

DOI:10.3969/j.issn.1000-1565.2024.04.013

基于自反馈阈值学习的半监督皮肤癌诊断模型

韩硕¹, 袁伟理¹, 杜泽宇²

(1. 河北医科大学 基础医学院, 河北 石家庄 050017; 2. 曼彻斯特大学 健康科学学院, 英格兰 曼彻斯特 M139PL)

摘要:为解决监督学习皮肤癌诊断模型的训练需要大量数据标注,且医学专家标注工作成本高、耗时长、易疲劳等问题,提出了一种基于自反馈阈值学习(Self-Feedback Threshold Learning, SFTL)的半监督皮肤癌诊断方法。在标注数据预训练的 ResNet 网络基础上,引入全局和局部类别间伪标签自反馈阈值学习机制动态筛选 ResNet 预测概率大于自反馈阈值的无标记样本,引入无监督阈值学习损失和分类交叉熵损失进行模型训练,在标记样本稀缺的情况下深入挖掘无标记数据的鉴别诊断信息,显著降低模型在无标记皮肤病变图像中的误判率。选取公开数据集 HAM10000 的皮肤病变图像展开实验验证,在仅需 50% 标记数据下实现了 0.8229 的准确率和 0.7651 的 F1 分数,证明所提出的 SFTL 模型在半监督场景下可有效解决皮肤癌诊断任务,相比其他同类方法具有更好的分类性能。

关键词:半监督皮肤癌诊断;自反馈阈值学习;卷积神经网络;半监督学习

中图分类号:U492.2;TP301.6

文献标志码:A

文章编号:1000-1565(2024)04-0441-08

Semi-supervised skin cancer diagnosis based on self-feedback threshold learning

HAN Shuo¹, YUAN Weicheng¹, DU Zeyu²

(1. College of Basic Medicine, Hebei Medical University, Shijiazhuang 050017, China;

2. School of Health Science, University of Manchester, Manchester M139PL, UK)

Abstract: To address the challenges associated with the need for a large amount of annotated data in supervised skin cancer diagnosis models, such as the high cost, time consumption, and fatigue experienced by medical experts during annotation, this study proposes a semi-supervised skin cancer diagnosis method based on Self-Feedback Threshold Learning (SFTL). Building upon the ResNet network pre-trained with labeled data, a global and local class pseudo-label self-feedback threshold learning mechanism is introduced to dynamically select unlabeled samples with ResNet prediction probabilities exceeding the self-feedback threshold. Unsupervised threshold learning loss and classification cross-entropy loss are incorporated for model training, thereby deeply mining the diagnostic information from unlabeled data when labeled samples are scarce and significantly reducing the misdiagnosis rate in unlabeled skin lesion images. Experimental validation was conducted using the publicly available HAM10000 skin lesion dataset, achieving an accuracy of 0.8229 and an F1 score of 0.7651 with only 50% of the data labeled. The results demonstrate that the proposed SFTL model effectively addresses the skin cancer diagnosis task in semi-supervised scenarios and outperforms other compared methods in terms of classification performance.

收稿日期:2024-01-08; **修回日期:**2024-05-23

基金项目:河北省自然科学基金资助项目(H2019206316)

第一作者:韩硕(1974—),男,河北医科大学讲师,博士,主要从事肿瘤及神经退行性疾病相关的基础研究及人工智能医疗应用。E-mail: hanshuo@hebm. edu. cn

Key words: semi-supervised skin cancer diagnosis; self-feedback threshold learning; convolutional neural network; semi-supervised learning

皮肤癌是一种全球普遍存在的癌症类型,其诊断对减少公共健康负担至关重要^[1-3].据世界卫生组织(WHO)的数据显示,皮肤癌的早期识别和治疗能够显著提高患者的生存率.然而,皮肤癌的发病率在全球范围内正以惊人的速度增长,尤其是在紫外线照射强度高的地区,皮肤癌的发病率和死亡率之高令人震惊.值得注意的是,皮肤癌的治疗费用也在持续上升,使得部分患者难以得到及时有效的治疗^[4-5].因此,准确及时地识别和诊断皮肤癌具有重大的临床意义.

