

中国科学技术大学

工程硕士研究生学位论文

开题报告

论文题目 果蝇电镜图像的全脑
血管三维重建与分析

学 生 姓 名 陈思远

学 生 学 号 SA19006157

学 校 导 师 吴枫、陈雪锦

企 业 导 师 中科类脑刘海峰

领 域 代 码 085208

研 究 方 向 计算机视觉

所 在 单 位 电子工程与信息科学系

中国科学技术大学研究生院

填表日期：2021 年 10 月 6 日

说 明

1. 工程硕士学位论文的开题报告是保证论文质量的一个重要环节，为了加强对工程硕士研究生培养的过程管理，规范其学位论文的开题报告，特制此表。
2. 工程硕士学位论文开题报告，应该在工程硕士学位授予点或培养单位组织的学术报告会上报告，听取意见，论证后再填写此表。
3. 此表一式两份经导师和培养单位负责人签字后，交培养单位研究生教学管理办公室存档。
4. 工程硕士研究生在申请学位论文答辩时，必须提交该学位论文开题报告。

一、 简况

研究生简况	学 号	SA19006157		姓 名	陈思远	姓名拼音	Chen Siyuan
	性 别	女	身份证号	340123199804070047		出生年月	1998 年 4 月
	工程领域	电子与通信工程			研究方向	计算机视觉	
	入学时间	2019 年 9 月		录取方式	联考 统考 自考	培养方式	脱产 不脱产
	本科毕业时间		2019 年 6 月		本科毕业学校	合肥工业大学	
论文类型与性质	名称	中文	果蝇电镜图像的全脑血管三维重建与分析				
		英文	Three-dimensional Reconstruction and Analysis of Blood Vessels from Whole Cerebral Drosophila Electron Microscopic Images				
	类别	1. 技术攻关研究 √ 2. 工程项目策划 3. 工程设计或技术改造 4. 新工艺、新材料、新设备、新产品的研制与开发					
		形式	1. 工程设计 2. 研究论文 √				
	性质		1. 应用技术研究 √ 2. 技术开发				
论文内容和意义	摘要	<p>全脑血管系统的重建对于理解生理和病理脑功能至关重要。电镜成像技术的发展使得对全脑组织进行高分辨率成像成为可能。然而，在不同脑区血管的形态巨大差异，给全脑血管自动分割和重建带来挑战。首先，基于深度学习的方法虽然有较强的特征表达能力，却需要大量有标签数据，然而人工标注代价高昂。其次全脑的数据特性差异大，易导致模型泛化性能差。面向 TB 级的全脑影像，急需自动、高效的重建和分析方法。同时，由于现有追踪算法不能完整追踪出超长距离的血管，因此在全脑血管追踪的结果中存在大量断裂血管，探究聚合断裂血管段的方法也是本课题需要解决的问题。本课题针对这些挑战，研究了一种高效精确的全脑血管重建方法，设计了血管分割、血管追踪、追踪片段聚合等重建流程，并基于全脑重建结果分析了血管分布，对理解生物脑功能有重要意义。</p>					
		<p>主题词数量不多于三个，主题词之间空一格（英文用“/”分隔）</p>					
	主题词	中文	全脑电镜、血管、三维重建				
		英文	Electron Microscopic Data of Whole Brain / Blood Vessels / Three-dimensional Reconstruction				

二、选题依据

1. 阐述该选题的研究意义，或工程设计的价值和意义，国内外概况和发展趋势，选题的先进性和实用性，技术难度及工作量。

1.1 研究意义

血管是生物运送血液的管道，按运输方向可分为动脉、静脉与微血管。动脉从心脏将血液带至身体组织，静脉将血液自组织间带回心脏，微血管则连接动脉与静脉，是血液与组织间物质交换的主要场所。生物体内各处血管病变会导致多种疾病。其中，脑血管系统的变化是影响大脑的许多疾病的一个关键特征。原发性血管病变、血管危险因素(如糖尿病)、创伤性脑损伤、血管闭塞、中风等均影响脑血管网络，干扰正常微循环和血管功能[1-3]。脑血管系统的改变也见于神经退行性疾病，如阿尔茨海默病、淀粉样病。这些阿尔茨海默病的特征会导致血管异常重塑。此外，毛细血管稀疏常被用作判别血管损伤的标志[4]。因此，对全脑血管进行高分辨率的重建具有重要意义。

同时，随着电镜成像技术的发展已能对全脑组织进行高分辨率成像，使得对全脑血管进行高分辨率的重建成为可能。然而，一个完整的果蝇大脑纳米级电镜扫描图像数据集 FAFB (Full Adult Fly Brain)，物理分辨率能达到 $(x,y,z) = (4, 4, 40)$ 纳米/体素，体量达到 40 万亿级像素，存储这样的数据通常需要数 TB 量级的空间[5]。面对如此超大规模的图像数据，依靠人工处理和分析已不可能，如何高效的分析超大规模脑数据，从中提取出有效的血管连接关系成为了一个亟需解决的问题。

综上所述，全脑血管系统的重建对于理解生理和病理脑功能至关重要，针对高效的全脑血管重建算法研究具有重要意义。

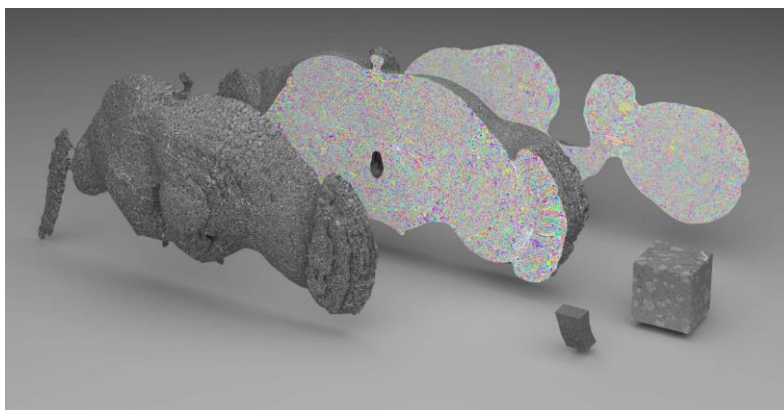


图 1 果蝇全脑电镜数据集[5]

