**Instance Credibility Inference for Few-Shot Learning**

会议：CVPR2020

代码：<https://github.com/Yikai-Wang/ICI-FSL>

摘要

小样本学习（FSL）的目的是通过每种类别的训练数据极其有限的方式来识别新对象。 先前的努力是通过利用元学习范式或数据扩充中的新颖原理来缓解这一极度数据稀缺的问题。 相比之下，本文提出了一种简单的统计方法，称为实例可信度推断（ICI），可以利用未标记实例的分布支持进行小样本学习。 具体来说，我们首先训练带有标记的小样本示例的线性分类器，然后使用它来推断未标记数据的伪标记。 为了衡量每个伪标记实例的可信度，我们然后建议通过增加附带参数的稀疏度来解决另一个线性回归假设，并以伪标记实例的稀疏度进行排名。 我们选择最可信赖的伪标记实例以及标记的实例，以重新训练线性分类器。 重复此过程，直到所有未标记的样本都包括在扩展的训练集中，即伪标记已收敛为未标记的数据池。 在两个FS设置下进行的大量实验表明，我们的简单方法可以在包括mini-ImageNet，tiered-ImageNet，CIFAR-FS和CUB在内的四个广泛使用的FSL基准数据集上建立新的技术水平。

**1.简介**

从一个或几个例子中学习是人类的一项重要能力。 例如，孩子们只要从书中的图片上瞥一眼，或听到它的描述就像长颈鹿一样，就可以毫无疑问地形成“长颈鹿”的概念[58]。 相反，最成功的识别系统[20、42、14、16]仍然高度依靠大量标注训练数据。 因此，这增加了稀有数据收集（例如，自动驾驶场景中的事故数据）和昂贵的数据注释（例如，用于医学诊断的疾病数据）的负担，并且从根本上限制了它们的可扩展性，从而无法扩展学习真实世界中长尾类别的开放式学习。

受这些观察结果的启发，近来对FSL的研究兴趣重新兴起[10,43,46,53]。 它旨在通过针对每个类别的训练数据来识别新对象。 基本上，FSL模型就有机会访问带有许多标记训练实例的源/基础数据集以进行模型训练，然后能够将其推广到仅具有稀缺标记数据的不相关但相关的目标/新颖数据集。 将学到的知识转移到新集的最简单基准是微调[57]。 但是，这将导致严重的过度拟合，因为一个或几个实例不足以对新颖类的数据分布进行建模。数据扩充和正则化技术可以减轻这种有限数据体制下的过度拟合，但无法解决。 通过模拟训练过程中的FSL场景，最近在利用学习学习或元学习范式方面做出了一些努力[24]。 然而，Chen等[6]从经验上讲，与线性分类器和深度特征提取器相结合的简单基线相比，这种学习范例通常会导致性能下降。

鉴于这种有限的数据机制（每个类别一个或几个带标签的示例），FSL的基本问题之一是在不引入归纳偏差的情况下几乎无法估计数据分布。 为了解决这个问题，除了传统的归纳式FSL之外，还有两种策略可以对新型类别的数据分布进行建模：（i）半监督FSL（SS-FSL）[28，37，45]假设我们可以 利用未标记的数据（大约是标记数据的十倍）来帮助学习模型； 此外，（ii）用于FSL的推论推理（TFSL）[18] [28，34]假设我们可以访问所有测试数据，而不必在推理过程中一一评估。 换句话说，FSL模型可以利用测试示例的数据分布。

自我学习[35]是利用未标记数据信息的最直接方法之一。通常，训练有素的分类器会推断未标记数据的标签，然后将其进一步用于更新分类器。 但是，推断出的伪标签可能并不总是可信赖的； 标注错误的实例可能会危害分类器的性能。 因此，必须调查每个未标记实例的标记可信度。

为此，我们提出了一种简单的统计方法，称为实例可信度推断（ICI），以利用未标记实例的分布支持进行FSL。 具体来说，我们首先训练带有标记的少量样本的线性分类器（例如，逻辑回归），然后使用它来推断未标记数据的伪标记。我们的模型旨在根据建议的ICI测量的可信度来迭代选择最可信赖的伪标签实例，以增强训练集。 因此，可以逐步更新分类器，并进一步推断未标记的数据。 我们重复此过程，直到所有未标记的样本都包含在扩展的训练集中，即对于未标记的数据池，伪标签已收敛。 示意图如图1所示。

