

附件 1:

南方科技大学

大学生创新创业训练计划项目 申请书

项目名称: 基于深度学习的眼科手术图像深度估计算法的开发

项目负责人: 周心怡

所在院系: 计算机科学与工程系

专业班级: 计算机科学与技术专业 18 级

联系电话: [REDACTED]

E - mail: [REDACTED]

指导教师: 刘江, 胡衍 职称 教授, 副研究员

申请日期: 2021 年 4 月 23 日

南方科技大学教学工作部 制

填 表 说 明

一、请严格按照表中要求填写各项。要求实事求是，表达明确、严谨。

二、项目只能由全日制本科生提出申请，原则上以二、三年级学生为主。申请者要品学兼优、学有余力，有较强的独立思考能力和创新意识，对科学研究、科技活动或社会实践有浓厚的兴趣。

三、“项目类别”指 A--创新训练项目；B--创业训练项目；C--创业实践项目。

四、申请书中第一次出现外文名词时，要写清全称和缩写，再出现同一词时可以使用缩写。

五、申请书用 A4 纸双面打印，于左侧装订成册。由指导教师和所在院系审查并签署意见后报送教学工作部，同时提交电子文档。

六、如表格不够，可以加附页。

项目 简 况	项目名称	基于深度学习的眼科手术图像深度估计算法的开发						
	项目类别	(A) A--创新训练项目; B--创业训练项目; C--创业实践项目						
	申请资助 经费	16000 元			项目起止时间	2021 年 5 月至 2022 年 5 月		
项目 负 责 人	姓 名	周心怡		性别	女		出生年月	2000.02
	专业年级	2018 级			学院 (系、部)	计算机科学与工程系		
	学分绩点 及专业排 名	[REDACTED]			电 话	[REDACTED]		
项目 组 主 要 成 员	姓 名	性 别	出 生 年 月	专业年级	所在学院 (系、部)	项目分工		签 字
	聂秋实	男	2000.11	2018 级	计算机科学与工程系	相机标定		
	周颖泉	女	2000.04	2018 级	计算机科学与工程系	显微镜影像深度估计		
	曾鸣	女	2000.02	2018 级	计算机科学与工程系	查找论文等资料		
指 导 教 师	姓 名	性 别		出生年月	职 称	最高学历	最后学位	研究方向
	刘江	男		1968.03	教授	博士	博士	眼科人工智能
	胡衍	女		1986.10	副研究员	博士	博士	手术辅助, 术中导航
	电 话	18758319186			E-mail	liuj@sustech.edu.cn		

一、立项依据（包括项目的意义，国内外研究现状与存在的问题，自身具备的知识条件,自己的特长、兴趣，相关经历，开展研究的前期准备工作等）

1.1 项目简介

通过图像获取深度信息是计算机视觉领域中的一个经典问题，简称深度估计。在自然场景中，深度估计可以被应用于 3D 建模，场景重建和机器人视觉计算等领域。在眼科手术场景中，尤其是眼底手术中，医生需要通过显微镜观察狭小漆黑的眼底，操控手术器械对眼底视网膜，比如黄斑病变区，实施手术。但显微镜所能看到的视野范围有限，加上眼睛的球状结构导致视野范围内存在焦面不均衡的现象，使得医生很难判断手术器械的操作深度，容易造成因器械触碰视网膜而引起的损伤甚至失明等后果，所以眼底视网膜手术对医生的手术经验等有极高的要求。

因此，我们拟基于深度学习，提出一套显微镜视野下的手术器械深度估计算法。首先通过传统的相机标定方法确定显微镜的内参，计算出目标器械的景深作为金标准，从而建立显微镜眼科手术场景数据库。为了提高深度估计的准确性，拟在基于建立的数据库获得全聚焦图像后，改进深度估计模型，估计显微镜图像的深度，用于辅助医生手术中进行器械深度的判断。最后拟基于离体猪眼及临床视频进行方法验证。

1.2 项目意义

1.2.1 社会意义

常见的眼科疾病包括青光眼、老年黄斑变性、糖尿病视网膜病变和高度近视导致的一系列眼底病变。其中老年黄斑变性是导致老年人失明的重要原因。同时，近年来电子屏幕的流行，导致近视人群的数量不断上涨，近视的发病呈现年龄小、进展快的特点，高度近视的人群也在不断扩大。高度近视的直接原因是眼轴的过度伸长，而眼轴过度伸长极易导致视网膜裂孔及脱裂、黄斑病变等眼底并发症。