1.2 国内外研究现状

近年来，众多研究人员对生物数据的血管重建开展了大量的研究工作，包括视网膜血管、心脏血管、肝脏血管、脑血管等。血管重建主要有以下步骤：分割血管得到分割图，追踪分割图得到血管的几何模型，基于形态分析修正拓扑结构。

1.2.1 血管分割发展现状

血管分割相关工作主要分为 2D CNN (Convolution Neural Network) 分割方法和 3D CNN

分割方法。

Wang 等人提出用 2D CNN 提取视网膜血管特征，并用机器学习方法如支持向量机、条件随机场做逐像素二分类[6, 7]。然而这种后处理方法由于计算成本可能很高，使得这类方法很难适用于大规模数据集。针对这一问题，Smistad 等人提出直接训练端到端的神经网络从而提升分割效率[8, 9]，但是网络层数太少，特征学习能力很弱。据此 R2U-Net [10]在编解码过程中将卷积层换成循环残差模块，从而有效增加网络深度，得到更具有表达性的特征。[11]则采用域适应方法促使分割层提取到“域不变”的特征，从而提升泛化性能。然而上述方法没有结合三维信息，所以会造成三维信息的缺失。为了突破二维信息的局限，Wu 等人考虑用不同角度的图像辅助分割，提出多路径的二维鼠脑血管分割方法[12, 13]，即用合并多角度的二维分割结果，以这种方式间接利用三维结构，如图 2。

由于脑数据集多为三维数据，因此直接使用三维卷积神经网络是很自然的。不同于多路径的二维分割，Todorov 等人提出直接训练 3D FCN [19]分割鼠脑光镜血管，之后还有一系列三维卷积神经网络分割鼠脑血管的工作[14-18]。这些工作通过使用 3D 卷积核分割三维数据块，能够结合更多相邻帧的信息，充分学习三维特征，从而获得连续性更好的分割结果。但是这些方法的计算量远超二维卷积网络，又受 GPU 内存限制，输入数据块的大小受限，导致训练和测试速度较慢。虽然速度较慢，但由于其优越的三维连续性，3D CNN 仍是脑血管分割的主流方法。针对速度慢的问题，本文拟修正卷积核来减少计算量，从而提升分割速度。

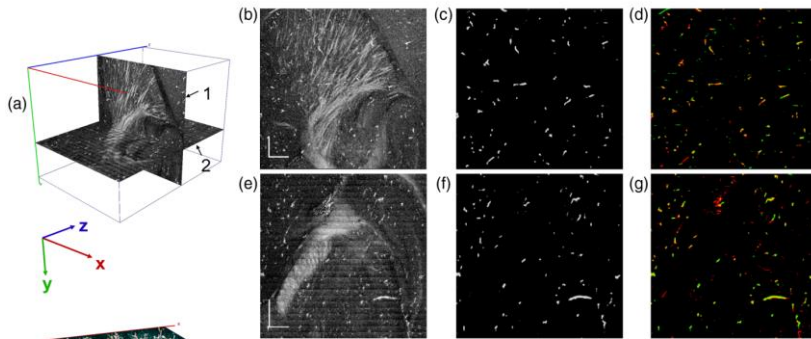


图 2 Wu 等人提出从不同角度分割鼠脑光镜血管[12]

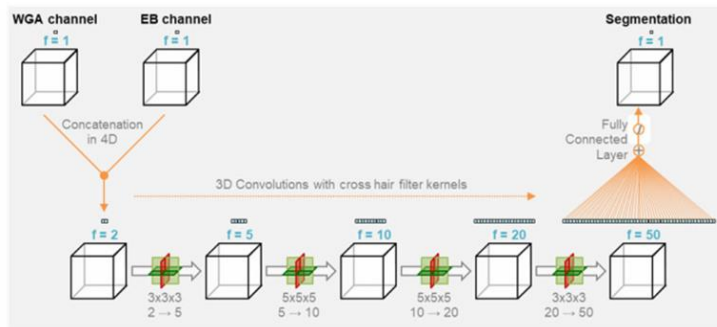


图 3 3D CNN 分割鼠脑光镜血管[14]

1.2.2 血管追踪和修正发展现状

血管追踪（Tracing）是创建其几何模型的过程。Zeng 等人通过提取中心线来得到肝脏血管的骨架[20]，此算法比较简洁，运算速度较快，但是易受噪声影响。尤其是提取骨架的物体有“尖峰”时，会导致骨架不沿着物体主干而产生断裂。MOST[21]使用三维局部信息，迭代地挖掘体素的邻域，并基于连通性进行聚类。然而 MOST 在追踪过程严重依赖于分割结果，并且只利用到局部分割信息，因此对分割质量有较高要求。在 MICCAI 2008 冠状动脉追踪挑战赛上，Wolterink 等人使用深度学习预测半径和固定 N 个方向的概率值，然而真实方向不在这 N 个方向时，会导致追踪方向误差[22]。上述方法都只利用了局部信息，因此易受噪声影响，因此本课题拟采用更全局的追踪方法进行全脑追踪。

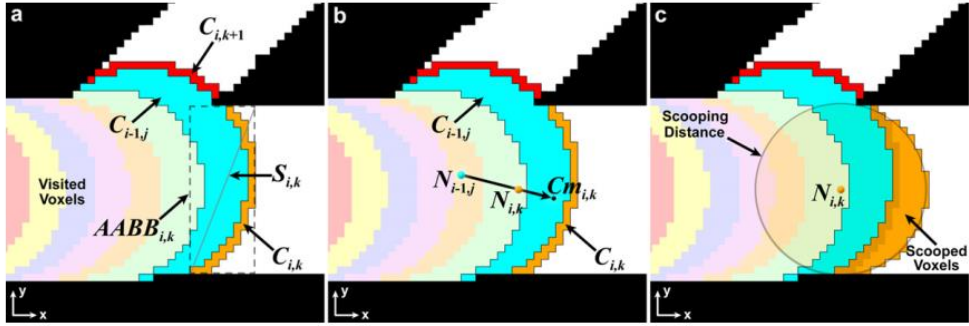


图 4 MOST[21]迭代挖掘邻域并聚类

由于分割和追踪过程中误差的存在，为了得到更准确的拓扑结构，常常需要修正错误骨架、补全断裂等。Dashtbozorg 等人先对分割图施加细化算法得到血管骨架，又定义了多种骨架错误连接类型，并根据连接点个数和分叉情况给节点归类，从而检测并纠正连接错误[23]。Mou 等人先引入血管局部服从二次曲线分布的先验，并模拟随机游走来连接分割图断裂[24]。Li 等人提出根据方向、粗细、距离计算相邻段之间的影响系数 λ ，以此更新置信度，多次迭代达到减少错误判别的目标[25]。然而上述方法只引入了简单的先验，修正后仍存在一定连接误差，所以本课题拟引入更多形态信息检测断裂，并将聚合断裂的任务转换为图割优化问题，从而优化全脑重建结果。

