

· 临床研究 ·

基于深度学习的角膜活体共聚焦显微镜图像辅助识别系统的构建及应用

颜瑜琳¹ 蒋维艳¹ 程思敏¹ 周奕文¹ 于慧¹ 郑碧清² 杨燕宁¹¹武汉大学人民医院眼科中心, 武汉 430060; ²武汉大学资源与环境学院, 武汉 430079

通信作者: 杨燕宁, Email: ophyyn@163.com

【摘要】 目的 构建基于深度学习的角膜活体共聚焦显微镜(IVCM)图像人工智能(AI)辅助识别系统,并评估其在临床上的应用价值。方法 采用诊断试验研究方法,收集2021年5月至2022年9月于武汉大学人民医院及武汉大学中南医院行IVCM检查的331例受试者18860张角膜IVCM图像,经资深角膜专家对图像进行筛选分类后用于模型的训练及测试。模型包括低质量图像过滤模型、角膜图像诊断模型及角膜上皮层、前弹力层、基质层和内皮层4个层次识别模型,判断正常及异常角膜图像与对应角膜层次。选取360张独立于数据库的IVCM图像进行人机大赛,比较3位角膜专业高年资医师与人工智能对图片识别的准确度及用时。另选取8位未经过IVCM培训且临床经验少于3年的低年资医师对360张图像进行无模型辅助及模型辅助下识图,对2次评估结果进行差异分析以评估模型辅助效果。结果 本诊断模型筛选高质量图像的准确率为0.954,在内部及外部测试集中,识别角膜正常/异常图像的总准确度分别为0.916和0.896;在识别正常及异常图像的角膜层次中,内部测试集的准确度分别达到了0.983及0.925,外部测试集分别达到了0.988及0.929。人机大赛中,模型总体的识别准确率为0.878,与3名高年资医师的平均准确度相近,且评估速度远高于高年资医师,约为其300倍。低年资医师经机器辅助后对图像正/异常及层次诊断的总体平均准确率为0.816±0.043,明显高于模型辅助前的0.669±0.061,差异有统计学意义($t=6.304, P<0.001$)。结论 成功构建了一种基于深度学习的IVCM角膜图像辅助识别系统,可以较准确地分辨正常/异常图像并诊断图像对应角膜层次,能够提高临床诊断效率并辅助医师训练学习。

【关键词】 显微镜, 共焦; 角膜; 人工智能; 深度学习**基金项目:** 国家自然科学基金(81770899); 湖北省重点研发计划(2020BCB055); 武汉大学人民医院交叉创新人才项目(JCRCZN-2022-007)

DOI:10.3760/cma.j.cn115989-20230101-00001

Construction and application of a deep learning-based assistant system for corneal *in vivo* confocal microscopy images recognition

Yan Yulin¹, Jiang Weiyang¹, Cheng Simin¹, Zhou Yiwen¹, Yu Yi¹, Zheng Biqing², Yang Yanning¹¹Department of Ophthalmology, Renmin Hospital of Wuhan University, Wuhan 430060, China; ²School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China

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

【Abstract】 Objective To construct an artificial intelligence (AI)-assisted system based on deep learning for corneal *in vivo* confocal microscopy (IVCM) image recognition and to evaluate its value in clinical applications.**Methods** A diagnostic study was conducted. A total of 18 860 corneal images were collected from 331 subjects who underwent IVCM examination at Renmin Hospital of Wuhan University and Zhongnan Hospital of Wuhan University from May 2021 to September 2022. The collected images were used for model training and testing after being reviewed and classified by corneal experts. The model design included a low-quality image filtering model, a corneal image diagnosis model, and a 4-layer identification model for corneal epithelium, Bowman membrane, stroma, and endothelium, to initially determine normal and abnormal corneal images and corresponding corneal layers. A human-machine competition was conducted with another 360 database-independent IVCM images to compare the accuracy and time spent on image recognition by three senior ophthalmologists and the AI system. In addition, 8 trainees without IVCM training and with less than three years of clinical experience were selected to recognize the same 360 images without and with model assistance to analyze the effectiveness of model assistance. This study adhered to the Declaration of Helsinki. The study protocol was approved by the Ethics Committee of Renmin Hospital of Wuhan University (No. WDRY2021-K148). **Results** The accuracy of this diagnostic model in screening high-quality images was 0.954. Its overall accuracy in identifying normal/abnormal corneal images was 0.916 and 0.896 in the internal and external test sets, respectively. Its accuracy reached 0.983, 0.925 in the internal test sets and 0.988, 0.929 in the external test sets in identifying corneal layers of normal and abnormal images, respectively. In the human-machine

competition, the overall recognition accuracy of the model was 0.878, which was similar to the average accuracy of the three senior physicians and was approximately 300 times faster than the experts in recognition speed. Trainees assisted by the system achieved an accuracy of 0.816 ± 0.043 in identifying corneal layers of normal and abnormal images, which was significantly higher than 0.669 ± 0.061 without model assistance ($t = 6.304, P < 0.001$). **Conclusions** A deep learning-based assistant system for corneal IVCM image recognition is successfully constructed. This system can discriminate normal/abnormal corneal images and diagnose the corresponding corneal layer of the images, which can improve the efficiency of clinical diagnosis and assist doctors in training and learning.