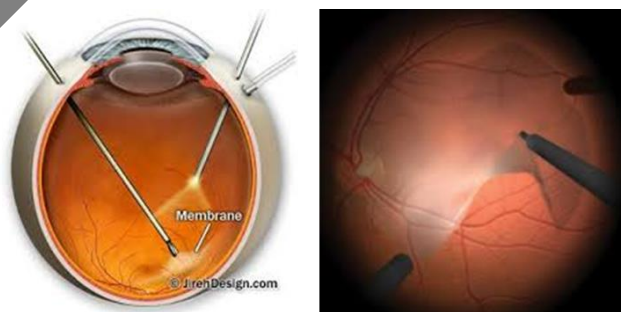


图 1：黄斑前膜剥离手术示意图

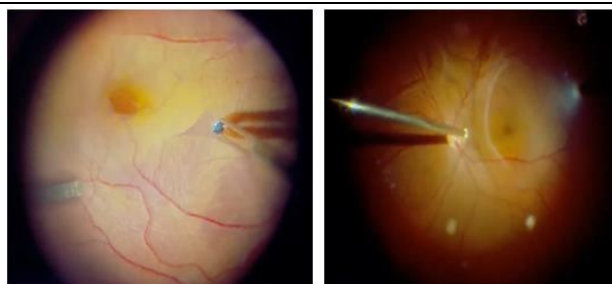


图 2：黄斑裂孔区内界膜剥离术

基于显微镜的眼科手术是治疗眼底视网膜疾病常用手段，如黄斑病变、黄斑前膜、黄斑裂孔等。黄斑裂孔是一类黄斑病变，常常由外伤、高度近视、囊性黄斑水肿导致，需要进行及时的手术治疗，否则会导致失明。黄斑裂孔的手术治疗方法包括采用手工剥除法剥除黄斑裂孔周围的视网膜内界膜。然而在进行剥离手术时，如果医生对手术刀深度的判断不准确，产生了范围过大的抖动，手术刀就容易误触黄斑裂孔周围视网膜的结构，造成视网膜的损伤[2]。这种问题不仅存在于黄斑区内界膜的剥离手术中，也存在于诸如白内障手术等其他眼科手术中。

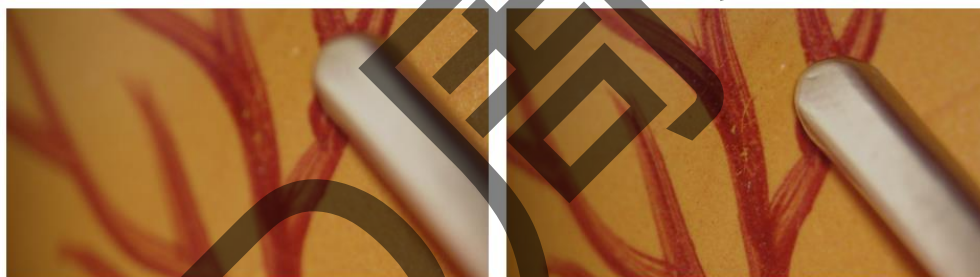


图 3：手术器械失焦现象

而造成医生难以判断深度的原因，除了视野范围小，手术空间拥挤之外，还有手术视野中存在的失焦现象。失焦现象让本身就难以辨明深度的眼科手术更加困难。因此，我们想设计开发一种算法，通过显微镜拍摄的图片得到手术场景的全聚焦图像，再用全聚焦图像提取手术场景的深度信息，让医生能够借助深度信息更好地判断手术器械移动的方向和幅度，降低眼科手术中因为难以估计器械深度而造成病人损伤的风险。

1.2.2 科学意义

现有的图像深度估计工作大部分是基于自然场景，比如建筑物，自然风景的大尺度场景深度估计，应用于自动驾驶等方面。而对于医学领域的应用场景，其应用场景尺度更小，需要的精度会更高，现有的方法不能被直接迁移使用。同时对于眼科手术，深度判断不清容易导致手术过程中器械碰伤眼底，造成医疗事故甚至失明，因此亟需

一种适用于此场景的深度估计算法。本项目拟用深度学习算法代替传统方法计算手术器械的深度，用于辅助医生手术。该项目不仅在小场景下的深度估计领域具有方法上的创新意义，同时还具有很大的医学意义。

1.3 国内外研究现状与存在的问题

1.3.1 全聚焦图像的获取

景深是指相机在焦点前后相对清晰的成像范围。在这个景深距离范围内，物体的像模糊的程度是肉眼无法察觉的，因此会被人眼当成边界清晰的图像。决定景深的因素有三个，分别是物距（镜头的对焦距离）、镜头焦距和镜头的光圈值。浅景深是导致图片中出现失焦、模糊区域的最直接的原因。在微距摄影和显微镜成像的场景中，由于物体离镜头很近，景深较小。因此显微镜使用场景中浅景深导致的图像区域性模糊问题一直存在。

在复原全聚焦图像的方法中，最常见的方法有两种：利用一系列聚焦在不同区域的图像组成全聚焦图像[3][4]，利用点扩散函数（Point Spread Function, PSF）的性质恢复失焦图像所对应的聚焦图像[5]。

在利用多幅图像组成全聚焦图像的方法中，Wang 等人首先采用将不同厚度的平板透镜安装在镜头前，利用折射的原理改变图像焦距，改变景深的方法获取同一场景，但是聚焦区域不同的图片。然后利用拉普拉斯边缘检测算子将边缘清晰的像素点，即在焦平面上，聚焦的像素点提取出来。最后把所有照片的像素点组合起来，得到一张原场景的全聚焦图像[3]。这种方法所采用的拉普拉斯算子属于传统的边缘检测算子，而如果场景没有清晰的边缘信息，其所得到的全聚焦图像结果并不理想。同时，这种方法并没有指出为了保证全聚焦图像的成功合成需要拍摄多少张照片。因此，我们需要解决这个问题。

