

辨别生物医学论文图片真伪的工具与审读方法探索

何学令 何琳 别明江 汤洁[†]

《四川大学学报(医学版)》编辑部,610041,成都

摘要 本研究在前人研究成果基础上,结合图片审查过程中的实际案例,利用 Forensically 和 Photoshop Droplets 图片检测工具对《四川大学学报(医学版)》来稿中图片重复使用、图片局部操纵以及图片数据伪造等典型学术不端案例进行审读,并总结出快速准确的图片审读方法,以期与国内生物医学期刊同行共同探讨,共同提升防范学术不端行为的能力,推进诚信学术生态的建设。

关键词 生物医学论文;图片;审读方法学;重复使用;局部操纵;伪造

An exploration of tools and review methods for distinguishing the authenticity of biomedical paper images//HE Xueling, HE Lin, BIE Mingjiang, TANG Jie

Abstract Based on the previous research results, this study combines therealcases in the process of image review, and uses Forensically and Photoshop Droplets image detection tools to review typical cases of academic misconduct such as image reuse, image local manipulation and image data falsification in manuscripts submitted to the *Journal of Sichuan University (Medical Edition)*, and summarizes some fast and accurate image review methodsto discuss with domestic biomedical journals colleagues, so as to jointly improve the ability to prevent academic misconduct and promote the construction of an honest academic ecology.

Keywords biomedical papers; pictures; review methodology; reuse; local manipulation; fraud

Authors' address Editorial Board of Journal of Sichuan University (Medical Sciences), 610041, Chengdu, China

DOI:10.16811/j.cnki.1001-4314.2023.05.020

图片是生物医学科技论文不可或缺的组成部分,也是生物医学科技论文核心实验数据可视化展示的一种重要方式^[1]。近年来,数字化图像科技在给生物医学研究带来极大便捷与智能分析优势的同时,成熟的数字化图像处理技术也让图片的操纵变得更加容易,图片学术不端问题受到前所未有的关注。解傲等^[2]对2019—2020年我国学者发表在SCIE收录期刊上的医学研究型论文撤稿原因进行了分析,其中图片问题成为撤稿的第二大原因(约占21.3%)。学术不端不仅给作者带来科研诚信危机,也会给期刊的学术影响力造成巨大的负面影响。加强期刊编辑对图片数据的辨别能力是科技期刊亟须研究的重要课题。

目前,国际上针对图片学术不端的检测软件和平台为期刊编辑对图片的审读带来了极大便利,如 Proofig、Figcheck、Forensically、Photoshop Droplets 及 Image J 等^[3-4]。不少国际优秀期刊专门增加了学术图片审读环节^[5-6]。图片学术不端问题也逐渐被我国期刊重视,国内不少专家学者针对图片学术问题进行了大量研究^[7-9]。但当期编辑面临大量稿件时,如何准确利用这些工具快速审读学术图片值得我们进一步去探索。本文作者通过选取审稿中典型的生物医学图片学术不端案例,采用常见审读工具探索不同类型图片学术不端的最佳审读方法,与国内生物医学期刊同行共同探讨,为生物医学图片审读提供参考。

1 学术图片审读常见工具

目前,我们常用的免费图片审读工具有 Forensically、Photoshop Droplets 等,本研究主要基于以上2款图片审读工具探索图片审读的方法。

1.1 Forensically

Forensically 是一套基于 Web 的免费图像取证工具(<https://29a.ch/photo-forensics>),它可用于检测克隆检测(Clone Detection)、错误等级分析(Error Level Analysis)、噪声分析(Noise Analysis)、水平扫描(Level Sweep)、亮度梯度(Luminance Gradient)、主成分分析(Principal Component Analysis)等。其中,Clone Detection 可检测图片是否存在重复像素的现象;Error Level Analysis 可使图片中被操作的区域被凸显出来(更暗或更亮);Noise Analysis 用于识别对图像的操作,如喷笔工具、变形、扭曲和透视校正克隆;Level Sweep 用于检查图片中复制粘贴区域,该功能可使复制粘贴引入部分的边缘更加明显;Luminance Gradient 可用于检测图像中不同部分的异常亮度;Principal Component Analysis 组件可以从不同的角度检测图像被修改的痕迹。

