

基于标签噪声对比学习的肺癌淋巴结转移鉴别^①



祁婧, 李子荣, 刘秀婷, 马露, 陈俊豪

(兰州财经大学 信息工程与人工智能学院, 兰州 730020)

通信作者: 李子荣, E-mail: ivanbear@163.com

摘要: 基于深度学习的人工智能诊断模型严重依赖于高质量地详尽注释数据进行算法训练, 但受到标签噪声信息的影响. 为了增强模型的鲁棒性并防止有噪声的标签记忆, 提出了一种标签噪声样本选择 (noise label sample selection, NLSS) 模型来充分挖掘噪声样本的隐藏信息, 减轻模型过拟合问题. 首先, 通过将混合增强图像作为输入, 提取图像分布式特征表示; 其次, 引入对比损失函数以及比较样本预测标签分布与其真实标签分布的相似性来评估样本, 进行样本选择; 最后, 通过标签重分配模块的伪标签提升策略在样本选择的基础上重新纠正噪声标签的监督信息. 以非小细胞肺癌 (non-small cell lung cancer, NSCLC) 患者的 PET/CT 数据集为例进行实验, 结果表明提出的模型均对比模型有一定的提升, 可降低淋巴结转移状态诊断中标签噪声的干扰.

关键词: 深度学习; 标签噪声; 样本选择; 对比学习

引用格式: 祁婧, 李子荣, 刘秀婷, 马露, 陈俊豪. 基于标签噪声对比学习的肺癌淋巴结转移鉴别. 计算机系统应用, 2024, 33(12): 177-184. <http://www.c-s-a.org.cn/1003-3254/9704.html>

Identification of Lung Cancer Lymph Node Metastasis Based on Label Noise Contrastive Learning

QI Jing, LI Zi-Rong, LIU Xiu-Ting, MA Lu, CHEN Jun-Hao

(School of Information Engineering and Artificial Intelligence, Lanzhou University of Finance and Economics, Lanzhou 730020, China)

Abstract: The AI diagnostic model based on deep learning relies heavily on high-quality detailed annotated data for algorithm training, but is affected by label noise information. To enhance the robustness of the model and prevent noisy label memory, a noise label sample selection (NLSS) model is proposed to fully mine the hidden information of noise samples and alleviate model overfitting. Firstly, distributed feature representations of the image are extracted by taking hybrid enhanced images as input. Secondly, the contrastive loss function is introduced to compare the similarity between the predicted label distribution of the sample and the real label distribution for sample evaluation and selection. Finally, based on sample selection, supervised information of the noisy label is re-corrected by the pseudo-label promotion strategy of the label redistribution module. Taking the PET/CT dataset of non-small cell lung cancer (NSCLC) patients as an example, results show that the proposed models outperform comparison models, reducing the interference of label noise in the diagnosis of lymph node metastasis.

Key words: deep learning; label noise; sample selection; contrastive learning

肺癌是全球发病人数第 2 位, 死亡人数第 1 位的癌症. 非小细胞肺癌 (non-small cell lung cancer, NSCLC) 是肺癌的主要类型之一, 约占肺癌病例的 85% 以上^[1].

对于非小细胞肺癌患者, 淋巴结转移是评判肿瘤的进展和恶性程度的指标之一, 并且淋巴结转移作为 TNM (tumor node metastasis) 分期系统中的重要内容, 直接

^① 基金项目: 甘肃省高等学校青年博士基金 (2022QB-118)

收稿时间: 2024-05-22; 修改时间: 2024-06-17; 采用时间: 2024-07-04; csa 在线出版时间: 2024-10-25

CNKI 网络首发时间: 2024-10-25

影响肺癌患者的生存预期^[2],所以准确识别淋巴结转移情况对于患者的治疗和预后评估具有重要意义。

目前临床上对于淋巴结转移常用的诊断方法为影像学诊断和病理学检测,其中病理学检测是淋巴结转移诊断的金标准,但由于淋巴结活检为有创检查,对于有些身体状况较差且不耐受的病人,无法进行穿刺活检,同时获取某些患者的病理检查较为困难的情况也时有发生。而无创影像学检查如^[18F]-FDG PET/CT,在提供解剖结构信息的基础上可显示肿瘤细胞的代谢活性,在无创性淋巴结转移的诊断工作中有一定的优势。但由于肿瘤的异质性、肿瘤大小深度的限制以及图像成像的质量等因素影响,经验丰富的医生在同一时期对同一影像的判断也存在分歧,标签往往存在噪声,为深度模型的训练带来了挑战。