同样是通过组合的方法得到全聚焦图像，Maximov 等人在利用聚焦像素点的同时，还利用了卷积神经网络模型输出的失焦程度图像，即图像上每个像素点为模型对原图对应像素点的弥散圆直径的估计值[4]。弥散圆，即当点光源在成像时，由于距离等因素无法汇聚成一个点，而形成的一个扩散的圆形投影。这种传统理论与先进深度学习技术相结合的全聚焦图合成方向刚刚起步，十分值得探索。

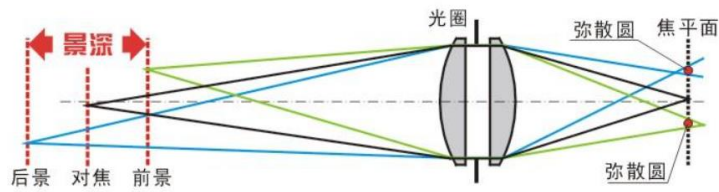


图 4：弥散圆和景深

点扩散函数 (PSF)，是用于描述成像系统对点光源的响应，即点光源输出像的光场分布。在数学上点光源可以用脉冲函数代表，输出的对光源的响应可以用脉冲响应函数代表。一张模糊的照片可以认为是由一张清晰的图像和 PSF 做卷积操作而得到的。Zhou 等人[5]尝试利用额外的卷积操作和比值法计算出造成原图模糊的 PSF，再将 PSF 转化为一张特殊的图像——图像的每个像素表示图像的模糊程度，最后利用图像处理方法，直接提取原图中的聚焦像素，合成全聚焦图像。但是这种方法在遮挡造成的模糊和失焦造成的模糊的区分上表现不佳。在眼科手术场景中，手术器械对眼睛的遮挡无法避免，这也说明我们在开发自己模型的时候也需要改进这方面的性能。

1.3.2 深度图像的获取

通过图像获取图像深度信息的方法大致可以分为三大类[6]。第一大类是多视觉几何方法，即通过多幅图像之间的特征点匹配，重建场景的三维结构来得到深度信息，代表方法为“运动恢复结构”(Structure from Motion, SfM) [7]。但是这一类方法所借助的特征点匹配算法需要较高质量的图片，否则无法准确地进行特征点匹配，从而导致在处理高倍显微镜图像时会有较大的误差。第二大类是基于深度传感器获取深度信息的方法，即通过深度相机 (RGB-D 相机) 或者激光雷达直接得到场景中的深度信息。深度相机可以借助将具有一定结构的光线投射到被拍摄物体上来采集深度信息，也可以通过安装多个镜头，得到不止一个视角的图像，结合双目立体视觉的知识计算出场景中物体的深度。但是由于对光线变化的高依赖性，这类方法对环境光线的变化特别敏感，对检测的距离和被检测场景的纹理情况也有较高的要求。激光雷达常用于无人驾驶领域的测距任务，但是激光雷达只能产生十分稀疏的深度图。因此，这两种方法也不适用于光线条件复杂，场景纹理多变，视野范围狭小的显微镜场景。第三大类是深度学习方法。比如利用物体六自由度位姿计算深度信息[8]，或是结合物体探测预测深度[9]，以及基于语义体素模型进行三维重构从而得到深度信息[10]。但

绝大多数方法都适用于如自动驾驶，机器人控制等大尺度场景下，与显微镜场景下物体以及相机允许的运动范围有很大不同。同时，这些模型大多数需要巨大的训练集以达到较高的精度，这使得数据采集相对困难的医学图像任务难以直接迁移现有的方法以达到较高的精度。因此这些现有的深度学习方法在医学图像的深度估计任务上，具有很大的局限性。

用于获取显微镜图像深度信息的传统方法，如使用极面图像[11]，焦堆栈[12]，角度或空间信息的方法[13]，大多鲁棒性不够强，并且各自特点比较明显，独立性较强。

然而，在深度估计方法领域，由 Griffin 等人提出了一种结合图像分割和相机运动的新型算法[14]。在进行深度预测过程中，根据几何学的原理，当相机与物体都位于同一坐标轴时（这里假设为 z 轴），物体运动时在相机成像平面上留下的投影面积与物体距相机的距离成比例地变化。因此，我们可以根据物体的移动范围与物体的成像面积得到物体深度[14]。

$$d_{object} = \frac{z_i \sqrt{a_i} - z_j \sqrt{a_j}}{\sqrt{a_i} - \sqrt{a_j}}$$

其中， d_{object} 表示物体的预测深度， z_i 与 z_j 表示物体在 i 时刻与 j 时刻的 z 轴坐标， a_i 与 a_j 表示物体在 i 时刻与 j 时刻的成像面积，由图像分割掩码计算得到。

