arXiv:2304.04059v2 [cs.CV] 2 Jul 2024

用于医学图像分类的通用半监督学习

LieJu[1,2,3],Yicheng Wu[3,4], Wei Feng[1,2,3], Zhen Yu[1,2,3],
Lin Wang[1,3,5],Zhuoting Zhu[6], 和 Zongyuan Ge[1,2,3](✉)

[1] 莫纳什-爱迪奥克研究中心，莫纳什大学，墨尔本，澳大利亚[2] 电子研究中心，莫纳什大学，墨尔本，澳大利亚[3] 莫纳什医学人工智能组，莫纳什大学，墨尔本，澳大利亚[4] 信息技术学院，莫纳什大学，墨尔本，澳大利亚[5] 哈尔滨工程大学，哈尔滨，中国[6] 澳大利亚眼科研究中心，墨尔本大学，墨尔本，澳大利亚
https://www.monash.edu/mmai-groupLie.Ju1@monash.edu,
zongyuan.ge@monash.edu

摘要。半监督学习（SSL）因降低了收集充足的高质量有标签训练数据的高昂成本，尤其是对于深度学习方法而言，而受到广泛关注。然而，传统半监督学习建立在这样一个假设之上：有标签数据和无标签数据应来自相同的分布 *e.g.*、类别和域。然而，在实际场景中，无标签数据可能来自未见类别或未见域，现有半监督学习方法仍难以利用这些数据。因此，本文提出了一个统一框架，以利用这些未见的无标签数据进行开放场景半监督医学图像分类。我们首先设计了一种新颖的评分机制，称为双路径离群样本估计，用于识别来自未见类别的样本。同时，为了提取未见域样本，我们进一步采用了有效的变分自编码器（VAE）预训练。之后，我们进行域适应，以充分挖掘所检测未见域样本的价值，从而提升半监督训练效果。我们在皮肤病学和眼科学任务上评估了所提出的框架。大量实验表明，我们的模型在各种医学半监督学习场景中均能取得优异的分类性能。代码实现可访问： https://github.com/PyJulie/USSL4MIC。

关键词： 半监督学习 · 开放集 · 皮肤病学 · 眼科学

1 引言

由于收集足够的高质量标注数据成本高昂，训练一个令人满意的深度模型来完成医学分类任务仍然极具挑战。因此，半监督学习（自监督学习）[15,1,25,16,19] 成为一种利用无标签数据、仅需少量标注数据的流行技术。从本质上说，大多数现有的自监督学习方法都基于这样一个假设：有标签

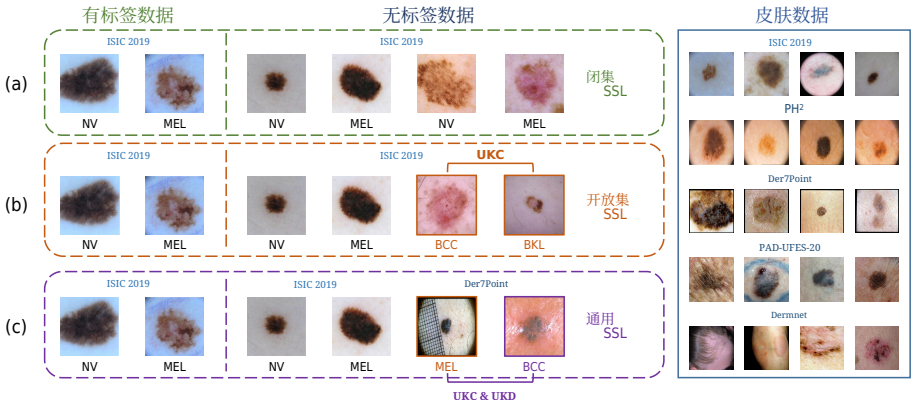

图1：问题示意图。(a) 封闭集半监督学习。有标签数据和无标签数据中的样本属于相同类别，且是在相同环境下采集的，即皮肤镜。(b) 开放集半监督学习。无标签数据中存在未知类别（UKC），例如基底细胞癌和 BKL。(c) 通用半监督学习。除了未知类别外，无标签数据中的样本还可能来自其他未知域（UKD），例如来自其他数据集、成像和条件设置不同的样本。