1.2 Photoshop Droplets

Droplets 插件是一个基于 Photoshop 软件行使功能的多用途工具包。可以根据图片类型选用相应的功能组件对图片进行分析,如用于分析电泳凝胶图和免疫印迹图的组件(Features in Dark or Light Areas)、用于分析流式细胞图和细胞免疫荧光图片的组件(Foren-

[†] 通信作者

sic-Gr Map)、用于分析检测图片重复使用的组件(Color Overlay Blots Advanced Overlay-AdjustmentLayers)等。Droplets 工具包下载以及详细使用方法见 <https://ori.hhs.gov/droplets>。

2 常见图片学术问题类型及审读方法

目前常见图片学术问题可分为4类,即图片剽窃,图片重复使用,图片局部操纵(包括复制粘贴、拼接等手段篡改图片),图片数据伪造^[10]。此外,在图片选用的科学性、一致性方面同样存在学术问题。接下来,我们以《四川大学学报(医学版)》来稿中典型图片学术不端为例,运用 Forensically 和 Photoshop Droplets 检测工具对稿件中的图片进行审读,探索常见类型图片学术不端快速的审读方法。

2.1 图片剽窃

图片剽窃包括将自己已发表的图片数据或将他人已发表的图片数据作为自己新的图片数据成果公开发表。对于该类图片学术不端审查主要依赖网络学术图片数据库相似度检测平台和相似图片检测算法^[11]。目前国内提供图片相似度检测服务的平台有百度 AI 平台,方正的图片检索系统等,其在一定程度上有助于图片学术不端的审查。而相似度检测算法技术需要具

备较高的专业背景,只有少数出版机构能够使用。

2.2 图片重复使用

图片重复使用是科技论文中常见的图片学术不端类型,在具有较高影响力的国际学术刊物上也屡见不鲜。图片重复使用常见类型有以下2种。

1)同一张图片素材重复使用于不同实验的结果。该类型主要包括原图重复使用,图片旋转、缩放后的重复使用等。由于同一张图片的重复使用较为明显,审读相对容易。图1-a~b为典型的 Western blot 蛋白条带重复使用;图1-c~d为典型的图片缩放后重复使用。对肉眼审读有一定难度的图片可使用 Forensically 中 Clone Detection 工具进行审读(图1-a、图1-d),Western blot 蛋白条带图片也可使用 Droplets 中 Forensic-Gr Map 工具进行审读(图1-b),组织细胞学图片可使用 Photoshop Droplets 中 Color Overlay Blots 工具进行审读(图1-c)。我们发现 Forensically 审读结果更直观(图1-a、图1-d 中线条连接区域)。Photoshop Droplets 可用于验证和确认 Forensically 对 Western blot 蛋白条带图的审读结果(方框标注条带)和组织细胞学图片审读的结果(图1-c 箭头指示红色区域),但 Photoshop Droplets 常出现将非重复区域标记为相似的情况,需编辑谨慎确认。

