization[C]//*2020 IEEE Conference on Computer Vision and Pattern Recognition*. Washington: IEEE, 2020: 13723–13732.
- [6] LI Junnan, SOCHER R, HOI S C H. DivideMix: learning with noisy labels as semi-supervised learning[EB/OL]. (2020-02-18)[2022-12-25]. <https://arxiv.org/abs/2002.07394.pdf>.
- [7] CORDEIRO F R, SACHDEVA R, BELAGIANNIS V, et al. LongReMix: robust learning with high confidence samples in a noisy label environment[J]. *Pattern recognition*, 2023, 133(1): 565–581.
- [8] KARIM N, KHALID U, ESMAEILI A, et al. CNLL: a semi-supervised approach for continual noisy label learning[C]//*2022 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. New Orleans: IEEE, 2022: 3877–3887.
- [9] ZAHEER M Z, LEE Jinha, ASTRID M, et al. Cleaning label noise with clusters for minimally supervised anomaly detection[EB/OL]. (2021-04-30)[2022-12-25]. <https://arxiv.org/abs/2104.14770.pdf>.
- [10] CHENG Hao, ZHU Zhaowei, LI Xingyu, et al. Learning with instance-dependent label noise: a sample sieve approach[EB/OL]. (2020-10-05)[2022-12-25]. <https://arxiv.org/abs/2010.02347.pdf>.
- [11] ZHANG Hongyi, CISSE M, DAUPHIN Y N, et al. Mixup: beyond empirical risk minimization[EB/OL]. (2018-04-27)[2022-12-25]. <https://arxiv.org/abs/1710.09412.pdf>.
- [12] LI Boyi, WU F, LIM S N, et al. On feature normalization and data augmentation[C]//*2021 IEEE Conference on Computer Vision and Pattern Recognition*. Kuala Lumpur: IEEE, 2021: 12378–12387.
- [13] BERTHELOT D, RAFFEL C, ROY A, et al. Understanding and improving interpolation in autoencoders via an adversarial regularizer[EB/OL]. (2018-07-23)[2022-12-25]. <https://arxiv.org/abs/1807.07543.pdf>.
- [14] SOHN K, BERTHELOT D, CARLINI N, et al. Fixmatch: simplifying semi-supervised learning with consistency and confidence[C]//*Advances in Neural Information Processing Systems*. Addis Ababa: NIPS, 2020: 596–608.
- [15] LIU Defu, ZHAO Jiayi, WU Jinzhao, et al. Multi-category classification with label noise by robust binary loss[J]. *Neurocomputing*, 2022, 482(16): 14–26.
- [16] WU Songhua, XIA Xiaobo, LIU Tongliang, et al. Class2Simi: a noise reduction perspective on learning with noisy labels[C]//*International Conference on Machine Learning*. London: ACM, 2021: 11285–11295.
- [17] SHARMA N, JAIN V, MISHRA A. An analysis of convolutional neural networks for image classification[J]. *Procedia computer science*, 2018, 132(9): 377–384.
- [18] DENG L. The mnist database of handwritten digit images for machine learning research[J]. *IEEE signal processing magazine*, 2012, 29(6): 141–142.
- [19] TAN C, XIA J, WU L, et al. Co-learning: Learning from noisy labels with self-supervision[C]//*Proceedings of the 29th ACM International Conference on Multimedia*. Chengdu: ACM, 2021: 1405–1413.

- [20] PATRINI G, ROZZA A, MENON A K, et al. Making deep neural networks robust to label noise: a loss correction approach[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2233–2241.
- [21] ZHANG Z, SABUNCU M. Generalized cross entropy loss for training deep neural networks with noisy labels[C]//Avances in Neural Information Processing Systems. Montreal: NIPS, 2018: 11400–11411.
- [22] ARAZO E, ORTEGO D, ALBERT P, et al. Unsupervised label noise modeling and loss correction[C]//International Conference on Machine Learning. Los Angeles: ACM, 2019: 312–321.
- [23] WANG Zhuowei, JIANG Jing, HAN Bo, et al. SemiNLL: a framework of noisy-label learning by semi-supervised learning[EB/OL]. (2020–11–02)[2022–12–27]. <https://arxiv.org/abs/2012.00925.pdf>.
- [24] ZHOU Xiong, LIU Xianming, WANG Chenyang, et al. Learning with noisy labels via sparse regularization[C]//2021 IEEE International Conference on Computer Vision. Montreal: IEEE, 2022: 72–81.
- [25] FENG Lei, SHU Senlin, LIN Zhuoyi, et al. Can cross entropy loss be robust to label noise[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. Yokohama: ACM, 2020: 2206–2212.
- [26] YI Li, LIU Sheng, SHE Qi, et al. On learning contrastive representations for learning with noisy labels[C]//2022 IEEE Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 16661–16670.
- [27] MENON A K, RAWAT A S, REDDI S J, et al. Can gradient clipping mitigate label noise[C]//International Conference on Learning Representations. Addis Ababa: ACM, 2020: 6204–6231.
- [28] SONG H, KIM M, LEE J G. Selfie: Refurbishing unclean samples for robust deep learning[C]//International Conference on Machine Learning, Los Angeles: ACM, 2019: 5907–5915.
- [29] ZHANG Yikai, ZHENG Songzhu, WU Pengxiang, et al. Learning with feature dependent label noise: a progressive approach[EB/OL]. (2021–05–27)[2022–12–25]. <https://arxiv.org/abs/2103.07756.pdf>.
- [30] CHEN Yingyi, SHEN Xi, HU S X, et al. Boosting co-teaching with compression regularization for label noise[C]//2021 IEEE Conference on Computer Vision and Pattern Recognition Kuala Lumpur, IEEE, 2021: 2682–2686.
- [31] RIPPEL O, GELBART M, ADAMS R. Learning ordered representations with nested dropout[C]//International Conference on Machine Learning. Beijing: ACM, 2014: 1746–1754.

#### 作者简介:



吕佳, 教授, 博士, 主要研究方向为机器学习、数据挖掘, 中国计算机学会会员。主持或参与的国家级、省部级科研项目 20 项, 发表学术论文 70 余篇。E-mail: lvjia@cqnu.edu.cn。



邱小龙, 硕士研究生, 主要研究方向为机器学习、数据挖掘、噪声标签学习算法。E-mail: 2021210516067@stu.cqnu.edu.cn。