而无标签数据应来自同一闭集分布，并忽视现实场景。然而，在实际临床任务中（例如，皮肤病变分类），无标签数据可能包含来自未知/开放集类别的样本，而这些类别未出现在训练集中，从而导致次优性能。

我们在图 1 中进一步说明这一问题。具体而言，图 1-(a) 展示了一个经典的封闭集半监督学习设置：来自 ISIC 2019 数据集的有标签数据和无标签数据 [10] 具有相同的类别，即黑素细胞痣（NV）和黑色素瘤（MEL）。图 1-(b) 展示了开放集半监督学习的情形，**开放集半监督学习**，其中引入了基底细胞癌（BCC）和良性角化病（BKL）这两个新类别。近期关于开放集半监督学习的研究 [23,17] 主要集中于在模型训练过程中识别这些离群样本。与之不同的是，本文进一步考虑了一种更为现实的场景，该场景同样严重违背了上述封闭集假设，如图 1-(c) 所示。无标签数据可能具有相同的类别，但来自截然不同的领域，例如，来自 Der7point 数据集的 MEL [5], ，该数据集除皮肤镜图像外，还包含临床图像。同时，还存在一些未知的新类别，例如，来自 Der7point 数据集的 BCC，从而导致已见/未见类别**和**领域**不匹配**。

为处理这一不匹配问题，*Huang* 等人 [9] 提出了 CAFA，将开放集识别（OSR）与域适应（DA）相结合，即**通用 SSL**。具体而言，他们提出评估一个样本属于未知类别（UKC）或未知域（UKD）的可能性，并利用该可能性对未标注样本进行重新加权。域适应项可以将未知域中的特征适应到已知域中，从而确保模型

可以充分挖掘未知域样本的价值。CAFA 的有效性在很大程度上依赖于开放集样本的检测，而所提出的技术在医学数据集上始终难以实现泛化。对于医学图像，例如皮肤数据，UKC 和未知域样本可能高度不可分（例如，图 1 - (a) 中的 MEL 与图 1 - (b) 中的基底细胞癌），尤其是在半监督设置下使用有限样本进行训练时。

因此，在本文中，我们针对类别和领域不匹配问题，提出了一种用于医学图像分类的新型通用半监督框架。具体而言，为了衡量一个未标注样本属于 UKC 的可能性，我们提出了一种双路径异常值估计技术，利用原型和预测置信度，分别在特征层面和分类器层面衡量未标注样本属于 UKC 的可能性。此外，我们首次提出了一种评分机制，通过预训练变分自编码器（VAE）模型来衡量未标注样本属于 UKD 的可能性；该模型更适用于医学图像域分离，并且所需的有标签样本更少。针对检测出的 UKD 样本，我们采用域适应方法对不同域进行特征匹配。随后，有标签样本和未标注样本（包括经过特征适应的 UKD 样本）即可利用传统的半监督学习技术进行优化。

我们的贡献可总结如下：(1) 我们提出了一种用于通用半监督医学图像分类的新型框架，使模型能够利用开放集识别和域适应技术从未知类别/域中学习。(2) 我们提出了一种新型评分机制，以提高从 UKC/UKD 中检测离群样本的可靠性，从而实现进一步的统一训练。(3) 在具有各种模态的数据集上的实验表明，我们提出的方法能够在不同的开放集半监督学习场景中取得良好表现。

2 方法

2.1 概述

我们提出的框架概览如图 2 所示。该框架主要包含一个特征提取器 $\mathcal{F}$、一个对抗判别器 $\mathcal{D}$、一个多类分类器 $\mathcal{C}$以及一个非对抗判别器 $\mathcal{D}'$。特征提取器将输入 $\mathcal{X}$ 编码为特征 $\mathcal{V}$。多类分类器 $\mathcal{C}$ 输出确切疾病的预测结果。非对抗判别器预测无标签数据中的实例属于 UKD 的可能性。对抗判别器对来自已知域和检测到的未知域的样本进行特征适应。总而言之，我们的目标是对来自未见类别/领域的未标注样本进行评分，以便进一步通过自监督学习和域适应进行训练。