[Key words] Microscopy, confocal; Cornea; Artificial intelligence; Deep learning

Fund program: National Natural Science Foundation of China (81770899); Major Project of Research and Development Program in Hubei Province (2020BCB055); Interdisciplinary Innovative Talents Foundation from Renmin Hospital of Wuhan University (JCRCZN-2022-007)

DOI: 10.3760/cma.j.cn115989-20230101-00001

角膜作为视觉形成的第一步,其透明度及屈光率可使光线折射进入眼内并聚焦于视网膜上,故角膜各层次的结构及功能损伤可导致视力下降,甚至致盲^[1]。角膜疾病,如颗粒状角膜营养不良、Fuchs 角膜内皮营养不良等可主要损害角膜单一层次,而感染性角膜炎、角膜机械性损伤及眼化学伤等则可造成多层次,甚至全层角膜损伤^[2]。同时,糖尿病及类风湿性关节炎等全身疾病亦可对角膜造成影响^[3]。活体共聚焦显微镜(*in vivo* confocal microscopy, IVCM)作为一种非侵入性的成像工具,可从细胞层面观测角膜及其在病理状态下的结构变化,拥有实时、无创、可反复检查及高分辨率等优点^[4],对多种角膜疾病的临床诊断具有重要参考价值,临床应用已日渐广泛。对角膜微观结构变化进行监测也有助于优化角膜病的针对性管理及评估患者全身疾病预后^[5-6]。但在实际工作中,由于 IVCM 镜头单次可拍摄面积较小($400 \mu\text{m} \times 400 \mu\text{m}$),详细评估角膜时需采集大量图像,人工分析非常费时费力并不可避免具有主观性^[7],且 IVCM 阅片也对医师的经验及相关专业知识具有一定要求,医师常常需要一定的培训周期才能区分角膜各层次形态及判断其是否正常。早期诊断对于角膜疾病的精准治疗及预防角膜盲均有重要意义^[8],据世界卫生组织统计,约 80% 的角膜盲可避免,而目前全球发达国家及发展中国家的眼科医师均存在短缺情况^[9],提高角膜图像阅片时的准确度以及诊断效率,可为临床及科研工作减负、提高医师工作效率并有望给予更多角膜病患者快速准确的诊疗。人工智能(artificial intelligence, AI)的不断进步正在改变各个医学领域的筛查、诊断及治疗方式^[10],AI 在眼科疾病中的应用也在过去 10 年中有着显著发展。目前,AI 在辅助 IVCM 图像中对角膜上皮细胞、角膜神经、角膜内皮细胞、真菌菌丝、树突状细胞及炎性细胞等多种结构的分割、量化及鉴别方面都取得了重大突破^[11-14],其阅片速度及准确度均表现出

优秀的性能,然而目前尚无对于角膜层次判断及识别角膜图像正/异常的相关研究。本研究拟构建 AI 辅助下 IVCM 图像的自动诊断模型,探讨其在临床应用中的效能及用于角膜疾病智能筛查的可行性。

1 资料与方法

1.1 一般资料

1.1.1 IVCM 图像来源 采用诊断试验研究方法,收集 2021 年 5 月至 2022 年 9 月于武汉大学人民医院眼科中心及武汉大学中南医院进行 IVCM 检查(HRT III/RCM,德国 Heidelberg Engineering 公司)患者的角膜图像。所有病例均来源于医院信息系统,所有 IVCM 图像均由工作经验超过 15 年的资深眼科 IVCM 检查医师经严格规范操作后采集,图像进行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审核(批文号:WDRY2021-K148)。

1.2 方法

1.2.1 IVCM 图像筛选、预处理及分类 共纳入武汉大学人民医院眼科中心 246 例患者 17 675 张及武汉大学中南医院 85 例患者 1 185 张 IVCM 图像。图像先经角膜专业高年资医师筛选,分为清晰的高质量图像及因过曝、光线不足、对焦不清晰、拍摄模糊或接触不佳等情况所致低质量图像。高质量图像经光学字符识别提取深度信息,先统一转换为 $384 \text{ 像素} \times 384 \text{ 像素}$ 大小再进行分类。依据 2014 年版《活体角膜激光共聚焦显微镜图谱》和 2021 年版《眼表活体共聚焦显微镜图谱》,同时参考 2 名临床经验超过 20 年的角膜医师意见,制定如下分类标准:(1)正常图像标准 正常上皮层图像为视野内出现的上皮细胞均结构完整,形态清晰,密度均匀;正常前弹力层图像背景为均质中度反光,可见粗细、弯曲度及密度适中的高反光神经纤维;正常基质层图像为无特征暗反光背景及边界清晰的基

质细胞核,可见少量粗大高反光基质神经;正常内皮层图像为排列规则的均匀 5~7 边形细胞,细胞边界清晰。(2)异常图像标准 异常上皮层图像可见上皮细胞水肿、结构不清、间隙增大和/或出现炎性细胞等;异常前弹力层图像中可见明显的神经纤维迂曲、变细及密度降低,有 10 个以上未活化的朗格汉斯细胞、有 1 个及以上活化的朗格汉斯细胞和/或椭圆形炎性细胞等;异常基质层图像可见基质细胞肿胀、活化、松针样高反光瘢痕及出现阿米巴包囊、真菌菌丝、真菌孢子、新生血管等异常结构;异常内皮细胞图像可见内皮细胞肿胀、变性、营养不良及任意角膜后沉积物等。该标准主要是为了提高灵敏度以尽可能多筛选出异常图像。图像首先由角膜专业高年资医师参照分类标准依图像特征分为正常及异常图像,然后分别依角膜层次分类为上皮层、前弹力层、基质层及内皮层。后弹力层因其厚度极薄并与角膜内皮细胞贴附紧密,故所收集图像较少,并且其临床意义相对有限,目前相关研究多集中于角膜移植领域,故不参与训练。将分类后图像划分为训练集、内部测试集、外部测试集及验证集,不同数据集分类详情见表 1,纳入病例的临床诊断及对应例数见表 2。