在医学图像分析领域，图像分割是一种用于识别，检测的常用技术。图像分割就是指把图像分成各具特性区域并提取出人们感兴趣的目标（前景）的技术过程。现阶段图像分割的方法有很多，比如基于神经网络的分割方法的基本思想是通过训练多层感知机来得到线性决策函数，然后用决策函数对像素进行分类来达到分割目的的方法。这种方法需要大量的训练数据。神经网络存在巨量的连接，容易引入空间信息，能较好地解决图像中的噪声和不均匀问题。Ronneberger 等人针对医学图像数据难以获得、训练数据少的特点，设计了一个用于图像分割的 U-net 卷积神经网络[16]。大致结构为：

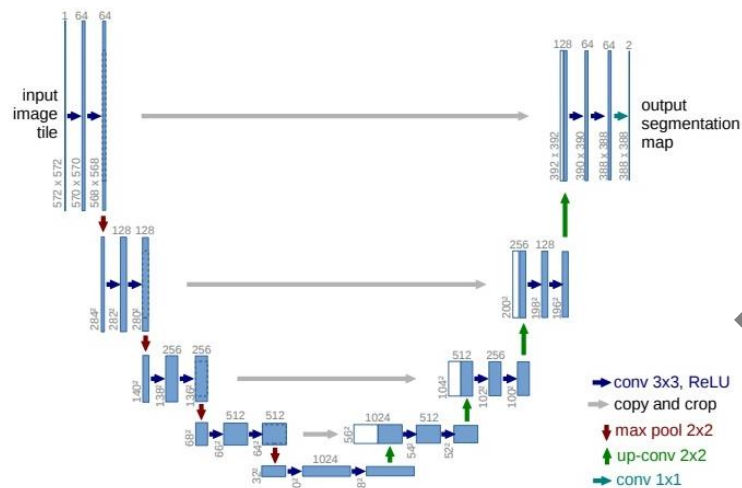


图 5：U-net 卷积神经网络示意图

这种网络采用 U 型对称结构，分为左侧：编码层，和右侧：解码层。编码层通过对输入图像做一系列的卷积池化操作，达到对有效区域的特征提取效果。对于解码层，每层解码层会将反卷积恢复的特征，与同层编码层提取的特征连接，再进行卷积，实现了编码层和解码层的特征融合，有助于特征恢复过程中保留图像的更多细节，使得到的分割图像轮廓更清晰。

由上述分析可知，基于图像分割和相机运动的深度估计算法在医学图像领域有很大的应用前景和发展潜力。我们希望通过进一步地研究，将医学图像领域的图像分割算法和基于图像分割的深度估计算法结合起来，设计出精度较高的眼科手术图像深度估计算法。

1.4 前期准备工作

已经完成相机标定版的采购和相机标定算法的设计，并且对多个放大倍数的实验场景进行了标定算法的实验，得到了可靠的结果。

已经完成相机景深测量算法的设计，用爱特蒙特光学的 5-15 景深测试板多次测量了各个放大倍数下的景深，并利用这个参数确定了每个场景下合成全聚焦图像需要拍摄的图片的数量。

已经完成手术器械图像分割、标定的准备工作，用于下阶段的深度预测工作。

1.4.1 相机标定

相机标定是计算机视觉领域研究的基础。在我们的眼科手术场景中，也需要相机的内外参数提供相机的信息，作为后续深度真实值计算，深度估计神经网络搭建的基础。

相机标定旨在计算相机的内参与外参。相机的内参有焦距、偏斜角、畸变系数、以及光学中心位置；相机的外参代表着它的位置以及朝向。许多双目以及单目深度预测研究都需要相机的内参与外参。普通相机的标定已经有大量文献及工具可以做到比较好的效果，而我们的研究是在显微镜下的，我们需要找出适用于的显微镜标定的方法。