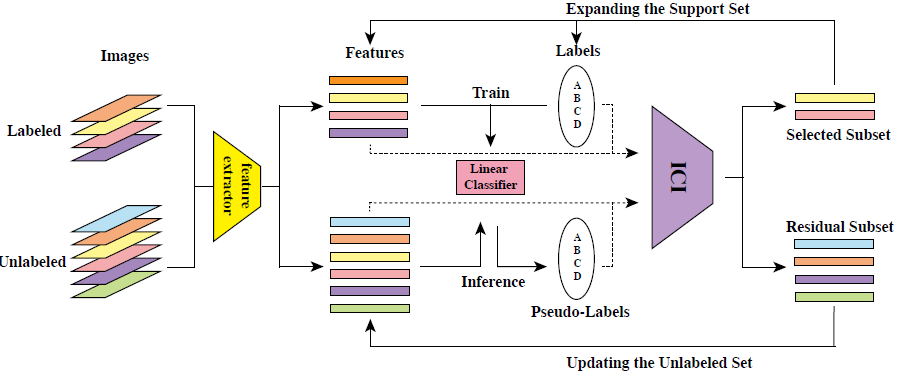


图1.我们提出的框架的示意图。 在使用未标记的数据进行N-shot FSL任务的推理过程中，我们嵌入每个实例，对每个未标记的数据进行推理，并使用ICI选择最可信赖的子集来扩展支持集。 重复此过程，直到所有未标记的数据都包括在支持集中。

基本上，我们通过ICI算法重新调整了标准自学式学习算法的用途。 如何选择伪标记的数据以排除错误预测的样本，即排除自学型学习策略引入的噪声？ 我们的直觉是，样本选择算法既不能仅依赖标签空间（例如根据分类器给出的每个类别的概率）或特征空间（例如，选择最类似于训练数据的样本）。 取而代之的是，我们通过将每个实例（带标签和带伪标签的）从特征回归到标签空间并引入偶发参数[9]的稀疏性直至消失来引入线性回归假设。 因此，我们可以将稀疏度作为伪标记实例的可信度。 我们对主要的FSL数据集进行了广泛的实验，以验证我们提出的算法的有效性。

这项工作的贡献如下：（i）我们提出了一种简单的统计方法，称为实例可信度推断（ICI），以利用未标记实例的分布支持进行FSL。 具体来说，我们的模型根据拟议的ICI用于分类器训练的可信度来迭代选择伪标记的实例。（ii）我们通过提出的ICI重新调整了标准的自学式学习算法[35]。 为了衡量每个伪标记实例的可信度，我们通过增加附带参数的稀疏性[9]来解决另一个线性回归假设，并将稀疏度作为每个伪标记实例的可信度进行排名。（iii）在两个FSL设置下进行的大量实验表明，我们的简单方法可以在包括mini-ImageNet，tiered-ImageNet，CIFARFS和CUB在内的四个广泛使用的FSL基准数据集上建立新的技术水平。

**2.相关工作**

半监督学习。 半监督学习（SSL）旨在通过以下方式提高学习成绩,通过利用大量未标记的数据来限制标记的数据。 常规方法着重于在标记和未标记的数据中找到低密度分离器[52、4、18]，并避免从未标记的数据中学习“错误”的知识[26]。 最近，具有深度学习模型的半监督学习使用一致性正则化[21]，移动平均技术[48]和对抗性扰动正则化[29]来训练带有大量未标记数据的模型。 半监督学习和使用未标记数据的FSL学习之间的主要区别在于，未标记数据在后者中仍然受到限制。 在某种程度上，在FSL情况下很难实现SSL中广泛使用的低密度假设，这使SSFSL成为一个更加困难的问题。