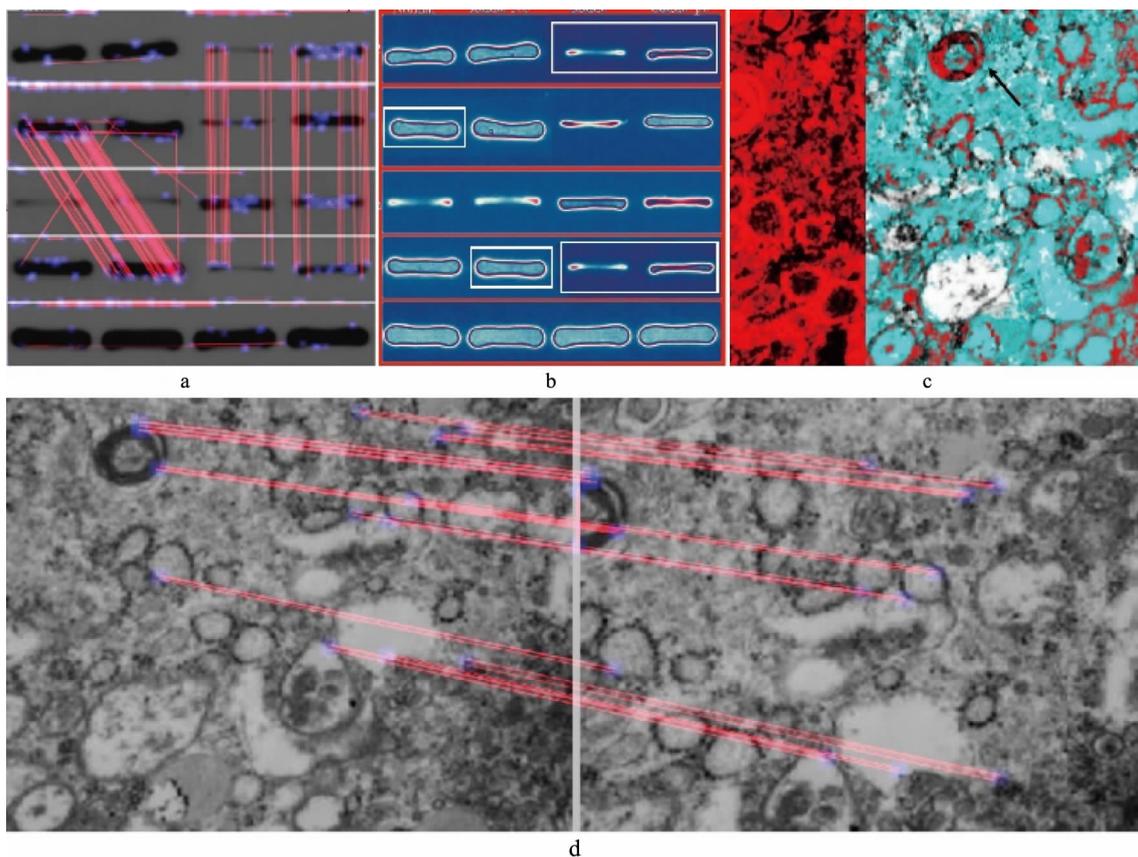


图1 蛋白条带、组织细胞学图片重复使用案例

2)不同图片之间部分区域重复使用。该类型常见于同一样本重复拍照等。对该类图片的审读比较困难,尤其对于内容具有天然相似形态结构的图片审读难度较大,如组织切片、体外培养细胞等。图2-a是本刊某篇来稿中的肺组织切片。作者拟观察某种纳米药物对小鼠肺腺癌的疗效,研究采用小鼠尾静脉注射A549肺腺癌细胞建立肺肿瘤模型,通过气管直接给予纳米药物,观察其抗肿瘤活性。在图2-a中a组、b组分别为对照组和无药物的纳米载体组,仅通过肉眼我们很难识别图片之间是否有重复的区域。我们先后采用Photoshop Droplets中Color Overlay Blots工具,Forensically中Clone Detection工具对图2-a中a组、b

组进行审读。结果发现,Photoshop Droplets检测2张图片重叠后出现较多标记为红色的相似部分(图2-b),但多数被标记的相似部分并非重复区域,这可能与组织结构的天然相似性有关。采用Forensically进行审读时,也同样检测出相似而非重复的区域,通过调整Minimal Similarity、Minimal Detail、Minimal Cluster Size检测参数,排除相似而非重复的区域,实际的重复区域就被线条显著标注出来(图2-c),再对标注区域进行人工鉴别和确认重复使用区域(图2-c中白色方框区域)。我们采用相同的方法对来稿中的类似图片进行检测,可快速准确地审读图片。

