

· 技术与应用 ·

基于人工智能硅藻自动化识别系统的实际案例应用

周圆圆^{1,2}, 曹永杰³, 杨越¹, 王亚丽¹, 邓恺飞², 马开军⁴, 陈忆九², 秦志强², 张建华², 黄平², 张吉², 陈丽琴¹

(1. 内蒙古医科大学法医学教研室, 内蒙古 呼和浩特 010030; 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. 南京医科大学法医学教研室, 江苏 南京 210000; 4. 上海市刑事科学技术研究院, 上海 200083)

摘要: 目的 探讨人工智能硅藻自动化识别系统在实际案例中的应用, 为应用该系统进行硅藻定量分析提供参考, 并对该系统所搭载的深度学习模型进行验证。方法 收集10例水中尸体的器官进行硅藻硝酸消解, 利用数字化切片扫描仪将涂片数字化扫描后, 使用人工智能硅藻自动化识别系统进行硅藻的定性定量检测。结果 该人工智能硅藻自动化识别系统所搭载的深度学习模型的受试者操作特征(receiver operator characteristic, ROC)曲线的曲线下面积(area under the curve, AUC)达到98.22%, 硅藻识别的查准率达到92.45%。结论 该人工智能硅藻自动化识别系统实现了硅藻的自动化识别, 可用于实际案例中硅藻的辅助检验, 并为水中尸体的死因鉴定提供参考依据。

关键词: 法医病理学; 人工智能; 硅藻类; 溺死

中图分类号: DF795.1 文献标志码: A doi: 10.12116/j.issn.1004-5619.2020.02.017

文章编号: 1004-5619(2020)02-0239-04



Application of Artificial Intelligence Automatic Diatom Identification System in Practical Cases

ZHOU Yuan-yuan^{1,2}, CAO Yong-jie³, YANG Yue¹, WANG Ya-li¹, DENG Kai-fei², MA Kai-jun⁴, CHEN Yi-jiu², QIN Zhi-qiang², ZHANG Jian-hua², HUANG Ping², ZHANG Ji², CHEN Li-qin¹

(1. Department of Forensic Medicine, Inner Mongolia Medical University, Hohhot 010030, 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. Department of Forensic Medicine, Nanjing Medical University, Nanjing 210000, China; 4. Shanghai Research Institute of Criminal Science and Technology, Shanghai 200083, China)

Abstract: Objective To discuss the application of artificial intelligence automatic diatom identification system in practical cases, to provide reference for quantitative diatom analysis using the system and to validate the deep learning model incorporated into the system. **Methods** Organs from 10 corpses in water were collected and digested with diatom nitric acid; then the smears were digitally scanned using a digital slide scanner and the diatoms were tested qualitatively and quantitatively by artificial intelligence automatic diatom identification system. **Results** The area under the curve (AUC) of the receiver operator characteristic (ROC) curve of the deep learning model incorporated into the artificial intelligence automatic diatom identification system, reached 98.22% and the precision of diatom identification reached 92.45%. **Conclusion** The artificial intelligence automatic diatom identification system is able to automatically identify diatoms, and can be used as an auxiliary tool in diatom testing in practical cases, to provide reference to drowning diagnosis.

Keywords: forensic pathology; artificial intelligence; diatoms; death from drowning

基金项目: 国家自然科学基金资助项目(81601645, 81722027, 81801873); “十三五”国家重点研发计划资助项目(2016YFC0800702); 上海市法医学重点实验室资助项目(17DZ2273200); 上海市司法鉴定专业技术服务平台资助项目(19DZ2290900); 上海市法医学重点实验室基金资助项目(KF1802)

