

· 实验研究 ·

基于人工智能引导的玻璃体内注射机器人系统 对注射点识别和运动控制精度评估

陈璟文¹ 庞义杰¹ 袁进² 唐晓颖^{1,3}¹南方科技大学电子与电气工程系,深圳 518055;²首都医科大学附属北京同仁医院眼科,北京 100730;³南方科技大学嘉兴研究院,嘉兴 314012

通信作者:唐晓颖,Email:tangxy@sustech.edu.cn

【摘要】 目的 开发基于人工智能(AI)引导的玻璃体内注射机器人系统,实现对眼表注射点精准检测,并通过 3D 位置解算引导机械臂完成玻璃体内注射定位任务。**方法** 纳入 TEyeD 数据集的 Dikablis 子集,采用等间隔时间采样策略构建训练集、验证集、测试集。系统利用 RGBD 相机读取眼表彩色 RGB 图像,接着使用基于 PatchCrop-Transformer 的注射点检测算法对图像进行瞳孔、虹膜、眼睑等关键点检测与定位,然后根据相机获取的深度信息提取注射点附近局部 3D 点云数据,通过对局部区域点云数据进行主成分分析(PCA),确定注射点和注射方向。将这些关键信息传递给机械臂系统,机械臂末端采用远程中心运动(RCM)机构,经过正逆运动学求解得到关节运动路径,控制机械臂移动至注射点上方 2 cm 处,再经过医师确认后完成插入、注射及拔出等操作,确保注射过程的稳定性和重复性。对本研究方法与 NFDP、SLPT 和 StarLoss 等前沿技术进行关键点定位的均方误差(MSE)和不同像素(2、5、10 个像素)误差允许范围内成功检测率(SDR)比较,并评估随机权重增强、固定权重增强和无增强方法对各关键点定位 MSE 的影响。评估机械臂系统的重复定位精度和绝对定位精度。**结果** 本研究模型在加入了随机权重增强之后,在均方误差和成功检测率方面均优于固定权重增强和无增强方法。对眼部整体、瞳孔和虹膜定位的 MSE 分别为 4.25、2.41 和 1.54,低于 NFDP、Starloss 和 SLPT 方法。在 5 像素和 10 像素的误差允许范围内,SDR 分别为 72.09% 和 92.68%,均高于 NFDP、Starloss 和 SLPT 方法。机械臂单轴的重复定位误差和绝对定位误差均在 $\pm 5 \mu\text{m}$ 范围内。**结论** 基于 AI 引导的玻璃体内注射机器人系统融合 RGBD 图像实现了眼部注射点的自动识别,并通过 RCM 机构设计与相应的运动学求解方法实现了高精度运动控制。

【关键词】 玻璃体内注射;人工智能;注射点检测;眼科手术操作臂;远程中心运动机构

基金项目: 国家重点研发计划(2023YFC2415400);国家自然科学基金(T2422012、62071210);广东省基础与应用基础研究计划(2024B1515020088);深圳市科技计划(RCYX20210609103056042);深圳南方科技大学高水平专项基金(G030230001、G03034K003)

DOI:10.3760/cma.j.cn115989-20250610-00194

Evaluation of injection point recognition and motion control accuracy of an intravitreal injection robot system guided by artificial intelligence

Chen Jingwen¹, Pang Yijie¹, Yuan Jin², Tang Xiaoying^{1,3}¹Southern University of Science and Technology, Shenzhen 518055, China; ²Ophthalmology Department of Beijing Tongren Hospital, Capital Medical University, Beijing 100730, China; ³Jiaxing Institute of Southern University of Science and Technology, Jiaxing 314012, China

Corresponding author: Tang Xiaoying, Email: tangxy@sustech.edu.cn

[Abstract] Objective To develop an artificial intelligence (AI)-guided intravitreal injection robot system to accurately detect the injection point on the ocular surface and guide the robotic arm to complete the intravitreal injection positioning task through 3D position calculation. **Methods** The Dikablis subset of the TEyeD dataset was used. Training set, testing set, and validation set were constructed by using equal interval sampling strategy. The system read the ocular surface color RGB image with an RGBD camera, then used a PatchCrop-Transformer-based injection point detection algorithm to detect and locate key points such as the pupil, iris, and eyelid in the image.

Next, it extracted the local 3D point cloud data near the injection point based on the depth information obtained by the camera. Through principal component analysis (PCA) of the local area point cloud data, the injection point and injection direction were determined. The key information was then passed to the robotic arm system. The end of the robotic arm adopted a remote center of motion (RCM) mechanism. After solving the forward and inverse kinematics, the joint movement path was obtained, and the robotic arm was controlled to move to 2 cm above the injection point. After confirmation by the doctor, the insertion, injection, and withdrawal operations were completed to ensure the stability and repeatability of the injection process. The mean square error (MSE) of key points localization and the success detection rate (SDR) within different pixel error ranges (2, 5, and 10 pixels) of the study method were compared with those of the NFDP, SLPT, and StarLoss methods, and the effects of random weight enhancement, fixed weight enhancement, and no enhancement methods on the MSE of key points localization were evaluated. The repeatability and absolute positioning accuracy of the robotic arm system were also evaluated. **Results** After adding random weight enhancement, the model of this study outperformed the fixed weight enhancement and no enhancement methods in both MSE and SDR. The MSEs of the model proposed in this study for overall eye, pupil, and iris localization were 4.25, 2.41, and 1.54, respectively, which were lower than those of the NFDP, StarLoss, and SLPT methods. Within the error ranges of 5 and 10 pixels, the SDRs of the model proposed in this study were 72.09% and 92.68%, respectively, which were higher than those of the NFDP, StarLoss, and SLPT methods. The single-axis repeatability errors and absolute positioning errors of the robotic arm were within $\pm 5 \mu\text{m}$. **Conclusions** The AI-guided intravitreal injection robot system integrates RGBD images to achieve automatic recognition of the ocular injection point and high-precision motion control through RCM mechanism design and corresponding kinematic solution methods.

[Key words] Intravitreal injection; Artificial intelligence; Injection point detection; Ophthalmic surgical manipulator; Remote center of motion mechanism

Fund program: National Key Research and Development Program (2023YFC2415400); National Natural Science Foundation of China (T2422012, 62071210); Guangdong Provincial Basic and Applied Basic Research Program (2024B1515020088); Shenzhen Science and Technology Program (RCYX20210609103056042); High-level Special Fund of Southern University of Science and Technology, Shenzhen (G030230001, G03034K003)

DOI: 10.3760/cma.j.cn115989-20250610-00194

玻璃体内注射是当前治疗眼底疾病的重要介入手段,可实现抗血管内皮生长因子药物、糖皮质激素等药物的精准递送,已被广泛应用于多种视网膜血管性疾病的临床治疗中,包括年龄相关性黄斑变性(age-related macular degeneration, AMD)^[1]、糖尿病性黄斑水肿(diabetic macular edema, DME)^[2-3]等。眼科手术对操作精度和稳定性的要求极高,临床规范要求注射针垂直入眼,注射位点需准确定位于角膜缘后 3.5~4.0 mm,以最大限度避免对晶状体的损伤^[4]。在传统操作中,医师可能因手部震颤或操作疲劳导致操作误差。近年来,随着精准医疗和智能技术的发展,手术机器人系统凭借其亚毫米级的定位精度和高稳定性的操作优势,正在推动玻璃体内注射手术流程向自动化、标准化方向演进。这一变革不仅提升了微创手术的一致性和安全性,也拓宽了眼科高精度操作的临床应用边界。

