

基于深度学习的睑板腺功能障碍图像分析模型研究和评价

张祖辉¹, 于新新¹, 林晓蕾², 傅亚娜¹, 戴琦¹

引用:张祖辉, 于新新, 林晓蕾, 等. 基于深度学习的睑板腺功能障碍图像分析模型研究和评价. 国际眼科杂志 2022; 22(5): 746-751

基金项目:浙江省医药卫生科技计划项目 (No. 2022PY074, 2022KY217); 浙江省教育厅科研项目 (No. Y202147994)

作者单位:¹(325027) 中国浙江省温州市, 温州医科大学附属眼视光医院角膜病与屈光手术中心;²(200126) 中国上海市, 复旦大学附属耳鼻喉科医院白内障中心

作者简介:张祖辉, 毕业于温州医科大学, 硕士, 住院医师, 研究方向: 干眼、角膜病及屈光手术。

通讯作者:戴琦, 毕业于浙江大学, 博士, 副主任医师, 研究方向: 干眼、角膜病及屈光手术. dq@mail.eye.ac.cn

收稿日期: 2021-11-09 修回日期: 2022-04-06

摘要

目的:构建一个基于卷积神经网络 (CNN) 的人工智能 (AI) 系统, 能够全自动地评价睑板腺功能障碍 (MGD) 患者的睑板腺形态变化。

方法:选取 2021-01/11 在温州医科大学附属眼视光医院杭州院区就诊的 145 名受试者右眼纳入研究。随机选择其中 60 名受试者的睑板腺照相用于 AI 训练。收集睑板腺图像后首先标注出睑板区域和每一根睑板腺腺体。使用残差神经网络 (ResNet) 结合 U-Net 模型进行数据训练, 获得成熟的 AI 系统; 85 名受试者包括阻塞性 MGD 患者 53 名和睑板腺正常的志愿者 32 名, 使用 AI 系统自动分析其各项睑板腺形态参数。同时观察临床指标包括眼表疾病指数 (OSDI)、泪河高度 (TMH)、泪膜破裂时间 (TBUT)、角膜荧光素染色 (CFS)、睑缘评分、睑板腺评分和睑板腺分泌能力评分。分析睑板腺参数与临床指标的相关性。

结果:通过多次版本迭代, 最终获得了交并比达 92.0% 的 AI 系统。使用该 AI 系统, 发现上眼睑的睑板腺密度与 OSDI ($r_s = -0.320$)、TBUT ($r_s = 0.484$)、睑缘评分 ($r_s = -0.350$)、睑板腺评分 ($r_s = -0.749$) 和睑板腺分泌能力评分 ($r_s = 0.425$) 存在显著相关性 (均 $P < 0.05$); 下眼睑的睑板腺密度与 OSDI ($r_s = -0.420$)、TBUT ($r_s = 0.598$)、睑缘评分 ($r_s = -0.396$)、睑板腺评分 ($r_s = -0.720$) 和睑板腺分泌能力评分 ($r_s = 0.438$) 存在显著相关性 (均 $P < 0.05$); 总眼睑的睑板腺密度与 OSDI ($r_s = -0.404$)、TBUT ($r_s = 0.601$)、睑缘评分 ($r_s = -0.416$)、睑板腺评分 ($r_s = -0.805$) 和睑板腺分泌能力评分 ($r_s = 0.480$) 存在显著相关性 (均 $P < 0.05$)。

结论:基于 CNN 的 AI 系统是一个准确、高效的睑板腺形态学评价系统, 能够方便地采用我们建立的睑板腺密度这

一指标对 MGD 患者的睑板腺形态进行快速准确地评价。睑板腺密度这一指标比目前通用的睑板腺评分更精确, 是评价睑板腺萎缩程度的全新定量指标。

关键词:卷积神经网络; 人工智能; 睑板腺功能障碍; 睑板腺密度

DOI: 10.3980/j.issn.1672-5123.2022.5.10

Study and evaluation of image analysis model of meibomian gland dysfunction based on deep learning

Zu-Hui Zhang¹, Xin-Xin Yu¹, Xiao-Lei Lin², Ya-Na Fu¹, Qi Dai¹

Foundation items: The Zhejiang Provincial Medical and Health Science Technology Program of Health and Family Planning Commission (No. 2022PY074, 2022KY217); A Project Supported by Scientific Research Fund of Zhejiang Provincial Education Department (No. Y202147994)

¹Corneal Disease and Refractive Surgery Center, Eye Hospital of Wenzhou Medical University, Wenzhou 325027, Zhejiang Province, China; ²Cataract Center, Eye and ENT Hospital of Fudan University, Shanghai 200126, China

Correspondence to: Qi Dai. Corneal Disease and Refractive Surgery Center, Eye Hospital of Wenzhou Medical University, Wenzhou 325027, Zhejiang Province, China. dq@mail.eye.ac.cn

Received: 2021-11-09 Accepted: 2022-04-06

Abstract

• **AIM:** To construct an artificial intelligence (AI) system based on convolutional neural network (CNN), which can automatically evaluate the morphological changes of meibomian gland (MG) in meibomian gland dysfunction (MGD) patients.

