

文章编号 1004-924X(2025)24-3931-09

卷积神经网络结合荧光光谱在乳腺癌分型上的应用研究

黄翰林¹, 官 韬¹, 秦艳利¹, 沈龙海¹, 齐东丽¹, 李宏达², 冯 瑜^{1*}

(1. 沈阳理工大学 理学院, 辽宁 沈阳 110159;

2. 沈阳理工大学 装备工程学院, 辽宁 沈阳 110159)

摘要: 乳腺癌是对女性健康威胁极大的肿瘤之一, 其分子亚型与患者临床治疗用药等关系密切, 准确判别乳腺癌的分子亚型对临床指导治疗和改善预后有着重要作用。本文提出一种基于一维卷积神经网络(1D-CNN)与荧光光谱技术相结合的端到端分析方法, 分析过程中无需光谱预处理, 流程简化且保留更多有效信息。自主设计搭建激光诱导荧光光谱系统, 采集 4 类乳腺癌亚型组织的自体荧光光谱, 通过高斯函数拟合光谱数据, 对乳腺癌组织生化信息进行定量与定性分析。利用 1D-CNN 模型进一步挖掘光谱数据深层特征, 实现对乳腺癌亚型的高精度识别。此外, 采用梯度加权类激活映射(Grad-CAM)技术对模型关注的关键光谱区域进行可视化分析, 增强了模型的可解释性。实验结果表明: 乳腺癌亚型组织的分类准确率达 98.7%, 召回率为 97.2%, 精确率为 98.1%, F1 分数为 97.7%。本研究所提出的方法展现出良好的临床应用前景, 可为乳腺癌分子分型的智能诊断提供有效的技术路径。

关键词: 光谱分析; 乳腺癌分型; 荧光光谱; 一维卷积神经网络

中图分类号: O433.4; **文献标识码:** A

doi: 10.37188/OPE.20253324.3931

CSTR: 32169.14.OPE.20253324.3931

Application of convolutional neural networks combined with fluorescence spectroscopy in breast cancer subtyping

HUANG Hanlin¹, GUAN Tao¹, QIN Yanli¹, SHEN Longhai¹, QI Dongli¹, LI Hongda², FENG Yu^{1*}

(1. School of Science, Shenyang Ligong University, Shenyang 110159, China;

2. Equipment Engineering College, Shenyang Ligong University, Shenyang 110159, China)

* Corresponding author, E-mail: 14025921@qq.com

Abstract: Breast cancer is one of the most significant threats to women's health. Its molecular subtypes are closely related to clinical treatment and medication choices for patients. Accurate identification of breast cancer molecular subtypes plays a crucial role in guiding clinical treatment and improving prognosis. This paper proposed an end-to-end analysis method integrating one-dimensional convolutional neural networks (1D-CNN) with fluorescence spectroscopy technology. The process required no spectral preprocessing, simplifying workflows while preserving more effective information. An independently designed and constructed laser-induced fluorescence spectroscopy system was used to collect autofluorescence spec-

收稿日期: 2025-09-15; 修订日期: 2025-10-17.

基金项目: 国家自然科学基金(No. 51974063); 辽宁省科学技术计划项目(No. 2023JH2/10700025)

tra from four breast cancer subtype tissues. Gaussian function fitting of the spectral data enabled quantitative and qualitative analysis of the biochemical information in breast cancer tissues. The 1D-CNN model further extracts deep features from spectral data to achieve high-precision identification of breast cancer subtypes. Additionally, Gradient-Weighted Class Activation Maps (Grad-CAM) technology was employed to visualize key spectral regions of interest, enhancing model interpretability. The experimental results show that the classification accuracy rate of breast cancer subtype tissues is 98.7%, the recall rate is 97.2%, the precision rate is 98.1%, and the F1 score is 97.7%. The method proposed by our institute demonstrates promising clinical application prospects and provides an effective technical pathway for the intelligent diagnosis of breast cancer molecular subtypes.

Key words: spectral analysis; breast cancer typing; fluorescence spectrum; 1D convolutional neural network

1 引言

乳腺癌是全球女性中发病率最高的恶性肿瘤,对女性健康构成了严重威胁。据世界卫生组织统计,乳腺癌的发病率和死亡率在全球范围内持续上升^[1]。乳腺癌的分子分型对于预测疾病进展、评估预后以及制定个体化治疗方案具有关键作用。1999年,美国的研究团队首次对乳腺癌进行了分子水平的分型界定^[2]。2011年 St. Gallen^[3]共识会议提出,将雌激素受体(ER)、孕激素受体(PR)、人类表皮生长因子受体2(HER-2)以及细胞增殖标志Ki-67的表达水平作为乳腺癌分子分型的重要依据。通过免疫组织化学(IHC)检测手段,可将乳腺癌分为4类主要分子亚型:Luminal A型、Luminal B型、HER-2过表达型和三阴型,不同亚型在治疗策略的选择和患者预后方面存在显著差异。