目前,国际上已有多个代表性注射机器人系统取得了突破性进展。例如,荷兰 Preceyes 主从式玻璃体视网膜手术机器人在视网膜下药物注射等操作中表现

出高精度表现,具备极高的位置分辨率,并集成术中光学相干断层扫描图像辅助功能,实现工具尖端位置的实时追踪^[9];哈尔滨工业大学团队则研发了面向玻璃体视网膜手术的主从式机器人,在猪眼和鸡胚模型中完成了视网膜血管插管与黄斑前膜剥离等模拟实验^[10];此外,多次迭代优化后的 Micron 手持式微创操作系统,在精度控制、视觉辅助与操作流畅性方面均取得了显著进展^[11-13]。在关键点检测方面,目前主流方法包括基于坐标回归^[14-15]和热图回归^[16-17]的算法,这 2 类方案在不同视觉任务中均取得了良好效果,尤其在人脸识别等通用视觉数据集上已得到广泛验证。

尽管现有的主从式与手持式眼科手术机器人技术均已取得一定成果,但仍存在局限性:其一,操作依赖医师持续遥控,系统自主性不足;其二,视觉感知与机械臂控制尚未实现深度融合,智能化水平有待进一步提升。因此,本研究旨在开发一种融合 AI 智能分析与 3D 视觉感知的注射机器人系统,通过关键点自动检测与空间位姿解算,引导具备远程中心运动(remote center of

motion, RCM) 机构的机械臂自主完成玻璃体内注射操作, 以提升眼科机器人系统的智能化与操作精度, 为未来复杂眼科手术的自动化提供技术基础与方法参考。

1 数据集与方法

1.1 实验数据集

由于眼部注射点周围纹理较弱, 系统难以稳定标注和模型训练, 本研究提出基于眼部关键点检测获取注射点位置的算法。在模型性能评估方面采用 TEyeD 数据集中的 Dikablis 子集对模型性能进行评估。该子集覆盖多种注视方向和光照条件, 提供了完整的眼部关键点标注和分割图, 支持对虹膜、瞳孔及眼睑等关键结构的精确定位。Dikablis 子集共标注 50 个关键点, 其中虹膜和瞳孔各 8 个, 眼睑 34 个。

鉴于原始数据体量较大, 且视频帧间存在高度时序冗余, 本研究引入等间隔时间采样策略以提升模型训练效率; 最终构建的数据集包括 3 个子集: 训练集 (图像 37 700 张)、验证集 (图像 4 328 张)、测试集 (图像 4 328 张)。

1.2 主要设备仪器

Microtrak II 激光测距仪 (美国 MTI Instruments 公司), 自主搭建的五轴机械臂、仿真人模型 (上海康为医疗科技发展有限公司) 及 LiDAR Camera L515 相机 (美国英特尔 RealSense 公司)。其中, Microtrak II 激光测距仪具备高精度测量能力, 可精确测定机械臂位置误差; 自主搭建的基于 RCM 机构的五轴机械臂操作灵活, 能精准完成多轴联动的手术操作任务; LiDAR Camera L515 可实时进行深度测量及三维成像, 为机械臂提供三维坐标信息, 引导机械臂完成相应的注射任务。

1.3 方法

1.3.1 系统工作流程 如图 1 所示, 本系统首先对训练集二维图像中的虹膜和瞳孔进行直方图匹配预处理, 随后输入 PatchCrop-Transformer 模型中进行训练, 得到注射点。模型训练完成之后, 通过测试集来检验模型定位注射点的性能。为进一步提升三维空间中注射点定位的准确性, 系统在模型预测的注射点周围 2 mm 范围内提取局部点云数据。该区域点云包含丰富的空间结构信息, 系统通过主成分分析 (principal components analysis, PCA) 精确确定注射点的三维坐标及其入射方向, 从而推算出最优的注射角度与位置。最后在仿真人实验中, 通过 3D 相机采集眼部三维图像数据, 并将其中的 RGB 通道彩色图像输入至已经训练好的基于 PatchCrop-Transformer 模型中, 得到注射点坐标, 再通过 PCA 推算出最优的注射角度与位置,

系统将上述获取的关键眼部注射点位置信息传输至机械臂系统。

机械臂末端采用 RCM 机构设计, 能够有效保障注射过程中的稳定性与精确性。系统基于正、逆运动学模型计算机械臂各关节运动轨迹, 确保其能够精准移动至目标注射位置。在正式进行注射操作前, 机械臂首先移动至注射点上方约 2 cm 处的预设待命位置, 此高度设置兼顾操作空间及安全性, 为后续注射过程提供缓冲与调整空间。系统在此位置暂停后, 医师可通过软件界面对注射角度与位置进行微调, 以确保操作的安全性及临床可控性。系统通过全过程的高精度控制, 最大限度降低人为误差, 确保注射操作的安全性与有效性。

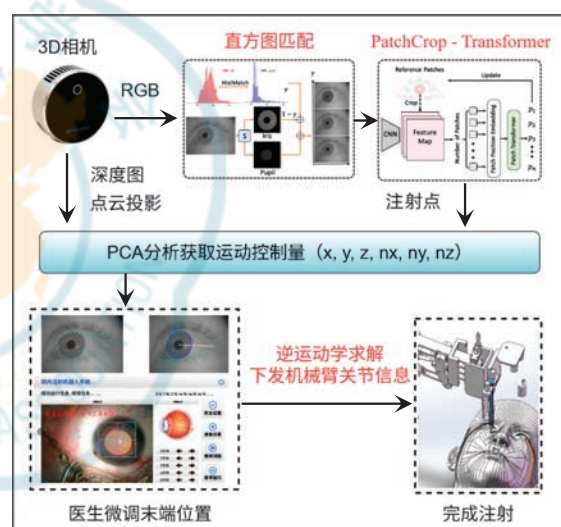


图 1 系统框架图 PCA: 主成分分析

Figure 1 System framework diagram PCA: principal components analysis

1.3.2 数据制作与迁移训练 TEyeD 数据集以欧洲血统个体眼部图像为主, 模型在其上训练后应用于中国血统个体眼部图像时性能欠佳, 主要由于中国人与欧洲人的虹膜和瞳孔颜色对比度存在差异, 导致模型在注射点检测方面存在差异。为解决此问题, 本研究团队创新性地运用直方图匹配技术进行色彩迁移以缩小欧美人群与中国人群眼球瞳孔和虹膜的色彩差异, 进而提升模型处理中国人群眼部图像的性能。