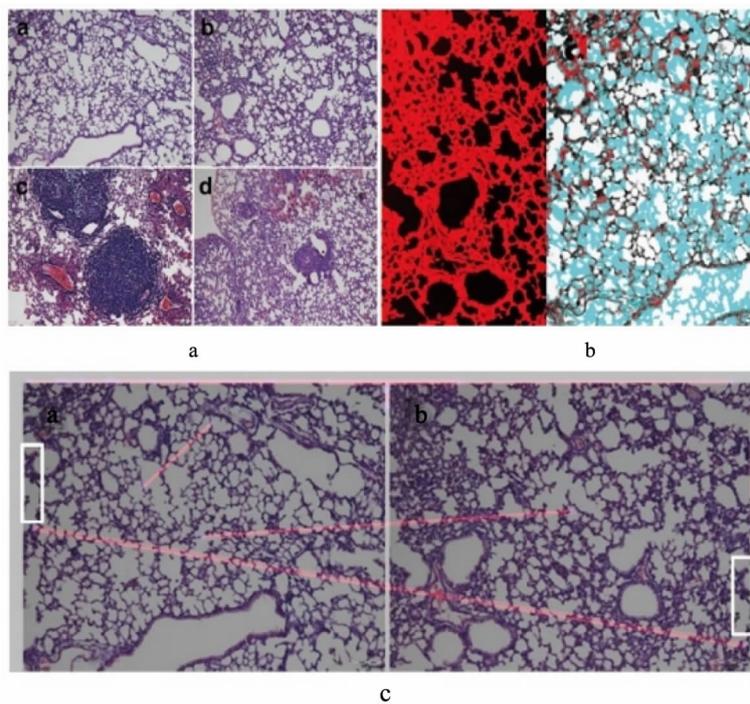


图2 组织学图片部分区域重复使用案例

2.3 图片局部操纵

近年来,因图片学术不端而撤稿的学术论文中,涉及图像局部操纵问题的占较大比例。图片局部操纵主要有局部复制粘贴、拼接、篡改等形式。这类图片有其典型特征,如对图片局部操纵后往往会破坏其像素连续性,图片背景不连续,亮度、颜色不能平稳过渡等。Forensically和Photoshop Droplets基本可满足审读该类图片的要求。图3与图1-a为同一张图片,我们已经确认作者将图片中蛋白条带重复使用的情况。为了验证作者确实对图片进行了条带复制粘贴,我们分别用Forensically中Principal Component Analysis工具(图3-a)、Photoshop Droplets中Dark or Light Areas工具进行审读(图3-b)。结果发现,2种审读方式结果显示

方式存在一定差异,但均明显证实图片有复制粘贴拼接的情况(方框标注条带)。

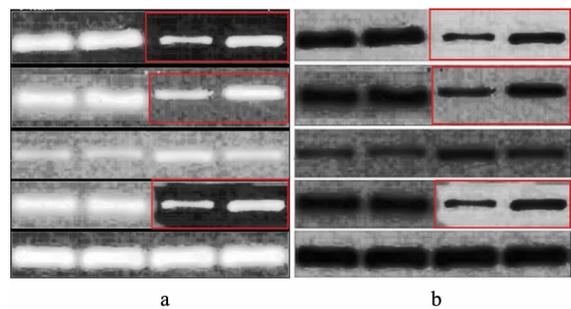


图3 蛋白条带图局部复制粘贴拼接案例

图4是本刊某篇来稿中的细胞荧光图片。采用

Forensically审读时发现图4-a左图以及左右图片之间均存在相似区域(部分线段标示区域)。为了确认图片是否存在局部操纵,我们采用Forensically中Principal Component Analysis->Mode->Component工具,Photoshop Droplets中Dark or Light Areas工具进行了审读,结果发现,左右图片背景曝光度不一致,同一张图片中背景不连续(图4-b中箭头指示区域)。为了验证审读结果,我们采用Forensically中Principal Component Analysis->Mode->Projection-Invert工具分析,进一步证实了图片有明显的篡改痕迹(图4-c中箭头指示区域)。我们建议判断该类荧光细胞图是否存在局部操纵,可采用Forensically中不同工具进行审读和验证,也可采用Forensically对Photoshop Droplets审读结果进行验证和确认。