目前,乳腺癌的临床分型主要采用IHC和基因表达谱分析等方法,但该过程操作复杂、耗时、易受检测人员主观影响。荧光光谱^[4, 5]技术作为一种非破坏性、快速、灵敏度高的检测方法,在生物医学领域受到广泛关注。生物样品的荧光光谱数据通常具有高维度、非线性和高噪声等缺点,导致传统数据分析方法难以有效挖掘数据特征,卷积神经网络(Convolutional Neural Network, CNN)作为一种深度学习模型,通过模拟人类大脑处理视觉信息的方式,利用卷积操作对数据进行特征提取,从而实现数据的高效分类和识别^[6]。CNN模型由于具备局部连接和权重共享这两个关键特征,在特征学习与分类性能上显著优于传统神经网络,展现出

更强的泛化能力^[7]。例如,2019年赵勇^[8]等将拉曼光谱与一维卷积神经网络(1D-CNN)相结合,对三种相似的雌激素粉末进行分类,分类准确率高达98.26%。2024年,于水^[9]等提出了一种基于并联卷积神经网络的建模分析方法,采用Inception网络块思想,通过增加网络宽度、减小网络深度,从不同尺度上提取光谱特征,减小光谱数据特征的损失,提高网络性能,实现对谷物、柴油等的近红外光谱数据精确分类。上述研究表明,将CNN与光谱学相结合进行分类具有广泛的应用前景。

本文提出将1D-CNN模型应用于乳腺癌组织分型的光谱分析方案,设计并搭建荧光光谱采集系统,获取4类乳腺癌组织的自体荧光光谱,设计数据增强方案,构建数据集,优化超参数并完成模型训练,实现对乳腺癌分子分型的快速精确判别。

2 实验及算法模型

2.1 样本准备

本研究所用乳腺组织切片由肿瘤医院提供,共计200片,涵盖4类不同分型的乳腺癌样本。其中,Luminal A型25片,Luminal B型131片,HER-2过表达型29片,三阴型15片。每片组织切片的厚度为10~20 μm ,年龄分布在33岁至72岁之间。

2.2 光谱数据采集

实验中采用波长为405 nm的连续激光器(MDL-III-405)作为激发光源,激光束首先经透镜1聚焦后,被二向色镜(SDM410S)反射,再经

透镜2聚焦至样品表面,激发样品产生荧光信号。荧光信号穿过二向色镜后,由透镜3进行二次聚焦,随后通过光纤(QP400-1-UV-YIS)将信号耦合传输至光谱仪(USB4000)进行数据采集与分析。

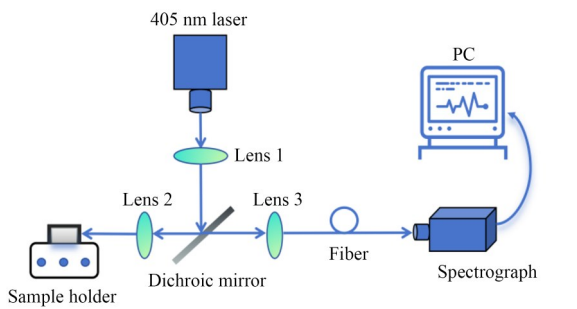


图1 实验装置示意图
Fig. 1 Schematic diagram of the experimental setup

2.3 光谱分析

研究表明,乳腺癌分子分型之间的内部生化环境存在显著差异^[10],如代谢通路、免疫微环境等。由于其组织中各种荧光基团的含量和分布不同,导致荧光光谱特征发生改变,具体表现为荧光特征峰的位移和强度发生改变。4类亚型组织的平均光谱如图2所示,可以看出不同亚型组织平均光谱的谱峰位置基本相似,但强度存在一定差异。

自体荧光光谱的变化主要源于光激发区域内多种内源性荧光基团含量的变化。这些荧光基团包括游离的还原型烟酰胺腺嘌呤二核苷酸

(NADH)、蛋白质结合 NADH、游离的黄素腺嘌呤二核苷酸(FAD)、中性和酸性形式的卟啉、色氨酸、胶原蛋白以及弹性蛋白^[11]。由于蛋白质结合 FAD 的荧光强度极弱,故本研究中暂不考虑;此外,色氨酸、胶原蛋白和弹性蛋白主要在中紫外和长波紫外光激发下才具有显著荧光效应,因此本研究所采用 405 nm 激发波长下亦不做分析。

表1 拟合参数:不同荧光基团的激发波长和发射峰^[12, 13]
Tab. 1 Fitted parameter: Excitation wavelengths and emission peaks of different fluorescent groups (nm)

Fluorophores	Excitation	Emission
Free NADH	405	487
Protein-bound NADH	405	501
Free FAD	405	544
Neutral porphyrins	405	590
Acidic porphyrins	405	630
Tryptophan	287	342
Collagen	339	380
Elastin	351	410