目前我们已经购得了适合显微镜各倍数下的高精度标定板，并且尝试发现使用 MATLAB 的计算机视觉工具包中的相机标定功能可以获得比较好的标定结果。

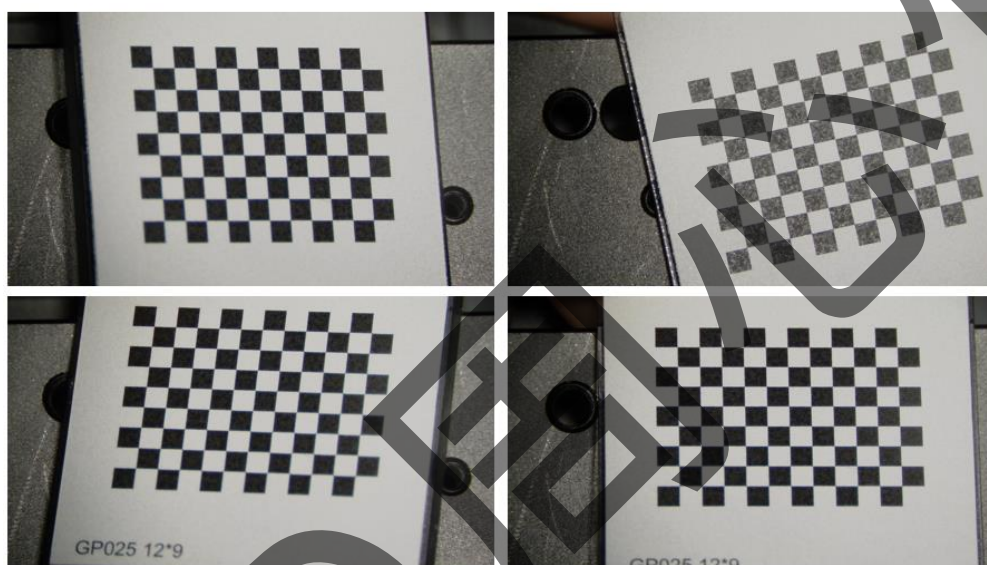


图 6: 手术显微镜 4 倍放大倍数下，标定板不同位姿的图像

MATLAB 的计算机视觉工具包（Open Camera Calibrator App）使用多张同一场景下，不同角度、位姿拍摄的棋盘格标定板图像，通过标定棋盘格角点来测算相机内外参数。同时，该工具包还可以将测算的角点结果和内外参输出，将角点标定结果重投影到棋盘格上（如下图所示），并计算重投影误差。

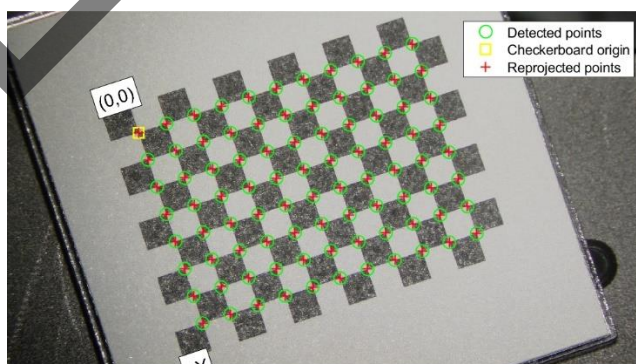


图 7: 棋盘格标定结果在原图上的重投影效果

MATLAB 工具包所提供的误差参数，说明该方法能够较好地通过标定测算显微镜相机的内外参，可以用于我们的显微镜眼科手术场景。

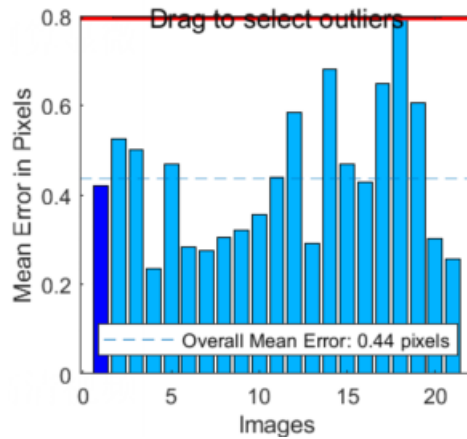


图 8: MATLAB 算法结果的重投影误差

1.4.2 景深测量

显微镜相机的型号: Canon 3CMOS camera 1080P, 纽康曼 NCVideoHD 高清视频系统。

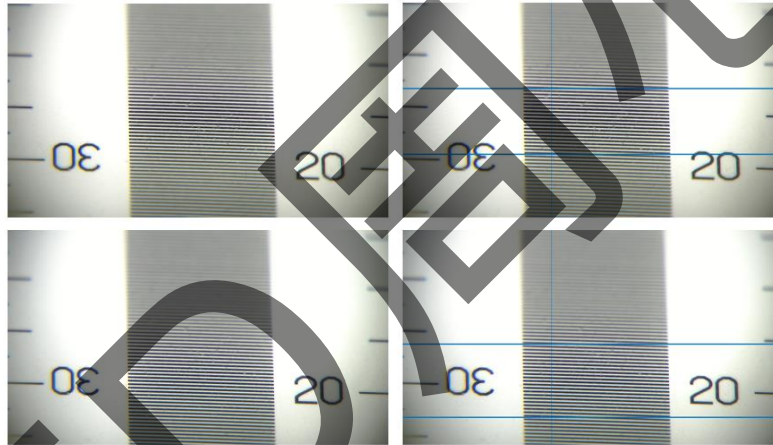


图 9: 在二十倍放大场景下, 景深测试板的成像, 其中两条蓝色横线之间的区域是图像的景深范围

放大倍数	4	6	10	16	20
平均景深 /mm	10.958	8.108	1.465	0.865	0.558

表格 1: 各个放大倍数下的平均景深

在这个部分, 我们验证了显微镜手术常用放大倍数下存在浅景深导致的部分聚焦现象 (如图 9 所示), 并用 MATLAB 代码计算出了每个放大倍数下的平均景深 (每个放大倍数下三组图像), 将用于计算每个放大倍数下每个场景需要采集的照片张数。