在实际诊断中,专家面临着多种挑战.首先,皮肤癌的病理形态多样,需要医生具有丰富的经验和精准的判断能力.由于皮肤癌的早期症状通常较为隐蔽,患者可能会忽视其重要性,从而错过治疗的最佳时机^[6-7].因此,提高皮肤癌的早期识别准确率成为了重要的研究方向.人工智能(AI)技术在皮肤癌诊断领域的应用显得尤为重要^[8].研究已经证明,AI在许多疾病的诊断中都展现出了相当高的精确度,其中就包括皮肤癌^[9-10].尽管AI技术在诊断皮肤癌方面有着巨大的潜力,但目前的应用仍然面临一些挑战.一方面,AI模型的训练通常需要大量的标注数据,而这些数据的获取过程往往需要经验丰富的医生进行手动标注,既耗时又费力,成为实际应用中的一大瓶颈.另一方面,部分算法虽然通过引入伪标签在一定程度上利用未标注数据与标注数据共同训练模型,但这些算法大多采用固定阈值来确定样本的伪标签.这种策略容易引入大量野点即错误标注数据,从而对模型的分类效果产生负面影响,限制了其在复杂医学任务中的适用性.因此,在实际临床应用中,需要一种能够充分利用已有资源,既能处理标注数据,又能充分有效利用未标注数据的人工智能模型.

为解决皮肤癌诊断中遇到的难题,本研究提出了一种基于自反馈阈值学习(self-feedback threshold learning, SFTL)的半监督皮肤癌诊断模型.此模型聚焦于半监督问题,利用标记数据和无标记数据共同训练网络,从而缓解当前有监督AI模型需要大量数据标注的问题.同时,本模型通过设置自反馈阈值学习模块,动态调整筛选无标记样本时的阈值,以筛选出更可靠的无标记样本帮助模型进一步训练.具体来说,模型在每轮迭代训练过程中,会根据未标注数据的平均置信度以及在各类别中的预测期望值,动态地调整筛选阈值,从而根据已有数据的分布和特性,逐步自我调整,增加伪标签的质量.此机制显著降低了模型在无标记数据上的误判率,提升了模型的鉴别能力和诊断准确性,从而有效解决了标注数据稀缺所带来的挑战.通过更有效地利用现有的医学图像资源,本模型不仅提高了诊断准确性,还能显著提高诊断过程的效率.

1 自反馈半监督学习模型

1.1 模型概述

以解决标记数据稀缺问题为出发点,提出的基于SFTL的半监督皮肤癌识别模型,旨在最大化利用未标记医学图像数据.首先,构建基于深度学习的ResNet^[11]网络并训练为初始识别模型;然后,对未标记的皮肤镜图像进行预测并生成伪标签.由于伪标签具有一定的噪声和不确定性,因此,设计了SFTL模块,它能有效筛选掉噪声样本,以提升模型的训练效果.

本模型的算法框架如图1所示,从图1可看出,模型的训练过程主要分为2个阶段.第1阶段,模型利用标记数据对分类器进行若干轮次的训练,使其获得初步的分类能力.第2阶段,模型引入无标签数据,通过自反馈阈值计算模块筛选出置信样本,并生成相应的伪标签,然后将这些带有伪标签的置信样本与原有的标记数据共同输入分类器进行联合训练,从而进一步提升分类器的性能.在阶段过渡时,仅需将置信样本与原有的标记样本一同输入分类器,并在损失函数中加入无监督学习损失,而无需进行其他额外操作.2个阶段的训练均完成后,模型训练即告终止.通过这种训练方式,模型能够充分利用未标记数据中的潜在信息,有效提升皮肤癌诊断的准确性和可靠性.