综上,本研究重点观察在 405 nm 激光激发条件下具有代表性的5类内源性荧光基团:游离 NADH、蛋白质结合 NADH、游离 FAD,以及卟啉的中性和酸性形式。如图3所示,对4类组织的自体荧光光谱进行高斯分峰拟合,结果表明采用5项高斯叠加可获得高度吻合的拟合效果(拟合校正系数 R^2 均在0.99以上),每个高斯峰分别对应于不同荧光基团的特征发射谱。

NADH和FAD的比值是反映细胞氧化还原状态和代谢活性的重要指标^[14-15],在本研究中,将光学氧化还原比(ORR)计算为NADH(游离和蛋白质结合)与游离FAD的荧光强度比即 $I_{\text{NADH}}/I_{\text{FAD}}$,并且研究表明由于肿瘤组织乳酸增多,局部pH值降低,使得卟啉更易质子化,质子化后卟啉的电子结构发生改变导致其荧光峰发生红移现象,并在630 nm附近增强^[16]。所以按照上述拟合程序对乳腺癌4类亚型组织的平均光谱进行分析并得到每类亚型组织的 $I_{\text{NADH}}/I_{\text{FAD}}$ 和

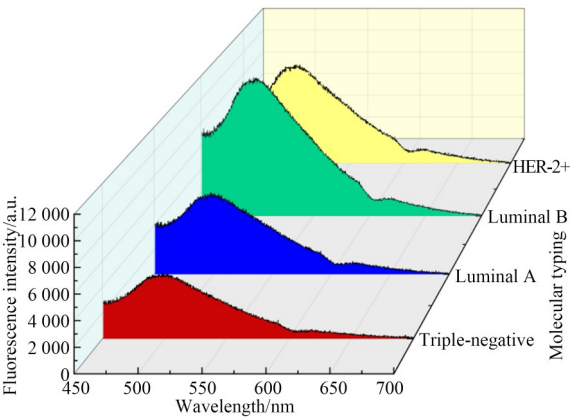


图2 不同亚型组织的平均荧光光谱
Fig. 2 Average fluorescence spectra of different subtypes of tissues

$I_{630\text{ nm}}/I_{590\text{ nm}}$,作为分型依据。

如表 2 所示,乳腺癌中恶性程度更高的 HER-2+型和三阴型组织中,NADH的含量相对较高,这是因为 NADH 水平是癌症干细胞代谢水平的关键指标,其显著提高表示癌症干细胞活跃^[17],并且肿瘤细胞内线粒体电子传递链酶受损会使游离 NADH 水平明显提高^[18]。此外,Masilamani . V^[19]的工作表明在癌症病发的所有阶段,590 nm 处的荧光强度都较为相似,但随着肿瘤恶性程度的进展,630 nm 处的荧光强度迅速增加,这与本文的实验结果相吻合。

上述研究通过对乳腺癌 4 类亚型组织的荧光光谱进行分析,发现这 4 类组织的荧光信号均可视为由 5 类内源性荧光基团的光谱叠加构成。进一步分析表明,不同亚型组织在化学成分上存在

表 2 乳腺癌组织代谢分析

Tab. 2 Metabolic analysis of breast cancer tissues

Molecular typing	$I_{\text{NADH}}/I_{\text{FAD}}$	$I_{630\text{ nm}}/I_{590\text{ nm}}$
Luminal A	1. 94	0. 87
Luminal B	1. 86	0. 92
HER-2+	2. 37	1. 12
Triple-negative	2. 51	1. 46

差异,利用荧光强度比值法可以在一定程度上实现对乳腺癌亚型组织的区分。为了更深入挖掘光谱数据中的潜在特征,引入 1D-CNN 模型,旨在实现对乳腺癌 4 类亚型组织的快速、准确判别。

