

· 临床研究 ·

基于深度学习的超声生物显微镜全景图像自动测量模型的构建与评估

朱剑¹ 颜瑜琳¹ 蒋维艳¹ 张少维¹ 牛晓光² 胡孝³ 郑碧清³ 杨燕宁¹¹武汉大学人民医院眼科中心, 武汉 430060; ²武汉大学附属爱尔眼科医院, 武汉 430000; ³楚精灵医疗科技有限公司, 武汉 430000

通信作者: 杨燕宁, Email: ophyyn@163.com

【摘要】 目的 构建一种基于深度学习的超声生物显微镜(UBM)全景图像自动测量的模型并对其进行评估。**方法** 采用诊断试验研究方法, 收集 2021 年 2 月至 2023 年 3 月在武汉大学人民医院眼科中心接受有晶状体眼后房型人工晶状体(ICL)植入术的 372 例 684 眼患者术前 UBM 检查结果, 共获得 1 368 张 UBM 全景图像, 建立图像数据库。并将数据集分为训练集 760 张、验证集 86 张和内部测试集 522 张。由 3 位眼科医师组成的专家小组对图像进行标记。利用 UNet++ 网络对图像中角膜、晶状体和虹膜等组织进行自动分割, 并开发一种图像处理技术和几何定位算法, 该算法可自动定位瞳孔直径(PD)、前房深度(ACD)、角-角距离(ATA)和沟-沟距离(STS)的解剖标志点并完成上述参数的测量。另选取来自武汉大学附属爱尔眼科医院 135 例 240 眼的 480 张 UBM 全景图像作为外部测试集, 进一步评估该模型在其他中心的性能。对模型测量结果与专家小组和 Pentacam 测量结果的一致性进行评估。最后从外部测试集中随机选择 150 张图像用于人机对比, 进一步评估模型性能。**结果** 在内部和外部测试集中, 以人工标注为参考标准, 模型的平均 Dice 系数均不小于 0.882; 解剖标志点定位欧氏距离差异在 250 μm 内的占比不小于 95.65%; PD、ACD、ATA 和 STS 的组内相关系数(ICC)均不小于 0.958, 平均相对误差均不大于 2.407%。以 Pentacam 测量结果为参考标注, 在内部测试集和外部测试集中, PD 测量结果的 ICC 分别为 0.540 和 0.466, ACD 测量结果的 ICC 分别为 0.946 和 0.908。人机对比中, 该模型测量结果与资深专家测量结果之间的 ICC 均不低于 0.969。**结论** 基于深度学习的模型可通过 ICL 植入术患者术前的 UBM 全景图像自动测量眼前节参数, 一致性与资深专家类似, 效率高于资深专家。在 ACD 的测量上, 模型测量结果与 Pentacam 测量结果具有良好的一致性。

【关键词】 人工智能; 深度学习; 超声生物显微镜; 有晶状体眼后房型人工晶状体**基金项目:** 国家自然科学基金(82371023); 湖北省重点研究计划(2020BCB055); 武汉大学人民医院交叉创新人才项目(JCRCZN-2022-007); 爱尔眼科集团科研基金项目(AGK2309D07)

DOI: 10.3760/cma.j.cn115989-20240618-00160

Construction and evaluation of automatic measurement model of panoramic ultrasound biomicroscopy images based on deep learning

Zhu Jian¹, Yan Yulin¹, Jiang Weiyan¹, Zhang Shaowei¹, Niu Xiaoguang², Hu Xiao³, Zheng Biqing³, Yang Yanning¹¹Eye Center, Renmin Hospital of Wuhan University, Wuhan 430060, China; ²Aier Eye Hospital of Wuhan University, Wuhan 430000, China; ³Wuhan EndoAngel Medical Technology Company, Wuhan 430000, China

Corresponding author: Yang Yanning, Email: ophyyn@163.com

[Abstract] Objective To develop and evaluate a deep learning-based automatic measurement model for panoramic ultrasound biomicroscopy (UBM) images. **Methods** A diagnostic test study was conducted. Preoperative UBM examination results of 372 patients who underwent implantable collamer lens (ICL) implantation were collected at the Eye Center of Renmin Hospital of Wuhan University between February 2021 and March 2023. A total of 1 368 panoramic UBM images were obtained to establish an image database. The dataset was divided into a training set (760 images), a validation set (86 images) and an internal test set (522 images). An expert panel consisting of three ophthalmologists annotated the images. The UNet++ network was used to automatically segment anterior segment tissues, such as the cornea, lens and iris. In addition, image processing techniques and geometric localization algorithms were developed to automatically identify the anatomical landmarks of pupil diameter (PD), anterior chamber depth (ACD), angle-to-angle distance (ATA) and sulcus-to-sulcus distance (STS) to complete the measurement of these parameters. Additionally, 480 panoramic UBM images of 135 patients (240 eyes) from Aier Eye Hospital of Wuhan University were used as an external test set to further evaluate the performance of the model in

different centers. The consistency between the measurements from the model and expert panel, the Pentacam system was assessed. Finally, 150 images were randomly selected from the external test set for a human-machine comparison to further evaluate the model's performance. This study adhered to the Declaration of Helsinki. The study protocol was approved by the Ethics Committee of Renmin Hospital of Wuhan University (No. WDRY-2022-K109) and Aier eye Hospital of Wuhan University (No. 2023IRBKY120903). Written informed consent was obtained from each subject. **Results** In the internal test dataset and external test dataset, with manual labeling as the reference standard, the model achieved a mean Dice coefficient of not less than 0.882. At least 95.65% of the anatomical landmark localization results had Euclidean distance differences within 250 μm . The intraclass correlation coefficients (ICCs) for the measurements of PD, ACD, angle-to-angle ATA, and STS were at least 0.958, with mean relative errors not exceeding 2.407%. With the Pentacam measurements as the reference standard, the ICCs for PD in the internal and external test sets were 0.540 and 0.466, respectively, while the ICCs for ACD were 0.946 and 0.908, respectively. In the human-machine comparison, the ICCs between the model's measurements and those of senior experts were all not lower than 0.969. **Conclusions** The deep learning-based model can automatically measure anterior segment parameters from preoperative panoramic UBM images of patients undergoing ICL surgery. The model demonstrates a consistency comparable to that of senior experts, while providing higher efficiency. In terms of ACD measurement, the model shows good agreement between the measurements obtained from the model and Pentacam system.