直方图匹配通过计算原图像与目标图像的颜色直方图, 依据累积分布函数等映射关系完成颜色迁移, 其过程可表示为 $I_m = HM(I_0, H(I_0), H(P_0))$, 其中 I_0 和 P_0 分别表示原始虹膜和瞳孔区域, $H(\cdot)$ 表示直方图分析, $HM(\cdot)$ 表示直方图匹配过程。为使数据集能模拟虹膜与瞳孔对比度的多种情形, 提出将原始虹膜

图像和经直方图匹配处理后的图像进行随机加权融合的方法,生成有平滑过渡效果的合成图像,可通过调整权重获得不同区分度的中国人群眼球图像,为模型训练提供更具代表性的样本。该方法的数学表达式为 $I_f = (1-\gamma)I_0 + \gamma I_m$, 其中 γ 是控制图像融合程度的随机参数,当 γ 为 1 时虹膜和瞳孔无色彩差异,当 γ 为 0 时则虹膜瞳孔色彩差异最大,该过程可以增强模型在不同虹膜瞳孔对比度条件下的鲁棒性。

1.3.3 模型设计 (1) 补丁特征选择 在视觉模型架构领域, Vision Transformer 作为一种新兴且备受瞩目的技术,通常会将输入图像划分为多个不重叠的图像块进行处理。然而,这种划分方式也面临着一个挑战:当增大图像块的行数(M)和列数(N)时,虽然有助于提升特征提取的粒度以及模型的预测性能,但也会显著增加计算复杂度。在关键点检测任务中,背景区域对于关键点定位的贡献相对较小,因此我们提出了一种策略,即仅保留与关键点直接相关的图像区域,从而得到更紧凑的特征表示。具体来说,首先将输入的图像通过 2 个级联的卷积神经网络模块进行预处理,生成高维特征图,随后依据参考关键点的位置信息,从高维特征图中裁剪获得相应局部区域。为了适应潜在的位置偏差,在初始阶段采用了较大的裁剪窗口。此外,为了保留眼部关键点间的相对位置信息,引入了可学习的位置参数,即结构编码。该结构编码的设计有助于提高关键点定位的准确性,同时使模型学习不同关键点间的相关性。通过这种补丁特征选择的方法,可在降低计算复杂度的同时提升关键点检测的性能。

(2) 内在关系层 内在关系层由多头自注意力模块和多头交叉注意力模块构成,保障模型对局部和全局特征的学习效果,提高关键点预测的准确性。注意力机制公式为 $A = \text{softmax} \left[\frac{(\hat{X}W^q)(XW^k)^T}{\sqrt{d_k}} \right]$, 其中 X 表示特征点切片的扁平化张量向量, $\hat{X}$ 在多头自注意力模块中与 X 相同,在多头交叉注意力模块中作为不同特征点切片的查询向量, d_k 为特征向量维度, W^q 和 W^k 是可学习投影矩阵。

(3) 预测头部分 预测头主要由 LayerNorm 层和多层感知器层组成。LayerNorm 层能够对输入数据进行归一化处理,使数据分布更稳定,从而有助于加快模型的收敛速度并提高训练稳定性。多层感知器层则通过其非线性映射能力,对归一化后的特征进行进一步的抽象和变换,生成最终的预测结果。在模型迭代优化过程中,预测头参数会依据特定的更新规则不断进

行调整和优化,该过程对模型性能的提升起到了至关重要的作用。随着迭代更新,预测头能够逐步学习到更准确和鲁棒的特征表示,进而不断提高模型对目标任务的预测能力和泛化性能。迭代优化过程依据以下更新规则调整优化: $x^i = x^{i-1} + w^i t^i$, 其中, x^{i-1} 表示前一阶段预测的关键点位置, w^i 为当前特征块大小, t^i 为位置偏移调整比例。

1.3.4 评估指标 为对本研究所提出医学关键点检测方法进行全面且深入的验证,本团队设计与当前领域内多种前沿技术的对比分析实验,涵盖了医学关键点检测领域中广受关注的 NFDP 以及在面部关键点检测方面表现出色的 SLPT、StarLoss 等方法。这些对比方法在学术界和工业界得到了广泛认可,代表了当前技术发展的较高水平,具有很强的对比价值和参考意义。这一过程旨在系统性地评估并充分证明本研究模型在有效性和优越性方面的优势。选取 2 个量化指标:均方误差(mean squared error, MSE)和在 2、5 和 10 像素误差范围内的成功检测率(successful detection rate, SDR)。MSE 通过计算预测值与真实值之间差异的平方平均值,能够精准地反映模型预测结果与真实值之间的偏离程度,是衡量模型预测准确性的一个重要参考。SDR 为在不同的误差允许范围内,统计模型成功检测到关键点的频率,从而直观地体现模型在不同精度要求下的可靠性和稳定性。具体而言,在 2 像素误差范围内的 SDR 能够有效评估模型在高精度场景下的表现;5 像素误差范围内的 SDR 则更符合一般精度要求的应用场景;10 像素误差范围内的 SDR 则可用于衡量模型在较低精度要求下的鲁棒性和适应性。

MSE 的定义公式为 $MSE(P, P_{gt}) = \frac{1}{N} \sum_{i=1}^N \|p^i - p_{gt}^i\|_2$, 其中 P 和 P_{gt} 分别表示预测和真实的关键点坐标, p^i 和 p_{gt}^i 分别表示 P 和 P_{gt} 中第 i 个地标点的坐标, N 表示关键点的数量。SDR 定义为在给定的误差允许范围(2、5 和 10 像素)内,模型成功检测到关键点的比例。

1.3.5 注射点 3D 位姿计算 为从二维入射点估计 3D 位姿,首先通过相机的深度传感器捕获眼表的深度信息,并投影至三维空间,生成点云数据(x, y, z)。随后,选取入射点附近的局部点云,通过局部点云分析估计注射方向和注射点位置。具体步骤为对局部点云进行 PCA,找到变化最小的方向,该方向通常就是点云平面的法向量方向,即所需的入射方向。因此,首先对局部点云进行中心化处理,即计算每个点与点云均值的差值。假设局部点云中的点集为 $P = \{p_1, p_2, \dots, p_n\}$,

每个点 $p_i = (x_i, y_i, z_i)$, 局部点云的均值 m 即注射点位置, 可以表示为 $m = \frac{1}{n} \sum_{i=1}^n p_i$; 对局部点云去中心化处理后, 即可计算局部点云的协方差矩阵 (Σ), 并对其进行特征值分解, 得到特征值和对应的特征向量。选择最小特征值对应的特征向量 v_{min} , 作为注射方向的估计。该特征向量反映了点云区域内的最小变化方向, 即玻璃体内注射方向。

1.3.6 真实场景验证 在测试集上初步验证模型效果之后, 将模型部署到仿真眼上进行模型效果的测试。与多种前沿技术进行对比分析, 包括 NFDP, 以及在面部关键点检测方面表现出色的 SLPT 和 StarLoss 等方法。采用 SDR 的指标来评价不同方法在仿真眼眼上的注射点识别效果。

1.3.7 实验过程 首先通过 LiDAR Camera L515 相机实时捕获仿真人脸图像, 再输入到不同模型方法中, 从而得到最终的注射点位置, 最后通过 SDR 指标来评判识别效果。