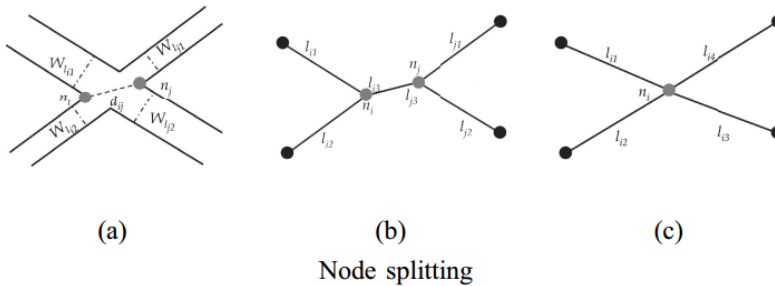


图 5 Mou 等人定义的一种错误骨架类型[23]

上文主要介绍了光镜脑血管和其他器官血管的分割与重建工作。目前脑血管分割的主流方法是三维卷积神经网络。此外拓扑结构的完整性、准确性也是各重建工作的焦点，为了获取准确的统计特征、拓扑结构，往往需要追踪、修正等处理。本课题针对以往方法存在的问

题，设计了高效准确的果蝇全脑血管重建流程。

1.3 选题先进性和实用性

先进性：本设计首次实现大规模电镜数据的血管重建。分割阶段我们构建了第一个电镜图像血管数据集。考虑到血管突出的延续性特征，采用三维卷积神经网络分割大规模数据，保证全脑拓扑较好的完整性。追踪阶段采用了更准确的追踪方法，并根据血管形态检测并聚合断裂，从而得到高质量的全脑血管几何模型。

实用性：本设计提出了第一个大规模电镜数据血管重建的完整流程，采用更高效的分割、追踪方法获得了更准确的重建结果，减少了全脑重建所需时间。此外还对重建结果做了丰富的可视化，有利于总结全脑血管分布规律。本课题是第一个电镜血管重建的工作，因此可作为后续电镜血管重建相关研究的基准。

1.4 技术难点

● 难点 1 数据无标注，数据量庞大，数据分布差异大

本课题基于 FAFB 进行全脑血管重建与分析。主流的血管分割算法采用监督式学习来训练神经网络，训练神经网络依赖有标注数据，而 FAFB 数据集并没有血管标签，且数据量庞大。此外不同脑区的血管特征存在明显的差异，这些都为血管分割带来了挑战。

● 难点 2 血管追踪存在断裂情况

现有追踪方法在追踪超长血管时表现不佳，追踪结果中仍存在大量断裂、不完整的情况。如何在追踪结果的基础上尽可能还原血管的连通性和完整性，探究聚合断裂血管段的方法也是本课题需要解决的问题。

2. 主要参考文献（列出作者、论文名称、期刊名称、出版年月）。

- [1] Bennett R E, Robbins A B, Hu M, et al. Tau induces blood vessel abnormalities and angiogenesis-related gene expression in P301L transgenic mice and human Alzheimer's disease[J]. Proceedings of the National Academy of Sciences, 2018, 115(6): E1289-E1298.
- [2] Joutel A, Monet-Leprêtre M, Gosele C, et al. Cerebrovascular dysfunction and microcirculation rarefaction precede white matter lesions in a mouse genetic model of cerebral ischemic small vessel disease[J]. The Journal of clinical investigation, 2010, 120(2): 433-445.
- [3] Obenaus A, Ng M, Orantes A M, et al. Traumatic brain injury results in acute rarefaction of the vascular network[J]. Scientific reports, 2017, 7(1): 1-14.
- [4] Edwards-Richards A, DeFreitas M, Katsoufis C P, et al. Capillary rarefaction: an early marker of microvascular disease in young hemodialysis patients[J]. Clinical kidney journal, 2014, 7(6): 569-574.
- [5] Zheng Z, Lauritzen J S, Perlman E, et al. A complete electron microscopy volume of the brain of adult *Drosophila melanogaster*[J]. Cell, 2018, 174(3): 730-743.
- [6] Wang S, Yin Y, Cao G, et al. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning[J]. Neurocomputing, 2015, 149: 708-717.
- [7] Xue D X, Zhang R, Feng H, et al. CNN-SVM for microvascular morphological type recognition

- with data augmentation[J]. *Journal of medical and biological engineering*, 2016, 36(6): 755-764.
- [8] Smistad E, Løvstakken L. Vessel detection in ultrasound images using deep convolutional neural networks[M]//*Deep Learning and Data Labeling for Medical Applications*. Springer, Cham, 2016: 30-38.
- [9] Prentašić P, Heisler M, Mammo Z, et al. Segmentation of the foveal microvasculature using deep learning networks[J]. *Journal of biomedical optics*, 2016, 21(7): 075008.
- [10] Alom M Z, Hasan M, Yakopcic C, et al. Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation[J]. 2018.
- [11] Javanmardi M, Tasdizen T. Domain adaptation for biomedical image segmentation using adversarial training[C]//*2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018: 554-558.
- [12] Wu Y, Xia Y, Song Y, et al. Vessel-Net: retinal vessel segmentation under multi-path supervision[C]//*International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2019: 264-272.
- [13] Li T, Liu C J, Akkin T. Contrast-enhanced serial optical coherence scanner with deep learning network reveals vasculature and white matter organization of mouse brain[J]. *Neurophotonics*, 2019, 6(3): 035004.
- [14] Todorov M I, Paetzold J C, Schoppe O, et al. Automated analysis of whole brain vasculature using machine learning[J]. *bioRxiv*, 2019: 613257.
- [15] Todorov M I, Paetzold J C, Schoppe O, et al. Machine learning analysis of whole mouse brain vasculature[J]. *Nature methods*, 2020, 17(4): 442-449.
- [16] Tahir W, Kura S, Zhu J, et al. Anatomical modeling of brain vasculature in two-photon microscopy by generalizable deep learning[J]. *BME Frontiers*, 2021, 2021.
- [17] Fu F, Wei J, Zhang M, et al. Rapid vessel segmentation and reconstruction of head and neck angiograms using 3D convolutional neural network[J]. *Nature communications*, 2020, 11(1): 1-12.
- [18] Haft-Javaherian M, Fang L, Muse V, et al. Deep convolutional neural networks for segmenting 3D in vivo multiphoton images of vasculature in Alzheimer disease mouse models[J]. *PloS one*, 2019, 14(3): e0213539.
- [19] Long, J., Shelhamer, E., & Darrell, T. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015: 3431-3440.
- [20] Zeng Y, Zhao Y, Tang P, et al. Liver vessel segmentation and identification based on oriented flux symmetry and graph cuts[J]. *Computer methods and programs in biomedicine*, 2017, 150: 31-39.
- [21] Rodriguez A, Ehlenberger D B, Hof P R, et al. Three-dimensional neuron tracing by voxel scooping[J]. *Journal of neuroscience methods*, 2009, 184(1): 169-175.
- [22] Wolterink J M, van Hamersvelt R W, Viergever M A, et al. Coronary artery centerline extraction in cardiac CT angiography using a CNN-based orientation classifier[J]. *Medical image analysis*, 2019, 51: 46-60.
- [23] Dashtbozorg B, Mendonça A M, Campilho A. An automatic graph-based approach for