[Key words] Artificial intelligence; Deep learning; Ultrasound biomicroscopy; Implantable collamer lens

Fund program: National Natural Science Foundation of China (82371023); Major Project of Research and Development Program in Hubei Province (2020BCB055); Interdisciplinary Innovative Talents Foundation from Renmin Hospital of Wuhan University (JCRCZN-2022-007); Aier Ophthalmology Group Research Fund Project (AGK2309D07)

DOI:10.3760/cma.j.cn115989-20240618-00160

超声生物显微镜 (ultrasound biomicroscopy, UBM) 扫描频率可达 50~100 MHz, 可以获得分辨率高达 50 μm 的眼部实时图像, 是目前唯一能够对虹膜后结构实时显示并实施定量测量的设备, 因此被广泛用于青光眼患者前房角形态的观察和有晶状体眼后房型人工晶状体 (implantable collamer lens, ICL) 植入术患者的术前评估。然而, UBM 在临床应用中也存在挑战, 一方面, UBM 检查是一项较耗时的接触性检查, 从而可能影响测量精确度; 另一方面, UBM 图像的判断和测量高度依赖于操作员的临床经验, 使得检查结果容易受到操作者主观影响。随着 ICL 植入矫正屈光不正的普及, 提高 UBM 在 ICL 术前检查中的准确性和效率具有重要临床意义。

近年来, 伴随着算力的提升以及医疗数据的不断积累, 人工智能在眼科医疗领域的应用日趋增多^[1], 常见的应用场景包括糖尿病视网膜病变和早产儿视网膜病变的辅助诊断和评估等^[2-3]。与人类相比, 人工智能在读取和分类图像方面具有显著优势, 成为临床实践中辅助诊断某些疾病的有力工具^[1-3]。在现有 UBM 图像识别模型中, 为了降低组织注释和自动分割的难度, 角膜、虹膜、巩膜和睫状体常被标记为一个类别^[4], 仅有少量模型能够根据眼部 UBM 图像中不同组织的解剖形态特征及其相邻关系对组织进行自动分割及识别^[5-6]。现有针对 UBM 图像的辅助读图及测量研究主要集中在青光眼方面, 训练模型自动定位巩

膜突, 并根据模型识别的巩膜突自动测量前房角相关参数, 从而对青光眼的病情进行评估, 而关于眼前节参数量化方面的模型研究少见。

本研究旨在开发一种基于人工智能的眼部 UBM 全景图像自动测量系统, 实现自动分割眼部 UBM 全景图像中的不同组织, 如角膜、虹膜、晶状体等, 准确定位瞳孔直径 (pupil diameter, PD)、前房深度 (anterior chamber depth, ACD)、角-角距离 (angle-to-angle distance, ATA) 和沟-沟距离 (sulcus-to-sulcus distance, STS) 的解剖标志点, 并进行相应的生物测量, 以期为医疗专业人员提高临床工作效率, 减少检查人员个人因素对测量结果的影响提供辅助。

1 资料与方法

1.1 一般资料

1.1.1 UBM 图像来源 采用诊断试验研究方法, 回顾性收集 2021 年 2 月至 2023 年 3 月在武汉大学人民医院眼科中心进行 ICL 植入术的 372 例 684 眼前 UBM 检查结果, 共获得 1 368 张 UBM 全景图像, 建立数据库。以 2022 年 5 月 31 日为时间节点将 1 368 张 UBM 全景图像进行划分, 在该时间节点之前共计 846 张图像, 按 9:1 进行随机划分, 其中 760 张图像作为训练集, 剩余 86 张图像作为验证集; 在该时间节点后的 522 张图像作为内部测试集。同时以 2023 年 4—5 月在武汉大学附属爱尔眼科医院进行 ICL 植入术 135 例

240 眼患者术前检查的 UBM 全景图像 480 张作为外部测试集。所有图像均由具有 20 年工作经验的资深眼科 UBM 检查医师使用 UBM (SW-2100, 天津索维公司) 按照操作规范^[7] 采集获得。所有图像经匿名处理后再用于标记及模型训练, 本研究遵循《赫尔辛基宣言》, 研究方案经武汉大学人民医院伦理委员会 (批文号: WDRY-2022-K109) 和武汉大学附属爱尔眼科医院伦理委员会审核批准 (批文号: 2023IRBKY120903), 并在医学研究登记备案信息系统中备案 (备案编号: MR-42-24-016241)。

1.1.2 UBM 全景图像筛选 临床 UBM 扫描模式包括全景模式 (超声探头频率 35 MHz, 扫描范围 15.25 mm×9.50 mm) 和半景模式 (超声波探头频率 50 MHz, 扫描范围 9.75 mm×6.00 mm)。由于只有全景图像才能在单张图像中显示前房、晶状体和双侧前房角, 从而测量 ATA 和 STS, 因此在本研究中, 仅选择在全景模式下拍摄的眼部 UBM 图像作为数据集。根据研究需要及操作规范^[8] 制定以下 UBM 全景图像的纳入标准: (1) 角膜前表面和后表面清晰可见; (2) 虹膜色素上皮层和晶状体前表面反射线清晰可见; (3) 晶状体前表面和虹膜后表面相切; (4) 每张全景图像中包含双侧前房角。排除标准: 由运动伪影、眼前节不完整或相关结构异常引起的非正常 UBM 图像。所收集的 UBM 全景图像为 BMP 格式, 大小为 1.91 MB, 尺寸为 1 024 像素×655 像素。

1.2 方法

1.2.1 UBM 全景图像数据集的建立 使用训练集和验证集对分割模型进行构建和优化, 使用内部测试集对模型性能进行评估。在完成分割模型的构建和评估后, 使用内部测试集的 UBM 全景图像开发并验证一种基于 Python 编程语言的图像处理程序及几何定位算法, 该算法可自动定位 PD、ACD、ATA 和 STS 的解剖标志点并完成上述参数的测量。使用外部测试集进一步评估该模型的性能。所有图像均使用以角膜中心为扫描中心在水平位置 (3:00~9:00) 或垂直位置 (6:00~12:00) 进行扫描获得。

1.2.2 UBM 图像的标注 稳健的参考标准是深度学习模型训练的基础, 为了确保注释结果的稳健性, 本研究建立由 2 名具有 10 年临床经验的眼科专家和 1 名具有 20 年临床经验的资深眼科专家组成的专家小组完成 UBM 图像标注。UBM 图像的标注过程包括 2 个步骤: (1) 角膜、晶状体、虹膜、角巩膜缘和睫状体的标注 2 名眼科专家使用 VGG 2.0.0 软件 (英国牛津大学) 对图像进行标注, 并由资深眼科专家进行检查和

校正。不同组织的边界根据角膜、晶状体、虹膜、角巩膜缘和睫状体结构的解剖特征和相邻关系确定。UBM 全景图像被分为 6 个区域, 即角膜中央区域 (a)、晶状体区域 (b)、双侧虹膜区域 (c 和 d)、双侧角巩膜缘睫状体区域 (e 和 f) (图 1A)。(2) 解剖标志点的标注 根据 PD、ACD、ATA 和 STS 的定义, 2 名眼科专家使用 VGG 软件对图像中相应的解剖标志点进行标注。使用标记坐标的平均值作为参考标准, 并由资深眼科专家进行检查和校正 为了便于记录, PD、ACD、ATA 和 STS 的起点和终点分别指定为点 1 和 2、点 3 和 4、点 5 和 6、点 7 和 8 (图 1B)。