1.3.7.1 RCM 机构设计与运动学求解 获取真实注射点的位置坐标及注射方向后, 将这些坐标信息反馈到机械臂系统进行注射操作。

在眼科手术中, 器械需要通过巩膜小切口进入眼内。RCM 机构设计可使手术器械的运动中心与巩膜刺入点实时重合, 从而避免器械在眼内的运动时对巩膜切口及周边组织造成额外压力和牵扯, 以降低手术风险。

如图 2 所示, 由于连杆是通过 C6 级别高精度机加工得到的, 连杆长度等尺寸指标的误差较小, 几乎无需标定, 如 c_1 和 c_5 ; 而 c_2 、 c_3 和 c_4 可能存在装配误差或连接误差, 需通过进一步测量进行误差分析。其中推杆长度 c_4 的正、逆推算公式为 $c_4^2 = (c_1 \cdot \sin(\theta) - c_3)^2 + (c_2 - c_5 - c_1 \cdot \cos(\theta))^2$, 连杆倾角 θ 的计算公式为 $\theta = \arcsin\left(\frac{A}{\sqrt{B^2 + C^2}}\right) - \arctan\left(\frac{C}{B}\right)$, 其中 $A = c_1^2 + c_3^2 + (c_2 - c_5)^2 - c_4^2$, $B = 2c_1 \times c_3$, $C = 2c_1(c_2 - c_5)$ 。 c_2 、 c_3 和 c_4 是需要进行标定的, 即这些参数被当作变量进行输入, 最终产生一个倾角。这个倾角主要影响 $\Delta\theta_5$, 不会对位置精度产生影响, 但在实验过程中需使用末端的态度信息作为监督信息完成对连杆的高精度标定。

正运动学求解是机械臂运动学分析的关键步骤之一。如图 3 所示, 0 号坐标系为基坐标系, 在整个坐标系中, 红色箭头代表 X 轴的正方向, 蓝色箭头代表 Z 轴的正方向。根据改进型 DH 建模法的定义, 先确定

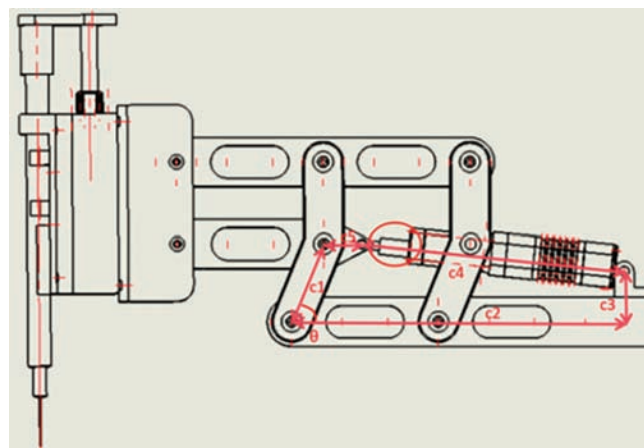


图 2 RCM 结构图 RCM: 远程中心运动

Figure 2 RCM structure diagram RCM: remote center of motion

Z 轴, 再定义 X 轴, X 轴沿过 Z_i 和 Z_{i+1} 轴的公法线方向, Y 轴则遵循右手定则, 构成一个标准的右手直角坐标系。这一过程确保了机械臂各个连杆和关节之间的相对位置和姿态关系能够被准确地描述和计算。

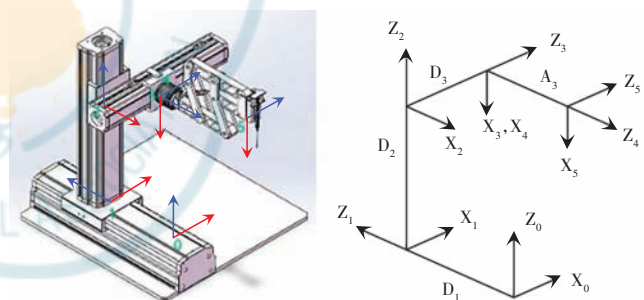


图 3 连杆坐标系

Figure 3 Connecting rod coordinate system

坐标系以及模型参数确定后, 需根据改进型 DH 建模法的定义, 先构建各关节的齐次变换矩阵, 再将这些矩阵依次相乘, 从而得到机械臂末端相对于基坐标系的总变换矩阵, 即末端变化矩阵。这个过程是机械臂正运动学求解的关键环节, 公式为 $T =$

$$\begin{bmatrix} \sin(\theta_4)\sin(\theta_5) & -\cos(\theta_4) & \sin(\theta_4)\cos(\theta_5) & d_3 \\ -\cos(\theta_5) & 0 & \sin(\theta_5) & d_1 - d_4 \\ -\sin(\theta_5)\cos(\theta_4) & -\sin(\theta_4) & -\cos(\theta_4)\cos(\theta_5) & d_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}。$$

逆运动学求解过程主要通过以下公式实现, 这些公式基于末端齐次坐标矩阵的元素来计算各个关节角度以及连杆偏距: 先计算 2 个旋转关节 (即 θ_4 和 θ_5) 的角度, 从末端齐次矩阵中提取元素 a_y , 再通过反正弦函数求得 $\theta_5 = \arcsin(a_y)$, 接着从末端齐次矩阵提取元素 a_x 以及 a_z , 结合 θ_5 的余弦值, 计算

$\theta_4 = \arctan2\left(\frac{a_x}{\cos(\theta_5)}, -\frac{a_z}{\cos(\theta_5)}\right)$, 最后通过从末端齐次

矩阵中提取位置信息来得到前 3 个关节连杆的偏距 (即 d_1 、 d_2 和 d_3)。

1.3.7.2 评价指标 本研究对机械臂的评价指标重点选取绝对定位精度与重复定位精度,分别从准确性与稳定性维度评价机械臂的精准性。

绝对定位精度指机械臂末端执行器按照预设指令到达的实际位置,与该位置在机械臂运动学模型或坐标系中的理论期望位置的偏差程度,它反映了机械臂在整个工作空间内对目标位置的“定位准确性”,是衡量机械臂能否精准匹配医学关键点空间坐标的核心指标。

重复定位精度指机械臂末端执行器在多次重复执行同一指令、前往同一目标位置时,其每次实际到达位置之间的一致性,它反映了机械臂定位的稳定性,与绝对定位精度的准确性形成互补——即使绝对定位精度存在偏差,若重复定位精度高,可通过预先校准偏差值来修正,保证每次操作的一致性。

1.3.7.3 测量方法 (1)重复定位精度测量 ①目标设定:将机械臂末端固定于单轴 100 μm 处目标位置,作为基准。②重复操作:控制机械臂末端离开该目标位置后,重复执行“返回目标位置”的动作,共进行 20 次测量。利用 Microtrak II 激光测距仪精确测定机械臂末端的实际位置。③数据处理:计算每次实际位置与目标位置的差值,得到重复定位误差。(2)绝对定位精度测量 ①目标设定:从 150~850 μm 区间等间距选取 15 个不同的目标位置。②定位操作:控制机械臂末端分别移动至每个预设的绝对目标位置,使用 Microtrak II 激光测距仪精确测定机械臂末端的实际位置。③数据处理:计算每次实际位置与绝对目标位置的差值,得到绝对定位误差。

