

·基础研究·

## 酶消化法结合人工智能技术在法医学溺死硅藻检验中的应用

于慧潇<sup>1,2</sup>, 朱永正<sup>2,3</sup>, 赵天琦<sup>1</sup>, 程奇<sup>2,4</sup>, 李周儒<sup>1</sup>, 殷文江<sup>1</sup>, 陈敏<sup>2</sup>, 蔡红星<sup>1</sup>

(1. 徐州医科大学法医学教研室, 江苏 徐州 221000; 2. 司法鉴定科学研究院//上海市法医学重点实验室//司法部司法鉴定重点实验室//上海市司法鉴定专业技术服务平台, 上海 200063; 3. 西安市公安局经济技术开发区分局刑事侦查大队, 陕西 西安 710018; 4. 苏州市公安局, 江苏 苏州 215000)

**摘要:**【目的】人工智能(AI)全涂片自动化硅藻检验技术能比人类专家更快速、高效进行法医病理学溺死硅藻检验。然而,该技术仅与硅藻提取率较低的传统消化法联用,本研究拟采用更加高效的蛋白酶K组织消解法(以下简称酶消化法)作为硅藻提取方法,探究该技术在其他硅藻提取方法中的泛化能力及可行性。【方法】收集6例溺死尸体的肺组织进行蛋白酶K消解并制成涂片,利用数字化图像矩阵切割方法将涂片进行数字化处理并据此建立硅藻-背景数据库,将数据集按照3:1:1的比例分为训练集、验证集和测试集,在ImageNet预训练基础上对卷积神经网络(CNN)模型进行训练和内部验证及外部测试。【结果】结果显示最佳模型外部测试的准确率达97.65%,且模型特征提取区域即为硅藻所在区域。实际应用中最佳的CNN模型对溺水尸体的硅藻检测精准率高达80%以上。【结论】研究表明,基于CNN模型的AI自动化硅藻检验技术和酶消化法联用能高效识别硅藻,可以作为溺水鉴定中硅藻检测的辅助方法。

**关键词:**法医病理学;硅藻;溺死;蛋白酶K;人工智能;卷积神经网络

**中图分类号:**R89 **文献标志码:**A **文章编号:**1672-3554(2023)03-0430-09

**DOI:**10.13471/j.cnki.j.sun.yat-sen.univ(med.sci).2023.0309

## Enzymatic Digestion Method Coupled with Artificial Intelligence Techniques in Forensic Drowning Diatom Detection

YU Hui-xiao<sup>1,2</sup>, ZHU Yong-zheng<sup>2,3</sup>, ZHAO Tian-qi<sup>1</sup>, CHENG Qi<sup>2,4</sup>, LI Zhou-ru<sup>1</sup>, YIN Wen-jiang<sup>1</sup>,  
CHEN Min<sup>2</sup>, CAI Hong-xing<sup>1</sup>

(1. Department of Forensic Medicine, Xuzhou Medical University, Xuzhou 221000, China; 2. Shanghai Key Laboratory of Forensic Medicine//Key Laboratory of Forensic Science, Ministry of Justice//Shanghai Forensic Service Platform//Academy of Forensic Science, Shanghai 200063, China; 3. Criminal Investigation Brigade, Economic and Technological Development Zone Subbureau, Xi'an Public Security Bureau, Xi'an, 710018, China; 4. Suzhou Public Security Bureau, Suzhou 21500, China)

Correspondence to: CAI Hong-xing, E-mail: xzmccaih@163.com; CHEN Min, E-mail: cminflames@163.com

**Abstract:**【Objective】Artificial intelligence (AI) full smear automated diatom detection technology can perform forensic pathology drowning diatom detection more quickly and efficiently than human experts. However, this technique was only used in conjunction with the strong acid digestion method, which has a low extraction rate of diatoms. In this study,

收稿日期:2022-10-25

基金项目:上海市法医学重点实验室,司法部司法鉴定重点实验室开放基金(KF202004);江苏省普通高校研究生科研、实践创新计划项目(KYCX22\_2876)

作者简介:于慧潇,硕士生,研究方向:法医病理学及人工智能算法, E-mail: yuhuixiao19552153521@163.com;蔡红星,第一通信作者,硕士,教授,研究方向:法医学基础及应用, E-mail: xzmccaih@163.com;陈敏,第二通信作者,主检法医师,研究方向:法医病理学, E-mail: cminflames@163.com.

we propose to use the more efficient proteinase K tissue digestion method (hereinafter referred to as enzyme digestion method) as a diatom extraction method to investigate the generalization ability and feasibility of this technique in other diatom extraction methods.【Methods】 Lung tissues from 6 drowned cadavers were collected for proteinase K ablation and made into smears, and the smears were digitized using the digital image matrix cutting method and a diatom and background database was established accordingly. The data set was divided into training set, validation set and test set in the ratio of 3:1:1, and the convolutional neural network (CNN) models were trained, internally validated, and externally tested on the basis of ImageNet pre-training.【Results】 The results showed that the accuracy rate of the external test of the best model was 97.65%, and the area where the model features were extracted was the area where the diatoms were located. The best CNN model in practice had a precision of more than 80% for diatom detection of drowned corpses.【Conclusion】 It is shown that the AI automated diatom detection technique based on CNN model and enzymatic digestion method in combination can efficiently identify diatoms and can be used as an auxiliary method for diatom detection in drowning identification.