• **METHODS:** The right eyes of 145 subjects who were treated at the Hangzhou Branch of the Eye Hospital of Wenzhou Medical University from January to November 2021 were selected for inclusion in the study. Meibography images of 60 of these subjects were randomly selected for AI training. The meibomian region and each MG in meibography were annotated and formed into datasets. The datasets were used for training and obtaining an AI system based on residual neural network (ResNet) combined with the U-Net model. The AI system was used to automatically analyze the MG morphological

parameters in 85 subjects, including 53 patients with obstructive MGD and 32 volunteers with normal meibomian glands. The clinical indices including ocular surface disease index (OSDI), tear meniscus height, tear film break-up time (TBUT), corneal fluorescein staining, lid margin score, meiboscore, and meibomian gland expressibility score were also observed. The correlation between MG morphological parameters and clinical indices were analyzed.

• **RESULTS:** After several iterations, we finally obtained an AI system with Intersection over Union of 92.0%. Using this AI system, we found that there was a significant correlation between the MG density in the upper eyelid with OSDI ($r_s = -0.320$), TBUT ($r_s = 0.484$), lid margin score ($r_s = -0.350$), meiboscore ($r_s = -0.749$), and meibum expressibility score ($r_s = 0.425$) (all $P < 0.05$). The MG density in the lower eyelid was significantly correlated with OSDI ($r_s = -0.420$), TBUT ($r_s = 0.598$), lid margin score ($r_s = -0.396$), meiboscore ($r_s = -0.720$), and meibum expressibility score ($r_s = 0.438$) (all $P < 0.05$). The MG density in the total eyelid was significantly correlated with OSDI ($r_s = -0.404$), TBUT ($r_s = 0.601$), lid margin score ($r_s = -0.416$), meiboscore ($r_s = -0.805$), and meibum expressibility score ($r_s = 0.480$) (all $P < 0.05$).

• **CONCLUSION:** The AI system based on CNN in this study is an accurate and efficient MG morphological evaluation system, which can be conveniently used to evaluate the MG morphology of MGD patients quickly and accurately by using the MG density index established by us. MG density is a new quantitative index to evaluate meibomian gland atrophy, which is more accurate than meiboscore.

• **KEYWORDS:** convolutional neural network; artificial intelligence; meibomian gland dysfunction; meibomian gland density

Citation: Zhang ZH, Yu XX, Lin XL, *et al.* Study and evaluation of image analysis model of meibomian gland dysfunction based on deep learning. *Guoji Yanke Zazhi (Int Eye Sci)* 2022; 22(5): 746-751

0 引言

睑板腺功能障碍 (meibomian gland dysfunction, MGD) 为一种慢性、弥漫性睑板腺病变,以睑板腺终末导管的阻塞和(或)睑板腺分泌物或量改变为主要病理特征,临床上可引起泪膜异常、眼部刺激症状、眼表炎症反应,从而影响患者的生活质量^[1]。我国的干眼患病率高达 31.40%^[2],而 Lemp 等^[3]发现超过 85%的干眼患者同时患有 MGD。因此,MGD 的正确诊断值得引起眼科医生的重视。睑板腺形态缺失作为 MGD 的诊断指标之一,具有重要意义^[4]。临床上主要通过睑板腺红外成像技术对睑板腺形态进行观察和评估,目前使用较多的评价睑板腺形态缺失的方法是 Arita 等^[5]的分级法,但其存在定性指标不够精确、人工分级工作量大、效率低、误差大的问题。

随着现代人工智能 (artificial intelligence, AI) 技术的

高速发展,AI 图像自动识别技术在不同领域的实践和应用已逐渐开展,医学领域也不例外,AI 图像识别技术在眼科中的应用日趋成熟^[6]。基于 CNN 的 AI 系统可以克服前述人工标注的问题,具有速度快、准确率高、重复性好的优势,已广泛应用于人脸识别、文字识别、图像检索等多个领域^[7-8],对于物体识别、图片分类的准确率可以接近甚至优于人工分析方式^[9-10],且成本大幅降低,是目前医学诊断当中最火热和最具有前景的发展方向。

