# Iterative Learning with Open-set Noisy Labels

2018-04-08

Yisen Wang1, 2 Weiyang Liu2 Xingjun Ma3 James Bailey3 Hongyuan Zha2 Le Song2 Shu-Tao Xia1

1Tsinghua University 2Georgia Institute of Technology 3The University of Melbourne

摘要：

拥有干净注释标签的大型数据集对于训练卷积神经网络（CNNs）是至关重要的。然而，对大规模数据进行标记非常昂贵且容易出错，甚至高质量的数据集也可能包含有噪声（不正确）的标签。现有的工作通常假设使用一个封闭数据集，即与噪声标签相关的样本拥有一个真正的类，包含在训练数据的已知类集合中。然而，对于许多应用程序来说，这样的假设太过严格，因为与噪声标签相关的样本实际上可能拥有一个在培训数据中不存在的真正类。我们将这个更复杂的场景称为开放式噪声标签问题，并表明为了做出准确的预测，它是非常重要的。为了解决这个问题，我们提出了一个新颖的迭代学习框架，用于在带有开放噪声标签的数据集上训练CNNs。我们的方法检测到噪声标签，并以迭代的方式学习深刻的辨别特征。为了从噪声标签检测中获益，我们设计了一个暹罗网络，以鼓励干净的标签和嘈杂的标签不一样。还应用了重权模块，同时强调从清洁标签中学习，减少噪声标签所造成的影响。在cifar 10、ImageNet和现实世界的嘈杂（网络搜索）数据集上的实验表明，我们所提出的模型可以在高比例的开集和封闭的噪声标签中，能够有力地训练CNNs。

## Iterative learning framework（迭代学习框架）

我们的目标是从带有噪声标签的数据集中学习辨别特征。我们提出了一个迭代学习框架，它可以逐渐从深层特征空间的干净样本中提取出噪声样本。如图3所示，我们的模型包含三个主要模块：1）迭代噪声标签检测，2）鉴别特征学习，3）权重更新。噪声标签检测使用网络的输出特性（虚线）将训练样本分成两个子集：干净的样本和噪声样本。为了从噪声标签检测中获益，我们使用一个Siamese网络来增加一个约束，强制要求干净的样本和噪声样本的表示尽可能地具有区分性。此外，根据噪声标签检测提供的置信度，对每个样品进行重加权的权值分配，以此用于强调干净样本，并在鉴别表征学习中弱化噪声样本。这种学习的鉴别表示将反过来使噪声标签检测受益。考虑到表示学习是一个迭代过程，我们进一步设计了迭代的噪声标签检测，这样就可以在迭代中联合改进识别特征学习和迭代噪声标签检测。对每个模块如何工作的简要描述如下：迭代噪声标签检测：我们基于网络的特性迭代地检测噪声标签，因为来自同一类的样本应该具有本质上的相似，而错误标签的样本通常不是。