- artery/vein classification in retinal images[J]. IEEE Transactions on Image Processing, 2013, 23(3): 1073-1083.
- [24] Mou L, Chen L, Cheng J, et al. Dense dilated network with probability regularized walk for vessel detection[J]. IEEE transactions on medical imaging, 2019, 39(5): 1392-1403.
- [25] Li L, Verma M, Nakashima Y, et al. Joint learning of vessel segmentation and artery/vein classification with post-processing[C]//Medical Imaging with Deep Learning. PMLR, 2020: 440-453.
- [26] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.
- [27] Çiçek Ö, Abdulkadir A, Lienkamp S S, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation[C]//International conference on medical image computing and computer-assisted intervention. Springer, Cham, 2016: 424-432.
- [28] Quan T, Zhou H, Li J, et al. NeuroGPS-Tree: automatic reconstruction of large-scale neuronal populations with dense neurites[J]. Nature methods, 2016, 13(1): 51-54.
- [29] Ming X, Li A, Wu J, et al. Rapid reconstruction of 3D neuronal morphology from light microscopy images with augmented rayburst sampling[J]. PloS one, 2013, 8(12): e84557.

三、课题内容及具体方案

拟采用的设计方法，技术路线，实验方案的可行性分析。

1 设计方法

为了实现果蝇电镜图像的全脑血管重建，本课题设计一种高效、精确的血管重建方法。血管分割和追踪得到全脑血管重建模型后，通过形态分析聚合断裂血管，继而统计全脑血管分布，从而探究其生物学意义。

针对数据无标注和数据分布差异大的难点，本课题制作了电镜血管数据集，充分考虑了血管特征的多样性，以提升模型的泛化性能；针对数据量庞大的难点，本课题改进了 3D U-Net，并使用更高效的追踪方法，以提升重建速度。

针对血管追踪存在断裂情况的问题，本课题根据断裂处的典型特征初步检测断裂，并将聚合血管断裂问题转换为图割优化问题，从而恢复全脑血管的拓扑连通性。

整体重建思路如图：

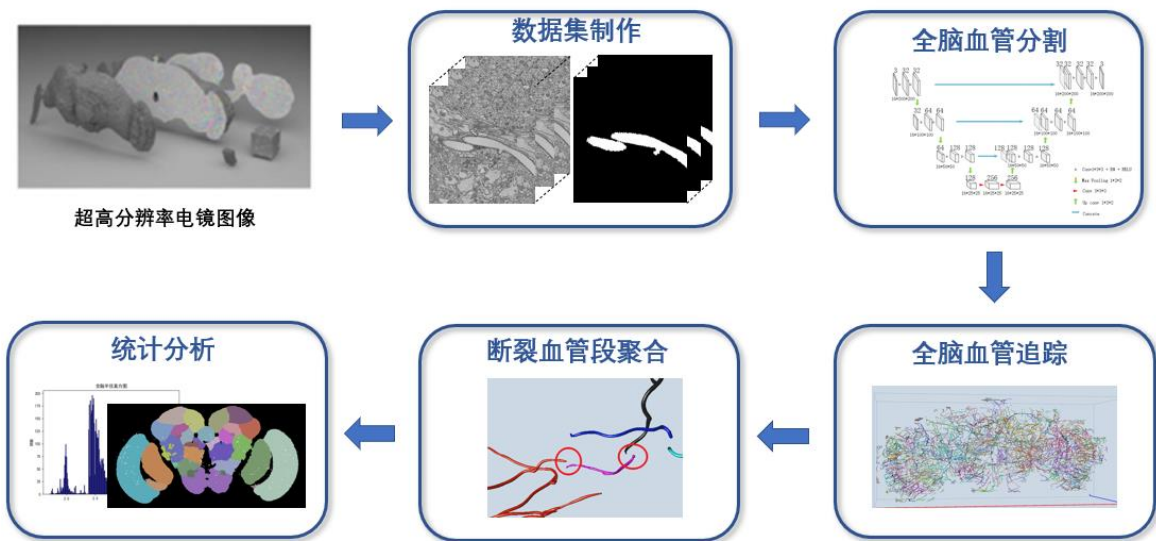


图 6 课题研究思路

2 技术路线

本章分为三部分依次阐述技术路线：血管标注与分割，血管追踪与断裂聚合，全脑血管分布的统计分析。

2.1 血管标注与分割

FAFB 不同区域的血管形态上存在统一性与多样性。其中统一性指血管三维连续性，多样性指二维平面特征的多样性。因此为了网络有更好的泛化性，我们先对血管形态分析归类，制作电镜血管数据集，再用三维卷积神经网络分割大规模数据。