本文拟探讨以基于卷积神经网络 (convolutional neural network, CNN) 的 AI 模型构建的图像识别系统应用于睑板腺红外照相图片处理。借助 AI 的运算能力优势,快速准确地识别出睑板腺的腺体。同时在 AI 的辅助下,提出全新的睑板腺萎缩形态定量指标——睑板腺密度 (MG density) 并验证其有效性。

1 对象和方法

1.1 对象 研究对象选取 2021-01/11 在温州医科大学附属眼视光医院杭州院区就诊的 145 名受试者,研究选取受试者的右眼。其中 60 名为随机选择,收集其睑板腺照用于 AI 训练。另招募 85 名受试者,包括阻塞性 MGD 患者 53 名和睑板腺正常的志愿者 32 名,以 AI 系统对其进行分析。阻塞性 MGD 患者均由两位经验丰富的眼科医生根据眼部症状、睑缘异常和睑板腺评分进行诊断,如果满足以下 3 个标准中的任何一个,则患者被诊断为阻塞性 MGD:(1)眼部症状评分 ≥ 3 分;(2)睑缘评分 ≥ 2 分;和(或)(3)睑板腺评分 ≥ 3 分^[11]。排除标准:(1)患有眼部疾病或已知影响眼前节解剖结构情况的患者,如有眼部炎症、眼部手术史、配戴隐形眼镜和(或)外伤史;(2)有影响睑板腺功能的全身用药史的患者;(3)患有任何其他已知会影响泪膜的眼部疾病或全身性疾病者。本研究经温州医科大学附属眼视光医院医学伦理委员会审批通过,所有受试者均签署知情同意书。

1.2 方法

1.2.1 图像处理 使用眼表分析仪拍摄 60 名随机受试者右眼上下眼睑的睑板腺图像作为 AI 模型训练。训练图像中的睑板腺均由两位高级职称医师手动标注,取其平均值。

1.2.2 AI 系统的建立 本研究使用的 CNN 结构是基于 U-Net 模型,并借助迁移学习的方法将 ImageNet 预训练好的模型和参数应用到我们的 U-Net 模型中。残差网络 (ResNet) 在图像分类和目标检测方面都有非常好的性能。通过改进原生 U-Net 模型,然后将 50 层的 ResNet (ResNet50) 替换 U-Net 的下采样部分,上采样部分保持不变,建立 ResNet50_U-Net 模型。网络结构见图 1。

由于原始训练图片数据集只有 60 名受试者,虽然达到了 U-net 网络的训练门槛,但是对于训练一个深度卷积神经网络而言数据仍显不足,因此我们调用了一个功能强大的开源数据增强库 imgaug (<https://github.com/aleju/imgaug#citation>) 来对训练图像进行多维增强。在训练中的每一次迭代均会在受试者上睑的原始训练图像中随机选取 4 张,调用 imgaug 随机使用算法和参数对输入图像进行增强,产生 4 张新图片。最终版本的模型在全部训练中共迭代 61440 次,产生 245760 张新图片作为训练数据

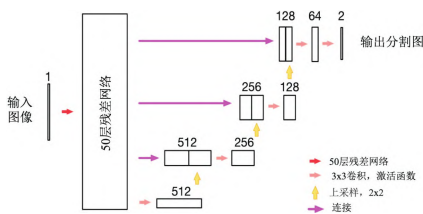


图1 AI模型的网络结构。

集,数据量完全满足了训练一个深度卷积神经网络的需求。训练完成后通过改变模型的目标区域提取,睑板腺分割AI模型无需重新训练即可应用于下睑睑板腺图像处理。

在模型的计算过程中,所有输入的图片都会被缩放到288×896像素的大小,然后输入到模型中,最终预测的图片会恢复到原始图片的大小。因此,如果原始图像的大小远远大于288×896像素,则预测结果图像中的腺体可能出现锯齿状。我们需要对结果图像进行高斯滤波,以获得更平滑的腺体。在得到图像的预测结果后,再通过OpenCV提供的传统图像处理方法提取预测图像中的每根睑板腺。提取步骤描述如下:(1)对图像进行中值滤波;(2)将图像二值化;(3)在图像中找到睑板腺的轮廓。如果两个等高线之间的垂直距离很小,则认为这些等高线属于同一个睑板腺并且是连接在一起的。睑板腺照相的原始图像经过图像预处理增强后,再经ResNet50_U-net预测滤波后得到平缓的睑板腺腺体。再与手工标注结果做比较(图2)。

1.2.3 睑板腺功能评价指标

1.2.3.1 眼表疾病指数 眼表疾病指数(ocular surface disease index, OSDI)^[11];所有受试者均接受OSDI问卷和眼部症状问卷调查。

1.2.3.2 症状评分 根据出现的MGD相关的眼部症状的数量(眼疲劳、分泌物、异物感、干燥、不适感、黏稠感、疼痛、溢泪、瘙痒、发红、沉重感、眩光、过度眨眼和睑板腺囊肿或睑腺炎),对症状从0~14分进行评分,分数越高,说明症状越明显。