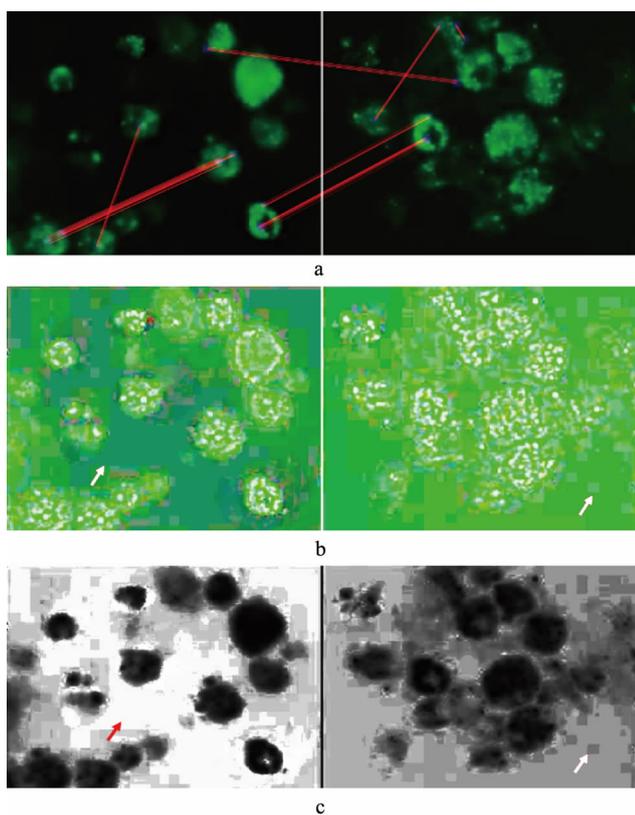


图4 荧光细胞图片局部操纵案例

图5是本刊某篇来稿中Western blot蛋白条带图,审稿时发现部分条带周边区域背景不连续,怀疑有篡改的痕迹。为了确认图片的真实性,我们分别用Forensically中Principal Component Analysis工具,Photoshop Droplets中Dark or Light Areas工具进行审读。结果显示,Forensically(图5-a)、Photoshop(图5-b)均清楚显示图片中第一行蛋白条带周围背景不连续,有明显的篡改痕迹(箭头指示区域)。我们进一步用Forensically中Luminance Gradient工具进行验证

(图5-c),同样清楚显示第一行以及第三行部分蛋白条带有篡改的痕迹(箭头指示区域)。针对该类图片的局部操纵,我们建议采用Forensically中不同工具进行审读或验证,也可采用Forensically对Photoshop Droplets的审读结果进行验证。