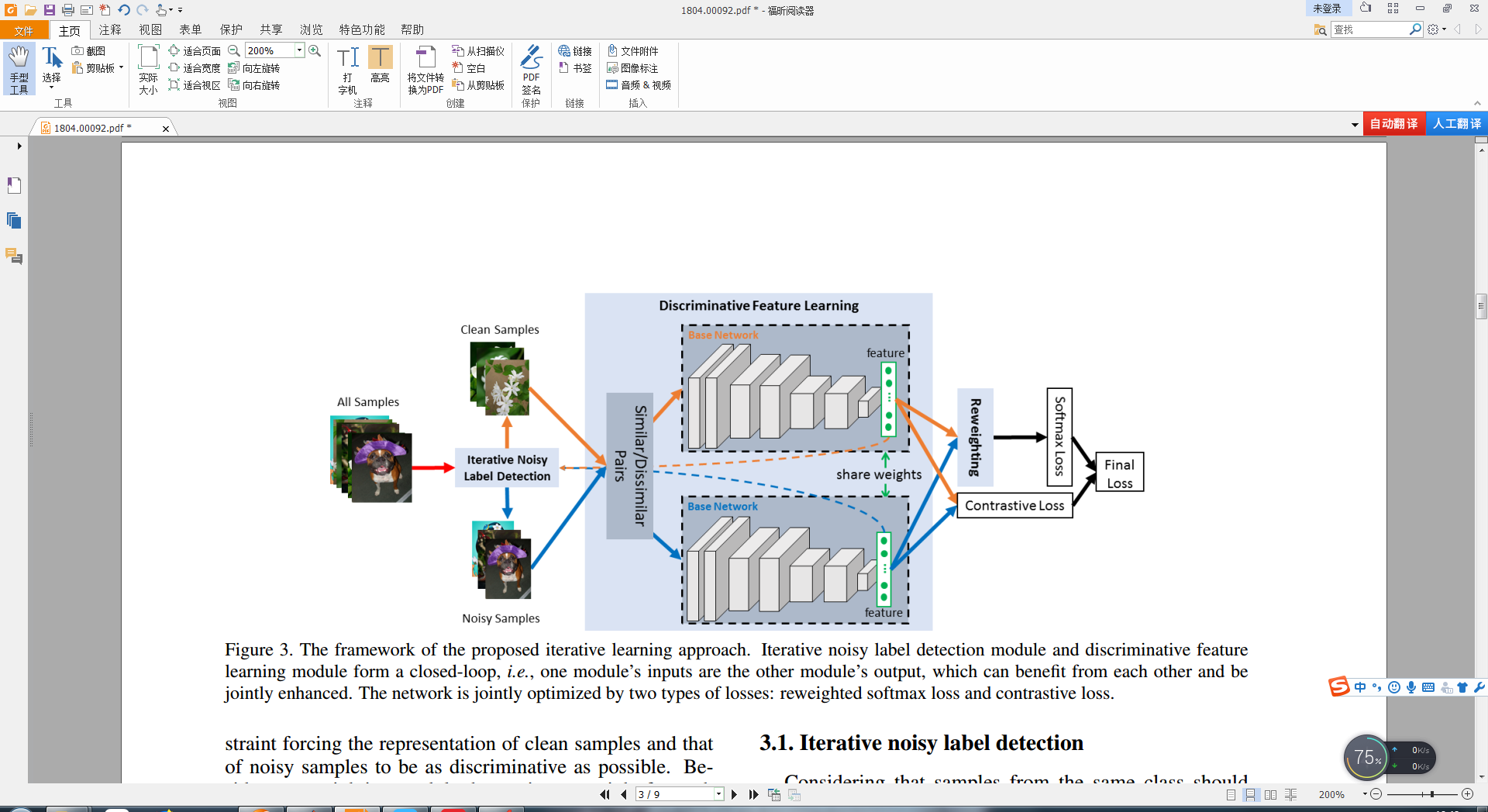
辨别特征学习：我们使用一个Siamese网络，有两个共享权重的子网络。它采用“相似”或“不同”的样本对作为输入，并使用对比损失来最小化相似样本之间的距离，并最大化不同样本之间的距离。它也可以被看作是一种表示约束。

重新加权：对于检测到的清洁样品，我们将它们的重量设置为1（没有重权），而对于检测到的噪声样本，我们根据一个样本的噪声的可能性，分别给它们分配更小的权重。为了避免误检，在决策边界附近的样品将对清洁样品具有重要的意义。

该框架由两个损失条件共同优化：



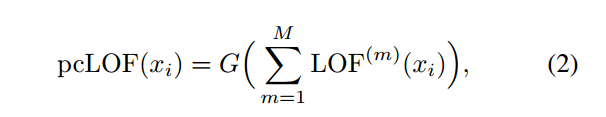
RSL是重加权的softmax损失，CL是相对的损失，是一个权衡参数。上面的目标包含了迭代的噪声标签检测，区分特征学习和重新加权，形成一个有效的学习框架，对嘈杂的标签是很有效的。



## Iterative noisy label detection（迭代噪声标签检测）

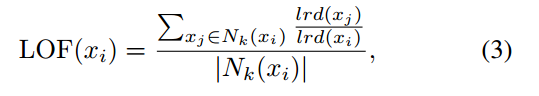
考虑到来自同一类的样本应该有类似的高级表示，但是同类中错误的标记样本则不是，我们根据pre-softmax层的表示来检测噪声标签。为了从表示的迭代学习过程中获益，我们每隔几个epoch就迭代地执行噪声标签检测。我们还使用基于所有以前的检测迭代的累积准则，以减少在一个特定迭代中随机性的影响，并进一步产生更稳定的检测结果。