1.2.3.3 泪河高度 泪河高度(tear meniscus height, TMH):使用眼表分析仪测量泪河高度并进行睑板腺红外拍照, TMH测量方法采用眨眼5s后测量下眼睑中央TMH,正常值在0.20mm左右。

1.2.3.4 泪膜破裂时间 泪膜破裂时间(tear film break-up time, TBUT)和角膜荧光素染色(corneal fluorescein staining, CFS)^[12]:滴入荧光素后进行CFS和测定TBUT。TBUT测定3次,取其平均值,正常值为>10s。CFS评分标准如下:在5个区域(中央、颞、鼻、上、下)评估角膜荧光素染色(0个点=0分,1~5个点=1分,6~15个点=2分,16~30个点=3分,>30个点=4分);如果发现融合染色,染料扩散到周围的上皮或基质,则根据融合区域的数量或角膜丝状物的存在进行评分(一个融合区域,加1分;两个或两个以上融合处,加2分),染色评分范围为0~20分,分数越高,说明角膜上皮损伤越重。

1.2.3.5 睑缘评分 根据黏膜交界处前后移位、血管充盈、睑板腺孔堵塞、眼睑边缘不规则4项参数对眼睑边缘异常进行评分,每项评分0~4分,分数越高说明睑缘异常程度越大。

1.2.3.6 睑板腺分泌能力评分^[13] 0分:清亮的睑脂容易排出;1分:轻微压力下排出混浊的睑脂;2分:超过中等压力下才能排出混浊的睑脂;3分:即使很大的压力下也无睑脂排出。睑板腺检查器(meibomian gland evaluator, MGE)评估下眼睑15个腺体分泌物的数量和质量。每个眼睑睑板腺分泌能力评分范围为0~45分,分数越低说明睑板腺分泌功能越接近正常。

1.2.3.7 睑板腺评分 睑板腺评分(meiboscore)^[5]:0分:无萎缩;1分:萎缩面积<1/3睑板总面积;2分:萎缩面积介于1/3~2/3的总睑板面积;3分:萎缩的>2/3的睑板总面积。将上下睑的睑板腺评分相加,得到每只眼睛的睑板腺评分范围从0到6,分数越高,说明睑板腺萎缩越严重。

1.2.4 睑板腺密度 睑板腺密度(MG density):计算每根睑板腺的面积(Smg)(以像素为单位),然后用每个睑板上所有腺体面积总和与睑板的面积(St)做对比,公式如下:

$$Density = \frac{\sum Sm g_i}{St}$$

统计学分析:采用SPSS23.0统计软件。符合正态分布的计量资料以 $\bar{x} \pm s$ 表示,组间比较采用两独立样本t检验;不符合正态分布的计量资料以 $M(P_{25}, P_{75})$ 表示,组间比较采用两独立样本Mann-Whitney U检验。由于MGD患者与正常组的年龄差异显著,故采用广义估计方程对年龄进行校正。采用Spearman秩相关分析睑板腺形态指标与功能指标的相关性。 $P < 0.05$ 为差异有统计学意义。

2 结果

本文采用IoU这一指标对睑板腺识别模型的准确率进行评价,最终,经过4h的训练,在同一份训练和原始验证样本中,IoU提升到了92%,训练完成。系统重复率达到100%,并且效率极高,在使用gtx1070 8G的GPU的情况下,分析一张睑板腺图像仅仅需要100ms。将85名受试者的85只右眼纳入AI自动分析研究。MGD患者及正常组的临床参数比较见表1。经年龄校正后的MGD组患者的MGD相关症状评分明显高于正常组,差异有统计学意义($P < 0.001$),OSDI评分明显高于正常组,差异有统计学意义($P < 0.001$)。MGD组TBUT明显低于正常组,差异有统计学意义($P < 0.001$)。与正常组相比,MGD组患者CFS实际参数值分布较高,但两组中位数都是0(0,0),然而矫正年龄之后,差异有统计学意义($P = 0.021$)。睑板腺状态,包括睑缘评分、睑板腺分泌能力评分和睑板腺评分,在MGD患者中明显比正常组更严重,差异均有统计学意义(均 $P < 0.001$)。两组受试者上眼睑、下眼睑和总眼睑的睑板腺密度和睑板腺评分比较,差异均有统计学意义(均 $P < 0.001$)。对于人眼难以区分的更复杂的图片,该模型显示出优秀的识别能力,见图2,表2。研究结果表明,上眼睑的睑板腺密度与OSDI($r_s = -0.320, P = 0.003$)、TBUT($r_s = 0.484, P < 0.001$)、睑缘评分($r_s = -0.350, P = 0.001$)、睑板腺评分($r_s = -0.749, P < 0.001$)和睑板腺分泌能力评分($r_s = 0.425, P < 0.001$)显著相关。下眼睑的睑板腺密度与OSDI($r_s = -0.420, P < 0.001$)、TBUT($r_s = 0.598, P < 0.001$)、睑缘评分($r_s = -0.396, P < 0.001$)、睑板腺评分($r_s = -0.720, P < 0.001$)和睑板腺分泌能力评分($r_s = 0.438, P < 0.001$)显著相关。总眼睑的睑板腺密度与OSDI($r_s = -0.404, P < 0.001$)、TBUT($r_s = 0.601, P < 0.001$)、睑缘评分