据此,本文提出一个对抗噪声标签的深度模型,根据PET/CT图像特征以及放射科医生的诊断,通过检测纠正医生标签的方法进行数据标签纠正,并进行模型训练。本文的主要贡献概括如下。

(1) 本文提出了一种对抗噪声标签的图像分类的方法,可对带有噪声标签的数据进行准确分类。

(2) 使用对比学习和样本选择相结合的策略筛选出干净样本和噪声样本,并采用伪标签提升策略,将分类网络预测结果作为噪声样本的监督信息,进一步减少标签噪声对训练的影响。

(3) 实验证明,本文所提出的框架能够有效地应对医学图像噪声标签,并且优于其他的模型方法和医生的诊断结果。

1 相关研究

1.1 噪声标签学习

由于医学图像成像质量、医生水平以及标注过程中的主观性,放射科医生存在一定程度的误诊和漏诊现象,不可避免地会产生医学图像上的噪声标签。尽管有许多应对噪声标签的方法,如估计噪声转移矩阵^[3]、样本选择^[4]和标签纠正^[5]等,此外,一些研究工作通过联合利用平均绝对误差和交叉熵损失^[6]、对噪声不敏感的互信息的广义版本^[7]或相互增强的鲁棒损失函数组合^[8]的优点来研究鲁棒损失函数。但以上成果主要针对的是自然图像噪声数据集,针对医学图像噪声标签的应用仍面临挑战。为了解决这个问题,Xue等人^[9]提出了一种使用样本选择模块和样本重加权策略的鲁棒

学习框架。Min等人^[10]提出了一种带噪声标签的半监督生物学分割网络。Xing等人^[11]提出了一种基于梯度一致性和特征一致性及基于可靠样本的混合增强的两阶段的方法,均用于处理医学图像中的噪声标签问题。

1.2 对比学习

近期研究表明,基于对比的相似性学习框架在表征学习中表现出良好的性能,该方法旨在最大化正对的相似性和最小化负对的相似性^[12]。其中正对是指具有相似性的样本对,通常是来自同一类别或具有相似属性的样本对;负对是指具有差异性的样本对,通常是来自不同类别或具有不同属性的样本对。李俊哲等人^[13]采用对比学习最大化两个增强视图的相似度,并引入了基于类原型的标签去噪模块来避免噪声标签的影响。Li等人^[14]提出一种选择性监督对比学习方法,从构建的样本对中学习相似度分布,进而识别出干净样本。Xu等人^[15]提出了一种新的NAC (noisy-as-clean)策略来训练自监督去噪网络。该策略将损坏的图像视为“干净目标”,将损坏图像和第2个类似损坏的图像的组合作为输入,在噪声微弱的情况下,学习监督网络的最佳参数。

2 NLSS 模型

2.1 网络架构

本文提出一种基于对比学习的噪声标签样本选择模型,简称NLSS (noise label sample selection)。模型分为特征提取、样本选择、标签重分配3个部分。模型的整体构成如图1所示。首先,输入图像经过特征提取模块提取特征向量。其次,采用样本选择模块对数据集进行划分。接着,针对噪声数据集的样本对其进行标签重分配,并且此部分数据参与训练。最后,将交叉熵损失函数和对比损失函数结合起来对网络进行训练。

2.2 特征提取

为训练模型使用的图像数据 x_i 和对应标签 y_i 表示为 $D_k = \{(x_i, y_i)\}_i^N$,其中 N 为训练Minibatch的大小,对 D_k 进行两种不同的增强方式,一种为随机翻转,另一种为随机增加噪声扰动,将 D_q , D_k 和 D_v 输入编码器 f 中提取特征 f_q , f_k , f_v 。该过程表示为:

$$F = f(x) \quad (1)$$

其中, f 代表特征提取器, x 表示输入图像, F 表示提取的特征,该模块旨在捕获图像抽象的特征表示。

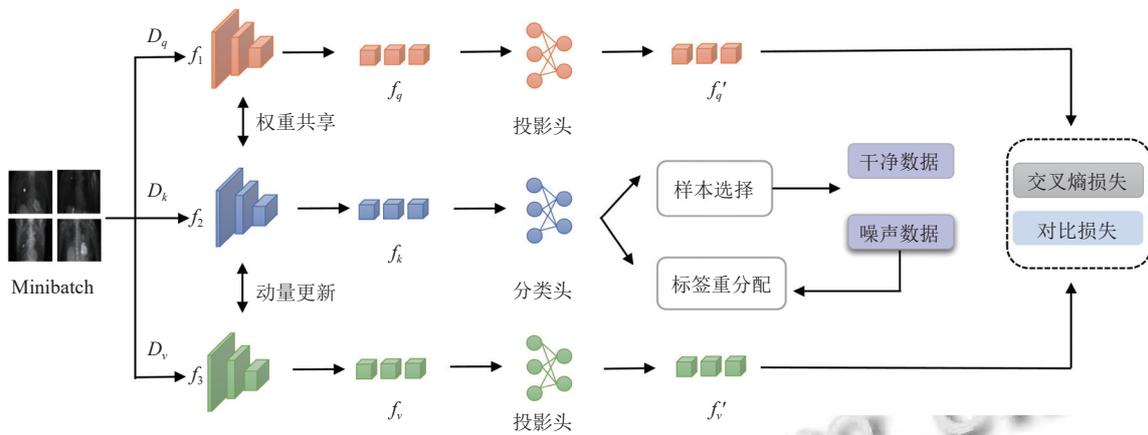


图1 网络整体框架

此外, f_3 中的模型参数由 f_2 中的模型参数进行动量更新, 动量更新机制可表示为:

$$\theta_v \leftarrow m\theta_v + (1 - m)\theta_k \quad (2)$$

其中, θ_v 表示 f_3 的模型参数, θ_k 表示 f_2 的模型参数, $m \in [0, 1)$ 为动量系数.

模型特征提取部分使用 ResNet-18^[16] 作为骨干网络提取图像的特征 f_q, f_k, f_v . f_q 经过投影头映射到相同维度的表示 f'_q , f_k 经过分类器预测分类输出, f_v 由动量编码器导出, 与 f_q 具有相同的维度. f'_v 是 f_v 经过投影头后提取的特征, 使用对比损失函数来学习 D_q 和 D_v 的鲁棒表示.

2.3 样本选择

一些研究^[17]将交叉熵损失小的样本视为干净样本, Co-teaching^[18]和 Co-teaching+^[19]在每个小批次中选择损失较小的样本视为干净样本, 虽然在一定程度上避免了噪声累积, 但是不同小批量中的噪声比会有波动. 为了缓解不同小批量中噪声比不平衡的问题, 本文采用 JS 散度 (Jensen-Shannon divergence)^[20]来选择干净样本. JS 散度是衡量两个概率分布之间的差异, 取值范围在 $[0, 1]$ 之间, 可作为全局选择度量. 样本选择模块的结构如图 2 所示.

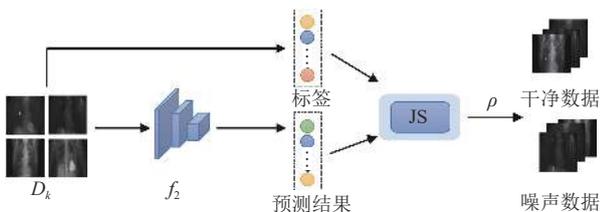


图2 样本选择模块结构

给定样本的真实标签分布 $y_i = [y^1, y^2, \dots, y^N]$ 和网络的预测结果分布 $p_i = [p_i^1, p_i^2, \dots, p_i^N]$, 首先根据以下

公式计算 d_i 的取值, 为了避免 $\log(0)$ 的情况, 本文采用标签平滑^[21]的方法对原始标签进行处理.

$$d_i = JS(p_i \parallel y_i) \quad (3)$$

$$JS(p_i \parallel y_i) = \frac{1}{2}KL\left(p_i \parallel \frac{p_i + y_i}{2}\right) + \frac{1}{2}KL\left(y_i \parallel \frac{p_i + y_i}{2}\right) \quad (4)$$

其中, $KL(\cdot \parallel \cdot)$ 是 KL (Kullback-Leibler) 散度函数^[22].

其次计算样本 x_i 的概率 ρ :

$$\rho(x_i) = 1 - d_i, \quad d_i \in [0, 1] \quad (5)$$

最后根据阈值, 对样本进行划分.

$$D_k = \begin{cases} D_{\text{clean}}, & \rho \geq \tau_{\text{clean}} \\ D_{\text{noisy}}, & \rho < \tau_{\text{clean}} \end{cases} \quad (6)$$

由于数据集存在噪声, 静态阈值无法适应变化, 对此本文使用可调整的动态阈值进行干净数据集和噪声数据集的划分.