**Key words:** forensic pathology; diatoms; drowning; proteinase K; artificial intelligence; convolutional neural network  
[J SUN Yat-sen Univ (Med Sci), 2023, 44(3): 430-438]

溺死是法医病理学常见死亡原因之一,据世界卫生组织统计,溺水是世界各地意外伤害死亡的第三大原因,其死亡率约占7%,并且世界各地每年约有23.6万人死于溺水<sup>[1]</sup>。在法医学实际检案中,当遇到溺死时间较长或高度腐败的尸体时,由于其溺死体征已不复存在,法医往往难以鉴定。因此,溺死诊断一直是法医鉴定的难点之一。作为水中常见浮游生物,硅藻在溺水过程中会随溺液进入至人体呼吸道,经肺泡壁-毛细血管屏障侵入血液循环,进而播散至肝、肾等实质器官<sup>[2]</sup>。由于硅藻的无机细胞骨架具有较强的耐酸性,法医工作者常常使用生物、化学等手段去除组织样本中的有机成分,最后在显微镜下确认硅藻的含量和种属从而为溺死诊断提供支持性证据,目前已经成为腐败尸体鉴定的最佳方法<sup>[3]</sup>。强酸消化法因其试剂成本低廉、操作步骤简便及组织消解充分等优点被广泛应用于法医学实践,但仍有部分学者出于该方法较低检出率的原因更加倾向于消解更为温和、硅藻破坏更小的酶消化法。Ming<sup>[4]</sup>将硝酸消化法、蛋白酶K消化法等几种常见硅藻消化方法的优缺点进行了详细比较,结果表明在实验可靠性和适用性方面酶消化法都是首选。然而,无论使用何种组织消解方法,传统形态学硅藻检验仍依靠人工肉眼在显微镜下逐个视野对硅藻进行识别、分类,该工作费时、费力且需要相关专家进行辨别和确认。近年来,人工智能(artificial intelligence, AI)深度学习算法在计算机视觉领域的应用愈发常见。目前,已有学者将深度学习技

术运用到环境科学领域中,主要从事水域硅藻种属分类、识别计数等工作<sup>[5-8]</sup>,所采用的算法多基于区域卷积人工神经网络的目标检测算法。然而,这类算法存在以下不足:①多候选区域图像提取占用磁盘空间较大;②图像提取、处理过程中会造成图片变形,从而导致重要信息丢失;③重叠区域运算造成了大量计算浪费,使得目标检测时间延长。因此,这类算法多用于单张高清显微或电镜小区域图片的目标识别,在技术手段、操作步骤、识别精度等方面难以满足法医学硅藻检验中需要全涂片或切片识别的实际需求。鉴于此,司法鉴定科学研究院法医病理学科研团队基于数字化图像矩阵切割方法将传统卷积人工神经网络(convolutional neural network, CNN)算法引入法医学溺死硅藻检验中。该团队涉及的深度学习模型<sup>[9-12]</sup>实现了快速、高效的全涂片硅藻识别,并在随后的硅藻属分类中获得了95.24%的外部验证平均准确率。与绝大多数研究<sup>[13-15]</sup>相一致的是,该模型结合相关向量比对算法进一步证实了根据现场水样和人体组织中硅藻种属及丰度的一致性分析来推断溺水地点的可行性。上述基于AI的全涂片自动化硅藻检验技术在硅藻识别及种属分类等方面具有较大的应用前景,但该团队所采用的样本处理方法仅局限于硝酸消化法。因此,本研究采用硅藻提取效率更高的酶消化法与AI全涂片自动化硅藻检验技术进行联用,进一步证实这种AI技术在其他组织消解方法中的泛化能力及可行性。

## 1 材料与方 法

### 1.1 仪器与试剂

电热恒温培养箱(型号:DHP-9082B,上海一恒科学仪器有限公司);磁力搅拌器(赛默飞世尔科技(中国)有限公司);移液器(德国艾本德公司);烘片机(型号:HI1220,美国Leica公司);自动化封片机(型号:CV5030,美国Leica公司);离心机(型号:5702,德国艾本德公司);高通量快速扫描仪(型号:APERIO AT2,美国Leica公司);Tris-HCL缓冲液(北京索莱宝科技有限公司);蛋白酶K溶液(德国Merck公司);SDS溶液(德国Merck公司)。

### 1.2 样品制备及硅藻提取

样品制备:在司法鉴定科学研究院搜集了6例溺死尸体的左肺上叶组织。取材时使用的器械经去离子水清洗以防止器官组织间交叉污染。提取的组织样本称重后放入-20℃冰箱内冷冻保存。为了增加硅藻种属丰富度在苏州河用取水器进行水样采集,将水样进行离心浓缩并避光密闭保存。