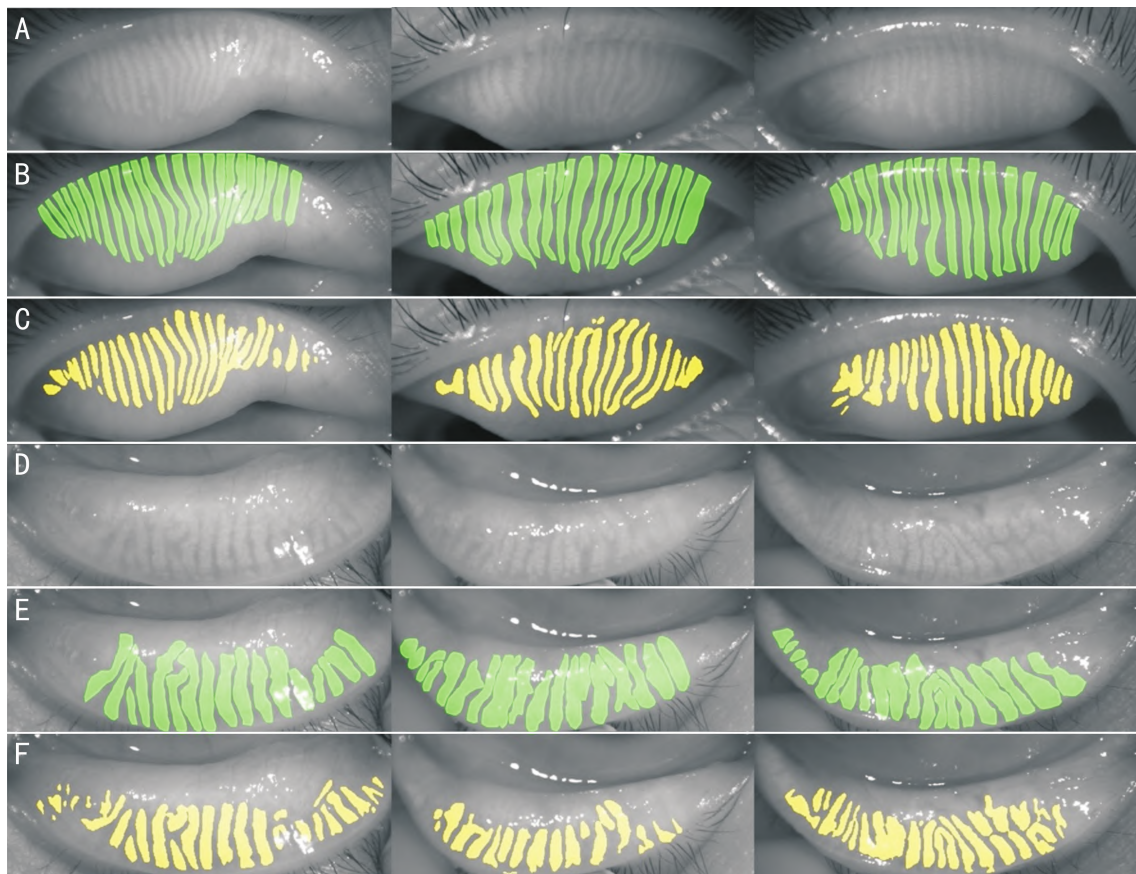


图2 睑板腺照相的原始图像经过图像预处理增强后,再经 ResNet50_U-net 预测滤波后得到平缓的睑板腺腺体与手工标注结果比较 A、D:原始图片;B、E:手工标注腺体(绿色部分);C、F:机器标注腺体(黄色部分)。