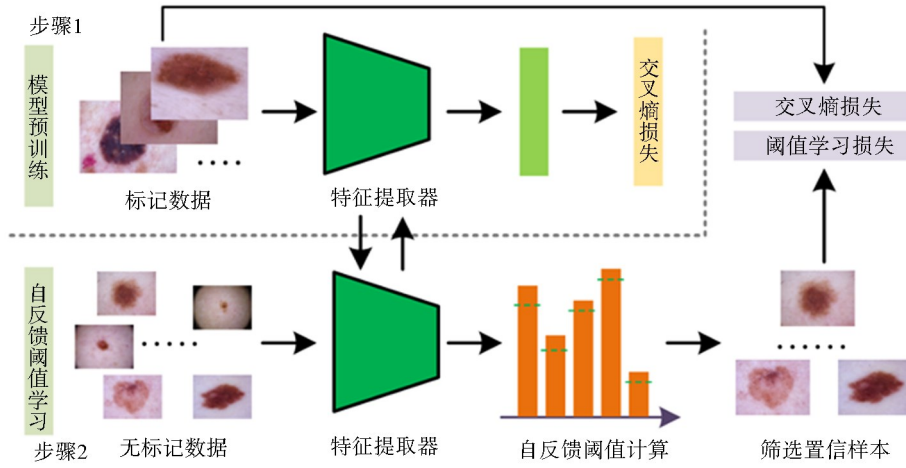


图 1 基于 SFTL 的半监督算法框架

Fig. 1 Semi-supervised algorithm framework based on SFTL model

1.2 问题定义

在半监督皮肤癌诊断任务中,训练数据包括已标记和无标记 2 类皮肤图像数据. 本文将标记样本定义为 $X_l = \{x_1^l, x_2^l, \dots, x_{N_l}^l\}$, 其中 N_l 代表标记样本的数量, 对应样本的病变类别标签集合定义为 $Y_l = \{y_1, y_2, \dots, y_{N_l}\}$, 其中每个标签 $y_i^l \in Y_l$ 均取自 $[1, N_c]$ 之间的整数, N_c 代表类别个数. 无标记皮肤病变图像集合定义为 $X_u = \{x_1^u, x_2^u, \dots, x_{N_u}^u\}$, N_u 代表无标记样本的数量. 对于特征提取, 本文采用 ResNet50 作为主干网络 $f(\cdot)$ 提取图像特征, 定义分类层 $p(\cdot)$ 预测特征类别.

1.3 基于标记样本的模型预训练

为了确保主干网络对皮肤病变图像的特征提取能力, 本模型首先基于标记样本 X_l 及其类别标签 Y_l , 引入交叉熵损失 (L_{ce}) 训练网络的特征提取能力. 损失定义如下:

$$L_{ce} = -E_{x_i^l \in X_l} y_i \log p(f(x_i^l)). \quad (1)$$

经过上述监督损失函数的初步训练, 特征提取器具备了一定的皮肤病变特征提取能力, 可用于预测无标记样本的概率分布, 基于此实现对无标记样本的伪标签生成过程. 给定未标记的皮肤病变图像样本 $x_j^u \in X_u$, 可以得到每个可能类别的预测分数 $q_j^u = p(f(x_j^u))$, 该分数代表了模型对样本属于每个类别的置信度, 而预测值最大概率对应的类别即为预测类别.

1.4 基于 SFTL 的无标记样本选择模型

在为无标记皮肤病变图像生成伪标签的过程中, 不可避免地会产生一些带有错误标签的噪声样本. 因此, 在本模型中, 移除这些带有噪声的伪标记样本变得至关重要. 本文主要通过全局自反馈阈值学习和局部自反馈阈值学习 2 个部分完成样本选择.

为了克服传统固定筛选阈值所带来的局限性, 提出了一种新颖的无标记样本模型选择策略, 这种策略基于自适应反馈阈值学习, 并涵盖了全局自适应反馈与类别特定自适应反馈 2 种机制. 本研究受到相关文献^[12]的影响, 认为在半监督学习 (SSL) 中确定阈值的关键在于阈值应能够反映出学习状态. 通过评估一个校准精准的模型预测置信度, 可以有效地判断学习效果. SFTL 方法利用训练过程中模型的预测结果, 自动地为每个类别定义并动态调整置信度阈值. SFTL 方法起初会估算出一个全局阈值, 作为模型置信度的指数移动平均值 (EMA)^[13], 接着利用模型对各个类别概率的 EMA 进行评估, 调整全局阈值.

SFTL 算法主要分为全局自反馈阈值和局部自反馈阈值 2 个模块. 首先, 全局自反馈阈值模块在模型迭代过程中, 根据未标注数据的平均置信度进行逐步优化. 具体而言, 该模块在每次迭代时, 通过计算所有无标

记样本的平均置信度,确定一个全局阈值,用于整体控制样本筛选过程.其次,局部自反馈阈值模块则更加细致地考虑了无标记数据对每个类别的诊断概率,全面评估各个类别的阈值.局部阈值依据每个类别的无标记样本的分类概率进行计算,确保对不同类别的样本进行差异化处理,从而更加全面地评估和筛选置信样本.最后,SFTL 算法对全局自反馈阈值和局部自反馈阈值进行综合计算,得到最终的自反馈阈值.这种动态、自适应的阈值调整机制,使得模型在利用无标记数据进行训练时,能够显著提升分类性能和预测精度.