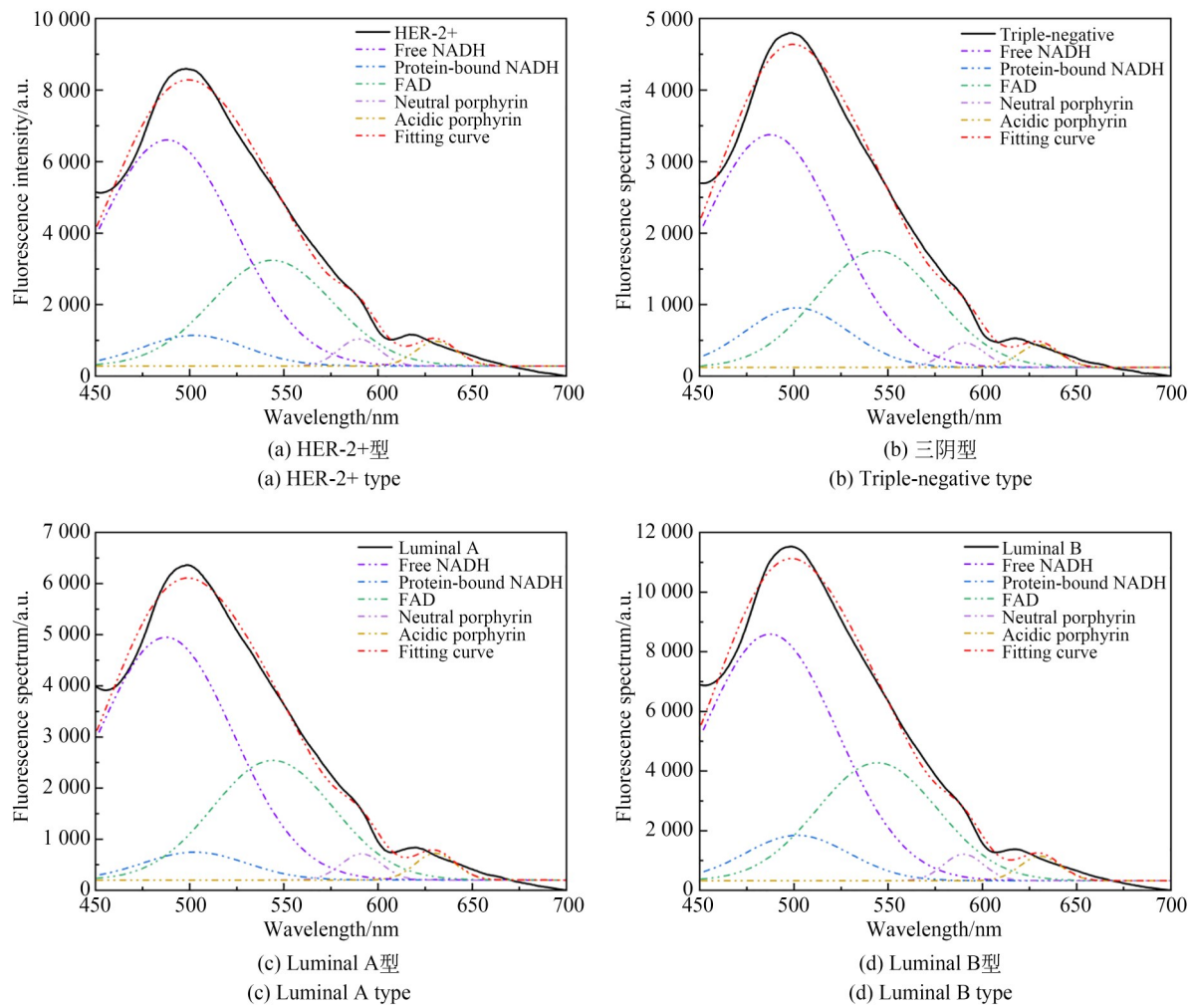


图 3 四种乳腺癌亚型光谱的高斯分解

Fig. 3 Gaussian decomposition of the spectra of four breast cancer subtypes

2.4 数据增广

卷积神经网络的优越性能得益于其对大规模数据集的学习能力,以便充分学习数据特征避免过拟合。在临床应用中,常因样本数量有限而难以满足深度学习对数据量的需求,同时本研究亦存在亚型样本数量不均衡问题。为此,本研究采用系统性的数据处理流程。首先,通过SMOTE过采样算法对少数类样本进行特征空间插值,使四类亚型的样本数量达到均衡。并在此基础上,采用三种数据增强策略:(1)在预设的位移扰动范围内,对原始光谱进行左右平移。(2)添加均值为0,标准差为0.01~0.02的高斯噪声。(3)线性叠加同一类多个光谱数据,使用为1的随机比例因子对每一类光谱数据进行扩展。三者均是机器学习中用于数据增强的常用技术,平移可以帮助提高模型针对仪器校准误差的鲁棒性^[20],添加噪声可提高模型处理有噪声或低噪声比数据的能力^[21],而线性叠加可以保留类别特征,模拟混合效应,提升模型对类内变化的鲁棒性。

2.5 模型建立

鉴于荧光光谱的数据特性,本文构建了如图4所示的1D-CNN模型用于乳腺癌亚型分类。该模型架构在保留传统CNN层次结构的基础上进行了针对性优化,模型包含三个卷积层、三个池化层和两个全连接层。与处理二维图像的常规CNN不同,本模型主要面向一维光谱信号进行分析,其卷积核沿波长维度进行特征提取,并相应调整了前向与反向传播的计算公式。

在模型输入阶段,将450~700 nm波段的荧光光谱数据以一维张量的形式输入至网络。卷积层位于输入层之后,由一组可学习参数的滤波器组成,又称为卷积核,每个卷积层的卷积核大小均为 5×1 ,一维卷积运算公式如式(1)所示:

$$y^j = f\left(\sum_i k^{ij} * x^i + b^j\right), \quad (1)$$

其中: $*$ 表示卷积运算, y^j 为第 j 个输出特征图, x^i 为第 i 个输入特征图, k^{ij} 为本层卷积运算所使用的卷积核, b^j 为第 j 个特征图的偏置。

