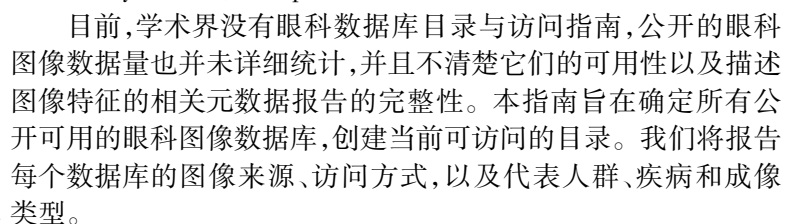


【述评】

中国医药教育协会数字影像与智能医疗专委会 中国医药教育协会智能医学专委会



1 数据资源获取方法

1.1 搜索策略和纳入标准 所有国家或地区任何形式的人类眼科图像数据库都在筛选范围内,且不受语言限制。非眼科图像、文本或仅数字数据和来自其他物种图像的数据库被排除在外,患者的年龄、性别或种族不包含在数据库纳入与排除的标准内^[19]。

搜索由两部分组成。首先,我们对 MEDLINE 进行文献检索,以确定描述可公开获得的眼科图像数据库的研究,然后尝试从源头获取数据库。该搜索将描述各种类型眼科图像的术语,如“眼(eye)”“眼底(fundus)”和“视网膜(retina)”,以及“光学相干断层扫描(optical coherence tomography)”“视网膜图像(retinal images)”“生物测量学(biometry)”和“地形图(topography)”,以及诸如“数据集(dataset)”和“数据库(database)”等术语结合起来。其次,我们使用 Google 数据集搜索和 Google 搜索引擎,以类似的搜索关键词进行了有针对性地搜索。Google 数据集搜索专为发现在线存储数据库而设计,支持查找表格、图像和文本数据库。此外,还为那些有兴趣使用元数据参考模式发布数据库的人提供索引。使用索引搜索后搜索返回的所有结果包括数据库内容、直接链接和文件格式的描述。Google 搜索还包括与眼部疾病和眼科图像相关的术语,以及与数据库相关的术语。我们系统地整理和筛选了每次 Google 搜索引擎和 Google 数据集搜索返回的前 10 页的结果。

除上述检索途径外,我们还通过手动筛选相关文章的参考文献、相关的会议论文和进展以及咨询该领域的智能眼科专家来补充。所有 3 种搜索途径来源的结果,包括 MEDLINE 数据库初始的所有结果,于 2019 年 12 月 3 日完成,MEDLINE 搜索结果于 2020 年 5 月 11 日更新。

1.2 眼科图像数据库的识别 MEDLINE 的搜索结果由 2 位主要智能眼科研究人员进行筛选,以确定所有相关数据库的名称和来源。在可用性状态不清楚的情况下,我们纳入了这些数据库并尝试访问其来源,报告了访问的任何一步出现的障碍。主要评审者还筛选了 Google 搜索引擎和 Google 数据集搜索结果,以直接识别相关数据库。如果数据库是否符合纳入标准存在模糊性,则由第 3 位智能眼科研究人员独立评审数据库,如果无法解决,则咨询第 4 位智能眼科研究人员。

1.3 数据库访问和描述 我们对数据库可访问性级别进行了分类,包括分层分级系统,其中数据库可访问性被描述为开放、安全或受控^[20-21]。在本次审查中,我们将数据库的可访问性定义为:(1)开放访问:对访问没有要求或仅有最低要求(如提交个人信息、电子邮件请求或创建账户等);(2)有障碍的开放访问:这些障碍是满足开放访问理论标准的数据库,

由于不可预测的原因(如对请求没有响应或超链接失效等)而无法访问;(3)规范访问:要求履行正式协议、批准或付款。对于数据库访问需要电子邮件请求的开放访问类别,允许电子邮件回复期限为 2 周;如果网站没有响应或下载链接不起作用,则在放弃访问尝试之前,由第 2 位和第 3 位审阅者进行检查。

1.4 数据库特征的提取 我们预先设计了数据库的信息提取表格,并在前 20 个数据库上进行了实验以验证其效率。记录表详细记录了每个数据库的信息,包括与数据源相关的直接链接、可访问性、数据库主要内容(人口、疾病种类和图像类型等)以及相关元数据(包括临床数据、图像标签等)。如果这些信息是从源头获取的(如在数据库网站上),或者有描述数据库的论文链接,则直接记录所提供的信息。此外,我们还介绍了所有开放获取数据库中关键临床元数据项目报告的完整性。如果信息在数据库文档、描述或引用的出版物中进行了描述,即使是在总体水平描述的,依然将这个目标标记为已报告。