硅藻提取:如图1所示,将去掉表面浆膜的组织样本切成若干小块,称取10g放入烧杯内,用搅拌机制成匀浆。加入2mL水样浓缩液、100mL Tris-HCL(0.01 mol/L)缓冲液(含20g/L的SDS溶液)和500μL蛋白酶K溶液(10g/L),并充分混合;用锡纸密封杯子,在50℃电热恒温培养箱中孵育过夜。第2天取出后,向烧杯中加入2g的SDS粉末,再用磁力搅拌器加热搅拌30min;将液体转移至另一个清洁烧杯内并加入分析纯硝酸10mL,继续加热搅拌30min,直至液体澄清透亮。取10mL消化液放入离心机中以30000×g离心5min,弃上清,加入超纯水至10mL,再次离心,弃上清。用移液器将离心后的沉淀反复吹打,吸取50μL进行涂片,涂片时尽量均匀。用烘片机烘干涂片并用自动化封片机封片,随后放入通风橱中晾干。将涂好的玻片在烘片机中烘干,用自动化封片机封片,然后放在通风橱中干燥。在样品制备和硅藻提取之前,确认所使用的去离子水、烧杯、搅拌器等设备以及取材工作接触面等不含硅藻,并通过设置空白对照试验来避免硅藻污染的存在。本研究经司法鉴定科学研究院伦理委员会审核批准,涉及的人体标本均已获得死者家属的知情同意。

### 1.3 建立硅藻数据库

使用莱卡扫描仪将制作的16张肺组织涂片进行扫描并转换为数字化格式,扫描倍数为400。数字化的图像采用矩阵切割方法被切割成255×255像素的小图像,并由经验丰富的法医专家进行注释、分类。基于酶消化法的涂片共标记出5598张轮廓、纹理清晰的“硅藻”图片和5000张“背景”图片,为方便数据集划分,保留“硅藻”和“背景”图片各5000张,形成一个包含5000张硅藻图像和5000张背景图像的硅藻-背景数据库。考虑到整张涂片中硅藻含量较少,而背景图像远多于硅藻图像,属于不平衡类分布,简单使用平衡训练不利于提高模型性能。因此本研究在数据采集上进行处理以获得数量相同的硅藻和背景图像,对硅藻进行随机过采样,对背景进行随机欠采样。最后随机选取60%的数据集作为模型参数优化的训练集,20%的数据集作为确定最佳CNN模型的验证集,其余20%的数据集作为模型性能独立评估的测试集。硅藻数据库包含一些常见的淡水藻类,如小环藻属、直链藻属等。

### 1.4 模型训练及验证

在本研究中,我们使用了GoogLeNet Inception V3<sup>[16]</sup>模型,重点研究了不同硅藻提取方法的可行性,排除了模型的干扰。为了更好地体现酶消化法与基于CNN模型的全涂片硅藻自动检测技术联用的可行性,本研究在模型训练方法上尽量与司法鉴定科学研究院法医病理学科研团队的Zhou等一致<sup>[9]</sup>,只使用ImageNet进行迁移学习。由于ImageNet数据库中含有常规的硅藻图片,因此将其作为预训练的模型大大减少了模型的训练时间。并且,我们最近的研究证实,与VGG16、ResNet50和Inception-ResNet-V2模型相比,InceptionV3模型在硅藻特征提取方面具有更强的指向性和靶向性<sup>[17]</sup>。该模型主要由卷积层、池化层和全连接层组成,其中卷积层主要对输入图像进行局部特征提取,池化层对卷积层输出的特征向量进行压缩并提取主要特征,全连接层将所有特征连接起来并将输出值送至softmax分类器,其模式图如图2所示。为了加速模型的收敛和提高模型的泛化能力,使用迁移学习的方法,将海量数据预训练的模型迁移到硅藻训练过程中,预训练数据集来自ImageNet数据库(<https://image-net.org/index.php>)<sup>[18]</sup>。使用Adadelta优化器<sup>[19]</sup>将训练集和验证集的数据输入到深度学习模型架构中进行迭代训练,学习率为0.01,单次

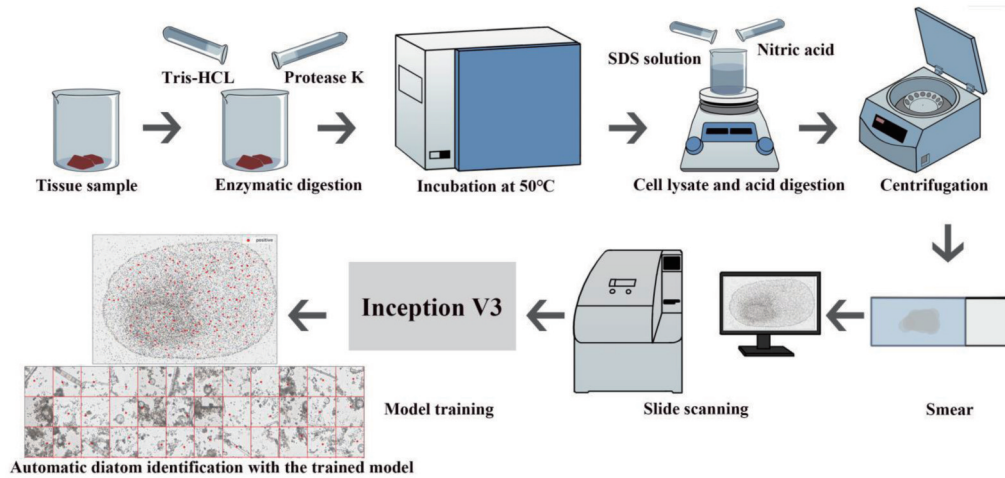


图1 实验流程图

Fig. 1 Experimental flowchart

训练样本量为64,迭代200次。模型训练时采用分批次反复学习的方法,对训练集数据进行分批训练,并将前一批训练的中间模型作为其下一批训练的输入,然后经监督学习及时调整模型的参数。同时,利用训练好的中间模型对训练数据进行预测,对模型预测错误的硅藻和背景数据进行人工修正后加入下一批次的训练集数据中。这种方法可以提高模型对硅藻和背景图像的鉴别能力。在训练