卷积层后对应池化层,用于对特征图进行下采样实现数据降维,在减少计算参数的同时增强模型对局部特征的鲁棒性。本研究采用最大池化方法,设置采样窗口大小为2,通过保留局部区

域内特征响应的最大值作为输出,其计算表达式为:

$$F(X_k) = \max_{a \in X_k} \{a_1, \dots, a_s\}. \quad (2)$$

式(2)是最大池化对信号降采样计算方法,对卷积层运算所得的一个特征映射将其划分为多个不重叠区域 $X_k, k=1, 2, \dots, K$,区域采样大小为 $s \times 1$ 。

经过卷积层和池化层的运算,特征图被展平(Flatten),展平后的维度为9 792。随后,该展平特征输入到两个全连接层,输出维度分别为128和4(对应乳腺癌的4类亚型)。为避免过拟合,提高模型泛化能力,每层网络的线性操作后加入批标准化层(Batch Normalization, BN),同时引入ReLU函数作为激活函数,并在第一个全连接层后应用比例为0.5的随机失活(Dropout)。

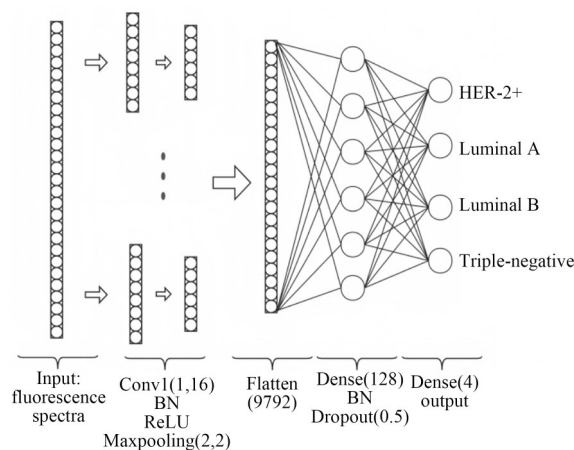


图4 分类模型结构图

Fig. 4 Classification model structure diagram

3 算法性能与分型效果分析

3.1 乳腺癌组织分类结果

将数据增强后所得到的乳腺癌荧光光谱数据集以 $1\,226 \times 1$ 的一维序列形式输入1D-CNN模型中,模型按照7:1.5:1.5划分训练集,测试集以及验证集。并引入准确率、精确率、召回率和F1分数四个度量标准来评估模型性能。表3展示了模型在各类亚型组织上的精确率、召回率和F1分数,从下表可以看出1D-CNN在四类亚型组织的判别中表现均衡,所有类别组织的评估指标均在96%以上,不存在明显短板。同时根据测试集

混淆矩阵的结果可得,模型仅存在有限误判,整体分类性能优异,在乳腺癌亚型鉴别任务中表现出色。

表 3 模型判别结果

Tab. 3 Model discrimination result

Type	Evaluation indicator		
	Recall	Precision	F1 score
HER-2+	98.3%	97.2%	97.8%
Luminal A	97.7%	96.1%	96.9%
Luminal B	97.3%	99.3%	98.7%
Triple-negative	100%	100%	100%

3.2 数据增强效果对比

表 4 为在分别采用数据增强和原始数据的情况下,模型对乳腺癌亚型组织的分类结果,通过加权平均计算总体分类准确率、召回率、精确率

表 4 数据增强对比实验结果

Tab. 4 Data-enhanced comparative experiment results

	Accuracy	Recall	Precision	F1-score
1D-CNN(Data-enhanced)	98.7%	97.2%	98.1%	97.7%
1D-CNN(Original data)	86.4%	64.1%	72.3%	68.3%

3.3 多种分类算法结果对比

使用 SVM,LDA,ResNet,Transformer 以及 LSTM 与本文 1D-CNN 模型进行对比,以上分类算法在乳腺癌光谱数据集上分别实验 10 次,并记录各自平均分类准确率。如图 6 所示,在六种分类模型中,1D-CNN 的表现最为出色,其分类准确率显著优于表现最差的 LSTM,高出 23.7%。

分析其原因在于:1D-CNN 主要得益于其结构特性与荧光光谱数据特性的高度契合,荧光光谱是一种连续且局部相关性强的一维信号,1D-CNN 通过卷积核可有效捕捉这些局部特征和谱形变化,在参数数量上也较为精简。相比之下,LDA 和 SVM 在处理高维非线性特征时能力有限,ResNet 和 Transformer 虽具强表征能力,但在分类任务中易过拟合,同时实验过程表明,Transformer 在处理高维光谱数据时计算开销较大,训练效率相对较低。LSTM 更擅长捕捉时间序列中的长期依赖关系,而荧光光谱信号的判别信息主要集中于局部特征。