1.5 从文献检索中识别的数据库 MEDLINE 搜索确定了 3542 篇文章,其中 2361 篇在标题和摘要筛选的基础上被排除。剩余 1181 篇文章获取全文后用于筛选文中提及的可公开获取的眼科图像数据库,其中 534 篇没有提及公用数据库,151 篇报告了非眼科数据库。在剩下的 496 篇文章中,我们确定了 161 个潜在的数据库。由于多项研究经常引用相同的数据库,我们使用 Google 数据集搜索和 Google 搜索引擎对数据库进行了有针对性地搜索,确定了 106 个数据库。对这些数据库对比后删除了 81 个重复数据库,并在此基础上进一步排除了 46 个数据库(35 个不符合纳入标准,11 个被包含于其他数据库中)。最终确定了 140 个独立的数据库,并将其纳入数据源进行进一步评估。

2 眼科图像数据库的特征

2.1 数据库访问类型 在 140 个独立的数据库中,94 个为开放访问,可以从中下载原始数据;27 个数据库被归类为有障碍的开放访问,无法从中下载数据;19 个数据库为规范访问,即具有受管制的访问权限(12 个需要许可协议,6 个需要道德委员会或机构批准,1 个需要支付 2250 英镑费用^[22])。只有 94 个开放获取数据库可以从中粗略表征以检查原始数据本身。

2.2 可访问数据库的地区分布特征 在 94 个开放获取数据库中,25 个来自亚洲(4 个来自南亚,21 个来自东南亚或东亚),9 个来自北非和中东,34 个来自欧洲,16 个来自北美,2 个来自南美,1 个来自非洲。其中 9 个数据库包含来自多个国家的图像,还有 13 个数据库来源地不详,此外没有来自大洋洲的数据库。从 2003 年到 2019 年,共有 47 个数据库注明了数据库的创建信息。全球 94 个开放获取数据

库特征的详细数据表请从《眼科新进展》官方网站获取,获取路径:<http://www.ykxjz.com/docs/tzgg/details.aspx?documentid=54&Nid=A967CBAD-BC53-4787-88ED-CD9D9ABAA7DE>。

2.3 患者和图片数量分布特征 从 94 个数据库中,我们能够访问来自 55 个数据库的至少122 364 名患者(39 个数据库未记录患者人数)的 507 724 张图片和 125 段视频。数据库中的患者人数从 2 人到 85 550 人不等(中位数 = 50;IQR = 371),图像数量从 8 张到 109 312 张不等(中位数 = 220;IQR = 1017)。我们无法确定 CASIA Iris Aging 数据库图像的确切数量,因此根据数据库描述计算保守估计值 $n = 26\,038$ 。在所有图像中,超过一半由 3 个最大的数据库贡献:Kermanny 及其同事(109 312 张图像)^[5]、EyePACS(88 702 张图像)^[23] 和 MRLEye(84 898 张图像)^[24]。有 68 个数据库的图像少于 1000 张,图像的数量从 8 张到 850 张不等(中位数 = 111;IQR = 245)。

2.4 数据库来源机构类型 图像采集的最常见原因是临床研究或临床试验(54/94,57.4%),其次是临床护理或筛查(23/94,24.5%)。5 个(5.3%)数据库来自初级卫生机构(包括筛查计划),45 个

(47.9%)来自二级卫生机构(医院或眼科诊所),18 个(19.1%)来自其他机构(如来自大学、研究机构或眼库),还有 1 个(1.1%)来自非医疗保健机构,另外有 25 个(26.6%)数据库未报告数据来源。只有 20 个(21.3%)数据库提供了患者知情同意证明,以及只有 26 个(27.7%)数据库提供了图像获取或分享的伦理批准。

2.5 数据库中图片来源疾病分布特征 数据库中包括了糖尿病视网膜病变(35/94,37.2%)、青光眼(19/94,20.2%)、年龄相关性黄斑变性(15/94,16.0%)、高血压性视网膜病变(6/94,6.4%)和白内障(4/94,4.3%)等,另外还有来自健康眼的数据(58/94,61.7%)(图 1)。此外,17 个数据库(18.1%)未明确其中包含的疾病。这些数据库可能包含健康眼的数据,但数据源没有给出具体的说明。94 个数据库中的 53 个(56.4%)包含一种以上的疾病或健康眼的数据。健康眼的图像旨在用于一系列生物医学应用,例如正常解剖结构的分析(包括内皮细胞密度、光感受器的检测、神经评估和血管形态)和技术用途(包括去噪图像、虹膜识别和眼动追踪)。