自学式学习[35]，也称为自训练[55]，是一种传统的半监督策​​略，利用未标记的数据来提高分类器的性能[1，12]。 通常，最初受过训练的分类器会预测未标记实例的类别标签； 进一步选择带有伪标签的未标签数据来更新分类器。 [22]。 当前基于自学学习的算法包括联合使用标记数据和伪标记数据来训练神经网络[22]，使用未标记数据和标记数据之间的混合来减少噪声的影响[2]，使用标记传播进行伪标记 基于最近邻图并使用熵[17]测量可信度，并基于特征空间上的聚类假设对伪标记的数据重新加权[40]。 不幸的是，预测的伪标签可能不可信。 与先前的重新加权或混合工作不同且正交的是，我们设计了一种统计算法来估计分配有其对应伪标签的每个实例的可信度。仅使用最自信的实例来更新分类器。

小样本学习。 关于FSL的最新努力是针对以下方面。 （1）度量学习方法着重于寻找更好的距离度量，包括加权最近邻分类器（例如，匹配网络[53]），为每个类别寻找原型（例如，原型网络[43]）或为每个任务学习特定的度量 （例如，TADAM [33]）； （2）元学习方法，例如Meta-Critic [47]，MAML [10]，Meta-SGD [27]，Reptile [32]和LEO [39]，优化了模型以快速适应新任务的能力。（3）数据增强算法扩大了可用数据，以减轻图像级别[7]或特征级别[37]中数据的缺乏。 另外，SNAIL [30]利用序列建模创建了一个新框架。 提出的统计算法是正交的，但可能对改进这些算法很有用–始终值得通过利用带有可靠预测的标记的未标记数据来增加训练集。

未标记数据的小样本学习。最近的方法通过诉诸其他未标记的数据来解决FSL问题。 具体而言，在半监督的FSL环境中，最近的工作[37，28]使来自相同类别的未标记数据可以更好地处理每个类别的真实分布。 此外，最近还考虑了转导设置。 例如，LST [45]以元学习的方式利用自学式学习策略。 与这些方法不同，本文提出了一种从自学式学习衍生而来的概念上简单的统计方法。 我们的方法仅使用非常简单的分类器（例如逻辑回归或支持向量机（SVM））就凭经验显着提高了FSL在多个基准数据集上的性能。

**3.方法论**

**3.1问题表述**

我们在这里介绍FSL的方式。 假定基本类别集Cbase和新类别集Cnovel的。 因此，基础数据集和新颖数据集分别为和。 在FSL中，应将Dbase上的识别模型推广到新颖类别Cnovel，每个类仅提供一个或几个训练示例。

为了进行评估，我们在Dnovel上采用了标准的N-way-m-shot分类[53]。 具体来说，在每个情节中，我们随机抽样N个类别L〜Cnovel； 分别在L中随机抽取每个类别的m和q个标记图像，分别构建支持集S和查询集Q。因此，我们有| S | = N×m和| Q | = N×q。 分类准确度是对许多元测试情节的查询集Q求平均的。 此外，我们还有未分类的新颖类别Unovel = {Iu}的数据。

**3.2从未标记的数据进行自学**

通常，获得用于机器学习的标记数据通常非常困难且昂贵，而未标记的数据可用于改善监督学习的性能。 因此，我们总结了自学式学习形式主义-FSL的最经典的半监督方法之一[35]。 特别地，假设f（·）是在基础数据集Dbase上训练的特征提取器。 可以在支持集S上训练监督分类器g（·），和带有分类器给出的相应置信度pi的伪标记未标记数据。 最有信心的未标记实例将进一步被用作支持集S中相应类的附加数据。因此，我们获得了更新的监督分类器。为此，FSL分类器获得了额外的训练实例，因此可以提高其性能。

但是，如果在单样本情况下直接利用self-taught，就会出现问题。 特别是，受监管的分类器g（·）仅受少数情况的训练。 具有高置信度的未标记实例可能未正确分类，分类器将因某些错误实例而更新。 更糟糕的是，不能假设未标记的实例遵循与标记数据相同的类标记或生成分布。 嘈杂的实例或离群值也可以用于更新分类器。 为此，我们提出了一种系统算法：实例可信度推理（ICI）以减少噪声。