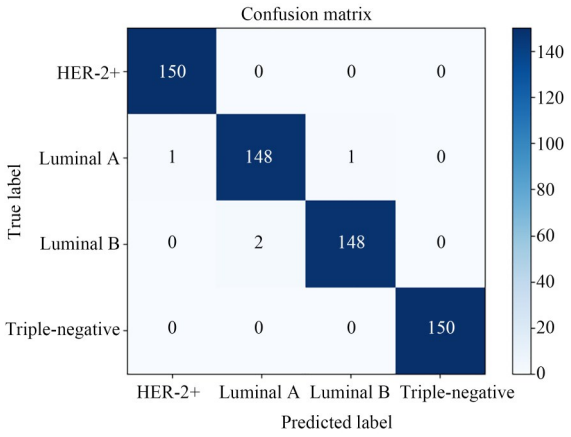


图 5 1D-CNN 模型用于乳腺癌亚型分类的混淆矩阵

Fig. 5 Confusion matrix of the 1D-CNN model for the classification of breast cancer subtypes

和 F1 分数作为评价标准。可以看出经过数据增强之后,模型的性能得到了一定程度的提高。

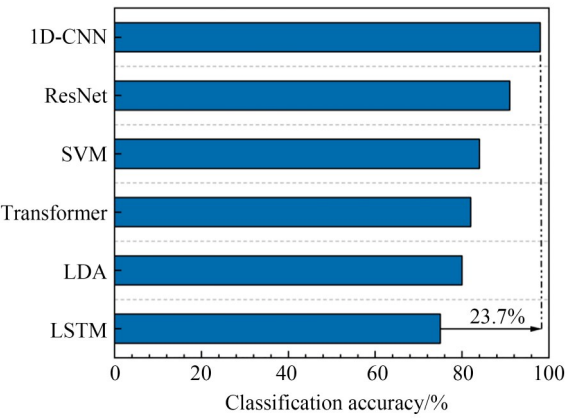


图 6 算法性能比较

Fig. 6 Comparison of algorithm performance

3.4 1D-CNN 模型决策可视化

在准确识别子类型的基础上,本研究为进一步深入了解 CNN 模型的内部机制,采用梯度加权类激活映射(Grad-CAM),重点探究卷积核是否有效捕获了与分型相关的特征性光谱波段。通过将测试集中 4 类亚型组织的平均光谱作为各

类别的代表性特征输入训练好的 CNN 模型,实验结果显示模型输出的预测标签与实际子类型一致,这一方面验证了所构建 CNN 模型具有良好的预测可靠性,另一方面也证实了通过 Grad-CAM 方法可视化的关键特征区域是可解释的。

不同亚型的 Grad-CAM 可视化区域如图 7 所示(图中标注区域为注意力强度最大值所对应的波长),对于 HER-2+ 型(图 7(a)),CNN 对 590~610 nm(中性卟啉)、620~630 nm(酸性卟啉)范围的注意力权重较高。卟啉是血红素合成途径的中间产物,癌变过程中,细胞中的胆色素脱氢基酶会产生大量卟啉物质,而亚铁螯合酶活力下

降,无法及时将卟啉转化为血红素,导致卟啉产物异常积累^[23]。而对于 Luminal A 型(图 7(d)),CNN 的注意力集中在 490~510 nm(游离 NADH)。NADH 是细胞代谢的关键辅酶,其荧光特性可用于评估肿瘤的氧化还原状态。Luminal B(图 7(b))和 Triple-negative(图 7(c))型的 Grad-CAM 可视化区域均集中在 580~590 nm(中性卟啉)。综上所述,CNN 预测亚型种类主要依赖于游离 NADH 以及中性和酸性卟啉的特征光谱,这与上文的研究相呼应,表明通过 CNN 模型反卷积得到的 Grad-CAM 可视化区域具有令人信服的生物学意义,进一步证明了利用荧光光谱和可视化策略挖掘生物标志物信息的可行性。