过程中实时记录每次迭代时模型对训练集和验证集的准确率和损失函数值的预测结果,当验证集的准确率不再随训练集的准确率增加,或者验证集的损失函数值不随迭代次数的增加而减少时,提前终止训练以防止模型的过度拟合。参考模型训练曲线图,将模型过拟合前准确度高、损失函数低的模型作为本次训练的最佳模型。

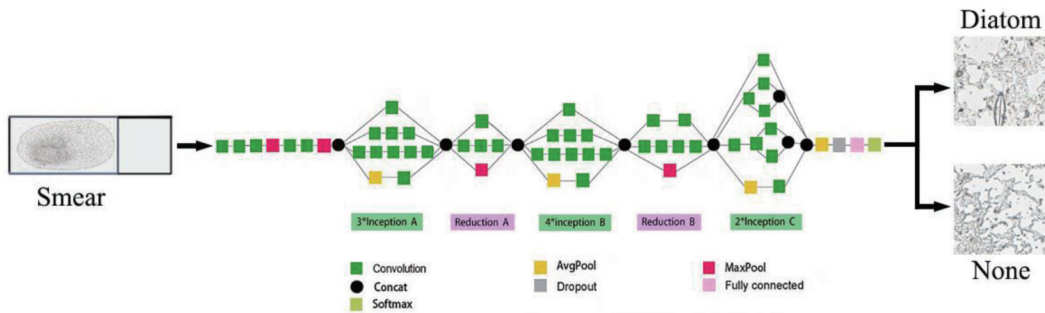


图2 CNN模型分类模式图

Fig. 2 CNN model classification pattern diagram

模型训练、验证在Ubuntu 16.04操作系统下完成,工作站硬件配置如下:CPU(Intel Xeon 6132)×2, GPU 显卡(Nvidia Geforce RTX2080ti)×2, 128 G 内存。深度学习架构采用Python 3.6.5 环境下运行,编写模块包括TensorFlow、NumPy、OpenCV 和sklearn。

### 1.5 模型测试与评估

利用训练的最佳模型对测试集的图片进行测试,根据模型测试结果绘制接受者操作特征(re-

ceiver operator characteristic, ROC)曲线并计算曲线下面积(area under the curve, AUC)值,ROC 概率阈值为0.5。对于模型二分类的预测结果,测试集的真实标签类别和通过模型预测的类别有TP(true positive)、FP(false positive)、TN(true negative)、FN(false negative)四种情形,通过统计模型识别硅藻和背景的图片数量计算TP、FP、TN、FN值,并绘制出混淆矩阵图。其中,TP代表真阳性数量,即模型正确识别硅藻图片的数量;TN代表真阴性数量,即

模型正确识别背景图片的数量;FP代表假阳性数量,即模型错误识别硅藻图片的数量;FN代表假阴性数量,即模型错误识别背景图片的数量。选择灵敏度(sensitivity)、特异度(specificity, S)及准确率(accuracy rate, AR)三个性能指标对模型的效能、泛化能力进行评估。灵敏度即召回率(re-call rate, RR),代表的是模型识别正确的“硅藻”数量占总硅藻数量的比例,反映了硅藻识别的检出率。特异度(S)代表的是模型识别正确的“背景”数量占总背景数量的比例,反映模型对背景样本的鉴别能力。准确率(AR)即所有正确识别的“硅藻”和“背景”占总样本的百分比,反映模型识别硅藻和背景的总体正确率。三个指标算式如下:

灵敏度

$$RR = \frac{TP}{TP + FN} \quad (1)$$

特异度

$$S = \frac{TN}{TN + FP} \quad (2)$$

准确率

$$AR = \frac{TP + TN}{TP + FN + TN + FP} \quad (3)$$

### 1.6 模型识别硅藻的可视化分析

虽然模型的识别性能已经用测试集数据和相关评价指标进行了验证,但深度学习模型特征提取和分类的内部运行机制还没有得到充分解释。目前已有一些方法被提出用于可视化内部运行机制<sup>[20-22]</sup>。通过算法将 CNN 模型预测情况可视化,即可使模型任意目标概念的梯度流入最后的卷积层并产生一个粗略的定位图,以此来突出图像中重要区域,即输出预测的依据<sup>[23]</sup>。换句话说,这使我们在识别硅藻时将模型的重点区域可视化。本研究采用两种 CNN 可视化算法,即引导反向传播(guided backpropagation, GBP)和梯度加权类激活映射(gradient-weighted class activation mapping, Grad-CAM)来可视化分析硅藻识别效果,并进一步确认模型识别区域就是硅藻所在的区域。

### 1.7 模型在溺死案例中的应用

搜集2例溺死案例,使用酶消化法将尸体中的肺组织进行消解并涂片,随后利用徕卡扫描仪将涂片扫描为数字化格式,最后使用训练的最佳模型对涂片中的硅藻进行自动化检测。通过法医学专家对最佳模型的硅藻识别结果进行人工校验,探究基于 CNN 模型的 AI 全涂片自动化硅藻检验技术和酶消化法联用在实际溺死案例中的应用效果。