2.4 标签重分配

输入网络的样本经过特征提取模块和样本选择模块后, 被划分为干净数据集和噪声数据集, 将噪声数据集中的样本视为无标签样本, 为了不丢弃这些噪声样本并充分挖掘其潜在的信息, 本文采用了伪标签提升策略来进行标签的重分配. 具体而言, 使用分类网络对这些噪声样本进行预测, 得到它们的伪标签. 将这些伪标签作为噪声样本的监督信息, 帮助网络进行训练和优化. 对于干净数据集, 则依旧使用原先的标签作为监督. 通过不断迭代训练, 网络可以利用伪标签提升策略逐渐学习到更准确的特征表示和分类决策. 标签重分配过程如式 (7) 所示:

$$\hat{y} = \begin{cases} y_i, & x_i \in D_{\text{clean}} \\ p_i, & x_i \in D_{\text{noisy}} \end{cases} \quad (7)$$

其中, y_i 表示给定的样本标签, p_i 表示网络的预测结果, \hat{y} 表示替代原始标签的伪标签。

2.5 损失函数

本文采用交叉熵损失函数和对比损失函数的组合训练模型。

2.5.1 交叉熵损失函数

交叉熵损失函数^[23]是机器学习和深度学习中常用的一种损失函数,通常用于分类问题的模型训练中。本文将非小细胞肺癌患者的转移状态视为二分类问题,构建二元交叉熵损失函数,实验数据中发生转移个数和未发生转移个数之比约为 1:5,为了减轻数据不平衡的影响,本文采用加权交叉熵损失函数^[24]。

$$L_{ce} = -\frac{1}{N} \sum_i w_p (y_i \log(\hat{y})) + w_n ((1 - y_i) \log(1 - \hat{y})) \quad (8)$$

$$w_p = \frac{1}{pos} \times total \quad (9)$$

$$w_n = \frac{1}{neg} \times total \quad (10)$$

其中, N 表示训练样本数量, y_i 表示给定样本的真实标签, \hat{y} 表示重新分配后的样本标签, pos 表示正例样本个数, neg 表示负例样本个数, $total$ 表示样本总数。

2.5.2 对比损失函数

对比学习已经成为自监督方法的重要组成部分。在 SimCLR^[25] 论文中,研究人员指出数据增强是至关重要的,而非线性投影头可以进一步提高学习到的表示质量。受此启发,本文使用了余弦相似度^[26]来计算两个不同增强视图提取的特征之间的对比损失。具体的计算公式如下:

$$L_q^{ctr} = -\log \frac{\exp(\text{sim}(z_q, z_v)/\tau)}{\sum_{m=1}^{2n} \exp(\text{sim}(z_q, z_m)/\tau)} \quad (11)$$

$$L_{ctr} = \sum_{q=1}^{2N} L_q^{ctr} \quad (12)$$

其中,经过两种不同增强方式增强后的样本数量为 $2N$, z_q 和 z_v 分别表示经过投影头后得到的特征向量, z_m 表示除了当前样本对应的增强视图之外的其他样本的增强视图, τ 是调节余弦相度量级的一个温度参数^[27],用于控制余弦相似度的量级, $\text{sim}(z_q, z_v) = \frac{z_q^T z_v}{\|z_q\| \cdot \|z_v\|}$ 表示特征向量 z_q, z_v 之间的相似度。最终的损失函数如下:

$$L_{total} = \alpha \times L_{ce} + \beta \times L_{ctr} \quad (13)$$

其中, α, β 表示权重系数,在本实验中 α 设置为 0.8, β 设置为 0.4。

3 实验分析

3.1 实验设置

本实验基于 Ubuntu 20.04 操作系统,使用 PyTorch 1.7.1 训练框架^[28],选取 ResNet-18 作为特征提取器,输出的特征经过投影头后为 128 维数据,优化方法为 SGD 梯度优化器^[29],动量值为 0.99,学习率为 0.001,网络总共设置的 epoch 为 50。