2.2 双路径离群值估计

近期的开放集半监督学习方法 [23,8] 主要关注 UKC 样本的检测，这被称为开放集识别任务。这些离群样本将在训练阶段被移除。在本节中，我们提出了一种新颖的开放集识别技术，即

基于特征相似度和分类器预测置信度评估 UKC 的双路径离群值估计（DOE）。形式上，给定标注样本 $\mathcal{X}_l$，我们首先使用标准交叉熵损失对模型进行预热。不同于 CAFA [9],（其计算实例级特征相似度），我们认为，来自已知类别的样本与中心表示（例如原型）之间的距离应小于离群样本与其之间的距离。一个类别的原型可以计算为其对应样本的平均输出$x_{l,\ i} \in \mathcal{X}_l$:

$$\mathbf{v}_{l,c_j} = \frac{\sum_{i=1, x_{l,i} \in \mathcal{X}_{l,c_j}}^{N_{c_j}} \mathcal{F}(x_{l,i})}{N_{c_j}}, \tag{1}$$

其中 N_{c_j} 表示类别 j 的实例数量，$\mathbf{v}^{c_j}$ 是经过全局池化层后形状为 $1\times$D 的向量。然后，可以计算实例 $x_{u,\ i} \in \mathcal{X}_U$ 与每个已知类别的特征相似度:

$$\mathbf{d} = \{d_{i,x_i \in c_j}\}_{j=1}^{N_{c_j}} = \{\left\|\mathcal{F}(x_{u,i|c_j}) - \mathbf{v}_{l,c_j}\right\|_2\}_{j=1}^{N_{c_j}}. \tag{2}$$

我们可以假设，如果某个样本与所有类别特定原型的距离相对较远，那么它应具有更大的平均值 d_{avg} 距离 $\mathbf{d}$ ，并可被视为潜在的离群样本 [13,20,28]。随后，我们对无标签输入进行强增强，并生成两个视图 $x_{u'_{i,1}}$ 和 $x_{u'_{i,2}}$，然后将它们一同输入预训练网络，并获得预测 $\mathbf{p}^{u_{i,1}}$ 和 $\mathbf{p}^{u_{i,2}}$。受一致性最大化原则 [24], 启发，可以根据这两个预测的一致性来确定样本是否为离群样本:

$$p_{ood}(x_i) = |\max\left(\mathbf{p}_{u_{i,1}}\right) - \max\left(\mathbf{p}_{u_{i,2}}\right)|. \tag{3}$$

最后，我们将基于原型的分数和基于预测的分数相结合:

$$w_{u,c} = 1 - \sigma(d_{avg} \cdot p_{ood}), \tag{4}$$

其中， σ 是一个将原始分布映射到（0,1]的归一化函数。

2.3 类别无关的域分离

尽管我们提出的 DOE 有助于检测潜在的未知类别样本，但一个显而易见的问题是，域差异很容易干扰对未知类别样本的估计，例如，未知域样本与已知域原型的距离更大。与检测未知类别样本不同，在无标签数据中区分未知域样本并不困难，因为不同域之间存在更大的环境差异，例如成像设备、模态或其他伪影。为此，我们采用变分自编码器，它不依赖监督信号，能够更加关注该域的全局风格 [26]。形式上，变分自编码器由编码器 $g(\cdot)$ 和解码器 $f(\cdot)$ 组成，其中编码器将高维输入特征 x_i 压缩到低维嵌入空间，而解码器则通过最小化误差来重构输入:

$$\mathcal{L}_{re} = \|x_i - f(g(x_i))\|_2^2. \tag{5}$$

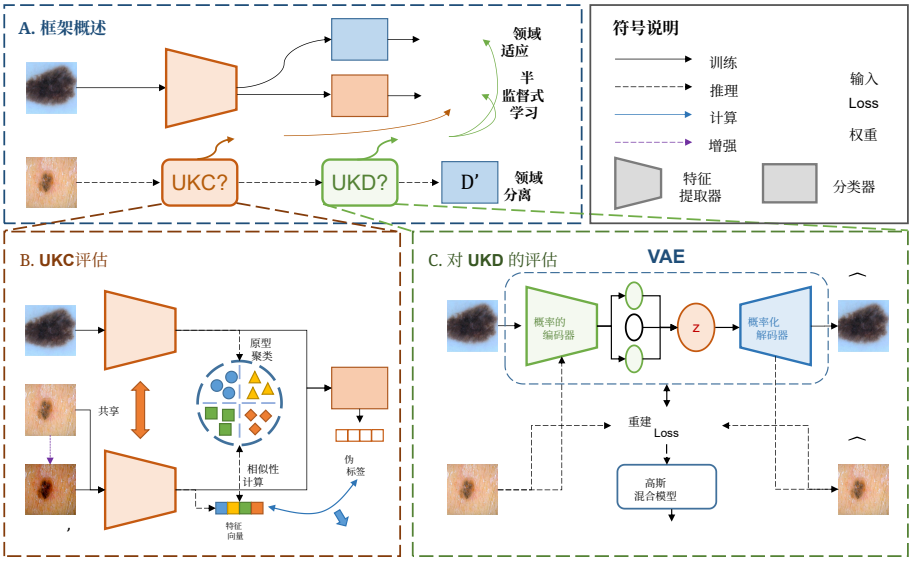

图 2：我们提出的框架概览。

在我们的场景中，我们使用有标签数据预训练变分自编码器模型，并在无标签数据上对其进行评估，以获得重构误差。随后，我们使用期望最大化算法拟合一个包含两个成分的高斯混合模型（GMM）。该模型能够灵活调整分布的尖锐程度，并且对低维分布更加敏感 [18], 即，重构损失 $\mathcal{L}_{re}$。对于每个样本 x_i，我们得到其后验概率 $w_{d,\,i}$ 用于领域分离。利用来自有标签数据的已知域样本（记为 $y_{l,\,d,\,i}=0$）以及来自无标签数据的潜在未知类别样本（记为 $y_{u,\,d,\,j}=w_{d,\,j}$），我们优化非对抗判别器的二元交叉熵损失 $\mathcal{D}'$:

$$\mathcal{L}_{dom}=-\frac{1}{N_l}\sum_{i=1}^{N_l}1-\log(\hat{y}_{l,d,i})-\frac{1}{N_u}\sum_{j=1}^{N_u}w_{d,j}\cdot\log(\hat{y}_{u,d,j})-(1-w_{d,j})\cdot(\log(1-\hat{y}_{u,d,j})). \tag{6}$$

2.4 对抗训练与自监督学习优化

为了对不同的未知域进行域适应，我们采用常规对抗训练方式 [2], ，将有标签数据视为目标领域，将无标签数据视为源领域。请注意，我们使用两个权重$w'_{d,\,u}$，分别来自非对抗判别器和 $w_{c,\,u}$ DOE，以确定要适应哪些样本。对抗损失可表述为:

$$\max_{\theta_F}\min_{\theta_D}\mathcal{L}_{adv}=-(1-y_t)\cdot\log(1-D(F(\mathcal{X}_l)))-y_s\cdot w'_{u,d}\cdot w_{u,c}\cdot\log D(F(\mathcal{X}_u)), \tag{7}$$

其中， θ 表示特定模块的参数， $y_s = 1$ 和 $y_t = 0$ 分别表示源领域和目标领域的初始领域标签。随后，我们可以在权重控制下，利用有标签数据和经选择性特征适配的无标签数据进行统一训练。总损失可表示为：

$$\mathcal{L}_{overall} = \mathcal{L}_{CE}(\mathcal{X}_l) - \alpha \cdot \mathcal{L}_{adv}(\mathcal{X}_u | w'_{u,d}, w_{u,c}) + \beta \cdot \mathcal{L}_{SSL}(\mathcal{X}_u | w_{u,c}), \quad (8)$$

其中 α 和 β 为系数。对于半监督项 $\mathcal{L}_{SSL}$，我们采用Π-model [15] 。因此，我们可以进行全局优化，更好地利用存在类别/领域不匹配的无标签数据。