1.2.2 角膜 IVCM 图像自动诊断模型建立 团队既往曾进行其他眼科图像的 AI 分类研究,经探索发现 Res-Net-50 作为目前稳定、简洁且高效的残差学习架构,十分适用于眼科图像分类,故本研究沿用了基于 Res-Net-50 的深度学习网络模型进行图像分类训练,使用 Early Stopping、Drop out 以及数据集扩增来降低过拟合风险。具有 4 个 NVIDIA Geforce GTX 2080 (GPU 内存 8GB) 的服务器被用于训练模型,计算机算法使用 Python (3.6.5 版) 编写,开源 TensorFlow 库 (1.12.2 版) 和 Keras 库 (2.2.5 版) 作为后端。

研究初期有针对层次及正/异常分类在模型中的顺序进行探索,依据模型时效性及准确度,最终建立模型训练流程如下:(1)深度卷积网络 1 (deep convolutional neural network 1,

DCNN1) 将图像分为高质量图像与低质量图像;(2)DCNN2 对高质量图像进行识别,将其分类为正常及异常图像;(3)DCNN3、DCNN4、DCNN5 及 DCNN6 由 DCNN3 及 DCNN4 分别将正常及异常图像分类为上皮/内皮层、前弹力层及基质层;再由 DCNN5 及 DCNN6 分别将正常及异常的上皮/内皮层图像分类为上皮层及内皮层图像。整体流程设计符合眼科模型建立的伦理要求^[15]。详细模型训练流程图见图 1。

表 1 收集数据的基本信息及数量
Table 1 Information and quantity of the data collected

| 数据集 | 患者例数 (n) | 图像质量 | 图像数量 (张) | 年龄 ($\bar{x} \pm s$, 岁) | 各层次图像数量(张) | | | |
|----------------|----------|------|----------|---------------------------|------------|-------|-------|-----|
| | | | | | 上皮层 | 前弹力层 | 基质层 | 内皮层 |
| 训练集 | 180 | 高质量 | 10 195 | 47.32±4.69 | 1 501 | 1 887 | 5 887 | 920 |
| | | 低质量 | 2 964 | | | | | |
| 内部测试集 | 45 | 高质量 | 3 260 | 49.65±5.77 | 560 | 1 103 | 964 | 633 |
| | | 低质量 | 896 | | | | | |
| 外部测试集 | 85 | 高质量 | 1 185 | 46.18±5.13 | 267 | 316 | 310 | 292 |
| 人机大赛及新手辅助识图数据集 | 21 | 高质量 | 360 | 50.24±6.05 | 62 | 89 | 106 | 103 |

注:/:因大部分低质量图像无法明确角膜层次,故各层次数量均不进行单独统计
Note:/: Because most of the low-quality images could not determine the corneal layers, the number of each layer was not counted separately

表 2 纳入病例的临床诊断及对应数量
Table 2 Clinical diagnosis and number of included cases

| 数据来源 | 正常及各类角膜疾病患者例数(n) | | | | | | | | |
|----------|------------------|--------|--------------|----|----------|-------|--------|-------|----|
| | 正常 | 感染性角膜炎 | Fuchs 角膜营养不良 | 干眼 | 神经营养性角膜炎 | 前葡萄膜炎 | 角膜移植术后 | 角膜化学伤 | 其他 |
| 武汉大学人民医院 | 93 | 53 | 18 | 15 | 12 | 10 | 8 | 8 | 29 |
| 武汉大学中南医院 | 31 | 15 | 8 | 8 | 3 | 5 | 2 | 3 | 11 |

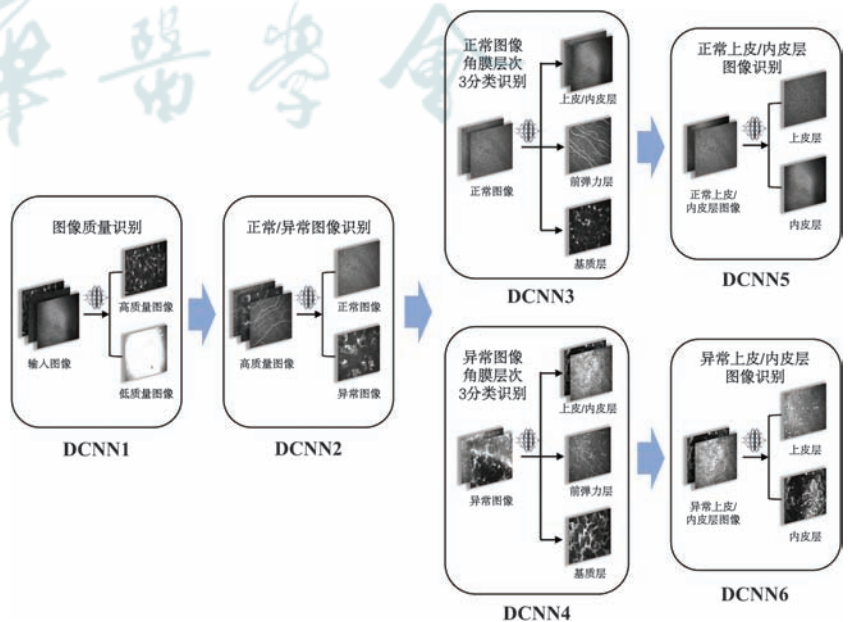


图 1 模型训练流程图 DCNN:深度卷积网络

Figure 1 Flow chart of model training DCNN: deep convolutional neural network

1.2.3 评价指标

1.2.3.1 模型准确性验证 1 个内部测试的数据集(武汉大学人民医院)及 1 个外部测试的数据集(武汉大学中南医院)被用于评估模型的性能。采用准确度、特异度、敏感度、受试者工作特征曲线(receiver operating characteristic curve, ROC)、ROC 曲线下面积(area under curve, AUC)、阳性预测值(positive predictive value, PPV)和阴性预测值(negative predictive value, NPV)评估模型准确性。