3.2 数据集

本文的实验数据来自芝加哥大学医学院 328 名 NSCLC 患者的 PET/CT 影像数据,112 例患者既有医生的诊断结果,又有病理学检测的金标准,216 例患者仅有医生的诊断结果。在此数据集中,328 名患者的医生诊断结果中含有错误的结果,视为部分含有噪声标签,而 112 例患者的金标准诊断结果视为正确标签。数据集划分策略为:将 112 例患者的数据集划分为 80% 的训练集和 20% 的测试集,216 名患者数据集全部加入训练集中,使用五折交叉验证^[30]的方法评估网络的性能。本实验在训练过程中全部使用医生标签进行训练,而使用 112 例患者的正确标签进行测试。

本文提出深度网络输入为 PET 图像正面最大强度投影生成 MIP (maximum intensity projection) 图像^[31]。如图 3 所示。MIP 是一种图像后处理技术,常用于医学影像学中,运用透视法获得二维图像,是通过计算沿着被扫描物每条射线上所遇到的最大密度像素而产生的。MIP 相比于处理完整的 3D 图像,图像的特征表示更加紧凑,具有较低的维度,使得模型更容易学习和泛化到新的未见过的数据,更容易捕捉到关键的结构和特征。

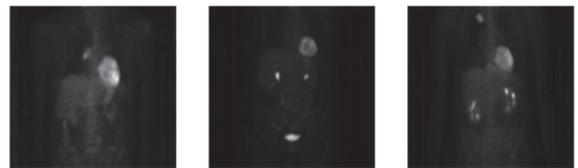


图 3 由 PET 图像投影的正面 MIP 图像

3.3 评价指标

本实验选取准确率 (Accuracy), 灵敏度 (Sensitivity), 特异性 (Specificity), AUC (area under the curve), 作为模型的评价指标。

$$Sensitivity = \frac{TP}{TP + FN} \quad (14)$$

$$Specificity = \frac{TN}{TN + FP} \quad (15)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (16)$$

其中, TP (true positive) 为将正样本正确地预测为正样本的个数, FP (false positive) 为将负样本错误地预测为正样本的个数, TN (true negative) 为将负样本正确地预测为负样本的个数, FN (false negative) 为将正样本错误地预测为负样本的个数. 本文中将病理检测结果为淋巴结转移的患者视为正样本.

3.4 实验结果分析

为了验证本模型样本选择模块和标签重分配模块的有效性, 进行如下实验.

由于 112 例患者有病理金标准, 故在训练阶段后输出网络认为干净样本的医生结果的标签与金标准的结果进行对比, 计算方法如式 (17) 所示, 对比结果的准确率如图 4 所示.

$$acc1 = \frac{n}{T} \times 100\% \quad (17)$$

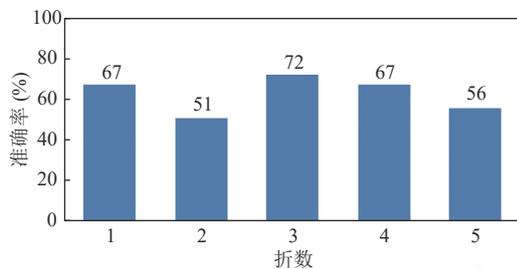


图 4 网络挑选结果较于金标准的准确率

五折实验中最高的准确率为 72%, 平均准确率为 62%. 同时对网络认为不干净的样本经过标签重分配后输出的结果, 进行标签二值化后与金标准的结果进行再次对比, 计算方法如式 (18) 所示, 对比结果的准确率如图 5 所示.

$$acc2 = \frac{m}{T - n} \times 100\% \quad (18)$$

式 (17) 和式 (18) 中, n 为网络认为干净的样本数, T 为训练阶段有金标准的样本数, m 为网络认为不干净的样本经过标签重分配后与金标准结果对比正确的个数.

五折实验中最高的准确率为 45%, 平均准确率为

41%. 综合以上二者, 即判断正确的样本和标签重分配修改正确的样本经过与金标准进行对比, 准确率如图 6 所示. 五折实验中最高的准确率为 83%, 平均准确率为 78%.