## 1.8 人工智能深度学习的数据统计方法

深度学习的本质是利用数据来训练模型,然后利用训练好的模型进行分析和预测。本研究的数据统计从硅藻、背景图片的数据处理开始(即图片的采集、标注和分类,建立硅藻数据库),首先通过监督学习对 Inception V3 模型进行建模(即模型的训练和验证,获得最优模型)、其次检验模型识别效果(即模型测试和评估),最后通过可视化算法对模型识别进行可解释性分析。

## 2 结果

### 2.1 最佳模型及相关指标

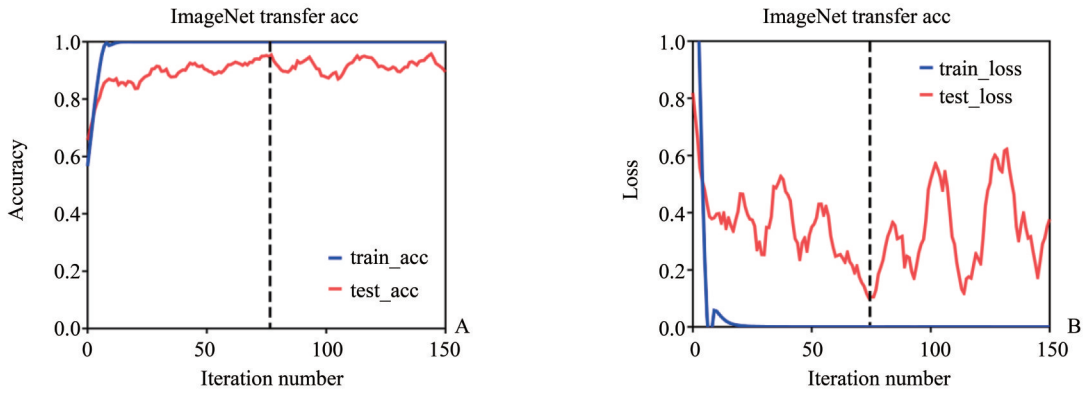
图3展示了基于酶消化法得到最佳模型时的训练曲线,可以看出,图中的黑色虚线位置代表训练结束的最佳时期,此时模型的准确度最高,损失函数值最低,迭代次数为73,即本次训练的最佳模型。而 Zhou 采用硝酸消化法作为硅藻提取方法,他们构建的数据库(训练集包含6585幅图像,验证集包含1545幅图像)训练的最佳模型与本研究获得的最佳模型的性能指标比较见表1。本研究得到的最佳模型在验证集中有更高的准确率(99.99%)和更低的损失函数值(0.0231),这表明本次训练得到的模型很可能具有更为优异的识别性能。

### 2.2 模型性能评估

图4展示了模型根据测试集(包括硅藻和背景图像各1000张)的预测结果绘制的 ROC 曲线图,其 AUC=0.9929(95%CI:0.9888~0.9971),说明模型具有良好的识别性能。图5是由模型预测结果的 TP、FP、TN 和 FN 值绘制的混淆矩阵,可以看出该模型在硅藻和背景识别中具有较高的均衡性和较高的识别正确率。其识别的灵敏度为97.8%、特异性为97.5%、准确率为97.65%,这些性能指标表明,该模型具有很高的硅藻检出率并且能够很好地识别涂片背景。这一结果与酶消化法的优良消化效果是分不开的,较清晰的硅藻图像为模型训练提供了良好的数据基础。

### 2.3 模型可视化分析

本研究中应用的 GBP 和 Grad-CAM 两种 CNN 可视化算法生成的硅藻热图如图6所示。可以看出,该模型能够很好地定位硅藻,几乎所有模型的兴趣区域都集中在硅藻的位置上,甚至不完整的硅



train\_acc: accuracy of the training set; test\_acc: accuracy of the validation set; train\_loss: loss function of the training set; test\_loss: loss function of the validation set; The blue line indicates the training set and the red line indicates the validation set.

图3 经二阶平滑处理的准确率、损失函数训练曲线

Fig. 3 The accuracy and loss curves after the second-order smoothing

表1 基于两种硅藻提取方法的最佳模型性能指标

Table 1 Performance indicators of the best model based on two diatom extraction methods

Diatom extraction methods	Training set		Validation set	
	Accuracy rate/%	Loss function	Accuracy rate/%	Loss function
Nitric acid digestion method	100	0.000 3	97.61	0.178 0
Enzyme digestion method	100	0.000 3	99.99	0.023 1