2.1.1 数据集制作

(1) 数据分块

FAFB 的物理分辨率可达到 (4, 4, 40) 纳米/体素, 体量达到 40 万亿级像素, 存储需要数 TB 量级的空间。面对如此超大规模的图像数据, 分割、存储分割结果会耗费大量资源。针对这个问题, 本设计采用牺牲分辨率的方式来换取时间资源。如果在较低分辨率的 Z 轴做降采样, 势必会模糊更多信息。因此选择牺牲 X、Y 轴分辨率, 在分辨率 (16, 16, 40) 纳米/体素的数据集上分割。除此之外, 由于 GPU 内存的限制, 需要将全脑数据分块输入。

(2) 分析数据

近年来已有很多基于 FAFB 的神经元分割、线粒体分割、突触分割的工作, 其中比较好的结果可以当作对应分割任务的标签/伪标签。然而目前并没有基于该数据集的血管分割结果, 标签也就无从谈起。

为获取训练 CNN 的数据集, 我们先对 FAFB 各区域的血管做了分析归类, 再选择数据块进行人工标记。有如下分类结果:

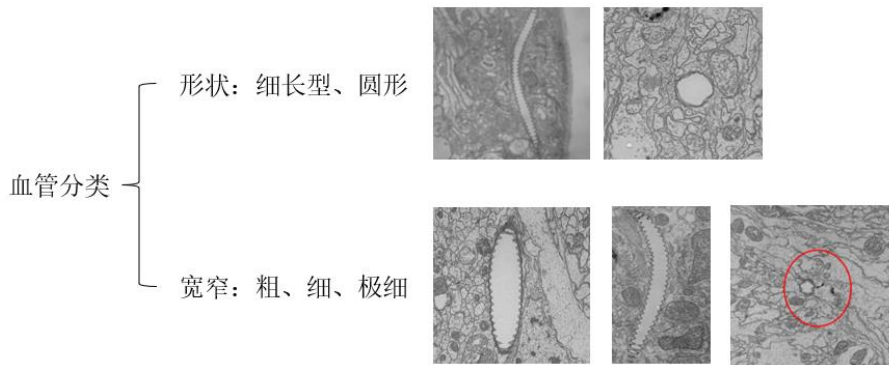


图 7 FAFB 全脑血管分类

全脑血管统一的平面特征是波浪状边缘和平滑的内部区域, 统一的三维特征则是具有较好的延续性。除此之外, 多样性具体表现在:

a) 形状多样性。

我们按图像中血管的平面形状将其分为细长形和圆形。一般来说, 细长型血管是当前血管段的中轴线与图像平面近似平行, 即纵剖面; 与之相反, 圆形血管则是血管段的横剖面。一般来说, 虽然圆形血管仍然保持着内部的平滑, 但是其周长却小于同半径的细长血管, 因此边缘的波浪特征相对较弱。分割这种横剖面血管就体现了用 3D CNN 学习血管三维延续性的优势。

b) 粗细多样性。

闭锁式循环 (closed circulation) 生物, 如哺乳类、鸟类、爬虫类、鱼类, 则由动脉连接毛细血管再接至静脉, 最后回归心脏。相较动、静脉, 毛细血管要细的多。所以在这些生物中血管的粗细变化较大。而昆虫虽然是开放式循环生物, 血管体系中只有动脉[23], 其全脑的血管粗细也在一定范围内波动。在这里我们按照观察经验, 初步将血管粗细分成三个等级: 极细血管 (<0.5 微米)、细血管 (0.5 微米-1.5 微米)、粗血管 (>1.5 微米)。

c) 数据分布多样性。

全脑各区域都有血管存在，而不同区域的细胞结构不尽相同，这就给分割网络的泛化性带来了不小的挑战。经过实验，发现确实有部分区域的平面特征和血管相似，容易被错检。如图 12，该区域的细胞也具有内部平滑的特点，因此很容易被误检。

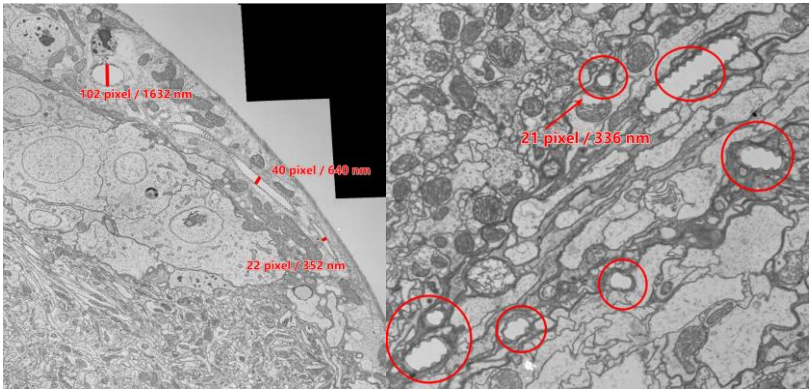


图 8 不同粗细的血管示例，左图是一个块中的血管示例，右图是放大后的小血管

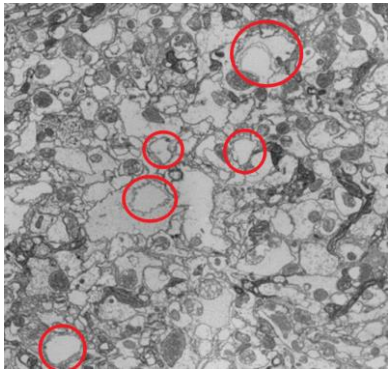


图 9 红圈内的细胞形态和血管形态相似，容易误检

（3）数据标记

基于上述对数据的探究分析，我们选取了各类别的数据块进行人工标记，构成首个电镜血管数据集。每个数据块大小是（1736，1736，26）体素，我们共选择 30 个数据块进行人工标记，详情如表 1。每种随机选择 75%做训练集，25%做测试集。

表 1 标记数据块详情	
数据特征	数据块个数
粗血管	连续 10 块 × 2
细血管	连续 4 块 × 1
圆血管	连续 2 块 × 1
相似块	不连续的 4 块

2.1.2 3D CNN 分割

血管的相邻帧连续性是血管最重要的特征之一，也是区别于其他细胞组织的主要特征之一，血管在 Z 轴上的跨度可多达数千帧。本工作是首个电镜图像的血管重建工作，所以我们认为最先需要保证的是全脑血管拓扑的完整性。要生成连续性较好的重建结果，可以考虑引入相邻切片的信息来帮助预测。