全局自反馈阈值学习:在 SFTL 方法中,全局阈值的设定需要与模型对未标注数据的置信度紧密相关,以便准确反映出整体的学习情况.此外,为了有效排除那些错误的伪标签,这一全局阈值在训练过程中应逐渐提升.在本研究中,将全局阈值定义为模型在第 t 次迭代时对未标注数据的平均置信度,全局阈值的优化依据文献[14]方式进行迭代更新

$$T_g(t) = \begin{cases} \frac{1}{N_c}, & \text{if } t = 0 \\ \lambda T_g(t-1) + (1-\lambda) \frac{1}{N_u} \sum_{j=1}^{N_u} \max(q_j^u), & \text{otherwise} \end{cases}, \quad (2)$$

其中, λ 是 EMA 的参数, $q_j^u = [q_j^u(1), q_j^u(2), \dots, q_j^u(N_c)]$ 代表无标记数据中第 j 个样本在各类别中的预测分数.

局部自反馈阈值学习:在 SFTL 方法中,局部自反馈阈值的目的在于以类别特定的方式调整全局自反馈阈值,以此来考虑类内多样性和潜在的类别邻近性.通过计算模型对每个类别的预测期望值,以评估该类别的具体学习状态.

$$T_r(t, c) = \begin{cases} \frac{1}{N_c}, & \text{if } t = 0 \\ \lambda T_r(t-1, c) + (1-\lambda) \frac{1}{N_u} \sum_{j=1}^{N_u} q_j^u(c), & \text{otherwise} \end{cases}, \quad (3)$$

由此,得到每个类别的自反馈阈值 $T_r(t) = [T_r(t, 1), \dots, T_r(t, c), \dots, T_r(t, N_c)]$. 式(3)中, $q_j^u(c)$ 代表无标记数据中第 j 个样本对第 c 个类别的预测分数.

通过全局自反馈和局部自反馈阈值学习,可以得到最终的自反馈阈值 T

$$T(t, c) = \text{MaxNorm}(T_r(t, c)) \cdot T_g(t) = \frac{T_r(t, c)}{\max\{T_r(t, c) : c \in [N_c]\}} \cdot T_g(t), \quad (4)$$

其中,MaxNorm 的计算形式为 $x' = \frac{x}{\max(x)}$. 根据上述计算的自反馈阈值计算公式,可以获得筛选后更为

可靠的无标记样本集合,即 \hat{X}_u .

最终,第 t 次迭代过程中用于无监督学习的阈值学习损失定义如下:

$$L_T = -E_{x_j^u \in X_u} \mathbf{1}(\max(q_j^u) > T(t, \arg \max(q_j^u))) \cdot (y_j^u \log p(f(x_j^u))), \quad (5)$$

其中, $\mathbf{1}$ 代表样本超过固定概率阈值的指示函数, y_j^u 代表无标记数据中第 j 个样本的伪标签,由此, L_T 中只保留预测概率大于自反馈阈值的无标记样本用于模型训练.

通过引入公式(5)中的阈值学习损失,本文模型的最终损失函数定义如下:

$$L = \lambda_1 L_{ce} + \lambda_2 L_T, \quad (6)$$

其中, λ_1 和 λ_2 是损失函数中平衡系数.

2 实验与结果分析

2.1 皮肤病变数据来源

实验验证基于公开的 HAM10000 皮肤癌诊断数据集^[15],这是目前最大的公开皮肤病变图像数据集,包含了奥地利维也纳医学大学的皮肤科和澳大利亚昆士兰的 Cliff Rosendahl 皮肤癌诊所,跨越 20 年时间收

集的 10 015 张涵盖 7 种皮肤病变的图像. 针对本研究的半监督自反馈阈值学习模型, 将数据集随机划分为 5 000 张训练集和 5 015 张的测试集, 其中训练集再细分为标记和未标记的子集, 进行不同标记样本量下的学习实验.

2.2 实验设计

实验选择了深度残差网络 ResNet50 作为特征提取器, 利用在 ImageNet 大规模视觉识别挑战赛上预先训练得到的模型参数对其进行初始化, 为模型的训练提供了良好的基础^[16]. 模型的训练采用了随机梯度下降优化器, 共进行 150 个迭代周期(Epoch)的训练. 在训练过程中, 模型的批处理大小设定为 64, 初始学习率定为 0.001, 平衡系数 λ_1 和 λ_2 分别设置为 1 和 0.3. 将所有输入图像统一调整为 224×224 的分辨率, 并通过随机水平翻转等数据增强技术, 提升了模型的鲁棒性. 本研究基于 Pytorch 框架进行开发, 并在 NVIDIA Geforce 4090 GPU 上进行模型的训练和验证. 为了全面评估模型的性能, 本文采用了包括准确率(Accuracy)、精确度(Precision)、灵敏度(Sensitivity)、特异性(Specificity)、Kappa 系数以及 AUC(area under the curve)值在内的多种评价指标.