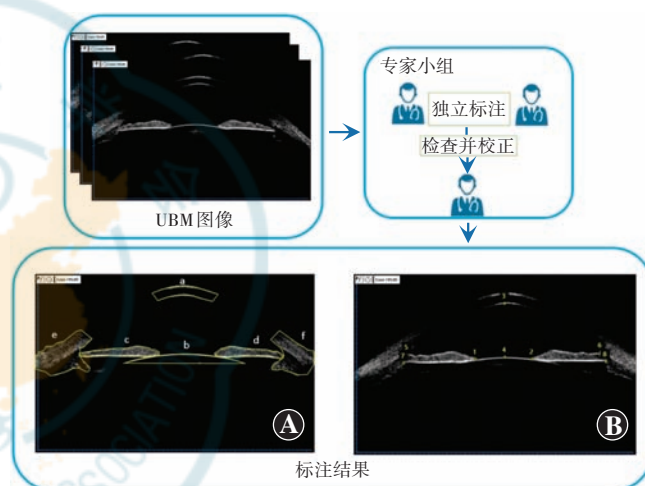


图 1 UBM 图像标注过程 A: 标记 UBM 图像中不同的组织 a 为角膜中央区域, b 为晶状体区域, c 和 d 为双侧虹膜区域, e 和 f 为双侧角巩膜缘睫状体区域 B: 标记 UBM 图像中解剖标志点 点 1 和 2、点 3 和 4、点 5 和 6、点 7 和 8 分别为 PD、ACD、ATA 和 STS 的起点和终点 UBM: 超声生物显微镜

Figure 1 UBM image labeling A: Labeling of different tissues in UBM images a represented the central corneal region, b represented the lens region, c and d represented the bilateral iris regions, and e and f represented the bilateral scleral spur-ciliary body regions B: Labeling of anatomical landmarks in UBM images Points 1 and 2, 3 and 4, 5 and 6, and 7 and 8 represented the start and end points of PD, ACD, ATA, and STS, respectively UBM: ultrasound biomicroscopy

1.2.3 模型的开发 计算机算法使用 Python3.9.12 编程语言编写, 开源 Keras 库 (2.9.0 版) 和 TensorFlow 库 (2.9.1 版) 作为后端。该模型在配备 NVIDIA GeForce RTX 3060 GPU (12GB GPU 内存) 的服务器上进行了训练。模型工作流程见图 2。

1.2.3.1 眼前节组织的分割与识别 通过深度卷积神经网络 (deep convolutional neural network, DCNN) 对 UBM 全景图像中的不同组织进行分割和识别, 以便于后续眼前节参数解剖标志点的定位和相应参数的测量。由于 UNet++ 网络^[8] 能够在小型数据集中对医学图像实现良好的分割性能, 本研究中 Unet++ 网络被用于 UBM 图像中不同眼部组织的分割。UBM 全景图像

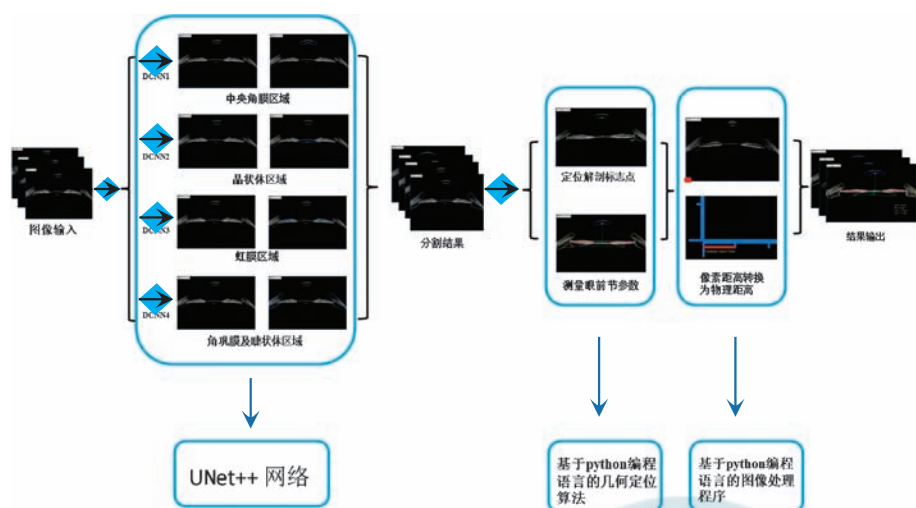


图 2 模型工作流程图 DCNN:深度卷积神经网络

Figure 2 Workflow of the model DCNN:deep convolutional neural network

中的眼部组织被分为 4 个区域,即角膜中央区域、晶状体区域、双侧虹膜区域和双侧角巩膜缘睫状体区域。分割模型结果的代表性图像见图 3。在训练 UNet++ 模型进行结构识别时,先统一将图像转化为 512 像素×512 像素再进行后续训练,并采用随机旋转、水平和垂直平移、错切、放大、水平和垂直翻转等数据增强方法以提升模型泛化能力。训练优化器使用 Adam,学习率为 0.000 1,损失函数为 binary cross entropy。训练中采用 early stopping 方法以避免过拟合,即监测验证集的损失函数值,当连续 10 轮该数值不再降低时,训练结束。训练曲线见图 4。

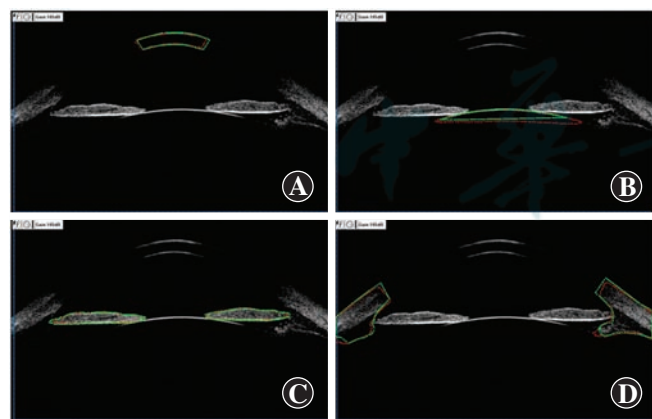


图 3 眼前节组织分割过程 A:DCNN1,分割识别中央角膜区域 B:DCNN2,分割识别晶状体区域 C:DCNN3,分割识别双侧虹膜区域 D:DCNN4,分割识别双侧角巩膜缘及睫状体区域 绿色区域表示手动标记的范围,红色区域表示深度学习模型识别的范围 DCNN:深度卷积神经网络

Figure 3 Anterior segment boundary segmentation A:DCNN1, recognizing and segmenting the central corneal region B:DCNN2, recognizing and segmenting the lens region C:DCNN3, recognizing and segmenting the bilateral iris region D:DCNN4, recognizing and segmenting the bilateral corneoscleral margin and ciliary body region The green area represented the manually labeled range and the red area represented the range recognized by the deep learning model DCNN: deep convolutional neural network

1.2.3.2 解剖标志点的自动定位 以第 1 步分割模型的结果为基础,开发基于 Python 语言的图像处理程序和几何定位算法实现对解剖标志点的自动定位,具体过程如下:(a) 双侧虹膜的最内侧点,分别标记为点 1 和 2;(b) 作线段 12 的中垂线与中央角膜区域的后表面和晶状体区域的前表面的交点分别为点 3 和 4;(c) 双侧虹膜区域的前表面与相应的角巩膜缘睫状体区域的交点为点 5 和 6;(d) 双侧虹膜区域的后表面与相应的角巩膜缘睫状体区域的