卷积神经网络 (CNN) 已广泛应用在众多计算机视觉任务中，包括医学图像分割。医学影像单个图像的手动分割可能需要数小时，而 CNN 可以在几秒钟内准确地分割目标结构。因此 CNN 已广泛应用于医学领域，以降低相关成本。

CNN 可以按所使用的卷积核的维度分类。2D CNN 使用 2D 卷积核来预测单个切片的分割图，利用切片长、宽两个维度上的信息，通过一次预测一个切片来预测整个体积的分割图。由于 2D CNN 将单个切片作为输入，所以本质上无法利用来自相邻切片的上下文，只能学到血管的平面特征，从而导致大量虚警、漏检。

针对这一问题，3D CNN 通过使用 3D 卷积核对三维数据块进行分割，不仅能学习二维平面特征，还可以结合相邻切片的信息。当 2D CNN 遇到连续多个困难的数据时，因为只有当前帧的信息，所以在这些图像上的分割结果都很差。但是这种情况下即使 2 维平面特征不好辨别，三维的连续性特征依然存在。如果采用 3D CNN 就能联系附近数据的信息判断是否符合连续性特征，从而优于 2D 网络的分割结构。

综上所述，我们采用 3D CNN 进行全脑血管的分割。

U-Net [26] 是比较早的使用全卷积网络进行医学图像语义分割的算法之一，论文中使用包含压缩路径和扩展路径的对称 U 形结构在当时非常具有创新性，可较好地应用于逐像素监督学习的分割场景。3D U-Net [27] 是 3D CNN 的典型代表，同时也是 2D U-Net 的扩展版本。由于其优异的性能，在医学图像的三维分割中得到了广泛的应用。

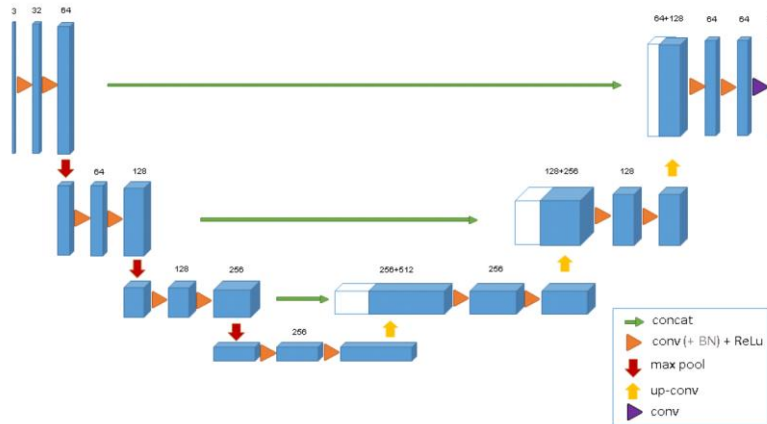


图 10 3D U-Net [27]

3D U-Net 由一个压缩路径(左)和一个扩展路径(右)组成。在压缩路径中，每个层包含两个 $3 \times 3 \times 3$ 卷积和一个 $2 \times 2 \times 2$ 池化层，分别用于提取特征和降采样。扩展路径中每一层先通过反卷积将特征图的尺寸乘 2，同时将其个数减半，然后和左侧对称的压缩路径的特征图合并，来自相同分辨率层的快速连接为合成图像提供了必要的高分辨率特征。在最后一层， $1 \times 1 \times 1$ 卷积将输出通道

的数量减少到标签的数量 3。每个卷积后面都跟着 ReLU 层和批处理规范化(batch normalization)。

2.1.3 网络调整

针对数据量大的问题，我们拟修改网络以更高效地分割。

3D CNN 利用切片间上下文的能力可以提高分割准确性，但由于使用的参数数量增加，计算成本也随之增加。为了减少计算量、提高分割速度，我们在网络的初始三层将 Z 方向的卷积核大小设为 1，也即使用二维卷积开始前三层的卷积，当网络的物理感受野接近各向同性后，将 z 方向的卷积核大小调整为与其他方向相同，也即使用各向同性的卷积核进行深层的卷积。通过这样的卷积过程，网络的参数量、计算量也显著减少，使得我们可以使用更大的图像块输入到网络，让网络使用随机梯度下降算法训练时，对梯度有更好的估计，加快网络的收敛速度，且更不易于陷于局部最优。

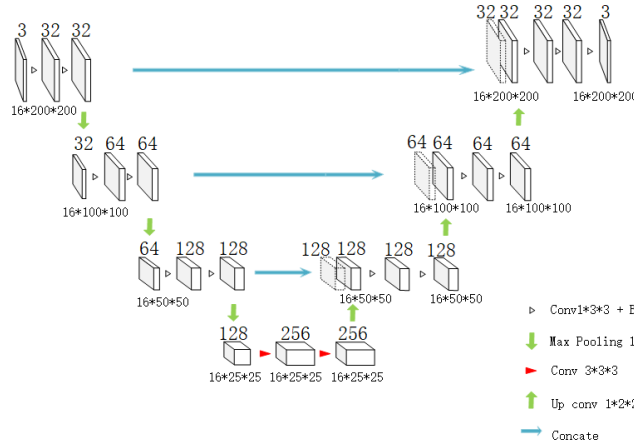


图 11 网络修正

2.2 血管追踪与断裂聚合

血管追踪 (Tracing) 是构建血管的数字模型的过程。常用血管追踪算法受限于当前位置附近很小范围内的信息，非常容易受分割图里的噪声影响产生分割漏检导致断裂。由于我们的数据量庞大，更需要采用一种考虑全局关系的准确高效追踪方法。

本节先介绍基于 NeuroGPS-Tree[28]的血管追踪方法，然后阐述基于图割优化的血管断裂聚合方法。

2.2.1 基于 NeuroGPS-Tree 的血管追踪

神经元重建和血管重建工作中都常用到追踪算法，其中神经元轴突树突和血管形态类似，都是管状结构，所以二者的追踪算法有相通之处。神经元追踪算法中的代表方法 NeuroGPS-Tree 算法在追踪时利用到更全局的信息，采用镭射采样算法 (Rayburst Sampling Method[29]) 判断当前点位置和计算半径，并采用丰富的后处理处理追踪错误，且高效便捷，适合大规模数据。因此，本课题采用 NeuroGPS-Tree 做基础血管追踪。