2.3 实验结果

2.3.1 结果分析

通过在 HAM10000 数据集上进行一系列深度实验, 充分验证了基于半监督自反馈阈值学习的皮肤癌诊断模型的效率和健壮性. 如表 1 所示, 在不同数量的标记样本上进行了详尽的性能评估和对比研究. 值得注意的是, 当仅使用 500 个标记样本时, 模型已经展示出显著的预测能力, 准确率高达 0.752 8, Kappa 系数为 0.624 1, $F1$ 分数达到 0.731 9, 召回率和精确度分别为 0.752 8 和 0.731 1, 这表明模型在标记数据稀缺的情况下仍具有强大的泛化和学习能力. 随着标记样本数量增加至 1 000~5 000, 模型的性能显著提升. 在 5 000 个标记样本的情况下, 各项性能指标达到了最高值: 准确率为 0.855 4, Kappa 系数为 0.813 6, $F1$ 分数为 0.851 8, 召回率和精确度分别为 0.851 8 和 0.851 0. 这些高效的性能指标, 特别是准确率和 Kappa 系数的显著提升, 有力证明了本文提出的模型在皮肤癌诊断中的准确性和一致性. 此外, $F1$ 分数、召回率和精确度的高水平表现, 进一步突显了该模型在平衡误诊与漏诊方面的出色性能. 这些实验结果表明 SFTL 模型在皮肤癌诊断任务中具有鲁棒性, 并为未来半监督皮肤癌研究提供了新的方向.

表 1 不同标记样本规模下 SFTL 实验结果

Tab. 1 Experimental results of SFTL with different labeled data scales

标记样本数量	准确率	Kappa 系数	$F1$	召回率	精确度
500	0.752 8	0.624 1	0.731 9	0.752 8	0.731 1
1 000	0.788 9	0.707 8	0.780 7	0.788 9	0.782 8
1 500	0.792 9	0.726 7	0.787 6	0.792 9	0.787 9
2 000	0.811 7	0.751 4	0.805 0	0.805 0	0.802 7
2 500	0.822 9	0.765 1	0.819 3	0.822 9	0.817 5
5 000	0.855 4	0.813 6	0.851 8	0.851 8	0.851 0

表 2 无 SFTL 模块的实验结果

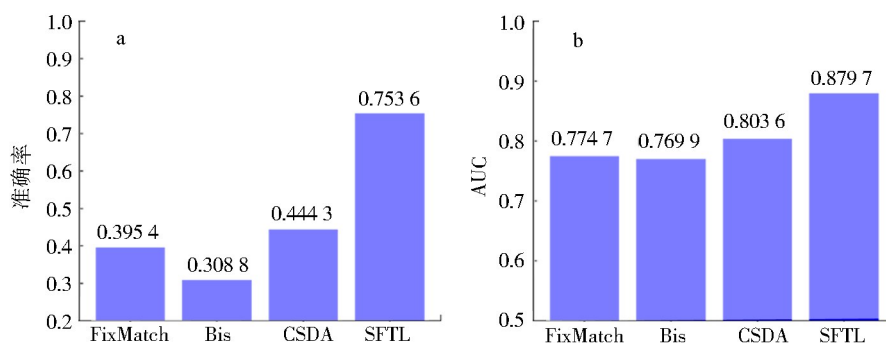
Tab. 2 Experimental results without SFTL module

标记样本数量	准确率	Kappa 系数	$F1$ 分数	召回率	精确度
500	0.747 8	0.641 0	0.641 0	0.747 8	0.747 8
1 000	0.788 7	0.712 4	0.781 3	0.788 7	0.788 7
1 500	0.794 1	0.794 1	0.787 7	0.787 7	0.783 8
2 000	0.813 7	0.783 8	0.809 3	0.813 7	0.806 7
2 500	0.820 3	0.760 5	0.815 1	0.815 1	0.812 0

同时,为了深入揭示无标记样本在模型训练过程中的影响,设计了一系列对比实验,对比了在监督学习和半监督学习框架下模型的表现.如表 2 所示,在去除 SFTL 模块后,模型的整体性能呈现出一定程度的下滑.以 500 个标记样本为例,模型的准确率由 0.752 8 下降至 0.747 8, Kappa 系数由 0.624 1 微降至 0.641 0,同时其他指标也都出现了相应的降低.这一趋势形象地揭示了 SFTL 模块在提升模型性能方面的重要作用.这一发现再次强调了 SFTL 模型的优势,即能够有效地利用未标记样本中的隐含信息,显著提升皮肤癌诊断的准确性和可靠性,凸显了该模块在标签资源有限场景中的巨大优势.