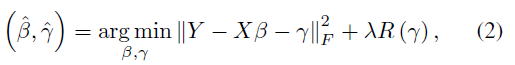
**3.3实例可信度推断（ICI）**

为了衡量未标记数据上预测标签的可信度，我们通过将每个实例从特征回归到标签空间来引入线性模型的假设。特别地，给定N个类的n个实例，S = {（Ii，yi，xi），yi∈Cnovel}，其中yi是当Ii来自支持集时的地面真理，或当Ii来自未标注时的伪标签集，我们使用一个简单的线性回归模型来“预测”类别标签，



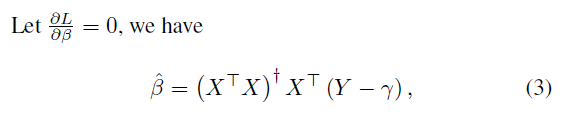
其中是分类系数矩阵；是实例i的特征向量； yi是N维one-hot向量，表示实例i的类标签。 注意，为方便计算，我们采用PCA [50]将提取的特征f（Ii）缩小为d。为零均值和σ方差的高斯噪声。 受附带参数[9]的启发，我们引入γi，j来修改实例i属于类yj的机会。越大，将实例i归于yj类的难度就越高。

写等式 1以矩阵形式表示所有实例，因此我们正在解决以下问题：

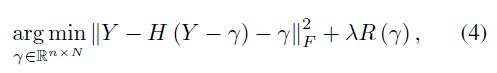


其中表示Frobenius范数。和分别表示标签和特征输入。 是偶数矩阵，惩罚。 λ是惩罚系数。 解决方程 2，我们将函数重写为

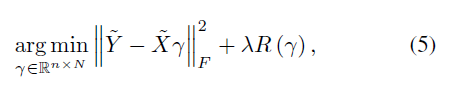




其中（·）†表示Moore-Penrose伪逆。注意（1）我们有兴趣利用γ来衡量每个实例沿其正则化路径的可信度，而不是估计ˆβ，因为线性回归模型一般而言，分类还不够好​​。 （2）ˆβ还依赖于γ的估计。 为此，我们采用等式3变成L（·）并解决问题，



其中是X的帽子矩阵。我们进一步定义。 然后可以将以上方程简化为



这是一个多响应回归问题。 我们通过检查正则化路径来寻找最佳子集，可以通过Glmnet [41]中实现的逐块下降算法轻松配置。 具体来说，我们有一个理论值 [41]来保证等式5的解全部为0。然后我们可以获得从0到λmax的λs列表。 我们解决一个特定带有每个λ的方程5，并求出γ的正则化路径。 特别地，我们认为γ是λ的函数。 当λ从0变为∞时，γ的稀疏度增加，直到所有元素都被迫消失。 此外，我们的惩罚R（γ）鼓励γ逐行消失，即逐个实例消失。此外，惩罚将趋向于以最小的偏差消失的子集，从而表明预测与基本事实之间的差异较小。 因此，当相应的γi消失时，我们可以通过伪标记数据的λ值对其进行排名。 如图2的一个玩具示例所示，用红线表示的实例的γ值首先消失，因此它是我们算法中最可信赖的样本。

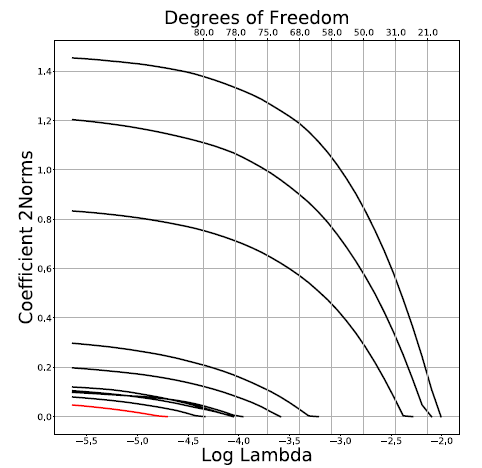
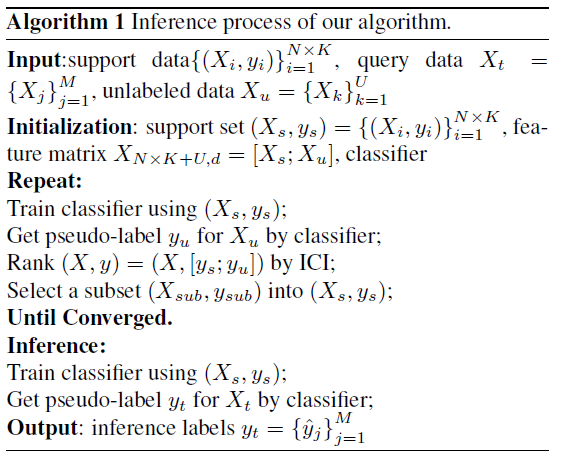