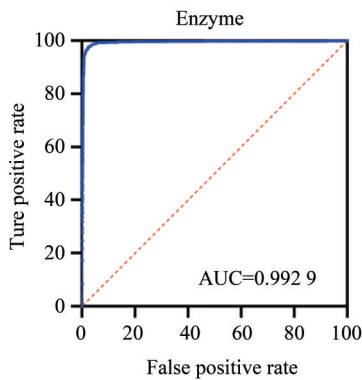


图4 ROC曲线图

Fig. 4 ROC curve diagram

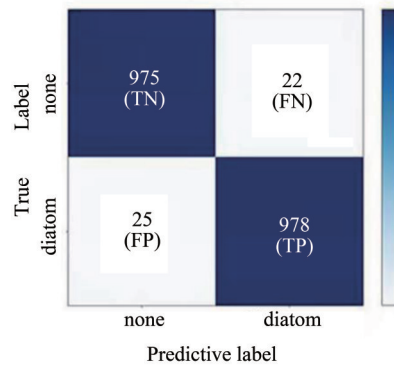


图5 混淆矩阵图

Fig. 5 Confusion matrix diagram

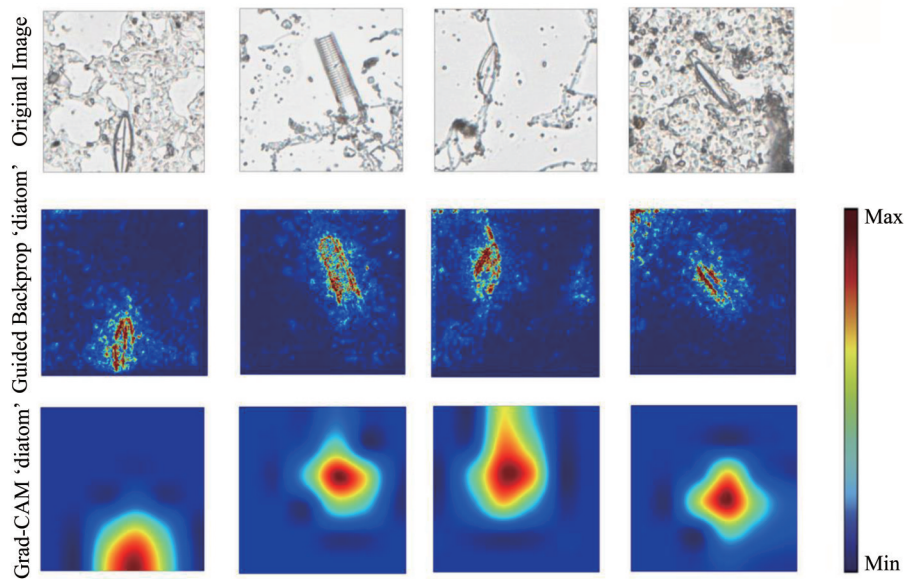
藻碎片也能被准确识别。此外,GBP硅藻热图具有更高的分辨率和更细的颗粒度,捕捉到了模型识别硅藻的形态和纹理特征,这与Grad-CAM硅藻热图相印证,证实了本研究中构建的模型的可靠性和模型识别的硅藻图像的准确性。

2.4 模型在实际案件中的应用效果

模型在实际溺水案件的肺组织涂片中均检测出较多硅藻,三位法医学专家对两个案例涂片中的硅藻进行了观察和计数,并计算了模型识别硅藻的精确率。模型预测的两个案件涂片中的硅藻数量

分别为39和104,而实际的硅藻数量分别为33和91。两个案例涂片中模型预测为硅藻的数量分别为39、104,而实际为硅藻的数量分别为33、91。如图7所示,红色散点代表模型正确预测的硅藻,绿色散点代表模型错误预测的硅藻。经计算,模型对两个案例硅藻预测的精确率分别为84.62%(图7左)和87.5%(图7右)。

我们的研究结果证明了基于酶消化法构建的CNN模型在溺死尸体鉴定中进行硅藻自动化识别的可行性。考虑到涂片样本保存的困难性以及人



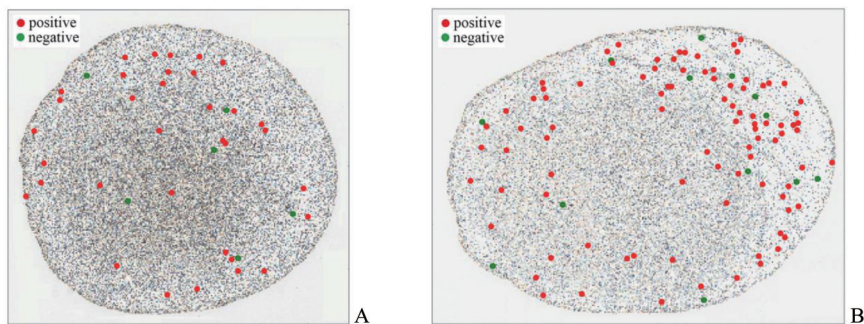
The red areas represent areas that the computer considers to be diatoms with a high probability; The blue areas are the ones that the computer thinks are low probability diatoms.

图6 硅藻可视化算法图

Fig. 6 Diatom visualization algorithm for diatoms map

工硅藻检查的耗时性,本研究中电子图像的应用更易于保存及日后复检,并且CNN模型在硅藻识别

效率和识别准确度上都更有优势,可以作为快速客观硅藻检验的替代方法。



Red dots: diatom identified by model and verified by manual; Green dots: diatom identified by model and non-diatom verified by manual, that red dots are true positives and green dots are false positives. A: red dot 33/ green dot 6; B: red 91/ green 13; magnification:0.8×10.