2 结果

2.1 注射点检测定位精度评价

对比 NFDP、SLPT 和 StarLoss 方法,本研究方法的 MSE 最低,尤其在对瞳孔和虹膜的定位精度方面,表明在处理复杂医学图像时,本研究方法能够更精准地识别并定位关键点。而在严格限定在 2 像素精度误差范围内的评估条件下,本研究方法 SDR 比 SLPT 高出近 10%,展现了其在精细关键点定位方面的突出能力。在 5 像素和 10 像素误差允许范围内,本研究方法 SDR 分别为 72.09% 和 92.68%,均高于其他方法,证明了本研究方法在不同精度要求下的稳定性和可靠性 (图 4,表 1)。

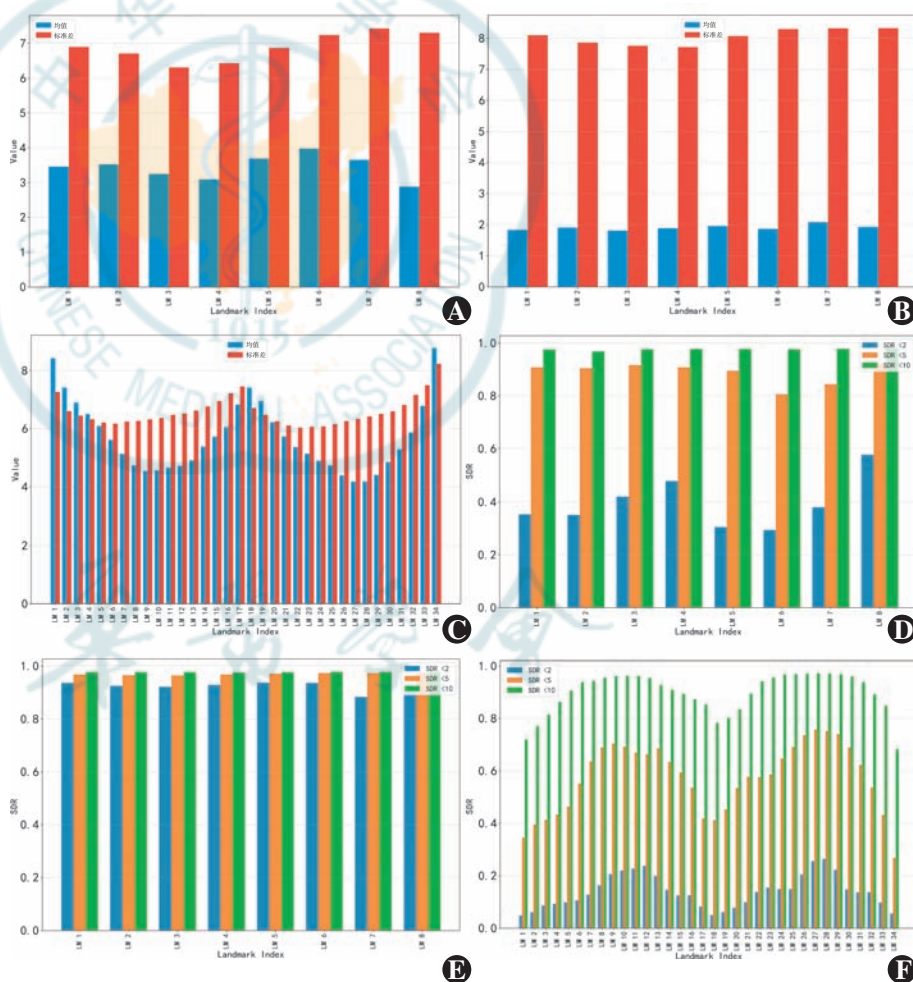


图 4 本研究模型眼部关键点检测指标结果 A:瞳孔检测误差的均值和标准差 B:虹膜检测误差的均值和标准差 C:眼睑检测误差的均值和标准差 D:不同像素误差范围内瞳孔的 SDR E:不同像素误差范围内虹膜的 SDR F:不同像素误差范围内眼睑的 SDR SDR:成功检测率

Figure 4 Results of eye key points detection indicators for our model A: Mean and standard deviation of pupil detection B: Mean and standard deviation of iris detection C: Mean and standard deviation of eyelid detection D: SDR of pupil detection within different error ranges E: SDR of iris detection within different error ranges F: SDR of eyelid detection within different error ranges SDR: success detection rate

表 1 不同方法注射点检测定位精度比较
Table 1 Comparison of injection point detection accuracy among different methods

方法	MSE (SD)				SDR (%)		
	整体	瞳孔	虹膜	眼睑	2 像素	5 像素	10 像素
HRnet ^[26]	4.95 (7.14)	3.98 (7.35)	2.13 (7.11)	5.89 (6.96)	27.82	68.28	92.10
NFDP ^[27]	5.25 (7.64)	4.91 (7.73)	2.29 (7.13)	6.02 (7.56)	26.20	65.78	91.11
Starloss ^[17]	4.85 (7.22)	3.82 (6.25)	1.95 (7.22)	5.77 (6.88)	28.35	67.33	92.15
SLPT ^[15]	4.72 (7.12)	3.43 (6.92)	1.90 (8.05)	5.68 (6.71)	30.60	68.81	92.40
本研究模型	4.25 (6.63)	2.41 (5.53)	1.54 (5.78)	5.32 (6.78)	40.37	72.09	92.68

注: MSE: 均方误差; SD: 标准差; SDR: 成功检测率
Note: MSE: mean squared error; SD: standard deviation; SDR: success detection rate

通过结合不同层次的量化指标以全方位、多角度地评估本研究所提出方法的性能, 确保其在各种实际应用场景中均能发挥出预期效能。通过实验验证, 本研究模型在所有眼部关键点检测的 MSE 为 4.25, 低于 NFDP、Starloss 和 SLPT 方法。不同类型的关键点在定位精度上存在显著差异。其中, 眼睑关键点的定位精度最低, 其次是瞳孔, 而虹膜的定位精度最高(图 5, 表 1)。在瞳孔和虹膜定位方面, MSE 分别达到 2.41 和 1.54, 显示出本模型在精细关键点定位方面的优势。此外, 在 5 像素和 10 像素的误差允许范围内, 本模型 SDR 分别为 72.09% 和 92.68%, 均优于其他方法。这些结果表明, 本研究模型能够有效提升玻璃体内注射点检测的稳定性和精度。