图2.十个样本的正则化路径。 红线对应于我们的ICI算法建议的最可信赖的样本。

**3.4. Self-taught learning with ICI**

所提出的ICI因此可以容易地集成以改进自学式学习算法。 特别地，初始化的分类器可以预测未标记实例的伪标签； 并且我们进一步采用ICI算法来选择未标记实例的最可信子集，以更新分类器。 整个算法可以迭代更新，如算法1所述。



**4.实验**

数据集。 我们的实验是针对通用对象识别和细粒度分类的几个广受欢迎的学习基准数据集进行的，包括mini-ImageNet [36]，tiered-ImageNet [37]，CIFAR-FS [8]和CUB [54]。 Mini-ImageNet由100个类别组成，每个类别中有600个带标签的实例。 我们遵循[36]提出的分割方法，使用64个类作为基础集来训练特征提取器，使用16个类作为验证集并报告由20个类组成的新颖集的性能。 与mini-ImageNet相比，tiered-ImageNet是一个更大的数据集，并且按类别结构选择了其类别，以在语义上拆分基础数据集和新数据集。我们遵循[37]中引入的拆分，其中包含20个超类（351个类）的基础集，6个超类（97个类）的验证集和8个超类（160个类）的新颖集。 每个类别平均包含1281张图像。 CUB是200个鸟类类别的细粒度数据集，共有11788张图像。 根据[15]中的先前设置，我们使用100个类作为基础集，使用50个类进行验证，使用50个新类。 为了进行公平的比较，我们使用[51]提供的边界框裁剪所有图像。 CIFAR-FS是从CIFAR100 [19]派生的具有较低分辨率图像的数据集。 它包含100个类，每个类中有600个实例。 我们遵循由[8]给出的拆分，使用64个类来构造基本集，使用16个类来进行验证，并使用20个类作为新集。

实验设置。 除非另有说明，否则我们将在实验中使用以下设置和实施方式来进行公平比较。与[30，33，23]中一样，我们在实验中使用具有4个残差块的ResNet-12 [13]作为特征提取器。每个块由三个3×3卷积层组成，每个卷积层之后是BatchNorm层和LeakyReLu（0.1）激活。 在每个块的末尾，使用2×2最大缓冲层来减小输出大小。 每个块中的滤波器数量分别为64、128、256和512。 具体来说，参考[23]，我们在前两个块中采用Dropout [44]来消除10％的输出，在后两个块中采用DropBlock [11]来消除通道级的10％的输出。 最后，采用平均池化层来生成输入特征嵌入。 我们从每个训练课程中选择90％的图像（例如mini-ImageNet的64个类别）以构建用于训练特征提取器的训练集，并使用剩余的10％作为验证集来选择最佳模型。 我们使用带有动量的SGD作为优化器来从头开始训练特征提取器。 动量因子和L2权重衰减分别设置为0.9和1e-4。 所有输入的大小都调整为84×84。我们将初始学习率设置为0.1，每30个周期后衰减10。 总训练时期为120个时期。 在我们所有的实验中，我们使用L2范数对特征进行归一化，并使用PCA [50]将特征尺寸减小为d = 5。 我们对模型和所有基线进行了超过600集的评估，每个类别有15个测试样本。