1.4.3 图像分割

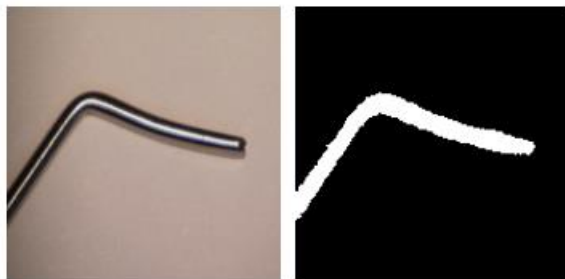


图 10：左侧为单色背景手术器械图像，右侧为分割结果

目前我们已经完成了基于 U-net（图 5）的图像分割算法的初步验证，得到了现有的手术器械的分割图像。

1.5 自身具备的知识条件

项目组成员系统地学习了大数据导论、计算机视觉、人工智能、数据库原理、算法设计等相关课程，有相关的算法开发和科研经验，能够熟练运用 MATLAB, Python（含 PyTorch、Numpy、scikit-learn 等深度学习，机器学习库），Java, SQL 等编程语言。目前我们已经完成了基于 U-net（图 5）的图像分割算法的初步验证，得到了现有的手术器械的分割图像。

项目成员所在的 iMED 深圳团队是一个专注眼科人工智能的国际化知名团队，指导老师胡衍博士也在手术辅助，术中导航等研究方向深耕多年，能够给团队成员提供有力的指导。

参考文献

- [1] “Surgery: Traumatic Macular Hole: Dr. Manish Nagpal,” YouTube, 06-Dec-2017. [Online]. Available: <https://www.youtube.com/watch?v=5oFnBgPAWy4>. [Accessed: 21-Apr-2021].
- [2] 王宁, 张小栋, 张政, 张红兵. 眼科显微手术机器人的研究与发展[J]. 机器人技术与应用, 2020(06):14-19.
- [3] L. Wang, J. Cui, S. Tabata, and M. Ishikawa, “Low-cost, readily available 3D microscopy imaging system with variable focus spinner,” Optics Express, vol. 26, no. 23, p. 30576, 2018.
- [4] M. Maximov, K. Galim, and L. Leal-Taixe, “Focus on Defocus: Bridging the Synthetic to Real Domain Gap for Depth Estimation,” 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

- [5] S. Zhuo and T. Sim, "Defocus map estimation from a single image," *Pattern Recognition*, vol. 44, no. 9, pp. 1852–1858, 2011.
- [6] C. Q. Zhao, Q. Y. Sun, C. Z. Zhang, Y. Tang, and F. Qian, "Monocular depth estimation based on deep learning: An overview," *Science China Technological Sciences*, vol. 63, no. 9, pp. 1612–1627, 2020.
- [7] Ullman, S.. "The interpretation of structure from motion." *Proceedings of the Royal Society of London. Series B. Biological Sciences* 203 (1979): 405 - 426.
- [8] V. Albiero, X. Chen, X. Yin, G. Pang, and T. Hassner, "img2pose: Face Alignment and Detection via 6DoF, Face Pose Estimation," *arXiv.org*, 14-Dec-2020. [Online]. Available: <https://arxiv.org/abs/2012.07791>. [Accessed: 21-Apr-2021].
- [9] B. A. Griffin and J. J. Corso, "Depth from Camera Motion and Object Detection," *arXiv.org*, 02-Mar-2021. [Online]. Available: <https://arxiv.org/abs/2103.01468>. [Accessed: 25-Mar-2021].
- [10] V. V. Kniaz, V. A. Knyaz, F. Remondino, A. Bordodimov, and P. Moshkantsev, "Image-to-Voxel Model Translation for 3D Scene Reconstruction and Segmentation," *Computer Vision – ECCV 2020*, pp. 105–124, 2020.
- [11] S. Wanner and B. Goldluecke, "Variational Light Field Analysis for Disparity Estimation and Super-Resolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 3, pp. 606–619, 2014.
- [12] H. Lin, C. Chen, S. B. Kang, and J. Yu, "Depth Recovery from Light Field Using Focal Stack Symmetry," *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015.
- [13] C. Chen, H. Lin, Z. Yu, S. B. Kang, and J. Yu, "Light Field Stereo Matching Using Bilateral Statistics of Surface Cameras," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
- [14] B. A. Griffin and J. J. Corso, "Learning Object Depth from Camera Motion and Video Object Segmentation," *Computer Vision – ECCV 2020*, pp. 295–312, 2020.
- [15] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Lecture Notes in Computer Science*, pp. 234–241, 2015.

二、项目研究内容（包括 1.具体研究内容和技术考核指标；2.拟解决的关键问题；3.项目可行性分析）

2.1 研究内容

在此项目中，首先拟建立一个眼科手术器械图像数据库，其次开发一个由含模糊区域的浅景深图像复原全聚焦图像的算法，最后设计一个深度学习网络模型，结合得到的全聚焦图像和图像分割方法，实现显微镜图像的深度估计。

2.1.1 深度信息金标准获取

我们考虑使用两种方法来实现这一点：第一种，采用将显微镜分光，使用两个摄像机达到双目摄像的效果，从而通过双目拍摄所形成的视差计算深度值；第二种，使用单目摄像机，移动物体，通过物体移动距离与图像坐标系下移动的距离的比值获取真实深度。我们将上述的误差更小的一种方法作为深度真实值的计算方法，用于计算每个所采集的图像的深度图（即每一个像素点为原图像素点深度值的图像）。