表1 两组受试者临床参数比较

参数	正常组(n=32)	MGD组(n=53)	统计量值	P	P*
年龄[M(P ₂₅ ,P ₇₅),岁]	25.00(16.25,32.75)	35.00(30.00,50.00)	376.500	<0.001	-
性别(男/女,名)	13/19	20/33	0.262	0.794	-
OSDI[M(P ₂₅ ,P ₇₅),分]	4.47(0.30,12.35)	25.00(13.24,37.80)	263.500	<0.001	<0.001
症状评分[M(P ₂₅ ,P ₇₅),分]	2.00(0,4.00)	7.00(5.00,8.00)	232.500	<0.001	<0.001
泪膜破裂时间[M(P ₂₅ ,P ₇₅),s]	5.00(5.00,7.75)	2.50(1.33,3.67)	252.500	<0.001	<0.001
角膜荧光素染色[M(P ₂₅ ,P ₇₅),分]	0(0,0)	0(0,0)	717.000	0.058	0.021
泪河高度[M(P ₂₅ ,P ₇₅),mm]	0.19(0.16,0.23)	0.20(0.17,0.24)	767.000	0.461	0.871
睑缘评分[M(P ₂₅ ,P ₇₅),分]	0(0,1.00)	2.00(1.00,2.00)	235.000	<0.001	<0.001
睑板腺评分[M(P ₂₅ ,P ₇₅),分]	2.00(1.00,2.00)	3.00(2.00,4.50)	221.000	<0.001	<0.001
睑板腺分泌能力评分[M(P ₂₅ ,P ₇₅),分]	38.50(30.00,45.00)	18.00(5.50,34.50)	335.500	<0.001	<0.001

注: * ;通过广义估计方程校正年龄差异后的P值。

表2 两组受试者睑板腺密度与睑板腺评分比较

睑板腺参数	正常组(n=32)	MGD组(n=53)	统计量值	P	M(P ₂₅ ,P ₇₅) P*	
上眼睑	睑板腺评分(分)	1.00(0,1.00)	1.00(1.00,2.00)	338.000	<0.001	<0.001
	睑板腺密度	0.30(0.27,0.32)	0.21(0.17,0.27)	273.000	<0.001	<0.001
下眼睑	睑板腺评分(分)	1.00(0,1.00)	2.00(1.00,2.50)	335.000	<0.001	<0.001
	睑板腺密度	0.32(0.28,0.41)	0.21(0.12,0.24)	7.299	<0.001	<0.001
总眼睑	睑板腺评分(分)	2.00(1.00,2.00)	3.00(2.00,4.50)	221.000	<0.001	<0.001
	睑板腺密度	0.31(0.29,0.34)	0.21(0.16,0.25)	8.124	<0.001	<0.001

注: * ;通过广义估计方程校正年龄差异后的P值。

($r_s = -0.416, P < 0.001$)、睑板腺评分($r_s = -0.805, P < 0.001$)和睑板腺分泌能力评分($r_s = 0.480, P < 0.001$)显著相关。其中上眼睑的睑板腺密度($r_s = -0.749$)、下眼睑的

睑板腺密度($r_s = -0.720$)及总眼睑的睑板腺密度($r_s = -0.805$)与睑板腺评分呈显著的负相关(均 $P < 0.001$)，见图3。

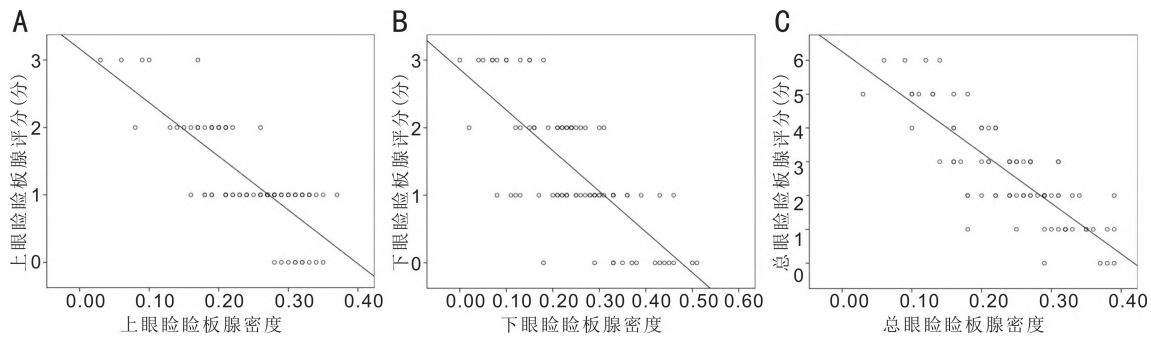


图3 睑板腺密度与睑板腺评分相关性 A:上眼睑睑板腺密度与上眼睑睑板腺评分的相关性;B:下眼睑睑板腺密度与下眼睑睑板腺评分的相关性;C:总眼睑睑板腺密度与总眼睑睑板腺评分的相关性。