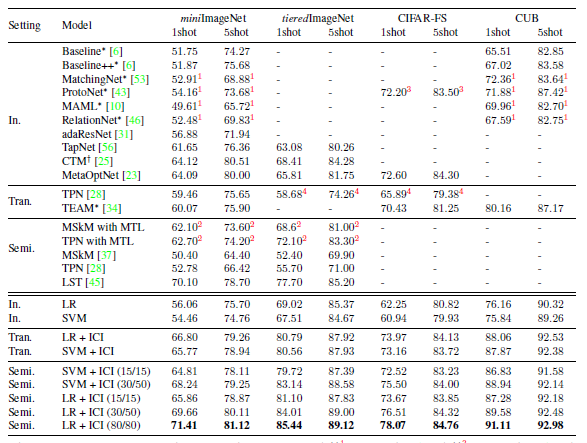


表1.在多个数据集上超过600集的测试准确性。 在[6]中报告了（·）1的结果，在[45]中报告了（·）2的结果，在[23]中报告了（·）3的结果。 （·）4是我们的实现，使用[28]的官方代码。 用（·）表示的方法表示输入大小为224×224的ResNet-18，而（·）†表示输入大小为84×84的ResNet-18。我们的方法和其他替代方法使用输入大小为84×84的ResNet-12。In和Tran分别表示归纳设置和传导设置。 semi表示半监督设置，其中（·/·）显示1次和5次实验中可用的未标记数据的数量。

**4.1. Semi-supervised few-shot learning**

设定 在推理过程中，来自相应类别池的未标记数据用于帮助FSL。 在我们的实验中，我们报告了以下SSFSL设置：（1）我们对每个类别使用15个未标记的样本（与TFSL相同），以比较我们在SSFSL和TFSL设置中的算法； （2）我们在1发任务中使用了30个未标记的样本，在5发任务中使用了50个未标记的样本，与当前的SSFSL方法相同[45]； （3）我们使用80个未标记的样本来显示ICI与具有更大网络和更高分辨率输入的FSL算法相比的有效性。我们在表1中将它们表示为（15/15），（30/50）和（80/80）。请注意，CUB是一个细粒度的数据集，没有在每个类别中有足够的样本，因此在后两个设置的5拍任务中，我们只需选择5个作为支持集，15个作为查询集，其他样本作为未标记集（平均约39个样本）。 对于所有设置，我们在每次迭代中为每个类选择5个样本。 从扩展支持集中排除每个类的最多五个实例时，该过程完成。 即总共选择（10/10），（25/45），（75/75）个未标记的实例。 此外，我们利用Logistic回归（表示为LR）和线性支持向量机（表示为SVM）来显示ICI对不同线性分类器的鲁棒性。

竞争对手。 我们将我们的算法与SS-FSL中的当前方法进行比较。 TPN [28]每次使用标签支持集和未标签集将标签传播到一个查询样本。 LST [45]还使用自学式学习策略来伪标记数据并选择可信的数据，但是它们是通过以元学习方式训练的神经网络进行多次迭代来实现的。 其他方法包括Masked Soft k-Means [37]和MTL与TPN和LST报告的Masked Soft k-Means的组合。

结果。 在表1中以“ Semi”表示在第一列中。 从实验结果分析，我们可以发现：（1）将相同数量的未标记数据与SSFSL和TFSL进行比较，我们可以看到我们的SSFSL结果仅略少与TFSL结果甚至超过TFSL结果，这表明信息 我们从未标记数据中获得的数据很健壮，实际上我们可以实际处理未标记数据的真实分布。 （2）我们获得的未标记数据越多，性能就越好。 因此，我们可以更多地从未标记数据中学习更多的知识。使用线性分类器（例如逻辑回归）。 当可以访问大量未标记的数据时，即使与使用更大网络和更高分辨率输入的竞争对手相比，ICI仍能在所有实验中达到最先进的水平。 （3）与其他SSFSL方法相比，ICI在几乎所有任务和数据集上均实现了不同程度的改进。 这些结果进一步表明了我们算法的鲁棒性。 将逻辑回归与SVM进行比较，ICI的鲁棒性仍然成立。