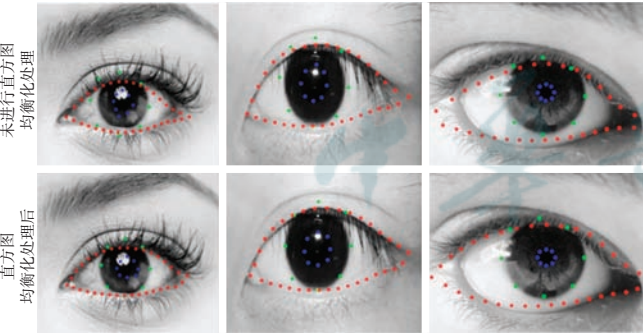


图 5 数据增强效果比较
Figure 5 Comparison of data enhancement effects

表 2 结果显示, 应用随机权重增强在所有关键类别中的 MSE 低于固定值权重增强和无增强, 且固定权重增强的 MSE 低于无增强。

2.2 真实场景下验证结果

从验证结果上可以看到本研究方法在不同像素误差允许范围下, 均优于其他方法, 特别是在 10 像素误差允许范围内, 实现 91.74% 的成功检测率(图 6, 表 3)。

表 2 数据增强替换权重的消融研究结果
Table 2 the ablation research results of data augmentation replacement weights

方法	MSE (SD)			
	整体	瞳孔	虹膜	眼睑
无增强	4.86 (7.73)	3.77 (8.51)	1.94 (8.74)	5.81 (7.05)
固定权重增强	4.61 (7.02)	3.12 (6.29)	1.78 (7.24)	5.56 (6.92)
随机权重增强	4.25 (6.63)	2.41 (5.53)	1.54 (5.78)	5.32 (6.78)

注: MSE: 均方误差; SD: 标准差
Note: MSE: mean squared error; SD: standard deviation

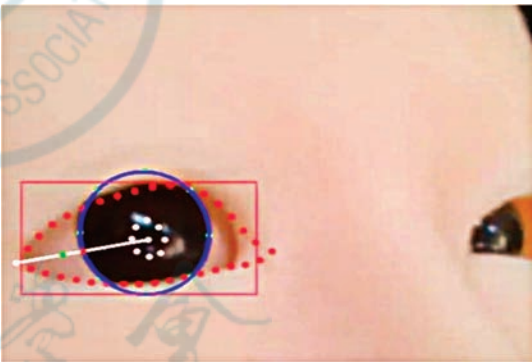


图 6 仿真入仰卧位注射点识别效果 红色方框表示初步确定眼睛位置, 红色点表示眼睑关键点, 蓝色圈表示虹膜关键点, 白色点表示瞳孔关键点, 白色线表示眼睑角点与拟合瞳孔中心点的连接线, 绿色点表示注射点
Figure 6 Recognition of injection points in simulated humans The red box indicated the preliminarily determined position of the eyes, the red dots indicated the key points of the eyelids, the blue circle indicated the key points of the iris, the white dots indicated the key points of the pupils, the white lines indicated the connecting lines between the eyelid corner points and the fitted pupil center points, and the green dots indicated the injection points

2.3 机械臂末端定位精度评价

在机械臂末端的重复定位误差始终稳定地控制在 $\pm 5 \mu\text{m}$ 之内, 表明机械臂末端展现出了优异的重复定位精度(图 7、8)。绝对定位误差控制在 $\pm 5 \mu\text{m}$ 之内, 表明机械臂具备稳定的绝对定位能力(图 9、10)。

表 3 不同方法在各像素误差允许范围内仿真眼上的注射点检测精度比较

Table 3 Comparison of injection point detection accuracy of different methods within different error ranges on the simulated human eye

方法	SDR (%)		
	2 像素误差	5 像素误差	10 像素误差
HRnet ^[26]	22.67	66.10	89.33
NFDP ^[27]	23.50	60.52	88.25
Starloss ^[17]	26.82	63.47	89.56
SLPT ^[15]	28.70	66.21	90.32
本研究方法	39.89	70.89	91.74

注:SDR:成功检测率

Note:SDR:succes detection rate

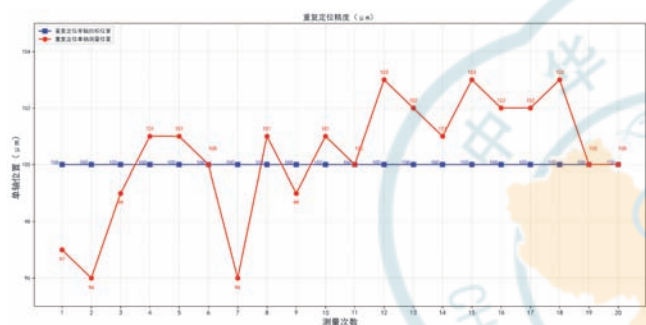


图 7 重复定位测试结果 测量次数为在同一目标位置下进行测量

Figure 7 Repeated positioning test results The measurements were measured at the same target position



图 8 重复定位误差 测量次数为在同一目标位置下进行测量

Figure 8 Repeated positioning error The measurements were measured at the same target position

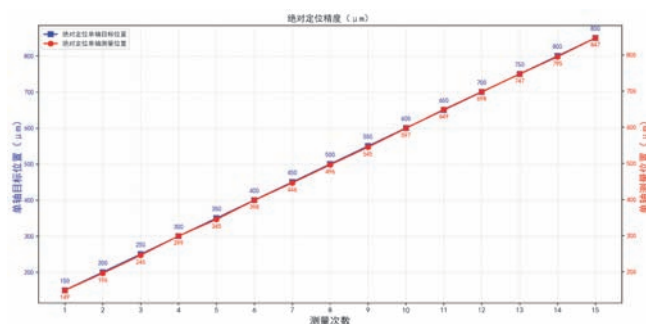


图 9 绝对定位测试结果 测量次数为在不同目标位置下进行测量

Figure 9 Absolute positioning test results The measurements were measured at different target positions

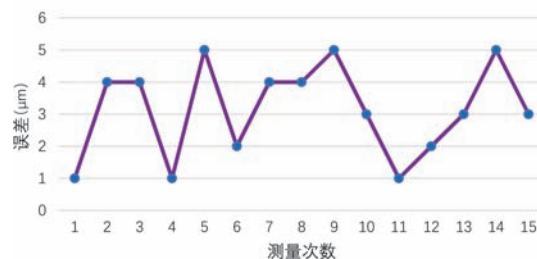


图 10 绝对定位误差 测量次数为在不同目标位置下进行测量

Figure 10 Absolute positioning error The measurements were measured at different target positions

3 讨论

本研究开发了一种基于 AI 引导的玻璃体内注射机器人系统,旨在通过 3D 视觉信息实现眼部入射点的自动检测和空间定位,并通过 RCM 机械臂结构和精确的运动学求解技术完成高精度的注射操作。系统利用 3D 相机获取眼表图像,并基于 PatchCrop-Transformer 模型分析注射点位置,结合 PCA 分析局部点云数据以确定注射点和注射方向。这种多步骤的方法不仅提高了注射点检测的准确性,还为机械臂提供了精确的导航支持。RCM 机械臂结构的设计确保了操作的稳定性和安全性,通过正、逆运动学求解得到关节运动路径可以用于机械臂执行具体的手术。实验结果显示,本研究设计的机械臂的重复定位误差和绝对定位误差均在 $\pm 5 \mu\text{m}$ 范围内,表明系统在定位精度和操作稳定性方面表现出色。