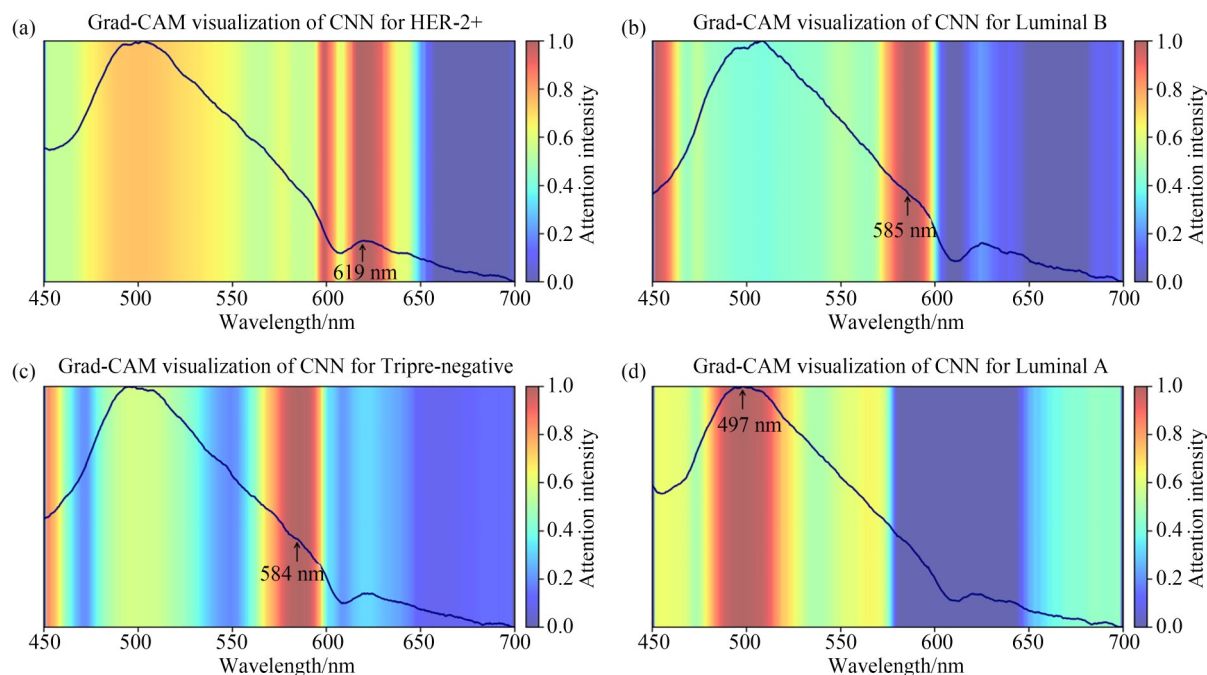


图 7 1D-CNN 模型决策可视化

Fig. 7 1D-CNN model decision visualization

4 结 论

本研究提出了一种融合荧光光谱技术与 1D-CNN 模型的乳腺癌分子分型方法。通过对不同亚型组织的荧光光谱进行分析,并结合代谢指标对 NADH, FAD 及卟啉等内源性荧光基团的分布特征进行量化评估,为分型提供生物学依据。所构建的 1D-CNN 模型结合多种数据增强策略,有效提升了小样本情况下的模型鲁棒性和泛化

能力,实现了对 4 类乳腺癌亚型组织的高精度识别,其分类准确率、召回率、精确率和 F1 分数分别达到 98.7%, 97.2%, 98.1% 和 97.7%。此外,通过 Grad-CAM 实现模型注意力可视化,进一步验证了网络识别特征与组织生物标志物的相关性,提升了模型的可信度和生物学解释力。

综上,本方法提供了一种高效、快捷的肿瘤检测手段,搭建的实验系统小巧、操作简单、分类算法性能优异,该方案具有良好的临床应用前

景。但我们深刻认识到,从方法提出到实际临床应用仍将面临诸多挑战。后续研究将围绕实验设备小型化、样本处理标准化以及数据验证多元化等方向展开。重点研究光谱采集模块与嵌入式硬件系统集成,开发一体式数据处理与智能分类检测系统;建立标准化样本采集与分析流程;构建不同来源和采集条件下的大规模肿瘤光谱数据库。进而加速成果转化,为乳腺癌的早期诊断和精准治疗提供可靠的技术支撑。

参考文献:

- [1] AHMAD A. *Breast Cancer Statistics: Recent Trends*[M]. Breast Cancer Metastasis and Drug Resistance: Challenges and Progress, 2nd Edition. 2019: 1-7.
- [2] GOLUB T R, SLONIM D K, TAMAYO P, *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring[J]. *Science*, 1999, 286(5439): 531-537.
- [3] GOLDBIRSH A, WOOD W C, COATES A S, *et al.* Strategies for subtypes: dealing with the diversity of breast cancer: highlights of the st Gallen international expert consensus on the primary therapy of early breast cancer 2011[J]. *Annals of Oncology*, 2011, 22(8): 1736-1747.
- [4] CABREDO S, PARRA A, ANZANO J. Bacteria spectra obtained by laser induced fluorescence[J]. *Journal of Fluorescence*, 2007, 17(2): 171-180.
- [5] JENKINS N C, EHRLICH K, KUFCSAK A, *et al.* Computational fluorescence suppression in shifted excitation Raman spectroscopy[J]. *IEEE Transactions on Bio-Medical Engineering*, 2023, 70(8): 2374-2383.
- [6] LEI X Y, PAN H G, HUANG X D. A dilated CNN model for image classification[J]. *IEEE Access*, 2019, 7: 124087-124095.
- [7] SHUJAAT M, WAHAB A, TAYARA H, *et al.* pcPromoter-CNN: a CNN-based prediction and classification of promoters[J]. *Genes*, 2020, 11(12): 1529.
- [8] 赵勇, 荣康, 谈爱玲. 基于一维卷积神经网络的雌激素粉末拉曼光谱定性分类[J]. *光谱学与光谱分析*, 2019, 39(12): 3755-3760.
- ZHAO Y, RONG K, TAN A L. Qualitative analysis method for Raman spectroscopy of estrogen