1.2.3.2 角膜 IVCN 图像诊断人机大赛 模型训练完成后,另选取独立于机器训练集及测试集的图像 360 张,邀请 3 位拥有 10 年以上工作经验的角膜专业高年资医师参与人机大赛。3 位医师分别独立对图像表现正常/异常及图像层次进行诊断,并在测试中由同一名研究人员记录耗时,同时图像经模型重复评估 3 次获得相应的模型准确度及用时,最后比较 3 位医师及模型评估结果的准确性及评估用时。

1.2.3.3 模型辅助低年资医师阅片 使用与人机大赛同一批图片,另邀请 8 位未经过 IVCN 培训且临床经验少于 3 年的眼科医师分别在无机器辅助及有机器辅助(图像先由计算机进行评估并以标签形式显示结果)的情况下独立对 360 张图片进行诊断并记录结果,比较 2 次评估的准确度。

1.3 统计学方法

采用 SPSS 25.0 软件(美国 IBM 公司)及 MedCalc 19.1-64 位软件(比利时 MedCalc Software 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。人机大赛中高年资医师及机器的诊断准确度比较采用独立样本 *t* 检验,计算机辅助前后低年资医师的诊断准确度比较采用配对样本 *t* 检验。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 模型识别准确性验证

(1)计算机模型 DCNN1 在内部测试集中对高质量图像分类识别的准确度为 0.954,敏感度为 0.960,特异度为 0.940,PPV 为 0.983,NPV 为 0.860。(2)DCNN2

在内部测试集中对异常图像分类识别的准确度为 0.916,敏感度为 0.910,特异度为 0.921,PPV 为 0.917,NPV 为 0.915;在外部测试集中准确度为 0.896,敏感度为 0.945,特异度为 0.847,PPV 为 0.862,NPV 为 0.938。(3)DCNN3 在内部测试集中对正常图像层次识别的总体准确度为 0.983,其中对上皮/内皮层识别的准确度为 1.000,对前弹力层识别的准确度为 0.958,对基质层识别的准确度为 0.986;在外部测试集中对正常图像层次识别的总体准确度为 0.988,其中对上皮/内皮层识别的准确度为 1.000,对前弹力层识别的准确度为 0.962,对基质层识别的准确度为 0.993。(4)DCNN4 在内部测试集中对异常图像层次识别的总体准确度为 0.925,其中对上皮/内皮层识别的准确度为 0.909,对前弹力层识别的准确度为 0.900,对基质层识别的准确度为 0.985;在外部测试集中对异常图像层次识别的总体准确度为 0.929,其中对上皮/内皮层识别的准确度为 0.916,对前弹力层识别的准确度为 0.893,对基质层识别的准确度为 0.988。(5)DCNN5 及 DCNN6 对正常及异常图像上皮及内皮层识别的准确度在内部及外部测试集中均为 1.000。DCNN1 和 DCNN2 对图像高/低质量及正/异常分类的混淆矩阵及 ROC 曲线见图 2;DCNN3、DCNN4、DCNN5 和 DCNN6 对图像层次识别的混淆矩阵见图 3。

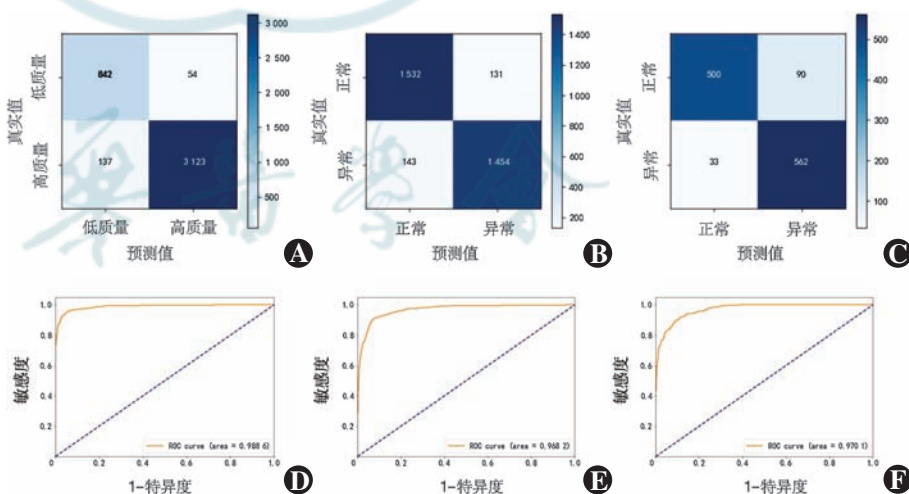


图 2 DCNN1、DCNN2 和 DCNN3 模型的混淆矩阵及 ROC 曲线 A:DCNN1 模型(内部测试)混淆矩阵 B:DCNN2 模型(内部测试)混淆矩阵 C:DCNN3 模型(内部测试)混淆矩阵 D:DCNN1 模型(内部测试)ROC 曲线 AUC = 0.989 E:DCNN2 模型(内部测试)ROC 曲线 AUC = 0.968 F:DCNN2 模型(外部测试)ROC 曲线 AUC = 0.970 ROC:受试者工作特征曲线