作者简介: 周圆圆(1993—), 女, 硕士研究生, 主要从事法医病理学研究; E-mail: zyy601263281@163.com

通信作者: 陈丽琴, 女, 教授, 硕士研究生导师, 主要从事法医遗传学和法医病理学研究; E-mail: lqchenyj@163.com

通信作者: 张吉, 男, 助理研究员, 主要从事法医病理学研究; E-mail: zhangj@ssfjd.cn

在法医病理学鉴定实践中,硅藻检验是鉴定溺死的重要手段之一。通过对尸体进行硅藻检验,结合案情的调查,可对疑似溺死者的死亡原因作出判断,对硅藻的种属研究甚至能够推断死者的溺亡地点^[1]。硅藻检验方法有很多,传统方法包括化学消化法^[2]、DNA测序^[3]等,近年来又在这些方法的基础上研发出其他新方法,例如微波消解-扫描电镜法^[4]等。其中强酸消化法因操作简单、试剂配制方便,是基层法医实验室中最常用的硅藻检验方法。

在基于强酸消化的传统观察方法中,计数需要有专业基础和经验的法医在光镜或电镜下肉眼观察,这一过程不仅耗时费力,且容易漏检较小的硅藻。因此,将人工智能方法结合数字化病理图像技术应用于法医病理学领域实现硅藻的自动化识别,以使硅藻检验定性定量分析过程更加智能,检验结果更加准确和客观,有利于提升检案效能。

本研究将硝酸破机法联合人工智能硅藻自动化识别系统应用于实际案例的硅藻检验,并对实验结果进行分析,探讨该方法可能面临的问题,为日后人工智能硅藻自动化识别系统的应用提供一定参考。

1 材料与方法

1.1 案例收集和准备

收集10例案例的内部器官,男性6例、女性4例,其中9例是溺死案例,1例是非溺死案例。2例解剖案例包含肺、肝和肾组织,2例送检案例包含肺和肾组织,另6例送检案例只有肺组织。以上器官与组织均来源于司法鉴定科学研究院,本实验已通过司法鉴定科学研究院伦理委员会批准。

组织样本处理:每个器官均取材20g,加入30mL分析纯硝酸和1mL无水乙醇,在60℃的恒温磁力搅拌器(上海梅颖浦仪器仪表制造有限公司)上消化至液体澄清无大颗粒杂质。取10mL澄清消化液,以3000×g离心5min,弃去上清液,将沉淀混匀后在载玻片上涂片。

水样本处理:取10mL水样(均为送检单位提供的现场水样)以3000×g离心5min,弃去上清液,沉淀混匀后制成水样涂片。

将器官涂片和水样涂片均使用Aperio AT2切片扫描仪(德国Leica Biosystems公司)扫描成400倍的数字化病理图像,使用人工智能硅藻自动化识别系统进行检测。

1.2 人工智能硅藻自动化识别系统

深度学习是近年来机器学习及人工智能领域新的研究热点,通过模拟人脑神经元结构对数据进行特

征提取,从而实现图像识别等功能^[5]。其中,卷积神经网络(convolutional neural network, CNN)就是一种常用于图像识别的深度学习算法^[6]。目前,深度学习构架已经被证实在数字化病理图像中疾病程度的自动分类方面具有非常大的应用价值^[7]。基于这些技术,本研究院团队自行开发了一套适用于法医硅藻检验的自动化识别系统。

该系统主要由数字化病理图像预处理模块、识别模块和结果展示模块组成(图1)。其中,数字化病理图像预处理模块包括图片切割和图片均值化处理。识别模块中涉及的模型为GoogLeNet深度学习模型。前期研究^[8]结果显示,该模型在硅藻识别方面准确率可达到97.67%,其通过逐区域扫描切片,自动提取包含硅藻信息的区片。结果展示模块是根据对提取法医实际检案中有效的辅助信息所设计的,主要可实现的功能有硅藻计数、硅藻图片的截取等。