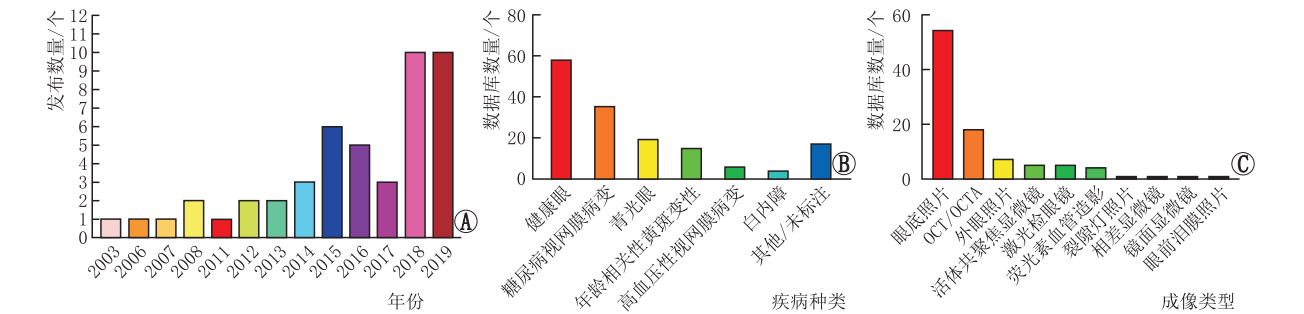


图 1 数据库的发布日期(A)、疾病类型(B)和成像类型(C)分布 注:如果数据集包含多种疾病,则会对其进行多次计数。

2.6 数据库中图片成像方式 数据库中的图片成像方式包括眼底照相(54/94,57.4%)、OCT 或 OCTA(18/94,19.1%)、外眼照片(7/94,7.4%)、活体共聚焦显微镜(5/94,5.3%)、扫描激光检眼镜和自适应光学扫描激光检眼镜(5/94,5.3%)、荧光素血管造影(4/94,4.3%)、裂隙灯照片(1/94,1.1%)、相差显微镜(1/94,1.1%)、镜面显微镜(1/94,1.1%)、眼前泪膜照片(1/94,1.1%)和视频(2/94,2.1%)(图 1)。其中,5 个数据库包含多模态图像。在基于 OCT 的 18 个图像数据库中,一半包含二维成像数据,另一半包含三维成像数据。在 94 个数据库中,大多数数据库(82/94,87.2%)以 PNG、TIFF、BIF 或 JPG 格式存储图像,10 个(10.6%)数据库提供了 Matlab 格式的文件,1 个数据库提供了 Python 格式的文件,1 个数据库提供了 Netpbm 格式的文件,1 个数据库提供了 HDF 格式的文件,9 个(9.6%)以多

种格式存储图像。在 94 个数据库中,55 个(58.5%)提供标签注释的图像(包括诊断标签,如糖尿病视网膜病变严重程度的等级或特征标签,又如标记为动脉或静脉的血管),33 个(35.1%)提供手动分割注释的图像,14 个(14.9%)既提供标签注释的图像又提供分割注释的图像。注释由眼科临床医生、普通医生和研究人员(包括医学生、验光师、操作员等)提供。此外,15 个数据库呈现的图像很容易分为训练子集或测试子集,但是数据库并没有说明是否在患者层面进行了区分。

2.7 元数据报告的完整性 图像文件及其采集相关的技术细节有比较详细的报告,但许多数据库缺少了相关的临床信息。所有的数据库均报告了成像方式、图像数量和图像格式;86.2%的数据库报告了国籍;85.1%的数据库报告了设备名称及制造商;81.9%的数据库报告了病种。但是包含患者特征

(如年龄、性别和种族)的数据库数量不足 20%, 还有 74.5% 的数据库没有报告任何的人口统计学数据, 且只有 14.9% 的数据库描述了纳入和排除标准, 只有 19.1% 的数据库报告了数据收集期。

3 数据库可访问性、透明度和代表性评价

这是第一项整理可公开获得的眼科图像数据库的指南。我们通过搜索医学文献和数据集搜索引擎, 发现了 94 个眼科数据库符合条件, 包含 50 多万张图片。这些图片中最多的是健康眼的数据, 其次是糖尿病视网膜病变、青光眼和年龄相关性黄斑变性。这些疾病是眼科临床和研究中最常见的疾病。许多国家有糖尿病视网膜病变筛查项目, 因此许多大型国家级数据库均包含糖尿病视网膜病变的资料^[25-27]。

由于眼底照相广泛的适用性与便捷性, 许多数据库 (54/94) 包含了这种类型的数据。第二种常见的成像方式是 OCT 和 OCTA (18/94, 其中 9 个包含三维 OCT 数据)。三维 OCT 数据有相邻的 B 扫描从而可以提供连续的信息, 因此有利于 ML 算法学习可能增强其性能的关键结构信息^[28-29]。

目前, 流行病学家、计算机科学家等一系列研究人员越来越多地使用公开数据库进行研究。这些数据库可以成为研究的有力工具, 但与任何数据源一样, 需要考虑该数据库的来源和局限性。在这一部分我们将重点介绍此类数据库三大重要特征: 可访问性、透明度以及不同人口的代表性。