3 讨论

目前,由于现有医疗资源在人力、物力方面等因素限制,我国 MGD 诊断仍存在部分医师诊疗水平较低、估算粗略片面等问题,不利于 MGD 和干眼诊疗精细化和规范化的开展^[14]。基层眼科医生需要一种简单方便、高效快捷、分析成本低,社会效益高的干眼和 MGD 相关指标的标准化评估方法。

其中的重要指标之一是睑板腺形态分析^[4,15-16]。本研究通过提出全新的睑板腺形态定量评价指标——睑板腺密度,比较正常组和 MGD 组睑板腺密度的差异。结果显示,MGD 组患者睑板腺密度较正常组明显减小。与之前的研究^[17-19]相似,本研究中 AI 系统显示睑板腺萎缩程度与 OSDI、眼表症状、泪膜稳定性、睑缘评分和睑板腺分泌能力均相关,说明睑板腺形态与功能存在密切联系。与以往的研究^[5,20-22]不同,本研究中 AI 系统能够准确地得到睑板腺的缺失面积,获得睑板腺密度这一连续性的定量指标,减少了由于睑板腺评分这一定性分级指标在分级过渡区附近的误差而导致的分级系统的不一致性和可变性。睑板腺密度与睑板腺评分高度线性相关,且 AI 系统分析得到的睑板腺密度与 OSDI、眼表症状、泪膜稳定性、睑缘评分和睑板腺分泌能力均相关,说明睑板腺密度可以代替睑板腺评分作为 MGD 诊断的有效指标。但是,如果想要在人工分析中采用我们的这种睑板腺分析方法是不现实的,因为这种分析方法需要大量的标注,不可能作为常规的评估指标。但是 AI 系统的特点恰恰适合这种需要大量分析工作的任务,因此我们的这种睑板腺分析方法只能在 AI 的帮助下才能实现并应用于临床。

本研究通过一个基于 CNN 的 AI 系统,达到了使睑板腺密度指标在临床工作中实用化的目的。该系统可以自动分析睑板腺的形态,极大地提高分析的准确性和效率,降低分析的成本,克服了人工标注主观上的误差。我们通过三代网络模型的迭代^[23-24],借助迁移学习的方法将 ImageNet 上预训练好的模型和参数应用到我们本研究中的第三代模型:ResNet50_U-Net 模型中。通过这一方法有效地降低了样本量的需求,尽管这个研究的样本数量相对较少,最终对睑板腺形态提取的准确性仍然较高 (IoU = 92%,重复性 = 100%)。目前我们仅使用一张 GTX1070 GPU 已经可以在 100ms 内处理一张睑板腺图片,随着算力的提高,使得数万幅图像在毫秒内同时完成分析成为可能。我们正在准备更大的训练数据集来支持更复杂的模

型,如 ResNet101^[10],FCN32^[25],SegNet^[26]和 PSPNet^[27],从而进一步提高腺体识别的准确性。

本研究也有局限性。研究中使用的 ResNet50_U-net 混合网络,是 CNN 的一种。CNN 倾向于将像素损失最小化。像素损失最小化策略补全矫正了腺体密度这一指标,较人工标注更为准确。但同时也导致靠的比较近的腺体之间无法自动分割,发生了粘连,从而被判断为一根腺体。这会导致除了睑板腺密度以外的、跟单根腺体形态有关的如腺体的数量、弯曲度等指标的失实,导致在我们前期研究中已被证实在早期 MGD 诊断中有效的睑板腺弯曲度 (MG tortuosity)^[28]在本研究中失去了作用。如果能够对算法进一步的发展,增强腺体的分割能力,则睑板腺形态识别有望获得更高的准确率。数据增强库 imgaug 能够部分改变原始图像的大量信息,是可以用于深度学习网络训练的,并且大大减轻了标注的工作量。但它仍然无法改变图片的一些基本信息,比如腺体的个数、密度等,因此不能完全替代全新的标注图片,未来我们的深度学习模型想要再进一步提升准确率,需要更多的全新图片进行训练。此外,本研究的样本量也较小,未来的研究中,我们团队将招募更多的受试者来训练和测试 AI 系统。

综上所述,基于 CNN 的 AI 系统是一个准确、高效的睑板腺形态学评价系统,能够方便地采用我们建立的睑板腺密度这一指标对 MGD 患者的睑板腺形态进行快速准确的评价。睑板腺密度这一指标比目前通用的睑板腺评分更精确,是评价睑板腺萎缩程度的全新定量指标。