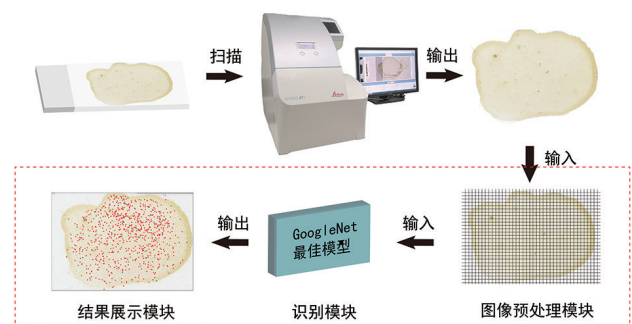


图1 人工智能硅藻自动化识别系统的操作示意图
Fig. 1 The operation diagram of artificial intelligence automatic diatom identification system

由于强烈的消化反应会产生硅藻碎片,影响计数结果,因此制定了硅藻计数规则,以保证计数的可靠性。此规则以硅藻面积为标准,面积大于50%的硅藻计数为1个,小于50%的硅藻则不计数。

1.3 实验平台与评价指标

该系统的硬件配置为:Inter Core i7-8700K@3.70 GHz,六核CPU,NVIDIA TITAN Xp 12 GB独立显卡,8 GB内存,Ubuntu 16.04操作系统,基于Python语言的PyCharm开发工具。

模型评价指标使用查准率(precision),并引入受试者操作特征(receiver operator characteristic, ROC)曲线和混淆矩阵,以便更直观且全面地对系统的性能进行验证。查准率是模型预测出真阳性的数量占所预测出阳性的总体数量的比例,代表了该模型对阳性结果的预测准确程度。ROC曲线以假阳性率(false positive rate, FPR)为横坐标、真阳性率(true positive rate, TPR)为纵坐标,曲线下面积(area under the curve, AUC)取值为0~1,值越接近1,说明该模型的性能越

好。各类指标的计算公式如下：

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (1)$$

$$\text{TPR} = \frac{TP}{TP + FN}, \quad (2)$$

$$\text{FPR} = \frac{FP}{TN + FP}, \quad (3)$$

上述公式中,TP代表真阳性,即专家(3位具有5年以上硅藻检验工作经验的法医,其中2位进行初检,1位进行复检)标识为硅藻、模型识别为硅藻的个数;FP代表假阳性,即专家标识为非硅藻、模型识别为硅藻的个数;FN代表假阴性,即专家标识为硅藻、模型识别为非硅藻的个数;TN代表真阴性,即专家标识为非硅藻、模型识别为非硅藻的个数。

2 结果

该人工智能硅藻自动化识别系统所搭载的深度学习模型的ROC曲线见图2,其AUC值达到98.22%,表明该模型在自动化识别硅藻方面具有较好的性能。经本系统识别实际案例中10 mL水样涂片10张,其硅藻检验结果均为阳性。肺涂片10张,肝涂片2张,肾涂片4张,共计检出硅藻6467个。器官涂片中假阳性结果有528个,由此得出硅藻识别的总体查准率为92.45%(表1)。其中,案件4虽然肺内硅藻只有16个,但在保证无污染的情况下,肝、肾内都检出1个硅藻,结合相关案情,该结果提示生前溺水;案件10肺内硅藻个数为0,不能确证生前溺水,结合案情调查及尸体检验,支持死后抛尸入水。结果表明,该系统灵敏度较高,可以辅助硅藻的检测识别工作。

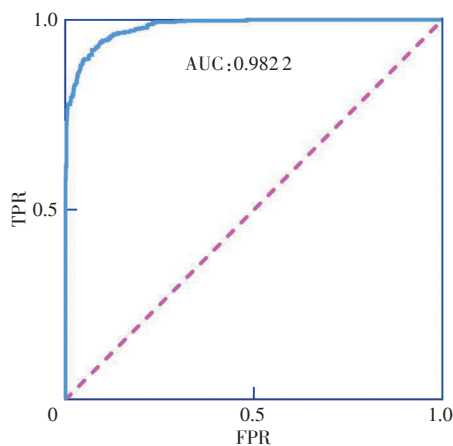


图2 人工智能硅藻自动化识别系统搭载的深度学习模型ROC曲线及AUC

Fig. 2 ROC curve and AUC of the deep learning model based on the artificial intelligence automatic diatom identification system