2.1.2 建立显微镜手术器械图像数据库

本项目是针对显微镜下的眼科手术场景展开的深度估计任务，因此在模型的设计、开发、验证过程中需要相关的手术图像及其深度图。但是由于这个场景下的深度估计本身是一项较新的工作，现存的眼科手术图像数据库并不包含深度信息的真实标注，我们需要在开始实验之前通过搭建的手术场景，采集相关图像、视频和深度图，用于后续深度估计神经网络模型训练的使用。

2.1.2.1 以单色为背景的手术器械图像



图 11：以单色为背景的手术器械图像

具体包括以单色为背景的单个眼科手术器械，多个手术器械在不同放大倍数和位置下的图像及其深度图像，也包括这种场景设置下手术器械向不同方向平移运动的视频及其深度标注。

2.1.2.2 以假眼模型、猪眼为背景的手术器械图像

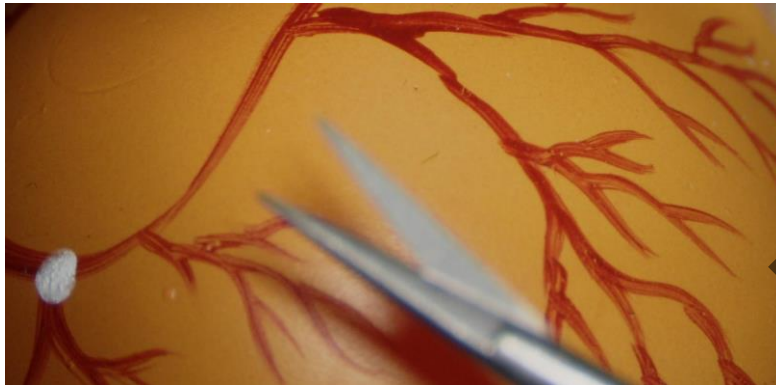


图 12：以假眼模型为背景的手术器械图像

具体包括将单个或多个手术器械作用于假眼模型、猪眼上的图像、视频及其深度标注。采集此类数据的目的，是为了提高图像中深度信息的复杂度，强化模型在真实眼科手术场景下性能。

2.1.3 获取全聚焦图像

针对高倍显微镜视野下出现的局部模糊现象，本项目拟在传统方法的基础上开发一种新模型，将局部失焦的显微镜图像复原成全景聚焦的图像，以方便医生在术中对场景的观察和判断，也用以提高后续通过图像分割获取深度图像任务的准确度。正如第一部分提到的，获取全聚焦图像的传统方法大致分为两类，一类是利用图像处理方法提取同一场景、不同聚焦区域的一组图像中的聚焦像素，拼凑成全聚焦图像，另一类是利用点扩散函数的数学性质将模糊图像复原成清晰图像。本项目将要提出的新模型将会立足于这两类传统方法，首先使用拼凑法进行尝试，然后分析精度，再尝试用点扩散函数的性质改进方法，并重点针对眼科手术图像的复原效果进行改进。

2.1.3 基于图像分割获取深度图像

图像分割是本项目用于深度预测采用的方法的基础。本项目首先将基于建立的眼科手术图像数据库，结合 U-net（如图 5 所示）进行图像分割。再利用图像分割结果和相机标定算法得到的相机内参，改进第一部分提及的结合图像分割和相机运动的新型深度估计算法。具体的改进方法，是将现有的纯数学方法跟深度神经网络（比如卷积神经网络，深度残差网络等具有良好性能的深度神经网络架构）结合在一起进行深度预测，设计出适用于眼科手术场景的，基于图像分割和深度学习的深度估计算法。最后将该算法运用在离体猪眼和临床视频上进行性能验证。

2.2 技术考核指标

- (1) 本项目所建立的眼科手术显微镜图像需包括常见的眼科手术器械，如镊子、剪子、针头、冲洗器等，每种各采集一套单色背景和假眼背景的图像以及移动视频。
- (2) 本项目所复原的全聚焦高倍眼科手术显微镜图像需为全景像素清晰的图像
- (3) 本项目所进行的深度估计项目计划以平均百分比误差衡量误差，且最终结果的平均百分比误差须不高于 30%。

2.3 拟解决的关键问题

- (1) 通过相机标定得到显微镜内参，通过景深标定板测量景深，借助相机标定计算出来的真实深度标注信息建立手术器械显微镜图像数据库
- (2) 利用浅景深图像复原全聚焦图像
- (3) 采用图像分割的方法，利用合成的全聚焦图像预测场景的深度图像

2.4 项目可行性分析

从原理上来说，全聚焦图的合成和深度估计在其他场景下均有大量相关的传统方法和深度学习方法。本项目立足于这些方法，特别着眼于像素提取和图像分割的系列方法进行改进，因此具有科学上的可行性。

从实验条件上来说，本项目数据采集所需的手术显微镜（上海铁德公司 SM2000J 实验手术显微镜），精密移动平台（大恒光电 GCM-T 系列精密平移台，大恒光电 GCM-V 系列精密侧升降台）都已经具备。