图7 模型识别效果图

Fig. 7 Diatom visualization with AI model

### 3 讨论

溺死在世界范围内均有较高的发生率。确定水中尸体是生前溺死还是死后入水一直是法医学死因诊断的难点。余海星等<sup>[24]</sup>采用文献计量学(bibliometrics)<sup>[25]</sup>方法统计了1991-2020年法医溺水领域的文献,包括英文文献630余篇,中文文献380余篇,发现英文文献更倾向于新技术、新方法

的研究,而硅藻检验是当前研究热点。多年来,学者们在硅藻提取上不断创新,并将人工智能深度学习算法引入硅藻检验,实现了硅藻的自动化识别。司法鉴定科学研究院法医病理学科团队<sup>[9]</sup>在法医学领域中首次开展了自动化硅藻检验研究,所采用的方法基于全切片图像扫描(whole slide image, WSI)技术,利用Inception V3模型,即图像分类算法,自动识别光学显微镜下的硅藻图像,并在几分

钟内完成了对整个硅藻涂片的检测。在实际案例应用中,总体召回率达到了92.45%<sup>[12]</sup>。随后,Yu<sup>[26]</sup>利用RetinaNet模型构建了一个基于电镜图像的硅藻自动识别模型,在0.5的阈值下,平均精确度达到0.82,平均召回率达到0.88。陈吉等<sup>[27]</sup>利用构建的YOLOv3模型对电镜图像中的硅藻自动检测,在不同的阈值设置下取得了0.94的平均精确度和0.81的平均召回率。就现有文献报道,基于光学显微镜的自动化硅藻检验方法其准确率相对较高、操作流程也更为简单,且在检验时间、仪器成本等方面具有较高的优势,便于在经费相对有限的基层法医中予以推广。

常规法医学硅藻检验采用硝酸消解法提取硅藻,而硝酸具有较强的腐蚀性,实验操作具有一定的危险性,实验过程中会产生有毒气体,损害实验人员的健康。实验过程中,发现硝酸涂片在观察及扫描时容易腐蚀显微镜的镜头。相比之下,酶解法更温和、安全,不产生有毒气体,能更好地保持硅藻结构和表面纹理的完整性,不仅能提高硅藻的提取率,还能为模型训练提供高质量的硅藻图像,便于模型特征提取,从而获得性能更好的识别模型。最初,我们担心由于酶消化法的涂片背景较为透明,AI识别会有困难,然而,研究结果表明,酶消化法获得的涂片图像特征可以被AI系统很好地捕捉和识别。我们的研究证实了酶消化法与AI自动化硅藻检验技术联用的可行性。

在AI系统建立过程中,通过对硅藻随机过采样、背景随机欠量采样的重采样方法建立数据库,解决了实际案件涂片中硅藻和背景图像分类不平衡的问题,从而提高了模型训练的精度。通过分批次反复学习,重新输入模型预测错误的结果进行反复训练,增强了模型的鲁棒性。最终得到的最佳模型在对测试集的预测中具有较好的表现,准确率高达97.65%、AUC值达0.9929。并且与Zhou等<sup>[9]</sup>基

于硝酸消化法训练的最优模型相比,其识别性能可能更为优异。我们推测,这一结果可能是由于在基于硝酸消化构建的模型中,其硅藻识别的性能可能受到硅藻图像轮廓的完整性和表面纹理的清晰度的影响。而本研究最佳模型综合识别能力的性能指标(灵敏度97.8%,特异性97.5%,准确率97.65%)都在97%以上。其中97.8%的灵敏度说明了模型能较好地识别硅藻,只有极少量的硅藻被漏检;97.5%的特异度说明该模型对背景具有较高的鉴别能力;97.65%的准确率更是体现了该模型在识别硅藻和背景方面的综合能力十分优异。同时,在实际案件的应用中,该模型表现出较高的硅藻预测精确率,能够满足日常检案需要。此外,酶消化法在检材量少等具有争议的案件中更有优势,能有效降低漏检的风险。在实际应用方面,司法鉴定科学研究院的法医病理学科团队基于硝酸消化法开发的人工智能自动化硅藻检验软件Diatom-Net V1.0也适用于本研究,可以将模型对接软件进行自动化硅藻检验的全国性推广,在实际应用中具有较大的价值。

我们的研究为未来的智能硅藻检验拓宽了道路,因为它将新的硅藻提取方法纳入AI全涂片自动化硅藻检验技术应用范畴中。人工智能技术的应用将原本需要2~3h的硅藻检验工作缩短到几分钟,大大提高了法医工作者的办案效率。酶消化法在自动化硅藻检验上的应用,进一步提高了硅藻检验的精准度,为生前溺死鉴定提供了更为精准的方法。本研究的实验样本多来自于江苏水域,未来计划多地区多水域取水对数据库中硅藻的种属及数量进行扩增,进一步扩展模型的应用范围。在本地区实际案例中的应用证实了该研究在实际应用中的高度可行性,为溺水诊断提供了更加准确和高效的技术手段。

#### 参考文献

- [1] World health organization global report on drowning: preventing a leading killer[J]. WHO, 2014.
- [2] Lunetta P, Penttila A, Hallfors G. Scanning and transmission electron microscopical evidence of the capacity of diatoms to penetrate the alveolo-capillary barrier in drowning [J]. Int J Legal Med, 1998, 111 (5) : 229-237.
- [3] Piette MH, De-letter EA. Drowning: still a difficult autopsy diagnosis[J]. Forensic Sci Int, 2006, 163(1/2): 1-9.
- [4] Ming M, Meng X, Wang E. Evaluation of four digestive methods for extracting diatoms [J]. Forensic Sci Int, 2007, 170(1): 29-34.
- [5] Pedraza A, Bueno G, Deniz O, et al. Lights and pit-