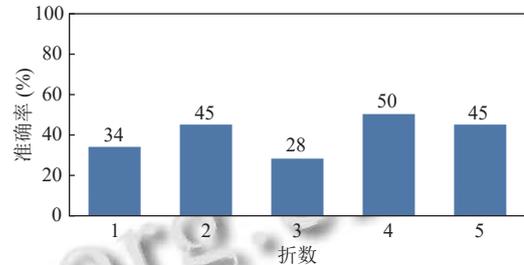


图 5 标签重分配结果较于金标准的准确率

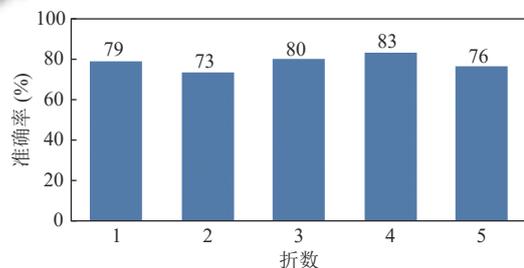


图 6 去噪模块较于金标准的准确率

为了证明本模型的有效性, 还将实验结果与放射科医生诊断结果以及近年来针对噪声标签任务较常用模型进行对比, 其中 Co-teaching^[18]方法是同时训练两个网络, 在每个批量数据中, 每个网络都将具有小损失的样本视为有用的知识, 并向其另一个网络传授这些样本以更新参数; JoCoR^[20]使用两个网络对相同的小批数据进行预测, 并计算每个训练示例的联合损失, 选择损失较小的样本, 同时更新两个网络的参数; Jo-SRC^[32]以对比学习的方式训练网络, 使用来自每个样本的两个不同视图的预测用于估计其干净或分布外的“可能性”, 通过引入一致性正则化来提高模型泛化性能; AsyCo^[33]采用多视图共识, 使用训练标签和模型预测之间的分歧来选择干净和有噪声的样本. 对比实验结果如表 1 所示. 由于放射科医生的结果是二元标签而不是概率的形式, 因此本文提出的模型仅与其他实验方法比较 AUC.

由表 1 可以看出, 本文提出的方法在各项指标上均优于医生的诊断结果, 其余各项指标较其他方法均有明显的提升, 准确率和特异性的提升较高, 与 Co-teaching, JoCoR, Jo-SRC, AsyCo 相比, 准确率分别提高了 0.06

(7.9%), 0.21 (34.4%), 0.12 (17.1%), 0.08 (10.8%), 特异性分别提高了 0.06 (7.3%), 0.1 (12.8%), 0.15 (20.5%), 0.08 (10%)。图 7 以可视化的形式展示了本实验方法与其他方法在各项评价指标的结果。

表 1 与其他方法和放射科医生诊断结果对比实验结果

Model	Accuracy	Sensitivity	Specificity	AUC
Co-teaching	0.76	0.44	0.82	0.63
JoCoR	0.61	0.44	0.78	0.62
Jo-SRC	0.70	0.50	0.73	0.62
AsyCo	0.74	0.44	0.80	0.62
医生诊断	0.79	0.44	0.86	—
Ours	0.82	0.50	0.88	0.69

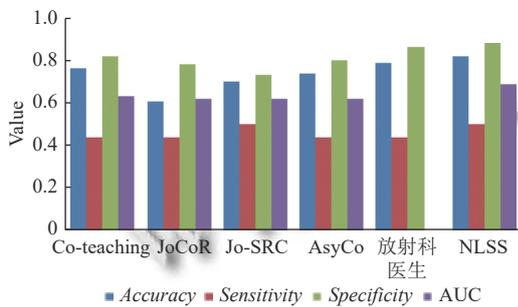


图 7 不同模型的评价指标对比结果图

3.5 消融实验

本文进行消融实验来验证模块的有效性, 消融实验策略有: 仅使用特征提取模块训练, 仅使用特征提取模块与对比损失训练, 仅使用特征提取模块和标签重分配模块进行训练, 使用本文方法进行训练, 将其实验结果进行对比, 如表 2 所示。

表 2 消融实验

特征提取	对比损失	标签重分配	Accuracy	Sensitivity	Specificity	AUC
√	—	—	0.69	0.33	0.76	0.54
√	√	—	0.78	0.28	0.87	0.58
√	—	√	0.78	0.39	0.85	0.62
√	√	√	0.82	0.50	0.88	0.69

上述结果说明了对比损失可利用内在结构的相似性促使模型学习到具有判别性的特征表示, 有效地应对噪声标签, 避免了小样本损失策略造成的噪声累积; 标签重分配模块引入了一种灵活的伪标签提升策略, 对噪声数据进行修正, 在初始分类预测的基础上细化带有伪标签的噪声数据的监督信息, 提高模型的性能和泛化能力。本文使用 3 种方式的组合, 最终达到在整体上分类效果最佳的成绩。