Figure 2 Confusion matrix diagrams and ROC curve of DCNN1, DCNN2 and DCNN3 models

A:Confusion matrix diagram of DCNN1 (Internal test) B:Confusion matrix diagram of DCNN2 (Internal test) C:Confusion matrix diagram of DCNN3 (Internal test) D:ROC curve of DCNN1 (Internal test) AUC = 0.989 E:ROC curve of DCNN2 (Internal test) AUC = 0.968 F:ROC curve of DCNN2 (External test) AUC = 0.970 ROC:receiver operating characteristic curve

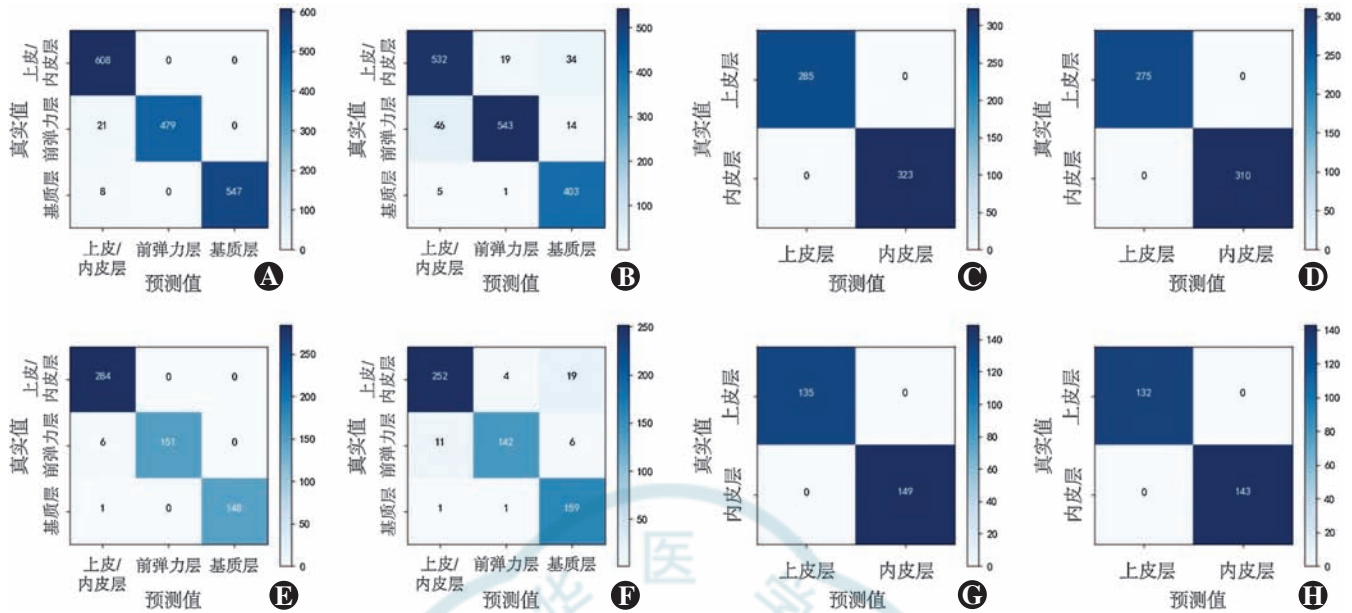


图 3 DCNN3、DCNN4、DCNN5 和 DCNN6 模型的混淆矩阵 A:DCNN3 模型(内部测试) B:DCNN4 模型(内部测试) C:DCNN5 模型(内部测试) D:DCNN6 模型(内部测试) E:DCNN3 模型(外部测试) F:DCNN4 模型(外部测试) G:DCNN5 模型(外部测试) H:DCNN6 模型(外部测试)

Figure 3 Confusion matrix diagrams of DCNN3, DCNN4, DCNN5 and DCNN6 models A:DCNN3 (Internal test) B:DCNN4 (Internal test) C:DCNN5 (Internal test) D:DCNN6 (Internal test) E:DCNN3 (External test) F:DCNN4 (External test) G:DCNN5 (External test) H:DCNN6 (External test)

2.2 人机诊断准确度比较

针对 360 张独立测试图像,模型及高年资医师对正常及异常图像分类识别的平均准确度分别为 0.922 ± 0.000 和 0.916 ± 0.008 ,对图像层次识别的平均准确度分别为 0.956 ± 0.000 和 0.951 ± 0.005 ,总体准确度分别为 0.878 ± 0.000 和 0.881 ± 0.009 。模型在正常/异常判断、层次识别及总体诊断的准确度均与 3 名高年资医师相近,差异均无统计学意义 ($t = 0.749, P = 0.495; t = 1.487, P = 0.300; t = -0.498, P = 0.645$)。高年资医师用时最短为 3 155 s,平均 8.764 s/张;最长用时 3 491 s,平均 9.697 s/张;模型评估用时 11 s,平均 0.031 s/张,评估速度远高于高年资医师,约为其 300 倍(表 3)。

表 3 人机识别 IVCM 角膜图像的准确度及用时比较
Table 3 Comparison of accuracy and time of identifying corneal IVCM images between the model and ophthalmologists

| 分类 | 人机识别 IVCM 角膜图像准确度 | | | 用时(s) |
|--------|-------------------|-------|-------|-------|
| | 正常/异常 | 层次 | 总体 | |
| 模型 | 0.922 | 0.956 | 0.878 | 11 |
| 眼科医师 1 | 0.917 | 0.953 | 0.869 | 3 155 |
| 眼科医师 2 | 0.928 | 0.944 | 0.892 | 3 491 |
| 眼科医师 3 | 0.908 | 0.956 | 0.883 | 3 420 |