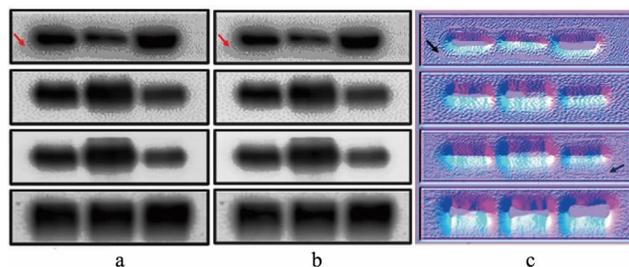


图5 蛋白条带图局部操纵案例

2.4 图片数据伪造

图片数据伪造主要是指图片数据不是源于实际实验取得的。相对于图片重复使用和图片局部操纵,图片数据伪造一般比较隐蔽,不容易被发现。图6是本刊某篇来稿中Western blot蛋白条带图。审稿时发现HO-1条带顶部与GPX4条带底部有疑似重叠区域(椭圆标注区域),考虑作者可能是为了清楚展示结果而有意将同一张聚偏二氟乙烯膜(PVDF膜)上的蛋白条带分割开。为了进一步对审读结果进行确认,我们利用Photoshop Droplets中Forensic-Gr Map工具进行审读,结果显示4个条带之间均有不同程度的重叠(方框和圆圈对应标注区域),证实这些蛋白条带均来自同一张PVDF膜。根据重叠区域线索还原了蛋白条带在PVDF膜上的原有排序为GPX4,HO-1,Nrf2, β -actin。然而,我们查阅蛋白的相对分子质量发现,GPX4、HO-1、Nrf2、 β -actin的相对分子质量分别约为22、33、110、42 kDa。根据聚丙烯酰胺电泳原理^[12],相对分子质量越小的蛋白迁移速率越快,因此Nrf2(110 kDa)条带在PVDF膜上应位于 β -actin(42 kDa)之后,而在该稿件结果中,Nrf2条带却位于 β -actin条带之前。因此,我们确认该稿件中的蛋白条带图片数据系伪造。

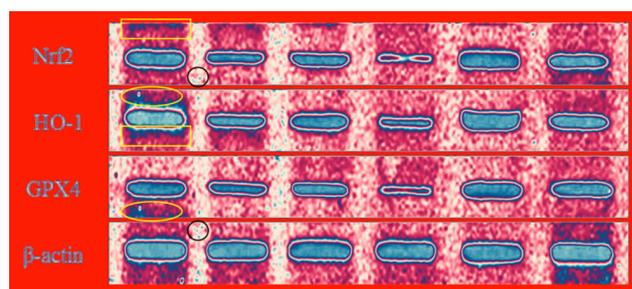


图6 蛋白条带图片数据伪造案例

2.5 涉及图片选用的科学性、一致性的学术问题

该类图片学术问题主要体现在作者选取图片的标准是否一致,选取的图片是否有误或不恰当等。由于该类图片学术不端涉及学科专业,审读难度较大。图7是本刊某篇来稿中肿瘤组织图片,作者拟观察某种药物对小鼠皮下移植瘤生长的影响,应用免疫组织化学的方法检测增殖细胞核抗原在肿瘤组织的表达,评价肿瘤的增殖情况(图7-a);采用原位末端标记法评价药物对肿瘤凋亡的影响(图7-b)。图中A组、B组、C组、D组分别代表模型组、空白组、低剂量组、高剂量组。作者得到的结论是高剂量组肿瘤细胞增殖能力受到明显抑制、药物能明显促进肿瘤细胞凋亡。我们审读时发现D组选取的组织区域与其余3组有一定差异,A组、B组、C组中肿瘤组织视野均选取肿瘤边缘,而D组却选取肿瘤组织中心区域(箭头指示区

域)。由于肿瘤组织边缘细胞和中心区域细胞增殖能力并不一致,且肿瘤中心由于缺氧等原因细胞会出现坏死,因此得到的结论可能不太准确^[13]。D组图片选取标准与其余3组不一致,在科学性方面存在学术不端的嫌疑。

在对肿瘤细胞凋亡图片审读时发现C组、D组荧光分布特征与B组有明显差异(箭头指示区域),通过调整图片曝光度发现,荧光较强区域并非均为肿瘤细胞凋亡产生的,大部分为肿瘤组织中的血管内或血管周围血细胞的自发荧光。该稿选取图片数据存在学术不端的嫌疑。

图片选用的科学性、一致性学术问题在很多情况下无法使用审读工具完成,只能通过编辑专业知识进行判断。因此,编辑在工作中需要加强学科的专业知识学习,不断提升自己的专业水平。