2.3.2 对比分析

根据 CSDA(class-specific distribution dignment)^[17]中汇报的多种先进半监督算法(包括 FixMatch^[18]、Bis^[19]、CSDA)在 HAM10000 数据集上的表现,系统地对比了不同方法的效果,并将结果汇总在图 2 中.为了保证实验条件的一致性,所有算法均在 600 个标记数据的基础上进行训练.结果显示,本文提出的基于 SFTL 模型取得了令人印象深刻的成绩,准确率达到 0.753 6, AUC 得分达到了 0.879 7.这一结果显著超越了其他对比方法的表现,与对比方法中性能最好的 CSDA 模型相比,准确率提高了 0.309 3, AUC 提高了 0.076 1.因此, SFTL 模型在少量标注数据情况下表现出卓越的诊断能力,显著优于其他先进半监督方法,表明 SFTL 模型在皮肤癌诊断领域的半监督学习中实现了实质性的进步,展示了其在临床应用中的巨大潜力.



a. 不同类型准确率对比; b. 不同类型 AUC 效果对比

图 2 SFTL 模型与其他先进半监督算法的对比

Fig. 2 Performance comparison between SFTL model with other state-of-the-art models

2.3.3 ROC 可视化分析

在模型性能评估阶段,使用接受者操作特性曲线(ROC)衡量模型在皮肤癌诊断中的表现.如图 3 所示,使用仅 500 个标记样本训练的模型在识别不同类型皮肤病变任务中展示了优秀的诊断性能,具体包括非色素性角化病(AKIEC)、基底细胞癌(BCC)、良性角化病样病变(BKL)、皮脂腺瘤(DF)、黑色素瘤(MEL)、黑色素痣(NV)和血管性病变(VASC).这些疾病的识别任务中,模型分别获得了 0.93、0.94、0.87、0.83、0.82、0.91 和 0.92 的 AUC 值.整体病种的宏观平均 ROC 曲线 AUC 值达到 0.89,进一步证明了模型在多元化和复杂皮肤癌诊断中的强大鲁棒性和稳定性,验证了自反馈阈值学习模型的卓越性能,特别是在标签稀缺情况下,其独特的半监督学习策略有效挖掘未标记样本信息,显著提升了模型的学习效果.

2.3.4 t -SNE 可视化分析

为了更直观地展示本研究所提出的半监督皮肤癌诊断模型在特征空间中对不同皮肤病变类型的类别划分,在图 4 中,采用 t -分布随机邻域嵌入(t -SNE)算法将模型生成的高维特征空间映射到二维平面上进行观察和分析.从可视化结果中可以清楚地看到,不同类型样本在二维空间中形成了清晰的簇.这表明本文模型在高维特征空间中能够有效区分并识别不同的皮肤病变类型.特别值得注意的是,基底细胞癌(BCC)和黑色素瘤(MEL)在特征空间中的分布较为分散,这可能揭示了这些病变在临床特征表现上的异质性.而黑色素痣(NV)的样本在图中聚集成一个显著区域,展示了模型对黑色素痣与其他类型皮肤病变的明确区分能力,这在降低诊断过程的假阳性率上具有重要价值,进一步突显了本文模型在特征抽取和分类任务中的卓越性能.