此外,本研究在数据处理方面也进行了创新。欧美人与中国人的瞳孔和虹膜存在显著色彩差异,传统基于欧美人群的虹膜和瞳孔定位技术在中国人眼球上的检测效果表现不佳,面对不同个体之间虹膜和瞳孔颜色的多样性情况,存在着一定的局限性。本研究利用直方图匹配技术对图像的灰度级分布进行调整,使得虹膜和瞳孔的颜色特征更接近中国人群眼部图像的目标分布。这不仅有效模拟了不同个体间的颜色变化,还显著提高了模型对中国人群眼部图像的适应性、在处理低对比度图像时的鲁棒性,从而在根本上提升了定位精度。

本研究结果显示,眼睑作为眼部的一个复杂结构,个体差异非常明显,不同的人可能具有双眼皮或单眼皮的形态差异,内眦的突出程度也不尽相同,此外,眼睑的遮挡情况也因人而异。这些复杂的因素相互叠加,使得眼睑关键点在手动标注时存在较大的模糊性。标注人员难以在这些复杂情况下精准地确定眼睑关键点的位置,导致标注错误的概率显著增加。而标注数据的准确性对于模型训练至关重要,由于眼睑关键点

标注错误较多,最终使得模型在定位眼睑关键点时的精度相对较低。

与眼睑关键点形成鲜明对比的是,虹膜关键点的定位精度最高,可能主要得益于虹膜与巩膜之间存在清晰且十分突出的颜色边界。这一显著的特征为模型提供了明确的定位依据,使得模型能够相对轻松地实现对虹膜关键点的高精度定位。模型可以很容易地捕捉到这一边界特征,并以此为基础准确地确定虹膜关键点的位置。

然而,瞳孔关键点的定位却并非易事,主要表现在不同人种眼球的情况模拟上,具体体现为瞳孔与虹膜的颜色对比度存在多样性。当对比度较低时,瞳孔与虹膜之间的颜色边界变得模糊不清,这给模型的定位带来了巨大的挑战。模型在识别瞳孔关键点时,往往难以准确判断其边界,从而导致定位精度受到一定影响。

但是,本研究团队通过不同数据增强操作,能够生成更多的训练样本。这些样本涵盖了更加广泛的虹膜和瞳孔颜色变化情况,使得模型在训练过程中能够学习到更为丰富的特征。当模型接触到大量多样化的样本后,模型在实际应用中面对各种复杂情况时,能够展现出更强的泛化能力和适应性。

本研究模型引入 3D 视觉信息,实现了对注射点的精准定位。与传统 2D 图像分析方法相比,3D 视觉信息能够提供更丰富的空间数据,帮助系统更准确地确定注射点和入射方向。该技术的引入,不仅提高了注射操作的精度,更为未来复杂眼科手术的自动化提供了新思路。相比之下,国外的 Preceyes 机器人系统虽然在视网膜下药物注射中表现出色,但主要采用主从式操作模式,仍需要术者长时间的远程操控。本系统实现了从图像获取到注射操作的全自动化,能够有效减少人为因素的干扰。

在仿真人眼实验中,虽然训练数据集图像为坐位拍摄,由于该数据集像素精度较高,在卧位的仿真人模型实验中,本研究方法的注射点检测成功率仍然优于其他方法。

在机械臂定位精度方面,误差始终稳定保持在 $\pm 5 \mu\text{m}$ 范围之内,充分反映了机械臂在重复定位过程中的高稳定性和精准性,还表明其内部的运动控制算法以及机械传动结构均具备高度可靠性。这种稳定性对于机械臂在实际应用中的可靠性至关重要,特别是在需要高精度和重复性操作的医疗领域。

本研究存在一些局限性。首先,本研究使用的 TEyeD 数据集主要涵盖欧洲个体的眼部图像,虽然通

过直方图匹配和加权融合策略提升了模型对国人眼部图像的适应性,但仍需在更广泛的人种和复杂环境下进一步验证;未来的研究可扩展测试数据集,涵盖更多人种和不同光照条件下的眼部图像,以提高模型的泛化能力。另外,考虑到安全性和伦理问题,本方案缺少真实场景下的测试,仅确定了一整套自动化注射机器人的控制流程;未来可进一步优化系统鲁棒性和安全性后进行临床测试,验证其真实临床效果。

综上所述,本研究通过引入 3D 视觉信息和 RCM 机械臂结构,显著提升了玻璃体内注射的精度和稳定性,为眼科手术机器人技术的发展提供了新的思路和方法。

利益冲突 所有作者均声明不存在利益冲突

作者贡献声明 陈璟文:酝酿与实验设计、研究实施、数据采集、数据分析与解释、论文撰写;庞义杰:研究实施、数据采集、数据分析与解释;唐晓颖:对文章的知识性内容进行批评性审阅、指导研究;袁进:对文章中眼科专业领域的知识性内容进行批评性审阅

参考文献

- [1] Starr MR, Xu D, Boucher N, et al. Characterizing progression to neovascular AMD in fellow eyes of patients treated with intravitreal anti-VEGF injections [J]. *Ophthalmic Surg Lasers Imaging Retina*, 2021, 52(3): 123-128. DOI: 10.3928/23258160-20210302-02.
- [2] Joao J, Arthur AN, Anna PAN, et al. Efficacy of ranibizumab plus aflibercept therapy with dexamethasone intravitreal injection in patients with DME [J]. *Invest Ophthalmol Vis Sci*, 2019, 60(9): 5400.
- [3] Agarwal A, Nagpal M. Intravitreal moxifloxacin injections in acute post-cataract surgery endophthalmitis: Efficacy and safety [J]. *Indian J Ophthalmol*, 2021, 69(2): 326-330. DOI: 10.4103/ijo. IJO_587_20.
- [4] Lai TY, Liu S, Das S, et al. Intravitreal injection—technique and safety [J]. *Asia Pac J Ophthalmol (Phila)*, 2015, 4(6): 321-328. DOI: 10.1097/APO. 0000000000000146.
- [5] Gheorghe Gruionu L, Constantinescu C, Iacob A, et al. Robotic system for catheter navigation during medical procedures [J]. *Appl Mech Mater*, 2020, 896: 211-217.
- [6] Elsayed AS, Aldhaam NA, Nitsche L, et al. Robot-assisted radical cystectomy: review of surgical technique, and perioperative, oncological and functional outcomes [J]. *Int J Urol*, 2020, 27(3): 194-205. DOI: 10.1111/iju. 14178.
- [7] Liang J, Wu J, Huang H, et al. Soft sensitive skin for safety control of a nursing robot using proximity and tactile sensors [J]. *IEEE Sens J*, 2020, 20(7): 3822-3830. DOI: 10.1109/JSEN. 2019. 2959311.
- [8] Guo Y, Yang Y, Feng M, et al. Review on development status and key technologies of surgical robots [C]. 2020 IEEE International Conference on Mechatronics and Automation (ICMA), 2020: 105-110. DOI: 10.1109/ICMA49215. 2020. 9233776.
- [9] Hendrix R. Robotically assisted eye surgery: a haptic master console [D]. Eindhoven University of Technology, S-Hertogenbosch, 2011.
- [10] Bai M, Zhang M, Zhang H, et al. An Error Compensation Method for Surgical Robot Based on RCM Mechanism [J]. *IEEE Access*, 2021, 9: 140747-140758. DOI: 10.1109/ACCESS. 2021. 3117350.
- [11] MacLachlan RA, Becker BC, Tabarés JC, et al. Micron: an actively stabilized handheld tool for microsurgery [J]. *IEEE Trans Robot*, 2012, 28(1): 195-212. DOI: 10.1109/TRO. 2011. 2169634.
- [12] Becker BC, MacLachlan RA, Lobes LA, et al. Vision-based retinal membrane peeling with a handheld robot [C]. 2012 IEEE International Conference on Robotics and Automation, 2012: 1075-1080. DOI: 10.