3 实验

3.1 数据集与实现细节

皮肤病学 对于皮肤病变识别，我们使用四个数据集评估所提方法：ISIC 2019 [10], PAD-UFES-20 [22], Derm7pt [14] 和 Dermnet [4]。四个数据集的统计信息见补充材料。ISIC2019 中的图像由皮肤镜拍摄。PAD-UFES-20 和 Dermnet 数据集中的图像采集自临床场景，而 Derm7pt 数据集则同时包含这两种来源。首先，我们将 ISIC2019 数据集划分为 4 个（NV、MEL、BCC、BKL） + 4 （AK、SCC、VASC、DF）类别，分别作为已知类别和未知类别。我们从已知类别中每类抽取 500 个样本，构建有标签数据集。随后，我们从已知类别中每类抽取 250 个样本，分别构建验证集和测试集。我们从剩余的 17,331 个样本中抽取 30% 的闭集样本和全部开放集样本，构成无标签数据集。对于其他三个数据集，我们将每个数据集分别与 ISIC 2019 无标签数据集混合，以验证所提方法在不同未知域数据上训练的有效性。

眼科学我们还在自建眼底数据集上评估了所提出的方法。这些数据集采集自常规眼底相机、手持式眼底相机和超广角眼底成像，分别覆盖 60°、45°和 200 °的视野。我们遵循 [12,11] ，将具有5个子类别（正常、轻度糖尿病视网膜病变、中度糖尿病视网膜病变、重度糖尿病视网膜病变和增生性糖尿病视网膜病变）的糖尿病视网膜病变（DR）分级作为已知类别。我们对每个类别分别采样1000 / 500/500个实例，以构建训练/验证/测试数据集。将存在年龄相关性黄斑变性（AMD）的样本作为4个未知类别引入，这些样本与DR具有相似的特征（小玻璃膜疣、大玻璃膜疣、干性AMD和湿性AMD）。更多详细信息请参阅我们的补充文件。

实现细节所有皮肤图像都被调整为 224×224 像素，所有眼底图像都被调整为 512×512 像素。我们采用ResNet-50 [7] 作为分类模型和变分自编码器训练的骨干网络。我们使用 Π-model [15]作为自监督学习正则化器。我们使用指数预热 [15] ，在200个训练周期中的前80 个周期对模型进行预热，以调整对抗训练 α 和自监督学习β的系数。我们使用学习率为 3×10^{-4} 、批量大小为32的SGD 优化器。我们应用了一些常规增强技术，例如随机裁剪和翻转，

表 1：皮肤数据集上的比较结果（5次试验平均准确率%）。

监督式	数据集	ISIC 2019	Derm7pt	PAD-UFES-20	Dermnet
	ERM	69.2 (± 0.89)			
封闭集半监督学习	PL [16]	69.4 (± 0.65)	70.1 (± 0.10)	70.2 (± 0.83)	66.9 (± 0.34)
	PI [15]	70.2 (± 0.33)	70.7 (± 0.64)	70.3 (± 0.22)	69.6 (± 0.87)
	MT [27]	69.6（± 0.45）	70.1（± 0.20）	68.7（± 0.36）	65.1（± 1.43）
	VAT [21]	70.3（± 0.29）	69.9（± 0.23）	69.2（± 0.48）	69.6（± 0.59）
	MM [1]	59.4（± 2.23）	60.2（± 0.55）	60.3（± 1.65）	41.7（± 3.97）
	FM [25]	59.8（± 1.88）	65.1（± 0.63）	63.0 (± 1.01)	53.2 (± 1.87)
开放集半监督学习	UASD [3]	70.0 (± 0.47)	67.5 (± 1.30)	68.5 (± 1.01)	61.2 (± 2.85)
	DS3L [6]	69.3 (± 0.87)	69.8 (± 0.55)	68.9 (± 0.76)	65.3 (± 1.14)
	MTCF [29]	65.4 (± 1.99)	69.2 (± 0.89)	66.3 (± 0.76)	66.8 (± 1.01)
	T2T [8]	60.2 (± 0.23)	61.7 (± 0.15)	60.3 (± 0.21)	60.9 (± 0.10)
	OM [23]	70.1 (± 0.29)	69.6 (± 0.54)	65.8 (± 0.46)	65.4 (± 0.54)
通用半监督学习	CAFA [9]	68.3 (± 1.08)	63.3 (± 1.02)	65.3 (± 1.70)	63.3 (± 1.88)
	Ours	**71.1**(± 1.31)	**70.9**(± 1.01)	**70.8**(± 0.98)	69.6 (± 1.28)