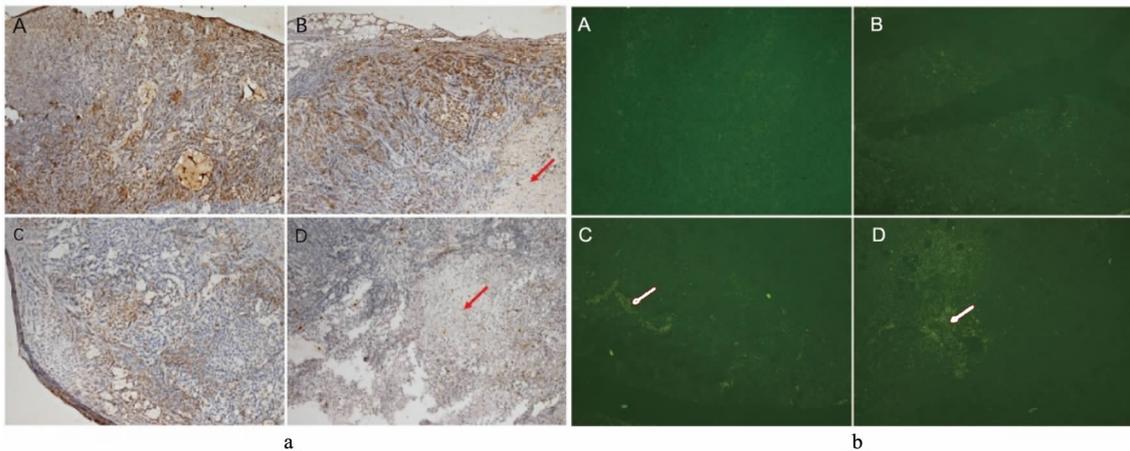


图7 图片选用的科学性、一致性存在问题的案例

3 思考与建议

3.1 图片数据分析岗位更需要复合型人才

数字化图像科技飞速发展给科学研究带来极大便捷性的同时,图片操纵方式也变得更加隐蔽。作为科研诚信守门人的科技期刊,对图片学术不端防范的重要性不言而喻。韩磊等^[14]建议国内生物医学科技期刊借鉴国际优秀生物医学期刊设立专职岗位对图片数据进行分析。欧洲分子生物学组织(EMBO)对岗位人员的专业背景侧重于图片软件技术,对生物医学专业背景没有较高要求。然而,从专业要求及实际工作出发,图片数据分析岗位更需要具有生物医学背景和熟悉图片软件技术运用的复合型人才。

3.2 编辑需要重视建立图片学术不端分析方法

虽然图片学术不端形式多样,但在稿件审读工作中我们可以根据其特征、手段等进行分类总结。通过

对稿件中真实案例图片学术不端行为类型及其特征的分析、总结,有助于图片学术不端行为的识别和防范。作为稿件初审编辑,可根据图片学术不端类型建立不同的检测方法,快速准确地对图片进行审读。

3.3 可采用多种方式审读图片数据并进行验证

叶青^[15]发现不同图片检测工具以及不同背景人员对结果的判定存在一定差异。本研究同样发现不同审读工具的检测结果存在一定差异。我们建议在建立图片学术不端分析方法时,针对某种类型图片学术不端建立不少于2种审读方法,将其中1种方法作为主要审读方式,再采用其余方法对审读结果进行验证,达到对图片数据的准确审读。

3.4 优化图片数据审读流程

EMBO出版社对图片数据进行分析的工作是在稿件经过同行评议、拟接收稿件后进行^[16]。我们建议图片数据审查环节可开始于稿件的初审阶段,对于存在

明显图片学术不端的稿件,如有不合理重复使用、拼接、伪造痕迹的图片,可直接作退稿处理;而稿件中的图片可疑,需要进一步确认,可要求作者提供原始数据并对图片数据进行说明,或在同行评议时邀请评审专家进行审查,这样可投入相对较少的时间将图片学术不端事件控制在发表前。