注:IVCM:活体共聚焦显微镜

Note:IVCM:in vivo confocal microscopy

2.3 模型辅助前后低年资医师诊断准确度比较

低年资医师经机器辅助后对图像正/异常及层次诊断的总体平均准确度为 0.816 ± 0.043 ,明显高于模型辅助前的 0.669 ± 0.061 ,差异有统计学意义 ($t = 6.304, P < 0.001$) (图 4)。

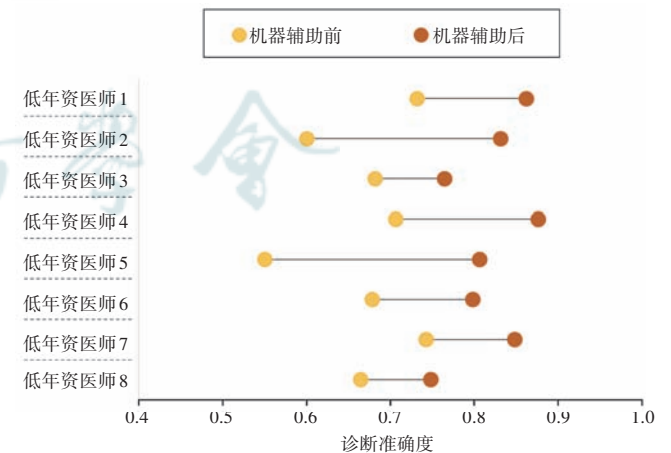


图 4 模型辅助前后低年资医师图像诊断准确度比较

Figure 4 Comparison of imaging diagnosis accuracy of trainees between without and with model assistance

3 讨论

本研究利用 2 个三分类及 4 个二分类模型对 9 种 IVCM 图像(低质量图像及高质量的正常上皮、异常上皮、正常前弹力层、异常前弹力层、正常基质、异常基

质、正常内皮、异常内皮)进行识别诊断。首先利用 DCNN1 进行高质量图像筛选;利用 DCNN2 进行正常及异常角膜图像识别;利用 DCNN3、DCNN5 进行正常角膜图像的层次识别;利用 DCNN4、DCNN6 进行异常角膜图像的层次识别。同时,虽然该网络为针对单张图片而非单个患者的诊断识别,实际临床操作中检查者会为每例患者采集几十到几百张不等的图片,只要将单个患者全部图像输入模型即可获得所采集图像的正/异常及层次识别结果,从而得到针对该患者 IVCM 检查的整体诊断。

近年来,深度学习技术已被广泛应用于眼科疾病的诊断、鉴别及防治^[16-19],美国食品药品监督管理局也于 2018 年批准了首个基于 AI 的糖尿病视网膜病变自主诊断工具^[20]。同时,机器学习在对裂隙灯显微镜图像、眼前节光学相干断层扫描(optical coherence tomography, OCT)、黄斑 OCT 及眼底造影等多种眼科前后节图像进行诊断、分割、量化及计算中均获得了令人瞩目的成果^[10, 21-22],如利用卷积神经网络构建了眼科 B 型超声的玻璃体视网膜疾病辅助诊断系统、基于深度学习的活动性角膜感染及角膜瘢痕区分模型等^[23-24]。在 IVCM 识图领域,目前也有相关研究发现 AI 可辅助角膜神经的量化与分割、神经纤维曲率分级、神经纤维与树突状细胞及真菌菌丝鉴别、活化树突状细胞与炎性细胞的识别及角膜内皮细胞分割与形态参数评估等^[7, 13, 25-26],充分说明深度学习适用于辅助 IVCM 图像进行多种结构识别。但现有研究多围绕单一层次的角膜图像或部分具体组织展开,而角膜由多个层次构成,因此开发一个适用范围更广并能全面评估角膜图像的辅助识图模型有利于弥补这一研究领域的空白。

Res-Net-50 是一种具有计算负担小、易优化等优点的残差学习框架,是基于现有深度网络训练基础上提出的,其包含 1 个全连接层和 49 个卷积层的优秀残差网络模型,可用来解决退化和梯度问题^[23]。本研究中采用 Res-Net-50 创建了基于 6 个深度神经网络的机器模型用于观察 AI 在筛选高质量图像、诊断正常/异常图像及识别角膜层次方面的能力,结果表明该模型在过滤低质量图像及识别图像表现正/异常方面具有较好的准确度、特异度及敏感度,并在区分角膜上皮层、前弹力层、基质层及内皮层 4 个层次上具有很高的诊断效能。人机比赛证明该模型具有与高年资角膜专家相当的准确度且识别速度远高于高年资医师,约为其 300 倍。同时,低年资医师在经模型辅助后对 IVCM 图像识别的准确度明显提高,其中部分低年资医师的

诊断准确度甚至接近高年资医师。上述结果表明,该模型在辅助评估 IVCM 图像中具有巨大的应用潜力,并有望助力实际临床科研工作中大批量图像的筛查分类,便于眼科医师快速集中地查看患者的异常图像或统一获取研究所需特定层次及类型的图像。