参考文献

- 1 Vu CHV, Kawashima M, Yamada M, et al. Influence of meibomian gland dysfunction and friction-related disease on the severity of dry eye. *Ophthalmology* 2018;125(8):1181-1188
- 2 Song PG, Xia W, Wang ML, et al. Variations of dry eye disease prevalence by age, sex and geographic characteristics in China: a systematic review and meta-analysis. *J Glob Health* 2018;8(2):020503
- 3 Lemp MA, Crews LA, Bron AJ, et al. Distribution of aqueous-deficient and evaporative dry eye in a clinic-based patient cohort: a retrospective study. *Cornea* 2012;31(5):472-478
- 4 亚洲干眼协会中国分会, 海峡两岸医药交流协会眼科专业委员会眼表与泪液病学组. 我国睑板腺功能障碍诊断与治疗专家共识 (2017 年). *中华眼科杂志* 2017;53(9):657-661
- 5 Arita R, Itoh K, Inoue K, et al. Noncontact infrared meibography to document age-related changes of the meibomian glands in a normal population. *Ophthalmology* 2008;115(5):911-915
- 6 赵乾, 沈琳琳, 赖铭莹. 基于机器学习的人工智能技术在眼科中的

应用进展. 国际眼科杂志 2018;18(9):1630-1634

7 Pan YH, Huang WM, Lin ZP, et al. Brain tumor grading based on Neural Networks and Convolutional Neural Networks. *Annu Int Conf IEEE Eng Med Biol Soc* 2015;2015:699-702

8 Chen XY, Xu YW, Wong DWK, et al. Glaucoma detection based on deep convolutional neural network. *Annu Int Conf IEEE Eng Med Biol Soc* 2015;2015:715-718

9 Taigman Y, Yang M, Ranzato M, et al. DeepFace: closing the gap to human-level performance in face verification. *2014 IEEE Conference on Computer Vision and Pattern Recognition* 2014:1701-1708

10 He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition*. 2016:770-778

11 Arita R, Itoh K, Maeda S, et al. Proposed diagnostic criteria for obstructive meibomian gland dysfunction. *Ophthalmology* 2009;116(11):2058-2063.e1

12 De Paiva CS, Pflugfelder SC. Corneal epitheliopathy of dry eye induces hyperesthesia to mechanical air jet stimulation. *Am J Ophthalmol* 2004;137(1):109-115

13 Lane SS, DuBiner HB, Epstein RJ, et al. A new system, the LipiFlow, for the treatment of meibomian gland dysfunction. *Cornea* 2012;31(4):396-404

14 刘祖国. 干眼的临床诊断与治疗亟待规范化和精细化. 中华眼科杂志 2017;53(9):641-644

15 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. 中华眼科杂志 2020;56(10):774-779

16 Deng YQ, Wang Q, Luo ZZ, et al. Quantitative analysis of morphological and functional features in Meibography for Meibomian Gland Dysfunction: diagnosis and Grading. *E Clin Med* 2021;40:101132

17 Pult H, Riede-Pult BH, Nichols JJ. Relation between upper and lower lids' meibomian gland morphology, tear film, and dry eye. *Optom Vis Sci* 2012;89(3):310-315

18 Green-Church KB, Butovich I, Willcox M, et al. The international workshop on meibomian gland dysfunction: report of the subcommittee on tear film lipids and lipid-protein interactions in health and disease. *Invest Ophthalmol Vis Sci* 2011;52(4):1979-1993

19 Nelson JD, Shimazaki J, Benitez-del-Castillo JM, et al. The international workshop on meibomian gland dysfunction: report of the definition and classification subcommittee. *Invest Ophthalmol Vis Sci* 2011;52(4):1930-1937

20 García-Marqués JV, García-Lázaro S, Talens-Estrelles C, et al. Diagnostic Capability of a New Objective Method to Assess Meibomian Gland Visibility. *Optom Vis Sci* 2021;98(9):1045-1055

21 Nichols JJ, Berntsen DA, Mitchell GL, et al. An assessment of grading scales for meibography images. *Cornea* 2005;24(4):382-388

22 Pult H, Nichols JJ. A review of meibography. *Optom Vis Sci* 2012;89(5):760-769

23 朱敏颖, 林晓蕾, 张祖辉, 等. 基于卷积神经网络的睑板腺形态人工智能分析系统的构建. 浙江医学 2021;43(18):1946-1952

24 Dai Q, Liu XY, Lin XL, et al. A novel meibomian gland morphology analytic system based on a convolutional neural network. *IEEE Access* 2021;9:23083-23094

25 Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017;39(4):640-651

26 Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017;39(12):2481-2495

27 Zhao HS, Shi JP, Qi XJ, et al. Pyramid scene parsing network. *2017 IEEE Conference on Computer Vision and Pattern Recognition* 2017:6230-6239

28 Lin XL, Fu YN, Li L, et al. A novel quantitative index of meibomian gland dysfunction, the meibomian gland tortuosity. *Transl Vis Sci Technol* 2020;9(9):34