**4.2. Transductive few-shot learning**

设定 在Transductive few-shot learning设置中，我们有机会在推理阶段访问查询数据。因此，未标记的集和查询数据集是相同的。在我们的实验中，我们在每次迭代中为每个类别选择5个实例，然后重复我们的算法，直到包括所有预期的查询样本，即每个类别最多可扩展15张图片。 我们还将Logistic回归和SVM分别用作分类器。

竞争对手。 我们将ICI与当前的TFSL方法进行了比较。TPN [28]构造了一个图，并使用标签传播将标签从支持样本转移到查询样本，并以元学习的方式学习其框架。TEAM [34]利用具有数据相关度量的类原型来推断查询样本的标签。

结果。 在表1中显示为Tran。在第一列中。 跨四个基准数据集的实验表明：（1）与基本线性分类器相比，ICI一直得到持续的改进，尤其是在1拍设置中，其中标记数据极为有限，无论使用哪种线性分类器，这种改进都是可靠的。 此外，通过对miniImageNet和tieredImageNet的比较结果，我们可以发现改进余量在相似的范围内，这表明ICI的改进不依赖于基础集和新颖集之间的语义关系。 因此，ICI的有效性和鲁棒性在实践中得到了证实。（2）与当前的TFSL方法相比，ICI还获得了最新的结果。

**4.3消融研究**

ICI的有效性。 为了显示ICI的有效性，我们在图3的一轮推理过程中可视化了γ的正则化路径，其中红线是正确预测的实例，而黑线是错误预测的实例。 显然，大多数正确预测的实例都位于左下角。 由于ICI将选择范数将在较低的λ中消失的样本。 与错误预测的实例相比，我们可以得到更高比例的正确预测的实例。

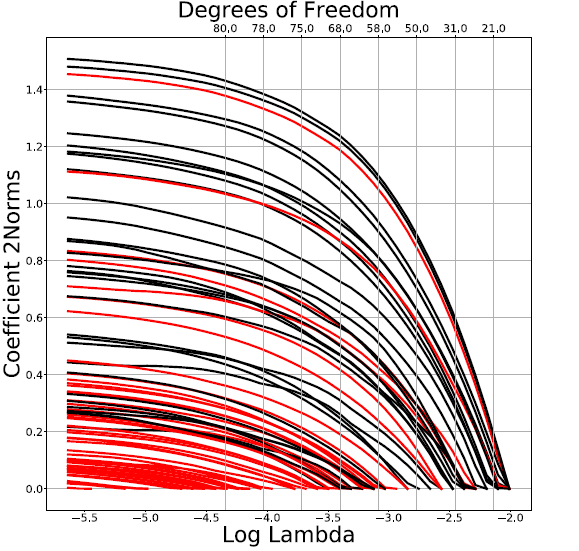


图3.的正则化路径。 红线是正确预测的实例，而黑线是错误预测的实例。 ICI将在左下子集中选择实例。

与基线进行比较。 为了进一步展示ICI的有效性，我们将ICI与自学式学习渠道下的其他样本选择策略进行了比较。一种简单的策略是在每次迭代中将未标记的数据随机采样到扩展的支持集中，表示为ra。 另一方法是基于分类器给出的置信度选择数据，用co表示。 在这种策略中，分类器对一个样本越有信心，该样本就越值得信赖。 最后一个方法是通过选择特征空间中每个类的最近邻居（表示为nn）来代替我们的可信度计算算法。 在此部分中，每个类有15个未标记的实例，并选择5个实例以通过Semi的不同方法重新训练分类器。 和特兰。 Mini-ImageNet上的任务。 从表2中可以看出，ICI在所有设置下均优于所有基准。