交点为点 7 和 8。图 5 展示了自动识别解剖标志点的代表性图像。所有解剖标志点的定位过程均遵循类似的方法,以点 8 的定位过程为例对前文提到的图像处理技术和几何定位算法进行说明。具体实现过程是首先获取右侧虹膜区域识别的黑白图和右侧角巩膜缘睫状体区域识别的黑白图,使用 cv2.bitwise_and 得到右侧虹膜区域和右侧角巩膜缘睫状体区域的交集,再使用 cv2.findContours 得到交集的轮廓点,最后识别这些轮廓点中具有最大 y 坐标的点,该点对应于点 8。

1.2.3.3 前房参数的测量 基于 1.2.3.2 识别的解剖标志点,采用平面坐标系中 2 个点间距离公式得到点 1 和 2、点 3 和 4、点 5 和 6 以及点 7 和 8 之间的距离(单位为像素),分别为 PD、ACD、ATA 和 STS。

1.2.3.4 像素距离和物理距离之间的转换 1.2.3.3 中测量结果是以像素为单位计算所得,为了方便描述称之为像素距离,但像素距离不能直接应用于临床,我们继续开发一种基于 Python 语言的图像处理程序将上述像素距离转换为物理距离。在每张 UBM 全景图像下方有 1 个比例尺(图 6A, B),比例尺中每一大格对应 1 mm,每一小格对应 0.25 mm,我们编写的图像处理程序可自动识别 UBM 图像下方比例尺上 0.25 mm 长度对应的像素长度值,从而建立像素长度与物理长度的对应关系。具体方法是以比例尺水平轴原点作为固定起始点,开始扫描红色水平线上每个像素的颜色值,直至检测到第 1 个非黑色像素。该第 1 个非黑色像素的 X 轴坐标减去起始位置 X 轴坐标的数值即对应于 0.25 mm 的像素数。经过计算全景模式中的 UBM 图像中 0.25 mm 对应 16.5 像素。根据上述对应关系将第 3 步测量得到的像素距离转换为物理距离,转换完成后将结果直接显示在 UBM 图像中(图 6C, D)。

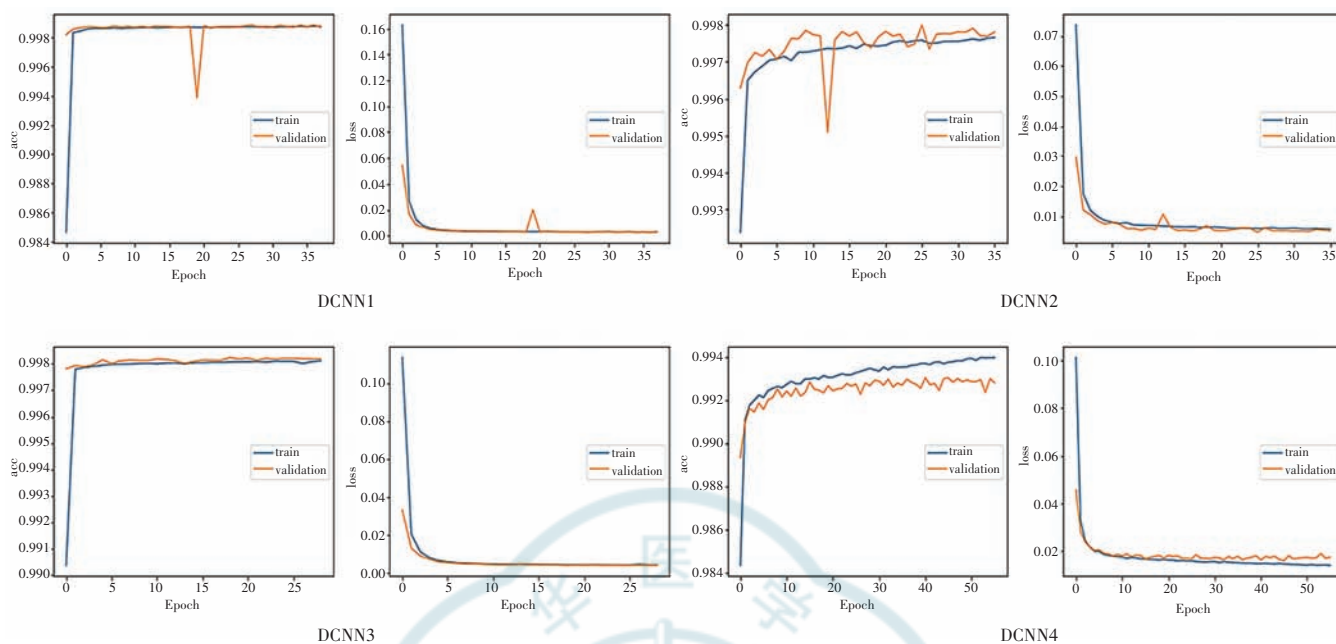


图4 DCNN1~DCNN4 训练曲线 DCNN:深度卷积神经网络

Figure 4 Training curves of the DCNN1-DCNN4 DCNN:deep convolutional neural network

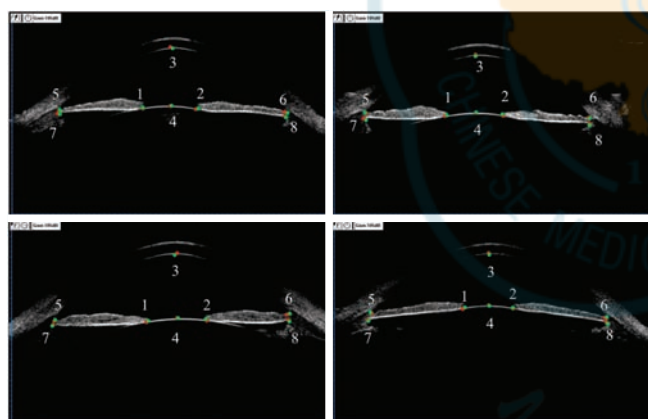


图5 专家小组标注(绿点)及模型识别(红点)的解剖标志点(点1~8)代表性图像 4幅图像为随机选择,旨在通过图像表明模型预测的解剖标志点和手工标注的解剖标志点

Figure 5 Representative images of anatomical landmarks (points 1~8) labeled by an expert panel (green dots) and identified by the model (red dots). The four images were randomly selected to indicate the anatomical landmark points predicted by the model and the manually labeled anatomical landmark points through the images.