3.6 可视化展示

为了增强模型的可解释性, 在数据集中随机选择

了 4 张 MIP 图片, 并绘制其注意力热图, 结果表明, 网络能够更好地关注到目标区域, 即肺部中的肿瘤部位, 对于高亮的器官部位, 并没有过度关注。可视化结果如图 8 所示。

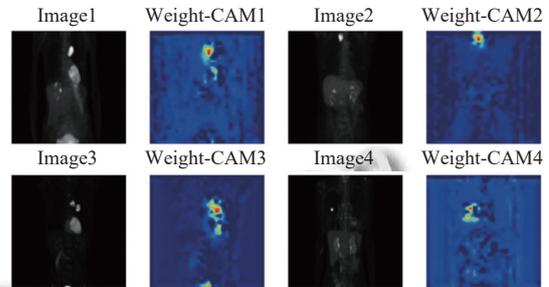


图 8 注意力热图

4 结论与展望

本研究探索了基于对比学习的标签噪声样本选择模型在非小细胞肺癌患者淋巴结转移诊断中的应用。通过对比学习的方法, 减少了噪声标签对于模型性能的负面影响, 学习到更具判别性和鲁棒性的特征表示, 以非小细胞患者数据集为例进行了验证, 实验结果表明模型取得了良好的性能。本研究还需要进一步验证和扩展, 以确保其在不同数据集和临床场景中的泛化能力和可靠性。

参考文献

- Sung H, Ferlay J, Siegel RL, *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 2021, 71(3): 209–249. [doi: 10.3322/caac.21660]
- 陈武飞, 马庄宣, 李铭. 影像组学在非小细胞肺癌诊疗中的应用进展. *中国医学影像学杂志*, 2023, 31(7): 781–784. [doi: 10.3969/j.issn.1005-5185.2023.07.021]
- Hendrycks D, Mazeika M, Wilson D, *et al.* Using trusted data to train deep networks on labels corrupted by severe noise. *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS 2018)*. Montreal: Curran Associates Inc., 2018. 10477–10486.
- Jiang L, Zhou ZY, Leung T, *et al.* MentorNet: Learning data-driven curriculum for very deep neural networks on corrupted labels. *Proceedings of the 35th International Conference on Machine Learning*. Stockholm: PMLR, 2018. 2304–2313.