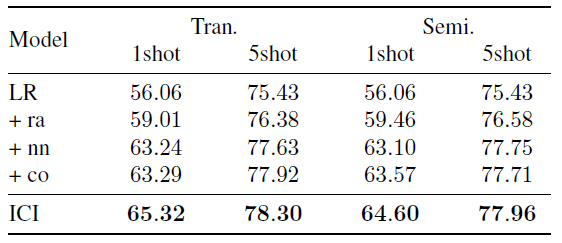


表2.在几种设置下与miniImageNet上的基线进行比较。

迭代方式的有效性。 我们的直觉是建议的ICI学会产生一套值得信赖的用于分类器训练的未标记数据。 一口气选择所有未标记的数据将无法考虑分布或未标记数据的可信度，从而产生更多的噪声标签，从而损害模型的性能。 因此，分类器将通过其预测进行训练，从而不会改善TFSL设置。 我们将其简短地验证为图4中的ICI（15），而ICI通过迭代选择方式获得了更好的精度。 例如，选择具有两次迭代（ICI（3））的6张图像优于一次迭代（ICI（8））选择8张图像。

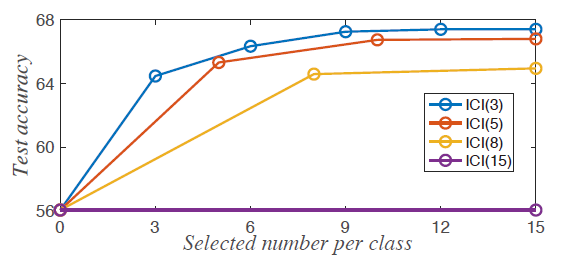


图4. miniImageNet上所选样本增加了600集以上时，准确性的变化。 “ ICI（n）”：在每次迭代中，每个类别选择n个样本。

针对初始分类器的鲁棒性。 初始线性分类器有哪些要求？ 是否需要满足初始线性分类器的精度高于50％甚至更高的要求？ 答案是不。 只要可以训练初始线性分类器，理论上我们的方法就可以工作。 因此，这是关于初始分类器如何影响的未来未解决的问题。 我们在表3中对此进行了简要验证。我们运行600个情节，每个情节训练一个初始分类器，其分类精度不同。表3显示，不管初始精度如何，大多数分类器都可以通过ICI进行改进（即使精度为3040％）。

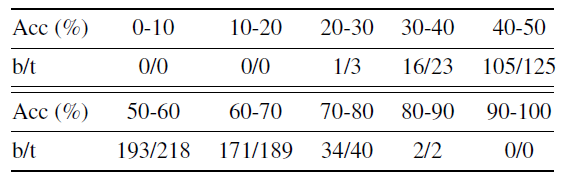


表3我们运行了600集，每集训练一个初始分类器。 我们将“ Acc”表示为精度间隔； 和“ b / T”，因为分类器数量经历了改进。 在此准确度区间内的所有分类器。

尺寸减小的影响。 在这一部分中，我们将研究算法中降维d对5次1次mini-ImageNet实验的影响。 表4中显示了具有减小的尺寸2、5、10、20、50且未减小尺寸（即d = 512）的结果。当减小尺寸维数远小于实例数（即d <<n）时，我们的算法可以获得更好的性能。这与理论性质一致[9]。 此外，我们可以观察到，当d = 5时，我们的模型可达到66.80％的最佳精度。实际上，我们在模型中采用d = 5。

降维算法的影响。 此外，我们研究了ICI对不同维度缩减算法的鲁棒性。 我们比较了Isomap [49]，主成分分析[50]（PCA），局部切线空间对齐[59]（LTSA），多维缩放[5]（MDS），局部线性嵌入[38]（LLE）和光谱 将[3]（SE）嵌入到5路1次mini-ImageNet实验中。从表4中，我们可以看到，除MDS（59.99％）外，ICI在大多数降维算法（从LTAS 64.61％到SE 67.7％）中均很可靠。 我们采用PCA来减少尺寸。

**5.结论**

在本文中，我们提出了一种简单的方法，称为实例可信度推断（ICI），以利用未标记实例的分布支持进行FSL。提出的ICI根据其可信度有效选择了最可信赖的伪标记实例，以扩大训练集。 为了衡量每个伪标记实例的可信度，我们建议通过增加附带参数的稀疏性来解决线性回归假设[9]，并用其稀疏度对伪标记实例进行排序。 大量的实验表明，我们的简单方法可以在包括mini-ImageNet，tiered-ImageNet，CIFARFS和CUB在内的四个广泛使用的少用学习基准数据集上建立新的技术水平。