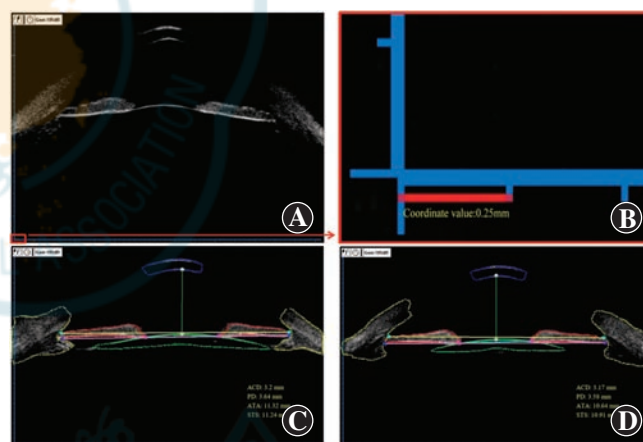


图6 长度单位转换及转换完成后模型输出结果 A:比例尺在UBM图像中位置(红色方框) B:图A中红色方框部分局部放大后的图像,以便显示更多细节 C、D:模型输出的测量结果

Figure 6 Length unit conversion and model output results after conversion A: Position of the scale bar (red box) in the UBM image B: Enlarged image within the red box in picture A to show more details C, D: Model output measurement results

1.2.4 模型的评估 (1)模型测量结果与专家小组测量结果的一致性评估 选取内部测试集中2022年6月1日至10月31日和外部测试集中2023年4月1—30日进行ICL植入手术患者的UBM全景图像对模型进行评估,最终内部测试集包含278张图像,外部测试集包含276张图像。以专家小组标注结果为参考标准,使用平均Dice系数评价模型的分割性能。使用模型预测的解剖标志点与手工标注解剖标志点间的欧式距离(Euclidean distance, ED),评估模型解剖标志点

的定位性能。使用组内相关系数(intra-class correlation coefficient, ICC)、平均相对误差(average relative error, ARE)和一致性界限(limits of agreement, LoA)评估眼前节参数模型与专家小组间测量结果的一致性。(2)模型测量结果与Pentacam(德国Oculus公司)测量结果的一致性评估 选取内部测试集中2022年11月1日至2023年3月31日和外部测试集中2023年5月1—31日进行ICL植入术患者的UBM全景图像对模型进行评估,最终内部测试集包含244张图像,外

部测试集包含 204 张图像。以 Pentacam 结果为参考标准,使用 ICC 评估模型与 Pentacam 测量结果的一致性。由于本研究选择的 4 个眼前节参数中 Pentacam 只能测量 ACD 和 PD,在这一部分只对 ACD 和 PD 2 个参数进行验证。(3)人机比较 采用随机数字法从外部测试集中随机选取 150 张 UBM 全景图像用于人机对比测试,2 名具有 20 年临床工作经验的资深屈光手术专家分别独立根据 UBM 图像测量 PD、ACD、ATA 和 STS,同时使用模型对该批次图像进行上述参数的测量,使用 ICC 评估模型测量结果与 2 名资深屈光手术专家测量结果的一致性。由同一位研究人员分别记录 2 名资深屈光手术专家和模型测量耗时,通过比较消耗时间评估资深屈光手术专家和模型的测量效率。

1.3 统计学方法

采用 SPSS 25.0 统计学软件(美国 IBM 公司)进行统计分析。计量资料数据经 Shapiro-Wilk 检验证实符合正态分布,以 $\bar{x} \pm s$ 表示。采用 ICC 用于评估模型预测与专家小组标记结果之间以及模型预测与 Pentacam 测量结果间的一致性。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 模型测量与专家小组测量结果的一致性评估

2.1.1 分割模型性能评估

使用 DCNN1~4 对来自内部测试集和外部测试集的 UBM 全景图像进行分割,图像被分为角膜中央区域、晶状体区域、双侧虹膜区域以及双侧角膜巩膜缘和睫状体区域。在内部测试集中,模型的平均 Dice 系数为 0.907~0.947。在外部测试集中,平均 Dice 系数为 0.882~0.933(表 1)。

2.1.2 解剖标志点定位性 以专家小组标注的解剖标志点为依据,在内部测试集中,所有解剖标志点模型定位结果和专家小组标注结果间 ED 小于 250 μm 的占比均不小于 97.48%,对应结果在外部测试集中为 95.65%(表 2)。

表 1 分割模型的平均 Dice 系数
Table 1 Average Dice coefficient of the segmentation model ($\bar{x} \pm s$)

模型	内部测试集	外部测试集
DCNN1	0.927	0.933
DCNN2	0.907	0.882
DCNN3	0.947	0.923
DCNN4	0.929	0.915

注:DCNN:深度卷积神经网络
Note:DCNN:deep convolutional neural network

2.1.3 眼前节参数测量结果的评估 以专家小组标注的解剖标志点为依据测量的眼前节参数(PD、ACD、ATA 和 STS)为参考标准,在内部测试集中,所有 ICC 值均超过 0.977,所有 ARE 值均不高于 2.120%。同样,在外部测试集中,所有 ICC 值均不小于 0.958,所有 ARE 值均不大于 2.407%(表 3,图 7)。

2.2 模型测量与 Pentacam 测量结果的一致性评估

使用 Pentacam 测量的 PD 和 ACD 值为参考标准,在内部测试集中,模型测得 PD 和 ACD 的 ICC 值分别为 0.540 和 0.946,在外部测试集中分别为 0.466 和 0.908(表 4)。

2.3 人机测量结果的对比分析

2 名资深屈光手术专家分别需要 1 027 和 1 116 s 才能完成所有测量,平均用时 1 072 s,该模型仅需 62 s 即能完成所有测量,模型测量结果与资深专家测量结果之间的 ICC 均不低于 0.969(表 5)。

表 2 不同测试数据集中解剖标志点定位的 ED 差异分布
Table 2 Differential distribution of ED at different anatomical landmarks in different test datasets ($\bar{x} \pm s$)

数据集	解剖标志点	ED($\bar{x} \pm s$, μm)	ED=0 (%)	ED $\leq$ 50 μm (%)	ED $\leq$ 100 μm (%)	ED $\leq$ 150 μm (%)	ED $\leq$ 200 μm (%)	ED $\leq$ 250 μm (%)
内部测试集	点 1	48.24 $\pm$ 27.29	22.30	82.37	98.92	100.00	100.00	100.00
	点 2	57.68 $\pm$ 40.06	18.35	73.02	93.17	97.84	100.00	100.00
	点 3	108.60 $\pm$ 69.04	6.83	35.25	67.99	84.17	92.81	97.84
	点 4	73.46 $\pm$ 48.58	12.95	59.35	87.05	95.32	98.92	99.64
	点 5	70.16 $\pm$ 59.85	18.71	61.87	89.21	93.88	96.04	98.92
	点 6	59.85 $\pm$ 46.49	20.14	72.30	91.37	97.48	98.56	99.64
	点 7	93.99 $\pm$ 61.43	5.76	38.85	79.86	92.09	95.32	97.48
	点 8	65.65 $\pm$ 43.83	15.83	62.95	91.01	97.84	99.64	99.64
外部测试集	点 1	69.88 $\pm$ 34.95	7.97	55.07	95.29	99.28	99.64	100.00
	点 2	70.71 $\pm$ 40.88	7.97	56.52	94.20	97.83	98.55	99.28
	点 3	124.20 $\pm$ 69.33	3.99	22.10	58.70	76.81	89.49	97.10
	点 4	87.87 $\pm$ 52.34	7.97	44.57	80.43	92.39	98.55	100.00
	点 5	67.80 $\pm$ 50.86	11.96	64.13	90.22	96.38	98.19	98.91
	点 6	73.85 $\pm$ 71.28	17.39	64.86	86.96	92.03	95.65	97.46
	点 7	123.30 $\pm$ 79.14	2.90	27.54	59.06	80.43	90.58	95.65
	点 8	65.72 $\pm$ 45.51	13.41	64.86	93.48	97.46	98.55	98.91

注:ED:欧式距离

Note:ED:Euclidean distance

表 3 模型测量结果与专家小组标注结果的一致性比较
Table 3 Comparison of the consistency between measurements from the model and the annotations from the expert panel