从项目组员上来说，项目组员均已在刘江老师课题组有一年多的课题研究经验，主要方向均集中在人工智能技术在眼科方向的应用，对该领域的课题研究思路和方法有足够的了解。

三、项目实施方案（包括 1.项目人员组成及分工；2. 项目研究进度安排：包括查阅资料、选题、自主设计项目研究方案、开题报告、实验研究、数据统计、处理与分析、研制开发、填写结题表、撰写研究论文和总结报告、参加结题答辩和成果推广；3.配套条件要求：包括仪器设备、场地、材料、资料、实验课时及指导教师要求等）

3.1 项目人员组成及分工

周心怡，项目负责人。主要负责全聚焦图像的合成和深度估计。

周颖泉，项目参与人。主要负责以图像分割为基础的深度估计。

聂秋实，项目参与人。主要负责相机标定和数据采集。

曾鸣，项目参与人。主要负责查找论文和相关资料。

3.2 研究进度安排

2021.4-5 设计项目研究方案，进行数据采集工作

2021.6-7 设计模型，进行实验，统计实验数据

2021.8-11 完成初步实验，进行对比试验，进一步改善模型性能

2021.12-2 撰写结题报告与论文，申请专利

3.3 配套条件要求

设备：上海轶德公司 SM2000J 实验手术显微镜，大恒光电 GCM-T 系列精密平移台，大恒光电 GCM-V 系列精密侧升降台。

场地：iMED 深圳团队南方科技大学实验室。

材料：假眼模型，猪眼模型

四、项目经费预算（包括大概支出科目、金额、计算根据及理由）

序号	支出科目	预算金额	计算根据及理由
1	调研	500 元	本项目相关文献的购买和下载。
2	资料、复印	500 元	相关文献资料、项目文件打印费。
3	实验材料	3000 元	用于购买眼科显微手术器械、猪眼。
6	发表论文	6000 元	计划投稿一篇国际/国内会议的论文。
7	文献/知识产权事务费	6000 元	计划申请两项专利，每项 3000 元。
合计		16000 元	

五、预期成果（研究论文、专著、调研报告、设计方案、专利、软件、产品、成果鉴定等）

项目成果为至少发表一篇学术论文，至少申请一项专利。

六、本项目的创新之处

（1） 创新地提出了一种眼科手术显微镜图像深度真实值的测量方法

测量深度的一般方法有 3D 激光雷达传感器，如 Velodyne 的 HDL-64E 激光雷达传感器，还有红外测距，如微软 Azure Kinect 技术，以及立体相机，如 STEREO LABS 的 ZED 2 立体相机。这些测量深度的方法被广泛地引用于大尺度自然场景，但是因为显微镜场景空间狭小，这些测量方法不适用于本项目针对的场景。

因此我们基于相机标定和显微镜本身的光学性质，提出了一种适用于眼科手术显微镜的深度真实值测量方法，具有创新性。

（2） 创新地将全聚焦图像的复原工作应用在眼科手术领域

本项目在实验中验证了，手术显微镜在高倍放大倍数下会出现由浅景深导致的局部模糊现象，同时立足于现有的全聚焦图像合成方法开发了一个性能更好的、适用于显微镜视角的全聚焦图像合成手法。

（3） 创新地建立了一个具有深度标注信息（深度图）的眼科手术显微镜图像数据库

常见的眼科手术显微镜图像数据库都是只包含 RGB 图像或者视频，并不包含深度标注信息。而我们收集的数据将会包含深度信息，这在以前的数据库内容的基础上，也具有很大的创新意义。

（4） 创新地设计了一个基于深度学习和图像分割方法的深度估计模型

在显微镜场景下的传统深度估计方法有很多，但是普遍鲁棒性不强，而现有的大尺度场景深度估计算法也因为不适用于图像数据较难获取，物体运动尺度很小的眼科手术场景而无法直接用来预测显微镜图像的深度，所以本项目基于深度学习和图像分割方法提出的显微镜图像深度估计算法具有很高的创新性。

七、项目诚信承诺

本项目负责人和全体成员郑重承诺，该项目研究不抄袭他人成果，不弄虚作假；按项目研究进度保质保量完成各项研究任务。

项目负责人签名：

项目组成员签名：

年 月 日

年 月 日

八、指导教师意见（包括项目研究的选题意义及研究方案的科学性、可行性等）

该项目将深度学习方法运用在眼科手术图像的深度估计任务上，具有一定的创新性。眼科手术中，显微镜为医生提供的手术视野有限，显示的手术图像也存在焦面不平衡所导致的失焦区域，这导致医生难以判断手术器械的深度，容易造成因器械触碰视网膜引起损伤甚至失明的医疗事故。在这样的问题背景下，设计一套显微镜视野下的手术器械深度估计算法，十分地有现实意义。

此计划书中的研究方案整体较为完整，课题组成员分工明确且有一定的前期研究基础，方案具有相当的可行性。

指导教师签字：

年 月 日

九、院、系初审意见	
<div>负责人签字：年 月 日</div>	
十、学校意见	
<div>负责人签字：</div>	<div>(盖 章)</div> <div>年 月 日</div>