从第一点可访问性来看, 我们的审查确定了 94 个可能是开放访问的数据库, 其中有一些数据库广为人知, 但许多数据库并非如此, 这就导致这些数据库较少被使用, 并可能由于过度使用一些较常用却代表性不足的数据库而导致研究结果的误差。我们整理的数据库目录将有效地改善这一现象, 通过提高数据库的影响力与使用率, 并提供一些关键元数据, 为研究人员的研究确定最合适的数据库。我们提供了一个初始访问路径, 以提高它们的访问率。此外, 我们建议数据库还应附带有其来源的充分信息, 以便研究人员能在符合伦理学和管理需求的框架下对这些数据进行访问。

第二个重要问题是数据库的透明度。数据库的价值不仅仅与它的大小有关, 也与其相关信息的透明度有关。当然, 数据库的大小也有价值, 例如在 ML 模型的开发中, 或者在异质人群中寻求检测适度信号时, 大的数据库具有重要价值。但数据库的可用性也将与数据的质量、深度和代表性相关。DRIVE 等小型数据库 (仅由 40 幅图像组成) 就是高质量、小数量数据库的代表。DRIVE 由于其丰富的注释和优秀的质量已经成为研究人员用于视网膜血管研究的热门资源^[30-31]。鉴于研究人员需要展示研究结果的普遍性及其临床适用性, 这些数字存储库

必须充分代表人口及其疾病的多样性, 同时应进行充分的报告, 以帮助用户做出适用性决策。然而, 我们发现关于数据特征的报告很少, 这有待于今后改进。虽然从数据库建立的技术角度来看, 这些是不重要的, 但它们几乎对于所有的临床应用都至关重要。如果没有关于人口和疾病的关键信息, 就不可能假设这些数据在现实世界中的普遍性。包括 Geburu 等^[32]在内的许多学者也呼吁数据库的报告应当提高透明度。

事实上, 提供标记丰富的数据存在公认的挑战。元数据项的管理要求很高, 成本也很高, 需要确保其准确性和完整性。过度包含详细元数据反而可能增加数据项重新识别的机会, 也会带来额外的隐私问题。因此, 保有、存储和访问这些数据库都需要适当的道德监督。当然, 这些风险应该与数据库广泛使用代表性可能不佳的数据和数据不可用所带来的潜在危害相平衡。另一方面, 遵守原始临床试验数据共享指南可以降低重新识别的投入和风险^[33]。符合要求的高质量数据库往往需要大量成本, 因此此类数据库不太可能免费提供。

最后一个关键问题是确保这些数据库有充分的代表性。目前存在的一个主要问题是在公共和其他数据库中, 特定群体的代表性可能不足, 从而对某些人群或疾病群体造成未知的偏倚。专门针对一个群体开发的极大似然算法在该群体之外的准确性可能会很差^[34]。如果 ML 算法在未充分描述的数据上运行不佳, 则很难确定原因所在^[35-37]。因此, 对所述群体的了解对于 ML 算法的开发非常重要, 对所述群体评估更是如此。由于在一些没有基础设施来建立合适的数据库而又希望开发相关算法应用程序来解决健康问题的国家, 往往只能将访问公共数据库作为第一选择, 因此从全球角度来看, 这是一个十分重要的问题。

疾病代表性不足也是一个问题。数据库可能反映与其来源国或地区特别相关的疾病; 来自常规临床护理的数据将反映就诊人群的疾病患病率; 队列数据将反映该研究的纳入标准, 但也反映该特定国家或地区的卫生优先事项。每个数据库中的疾病类型及其在数据库中的患病率将影响该数据库在全球其他国家或地区的通用性。值得注意的是, 大多数数据库中含有大量的糖尿病视网膜病变、青光眼和年龄相关性黄斑变性的数据。这三种疾病加起来估计占全球失明的 15% 和中重度视力损害的 5%, 而白内障 (4 个数据库)、沙眼 (1 个数据库) 和屈光不正 (3 个数据库) 等导致了 53% 的失明和 79% 的中重度视力损害^[38]。这种不匹配可能归因于许多因素, 包括成像技术在疾病管理中的相对重要性、针对最常见疾病 (如糖尿病视网膜病变) 的完善筛查计划以及特定研究领域的可用资金等。与白内障、沙眼和屈光不正相比, 糖尿病视网膜病变、青光眼和年龄

相关性黄斑变性更常作为标准护理的一部分来进行眼科成像检查。如果潜在的基于成像的解决方案能够通过任务共享的方式改善非专家团队对白内障、屈光不正等疾病患者的治疗和管理,那么也许需要有针对性的全球努力来支持这些疾病领域在眼科图像数据库中的发展。