本研究仍存在一定的局限性。首先,由于模型针对图像的识别结果是定性的,对实际异常的特征,如上皮层炎性细胞、基质层菌丝及内皮层角膜后沉着物等并不能进一步做定量或分级评估,本课题组计划在下一步研究中扩大训练的样本量,尝试针对不同异常特征进行分割识别。其次,虽然本研究已尽可能多地纳入临床角膜疾病的图像,但由于数据来源均为固定时间内的回顾性收集,仍有临床较少见的角膜疾病未能纳入;同时,模型训练依赖足够的样本数据,而部分纳入研究的疾病图像数量相对有限,不足以针对异常图像做进一步的病种识别,对于此问题,本课题组拟后期联合多家医院以建立病种更多、样本量更大的数据库用于模型训练优化,可综合分析患者各层次 IVCM 图像表现,增加对疾病病种的识别同时提高模型识别精度,以期将 AI 技术更好地应用于临床及科研,利于眼科智能医疗的普及。

综上,本研究开发了基于深度学习的角膜 IVCM 图像智能诊断模型,结果表明其拥有较高的准确度、特异度及敏感度,可辅助临床医师更快、更好地识别角膜 IVCM 图像。该模型有利于减轻人工阅片及诊断的工作量,后续有望结合移动终端帮助临床经验较少或角膜专业医师缺乏的社区和基层医院学习、识别角膜疾病的 IVCM 图像,也有助于针对大批量患者角膜疾病的筛查及获取研究所需特定角膜层次图像。

利益冲突 所有作者均声明不存在利益冲突

作者贡献声明 颜瑜琳:设计试验、实施研究、采集数据、分析/解释数据、文章撰写;蒋维艳:实施研究、采集数据、分析数据;程思敏:实施研究、采集数据;周奕文、于慧:设计试验、实施研究、采集数据;郑碧清:参与人工智能模型搭建及测试;杨燕宁:参与试验设计、文章审阅及定稿

参考文献

- [1] Kumar A, Yun H, Funderburgh ML, et al. Regenerative therapy for the cornea [J/OL]. *Prog Retin Eye Res*, 2022, 87 : 101011 [2023 - 06 - 10]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8918435/>. DOI:10.1016/j.preteyeres.2021.101011.
- [2] Durand ML, Barshak MB, Chodosh J. Infectious keratitis in 2021 [J]. *JAMA*, 2021, 326(13) : 1319-1320. DOI:10.1001/jama.2021.0424.
- [3] Wang EF, Misra SL, Patel DV. *In vivo* confocal microscopy of the human cornea in the assessment of peripheral neuropathy and systemic diseases [J/OL]. *Biomed Res Int*, 2015, 2015 : 951081 [2023 - 06 - 10]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4685107/>. DOI:10.1155/2015/951081.
- [4] 顾绍峰, 彭荣梅, 肖格格, 等. 共聚焦显微镜下后部多形性角膜营养