追踪前先根据概率值选择概率值较大的点为初始种子点。从初始种子点开时，设置初始切向为当前点指向附近球形区域最亮点的向量，选择附近方向一致的其它点，用这些点拟合曲线，并微调当前点位置。确定当前点位置后用镭射采样算法计算半径，并判断当前是否处于特殊位置，从而选择继续或终止追踪。

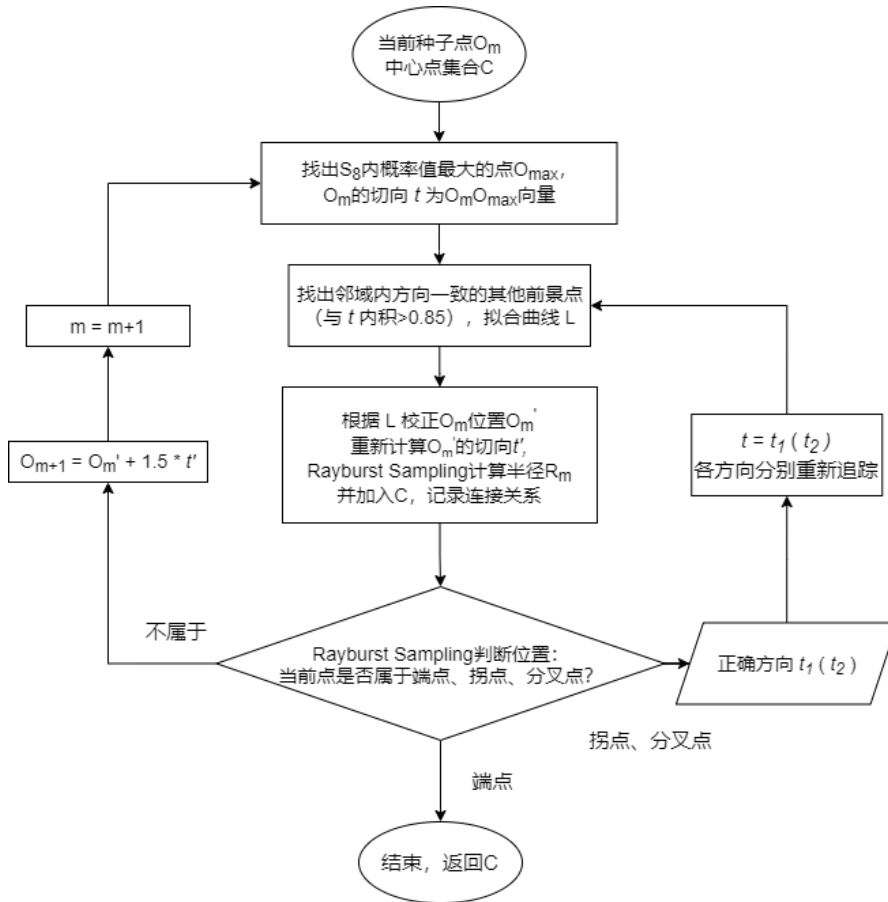


图 12 NeuroGPS-Tree 追踪步骤

2.2.2 断裂血管段聚合

在全脑血管重建的过程中, NeuroGPS-Tree 算法在追踪超长距离的血管时会产生大量血管断裂。为解决断裂问题, 本课题先根据断裂处的典型特征初步检测断裂, 接着利用位置关系和形态特征判断血管段属于同一根血管的概率, 最后将断裂聚合问题转换为图割优化问题, 从而恢复全脑血管的连通性。

(1) 初步检测断裂

为归纳断裂处血管对的特征, 观察分析原图、分割图和追踪结果, 得到如下分布规律:

- 断裂处的血管对距离很近
- 断裂处距离最近的两个点中至少一个点是待连接片段末端
- 存在两种断裂类别: 一是血管沿主干分成两半(图 13 左), 二是在分叉处断裂(图 13 右)。

如图 13 是两种断裂情况的示意图, 其中不同颜色表示不同的血管段, 血管段上的不同颜色点表示两端待连接片段距离彼此最近的两个点。如左图, 类别一中的两根血管距离近, 且方向有一致性, 而类别二的血管对虽然方向并不一致, 但仍呈逼近趋势。

基于上述规律, 本设计拟先沿端点查找所有血管段存在可能连接的候选段, 并建立无向无权图来表示血管段之间的可能连接关系。

为找到可能与血管段 S_i 存在连接的所有候选段, 首先从追踪结果中找出 S_i 的所有端点

$E_i = \{e_1, \dots, e_N\}$ ，寻找端点一定距离内的血管段为候选段，得到各端点和对应候选段的集合 $P_i = \{(e_{x1}, S_{y1}), \dots, (e_{xN}, S_{yM})\}$ 和相应距离集合 $D_i = \{dist_1, \dots, dist_{yM}\}$ ，其中 $dist_n$ 表示第 n 个血管对的距离。如果一个点同时和多个段有可能连接关系，只保留最短距离的连接，得到 S_i 最终的候选段集合 $C_i = \{(S_i, S_{y1}), \dots, (S_i, S_{yM})\}$ 。

我们建立无向图 $G = \{V, E\}$ ，其中结点 $V = \{V_1, \dots, V_I\}$ 对应着血管段 $S = \{S_1, \dots, S_I\}$ ，边 $E = \{E_1, \dots, E_L\} \subseteq V \times V$ 对应血管段的待连接关系。基于所有血管段的候选段集合 $C = \{C_1, \dots, C_I\}$ ，对任意 $i \in \{1, \dots, I\}$ 、 $j \in \{1, \dots, I\}$ 且 $i \neq j$ ，如果 $(S_i, S_j) \in C_i$ 且 V_i 、 V_j 未连接，则连接 V_i 、 V_j 得到边 E_{ij} ，