本指南中确定的公开可用数据库在全球分布不均。172个国家(相当于近35亿人,或全球人口的45%)没有已知的公开眼科图像数据库。如果考虑到受惠人群,其比例甚至更低。20个国家(占全球人口的54%)的58个数据库有健康眼的数据;6个国家(占全球人口的44%)的15个数据库有年龄相关性黄斑变性的数据;14个国家(占全球人口的50%)的35个数据库有糖尿病视网膜病变的数据。从数据中推断出的结果不能被假定为在不同的人群中普遍存在,也可能无法在其他人群中使用。

4 眼科图像数据库的不足

几乎所有数据库均存在不同程度临床信息报告不足的问题。大多数(74%)数据库没有报告包括年龄、性别和种族在内的人口统计学数据。此外,仅有15%的数据库报告了纳入和排除标准。因此,我们目前还不清楚数据中是否有适当的人口群体代表,研究人员使用这些数据而获得的研究结果的适用性将受到严重限制。

对于带有图像标签的数据库(例如诊断标签或功能标签),标注过程也缺乏详细报告。许多ML模型的位点假设是在认定标签正确的基础上提出的,因此确保标签的准确性非常重要。标签的准确性与进行标签标注的学者的经验和专业知识储备、多个标注标签学者的共识以及如何解决争论都密切相关^[28]。在报告这些信息的少数数据库中,标签标注者有医学生也有眼科专科医生,但在大多数情况下,不能确定进行标签标注者的专业水平。不可否认,对公共数据库数据进行详细的标注可能非常有用,补充其他信息可能会大大提高数据的实用性,也可能实现多个数据库的合并。

5 研究的局限性

我们进行这项工作的主要目的是系统整理公开的眼科图像数据库。在进行这项工作时,我们使用了多种检索方式,同时我们也发现通过论文等学术出版物并不能找到所有相关数据库。此外,我们验证了所有数据库可以开放访问的声明,以真正确定这些数据库开放访问的程度以及用户体验。我们采取合理措施获取数据,同时对其内容进行检查和验证。在这一过程中我们遇到了一些阻碍,即许多声称开放访问的数据库在访问时却存在多重阻碍。因此我们增添了关键元数据的可用程度。

我们的研究有几个局限:首先,只对Google数据

集搜索和Google搜索引擎返回的前10页结果进行了筛选。根据可获得的信息表明,Google数据集搜索中数据库排名是根据元数据质量、引用数量以及其他因素的组合确定的,但具体不详^[29]。由于Google搜索引擎的性质,我们无法更新搜索结果,而MEDLINE的搜索结果(2019年12月取得)在2020年5月进行了更新。但是考虑到数据库公开与研究发表并在MEDLINE上可见之间可能存在时间差,我们也不能确定之前所有的数据库均被包括在内。其次,这一领域正在快速发展,这次审查只是基于一个时间点的回顾。与学术论文不同,数据库可以多次进行编辑、更新或删除,因此这些数据的可用性也会随时间而变化。再次,还有其他数据库检索方式,比如Kaggle,但本指南并未对这些来源进行探讨(Google数据集搜索工具显示其索引了Kaggle)。后期条件允许我们将采用更多的搜索方式。最后,我们的目标是确定开放获取的数据库,如果某些数据库没有将自身定义为开放访问,我们也没有将其纳入研究。如果对访问数据有一些要求,我们会采取合理措施获取数据,比如按要求向数据的作者或所有者发送了电子邮件,但我们没有签订许可协议或提供伦理委员会的批准。对这些数据库的访问权限进行规范化是很合理的行为,但这样也会给数据库的运行机构带来额外的负担。在研究中,我们发现一些受监管访问的数据库质量可能更好,而且这种更严格的访问可能侧面反映了重视质量和元数据报告相关的管理流程。但是在研究中,研究人员十分看重数据库访问的便利性和速率,这可能导致这些访问较严格的高质量数据库被忽视。

6 总结

公用数据库是医疗保健研究和创新的潜在宝贵资源。然而由于元数据不完整,使这些数据库存在可见性差、访问受限或可用性有限等缺陷。元数据不完整包括缺乏评估来源、数据质量评价和抽样人群多样性所必需的关键参数。由于少数数据库知名度高或易于访问而被频繁使用,研究人员存在使用小而偏斜的数据池的危险,这可能会导致严重的偏倚。在运用人工智能进行研究时,可能会导致普遍性差,在环境和人群之间转移时可能会出现性能不佳甚至不可用的风险。我们认为这是一种数据贫困,数据库(公共和其他数据库)的代表性较低将限制一些个人甚至整个地区人口从数字健康解决方案和人工智能系统中受益。本指南旨在努力提高公用数据库的可见性、可访问性和可用性,以支持目前人工智能或机器学习领域的研究和创新。

形成指南专家组成员

执笔专家:

邵毅 南昌大学第一附属医院

杨卫华	暨南大学附属深圳眼科医院
陈蔚	温州医科大学附属眼视光医院
谭钢	南华大学附属第一医院
杨文利	首都医科大学附属北京同仁医院
黄永志	四川大学附属华西医院
徐积兄	南昌大学第一附属医院
张慧	昆明医科大学第一附属医院
黄锦海	复旦大学附属眼耳鼻喉科医院
赵慧	上海交通大学医学院附属第一人民医院

参与起草的专家(按姓氏拼音排列)

白燕慧	郑州大学第一附属医院
陈波	四川省人民医院
陈俊	江西中医药大学
陈序	荷兰马斯特里赫特大学
陈雨	上海大学
程琪	北京心联光电科技有限公司
旦增仁青	西藏自治区藏医院 西藏眼科中心
邓宇晴	中山大学中山眼科中心
巩倩文	温州医科大学附属眼视光医院
韩忆	厦门大学眼科研究所
贺佳	济宁医学院
何媛	西安医学院第二附属医院
胡守龙	首都医科大学附属北京儿童医院
黄彩虹	厦门大学眼科研究所
黄晓明	四川眼科医院
黄歆	江西省人民医院
贾艳妮	山东省眼科医院
姜楠	温州医科大学附属眼视光医院
黎颖莉	南方医科大学附属珠江医院
李洁	温州医科大学附属眼视光医院
李娟	陕西省眼科医院
李乃洋	中山市人民医院
李清坚	厦门大学眼科研究所
李祥春	天津市肿瘤医院
李植源	郴州市第一人民医院
李中文	温州医科大学附属宁波市眼科医院
廖许琳	香港中文大学
林松	天津医科大学眼科医院
刘光辉	福建中医药大学附属人民医院
刘华	锦州医科大学第三医院
刘秋平	南华大学附属第一医院
刘玉姬	江西省广丰人民医院
刘祖国	厦门大学眼科研究所
娄岩	中国医科大学
欧阳君	九江市第一人民医院
欧阳维杰	厦门大学眼科研究所
潘红飙	中国科学技术大学附属第一医院
彭娟	广州医科大学附属第二医院
彭志优	南昌大学第一附属医院

邱昆良	汕头国际眼科中心
任胜卫	河南省立眼科医院
容蓉	中南大学湘雅医院
邵婷婷	复旦大学附属眼耳鼻喉科医院
石文卿	复旦大学附属金山医院
宋艳	新疆维吾尔自治区医院
苏婷	武汉大学人民医院
苏兆安	浙江大学医学院附属第二医院
谭叶辉	南昌大学第一附属医院
唐丽颖	厦门大学附属中山医院
佟莉杨	温州医科大学附属宁波市眼科医院
王海燕	陕西省眼科医院
王乐韵	中科院动物研究所
王桑	北京茗视光眼科
王晓刚	山西省眼科医院
王雪林	江西医专第一附属医院
王岩	内蒙古医科大学附属医院
王怡欣	英国卡迪夫大学
魏雁涛	中山大学中山眼科中心
吴洁丽	长沙爱尔眼科医院
许言午	北京百度网讯科技有限公司
杨启晨	四川大学附属华西医院
杨舒	昆明市第一人民医院
杨青华	解放军总医院
杨玮枫	海南大学
杨于力	陆军军医大学第一附属医院
俞益丰	南昌大学第二附属医院
余瑶	南昌大学第一附属医院
袁晴	九江市第一人民医院
张冰	杭州市儿童医院
张妍	吉林大学第二医院
张艳艳	温州医科大学附属宁波市眼科医院
张雨晴	重庆大学附属第二医院
赵一天	中科院宁波医工所
郑钦象	温州医科大学附属眼视光医院
钟菁	中山大学中山眼科中心
朱佩文	复旦大学附属眼耳鼻喉科医院

利益冲突:

所有作者均声明不存在利益冲突。本指南的制定未接受任何企业的赞助。

指南声明:

所有参与本指南制定的专家均声明,坚持客观的立场,以专业知识、研究数据和临床经验为依据,经过充分讨论,全体专家一致同意后形成本共识,本指南为中国医药教育协会数字影像与智能医疗专委会眼科影像与智能医疗分会(筹)和中国医药教育协会智能医学专委会智能眼科分会部分专家起草。

免责声明:

本指南的内容仅代表参与制定的专家对部分公

开可用的眼科图像数据库的使用指导意见,供临床医师参考。尽管专家们进行了广泛的意见征询和讨论,但仍有不全面之处。本指南所提供的建议并非强制性意见,与本指南不一致的做法并不意味着错误或不当。临床实践中仍存在诸多问题需要探索,正在进行和未来开展的临床研究将提供进一步的证据。随着临床经验的积累和治疗手段的涌现,未来需要对本指南定期修订、更新,为患者带来更多临床获益。

参考文献

[1]

PARIK KIM H R B, GDOWSKI A, PATT D A, HERTLER A, MERMEL C, BEKELMAN J E. Using big data and predictive analytics to determine patient risk in oncology[J]. *Am Soc Clin Oncol Educ Book*, 2019, 39: e53-e58.