- 不良的影像学特征[J]. 中华眼科杂志, 2022, 58(2): 103-111. DOI: 10. 3760/cma. j. cn112142-20210228-00099.
- Gu SF, Peng RM, Xiao GG, et al. Imaging features of posterior polymorphous corneal dystrophy observed by *in vivo* confocal microscopy [J]. Chin J Ophthalmol, 2022, 58(2): 103-111. DOI: 10. 3760/cma. j. cn112142-20210228-00099.
- [5] Niederer RL, McGhee CN. Clinical *in vivo* confocal microscopy of the human cornea in health and disease [J]. Prog Retin Eye Res, 2010, 29(1): 30-58. DOI: 10. 1016/j. pretyeres. 2009. 11. 001.
- [6] Wajnsztajn D, Nche E, Solomon A. Corneal complications of rheumatoid arthritis [J]. Curr Opin Allergy Clin Immunol, 2022, 22(5): 304-313. DOI: 10. 1097/ACI. 0000000000000844.
- [7] Xu F, Qin Y, He W, et al. A deep transfer learning framework for the automated assessment of corneal inflammation on *in vivo* confocal microscopy images [J/OL]. PLoS One, 2021, 16(6): e0252653 [2023-06-12]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8174724/>. DOI: 10. 1371/journal.pone. 0252653.
- [8] Jeng BH, Ahmad S. In pursuit of the elimination of corneal blindness: is establishing eye banks and training surgeons enough? [J]. Ophthalmology, 2021, 128(6): 813-815. DOI: 10. 1016/j. ophtha. 2020. 06. 042.
- [9] Li Z, Jiang J, Chen K, et al. Preventing corneal blindness caused by keratitis using artificial intelligence [J/OL]. Nat Commun, 2021, 12(1): 3738 [2023-06-12]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8213803/>. DOI: 10. 1038/s41467-021-24116-6.
- [10] Kang L, Ballouz D, Woodward MA. Artificial intelligence and corneal diseases [J]. Curr Opin Ophthalmol, 2022, 33(5): 407-417. DOI: 10. 1097/ICU. 0000000000000885.
- [11] Qu JH, Qin XR, Peng RM, et al. A fully automated segmentation and morphometric parameter estimation system for assessing corneal endothelial cell images [J]. Am J Ophthalmol, 2022, 239: 142-153. DOI: 10. 1016/j. ajo. 2022. 02. 026.
- [12] McCarron ME, Weinberg RL, Izzi JM, et al. Combining *in vivo* corneal confocal microscopy with deep learning-based analysis reveals sensory nerve fiber loss in acute simian immunodeficiency virus infection [J]. Cornea, 2021, 40(5): 635-642. DOI: 10. 1097/ICO. 0000000000002661.
- [13] Mou L, Qi H, Liu Y, et al. DeepGrading: deep learning grading of corneal nerve tortuosity [J]. IEEE Trans Med Imaging, 2022, 41(8): 2079-2091. DOI: 10. 1109/TMI. 2022. 3156906.
- [14] Xu F, Jiang L, He W, et al. The clinical value of explainable deep learning for diagnosing fungal keratitis using *in vivo* confocal microscopy images [J/OL]. Front Med (Lausanne), 2021, 8: 797616 [2023-06-15]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8712475/>. DOI: 10. 3389/fmed. 2021. 797616.
- [15]《眼科人工智能临床应用伦理专家共识(2023)》专家组, 中国医药教育协会数字影像与智能医疗分会, 中国医药教育协会智能医学专业委员会. 眼科人工智能临床应用伦理专家共识(2023) [J]. 中华实验眼科杂志, 2023, 41(1): 1-7. DOI: 10. 3760/cma. j. cn115989-20220905-00414.
- Expert Workgroup of Expert consensus for ethics of clinical application of artificial intelligence in ophthalmology (2023), Digital Imaging and Intelligent Medicine Branch of China Medical Education Association, Intelligent Medicine Special Committee of China Medical Education Association. Expert consensus for ethics of clinical application of artificial intelligence in ophthalmology (2023) [J]. Chin J Exp Ophthalmol, 2023, 41(1): 1-7. DOI: 10. 3760/cma. j. cn115989-20220905-00414.
- [16] 周奕文, 于慧, 周亚标, 等. 睑板腺缺失面积的图像深度处理分析研究 [J]. 中华实验眼科杂志, 2020, 56(10): 774-779. DOI: 10. 3760/cma. j. cn112142-20200415-00272.
- Zhou YW, Yu Y, Zhou YB, et al. An advanced imaging method for measuring and assessing meibomian glands based on deep learning [J]. Chin J Ophthalmol, 2020, 56(10): 774-779. DOI: 10. 3760/cma. j. cn112142-20200415-00272.
- [17] Xie Y, Zhao L, Yang X, et al. Screening candidates for refractive surgery with corneal tomographic-based deep learning [J]. JAMA Ophthalmol, 2020, 138(5): 519-526. DOI: 10. 1001/jamaophthalmol. 2020. 0507.
- [18] Li Z, Guo C, Nie D, et al. Deep learning for detecting retinal detachment and discerning macular status using ultra-widefield fundus images [J/OL]. Commun Biol, 2020, 3(1): 15 [2023-06-20]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6949241/>. DOI: 10. 1038/s42003-019-0730-x.
- [19] 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.
- Dong XQ, Du SL, Liu HX, et al. Application of artificial intelligence for community-based diabetic retinopathy detection and referral [J]. Chin J Exp Ophthalmol, 2022, 40(12): 1158-1163. DOI: 10. 3760/cma. j. cn115989-20220316-00108.
- [20] Wong TY, Bressler NM. Artificial intelligence with deep learning technology looks into diabetic retinopathy screening [J]. JAMA, 2016, 316(22): 2366-2367. DOI: 10. 1001/jama. 2016. 17563.
- [21] 张弛, 叶子, 李朝辉. 高度近视白内障患者人工晶状体计算公式的研究进展 [J]. 中华实验眼科杂志, 2022, 40(5): 466-469. DOI: 10. 3760/cma. j. cn115989-20200512-00341.
- Zhang C, Ye Z, Li ZH. Advances in intraocular lens power calculation formulas in high myopia [J]. Chin J Exp Ophthalmol, 2022, 40(5): 466-469. DOI: 10. 3760/cma. j. cn115989-20200512-00341.
- [22] Yu Y, Zhou Y, Tian M, et al. Automatic identification of meibomian gland dysfunction with meibography images using deep learning [J]. Int Ophthalmol, 2022, 42(11): 3275-3284. DOI: 10. 1007/s10792-022-02262-0.
- [23] 于慧, 周奕文, 陈弟, 等. 基于深层卷积神经网络的眼科 B 型超声图像玻璃体视网膜疾病辅助诊断系统应用评估 [J]. 中华实验眼科杂志, 2021, 39(9): 792-797. DOI: 10. 3760/cma. j. cn115989-20210114-00037.
- Yu Y, Zhou YW, Chen D, et al. Evaluation of an aided diagnosis system for vitreous and retinal diseases by analyzing B-scan ultrasound images based on deep convolutional neural network [J]. Chin J Exp Ophthalmol, 2021, 39(9): 792-797. DOI: 10. 3760/cma. j. cn115989-20210114-00037.
- [24] Tiwari M, Piech C, Baitemirova M, et al. Differentiation of active corneal infections from healed scars using deep learning [J]. Ophthalmology, 2022, 129(2): 139-146. DOI: 10. 1016/j. ophtha. 2021. 07. 033.
- [25] Wei S, Shi F, Wang Y, et al. A deep learning model for automated sub-basal corneal nerve segmentation and evaluation using *in vivo* confocal microscopy [J/OL]. Transl Vis Sci Technol, 2020, 9(2): 32 [2023-06-20]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7414615/>. DOI: 10. 1167/tvst. 9. 2. 32.
- [26] Setu M, Schmidt S, Musial G, et al. Segmentation and evaluation of corneal nerves and dendritic cells from *in vivo* confocal microscopy images using deep learning [J/OL]. Transl Vis Sci Technol, 2022, 11(6): 24 [2022-06-23]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9251793/>. DOI: 10. 1167/tvst. 11. 6. 24.

(收稿日期: 2023-08-10 修回日期: 2023-12-27)

(本文编辑: 刘艳 施晓萌)