我们的检测方法是局部离群因子算法（pcLOF）的一种概率及累积的版本，它继承了LOF的优点。它是一种无人监督的算法，它在高维数据上表现良好，不需要假设底层数据的分布。在形式上，pcLOF被定义为：

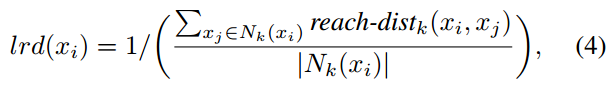


M是当前的迭代次数，G是一个局部高斯统计变换，它将累积的LOF分数扩展到[0,1]的概率值。pcLOF分数可以直接解释为样本离群值的概率。在嘈杂的标签检测设置中，pcLOF得分接近0表示一个干净的样本，而一个接近1的分数表示一个噪声样本。

LOF是一种基于密度的离群值检测算法，一个样本的LOF分数定义如下：



数knn中的集合，是的局部可达密度（lrd）：



是和间的可达距离。直观的，如果距离很远，则可达距离为简单的两者实际距离。然而，如果他们两者足够的近，实际距离将被代替（与它的第k个最近邻居样本间的距离），这意味着样本与k最近的间的距离是相等的。

注意，噪声标签检测是迭代的，因此我们不需要复杂的检测算法。由于这种表示法变得更加具有区别性，只要迭代足够长，它们就可以收敛到几乎相同的结果。为了平衡训练的效率和效果，我们在实验中在2-epoch的网络初始化后每隔10个epoch就执行基于pcLOF的迭代噪声标签检测。

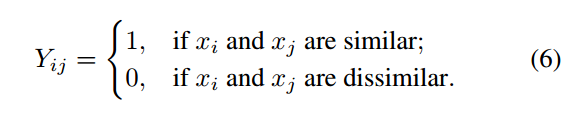
## Discriminative feature learning（辨别特征学习）

我们实现了一个有两个相同基础网络和共享权重通道的Siamese网络。它基于清洁和嘈杂的噪音标签检测模块对噪声标签样本产生了“相同”和“不同”样本,并用contrastive loss最小化同类样本之间的距离，最大化不同类别的样本之间以及清洁样本和噪声样本之间的距离。

与在深度表达空间中的欧氏距离定义为：

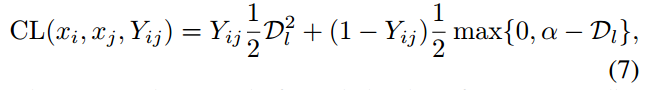


其中，为在参数下的网络中第第l层的输出。定义噪声标签检测的输出为相似标记：



两个样本被认为是“相似的”，当且仅当两个样本来自同一类并且两个样本都被正确地标记。两个样本被认为是“不同的”，如果两个样本来自不同的类，或者一个是干净的样本，另一个是一个噪声样本。注意，我们没有定义两个噪声样本之间的关系，因为它们的真实类没有包含在训练数据中，因此不能简单地定义为相似或不同。在噪声标签检测的第一次迭代之前没有噪声的样本时，不同的对只包含来自不同类的样本。

辨别特征学习的Contrastive loss 定义为：



0是一种边缘，它可以确定两个不同的样本之间的距离是多少。这种constrastive loss将关注不同的配对间的距离，例如，干净的样品和噪声的样本间的距离比边缘a更大，来自同类的相似的样本和干净的样本是聚集在一起，距离小于边缘a。

尽管可能的相似样本对数量非常巨大，但是其中的一些对很容易区分或在训练中不起作用的（例如两个不同的样本之间的距离已经大于边缘a了）。因此,我们实现了广泛使用样本最小策略，在网络中能够包含不同的样本和大部分相似样本。基本网络可以是任何类型的架构，如VGG、ResNet和Inception。

## Reweighting

为了确保有效和准确的表示学习，我们在softmax loss前设计了一个reweighting模块，以不同的信心使用标签信息。在干净的样品上应用软最大的损失是凭直觉的，利用他们值得信赖的标签信息。在噪声采样中应用软最大损失的原因是，检测到的噪声样品可能含有一些干净的样品，特别是在训练开始的时候。这些样本接近于决策边界，并且通常对13中所指出的表示学习提供了非常有用的信息，这不能被简单地忽略。把它们的权值设为0。这也通过第4.4.3节中的实验验证。因此，将一个重加权模块与软最大损失结合在一起，用于自适应地处理检测到的干净和噪声的样本