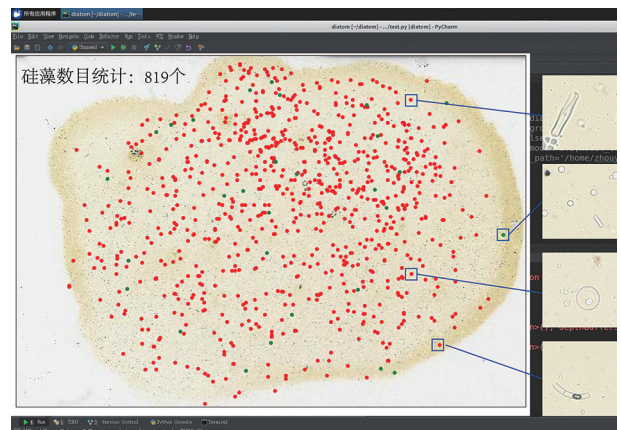
表1 人工智能硅藻自动化识别系统识别10个实际案例的硅藻数量

Tab. 1 The number of diatoms in 10 actual cases identified by the artificial intelligence automatic diatom identification system (个)

案例编号	水样	肺	肝	肾	器官涂片假阳性数量
1	1555	832	-	-	58
2	1845	2026	-	-	152
3	304	819	1	0	28
4	379	16	1	1	21
5	4031	560	-	10	40
6	388	84	-	-	13
7	874	484	-	1	56
8	743	122	-	-	3
9	1311	1519	-	-	131
10	314	0	-	-	26
合计	-	6453	2	12	528

注:“-”表示无数据。

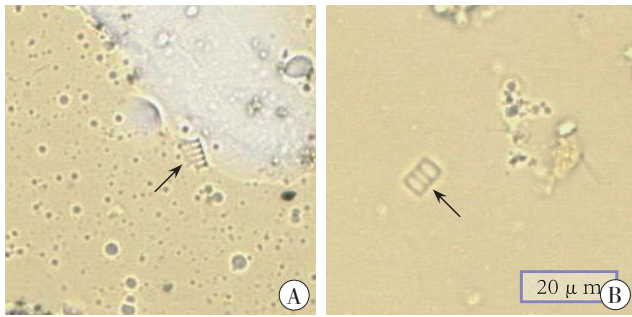
该人工智能硅藻自动化识别系统的结果展示模块见图3。由于计算机识别存在一定的误差,在这一步需要人工核实后出具结果(人工确认用时约5 min)。如图所示,真阳性结果标注为红色圆点,假阳性结果标注为绿色圆点。该模块结合法医学鉴定中的实际需求,不但可以直观地观察硅藻的分布情况并进行图片截取,还能够实现自动化硅藻计数。该系统对像素变化十分敏感,相较于人眼,计算机仅在400倍的光镜视野下就可以在复杂的背景中发现肉眼不易发现的组织样本中直径20 μm左右的微型硅藻及硅藻碎片(图4)。



红色圆点为真阳性结果,绿色圆点为假阳性结果。

图3 人工智能硅藻自动化识别系统的结果展示模块

Fig. 3 The result display module of the artificial intelligence automatic diatom identification system



A:硅藻碎片(<20 μm,箭头所示);B:小硅藻(<20 μm,箭头所示)。

图4 人工智能硅藻自动化识别系统识别出的小硅藻及硅藻碎片(×400)

Fig. 4 The small diatoms and diatoms fragments identified by the artificial intelligence automatic diatom identification system (×400)

3 讨论

硅藻检验是法医诊断溺死较为常用的方法。目前,用于硅藻检验的方法较多,但在现有的常用方法中,基于DNA序列的硅藻检验因其仪器、试剂较为昂贵,不适用于基层的法医学实验室。而基于形态学方法的硅藻检验常使用各类试剂进行消解,涂片后镜下观察过程耗时费力,且需要具备一定的专业知识和经验,为法医工作者带来一定的挑战。鉴于此,我们开发了基于人工智能的硅藻自动化识别系统^[8]。本研究验证了该系统在实际检案中的可行性,为溺死的法医学鉴定提供支持。