[2]

WONG Z S Y, ZHOU J, ZHANG Q. Artificial intelligence for infectious disease big data analytics[J]. *Infect Dis Health*, 2019, 24(1): 44-48.

[3]

KIM H E, KIM H H, HAN B K, KIM K H, HAN K, NAM H, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study[J]. *Lancet Digit Health*, 2020, 2(3): e138-e148.

[4]

ESTEVA A, KUPREL B, NOVOA R A, KO J, SWETTER S M, BLAU H M, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. *Nature*, 2017, 542(7639): 115-118.

[5]

KERMANY D S, GOLDBAUM M, CAI W, VALENTIM C C S, LIANG H, BAXTER S L, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning[J]. *Cell*, 2018, 172(5): 1122-1131.

[6]

LIU X, FAES L, KALE A U, WAGNER S K, FU D J, BRUYN-SEELS A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis[J]. *Lancet Digit Health*, 2019, 1(6): e271-e297.

[7]

MCKINNEY S M, SIENIEK M, GODBOLE V, GODWIN J, ANTROPOVA N, ASHRAFIAN H, et al. International evaluation of an AI system for breast cancer screening[J]. *Nature*, 2020, 577(7788): 89-94.

[8]

ARCADU F, BENMANSOUR F, MAUNZ A, WILLIS J, HASKOVA Z, PRUNOTTO M. Deep learning algorithm predicts diabetic retinopathy progression in individual patients[J]. *NPJ Digit Med*, 2019, 2: 92.

[9]

YIM J, CHOPRA R, SPITZ T, WINKENS J, OBIKA A, KELLY C, et al. Predicting conversion to wet age-related macular degeneration using deep learning[J]. *Nat Med*, 2020, 26(6): 892-899.

[10]

MILEA D, NAJJAR R P, ZHUBO J, TING D, VASSENEIX C, XU X, et al. Artificial intelligence to detect papilledema from ocular fundus photographs[J]. *N Engl J Med*, 2020, 382(18): 1687-1695.

[11]

DEFAUW J, LEDSAM J R, ROMERA-PAREDES B, NIKOLOV S, TOMASEV N, BLACKWELL S, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease[J]. *Nat Med*, 2018, 24(9): 1342-1350.

[12]

TING D S W, PASQUALE L R, PENG L, CAMPBELL J P, LEE A Y, RAMAN R, et al. Artificial intelligence and deep learning in ophthalmology[J]. *Br J Ophthalmol*, 2019, 103(2): 167-175.

[13]

ABRAMOFF M D, LAVIN P T, BIRCH M, SHAH N, FOLK J C. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices[J]. *NPJ Digit Med*, 2018, 1: 39.

[14]

WEN J C, LEE C S, KEANE P A, XIAO S, ROKEM A S, CHEN P P, et al. Forecasting future humphrey visual fields using deep learning[J]. *PLoS One*, 2019, 14(4): e0214875.

[15]

DECENCIÈRE E, ZHANG X, CAZUGUEL G, LAÏ B, COCHENER B, TRONE C, et al. Feedback on a publicly distributed image database: the messidor database[J]. *Image Ana-*

ly Stereol

, 2014, 33(3): 231-234.

[16]

NIEMEIJER M, STAAL J J, GINNEKEN B V, LOOG M, ABRAMOFF M D. DRIVE: digital retinal Images for vessel extraction[DB/OL]. [2020-10-15]. <https://drive.grand-challenge.org>.

[17]

GULSHAN V, PENG L, CORAM M, STUMPE M C, WU D, NARAYANASWAMY A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs[J]. *JAMA*, 2016, 316(22): 2402-2410.

[18]

DECENCIÈRE E, CAZUGUEL G, ZHANG X, THIBAUT G, KLEIN J C, MEYER F, et al. TeleOphta: machine learning and image processing methods for teleophthalmology[J]. *IRBM*, 2013, 34(2): 196-203.

[19]

BURTON M J, FAAL H B, RAMKE J, RAVILLA T, HOLLAND P, WANG N, et al. Announcing the lancet global health commission on global eye health[J]. *Lanc Glob Health*, 2019, 7(12): e1612-e1613.

[20]

Service U D. Data access policy[EB/OL]. [2022-10-15]. <https://www.ukdataservice.ac.uk/get-data/data-access-policy/open-data.aspx>.

[21]

MILEA D, AZMI S, REGINALD P, VERPILLAT P, FRANCOIS C. A review of accessibility of administrative healthcare databases in the Asia-Pacific region[J]. *J Mark Access Health Policy*, 2015, 3: PMID27123180.

[22]

Biobank[DB/OL]. [2022-10-15]. <http://www.ukbiobank.ac.uk/scientists-3>.

