摘要

随着生物医学研究的深入，电子显微镜（Electron Microscopy, EM）已成为解析细胞器超微结构及其生物学功能的核心技术。近年来，高分辨率电镜成像技术的快速发展，使得大规模纳米级图像数据的获取成为可能。然而，这些海量数据的自动化解析，尤其是精准、高效的图像分割，仍然面临诸多挑战，包括图像噪声较高、对比度较低、生物结构复杂性强，以及高质量标注数据的稀缺性。因此，如何利用先进的计算方法，特别是深度学习技术，提升电镜图像分割的准确性、鲁棒性和泛化能力，已成为当前生物医学图像分析领域的重要研究方向。

为应对上述挑战，本文围绕电子显微镜图像分割展开研究，面向生物细胞结构从专用分割模型到通用分割模型，提出了一系列创新方法，包括主观一致性感知分割、证据不确定性修正、多语义适配及自监督预训练，并针对不同层面的关键问题进行了深入探索，具体贡献如下：

第一，针对细胞膜分割中的结构缺失问题，提出了一种基于主观一致性评价的细胞膜分割方法。本文引入感知豪斯多夫距离（Perceptual Hausdorff Distance, PHD），并基于此构建 PS-Net 分割网络。通过全局-局部协同学习策略，有效提升细胞膜分割的完整性和精准度。此外，本文开展眼动追踪实验，深入分析人类视觉在分割评估过程中的注意力分布规律，以此优化分割质量评价标准。实验结果表明，该方法在多个公开电镜图像数据集上提升了细胞膜分割的拓扑完整性和边界精确度。

第二，针对三维线粒体分割的不确定性问题，提出了一种基于证据不确定性修正的线粒体分割方法。为提升分割的可靠性，本文采用证据理论（Dempster-Shafer Theory）量化模型预测的不确定性，并设计多尺度邻域注意力模块来抑制不确定性传播。此外，引入形态学约束机制，增强对细长结构的连续性建模，从而提高分割的连贯性与稳定性。实验结果表明，该方法能够有效减少误分割区域，在多个具有挑战性的电镜图像数据集上均取得优异的分割性能，并在不同成像条件下表现出更强的泛化能力。

第三，针对多语义分割任务的泛化性瓶颈，提出了一种形态引导的基于分割基础模型的微调方法。鉴于电镜图像中多种结构并存且尺度变化显著，本文提出基于 3D Mamba-2 结构的长程依赖建模框架，结合局部曲率与纹理描述符，增强模型的域适配能力，从而优化多语义结构（如细胞膜、线粒体、细胞核等）的分割性能。实验结果表明，该方法在多个电镜图像数据集上均取得最优性能，尤其在跨数据集泛化方面展现出适应性，可有效应对不同生物组织的分割需求。

第四，针对大规模数据异质性挑战，构建了首个面向电子显微镜图像的多任务通

用基础模型 Omni-EM。该模型基于大规模异质性电镜数据集，利用自监督学习预训练构建出视觉编码器 EM-ViT，并结合 U 形结构实现多种分割任务的统一建模。在多个公开数据集上，Omni-EM 展现出优异的零样本泛化能力和任务适应性。此外，本章还引入无监督微调策略，有效提升模型在新数据上的表现。Omni-EM 为电镜图像分析提供了统一、高效的解决方案，展现出广泛的应用潜力。

本文在多个公开电镜图像数据集上进行了系统实验评估，实验结果表明所提出的方法在分割准确性、泛化能力和计算效率方面均优于现有先进方法，提升了电镜图像分析的自动化水平，并降低了生物医学研究对手动标注的依赖。研究成果有助于加速生物医学领域的科学发现，为神经科学、细胞生物学、病理学等方向的深入研究提供了强有力的计算支持。未来，随着大规模预训练模型、自监督学习、多模态融合及生成式人工智能等技术的发展，电镜图像分割有望进一步迈向更高效、更智能的自动化分析体系，为生命科学研究提供更精准、更全面的数据解析能力。

关键词：电子显微镜，图像分割，深度学习，不确定性修正，基础模型

Research on Electron Microscopy Image Segmentation Methods for Cellular Structures

Ruohua Shi (Major in Computer Applied Technology)
Supervised by Prof. Lingyu Duan and A.P. Tingting Jiang