硅藻作为一种微生物藻类,形体极为微小,一般只有十几微米到几十微米。在溺死案例中,可以穿过气血屏障经血液循环到达肝、肾的硅藻更加微小,且消化过程中也会产生硅藻碎片,这些小硅藻及硅藻碎片在光镜下通过肉眼难以发现,而本研究所使用的卷积神经网络通过对图像像素进行特征提取从而训练出的自动化识别模型,对像素变化的灵敏度极高,在400倍的光镜视野下就可以发现肉眼不易发现的小硅藻及硅藻碎片。该系统工作效率远高于经验丰富的专家^[8],而且针对体积微小的硅藻更具有优势,不仅提高了硅藻的检出率和检案效率,结果也更准确、直观。

当然,该系统也存在一定的问题。第一,所使用的深度学习构架,其数据量的大小直接影响训练模型的性能^[5]。小的数据集训练容易过拟合,需要海量的样本进行训练才能达到较好的泛化能力。其二,该系统的识别能力取决于深度学习模型的学习量,即新背景的图片对系统的识别能力存在影响。针对上述问题,可以在检案的同时,不断对现有的硅藻数据库进行补充与扩建,在现阶段最佳模型上继续进行训练。

综上所述,本研究通过自主研发的人工智能硅藻自动化识别系统对10例实际案件中的水中尸体进行

硅藻检验,证明了该系统在法医学溺死案件中应用具有可行性。该系统在降低人工干预方面具有独特的优势,并可与诸多硅藻前处理方法(如酶消化法、真空过滤等)联合使用,为未来实现智能化硅藻检验奠定了基础。由于本系统仍处于初步研究阶段,目前仅具有自动化识别、计数等少数功能,未来可通过硅藻训练数据库的扩充、深度学习算法的优化以及视窗界面的完善等进一步提升系统的功能。此外,我们将在该硅藻自动化识别系统的基础上建立更为准确、客观、高效的硅藻种属分类模型,从而挖掘深度学习技术在溺水地点推断中的潜在应用价值。

参考文献:

- [1] 赵子琴. 法医病理学[M].4版.北京:人民卫生出版社, 2009:408-426.
ZHAO Z Q. Forensic pathology[M]. 4th ed. Beijing: People's Medical Publishing House, 2009:408-426.
- [2] TIMPERMAN J. The diagnosis of drowning. A review[J]. Forensic Sci, 1972, 1(4): 397-409. doi: 10.1016/0300-9432(72)90015-5.
- [3] CHEN X G, ZHANG J, HUANG Y, et al. Diatom taxa identification based on single-cell isolation and rDNA sequencing[J]. Forensic Sci Int Genet Suppl Ser, 2013, 4(1): e308-e309. doi: 10.1016/j.fsigss.2013.10.157.
- [4] 苏会芳,刘超,胡孙林,等. 微波消解-扫描电镜联用法在溺死诊断中的应用[J]. 中国法医学杂志, 2012, 27(2): 97-100, 107. doi:10.3969/j.issn.1001-5728.2012.02.002.
SU H F, LIU C, HU S L, et al. The application of microwave digestion technique and scanning electron microscopy to diatom test in diagnosis of drowning[J]. Zhongguo Fayixue Zazhi, 2012, 27(2): 97-100, 107.
- [5] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444. doi: 10.1038/nature14539.
- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// Proceedings of the 25th International Conference on Neural Information Processing Systems: Volume 1. New York: Curran Associates Inc., 2012:1097-1105.
- [7] ERTOSUN M G, RUBIN D L. Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks[J]. AMIA Annu Symp Proc, 2015, 2015: 1899-1908.
- [8] 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. doi: 10.1016/j.forsciint.2019.109922.

(收稿日期:2019-11-18)
(本文编辑:李正东)