- 5 Liu JR, Li RR, Sun C. Co-correcting: Noise-tolerant medical image classification via mutual label correction. *IEEE Transactions on Medical Imaging*, 2021, 40(12): 3580–3592. [doi: [10.1109/TMI.2021.3091178](https://doi.org/10.1109/TMI.2021.3091178)]
- 6 Zhang ZL, Sabuncu MR. Generalized cross entropy loss for training deep neural networks with noisy labels. *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS 2018)*. Montreal: Curran Associates Inc., 2018. 8792–8802.
- 7 Xu YL, Cao P, Kong YQ, *et al.* L_{DMI} : A novel information-theoretic loss function for training deep nets robust to label noise. *Proceedings of the 33rd International Conference on Neural Information Processing Systems (NeurIPS)*. Vancouver: Curran Associates Inc., 2019. 559.
- 8 Ma XJ, Huang HX, Wang YS, *et al.* Normalized loss functions for deep learning with noisy labels. *Proceedings of the 37th International Conference on Machine Learning*. PMLR, 2020. 6543–6553.
- 9 Xue C, Dou Q, Shi XY, *et al.* Robust learning at noisy labeled medical images: Applied to skin lesion classification. *Proceedings of the 16th IEEE International Symposium on Biomedical Imaging (ISBI)*. Venice: IEEE, 2019. 1280–1283. [doi: [10.1109/ISBI.2019.8759203](https://doi.org/10.1109/ISBI.2019.8759203)]
- 10 Min SB, Chen XJ, Zha ZJ, *et al.* A two-stream mutual attention network for semi-supervised biomedical segmentation with noisy labels. *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. Honolulu: AAAI, 2019. 4578–4585.
- 11 Xing XH, Chen Z, Gao ZF, *et al.* Gradient and feature conformity-steered medical image classification with noisy labels. *Proceedings of the 26th International Conference Medical Image Computing and Computer Assisted Intervention*. Vancouver: Springer, 2023. 75–84.
- 12 Chen T, Kornblith S, Norouzi M, *et al.* A simple framework for contrastive learning of visual representations. *Proceedings of the 37th International Conference on Machine Learning (ICML)*. PMLR, 2020. 1597–1607.
- 13 李俊哲, 曹国. 基于对比学习的标签带噪图像分类. *计算机系统应用*, 2023, 32(12): 104–111. [doi: [10.15888/j.cnki.csa.009339](https://doi.org/10.15888/j.cnki.csa.009339)]
- 14 Li SK, Xia XB, Ge SM, *et al.* Selective-supervised contrastive learning with noisy labels. *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE, 2022. 316–325.
- 15 Xu J, Huang Y, Cheng MM, *et al.* Noisy-as-clean: Learning self-supervised denoising from corrupted image. *IEEE Transactions on Image Processing*, 2020, 29: 9316–9329. [doi: [10.1109/TIP.2020.3026622](https://doi.org/10.1109/TIP.2020.3026622)]
- 16 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE, 2016. 770–778. [doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90)]
- 17 Song H, Kim M, Park D, *et al.* Learning from noisy labels with deep neural networks: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 34(11): 8135–8153. [doi: [10.1109/TNNLS.2022.3152527](https://doi.org/10.1109/TNNLS.2022.3152527)]
- 18 Han B, Yao QM, Yu XR, *et al.* Co-teaching: Robust training of deep neural networks with extremely noisy labels. *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS 2018)*. Montreal: Curran Associates Inc., 2018. 8536–8546.
- 19 Yu XR, Han B, Yao JC, *et al.* How does disagreement help generalization against label corruption? *Proceedings of the 36th International Conference on Machine Learning*. Long Beach: PMLR, 2019. 7164–7173.
- 20 Wei HX, Feng L, Chen XY, *et al.* Combating noisy labels by agreement: A joint training method with co-regularization. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2020. 13723–13732.
- 21 Szegedy C, Vanhoucke V, Ioffe S, *et al.* Rethinking the inception architecture for computer vision. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE, 2016. 2818–2826.
- 22 李晓艳, 张子刚, 张逸石, 等. 一种基于KL散度和类分离策略的特征选择算法. *计算机科学*, 2012, 39(12): 224–227. [doi: [10.3969/j.issn.1002-137X.2012.12.053](https://doi.org/10.3969/j.issn.1002-137X.2012.12.053)]
- 23 Le PB, Nguyen ZT. ROC curves, loss functions, and distorted probabilities in binary classification. *Mathematics*, 2022, 10(9): 1410. [doi: [10.3390/math10091410](https://doi.org/10.3390/math10091410)]
- 24 程磊, 吴晓富, 张索非. 数据集类别不平衡性对迁移学习的影响分析. *信号处理*, 2020, 36(1): 110–117.
- 25 Zeng H, Cui XH. SimCLRT: A simple framework for contrastive learning of rumor tracking. *Engineering Applications of Artificial Intelligence*, 2022, 110: 104757. [doi: [10.1016/j.engappai.2022.104757](https://doi.org/10.1016/j.engappai.2022.104757)]
- 26 Yi L, Liu S, She Q, *et al.* On learning contrastive representations for learning with noisy labels. *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans: IEEE, 2022. 16661–16670.

- 27 He KM, Fan HQ, Wu YX, *et al.* Momentum contrast for unsupervised visual representation learning. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 9726–9735.
- 28 Paszke A, Gross S, Massa F, *et al.* PyTorch: An imperative style, high-performance deep learning library. Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2019. 721.
- 29 常禧龙, 梁琨, 李文涛. 深度学习优化器进展综述. 计算机工程与应用, 2024, 60(7): 1–12. [doi: [10.3778/j.issn.1002-8331.2307-0370](https://doi.org/10.3778/j.issn.1002-8331.2307-0370)]
- 30 张剑飞, 王真, 崔文升, 等. 基于交叉验证和神经网络融合
的医学数据分类. 齐齐哈尔大学学报(自然科学版), 2019, 35(4): 1–5.
- 31 Aung N, Kechadi T, Chen LM, *et al.* IP-UNet: Intensity projection UNet architecture for 3D medical volume segmentation. arXiv:2308.12761, 2023.
- 32 Yao YZ, Sun ZR, Zhang CY, *et al.* Jo-SRC: A contrastive approach for combating noisy labels. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021. 5188–5197.
- 33 Liu FB, Chen YH, Wang C, *et al.* Asymmetric co-teaching with multi-view consensus for noisy label learning. arXiv: 2301.01143, 2023.

(校对责编: 张重毅)