参数	内部测试集				外部测试集			
	ICC (95% CI)	P 值	ARE (%)	差异值 (LoA)	ICC (95% CI)	P 值	ARE (%)	差异值 (LoA)
PD	0.993 (0.990–0.995)	<0.01	2.120	27.386 (–120.264–175.037)	0.992 (0.986–0.995)	<0.01	2.407	38.253 (–135.727–212.232)
ACD	0.990 (0.977–0.995)	<0.01	1.075	21.009 (–51.044–93.061)	0.982 (0.952–0.991)	<0.01	1.529	30.710 (–64.920–126.340)
ATA	0.981 (0.975–0.985)	<0.01	0.731	–26.237 (–236.729–184.255)	0.976 (0.970–0.981)	<0.01	0.713	–2.757 (–219.071–213.556)
STS	0.977 (0.967–0.983)	<0.01	0.821	38.137 (–192.266–268.540)	0.958 (0.911–0.976)	<0.01	1.000	72.658 (–191.035–336.351)

注: ICC: 组内相关系数; CI: 置信区间; ARE: 平均相对误差; LoA: 一致性界限; PD: 瞳孔直径; ACD: 前房深度; ATA: 角-角距离; STS: 沟-沟距离
Note: ICC: intra-class correlation coefficient; CI: confidence interval; ARE: average relative error; LoA: limits of agreement; PD: pupil diameter; ACD: anterior chamber depth; ATA: angle-to-angle distance; STS: sulcus-to-sulcus distance

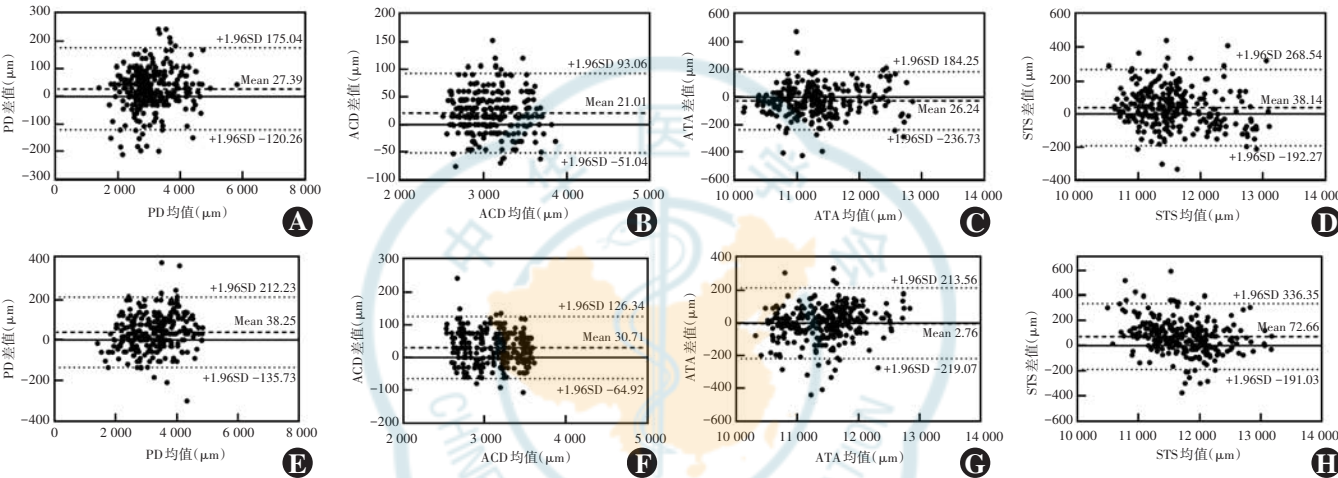


图 7 眼前节参数模型测量结果与专家小组测量结果的 Bland-Altman 图 A~D: 内部测试集中 PD、ACD、ATA、STS 结果 E~H: 外部测试集中 PD、ACD、ATA、STS 结果 PD: 瞳孔直径; ACD: 前房深度; ATA: 角-角距离; STS: 沟-沟距离
Figure 7 Bland-Altman plot comparing the measurements of anterior segment parameters from the model and expert panel A~D: PD, ACD, ATA and STS results in internal testing set E~H: PD, ACD, ATA and STS results in external testing set PD: pupil diameter; ACD: anterior chamber depth; ATA: angle-to-angle distance; STS: sulcus-to-sulcus distance

表 4 模型测量结果与 Pentacam 测量结果的一致性比较
Table 4 Comparison of the consistency between measurements from the model and Pentacam

数据集	参数	P ($\bar{x} \pm s, \mu\text{m}$)	UV ($\bar{x} \pm s, \mu\text{m}$)	UH ($\bar{x} \pm s, \mu\text{m}$)	ICC (95% CI)	P 值
内部测试集	PD	3 284.18 ± 606.92	3 262.97 ± 737.50	3 249.89 ± 680.68	0.540 (0.469–0.607)	<0.01
	ACD	3 149.30 ± 231.52	3 148.60 ± 235.16	3 139.47 ± 230.70	0.946 (0.933–0.956)	<0.01
外部测试集	PD	3 579.90 ± 751.89	4 316.35 ± 1 419.32	4 157.60 ± 1 515.31	0.466 (0.351–0.567)	<0.01
	ACD	3 193.38 ± 253.75	3 245.54 ± 243.77	3 223.11 ± 254.32	0.908 (0.875–0.931)	<0.01

注: P: Pentacam 测量结果; UV: UBM 垂直位扫描测量结果; UH: UBM 水平位扫描测量结果; ICC: 组内相关系数; CI: 置信区间; PD: 瞳孔直径; ACD: 前房深度
Note: P: Pentacam measurement results; UV: vertical scan measurement results of UBM; UH: horizontal scan measurement results of UBM; ICC: intra-class correlation coefficient; CI: confidence interval; PD: pupil diameter; ACD: anterior chamber depth

表 5 模型测量结果与资深专家测量结果的一致性比较
Table 5 Comparison of the consistency between measurements from the model and senior experts

参数	模型测量结果 ($\bar{x} \pm s, \mu\text{m}$)	资深专家 A 测量结果 ($\bar{x} \pm s, \mu\text{m}$)	资深专家 B 测量结果 ($\bar{x} \pm s, \mu\text{m}$)	ICC (95% CI)	P 值
PD	2 956.10 ± 540.39	2 935.75 ± 541.50	2 922.60 ± 564.34	0.987 (0.982–0.990)	<0.01
ACD	3 094.61 ± 628.09	3 082.88 ± 629.47	3 066.54 ± 659.06	0.992 (0.989–0.994)	<0.01
ATA	11 337.02 ± 552.16	11 344.19 ± 558.45	11 396.27 ± 592.03	0.969 (0.957–0.978)	<0.01
STS	11 668.16 ± 556.22	11 578.62 ± 598.26	11 589.02 ± 603.46	0.970 (0.948–0.981)	<0.01

注: ICC: 组内相关系数; CI: 置信区间; PD: 瞳孔直径; ACD: 前房深度; ATA: 角-角距离; STS: 沟-沟距离
Note: ICC: intra-class correlation coefficient; CI: confidence interval; PD: pupil diameter; ACD: anterior chamber depth; ATA: angle-to-angle distance; STS: sulcus-to-sulcus distance