表2：眼底数据集上的比较结果（5次试验平均准确率%）。

监督式	数据集	常规	手持式	UWF
	ERM	70.96（± 0.98）		
封闭集半监督学习	PI [15]	75.01 (± 0.86)	74.51 (± 1.20)	73.85(± 1.65)
开放集半监督学习	UASD [3]	73.25(± 2.51)	73.05(± 1.35)	73.96 (± 1.77)
通用半监督学习	CAFA [9]	72.00(± 2.34)	73.52(± 1.76)	65.21(± 4.21)
	Ours	**77.55**(± 2.85)	**78.02**(± 2.01)	**74.10**(± 2.36)

采用颜色抖动和高斯模糊作为强增强，以评估 UKC。为确保比较公平，我们在各对比方法中保持所有基本超参数一致，例如增强方式、批量大小和学习率。

3.2 对比研究

皮肤数据集上的性能 如表 1所示，我们将提出的方法与现有的自监督学习方法进行了比较，并根据不同的现实场景对这些方法进行了分组。可以看出，在具有不同领域偏移的所有数据集上，我们提出的方法均取得了有竞争力的结果。一个有趣的发现是，现有的开放集半监督学习方法（例如 MTCF） [29], 在我们的设置下表现不佳。这是因为这些方法严重依赖于对未知类别样本的估计和移除。然而，所采用的开放集识别技术是为传统图像分类设计的，并不适用于医学图像。**眼底数据集上的性能** 表2 报告了眼底数据集上的对比结果。我们从每组中选取一种在不同自监督学习场景下表现最佳的技术进行比较。与其他方法相比，我们提出的方法取得了最佳性能，并且在所有设置下均显著优于基线模型。与皮肤数据集上的结果不同，

表3：领域分离 AUC%。

数据集	ISIC	D7pt	PAD	DN
置信度	59.16	61.94	63.01	68.41
扰动 [9]	61.12	62.55	61.32	60.17
VAE	66.02	66.32	63.28	79.09
OVA-Net [23]	58.17	59.82	56.71	79.46
Ours	**67.99**	**66.99**	**65.32**	**83.21**

表 4：消融实验结果。

数据集	ISIC	D7pt	PAD	DN
不使用自监督学习	69.2			
不含 DOE	68.7	68.2	67.9	66.0
不使用 CDS	55.6	59.3	61.1	51.9
不使用数据增强	70.5	69.3	68.5	65.6
Ours	71.1	70.9	70.8	69.6

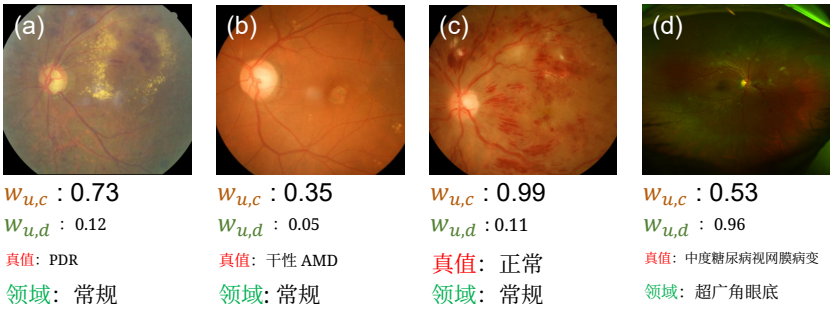

图3：带有归一化分数的无标签数据可视化示例。

CAFA取得了令人满意的结果，但超广角眼底（UWF）除外。UASD的性能优于经验风险最小化（ERM）基线模型。这可能是因为DR和AMD具有相似的特征或语义信息，例如出血和渗出物，这些信息能够有效增强特征学习 [12]。