ABSTRACT

With the rapid advancement of biomedical research, electron microscopy (EM) has emerged as a core technique for resolving the ultrastructure of organelles and elucidating their biological functions. In recent years, the development of high-resolution EM imaging technologies has enabled the acquisition of large-scale, nanoscale image datasets. However, the automated analysis of such massive datasets—particularly the accurate and efficient segmentation of EM images—continues to face significant challenges. These include high image noise, low contrast, structural complexity of biological tissues, and the scarcity of high-quality annotated data. Consequently, leveraging advanced computational methods—especially deep learning techniques—to enhance the accuracy, robustness, and generalization ability of EM image segmentation has become a critical research direction in the field of biomedical image analysis.

To address these challenges, this study focuses on the segmentation of EM images, proposing a series of innovative approaches ranging from task-specific to general-purpose segmentation models for biological cellular structures. These include perceptual consistency-aware segmentation, uncertainty correction based on evidential theory, multi-semantic adaptation, and self-supervised pretraining. Key contributions are as follows:

First, to tackle the issue of structural discontinuity in cell membrane segmentation, a method based on perceptual consistency evaluation is proposed. A novel metric, the Perceptual Hausdorff Distance (PHD), is introduced and integrated into a segmentation network called PS-Net. By employing a global-local collaborative learning strategy, the method significantly improves the completeness and accuracy of membrane segmentation. Moreover, eye-tracking experiments are conducted to analyze human visual attention patterns during segmentation assessment, leading to refined evaluation standards. Experimental results demonstrate that the proposed method substantially enhances segmentation completeness across several public EM datasets and more accurately captures biological tissue boundaries compared to traditional

approaches.

Second, to address the uncertainty in 3D mitochondrial segmentation, a method based on evidential uncertainty correction is developed. Using Dempster-Shafer theory, the method quantifies model prediction uncertainty and introduces a multi-scale neighborhood attention module to suppress uncertainty propagation. Additionally, morphological constraints are incorporated to improve continuity modeling of elongated structures, thereby enhancing segmentation coherence and stability. Experimental evaluations reveal that this method effectively reduces mis-segmentation areas and achieves state-of-the-art performance across multiple challenging EM datasets, exhibiting strong generalization under varied imaging conditions.

Third, to overcome generalization bottlenecks in multi-semantic segmentation tasks, a morphology-guided fine-tuning approach is proposed based on a segmentation foundation model. Given the coexistence of diverse structures and significant scale variation in EM images, a long-range dependency modeling framework is built upon the 3D Mamba-2 architecture. By integrating local curvature and texture descriptors, the method enhances domain adaptability and optimizes segmentation of various semantic structures such as membranes, mitochondria, and nuclei. Experimental results show superior performance across multiple EM datasets, particularly in cross-dataset generalization, effectively addressing segmentation needs for different biological tissues.

Fourth, in response to the challenge of large-scale data heterogeneity, the first general-purpose multi-task foundation model for EM images, named Omni-EM, is proposed. Trained on a large-scale heterogeneous EM dataset using self-supervised learning, Omni-EM builds a vision encoder (EM-ViT) and integrates a U-shaped architecture to unify diverse segmentation tasks. The model demonstrates strong zero-shot generalization and task adaptability across several public datasets. Additionally, an unsupervised fine-tuning strategy is introduced to further improve performance on novel datasets. Omni-EM offers a unified and efficient solution for EM image analysis and exhibits broad application potential.

Comprehensive experimental evaluations on multiple public EM datasets confirm that the proposed methods outperform existing state-of-the-art approaches in terms of segmentation accuracy, generalization, and computational efficiency. These advancements significantly improve the automation of EM image analysis and reduce the reliance on manual annotation in biomedical research. The findings contribute to accelerating scientific discovery in fields such as neuroscience, cell biology, and pathology by providing robust computational support. Looking ahead, the integration of large-scale pretraining, self-supervised learning, multimodal

fusion, and generative AI is expected to drive EM image segmentation toward a more intelligent and efficient analytical paradigm, enabling more precise and comprehensive data interpretation for life science research.

KEY WORDS: Electron microscopy, image segmentation, deep learning, uncertainty rectification, foundation model