3 讨论

本研究构建并评估了一种基于深度学习的模型,该模型可以根据 ICL 植入术患者术前的 UBM 检查图像自动测量眼前节参数。以专家小组标注结果为标注,在内部和外部测试集中,分割模型平均 Dice 系数为 0.882~0.947;所有解剖标志点模型定位结果与专家小组标注结果间 ED 小于 250 μm 的占比均不小于 95.65%;PD、ACD、ATA 和 STS 的 ICC 均不小于 0.958,ARE 均不大于 2.407%。以上结果表明,模型测量的眼前节参数与专家小组手动测量结果具有良好一致性。Bland-Altman 图进一步证实了二者测得前房参数之间的良好一致性。在与 Pentacam 测量结果的比较中发现,在内部测试集和外部测试集中,PD 的 ICC 分别为 0.540 和 0.466,而 ACD 的 ICC 分别为 0.946 和 0.908,表明模型与 Pentacam 测得 ACD 结果具有良好的一致性,而二者测得 PD 结果一致性较差。在人机对比中,4 个参数中 ICC 值最小的是 0.969,结果表明模型测量与资深专家测量结果之间具有高度一致性。值得注意的是,与眼科专家相比,该模型的测量速度更快,大约是眼科专家的 20 倍,与既往研究结果类似^[9]。

眼前节参数 ACD、PD、ATA 和 STS 在 ICL 尺寸选择中具有重要作用,尤其是 STS^[10-12]。在临床实践中,通常使用眼前节光学相干断层扫描 (anterior segment optical coherence tomography, AS-OCT) 和/或 UBM 对眼前节进行定量评估^[13-14]。最初通过 UBM 图像进行相关测量是通过机器内置测量软件,如 UBM Pro 2000 程序^[15]和 Lin 等^[16]开发的 UBM 图像测量软件实现的,上述测量软件需要操作者根据自身经验手动标注测量的起止点。这种半自动测量方法不可避免地会受到操作者临床经验、工作状态和条件的影响。即使在遵循相同的定位标准时,由于图像分辨率和对比度的差异,对相同的解剖标志点不同操作员标注的位置也可能存在较大差异,从而影响最终的测量结果。近年来,一些研究报道了 UBM 图像的自动评估系统。Wang 等^[17]开发了一种通过 UBM 图像对巩膜突自动定位的模型,该模型预测的巩膜突与手动标记巩膜突之间的 ED 差异在 150 μm 以内的占比为 95.35%。本研究模型在内部测试集和外部测试集中解剖标志点定位差异在 250 μm 及以下的占比分别大于 97.48% 和 95.65%,明显高于 Wang 等^[17]的结果。Wang 等^[17]开发的 UBM 图像自动评估模型是在半景模式下通过前房角 UBM 图像进行巩膜突的定位。根据设备相关性

能参数,本研究中使用的 UBM 设备和 Wang 等^[17]研究中使用的设备性能相当,参考本研究中物理距离和像素距离之间转换的方法,在半景模式中 150 μm 对应 15.6 像素长度,在全景模式中 250 μm 对应 16.5 像素长度,可认为本模型对解剖标志点的定位性能与 Wang 等^[17]开发的模型相当。Li 等^[18]、Wang 等^[4]和 Jiang 等^[5]开发的前房角自动测量系统可通过 UBM 图像对前房角自动进行量化评估,可测量小梁虹膜角和房角开放距离,模型测量结果与手工标注结果的 ICC 在 0.911~0.992。本研究中,模型测量结果与专家小组标注结果之间眼前节参数的 ICC 在 0.958~0.993。但在与 Pentacam 测量的比较中发现,PD 测量结果的一致性较差,ACD 测量结果的一致性良好。可能是因为 PD 大小易受环境的照度影响,Pentacam 检查是在暗室中进行,而 UBM 检查在相对明亮的环境中进行,导致测量出的 PD 一致性较差,而 ACD 不受上述因素影响,模型测量结果与 Pentacam 测量结果表现出良好的一致性。本研究中模型的性能与上述模型相当,ACD 测量结果可代替 Pentacam 测量结果。然而上述所有 UBM 图像分析系统均用于对前房角进行评估,但本研究是对眼前节进行评估,类似报道少见。

本研究中开发的基于 Python 编程语言的算法能够自动识别分割 UBM 图像中的不同组织,再根据不同组织间的解剖学关系自动定位解剖标志点。Pham 等^[6]开发了一种卷积神经网络模型,该模型能够将前房角的 AS-OCT 图像分割为虹膜、角膜巩膜和前房区域,并自动量化前房角参数。Pham 等^[6]研究中模型分割结果的平均 Dice 系数为 95.7%,本研究模型在内部测试集中对虹膜和角巩膜区域分割结果的平均 Dice 系数分别为 94.7% 和 92.9%,在外部测试数据集中分别为 92.3% 和 91.5%。分析原因,AS-OCT 的分辨率远高于 UBM,这种分辨率差异是导致本研究模型中的 Dice 系数较 Pham 等^[6]的模型低的主要原因。但 UBM 是目前唯一可以显示虹膜后睫状体和晶状体悬韧带的成像技术,由于 ICL 的植入位置为后房,本研究认为与通过测量水平角膜直径或 ATA 选择 ICL 尺寸的方法相比,通过 UBM 直接测量 STS 更优,这一观点在 Chen 等^[19]的研究中得到证实。因此,本研究认为 UBM 检查在 ICL 手术的术前评估中具有重要作用。

本研究也存在一定局限性。首先,尽管模型中使用的 UBM 图像由具有丰富临床经验的专家小组完成注释,但在注释过程中仍可能会引入人为错误,模型在训练过程中不可避免地会学习到这些错误。其次,模型与其他测量设备的比较中仅选择了 Pentacam,虽然

STS 的测量只能通过 UBM 实现,但 ATA 测量结果可通过 AS-OCT 获得,本研究选取的 ICL 患者中 AS-OCT 检查未作为术前常规检查项目,导致在本研究中无法对模型测量的 ATA 结果进行进一步验证。此外,本研究构建模型的目的是辅助 ICL 植入术患者的术前评估,术者选择合适尺寸的 ICL。然而,本模型仅测量了 STS、ATA、PD 和 ACD,根据相关研究还有其他因素与人工晶状体的选择有较大影响,如睫状体的形态和晶状体矢高^[20-21],未来可进一步对其他参数进行分析。

综上,本研究构建的基于人工智能的 UBM 图像模型可自动对眼前节进行可靠的量化评估,测量精度与资深眼科专家相当,且速度更快,不受工作时长影响,可辅助 UBM 操作人员更高效地完成检查。然而,还需要进一步研究改进和增强该模型临床性能的方法,为临床上选择合适尺寸的 ICL 提供更多决策支持信息。

利益冲突 所有作者均声明不存在利益冲突

作者贡献声明 朱剑:设计试验、实施研究、采集数据、分析/解释数据、文章撰写;颜瑜琳:实施研究、采集数据、分析数据;蒋维艳:实施研究、采集数据;张少维、牛晓光:实施研究、采集数据、参与人工智能模型测试;胡孝、郑碧清:参与人工智能模型的构建及测试;杨燕宁:参与设计试验、文章审阅及定稿