新颖领域检测 由于我们声称所提出的评分机制能够准确识别未知域样本，因此我们采用不同技术对未知领域分离进行了实验。我们提出的 CDS 评分机制在未知领域分离方面取得了最佳结果。此外，我们发现未知域样本对原型距离也很敏感，例如，在 Dermnet（DN）上取得了 80.79% 的高 AUC，这证实了解耦领域信息对于检测 UKC 的重要性和必要性。

3.3 消融实验

各组成部分的分析我们对各个组成部分进行了消融研究，结果如表3所示。“w/o DOE”表示我们使用经 w_d 重新加权的无标签数据来训练 Π模型，而不考虑类别不匹配。“w/o CDS”表示我们直接将全部无标签数据作为未知域进行域适应，而不考虑域混合的条件，例如ISIC 和 Dermnet。“w/o DA”表示我们利用 w_c 和 w_d 对无标签样本进行重新加权，但不使用对抗训练项。结果发现，领域差异较大的数据集（如 Dermnet）性能下降更明显（从 69.6% 降至 65.6%）。

开放集样本的语义相关性为了探究存在类别/领域不匹配的样本之间的语义相关性，我们将无标签数据中的样本及其对应的归一化 UKC/UKD 分数可视化，如图3所示。值得注意的是，尽管图3-(c)中没有与 AMD 相关的病变，但其 $w_{u,\ c}$ 最高，表明这些样本在特征空间中具有更高的语义相似性，例如出血。将此类样本纳入统一训练，可以通过丰富有用特征有效提升模型性能。

4 结论

在本研究中，我们提出了一种用于医学图像分类的新型通用半监督学习框架。我们提出了两种新型评分机制，用于区分来自未知类别和未知域的样本。随后，我们对重新加权的未标注样本采用常规半监督学习和域适应，以进行统一训练。实验结果表明，在不同的现实场景下，我们提出的方法均能表现良好。

致谢。 本研究部分得到了 Airdoc 医疗人工智能项目捐赠第二阶段、莫纳什-爱迪奥克研究中心的支持，并部分得到了 MRFF NCRI GA89126 的资助。

利益披露。 作者声明不存在任何利益冲突。

参考文献

1. Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A., Raffel, C.A.: Mixmatch：一种整体性的半监督学习方法。神经信息处理系统进展**32** （2019）2. Cao, Z., Ma, L., Long, M., Wang, J.：部分对抗域适应。载于：欧洲计算机视觉会议。第135–150页（2018）3. Chen, Y., Zhu, X., Li, W., Gong, S.：类别分布不匹配条件下的半监督学习。载于：AAAI人工智能会议论文集。第34(4)卷，第3569–3576页（2020）4. Dermnet：Dermnet（2023），https://dermnet.com/ 5. Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S.：利用深度神经网络实现皮肤癌的皮肤科医生级分类。Nature**542**（7639），115–118（2017）6. Guo, L.Z., Zhang, Z.Y., Jiang, Y., Li, Y.F., Zhou, Z.H.：针对未见类别无标签数据的安全深度半监督学习。载于：国际机器学习会议。第3897–3906页。PMLR（2020）7. He, K., Zhang, X., Ren, S., Sun, J.：用于图像识别的深度残差学习。载于：IEEE/CVF计算机视觉与模式识别会议。第770–778页（2016）