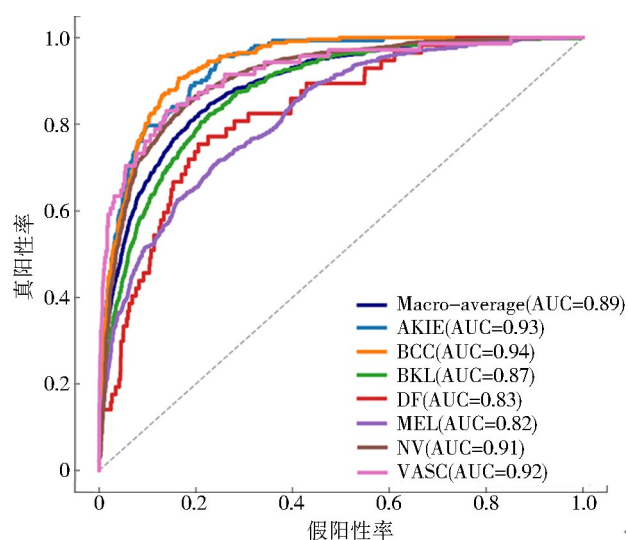


图3 SFTL模型在500例标注样本情况下每个类别的ROC曲线

Fig. 3 ROC curve of each category in SFTL model under 500 annotated samples

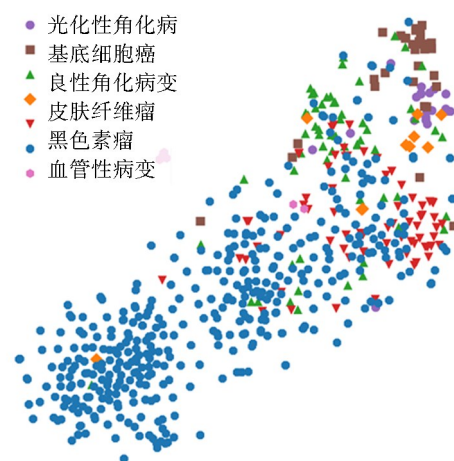


图4 本文模型在2500例标注样本情况下的t-SNE可视化

Fig. 4 t-SNE visualization in this model under 2500 annotated samples

2.3.5 损失函数中平衡系数的消融实验分析

为了验证损失函数使用不同平衡参数比例下模型的效果,设置了一系列消融实验,并将结果汇总在图5中.消融实验过程中,本文分别将其中一个平衡参数固定为最优值的情况下对另一个参数进行一系列实验来验证最优参数.图5中的结果表明,当且仅当 λ_1 和 λ_2 的参数分别设置为1和0.3时,模型的结果达到最优.这一结果表明用于有监督学习的交叉熵损失仍然需要在模型训练过程中占据较大比重,但用于无监督学习的阈值学习损失的比例同样不能过低,否则都会影响模型诊断皮肤癌的能力.该实验结果进一步说明了自反馈阈值学习这一无监督学习过程在模型中占据着不可替代的作用.因此,SFTL模型不仅能够帮助减少医疗人员在数据标注方面的工作压力,还能推进皮肤癌诊治研究的进一步发展,从而显著改善皮肤癌患者的治疗效果.

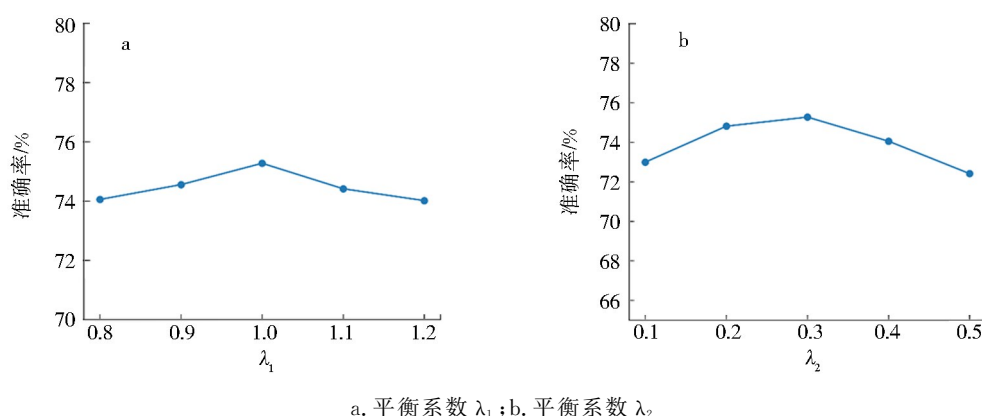
a. 平衡系数 λ_1 ; b. 平衡系数 λ_2

图5 不同平衡参数下SFTL的消融实验结果

Fig. 5 Ablation results of SFTL under different equilibrium parameters

3 结论

提出了一个创新的半监督学习框架——SFTL模型.这一模型聚焦半监督学习任务,使用少量标记数据与大量的未标记数据共同训练,有效缓解了当前皮肤癌诊断任务中经常遇到的标记数据不足的问题. SFTL模型的创新在于其在每次迭代过程中,能够根据数据的全局置信度和类内预测期望动态地调整筛选置信样

本时的阈值,从而增加了对无标记数据的利用效率和筛选置信样本的精确度.这种机制不仅提升了模型的训练效果,同时显著提高了其皮肤癌诊断能力.在 HAM10000 数据集上的实验结果表明,即使在标注样本稀缺的情况下,SFTL 模型依然能够达到较高的准确率,从而验证了其在标签稀缺环境下的强大诊断能力和鲁棒性.综合来看,本研究所提出的 SFTL 模型为解决医学图像分析中标注数据不足的问题提供了新的思路和方法,对提升皮肤癌早期诊断的准确性和效率具有重要参考价值.