作者贡献声明:

黄翰林:实验方案设计与实施,数据分析及论文撰写;

官韬:实验数据获取及整理;

秦艳利:光谱数据采集指导;

沈龙海:项目总体管理;

齐东丽:实验方法指导;

李宏达:分析算法指导;

冯瑜:项目提出及论文审核。

based on one-dimensional convolutional neural network [J]. *Spectroscopy and Spectral Analysis*, 2019, 39(12): 3755-3760. (in Chinese)

- [9] 于水, 宦克为, 刘小溪, 等. 并联卷积神经网络的近红外光谱定量分析模型[J]. *光谱学与光谱分析*, 2024, 44(6): 1627-1635.

YU S, HUAN K W, LIU X X, *et al.* Quantitative analysis modeling of near infrared spectroscopy with parallel convolution neural network[J]. *Spectroscopy and Spectral Analysis*, 2024, 44(6): 1627-1635. (in Chinese)

- [10] SMITH S E, MELLOR P, WARD A K, *et al.* Molecular characterization of breast cancer cell lines through multiple omic approaches[J]. *Breast Cancer Research*, 2017, 19(1): 65.

- [11] WANG M Y, LONG F, TANG F, *et al.* Auto-fluorescence imaging and spectroscopy of human lung cancer [J]. *Applied Sciences*, 2017, 7(1): 32.

- [12] JING Y Y, WANG Y L, WANG X Y, *et al.* Label-free imaging and spectroscopy for early detection of cervical cancer[J]. *Journal of Biophotonics*, 2018, 11(5): e201700245.

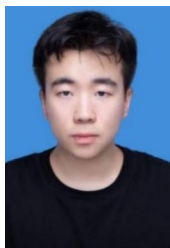
- [13] PU Y, TANG G C, WANG W B, *et al.* Native fluorescence spectroscopic evaluation of chemotherapeutic effects on malignant cells using nonnegative matrix factorization analysis [J]. *Technology in Cancer Research & Treatment*, 2011, 10(2): 113-120.

- [14] OSTRANDER J H, MCMAHON C M, LEM S, *et al.* Optical redox ratio differentiates breast cancer cell lines based on estrogen receptor status[J]. *Cancer Research*, 2010, 70(11): 4759-4766.

- [15] ZHUO S M, YAN J, CHEN G, *et al.* Label-free

- monitoring of colonic cancer progression using multiphoton microscopy [J]. *Biomedical Optics Express*, 2011, 2(3): 615-619.
- [16] WIKE-HOOLEY J L, HAVEMAN J, REINHOLD H S. The relevance of tumour pH to the treatment of malignant disease [J]. *Radiotherapy and Oncology*, 1984, 2(4): 343-366.
- [17] BONUCCELLI G, DE FRANCESCO E M, DE BOER R, *et al.* NADH autofluorescence, a new metabolic biomarker for cancer stem cells: Identification of Vitamin C and CAPE as natural products targeting “stemness” [J]. *Oncotarget*, 2017, 8(13): 20667-20678.
- [18] YU Q R, HEIKAL A A. Two-photon autofluorescence dynamics imaging reveals sensitivity of intracellular NADH concentration and conformation to cell physiology at the single-cell level [J]. *Journal of Photochemistry and Photobiology B: Biology*, 2009, 95(1): 46-57.
- [19] MASILAMANI V, AL-ZHRANI K, AL-SALHI M, *et al.* Cancer diagnosis by autofluorescence of blood components [J]. *Journal of Luminescence*, 2004, 109(3): 143-154.
- [20] SUBBASWAMY A, SARIA S. From development to deployment: dataset shift, causality, and shift-stable models in health AI [J]. *Biostatistics*, 2020, 21(2): 345-352.
- [21] KIM E, KIM J, LEE H, *et al.* Adaptive data augmentation to achieve noise robustness and overcome data deficiency for deep learning [J]. *Applied Sciences*, 2021, 11(12): 5586.
- [22] MA D Y, SHANG L W, TANG J L, *et al.* Classifying breast cancer tissue by Raman spectroscopy with one-dimensional convolutional neural network [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2021, 256: 119732.
- [23] BALSUKURI N, DAS S, GUPTA I. Carbazole-corrole and carbazole-prophyrin dyads: synthesis, fluorescence and electrochemical studies [J]. *New Journal of Chemistry*, 2015, 39(1): 482-491.

作者简介:



黄翰林(2001—),男,安徽淮北人,硕士研究生,主要从事激光光谱技术在肿瘤检测方面的研究。
E-mail:hhl200188@163.com

通讯作者:



冯 瑜(1977—),男,辽宁抚顺人,博士,副教授,硕士生导师,主要从事激光光谱技术及应用领域的研究。
E-mail:14025921@qq.com