[23]

Diabetic Retinopathy Detection. Diabetic retinopathy detection[M/OL]. [2022-10-15]. <https://www.kaggle.com/c/diabetic-retinopathy-detection>; California Healthcare Foundation.

[24]

FUSEK R. MRL eye dataset[DB/OL]. [2022-10-15]. <http://mrl.cs.vsb.cz/eyedataset>.

[25]

SCANLON P H. The English national screening programme for diabetic retinopathy 2003 – 2016[J]. *Acta Diabetol*, 2017, 54(6): 515-525.

[26]

RIBEIRO L, OLIVEIRA C M, NEVES C, RAMOS J D, FERREIRA H, CUNHA-VAZ J. Screening for diabetic retinopathy in the Central Region of Portugal. Added value of automated ‘disease/no disease’ grading[J]. *Ophthalmologica*, 2015, 233(2): 96-103.

[27]

MASSIN P, CHABOUIS A, ERGINAY A, VIENS-BITKER C, LECLEIRE-COLLET A, MEAS T, et al. A telemedical network screening system for diabetic retinopathy in the Île-de-France[J]. *Diabetes Metab*, 2008, 34(3): 227-234.

[28]

KRAUSE J, GULSHAN V, RAHIMY E, KARTH P, WIDNER K, CORRADO G S, et al. Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy[J]. *Ophthalmology*, 2018, 125(8): 1264-1272.

[29]

BLOG G A. Building Google dataset search and fostering an open data ecosystem[EB/OL]. [2022-10-15]. <https://ai.googleblog.com/2018/09/building-google-dataset-search-and.html>.

[30]

CHALAKKAL R J, ABDULLA W. Automatic segmentation of retinal vasculature; ICASSP, IEEE international conference on acoustics, speech and signal processing-proceedings[C]. New Orleans; ICASSP, 2017.

[31]

XU X, DING W, WANG X, CAO R, ZHANG M, LÜ P, et al. Smartphone-based accurate analysis of retinal vasculature towards point-of-care diagnostics[J]. *Sci Rep*, 2016, 6: 34603.

[32]

GEBRU T, MORGENSTERN J, VECCHIONE B, VAUGHAN J W, WALLACH H, DAUMÉE H, III, et al. Datasheets for datasets: Proceedings of the 5th Workshop on Fairness, Accountability, and Transparency in Machine Learning[C]. Sweden: Stockholm, 2018.

[33]

HRYNASZKIEWICZ I, NORTON M L, VICKERS A J, ALTMAN D G. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers[J]. *BMJ*, 2010, 340: e181.

[34]

YIM J, CHOPRA R, SPITZ T, WINKENS J, OBIKA A, KELLY C, et al. Predicting conversion to wet age-related macular degeneration using deep learning[J]. *Nat Med*, 2020, 26(6): 892-899.

[35]

LARRAZABAL A J, NIETO N, PETERSON V, MILONE D H,

FERRANTE E. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis[J]. *Proc Natl Acad Sci U S A*, 2020, 117 (23) : 12592-12594.

[36] ADAMSON A S, SMITH A. Machine learning and health care disparities in dermatology[J]. *JAMA Dermatology*, 2018, 154 (11) : 1247-1248.

[37] CHEN I Y, JOSHI S, GHASSEMI M. Treating health disparities with artificial intelligence[J]. *Nat Med*, 2020, 26 (1) : 16-17.

[38] FLAXMAN S R, BOURNE R R A, RESNIKOFF S, ACKLAND P, BRAITHWAITE T, CICINELLI M V, *et al.* Global causes of blindness and distance vision impairment 1990 – 2020 : a systematic review and meta-analysis[J]. *Lancet Glob Health*, 2017, 5 (12) : e1221-e1234.

Guidelines for the use of global public databases on ophthalmic images (2022)

Digital Imaging and Intelligent Medical Branch of China Medical Education Association, Intelligent Medical Special Committee of China Medical Education Association

Corresponding author: SHAO Yi, E-mail: freebee99@163.com; LIU Zuguo, E-mail: zuguoliu@xmu.edu.cn; SHAO Tingting, E-mail: mamabq12@163.com

[Abstract] Publicly accessible medical data are valuable for digital health research. Some public databases containing ophthalmic images are commonly used in machine learning studies, but there have been no criteria for standard and efficient use of these databases. These guidelines aim to identify some publicly available ophthalmic image databases with a detailed description of the disease types, image sources and imaging modalities. Using MEDLINE, Google’s search engine, and Google dataset search, we identified 94 open-access databases containing 507 724 images and 125 videos from 122 364 patients. These guidelines provide support and reference for ophthalmic database-related research. In addition, increasing disparities in different populations and disease groups in the databases were found. The improved metadata reports enable researchers to access the most appropriate databases as required and maximize the role of image resources.

[Key words] ophthalmic images; ophthalmic databases; image databases; guidelines for use