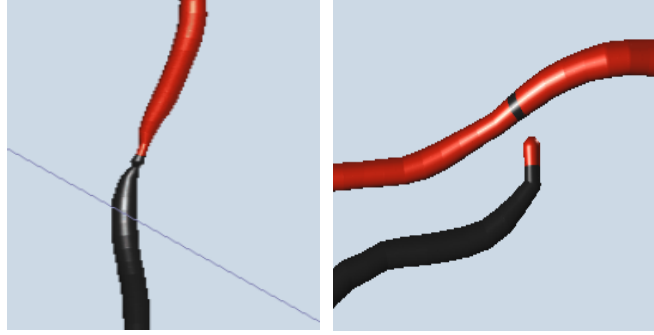


图 13 断裂的两种情况，左图为类别一，右图为类别二

(2) 判断聚合概率

为了利用图割理论聚合断裂血管段，本课题对无向图的每个边赋予权重表示对应的聚合概率值，形式化地，我们需要权重函数 $f(E) \in [0, 1]$ 。本文基于血管骨架训练了一个新的三维卷积神经网络来拟合 f ，利用三维位置关系和三维形状信息来预测血管对的聚合概率作为对应边的权值，从而构建无向有权图 $G = \{V, E, f\}$ ，如图 14。其中骨架是用追踪结果的中心线和半径生成的。

网络结构如图 15，分三个通道输入骨架 A、B 待聚合处的三维数据块，分别是只有骨架 A 的块、只有骨架 B 的块、以及 AB 同时存在的块，输出骨架 A、B 的聚合概率值。训练这个网络需要制作分类数据集。已经观察到的断裂血管对可以作为正样本，为增加正样本数量，本课题拟进行人工截断连续血管，来增加断裂血管对。然而人工截断血管的特征分布不一定和追踪断裂处一致，所以需要施加额外的数据增强增加数据多样性。负样本可以用距离很近的不同血管构成。

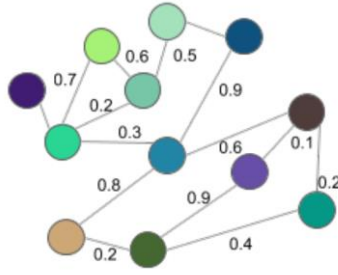


图 14 用无向有权图 $G = \{V, E, f\}$ 表征待连接血管段

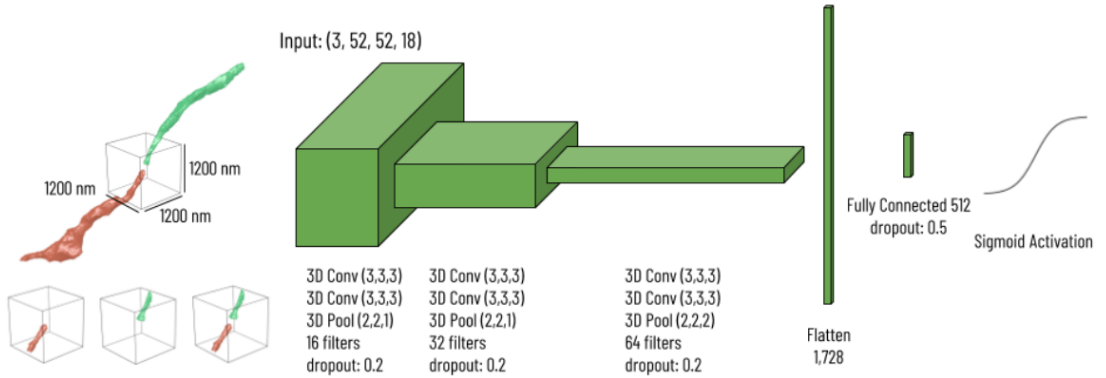


图 15 3D CNN 结合三维位置、形状信息预测聚合概率

(3) 图割优化聚合断裂

不局限于相邻段之间的联系，为获取更远距离的血管对的可能连接关系，本课题把判定聚合断裂问题转换成图割问题来优化。在得到图 15 所示的图网络后，用贪心算法得到图的最佳分割，通过聚合每个子图中边的对应血管段，来聚合全脑血管断裂。

2.3 全脑血管分布的统计分析

基于上述方法，我们可以得到并修正三维果蝇全脑血管的重建结果。为了研究血管形态、统计分布的生物学意义，我们基于这一重建结果对全脑范围的血管进行生物学分析。为了探究血管分布与各功能脑区之间的关系，本课题拟将重建结果匹配到标准脑区模板，并统计不同脑区中血管的分布情况，包括长度、粗细、分叉点密度、形态等。

3 可行性分析

该方法的总体思路建立在对相关文献的调研基础之上，首先三维信息对生物电镜的重要性是不言而喻的，符合认知常识，3D CNN 充分利用了三维信息，并且很多相关研究也证明了其正确性[14-19]。其次，为重建全脑血管的数字模型，本设计中使用了一种更精确高效的追踪方法。接着为了修正分割和追踪带来的误差，进一步连接断裂，最后统计全脑血管分布规律。总体方案理论上是切实可行的。

总体来说，本文提供的方案在理论上是有充足支撑，整体思路是可行的，并基于实践确保方案在工程上是可以实现的。

四、工作进度的大致安排

应包括文献调研，工程设计，新工艺、新材料、新设备、新产品的研制和调试，实验操作，实验数据的分析处理，撰写论文等。

2020 年 05 月-2021 年 02 月：调研血管重建的相关工作；

2021 年 03 月-2021 年 04 月：熟悉 FAFB 数据集，制作血管数据集；

2021 年 05 月-2021 年 07 月：实现全脑血管分割，得到血管分割图；

2021 年 08 月-2021 年 10 月：实现全脑血管追踪，得到血管数字化模型；

2021 年 11 月-2022 年 02 月：聚合断裂血管段，并统计分析，探究生物学意义；

2022 年 02 月-2022 年 05 月：完成硕士毕业论文的撰写，准备硕士论文答辩。

预期成果

完成全脑血管重建，即制作电镜血管数据集，全脑血管分割得到分割图，全脑血管追踪得到数字模型，并在此基础上通过图割优化来聚合断裂血管段。基于最后的重建结果统计血管的分布规律，探究其生物学意义。

导师意见

导师签名：（学校）

（企业）

年 月 日

评审小组意见

开题报告符合要求，通过答辩

评审小组组长签字：

孙光艳

年 月 日

预期成果

完成全脑血管重建，即制作电镜血管数据集，全脑血管分割得到分割图，全脑血管追踪得到数字模型，并在此基础上通过图割优化来聚合断裂血管段。基于最后的重建结果统计血管的分布规律，探究其生物学意义。

导师意见

同意开题

导师签名： (学校) 陈玉锦

(企业) 刘海峰
2021年11月12日

评审小组意见

开题报告符合要求，通过答辩

评审小组组长签字： 孙艳

2021年11月12日