- 1109/ICRA. 2012. 6224844.
- [13] Becker BC, Yang S, Maclachlan RA, et al. Towards vision-based control of a handheld micromanipulator for retinal cannulation in an eyeball phantom [J]. Proc IEEE RAS EMBS Int Conf Biomed Robot Biomechatron, 2012, 2012 : 44 - 49. DOI: 10. 1109/BioRob. 2012. 6290813.
- [14] Zhu M, Shi D, Zheng M, et al. Robust facial landmark detection via occlusion-adaptive deep networks [C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019 : 3481-3491. DOI:10. 1109/CVPR. 2019. 00360.
- [15] Xia J, Qu W, Huang W et al. Sparse local patch transformer for robust face alignment and landmarks inherent relation learning [J]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022 : 4042 - 4051. DOI: 10. 1109/CVPR52688. 2022. 00402.
- [16] Wu W, Qian C, Yang S, et al. Look at Boundary: a boundary-aware face alignment algorithm [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018 : 2129-2138. DOI:10. 1109/CVPR. 2018. 00227.
- [17] Zhou Z, Li H, Liu N, Wang G, Yu R, Ji S. STARloss: reducing semantic ambiguity in facial landmark detection [C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023 : 15475-15484. DOI:10. 1109/CVPR52729. 2023. 01485.
- [18] Tuan H N A, Hai N D X, Thinh N T. The improved faster r-cnn for detecting small facial landmarks on vietnamese human face based on clinical diagnosis [J]. J Imaging Graph, 2022, 10 (2) : 76-81.
- [19] Sun Z, Feng C, Patras I, Tzimiropoulos G. LAFS: landmark-based facial self-supervised learning for face recognition [C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024 : 1639-1649. DOI:10. 1109/CVPR52733. 2024. 00162.
- [20] Yang J, Liu Q, Zhang K. Stacked hourglass network for robust facial landmark localisation [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 : 2025-2033. DOI:10. 1109/CVPRW. 2017. 253.
- [21] Dong X, Yan Y, Ouyang W, Yang Y. Style aggregated network for facial landmark detection [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018 : 379-388. DOI:10. 1109/CVPR. 2018. 00047.
- [22] Liu Z, Zhu X, Hu G, et al. Semantic alignment: finding semantically consistent ground-truth for facial landmark detection [C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019 : 3462-3471. DOI:10. 1109/CVPR. 2019. 00358.
- [23] Bulat A, Tzimiropoulos G. Binarized convolutional landmark localizers for human pose estimation and face alignment with limited resources [C]. 2017 IEEE International Conference on Computer Vision (ICCV), 2017 : 3726-3734. DOI:10. 1109/ICCV. 2017. 400.
- [24] Stern A, Sharan L, Romano G, et al. Heatmap-based 2d landmark detection with a varying number of landmarks [J/OL]. arXiv ePrints, 2021 : 2101. 02737 [2025 - 05 - 18]. <https://doi.org/10.48550/arXiv.2101.02737>.
- [25] Ma L, Xiao D, Kim D, et al. Simulation of postoperative facial appearances via geometric deep learning for efficient orthognathic surgical planning [J]. IEEE Trans Med Imaging, 2023, 42 (2) : 336-345. DOI:10. 1109/TMI. 2022. 3180078.
- [26] Wang J, Sun K, Cheng T, et al. Deep high-resolution representation learning for visual recognition [J]. IEEE Trans Pattern Anal Mach Intell, 2021, 43 (10) : 3349 - 3364. DOI: 10. 1109/TPAMI. 2020. 2983686.
- [27] Huang Z, Zhao R, Leung F, et al. Landmark localization from medical images with generative distribution prior [J]. IEEE Trans Med Imaging, 2024, 43 (7) : 2679-2692. DOI:10. 1109/TMI. 2024. 3371948.

(收稿日期:2025-09-10 修回日期:2025-10-26)

(本文编辑:张宇 骆世平)

消 息

经典译著《眼科学(第6版)》出版发行

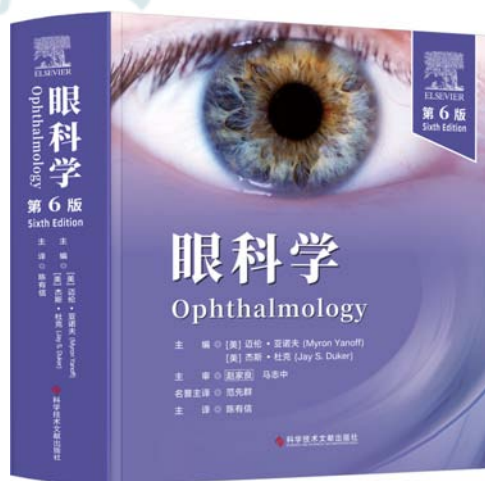
由美国迈伦·亚诺夫(Myron Yanoff)、杰斯·杜克(Jay S. Duker)博士主编,赵家良教授、马志中教授主审,范先群院士名誉主译,陈有信教授主译的《眼科学(第6版)》,2025年9月由科学技术文献出版社重磅出版发行!

精准弥合知识断层 全景覆盖眼科学领域:本书系统全面,开篇阐释眼的解剖、生理、遗传基础,筑牢学习根基;继而深入光学与屈光手术,详解屈光不正及矫正方案;疾病部分覆盖眼球结构、眼附属器、视觉通路病症,明晰病因、症状、诊断与治疗;专设眼部药物、激光及冷冻疗法章节,兼顾全身疾病眼科关联与社区眼科学内容;临床实操层面,规范眼科检查、暗室操作及仪器使用要点等。

创新视角赋能临床 前沿聚焦单卷精要:本书为兼具全新视角、全面内容与临床关联的单卷本眼科教科书,紧跟前沿新增光治疗性角膜切除术等章节,布局精简直观,配超2 000幅高质量插图及70余个手术视频,助力同步掌握基础理论与前沿进展,精准填补多卷本与专科书籍空白。

创新编码高效检索 案头必备赋能医研:本书每章采用颜色编码体系,便于跨章查阅与全书浏览,编排呈现突破传统,助力快速抓取关键信息,适配眼科医教研人群,可作为临床案头常备参考书。

本书开本为889×1 194 mm、1/12开,共160印张、1 920页,精装彩印,定价980.00元,可通过“中国医学临床百家”微店购买。



(陈有信)



中华医学杂志社
Chinese Medical Association Publishing House

版权所有 违者必究