4 结束语

本研究采用不同的图片审读工具对本刊来稿中学术不端图片进行了检测,主要涉及图片重复使用、图片局部操纵以及图片数据伪造3种类型,为生物医学科技期刊图片初审编辑快速审读图片提供参考。同时对来稿中涉及图片选取标准缺乏一致性、科学性案例进行了简单分析和审读。由于生物医学稿件专业性较强,虽然多数编辑具有一定的专业背景,但由于长期脱离科研工作,编辑往往在审稿时显得能力不足。因此,科技期刊图片数据分析编辑在熟练掌握图片软件相关技术的同时,也需要在工作中不断更新自己的专业知识,积极参加学术会议和相关科研工作,力争成为复合型编辑。

5 参考文献

- [1] 沙力妮. 科技论文图表问题分析与对策[J]. 新闻传播, 2022, 430(13): 87
- [2] 解傲, 袁路, 汪伟. 我国医学 SCIE 研究型论文被撤销的原因分析[J]. 中国科技期刊研究, 2022, 33(5): 554
- [3] VAN NOORDEN R. Journals adopt AI to spot duplicated images in manuscripts[J]. Nature, 2022, 601(7891): 14
- [4] BAKER M. Problematic images found in 4% of biomedical papers [J/OL]. [2023-05-01]. <https://doi.org/10.1038/nature.2016.19802>

- [5] Image doctoring must be halted [J]. Nature, 2017, 546(7660): 575
- [6] POLKA J K, KILEY R, KONFORTI B, et al. Publish peer reviews [J]. Nature, 2018, 560(7720): 545
- [7] 张维, 邹仲敏, 汪勤俭, 等. 生物医学论文典型学术造假图片辨析及防范措施探讨 [J]. 编辑学报, 2021, 33(3): 280
- [8] 陈秀妍, 张梦狄, 韩向娣, 等. 图表数据学术不端案例调研与防范研究 [J]. 中国科技期刊研究, 2021, 32(5): 555
- [9] 余菁, 邬加佳, 刘清海, 等. 医学论文图片篡改实例分析及对策 [J]. 传播与版权, 2019(5): 51
- [10] 叶青, 林汉枫, 张月红. 图片中学术不端的类型与防范措施 [J]. 编辑学报, 2019, 31(1): 45
- [11] 孙力炜, 贺郝钰, 迟秀丽, 等. 防范图片学术不端的举措研究 [J]. 中国科技期刊研究, 2021, 32(5): 563
- [12] 陈彩萍, 冯志奇, 宋壮, 等. 蛋白免疫印迹法同时检测大、小分子蛋白的实验条件改进 [J]. 现代生物医学进展, 2016, 16(24): 4618
- [13] ZHAO Y, FU X, LOPEZ J I, et al. Selection of metastasis competent subclones in the tumour interior [J]. Nature Ecology & Evolution, 2021, 5(7): 1033
- [14] 韩磊, 叶青, 郑云飞, 等. 国际优秀生物医学期刊深度查证学术图片的特色流程: data integrity analysis 专岗审核 [J]. 编辑学报, 2021, 33(2): 231
- [15] 叶青. 英文科技期刊论文图片学术不端审读方法探索与思考 [J]. 中国科技期刊研究, 2022, 33(5): 591
- [16] ROSSNER M. Figure manipulation: assessing what is acceptable [J]. The Journal of Cell Biology, 2002, 158(7): 151

(2023-05-04收稿;2023-07-22修回)

满庭芳·浙江高职高专期刊专委会年会感怀

谭 华/《南通职业大学学报》编辑部,226007,江苏南通

2023年5月,浙江省高等教育学会高校期刊分会高职高专委员会学术年会在温州召开,应邀参会,与业界同人学习交流甚欢,又《文史哲》前副主编刘京希教

授、《浙江理工大学学报》编辑部主任康锋博士等报告精彩之至,收获良多,因作此词记之。

丽水街幽,楠溪水碧,瓯城欢事重重。隔屏赏忆,欣记此相逢。
康博京希授道,珠联璧,学贯西中。咖云集,道来娓娓,处处见神功。

群贤风采美,豪情智趣,清雅仪容。漫留得,欢歌笑语随风。
此乐谁堪共说,唯记取,浅聚情浓。佳期待,来时相拥,锦字寄归鸿。