参考文献

- [1] 王耿媛,袁进. 人工智能在眼科学领域的应用研究进展[J]. 中华实验眼科杂志, 2019, 37(8): 680-683. DOI: 10. 3760/cma. j. issn. 2095-0160. 2019. 08. 018.
- [2] 董秀清,杜绍林,刘华秀,等. 人工智能在社区糖尿病视网膜病变诊断及转诊中的应用[J]. 中华实验眼科杂志, 2022, 40(12): 1158-1163. DOI: 10. 3760/cma. j. cn115989-20220316-00108.
- [3] Brown JM, Campbell JP, Beers A, et al. Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks[J]. JAMA Ophthalmol, 2018, 136(7): 803-810. DOI: 10. 1001/jamaophthalmol. 2018. 1934.
- [4] Wang W, Wang L, Wang X, et al. A deep learning system for automatic assessment of anterior chamber angle in ultrasound biomicroscopy images[J/OL]. Transl Vis Sci Technol, 2021, 10(11): 21[2024-10-20]. https://pubmed. ncbi..nlm.nih.gov/34570190/. DOI: 10. 1167/tvst. 10. 11. 21.
- [5] Jiang W, Yan Y, Cheng S, et al. Deep learning-based model for automatic assessment of anterior angle chamber in ultrasound biomicroscopy[J]. Ultrasound Med Biol, 2023, 49(12): 2497-2509. DOI: 10. 1016/j. ultrasmedbio. 2023. 08. 013.
- [6] Pham TH, Devalla SK, Ang A, et al. Deep learning algorithms to isolate and quantify the structures of the anterior segment in optical coherence tomography images[J]. Br J Ophthalmol, 2021, 105(9): 1231-1237. DOI: 10. 1136/bjophthalmol-2019-315723.
- [7] 王宁利,刘文. 活体超声生物显微镜眼科学[M]. 2版. 北京: 科学出版社, 2010: 39.
- [8] Wang NL, Liu W. Ultrasound biomicroscopic ophthalmology[M]. 2nd ed. Beijing: Science Press, 2010: 39.
- [9] Zhou Z, Siddiquee M, Tajbakhsh N, et al. UNet++: a nested U-Net architecture for medical image segmentation[J]. Deep Learn Med Image Anal Multimodal Learn Clin Decis Support, 2018, 11045: 3-11. DOI: 10. 1007/978-3-030-00889-5_1.
- [10] 林浩添,吴晓航. 加快基于眼科图像数据库的眼病人工智能辅助诊断平台建设[J]. 中华实验眼科杂志, 2018, 36(8): 577-580. DOI: 10. 3760/cma. j. issn. 2095-0160. 2018. 08. 001.
- [11] Lin HT, Wu XH. Accelerating the construction of artificial intelligence diagnostic platform based on ophthalmic imaging database[J]. Chin J Exp Ophthalmol, 2018, 36(8): 577-580. DOI: 10. 3760/cma. j. issn. 2095-0160. 2018. 08. 001.
- [12] Lee H, Kang D, Choi JY, et al. Analysis of pre-operative factors affecting range of optimal vaulting after implantation of 12. 6-mm V4c implantable collamer lens in myopic eyes[J/OL]. BMC Ophthalmol, 2018, 18(1): 163[2024-10-20]. https://pubmed. ncbi. nlm. nih. gov/29980187/. DOI: 10. 1186/s12886-018-0835-x.
- [13] Trancón AS, Manito SC, Sierra OT, et al. Determining vault size in implantable collamer lenses: preoperative anatomy and lens parameters[J]. J Cataract Refract Surg, 2020, 46(5): 728-736. DOI: 10. 1097/j. jcrs. 0000000000000146.
- [14] Reinstein DZ, Lovisolo CF, Archer TJ, et al. Comparison of postoperative vault height predictability using white-to-white or sulcus diameter-based sizing for the visian implantable collamer lens[J]. J Refract Surg, 2013, 29(1): 30-35. DOI: 10. 3928/1081597X-20121210-02.
- [15] Janssens R, van Rijn LJ, Eggink CA, et al. Ultrasound biomicroscopy of the anterior segment in patients with primary congenital glaucoma: a review of the literature[J]. Acta Ophthalmol, 2022, 100(6): 605-613. DOI: 10. 1111/aos. 15082.
- [16] Mirzayev I, Gündüz AK, Aydın Ellialtıoğlu P, et al. Clinical applications of anterior segment swept-source optical coherence tomography: a systematic review[J/OL]. Photodiagnosis Photodyn Ther, 2023, 42: 103334[2024-10-20]. https://pubmed. ncbi. nlm. nih. gov/36764640/. DOI: 10. 1016/j. pdpdt. 2023. 103334.
- [17] Ishikawa H, Liebmann JM, Ritch R. Quantitative assessment of the anterior segment using ultrasound biomicroscopy[J]. Curr Opin Ophthalmol, 2000, 11(2): 133-139. DOI: 10. 1097/00055735-200004000-00012.
- [18] Lin Z, Mou da P, Liang YB, et al. Reproducibility of anterior chamber angle measurement using the Tongren ultrasound biomicroscopy analysis system[J]. J Glaucoma, 2014, 23(2): 61-68. DOI: 10. 1097/IJG. 0b013e3182698094.
- [19] Wang W, Wang L, Wang T, et al. Automatic localization of the scleral spur using deep learning and ultrasound biomicroscopy[J/OL]. Transl Vis Sci Technol, 2021, 10(9): 28[2024-10-21]. https://pubmed. ncbi. nlm. nih. gov/34427626/. DOI: 10. 1167/tvst. 10. 9. 28.
- [20] Li W, Chen Q, Jiang Z, et al. Automatic anterior chamber angle measurement for ultrasound biomicroscopy using deep learning[J]. J Glaucoma, 2020, 29(2): 81-85. DOI: 10. 1097/IJG. 0000000000001411.
- [21] Chen X, Ye Y, Yao H, et al. Predicting post-operative vault and optimal implantable collamer lens size using machine learning based on various ophthalmic device combinations[J/OL]. Biomed Eng Online, 2023, 22(1): 59[2024-10-21]. https://pubmed. ncbi. nlm. nih. gov/37322471/. DOI: 10. 1186/s12938-023-01123-w.
- [22] Chen Q, Tan W, Lei X, et al. Clinical prediction of excessive vault after implantable collamer lens implantation using ciliary body morphology[J]. J Refract Surg, 2020, 36(6): 380-387. DOI: 10. 3928/1081597X-20200513-02.
- [23] Cerpa Manito S, Sánchez Trancón A, Torrado Sierra O, et al. Biometric and ICL-related risk factors associated to sub-optimal vaults in eyes implanted with implantable collamer lenses[J/OL]. Eye Vis (Lond), 2021, 8(1): 26[2024-10-21]. https://pubmed. ncbi. nlm. nih. gov/34225809/. DOI: 10. 1186/s40662-021-00250-6.

(收稿日期: 2024-12-10 修回日期: 2025-05-10)

(本文编辑: 刘艳 施晓萌)