- falls of convolutional neural networks for diatom identification[C]. SPIE, 2018, 10679: 88–96.
- [6] Pedraza A, Bueno G, Deniz O, et al. Automated diatom classification (part B): a deep learning approach [J]. Appl Sci, 2017, 7(5): 460.
- [7] Salido J, Sánchez C, Ruiz-santaquiteria J, et al. A low-cost automated digital microscopy platform for automatic identification of diatoms [J]. Appl Sci, 2020, 10(17): 6033.
- [8] Krause LMK, Koc J, Rosenhahn B, et al. Fully convolutional neural network for detection and counting of diatoms on coatings after short-term field exposure [J]. Environ Sci Technol, 2020, 54(16): 10022–10030.
- [9] Zhou Y, Zhang J, Huang J, et al. Digital whole-slide image analysis for automated diatom test in forensic cases of drowning using a convolutional neural network algorithm[J]. Forensic Sci Int, 2019, 302: 109922.
- [10] Zhang J, Zhou Y, Vieira DN, et al. An efficient method for building a database of diatom populations for drowning site inference using a deep learning algorithm[J]. Int J Legal Med, 2021, 135(3): 817–827.
- [11] Zhou Y, Cao Y, Huang J, et al. Research advances in forensic diatom testing [J]. Forensic Sci Res, 2020, 5(2): 98–105.
- [12] 周圆圆, 曹永杰, 杨越, 等. 基于人工智能硅藻自动化识别系统的实际案例应用[J]. 法医学杂志, 2020, 36(2): 239–242.
- Zhou YY, Cao YJ, Yang Y, et al. Application of artificial intelligence automatic diatom identification system in practical cases[J]. J Forensic Med, 2020, 36(2): 239–242.
- [13] Ludes B, Coste M, North N, et al. Diatom analysis in victim's tissues as an indicator of the site of drowning [J]. Int J Legal Med, 1999, 112(3): 163–166.
- [14] Coelho S, Ramos P, Ribeiro C, et al. Contribution to the determination of the place of death by drowning—A study of diatoms' biodiversity in douro river estuary[J]. J Forensic Leg Med, 2016, 41: 58–64.
- [15] Carballeira R, Vieira DN, Febrero-bande M, et al. A valid method to determine the site of drowning[J]. Int J Legal Med, 2018, 132(2): 487–497.
- [16] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]. AAAI, 2017: 4278–4284.
- [17] 朱永正, 张吉, 程奇, 等. 4种深度学习图像分类算法在人工智能硅藻检验中的比较[J]. 法医学杂志, 2022, 38(1): 31–39.
- Zhu YZ, Zhang J, Cheng Q, et al. Comparison among four deep learning image classification algorithms in AI-based diatom test [J]. J Forensic Med, 2022, 38(1): 31–39.
- [18] Deng J, Dong W, Socher R, et al. ImageNet: a large-scale hierarchical image database [C]. IEEE, 2009: 248–255. <http://dx.doi.org/10.1109/CVPR.2009.5206848>
- [19] Zeiler MD. Adadelta: an adaptive learning rate method [C]. ArXiv, 2012. <https://arxiv.org/abs/1212.5701>
- [20] Simonyan K, Vedaldi A, Zisserman A. Deep inside convolutional networks: visualising image classification models and saliency maps [J]. CoRR, 2014. <https://arxiv.org/abs/1312.6034>
- [21] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks [M]. Springer, 2014: 818–833. [https://doi.org/10.1007/978-3-319-10590-1\\_53](https://doi.org/10.1007/978-3-319-10590-1_53)
- [22] Bach S, Binder A, Montavon G, et al. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation [J]. PloS One, 2015, 10(7): e0130140.
- [23] Selvaraju RR, Cogswell M, Das A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization [C]. IEEE, 2017: 618–626.
- [24] 余海星, 张吉, 朱永正, 等. 1991–2020年法医学溺死领域中英文文献计量学分析[J]. 法医学杂志, 2022, 38(1): 98–109.
- Yu HS, Zhang J, Zhu YC, et al. Bibliometrics analysis in English and Chinese literature on drowning in forensic medicine from 1991 to 2020 [J]. J Forensic Med, 2022, 38(1): 98–109.
- [25] Jacomy M, Venturini T, Heymann S, et al. Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software [J]. PloS one, 2014, 9(6): e98679.
- [26] Yu W, Xue Y, Knoops R, et al. Automated diatom searching in the digital scanning electron microscopy images of drowning cases using the deep neural networks[J]. Int J Legal Med, 2021, 135(2): 497–508.
- [27] 陈吉, 刘晓荣, 杨佳雯, 等. 基于YOLOv3的硅藻电子显微镜图像识别模型的构建及应用[J]. 法医学杂志, 2022, 38(1): 46–52.
- Chen J, Liu XR, Yang JW, et al. Construction and application of YOLOv3-based diatom identification model of scanning electron microscope images [J]. J Forensic Med, 2022, 38(1): 46–52.

(编辑 孙慧兰, 祁方昉)