8. Huang, J., Fang, C., Chen, W., Chai, Z., Wei, X., Wei, P., Lin, L., Li, G.：变废为宝：利用跨模态匹配挖掘分布外数据，用于开放集半监督学习。载于：IEEE/CVF国际计算机视觉会议。第8310–8319页（2021）9. Huang, Z., Xue, C., Han, B., Yang, J., Gong, C.：通用半监督学习。神经信息处理系统进展**34**，26714–26725（2021）10. ISIC：ISIC档案库（2023），https://www.isic-archive.com/ 11. Ju, L., Wang, X., Zhao, X., Bonnington, P., Drummond, T., Ge, Z.：通过对抗学习和伪标签，利用常规眼底图像训练超广角眼底诊断模型。IEEE医学影像学汇刊**40**（10），2911–2925（2021）12. Ju, L., Wang, X., Zhao, X., Lu, H., Mahapatra, D., Bonnington, P., Ge, Z.：通过知识蒸馏和多任务学习进行视网膜疾病分级的协同对抗标签学习。IEEE生物医学与健康信息学杂志**25**（10），3709–3720（2021）13. Ju, L., Wu, Y., Wang, L., Yu, Z., Zhao, X., Wang, X., Bonnington, P., Ge, Z.：长尾皮肤病变分类的灵活采样。载于：医学图像计算与计算机辅助干预——MICCAI 2022。第462–471页。施普林格（2022）14. Kawahara, J., Daneshvar, S., Argenziano, G., Hamarneh, G.：使用多任务多模态神经网络进行七点检查表与皮肤病变分类。IEEE生物医学与健康信息学杂志**23**（2），538–546（2018）15. Laine, S., Aila, T.：用于半监督学习的时间集成。arXiv预印本 arXiv:1610.02242（2016）16. Lee, D.H., 等：伪标签：一种简单高效的深度神经网络半监督学习方法。载于：ICML表征学习挑战研讨会。第3(2)卷，第896页（2013）17. Lee, D., Kim, S., Kim, I., Cheon, Y., Cho, M., Han, W.S.：用于半监督学习的对比正则化。载于：IEEE/CVF计算机视觉与模式识别会议。第3911–3920页（2022）18. Li, J., Socher, R., Hoi, S.C.：DivideMix：将噪声标签学习作为半监督学习。载于：国际学习表征会议（2020）19. Liu, Q., Yu, L., Luo, L., Dou, Q., Heng, P.A.：基于关系驱动自集成模型的半监督医学图像分类。IEEE医学影像学汇刊**39**（11），3429–3440（2020）20. Ming, Y., Sun, Y., Dia, O., Li, Y.：如何利用超球面嵌入进行分布外检测？arXiv预印本 arXiv:2203.04450（2022）21. Miyato, T., Maeda, S.i., Koyama, M., Ishii, S.：虚拟对抗训练：一种用于监督学习和半监督学习的正则化方法。IEEE模式分析与机器智能汇刊 **41**（8），1979–1993（2018）22. Pacheco, A.G., Lima, G.R., Salomao, A.S., Krohling, B., Biral, I.P., de Angelo, G.G., Alves Jr, F.C., Esgario, J.G., Simora, A.C., Castro, P.B., 等：PAD-UFES-20：一个由患者数据和通过智能手机采集的临床图像组成的皮肤病变数据集。Data in Brief**32**，106221（2020）23. Saito, K., Kim, D., Sae nko, K.：OpenMatch：采用开放集一致性正则化的开放集半监督学习。神经信息处理系统进展**34**，25956–25967（2021）24. Sindhwani, V., Niyogi, P., Belkin, M.：一种基于协同正则化的多视图半监督学习方法。载于：ICML多视图学习研讨会论文集。第2005卷，第74–79页（2005）

25. Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C.A., Cubuk, E.D., Kurakin, A., Li, C.L.：FixMatch：通过一致性与置信度简化半监督学习。神经信息处理系统进展**33**，596–608（2020） 26. Sun, X., Yang, Z., Zhang, C., Ling, K.V., Peng, G.：用于开放集识别的条件高斯分布学习。在：IEEE/CVF计算机视觉与模式识别会议论文集。第13480–13489页（2020） 27. Tarvainen, A., Valpola, H.：平均教师是更好的榜样：权重平均的一致性目标改善半监督深度学习结果。神经信息处理系统进展 **30** （2017） 28. Ye, H., Xie, C., Cai, T., Li, R., Li, Z., Wang, L.：迈向分布外泛化的理论框架。神经信息处理系统进展 **34**，23519–23531（2021） 29. Yu, Q., Ikami, D., Irie, G., Aizawa, K.：用于开放集半监督学习的多任务课程学习框架。在：欧洲计算机视觉会议。第438–454页。施普林格（2020）