参 考 文 献:

- [1] 蔡桂月,李思锐,刘颖,等.皮肤肿瘤患者外周血循环肿瘤细胞与循环肿瘤血管内皮细胞检测的临床价值[J].皮肤性病诊疗学杂志,2023,30(3):198-204. DOI: 10.3969/j.issn.1674-8468.2023.03.003.
- [2] ZENG L, JASWANTH GOWDA B H J, AHMED M G, et al. Advancements in nanoparticle-based treatment approaches for skin cancer therapy[J]. Mol Cancer, 2023, 22(1): 10. DOI: 10.1186/s12943-022-01708-4.
- [3] KING C, FOWLER J C, ABNIZOVA I, et al. Somatic mutations in facial skin from countries of contrasting skin cancer risk[J]. Nat Genet, 2023, 55(9): 1440-1447. DOI: 10.1038/s41588-023-01468-x.
- [4] WU Y H, CHOU C L, CHANG H C. Risk of skin cancer after ultraviolet phototherapy in patients with vitiligo: a systematic review and meta-analysis[J]. Clin Exp Dermatol, 2022, 47(4): 692-699. DOI: 10.1111/ced.15010.
- [5] 唐洪波,梁俊琴.紫外线致非黑色素瘤皮肤癌的作用机制研究进展[J].肿瘤预防与治疗,2019,32(10):940-944. DOI: 10.3969/j.issn.1674-0904.2019.10.014.
- [6] 崔灿,张芳,孙圆圆,等. Sirtuins 在皮肤恶性肿瘤发病机制中的作用及研究进展[J].皮肤病与性病,2023,45(4):238-241. DOI: 10.3969/j.issn.1002-1310.2023.04.003.
- [7] 万春雷.皮肤癌的发病机制及在不同种群的发病差异[J].皮肤科学通报,2022,39(3):173-176.
- [8] TSCHANDL P, RINNER C, APALLA Z, et al. Human-computer collaboration for skin cancer recognition[J]. Nat Med, 2020, 26(8): 1229-1234. DOI: 10.1038/s41591-020-0942-0.
- [9] ESTEVA A, KUPREL B, NOVOA R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. nature, 2017, 542(7639): 115-118. DOI: 10.1038/nature21056.
- [10] GULSHAN V, PENG L, CORAM M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs[J]. jama, 2016, 316(22): 2402-2410. DOI: 10.1001/jama.2016.17216.
- [11] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA, IEEE, 2016: 770-778. DOI: 10.1109/cvpr.2016.90.
- [12] KLINKER F. Exponential moving average versus movingexponentialaverage[J]. Math Semesterber, 2011, 58(1): 97-107. DOI: 10.1007/s00591-010-0080-8.
- [13] WANG Y D, CHEN H, HENG Q, et al. FreeMatch: self-adaptive thresholding for semi-supervised learning[EB/OL]. 2022; arXiv: 2205.07246. <http://arxiv.org/abs/2205.07246>.
- [14] WANG Y, CHEN H, HENG Q, et al. Freematch: Self-adaptive thresholding for semi-supervised learning[J]. arxiv preprint arxiv:2205.07246, 2022. DOI: <https://doi.org/10.48550/arXiv.2205.07246>.
- [15] TSCHANDL P, ROSENDAHL C, KITTLER H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions[J]. Sci Data, 2018, 5: 180161. DOI: 10.1038/sdata.2018.161.
- [16] AMARI S I. Backpropagation and stochastic gradient descent method[J]. Neurocomputing, 1993, 5(4/5): 185-196. DOI: 10.1016/0925-2312(93)90006-o.
- [17] HUANG Z, WU J, WANG T, et al. Class-specific distribution alignment for semi-supervised medical image classification[J]. Computers in Biology and Medicine, 2023, 164: 107280. DOI: 10.1016/j.compbimed.2023.107280.
- [18] SOHN K, BERTHELOT D, CARLINI N, et al. Fixmatch: Simplifying semi-supervised learning with consistency and confidence[J]. Advances in neural information processing systems, 2020, 33: 596-608. DOI: <https://doi.org/10.48550/arXiv.2001.07685>.
- [19] HE J, KORTYLEWSKI A, YANG S, et al. Rethinking re-sampling in imbalanced semi-supervised learning[J]. arxiv preprint arxiv:2106.00209, 2021. DOI: <https://doi.org/10.48550/arXiv.2106.00209>.

(责任编辑:孟素兰)